

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2017285211 B2**

(54) Title
Treatment of clostridium difficile infection

(51) International Patent Classification(s)
A61K 35/74 (2006.01) **C12N 1/20** (2006.01)
A61K 35/742 (2015.01)

(21) Application No: **2017285211** (22) Date of Filing: **2017.06.14**

(87) WIPO No: **WO17/218680**

(30) Priority Data

(31)	Number	(32)	Date	(33)	Country
	62/349,914		2016.06.14		US

(43) Publication Date: **2017.12.21**

(44) Accepted Journal Date: **2024.06.27**

(71) Applicant(s)
Vedanta Biosciences, Inc.

(72) Inventor(s)
Schneider, Jessica;Kim, Yun-Gi;Olle, Bernat;Reddy, Shilpa;Norman, Jason;Patarroyo, Juan

(74) Agent / Attorney
Spruson & Ferguson, GPO Box 3898, Sydney, NSW, 2001, AU

(56) Related Art
SEIKO NARUSHIMA ET AL: "Characterization of the 17 strains of regulatory T cell-inducing human-derived Clostridia", GUT MICROBES, vol. 5, no. 3, 1 May 2014 (2014-05-01), pages 333 - 339, DOI: 10.4161/gmic.28572
WO 2014/121301 A1
CHARLIE G. BUFFIE ET AL: "Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile", HHS PUBLIC ACCESS AUTHOR MANUSCRIPT, vol. 517, no. 7533, 2014-10-22, pages 1 - 25, DOI: 10.1038/nature13828



(51) International Patent Classification:

A61K 35/74 (2015.01) *C12N 1/20* (2006.01)
A61K 35/742 (2015.01)

(21) International Application Number:

PCT/US2017/037498

(22) International Filing Date:

14 June 2017 (14.06.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/349,914 14 June 2016 (14.06.2016) US

(71) Applicant: VEDANTA BIOSCIENCES, INC. [US/US];
19 Blackstone Street, Cambridge, MA 02139 (US).(72) Inventors: SCHNEIDER, Jessica; 444 Windsor Street,
Unit 2, Cambridge, MA 02141 (US). KIM, Yun-Gi; 77
Waverley Avenue, Watertown, MA 02472 (US). OLLE,
Bernat; 68 Brookline Street, Apt. 104, Cambridge, MA
02139 (US). REDDY, Shilpa; 375 Mt Auburn St., Apt.
2R, Watertown, MA 02472 (US). NORMAN, Jason; 179
Norton St., North Weymouth, MA 02191 (US). PATAR-

ROYO, Juan; 1073 Massachusetts Ave., Lexington, MA
02420 (US).

(74) Agent: WITTE-GARCIA, Chelsea, E.; Wolf, Green-
field & Sacks, P.C., 600 Atlantic Avenue, Boston, MA
02210-2206 (US).

(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, JO, 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,

(54) Title: TREATMENT OF CLOSTRIDIUM DIFFICILE INFECTION

Figure 1

Composition A	Composition B	Composition C	Composition D
SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*	SEQ_10 - 211 - Flavonifractor_plautii (IV)	SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*	SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*
SEQ_04 - 7 - Blautia_hansenii (XIVa)*	SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)	SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*	SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*
SEQ_05 - 10 - Blautia_hansenii (XIVa)*	SEQ_15 - VE202-14 - Eubacterium_fissicatena (XIVa)	SEQ_05 - 10 - Blautia_hansenii (XIVa)*	SEQ_05 - 10 - Blautia_hansenii (XIVa)*
SEQ_07 - 59 - Blautia_producta / Blautia_coccoides (XIVa)	SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)	SEQ_01 - 71 - Blautia_wexlerae (XIVa)*	SEQ_01 - 71 - Blautia_wexlerae (XIVa)*
SEQ_08 - 79 - Blautia_hansenii (XIVa)*	SEQ_17 - VE202-7 - Clostridium_bolteae (XIVa)	SEQ_07 - 59 - Blautia_producta/Blautia_coccoides (XIVa)*	SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)
SEQ_09 - VE202-21 - Eubacterium_contortum / Eubacterium_fissicatena (XIVa)*	SEQ_20 - 170 - Dorea_longicatena (XIVa)	SEQ_18 - 148 - Dorea_longicatena (XIVa)	SEQ_18 - 148 - Dorea_longicatena (XIVa)
SEQ_11 - VE202-9 - Anaerostipes_caccae (XIVa)*	SEQ_19 - 16 - Blautia_producta (XIVa)	SEQ_21 - 189 - Clostridium_innocuum (XVII)	SEQ_21 - 189 - Clostridium_innocuum (XVII)
SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*	SEQ_21 - 189 - Clostridium_innocuum (XVII)	SEQ_10 - 211 - Flavonifractor_plautii (IV) /	SEQ_10 - 211 - Flavonifractor_plautii (IV) /
SEQ_13 - 136 - Marvinbryantia_formatexigens (XIVa)*		SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)	SEQ_02 - 102 - Turicibacter_sanguinis (non-Clostridium)
SEQ_23 - VE202-29 - Eisenbergiella_tayi (XIVa)*		SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)	SEQ_06 - 40 - Lactobacillus_mucosae (non-Clostridium)

* = BaiCD*

bolding indicates strains other than Clostridium cluster XIVa

(57) Abstract: Provided herein are compositions and methods for the treatment or prevention of pathogenic infections.

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

TREATMENT OF CLOSTRIDIUM DIFFICILE INFECTION

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional
5 application number 62/349,914, filed June 14, 2016, which is incorporated by reference
herein in its entirety.

FIELD OF INVENTION

The disclosure relates to compositions of purified bacterial strains, and methods for
10 treating pathogenic infections, such as *Clostridium difficile* infections, by administering the
compositions to a subject having a pathogenic infection.

BACKGROUND OF THE INVENTION

The collection of bacterial, viral, and fungal commensal microorganisms that reside
15 within and on the human body are collectively known as the human microbiome. The
bacterial subset of the human microbiome plays an important role in host nutrient acquisition,
development, immunological homeostasis, neurological health, and protection against
pathogens (LeBlanc et al. *Curr. Opin. Biotechnol.* (2013) 24(2): 160-168; Hooper et al.
Science (2012) 336(6086): 1268-1273; Hughes et al. *Am. J. Gastroenterol.* (2013) 108(7):
20 1066-1074). As the largest reservoir of mammalian commensals, bacteria residing in the
gastrointestinal (GI) tract influence nearly all of these aspects of human biology (Blaser *J.*
Clin. Invest. (2014) 124(10): 4162-4165). Consequently, perturbation of the normal bacterial
populations within the GI niche, a state known as dysbiosis, can predispose humans to a
variety of diseases.

25 *Clostridium difficile* infection (CDI) arises after intestinal colonization by the
anaerobic spore-forming Gram-positive pathogen *Clostridium difficile*. Upon colonization of
the GI tract, *C. difficile* produces toxins which causes diarrhea and may ultimately lead to
death. This illness is the most common identifiable cause of nosocomial diarrhea and is
thought to arise as a direct result of dysbiosis (Calfee *Geriatrics* (2008) 63: 10-21; Shannon-
30 Lowe et al *BMJ* (2010) 340: c1296). Not surprisingly, usage of nearly all classes of
antibiotics has been associated with CDI, presumably by inducing dysbiosis in the GI tract
and thereby enabling *C. difficile* outgrowth. The Center for Disease Control currently
classifies CDI as a public health threat requiring immediate and aggressive action because of
its natural resistance to many drugs and the emergence of a fluoroquinolone-resistant strain

that is now prevalent throughout North America and Europe. *C. difficile* was responsible for almost half a million infections and was associated with approximately 29,000 deaths in 2011 (Lessa et al. *NEJM* 2015, 372: 825-834).

The antibiotics metronidazole, vancomycin, and fidaxomicin are the current therapeutic options for treatment of CDI. However, metronidazole is inadequate because of decreased response rates and neither metronidazole nor vancomycin prevent disease recurrence, with up to 30% of patients initially responding experiencing a clinical recurrence after antibiotic cessation (Miller *Expert Opin. Pharmacother.* (2010) 11: 1569-1578).

Fidaxomicin has been shown to be superior to vancomycin in preventing recurrent CDI (Mullane *Ther. Adv. Chronic Dis.* (2014) 5(2): 69-84). Because of its narrow spectrum of activity, fidaxomicin is thought to enable normal microbiome repopulation of the gut following dysbiosis and CDI, thereby lowering the likelihood of recurrent disease (Tannock et al. *Microbiology* (2010) 156 (Pt 11): 3354-3359; Louie et al. *Clin. Infect. Dis.* (2012) 55 Suppl. 2: S132-142). Nonetheless, 14% of fidaxomicin-treated patients experience CDI relapse and mutations conferring reduced sensitivity have already been reported (Eyre et al. *J. Infect. Dis.* (2014) 209(9): 1446-1451).

Because the risk of recurrent CDI is heightened by antibiotic use and *C. difficile* spores are inherently recalcitrant to the available chemotherapeutic arsenal, alternative therapeutic modalities are being pursued for the treatment of CDI. Fecal microbiota transplantation (FMT) is one such modality that has shown efficacy against CDI (Khoruts et al. *Immunol. Lett.* (2014) 162(2): 77-81; van Nood et al. *N. Engl. J. Med.* (2013) 368(5): 407-415). To date, results of FMT studies for the treatment of CDI, have reported cure rates up to 90% in three randomized controlled studies (Cammarota et al. *Alimen. Pharmacol. Therap.* (2015) 41(9): 835-843; Kassam et al. *Am. J. Gastroenterol.* (2013) 108(4): 500-508; van Nood et al. *N. Engl. J. Med.* (2013) 368(5): 407-415; Youngster et al. *Infect. Dis. Soc. Am.* (2014) 58(11): 1515-1522).

Despite the success of FMT, this therapeutic approach is not without risks and logistical concerns. Selection of FMT donors is critical and challenging. When FMT donor recruitment is performed with stringent screening and standardization protocols, most prospective donors fail this process. Only 6-10% of prospective FMT donors qualify, with the majority of failures arising from asymptomatic carriage of GI pathogens (Paramsothy et al. *Inflamm. Bowel Dis.* (2015) 21(7): 1600-1606; Borody et al. *Curr. Opin. Gastroenterol.* (2014) 30(10): 97-105; Burns et al. *Gastroenterology* (2015) 148: S96-S97; Surawicz *Ann. Intern. Med.* (2015) 162(9): 662-663). Furthermore, variation between donors may lead to

variation in FMT efficacy. In addition, the risk of transmission of even non-infectious illnesses may be heightened by FMT. Indeed, significant weight gain has been reported in a patient who received an FMT from an overweight stool donor (Alang *et al. Open Forum Infect. Dis.* (Winter 2015) 2(1)).

Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of common general knowledge in the field.

SUMMARY OF THE INVENTION

Provided herein are compositions and methods for the treatment or prevention of pathogenic infections including *C. difficile*.

In one aspect, the present disclosure provides a composition comprising:

- (i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
- (ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
- (iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
- (iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
- (v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
- (vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;
- (vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and
- (viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

In another aspect, the present disclosure provides a composition comprising:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;

- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

In another aspect, the present disclosure provides a composition comprising:

- (i) an isolated bacterial strain belonging to the species *Clostridium boltea*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

In another aspect, the present disclosure provides food product comprising a composition of the invention and a nutrient.

In another aspect, the present disclosure provides a pharmaceutical composition comprising a composition of the invention and a pharmaceutically acceptable excipient.

In another aspect, the present disclosure provides method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:

- (i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
- (ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
- (iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
- (iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
- (v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
- (vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;

(vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and

(viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

In another aspect, the present disclosure provides a method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:

(i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

(ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;

(iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;

(iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;

(v) an isolated bacterial strain belonging to the species *Blautia producta*;

(vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;

(vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and

(viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

In another aspect, the present disclosure provides a method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:

(i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

(ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;

(iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;

(iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;

(v) an isolated bacterial strain belonging to the species *Blautia producta*;

(vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;

(vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and

(viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

In another aspect, the present disclosure provides a kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

(i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;

(ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;

(iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;

(iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;

(v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;

(vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;

(vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and

(viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

In another aspect, the present disclosure provides a kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

(i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

(ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;

(iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;

(iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;

(v) an isolated bacterial strain belonging to the species *Blautia producta*;

(vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;

(vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and

(viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

In another aspect, the present disclosure provides a kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

(i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

(ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;

(iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;

(iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;

(v) an isolated bacterial strain belonging to the species *Blautia producta*;

- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavinifractor plautii*.

In another aspect, the disclosure provides compositions comprising two or more purified bacterial strains of species selected from the group consisting of: *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia producta* ATCC 27340, *Clostridium bacterium* UC5.1-1D4, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimonas intestinalis*, *Dracourtella massiliensis*, *Dracourtella massilinesis* GD1, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella tayi*, *Flavinifractor plautii*, *Clostridium orbiscindens* 1_3_50AFAA, *Lachnospiraceae bacterium* 7_1_58FAA, *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus colihominis* DSM 17241, *Clostridium symbiosum*, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae*, *Clostridium bolteae* 90A9, *Dorea longicatena*, *Dorea longicatena* CAG:42, *Clostridium innocuum*, *Erysipelatoclostridium ramosum*, *Blautia wexlerae*, *Clostridium disporicum*, *Erysipelatoclostridium ramosum*, *Pseudoflavinifractor capillosus*, *Turicibacter sanguinis*, *Lactobacillus mucosae*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter saccharivorans*, *Roseburia faecis*, *Blautia faecis*, *Dorea formicigenerans* and *Bacteroides ovatus*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia producta* ATCC 27340, *Clostridium bacterium* UC5.1-1D4, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimonas intestinalis*, *Dracourtella massiliensis*, *Dracourtella massilinesis* GD1, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella tayi*, *Flavinifractor plautii*, *Clostridium orbiscindens* 1_3_50AFAA, *Lachnospiraceae bacterium* 7_1_58FAA, *Subdoligranulum*,

Anaerotruncus colihominis, *Anaerotruncus colihominis* DSM 17241, *Clostridium symbiosum*, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae*, *Clostridium bolteae* 90A9, *Dorea longicatena*, *Dorea longicatena* CAG:42, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Blautia wexlerae*, *Turicibacter sanguinis*,
 5 *Lactobacillus mucosae*, and *Bacteroides ovatus*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Anaerostipes caccae*, *Clostridium*
 10 *scindens*, *Marvinbryanta formatexigens*, and *Eisenbergiella tayi*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Flavinofractor plautii*, *Clostridium orbiscindens 1_3_50AFAA*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus*
 15 *colihominis* DSM 17241, *Eubacterium fissicatena*, *Sellimonas intestinalis*, *Dracourtella massiliensis*, *Dracourtella massilinesis GD1*, *Ruminococcus torques*, *Clostridium symbiosum*, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae*, *Clostridium bolteae* 90A9, *Dorea longicatena*, *Dorea longicatena* CAG:42, *Blautia producta*, *Blautia producta* ATCC 27340, *Clostridium bacterium UC5.1-1D4*, *Clostridium innocuum*, and
 20 *Erysipelotrichaceae_bacterium_21-3*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium orbiscindens 1_3_50AFAA*, *Anaerotruncus colihominis DSM 17241*, *Dracourtella massilinesis GD1*, *Clostridium symbiosum WAL-14163*, *Clostridium bolteae*
 25 *90A9*, *Dorea longicatena CAG:42*, *Clostridium bacterium UC5.1-1D4*, and *Erysipelotrichaceae_bacterium_21-3*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium orbiscindens 1_3_50AFAA*, *Anaerotruncus colihominis DSM 17241*,
 30 *Sellimonas intestinalis*, *Clostridium symbiosum WAL-14163*, *Clostridium bolteae 90A9*, *Dorea longicatena CAG:42*, *Clostridium bacterium UC5.1-1D4*, and *Erysipelotrichaceae_bacterium_21_3*.

In some embodiments of the compositions provided herein, the composition comprises purified bacterial strains *Clostridium orbiscindens 1_3_50AFAA*, *Anaerotruncus*

colihominis DSM 17241, *Dracourtella massilinesis* GD1, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae* 90A9, *Dorea longicatena* CAG:42, *Clostridium bacterium* UC5.1-1D4, and *Erysipelotrichaceae_bacterium_21-3*.

In some embodiments of the compositions provided herein, the composition
5 comprises purified bacterial strains *Clostridium orbiscindens* 1_3_50AFAA, *Anaerotruncus colihominis* DSM 17241, *Sellimonas intestinalis* GD1, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae* 90A9, *Dorea longicatena* CAG:42, *Clostridium bacterium* UC5.1-1D4, and *Erysipelotrichaceae_bacterium_21_3*.

In some embodiments of the compositions provided herein, the composition
10 comprises two or more purified bacterial strains of species selected from the group consisting of: *Flavinofractor plautii*, *Anaerotruncus colihominis*, *Dracourtella massiliensis*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

In some embodiments of the compositions provided herein, the composition
15 comprises purified bacterial strains *Flavinofractor plautii*, *Anaerotruncus colihominis*, *Dracourtella massiliensis*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

In some embodiments of the compositions provided herein, the composition
20 comprises two or more purified bacterial strains of species selected from the group consisting of: *Flavinofractor plautii*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

In some embodiments of the compositions provided herein, the composition
25 comprises purified bacterial strains *Flavinofractor plautii*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

In some embodiments of the compositions provided herein, the composition
30 comprises two or more purified bacterial strains of species selected from the group consisting of: *Flavinofractor plautii*, *Lachnospiraceae bacterium* 7_1_58FAA, *Subdoligranulum*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Ruminococcus torques*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, and *Bacteroides ovatus*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium*

orbiscindens 1_3_50AFAA, *Anaerotruncus colihominis* DSM 17241, *Dracourtella massiliensis* GD1, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae* 90A9, *Dorea longicatena* CAG:42, *Clostridium bacterium* UC5.1-1D4, *Erysipelotrichaceae_bacterium_21-3*, and *Bacteroides ovatus*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium orbiscindens 1_3_50AFAA*, *Anaerotruncus colihominis* DSM 17241, *Sellimonas intestinalis*, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae* 90A9, *Dorea longicatena* CAG:42, *Clostridium bacterium* UC5.1-1D4, *Erysipelotrichaceae_bacterium_21-3*, and *Bacteroides ovatus*.

In some embodiments of the compositions provided herein, the composition does not include a bacterial strain of the species *Flavinofractor plautii*, *Subdoligranulum*, or *Lachnospiraceae bacterium 7_1_58FAA*. In some embodiments of the compositions provided herein, the composition does not include a bacterial strain of the species *Bacteroides ovatus*. The composition of any one of claims 4-12, wherein the composition does not include a bacterial strain of the species *Flavinofractor plautii*, *Subdoligranulum*, *Clostridium orbiscindens 1_3_50AFAA*, or *Lachnospiraceae bacterium 7_1_58FAA*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Blautia producta*, *Blautia coccoides*, *Dorea longicatena*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Flavinofractor plautii*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, and *Clostridium symbiosum*. In some embodiments of the compositions provided herein, the composition does not include a bacterial strain of the species *Flavinofractor plautii*, *Subdoligranulum*, or *Lachnospiraceae bacterium 7_1_58FAA*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Anaerotruncus colihominis*, *Dorea longicatena*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Flavinofractor plautii*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Turicibacter sanguinis*, and *Lactobacillus mucosae*. In some embodiments of the compositions provided herein, the composition does not include a bacterial strain of the species *Flavinofractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium 7_1_58FAA*.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Dorea longicatena*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*,
5 *Anaerostipes hadrus*, *Flavinofractor plautii*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Ruminococcus faecis*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter saccharivorans*, *Blautia faecis*, *Dorea formicigenerans*, and *Blautia hansenii*.

10 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains of species selected from the group consisting of: *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Anaerostipes hadrus*, *Blautia faecis*, *Blautia hansenii*, *Dorea formicigenerans*, *Dorea longicatena*, *Eubacterium rectale*, *Flavinofractor plautii*, *Fusicatenibacter saccharivorans*, *Megasphaera elsdenii*, *Roseburia faecis*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Ruminococcus faecis*, and
15 *Ruminococcus obeum*.

In one aspect the disclosure provides compositions comprising two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-83 and 124-159.

20 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-23, SEQ ID NO:83, SEQ ID NOs: 124-159.

25 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ
30 ID NO:13, and SEQ ID NO:23.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID

NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21. In some embodiments of the compositions provided herein, the composition comprises purified bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO: 129, SEQ ID NO: 132, SEQ ID NO: 137, SEQ ID NO: 141, SEQ ID NO: 146, SEQ ID NO: 152, and SEQ ID NO: 157. In some embodiments of the compositions provided herein, the composition comprises purified bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO: 129, SEQ ID NO: 132, SEQ ID NO: 137, SEQ ID NO: 141, SEQ ID NO: 146, SEQ ID NO: 152, and SEQ ID NO: 157.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20 and SEQ ID NO:21, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from

the group consisting of SEQ ID NOs:124-156, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

In some embodiments of the compositions provided herein, the composition
5 comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21 and SEQ ID NO:22. In some embodiments of the compositions provided herein, the composition

10 comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequence selected from the group consisting of SEQ ID NO: 124-145, SEQ ID NO: 152-159, SEQ ID NO: 18, and SEQ ID NO: 22.

In some embodiments of the compositions provided herein, the composition
15 comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21 and SEQ ID NO:22, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having
20 at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124-145 and SEQ ID NO: 152-156, and wherein the
25 composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

In some embodiments of the compositions provided herein, the composition
30 comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10 and SEQ ID NOs:14-22. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise

16S rDNA sequences having at least 97% homology with nucleic acid sequence selected from the group consisting of SEQ ID NOs: 124-159, SEQ ID NO: 18, and SEQ ID NO: 22.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:14-22 and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 129-156, SEQ ID NO: 18, SEQ ID NO: 22, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21 and SEQ ID NO:83. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID Nos: 124-159 and SEQ ID NO: 83.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21 and SEQ ID NO:83, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial

strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-156 and SEQ ID NO: 83, and wherein composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:83. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, SEQ ID NO: 22, and SEQ ID NO: 83,

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:83, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-145, SEQ ID NOs: 152-156, SEQ ID NO: 22, and SEQ ID NO: 83, wherein composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NOs:14-22, and SEQ ID NO:83. In some embodiments of the compositions provided herein, the composition

comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, SEQ ID NO: 18, SEQ ID NO:22, and SEQ ID NO: 83.

5 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:14-22 and SEQ ID NO:83, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence
10 having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124-15, SEQ ID NO: 18, SEQ ID NO:22, and SEQ ID
15 NO: 83, wherein composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:157-159.

 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial
20 strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, and SEQ ID NO:21.

 In some embodiments of the compositions provided herein, the composition
25 comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:18, and SEQ ID NO:21.

30 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:24-79.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:24-27, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:51, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:76 and SEQ ID NO:77.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:21, and SEQ ID NOs:80-82.

In one aspect the disclosure provides compositions comprising two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:84-123.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:121, and SEQ ID NO:122.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:109, and SEQ ID NO:121.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid

sequences selected from the group consisting of SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122.

5 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence
10 having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:93, SEQ ID NO:95, SEQ ID
15 NO:97, SEQ ID NO:98, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid
20 sequences selected from the group consisting of SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122, and wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

25 In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:103,
30 SEQ ID NO:105, SEQ ID NO:106, and SEQ ID NO:122.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:90, SEQ ID

NO:91, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, and SEQ ID NO:105.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:104, SEQ ID NO:107, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, and SEQ ID NO:120.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:104, SEQ ID NO:107, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, and SEQ ID NO:119.

In some embodiments of the compositions provided herein, the composition comprises two or more purified bacterial strains, wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:86, SEQ ID NO:95, SEQ ID NO:98, SEQ ID NO:110, SEQ ID NO:122, and SEQ ID NO:123.

In some embodiments of the compositions provided herein, the composition comprises at least one bacterial strain from Clostridium cluster XIVa and at least one bacterial strain from Clostridium cluster XVII. In some embodiments of the compositions provided herein, the composition comprises at least one bacterial strain from Clostridium cluster IV and at least one bacterial strain from Clostridium cluster XVII. In some embodiments of the compositions provided herein, the composition comprises at least one bacterial strain from Clostridium cluster XIVa, at least one strain from Clostridium cluster IV and at least one bacterial strain from Clostridium cluster XVII.

In some embodiments of the compositions provided herein, the composition comprises at least one *Bacteroides* strain. In some embodiments of the compositions provided herein, the composition does not include *Clostridium scindens*.

In some embodiments of the compositions provided herein, the composition comprises at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or at least 20 purified bacterial strains.

5 In some embodiments of the compositions provided herein, one or more of the bacterial strains are spore formers. In some embodiments of the compositions provided herein, one or more of the bacterial strains are in spore form. In some embodiments of the compositions provided herein, each of the bacterial strains is in spore form.

10 In some embodiments of the compositions provided herein, one or more of the bacterial strains is in vegetative form. In some embodiments of the compositions provided herein, each of the bacterial strains is in vegetative form.

In some embodiments of the compositions provided herein, the composition comprises only obligate anaerobic bacterial strains. In some embodiments of the compositions provided herein, the composition comprises bacterial strains that originate from
15 more than one human donor.

In some embodiments of the compositions provided herein, one or more of the bacterial strains are *baiCD*⁻. In some embodiments of the compositions provided herein, each of the bacterial strains is *baiCD*⁻. In some embodiments of the compositions provided herein, the composition does not mediate bile acid 7- α -dehydroxylation. In some
20 embodiments of the compositions provided herein, the composition inhibits *C. difficile* toxin production. In some embodiments of the compositions provided herein, the composition inhibits *C. difficile* replication and/or survival.

In some embodiments of the compositions provided herein, the bacterial strains are lyophilized.

25 In some embodiments of the compositions provided herein, the composition induces the proliferation and/or accumulation of regulatory T cells (Tregs).

In one aspect, the disclosure provides compositions comprising two or more purified bacterial strains, wherein the composition comprises at least one bacterial strain from Clostridium cluster XIVa and at least one bacterial strain from Clostridium cluster XVII. In
30 one aspect, the disclosure provides compositions comprising two or more purified bacterial strains, wherein the composition comprises at least one bacterial strain from Clostridium cluster IV and at least one bacterial strain from Clostridium cluster XVII. In one aspect, the disclosure provides compositions comprising two or more purified bacterial strains, wherein the composition comprises at least one bacterial strain from Clostridium cluster IV, at least

one bacterial strain from Clostridium cluster XIVa and at least one bacterial strain from Clostridium cluster XVII.

In some embodiments of the compositions provided herein, the composition comprises at least one *Bacteroides* strain. In some embodiments of the compositions provided herein, the composition does not include *Clostridium scindens*.

In some embodiments of the compositions provided herein, the composition comprises at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or at least 20 purified bacterial strains.

In some embodiments of the compositions provided herein, one or more of the bacterial strains are spore formers. In some embodiments of the compositions provided herein, one or more of the bacterial strains are in spore form. In some embodiments of the compositions provided herein, each of the bacterial strains is in spore form.

In some embodiments of the compositions provided herein, one or more of the bacterial strains is in vegetative form. In some embodiments of the compositions provided herein, each of the bacterial strains is in vegetative form.

In some embodiments of the compositions provided herein, the composition comprises only obligate anaerobic bacterial strains.

In some embodiments of the compositions provided herein, the composition comprises bacterial strains that originate from more than one human donor.

In some embodiments of the compositions provided herein, one or more of the bacterial strains are *baiCD*⁻. In some embodiments of the compositions provided herein, each of the bacterial strains is *baiCD*⁻. In some embodiments of the compositions provided herein, the composition does not mediate bile acid 7- α -dehydroxylation. In some embodiments of the compositions provided herein, the composition inhibits *C. difficile* toxin production. . In some embodiments of the compositions provided herein, the composition inhibits *C. difficile* replication and/or survival.

In some embodiments of the compositions provided herein, the bacterial strains are lyophilized.

In some embodiments of the compositions provided herein, the composition induces the proliferation and/or accumulation of regulatory T cells (Tregs).

In one aspect, the disclosure provides a pharmaceutical composition comprising any of the compositions provided herein further comprising a pharmaceutically acceptable excipient. In some embodiments of the pharmaceutical compositions provided herein, the

pharmaceutical composition is formulated for oral delivery. In some embodiments of the pharmaceutical compositions provided herein, the pharmaceutical composition is formulated for rectal delivery. In some embodiments of the pharmaceutical compositions provided herein, the pharmaceutical composition is formulated for delivery to the intestine. In some
5 embodiments of the pharmaceutical compositions provided herein, the pharmaceutical composition is formulated for delivery to the colon. In one aspect, the disclosure provides a food product comprising any of the compositions provided herein further comprising a nutrient.

In one aspect, the disclosure provides a method of treating a pathogenic infection in a
10 subject, comprising administering to the subject a therapeutically effective amount of any of the compositions or food products provided herein to treat the pathogenic infection.

In some embodiments of the methods provided herein, the pathogenic infection is *C. difficile*, Vancomycin Resistant *Enterococci* (VRE), Carbapenem Resistant *Enterobacteriaceae* (CRE), *Neisseria gonorrhoeae*, Multidrug Resistant *Acinetobacter*,
15 *Campylobacter*, Extended spectrum beta-lactamase (ESBL) producing *Enterobacteriaceae*, Multidrug Resistant *Pseudomonas aeruginosa*, *Salmonella*, Drug resistant non-typhoid *Salmonella*, Drug resistant *Salmonella Typhi*, Drug resistant *Shigella*, Methicillin Resistant *Staphylococcus aureus*, Drug resistant *Streptococcus pneumoniae*, Drug resistant Tuberculosis, Vancomycin resistant *Staphylococcus aureus*, Erythromycin Resistant Group A
20 *Streptococcus*, Clindamycin resistant Group B *Streptococcus*, and combinations thereof. In some embodiments of the methods provided herein, the pathogenic infection is *C. difficile*. In some embodiments of the methods provided herein, the pathogenic infection is Vancomycin-Resistant *Enterococci*.

In some embodiments of the methods provided herein, the subject is human. In some
25 embodiments of the methods provided herein, the subject is an asymptotic carrier.

In some embodiments of the methods provided herein, the subject is administered a dose of an antibiotic prior to administration of the composition. In some embodiments of the methods provided herein, the subject is administered more than one dose of the antibiotic prior to administration of the composition. In some embodiments of the methods provided
30 herein, the subject has not been administered an antibiotic prior to administration of the composition.

In some embodiments of the methods provided herein, the composition is administered to the subject by oral administration. In some embodiments of the methods provided herein, the composition is administered to the subject by rectal administration.

In some embodiments of the methods provided herein, the administering results in proliferation and/or accumulation of regulatory T cells (Tregs).

Each of the limitations of the invention can encompass various embodiments of the invention. It is, therefore, anticipated that each of the limitations of the invention involving any one element or combinations of elements can be included in each aspect of the invention. This invention is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced or of being carried out in various ways.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings are not intended to be drawn to scale. The figures are illustrative only and are not required for enablement of the disclosure. For purposes of clarity, not every component may be labeled in every drawing. In the drawings:

Figure 1 shows the strains of Compositions A-D. Each entry includes the SEQ ID NO of the 16S rDNA sequence of the strain, a strain identifier, and the species with the closest known homology (can be more than one species). The bracketed roman numeral indicates the *Clostridium* cluster classification of each strain based on the closest species homology. Strains that are not classified in Cluster XIVa are highlighted in bold. The two non-clostridial strains (SEQ ID NO:2, closest known species *Turicibacter sanguinis*, and SEQ ID NO:6, closest known species *Lactobacillus mucosae*) do not belong to the *Clostridium* genus.

Figure 2 shows various *Clostridium difficile* infection models. Timelines indicate antibiotic type, duration of treatment, as well as exposure to *C. difficile* spores. The top panel shows an antibiotic cocktail treatment model in which the antibiotic cocktail is provided in the drinking water from day -10 to day -3 followed by intraperitoneal clindamycin on day -1. The middle panel shows a clindamycin IP injection model, in which clindamycin is administered by intraperitoneal injection on day -1. The bottom panel shows the cefoperazone treatment model, in which cefoperazone is provided in the drinking water from day -12 to day -2, followed by administration of a live biotherapeutic product (LBP) on day -1.

Figure 3 shows the experimental conditions described in Example 1. The groups of mice were divided based on the antibiotic regimen received prior to administration of the indicated amount of *C. difficile* spores. “Abx” refers to treatment with any of the antibiotic regimens.

Figures 4A-4L show data obtained in Example 1. Figures 4A-4D show survival of mice that received no treatment (Figure 4A), antibiotic cocktail (Figure 4B), clindamycin (Figure 4C), or cefoperazone (Figure 4D) prior to *C. difficile* infection. Figures 4E-4H show body weight of mice that received no treatment (Figure 4E), antibiotic cocktail (Figure 4F), clindamycin (Figure 4G), or cefoperazone (Figure 4H) prior to *C. difficile* infection. Figures 4I-4L show *C. difficile* burden (CFU) per gram of feces from mice that received no treatment (Figure 4I), antibiotic cocktail (Figure 4J), clindamycin (Figure 4K), or cefoperazone (Figure 4L) prior to *C. difficile* infection. Open circles indicate infection with 10 *C. difficile* spores; closed squares indicate infection with 10,000 *C. difficile* spores. Black triangles in Figure 4J indicate an additional experimental arm in which mice were treated with vancomycin following *C. difficile* infection.

Figure 5 shows experimental conditions evaluated in Example 2, the results for which are presented in Figures 7-9. Composition E corresponds to a mixture of 17 bacterial strains (See *e.g.*, Narushima *et al.*, Gut Microbes 5: 3, 333-339). Composition I corresponds to a mixture of *Clostridium scindens*, *Pseudoflavonifractor capillosus*, and *Blautia hansenii*. “Abx” refers to treatment with any of the antibiotic regimens.

Figure 6 shows survival of mice over time post infection with *C. difficile* spores, according to the experimental conditions shown in Figure 5. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves.

Figures 7A-7I show weight of the mice at various times post infection with *C. difficile* spores. Groups of mice received cefoperazone (Abx) treatment followed by the indicated composition, or no cefoperazone (no Abx), then were administered *C. difficile* spores. Figure 7A shows weight of the mice that received no antibiotic treatment. Figure 7B shows weight of the mice that received cefoperazone treatment. Figure 7C shows weight of the mice that received cefoperazone treatment followed by vancomycin. Figure 7D shows weight of the

mice that received cefoperazone treatment followed by Composition I. Figure 7E shows weight of the mice that received cefoperazone treatment followed by Composition E. Figure 7F shows weight of the mice that received cefoperazone treatment followed by composition A. Figure 7G shows weight of the mice that received cefoperazone treatment followed by composition B. Figure 7H shows weight of the mice that received cefoperazone treatment followed by composition C. Figure 7I shows weight of the mice that received cefoperazone treatment followed by composition D.

Figures 8A-8C show the load of *C. difficile* in colony forming units (CFUs) in fecal pellets at various times post infection with *C. difficile*. Figure 8A shows *C. difficile* CFU/g feces one-day post infection. Figure 8B shows *C. difficile* CFU/g feces 3 days post infection. Figure 8C shows *C. difficile* CFU/g feces 8 days post infection.

Figure 9 shows experimental conditions evaluated in Example 3, the results for which are presented in Figures 10-12.

Figure 10 shows survival of the mice over time post infection with *C. difficile* spores, according to the experimental conditions shown in Figure 9. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves.

Figure 11 shows weight of the mice at various times post infection with *C. difficile* spores.

Figure 12 shows the *C. difficile* burden in colony forming units (CFUs) in fecal pellets collected from mice 1, 3, and 8 days post infection with *C. difficile*.

Figure 13 shows the strains of Composition F. The genus-species notation indicates the closest species based on the sequence of the isolated strain.

Figure 14 shows the classification by Clostridium cluster of the strains in Composition F and their short-chain fatty acid producing abilities.

Figure 15 shows experimental conditions evaluated in Example 4, the results for which are presented in Figures 16-18. The dosing days are relative to *C. difficile* infection. FMT refers to Fecal Matter Transplant with fecal matter isolated from mice or from humans.

Figures 16 shows survival of the mice over time post infection with *C. difficile* spores, according to the experimental conditions shown in Figure 15. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves.

Figures 17A-17H show weight of the mice at various times post infection with *C. difficile* spores. Groups of mice received cefoperazone (Abx) treatment followed by the indicated composition, then were administered *C. difficile* spores. Figure 17A shows weight of the mice that received cefoperazone treatment. Figure 17B shows weight of the mice that received cefoperazone treatment followed by FMT with fecal matter from a human. Figure 17C shows weight of the mice that received cefoperazone treatment followed by FMT with fecal matter from a mouse. Figure 17D shows weight of the mice that received cefoperazone treatment followed by Composition B on day -1. Figure 17E shows weight of the mice that received cefoperazone treatment followed by Composition B on days -2 and -1. Figure 17F shows weight of the mice that received cefoperazone treatment followed by Composition B on days -2, -1, 1, 2, and 3. Figure 17G shows weight of the mice that received cefoperazone treatment followed by Composition F on day -1. Figure 17H shows weight of the mice that received cefoperazone treatment followed by Composition F on days -2, -1, 1, 2, and 3.

Figures 18A-18B show the load of *C. difficile* in colony forming units (CFUs) in fecal pellets at various times post infection with *C. difficile*. Figure 18A shows *C. difficile* CFU/g feces 8 days post infection. Figure 18B shows *C. difficile* CFU/g feces 17 days post infection.

Figure 19 shows the strains of Composition G. The genus-species notation indicates the closest species based on the sequence of the isolated strain.

Figure 20 shows experimental conditions evaluated in Example 5, the results for which are presented in Figures 21-23. Composition B1 = Composition B with *Bacteroides*; Composition B2 = Composition B with *Bacteroides* but without *Flavonifractor plautii*.

Figure 21 shows survival of the mice over time post infection with *C. difficile* spores, according to the experimental conditions shown in Figure 20. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves.

5 Figures 22A-22J show weight of the mice at various times post infection with *C. difficile* spores. Figure 22A shows weight of the mice that received vehicle control. Figure 22B shows weight of the mice that received Composition F. Figure 22C shows weight of the mice that received Composition G. Figure 22D shows weight of the mice that received cefoperazone treatment followed by Composition B. Figure 22E shows weight of the mice
10 that received cefoperazone treatment followed by Composition B2 (= Composition B without *Flavonifractor plautii* and with added *Bacteroides*). Figure 22F shows weight of the mice that received cefoperazone treatment followed by Composition B1 (= Composition B with *Bacteroides* added). Figure 22G shows weight of the mice that received cefoperazone treatment followed by frozen Composition B. Figure 22H shows weight of the mice that
15 received cefoperazone treatment followed by ethanol treated human fecal samples. Figure 22I shows weight of the mice that received cefoperazone treatment followed by ethanol treated Composition B. Figure 22J shows weight of the mice that received cefoperazone treatment followed by Composition J.

20 Figure 23 shows the load of *C. difficile* in colony forming units (CFUs) in fecal pellets at various times post infection with *C. difficile*.

Figure 24 shows weight of the indicated groups of mice at various times post infection with *C. difficile* spores.

25

Figure 25 shows experimental conditions evaluated in Example 6, the results of which are presented in Figures 27-29.

Figure 26 shows the strains in Composition H (SEQ ID NO:14 - VE202-13 –
30 *Anaerotruncus colihominis* (Cluster IV); SEQ ID NO:16 - VE202-16 – *Clostridium symbiosum* (Cluster XIVa); SEQ ID NO:21 - 189 – *Clostridium innocuum* (Cluster XVII); SEQ ID NO:82 - PE9 – *Clostridium disporicum* (Cluster I); SEQ ID NO:81 – PE5 – *Clostridium bolteae* (Cluster XIVa); SEQ ID NO:80 – VE202-18 – *Erysipelatoclostridium ramosum* (Cluster XVIII).

Figures 27A and 27B shows survival and weight loss of the mice over time post infection with *C. difficile* spores, according to the experimental conditions shown in Figure 25. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves. Figure 29A shows survival/mortality of mice that received the indicated treatment prior to *C. difficile* infection. Figure 29B shows the weight over time of mice that received the indicated treatment prior to *C. difficile* infection.

Figures 28A and 28B show results from the experimental conditions shown in Figure 25. Figure 28A shows survival/mortality of mice that received the indicated treatment prior to *C. difficile* infection. Figure 28B shows the weight over time of mice that received the indicated treatment prior to *C. difficile* infection.

Figures 29A and 29B show the *C. difficile* burden in CFU/gram feces collected from mice that received the indicated treatment prior to *C. difficile*. Figure 29A shows *C. difficile* burden at one-day post *C. difficile* infection. Figure 29B shows *C. difficile* burden at 4 days post *C. difficile* infection. Figure 29C shows *C. difficile* burden at 19 days post *C. difficile* infection.

Figure 30 shows that Composition B reduced the amount of *C. difficile* Toxin B compared to no treatment controls: “2-1 (Cdiff)” and “2-4 (Cdiff)” and FMT. In addition, Composition B reduced the amount of *C. difficile* Toxin B compared to Composition B with additional spores.

Figure 31 shows Composition B reduced *C. difficile* growth in *in vitro* competition experiments. Cultures of *C. difficile* were incubated in the presence of *B. thetaiotaomicron*, *C. bifermentans*, or Composition B, or in the absence of a competing strain(s) (*C. diff* only). The quantity of *C. difficile* is presented as the percentage of the control (*C. diff* only).

Figure 32 shows that inoculation with Composition B induced the percentage of FoxP3+ CD4+ cells (regulatory T cells) in the intestine of germ-free mice as compared to control mice (“GF”).

DETAILED DESCRIPTION OF THE INVENTION

Disclosed herein are compositions comprising purified bacterial strains and pharmaceutical compositions and food products containing such compositions and bacterial strains. Also disclosed are methods of treating a pathogenic infection, such as *Clostridium*
5 *difficile* (*C. difficile*) infection, in a subject by administering said compositions to the subject.

Various factors including antibiotic usage can induce dysbiosis of the gastrointestinal tract, which may allow for colonization by pathogenic microorganisms, such as *C. difficile*. Such colonization or pathogenic infection can lead to a variety of adverse effects in the subject including diarrhea, which is one of the primary symptoms characteristic of *C. difficile*
10 infection (CDI). In the case of CDI, diarrhea is thought to be a result of *C. difficile* production of Toxin B (also referred to as cytotoxin TcdB), which results in opening of the tight junctions between intestinal epithelial cells, increasing vascular permeability, hemorrhage, and inflammation.

The compositions described herein are effective in the treatment of *C. difficile*
15 infection. As shown herein, the disclosed compositions are effective in suppressing the pathogenic effects of *C. difficile* infection. The compositions provided herein reduce the amount of *C. difficile* after infection and thereby provide an effective method for eliminating *C. difficile* from the body (*e.g.*, the gut). The compositions provided herein induce the proliferation and/or accumulation of regulatory T cells (Tregs), for example when
20 administered to a subject. Remarkably, the compositions disclosed herein have been found to reduce or inhibit production or activity of *C. difficile* Toxin B and thereby represent effective compositions for the treatment or prevention of CDI. The compositions disclosed herein have also been found to inhibit the growth and/or survival of *C. difficile*.

The present disclosure provides compositions comprising purified bacterial strains
25 that can be administered to subjects experiencing or having experienced a pathogenic infection to treat the infection. In some embodiments, the compositions may be administered to subjects who may be at risk for a pathogenic infection. Such subjects include subjects who previously had pathogenic infections, subjects who have been treated with antibiotics and subjects who will undergo a procedure that will put them at an increased risk for a pathogenic
30 infection (*e.g.*, surgery and/or hospitalization). In some embodiments, the pathogenic infection, is infection by a pathogen that is present predominantly in the gut or the intestine. In some embodiments, the pathogen that is present predominantly in the gut or the intestine is *Clostridium difficile*.

In some embodiments, the one or more of the bacterial strains of the compositions provided herein colonize or recolonize the intestinal tract or parts of the intestinal tract (*e.g.*, the colon or the cecum) of the subject. Such colonization or recolonization may also be referred to as grafting. In some embodiments, the one or more of the bacterial strains of the compositions recolonize the intestinal tract (*e.g.*, the colon or the cecum) of the subject after the naturally present microbiome has been partially or completely removed, *e.g.*, because of administration of antibiotics. In some embodiments, the one or more of the bacterial strains of the compositions colonize a dysbiotic gastrointestinal tract.

In some embodiments, the one or more of the bacterial strains of the compositions can “outgrow” a pathogen, such as *C. difficile*. Thus, in some embodiments, if a pathogen (*e.g.*, *C. difficile*) and one or more bacteria of compositions provided herein are both present in the intestinal tract (*e.g.*, the colon or the cecum), the one or more bacteria of compositions provided herein grow faster (*e.g.*, have a shorter doubling time) than the pathogen, thereby preventing the pathogen from accumulating in the intestinal tract (*e.g.*, the colon or the cecum). In some embodiments, the faster growth results because the one or more bacteria of the compositions provided herein are better at grafting in the intestinal tract (*e.g.*, the colon or the cecum). In some embodiments, the faster growth results because the one or more bacteria of the compositions provided herein are better at metabolizing nutrients present in the intestinal tract (*e.g.*, the colon or the cecum). In some embodiments, the compositions of bacterial strains provided herein prevent or inhibit production of bacterial toxins by the pathogenic infection, or prevent or inhibit the cytopathic or cytotoxic effects of such bacterial toxins. In some embodiments, the bacterial strains of the compositions provided herein can treat pathogenic infections, because of the synergy between the bacterial strains. Thus, without being limiting, in some embodiments, the combination of the bacterial strains of the compositions provided herein act synergistically because the combination of the strains is particularly well-suited to use nutrients in the intestinal tract (*e.g.*, the colon or the cecum), or instance through metabolic interactions, and/or because the combination is superior in grafting (*e.g.*, by providing a favorable microenvironment).

In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein is superior in the use of nutrients when compared to the pathogen such as *C. difficile*, thereby suppressing the growth of the pathogen such as *C. difficile*. In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein is superior in grafting when compared to the pathogen such as *C. difficile*,

thereby suppressing the growth of the pathogen such as *C. difficile*. In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein is superior in the use of nutrients and in grafting when compared to the pathogen such as *C. difficile*, thereby suppressing the growth of the pathogen such as *C. difficile*. In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein inhibits the growth and/or survival of the pathogen such as *C. difficile*. In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein induces regulatory T cells (Tregs) in the subject that results in reduction or elimination of the pathogen such as *C. difficile*. In some embodiments, a pathogenic infection such as *C. difficile* is treated because the combination of bacterial strains of the compositions provided herein inhibits the growth and/or survival of the pathogen and induces regulatory T cells (Tregs) in the subject that results in reduction or elimination of the pathogen such as *C. difficile*.

In some embodiments, the synergistic effect is provided by the capacity of the combination to colonize specific niches in the intestinal tract (*e.g.*, the colon or the cecum). In some embodiments, the synergistic effect is provided by the capacity of the combination to metabolize specific nutrients. In some embodiments, the synergistic effect is provided by the capacity of the combination to provide specific metabolites to the environment. Such specific metabolites may suppress growth of the pathogen and/or stimulate growth of non-pathogens. In some embodiments, the synergistic effect is provided by the capacity of the combination to provide short-chain fatty acids to the environment. In some embodiments, the synergistic effect is provided by the capacity of the combination to provide specific short-chain fatty acids to the environment. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce butyrate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce acetate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce lactate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce propionate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce succinate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce multiple metabolites. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce multiple short-chain fatty acids. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce both butyrate and acetate. In some embodiments, the

synergistic effect is provided by the capacity of the combination to produce both butyrate and lactate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce both butyrate and propionate. In some embodiments, the synergistic effect is provided by the capacity of the combination to produce both butyrate and succinate.

5 In some embodiments, the synergistic effect is provided by the capacity of the combination to produce butyrate, acetate and additional short-chain fatty acids.

The bacterial strains used in the compositions provided herein generally are isolated from the microbiome of healthy individuals. In some embodiments, the compositions include strains originating from a single individual. In some embodiments, the compositions include
10 strains originating from multiple individuals. In some embodiments, the bacterial strains are obtained from multiple individuals, isolated and grown up individually. The bacterial compositions that are grown up individually may subsequently be combined to provide the compositions of the disclosure. It should be appreciated that the origin of the bacterial strains of the compositions provided herein is not limited to the human microbiome from a healthy
15 individual. In some embodiments, the bacterial strains originate from a human with a microbiome in dysbiosis. In some embodiments, the bacterial strains originate from non-human animals or the environment (*e.g.*, soil or surface water). In some embodiments, the combinations of bacterial strains provided herein originate from multiple sources (*e.g.*, human and non-human animals).

20 In some embodiments, the bacteria of the compositions provided herein are anaerobic bacteria. In some embodiments, the bacteria of the compositions provided herein are obligate anaerobic bacteria. In some embodiments, the bacteria of the compositions provided herein are *clostridia*. *Clostridia* may be classified into phylogenetic clusters with other closely related strains and species. (See *e.g.*, Rajilic-Stojanovic, M., and de Vos, W.M. *FEMS Microbiol Rev* 38, (2014) 996-1047). In general, *clostridia* are classified as belonging to a
25 specific cluster based on their 16S rRNA (or 16S rDNA) nucleic acid sequence. Methods for determining the identity of specific bacterial species based on their 16S rRNA (or 16S rDNA) nucleic acid sequence are well known in the art (See *e.g.*, *Jumpstart Consortium Human Microbiome Project Data Generation Working, G. PLoS One* (2012) 7, e39315).

30 Provided herein are compositions comprising bacterial strains belonging to specific *Clostridium* clusters that have been found to be effective in treating and/or preventing pathogenic infection (*e.g.*, *C. difficile* infection). In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster IV. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XIVa.

In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster XVII. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster I. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IX. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster XIVa and at least one of the bacterial strains belongs to Clostridium cluster XVII. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV and at least one of the bacterial strains belongs to Clostridium cluster XVII. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV, at least one of the bacterial strains belongs to Clostridium cluster XIVa, and at least one of the bacterial strains belongs to Clostridium cluster XVII.

In some embodiments, the composition has at least twice as many bacterial strains that belong to Clostridium cluster XIVa when compared to the bacterial strains that belong to Clostridium cluster IV. In some embodiments, at least two of the bacterial strains of the composition belong to Clostridium cluster IV and at least five of the bacterial strains belong to Clostridium cluster XIVa. In some embodiments, the composition has at least twice as many bacterial strains that belong to Clostridium cluster XIVa when compared to the bacterial strains that belong to Clostridium cluster IV, and the composition has at least one strain that belongs to Clostridium cluster XVII. In some embodiments, at least two of the bacterial strains of the composition belong to Clostridium cluster IV, at least five of the bacterial strains belongs to Clostridium cluster XIVa, and at least one of the bacterial strains belongs to Clostridium cluster XVII.

In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster XVIII. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster XVI. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster XI. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster I.

In one aspect, the disclosure provides bacterial strains comprising a 16S rDNA sequence with a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-83 and 124-159. It should be appreciated that SEQ ID NOs: 1-83 and 124-159 may include both full length and partial 16S rDNA sequences.

In one aspect, the disclosure provides compositions comprising a bacterial strain comprising a 16S rDNA sequence with a nucleic acid sequence selected from the group

consisting of SEQ ID NOs: 1-83 and 124-159. In one aspect, the disclosure provides compositions comprising as an active ingredient a bacterial strain comprising a 16S rDNA sequence with a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-83 and 124-159. It should be appreciated that for all compositions provided herein, in some
5 embodiments, the bacterial strain or the bacterial strains are the active ingredient of the composition.

It should be appreciated that for all compositions provided herein, in some embodiments, the bacterial strains are purified. Thus, for example the disclosure provides purified bacterial strains comprising a 16S rDNA sequence with a nucleic acid sequence
10 selected from the group consisting of SEQ ID NOs: 1-83 and 124-159. In addition, for example, the disclosure provides compositions comprising purified bacterial strains comprising a 16S rDNA sequence with a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-83 and 124-159. The bacterial strains disclosed herein originally may have been obtained and purified from the microbiota of one or more human
15 individuals or obtained from sources other than the human microbiota, including soil and non-human microbiota. As provided herein, in some embodiments, bacteria isolated from the human microbiota, non-human microbiota, soil, or any alternative source are purified prior to use in the compositions and methods provided herein.

In one aspect, the disclosure provides compositions comprising one or more bacterial
20 strains, wherein the one or more bacterial strains comprise a 16S rDNA sequence with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1-83 and 124-159. In one aspect, the disclosure provides compositions comprising one or more bacterial strains wherein the one or more bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-
25 83 and 124-159. As discussed previously, in some embodiments, the bacterial strains are purified. Thus, in one aspect, the disclosure provides compositions comprising one or more purified bacterial strains wherein the one or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-83 and 124-159.

30 In one aspect, the disclosure provides compositions comprising two or more purified bacterial strains wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-83 and 124-159. As discussed above, in some embodiments, the bacterial strains are the active ingredient of the composition. Thus, in some

embodiments, the disclosure provides compositions comprising as an active ingredient two or more purified bacterial strains wherein the two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-83 and 124-159.

5 In one aspect, the disclosure provides bacterial strains and combinations of bacterial strains that are homologous or have a high percent of homology with bacterial strains comprising 16S rDNA sequences selected from the group consisting of SEQ ID NOs:1-83 and 124-159. As discussed previously, in some embodiments, the bacterial strains are purified. The bacterial strains disclosed herein that have a 16S rDNA sequence with a
10 nucleic acid sequence selected from the group consisting of SEQ ID NOs:1-83 and 124-159 have a high percent of homology (*e.g.*, greater than 90%) with 16S rDNA sequences of bacterial strains that have been described in various databases (See *e.g.*, the National Center for Biotechnology Information). Table 1 and Table 3 provides the closest known species by homology when the 16S rDNA sequences comprising SEQ ID NOs:1-83 and 124-159 are
15 compared to 16S rDNA sequences of bacterial species available in public databases. By way of example, the bacterial strain comprising a 16S rDNA sequence with SEQ ID NO:1 (also referred to herein as “Strain 71”) disclosed herein has the highest homology with a bacterial strain of the species *Blautia wexlerae* as defined by Accession # NR_044054 (having 16S rDNA sequence SEQ ID NO:94). While the bacterial strain with SEQ ID NO:1 has
20 homology with other published bacterial strains as well, the highest homology is with a bacterial strain of the species *Blautia wexlerae* as defined by Accession # NR_044054. In this particular example the homology of SEQ ID NO:1 is 96.6% with SEQ ID NO:94 (corresponding to *Blautia wexlerae*). It should be appreciated that multiple bacterial strains disclosed herein may have the highest homology with the same species. (*e.g.*, both SEQ ID
25 NO:4 and SEQ ID NO:5 have the highest homology with a 16S rDNA sequence of a strain of the species *Blautia hansenii*).

It should further be appreciated that the bacterial strains disclosed herein that have a 16S rDNA sequence with a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1-83 and 124-159, are also homologous to other strains based on their whole genome
30 sequence, or subset of their whole genome sequence. Homologies based on whole genome analysis are provided in Table 2 and Table 3.

In one aspect, the disclosure provides compositions comprising one or more bacterial strains wherein the one or more bacterial strains are of species selected from the group consisting of *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia producta*

- ATCC 27340, *Clostridium bacterium UC5.1-1D4*, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimona intestinalis*, *Dracourtella massiliensis*, *Dracourtella massiliensis GD1*, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella tayi*, *Flavinofractor plautii*,
- 5 *Clostridium orbiscindens 1_3_50AFAA*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus colihominis DSM 17241*, *Clostridium symbiosum*, *Clostridium symbiosum WAL-14163*, *Clostridium bolteae*, *Clostridium bolteae 90A9*, *Dorea longicatena*, *Dorea longicatena CAG:42*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Blautia wexlerae*, *Clostridium disporicum*,
- 10 *Erysipelatoclostridium ramosum*, *Pseudoflavinofractor capillosus*, *Turicibacter sanguinis*, *Lactobacillus mucosae*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter*
- 15 *saccharivorans*, *Roseburia faecis*, *Blautia faecis*, *Dorea formicigenerans* and *Bacteroides ovatus*.

- In some embodiments, the disclosure provides compositions comprising two or more bacterial strains, wherein the two or more bacterial strains are of species selected from the group consisting of *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia*
- 20 *producta ATCC 27340*, *Clostridium bacterium UC5.1-1D4*, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimona intestinalis*, *Dracourtella massiliensis*, *Dracourtella massiliensis GD1*, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella tayi*, *Flavinofractor plautii*, *Clostridium orbiscindens 1_3_50AFAA*, *Lachnospiraceae bacterium 7_1_58FAA*,
- 25 *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus colihominis DSM 17241*, *Clostridium symbiosum*, *Clostridium symbiosum WAL-14163*, *Clostridium bolteae*, *Clostridium bolteae 90A9*, *Dorea longicatena*, *Dorea longicatena CAG:42*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Blautia wexlerae*, *Clostridium disporicum*, *Erysipelatoclostridium ramosum*, *Pseudoflavinofractor capillosus*, *Turicibacter sanguinis*,
- 30 *Lactobacillus mucosae*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter*

saccharivorans, *Roseburia faecis*, *Blautia faecis*, *Dorea formicigenerans* and *Bacteroides ovatus*.

It should be appreciated that the compositions may include multiple strains of a particular species. Thus, for illustration, a non-limiting example of the compositions disclosed herein, comprises one strain of *Clostridium hathewayi* and two strains of *Blautia hansenii*.

The invention also encompasses compositions comprising bacterial strains that are close in homology to and/or fall within the species *Clostridium hathewayi*, *Blautia hansenii*, *Blautia producta*, *Blautia producta* ATCC 27340, *Clostridia bacteria UC5.1-1D4*, *Blautia coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimona intestinalis*, *Dracourtella massiliensis*, *Dracourtella massiliensis GD1*, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella tayi*, *Flavinofractor plautii*, *Clostridium orbiscindens 1_3_50AFAA*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus colihominis DSM 17241*, *Clostridium symbiosum*, *Clostridium symbiosum WAL-14163*, *Clostridium bolteae*, *Clostridium bolteae 90A9*, *Dorea longicatena*, *Dorea longicatena CAG:42*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Blautia wexlerae*, *Clostridium disporicum*, *Erysipelatoclostridium ramosum*, *Pseudoflavinofractor capillosus*, *Turicibacter sanguinis*, *Lactobacillus mucosae*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter saccharivorans*, *Roseburia faecis*, *Blautia faecis*, *Dorea formicigenerans* and *Bacteroides ovatus*. Thus, in one embodiment, the compositions of the disclosure include one or more bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 84-123. In some embodiments, the compositions of the disclosure include two or more bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:84-123.

In one aspect, the compositions of the disclosure include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:1-23 and 124-159. In some embodiments, the compositions of the disclosure include two or more bacterial strains of species selected from the group consisting of *Clostridium hathewayi*, *Blautia hansenii*,

Blautia producta, *Blautia producta* ATCC 27340, *Clostridia* bacteria UC5.1-1D4, *Blautia* *coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Sellimona intestinalis*, *Dracourtella massiliensis*, *Dracourtella massiliensis* GD1, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryanta formatexigens*, *Eisenbergiella* *tayi*, *Flavinofractor plautii*, *Clostridium orbiscindens* 1_3_50AFAA, *Lachnospiraceae* *bacterium* 7_1_58FAA, *Subdoligranulum*, *Anaerotruncus colihominis*, *Anaerotruncus* *colihominis* DSM 17241, *Clostridium symbiosum*, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae*, *Clostridium bolteae* 90A9, *Dorea longicatena*, *Dorea longicatena* CAG:42, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Blautia wexlerae*, *Turicibacter sanguinis*, *Lactobacillus mucosae*, and *Bacteroides ovatus*. In some embodiments, the compositions of the disclosure include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:121, and SEQ ID NO:122.

In one aspect, the disclosure provides Composition A (See *e.g.*, Figure 1, Table A). As shown in Figure 1, Composition A contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with five or more purified bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with at least ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively. In some embodiments, the disclosure provides a composition

consisting of ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively. In some embodiments, the disclosure provides a composition essentially consisting of ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively. As used herein, essentially consisting of refers to a composition that includes no additional bacterial strains.

In some embodiments, the disclosure provides compositions with bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with five or more purified bacterial strains that comprise 16S rDNA having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the disclosure provides compositions with at least ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively. In some embodiments, the disclosure provides a composition consisting of ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively. In some embodiments, the disclosure provides a composition essentially consisting of ten purified bacterial strains, wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7,

SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, respectively.

The bacterial strains in Composition A are related to the following bacterial species:

Clostridium hathewayi, *Blautia hansenii*, *Blautia producta*, *Blautia coccooides*, *Eubacterium*
5 *contortum*, *Eubacterium fissicatena*, *Anaerostipes caccae*, *Clostridium scindens*,

Marvinbryanta formatexigens, and *Eisenbergiella tayi* (See *e.g.*, Table 1). It should be appreciated that multiple bacterial strains of the compositions disclosed herein can have the same related bacterial species. For instance, the bacterial strains having 16S rDNA

sequences with nucleic acid sequences SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:7 all

10 have *Blautia hansenii* as related species. In some embodiments, the disclosure provides

compositions with two or more bacteria of species selected from the group consisting of

Clostridium hathewayi, *Blautia hansenii*, *Blautia producta*, *Blautia coccooides*, *Eubacterium*
contortum, *Eubacterium fissicatena*, *Anaerostipes caccae*, *Clostridium scindens*,

Marvinbryanta formatexigens, and *Eisenbergiella tayi*. In some embodiments, the disclosure

15 provides compositions with two or more purified bacterial strains comprising 16S rDNA

sequences having at least 97% homology with nucleic acid sequences selected from the group

consisting of SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:99, SEQ ID

NO:103, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:109, and SEQ ID
NO:121.

20 In some embodiments, the composition comprises two or more purified bacterial
strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid

sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID

NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and

SEQ ID NO:23. In some embodiments, the composition comprises two or more purified

25 bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic

acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:5, SEQ ID

NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and

SEQ ID NO:23. In some embodiments, the composition comprises two or more purified

bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic

30 acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID

NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and

SEQ ID NO:23. In some embodiments, the composition comprises two or more purified

bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic

acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID

NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID

NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23. In some

embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:8, and SEQ ID NO:13.

Each of the bacterial strains of Composition A are *BaiCD*+, meaning that the bacterial strains encode, or are predicted to encode, the bile inducible operon gene *BaiCD* and/or a protein with stereospecific NAD(H)-dependent 3-oxo- Δ^4 -cholenoic acid oxidoreductase activity. The *BaiCD* status of a bacterial strain can be determined for instance by PCR (See *e.g.*, Wells et al. *Clin Chim Acta* (2003) May; 331(1-2):127-34). Furthermore, each of the strains of Composition A are classified as belonging to Clostridium cluster XIVa. In some embodiments, the disclosure provides compositions comprising two or more bacterial strains, wherein the bacterial strains are *BaiCD*+ strains. In some embodiments, the disclosure provides compositions comprising two or more bacterial strains, wherein the bacterial strains are *BaiCD*+ and belong to Clostridium cluster XIVa. In some embodiments of the compositions comprising two or more bacterial strains that are *BaiCD*+ strains and that belong to Clostridium cluster XIVa, the compositions do not include bacterial strains that belong to Clostridium cluster IV.

In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and SEQ ID NO:23, wherein all the bacterial strains belong to Clostridium cluster XIVa.

Table A

Composition A	
SEQ_03 - 5 -	Clostridium_hathewayi (XIVa)*
SEQ_04 - 7 -	Blautia_hansenii (XIVa)*
SEQ_05 - 10 -	Blautia_hansenii (XIVa)*
SEQ_07 - 59 -	Blautia_producta / Blautia_coccoides (XIVa)
SEQ_08 - 79 -	Blautia_hansenii (XIVa)*
SEQ_09 - VE202-21 -	Eubacterium_contortum / Eubacterium_fissicatena (XIVa)*
SEQ_11 - VE202-9 -	Anaerostipes_caccae (XIVa)*
SEQ_12 - VE202-26 -	Clostridium_scindens (XIVa)*
SEQ_13 - 136 -	Marvinbryantia_formatexigens (XIVa)*
SEQ_23 - VE202-29 -	Eisenbergiella_tayi (XIVa)*

*= BaiCD⁺

In one aspect, the disclosure provides Composition B (See *e.g.*, Figure 1, Table B).

As shown in Figure 1, Composition B contains bacterial strains that comprise 16S rDNA

5 sequences with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In

5 some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In
10 some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

15 In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the compositions include at least eight purified bacterial
20 strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ
25 ID NOs: 124-159, respectively. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition consists of at least eight purified bacterial
30 strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID

NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs:124-159, respectively. In some
5 embodiments, the compositions consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition essentially consists of at least eight purified
10 bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ
15 ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs:124-159, respectively. In some embodiments, the composition essentially consists of at least eight
20 purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group
25 consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the compositions include eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID
30 NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the compositions include eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, respectively. In some embodiments, the compositions include eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from

the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the composition consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs:124-159, respectively. In some embodiments, the compositions consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition essentially consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the composition essentially consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition essentially consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NOs:124-159, respectively. In some embodiments, the composition essentially consists of eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157.

In one aspect, the disclosure provides a composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-159. In one aspect, the disclosure provides a composition that

contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22. In one aspect, the disclosure provides a

composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic

acid sequences: SEQ ID NOs: 124-159. In one aspect, the disclosure provides a composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group

consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the compositions include two or more purified

bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from

the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22. In some

embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the

compositions include two or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157.

In some embodiments, the compositions include five or more purified bacterial strains

comprising 16S rDNA sequences with nucleic acid sequences selected from the group

consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-145, and

SEQ ID NOs: 152-159. In some embodiments, the compositions include five or more

purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences

selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22. In some

embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the

compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

5 In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the compositions include at least
10 eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group
15 consisting of SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. . In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

20 In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the composition consists of at least
25 eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the composition
30 consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO: 22, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In one aspect, the disclosure provides compositions that contain bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-159. In one aspect, the disclosure provides compositions that contain bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In one aspect, the disclosure provides compositions that contain bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of: SEQ ID NOs: 124-159. In one aspect, the disclosure provides compositions that contain bacterial strains that comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In some embodiments, the compositions include two or more purified

bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-159, respectively. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the compositions include at least

eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NOs: 124-159, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs:124-159, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-159, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, respectively. In some embodiments, the composition essentially consists of at least

eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In one aspect, the disclosure provides a composition that contains bacterial strains that
5 comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences :
SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID
NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159.

In one aspect, the disclosure provides a composition that contains bacterial strains that
comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences:
10 SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID
NO:19, SEQ ID NO:21, and SEQ ID NO:22. In one aspect, the disclosure provides a
composition that contains bacterial strains that comprise 16S rDNA sequences having at least
97% homology with nucleic acid sequences: SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ
ID NOs: 152-159.

In some embodiments, the compositions include two or more purified bacterial strains
15 comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences
selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ
ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID
NO: 124-145, and SEQ ID NOs: 152-159. In some embodiments, the compositions include
20 two or more purified bacterial strains comprising 16S rDNA sequences having at least 97%
homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10,
SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID
NO:21, and SEQ ID NO:22. In some embodiments, the compositions include two or more
purified bacterial strains comprising 16S rDNA sequences having at least 97% homology
25 with nucleic acid sequences selected from the group consisting of SEQ ID NO: 22, SEQ ID
NO: 124-145, and SEQ ID NOs: 152-159.

In some embodiments, the compositions include five or more purified bacterial strains
comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences
selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ
30 ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:
124-145, and SEQ ID NOs: 152-159. In some embodiments, the compositions include five
or more purified bacterial strains comprising 16S rDNA sequences having at least 97%
homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10,
SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID

NO:21, and SEQ ID NO:22. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 20, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159.

5 In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the compositions
10 include at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the compositions include at least eight purified bacterial strains comprising 16S rDNA sequences having at
15 least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159.

In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17,
20 SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively.
25 In some embodiments, the composition consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO: 22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159.

In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic
30 acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ

ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the composition essentially consists of at least eight purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO: 22, SEQ ID NO: 124-145, and SEQ ID NOs: 152-159.

In one aspect, the disclosure provides a composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-159. In one aspect, the disclosure provides a composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. In one aspect, the disclosure provides a composition that contains bacterial strains that comprise 16S rDNA sequences with nucleic acid sequences: SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NO: 124-159.

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-159. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NO: 124-159.

In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-159. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10,

SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the compositions include five or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NOs: 124-159.

In some embodiments, the compositions include at least ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-159, respectively. In some embodiments, the compositions include at least ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, respectively.

In some embodiments, the compositions include at least ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 18, SEQ ID NO: 22, and SEQ ID NOs: 124-159.

In some embodiments, the composition consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-159, respectively. In some embodiments, the composition consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the composition consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, SEQ ID NOs: 124-159, respectively.

In some embodiments, the composition essentially consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20,

SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-159, respectively. In some embodiments, the composition essentially consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, respectively. In some embodiments, the composition essentially consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NO: 124-159, respectively.

The bacterial strains in Composition B are related to the following bacterial species: *Flavinofractor plautii*, *Lachnospiraceae*, *bacterium 7_1_58FAA*, *Subdoligranulum Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Ruminococcus torques* *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, *Clostridium innocuum*, and *Erysipelotrichaceae_bacterium_21-3* (See e.g., Table 2).

Selected strains were subjected to whole genome sequencing using a PacBio Biosciences platform (Menlo Park, CA) and sequences were assembled into whole genomes (Table 3). The 16S rDNA sequences were identified using Prokka and Barnap. It was found that several strains contained more than one 16S sequence. All identified 16S rRNA gene nucleotide sequences for each strain were then clustered at 97% identity using the usearch (v 5.2.236) algorithm and the cluster seed sequence was selected as the representative sequence for each Composition B strain (The Consensus 16S sequence: column labeled “*Consensus SEQ ID # of 16S region as determined by WGS” in Table 3). Table 3 provides identification of the indicated strains included in Composition B based on Sanger sequencing of the 16S region as well as on whole genome sequencing (WGS). The closest species of the bacterial strains were identified both by comparison to a 16S database (column labeled: “Closest species based on Consensus SEQ ID # of 16S region as compared with 16S database”) and to whole genome databases (column labeled: “Closest species based on WGS compared versus WG databases”).

Based on identification of 16S sequences through whole genome sequencing, and by comparing these sequences with 16S databases, the bacterial strains in Composition B are related to the following bacterial species: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Dracourtella massiliensis*, *Clostridium symbiosum* *Blautia producta*, *Dorea longicatena* *Clostridium innocuum* and *Flavinofractor plautii* (see, e.g., Table 3).

Based on whole genome sequencing and comparing of the whole genome to whole genome databases, the bacterial strains in Composition B are most closely related to the following bacterial species: *Clostridium bolteae* 90A9, *Anaerotruncus colihominis* DSM 17241, *Dracourtella massiliensis* GD1, *Clostridium symbiosum* WAL-14163, *Clostridium*
 5 *bacterium* UC5.1-1D4, *Dorea longicatena* CAG:42, *Erysipelotrichaceae bacterium* 21_3, and *Clostridium orbiscindens* 1_3_50AFAA (see, e.g., Table 3).

It should be appreciated that multiple strains of the compositions disclosed herein can have the same related bacterial species. For instance, the bacterial strains comprising 16S
 10 rDNA sequences with nucleic acid sequences SEQ ID NO 18, SEQ ID NO:20 and SEQ ID NO:22 all have *Dorea longicatena* as related bacterial species. In some embodiments, the disclosure provides compositions with two or more bacteria selected from the group consisting of *Flavinofractor plautii*, *Lachnospiraceae bacterium* 7_1_58FAA, *Subdoligranulum Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Ruminococcus*
 15 *torques Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, *Clostridium innocuum* and *Erysipelotrichaceae bacterium* 21-3. In some embodiments, the disclosure provides compositions with two or more bacteria selected from the group consisting of *Flavinofractor plautii*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and
 20 *Clostridium innocuum*. In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-159. In some embodiments, the composition
 25 comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA

sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:

5 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152 and SEQ ID NO:157.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID
10 NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18,
15 SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NO: 124-159. In some
20 embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157.

In some embodiments, the composition comprises two or more purified bacterial
25 strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at
30 least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group

consisting of SEQ ID NO:18 and SEQ ID NO: 124-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO: 124-145, and SEQ ID NO: 151-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:22, SEQ ID NO: 124-145, and SEQ ID NO: 151-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:152, and SEQ ID NO:157

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid

sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NOS: 124-145, and SEQ ID NOS: 152-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, and SEQ ID NO:21. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO: 18, SEQ ID NOS: 124-145, and SEQ ID NOS: 152-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NOS: 157-159, and SEQ ID NOS:141-156. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, SEQ ID NOS: 157-159, and SEQ ID NO:141-156.

Each of the bacteria of Composition B are *BaiCD*- strains, meaning that the strains do not encode and/or are not predicted to encode the bile inducible operon gene *baiCD* and/or a protein with stereospecific NAD(H)-dependent 3-oxo- Δ^4 -cholenoic acid oxidoreductase activity. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the bacteria are *BaiCD*- strains. The strains of Composition B are classified as belonging to Clostridium clusters IV, XIVa, and XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters IV, XIVa, or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters IV or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters

XIVa or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters IV or XIVa. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium cluster IV. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium cluster XIVa. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*- strains and belong to Clostridium clusters IV, XIVa, and XVII and do not belong to Clostridium clusters XVI or XVIII.

In some embodiments, the disclosure provides two or more bacterial strains wherein the bacterial strains are spore forming bacterial strains. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria are spore formers and wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:21, SEQ ID NOs 124-140, and SEQ ID NO: 152-156. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria are spore formers and wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:17, and SEQ ID NO:21. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria are spore formers and wherein the bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-140, and SEQ ID NOs: 152-156.

In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria include both spore formers and non-spore formers. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria include both spore formers and non-spore formers, and wherein the spore forming bacterial strains comprise two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:21, SEQ ID NOs: 124-140, and SEQ ID NOs: 152-156. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria include both spore formers and non-spore formers, and wherein the

spore forming bacterial strains comprise two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:17, and SEQ ID NO:21. In some embodiments, the disclosure provides two or more bacterial strains wherein the bacteria include both spore formers and non-spore formers, and wherein the spore forming bacterial strains comprise two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-140 and SEQ ID NO: 152-156.

Table B**Composition B****SEQ_10 - 211 - Flavonifractor_plautii (IV)****SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)**

SEQ_15 - VE202-14 - Eubacterium_fissicatena (XIVa)

SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)

SEQ_17 - VE202-7 - Clostridium_bolteae (XIVa)

SEQ_19 - 16 - Blautia_producta (XIVa)

SEQ_20 - 170 - Dorea_longicatena (XIVa)

SEQ_21 - 189 - Clostridium_innocuum (XVII)

In some embodiments, the compositions include one or more bacterial species from the *Bacteroides* genus. In some embodiments, the compositions include one or more bacterial species selected from the group consisting of *B. acidifaciens*, *B. caccae*, *B. coprocola*, *B. coprosuis*, *B. eggerthii*, *B. finegoldii*, *B. fragilis*, *B. helcogenes*, *B. intestinalis*, *B. massiliensis*, *B. nordii*, *B. ovatus*, *B. thetaiotaomicron*, *B. vulgatus*, *B. plebeius*, *B. uniformis*, *B. salyersai*, *B. pyogenes*, *B. goldsteinii*, *B. dorei*, and *B. johnsonii*. In some embodiments, the compositions include *Bacteroides ovatus*. In some embodiments, the *Bacteroides ovatus* has a 16S rDNA sequence comprising SEQ ID NO:83. In some embodiments, the *Bacteroides ovatus* has a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence comprising SEQ ID NO:83. In some embodiments, the *Bacteroides ovatus* has a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence comprising SEQ ID NO:101.

While not being limited to a specific mechanism it is thought that the inclusion of a *Bacteroides* species in the bacterial compositions disclosed herein increases the ability to

sense and adapt to nutrient availability or influence the host immune system so that it becomes more effective in fighting pathogens (*e.g.*, *C. difficile*). In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO: 124-159, and SEQ ID NO:83. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:83. (Composition B1, See *e.g.*, Table B1). In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, and SEQ ID NO: 83.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-145, SEQ ID NO: 152-159, and SEQ ID NO:83. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NO:83. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-145, SEQ ID NO: 152-159, and SEQ ID NO: 83.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO: 124-159, and SEQ ID NO:83. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected

from the group consisting of SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:83. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97%

5 homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-159, SEQ ID NO: 18, SEQ ID NO: 22, and SEQ ID NO: 83.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16s rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122. It should also be appreciated that in some embodiments, the compositions disclosed herein do not include bacterial species from the *Bacteroides* genus.

Table B1

Composition B1

SEQ_10 - 211 - Flavonifractor_plautii (IV)

SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)

SEQ_15 - VE202-14 - Eubacterium_fissicatena (XIVa)

SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)

SEQ_17 - VE202-7 - Clostridium_bolteae (XIVa)

SEQ_20 - 170 - Dorea_longicatena (XIVa)

SEQ_19 - 16 - Blautia_producta (XIVa)

SEQ_21 - 189 - Clostridium_innocuum (XVII)

SEQ_83 *Bacteroides ovatus*

In some embodiments, the compositions disclosed herein do not include *Clostridium orbiscindens* 1_3_50AFAA, *Flavinofractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium* 7_1_58FAA. In some embodiments, the compositions disclosed herein do not include *Clostridium orbiscindens* 1_3_50AFAA. In some embodiments, the compositions disclosed herein do not include *Flavinofractor plautii*. In some embodiments, the compositions disclosed herein do not include *Subdoligranulum*. In some embodiments, the compositions disclosed herein do not include *Lachnospiraceae bacterium* 7_1_58FAA.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO: 15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO: 124-156, wherein the composition does not include a bacterial strain comprising a 16s rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO: 15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, wherein the composition does not include a bacterial strain comprising a 16s rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NOs: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs:157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-146 and SEQ ID NO: 152-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, and SEQ ID NO:22, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:22, SEQ ID NOs: 124-146, and SEQ ID NOs: 152-156, wherein the composition does not

include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs:157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, and SEQ ID NOs: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:22, and SEQ ID NOs: 124-159, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs: 157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

In some embodiments, the compositions include one or more bacterial species from the *Bacteroides* genus and do not include *Clostridium orbiscindens* 1_3_50AFAA, *Flavinofractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium 7_1_58FAA*. (Composition B2, See *e.g.*, Table B2). In some embodiments, the compositions include *Bacteroides ovatus* and do not include *Clostridium orbiscindens* 1_3_50AFAA, *Flavinofractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium 7_1_58FAA*.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid

sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21 SEQ ID NO:83, and SEQ ID NOs: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21 and SEQ ID NO:83, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:83 and SEQ ID NOs: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs: 157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:83, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:83, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:22 SEQ ID NO:83, SEQ ID NOs: 124-145, and SEQ ID NOs: 152-156,, wherein the composition does not include a bacterial strain comprising a 16S rDNA

sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs: 157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:83, and SEQ ID NO: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10 and SEQ ID NOs: 157-159. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:83, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:22, SEQ ID NO:83, and SEQ ID NO: 124-156, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NOs: 157-159.

In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:106, SEQ ID NO:110, and SEQ ID NO:122, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

Table B2

Composition B2

SEQ_14 - VE202-13 - *Anaerotruncus_colihominis* (IV)

SEQ_15 - VE202-14 - *Eubacterium_fissicatena* (XIVa)

SEQ_16 - VE202-16 - *Clostridium_symbiosum* (XIVa)

SEQ_17 - VE202-7 - *Clostridium bolteae* (XIVa)

SEQ_20 - 170 - *Dorea longicatena* (XIVa)

SEQ_19 - 16 - *Blautia producta* (XIVa)

SEQ_21 - 189 - *Clostridium innocuum* (XVII)

SEQ_83 *Bacteroides ovatus*

In one aspect, the disclosure provides Composition C (See *e.g.*, Figure 1, Table C). As shown in Figure 1, Composition C contains bacteria that have the following 16S rDNA sequences: SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that have 16S rDNA sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the compositions include four or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the compositions include at least ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the composition consists of ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16, respectively. In some embodiments, the composition consists essentially of ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16, respectively.

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and

SEQ ID NO:16. In some embodiments, the compositions include four or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the compositions include at least ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments, the composition consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16, respectively. In some embodiments, the composition essentially consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16, respectively.

The bacterial strains in Composition C are related to the following species:

Clostridium scindens, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Blautia producta*, *Blautia coccoides*, *Dorea longicatena*, *Clostridium innocuum*, *Flavonifractor plautii*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, and *Clostridium symbiosum*. In some embodiments, the disclosure provides compositions with two or more bacterial strains of species selected from the group consisting of *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Blautia product*, *Blautia coccoides*, *Dorea longicatena*, *Clostridium innocuum*, *Flavonifractor plautii*, *Lachnospiraceae bacterium 7_1_58FAA*, *Subdoligranulum*, *Anaerotruncus colihominis*, and *Clostridium symbiosum*. In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, and SEQ ID NO:122.

In some embodiments, the compositions disclosed herein do not include *Flavonifractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium 7_1_58FAA*. In

some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:14, and SEQ ID NO:16, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:106, and SEQ ID NO:122, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

The strains of Composition C include both *BaiCD*⁺ strains and *BaiCD*⁻ strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein one or more bacteria are *BaiCD*⁺ strains and one or more bacteria are *BaiCD*⁻ strains. In some embodiments of the one or more bacteria that are *BaiCD*⁺ strains are selected from bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, and SEQ ID NO:7. In some embodiments the one or more bacteria that are *BaiCD*⁻ strains are selected from bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:14, and SEQ ID NO:16. In some embodiments of the one or more bacteria that are *BaiCD*⁺ strains are selected from the bacterial species *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Blautia product*, and *Blautia coccooides*. In some embodiments of the one or more bacteria that are *BaiCD*⁻ strains are selected from the bacterial species *Dorea longicatena*, *Clostridium innocuum*, *Flavonifractor plautii*, or *Lachnospiraceae bacterium 7_1_58FAA*, *Anaerotruncus colihominis*, and *Clostridium symbiosum*. The clostridial strains of Composition C are classified as belonging to Clostridium clusters IV, XIVa, and XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺ strains and belong to Clostridium clusters IV, XIVa, or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺

strains and belong to Clostridium clusters XIVa or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺ strains and belong to Clostridium clusters IV or XIVa.

Table C**Composition C**SEQ_12 - VE202-26 - *Clostridium_scindens* (XIVa)*SEQ_03 - 5 - *Clostridium_hathewayi* (XIVa)*SEQ_05 - 10 - *Blautia_hansenii* (XIVa)*SEQ_01 - 71 - *Blautia_wexlerae* (XIVa)*SEQ_07 - 59 - *Blautia_producta*/*Blautia_coccoides* (XIVa)*SEQ_18 - 148 - *Dorea_longicatena* (XIVa)**SEQ_21 - 189 - *Clostridium_innocuum* (XVII)****SEQ_10 - 211 - *Flavonifractor_plautii* (IV)****SEQ_14 - VE202-13 - *Anaerotruncus_colihominis* (IV)**SEQ_16 - VE202-16 - *Clostridium_symbiosum* (XIVa)*= *BaiCD*⁺

5

In one aspect, the disclosure provides Composition D (See *e.g.*, Figure 1, Table D). As shown in Figure 1, Composition D contains bacteria that have the following 16S rDNA sequences: SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some

10 embodiments, the disclosure provides compositions with two or more purified bacterial strains that have 16S rDNA sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some embodiments, the disclosure provides compositions that include three or more purified bacterial strains

15 comprising 16S rDNA sequences with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some embodiments, the disclosure provides compositions that include at least ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences selected from the group

20 consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some

embodiments, the disclosure provides a composition that consists of ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6, respectively. In some embodiments, the disclosure provides a composition that consists essentially of ten purified bacterial strains comprising 16S rDNA sequences with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6, respectively. In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some embodiments, the disclosure provides compositions that include three or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some embodiments, the disclosure provides compositions that include at least ten more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6. In some embodiments, the disclosure provides a composition that consists of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6, respectively. In some embodiments, the disclosure provides a composition that consists essentially of ten purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, SEQ ID NO:2, and SEQ ID NO:6, respectively.

The bacterial strains in Composition D are related to the following bacteria:

Clostridium scindens, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Anaerotruncus colihominis*, *Dorea longicatena*, *Clostridium innocuum*, *Flavonifractor*

plautii, *Lachnospiraceae* bacterium 7_1_58FAA, *Subdoligranulum*, *Turicibacter sanguinis*, and *Lactobacillus mucosae*. In some embodiments, the disclosure provides compositions with two or more bacterial strains of species selected from the group consisting of *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*,

5 *Anaerotruncus colihominis*, *Dorea longicatena*, *Clostridium innocuum*, *Erysipelotrichaceae* bacterium 21-3, *Flavonifractor plautii*, *Lachnospiraceae* bacterium 7_1_58FAA, *Turicibacter sanguinis*, and *Lactobacillus mucosae*. In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected
10 from the group consisting of SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, and SEQ ID NO:105.

In some embodiments, the compositions disclosed herein do not include *Flavonifractor plautii*, *Subdoligranulum* or *Lachnospiraceae* bacterium 7_1_58FAA. In
15 some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:1, SEQ ID NO:14, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:2, and SEQ ID NO:6, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence
20 having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10. In some embodiments, the composition comprises two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, and SEQ ID
25 NO:105, wherein the composition does not include a bacterial strain comprising a 16S rDNA sequence having at least 97% homology with a nucleic acid sequence of SEQ ID NO:93.

The strains of Composition D include both *BaiCD*⁺ strains and *BaiCD*⁻ strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein one or more bacteria are *BaiCD*⁺ strains and one or more bacteria are *BaiCD*⁻
30 strains. In some embodiments of the one or more bacteria that are *BaiCD*⁺ strains are selected from bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:12, SEQ ID NO:3, SEQ ID NO:5, and SEQ ID NO:1. In some embodiments the one or more bacteria that are *BaiCD*⁻ strains are selected from bacterial strains comprising 16S rDNA

sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:10, and SEQ ID NO:14. In some embodiments of the one or more bacteria that are *BaiCD*⁺ strains are selected from the bacterial species *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, and *Blautia wexlerae*. In some embodiments of the one or more bacteria that are *BaiCD*⁻ strains are selected from the bacterial species *Dorea longicatena*, *Clostridium innocuum*, *Flavonifractor plautii*, and *Anaerotruncus colihominis*. The Clostridial strains of Composition D are classified as belonging to Clostridium clusters IV, XIVa, and XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺ strains and belong to Clostridium clusters IV, XIVa, or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺ strains and belong to Clostridium clusters XIVa or XVII. In some embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are *BaiCD*⁻ strains and *BaiCD*⁺ strains and belong to Clostridium clusters IV or XIVa.

Composition D includes the non-Clostridium strains *Turicibacter sanguinis* and *Lactobacillus mucosae*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition includes both *Clostridium* strains and non-*Clostridium* strains. In some embodiments, the non-clostridium strains are the members of the genus *Lactobacillus*. Members of the genus *Lactobacillus* include, without limitation *L. acetotolerans*, *L. acidifarinae*, *L. acidipiscis*, *L. acidophilus*, *L. agilis*, *L. algidus*, *L. alimentarius*, *L. amylolyticus*, *L. amylophilus*, *L. amylophobicus*, *L. amylovorus*, *L. animalis*, *L. antri*, *L. apodemi*, *L. aviarius*, *L. bifementans*, *L. brevis*, *L. buchneri*, *L. camelliae*, *L. casei*, *L. catenaformis*, *L. ceti*, *L. coleohominis*, *L. collinoides*, *L. composti*, *L. concavus*, *L. coryniformis*, *L. crispatus*, *L. crustorum*, *L. curvatus*, *L. delbrueckii* subsp. *bulgaricus*, *L. delbrueckii* subsp. *delbrueckii*, *L. delbrueckii* subsp. *lactis*, *L. dextrinicus*, *L. diolivorans*, *L. equi*, *L. equigenerosi*, *L. farraginis*, *L. farciminis*, *L. fermentum*, *L. fornicalis*, *L. fructivorans*, *L. frumenti*, *L. fuchuensis*, *L. gallinarum*, *L. gasseri*, *L. gastricus*, *L. ghanensis*, *L. graminis*, *L. hammesii*, *L. hamsteri*, *L. harbinensis*, *L. hayakitensis*, *L. helveticus*, *L. hilgardii*, *L. homohiochii*, *L. iners*, *L. ingluviei*, *L. intestinalis*, *L. jensenii*, *L. johnsonii*, *L. kalixensis*, *L. kefiranoferens*, *L. kefir*, *L. kimchii*, *L. kitasatonis*, *L. kunkeei*, *L. leichmannii*, *L. lindneri*, *L. malefermentans*, *L. mali*, *L. manihotivorans*, *L. mindensis*, *L. mucosae*, *L. murinus*, *L. nagelii*, *L. namurensis*, *L. nantensis*, *L. oligofermentans*, *L. oris*, *L. panis*, *L. pantheris*, *L. parabrevis*, *L. parabuchneri*, *L. paracasei*, *L. paracollinoides*, *L.*

parafarraginis, *L. parakefiri*, *L. paralimentarius*, *L. paraplantarum*, *L. pentosus*, *L. perolens*, *L. plantarum*, *L. pontis*, *L. protectus*, *L. psittaci*, *L. rennini*, *L. reuteri*, *L. rhamnosus*, *L. rimae*, *L. rogosa*, *L. rossiae*, *L. ruminis*, *L. saerimneri*, *L. sakei*, *L. salivarius*, *L. sanfranciscensis*, *L. satsumensis*, *L. secaliphilus*, *L. sharpeae*, *L. siliginis*, *L. spicheri*, *L. suebicus*, *L. thailandensis*, *L. ultunensis*, *L. vaccinostercus*, *L. vaginalis*, *L. versmoldensis*, *L. vini*, *L. vitulinus*, *L. zaeae*, and *L. zymae*. In some embodiments, the non-clostridium strain is *Lactobacillus mucosae*. In some embodiments, the non-clostridium strain is *Lactobacillus mucosae*. In some embodiments, the *Lactobacillus mucosae* has a 16S rDNA sequence comprising SEQ ID NO:2. In some embodiments, the *Lactobacillus mucosae* has a 16S rDNA sequence having at least 97% homology with a nucleic acid comprising SEQ ID NO:2. In some embodiments, the *Lactobacillus mucosae* has a 16S rDNA sequence having at least 97% homology with a nucleic acid comprising SEQ ID NO:91.

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition includes both *Clostridium* strains and non-*Clostridium* strains. In some embodiments, the non-clostridium strains are members of the genus *Turicibacter*. In some embodiments, the non-clostridium strain is *Turicibacter sanguinis*. In some embodiments, the *Turicibacter sanguinis* has a 16S rDNA sequence comprising SEQ ID NO:6. In some embodiments, the *Turicibacter sanguinis* has a 16S rDNA sequence having at least 97% homology with a nucleic acid comprising SEQ ID NO:6. In some embodiments, the *Turicibacter sanguinis* has a 16S rDNA sequence having at least 97% homology with a nucleic acid comprising SEQ ID NO:90.

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition includes both *Clostridium* strains and non-*Clostridium* strains. In some embodiments, the non-*Clostridium* strains are *Lactobacillus mucosae* and *Turicibacter sanguinis*.

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Lactobacillus*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Turicibacter*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Lactobacillus* or *Turicibacter*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition only includes *clostridia* strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition only includes *clostridia* strains belonging to *Clostridium* cluster IV,

XIVa or XVII strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include Clostridium cluster XI strains.

In some embodiments, the disclosure provides compositions comprising two or more purified bacterial strains selected from the group consisting of: *Clostridium scindens*, *Pseudoflavonifractor capillosus*, and *Blautia hansenii*. In some embodiments, the compositions disclosed herein do not include *Clostridium scindens*, *Pseudoflavonifractor capillosus*, or *Blautia hansenii*.

Table D

Composition D

SEQ_12 - VE202-26 - *Clostridium_scindens* (XIVa)*

SEQ_03 - 5 - *Clostridium_hathewayi* (XIVa)*

SEQ_05 - 10 - *Blautia_hansenii* (XIVa)*

SEQ_01 - 71 - *Blautia_wexlerae* (XIVa)*

SEQ_14 - VE202-13 - *Anaerotruncus_colihominis* (IV)

SEQ_18 - 148 - *Dorea_longicatena* (XIVa)

SEQ_21 - 189 - *Clostridium_innocuum* (XVII)

SEQ_10 - 211 - *Flavonifractor_plautii* (IV)

SEQ_02 - 102 - *Turicibacter_sanguinis* (non-Clostridium)

SEQ_06 - 40 - *Lactobacillus_mucosae* (non-Clostridium)

*= BaiCD⁺

10

In one aspect, the disclosure provides Composition F (See *e.g.*, Figures 13 and 14, and Tables F1 and F2). As shown in Figure 13, Composition F contains bacteria that have the following 16S rDNA sequences: SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID

15

NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, and SEQ ID NO:79.

In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that have 16S rDNA sequences selected from the group consisting of SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, and SEQ ID NO:79.

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, and SEQ ID NO:79.

The bacterial strains in Composition F are related to the following bacteria: *Dorea longicatena*, *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*,

Acidaminococcus intestine, *Megasphaera elsdenii*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Ruminococcus obeum*, *Flavonifractor plautii*, *Eubacterium rectale*, *Flavonifractor plautii*, *Megasphaera elsdenii*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Ruminococcus champanellensis*,
5 *Ruminococcus faecis*, *Bifidobacterium bifidum*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus faecis*, *Blautia luti*, *Ruminococcus faecis*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Ruminococcus faecis*, *Eubacterium rectale*, *Eubacterium rectale*, *Anaerostipes hadrus*, *Ruminococcus faecis*, *Ruminococcus faecis*, *Dorea longicatena*, *Roseburia faecis*, *Blautia luti*, *Fusicatenibacter saccharivorans*,
10 *Fusicatenibacter saccharivorans*, *Roseburia faecis*, *Megasphaera elsdenii*, *Eubacterium rectale*, *Eubacterium rectale*, *Roseburia faecis*, *Blautia faecis*, *Fusicatenibacter saccharivorans*, and *Dorea formicigenerans*.

In some embodiments, the disclosure provides compositions with two or more bacterial strains of species selected from the group consisting of *Dorea longicatena*,
15 *Ruminococcus obeum*, *Megasphaera elsdenii*, *Acidaminococcus fermentans*, *Acidaminococcus intestine*, *Megasphaera elsdenii*, *Ruminococcus faecis*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Ruminococcus obeum*, *Flavonifractor plautii*, *Eubacterium rectale*, *Flavonifractor plautii*, *Megasphaera elsdenii*, *Eubacterium rectale*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Ruminococcus champanellensis*,
20 *Ruminococcus faecis*, *Bifidobacterium bifidum*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus faecis*, *Blautia luti*, *Ruminococcus faecis*, *Anaerostipes hadrus*, *Anaerostipes hadrus*, *Ruminococcus faecis*, *Eubacterium rectale*, *Eubacterium rectale*, *Anaerostipes hadrus*, *Ruminococcus faecis*, *Ruminococcus faecis*, *Dorea longicatena*, *Roseburia faecis*, *Blautia luti*, *Fusicatenibacter saccharivorans*,
25 *Fusicatenibacter saccharivorans*, *Roseburia faecis*, *Megasphaera elsdenii*, *Eubacterium rectale*, *Eubacterium rectale*, *Roseburia faecis*, *Blautia faecis*, *Fusicatenibacter saccharivorans*, and *Dorea formicigenerans*.

In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with
30 nucleic acid sequences selected from the group consisting of SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:104, SEQ ID NO:107, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, and SEQ ID NO:120. It should be appreciated that multiple

strains of the compositions disclosed herein can have the same related bacterial species. For instance, Composition F includes 12 strains that have *Eubacterium rectale* as the closest related species.

In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster XIVa. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IX. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster XIVa and at least one of the bacterial strains belongs to Clostridium cluster IX. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV and at least one of the bacterial strains belongs to Clostridium cluster IX. In some embodiments, at least one of the bacterial strains of the composition belongs to Clostridium cluster IV, at least one of the bacterial strains belongs to Clostridium cluster XIVa, and at least one of the bacterial strains belongs to Clostridium cluster IX. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster XVIII. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster XVI or XVIII.

Composition F includes non-clostridium bacterial strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition includes both Clostridium strains and non-clostridium strains. In some embodiments, the non-clostridium strains are the members of the genus *Bacteroides*. In some embodiments, the non-clostridium strain is *Bacteroides cellulosilyticus*. In some embodiments, the non-clostridium strains are the members of the genus *Bifidobacterium*. In some embodiments, the non-clostridium strain is *Bifidobacterium bifidum*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition includes both Clostridium strains and non-Clostridium strains, and wherein the non-Clostridium strains are *Bacteroides cellulosilyticus* and *Bifidobacterium bifidum*.

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Bacteroides*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Bifidobacterium*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include

Bacteroides and does not include *Bifidobacterium*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include non-*Clostridium* strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition only includes *clostridia* strains belonging to Clostridium cluster IV, XIVa or XVII strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include Clostridium cluster XI strains.

Table F1

Composition F

10

SEQ_NO	StrainID	Genus_species	SEQ_NO	StrainID	Genus_species
SEQ_24	YK96	Dorea_longicatena	SEQ_52	YK51	Eubacterium_rectale
SEQ_25	YK101	Ruminococcus_obeum	SEQ_53	YK52	Eubacterium_rectale
SEQ_26	YK110	Megasphaera_elsdenii	SEQ_54	YK54	Anaerostipes_hadrus
		Acidaminococcus_fermentans /			
SEQ_27	YK149	Acidaminococcus_intestini	SEQ_55	YK56	Ruminococcus_faecis
SEQ_28	YK154	Megasphaera_elsdenii	SEQ_56	YK57	Ruminococcus_faecis
SEQ_29	YK36	Ruminococcus_faecis	SEQ_57	YK58	Dorea_longicatena
SEQ_30	YK95	Bacteroides_cellulosilyticus	SEQ_58	YK65	Roseburia_faecis
SEQ_31	YK32	Anaerostipes_hadrus	SEQ_59	YK67	Blautia_luti
SEQ_32	YK64	Ruminococcus_obeum	SEQ_60	YK69	Fusicatenibacter_saccharivorans
SEQ_33	YK73	Flavonifractor_plautii	SEQ_61	YK70	Fusicatenibacter_saccharivorans
SEQ_34	YK87	Eubacterium_rectale	SEQ_62	YK71	Roseburia_faecis
SEQ_35	YK105	Flavonifractor_plautii	SEQ_63	YK74	Megasphaera_elsdenii
SEQ_36	YK153	Megasphaera_elsdenii	SEQ_64	YK88	Eubacterium_rectale
SEQ_37	YK163	Eubacterium_rectale	SEQ_65	YK89	Eubacterium_rectale
		Ruminococcus_champanellensis /			
SEQ_38	YK191	Ruminococcus_albus	SEQ_66	YK97	Roseburia_faecis
SEQ_39	YK99	Ruminococcus_champanellensis	SEQ_67	YK98	Blautia_faecis
SEQ_40	YK55	Ruminococcus_faecis	SEQ_68	YK139	Fusicatenibacter_saccharivorans
SEQ_41	YK75	Bifidobacterium_bifidum	SEQ_69	YK141	Dorea_formicigenerans
SEQ_42	YK90	Anaerostipes_hadrus	SEQ_70	YK142	Ruminococcus_faecis
SEQ_43	YK30	Anaerostipes_hadrus	SEQ_71	YK152	Blautia_hansenii
SEQ_44	YK31	Anaerostipes_hadrus	SEQ_72	YK155	Blautia_hansenii
SEQ_45	YK12	Eubacterium_rectale	SEQ_73	YK157	Eubacterium_rectale
SEQ_46	YK27	Ruminococcus_faecis	SEQ_74	YK160	Roseburia_faecis
SEQ_47	YK28	Blautia_luti	SEQ_75	YK166	Eubacterium_rectale
SEQ_48	YK29	Ruminococcus_faecis	SEQ_76	YK168	Eubacterium_rectale
SEQ_49	YK33	Anaerostipes_hadrus	SEQ_77	YK169	Eubacterium_rectale
SEQ_50	YK34	Anaerostipes_hadrus	SEQ_78	YK171	Eubacterium_rectale
SEQ_51	YK35	Ruminococcus_faecis	SEQ_79	YK192	Roseburia_faecis

Table F2

Composition F, strain groupings

15

Cluster	Composition F	*SCFAs
XIVa	<i>Eubacterium rectale</i> 12	A, B, L
	<i>Ruminococcus faecis</i> 8	A, L
	<i>Ruminococcus obeum</i> 2	A, L
	<i>Blautia faecis</i> 1	A, L
	<i>Blautia hansenii</i> 2	A, L
	<i>Blautia luti</i> 2	A, L
	<i>Anaerostipes hadrus</i> 7	B
	<i>Roseburia faecis</i> 5	A, B
	<i>Fusicatenibacter saccharivorans</i> 3	A, L
	<i>Dorea formicigenerans</i> 1	A
	<i>Dorea longicatena</i> 2	A
IV	<i>Flavonifractor plautii</i> 2	A, B
	<i>Ruminococcus champanellensis</i> 2	A
IX	<i>Acidaminococcus fermentans</i> 1	A, B, P
	<i>Megasphaera elsdeni</i> 4	P
other	<i>Bacteroides cellulosilyticus</i> 1	A, S
	<i>Bifiidobacterium Bifidum</i>	L, A

*Short chain fatty acid legend:

A, acetate; B, Butyrate; L, lactate;
P, propionate; S, succinate

In one aspect, the disclosure provides Composition G (See *e.g.*, Figure 19; Table G). As shown in Figure 19, Composition G contains bacteria that have the following 16S rDNA sequences: SEQ ID NO:27, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:55, SEQ ID NO:68, SEQ ID NO:72, SEQ ID NO:70, SEQ ID NO:24, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:46, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:35, SEQ ID

NO:62, SEQ ID NO:26, SEQ ID NO:63, SEQ ID NO:67, SEQ ID NO:40, SEQ ID NO:38, SEQ ID NO:47, SEQ ID NO:56, SEQ ID NO:25, and SEQ ID NO: 32.

In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that have 16S rDNA sequences selected from the group consisting of
5 SEQ ID NO:27, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:55, SEQ ID NO:68, SEQ ID NO:72, SEQ ID NO:70, SEQ ID NO:24, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:46, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:35, SEQ ID NO:62, SEQ ID NO:26, SEQ ID NO:63, SEQ ID NO:67, SEQ ID NO:40, SEQ ID NO:38, SEQ ID NO:47, SEQ ID NO:56, SEQ ID NO:25, and SEQ ID NO: 32.

10 In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:27, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:55, SEQ ID NO:68, SEQ ID NO:72, SEQ ID NO:70, SEQ ID NO:24, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:46, SEQ ID NO:76, SEQ ID NO:77, SEQ ID
15 NO:35, SEQ ID NO:62, SEQ ID NO:26, SEQ ID NO:63, SEQ ID NO:67, SEQ ID NO:40, SEQ ID NO:38, SEQ ID NO:47, SEQ ID NO:56, SEQ ID NO:25, and SEQ ID NO: 32.

The bacterial strains in Composition G are related to the following bacteria:

Acidaminococcus fermentans, *Acidaminococcus intestine*, *Anaerostipes hadrus*, *Blautia faecis*, *Blautia hansenii*, *Dorea formicigenerans*, *Dorea longicatena*, *Eubacterium rectale*,
20 *Flavonifractor plautii*, *Fusicatenibacter saccharivorans*, *Megasphaera elsdenii*, *Roseburia faecis*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Ruminococcus faecis*, and *Ruminococcus obeum*.

In some embodiments, the disclosure provides compositions with two or more bacterial strains of species selected from the group consisting of *Acidaminococcus*
25 *fermentans*, *Acidaminococcus intestine*, *Anaerostipes hadrus*, *Blautia faecis*, *Blautia hansenii*, *Dorea formicigenerans*, *Dorea longicatena*, *Eubacterium rectale*, *Flavonifractor plautii*, *Fusicatenibacter saccharivorans*, *Megasphaera elsdenii*, *Roseburia faecis*, *Ruminococcus champanellensis*, *Ruminococcus albus*, *Ruminococcus faecis*, and *Ruminococcus obeum*.

30 In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:99, SEQ ID

NO:104, SEQ ID NO:107, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, and SEQ ID NO:119.

Table G

5

Composition G

SEQ_27	YK149	Acidaminococcus_fermentans/Acidaminococcus_intesti
SEQ_43	YK90	Anaerostipes_hadrus
SEQ_44	YK30	Anaerostipes_hadrus
SEQ_51	YK34	Anaerostipes_hadrus
SEQ_55	YK54	Anaerostipes_hadrus
SEQ_68	YK98	Blautia_faecis
SEQ_72	YK152	Blautia_hansenii
SEQ_70	YK141	Dorea_formicigenerans
SEQ_24	YK96	Dorea_longicatena
SEQ_34	YK87	Eubacterium_rectale
SEQ_37	YK163	Eubacterium_rectale
SEQ_46	YK12	Eubacterium_rectale
SEQ_76	YK166	Eubacterium_rectale
SEQ_77	YK168	Eubacterium_rectale
SEQ_35	YK105	Flavonifractor_plautii
SEQ_62	YK70	Fusicatenibacter_saccharivorans
SEQ_26	YK110	Megasphaera_elsdenii
SEQ_63	YK71	Roseburia_faecis
SEQ_67	YK97	Roseburia_faecis
SEQ_40	YK99	Ruminococcus_champanellensis
SEQ_38	YK191	Ruminococcus_champanellensis/Ruminococcus_albus
SEQ_47	YK27	Ruminococcus_faecis
SEQ_56	YK56	Ruminococcus_faecis
SEQ_25	YK101	Ruminococcus_obeum
SEQ_32	YK64	Ruminococcus_obeum

In one aspect, the disclosure provides Composition H (See *e.g.*, Figure 26, Table H). As shown in Figure 26, Composition H contains bacteria that have the following 16S rDNA sequences: SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:21, SEQ ID NO:82, SEQ ID NO:81, and SEQ ID NO:80. In some embodiments, the disclosure provides compositions with two or more purified bacterial strains that have 16S rDNA sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:21, SEQ ID NO:82, SEQ ID NO:81, and SEQ ID NO:80.

15

In some embodiments, the compositions include two or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:21, SEQ

ID NO:82, SEQ ID NO:81, and SEQ ID NO:80. In some embodiments, the compositions include four or more purified bacterial strains comprising 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:21, SEQ ID NO:82, SEQ ID NO:81, and SEQ ID NO:80.

The bacterial strains in Composition H are related to the following bacteria:

Anaerotruncus colihominis, *Clostridium symbiosum*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Clostridium disporicum*, *Clostridium bolteae*, and *Erysipelatoclostridium ramosum*. In some embodiments, the disclosure provides compositions with two or more bacterial strains selected from the group consisting of *Anaerotruncus colihominis*, *Clostridium symbiosum*, *Clostridium innocuum*, *Erysipelotrichaceae_bacterium_21-3*, *Clostridium disporicum*, *Clostridium bolteae*, and *Erysipelatoclostridium ramosum*.

In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains comprise 16S rDNA sequences having at least 97% homology with nucleic acid sequences selected from the group consisting of SEQ ID NO:86, SEQ ID NO:95, SEQ ID NO:98, SEQ ID NO:110, SEQ ID NO:122, and SEQ ID NO:123.

Composition H includes bacteria from *Clostridium* cluster I, IV, XIVa, XVII and XVIII. In some embodiments, the disclosure provides compositions that include two or more purified bacterial strains from *Clostridium* cluster I, IV, XIVa, XVII and XVIII. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster IV. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XIVa. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XVII. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster I. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XVIII. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XIVa and at least one of the bacterial strains belongs to *Clostridium* cluster IV. In some embodiments, at least one of the bacterial strains of the composition belongs to *Clostridium* cluster XIVa and at least one of the bacterial strains belongs to *Clostridium* cluster XVII.

Table H

Composition H			
SEQ ID NO	Strain	Closest species	Cluster
SEQ ID NO: 14	VE202-13	<i>Anaerotruncus colihominis</i>	Cluster IV
SEQ ID NO: 16	VE202-16	<i>Clostridium symbiosum</i> WAL-14163	Cluster XIVa
SEQ ID NO: 21	189	<i>Clostridium innocuum</i>	Cluster XVII
SEQ ID NO: 82	PE9	<i>Clostridium disporicum</i>	Cluster I
SEQ ID NO: 81	PE5	<i>Clostridium bolteae</i>	Cluster XIVa
SEQ ID NO: 80	VE202-18	<i>Erysipelatoclostridium</i> <i>ramosum</i>	Cluster XVIII

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include bacteria from *Clostridium* cluster I. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include bacteria from *Clostridium* cluster XVIII. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include bacteria from *Clostridium* cluster I and does not include bacteria from *Clostridium* cluster XVIII.

In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein all the bacteria are anaerobic bacteria. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein all the bacteria are obligate anaerobic bacteria.

In some embodiments, the disclosure provides compositions comprising two or more bacteria (*e.g.*, purified bacterial strains), wherein the composition does not include *Clostridium scindens*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Flavonifractor plautii*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Parabacteroides*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Lactobacillus*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include

Colinsella. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Dialister*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Raoultella*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Streptococcus*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Staphylococcus*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Microbacterium*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Proteobacteria*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Peptostreptococcaceae*. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein the composition does not include *Oscillospiraceae*.

In one aspect, the disclosure provides bacterial strains with 16S rDNA sequences that have homology to a nucleic acid sequence of any one of the sequences of the bacterial strains or species described herein. In some embodiments, the bacterial strain has at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8% or 99.9% homology relative to any of the strains or bacterial species described herein over a specified region or over the entire sequence. It would be appreciated by one of skill in the art that the term "homology" or "percent homology," in the context of two or more nucleic acid sequences or amino acid sequences, refers to a measure of similarity between two or more sequences or portion(s) thereof. The homology may exist over a region of a sequence that is at least about 50 nucleotides in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides in length. In some embodiments, the homology exists over the length the 16S rRNA or 16S rDNA sequence, or a portion thereof.

Additionally, or alternatively, two or more sequences may be assessed for the identity between the sequences. The terms "identical" or percent "identity" in the context of two or more nucleic acids or amino acid sequences, refer to two or more sequences or subsequences that are the same. Two sequences are "substantially identical" if two sequences have a specified percentage of amino acid residues or nucleotides that are the same (*e.g.*, at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8% or 99.9% identical) over a specified region or over the entire sequence, when compared and aligned for

maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Optionally, the identity exists over a region that is at least about 50 nucleotides in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides in length. In some embodiments, the identity exists over the length the 16S rRNA or 16S rDNA sequence.

Additionally, or alternatively, two or more sequences may be assessed for the alignment between the sequences. The terms "alignment" or percent "alignment" in the context of two or more nucleic acids or amino acid sequences, refer to two or more sequences or subsequences that are the same. Two sequences are "substantially aligned" if two sequences have a specified percentage of amino acid residues or nucleotides that are the same (e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8% or 99.9% identical) over a specified region or over the entire sequence, when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Optionally, the alignment exists over a region that is at least about 50 nucleotides in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides in length. In some embodiments, the identity exists over the length the 16S rRNA or 16S rDNA sequence.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. Methods of alignment of sequences for comparison are well known in the art. See, e.g., by the local homology algorithm of Smith and Waterman (1970) *Adv. Appl. Math.* 2:482c, by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* (1970) 48:443, by the search for similarity method of Pearson and Lipman. *Proc. Natl. Acad. Sci. USA* (1998) 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, Madison, WI), or by manual alignment and visual inspection (see, e.g., Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (Ringbou ed., 2003)). Two examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., *Nuc. Acids Res.* (1977) 25:3389-3402, and Altschul et al., *J. Mol. Biol.* (1990) 215:403-410, respectively.

In one aspect, the disclosure provides compositions comprising multiple purified bacterial strains (e.g., Compositions A-J). For instance, Figures 1, 13, 19, and 26 present

several example compositions comprising multiple bacterial strains. In one aspect, the 16S rDNA sequences of purified bacterial strains of the compositions were compared to 16S rDNA sequences of known bacterial species/strains in a bacterial genome database to identify the closest known related bacterial species to the bacterial strains disclosed herein (See *e.g.*,
5 Table 1). It should be appreciated that multiple bacterial strains of the compositions disclosed herein may have the same closest related bacterial species. In one aspect, the disclosure provides compositions comprising one or more bacterial strains or species with 16S rDNA sequences that have homology to a nucleic acid sequence of any one of the sequences provided by SEQ ID NOs:1-83 and 124-159. In some embodiments, the species
10 with 16S rDNA sequences with homology to a nucleic acid sequence of any one of the closest related species to any of the strains described herein, correspond to bacterial strains with 16S rDNA sequences provided by SEQ ID NOs:84-123.

In some embodiments, the compositions disclosed herein provide at least one of the bacterial strains (*e.g.*, purified bacterial strains) described herein. In some embodiments, the
15 compositions that comprise at least one bacterial strain, comprise at least one bacterial strain with a 16S rDNA sequence selected from any one of SEQ ID NOs:1-122 and 124-159. In some embodiments, the compositions that comprise at least one bacterial strain, comprise at least one bacterial strain with a 97% homology to 16S rDNA sequence selected from any one of SEQ ID NOs:1-122 and 124-159.

20 In some embodiments, the compositions disclosed herein comprise two or more bacterial strains. In some embodiments, the compositions described herein comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or at least 20 or more bacterial strains (*e.g.*, purified bacterial strains).

25 It should be appreciated the compositions and methods provided herein can be distinguished from compositions and methods associated with the treatment of *C. difficile* infection that are available. For instance, it has been proposed that non-toxigenic *C. difficile* strains, *i.e.*, strains that do not produce *C. difficile* toxins, may be used to treat *C. difficile* infection (See, *e.g.*, US 6,635,260). The compositions disclosed herein can be distinguished
30 at least because the compositions described herein do not comprise non-toxigenic strains of *C. difficile*. Thus, in some embodiments, the compositions herein do not include comprise non-toxigenic strains of *C. difficile*. *C. difficile* belongs to Clostridium cluster XI. In some embodiments, the compositions herein do not include bacterial strains belonging to Clostridium cluster XI.

It is also considered in the art that bacterial strains expressing a bile inducible 7 α / β -dehydroxylation operon can be used in the treatment of *C. difficile* (see, *e.g.*, Buffie et al. *Nature* (2015) 517:205-208). The catalysis of bile acid 7 α dihydroxylation is mediated by a stereo-specific NAD(H)-dependent 3-oxo- Δ^4 -cholenoic acid oxidoreductase encoded by the gene *baiCD*. In some embodiments, the compositions provided herein do not mediate bile acid 7-alpha-dehydroxylation.

In contrast to the findings in the art, in some embodiments, as shown herein, combinations of bacterial strains that do not encode *baiCD* (or a homolog thereof), or encode a *baiCD* that comprises one or more mutations that result in a non-functional BaiCD protein (“baiCD-”), are more effective at treating *C. difficile* infection and/or reducing or inhibiting production of Toxin B by *C. difficile* than combinations of bacterial strains that have a functional BaiCD protein (“baiCD+”). Thus, in some embodiments, the compositions of bacterial strains provided herein are baiCD- (*i.e.*, the combination of the bacteria has no effective baiCD+ function). In some embodiments, all of bacterial strains in the compositions provided herein are baiCD-. In some embodiments, the majority (*i.e.*, 50% or greater) of the bacterial strains in the compositions are baiCD-. In some embodiments, the majority (*i.e.*, 50% or greater) of the bacterial strains in the compositions are baiCD- and the composition has no effective BaiCD function. In some embodiments, the minority (*i.e.*, 50% or less) of the bacterial strains in the compositions are baiCD- and the composition has no effective BaiCD function. In some embodiments, bacterial strains for the compositions are selected based on the absence (or presence) of a *baiCD* gene or a predicted *baiCD* gene. In some embodiments, bacterial strains may be modified (*e.g.*, genetically engineered) to prevent or reduce expression of a *baiCD* gene and/or to reduce or eliminate NAD(H)-dependent 3-oxo- Δ^4 -cholenoic acid oxidoreductase activity of BaiCD protein. The NAD(H)-dependent 3-oxo- Δ^4 -cholenoic acid oxidoreductase activity of a bacterial strain may be assessed by methods such as measuring the amount of 7 α -dehydroxylated bile acid. In some embodiments, the compositions described herein comprise bacterial strains without the *baiCD* operon (*baiCD*⁻) or *baiCD* function.

In some embodiments, the compositions described herein do not include *Clostridium scindens*. In some embodiments, the compositions described herein do not include *Barnesiella intestihominis*. In some embodiments, the compositions described herein do not include *Blautia hansenii*. In some embodiments, the compositions described herein do not include *Pseudoflavinoxylator capillosus*. In some embodiments, the compositions described

herein do not include *Clostridium scindens*, *Barnesiella intestihominis*, *Blautia hansenii* or *Pseudoflavinofractor capillosus*.

In some embodiments, the compositions provided herein do not include *Colinsella aerofaciens*. In some embodiments, the compositions provided herein do not include
5 *Acetovibrio ethanolgignens*. In some embodiments, the compositions provided herein do not include bacterial strains belonging to Clostridium cluster I. In some embodiments, the compositions provided herein do not include *Clostridium butyricum*. In some embodiments, the compositions provided herein do not include *Clostridium disporicum*. In some embodiments, the compositions provided herein do not include strains belonging to Clostridium cluster XI.
10 In some embodiments, the compositions provided herein do not include *Clostridium glycolicum*. In some embodiments, the compositions provided herein do not include *Faecalibacterium prausnitzii*. In some embodiments, the compositions provided herein do not include *Turicibacter sanguinis*. In some embodiments, the compositions provided herein do not include *Eubacterium rectale*. In some embodiments, the compositions provided herein
15 do not include *Eubacterium ventriosum*. In some embodiments, the compositions provided herein do not include *Ruminococcus obeum*. In some embodiments, the compositions provided herein do not include *Pseudobutyrvibrio*. In some embodiments, the compositions provided herein do not include *Christensenellaceae*. In some embodiments, the compositions do not comprise gram-negative bacteria. In some embodiments, the
20 compositions do not comprise *E. coli*. In some embodiments, the compositions do not comprise fungi, such as *Monilla* species.

In some embodiments of the compositions provided herein, the compositions do not include bacterial strains that are resistant to one or more antibiotics. It should be appreciated that it may be desirable to have a mechanism to remove the bacterial compositions provided
25 herein from the body of the subject after administration. One such mechanism is to remove the bacterial compositions by antibiotic treatment. Thus, in some embodiments, the compositions do not include bacterial strains that are resistant to one or more antibiotics. In some embodiments, the compositions do not include bacterial strains that are resistant to one or more antibiotics selected from the group consisting of penicillin, benzylpenicillin,
30 ampicillin, sulbactam, amoxicillin, clavulanate, tazobactam, piperacillin, cefmetazole, vancomycin, imipenem, meropenem, metronidazole and clindamycin. In some embodiments, the compositions do not include bacterial strains that are resistant to vancomycin.

In some embodiments, the compositions include bacterial strains that are susceptible to at least four antibiotics that are efficacious in humans. In some embodiments, the

compositions include bacterial strains that are susceptible to at least three antibiotics that are efficacious in humans. In some embodiments, the compositions include bacterial strains that are susceptible to at least two antibiotics that are efficacious in humans. In some
5 embodiments, the compositions include bacterial strains that are susceptible to at least one antibiotic that is efficacious in humans. In some embodiments, the compositions include only bacterial strains that are susceptible to at least four antibiotics that are efficacious in humans. In some embodiments, the compositions include only bacterial strains that are susceptible to at least three antibiotics that are efficacious in humans. In some embodiments, the
10 compositions include only bacterial strains that are susceptible to at least two antibiotics that are efficacious in humans. In some embodiments, the compositions include bacterial strains that are susceptible to at least one antibiotic that is efficacious in humans. As used herein, an “antibiotic that is efficacious in a human” refers to an antibiotic that has been used to successfully treat bacterial infections in a human.

In some embodiments, the compositions described herein comprise spore forming and
15 non-spore forming bacterial strains. In some embodiments, the compositions described herein comprise spore forming bacterial strains. In some embodiments, the compositions described herein comprise only spore forming bacterial strains. In some embodiments, the compositions described herein comprise only non-spore forming bacterial strains. The spore-forming bacteria can be in spore form (*i.e.*, as spores) or in vegetative form (*i.e.*, as
20 vegetative cells). In spore form, bacteria are generally more resistant to environmental conditions, such as heat, acid, radiation, oxygen, chemicals, and antibiotics. In contrast, in the vegetative state or actively growing state, bacteria are more susceptible to such environmental conditions, compared to in the spore form. In general, bacterial spores are able to germinate from the spore form into a vegetative/actively growing state, under
25 appropriate conditions. For instance, bacteria in spore format may germinate when they are introduced in the intestine.

In some embodiments, at least one (*e.g.*, 1, 2, 3, 4, 5, or more) of the bacterial strains in the composition is a spore former. In some embodiments, at least one (*e.g.*, 1, 2, 3, 4, 5, or more) of the bacterial strains in the composition is in spore form. In some embodiments, at
30 least one (*e.g.*, 1, 2, 3, 4, 5, or more) of the bacterial strains in the composition is a non-spore former. In some embodiments, at least one (*e.g.*, 1, 2, 3, 4, 5, or more) of the bacterial strains in the composition is in vegetative form (As discussed above, spore forming bacteria can also be in vegetative form). In some embodiments, at least one (*e.g.*, 1, 2, 3, 4, 5, or more) of the bacterial strains in the composition is in spore form and at least one (*e.g.*, 1, 2, 3, 4, 5, or

more) of the bacterial strains in the composition is in vegetative form. In some embodiments, at least one bacterial strain that is considered able to form spores (*i.e.*, a spore-former) but is present in the composition in vegetative form. In some embodiments, at least one bacterial strain that is considered able to form spores is present in the composition both in spore form and in vegetative form.

In some embodiments, the disclosure provides compositions wherein the compositions comprise bacterial strains that are spore forming bacterial strains. In some embodiments, the disclosure provides compositions wherein the compositions comprise bacterial strains that are non-spore forming bacterial strains. In some embodiments, the disclosure provides compositions wherein the compositions comprise bacterial strains that are spore forming bacterial strains and bacterial strains that are non-spore forming bacterial strains. In some embodiments, the disclosure provides compositions, wherein the compositions comprise a mixture of bacterial strains wherein at least 10% of the bacterial strains are spore forming bacterial strains, at least 20% of the bacterial strains are spore forming bacterial strains, at least 30% of the bacterial strains are spore forming bacterial strains, at least 40% of the bacterial strains are spore forming bacterial strains, at least 50% of the bacterial strains are spore forming bacterial strains, at least 60% of the bacterial strains are spore forming bacterial strains, at least 70% of the bacterial strains are spore forming bacterial strains, at least 80% of the bacterial strains are spore forming bacterial strains, at least 90% of the bacterial strains are spore forming bacterial strains bacteria up to 100% spore forming bacterial strains. Whether a bacterial strain is a spore forming strain can be determined for instance by evaluating the genome of the bacterial strain for the presence of sporulation genes. However, it should be appreciated that not all bacteria that are predicted to encode spore forming genes can be made to sporulate. In addition, whether a bacterial strain is a spore forming strain can be determined by exposing the bacterial strain to stress conditions, *e.g.*, heat or exposure to chemicals (*e.g.*, ethanol or chloroform), that are known to induce sporulation.

It should be appreciated that spore forming bacteria can be in spore form or in vegetative form. In some embodiments of the compositions provided herein, the spore forming bacteria are in spore form. In some embodiments of the compositions provided herein, the spore forming bacteria are in vegetative form. In some embodiments of the compositions provided herein, the spore forming bacteria are both present in spore form and in vegetative form. In some embodiments, the disclosure provides compositions, wherein the compositions comprise spore forming bacteria at least 10% of the spore forming bacteria are

in spore format, at least 20% of the spore forming bacteria are in spore format, at least 30% of the spore forming bacteria are in spore format, at least 40% of the spore forming bacteria are in spore format, at least 50% of the spore forming bacteria are in spore format, at least 60% of the spore forming bacteria are in spore format, at least 70% of the spore forming bacteria are in spore format, at least 80% of the spore forming bacteria are in spore format, at least 90% of the spore forming bacteria are in spore format, up to 100% in spore format.

It is envisioned that the bacterial strains of the compositions provided herein are alive and will be alive when they reach the target area (*e.g.*, the intestines). Bacterial spores are considered to be alive in this regards. In some embodiments, bacteria that are administered as spores may germinate in the target area (*e.g.*, the intestines). It should further be appreciated that not all of the bacteria are alive and the compositions can include a percentage (*e.g.*, by weight) that is not alive. In addition, in some embodiments, the compositions include bacterial strains that are not alive when administered or at the time when the composition reaches the target area (*e.g.*, the intestines). It is envisioned that non-living bacteria may still be useful by providing some nutrients and metabolites for the other bacterial strains in the composition.

Methods of inducing sporulation of spore-forming bacterial strains are well known in the art (See *e.g.*, Paredes-Sabja *et al.*, *Trends Microbiol.* (2011) 19(2):85-94). Generally, bacterial strains that are spore-formers can be made to go into spore form by stressing the bacterial strains. Non-limiting examples of stresses that can induce sporulation are an increase in temperature, change in the nutrients available and/or exposure to chemicals (*e.g.*, ethanol or chloroform). It should be noted that bacteria that are non-spore formers, for instance because they are missing sporulation genes, cannot be made to sporulate by stress. To prepare compositions in which all the bacterial strains are in the spore form, the composition or bacterial cultures used to prepare the composition may be subjected to treatment to kill any bacteria not in spore form (*e.g.*, in vegetative form), for example by exposing the composition to heat and are chemically breaking down the non-spore bacteria. The bacteria in spore format can subsequently be separated from the non-spore bacteria for instance by filtration.

The amount of spores can be quantified using techniques know in the art. These techniques include phase contrast microscopy for enumerating spores using a hemocytometer. In addition, the viability of spores can be determined by plating the spores and growing the spores. For instance, spores can be plated in appropriate media and incubated in the anaerobic chamber for a period of time (*e.g.*, 48-96 hrs.). Viability can

subsequently be determined by quantifying the colony forming units which correspond to spores that germinated. For instance, spores can be plated on TCCFA plates (Taurocholate, cycloserine, cefoxitin, fructose agar plates), in which taurocholate helps the spores to germinate. In addition, spores can be quantified using the dipicolinic assay (DPA assay).

5 DPA is an agent that allows for spore selection and is a clear indicator of endospores. When complexed with terbium, bright green luminescence is observed.

In any of the compositions provided herein, in some embodiments, the bacterial strains are purified. In any of the compositions provided herein, in some embodiments, the bacterial strains are isolated. Any of the bacterial strains described herein may be isolated
10 and/or purified, for example, from a source such as a culture or a microbiota sample (*e.g.*, fecal matter). The bacterial strains used in the compositions provided herein generally are isolated from the microbiome of healthy individuals. However, bacterial strains can also be isolated from individuals that are considered not to be healthy. In some embodiments, the compositions include strains originating from multiple individuals.

15 As used herein, the term “isolated” bacteria that have been separated from one or more undesired component, such as another bacterium or bacterial strain, one or more component of a growth medium, and/or one or more component of a sample, such as a fecal sample. In some embodiments, the bacteria are substantially isolated from a source such that other components of the source are not detected.

20 As also used herein, the term “purified” refers to a bacterial strain or composition comprising such that has been separated from one or more components, such as contaminants. In some embodiments, the bacterial strain is substantially free of contaminants. In some embodiments, one or more bacterial strains of a composition may be independently purified from one or more other bacteria produced and/or present in a culture
25 or a sample containing the bacterial strain. In some embodiments, a bacterial strain is isolated or purified from a sample and then cultured under the appropriate conditions for bacterial replication, *e.g.*, under anaerobic culture conditions. The bacteria that is grown under appropriate conditions for bacterial replication can subsequently be isolated/purified from the culture in which it is grown.

30 In some embodiments, the bacterial strains of the compositions provided herein are obligate anaerobes. In some embodiments, the bacterial strains of the compositions provided herein are facultative anaerobes.

Aspects of the present disclosure are related to methods for treating a pathogenic infection in a subject by administering a therapeutically effective amount of any of the

compositions described herein. In some embodiments, the subject is a mammalian subject, such as a human, non-human primate, rodent, rabbit, sheep, pig, dog, cat, horse, or cow. In some embodiments, the subject is a human subject. In some embodiments, the subject is a pig.

5 In some embodiments, the subject is a carrier of a pathogenic organism and is suffering from the effects of the infection (*e.g.*, diarrhea caused by *C. difficile* toxins). In some embodiments the subject is an asymptomatic carrier of a pathogen. In some embodiments, the subject is a carrier of *C. difficile*. In some embodiments the subject is an asymptomatic *C. difficile* carrier. In some embodiments, the subject has experienced
10 recurrent or chronic pathogenic infections. In some embodiments, the subject is suffering from a first occurrence of a particular pathogenic infection. In some embodiments, the subject has been treated with antibiotics which resulted in the recurrence of the pathogenic infection. In some embodiments, the subject has been treated with antibiotics which resulted in a first occurrence of a pathogenic infection. In some embodiments, the subject is to
15 undergo a procedure that puts the subject at a higher risk of infection. In some embodiments, the compositions provided herein are administered to a subject to lower the risk of becoming infected by a pathogen.

In some embodiments, the compositions provided herein are administered to a subject if the subject has a dysbiosis (*e.g.*, has as microbiome associated with a disease state). In
20 some embodiments, treatment with the compositions provided herein results in the change in the microbiome of the subject. In some embodiments, treatment with the compositions provided herein removes the dysbiosis in the subject resulting in a healthy microbiome. In some embodiments, treatment with the compositions provided herein removes the dysbiosis in the subject resulting in microbiome refractory or less susceptible to infection by a
25 pathogen.

As used herein, the term “pathogen” in regard to a pathogenic infection refers to a microorganism (*e.g.*, a bacterium) that causes a disease or a disease state in a subject. In some embodiments, the disease or disease state of the subject may include symptoms such as colitis, diarrhea, watery diarrhea, abdominal cramping, fever, blood or pus in the stool,
30 nausea, dehydration, loss of appetite, chills, weight loss, and/or kidney failure. In some embodiments, the pathogenic infection may be diagnosed, for example, by detecting a pathogen (or protein or nucleic acid associated with a pathogen) in a fecal sample collected from the subject. In some embodiments, the pathogenic infection may be diagnosed, for

example, by comparing the microbiota of a fecal sample of the subject with the microbiota in a fecal sample of a healthy subject.

- In some embodiments, the pathogenic infection is *C. difficile*; *Clostridium perfringens*; *Clostridium botulinum*; *Clostridium tributrycum*; *Clostridium sporogenes*;
 5 *Escherichia coli*; *Pseudomonas aeruginosa*, such as Multidrug Resistant *Pseudomonas aeruginosa*; Vancomycin Resistant *Enterococci* (VRE); Carbapenem Resistant *Enterobacteriaceae* (CRE); *Neisseria gonorrhoeae*; *Acinetobacter*; Multidrug Resistant *Acinetobacter*; *Campylobacter*; Multi-drug resistant *Campylobacter*; *Candida*; Fluconazole-resistant *Candida*; Extended spectrum beta-lactamase (ESBL) producing *Enterobacteriaceae*;
 10 *Salmonella*, *Salmonella Typhimurium*, Drug resistant non-typhoid *Salmonella spp.*; Drug resistant *Salmonella Typhi*; Drug resistant *Shigella*; *Staphylococcus aureus*, such as Methicillin Resistant *S. aureus* or vancomycin resistant *S. aureus*; Drug resistant *Streptococcus pneumoniae*; Drug resistant Tuberculosis; Erythromycin Resistant Group A *Streptococcus*; Clindamycin resistant Group B *Streptococcus*, and any combinations thereof.
 15 In some embodiments, the pathogenic infection is *C. difficile*. In some embodiments, the *C. difficile* is an antibiotic-resistant *C. difficile*, e.g., fluoroquinolone resistant *C. difficile*. In some embodiments, the pathogenic infection is vancomycin-resistant *Enterococci*.

- Additional non-limiting examples of pathogens responsible for pathogenic infection that can be treated according to the methods provided herein are *Leishmania*, *Staphylococcus*
 20 *epidermis*, *Staphylococcus saprophyticus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Enterococcus faecalis*, *Corynebacterium diphtheriae*, *Bacillus anthracis*, *Listeria monocytogenes*, *Clostridium perfringens*, *Clostridium tetanus*, *Clostridium botulinum*, *Clostridium difficile*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Escherichia coli*, *Salmonella typhimurium*, *Salmonella choleraesuis*, *Salmonella enterica*,
 25 *Salmonella enteritidis*, *Yersinia pestis*, *Yersinia pseudotuberculosis*, *Yersinia enterocolitica*, *Vibrio cholerae*, *Campylobacter jejuni*, *Campylobacter fetus*, *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Pseudomonas mallei*, *Haemophilus influenzae*, *Bordetella pertussis*, *Mycoplasma pneumoniae*, *Ureaplasma urealyticum*, *Legionella pneumophila*, *Treponema pallidum*, *Leptospira interrogans*, *Borrelia burgdorferi*, *Mycobacterium*
 30 *tuberculosis*, *Mycobacterium leprae*, *Chlamydia psittaci*, *Chlamydia trachomatis*, *Chlamydia pneumoniae*, *Rickettsia rickettsii*, *Rickettsia akari*, *Rickettsia prowazekii*, *Brucella abortus*, *Brucella melitens*, *Brucella suis*, and *Francisella tularensis*. In general, any bacterium that is capable of inducing a disease in a subject and/or that is not present in healthy individual is

considered a pathogen herein. It should be appreciated that a subject may carry multiple pathogens and/or have multiple pathogenic infections.

Any of the compositions described herein may be administered to a subject in a therapeutically effective amount or a dose of a therapeutically effective amount to treat or prevent a pathogenic infection (*e.g.*, one or more pathogenic infections). The terms “treat” or “treatment” refer to reducing or alleviating one or more of the symptoms associated with a pathogenic infection, reducing the amount of bacterial toxin produced by the pathogenic infection, and/or reducing the bacterial load of the pathogenic infection. The terms “prevent” or “prevention” encompass prophylactic administration and may reduce the incidence or likelihood of pathogenic infection or a recurrent or chronic pathogenic infection. For instance, in some embodiments, administration of the compositions provided herein result in a healthy microbiome that is refractory to pathogenic infection, thereby preventing the pathogenic infection.

As used herein, a “therapeutically effective amount” of composition, such as a pharmaceutical composition, is any amount that results in a desired response or outcome in a subject, such as those described herein, including but not limited to prevention of infection, an immune response or an enhanced immune response to the pathogenic infection, prevention or reduction of symptoms associated with pathogenic infection, and/or a reduction or inhibition of toxin production by the pathogenic infection. It should be appreciated that the term effective amount may be expressed in number of bacteria or bacterial spores to be administered. It should further be appreciated that the bacteria can multiply once administered. Thus, administration of even a relatively small amount of bacteria may have therapeutic effects.

In some embodiments, the therapeutically effective amount of any of the compositions described herein is an amount sufficient to enhance survival of the subject, reduce the bacterial burden of the pathogenic infection in the subject, and/or reduce or inhibit toxin production by the pathogenic infection. In some embodiments, the therapeutically effective amount is an amount sufficient to reduce the bacterial burden of the pathogenic infection in a fecal sample from the subject by at least 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 100-fold, 1000-fold, 10^4 -fold, 10^5 -fold or more, as compared to the bacterial burden in a subject with a pathogenic infection that has not received any of the compositions described herein, or as compared to a fecal sample from the same subject that was collected prior to administration of any of the compositions.

In some embodiments, the compositions provided herein inhibit the production of a bacterial toxin, *e.g.*, *C. difficile* Toxin B. In some embodiments, the therapeutically effective amount is an amount sufficient to reduce or inhibit the amount of bacterial toxin (*e.g.*, *C. difficile* Toxin B) produced by pathogenic infection in a fecal sample from the subject by at least 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 100-fold, 150-fold, 200-fold, 500-fold or more, as compared to the amount of the bacterial toxin in a subject with a pathogenic infection that has not received any of the compositions described herein or as compared to a fecal sample from the same subject that was collected prior to administration of any of the compositions.

In some embodiments, the compositions provided herein induce the proliferation and/or accumulation of regulatory T cells in the subject. As will be evident to one of ordinary skill in the art, regulatory T cells, also referred to as "Tregs," are a subset of T lymphocytes that are generally thought to suppress an abnormal or excessive immune response and play a role in immune tolerance. Regulatory T cells may be identified based expression of the markers Foxp3 and CD4 (Foxp3+ CD4+). The term regulatory T cells may also include Foxp3-negative regulatory T cells that are IL-10-producing CD4-positive T cells.

In some embodiments, the therapeutically effective amount is an amount sufficient to induce the proliferation and/or accumulation of Tregs in the subject (or in a sample obtained from a subject) by at least 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 100-fold, 150-fold, 200-fold, 500-fold or more, as compared to the amount of Tregs in a subject (*e.g.*, a subject with a pathogenic infection) that has not received any of the compositions described herein or as compared to a fecal sample from the same subject that was collected prior to administration of any of the compositions.

As used herein, the phrase "induces proliferation and/or accumulation of regulatory T cells" refers to an effect of inducing the differentiation of immature T cells into regulatory T cells, which differentiation leads to the proliferation and/or the accumulation of regulatory T cells. Further, the meaning of "induces proliferation and/or accumulation of regulatory T cells" includes *in vivo* effects, *in vitro* effects, and *ex vivo* effects. In some embodiments, the proliferation and/or accumulation of regulatory T cells may be assessed by detecting and/or quantifying the number of cells that express markers of regulatory T cells (*e.g.*, Foxp3 and CD4), for example by flow cytometry. In some embodiments, the proliferation and/or accumulation of regulatory T cells may be assessed by determining the activity of the regulatory T cells, such as the production of cytokines (*e.g.*, IL-10).

In some embodiments, the therapeutically effective amount is an amount sufficient to recolonize or repopulate the gastrointestinal tract of the subject with non-pathogenic bacteria. In some embodiments, the therapeutically effective amount is an amount sufficient to graft one or more of the bacterial strains of the composition in the gastrointestinal tract of the subject. In some embodiments, a fecal sample is obtained from the subject to assess the bacterial burden of the pathogenic infection and/or evaluate the efficacy of administration of the bacterial compositions described herein. In some embodiments, the microbiota of the subject (*e.g.*, the identity and abundance of strains and/or species of the microbiota) may be assessed to determine a disease state of the subject and/or assess progress of the treatment. In some embodiments, the microbiota of the subject having a pathogenic infection is compared to the microbiota of a healthy subject, such as a subject that is not experiencing or has not experienced the pathogenic infection. In some embodiments, the microbiota of the subject having a pathogenic infection is compared to the microbiota of the same subject from a fecal sample obtained from the subject prior to the pathogenic infection.

Any of the compositions described herein, including the pharmaceutical compositions and food products comprising the compositions, may contain bacterial strains in any form, for example in an aqueous form, such as a solution or a suspension, embedded in a semi-solid form, in a powdered form or freeze dried form. In some embodiments, the composition or the bacterial strains of the composition are lyophilized. In some embodiments, a subset of the bacterial strains in a composition is lyophilized. Methods of lyophilizing compositions, specifically compositions comprising bacteria, are well known in the art. See, *e.g.*, US 3,261,761; US 4,205, 132; PCT Publications WO 2014/029578 and WO 2012/098358, herein incorporated by reference in their entirety. The bacteria may be lyophilized as a combination and/or the bacteria may be lyophilized separately and combined prior to administration. A bacterial strain may be combined with a pharmaceutical excipient prior to combining it with the other bacterial strain or multiple lyophilized bacteria may be combined while in lyophilized form and the mixture of bacteria, once combined may be subsequently be combined with a pharmaceutical excipient. In some embodiments, the bacterial strain is a lyophilized cake. In some embodiments, the compositions comprising the one or more bacterial strains are a lyophilized cake.

The bacterial strains of the composition can be manufactured using fermentation techniques well known in the art. In some embodiments, the active ingredients are manufactured using anaerobic fermenters, which can support the rapid growth of anaerobic bacterial species. The anaerobic fermenters may be, for example, stirred tank reactors or

disposable wave bioreactors. Culture media such as BL media and EG media, or similar versions of these media devoid of animal components, can be used to support the growth of the bacterial species. The bacterial product can be purified and concentrated from the fermentation broth by traditional techniques, such as centrifugation and filtration, and can optionally be dried and lyophilized by techniques well known in the art.

In some embodiments, the composition of bacterial strains may be formulated for administration as a pharmaceutical composition. The term “pharmaceutical composition” as used herein means a product that results from the mixing or combining of at least one active ingredient, such as any two or more purified bacterial strains described herein, and one or more inactive ingredients, which may include one or more pharmaceutically acceptable excipient.

An “acceptable” excipient refers to an excipient that must be compatible with the active ingredient and not deleterious to the subject to which it is administered. In some embodiments, the pharmaceutically acceptable excipient is selected based on the intended route of administration of the composition, for example a composition for oral or nasal administration may comprise a different pharmaceutically acceptable excipient than a composition for rectal administration. Examples of excipients include sterile water, physiological saline, solvent, a base material, an emulsifier, a suspending agent, a surfactant, a stabilizer, a flavoring agent, an aromatic, an excipient, a vehicle, a preservative, a binder, a diluent, a tonicity adjusting agent, a soothing agent, a bulking agent, a disintegrating agent, a buffer agent, a coating agent, a lubricant, a colorant, a sweetener, a thickening agent, and a solubilizer.

Pharmaceutical compositions of the invention can be prepared in accordance with methods well known and routinely practiced in the art (see *e.g.*, Remington: The Science and Practice of Pharmacy, Mack Publishing Co. 20th ed. 2000). The pharmaceutical compositions described herein may further comprise any carriers or stabilizers in the form of a lyophilized formulation or an aqueous solution. Acceptable excipients, carriers, or stabilizers may include, for example, buffers, antioxidants, preservatives, polymers, chelating reagents, and/or surfactants. Pharmaceutical compositions are preferably manufactured under GMP conditions. The pharmaceutical compositions can be used orally, nasally or parenterally, for instance, in the form of capsules, tablets, pills, sachets, liquids, powders, granules, fine granules, film-coated preparations, pellets, troches, sublingual preparations, chewables, buccal preparations, pastes, syrups, suspensions, elixirs, emulsions, liniments,

ointments, plasters, cataplasms, transdermal absorption systems, lotions, inhalations, aerosols, injections, suppositories, and the like.

In some embodiments, the bacteria are formulated for delivery to the intestines (*e.g.*, the small intestine and/or the colon). In some embodiments, the bacteria are formulated with an enteric coating that increases the survival of the bacteria through the harsh environment in the stomach. The enteric coating is one which resists the action of gastric juices in the stomach so that the bacteria which are incorporated therein will pass through the stomach and into the intestines. The enteric coating may readily dissolve when in contact with intestinal fluids, so that the bacteria enclosed in the coating will be released in the intestinal tract. Enteric coatings may consist of polymer and copolymers well known in the art, such as commercially available EUDRAGIT (Evonik Industries). (See *e.g.*, Zhang, *AAPS PharmSciTech*, (2016) 17 (1), 56-67).

The bacteria may also be formulated for rectal delivery to the intestine (*e.g.*, the colon). Thus, in some embodiments, the bacterial compositions may be formulated for delivery by suppository, colonoscopy, endoscopy, sigmoidoscopy or enema. A pharmaceutical preparation or formulation and particularly a pharmaceutical preparation for oral administration, may include an additional component that enables efficient delivery of the compositions of the disclosure to the intestine (*e.g.*, the colon). A variety of pharmaceutical preparations that allow for the delivery of the compositions to the intestine (*e.g.*, the colon) can be used. Examples thereof include pH sensitive compositions, more specifically, buffered sachet formulations or enteric polymers that release their contents when the pH becomes alkaline after the enteric polymers pass through the stomach. When a pH sensitive composition is used for formulating the pharmaceutical preparation, the pH sensitive composition is preferably a polymer whose pH threshold of the decomposition of the composition is between about 6.8 and about 7.5. Such a numeric value range is a range in which the pH shifts toward the alkaline side at a distal portion of the stomach, and hence is a suitable range for use in the delivery to the colon. It should further be appreciated that each part of the intestine (*e.g.*, the duodenum, jejunum, ileum, cecum, colon and rectum), has different biochemical and chemical environment. For instance, parts of the intestines have different pHs, allowing for targeted delivery by compositions that have a specific pH sensitivity. Thus, the compositions provided herein may be formulated for delivery to the intestine or specific parts of the intestine (*e.g.*, the duodenum, jejunum, ileum, cecum, colon and rectum) by providing formulations with the appropriate pH sensitivity. (See *e.g.*, Villena et al., *Int J Pharm* 2015, 487 (1-2): 314-9).

Another embodiment of a pharmaceutical preparation useful for delivery of the compositions to the intestine (*e.g.*, the colon) is one that ensures the delivery to the colon by delaying the release of the contents (*e.g.*, the bacterial strains) by approximately 3 to 5 hours, which corresponds to the small intestinal transit time. In one embodiment of a

5 pharmaceutical preparation for delayed release, a hydrogel is used as a shell. The hydrogel is hydrated and swells upon contact with gastrointestinal fluid, with the result that the contents are effectively released (released predominantly in the colon). Delayed release dosage units include drug-containing compositions having a material which coats or selectively coats a drug or active ingredient to be administered. Examples of such a selective coating material
10 include in vivo degradable polymers, gradually hydrolyzable polymers, gradually water-soluble polymers, and/or enzyme degradable polymers. A wide variety of coating materials for efficiently delaying the release is available and includes, for example, cellulose-based polymers such as hydroxypropyl cellulose, acrylic acid polymers and copolymers such as methacrylic acid polymers and copolymers, and vinyl polymers and copolymers such as
15 polyvinylpyrrolidone.

Additional examples of pharmaceutical compositions that allow for the delivery to the intestine (*e.g.*, the colon) include bioadhesive compositions which specifically adhere to the colonic mucosal membrane (for example, a polymer described in the specification of US Patent No. 6,368,586) and compositions into which a protease inhibitor is incorporated for
20 protecting particularly a biopharmaceutical preparation in the gastrointestinal tracts from decomposition due to an activity of a protease.

Another example of a system enabling the delivery to the intestine (*e.g.*, the colon) is a system of delivering a composition to the colon by pressure change in such a way that the contents are released by utilizing pressure change caused by generation of gas in bacterial
25 fermentation at a distal portion of the stomach. Such a system is not particularly limited, and a more specific example thereof is a capsule which has contents dispersed in a suppository base and which is coated with a hydrophobic polymer (for example, ethyl cellulose).

A further example of a system enabling the delivery of a composition to the intestine (*e.g.*, the colon), is a composition that includes a coating that can be removed by an enzyme
30 present in the gut (*e.g.*, the colon), such as, for example, a carbohydrate hydrolase or a carbohydrate reductase. Such a system is not particularly limited, and more specific examples thereof include systems which use food components such as non-starch polysaccharides, amylose, xanthan gum, and azopolymers.

The compositions provided herein can also be delivered to specific target areas, such as the intestine, by delivery through an orifice (*e.g.*, a nasal tube) or through surgery. In addition, the compositions provided herein that are formulated for delivery to a specific area (*e.g.*, the cecum or the colon), may be administered by a tube (*e.g.*, directly into the small intestine). Combining mechanical delivery methods such as tubes with chemical delivery methods such as pH specific coatings, allow for the delivery of the compositions provided herein to a desired target area (*e.g.*, the cecum or the colon).

The compositions comprising bacterial strains are formulated into pharmaceutically acceptable dosage forms by conventional methods known to those of skill in the art. Dosage regimens are adjusted to provide the optimum desired response (*e.g.*, the prophylactic or therapeutic effect). In some embodiments, the dosage form of the composition is a tablet, pill, capsule, powder, granules, solution, or suppository. In some embodiments, the pharmaceutical composition is formulated for oral administration. In some embodiments, the pharmaceutical composition is formulated such that the bacteria of the composition, or a portion thereof, remain viable after passage through the stomach of the subject. In some embodiments, the pharmaceutical composition is formulated for rectal administration, *e.g.* as a suppository. In some embodiments, the pharmaceutical composition is formulated for delivery to the intestine or a specific area of the intestine (*e.g.*, the colon) by providing an appropriate coating (*e.g.*, a pH specific coating, a coating that can be degraded by target area specific enzymes, or a coating that can bind to receptors that are present in a target area).

Dosages of the active ingredients in the pharmaceutical compositions of the present invention can be varied so as to obtain an amount of the active ingredient which is effective to achieve the desired pharmaceutical response for a particular subject, composition, and mode of administration, without being toxic or having an adverse effect on the subject. The selected dosage level depends upon a variety of factors including the activity of the particular compositions of the present invention employed, the route of administration, the time of administration, the duration of the treatment, other drugs, compounds and/or materials used in combination with the particular compositions employed, the age, sex, weight, condition, general health and prior medical history of the subject being treated, and like factors.

A physician, veterinarian or other trained practitioner, can start doses of the pharmaceutical composition at levels lower than that required to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect (*e.g.*, treatment of a pathogenic infection, reduction of bacterial burden of pathogenic infection, reduction or inhibition of toxin production) is achieved. In general, effective doses of the compositions of

the present invention, for the prophylactic treatment of groups of people as described herein vary depending upon many different factors, including routes of administration, physiological state of the subject, whether the subject is human or an animal, other medications administered, and the therapeutic effect desired. Dosages need to be titrated to optimize safety and efficacy. In some embodiments, the dosing regimen entails oral administration of a dose of any of the compositions described herein. In some embodiments, the dosing regimen entails oral administration of multiple doses of any of the compositions described herein. In some embodiments, the composition is administered orally the subject once, twice, 3 times, 4 times, 5 times, 6 times, 7 times, 8 times, 9 times, or at least 10 times.

The compositions, including the pharmaceutical compositions disclosed herein, include compositions with a range of active ingredients (*e.g.*, live bacteria, bacteria in spore format). The amount of bacteria in the compositions may be expressed in weight, number of bacteria and/or CFUs (colony forming units). In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^1 , about 10^2 , about 10^3 , about 10^4 , about 10^5 , about 10^6 , about 10^7 , about 10^8 , about 10^9 , about 10^{10} , about 10^{11} , about 10^{12} , about 10^{13} or more of each of the bacteria of the composition per dosage amount. In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^1 , about 10^2 , about 10^3 , about 10^4 , about 10^5 , about 10^6 , about 10^7 , about 10^8 , about 10^9 , about 10^{10} , about 10^{11} , about 10^{12} , about 10^{13} or more total bacteria per dosage amount. It should further be appreciated that the bacteria of the compositions may be present in different amounts. Thus, for instance, as a non-limiting example, a composition may include 10^3 of bacteria A, 10^4 of bacteria B and 10^6 of bacteria C. In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^1 , about 10^2 , about 10^3 , about 10^4 , about 10^5 , about 10^6 , about 10^7 , about 10^8 , about 10^9 , about 10^{10} , about 10^{11} , about 10^{12} , about 10^{13} or more CFUs of each of the bacteria in the composition per dosage amount. In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^1 , about 10^2 , about 10^3 , about 10^4 , about 10^5 , about 10^6 , about 10^7 , about 10^8 , about 10^9 , about 10^{10} , about 10^{11} , about 10^{12} , about 10^{13} or more CFUs in total for all of the bacteria combined per dosage amount. As discussed above, bacteria of the compositions may be present in different amounts. In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^{-7} , about 10^{-6} , about 10^{-5} , about 10^{-4} , about 10^{-3} , about 10^{-2} , about 10^{-1} or more grams of each of the bacteria in the composition per dosage amount. In some embodiments, the pharmaceutical compositions disclosed herein contain about 10^{-7} , about 10^{-6} , about 10^{-5} , about 10^{-4} , about 10^{-3} , about 10^{-2} , about 10^{-1} or more grams in total for all of the bacteria combined per dosage

amount. In some embodiment, the dosage amount is one administration device (*e.g.*, one table, pill or capsule). In some embodiment, the dosage amount is the amount that is administered in a particular period (*e.g.*, one day or one week).

In some embodiments, the pharmaceutical compositions disclosed herein contain

5 between 10 and 10^{13} , between 10^2 and 10^{13} , between 10^3 and 10^{13} , between 10^4 and 10^{13} , between 10^5 and 10^{13} , between 10^6 and 10^{13} , between 10^7 and 10^{13} , between 10^8 and 10^{13} , between 10^9 and 10^{13} , between 10^{10} and 10^{13} , between 10^{11} and 10^{13} , between 10^{12} and 10^{13} , between 10 and 10^{12} , between 10^2 and 10^{12} , between 10^3 and 10^{12} , between 10^4 and 10^{12} , between 10^5 and 10^{12} , between 10^6 and 10^{12} , between 10^7 and 10^{12} , between 10^8 and 10^{12} ,

10 between 10^9 and 10^{12} , between 10^{10} and 10^{12} , between 10^{11} and 10^{12} , between 10 and 10^{11} , between 10^2 and 10^{11} , between 10^3 and 10^{11} , between 10^4 and 10^{11} , between 10^5 and 10^{11} , between 10^6 and 10^{11} , between 10^7 and 10^{11} , between 10^8 and 10^{11} , between 10^9 and 10^{11} , between 10^{10} and 10^{11} , between 10 and 10^{10} , between 10^2 and 10^{10} , between 10^3 and 10^{10} , between 10^4 and 10^{10} , between 10^5 and 10^{10} , between 10^6 and 10^{10} , between 10^7 and 10^{10} ,

15 between 10^8 and 10^{10} , between 10^9 and 10^{10} , between 10 and 10^9 , between 10^2 and 10^9 , between 10^3 and 10^9 , between 10^4 and 10^9 , between 10^5 and 10^9 , between 10^6 and 10^9 , between 10^7 and 10^9 , between 10^8 and 10^9 , between 10 and 10^8 , between 10^2 and 10^8 , between 10^3 and 10^8 , between 10^4 and 10^8 , between 10^5 and 10^8 , between 10^6 and 10^8 , between 10^7 and 10^8 , between 10 and 10^7 , between 10^2 and 10^7 , between 10^3 and 10^7 ,

20 between 10^4 and 10^7 , between 10^5 and 10^7 , between 10^6 and 10^7 , between 10 and 10^6 , between 10^2 and 10^6 , between 10^3 and 10^6 , between 10^4 and 10^6 , between 10^5 and 10^6 , between 10 and 10^5 , between 10^2 and 10^5 , between 10^3 and 10^5 , between 10^4 and 10^5 , between 10 and 10^4 , between 10^2 and 10^4 , between 10^3 and 10^4 , between 10 and 10^3 , between 10^2 and 10^3 , or between 10 and 10^2 of each of the bacteria of the composition per dosage

25 amount. In some embodiments, the pharmaceutical compositions disclosed herein contain between 10 and 10^{13} , between 10^2 and 10^{13} , between 10^3 and 10^{13} , between 10^4 and 10^{13} , between 10^5 and 10^{13} , between 10^6 and 10^{13} , between 10^7 and 10^{13} , between 10^8 and 10^{13} , between 10^9 and 10^{13} , between 10^{10} and 10^{13} , between 10^{11} and 10^{13} , between 10^{12} and 10^{13} , between 10 and 10^{12} , between 10^2 and 10^{12} , between 10^3 and 10^{12} , between 10^4 and 10^{12} ,

30 between 10^5 and 10^{12} , between 10^6 and 10^{12} , between 10^7 and 10^{12} , between 10^8 and 10^{12} , between 10^9 and 10^{12} , between 10^{10} and 10^{12} , between 10^{11} and 10^{12} , between 10 and 10^{11} , between 10^2 and 10^{11} , between 10^3 and 10^{11} , between 10^4 and 10^{11} , between 10^5 and 10^{11} , between 10^6 and 10^{11} , between 10^7 and 10^{11} , between 10^8 and 10^{11} , between 10^9 and 10^{11} , between 10^{10} and 10^{11} , between 10 and 10^{10} , between 10^2 and 10^{10} , between 10^3 and 10^{10} ,

between 10^4 and 10^{10} , between 10^5 and 10^{10} , between 10^6 and 10^{10} , between 10^7 and 10^{10} ,
 between 10^8 and 10^{10} , between 10^9 and 10^{10} , between 10 and 10^9 , between 10^2 and 10^9 ,
 between 10^3 and 10^9 , between 10^4 and 10^9 , between 10^5 and 10^9 , between 10^6 and 10^9 ,
 between 10^7 and 10^9 , between 10^8 and 10^9 , between 10 and 10^8 , between 10^2 and 10^8 ,
 5 between 10^3 and 10^8 , between 10^4 and 10^8 , between 10^5 and 10^8 , between 10^6 and 10^8 ,
 between 10^7 and 10^8 , between 10 and 10^7 , between 10^2 and 10^7 , between 10^3 and 10^7 ,
 between 10^4 and 10^7 , between 10^5 and 10^7 , between 10^6 and 10^7 , between 10 and 10^6 ,
 between 10^2 and 10^6 , between 10^3 and 10^6 , between 10^4 and 10^6 , between 10^5 and 10^6 ,
 between 10 and 10^5 , between 10^2 and 10^5 , between 10^3 and 10^5 , between 10^4 and 10^5 ,
 10 between 10 and 10^4 , between 10^2 and 10^4 , between 10^3 and 10^4 , between 10 and 10^3 , between
 10^2 and 10^3 , or between 10 and 10^2 total bacteria per dosage amount.

In some embodiments, the pharmaceutical compositions disclosed herein contain
 between 10 and 10^{13} , between 10^2 and 10^{13} , between 10^3 and 10^{13} , between 10^4 and 10^{13} ,
 between 10^5 and 10^{13} , between 10^6 and 10^{13} , between 10^7 and 10^{13} , between 10^8 and 10^{13} ,
 15 between 10^9 and 10^{13} , between 10^{10} and 10^{13} , between 10^{11} and 10^{13} , between 10^{12} and 10^{13} ,
 between 10 and 10^{12} , between 10^2 and 10^{12} , between 10^3 and 10^{12} , between 10^4 and 10^{12} ,
 between 10^5 and 10^{12} , between 10^6 and 10^{12} , between 10^7 and 10^{12} , between 10^8 and 10^{12} ,
 between 10^9 and 10^{12} , between 10^{10} and 10^{12} , between 10^{11} and 10^{12} , between 10 and 10^{11} ,
 between 10^2 and 10^{11} , between 10^3 and 10^{13} , between 10^4 and 10^{13} , between 10^5 and 10^{13} ,
 20 between 10^6 and 10^{13} , between 10^7 and 10^{11} , between 10^8 and 10^{11} , between 10^9 and 10^{11} ,
 between 10^{10} and 10^{11} , between 10 and 10^{10} , between 10^2 and 10^{10} , between 10^3 and 10^{10} ,
 between 10^4 and 10^{10} , between 10^5 and 10^{10} , between 10^6 and 10^{10} , between 10^7 and 10^{10} ,
 between 10^8 and 10^{10} , between 10^9 and 10^{10} , between 10 and 10^9 , between 10^2 and 10^9 ,
 between 10^3 and 10^9 , between 10^4 and 10^9 , between 10^5 and 10^9 , between 10^6 and 10^9 ,
 25 between 10^7 and 10^9 , between 10^8 and 10^9 , between 10 and 10^8 , between 10^2 and 10^8 ,
 between 10^3 and 10^8 , between 10^4 and 10^8 , between 10^5 and 10^8 , between 10^6 and 10^8 ,
 between 10^7 and 10^8 , between 10 and 10^7 , between 10^2 and 10^7 , between 10^3 and 10^7 ,
 between 10^4 and 10^7 , between 10^5 and 10^7 , between 10^6 and 10^7 , between 10 and 10^6 ,
 between 10^2 and 10^6 , between 10^3 and 10^6 , between 10^4 and 10^6 , between 10^5 and 10^6 ,
 30 between 10 and 10^5 , between 10^2 and 10^5 , between 10^3 and 10^5 , between 10^4 and 10^5 ,
 between 10 and 10^4 , between 10^2 and 10^4 , between 10^3 and 10^4 , between 10 and 10^3 , between
 10^2 and 10^3 , or between 10 and 10^2 CFUs of each of the bacteria of the composition per
 dosage amount. In some embodiments, the pharmaceutical compositions disclosed herein
 contain between 10 and 10^{13} , between 10^2 and 10^{13} , between 10^3 and 10^{13} , between 10^4 and

10^{13} , between 10^5 and 10^{13} , between 10^6 and 10^{13} , between 10^7 and 10^{13} , between 10^8 and 10^{13} , between 10^9 and 10^{13} , between 10^{10} and 10^{13} , between 10^{11} and 10^{13} , between 10^{12} and 10^{13} , between 10 and 10^{12} , between 10^2 and 10^{12} , between 10^3 and 10^{12} , between 10^4 and 10^{12} , between 10^5 and 10^{12} , between 10^6 and 10^{12} , between 10^7 and 10^{12} , between 10^8 and 10^{12} , between 10^9 and 10^{12} , between 10^{10} and 10^{12} , between 10^{11} and 10^{12} , between 10 and 10^{11} , between 10^2 and 10^{11} , between 10^3 and 10^{13} , between 10^4 and 10^{13} , between 10^5 and 10^{13} , between 10^6 and 10^{13} , between 10^7 and 10^{11} , between 10^8 and 10^{11} , between 10^9 and 10^{11} , between 10^{10} and 10^{11} , between 10 and 10^{10} , between 10^2 and 10^{10} , between 10^3 and 10^{10} , between 10^4 and 10^{10} , between 10^5 and 10^{10} , between 10^6 and 10^{10} , between 10^7 and 10^{10} , between 10^8 and 10^{10} , between 10^9 and 10^{10} , between 10 and 10^9 , between 10^2 and 10^9 , between 10^3 and 10^9 , between 10^4 and 10^9 , between 10^5 and 10^9 , between 10^6 and 10^9 , between 10^7 and 10^9 , between 10^8 and 10^9 , between 10 and 10^8 , between 10^2 and 10^8 , between 10^3 and 10^8 , between 10^4 and 10^8 , between 10^5 and 10^8 , between 10^6 and 10^8 , between 10^7 and 10^8 , between 10 and 10^7 , between 10^2 and 10^7 , between 10^3 and 10^7 , between 10^4 and 10^7 , between 10^5 and 10^7 , between 10^6 and 10^7 , between 10 and 10^6 , between 10^2 and 10^6 , between 10^3 and 10^6 , between 10^4 and 10^6 , between 10^5 and 10^6 , between 10 and 10^5 , between 10^2 and 10^5 , between 10^3 and 10^5 , between 10^4 and 10^5 , between 10 and 10^4 , between 10^2 and 10^4 , between 10^3 and 10^4 , between 10 and 10^3 , between 10^2 and 10^3 , or between 10 and 10^2 total CFUs per dosage amount.

In some embodiments, the pharmaceutical compositions disclosed herein contain between 10^{-7} and 10^{-1} , between 10^{-6} and 10^{-1} , between 10^{-5} and 10^{-1} , between 10^{-4} and 10^{-1} , between 10^{-3} and 10^{-1} , between 10^{-2} and 10^{-1} , between 10^{-7} and 10^{-2} , between 10^{-6} and 10^{-2} , between 10^{-5} and 10^{-2} , between 10^{-4} and 10^{-2} , between 10^{-3} and 10^{-2} , between 10^{-7} and 10^{-3} , between 10^{-6} and 10^{-3} , between 10^{-5} and 10^{-3} , between 10^{-4} and 10^{-3} , between 10^{-7} and 10^{-4} , between 10^{-6} and 10^{-4} , between 10^{-5} and 10^{-4} , between 10^{-7} and 10^{-5} , between 10^{-6} and 10^{-5} , or between 10^{-7} and 10^{-6} grams of each of the bacteria in the composition per dosage amount. In some embodiments, the pharmaceutical compositions disclosed herein contain between 10^{-7} and 10^{-1} , between 10^{-6} and 10^{-1} , between 10^{-5} and 10^{-1} , between 10^{-4} and 10^{-1} , between 10^{-3} and 10^{-1} , between 10^{-2} and 10^{-1} , between 10^{-7} and 10^{-2} , between 10^{-6} and 10^{-2} , between 10^{-5} and 10^{-2} , between 10^{-4} and 10^{-2} , between 10^{-3} and 10^{-2} , between 10^{-7} and 10^{-3} , between 10^{-6} and 10^{-3} , between 10^{-5} and 10^{-3} , between 10^{-4} and 10^{-3} , between 10^{-7} and 10^{-4} , between 10^{-6} and 10^{-4} , between 10^{-5} and 10^{-4} , between 10^{-7} and 10^{-5} , between 10^{-6} and 10^{-5} , or between 10^{-7} and 10^{-6} grams of all of the bacteria combined per dosage amount.

Also with the scope of the present disclosure are food products comprising any of the bacterial strains described herein and a nutrient. Food products are, in general, intended for the consumption of a human or an animal. Any of the bacterial strains described herein may be formulated as a food product. In some embodiments, the bacterial strains are formulated as a food product in spore form. In some embodiments, the bacterial strains are formulated as a food product in vegetative form. In some embodiments, the food product comprises both vegetative bacteria and bacteria in spore form. The compositions disclosed herein can be used in a food or beverage, such as a health food or beverage, a food or beverage for infants, a food or beverage for pregnant women, athletes, senior citizens or other specified group, a functional food, a beverage, a food or beverage for specified health use, a dietary supplement, a food or beverage for patients, or an animal feed. Non-limiting examples of the foods and beverages include various beverages such as juices, refreshing beverages, tea beverages, drink preparations, jelly beverages, and functional beverages; alcoholic beverages such as beers; carbohydrate-containing foods such as rice food products, noodles, breads, and pastas; paste products such as fish hams, sausages, paste products of seafood; retort pouch products such as curries, food dressed with a thick starchy sauces, soups; dairy products such as milk, dairy beverages, ice creams, cheeses, and yogurts; fermented products such as fermented soybean pastes, yogurts, fermented beverages, and pickles; bean products; various confectionery products such as Western confectionery products including biscuits, cookies, and the like, Japanese confectionery products including steamed bean-jam buns, soft adzuki-bean jellies, and the like, candies, chewing gums, gummies, cold desserts including jellies, cream caramels, and frozen desserts; instant foods such as instant soups and instant soy-bean soups; microwavable foods; and the like. Further, the examples also include health foods and beverages prepared in the forms of powders, granules, tablets, capsules, liquids, pastes, and jellies.

Food products containing bacterial strains described herein may be produced using methods known in the art and may contain the same amount of bacteria (*e.g.*, by weight, amount or CFU) as the pharmaceutical compositions provided herein. Selection of an appropriate amount of bacteria in the food product may depend on various factors, including for example, the serving size of the food product, the frequency of consumption of the food product, the specific bacterial strains contained in the food product, the amount of water in the food product, and/or additional conditions for survival of the bacteria in the food product.

Examples of food products which may be formulated to contain any of the bacterial strains described herein include, without limitation, a beverage, a drink, a bar, a snack, a

dairy product, a confectionery product, a cereal product, a ready-to-eat product, a nutritional formula, such as a nutritional supplementary formulation, a food or beverage additive.

In some embodiments, the subject has not received a dose of an antibiotic prior to administration of the bacterial composition. In some embodiments, the subject has not been administered an antibiotic at least 1, at least 2, at least 3, at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 60, at least 90, at least 120, at least 180 or at least 360 days prior to administration of the compositions provided herein. In some embodiments, the person has not been administered an antibiotic to treat the pathogenic infection. In some embodiments, the compositions provided herein comprise the first treatment of the pathogenic infection.

In some embodiments, the subject may be administered one or more doses of an antibiotic prior to or concurrently with a bacterial composition. Generally, the first line of defense in the treatment of a pathogenic infection is the administration of an antibiotic. In some embodiments, the subject is administered a single dose of an antibiotic prior to the bacterial composition. In some embodiments, the subject is administered multiple doses of an antibiotic prior to the bacterial composition. In some embodiments, the subject is administered at least 2, 3, 4, 5 or more doses of an antibiotic prior to the bacterial composition. In some embodiments, the subject is administered a dose of an antibiotic at substantially the same time as the bacterial composition. Examples of antibiotics that can be administered include, without limitation, kanamycin, gentamicin, colistin, metronidazole, vancomycin, clindamycin, fidaxomicin, and cefoperazone.

Table 1 below provides sequence identifier numbers (SEQ ID NOs) used in the compositions of the experiments disclosed herein, along with the accompanying strain identification number (Strain ID). The closest bacterial species to the indicated strain is presented by genus-species. The 16S rDNA sequence associated with each genus species identified as the closest related genus species is also provided. The percent alignment presents the percent identity between the sequence of the indicated strain with the sequence from the closest genus species and the length of the alignment. The GenBank Accession Number of the closest related species is provided in the last column.

Table 1: Closest bacterial species to the strains described herein

SEQ ID	Strain ID	Closest Genus species	SEQ ID NO. of closest species	Percent alignment	Alignment length	Accession # of closest species
SEQ ID NO: 01	71	Blautia wexlerae	SEQ_94	96.62	207	NR_044054
SEQ ID NO: 02	102	Turicibacter_sanguinis	SEQ_91	97.81	183	NR_028816
SEQ ID NO: 03	5	Clostridium_hathewayi	SEQ_105	92.42	198	NR_036928
SEQ ID NO: 04	7	Blautia_hansenii	SEQ_99	96.62	207	NR_104687
SEQ ID NO: 05	10	Blautia_hansenii	SEQ_99	98.06	206	NR_104687
SEQ ID NO: 06	40	Lactobacillus_mucosae	SEQ_90	87.57	185	NR_024994
SEQ ID NO: 07	59	Blautia_producta	SEQ_106	98.54	206	NR_113270
SEQ ID NO: 07	59	Blautia_coccoides	SEQ_103	98.54	206	NR_104700
SEQ ID NO: 08	79	Blautia_hansenii	SEQ_99	100	194	NR_104687
SEQ ID NO: 09	VE202-21	Eubacterium_contortum	SEQ_109	94.59	296	NR_117147
SEQ ID NO: 09	VE202-21	Eubacterium_fissicatena	SEQ_108	94.59	296	NR_117142
SEQ ID NO: 10	211	Flavonifractor_plautii	SEQ_93	98.49	199	NR_043142
SEQ ID NO: 11	VE202-9	Anaerostipes_caccae	SEQ_88	99.5	399	NR_028915
SEQ ID NO: 12	VE202-26	Clostridium_scindens	SEQ_87	95.76	354	NR_028785
SEQ ID NO: 13	136	Marvinbryantia_formatexigens	SEQ_89	94.66	131	NR_042152
SEQ ID NO: 14	VE202-13	Anaerotruncus_colihominis	SEQ_95	99.34	1365	NR_027558
SEQ ID NO: 15	VE202-14	Eubacterium_fissicatena	SEQ_102	93.33	1530	NR_117563
SEQ ID NO: 16	VE202-16	Clostridium_symbiosum	SEQ_122	98.43	1469	NR_118730
SEQ ID NO: 17	VE202-7	Clostridium_bolteae	SEQ_110	99.86	1390	NR_113410
SEQ ID NO: 18	148	Dorea_longicatena	SEQ_97	99.7	1318	NR_028883
SEQ ID NO: 19	16	Blautia_producta	SEQ_106	98.33	1493	NR_113270
SEQ ID NO: 20	170	Dorea_longicatena	SEQ_97	99.7	1318	NR_028883
SEQ ID NO: 21	189	Clostridium_innocuum	SEQ_98	98.64	1476	NR_029164
SEQ ID NO: 22	169	Dorea_longicatena	SEQ_97	99.58	475	NR_028883

SEQ ID NO: 23	VE202- 29	Eisenbergiella_tayi	SEQ_121	100	354	NR_118643
SEQ ID NO: 24	YK96	Dorea_longicatena	SEQ_97	99.48	191	NR_028883
SEQ ID NO: 25	YK101	Ruminococcus_obeum	SEQ_85	96.81	188	NR_118692
SEQ ID NO: 26	YK110	Megasphaera_elsdenii	SEQ_119	96.62	207	NR_102980
SEQ ID NO: 27	YK149	Acidaminococcus_fermentans	SEQ_115	99.48	192	NR_074928
SEQ ID NO: 27	YK149	Acidaminococcus_intestini	SEQ_112	99.48	192	NR_074306
SEQ ID NO: 28	YK154	Megasphaera_elsdenii	SEQ_119	96.12	206	NR_102980
SEQ ID NO: 29	YK36	Ruminococcus_faecis	SEQ_96	99.29	425	NR_116747
SEQ ID NO: 30	YK95	Bacteroides_cellulosilyticus	SEQ_100	99.54	437	NR_112933
SEQ ID NO: 31	YK32	Anaerostipes_hadrus	SEQ_107	98.8	415	NR_104799
SEQ ID NO: 32	YK64	Ruminococcus_obeum	SEQ_84	99.04	415	NR_119185
SEQ ID NO: 33	YK73	Flavonifractor_plautii	SEQ_93	98.56	418	NR_043142
SEQ ID NO: 34	YK87	Eubacterium_rectale	SEQ_114	99.52	416	NR_074634
SEQ ID NO: 35	YK105	Flavonifractor_plautii	SEQ_93	99.26	407	NR_043142
SEQ ID NO: 36	YK153	Megasphaera_elsdenii	SEQ_119	96.04	429	NR_102980
SEQ ID NO: 37	YK163	Eubacterium_rectale	SEQ_114	99.76	415	NR_074634
SEQ ID NO: 38	YK191	Ruminococcus_champanellensis	SEQ_117	94.47	416	NR_102884
SEQ ID NO: 38	YK191	Ruminococcus_albus	SEQ_113	94.47	416	NR_074399
SEQ ID NO: 39	YK99	Ruminococcus_champanellensis	SEQ_117	97.28	184	NR_102884
SEQ ID NO: 40	YK55	Ruminococcus_faecis	SEQ_96	99.02	408	NR_116747
SEQ ID NO: 41	YK75	Bifidobacterium_bifidum	SEQ_118	99.45	183	NR_102971
SEQ ID NO: 42	YK90	Anaerostipes_hadrus	SEQ_107	98.97	194	NR_104799
SEQ ID NO: 43	YK30	Anaerostipes_hadrus	SEQ_107	99.48	191	NR_104799
SEQ ID NO: 44	YK31	Anaerostipes_hadrus	SEQ_107	98.97	194	NR_104799
SEQ ID NO: 45	YK12	Eubacterium_rectale	SEQ_114	99.27	412	NR_074634
SEQ ID NO: 46	YK27	Ruminococcus_faecis	SEQ_96	99.51	412	NR_116747
SEQ ID NO: 47	YK28	Blautia_luti	SEQ_111	99.5	400	NR_041960

SEQ ID NO: 48	YK29	Ruminococcus_faecis	SEQ_96	99.03	413	NR_116747
SEQ ID NO: 49	YK33	Anaerostipes_hadrus	SEQ_107	99.27	413	NR_104799
SEQ ID NO: 50	YK34	Anaerostipes_hadrus	SEQ_107	99.51	410	NR_104799
SEQ ID NO: 51	YK35	Ruminococcus_faecis	SEQ_96	99.51	409	NR_116747
SEQ ID NO: 52	YK51	Eubacterium_rectale	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 53	YK52	Eubacterium_rectale	SEQ_114	99.03	413	NR_074634
SEQ ID NO: 54	YK54	Anaerostipes_hadrus	SEQ_107	85.82	409	NR_104799
SEQ ID NO: 55	YK56	Ruminococcus_faecis	SEQ_96	99.03	413	NR_116747
SEQ ID NO: 56	YK57	Ruminococcus_faecis	SEQ_96	98.79	413	NR_116747
SEQ ID NO: 57	YK58	Dorea_longicatena	SEQ_97	98.8	417	NR_028883
SEQ ID NO: 58	YK65	Roseburia_faecis	SEQ_92	99.27	413	NR_042832
SEQ ID NO: 59	YK67	Blautia_luti	SEQ_111	98.57	419	NR_041960
SEQ ID NO: 60	YK69	Fusicatenibacter_saccharivorans	SEQ_116	99.27	413	NR_114326
SEQ ID NO: 61	YK70	Fusicatenibacter_saccharivorans	SEQ_116	98.79	414	NR_114326
SEQ ID NO: 62	YK71	Roseburia_faecis	SEQ_92	99.28	414	NR_042832
SEQ ID NO: 63	YK74	Megasphaera_elsdenii	SEQ_119	96.06	431	NR_102980
SEQ ID NO: 64	YK88	Eubacterium_rectale	SEQ_114	99.28	415	NR_074634
SEQ ID NO: 65	YK89	Eubacterium_rectale	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 66	YK97	Roseburia_faecis	SEQ_92	99.28	414	NR_042832
SEQ ID NO: 67	YK98	Blautia_faecis	SEQ_104	98.02	405	NR_109014
SEQ ID NO: 68	YK139	Fusicatenibacter_saccharivorans	SEQ_116	99.03	412	NR_114326
SEQ ID NO: 69	YK141	Dorea_formicigenerans	SEQ_120	98.51	402	NR_044645
SEQ ID NO: 70	YK142	Ruminococcus_faecis	SEQ_96	98.79	413	NR_116747
SEQ ID NO: 71	YK152	Blautia_hansenii	SEQ_99	99.5	401	NR_104687
SEQ ID NO: 72	YK155	Blautia_hansenii	SEQ_99	98.79	413	NR_104687
SEQ ID NO: 73	YK157	Eubacterium_rectale	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 74	YK160	Roseburia_faecis	SEQ_92	99.03	414	NR_042832
SEQ ID NO: 75	YK166	Eubacterium_rectale	SEQ_114	99.27	409	NR_074634

NO: 75						
SEQ ID NO: 76	YK168	Eubacterium_rectale	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 77	YK169	Eubacterium_rectale	SEQ_114	99.28	416	NR_074634
SEQ ID NO: 78	YK171	Eubacterium_rectale	SEQ_114	97.87	188	NR_074634
SEQ ID NO: 79	YK192	Roseburia_faecis	SEQ_92	99.03	414	NR_042832
SEQ ID NO: 80	VE202-18	Erysipelatoclostridium_amosum	SEQ_123	100	1485	NR_113243
SEQ ID NO: 81	PE5	Clostridium_bolteae	SEQ_110	100	1385	NR_113410
SEQ ID NO: 82	PE9	Clostridium_disporicum	SEQ_86	99.21	382	NR_026491
SEQ ID NO: 83	211-B	Bacteroides_ovatus	SEQ_101	95.64	436	NR_112940

Table 2: Bacterial species with a high degree of homology based on whole genome analysis:

Strain	<u>Whole genome homology</u>
SEQ_10 - 211	<i>Lachnospiraceae bacterium 7_I_58FAA</i>
SEQ_14 - VE202-13	<i>Subdoligranulum</i>
	<i>Flavinofractor plautii</i>
	<i>Anaerotruncus colihominis</i>
SEQ_15 - VE202-14	<i>Eubacterium fissicatena</i>
	<i>Ruminococcus torques</i>
SEQ_16 - VE202-16	<i>Clostridium symbiosum</i>
SEQ_17 - VE202-7	<i>Clostridium bolteae</i>
SEQ_22 - 169 / SEQ_20 - 170	<i>Dorea longicatena</i>
SEQ_19 - 16	<i>Blautia producta</i>
	<i>Clostridium innocuum</i>
SEQ_21 - 189	<i>Erysipelotrichaceae bacterium 21_3</i>

Table 3: Bacterial species with highest degree of homology based on whole genome analysis

Composition B strain number	Strain identifier	SEQ ID # of 16S region as determined by Sanger sequencing	Closest species based on Sanger sequencing of 16S region	SEQ ID # of 16S regions as determined by WGS [^]	*Consensus SEQ ID # of 16S region as determined by WGS	Closest species based on Consensus SEQ ID # of 16S region as compared with 16S database	Closest species based on WGS compared versus WG databases	Additional closely related sequences	Clostridium cluster
1	VE202-7	17	Clostridium bolteae	124, 125, 126, 127, 128	124	Clostridium bolteae	Clostridium bolteae 90A9		XIVa
2	VE202-13	14	Anaerotruncus colihominis	129,130, 131	129	Anaerotruncus colihominis	Anaerotruncus colihominis DSM 17241		IV
3	VE202-14	15	Eubacterium fissicatena	132, 133, 134, 135, 136	132	Dracourtella massiliensis	Dracourtella massiliensis GD1	Ruminococcus torques; Sellimonas intestinalis	XIVa
4	VE202-16	16	Clostridium symbiosum	137, 138, 139, 140	137	Clostridium symbiosum	Clostridium symbiosum WAL- 14163		XIVa
5	strain #16	19	Blautia producta	141, 142, 143, 144, 145	141	Blautia producta	Clostridium bacterium UC5.1- 1D4	Blautia product ATCC 27340	XIVa
6	strain #170	20	Dorea longicatena	146, 147, 148, 149, 150, 151	146	Dorea longicatena	Dorea longicatena CAG:42		XIVa
7	strain #189	21	Clostridium innocuum	152, 153, 154, 155, 156	152	Clostridium innocuum	Erysipelotrichaceae bacterium 21_3		XVII
8	strain #211	10	Flavinofractor plautii	157, 158, 159	157	Flavinofractor plautii	Clostridium orbiscindens 1_3_50AFAA	Subdoligranulum	IV

[^]WGS refers to Whole Genome Sequencing performed on a PacBio Biosciences platform (Menlo Park, CA).

*Consensus sequence is defined as the 16S sequence that has the most overlap with all other identified 16S sequences.

In some embodiments, in any of the compositions described herein, *Clostridium bolteae* can be replaced with *Clostridium bolteae* 90A9 . In some embodiments, in any of the compositions described herein, *Anaerotruncus colihominis* can be replaced with *Anaerotruncus colihominis* DSM 17241. In some embodiments, in any of the compositions described herein, *Eubacterium fissicatena* can be replaced with *Sellimonas instestinalis*, *Drancourtella massiliensis* or *Drancourtella massiliensis GPI*. In some embodiments, in any of the compositions described herein, *Clostridium symbiosum* can be replaced with *Clostridium symbiosum* WAL-14163. In some embodiments, in any of the compositions described herein, *Blautia producta* can be replaced with *Clostridium bacterium* CD5.1-1D4 or *Blautia product* ATCC27340. In some embodiments, in any of the compositions described herein, *Dorea longicatena* can be replaced with *Dorea longicatena* CAG:42. In some embodiments, in any of the compositions described herein, *Clostridium innocuum* can be replaced with *Erysipelotrichaceae bacterium* 21_3. In some embodiments, in any of the compositions described herein, *Flavonifractor plautii* can be replaced with *Clostridium orbiscindens* 1_3_50AFAA.

Aspects described herein provide pharmaceutical composition comprising a purified bacterial mixture consisting of bacterial strains comprising 16S rDNA sequences of at least 95% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 97% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 98% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 99% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21.

In some aspects, at least a portion of the bacteria of the pharmaceutical composition are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In

some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide pharmaceutical compositions comprising a purified bacterial mixture consisting of bacterial strains comprising 16S rDNA sequences of at least 95% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 97% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 98% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 99% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157.

In some aspects, at least a portion of the bacterial strains are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide pharmaceutical compositions comprising a purified bacterial mixture comprising bacterial strains comprising 16S rDNA sequences of at least 95% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 97% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 98% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21. In some aspects, the bacterial strains have at least 99% homology to SEQ ID NO:10, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21.

In some aspects, at least a portion of the bacterial strains are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide pharmaceutical compositions comprising a purified bacterial mixture comprising bacterial strains comprising 16S rDNA sequences of at least 95% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 97% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 98% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, the bacterial strains have at least 99% homology to SEQ ID NO:124, SEQ ID NO:129, SEQ ID NO:132, SEQ ID NO:137, SEQ ID NO:141, SEQ ID NO:146, SEQ ID NO:152, and SEQ ID NO:157. In some aspects, at least a portion of the bacterial strains are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide pharmaceutical compositions comprising a purified bacterial mixture consisting of the following bacterial strains: *Clostridium bolteae*, *Anaerostruncus colihominis*, *Sellimonas intestinalis*, *Clostridium symbiosum*, *Blautia producta*, *Dorea Longicatena*, *Erysipelotrichaceae bacterium*, and *Clostridium orbiscindens*.

In some aspects, at least a portion of the bacterial strains are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide pharmaceutical compositions comprising a purified bacterial mixture comprising the following bacterial strains: *Clostridium bolteae*, *Anaerostruncus colihominis*, *Sellimonas intestinalis*, *Clostridium symbiosum*, *Blautia producta*, *Dorea Longicatena*, *Erysipelotrichaceae bacterium*, and *Clostridium orbiscindens*.

In some aspects, at least a portion of the bacterial strains are in spore-form. In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

In some aspects, the pharmaceutical composition is formulated for oral administration. In some aspects, the pharmaceutical composition is in the form of a capsule. In some aspects, the pharmaceutical composition is formulated for delivery to the colon. In some aspects, the pharmaceutical composition further comprises a pH sensitive composition comprising one or more enteric polymers.

Aspects described herein provide methods of treating an infectious disease in a subject, the method comprising administering the pharmaceutical composition of any of the aspects described herein to the subject in an amount sufficient to treat the infectious disease. In some aspects, the infectious disease is *Clostridium difficile* infection.

The nucleic acid sequences of the 16S rDNA, or portion thereof, for the bacterial strains described herein are provided below:

> SEQ ID NO: 01|71|

GCCCCGAGCAGTTGATGTGAAGGATGGGTCACCTGTGGACTGCATTGGAACTGTCATACTTGAGTGCCGGAGGGT
AAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGA
CGGTAAGTACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAA

> SEQ ID NO: 02|102|

CTAACCGTGGAGGTCATTGGAACTGGTCAACTTGAGTGCAGAAGAGGGAAGTGGAAATTCATGTGTAGCGGTGA
AATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGGCTTCCTGGTCTGTAACTGACACTGAGGCGCGAAAG
CGTGGGGGGCAAACAGGATTAGATCCCCCGGTAA

> SEQ ID NO: 03151

ATGAAAGCCGGGGCTCAACCCCGGTACTGCTTTGGAAACTGTTTGACTTGAGTGCTTGAGAGGTAAGTGGAATTC
CTAGTGTAGCGGGAAATGTTTAGATATTAGGAGGACACCAGTGGCGAAGGCGGCTTACTGGACTGTAAGTACGCT
TGTGGCTCGATTTGTGGGGAGCAAACAGGATTATATCCCTGGTAA

5

> SEQ ID NO: 04171

CGGAAGGTCTGATGTGAAGGTTGGGGCTTACCCCGGACTGCATTGGAAACTGTTTTTCTAGAGTGCCCGAGAGGT
AAGCGGAATTCCTAGTGTAGCGGTGAAATGCTTTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGG
ACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAA

10

> SEQ ID NO: 051101

CGATGTCTGAGTGAAGGCTGGGGCTTACCCAGGACTGCATTGGAACTGTTTTTCTAGAGTGCCGGAGAGGTAAG
CGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGG
TAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAA

15

> SEQ ID NO: 061401

TTAACCAAGAAGTGCATTGGAACTGTCAGACTTGGGGGAAAAAAGACAGTGCAACTCCATGTGTAGCGGTGGAA
TGCTCCATATATATGGAAGAACACCAGTGGCGAAGGCGGCTGTCTGGTCTGCAACTGACGCTGAGGCTCGAATTC
ATGGGTAAGAAAGTATTAGTCCCTTGTA

20

> SEQ ID NO: 071591

ACCCGCTTGGTCTGAGGTGAGGCTGGGGCTTAACCCAGGACTGCATTGGAACTGTTGTTCTAGAGTGCCGGAG
AGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTAC
TGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAA

25

> SEQ ID NO: 081791

TAGGCTGGGGCTTAACCCAGGACTGCATTGGAACTGTTTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTA
GTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGACGTT
GAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAA

30

> SEQ ID NO: 091VE202-211

TTGCATTGGACACTATGTCAGCTGAGTGTGCGAGAGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGAT
ATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGCACGTTTTCTGACGTTGAGGCTCGAAATCGTGGGGAGCA
AACAAAAATAGATACCCTGGTAGTCCACGCCGTAAACGATGCATACTAGGTGTCGGGTGGCAAAGCCATTGCGGTG
CCGCAGCAAACGCAATAAGTATGCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAATAAATTGACGGA

35

> SEQ ID NO: 1012111

CCCGTCGTAGATGTGAACTGGGGGCTCACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCA
ATCGGAATTCGTTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGATTGCTGGAC
AGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTCATAA

40

> SEQ ID NO: 111VE202-91

ACCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGTATGTAAAGCTCTATCAGCAGGGAAGAAAAAGA
CGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGGCAAGCGTTATCCG
GAATTACTGGGTGTAAAGGGTGCCTAGGTGGCATGGTAAAGTCAGAAAGTAAAAGCCCGGGGCTTAACCCCGGGACT
GCTTTTGAAACTGTCATGCTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATAT

45

TAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACTGTCACTGACACTGATGCACGAAAGCGTGGGGAGCAAA
CAGGATTAGATACCCTGGAAGTCCAT

> SEQ ID NO: 12|VE202-26|

5 ATGGGAGCGTAGATGGCGACTGGGCCATATGTGACAGCCCTGGTCTCAACCCCTTAACTGCATTTGGAAC TGAGT
GGCTGGAGTGTTCGGAGAGGCAGGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGT
GGCGAAGGCGGCCTGCTGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCT
GGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCTGGGTGGCAAGGACATTCCGTGCCGCAGCAAACGCAATAA
GTAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAAATTGACGGA

10

> SEQ ID NO: 13|136|

CGCAGCGGAGTGTATCCTAGGCTCACCTGGCTGCTTTCGAACTGGTTTTCTAGATCGTGTAGAGGGGGAGATTCC
TGGTGTAGCGTGAAATGCGTAGATATCTGGAGGAACACCAGTGGCGAAGGCGGCCTCCTGGACGGCAACTGACGT
TGAGGCTCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAA

15

> SEQ ID NO: 14|VE202-13|

20 TCAAAGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGCGCCTAACACATGCAAGTCGAACGGAGCTTACGT
TTTGAAGTTTTTCGGATGGACGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTTCAG
AGGGGGATAACAGCCGGAACCGGTGCTAATACCGCATGATGTTGCGGGGGCACATGCCCCTGCAACCAAAGGAG
CAATCCGCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCGGTA
GCCGGACTGAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGG
ATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGAAGACGGTCTTCGGATTGTAAACCTCTG
TCTTTGGGGAAGAAAATGACGGTACCCAAAGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
25 AGGGAGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTAGAATGTTAAATCCA
TCGGCTCAACCGGTGGCTGCGTTCTAAACTGCCGTTCTTGAGTGAAAGTAGAGGCAGGCGGAATTCCTAGTGTAGC
GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGGGCTTTAACTGACGCTGAGGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGGG
ACTGACCCCTTCCGTGCCGCAGTTAAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAA
AGGAATTGACGGGGGCCCGCACAAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGT
30 CTTGACATCGGATGCATAGCCTAGAGATAGGTGAAGCCCTTCGGGGCATCCAGACAGGTGGTGCATGGTTGTCGT
CAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATTATTAGTTGCTACGCAAGAGCA
CTCTAATGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGG
GCTACACACGTACTACAATGGCACTAAAACAGAGGGCGGCGACACCGCGAGGTGAAGCGAATCCCGAAAAAGTGT
35 CTCAGTTCAGATTGCAGGCTGCAACCCGCTGCATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCG
GTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTCGGTAACACCCGAAGCCAGTAGCCT
AACC GCAAGGGGGGCGCTGTGCAAGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAG
GTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 15|VE202-14|

40 TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
TTCAGAATCTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCACCAAGCCGACGATCAGTAG
45 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAA
TATTGCACAATGGGGGAAACCTTGATGCAGCGACGCCGCTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT
CAGCAGGGAAGAAGATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTA
TGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAGCCCA
GGGCTCAACCCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGAGAGGTAAGCGGAATTCCTAGTGTAGC
GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
50 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCTGGTGT
GCAAAGCATCCGTTCCCGCAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAA
AGGAATTGACGGGGACCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGGT
CTTGACATCCGGATGACGGGCGAGTAATGTGCGCCGTCCCTTCGGGGCATCCGAGACAGGTGGTGCATGGTTGTCG
TCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCTTCAGTAGCCAGCATATAAG

GTGGGCACTCTGGAGAGACTGCCAGGGAGAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTAT
 GGCCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAATCCCAAAAA
 TAACGTCTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT
 GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCCATGGGAGTCAGTAACGCCCGAAGCCAG
 TGACCCAACCTTAGAGGAGGGAGCTGTGCAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 16|VE202-16|

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTAA
 CGGAAGTTTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTGATAC
 TGGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGCATGATACAGTGTGAAAACTC
 CGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGCGAAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAAGCCCCGC
 GGCTCAACTGCGGGACTGCTTTGGAACTGTTTAACTGGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAG
 CAAAGCTCTTCGGTGCCGTGCGAAACGCAGTAAGTATTCACCTGGGGAGTACGTTGCAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCAACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAGGTC
 TTGACATCGATCCGACGGGGGAGTAACGTCCCTTCCCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTGTTCGT
 CAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATTCCTAAGTAGCCAGCGGTTCCGGC
 CGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTTATG
 ATCTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAAAT
 AACGTCTCAGTTCCGACTGCAGGCTGCAACTCGCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
 TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGT
 GACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
 GGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 17|VE202-7|

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 ATGAAGTTTTTCGGATGGATTTTTGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACAC
 TGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCAGCATGGTACGGTGTGAAAACTC
 CGGTGGTGTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAACGGCCCAACAAAGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGCGAAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAG
 GGCTCAACCCTGGGACTGCTTTGGAACTGTTTTGCTAGAGTGTGCGAGAGGTAAAGTGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGG
 CAAAGCCCTTCGGTGCCGTGCGAAACGCAGTAAGCATTCACCTGGGGAGTACGTTGCAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCAACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTC
 TTGACATCCTCTTGACCGGCGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTTCGT
 CAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCCTTAGTAGCCAGCAGGTAAAG
 CTGGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTAT
 GATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAA
 TAACGTCCAGTTCCGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAAT
 GTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGCAACGCCCGAAGTCAG
 TGACCCAACCTCGCAAGAGAGGGAGCTGCCGAAGGCGGGCAGGTAACCTGGGGTGAAGTCGTAACAAGGTAGCCGT
 ATCGGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 18|148|

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCACTTTG
 GAAGATTCTTCGGATGAAGACTTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATAC

AGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACTC
 CGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAGC
 CGACCTGAGAGGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTTCGGTATGTAAACTTCTATC
 5 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCCGG
 GGCTCAACCCCGGGACTGCATTTGGAAGTCTGAGCTAGAGTGTGCGGAGAGGCAAGTGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCACTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTGAGGCTCG
 10 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTGCGGTGG
 CAAAGCCATTTCGGTGCCGACGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCTGATC
 TTGACATCCCGATGACCGCTTCGTAATGGAAGCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCTGT
 CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAAG
 15 CTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 GACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAAGAGAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAAA
 TAACGTCTCAGTTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAT
 GCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGTAACGCCCGAAGTCAG
 TGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
 CGGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 19|16|

ATCAGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
 AGTGGATCTCTTCGGATTGAAACTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCAT
 25 ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
 TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAAACGGCCACCAAGGCGACGATCAGTA
 GCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGA
 ATATTGCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAAACTTCTA
 TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
 30 AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
 GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTCTAGAGTGCCGGAGAGGTAAAGCGGAATTCCTAGTGTAG
 CGGTGAATGCGTAGATATTAGGAGGAACACCACTGGCGAAGGCGGCTTACTGGACGGTAACGACGTTGAGGCT
 CGAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTGCGGT
 GGCAAAGCCATTTCGGTGCCGACGCAAACGCAATAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCA
 35 AAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
 TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
 GTCAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGGTGA
 AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT
 ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCAAATCCCAAA
 40 AATAACTCCAGTCCGACTGCGACTGCAGTGCAGTGCACGAAGCTGGAATCGTAGTAATCGCGAATCCAGA
 ATGTGCGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGTAACGCCCGAAGTC
 AGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
 TATCGGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 20|170|

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTTG
 GAAGATTCTTCGGATGATTTCCCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATAC
 AGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACTC
 CGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
 50 ATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTTCGGTATGTAAACTTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCCGG
 GGCTCAACCCCGGGACTGCATTTGGAAGTCTGAGCTAGAGTGTGCGGAGAGGCAAGTGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCACTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTGAGGCTCG
 55 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTGCGGTGG
 CAAAGCCATTTCGGTGCCGACGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCTGATC
 TTGACATCCCGATGACCGCTTCGTAATGGAAGCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCTGT

CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAAG
CTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
GACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAAA
5 TAACGTCTCAGTTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAT
GCTGCGGTGAATACGTTCCCCGGGTCTTGTACACACCGCCCCGTACACCCATGGGAGTCAGTAACGCCCGAAGTCAG
TGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
CGGAAGGTGCGGCTGGATCACCTCCTTT

> SEQ ID NO: 21|189|

10 ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGATGATGCGTAGCCG
15 GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
TCGTCAATGGGGGAAACCCCTGAACGAGCAATGCCCGGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
AAGTGAAGAACGGCTCATAGAGGAAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGTGGCGTA
CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAACTGGTATGCTGGAGTGCAGAAGAGGG
CGATGGAATTCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTGCGCTGGT
20 CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
ATGAGAACCTAAGTGTGGAGGAATTCAGTGTGCTGAGTTAACGCAATAAGTTCTCCGCCTGGGGAGTATGCACGCA
AGTGTGAAACCTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
ATGGTTGTCTGCTCAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTGTGCGATGTTACC
25 AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGGCCACAAAGAGCAGCGACACAGTGATGTGAAGCGAA
TCTCATAAAGGTCTGCTCAGTTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGACACACCGCCCCGTCAAACCATGGGAGTCAGTAATACC
CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
30 TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

> SEQ ID NO: 22|169|

AGTAACGCGTGGGTAAACCTGCCTCATAACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACG
GTACCGCATGGTACAGTGGTAAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTA
35 ACGGCCTACCAAGCCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAG
ACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCTGAAGGAT
GAAGTATTTTCGGTATGTAACTTCTATCAGCAGGGAAGAAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACT
ACGTGCCAGCAGCCGCGGTAAATACGTAGGGGGCAAGCGTTATCCGATTACTGGGTGTAAAGGGAGCGTAGACG
GCACGGCAAGCCAGATGTGAAAGCCC

40

> SEQ ID NO: 23|VE202-29|

CAGGCTGGAGTGCAGGAGAGGTAAGCGGAATTCCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCA
GTGGCGAAGGCGGCTTACTGGACTGTAAGTACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACC
CTGGTAGTCCACGCGGTAAACGATGATTGCTAGGTGTAGGTGGGTATGGACCCATCGGTGCCGCAGCTAACGCAA
45 TAAGCAATCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGG
AGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCC

> SEQ ID NO: 24|YK96|

CCGGGGCTCACCCCGGGACTGCATTTGGAAGTGTGAGCTAGAGTGTGCGAGAGGCAAGTGGAATTCCTAGTGTA
GCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTGAGGC
TCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTA

50

> SEQ ID NO: 25|YK101|

AGGGTCAACCCCTGGACTGCATTGGAACTGTCAGGCTGGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAGC
GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGACGTTGATGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGTTAAA

5 > SEQ ID NO: 26|YK110|

GGGAAGTCGGTCTTAAGTGCAGGGGCTTAACCCCGTGAGGGGACCGAACTGTGAAGCTCGAGTGTGGAGAGGAA
AGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGAC
GACAACTGACGCTGAGGCGCGAAAGCCAGGGGAGCAAACGGGATTAGATAACCCAGTAA

10 > SEQ ID NO: 27|YK149|

TAGTCTGAGTGATGCGGGGCTTAACCCCGTATGGCGTTGGATACTGGAAGTCTTGAGTGCAGGAGAGGAAAGGGG
AATTCACAGTGTAGCGGTGAAATGCGTAGATATTGGGAGGAACACCAGTGGCGAAGGCGCCTTTCTGGACTGTGT
CTGACGCTGAGATGCGAAAGCCAGGGTAGCAAACGGGATTAGATAACACGGTA

15 > SEQ ID NO: 28|YK154|

GATAGTCGGTCTTAAGTGCAGGGGCTTACCCCGTGAGGGGACCGAACTGTGAAGCTCGAGTGTGGAGAGGAAAG
CGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGACGA
CAACTGACGCTGAGGCGCGAAAGCCAGGGGAGCAAACGGGATTAGATAACACGGTAA

20 > SEQ ID NO: 29|YK36|

CGTTTGCTCCACGCTTTTCGAGCCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAA
TATCTACGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCA
GTCCCGGGGTTGAGCCCCGGGTTTTACATCAGACTTGCCACTCCGCTCTACGCTCCCTTTACACCCAGTAAATCC
GGATAACGCTTGACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGT
25 CATTTTCTTCCCTGCTGATAGAGCTTTACATAACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTT
TCCCCCATTTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGTTTGA

> SEQ ID NO: 30|YK95|

30 TGTCACACTTTTCGAGCATCAGCGTCAGTTACAGTCCAGTAAGCTGCCTTCGCAATCGGAGTTCTTCGTGATATCT
AAGCATTTACCGCTACACCACGAATTCGCTTACCTCTACTGCACTCAAGACGACCAGTATCAACTGCAATTTT
ACGGTTGAGCCGCAAACTTTACAGCTGACTTAATAGTCCGCCTACGCTCCCTTTAAACCCAATAAATCCGGATA
ACGCTTGATCCTCCGTATTACCGCGGCTGCTGGCACGGAGTTAGCCGATCCTTATTCGTATGGTACATACAAA
AGCCACACGTGGCTCACTTTATTCCCATATAAAAGAAGTTTACAACCCATAGGGCAGTCATCTTCACGCTACTT
GGCTGGTTCAGACTCTCGTCCATTGACCAATATTCTCACTGCTGCCTCCCGTAGGTAGTTTGGAA

35

> SEQ ID NO: 31|YK32|

CCGTTGTCACGCTTTTCGTGCTCAGTGTCAGTTTCAGTCCAGTAAGCCGCCTTCGCCACTGATGTTCTCCTAATA
TCTACGCATTTACCGCTACACTAGGAATTCGCTTACCTCTCCTGCACTCCAGTCTGACAGTTTCAAAAGCAGT
CCCAGAGTTAAGCCCTGGGTTTTCACTTCTGACTTGCCATACCACTACGCACCTTTACACCCAGTAATTCGG
40 ATAACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCA
TTTTCTTCCCTGCTGATAGAGCTTTACATAACCGAAATACTTCTTCACTCACGCGGCGTCGCTGCATCAGGGTTCC
CCCCATTGTGCAATATTCCCCACTGCTGCCTCCCGTGGAGTTTGA

> SEQ ID NO: 32|YK64|

45 GCGAATGTACGCATTCGAGCCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAAT
ATCTACGCATTTACCGCTACACTAGGAATTCGCTTACCTCTCCGGCACTCAAGACTAACAGTTTCCAATGCAG
TCCAGGGGTTGAGCCCCCGCTTTACATCAGACTTGCCAGTCCGCTTACGCTCCCTTTACACCCAGTAAATCCG
GATAACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTC
ACTATCTTCCCTGCTGATAGAAGTTTACATAACCGAGATACTTCTTCTTCACGCGGCGTCGCTGCATCAGGGTTT
50 CCCCCATTGTGCAATATTCCCCACTGCTGCCTCCCGTGGGAGTTTGGAA

> SEQ ID NO: 33|YK73|

5 TGCTCACGCTTTTCGCGCTCAGCGTCAGTTACTGTCCAGCAATCCGCCTTCGCCACTGGTGTTCCTCCGTATATCT
ACGCATTTACCGCTACACACGGAATTCCGATTGCCTCTCCAGCACTCAAGAACTACAGTTTCAAATGCAGGCTG
GAGGTTGAGCCCCCAGTTTTACATCTGACTTGCAATCCCGCTACACGCCCTTTACACCCAGTAAATCCGGATA
ACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTATTTCGTACAGGTACCGTCATTT
GTTTCGTCCCTGACAAAAGAAGTTTACAACCCGAAAGCCTTCTTCCTTCACGCGGCGTTGCTGGGTACAGGCTTGC
GCCCATTGCCCAATATTCCCCACTGCTGCCTCCCGTGGTAGTTTGGA

10 > SEQ ID NO: 34|YK87|

15 TGTCCACGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCT
ACGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACC
GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATA
ACGCTTGCCACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGGCTTCTTAGTCAGGTACCGTCATTA
TCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCAGCGGCGTTCGCTGCATCAGGCTTTCGCC
CATTGTGCAATATTCCCCACTGCTGACTCCCGTAGGAGTTTGGA

> SEQ ID NO: 35|YK105|

20 CGTTTCTCCACGCTTCGCGCTCAGCGTCAGTTACTGTCCAGCAATCCGCCTTCGCCACTGGTGTTCCTCCGTATA
TCTACGCATTTACCGCTACACACGGAATTCCGATTGCCTCTCCAGCACTCAAGAACTACAGTTTCAAATGCAGG
CTGGAGGTTGAGCCCCCAGTTTTACATCTGACTTGCAATCCCGCTACACGCCCTTTACACCCAGTAAATCCGG
ATAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTATTTCGTACAGGTACCGTCA
TTTGTTCGTCCCCGACAAAAGAAGTTTACAACCCGAAAGCCTTCTTCCTTCACGCGGCGTTGCTGGGTACAGGCT
TGCGCCCATTTGCCCAATATTCCCCACTGCTGCCTCCCTGGGAAGTTTGGA

25

> SEQ ID NO: 36|YK153|

30 ATGTCTGACTTCGCGCCTCAGCGTCAGTTGTGCTCCAGAAAGCCGCTTTCGCCACTGGTGTTCCTCCTAATATC
TACGCATTTACCGCTACACTAGGAATTCCGCTTTCTCTCCGACACTCGAGCTTCACAGTTTCGGTCCCTCAC
GGGGTTAAGCCCCGCACTTTAAGACCGACTTGCGATGCCGCTGCGCGCCCTTTACGCCCAATAATTCCGGACA
ACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTCTCTTACGGTACCGTCAGGG
ATAACGGGTATTGACCGCTATCCTGTTTCGTCCCATATAACAGAACTTTACAACCCGAAGGCCGTATCGTTCACG
CGGCGTTGCTCCGTACAGCTTTTCGTCCATTGCGGAAGATTCCCCACTGCTGCCTCCCTGGGAAGTTTGGA

> SEQ ID NO: 37|YK163|

35 GTTTGCTCAGCCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATA
TCTACGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGT
ACCGGGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGG
ATAACGCTTGCCACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGGCTTCTTAGTCAGGTACCGTCA
TTATCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCAGCGGCGTTCGCTGCATCAGGCTTTC
40 GCCCATTTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGTTTGGA

> SEQ ID NO: 38|YK191|

45 CGTTGCTCAGCATTTCGAGCCTCAGCGTCAGTTAAGCCCAGTAAGCCGCCTTCGCCACTGATGTTCCTCCTAATA
TCTACGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTACTTCACTCAAGAACCACAGTTTCAAATGCAGT
TTATGGGTTAAGCCCATAGTTTTACATCTGACTTGCGATCCCGCTACGCTCCCTTTACACCCAGTAAATCCGG
ACAACGCTCGCTCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGAGCTTCTTCCTCAGGTACCGTCT
TTTTTTCGTCCCTGAAGACAGAGGTTTACAATCCTAAACCTTCTTCCCTCAGCGGCGTTCGCTGCATCAGAGTTT
CCTCCATTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGTTTGGA

50 > SEQ ID NO: 39|YK99|

TGGGCTTACCCATAAACTGCATTTGAAACTGTGGTTCCTTGAGTGAAAGTAGAGGTAAGCGGAATTCCTAGTGTAGC
GGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGGCTTTAACTGACGCTGAGGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCCAAGTAA

5 > SEQ ID NO: 40|YK55|

GTCAGCATCGAGCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTACGCA
TTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGGGGT
TGAGCCCCGGGTTTTACATCAGACTTGCCACTCCGCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGCT
TGCACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTCTTC
10 CCTGCTGATAGAGCTTTACATACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCATG
TGCAATATTCCCCACTGCTGCCTCCCGAGGGAGTTTGGA

> SEQ ID NO: 41|YK75|

TCATCGCTTACGGTGGAATCTGCGCCGGGTACGGGCGGGCTGGAGTGCGGTAGGGGAGACTGGAATTCGCGGTGTA
15 ACGGTGGAATGTGTAGATATCGGGAAGAACACCGATGGCGAAGGCAGGTCTCTGGGCCGTCACTGACGCTGAGGA
GCGAAAGCGTGGGGAGCGAACAGGATTAGATACAACGGTAA

> SEQ ID NO: 42|YK90|

TGAACCCAGGGCTTAACTCTGGGACTGCTTTTGAAGTGTGAGTGCAGGAGAGGTAAGCGGAATTCCTA
20 GTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGACTGAAACTGACACT
GAGGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCATGGTAA

> SEQ ID NO: 43|YK30|

ACCAGGGCTTAACTCTGGGACTGCTTTTGAAGTGTGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGT
25 AGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGACTGAAACTGACACTGAGG
CACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAA

> SEQ ID NO: 44|YK31|

GAACCCAGGGCTTAACTCTGGGACTGCTTTTGAAGTGTGAGTGCAGGAGAGGTAAGCGGAATTCCTAG
30 TGTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGACTGAAACTGACACTG
AGGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCGGTAA

> SEQ ID NO: 45|YK12|

GAGTCAGCTTTCGAGCTCACGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTA
35 CGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCG
GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTAT
CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGAGTTTGGA
40

> SEQ ID NO: 46|YK27|

TGTCAGCTTTCGAGCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTACG
CATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGGG
45 GTTGAGCCCCGGGTTTTACATCAGACTTGCCACTCCGCTACGCTCCCTTTACACCCAGTAAATCCGGATAACG
CTTGACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTCT
TCCCTGCTGATAGAGCTTTACATACCGAAATACTTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCAT
TGTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGTTTGGA

> SEQ ID NO: 47|YK28|

5 CACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTACGCATTTACCGCTACAC
TAGGAATTCCGCTTACCTCTCCGGCACTCAAGACGGGCAGTTTCCAATGCAGTCCCGGGGTTGAGCCCCAGCCTT
TCACATCAGACTTGTCCATCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGCTTGCCCCCTACGTAT
TACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTTCTTCCCTGCTGATAGAAG
TTTACATACCGAGATACTTCTTCCCTTACGCGGGCGTCGCTGCATCAGGGTTTCCCCCATTGTGCAATATTCCCCA
CTGCTGCCTCCCGTAGGAGTTTGGG

> SEQ ID NO: 48|YK29|

10 GTCAGCTTTTCGAGCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTACGC
ATTTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGGGG
TTGAGCCCCGGGTTTTTCACATCAGACTTGCCACTCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGC
TTGCACCATAACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTCTT
CCCTGCTGATAGAGCTTTACATACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCATT
15 GTGCAATATTCCCCACTGCTGCCTCCCGTGGGGAGTTTGGG

> SEQ ID NO: 49|YK33|

20 GATGCTCAGCTTTTCGTGCTCAGTGTCAGTTTTCAGTCCAGTAAGCCGCCTTCGCCACTGATGTTCCCTCCTAATATC
TACGCATTTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCTGCACTCCAGTCTGACAGTTTCAAAGCAGTCC
CAGAGTTAAGCCCTGGGTTTTCACTTCTGACTTGCCATACCACTACGCACCCCTTTACACCCAGTAAATCCGGAT
AACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATT
TTCTTCCCTGCTGATAGAGCTTTACATACCGAGATACTTCTTCACTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
CCATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGAAGTTTGGG

> SEQ ID NO: 50|YK34|70A_009_YK34_A1_A02

25 GTGTGAGCTTCGTGCTCAGTGTCAGTTTTCAGTCCAGTAAGCCGCCTTCGCCACTGATGTTCCCTCCTAATATCTAC
GCATTTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCTGCACTCCAGTCTGACAGTTTCAAAGCAGTCCCAG
AGTTAAGCCCTGGGTTTTCACTTCTGACTTGCCATACCACTACGCACCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTTC
30 TTCCCTGCTGATAGAGCTTTACATACCGAGATACTTCTTCACTCACGCGGCGTCGCTGCATCAGGGTTTCCCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGGAGTTTGGG

> SEQ ID NO: 51|YK35|

35 GTCAGCTTCGAGCCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTACGC
ATTTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGGGG
TTGAGCCCCGGGTTTTTCACATCAGACTTGCCACTCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGC
TTGCACCATAACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTCTT
CCCTGCTGATAGAGCTTTACATACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCATT
40 GTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGTTTGGG

> SEQ ID NO: 52|YK51|

45 TGTGAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCCTCCTAATATCTAC
GCATTTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGG
GGTTGAGCCCCGGGTTTTTCACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGACCATAACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTATC
TTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTGCCCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGGG

> SEQ ID NO: 53|YK52|

TTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACG
CATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGGG
GTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCTGCGCTCCCTTTACACCCAGTAAATCCGGATAACG
5 CTTGCACCATAACGTATTACCGCGGCTGCTGGCAGCTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTATCT
TCCCTGCTGATAGAGCTTTACATAACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCCAT
TGTGCAATATTCCCCACTGCTGCCTCCCCGAGGGGAGTTTG

> SEQ ID NO: 54|YK54|

10 TTCGGTCTGCTTTCCCTTCTCGCGCTCAGTGTGAGTTTCTGTCTAGTAAGCCGCCTTCGCCACTGATGTTTCCT
CCTAATATCTACGCACTTCACCGCTCCACAATGAATTCCGCTTACCCCTCCCGCGCTCTAGTCTGACAGTTTAA
AAAACTCCCCGAGAGAAACCTGGGTTTTTCTTCTGACATGCGATATCCACCCCCACCTTTTATACACCCAA
AAATCGGATAAAAGGTGCGACCTACGTATTATACCGGCTGCTGGGGCGTAGATAGCCGGGGTTCCTTATACAGGG
ACCGTCATTTTCTTTCCCGCTGATACAGCTTTACATAACCGAAATACTTCTTTCTCACGCGGCGTCGCTGCATCAG
15 GGTTCCTCCCCATTGTGCAATATTCCCCACTGCTGCCTCCCCGAGGGGAAGTTGGGGGAAA

> SEQ ID NO: 55|YK56|

20 GTTCAGCTTTTCGAGCCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGG
GGTTGAGCCCCGGGTTTTACATCAGACTTGCCACTCCGCTACGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGCACCATAACGTATTACCGCGGCTGCTGGCAGCTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTC
TTCCCTGCTGATAGAGCTTTACATAACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAGTTTGA

> SEQ ID NO: 56|YK57|

25 GTCAGCTTTTCGAGCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACGC
ATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGGGG
TTGAGCCCCGGGTTTTACATCAGACTTGCCACTCCGCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGC
TTGCACCATAACGTATTACCGCGGCTGCTGGCAGCTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTCTT
30 CCCTGCTGATAGAGCTTTACATAACCGAAATACTTCATCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCATT
GTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAGTTTGA

> SEQ ID NO: 57|YK58|

35 TCTCACGCTTTTCGAGCTCACGTCAGTCATCGTCCAGCAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTACCGCTACACTAGGAATTCCACTTGCTCTCCGACACTCTAGCTCAGCAGTTCCAAATGCAGTCCCGG
GGTTGAGCCCCGGGCTTTACATCTGGCTTGCCGTGCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGGCCCCCTACGTATTACCGCGGCTGCTGGCAGCTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTTC
TTCCCTGCTGATAGAAGTTTACATAACCGAAATACTTCATCCTTCACGCGGCGTCGCTGCATCAGAGTTTCTCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGGAGTTTG

> SEQ ID NO: 58|YK65|

40 GTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACG
CATTTACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGGG
GTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCTGCGCTCCCTTTACACCCAGTAAATCCGGATAACG
45 CTTGCACCATAACGTATTACCGCGGCTGCTGGCAGCTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTCTTC
TTCCCTGCTGATAGAGCTTTACATAACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAGTTTGA

> SEQ ID NO: 59|YK67|

50 AGCCCCGCTTTTCGAGCCTCACGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCAAGACGGGCAGTTTCCAATGCAGTCCCG
GGGTTGAGCCCCAGCCTTTACATCAGACTTGTCATCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAA

CGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTT
CTTCCCTGCTGATAGAAGTTTACATACCGAGATACTTCTTCCCTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGAAGTTTGGA

5 > SEQ ID NO: 60|YK69|

TGCTCAGCTTTTCGAGCCTCAGTTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCGAGCCAGACAGTTTCCAATGCAGTCCCA
GGGTTAAGCCCTGGGTTTTACATCAGACTTGCCCTGCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTAT
10 CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGGA

> SEQ ID NO: 61|YK70|

GTTGCTCAGCTTTTCGAGCTCAGTTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCT
15 ACGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCGAGCCAGACAGTTTCCAATGCAGTCCC
AGGGTTAAGCCCTGGGTTTTACATCAGACTTGCCCTGCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATA
ACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTA
TCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
CATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGAAAGTTTGGA

20

> SEQ ID NO: 62|YK71|

TGCTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCG
25 GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGCAACATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTCT
TCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
CATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGGA

> SEQ ID NO: 63|YK74|

GATGCCCTGGCTTCGCGCTCAGCGTCAGTTGTCGTCCAGAAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATC
TACGCATTTACCGCTACACTAGGAATTCCGCTTTCCCTCTCCGACACTCGAGCTTCACAGTTTCGGTCCCCTCAC
GGGTTAAGCCCCGCACTTTAAGACCGACTTGCGATGCCGCCTGCGCGCCCTTTACGCCAATAAATCCGGACA
ACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTCTTACGGTACCGTCAGGG
35 ATAACGGGTATTGACCGCTATCCTGTTTCGTCCATATAACAGAACTTTACAACCCGAAGGCCGTATCGTTCACG
CGGCGTTGCTCCGTCAGACTTTTCGTCCATTGCGGAAGATTCCCCACTGCTGCCTCCCGGGGGAGTTTGGA

> SEQ ID NO: 64|YK88|

GTCCCGCTTTTCGAGCCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCG
40 GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGCAACATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTAT
CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAAGTTTGGA

45 > SEQ ID NO: 65|YK89|

TGTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGG
GGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCTGCGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGCAACATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTATC
50 TTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCCC
TTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGGA

> SEQ ID NO: 66|YK97|

5 TGCTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCACTTACCCCTCCGCACTCTAGTACGACAGTTTCCAATGCAGTACCG
GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTCT
TCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCC
CATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGGA

10 > SEQ ID NO: 67|YK98|

ATTTCAGCTTTTCGAGCTCACGTACGTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACG
CATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGGCACTCAAGCATACCAGTTTCCAATGCAGTCCAGGG
GTTAAGCCCCCTGCCTTTTACATCAGACTTGATACGCGCTCTACGCTCCCTTTACACCCAGTAAATCCGGATAACG
CTCGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTATCT
15 TCCCTGCTGATAGAAGTTTACATACCGAGATACTTCTTCCCTTACGCGGCGTCGCTGCATCAGGGTTTCCCCAT
TGTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAAGTTTGGA

> SEQ ID NO: 68|YK139|

20 GTGTCAGCTTTTCGAGCTCACGTACGTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCGAGCCAGACAGTTTCCAATGCAGTCCCAG
GGTTAAGCCCTGGGTTTTTACATCAGACTTGCCCTTGCCGCTCTACGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTATC
TTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCA
TTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGAGTTTGGA

25

> SEQ ID NO: 69|YK141|

GCCAGCTTCGAGCCTCACGTACGTACGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACGC
ATTTACCGCTACACTAGGAATTCCACTTACCTCTCCGCACTCTAGCTGCACAGTTTCCAAAGCAGTCCACAGG
TTGAGCCCATGCCTTTTCACTTACAGACTTGACACAGCCGCTACGCTCCCTTTACACCCAGTAAATCCGGATAACGC
30 TTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTTCTT
CCCTGCTGATAGAAGTTTACATACCGAAATACTTCACTCCTTACGCGGCGTCGCTGCATCAGGCTTTCGCCCATT
GTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAAGTTTGGA

> SEQ ID NO: 70|YK142|

35 TGATCAGCTTTTCGAGCTCACGTACGTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTTCCAATGCAGTCCCGG
GGTTGAGCCCCGGGTTTTTACATCAGACTTGCCACTCCGCTCTACGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGACCATACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTTTC
TTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCACTCCTTACGCGGCGTCGCTGCATCAGGGTTTCCCCCA
40 TTGTGCAATATTCCCCACTGCTGCCTCCCGGGGGAGTTTGGA

> SEQ ID NO: 71|YK152|

GATGATCAGCTTTTCGAGCTCACGTACGTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCT
ACGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCTAGAAAAACAGTTTCCAATGCAGTCCCT
45 GGGGTTAAGCCCCAGCCTTTTACATCAGACTTGCTCTTCCGCTCTACGCTCCCTTTACACCCAGTAAATCCGGATA
ACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTT
TCTTCCCTGCTGATAGAAGTTTACATACCGAGATACTTCTTCCCTTACGCGGCGTCGCTGCATCAGGGTTTCCCC
CATTGTGCAATATTCCCCACTGCTGCCTCCCGGGGGAAGTTTGGA

50 > SEQ ID NO: 72|YK155|

TTGATCAGCTTTTCGAGCTCAGTCAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCTAGAAAAACAGTTTCCAATGCAGTCCTG
GGGTTAAGCCCCAGCCTTTACATCAGACTTGCTCTTCCGTCTACGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGCCCCCTACGTATTACCGCGGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTT
5 CTTCCCTGCTGATAGAAAGTTTACATACCGAGATACTTCTTCCCTTCACGCGGCGTCGCTGCATCAGGGTTTCCCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGG

> SEQ ID NO: 73|YK157|

GTATTTTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATC
10 TACGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTAC
CGGGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGAT
AACGCTTGACCATACGTATTACCGCGGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATT
ATCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTTCGC
15 CCATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGG

> SEQ ID NO: 74|YK160|

GCTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTTCACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGG
20 GGTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGACCATACGTATTACCGCGGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTCTT
CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGGTTTCCCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAGTTTGG

> SEQ ID NO: 75|YK166|

TTTCAGCTTCGAGCCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTAC
GCATTTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCGG
GGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAAC
GCTTGACCATACGTATTACCGCGGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTATC
25 TTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCCA
30 TTGTGCAATATTCCCCACTGCTAGCTCCCGAAGGAGTTTGG

> SEQ ID NO: 76|YK168|

AGCTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCG
35 GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGACCATACGTATTACCGCGGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTAT
CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCC
ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGG

> SEQ ID NO: 77|YK169|

GTCCAGCTTTTCGAGCCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTA
CGCATTTACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGTACCG
GGGTTGAGCCCCGGGCTTTACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGGATAA
CGCTTGACCATACGTATTACCGCGGGCTGCTGGCACGTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCATTAT
40 CTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTCGCTCACGCGGCGTCGCTGCATCAGGCTTTCGCCC
45 ATTGTGCAATATTCCCCACTGCTGCCTCCCGAAGGGAGTTTGG

> SEQ ID NO: 78|YK171|

TGAGCCGGGCTCACCCCGGTACTGCATTGGAAGTGTCTACTAGAGTGTCCGAGGGGTAAGCGGAATTCCTAGTG
TAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGCTGAG
50 GCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACACCGGTAA

> SEQ ID NO: 79|YK192|

5 CACGATGTCAGCTTTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATA
TCTACGCATTTACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTTCCAATGCAGT
ACCGGGGTTGAGCCCCGGGCTTTCACATCAGACTTGCCGCACCGCCTGCGCTCCCTTTACACCCAGTAAATCCGG
ATAACGCTTGACCATACGTATTACCGCGGCTGCTGGCAGCTATTTAGCCGGTGCTTCTTAGTCAGGTACCGTCA
TTCTTCTTCCCTGCTGATAGAGCTTTACATAACGAAATACTTCTTCGCTCACGCGGCTGCTGCATCAGGGTTT
CCCCATTGTGCAATATTCCCCACTGCTGCCTCCCGAGGGGAGTTTGA

10 > SEQ ID NO: 80|VE202-18|

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAACGCGAGCACTTG
TGCTCGAGTGGCGAACGGGTGAGTAATACATAAGTAACCTGCCCTAGACAGGGGGATAACTATTGGAAACGATAG
CTAAGACCGCATAGGTACGGACACTGCATGGTGACCGTATTAAAAGTGCCCTCAAAGCACTGGTAGAGGATGGACT
TATGG
15 CGCATTAGCTGGTTGGCGGGGTAACGGCCCCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGACCGGCCAC
ACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGGGGAAACCCTGA
CCGAGCAACGCCGCGTGAAGGAAGAAGGTTTTCGGATTGTAAACTTCTGTTATAAAGGAAGAACGGCGGCTACAG
GAAAT
GGTAGCCGAGTGACGGTACTTTATTAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGG
20 CAAGCGTTATCCGGAATTATTGGGCGTAAAGAGGGAGCAGGCGGCAGCAAGGTCTGTGGTGAAAGCCTGAAGCT
TAACTTCAGTAAGCCATAGAAACCAGGCAGCTAGAGTGCAGGAGAGGATCGTGGAATTCCATGTGTAGCGGTGAA
ATGCG
TAGATATATGGAGGAACACCACTGGCGAAGGCGACGATCTGGCCTGCAACTGACGCTCAGTCCCGAAAGCGTGGG
GAGCAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACGATGAGTACTAAGTGTGGATGTCAAAGTTTCAG
25 TGCTGCAGTTAACGCAATAAGTACTCCGCTGAGTAGTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGG
GCCCC
CACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGACATACTCATAAAGG
CTCCAGAGATGGAGAGATAGCTATATGAGATACAGGTGGTGCATGGTTGTCGTGAGCTCGTGTCTGAGATGTTG
GGTTAAGTCCCGCAACGAGCGCAACCCCTTATCGTTAGTTACCATCATTAAGTTGGGGACTCTAGCGAGACTGCCA
30 GTGAC
AAGCTGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGA
TGGTGCAGAGGGAAGCGAAGCCGCGAGGTGAAGCAAAACCCATAAAACCATTTCTCAGTTCGGATTGTAGTCTGCA
ACTCGACTACATGAAGTTGGAATCGCTAGTAATCGCGAATCAGCATGTGCGGGTGAATACGTTCTCGGGCCTTGT
ACACA
35 CCGCCCGTCACACCACGAGAGTTGATAACACCCGAAGCCGGTGGCCTAACCGCAAGGAAGGAGCTGTCTAAGGTG
GGATTGATGATTGGGGTGAAGTCGTAACAAGGTATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

> SEQ ID NO: 81|PE5|

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTGAA
40 GGAAGTTTTTCGGATGGAATTCGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACACTG
GGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACAGTGTGAAAACTCCG
GTGGT
GTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAACGGCCCCACCAAGCGACGATCAGTAGCCGACCTG
AGAGGGTGACCGCCACATTGGGACTGAGACACGGCCCCAAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCAC
45 AATGGGCGAAAGCCTGATGCAGCGACCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATCAGCAGGG
AAGAA
AATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTT
ATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTAAAACCCAGGGCTCAACCCTG
GGACTGCTTTGGAACTGTTTTGCTAGAGTGTGCGGAGAGGTAAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTA
50 GATAT
TAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAAA
CAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGGCAAAGCCCTTCGGTGCC
GTCGCAACCGCAGTAAGCATTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGGACCC
GCACA
55 AGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTGACATCCTCTTGACCGGCGT
GTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCGTGAGCTCGTGTCTGAGATGTTG

GGTTAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCAGGTAAAGCTGGGCACTCTAGGGAGACTGC
CAGGG
ATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGTGCTACAATG
GCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAAATAACGTCCCAGTTCGGACTGTAGTCTG
5 CAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTGCGGTTGAATACGTTCCCGGGTCTT
GTACA
CACCGCCCGTCACACCATGGGAGTCAGCAACGCCCGAAGTCAGTGACCCAACTCGCAAGAGAGGGAGCTGCCGAA
GGCGGGGCAAGTAAGTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTTT

10 > SEQ ID NO: 82|PE9|

AATTCGACGTTGTCCGGATTACTGGGCGTAAAGGGAGCGTAGGCGGACTTTTAAAGTGAGATGTGAAATACCCGGG
CTCAACTTGGGTGCTGCATTTCAAAGTGAAGTCTAGAGTGCAGGAGAGGAGAATGGAATTCCTAGTGTAGCGGT
GAAATGCGTAGAGATTAGGAAGAACACCAAGTGGCGAAGGCGATTCTCTGGACTGTAACCTGACGCTGAGGCTCGAA
AGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTAGGGGTTGT
15 CATGACCTCTGTGCCGCCGCTAACGCATTAAGTATTCCGCCTGGGGAGTACGGTGCAGATTAACTCAAAGA
AATTGACGGA

> SEQ ID NO: 83|211-B|

ACGAGCGTATCGGATTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAGTTTTCGGCTCAA
20 CCGTAAATTTGCAGTTGAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGGTGTAGCGGTGAAAT
GCTTAGATATACGAAGAACTCCGATTGCGAAGGCAGCTCACTAGACTGTCACTGACACTGATGCTCGAAAGTGT
GGGTATCAAACAGGATTAGATACCCTGGTAGTCCACACAGTAAACGATGAATACTCGCTGTTTTCGATATACAGT
AAGCGGCCAAGCGAAAGCATTAAGTATTCCACCTGGGGAGTACGCCGGCAACGGTGAACCTCAAAGAAATTGACG
GAAGCCCGCCAGGGGGGAAAAACATGGGGTTTAGTTGGATGATACGGGGAGGAACCTC

25 > SEQ ID NO: 84|NR_119185.1|Ruminococcus obeum 16S ribosomal RNA gene, complete
sequence

GGCGGCGTGCTTAACACATGCAAGTCGAACGGGAAAACCTTTTCATTGAAGCTTCGGCAGATTTGGNNTGTTTCTAG
TGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTTATACAGGGGGGATAACAACCAGAAATGGTTGCTAATACC
30 GCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCT
AGTTGGCAGGTAACGGCCTACCAAGGCGACGATCCATAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTG
AGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCCTGATGCAGCGACG
CCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGCAGGGAAGATAGTGACGGTACCTGACTAAGAA
GCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGGCAAGCGTTATCCGGATTTACTGGGTGTA
35 GGGAGCGTAGACGGACTGGCAAGTCTGATGTGAAAGGCGGGGGCTCAACCCCTGGACTGCATTGGAACTGTTAG
TCTTGAGTGCCGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAAGT
GCGAAGGCGGCTTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTG
GTAGTCCACGCCGTAAACGATGATTACTAGGTGTTGGGGAGCAAAGCTCTTCGGTGCCGCCGCAAACGCATTAAG
TATTCCACCTGGGGAGTACGTTGCAAGAATGAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCA
40 TGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCCCTCTGACCGNCCCTTAACCGGATCT
TTCCTTCGGGACAGGGGAGACAGGTGGTGCATGGTTGTCGTGAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCG
CAACGAGCGCAACCCCTATCCCCAGTAGCCAGCAGTCCGGCTGGGCACTCTGAGGAGACTGCCAGGGATAACCTG
GAGGAAGGCGGGGATGACGTCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAAC
AAAGGGAAGCAAGCCTGCGAAGGTAAGCAAAATCCCAAAAAATAACGTCCCAGTTCGGACTGCAGTCTGCAACTCGA
45 CTGCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACAC
CGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACCTAACTGCAAAGAAGGAGCTGCCGAAGGCGG
GACCGATGACTGGGGTGAAGTCGTAACAAGGT

50 > SEQ ID NO: 85|NR_118692.1|Ruminococcus obeum strain ATCC 29174 16S ribosomal
RNA gene, complete sequence

GGCGTGCTTAACACATGCAAGTCGAACGGGAAAACCTTTTCATTGAAGCTTCGGCAGATTTGGTCTGTTTCTAGTGG
CGGACGGGTGAGTAACGCGTGGGTAACTGCCTTATACAGGGGGGATAACAACCAGAAATGGTTGCTAATACCGCA
TAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCTAGT

TGGCAGGGTAACGGCCTACCAAGGCGACGATCCATAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGA
 CACGCCCCAGACTCCTCGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGC
 GTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGCAGGGAAGATAGTGACGGTACCTGACTAAGAAGCCC
 CGKCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGA
 5 GCGTAGACGGACTGGCAAGTCTGATGTGAAAGGCGGGGGCTCAACCCCTGGACTGCATTGGAAGTGTAGTCTT
 GAGTGCCCGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGA
 AGGCGGCTTACTGGACGGTAACTGACGTTGAGGCTCGAAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAG
 TCCACGCCGCAAACGATGAATACTAGGTGTTGGGGAGCAAAGCTCTTCGGTGCCGCCGCAAACGCATTAAGTATT
 CCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTG
 10 GTTTAATTTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTGACATCCCTCTGACCGTCCCTTAACCGGATCTTTCC
 TTCGGGACAGGGGAGACAGGTGGTGCATGGTTGTCTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAAC
 GAGCGCAACCCCTATCCCCAGTAGCCAGCAGTNCGGCTGGGCACCTCTGAGGAGACTGCCAGGGGATAACCTGGAGG
 AAGGCGGGGATGACGTCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAG
 GGAAGCNAGCCTKCGRAGGTAAGCAAATCCCANAAAAAACGTCCAGTTCGGACTGCAGTCTGCAACTCGACTGC
 15 ACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCC
 CGTCACACCATGGGAGTCAGTAACGCCCCGAAGTCAGTGACCTAACTGC

> SEQ ID NO: 86|NR_026491.1|Clostridium disporicum strain DS1 16S ribosomal RNA
gene, partial sequence

GCTCAGGACGAACGCTGGCGGCGTGCCCTAACACATGCAAGTCGAGCGAGTTGATTCTCTTCGGAGATGAAGCTAG
 CGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCTCATAGAGGGGAATAGCCTCCCGAAAGGGAGATTAATACC
 GCATAAGATTGTAGCTTCGCATGAAGTAGCAATTAAGGAGCAATCCGCTATGAGATGGGCCCGCGGCGCATTAG
 CTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACATTGGGAC
 TGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCCTGATGCAGCAA
 25 CGCCGCGTGAGTGATGACGGCCTTCGGGTGTAAAGCTCTGTCTTCAGGGACGATAATGACGGTACCTGAGGAGG
 AAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCGAGCGTTGTCCGGATTTACTGGGCGTA
 AAGGGAGCGTAGGCGGACTTTTAAGTGAGATGTGAAATACCCGGGCTCAACTTGGGTGCTGCATTTCAAACCTGGA
 AGTCTAGAGTGACAGGAGAGGAGAATGGAATTCCTAGTGTAGCGGTGAAATGCGTAGAGATTAGGAAGAACCAG
 TGGCGAAGGCGATTCTCTGGACTGTAACGTGACGCTGAGGCTCGAAAAGCGTGGGGAGCAAACAGGATTAGATACCC
 30 TGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTAGGGGTGTGCATGACCTCTGTGCCGCCGCTAACGCATTA
 AGTATTCGCGCTGGGGAGTACGGTCGCAAGATTAACTCAAAGGAATTGACGGGGGCGCACAAGCAGCGGAG
 CATGTGGTTTAATTTCGAAGCAACGCGAAGAACCCTTACCTAGACTTGACATCTCTGAATTACCCGTAACCTGGGGA
 AGCCACTTCGGTGGCAGGAAGACAGGTGGTGCATGGTTGTCTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCC
 CGCAACGAGCGCAACCCCTTATTGTTAGTTGCTACCATTTAGTTGAGCACTCTAGCGAGACTGCCCGGGTTAACCG
 35 GGAGGAAGGTGGGATGACGTCAAATCATCATGCCCCCTTATGTCTAGGGCTACACACGTGCTACAATGGCAAGTA
 CAAAGAGAAGCAAGACCGCGAGGTGGAGCAAACTCAAAACTTGTCTCAGTTTCGGATTGTAGGCTGAAACTCGC
 CTACATGAAGCTGGAGTTGCTAGTAATCGCGAATCAGCATGTCTCGCGGTGAATACGTTCCCGGGCCTTGTACACAC
 CGCCCGTCACACCATGAGAGTTGGCAATACCAACGTACGTGATCTAACCCGCAAGGGAGGAAGCGTCTAAGGT
 40 AGGGTCAGCGATTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAA

> SEQ ID NO: 87|NR_028785.1|Clostridium scindens strain ATCC 35704 16S ribosomal
RNA gene, complete sequence

GAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCCTAACACATGCAAGTCGAACGAAGCGCCTGGCCC
 CGACTTCTTCGGAACGAGGAGCCTTTCGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTTGCAC
 45 TGGGGGATAACAGCCAGAAATGGCTGCTAATACCGCATAAGACCGAAGCGCCGCATGGCGCGGCGGCCAAAGCCC
 CGGCGGTGCAAGATGGGCCCGCGTCTGATTAGGTAGTTGGCGGGTAAACGGGCCACCAAGCCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCTGAAGGATGAAGTATTTCCGTATGTAACTTCTATC
 AGCAGGGAAGAAGATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 50 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGAGCGTAGACGGCGATGCAAGCCAGATGTGAAAGCCCCGG
 GGCTCAACCCCGGACTGCATTTGGAAGTGCCTGGCTGGAGTGTCTGAGAGGAGGCGGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCCTGCTGGACGATGACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTGGGTGG
 CAAGGCCATTTCGGTGCCGCGAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAA
 55 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTTCGAAGCAACGCCAAGAACCCTTACCTGATC
 TTGACATCCCGATGCCAAAGCGCGTAACGCGCTCTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCTGT

CAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCATTGGA
 TGGGCACTCTGGAGAGACTGCCAGGGAGAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
 ACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCGAACCCGCGAGGGTGGGCAAATCCCAAAAAT
 AACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGTTGGAATCGCTAGTAATCGCGAATCAGAATG
 TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCCGCCGTCACACCATGGGAGTCAGTAACGCCCGAAGCCGGT
 GACCCAACCCGTAAGGGAGGGAGCCGTCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTC

> SEQ ID NO: 88|NR_028915.1|Anaerostipes caccae strain L1-92 16S ribosomal RNA gene,
 partial sequence

GCGCTTAATACATGTCAAGTCGAACGAAGCATTAGGATTGAAGTTTTTCGGATGGATTTCTATATGACTGAGTG
 GCGGACGGGTGAGTAACGCGTGGGGAACCTGCCCTATACAGGGGGATAACAGCTGGAAACGGCTGCTAATACCGC
 ATAAGCGCACAGAATCGCATGATTGAGTGTAAAAAGCCCTGGCAGTATAGGATGGTCCCGCGTCTGATTAGCTGG
 TTGGTGAGGTAACGGCTCACCAAGGCGACGATCAGTAGCCGGCTTGAGAGAGTGAACGGCCACATTGGGACTGAG
 ACACGGCCCAAACCTCTACGGGAGGCGAGCAGTGGGGAATATTGCACAATGGGGGTAAACCTGATGCAGCGACGC
 CGCGTGAGTGAAGAAGTATTTCCGTATGTAAAGCTCTATCAGCAGGGAAGAAAACAGACGGTACCTGACTAAGAA
 GCCCCGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGGCAAGCGTTATCCGGAATTACTGGGTGTAAA
 GGGTGCCTAGGTGGCATGGTAAGTCAGAACTGAAAGCCCGGGGCTTAACCCCGGACTGCTTTTGAACCTGTCTAT
 GCTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAAGTG
 GCGAAGGCGGCTTACTGGACTGTCACTGACACTGATGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTG
 GTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGGCCGTAGAGGCTTCGGTGCCGAGCAAACGCAGTAAG
 TATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCA
 TGTGGTTTAATTGCAAGCAACGCGAAGAACCTTACCTGGTCTTGACATCCCAATGACCGAACCTTAACCGGTTTT
 TTCTTTTCGAGACATTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCG
 CAACGAGCGCAACCCCTATCTTTAGTAGCCAGCATTTAAGGTGGGCACTCTAGAGAGACTGCCAGGGATAACCTG
 GAGGAAGGTGGGGACGACGTCAAATCATCATGCCCTTATGGCCAGGGCTACACACGTGCTACAATGGCGTAAAC
 AAAGGGAAGCGAAGTCGTGAGGCGAAGCAAATCCAGAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGA
 CTACATGAAGCTGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTTCCCGGGTCTTGTACACAC
 CGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACCCAACCGCAAGGAGGGAGCTGCCGAAGGTGG
 GACCGATAACTGGGGTGAAGTCGTAACAAGG

> SEQ ID NO: 89|NR_042152.1|Marvinbryantia formatexigens strain I-52 16S ribosomal
 RNA gene, partial sequence >gil636558750|refl|NR_114807.1|Marvinbryantia formatexigens
 strain I-52 16S ribosomal RNA gene, complete sequence

TGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCATTTTAAATGAAGTTTTTCGGACGGAATTTAAATGACTG
 AGCGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTTATACAGGGGGATAACAGCCAGAAATGGCTGCTAATA
 CCGCATAAGCGCAGGTAACCGCATGGTACAGTGTGAAACTCCGGTGGTATAAGATGGGTCCGCTTGGATTAG
 GCAGTTGGCGGGGTAAAGGCCACCAACCGACGATCCATAGCCGGCCTGAGAGGGTGGACGGCCACATTGGGAC
 TGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCCTGATGCAGCGA
 CGCCGCGTGGGTGAAGAAGTATTTCCGTATGTAAAGCCCTATCAGCAGGGAAGAAAATGACGGTACCTGACCAAG
 AAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTA
 AAGGGAGCGTAGACGGCCATGCAAGTCTGGTGTGAAAGGCGGGGCTCAACCCCGGACTGCATTGGAACCTGTA
 TGGCTTGAGTGCCGGAGAGGTAAGCGGAATTCCTGGTGTAGCGGTGAAATGCGTAGATATCAGGAGGAACACCAG
 TGGCGAAGGCGGCTTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCC
 TGGTAGTCCACGCCGTAAACGATGAATAACAGGTGTGGGGGACACGGTCCTTCGGTGCCGAGCAAACGCACTA
 AGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAG
 CATGTGGTTTAATTGCAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCGACGACCGGACAGTAACGTGTC
 CTTCCCTTCGGGGCGTCCGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCC
 CGCAACGAGCGCAACCCCTGTTCCAGTAGCCAGCATTCAGGATGGGCACTCTGGGGAGACTGCCAGGGATAACC
 TGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCTTATGATCTGGGCTACACACGTGCTACAATGGCGTGA
 ACAGAGGGAAGCGAACCCGCGAGGGGGAGCAAAATCCAGAAATAACGTCCAGTTCCGATTGTAGTCTGCAACCC
 GGCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTTGTACAC
 ACCGCCCCGTACACCATGGGAGTCGGAATGCCCGAAGTCAGTGACCCAACCGAAGGAGGGAGCTGCCGAAGGC
 GGGGCCGGTAACTGGGGTGAAGTCGTAACAA

> SEQ ID NO: 90|NR_024994.1|*Lactobacillus mucosae* strain S32 16S ribosomal RNA gene, complete sequence

```

5  AGAGTTTGATCCTGGCTCAGGATGAACGCCGCGGTGTGCCTAATACATGCAAGTCGAACGCGTTGGCCCAACTG
  ATTGAACGTGCTTGCACGGACTTGACGTTGGTTTACCAGCGAGTGGCGGACGGGTGAGTAACACGTAGGTAACCT
10  GCCCCAAAGCGGGGGATAACATTTGGAAACAGATGCTAATACCGCATAACAATTTGAATCGCATGATTCAAATTT
  AAAAGATGGCTTCGGCTATCACTTTGGGATGGACCTGCGGCGCATTAGCTTGTGGTAGGGTAACGGCCTACCAA
  GGCTGTGATGCGTAGCCGAGTTGAGAGACTGATCGGCCACAATGGAAC TGAGACACGGTCCATACTCCTACGGGA
  GGCAGCAGTAGGGAATCTTCCACAATGGGCGCAAGCCTGATGGAGCAACACCGCGTGAGTGAAGAAGGGTTTCGG
15  CTCGTAAAGCTCTGTTGTTAGAGAAGAACGTGCGTGAGAGCAACTGTTACGCGAGTGACGGTATCTAACCAGAAA
  GTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAA
  GCGAGCGCAGGCGGTTTGATAAGTCTGATGTGAAAGCCTTTGGCTTAACCAAAGAAGTGCATCGGAAACTGTCAG
  ACTTGAGTGCAGAAGAGGACAGTGGAACCTCCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCACTG
  GCGAAGGCGGCTGTCTGGTCTGCAACTGACGCTGAGGCTCGAAAGCATGGGTAGCGAACAGGATTAGATACCCTG
20  GTAGTCCATGCCGTAAACGATGAGTGCTAGGTGTTGGAGGGTTTCCGCCCTTCAGTGCCGCAGCTAACGCATTAA
  GCACTCCGCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCGCACAAGCGGTGGAGC
  ATGTGGTTTAATTCGAAGCTACGCGAAGAACCCTTACCAGGTCTTGACATCTTGCGCCAACCCCTAGAGATAGGGCG
  TTTCTTTCGGGAACGCAATGACAGGTGGTGATGGTCGTCGTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCC
  GCAACGAGCGCAACCCCTTGTTACTAGTTGCCAGCATTCAGTTGGGCACTCTAGTGAGACTGCCGGTGACAAACCG
25  GAGGAAGGTGGGGACGACGTCAGATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGACGGTAC
  AACGAGTCGCGAACTCGCGAGGGCAAGCTAATCTCTTAAACCGTTCTCAGTTCGGACTGCAGGCTGCAACTCGC
  CTGACGAAGTCCGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACAC
  CGCCCGTCACACCATGAGAGTTTGCAACACCCAAAGTCGTTGGGGTAACCCCTTCGGGGAGCTAGCCCGCTAAGGT
  GGGGCAGATGATTAGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCTGGATCACCTCCT

```

25 > SEQ ID NO: 91|NR_028816.1|*Turicibacter sanguinis* strain MOL361 16S ribosomal RNA gene, complete sequence

```

30  AGAGTTTGATCATGGCTCAGGATGAACGCTGGCGGCGTGCCCTAATACATGCAAGTCGAGCGAACCACTTCGGTGG
  TGAGCGGCGAACGGGTGAGTAACACGTAGGTTATCTGCCCATCAGACGGGGACAACGATTGGAAACGATCGCTAA
  TACCGGATAGGACGAAAGTTTAAAGGTGCTTCGGCACCACTGATGGATGAGCCTGCGGCGCATTAGCTAGTTGGT
  AGGGTAAAGGCCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGAACGGCCACACTGGGACTGAGACACG
  GCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGCGAAAGCCTGACCGAGCAACGCCGCGTG
  AATGATGAAGGCCTTCGGGTTGTAAAAATCTGTTATAAGGGAAGAAATGGCTCTAGTAGGAAATGGCTAGAGTGTG
  ACGGTACCTTATGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCGAGCGTTATCC
35  GGAATTATTGGGCGTAAAGAGCGCGCAGGTGGTTGATTAAGTCTGATGTGAAAGCCCACGGCTTAACCGTGGAGG
  GTCATTGGAACTGGTCAACTTGAGTGCAGAAGAGGGAAGTGGAATTCATGTGTAGCGGTGAAATGCGTAGAGA
  TATGGAGGAACACCACTGGCGAAGGCGGCTTCCTGGTCTGTAACCTGACACTGAGGCGCGAAAGCGTGGGGAGCAA
  ACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTGGGGGTGCAACCTCAGTGCTGA
  AGTTAACGCATTAAAGCACTCCGCTGGGGAGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGACCCGC
  ACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATACCAGTGACCGT
40  CCTAGAGATAGGATTTTCCCTTCGGGGACAATGGATACAGGTGGTGATGGTTGTGTCGTCAGCTCGTGTGTCGAGA
  TGTTGGGTTAAGTCCC GCAACGAGCGCAACCCCTGTCGTTAGTTGCCAGCATTCAGTTGGGGACTCTAACGAGAC
  TGCCAGTGACAACTGGAGGAAGGTGGGGATGACGTCAAAATCATCATGCCCTTATGACCTGGGCTACACACGTG
  CTACAATGGTTGGTACAAAGAGAAGCGAAGCGGTGACGTGGAGCAAACCTCATAAAGCCAATCTCAGTTCGGATT
  GTAGGCTGCAACTCGCCTACATGAAGTTGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGTGAATACGTTCC
45  CGGGTCTTGATACACACCGCCGTCACACCACGAGAGTTTACAACACCCGAAGTCAGTGGCCTAACCGCAAGGAGG
  GAGCTGCCTAAGGTGGGGTAGATGATTGGGGTGAAGTCGTAACAAGGTATCCCTACCGGAAGGTGGGGTTGGATC
  ACCTCCTT

```

50 > SEQ ID NO: 92|NR_042832.1|*Roseburia faecis* strain M72/1 16S ribosomal RNA gene, partial sequence

```

  GATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGAAGCACTCTATTTGATTTTCTTCGGAAATGAAGA
  TTTTGTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATACAGGGGGATAACAGTTGGAAAC
  GACTGCTAATACCGCATAAGCGCACAGGATCGCATGATCCGGTGTGAAAACTCCGGTGGTATGAGATGGACCCG
  CGTCTGATTAGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGC

```


CACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCC
 TGATGCAGCGACGCCCGGTGAGCGAAGAAGTATTTCCGTATGTAAAGCTCTATCAGCAGGGAAGAAGAATGACGG
 TACCTGACTAAGAAGCACC GGCTAAATACGTGCCAGCAGCCGCGTAAATACGTATGGTGCAAGCGTTATCCGGAT
 5 TTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAAGCCCCGGGGCTCAACCCCGGTACTGCA
 TTGGAAACTGTCTGCTACTAGAGTGTCTGGAGGGGTAAAGTGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAG
 GAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAG
 GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCTGGGGAGCATTTGCTCTTCGGTGCCGCA
 GCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCA
 CAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTGACATCCCGATGACAGAG
 10 TATGTAATGTACYTTCTCTTCGGAGCATCGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCTGAGATGT
 TGGGTTAAGTCCCGCAACGAGCGCAACCCCTGTCTTAGTAGCCAGCGGTTTCGGCCGGGCACCTTAGGAGAGACTG
 CCAGGGATAACCTGGAGGAAGCGGGGATGACGTCAAATCATCATGCCCTTATGACTTGGGCTACACACGTGCT
 ACAATGGCGTAAACAAAGGGAAGCGGAGCCGTGAGGCCGAGCAAAATCTCAAAAATAACGTCTCAGTTTCGGACTGT
 AGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCAGATCAGAATGCTGCGGTGAATACGTTCCCG
 15 GGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGAAATGCCCCGAAGTCAGTGACCCAACCGCAAGGAGGGA
 GCTGCCGAAGGCAGGTTTCGATAACTGGGGTG

> SEQ ID NO: 93|NR_043142.1|Flavonifractor plautii strain Prevot S1 16S ribosomal RNA
gene, partial sequence

CGCTGGCGCGTGTCTTAACACATGCAAGTCGAACGGGGTGCTCATGACGGAGGATTTCGTCCAATGGATTGAGTTA
 CCTAGTGGCGGACGGGTGAGTAACCGGTGAGGAACCTGCCCTTGAGAGGGGAATAACACTCCGAAAGGAGTGCTA
 ATACCGCATGAAGCAGTTGGGTGCGATGGCTCTGACTGCCAAAGATTTATCGCTCTGAGATGGCCTCGCGTCTGA
 TTAGCTAGTAGGCGGGTAACGGCCACCTAGGCGACGATCAGTAGCCGACTGAGAGGTTGACCGGCCACATTG
 GGA CTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGCGCAAGCCTGACCCA
 25 GCAACGCCCGGTGAAGGAAGAAGGCTTTTCGGGTGTAACTTCTTTTGTCTGGGGACGAAACAAATGACGGTACCC
 GACGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAAATACGTAGGTGGCAAGCGTTATCCGGATTTACT
 GGGTGTAAAGGGCGTGTAGGCGGGATTGCAAGTCAGATGTGAAAACCTGGGGGCTCAACCTCCAGCCTGCATTTGA
 AACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAATTCGGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGA
 ACACCAGTGGCGAAGGCGGATTGCTGGACAGTAACCTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTA
 30 GATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGGTGTGGGGGGTCTGACCCCTCCGTGCCGCAGTTAA
 CACAATAAGTATCCACCTGGGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGC
 GGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGGCTTGACATCCCACTAACGAGGCAGAG
 ATGCGTTAGGTGCCCTTCGGGGAAAGTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCTGAGATGTTG
 GGTAAAGTCCCGCAACGAGCGCAACCCCTTATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGAC
 35 AAAACGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCTTATGTCTGGGCCACACACGTACTACAATGGT
 CGTTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAAATCCCTAAAAGCCATCCAGTTTCGGATTGCAGGCTGAA
 ACCCGCCTGTATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGT
 ACACACCGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTCCGTAGCCTAACCGCAAGGAGGGCGCGGCCGA
 AGGTGGGTTCGATAATTGGGGTGAAGTCGTAACAAGGTAG

40

> SEQ ID NO: 94|NR_044054.1|Blautia wexlerae strain DSM 19850 16S ribosomal RNA
gene, partial sequence

CAAGTCGAACGGGAATTANTTTATTGAAACTTCGGTTCGATTTAATTTAATTCTAGTGGCGGACGGGTGAGTAACG
 CGTGGGTAACTGCCTTATACAGGGGGATAACAGTCAGAAATGGCTGCTAATACCGCATAAGCGCACAGAGCTGC
 45 ATGGCTCAGTGTGAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCTTGTGGTGGGGTAACGGCCC
 ACCAAGGCGACGATCCATAGCCGGCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTA
 CGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTAT
 CTCGGTATGTAACTTCTATCAGCAGGGAAGATAGTGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCC
 AGCAGCCGCGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTGTGGC
 50 AAGTCTGATGTGAAAGGCATGGGCTCAACCTGTGACTGCATTGGAACCTGTCATACTTGAGTGCCGGAGGGGTA
 AGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATAATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGAC
 GGTAAC TGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGA
 TGAATAACTAGGTGTCTGGGTGGCAAAGCCATTTCGGTGCCGTGCGAAACGCAGTAAGTATTCCACCTGGGGAGTAC
 GTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCA
 55 ACGCGAAGAACCCTTACCAAGTCTTGACATCCGCTGACCGATCCTTAACCGGATCTTTCTTCGGGACAGGCGAG
 ACAGGTGGTGCATGGTTGTCTGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTAT

CCTCAGTAGCCAGCATTTTAAGGTGGGCACTCTGGGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACG
TCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGATTGTG
AGATGGAGCAAATCCCCAAAATAACGTCCCAGTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCG
CTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGA
5 GTACAGTAACGCCCCGAAGTCAGTGACCTAACTGCAAGAAGGAGCTGCCGAAGGCGGGACCGATGACTGGGGTGAA
GTCGTAACAAGGT

> SEQ ID NO: 95|NR_027558.1|Anaerotruncus colihominis strain WAL 14565 16S

ribosomal RNA gene, partial sequence

10 AACGGAGCTTACGTTTTGAAGTTTTTCGGATGGATGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGC
AACCTGCCTTTTCAGAGGGGGATAACAGCCGGAAACGGCTGCTAATACCGCATGATGTTGCGGGGGCACATGCCCC
TGCAACCAAAGGAGCAATCCGCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCACCAA
AGCGACGATCGGTAGCCGGACTGAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGA
GGCAGCAGTGGGGGATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGAAGACGGTCTTCGG
15 ATTGTAAACCTCTGTCTTTGGGGAAGAAAATGACGGTACCCAAAAGAGGAAGCTCCGGCTAACTACGTGCCAGCAG
CCGCGGTAATACGTAGGGAGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTA
GAATGTTAAATCCATCGGCTCAACCGGTGGCTGCGTTCTAACTGCCGTTCTTGAGTGAAGTAGAGCGAGCGGA
ATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCGTGCTGGGCTTTAAG
TGACGCTGAGGCTCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGATTA
20 CTAGGTGTGGGGGACTGACCCCTTCCGTGCCGAGTTAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCA
AGGTTGAAACTCAAAGGAATTGACGGGGGCGGCACAAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGTCTTGACATCGGCGTAATAGCCTAGAGAGTAGGTGAAGCCCTTCGGGGCATCCAGACAGGTG
GTGCATGGTTGTCTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTATTATTAGT
TGCTACGCAAGAGCACTCTAATGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATG
25 CCCCCTTATGACCTGGGCTACACACGTACTACAATGGCACTAAAACAGAGGGCGGCGACACCGCGAGGTGAAGCGA
ATCCCAGAAAAAGTGTCTCAGTTCAGATTGCAGGCTGCAACCCGCTGCATGAAGTCGGAATTGCTAGTAATCGC
GGATCAGCATGCCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCCGTCACACCATGGGAGTCCGGGTAAC
ACCCGAAGCCAGTAG

30 > SEQ ID NO: 96|NR_116747.1|Ruminococcus faecis strain Eg2 16S ribosomal RNA gene,
partial sequence

ATGCAAGTCGAACGAAGCACCTTGATTTGATTCTTCGGATGAAGATCTTGGTGAAGTGGCGGACGGGTGAGT
AACCGTGGGTAACTGCCTCATACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACAGCA
CCGCATGGTGCAGGGGTAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACG
35 GCCTACCAAGCCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAACT
CCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAGCGATGAA
GTATTTCCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAAATGACGGTACCTGACTAAGAAGCACC GGCTAAATACG
TGCCAGCAGCCGCGGTAATACGTATGGTGAAGCGTTATCCGGATTACTGGGTGTAAAGGGAGCGTAGACGGAG
TGGCAAGTCTGATGTGAAAACCCGGGGCTCAACCCCGGGACTGCATTGGAACTGTCAATCTAGAGTACCGGAGA
40 GGTAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACT
GGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAA
ACGATGACTACTAGGTGTGCGGCAGCAAAGCTGTTGCGTGCCGAGCAAACGCAATAAGTAGTCCACCTGGGGAG
TACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAA
GCAACGCGAAGAACCTTACCTGCTCTTGACATCTCCCTGACCGGCAAGTAATGTTGCCCTTCCTTCGGGACAGGG
45 ATGACAGGTGGTGCATGGTTGTCTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCC
TATCTTTAGTAGCCAGCGGTTTGGCCGGGCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATG
ACGTCAAATCATCATGCCCCCTTATGAGCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCAGAACC
GCGAGGTGAGCAAAATCCCCAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAA
TCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATG
50 GGAGTCAGTAACGCCCCGAAGTCAGTGACCCAACCGTAAGGAGGAGCTGCCGAAG

> SEQ ID NO: 97|NR_028883.1|Dorea longicatena strain 111-35 16S ribosomal RNA gene,
partial sequence

TAACGCGTGGGTAACTGCCTCATACAGGGGGATAACAGTTAGAAAATGACTGCTAATACCGCATAAGACCACGTA
 CCGCATGGTACAGTGGTAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACG
 GCCTACCAAGCCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAGACT
 CCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCGTGAAGGATGAA
 5 GTATTTTCGGTATGTAACTTCTATCAGCAGGGAAGAAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACG
 TGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTACTGGGTGTAAAGGGAGCGTAGACGGCA
 CGGCAAGCCAGATGTGAAAAGCCCCGGGGCTCAACCCCGGGACTGCATTGGAACTGCTGAGCTAGAGTGTCCGAG
 AGGCAAGTGGAAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGC
 TGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCTGGTAGTCCACGCCGTA
 10 AACGATGACTGCTAGGTGTGGGTGGCAAAGCCATTCCGGTGCCGACGTAACGCAATAAGCAGTCCACCTGGGGA
 GTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCTGA
 AGCAACGCGAAGAACCTTACCTGATCTTGACATCCCGATGACCGCTTCGTAATGGAAGTTTTTCTTCGGAACATCA
 GGTGACAGGTGGTGCATGGTTGTCTGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCC
 CTATCTTCAGTAGCCAGCAGGTTAAGCTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGA
 15 TGACGTCAAATCATCATGCCCCTTATGACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAAC
 TCGCGAGGGTAAGCAAATCTCAAAAAATAACGTCTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGG
 AATCGCTAGTAATCGCAGATCAGAATGCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCA
 TGGGAGTCATAACGCCCCGAAGTCAGTGACCCAACCGTAAG

20 > SEQ ID NO: 98|NR_029164.1|Clostridium innocuum strain B-3 16S ribosomal RNA gene,
 partial sequence

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTCTTCAG
 GAAGCTTGCTTCCAAAAAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
 ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACGGAGCGCATGCTCTGTATATTAAAGCGCCCTTCAAGGC
 25 GTGAACATGGATGGACCTGCGACGCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGATGATGCGTAGCCG
 GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTAGGGAATTT
 TCGTCAATGGGGGAAACCTGAACGAGCAATGCCGCGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
 AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
 CCAGYAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAAATCATTGGGCGTAAAGGGTGCCTAGGTGGCGTA
 30 CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAACTGGTATGCTGGAGTGCAGAAGAGGG
 CGATGGAATTCATGTGTAGCGGTAAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTTCGCTGGT
 CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAAATAGGATTAGATACCTTAGTCCACGCCGTAAACG
 ATGAGAACTAAGTGTGGAGAAATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCCCTGGGGAGTATGCACGCA
 AGTTNGAACTCAAAGGAATTGACGGGGGCCCCGACAAGCGNTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
 35 GAACCTTACCAGGCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
 ATGGTTGTGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTCTGCATGTTACC
 AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
 GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGACCACAAAGAGCAGCGACTTGGTGACAAGAAGCGAA
 TCTCATAAAGATCGTCTCAGTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
 40 GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGACACACCGCCCGTCAAACCATGGGAGTCAGTAATACC
 CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGA

> SEQ ID NO: 99|NR_104687.1|Blautia hansenii strain JCM 14655 16S ribosomal RNA
 gene, partial sequence

AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTATCATT
 GACTCTTCGGAAGATTTGATATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATACAGG
 GGAATAACAGTTAGAAAATGGCTGCTAATGCCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACTGAGG
 TGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCTTACCAAGCCGACGATCAGTAGCCGG
 CCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATT
 50 GCACAATGGGGGAAACCTGATGCAGCGACGCCGCTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGC
 AGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGG
 GCAAGCGTTATCCGGATTACTGGGTGTAAAGGGAGCGTAGACGGAAAGAGCAAGTCTGATGTGAAAGGCTGGGGC
 TTAACCCCAGGACTGCATTGGAACTGTTTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTG
 AAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACCTGACGTTGAGGCTCGAAA
 55 GCGTGGGGAGCAAACAGGATTAGATACCTGGTAGTCCACGCCGTAAACGATGAATACAGGTGTCGGGGTGCAA
 AGCAGTTCCGTGCCGACGAAACGCAATAAGTATTCACCTGGGGAGTACGTTCCGAAGAATGAACTCAAAGGA

ATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTG
 ACATCTGCCTGACCGTTCCCTTAACCGGAGCTTTCCCTTCGGGACAGGCAAGACAGGTGGTGCATGGTTGTCGTCAG
 CTCGTGTCGTGAGATGTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGTCCGGCTGG
 GCACTCTAGGGGAGACTGCCGGGGGATAACCCGGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTTATGATT
 TGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAAGCGGTGACGCTTAGCAAAATCTCAAAAATAAC
 GTCCCAGTTCCGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCG
 CGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCAACCATGGGAGTCAGTAACGCCCGAAGTCAGTGAC
 CCAACCTTATGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAAC

10 > SEQ ID NO: 100|NR_112933.1|Bacteroides cellulosilyticus strain JCM 15632 16S
 ribosomal RNA gene, partial sequence

AGAGTTTGATCCTGGCTCAGGATGAACGCTAGCTACAGGCTTAACACATGCAAGTCGAGGGGCAGCATGACCTAG
 CAATAGGTTGATGGCGACCGGCGCACGGGTGAGTAACACGTATCCAACCTACCGGTTATTCGGGATAGCCTTTC
 GAAAGAAAGATTAATACCGGATAGTATAACGAGAAGGCATCTTTTTGTTATTAAAGAAATTCGATAACCGATGGG
 GATGCGTTCCATTAGTTTGTGGCGGGGTAACGGCCACCAAGACATCGATGGATAGGGGTTCTGAGAGGAAGGT
 CCCCCACATTGGAAGTGAACACGGTCCAACTCCTACGGGAGGCAGCAGTGAGGAATATTGGTCAATGGACGAG
 AGTCTGAACCAAGCAAGTAGCGTGAAGGATGACTGCCCTATGGGTTGTAACTTCTTTTATATGGGAATAAAGTG
 AGCCACGTGTGGCTTTTGTATGTACCATACGAATAAGGATCGGCTAACTCCGTGCCAGCAGCCGCGTAAATACG
 GAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACTATTAAGTCAGCTGTGAAAGTT
 TGCGGCTCAACCGTAAATTTGCAGTTGATACTGGTCTGCTTGGAGTGCAGTAGAGGTAGGCGGAATTCGTGGTGTG
 GCGGTGAAATGCTTAGATATCACGAAGAAGTCCGATTGCGAAGGCAGCTTACTGGACTGTAAGTACGCTGATGC
 TCGAAAGTGTGGGTATCAAACAGGATTAGATACCCCTGGTAGTCCACACAGTAAACGATGAATACTCGCTGTTTGC
 GATATACAGCAAGCGGCCAAGCGAAAGCATTAAGTATTCACCTGGGGAGTACGCCGGCAACGGTGAAACTCAAA
 GGAATTGACGGGGGCCCCGCACAAGCGGAGGAACATGTGGTTTAAATTCGATGATACGCGAGGAACCTTACCCGGGC
 TTAAATTGCATCTGAATAATTTGGAACAGATTAGCCGCAAGGCAGATGTGAAGGTGCTGCATGGTTGTCGTCAG
 CTCGTGCCGTGAGGTGTGCGCTTAAGTGCCATAACGAGCGCAACCCCTATCTTTAGTTACTAACAGGTCATGCTG
 AGGACTCTAGAGAGACTGCCGTGTAAGATGTGAGGAAGGTGGGGATGACGTCAAATCAGCACGGCCCTTACGTC
 CGGGGCTACACACGTGTTACAATGGGGGGTACAGAAGGCAGCTACACAGCGATGTGATGCTAATCCCAAAAGCCT
 CTCTCAGTTCCGATTGGAGTCTGCAACCCGACTCCATGAAGCTGGATTTCGCTAGTAATCGCGCATCAGCCACGGC
 GCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCAAGCCATGAAAGCCGGGGGTACCTGAAGTCCGTAA
 CCGCAAGGAGCGGCCTAGGGTAAACTGGTAATTGGGGCTAAGTCGTA

> SEQ ID NO: 101|NR_112940.1|Bacteroides ovatus strain JCM 5824 16S ribosomal RNA
 gene, partial sequence

GGCTCAGGATGAACGCTAGCTACAGGCTTAACACATGCAAGTCGAGGGGCAGCATTTTGTGTTGCTGCAAACTG
 AAGATGGCGACCGGCGCACGGGTGAGTAACACGTATCCAACCTGCCGATAACTCCGGAATAGCCTTTCGAAAGAA
 AGATTAATACCGGATAGCATACGAATATCGCATGATATTTTATTAAAGAAATTCGGTTATCGATGGGGATGCGT
 TCCATTAGTTTGTGGCGGGGTAACGGCCACCAAGACTACGATGGATAGGGGTTCTGAGAGGAAGGTCCCCAC
 ATTGGAAGTGAACACGGTCCAACTCCTACGGGAGGCAGCAGTGAGGAATATTGGTCAATGGGCGAGAGCCTGA
 ACCAGCCAAGTAGCGTGAAGGATGAAGGCTCTATGGGTGCTAACTTCTTTTATATGGGAATAAAGTTTCCACG
 TGTGGAATTTTGTATGTACCATATGAATAAGGATCGGCTAACTCCGTGCCAGCAGCCGCGTAAATACGGAGGATC
 CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGCTGGATTGTTAAGTCAGTTGTGAAAGTTTGC GGCT
 CAACCGTAAATTTGCAGTTGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGGTGTAGCGGTGA
 AATGCTTAGATATCACGAAGAAGTCCGATTGCGAAGGCAGCTCACTAGACTGTTACTGACACTGATGCTCGAAAG
 TGTGGGTATCAAACAGGATTAGATACCCCTGGTAGTCCACACAGTAAACGATGAATACTCGCTGTTTGCATATAC
 AGTAAGCGGCCAAGCGAAAGCATTAAGTATTCACCTGGGGAGTACGCCGGCAACGGTGAAACTCAAAGGAATTG
 ACGGGGGCCCCGCACAAGCGGAGGAACATGTGGTTTAAATTCGATGATACGCGAGGAACCTTACCCGGGCTTAAATT
 GCAACAGAATATATTGGAACAGTATAGCCGTAAGGCTGTTGTGAAGGTGCTGCATGGTTGTCGTCAGCTCGTGC
 CGTGAGGTGTGCGCTTAAGTGCCATAACGAGCGCAACCCCTTATCTTTAGTTACTAACAGGTTATGCTGAGGACTC
 TAGAGAGACTGCCGTGTAAGATGTGAGGAAGGTGGGGATGACGTCAAATCAGCACGGCCCTTACGTCCGGGGCT
 ACACACGTGTTACAATGGGGGGTACAGAAGGCAGCTACCTGGCGACAGGATGCTAATCCCAAAACCTCTCTCAG
 TTCGGATCGAAGTCTGCAACCCGACTTCGTGAAGCTGGATTTCGCTAGTAATCGCGCATCAGCCATGGCGCGGTGA
 ATACGTTCCCGGGCCTTGTACACACCGCCCGTCAAGCCATGAAAGCCGGGGGTACCTGAAGTACGTAACCGCAAG
 GAGCGTCTAGGGTAAACTGGTAATTGGGGCTA

> SEQ ID NO: 102|NR_117563.1|Eubacterium fissicatena 16S ribosomal RNA gene, partial sequence

5 TAGAGTTTGGATCCTGGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCGCTTTACTT
AGATTTCTTCGGATTGAAGAGTTTTTGGGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATAC
AGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACAGTACCGCATGGTACAGTGGGAAAAACTC
CGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAAT
ATTGCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCCGTATGTAACTTCTATC
10 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTTATGTAAGTCTGATGTGAAAACCCGG
GGCTCAACCCCGGGACTGCATTGGAACTATGTAAC TAGAGTGTCCGAGAGGTAAGTGAATTCCTAGTGTAGCG
GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATCACTGACGTTGAGGCTCG
AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATAC TAGGTGTCGGGTGG
CAAAGCCATTCCGGTGCCGCAGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTCCGAAGAATGAAACTCAAAA
15 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCCGAAGCAACGCGAAGAACCTTACCTGCTC
TTGACATCCCCTGACCGGCGTGTAATGGCGCCTTCCCTTCGGGGCAGTGGAGACAGGTGGTGCATGGTTGTCGT
CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCTTTAGTAGCCAGCGGTTTGGC
CGGGCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
AGCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCAATACCGCGAGGTTGAGCAAATCCCAAAAAT
20 AACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCGAAGTCAGT
GACCAACCCGTAAGGAGGGAGCTGCCGAAGGCGGGATCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
GGAAGGTGCGGCTGGATCACCTCCTT

25 > SEQ ID NO: 103|NR_104700.1|Blautia coccoides strain JCM 1395 16S ribosomal RNA gene, partial sequence

AGAGTTTGGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCGCTAAGACAG
ATTTCTTCGGATTGAAGTCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATACAGG
30 GGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAACTCCGG
TGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAAACGGGCCACCAAGGCGACGATCAGTAGCCGG
CCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAAGACTCCTACGGGAGGCAGCAGTGGGGAATATT
GCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGC
AGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGG
35 GCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCTGGGGC
TTAACCCCAAGGACTGCATTGGAACTGTTGTTCTAGAGTGCCGAGAGGTAAGCGGAATTCTAGTGTAGCGGTG
AAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGACGTTGAGGCTCGAAA
GCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATAC TAGGTGTCGGGTGGCAA
AGCCATTCCGTGCCGCAGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTCCGAAGAATGAAACTCAAAGGA
ATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCCGAAGCAACGCGAAGAACCTTACCAAGTCTTG
40 ACATCCCTCTGACCGTCCCGTAACGGGGGCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTCGTGAG
CTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCACATGATGGTG
GGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTTATGAT
TTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCGAATCCCAAAAATAA
CGTCCCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCC
45 GCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGA
CCTAACCGAAAGGAAGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAACC

> SEQ ID NO: 104|NR_109014.1|Blautia faecis strain M25 16S ribosomal RNA gene, partial sequence

50 ATAACAGCCAGAAATGACTGCTAATACCGCATAAGCGCACAGAACCGCATGGTTCGGTGTGAAAAACTCCGGTGG
TATAAGATGGACCCGCGTTGGATTAGCTAGTTGGCAGGGCAGCGGCCCTACCAAGGCGACGATCCATAGCCGGCCT
GAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCA
CAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGCAGG
GAAGATAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCA

AGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGCAGCAAGTCTGATGTGAAAGGCAGGGGCTTA
 ACCCCTGGACTGCATTGGAACTGCTGTGCTTGAGTGCCGGAGGGGTAAGCGGAATTCCTAGTGTAGCGGTGAAA
 TGGTAGATATTAGGAGGAACACCACTGGCGAAGGCGGCTTACTGGACGGTAACGACGTTGAGGCTCGAAAGCG
 TGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCAGGGAGCACAGC
 5 TCTTTGGTGCCGCCGCAAACGCATTAAGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATT
 GACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAATCTTGACA
 TCCCTCTGACCGGGACTTAACCGTCCCTTTCTTCGGGACAGGGGAGACAGGTGGTGCATGGTTGTCGTCAGCTC
 GTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCACGCARTGGTGGGC
 ACTCTGAGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCCCTTATGATTTG
 10 GGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAACCCGCGAGGGTGGGCATAATCTCAAAAATAACGT
 CCCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCCGG
 GTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCG

> SEQ ID NO: 105|NR_036928.1|Clostridium hathewayi strain 1313 16S ribosomal RNA

gene, partial sequence

CTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCGGTTTCAATGAAGTTTTCGGATGGA
 TTTGAAATTGACTTAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTTACACTGGGGGATAACAGTTAGA
 AATGACTGCTAATACCGCATAAGCGCACAGGCCCGCATGGNCTGGTGTGAAAAACTCCGGNGGTGTAAGATGGAC
 CCGCGTCTGATTAGGTAGTTGGNNGGGTAACGGCCCAACGCGGACGATCAGTAGCCGACCTGAGAGGGTGACC
 20 GGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAA
 GCCTGATCCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGAC
 GGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGG
 ATTTACTGGGTGTAAAGGGAGCGTAGACGGTTTAGCAAGTCTGAAGTGAAAGCCCCGGGGCTCAACCCCGGTACTG
 CTTTGGAACTGTTAGACTTGAGTGCAGGAGAGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATT
 25 AGGAGGAACACCACTGGCGAAGGCGGCTTACTGGACTGTAACGACGTTGAGGCTCGAAAGCGTGGGGAGCAAAC
 AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGGGGCAAAGCCCTTCGGTGCCG
 CCGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCG
 CACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTGACATCCCCTGAAAA
 CACNTTAACCGTGATCCCTCTTCGGAGCAGTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCTGTGAGAT
 30 GTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCGAGTAGAGTCGGGCACTCTGGGGAGA
 CTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGT
 GCTACAATGGCGTAAACAAAGGGAGGCAAGGAGCGATCTGGAGCAAAACCCCAAAAATAACGTCTCAGTTCGGAT
 TGCAGGCTGCAACTCGCTGCATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTT
 CCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCGAAGTCAGTGACCAACCGAAAGGAG
 35 GGAGCT

> SEQ ID NO: 106|NR_113270.1|Blautia producta strain JCM 1471 16S ribosomal RNA

gene, partial sequence

AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTAAGACGG
 40 ATTTCTTCGGATTGAAGTCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATAACAGG
 GGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAACTCCGG
 TGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGAGGGGTAAACGGCCACCAAGGCGACGATCAGTAGCCGG
 CCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCACAGTCTCTACGGGAGGCAGCAGTGGGGAATATT
 GCACAATGGGGGAAACCCGTATGCAGCGACGCCGCTGAAGGAAGAAGTATCTCGGTATGTAACTTCTATCAGC
 45 AGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGG
 GCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGGAAGAGCAAGTCTGATGTGAAAGGCTGGGGC
 TTAACCCCAAGGACTGCATTGGAACTGTTGTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTG
 AAATGCGTAGATATTAGGAGGAACACCACTGGCGAAGGCGGCTTACTGGACGGTAACGACGTTGAGGCTCGAAA
 GCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGTGGCAA
 50 AGCCATTCGGTGCCGACGAAACGCAATAAGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGA
 ATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTTG
 ACATCCCTCTGACCGTCCCGTAACGGGGACTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTCGTCAG
 CTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCACATGATGGTG
 GGCACCTCTAGGGAGACTGCCGGGGATAACCCGGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCCCTTATGAT
 55 TTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCGAATCCCAAAAATAA
 CGTCCCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCC

> SEQ ID NO: 107|NR_104799.1|Anaerostipes hadrus strain DSM 3319 16S ribosomal RNA

TGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGAAGCTGCTTAACTGATCTTCTTCGGA
ATTGACGTTTTGTAGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCCTGTACAGGGGGATAACAG
TCAGAAATGACTGCTAATACCGCATAAGACCACAGCACC GCATGGTGCAGGGGTAAAAACTCCGGTGGTACAGGA
TGGACCCGCGTCTGATTAGCTGGTTGGTGAGGTAACGGCTCACCAAGGCAGCATCAGTAGCCGGCTTGAGAGAG
TGAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGG
GGAAACCCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATCTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAA
ATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTA
TCCGGAATTTACTGGGTGTAAAGGGTGCGTAGGTGGTATGGCAAGTCAGAAGTGAAAAACCCAGGGCTTAACTCTGG
GACTGCTTTTTGAAACTGTCTAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAG
ATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGACTGAAACTGACACTGAGGCACGAAAGCGTGGGGAG
CAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCTGGGGCCCGTAGAGGCTTCGG
TGCCGCAGCCAAACGCAGTAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGG
ACCCGCACAAGCGGTGGAGCATGTGGTTTAATTTCGAAGCAACGCGAAGAACCCTTACCTGGTCTTGACATCCTTCT
GACCGGTCTTTAACC GGACCTTTCTTCGGGACAGGAGAGACAGGTGGTGCATGGTTGTCTGTAGCTCGTGTCTGT
GAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTTAGTAGCCAGCATTTCTAGTGGGCAGCTCTAGA
GAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGACGACGCTCAAAATCATCATGCCCTTATGACCACGGCTACAC
ACGTGCTACAATTGGCGTAAACAGAGAGGGAAGCAGCCTCGTGAGAGTGAGCAAAATCCCCAAAAATAACGCTCTCAGTTT
GGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAAATGTCGCGGTGAATAC
GTTCCCGGGTCTTGTTACACACCGCCCGTACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACCCCAACCGTAA
GGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAGGTAGCCGTATCGGAAGGTGCGGCT
GGATCACCTCCTTTT

RNA gene, partial sequence

GTTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGC'TAACACATGCAAGTCGAGCGAAGCGCTTTACT'TAGAT
 TTCTTCGGATTGAAGAGT'TTTGCGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTTGCCTCATACAGG
 GGATAACAGT'TAGAAATGACTGCTAATACCGCATAAGACCACAGTACCGCATGGTACAGTGGGAAAAAACTCCGGT
 GGTATGAGATGGACCCGCGTCTGATTAGCTAGT'TGGTAAGGTAACGGCTTACCAAGGCAACGATCAGTAGCCGAC
 CTGAGAGGGTGACCGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAATATTG
 CACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTTCGGTATGTAAACTTCTATCAGCA
 GGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAATACGTAGGGGG
 CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTTATGTAAAGTCTGATGTGAAAACCCGGGGCT
 CAACCCCGGGACTGCATTGGAAACTATGTAACTAGAGTGTTCGGAGAGGTAAGTGGAAATTCCTAGTGTAGCGGTGA
 AATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATCACTGACGTTGAGGCTCGAAA
 CGTGGGGAGCAAACAGGAT'TAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTCGGGTGGCAAA
 GCCAT'TCGGTGCCGCAGCAAACGCAATAAGTAT'TCCACCTGGGGAGTACGTTTCGCAAGAAATGAAACTCAAAGGAA
 TTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGT'TTAATTCGAAGCAACGCGAAGAACC'TTACCTGCTCTTGA
 CATCCCACTGACCGGCGTGTAATGGCGCCTTCCCTTCGGGGCAGTGGAGACAGGTGGTGCATGGTTGTCGTCAGC
 TCGTGTCTGAGATGTTGGGT'TAAGTCCCGCAACGAGCGCAACCCTTATCTTTAGTAGCCAGCGGTTTGGCCGGG
 CACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAAATCATCATGCCCTTATGAGCA
 GGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCAATACCGCGAGGTTGAGCAAAATCCCAAAAAATAACG
 TCTCAGATTTCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGC
 GGTGAAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCCAAGTCAGTGACC
 CAACCGTAAGGTAGGGAGGCTGGCGAAGGCGGGATCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAA
 GGTGCGGCTGGATCACCTCCTTT

RNA gene, partial sequence

TTTGATCCTGGCTCAGGATGAACGCTGGCGACGTGCTTAACACATGCAAGTCGAGCGAAGCACTTTACTTTGATT
 TCTTCGGAATGAAAGGTTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATACAGGGG
 GATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACAGTACCGCATGGTACAGTGGGAAAACTCCGGTG
 GTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGCCGACC
 5 TGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATATTGC
 ACAATGGGGGAAACCTTGATGCAGCGACGCCGCTGAAGGATGAAGTATTTCCGGTATGTAACTTCTATCAGCAG
 GGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGC
 AAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTTATGTAAAGTCTGATGTGAAAACCCGGGGCTC
 AACCCCGGGACTGCATTGGAACTATGTAAGTACGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGTAGCGGTGAA
 10 ATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATGACTGACGTTGAGGCTCGAAAGC
 GTGGGGAGCAACACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATACAGGTGTCGGGTGGCAAAG
 CCATTCCGGTGCCGACGCAACGCAATAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAAT
 TGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCTGCTCTTGAC
 ATCCCCCTGACCGGCGTGTAATGGTGCCTTTCCTTCGGGACAGGGGAGACAGGTGGTGCATGGTTGTCGTCAGCT
 15 CGTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCTTTAGTAGCCAGCGGTTTGGCCGGGC
 ACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGAGCAG
 GGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCGAAGCCGTGAGGTGGAGCAAATCCCAAAAATAACGT
 CTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCG
 GTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCGAAGTCAGTGACCC
 20 AACCGCAAGGAGGGAGCTGCCGAGGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAG
 GTGCGGCTGGATCACCTCCTTTCT

> SEQ ID NO: 110|NR_113410.1|Clostridium bolteae strain JCM 12243 16S ribosomal RNA gene, partial sequence

TTTTAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACACTGGGGGATAACAGTTAGAA
 ATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACAGTGTGAAAACTCCGGTGGTGTGAGATGGATC
 CGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGCCGACCTGAGAGGGTGACCG
 GCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAAG
 CCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACG
 30 GTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGA
 TTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAGGGCTCAACCCCTGGGACTGC
 TTTGGAACTGTTTTGCTAGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTA
 GGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACA
 GGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGGCAAAGCCCTTCGGTCCCGT
 35 CGCAAACGCAGTAAGCATTCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGC
 ACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCCTCTTGACCGG
 CGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATG
 TTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCAGGTAAAGCTGGGCACCTTAGGGAGAC
 TGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGATTTGGGCTACACACGTG
 40 CTACAATGGCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAATAACGTCCCAGTTCGGACT
 GTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCC
 CGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGCAACGCCCGAAGTCAGTGACCCAACCTCGCAAGAGA
 GGGAGCTGCCGAAGGCGGGGAGGTAACCTGGGGTGAAGTC

> SEQ ID NO: 111|NR_041960.1|Blautia luti strain BInIX 16S ribosomal RNA gene, complete sequence

GTGGGTAACTGCCTTATACAGGGGGATAACAGTCAGAAATGACTGCTAATACCGCATAAGCGCACAGAGCTGCA
 TGGCTCCGGTGTGAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCTAGTTGGTGAGGTAACGGCCC
 50 ACCAAGGCGACGATCCATAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTA
 CGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTAT
 CTCGGTATGTAACTTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCC
 AGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCATGGA
 CAAGTCTGATGTGAAAGGCTGGGGCTCAACCCCGGGACTGCATTGGAACTGCCCGTCTTGAGTGCCGGAGAGGT
 AAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGA
 55 CGGTAACCTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACG
 ATGAATCCTAGGTGTGCGGGAGCAAANNNTTCGGTGCCGCCGCAAACGCATTAAGCATTCACCTGGGGAGTAC

GTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCA
 ACGCGAAGAACCTTACCAAGTCTTGACATCCCTCTGACCGAGTATGTATGGTACTTTTTCCTTCGGGAGAGAGAGG
 AGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCT
 ATCCCCAGTAGCCAGCGGTTTCGGCCGGGCACCTCTGAGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGA
 5 CGTCAAATCATCATGCCCCCTTATGATTTGGGCTACACACGTGCTACAAATGGCGTAAACAAAGGGAAGCAAGCCTG
 CGAGGGTGGGCAAAATCCCCAAAAATAACGTCCCAGTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAAT
 CGCTAGTAATCGCGGATCAGAAATGCCCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGG
 GAGTCAGTAACGCCGAAGTCAGTGACCTAACT

10 > SEQ ID NO: 112|NR_074306.1|Acidaminococcus intestini RyC-MR95 strain RyC-MR95
 16S ribosomal RNA, complete sequence

CTGGCGGCGTGCTTAACACATGCAAGTCGAACGGAGAACTTATTTCCGGTAAGTTCTTAGTGGCGAACGGGTGAGT
 AACCGGTGGGCAACCTGCCCTCCAGTTGGGGACAACATTCCGAAAGGGATGCTAATACCGAATGTCTCCCTCCT
 CCGCATGGAGGAGGGAGGAAAGATGGCCTCTGCTTGCAAGCTATCGCTGGAAGATGGGCCCGCGTCTGATTAGCT
 15 AGTTGGTGGGGTAACGGCTCACCAAGGCGATGATCAGTAGCCGGTCTGAGAGGATGAACGGCCACATTGGGACTG
 AGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACG
 CCGCGTGAGTGATGAAGTCTTCCGATTGTAAACTCTGTTGTTAGGGACGAAAGCACCGTGTTCGAACAGGTCA
 TGGTGTGACGGTACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAAATACGTAGGTGGCAAG
 CGTTGTCCGGAATTATTGGGCGTAAAGAGCATGTAGGCGGGCTTTTAAAGTCTGACGTGAAAATGCGGGGCTTAAC
 20 CCCGTATGGCGTTGGATACTGGAAGTCTTGAGTGCAGGAGAGGAAAAGGGGAATTCCCAGTGTAGCGGTGAAATGC
 GTAGATATTGGGAGGAACACCAGTGGCGAAGGCGCCTTTCTGGACTGTGTCTGACGCTGAGATGCGAAAGCCAGG
 GTAGCAAACGGGATTAGATACCCCGGTAGTCTTGCCGTAAACGATGGATACTAGGTGTAGGAGGTATCGACCCC
 TTCTGTGCCGAGTTAACGCAATAAGTATCCCGCTGGGGACTACGATCGCAAGATTGAAACTCAAAGGAATTGA
 CGGGGGCCCGCACAAGCGGTGGAGTATGTGGTTTAAATTCGACGCAACGCGAAGAACCCTTACCAAGGCTTGACATT
 25 GAGTGAAAGACCTAGAGATAGGTCCCTCCCTTCGGGGACACGAAAACAGGTGGTGCATGGCTGTCGTAGCTCGT
 GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCCTATGTTACCAGCGCGTAAAGGCGGGGAC
 TCATAGGAGACTGCCAGGGATAACTTGGAGGAAGGCGGGGATGACGTCAAGTCATCATGCCCTTATGTCTTGGG
 CTACACACGTACTACAATGGTCCGCAACAAAGGGCAGCGAAACCGCGAGGTGGAGCAAAATCCCAGAAACCCGACC
 CCAGTTCGGATCGTAGGCTGCAACCCGCCTACGTGAAGTTGGAATCGCTAGTAATCGCAGGTGAGCATACTGCGG
 30 TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAAAGTTGGTAACACCCGAAGCCGGTGAGATA
 ACCTTTTAGGAGTCAGCTGTCTAAGGTGGGGCCGATGATTTGGGGTGAAGTCGTAACAAGGTAGC

> SEQ ID NO: 113|NR_074399.1|Ruminococcus albus strain 7 16S ribosomal RNA gene,
 complete sequence

AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCACGCTTAACACATGCAAGTCGAACGAGCGAAAGAGTGCT
 TGCACCTCTCTAGCTAGTGGCGGACGGGTGAGTAACACGTGAGCAATCTGCCTTTTCGGAGAGGGATACCAATTGGA
 AACGATTGTTAATACCTCATAACATAACGAAGCCGCATGACTTTGTTATCAAATGAATTTCCGCCAAAGATGAGC
 TCGCGTCTGATTAGGTAGTTGGTGAGGTAACGGCCACCAAGCCGACGATCAGTAGCCGGACTGAGAGGTTGAAC
 35 GGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAA
 GCCTGATGCAGCGATGCCGCGTGAGGGAAGAAGGTTTTAGGATTGTAAACCTCTGTCTTTGGGGACGATAATGAC
 40 GGTACCCAAGGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGTAAATACGTAGGGAGCGAGCGTTGTCGG
 AATTACTGGGTGTAAAGGGAGCGTAGGCGGGATTGCCAGTCAAGGTGTGAAATTTAGGGGCTTAACCCCTGAACGT
 CACTTGAAACTGTAGTTCTTGAGTGAAGTAGAGGTAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATT
 AGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGGCTTTAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAAC
 45 AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCG
 CAGTTAAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGACCCG
 CACAAGCAGTGGAGTATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGACATCGTACGCATAG
 CATAGAGATATGTGAAATCCCTTCGGGGACGTATAGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTCTGAGA
 TGTGTTGGGGTTAAGTCCCGCAACGAGCGCAACCCCTACTGTTAGTTGCTACGCAAGAGCACTCTAGCAGGACTGCC
 50 GTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTACTAC
 AATGGCTGTAAACAGAGGGAAGCAAAACAGTGATGTGGAGCAAAACCCCTAAAAGCAGTCTTAGTTCCGATTGTAG
 GCTGCAACCCGCCTACATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGG
 CCTTGTACACACCGCCCGTCACGCCATGGGAGTCGGTAACACCCGAAGCCTGTGTTCTAACCAGCAAGGAGGAAGC
 AGTCGAAGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT
 55

> SEQ ID NO: 114|NR_074634.1|*Eubacterium rectale* strain ATCC 33656 16S ribosomal RNA gene, complete sequence

5 AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGAAGCACTTTATTTG
 ATTTCTTCGGGACTGATTATTTTGTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTTGTACA
 10 GGGGGATAACAGTTGGAAACGGCTGCTAATACCGCATAAGCGCACGGCATCGCATGATGCAGTGTGAAAACTCC
 GGTGGTATAAGATGGACCCGCGTTGGATTAGCTAGTTGGTGAGGTAACGGGCCACCAAGGCGACGATCCATAGCC
 GACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATA
 TTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGAGCGAAGAAGTATTTCCGGTATGTAAAGCTCTATCA
 15 GCAGGGAAGATAATGACGGTACCTGACTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTATG
 GTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAAGCCCGGG
 GCTCAACCCCGGTACTGCATTGGAACTGTCTACTAGAGTGTGCGAGGGGTAAAGCGGAATTCCTAGTGTAGCGG
 TGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGCTGAGGCTCGA
 AAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATACCTAGGTGTTGGGAAGC
 20 ATTGCTTCTCGGTGCCGTGCGCAAACGCAGTAAGTATTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAAG
 GAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCT
 TGACATCCTTCTGACCGGTACTTAACCGTACCTTCTCTTCGGAGCAGGAGTGACAGGTGGTGCATGGTTGTCGTC
 AGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTTAGTAGCCAGCGGTTCCGGC
 GGGCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCTTATGA
 25 CTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCAAAGCTGTGAAGCCGAGCAAATCTCAAAAATA
 ACGTCTCAGTTCCGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAATGC
 TGGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTTGGGAATGCCCGAAGCCAGTG
 ACCTAACCGAAAGGAAGGAGCTGTGCAAGGCAGGCTCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCG
 GAAGGTGCGGCTGGATCACCT

25 > SEQ ID NO: 115|NR_074928.1|*Acidaminococcus fermentans* strain DSM 20731 16S ribosomal RNA gene, complete sequence

30 AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGAGAACCTTTCTTCG
 GAATGTTCTTAGTGCGAACGGGTGAGTAACGCGTAGGCAACCTGCCCTCTGGTTGGGGACAACATTCGGAAAGG
 GATGCTAATACCGAATGAGATCCTCTTTCCGCATGGAGAGAGGATGAAAGATGGCCTCTACTTGTAAGCTATCGC
 CAGAAGATGGGCCTGCGTCTGATTAGCTAGTAGGTGAGGTAACGGCTCACCTAGGCGATGATCAGTAGCCGGTCT
 GAGAGGATGAACGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATCTTCCG
 CAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGTGATGAAGGCCCTTCGGGTGTAAAACTCTGTTGTCAGG
 35 GACGAAAGCACCGATCTATAATACATTTTGGTGTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCA
 GCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCATGTAGGCGGGCTTTTA
 AGTCCGACGTGAAAATGCGGGGCTTAACCCGATGCGGCTTGATACTGGAAGTCTTGAGTGCAGGAGAGGAAAG
 GGGAAATTCAGTGTAGCGGTGAAATGCGTAGATATTGGGAGGAACACCAGTGGCGAAGGCGCCTTTCTGGACTG
 TGTCTGACGCTGAGATGCGAAAGCCAGGGTAGCAAACGGGATTAGATACCCCGGTAGTCTGGCCGTAAACGATG
 40 GGTACTAGGTGTAGGAGGTATCGACCCCTTCTGTGCCGGAGTTAACGCAATAAGTACCCCGCTGGGGACTACGA
 TCGCAAGATTGAAACTCAAAGGAATTGACGGGGGGCCGCACAAGCGGTGGAGTATGTGGTTTAATTCGACGCAAC
 GCGAAGAACCTTACCAAGGCTTGACATTGAGTGAAAAGACCCAGAGATGGGTCCCCTTCTTCGGAAGCACGAAAAC
 AGGTGGTGCATGGCTGTGCTCAGCTCGTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCC
 TATGTTACCAGCACGTAATGGTGGGACTCATAGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACGT
 CAAGTCATCATGCCCTTATGTCTTGGGCTACACACGTACTACAATGGTCGGCAACAAAGGGCAGCGAAGCCGCG
 45 AGGCGGAGCCAATCCCAGAAACCCGACCCAGTTTCGGATCGCAGGCTGCAACCCGCCTGCGTGAAGTTGGAATCG
 CTAGTAATCGCAGGTCAGCATACTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTACACCACGAAA
 GTTGGTAACACCCGAAGCCGGTGAGATAACCTTTTAGGAGTCAGCTGTCTAAGGTGGGGCCGATGATTGGGGTGA
 AGTCGTAACAAGGTAGCCGTTTCGAGAACGAGCGGCTGGATCACCT

> SEQ ID NO: 116|NR_114326.1|*Fusicatenibacter saccharivorans* strain HT03-11 16S

50 ribosomal RNA gene, partial sequence

TGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCAGTTAAGAAGATTYTTTCGGATG
 ATTTCTTACTGACTGAGCGGGCGGACGGGTGAGTAACGCGTGGGTGACCTGCCCCATACCGGGGGATAACAGCTGG
 AAACGGCTGCTAATACCGCATAAGCGCACAGAGCTGCATGGCTCGGTGTGAAAACTCCGGTGGTATGGGATGGG
 CCCGCGTCTGATTAGGCAGTTGGCGGGGTAAACGGCCCCACAAACCGACGATCAGTAGCCGGCCTGAGAGGGCGAC

CGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAA
 ACCCTGATGCAGCGACGCCGCTGAGCGAAGAAGTATTTTCGGTATGTAAAGCTCTATCAGCAGGGAAGATAATGA
 CGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGTAACTACGTAGGGGGCAAGCGTTATCCG
 GATTTACTGGGTGTAAAGGGAGCGTAGACGGCAAGGCAAGTCTGATGTGAAAACCCAGGGCTTAACCCCTGGGACT
 5 GCATTGGAAACTGTCTGGCTCGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATAT
 TAGGAAGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAA
 CAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGAGCAAAGCTCTTCGGTGCC
 GCCGCAAACGCATTAAAGCATTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAAGGAATTGACGGGGACCC
 GCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGACATCCCGATGACC
 10 GGCCCGTAACGGGGCCTTCTCTTCGGAGCATTGGAGACAGGTGGTGCATGGTTGTCGTACGCTCGTGTCTGAGA
 TGTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCCTCAGTAGCCAGCAGGTAAAGCTGGGCACCTCTGTGGAG
 ACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGATCTGGGCTACACACG
 TGCTACAATGGCGTAAACAAAGGGAGGCAAGCCGCGAGGTGGAGCAAATCCCAAAAAATAACGTCTCAGTTCGGA
 CTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTT
 15 CCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCCGAAGTCAGTGACCCAACCTTTTA

> SEQ ID NO: 117|NR_102884.1|Ruminococcus champanellensis strain 18P13 16S

ribosomal RNA gene, complete sequence

AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCACGCCTAACACATGCAAGTCGAACGGAGATAAAGACTTC
 20 GGTTTTTATCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTCTGAGAGAGGGATAGCTTCTGGAAA
 CGGATGGTAATACCTCATAACATAGCGGTACCGCATGATACTGCTATCAAAGATTTATCGCTCAGAGATGGGCTC
 GCGTCTGATTAGCTAGATGGTGAGGTAACGGCTCACCATGGCGACGATCAGTAGCCGGACTGAGAGGTTGAACGG
 CCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGC
 CTGATGCAGCGATGCCGCGTGGAGGAAGAAGGTTTTTCGGATTGTAAACTCCTGTCTTAAGGGACGATAATGACGG
 25 TACCTTAGGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTGTCCGGAA
 TTACTGGGTGTAAAGGGAGCGTAGGCGGGATTGCAAGTCAGATGTGAAAACCTATGGGCTTAACCCATAGACTGCA
 TTTGAAACTGTAGTTCTTGAGTGAAGTAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAG
 GAGGAACATCGGTGGCGAAGGCGGCTTACTGGGCTTTTACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAG
 GATTAGATACCCTGGTAGTCCACGCTGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCGCA
 30 GTTAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGGCCCGCA
 CAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAAAACCTTACCAGGTCTTGACATCGAGTGAATGATC
 TAGAGATAGATCAGTCCTTCGGGACACAAAGACAGGTGGTGCATGGTTGTCGTACGCTCGTGTCTGAGATGTTG
 GGTAAAGTCCCGCAACGAGCGCAACCCCTTACCTTTAGTTGCTACGCAAGAGCACTCTAGAGGGACTGCCGTTGAC
 AAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTACTACAATGGC
 35 AATGAACAGAGGAAGCAATACAGTGTGAGGCAAAATCCCAAAAAATTGTCCAGTTTCAAGTTGAGGCTGCA
 ACTCGCTACATGAAGTCGGAATTGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACGTTCCCGGGCCTTGT
 ACACACCGCCCGTCACACCATGGGAGTCGGTAACACCCGAAGCCAGTAGCCTAACCAGGAAGGAGGGCGCTGTGCA
 AGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT

40 > SEQ ID NO: 118|NR_102971.1|Bifidobacterium bifidum S17 strain S17 16S ribosomal
 RNA, complete sequence

TTTTTGTGGAGGGTTTCGATTCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAAACATGCAAGTCGAACGGGATC
 CATCGGGCTTTTGCTTGGTGGTGAGAGTGGCGAACGGGTGAGTAATGCGTGACCGACCTGCCCATGCTCCGGAAT
 45 AGCTCCTGGAAACGGGTGGTAATGCCGGATGTTCCACATGATCGCATGTGATTGTGGGAAAGATTCTATCGGCGT
 GGGATGGGGTCGCGTCTATCAGCTTGTGGTGAGGTAAACGGCTCACCAAGGCTTCGACGGGTAGCCGGCCTGAG
 AGGGCGACCGCCACATTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAA
 TGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGGAGGCCTTCGGGTTGTAAACCTCTTTTGTGGGAG
 CAAGCCTTCGGGTGAGTGTACCTTTCAATAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGG
 50 CGCAAGCGTTATCCGGATTTATTGGGCGTAAAGGGCTCGTAGGCGGCTCGTCGCGTCCGGTGTGAAAGTCCATCG
 CTTAACGGTGGATCTGCGCCGGGTACGGGCGGGCTGGAGTGCAGTAGGGGAGACTGGAATTCCTGGTGTAAACGGT
 GGAATGTGTAGATATCGGGAAGAACACCGATGGCGAAGGCAGGTCTCTGGGCGGTCACTGACGCTGAGGAGCGAA
 AGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGACGCTGGATGTGGGGCACGT
 TCCACGTGTTCCGTGTGCGGAGCTAACCGGTTAAGCGTCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACCTCAA
 55 GAAATTGACGGGGGGCCCGCAAGCGGCGGAGCATGCGGATTAATTCGATGCAACGCGAAGAACCCTTACCTGGGC
 TTGACATGTTCCCGACGACGCCAGAGATGGCGTTTCCCTTCGGGGCGGGTTCACAGGTGGTGCATGGTGTCTGTC
 AGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTCGCCCCGTGTTGCCAGCACGTTATGG

TGGGAACCTCACGGGGGACCGCCGGGGTTAACTCGGAGGAAGGTGGGGATGACGTCAGATCATCATGCCCCCTTACG
TCCAGGGCTTCACGCATGCTACAATGGCCGGTACAGCGGGATGCGACATGGCGACATGGAGCGGATCCCTGAAAA
CCGGTCTCAGTTCGGATCGGAGCCTGCAACCCGGCTCCGTGAAGGCGGAGTCGCTAGTAATCGCGGATCAGCAAC
5 GCCGCGGTGAATGCGTTCCCGGGCCTTGTACACACCGCCCGTCAAGTCATGAAAGTGGGCAGCACCCGAAGCCGG
TGGCCTAACCCCTTGTGGGATGGAGCCGTCTAAGGTGAGGCTCGTGATTGGGACTAAGTCGTAACAAGGTAGCCG
TACCGGAAGGTGCGGCTGGATCACCTCCTTTCT

> SEQ ID NO: 119|NR_102980.1|Megasphaera elsdenii strain DSM 20460 16S ribosomal
RNA gene, complete sequence

10 AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGAGAAGAGATGAGAA
GCTTGCTTCTTATCAATTTCGAGTGGCAAACGGGTGAGTAACGCGTAAGCAACCTGCCCTTCAGATGGGGACAACA
GCTGGAAACGGCTGCTAATACCGAATACGTTCTTTTTGTGCGCATGGCAGAGGGAAGAAAGGGAGGCTCTTCGGAG
CTTTTCGCTGAAGGAGGGGCTTGGCTCTGATTAGCTAGTTGGAGGGGTAAACGGCCACCAAGGCGACGATCAGTAG
15 CCGGTCTGAGAGGATGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
TCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAACGATGACGGCCTTCGGGTTGTAAAGTTCTGT
TATACGGGACGAATGGCGTAGCGGTCAATACCCGTTACGAGTGACGGTACCGTAAGAGAAAGCCACGGCTAACTA
CGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCGCAGGCGG
CGTCGTAAGTCGGTCTTAAAGTGCAGGGCTTAACCCCGTGAGGGGACCGAAACTGCGATGCTAGAGTATCGGAG
20 AGGAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAAGCGGCTTTC
TGGACGACAACCTGACGCTGAGGCGCGAAAGCCAGGGGAGCAAACGGGATTAGATACCCCGGTAGTCTTGCCCGTA
AACGATGGATACTAGGTGTAGGAGGTATCGACCCCTTCTGTGCCGGAGTTAACGCAATAAGTATCCCGCCTGGGG
AGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGGTGGAGTATGTGGTTTAATTCTG
ACGCAACGCGAAGAACCTTACCAAGCCTTGACATTGATTGCTATGGATAGAGATATCCAGTTCTCTTCGGAGGA
25 CAAGAAAACAGGTGGTGCACGGCTGTCTGTCAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAA
CCCCATCTTCTGTTACCAGCGGTTTCGGCCGGGACTCAGGAGAGACTGCCGCAGACAATGCGGAGGAAGGCGGG
GATGACGTCAAGTCATCATGCCCCCTTATGGCTTGGGCTACACACGTACTACAATGGCTCTTAATAGAGGGAAGCG
AAGGAGCGATCCGGAGCAAACCCCAAAAACAGAGTCCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGC
AGGAATCGCTAGTAATCGCAGGTGAGCATACTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACA
30 CCACGAAAGTCATTACACCCGAAGCCGGTGAGGTAAACCTTTTGAGCCAGCCGTGCAAGGTGGGGGCGATGATT
GGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT

> SEQ ID NO: 120|NR_044645.2|Dorea formicigenerans strain ATCC 27755 16S ribosomal
RNA gene, complete sequence

35 TTAAACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACAA
TAAGTTTGATTCTTCGGATGAAGACTTTTGTGACTGAGCGGCGGACGNNNGAGTAACGCGTGGGTAACCTGCCTC
ATACAGGGGGATAACAGYTAGAAATGGCTGCTAATACCGCATAAGACCACAGTACTGCATGGTACAGTGNNNAAA
ACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGAGGTAACGGCCACCNAGCCGACGATCAG
TAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCNNGACTCCTACGGGAGGCAGCAGTGGG
40 GAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCCGTATGTAACTTC
TATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGNGGTAATAC
GTAGGGGNNAGCGTTATCCGGATTTACTGGGTGTAAAGGAGCGTAGACGGCTGTGCAAGTCTGAAGTGAAGG
CATGGGCTCAACCTGTGGACTGCTTTGGAACTGTACAGCTAGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGT
AGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCNTACTGGACGATGACTGACGTTGAGG
45 CTCGAAAGCGTGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTGCG
GTAGCAAAGCTATTTCGGTGCCGCGAGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACT
CAAAGGAATTGACGGGGNCCNGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAANNAACGCAAGAACCTTACCT
GATCTTGACATCCCGATGACCGCTTCGTAATGGAAGYTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTG
TCGTCAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCTTCAGTAGCCAGCATTT
50 AGGATGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTNNAATCATCATGCCCT
TATGACCAGGGCTACACACGTGCTACAATGGCGTAAACAGAGGGAGGCAGAGCCGCGAGGCCGAGCAAATCTCAA
AAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAG
AATGCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGT
CAGTGACCCAACCGAAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGT

> SEQ ID NO: 121|NR_118643.1|Eisenbergiella tayi strain B086562 16S ribosomal RNA gene, partial sequence

```

5  GGTATAACTTAGTGGCGGACGGGTGAGTAACGCGTGGGAAACCTGCCCTGTACCGGGGGATAACACTTAGAAATA
  GGTGCTAATACCGCATAAGCGCACGGAACCGCATGGTTCCGTGTGAAAACTCCGGTGGTACAGGATGGTCCCGC
10 GTCTGATTAGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCAGTAGCCGGCCTGAGAGGGTGAACGGCC
  ACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCT
  GATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTA
  CCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTT
15 ACTGGGTGTAAAGGGAGCGTAGACGGCATGGCAAGCCAGATGTGAAAACCCAGGGCTCAACCTTGGGATTGCATT
  TGGAACTGCCAGGCTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGA
  GGAACACCAGTGGCGAAGGCGGCTTACTGGACTGTAACGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGA
  TTAGATAACCTGGTAGTCCACGCGGTAAACGATGATTGCTAGGTGTAGGTGGGTATGGACCCATCGGTGCCGCAG
  CTAACGCAATAAGCAATCCACCTGGGGAGTACGTTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCAC
20 AAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCCCAATGACGCACC
  TGTAAGAGGTGTTCCCTTCGGGGCATTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGCTGAGATGTT
  GGGTTAAGTCCCGCAACGAGCGCAACCCCTTATTCTTAGTAGCCAGCAGGTAAAGCTGGGCACCTAAGGAGACTG
  CCGGGGATAACCCGGAGGAAGGCGGGGATGACGTCAAAATCATCATGCCCTTATGATTTGGGCTACACACGTGCT
  ACAATGGCGTAAACAAAGGGAAGCGAGACAGTGATGTGGAGCAAATCYCAGAAATAACGTCTCAGTTCGGATTGT
25 AGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGCATGTGCGGTGAATACGTTCCCG
  GGTCTTGACACACCGCCCGTCACACCATGGGAGTTGAAAATGCCCCAAGTCTGTGACCTAACCGAAAGGGAGGA
  GCAGCCGAAGGCAGGTCTGATAACTGGGGTGAAGTCGTAA

```

> SEQ ID NO: 122|NR_118730.1|Clostridium symbiosum strain ATCC 14940 16S ribosomal RNA gene, partial sequence

```

25  AAACATGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCGAT
  TTAACGGAAGTTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTT
  GTACTGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATTGCATGATACAGTGTGAAAA
  ACTCCGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAG
30 TAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCNNACTCCTACGGGAGGCAGCAGTGGG
  GAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTC
  TATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATAC
  GTAGGGGNNAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAGC
  CCGCGNCTCAACTGCGGNNCTGCTTTGAAAACCTGTTAACTGGAGTGTGCGGAGAGGTAAGTGAATTCCTAGTGT
35 AGCGGTGAAATGCGTAGATATTAGGAGGAACACNAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGG
  CTCGAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGG
  GGAGCAAAGCTCTTCGGTGCCGTCGCAAACGCAGTAAGTATTCCACCTGGGGAGTACGTTTCGAAGAATGAAACT
  CAAAGGAATTGACGGGGACCNGCACAAGCGGTGGAGCATGTGGTTTAATTCGAANNAACGCGAAGAACCTTACCA
  GGTCTTGACATCGACTCGACGGGGGAGTAACGTCCCNNTNCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTG
40 TCGTCAGCTCGTGTGCTGAGATGTTGGGTTNAGTCCCGCAACGAGCGCAACCCCTTATTCTAAGTAGCCAGCGGTT
  CGGCCGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAATCATCATGCCCTT
  TATGATCTGGGCTACACACGTGCTACAATGGCGTAAACANAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAA
  AAATAACGTCTCAGTTCGGAAGTGCAGGCTGCAACTCGCCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAG
  AATGTGCGGTGAATACGTTCCCGGGTCTTGATACACACCGNCGTCACACCATGGGAGTCAGTAACGCCCGAAGT
45 CAGTGACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGANAAACNNGG

```

> SEQ ID NO: 123|NR_113243.1|Erysipelatoclostridium ramosum strain JCM 1298 16S ribosomal RNA gene, partial sequence

```

50  AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAACGCGAGCACTTGTGCT
  CGAGTGGCGAACGGGTGAGTAATACATAAGTAACCTGCCCTAGACAGGGGGATAACTATTGGAAACGATAGCTAA
  GACCGCATAGGTACGGACACTGCATGGTGACCGTATTAAGTGCCTCAAAGCACTGGTAGAGGATGGACTTATG
  GCGCATTAGCTGGTTGGCGGGGTAAACGGCCCAACGAGCGACGATGCGTAGCCGACCTGAGAGGGTGACCGGCA
  CACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGGGGAAACCCTG
  ACCGAGCAACGCCGCGTGAAGGAAGAAGTTTTTCGGATTGTAAACTTCTGTTATAAAGGAAGAAGCGCGGTACA
  GGAAATGGTAGCCGAGTGACGGTACTTTATTAGAAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT

```

AGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGAGGGAGCAGGCGGCAGCAAGGGTCTGTGGTGAAAGCCT
 GAAGCTTAACTTCAGTAAGCCATAGAAACCAGGCAGCTAGAGTGCAGGAGAGGATCGTGGAATTCATGTGTAGC
 GGTGAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGACGATCTGGCCTGCAACTGACGCTCAGTCCC
 5 GAAAGCGTGGGGAGCAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACGATGAGTACTAAGTGTTGGATG
 TCAAAGTTTCAGTGCTGCAGTTAACGCAATAAGTACTCCGCCTGAGTAGTACGTTTCGCAAGAATGAAACTCAAAGG
 AATTGACGGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTT
 GACATACTCATAAAGGCTCCAGAGATGGAGAGATAGCTATATGAGATACAGGTGGTGCATGGTTGTCTGTCAGCTC
 GTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCGTTAGTTACCATCATTAAGTTGGGGAC
 10 TCTAGCGAGACTGCCAGTGACAAGCTGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGG
 CTACACACGTGCTACAATGGATGGTGCAGAGGGAAGCGAAGCCGCGAGGTGAAGCAAAACCCATAAAACCATTCT
 CAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGTTGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGT
 GAATACGTTCTCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTGATAACACCCGAAGCCGGTGGCCTAA
 CCGCAAGGAAGGAGCTGTCTAAGGTGGGATTGATGATTGGGGTGAAGTCGTAACAAGGTAAAC

15 >SEQ ID NO: 124 |PROKKA_00507 16S ribosomal RNA gene | VE202-7

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 GGAAGTTTTTCGGATGGAATTTGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACACTG
 GGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCCG
 20 GTGGTGTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGCCG
 ACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAATAT
 TGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATCAG
 CAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGTAAACGTAGGG
 GGCAAGCGTTATCCGGATTCACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAGGG
 CTCAACCCTGGGACTGCTTTGGAAACTGTTTTGCTAGAGTGTCCGGAGAGGTAAGTGGAATTCCTAGTGTAGCGGT
 25 GAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCGAA
 AGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGGCA
 AAGCCCTTCGGTGCCGTCGCAAACGCAGTAAGCATTCACCTGGGGAGTACGTTCCGAAGAATGAAACTCAAAGG
 AATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCTT
 GACATCCTCTTGACCGGCGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCTCA
 30 GCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCCTTAGTAGCCAGCAGGTAGAGCT
 GGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGA
 TTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAATA
 ACGTCCCAGTTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGT
 CGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGCAACGCCCGAAGTCAGTG
 35 ACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCGGGGCAGGTAACCTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
 CGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 125 |PROKKA_00709 16S ribosomal RNA gene | VE202-7

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 40 ATGAAGTTTTTCGGATGGATTTTTGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACAC
 TGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTC
 CGGTGGTGTGGGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATC
 45 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAG
 GGCTCAACCCTGGGACTGCTTTGGAAACTGTTTTGCTAGAGTGTCCGGAGAGGTAAGTGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGG
 50 CAAAGCCCTTCGGTGCCGTCGCAAACGCAGTAAGCATTCACCTGGGGAGTACGTTCCGAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCCTTACCAAGTC
 TTGACATCCTCTTGACCGGCGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCTG
 CAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTTATCCTTAGTAGCCAGCAGGTAAAG
 CTGGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 55 GATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAA
 TAACGTCCCAGTTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAAT
 GTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGCAACGCCCGAAGTCAG
 TGACCCAACCTCGCAAGAGAGGGAGCTGCCGAAGGCGGGGCAGGTAACCTGGGGTGAAGTCGTAACAAGGTAGCCGT
 ATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 126 lPROKKA_01766 16S ribosomal RNA gene | VE202-7

5 ATGAGAGTTTGGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 ATGAAGTTTTTCGGATGGATTTTTGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACAC
 TGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTC
 CGGTGGTGTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAAT
 10 ATTGCACAATGGGCGAAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGG
 GGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAGAACCCAGG
 GCTCAACCCTGGGACTGCTTTGGAAACTGTTTTGCTAGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGTAGCGG
 TGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCGA
 15 AAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGCA
 AAGCCCTTCGGTGCCGTCGCAAACGCAGTAAGCATTCCACCTGGGGAGTACGTTGCGAAGAATGAAACTCAAAGG
 AATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGAAGAACCTTACCAAGTCTT
 GACATCCTCTTGACCGGCGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCGTCA
 GCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCAGGTAAAGCT
 20 GGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTTATGA
 TTTGGGCTACACACGTGCTACAATGGCGTAAACAAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAATA
 ACGTCCCGAGTTCCGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGT
 CGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGCAACGCCCGAAGTCAGTG
 ACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCGGGGCAGGTAACCTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
 25 CGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 127 lPROKKA_01779 16S ribosomal RNA gene | VE202-7

ATGAGAGTTTGGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 ATGAAGTTTTTCGGATGGATTTTTGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACAC
 30 TGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTC
 CGGTGGTGTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGCGAAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAGAACCCAG
 35 GGCTCAACCCCTGGGACTGCTTTGGAACACTGTTTTGCTAGAGTGTGCGGAGAGGTAAAGTGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGGCTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGC
 AAAGCCCTTCGGTGCCGTCGCAAACGCAGTAAGCATTCCACCTGGGGAGTACGTTGCGAAGAATGAAACTCAAAG
 GAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGAAGAACCTTACCAAGTCT
 40 TGACATCCTCTTGACCGGCGTGTAACGGCGCCTTCCCTTCGGGGCAGGAGAGACAGGTGGTGCATGGTTGTCGTG
 AGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCAGGTAGAGC
 TGGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTTATG
 ATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAAATA
 ACGTCCCGAGTTCCGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGT
 45 CGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGCAACGCCCGAAGTCAGTG
 ACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCGGGGCAGGTAACCTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
 CGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 128 lPROKKA_05926 16S ribosomal RNA gene | VE202-7

50 ATGAGAGTTTGGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAA
 ATGAAGTTTTTCGGATGGATTTTAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACACT
 GGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCC
 GGTGGTGTGAGATGGATCCGCGTCTGATTAGCCAGTTGGCGGGGTAAACGGCCACCAAAGCGACGATCAGTAGCC
 GACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAATA
 55 TTGCACAATGGGCGAAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATCA
 GCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGG
 GGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAGAACCCAGG

GCTCAACCCCTGGGACTGCTTTGGAACTGTTTTGCTAGAGTGTCGGAGAGGTAAGTGGAAATTCCTAGTGTAGCGG
 TGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGTTGAGGCTCGA
 AAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATGCTAGGTGTTGGGGGGC
 AAAGCCCTTCGGTGCCGTGCGCAAACGCAGTAAGCATTCCACCTGGGGAGTACGTTGCAAGAATGAAACTCAAAG
 5 GAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAAGTCT
 TGACATCCTCTTGACCGGCGGTGTAACGGCGCCTTCCCTTCGGGGCAAGAGAGACAGGTGGTGCATGGTTGTCGTC
 AGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCCTTAGTAGCCAGCAGGTAAAGC
 TGGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
 10 ATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCAAGACAGTGATGTGGAGCAAATCCCAAAATA
 ACGTCCCAGTTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGT
 CGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGCAACGCCCGAAGTCAGTG
 ACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCGGGGAGGTAAGTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
 CGGAAGGTGCGGCTGGATCACCTCCTTT

15 >SEQ ID NO: 129 | PROKKA_01784 16S ribosomal RNA gene

TCAAAGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGCGCCTAACACATGCAAGTCGAACGGAGCTTACGT
 TTTGAAGTTTTTCGGATGGATGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTTCAG
 AGGGGGATAACAGCCGGAACCGGTGCTAATACCGCATGATGTTGCGGGGGCACATGCCCCTGCAACCAAAGGAG
 CAATCCGCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCGACGATCGGTA
 20 GCCGACTGAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGG
 ATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGAAGACGGTCTTCGGATTGTAAACCTCTG
 TCTTTGGGGAAGAAAATGACGGTACCCAAAGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
 AGGGAGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTAGAATGTTAAATCCA
 TCGGCTCAACCGGTGGCTGCGTTCTAAACTGCCGTTCTTGAGTGAAGTAGAGGCAGGCGGAATTCCTAGTGTAGC
 25 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGGGCTTTAACTGACGCTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGG
 ACTGACCCCTTCCGTGCCGAGTTAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAA
 AGGAATTGACGGGGGCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGT
 CTTGACATCGGATGCATAGCCTAGAGATAGGTGAAGCCCTTCGGGGCATCCAGACAGGTGGTGCATGGTTGTCGT
 30 CAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTATTAGTTGCTACGCAAGAGCA
 CTCTAATGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGG
 GCTACACACGTACTACAATGGCACTAAAACAGAGGGCGGCGACACCGCGAGGTGAAGCGAATCCCGAAAAAGTGT
 CTCAGTTTCAAGATTGCAGGCTGCAACCCGCTGTCATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCG
 GTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTACACCATGGGAGTCGGTAACACCCGAAGCCAGTAGCCT
 35 AACC GCAAGGGGGGCGCTGTGCAAGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAG
 GTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 130 | PROKKA_01864 16S ribosomal RNA gene

TCAAAGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGCGCCTAACACATGCAAGTCGAACGGAGCTTACGT
 TTTGAAGTTTTTCGGATGGACGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTTCAG
 AGGGGATAACAGCCGGAACCGGTGCTAATACCGCATGATGTTGCGGGGGCACATGCCCCTGCAACCAAAGGAGC
 AATCCGCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCGACGATCGGTAG
 CCGGACTGAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGGA
 TATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGAGGGAAGACGGTCTTCGGATTGTAAACCTCTGT
 40 CTTTGGGGAAGAAAATGACGGTACCCAAAGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 GGGAGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTAGAATGTTAAATCCAT
 CGGCTCAACCGGTGGCTGCGTTCTAAACTGCCGTTCTTGAGTGAAGTAGAGGCAGGCGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTGCTGGGCTTTAACTGACGCTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGGA
 50 CTGACCCCTTCCGTGCCGAGTTAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAA
 GGAATTGACGGGGGCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTC
 TTGACATCGGATGCATAGCCTAGAGATAGGTGAAGCCCTTCGGGGCATCCAGACAGGTGGTGCATGGTTGTCGTC
 AGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTATTAGTTGCTACGCAAGAGCAC
 TCTAATGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGG
 55 CTACACACGTACTACAATGGCACTAAAACAGAGGGCGGCGACACCGCGAGGTGAAGCGAATCCCGAAAAAGTGT
 TCAGTTTCAAGATTGCAGGCTGCAACCCGCTGTCATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCGG
 TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTACACCATGGGAGTCGGTAACACCCGAAGCCAGTAGCCTA

ACCGCAAGGGGGGCGCTGTCTGAAGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGG
TGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 131 | PROKKA_02671 16S ribosomal RNA gene

5 TCAAAGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGCGCTAACACATGCAAGTCGAACGGAGCTTACGT
TTTGAAGTTTTTCGGATGGATGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTTCAG
AGGGGGATAACAGCCGGAACCGCTGCTAATACCGCATGATGTTGCGGGGGCACATGCCCCGCAACCAAAGGAG
CAATCCGCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCACCAAAGCGACGATCGGTA
10 GCCGGAAGTGAAGGTTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGG
ATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGAGGGAAGACGGTCTTCGGATTGTAAACCTCTG
TCTTTGGGGAAGAAAATGACGGTACCCAAAGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGTAATACGT
AGGGAGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTAGAATGTTAAATCCA
TCGGCTCAACCGGTGGCTGCGTTCTAAACTGCCGTTCTTGAGTGAAGTAGAGGCAGGCGGAATTCCTAGTGTAGC
15 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCCTGCTGGGCTTTAACTGACGCTGAGGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGGG
ACTGACCCCTTCCGTGCCGAGTTAACACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAA
AGGAATTGACGGGGGCGCACAAGCAGTGGAGTATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGT
CTTGACATCGGATGCATAGCCTAGAGATAGGTGAAGCCCTTCGGGGCATCCAGACAGGTGGTGCATGGTTGTCGT
20 CAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTATTAGTTGCTACGCAAGAGCA
CTCTAATGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGG
GCTACACACGTACTACAATGGCACTAAAACAGAGGGCGGCGACACCGCGAGGTGAAGCGAATCCCGAAAAAGTGT
CTCAGTTGAGATTGACGGCTGCAACCCGCGCTGCATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCG
GTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTACACCATGGGAGTCGGTAACACCCGAAGCCAGTAGCCCT
25 AACC CGCAAGGGGGGCGCTGTCTGAAGGTGGGATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAG
GTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 132 | PROKKA_00690 16S ribosomal RNA gene

TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
TTCAGAATCTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
30 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCACCAAGCCGACGATCAGTAG
CCGACCTGAGAGGGTGACCGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGTGGGGAA
TATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT
35 CAGCAGGGAAGAAAAATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGTAATACGTA
TGGTGAAGCGTTATCCGGATTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAGACCCA
GGGCTCAACCCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGAGAGGTAAAGCGGAATTCCTAGTGTAGC
GGTGAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGT
40 GCAAAGCACATCGGTGCCGAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAA
AGGAATTGACGGGGACCCGACAAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCTGGT
CTTGACATCCGGATGACGGGCGAGTAATGTCGCCGTCCTTCGGGGCGTCCGAGACAGGTGGTGCATGGTTGTCG
TCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATCTTCAGTAGCCAGCATATAAG
GTGGGCACTCTGGAGAGACTGCCAGGGGAGAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
45 GGCCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAATCCCAAAA
TAACGTCTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT
GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTACACCATGGGAGTCAGTAACGCCCGAAGCCAG
TGACCAACCTTAGAGGAGGGAGCTGTCTGAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
TCGGAAGGTGCGGCTGGATCACCTCCTTT

50 >SEQ ID NO: 133 | PROKKA_00991 16S ribosomal RNA gene

TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
TTCAGAATCTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
55 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCACCAAGCCGACGATCAGTAG
CCGACCTGAGAGGGTGACCGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTGGGGAA
TATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT

CAGCAGGGAAGAAGATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTA
 TGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAAACCCA
 GGGCTCAACCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGAGAGGTAAGCGGAATTCCTAGTGTAGC
 5 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGT
 GCAAAGCACATCGGTGCCGAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTGCAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGGT
 CTTGACATCCGGATGACGGGCGAGTAATGTCGCCGTCCCTTCGGGGCATCCGAGACAGGTGGTGCATGGTTGTCG
 10 TCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCAGAACGAGCGCAACCCCTTATCTTCAGTAGCCAGCATATAAG
 GTGGGCACTCTGGAGAGACTGCCAGGGAGAACCCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 GGCCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGAAGCGAATCCCAAAAA
 TAACGTCTCAGTTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT
 GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGCCAG
 15 TGACCCAACCTTAGAGGAGGGAGCTGTGCAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 134 | PROKKA_01948 16S ribosomal RNA gene

TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
 TTCAGAATCTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
 20 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAGATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTA
 25 TGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAAACCCA
 GGGCTCAACCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGAGAGGTAAGCGGAATTCCTAGTGTAGC
 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGT
 30 GCAAAGCACATCGGTGCCGAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTGCAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGGT
 CTTGACATCCGGATGACGGGCGAGTAATGTCGCCGTCCCTTCGGGGCATCCGAGACAGGTGGTGCATGGTTGTCG
 TCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCAGAACGAGCGCAACCCCTTATCTTCAGTAGCCAGCATATAAG
 GTGGGCACTCTGGAGAGACTGCCAGGGAGAACCCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 35 GGCCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAATCCCAAAAA
 TAACGTCTCAGTTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT
 GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGCCAG
 TGACCCAACCTTAGAGGAGGGAGCTGTGCAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

40 >SEQ ID NO: 135 | PROKKA_02310 16S ribosomal RNA gene

TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
 TTCAGAATCTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
 45 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAGATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTA
 TGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAAACCCA
 50 GGGCTCAACCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGAGAGGTAAGCGGAATTCCTAGTGTAGC
 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGT
 GCAAAGCACATCGGTGCCGAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTGCAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGGT
 CTTGACATCCGGATGACGGGCGAGTAATGTCGCCGTCCCTTCGGGGCATCCGAGACAGGTGGTGCATGGTTGTCG
 55 TCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCAGAACGAGCGCAACCCCTTATCTTCAGTAGCCAGCATATAAG
 GTGGGCACTCTGGAGAGACTGCCAGGGAGAACCCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 GGCCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGAAGCGAATCCCAAAAA
 TAACGTCTCAGTTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT

GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGCCAG
TGACCCAACTTAGAGGAGGGAGCTGTCTGAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
TCGGAAGGTGCGGCTGGATCACCTCCTTT

5 >SEQ ID NO: 136 | PROKKA_02993 16S ribosomal RNA gene

TACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTT
TTCAGAATCTTTCGGAGGAAGAGGACAGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGCAACCTGCCTCATA
CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTAAAAACT
10 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAAAGGCCACCAAGCCGACGATCAGTAG
CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAA
TATTGCACAATGGGGGAAACCTTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATTTTCGGTATGTAACTTCTAT
CAGCAGGGAAGAAGATGACGGTACCTGAGTAAGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTA
TGGTGC AAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAAACCCA
GGGCTCAACCCTGGGACTGCTTTGGAAACTGCAGATCTGGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAGC
15 GGTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGGCTTACTGGACGGTGACTGACGTTGAGGCTC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCTTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGT
GCAAAGCACATCGGTGCCGACGAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAA
AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGGT
CTTGACATCCGGATGACGGGCGAGTAATGTCGCCGTCCCTTCGGGGCATCCGAGACAGGTGGTGCATGGTTGTCG
20 TCAGCTCGTGTCTGTAGATGTTGGGTTAAGTCCC GCAACGAGCGCAACCCTTATCTTCAGTAGCCAGCATATAAG
GTGGGCACTCTGGAGAGACTGCCAGGGAGAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
GGCGAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAATCCCAAAAA
TAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCAT
GCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGCCAG
25 TGACCCAACTTAGAGGAGGGAGCTGTCTGAAGGCGGGACGGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
TCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 137 | PROKKA_00436 16S ribosomal RNA gene

ATGAGAGTTTGATCCTAGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTTAA
30 CGGAAGTTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTGTAC
TGGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGCATGATACAGTGTGAAAAACTC
CGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGGCCAAACTCCTACGGGAGGCAGCAGTGGGGAAT
ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTTCGGTATGTAAAGCTCTATC
35 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAAGCCCGC
GGCTCAACTGCGGGACTGCTTTGGAAACTGTTAACTGGAGTGTCGGAGAGGTAAGTGGAATTCCTAGTGTAGCG
GTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGGCTCG
AAAGCGTGGGGAGCAAACAGGATTAGATACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAG
40 CAAAGCTCTTCGGTGCCGTCGCAAACGCAGTAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAA
GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTC
TTGACATCGATCCGACGGGGGAGTAACGTCCCCCTTCCCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTGTCGT
CAGCTCGTGTCTGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTC TAAGTAGCCAGCGGTTCCGGC
CGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
45 ATCTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAAAT
AACGTCTCAGTTCGGACTGCAGGCTGCAACTCGCCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGT
GACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
GGAAGGTGCGGCTGGATCACCTCCTTT

50

>SEQ ID NO: 138 | PROKKA_00685 16S ribosomal RNA gene

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTTAA
CGGAAATTTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTGTAC
TGGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGCATGATACAGTGTGAAAAACTC
55 CCGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGGAAAT

ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAAGCCCCGC
 5 GGCTCAACTGCGGGACTGCTTTGGAAACTGTTTAACTGGAGTGTCGGAGAGGTAAGTGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAG
 CAAAGCTCTTCGGTGCCGTGCGAAACGCAGTAAGTATTCACCTGGGGAGTACGTTTCGAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTC
 10 TTGACATCGATCCGACGGGGAGTAACGTCCCCTTCCCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTGTCGT
 CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTCTAAGTAGCCAGCGGTTCCGC
 CGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
 ATCTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAAAT
 AACGTCTCAGTTCGGACTGCAGGCTGCAACTCGCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
 15 TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGT
 GACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
 GGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 139 | PROKKA_01171 16S ribosomal RNA gene

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTTAA
 20 CGGAAGTTTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTTGTAC
 TGGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGCATGATACAGTGTGAAAACTC
 CGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAAT
 25 ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAAGCCCCGC
 GGCTCAACTGCGGGACTGCTTTGGAAACTGTTTAACTGGAGTGTCGGAGAGGTAAGTGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGGCTCG
 30 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAG
 CAAAGCTCTTCGGTGCCGTGCGAAACGCAGTAAGTATTCACCTGGGGAGTACGTTTCGAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTC
 TTGACATCGATCCGACGGGGAGTAACGTCCCCTTCCCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTGTCGT
 CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTCTAAGTAGCCAGCGGTTCCGC
 35 CGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
 ATCTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAAAT
 AACGTCTCAGTTCGGACTGCAGGCTGCAACTCGCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
 TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGT
 GACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
 40 GGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 140 | PROKKA_05969 16S ribosomal RNA gene

ATGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTTAA
 45 CGGAAGTTTTTCGGATGGAAGTTGGATTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTTGTAC
 TGGGGGACAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGCATGATACAGTGTGAAAACTC
 CGGTGGTACAAGATGGACCCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGC
 CGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAAT
 ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCCGGTATGTAAAGCTCTATC
 AGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAG
 50 GGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAAGCCCCGC
 GGCTCAACTGCGGGACTGCTTTGGAAACTGTTTAACTGGAGTGTCGGAGAGGTAAGTGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGGACGATAACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAG
 CAAAGCTCTTCGGTGCCGTGCGAAACGCAGTAAGTATTCACCTGGGGAGTACGTTTCGAAGAATGAAACTCAAA
 GGAATTGACGGGGACCCGCAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTC
 55 TTGACATCGATCCGACGGGGAGTAACGTCCCCTTCCCTTCGGGGCGGAGAAGACAGGTGGTGCATGGTTGTCGT
 CAGCTCGTGTCTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTCTAAGTAGCCAGCGGTTCCGC
 CGGGAACCTCTTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG
 ATCTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAAAT

AACGTCTCAGTTCGGACTGCAGGCTGCAACTCGCCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATG
TCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGT
GACCCAACCGCAAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTATC
GGAAGGTGCGGCTGGATCACCTCCTTT

5

>SEQ ID NO: 141 | PROKKA_00279 16S ribosomal RNA gene

ATCAGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
AGTGGATCTCTTCGGATTGAAACTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCAT
ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
10 TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAACGGCCACCAAGGCGACGATCAGTA
GCCCGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGA
ATATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTA
TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
15 GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAG
CGGTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGGCTTACTGGACGGTAACGTGACGTTGAGGCT
CGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTGCGGT
GGCAAAGCCATTTCGGTGCCGACGAAACGCAATAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAACTCA
AAGGAATTGACGGGGACCCGACAAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
20 TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
GTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGGTGA
AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT
ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCAAATCCCAAA
AATAACGTCCAGTTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGA
25 ATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTC
AGTGACCCAACCTTTTAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
TATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 142 | PROKKA_01221 16S ribosomal RNA gene

ATCAGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
AGTGGATCTCTTCGGATTGAAACTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCAT
ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAACGGCCACCAAGGCGACGATCAGTA
GCCCGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGA
35 ATATTGCACAATGGGGGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTA
TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTTCTAGAGTGCCGGAGAGGTAAGTGAATTCCTAGTGTAG
CGGTGAAATGCGTAGATATTAGGAGGAACACCAAGTGGCGAAGGCGGCTTACTGGACGGTAACGTGACGTTGAGGCT
40 CGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTGCGGT
GGCAAAGCCATTTCGGTGCCGACGAAACGCAATAAGTATTCACCTGGGGAGTACGTTTCGCAAGAATGAACTCA
AAGGAATTGACGGGGACCCGACAAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
GTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGGTGA
45 AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT
ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGGGACAGCGATGTTGAGCAAATCCCAAA
AATAACGTCCAGTTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGA
ATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTC
AGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
50 TATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 143 | PROKKA_02318 16S ribosomal RNA gene

ATCAGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
AGTGGATCTCTTCGGATTGAAGCTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCAT
ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAACGGCCACCAAGGCGACGATCAGTA

55

GCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGA
 ATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTA
 TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAATACGT
 5 AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
 GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAG
 CCGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACGTGACGTTGAGGCT
 CGAAAGCGTGCGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCCGGT
 GGCAAAGCCATTCCGGTGCCGAGCAAACGCAATAAGTATTCACCTGGGGAGTACGTTCCGAAGAATGAAACTCA
 10 AAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
 TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
 GTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTACCCAGCAGGTGA
 AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT
 ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCAAATCCCAAA
 15 AATAACGTCCCAGTTCCGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGA
 ATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTC
 AGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
 TATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 144 | PROKKA_02336 16S ribosomal RNA gene

20 ATCAGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
 AGCGGATCTCTTCGGATTGAAACTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCAT
 ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
 TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAACGGGCCACCAAGGCGACGATCAGTA
 25 GCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGA
 ATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTA
 TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAATACGT
 AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
 GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAG
 30 CCGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACGTGACGTTGAGGCT
 CGAAAGCGTGCGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCCGGT
 GGCAAAGCCATTCCGGTGCCGAGCAAACGCAATAAGTATTCACCTGGGGAGTACGTTCCGAAGAATGAAACTCA
 AAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
 TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
 35 GTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGGTGA
 AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT
 ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCAAATCCCAAA
 AATAACGTCCCAGTTCCGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGA
 ATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTC
 40 AGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
 TATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 145 | PROKKA_04947 16S ribosomal RNA gene

ATCAGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTA
 45 AGTGATCTCTTCGGATTGAAACTTATTTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCAT
 ACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAAC
 TCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGAGGGGTAACGGGCCACCAAGGCGACGATCAGTA
 GCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGA
 ATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACTTCTA
 TCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAATACGT
 50 AGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCT
 GGGGCTTAACCCAGGACTGCATTGGAAGTGTCTTTCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAGTGTAG
 CCGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGACGGTAACGTGACGTTGAGGCT
 CGAAAGCGTGCGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCCGGT
 GGCAAAGCCATTCCGGTGCCGAGCAAACGCAATAAGTATTCACCTGGGGAGTACGTTCCGAAGAATGAAACTCA
 55 AAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAAG
 TCTTGACATCCCTCTGACCGGCCCGTAACGGGGCCTTCCCTTCGGGGCAGAGGAGACAGGTGGTGCATGGTTGTC
 GTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCCAGCAGGTGA
 AGCTGGGCACTCTAGGGAGACTGCCGGGGATAACCCGAGGAAGGCGGGGACGACGTCAAATCATCATGCCCTT

ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGCGATGTTGAGCAAATCCCAAA
 AATAACGTCTTAGTTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGA
 ATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTC
 AGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGCGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCG
 TATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 146 | PROKKA_00208 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTAA
 GTTTGATTCTTCGGATGAAGACTTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACT
 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCTGAAGGATGAAGTATTTTCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
 GGGCTCAACCCCGGGACTGCATTTGGAAGTGTGAGCTAGAGTGTGCGAGAGGCAAGTGGAATTCCTAGTGTAGC
 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGTGGACGATGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCGGGTG
 GCAAAGCCATTCCGTGCCGACGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGAT
 CTTGACATCCCGATGACCGCTTCGTAATGGAAGTTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCG
 TCAGTCTGTCGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAA
 GCTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTA
 TGACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAA
 ATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAA
 TGCTGCGGTGAATACGTTCCCGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCA
 GTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 147 | PROKKA_00340 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTTG
 GAAAGATTCTTCGGATGATTTCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACT
 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCTGAAGGATGAAGTATTTTCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
 GGGCTCAACCCCGGACTGCATTTGGAAGTGTGAGCTAGAGTGTGCGAGAGGCAAGTGGAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGTGACGATGACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCGGGTGG
 CAAAGCCATTCCGGTGCCGACGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAAA
 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGATC
 TTGACATCCCGATGACCGCTTCGTAATGGAAGCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCGT
 CAGCTCGTGTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAAG
 CTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
 GACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAA
 TAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAT
 GCTGCGGTGAATACGTTCCCGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAG
 TGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 CCGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 148 | PROKKA_01031 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTAA
 GTTTGATTCTTCGGATGAAGACTTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACT

CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 5 GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
 GGGCTCAACCCCGGGACTGCATTTGGAAGTGTGAGCTAGAGTGTCCGAGAGGCAAGTGGAAATTCCTAGTGTAGC
 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGTGGACGATGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCCGGTG
 10 GCAAAGCCATTCCGTGCCGCAGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTCCGAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGAT
 CTTGACATCCCGATGACCGCTTCGTAATGGAAGCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCG
 TCAGCTCGTGTCTGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTA
 GCTGGGCACCTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTA
 15 TGACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAA
 ATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGTAGTAATCGCAGATCAGAA
 TGCTGCGGTGAATACGTTCCCGGGTCTTGATACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCA
 GTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

20 >SEQ ID NO: 149 | PROKKA_01840 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCACTTTG
 GAAAGATTCTTCGGATGATTTCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAAACT
 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
 25 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCCGGTATGTAACTTCTAT
 CAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
 GGGCTCAACCCCGGGACTGCATTTGGAAGTGTGAGCTAGAGTGTCCGAGAGGCAAGTGGAAATTCCTAGTGTAGC
 30 GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGTCTGGACGATGACTGACGTTGAGGCTC
 GAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCCGGTG
 GCAAAGCCATTCCGTGCCGCAGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTCCGAAGAATGAACTCAA
 AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGAT
 CTTGACATCCCGATGACCGCTTCGTAATGGAAGCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCG
 35 TCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAA
 GCTGGGCACCTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTA
 TGACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAA
 ATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAA
 TGCTGCGGTGAATACGTTCCCGGGTCTTGATACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCA
 40 GTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 150 | PROKKA_02944 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCGCTTAA
 45 GTTTGATTCTTCGGATGAAGACTTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACTGCCTCATA
 CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAAACT
 CCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
 CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
 TATTGCACAATGGAGGAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCCGGTATGTAACTTCTAT
 50 CAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
 GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
 GGCTCAACCCCGGGACTGCATTTGGAAGTGTGAGCTAGAGTGTCCGAGAGGCAAGTGGAAATTCCTAGTGTAGCG
 GTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGTGGACGATGACTGACGTTGAGGCTCG
 AAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCCGGTGG
 55 CAAAGCCATTCCGTGCCGCAGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTCCGAAGAATGAACTCAAA
 GGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTACCTGATC
 TTGACATCCCGATGACCGCTTCGTAATGGAAGTTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCGT
 CAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAAG

CTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTAT
GACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAGGCAAACCTCGCGAGGGTAAGCAAATCTCAAAAA
TAACGTCTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAT
5 GCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAG
TGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTAT
CGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 151 | PROKKA_04036 16S ribosomal RNA gene

AACGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGAAGCGCTTTG
10 GGAAGATTCTTCGGATGATTTCTTTGTGACTGAGCGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATA
CAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGTAAAACT
CCGGTGGTATGAGATGGACCCGCGTTTGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGACGATCAGTAG
CCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
15 TATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTTCGGTATGTAAACTTCTAT
CAGCAGGGAAGAAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTA
GGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCG
GGGCTCAACCCCGGGACTGCATTTGGAACCTGCTGAGCTAGAGTGTGCGAGAGGCAAGTGAATTTCCTAGTGTAGC
GGTGAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTGAGGCTC
20 GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCGGGTG
GCAAAGCCATTCCGTGCCGACGCTAACGCAATAAGCAGTCCACCTGGGGAGTACGTTTCGCAAGAATGAACTCAA
AGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGAT
CTTGACATCCCGATGACTGCTTCGTAATGGAAGTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTGCG
TCAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAA
25 GCTGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTA
TGACCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAACTCGCGAGGGTAAGCAAATCTCAAAA
ATAACGTCTCAGTTCCGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAA
TGCTGCGGTGAATACGTTCCCGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCA
GTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAGCCGTA
30 TCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 152 | PROKKA_00437 16S ribosomal RNA gene

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
35 ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGATGATGCGTAGCCG
GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
TCGTCAATGGGGAAACCTGAACGAGCAATGCCGCGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
40 CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCCTAGGTGGCGTA
CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGAAGCTATGGAACTGGTATGCTGGAGTGCAGAAGAGGG
CGATGGAATTCCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTGCGCTGGT
CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
ATGAGAACTAAGTGTGGAGGAATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCTGGGGAGTATGCACGCA
45 AGTGTGAAACTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
ATGGTTGTCGTCAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTGTCGCATGTTACC
AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
GCCCCCTTATGGCTGGGCTACACACGTACTACAATGGCGGCCACAAAGAGCAGCGACACAGTGTGTGAAGCGAA
50 TCTCATAAAGGTCGCTCAGTTCGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGTACACACCGCCCGTCAAACCATGGGAGTCAGTAATACC
CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

>SEQ ID NO: 153 | PROKKA_00896 16S ribosomal RNA gene

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
55 GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA

ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGATGATGCGTAGCCG
GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
TCGTCAATGGGGGAAACCCCTGAACGAGCAATGCCGCGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
5 AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGTGGCGTA
CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAAACTGGTATGCTGGAGTGCAGAAGAGGG
CGATGGAATTCCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTTCGCTGGT
10 CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
ATGAGAACTAAGTGTGGAGGAATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCCTGGGGAGTATGCACGCA
AGTGTGAAACTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
ATGGTTGTCTGCTCAGCTCGTGTCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGTTCGCATGTTACC
15 AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGGCCACAAAGAGCAGCGACACAGTGTGAAGCGAA
TCTCATAAAGGTCTGCTCAGTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGACACACCGCCCGTCAAACCATGGGAGTCAGTAATACC
CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

>SEQ ID NO: 154 | PROKKA_02845 16S ribosomal RNA gene

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
25 GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCTCACCAGGCGATGATGCGTAGCCG
GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
TCGTCAATGGGGGAAACCCCTGAACGAGCAATGCCGCGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGTGGCGTA
30 CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAACTGGTATGCTGGAGTGCAGAAGAGGG
CGATGGAATTCCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTTCGCTGGT
CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
ATGAGAACTAAGTGTGGAGGAATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCCTGGGGAGTATGCACGCA
AGTGTGAAACTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
35 GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
ATGGTTGTCTGCTCAGCTCGTGTCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGTTCGCATGTTACC
AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGGCCACAAAGAGCAGCGACACAGTGTGAAGCGAA
TCTCATAAAGGTCTGCTCAGTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
40 GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGACACACCGCCCGTCAAACCATGGGAGTCAGTAATACC
CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

>SEQ ID NO: 155 | PROKKA_04164 16S ribosomal RNA gene

ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
ACTGCTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
45 GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGATGATGCGTAGCCG
GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
TCGTCAATGGGGGAAACCCCTGAACGAGCAATGCCGCGTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGTGGCGTA
CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAAACTGGTATGCTGGAGTGCAGAAGAGGG
CGATGGAATTCCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTTCGCTGGT
50 CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
ATGAGAACTAAGTGTGGAGGAATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCCTGGGGAGTATGCACGCA
AGTGTGAAACTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC

ATGGTTGTCGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTGTCGCATGTTACC
 AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
 GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGACCACAAAGAGCAGCGACACAGTGTATGTAAGCGAA
 TCTCATAAAGGTCGTCTCAGTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
 5 GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGTTACACACCGCCCCGTCAAACCATGGGAGTCAGTAATACC
 CGAAGCCGGTGGCATAACCGCAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
 TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

>SEQ ID NO: 156 | PROKKA_04921 16S ribosomal RNA gene

10 ATGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCATGCCTAATACATGCAAGTCGAACGAAGTTTCGAG
 GAAGCTTGCTTCCAAAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCCATGTGTCCGGGATA
 ACTGCTGGAACCGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGCGCCCATCAAGGC
 GTGAACATGGATGGACCTGCGGCGCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGATGATGCGTAGCCG
 GCCTGAGAGGGTAAACGGCCACATTGGGACTGAGACACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTT
 15 TCGTCAATGGGGGAAACCCCTGAACGAGCAATGCCCGCTGAGTGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGT
 AAGTGAAGAACGGCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACCAGAAAGCCACGGCTAACTACGTG
 CCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGTGGCGTA
 CTAAGTCTGTAGTAAAAGGCAATGGCTCAACCATTTGTAAGCTATGGAACTGGTATGCTGGAGTGCAGAAGAGGG
 CGATGGAATTCATGTGTAGCGGTAAAATGCGTAGATATATGGAGGAACACCAGTGGCGAAGGCGGTGCGCTGGT
 20 CTGTAACCTGACACTGAGGCACGAAAGCGTGGGGAGCAAATAGGATTAGATACCCTAGTAGTCCACGCCGTAAACG
 ATGAGAATAAGTGTGGAGGAATTCAGTGTGCTGAGTTAACGCAATAAGTTCTCCGCCCTGGGAGTATGCACGCA
 AGTGTGAAACTCAAAGGAATTGACGGGGGCCCCGACAAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
 GAACCTTACCAGGCCTTGACATGGAAACAAATACCCTAGAGATAGGGGGATAATTATGGATCACACAGGTGGTGC
 ATGGTTGTCGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTGTCGCATGTTACC
 25 AGCATCAAGTTGGGGACTCATGCGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
 GCCCCCTTATGGCCTGGGCTACACACGTACTACAATGGCGGCCACAAAGAGCAGCGACACAGTGTATGTAAGCGAA
 TCTCATAAAGGTCGTCTCAGTTCGGATTGAAGTCTGCAACTCGACTTCATGAAGTCGGAATCGCTAGTAATCGCA
 GATCAGCATGCTGCGGTGAATACGTTCTCGGGCCTTGTTACACACCGCCCCGTCAAACCATGGGAGTCAGTAATACC
 CGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCAAGGTAGGACCGATGACTGGGGTTAAGTCGTAACAAGG
 30 TATCCCTACGGGAACGTGGGGATGGATCACCTCCTTT

>SEQ ID NO: 157 | PROKKA_00199 16S ribosomal RNA gene

TATTGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGGGTGCTCA
 TGACGGAGGATTTCGTCCAACGGATTGAGTTACCTAGTGGCGGACGGGTGAGTAACGCGTGAGGAACCTGCCTTGG
 35 AGAGGGGAATAACACTCCGAAAGGAGTGCTAATACCGCATGATGCAGTTGGGTGCGATGGCTCTGACTGCCAAAG
 ATTTATCGCTCTGAGATGGCCTCGCGTCTGATTAGCTAGTAGGCGGGGTAACGGCCACCTAGGCGACGATCAGT
 AGCCGGACTGAGAGGTTGACCGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGG
 AATATTGGGCAATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTTGTAACTTCT
 TTTGTGCGGGACGAAACAAATGACGGTACCCGACGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAAT
 40 ACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGATTGCAAGTCAGATGTGAAA
 ACTGGGGGCTCAACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAATTCGGTGT
 GTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGATTGCTGGACAGTAACCTGACGCTGA
 GCGCGGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACCTAGGTGTG
 GGGGTCTGACCCCTCCGTGCCGAGTTAACACAAATAAGTATCCCACTGGGGAGTACGATCGCAAGGTTGAAAC
 45 TCAAAGGAATTGACGGGGGCGCACAAGCGGTGGAGTATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACC
 AGGGCTTGACATCCCACTAACGAAGCAGAGATGCATTAGGTGCCCTTCGGGGAAAGTGGAGACAGGTGGTGCATG
 GTTGTGCTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTGTTAGTTGCTACG
 CAAGAGCACTCTAGCGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCTTA
 TGTCTTGGGCCACACAGTACTACAATGGTGGTTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAATCCCTAA
 50 AAGCCATCCAGTTCGGATTGACAGGTGAAACCCGCTGTATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCA
 TGCCGCGGTGAATACGTTCCCGGGCCTTGTTACACACCGCCCCGTACACCATGAGAGTCGGGAACACCCGAAGTCC
 GTAGCCTAACCGCAAGGAGGGCGCGGCCGAAGGTGGGTTCGATAATTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 TCGGAAGGTGCGGCTGGATCACCTCCTTT

55 >SEQ ID NO: 158 | PROKKA_00208 16S ribosomal RNA gene

TATTGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGGGTGCTCA
 TGACGGAGGATTTCGTCCAACGGATTGAGTTACCTAGTGGCGGACGGGTGAGTAACGCGTGAGGAACCTGCCTTGG
 AGAGGGGAATAAACAACCTCCGAAAGGAGTGCTAATACCGCATGATGCAGTTGGGTGCGATGGCTCTGACTGCCAAAG
 ATTTATCGCTCTGAGATGGCCTCGCGTCTGATTAGCTAGTAGGCGGGGTAACGGCCCACCTAGGCGACGATCAGT
 5 AGCCGGACTGAGAGGTTGACCGGCCACATTGGGACTGAGACACGGGCCAGACTCCTACGGGAGGCAGCAGTGGGG
 AATATTGGGCAATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTTGTAACTTCT
 TTTGTGGGGACGAAACAAATGACGGTACCCGACGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAAT
 ACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTTAGGCGGGATTGCAAGTCAGATGTGAAA
 10 ACTGGGGGCTCAACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAATTCCTGT
 GTAGCGGTGAAATGCGTAGATATACGGAGGAACACCACTGGCGAAGGCGGATTGCTGGACAGTAACCTGACGCTGA
 GCGCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACCTAGGTGTG
 GGGGCTCTGACCCCTCCGTGCCGCGAGTTAACACAATAAGTATCCACCTGGGGAGTACGATCGCAAGGTTGAAAC
 CTCAAAGGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGTATGTGGTTTAATTCTGAAGCAACGCCGAAGAACCTTAC
 CAGGGCTTGACATCCCACTAACGAAGCAGAGATGCATTAGGTGCCCTTCGGGGAAAGTGGAGACAGGTGGTGCAT
 15 GGTGTGTCGTCAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTGTTAGTTGCTAC
 GCAAGAGCACTCTAGCGAGACTGCCGTTGACAAAAACGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCTT
 ATGTCTGGGCCACACACGTACTACAATGGTGGTTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAATCCCTA
 AAAGCCATCCCAGTTCCGATTGCAGGCTGAAACCCGCTGTATGAAGTTGGAATCGCTAGTAATCGCGGATCAGC
 ATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTC
 20 CGTAGCCTAACCGCAAGGAGGGCGCGGCCGAAGGTGGGTTCGATAATTGGGGTGAAGTCGTAACAAGGTAGCCGT
 ATCGGAAGGTGCGGCTGGATCACCTCCTTT

>SEQ ID NO: 159 | PROKKA_04460 16S ribosomal RNA gene

TATTGAGAGTTTGTATCCTGGCTCAGGATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGGGTGCTCA
 25 TGACGGAGGATTTCGTCCAACGGATTGAGTTACCTAGTGGCGGACGGGTGAGTAACGCGTGAGGAACCTGCCTTGG
 AGAGGGGAATAAACAACCTCCGAAAGGAGTGCTAATACCGCATAATGCAGTTGGGTGCGATGGCTCTGACTGCCAAAG
 ATTTATCGCTCTGAGATGGCCTCGCGTCTGATTAGCTAGTAGGCGGGGTAACGGCCCACCTAGGCGACGATCAGT
 AGCCGGACTGAGAGGTTGACCGGCCACATTGGGACTGAGACACGGGCCAGACTCCTACGGGAGGCAGCAGTGGGG
 AATATTGGGCAATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTTGTAACTTCT
 30 TTTGTGAGGGACGAAACAAATGACGGTACCTGACGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAAT
 ACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTTAGGCGGGATTGCAAGTCAGATGTGAAA
 ACTGGGGGCTCAACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAATTCCTGT
 GTAGCGGTGAAATGCGTAGATATACGGAGGAACACCACTGGCGAAGGCGGATTGCTGGACAGTAACCTGACGCTGA
 35 GCGCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGGTGTG
 GGGGGTCTGACCCCTCCGTGCCGCGAGTTAACACAATAAGTATCCACCTGGGGAGTACGATCGCAAGGTTGAAAC
 TCAAAGGAATTGACGGGGGCGCCGACAAAGCGGTGGAGTATGTGGTTTAATTCTGAAGCAACGCCGAAGAACCTTACC
 AGGGCTTGACATCCCACTAACGAAGCAGAGATGCATTAGGTGCCCTTCGGGGAAAGTGGAGACAGGTGGTGCATG
 GTTGTGCTGACGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTTATTGTTAGTTGCTACG
 CATTAGCACTCTAGCGAGACTGCCGTTGACAAAAACGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCTTAA
 40 TGTCTGGGCCACACACGTACTACAATGGTGGTTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAAATCCCTAA
 AAGCCATCCCAGTTCCGATTGCAGGCTGAAACCCGCTGTATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCA
 TGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTCC
 GTAGCCTAACCGCAAGGAGGGCGCGGCCGAAGGTGGGTTCGATAATTGGGGTGAAGTCGTAACAAGGTAGCCGTA
 45 TCGGAAGGTGCGGCTGGATCACCTCCTTT

This invention is not limited in its application to the details of construction and the
 arrangement of components set forth in the following description or illustrated in the
 drawings. The invention is capable of other embodiments and of being practiced or of being
 carried out in various ways. Also, the phraseology and terminology used herein is for the
 purpose of description and should not be regarded as limiting. The use of “including,”
 50 “comprising,” or “having,” “containing,” “involving,” and variations thereof herein, is meant
 to encompass the items listed thereafter and equivalents thereof as well as additional items.

Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. The methods and techniques of the present disclosure are generally performed according to conventional methods well-known in the art. Generally, nomenclatures used in connection with, and techniques of biochemistry, enzymology, molecular and cellular biology, microbiology, virology, cell or tissue culture, genetics and protein and nucleic chemistry described herein are those well-known and commonly used in the art. The methods and techniques of the present disclosure are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated.

The present invention is further illustrated by the following Examples, which in no way should be construed as further limiting. The entire contents of all of the references (including literature references, issued patents, published patent applications, and co-pending patent applications) cited throughout this application are hereby expressly incorporated by reference, in particular for the teaching that is referenced hereinabove. However, the citation of any reference is not intended to be an admission that the reference is prior art.

EXAMPLES

Example 1: Mouse model of *C. difficile* infection5 Mouse husbandry

Experiments were performed using C57BL/6J female mice purchased from Jackson Laboratories (Bar Harbor, ME) and housed in ventilated sterile cages. All animals were maintained in a specific-pathogen-free facility. Animals were acclimated to the vivarium for at least 3 days prior to study (*i.e.*, commencing antibiotic courses). For experiments involving *C. difficile* infection, mice were administered $10\text{-}10^4$ *C. difficile* VPI 10463 spores in 200 μl PBS by oral gavage. Experiments were performed in compliance with institutional guidelines and approved by the institution's Institutional Animal Care and Use Committee. Sterile food and drinking water were provided to the animals

15 Live Biotherapeutic Product (LBP) Preparation

Individual bacterial strains were isolated from fecal material obtained from healthy donors. The individual strains were struck out from 15% glycerol freezer stocks onto EG (Eggerth Gagnon) agar plates containing 5% horse blood in an anaerobic chamber and incubated for 24-48 hours at 37°C. Colonies were inoculated into pre-reduced liquid Peptone Yeast Glucose (PYG) media and grown for 24-48 hours until dense (static in the anaerobic chamber). Optical density (OD₆₀₀) of the cultures was assessed and live biotherapeutic product (LBP) cocktails were prepared inside an anaerobic chamber adjusting inputs based upon OD₆₀₀ for equal CFU ratio cocktails in PBS (sterile, pre-treated).

25 *C. difficile* colony forming unit (CFU) determination

Fecal pellets were collected, transported to an anaerobic chamber (<2 hours), and manually homogenized in 500 μL of pre-reduced PBS using a pipette tip and through repeated pipetting. Serial dilutions of fecal homogenates were prepared in pre-reduced PBS, 100 μL of which was spread onto cycloserine-cefoxitin-fructose agar with sodium taurocholate (TCCFA) plates, and incubated anaerobically at 37°C. *C. difficile* CFUs were enumerated at 48 hours.

Murine susceptibility to *C. difficile* infection

Groups of mice were evaluated for susceptibility to *C. difficile* using three antibiotic regimen protocols: (1) an antibiotic cocktail, (2) clindamycin administration, or (3) cefoperazone administration (Figures 2 and 3). The antibiotic cocktail consisted of kanamycin (0.4 mg/ml), gentamicin (0.035 mg/ml), colistin (0.056 mg/ml), metronidazole (0.215 mg/ml), vancomycin (0.045 mg/ml) in the drinking water from day -10 to day -3, followed by a single intraperitoneal clindamycin injection (200 µg/mouse). The clindamycin administration involved a single intraperitoneal injection of clindamycin (200 µg/mouse) on day -1. The notation of days is relative to day 0, the day of *C. difficile* infection.

Mice were treated with the indicated antibiotic regimen as described above and then infected with either 10 or 10⁴ *C. difficile* spores by oral gavage on day 0 (Figures 2 and 3). An additional experimental arm was added to the antibiotic treatment model in which mice were treated with vancomycin after *C. difficile* infection (Figure 4J; black triangles).

Mice were monitored daily following infection for mortality/survival (Figures 4A-4D) and weight (Figures 4E-4H). Fecal pellets were also collected daily and used for *C. difficile* CFU enumeration, presented as CFU/gram feces (Figures 4I-4L).

The groups of mice that received cefoperazone treatment had a significant change in weight (Figure 4H) and substantial *C. difficile* bacterial load in the fecal pellets (Figure 4L), even following administration with 10 *C. difficile* spores. These results indicated that the cefoperazone pre-treatment regimen provided a good model for *C. difficile* infection and for evaluating protection and/or treatment of *C. difficile* infection. In the absence of antibiotic treatment prior to infection, *C. difficile* infection was not established (Figure 4I) and all mice survived (Figure 4A) without significant change in body weight (Figure 4E).

Example 2: Live Biotherapeutic Product (LBP) preparations protect against *C. difficile* infection.

The following LBP compositions were evaluated for their capacity to protect and/or treat *C. difficile* infection:

Composition A,

Composition B,

Composition C,

Composition D,

Composition E (See *e.g.*, Narushima *et al.*, Gut Microbes (2014) 5(3) 333-339), and

Composition I: a mixture of *Clostridium scindens*, *Pseudoflavonifractor capillosus* and *Blautia hansenii* (Figure 5).

In general, LBP cocktails were mixed in PYG media, and each mouse was administered a dose by oral gavage in 250 μ L pre-reduced PBS (media-free). For composition E, bacteria were mixed in equal volumes (not equal ratios/CFUs) and administered in a 250 μ L dose. Each LBP of Compositions A-D contained 10^8 CFUs total in a 250 μ L dose, comprised of 10^7 CFU of each of the bacterial strains (Figure 1), for a total of 10^8 CFU administered to each animal. Composition I contained a total of 10^6 CFUs in a 250 μ L dose (approximately 333,000 of each of the 3 bacteria mixed).

Groups of mice were subjected to cefoperazone treatment, as described in Example 1, and were administered the indicated composition by oral gavage 2 days after the cessation of cefoperazone treatment. Twenty-four hours later, the mice were subjected to infection with 10^4 *C. difficile* spores (Figure 5). Mice evaluated for survival/mortality (Figure 6), weight (Figures 7A-7I, and *C. difficile* CFUs (Figures 8A-8C). The results show that administration of Composition B prior to *C. difficile* infection is an effective protection and/or treatment against *C. difficile* infection.

Example 3: Composition B protects against and/or treats *C. difficile* infection.

Groups of 10-12 week old mice were used in the *C. difficile* mouse model (Figure 9). Mice were subjected to cefoperazone treatment as described in Example 1. One group of mice was then administered Composition B (10^8 CFU per mouse) administered by oral gavage, as described in Examples 1 and 2, 2 days after the cessation of cefoperazone treatment. The other group of mice did not receive a live biotherapeutic product after cefoperazone treatment (control). Twenty-four hours later, the mice were subjected to *C. difficile* infection (10^4 *C. difficile* spores) and then evaluated for survival/mortality (Figure 10), weight (Figure 11), and *C. difficile* burden (CFUs per gram feces; Figure 12). These results confirm the results of Example 2 that demonstrate treatment with Composition B prior to *C. difficile* infection is an effective protection and/or treatment against *C. difficile* infection.

Example 4: LBP Composition F protects against and/or treat *C. difficile* infection.

Figure 13 shows the strains of live biotherapeutic product (LBP) Composition F. The genus-species classification indicates the closest species based on the sequence of the isolated strain. Figure 14 shows the classification by *Clostridium* cluster of the strains in Composition F.

Groups of mice were administered cefoperazone, as described in the Examples above, then administered LBPs or fecal matter transplant (FMT) from mice or human (Figure 15). Composition B was administered to the indicated groups on day -1; days -2 and -1; or on days -2, -1, 1, 2, and 3, relative to infection with 10^4 *C. difficile* spores. Composition F was administered to the indicated groups on day -1 or on days -2, -1, 1, 2, and 3, relative to administration of *C. difficile* spores. Additional groups received FMT from mice or from humans (200 μ L of a 10% fecal sample s per mouse). Mice were then evaluated for survival/mortality (Figure 16), weight (Figures 17A-17H), and *C. difficile* burden (CFU/gram feces) on days 1, 3, 8 and 17 after infection (Figures 18A and 18B). The data demonstrate that Composition B, Composition F, and FMT protect against and/or treat *C. difficile* infection.

Example 5: LBP compositions protect against and/or treat *C. difficile* infection.

Figure 19 shows the strains of LBP Composition G. The genus-species notation indicates the closest species based on the sequence of the isolated strain. Composition G includes a subset of the strains of Composition F. Groups of mice were administered cefoperazone, as described in the Examples above, then administered the LBP:

Composition B;

Composition B-1 (Composition B with *Bacteroides* added);

Composition B-2 (Composition B from which *Flavonifractor plautii* was removed and *Bacteroides* added);

Composition F;

Composition G;

Human fecal samples subjected to ethanol treatment;

Composition B subjected to ethanol treatment;

Composition B that had been frozen; or

Composition J: *Clostridium innocuum*, *Clostridium bolteae* and *Clostridium symbiosum* subjected to ethanol treatment;

(See also Figure 20).

The *Bacteroides* strain used in Composition B-1 and B-2 was *Bacteroides ovatus* (strain identifier 211-B; SEQ ID NO: 83).

Mice were challenged with *C. difficile* VPI 10463 spores (10^4) and monitored daily (Day 0 to Day 7 post *C. difficile* infection) for survival/mortality (Figures 21 and 23) and

change in weight (Figures 22A-22J and 24). These data show that the compositions protect against and/or treat *C. difficile* infection.

Example 6: LBP compositions protect against and/or treat *C. difficile* infection.

Groups of mice were subjected to cefoperazone treatment, as described above, then administered human fecal matter transplant, Composition B, Composition B + 4 spores, or Composition H (Figure 25). “Composition B + 4 spores” refers to Composition B plus the following four strains in spore form: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum* and *Clostridium innocuum*. Composition H contains the following six strains in spore form: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum*, *Clostridium innocuum*, *Clostridium disporicum* and *Erysipelatoclostridium ramosum* (Figure 26).

Mice were then challenged with *C. difficile* infection with 10^4 *C. difficile* VPI 10463 spores and monitored for survival/mortality (Figures 27A and 28A), weight (Figures 27B and 28B). Mice that lost more than 20% body weight relative to baseline were included in mortality numbers in survival curves. The *C. difficile* burden was assessed by CFU in fecal pellets on days 1, 4 and 19 after infection (Figures 29A-29C).

These data indicate that Composition B as well as other compositions can improve survival in the cefoperazone-induced *C. difficile* mouse model and protect against and/or treat *C. difficile* infection.

Example 7: *C. difficile* toxin experiment

Vero cells, epithelial cells derived from African Green Monkey kidney epithelium, are sensitive to a variety of bacterial toxins, including *C. difficile* Toxin B. Exposure of cells to *C. difficile* Toxin B results in inhibition of the function of Rho, Rac, and Cdc42 leading to a decline in F-actin, a change in cell morphology (e.g., cell rounding), and eventually apoptosis.

To determine whether administration of bacterial compositions described herein has an effect on the production or activity of *C. difficile* Toxin B, a cellular assay was performed. Briefly, groups of mice were treated with cefoperazone, as described above, and administered human fecal matter transplant (FMT) (“4-3”); Composition B (“5-3”); Composition B plus four strains in spore form: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum* and *Clostridium innocuum* (“7-4”), or no treatment. Each of the groups of mice were then exposed to *C. difficile* infection with 10^4 *C. difficile* spores. The groups of mice

that did not receive a treatment after cefoperazone administration and prior to *C. difficile* infection are referred to as “2-1 (Cdiff)” and “2-4 (Cdiff).” An additional group of mice was not exposed to *C. difficile* as indicated by “N3 (Healthy)”.

Fecal pellets were collected from each of the groups of mice, weighed, and homogenized in PBS and normalized to a fixed concentration (~25 mg/mL). The samples were centrifuged to prepare a clarified supernatant, which was then diluted in 10-fold serial dilutions to produce a range from 1:10 to 1:10⁻⁶ dilutions of clarified pellet supernatant. Vero cell cultures were exposed to the diluted samples for approximately 18 hours, then visualized by phase contract microscopy to assess morphological changes (*i.e.*, cell rounding) associated with *C. difficile* toxin exposure. The cells were scored based on the highest concentration of supernatant that did not yield a change in morphology (Figure 30). The samples from mice that had been treated with Composition B prior to *C. difficile* infection had reduced amounts of *C. difficile* Toxin B, as compared to samples from control mice that did not receive a treatment after cefoperazone administration and prior to *C. difficile* infection (“2-1 (Cdiff)” and “2-4 (Cdiff)”) as well as compared to samples from mice that received FMT. Notably, the samples from mice that had been treated with Composition B also had reduced amounts of *C. difficile* Toxin B, as compared to samples from mice that had been treated with Composition B with additional spores.

Example 8: *In vitro* Competition Between Compositions B and *C. difficile*

Composition B was assessed for its ability to suppress *Clostridium difficile* growth by an *in vitro* mixed culture competition assay. From glycerol freezer stocks, individual strains of Composition B, *C. difficile* (Cdiff), *Clostridium bifermentans*, and *Bacteroides thetaiotaomicron* were struck out onto Eggerth-Gagnon agar plates with horse blood (EG+HB). Single colonies of each of the strains were subsequently inoculated into brain heart infusion (BHI) liquid media and allowed to grow in pure culture for 24-48 hours. Turbid cultures were sub-cultured then grown to exponential phase and finally diluted and combined to prepare a mixed culture with an optical density (OD₆₀₀) of 0.1. Exponential phase Cdiff culture was added to the mixed culture at a final concentration with an OD of 0.1. After the cultures were combined and incubated for 2-3 hours, samples were collected, serially diluted, and plated on Taurocholate-Cycloserine-Cefoxitin-Fructose Agar (TCCFA) plates to select for Cdiff growth. After 48-72 hours, the colony forming units (CFUs) of Cdiff in each competition experiment were determined by manual colony counting.

EG+HB agar plates were prepared according to standard procedures and reduced in an anaerobic environment for at least 6-8 hours prior to use. Liquid BHI medium was obtained from BD Biosciences (Catalog # 211059, San Jose, CA), prepared according to the manufacturer's instructions, and reduced in an anaerobic environment for at least 18-24 hours prior to use. TCCFA plates were prepared according to standard procedures and reduced in an anaerobic environment for at least 6-8 hours prior to use. *Clostridium difficile* strain used in the experiments: American Type Culture Collection (ATCC) 43255.

Table 4: Composition B strains

Composition B
VE202-7
VE202-13
VE202-14
VE202-16
Strain #16
Strain #170
Strain #189
Strain #211

Strains were struck out onto EG+HB agar plates from frozen glycerol stocks inside an anaerobic chamber for 48-72 hours. Single colonies were inoculated into 10 mL of BHI media and grown 24-48 hours at 37°C in the anaerobic chamber. Turbid cultures were then diluted to an OD of 0.1 and grown for 2-3 hours at 37°C in the anaerobic chamber.

Exponential phase cultures were diluted and combined at equivalent ODs. For the competition assay, each of the strains of Combination B (Table 4) were combined in equal parts, based on OD₆₀₀, to reach a final consortium OD₆₀₀ of 0.1. *C. bifermentans* and *B. thetaiotaomicron* were setup to compete with Cdiff individually at an OD of 0.1. The OD₆₀₀ for Cdiff in each of the mixed culture competition experiments was 0.1. After combination, the cultures were incubated for 2-3 hours at 37°C in the anaerobic chamber, then prepared for enumerations on Cdiff selective plates.

TCCFA plates are selective for Cdiff growth, and none of the Combination B strains, nor either of the control strains (*C. bifermentans* and *B. thetaiotaomicron*), grow on these plates. Inside an anaerobic chamber, a 100 µL sample of each competition culture was collected and serially diluted 1:10 to reach a final dilution of 1x10⁻⁶. Plates for CFU enumeration were prepared by spreading 100 µL of each of the 1x10⁻⁴ through 1x10⁻⁶

dilutions on TCCFA plates using sterile spreading loops. CFU plates were incubated for 48-72 hours at 37°C in the anaerobic chamber. CFU enumeration was completed by manually counting colonies.

To determine the effect of competition, the ratio of CFUs determined for the competition samples and Cdiff alone was calculated and expressed as a percentage. Inhibition of Cdiff growth by the Composition B cocktail was compared to the responses of *B. thetaiotaomicron* (negative control) and *C. bifermentans* (positive control). The results are shown in Table 5 and Figure 31.

Table 5: Summary Results for *In Vitro* Competition

Experiment Number	No Competing Strain(s)	Competition with <i>B. thetaiotaomicron</i>	Competition with <i>C. bifermentans</i>	Competition with Composition B
n=1	100			33.8
n=2	100	9.90	0.1	0.5
n=3	100	115	39.5	33.1
n=4	100	41.3	0.7	0.7
n=5	100	105	14.1	20.9
n=6	100	57.4	4.1	1.6
Mean	100	65.6	11.7	15.1
Std. Dev.	0	43.8	16.5	16.2
Total N	6	5	5	6

Data is expressed as Cdiff CFU as a percentage of control. Each n is representative of a single biological replicate, independent of other measurements.

In *in vitro* competition, Composition B inhibited Cdiff growth to 15.1 ± 16.2 % of control (absence of competing strain(s)). This result is consistent with the inhibition observed by the positive control, *C. bifermentans*, of 11.7 ± 16.5 % of control. *B. thetaiotaomicron*, a negative control, yielded a negligible effect on Cdiff growth at 65.6 ± 43.8 % of control. Given the variability inherent in the assessment of CFU, inhibition of growth to < 25 % of control is considered to be significant inhibition and both the positive control and Composition B cocktail meet this threshold of activity. The Composition B consortium attenuated Cdiff growth *in vitro* comparable to the direct competition observed by *C. bifermentans*. Direct competition with *B. thetaiotaomicron* did not significantly inhibit Cdiff growth.

Example 9: Determination of *In Vitro* Short-Chain Fatty Acid Production

Each strain of Composition B was assessed for individual short-chain fatty acid (SCFA) production *in vitro*. Composition B strains were grown in pure cultures inside an anaerobic chamber. Spent supernatant from liquid media cultures was harvested by centrifugation, filter sterilized, and then stored at $< -70^{\circ}\text{C}$. Frozen clarified supernatant specimens were analyzed for short-chain fatty acids (SCFAs).

EG+HB agar plates (Eggerth-Gagnon agar plates with horse blood) were prepared according to standard methods and reduced in an anaerobic environment for at least 6-8 hours prior to use. Liquid PYG medium (pre-formulated, pre-reduced) was obtained from Anaerobe Systems (Catalog#AS-822; Morgan Hill, CA).

Strains were struck out onto EG+HB agar plates from frozen 15% glycerol stocks inside an anaerobic chamber for 48 – 72 hours. Single colonies were inoculated into 7 mL PYG media and grown 24-48 hours at 37°C in the anaerobic chamber. Unless otherwise noted, when the optical density (OD) was ≥ 0.2 , samples were collected for CFU enumeration and filtration. Inside an anaerobic chamber, a 100 μL sample of turbid culture was collected and serially diluted 1:10 to reach a final dilution of 1×10^{-6} . Plates for CFU enumeration were prepared by spreading 100 μL /dilution for the 1×10^{-4} through 1×10^{-6} dilutions on EG+HB agar plates using sterile glass beads. CFU plates were incubated for 48-72 hours in the anaerobic chamber. CFU enumeration was completed using the EasyCount 2 (bioMérieux SA, Marcy-l'Étoile, France). Immediately after samples of turbid cultures were collected for CFU enumeration, the remaining turbid cultures were centrifuged at approximately 1000 RCF for 10 minutes to pellet cellular debris. The clarified supernatants were transferred to a 0.2 μm plate filter and vacuum filtered to remove any remaining particulates prior to bioanalysis. In the event of blockage in the filter plate, clarified supernatants were manually filtered using 0.2 μm syringe filters. Filtered supernatants were aliquoted and stored at $< -70^{\circ}\text{C}$ prior to bioanalysis of SCFAs.

To facilitate easier comparisons between samples, raw SCFA data ($\mu\text{g/mL}$) was normalized by the $\log_{10}$ of corresponding determined/estimated CFU for the culture. The results are depicted in Table 6 and Table 7 below.

Table 6: Enumerated CFUs for Composition B Strains

Sample ID	OD600	Enumerated CFU (CFU/mL)
VE202-7	> 2	6.11E+08
VE202-13	0.8	4.00E+08
VE202-14	> 2	1.60E+09
VE202-16	1.92	1.28E+09
#16	1.97	1.69E+08
# 170	1.8	1.08E+08
# 189	1.03	1.74E+09
# 211	0.35	3.71E+08

Table 7: SFCAs produced by individual Composition B strains

Sample ID	Normalized ($\mu\text{g}/\text{Log}(\text{CFU}/\text{mL}) \cdot \text{mL}$)							
	Acetate	Propionate	Iso-butyrate	Butyrate	2-Methyl-butyrate	Iso-valerate	Valerate	Hexanoate
VE202-7	123.7	0.077	0.102	0.208	0.015	0.056	BLOQ	0.031
VE202-13	30.1	0.545	0.116	34.452	0.288	0.188	0.097	0.034
VE202-14	110.5	0.054	0.022	0.248	0.011	0.014	BLOQ	0.009
VE202-16	313.2	0.000	0.000	0.280	0.004	0.000	BLOQ	0.009
#16	104.0	0.005	0.000	50.988	0.014	0.033	BLOQ	0.009
# 170	87.1	0.055	0.025	0.215	0.011	0.039	BLOQ	0.016
# 189	0.0	BLOQ	0.000	35.751	0.005	0.019	0.359	0.587
# 211	57.6	5.289	0.000	78.227	0.028	0.050	0.053	0.095

Seven strains of Composition B were found to produce significant quantities ($> 1 \mu\text{g}/\text{Log}(\text{CFU}/\text{mL}) \cdot \text{mL}$) of the 2-carbon SCFA, acetate. One strain, (#211), produced substantial quantities of the 3-carbon SCFA, propionate. Four strains of Composition B produced substantial quantities of the 4-carbon SCFA, butyrate. Trace quantities ($< 1 \mu\text{g}/\text{Log}(\text{CFU}/\text{mL}) \cdot \text{mL}$) of other SCFAs were also produced by the Composition B strains.

Example 10: Composition B induces regulatory T cells (Tregs)

Each of the bacterial strains of Composition B were grown to log phase, combined to a total dose of $\sim 10^8$ cfu per mouse. Germ-free mice were inoculated with Composition B or a negative control by oral gavage and sacrificed following four weeks of colonization. Lamina propria leukocytes were isolated from colonic tissue of individual mice by standard procedures and assessed by flow cytometry. The regulatory T cell content was evaluated as the percentage of Foxp3-positive cells among CD4+ T cells.

As shown in Figure 32, mice that were inoculated with Composition B were found to have significantly more regulatory T cells as compared to mice that were inoculated with the control.

What is claimed is:

CLAIMS

1. A composition comprising:
 - (i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
 - 5 (ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
 - (iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
 - (iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
 - 0 (v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
 - (vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;
 - 5 (vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and
 - (viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.
- 10 2. The composition of claim 1, wherein:
 - (i) the isolated bacterial strain of (i) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
 - (ii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
 - 25 (iii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
 - (iv) the isolated bacterial strain of (iv) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
 - (v) the isolated bacterial strain of (v) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
 - 30 (vi) the isolated bacterial strain of (vi) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 146;
 - (vii) the isolated bacterial strain of (vii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and

(viii) the isolated bacterial strain of (viii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

3. A composition comprising:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

4. A composition comprising:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

5. The composition of any one of claims 1–4, wherein the isolated bacterial strains are lyophilized.

6. The composition of any one of claims 1–5, wherein one or more of the isolated bacterial strains are in spore form.

7. A food product comprising the composition of any one of claims 1–6 and a nutrient.

8. A pharmaceutical composition comprising the composition of any one of claims 1–6 and a pharmaceutically acceptable excipient.
9. The pharmaceutical composition of claim 8, wherein the pharmaceutical composition is formulated for delivery to the colon.
10. The pharmaceutical composition of claim 8 or 9, wherein the pharmaceutical composition is formulated for oral delivery.
11. The pharmaceutical composition of any one of claims 8–10, wherein the pharmaceutical composition comprises a pH-sensitive composition comprising one or more enteric polymers.
12. A method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:
 - (i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
 - (ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
 - (iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
 - (iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
 - (v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
 - (vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;
 - (vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and
 - (viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.
13. The method of claim 12, wherein:

- (i) the isolated bacterial strain of (i) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
- (ii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
- (iii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
- (iv) the isolated bacterial strain of (iv) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 137;
- (v) the isolated bacterial strain of (v) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 141;
- (vi) the isolated bacterial strain of (vi) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 146;
- (vii) the isolated bacterial strain of (vii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and
- (viii) the isolated bacterial strain of (viii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

14. A method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

15. A method of treating or preventing *Clostridium difficile* infection in a subject, the method comprising administering to the subject:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

16. The method of any one of claims 12–15, wherein the isolated bacterial strains are administered in more than one composition, each composition comprising at least one of the isolated bacterial strains.

17. The method of any one of claims 12–16, wherein the subject is administered an antibiotic prior to administration of the isolated bacterial strains.

18. The method of claim 17, wherein the antibiotic is vancomycin, kanamycin, gentamicin, colistin, metronidazole, clindamycin, fidaxomicin, or cefoperazone.

19. The method of claim 17 or 18, wherein the antibiotic is vancomycin.

20. The method of any one of claims 12–19, wherein the method is a method of preventing recurrence of *Clostridium difficile* infection in the subject.

21. A kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

- (i) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 124;
- (ii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 129;
- (iii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 132;
- (iv) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 137;

(v) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 141;

(vi) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 146;

(vii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and

(viii) an isolated bacterial strain comprising a 16S rDNA sequence with at least 97% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

22. The kit of claim 21, wherein:

(i) the isolated bacterial strain of (i) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 124;

(ii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 129;

(iii) the isolated bacterial strain of (ii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 132;

(iv) the isolated bacterial strain of (iv) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 137;

(v) the isolated bacterial strain of (v) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 141;

(vi) the isolated bacterial strain of (vi) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 146;

(vii) the isolated bacterial strain of (vii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 152; and

(viii) the isolated bacterial strain of (viii) comprises a 16S rDNA sequence with at least 99% sequence identity to the nucleotide sequence of SEQ ID NO: 157.

23. A kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

(i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;

(ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;

(iii) an isolated bacterial strain belonging to the species *Drancourtella massiliensis*;

- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

24. A kit when used in treating or preventing *Clostridium difficile* infection, the kit comprising:

- (i) an isolated bacterial strain belonging to the species *Clostridium bolteae*;
- (ii) an isolated bacterial strain belonging to the species *Anaerotruncus colihominis*;
- (iii) an isolated bacterial strain belonging to the species *Sellimonas intestinalis*;
- (iv) an isolated bacterial strain belonging to the species *Clostridium symbiosum*;
- (v) an isolated bacterial strain belonging to the species *Blautia producta*;
- (vi) an isolated bacterial strain belonging to the species *Dorea longicatena*;
- (vii) an isolated bacterial strain belonging to the species *Clostridium innocuum*; and
- (viii) an isolated bacterial strain belonging to the species *Flavonifractor plautii*.

25. The kit of any one of claims 21–24, wherein the kit comprises more than one dosage form, each dosage form comprising at least one of the isolated bacterial strains.

Figure 1

Composition A	Composition B	Composition C	Composition D
SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*	SEQ_10 - 211 - Flavonifractor_plautii (IV)	SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*	SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*
SEQ_04 - 7 - Blautia_hansenii (XIVa)*	SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)	SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*	SEQ_03 - 5 - Clostridium_hathewayi (XIVa)*
SEQ_05 - 10 - Blautia_hansenii (XIVa)*	SEQ_15 - VE202-14 - Eubacterium_fissicatena (XIVa)	SEQ_05 - 10 - Blautia_hansenii (XIVa)*	SEQ_05 - 10 - Blautia_hansenii (XIVa)*
SEQ_07 - 59 - Blautia_producta / Blautia_coccoides (XIVa)	SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)	SEQ_01 - 71 - Blautia_wexlerae (XIVa)*	SEQ_01 - 71 - Blautia_wexlerae (XIVa)*
SEQ_08 - 79 - Blautia_hansenii (XIVa)*	SEQ_17 - VE202-7 - Clostridium_bolteae (XIVa)	SEQ_07 - 59 - Blautia_producta/Blautia_coccoides (XIVa)*	SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)
SEQ_09 - VE202-21 - Eubacterium_contortum / Eubacterium_fissicatena (XIVa)*	SEQ_20 - 170 -- Dorea longicatena (XIVa)	SEQ_18 - 148 - Dorea_longicatena (XIVa)	SEQ_18 - 148 - Dorea_longicatena (XIVa)
SEQ_11 - VE202-9 - Anaerostipes_caccae (XIVa)*	SEQ_19 - 16 - Blautia_producta (XIVa)	SEQ_21 - 189 - Clostridium_innocuum (XVII)	SEQ_21 - 189 - Clostridium_innocuum (XVII)
SEQ_12 - VE202-26 - Clostridium_scindens (XIVa)*	SEQ_21 - 189 - Clostridium_innocuum (XVII)	SEQ_10 - 211 - Flavonifractor_plautii (IV) / (IV) /	SEQ_10 - 211 - Flavonifractor_plautii (IV) /
SEQ_13 - 136 - Marvinbryantia_formatexigens (XIVa)*		SEQ_14 - VE202-13 - Anaerotruncus_colihominis (IV)	SEQ_02 - 102 - Turicibacter_sanguinis (non-Clostridium)
SEQ_23 - VE202-29 - Eisenbergiella_tayi (XIVa)*		SEQ_16 - VE202-16 - Clostridium_symbiosum (XIVa)	SEQ_06 - 40 - Lactobacillus_mucosae (non-Clostridium)

* = BaiCD⁺***bolding indicates strains other than Clostridium cluster XIVa***

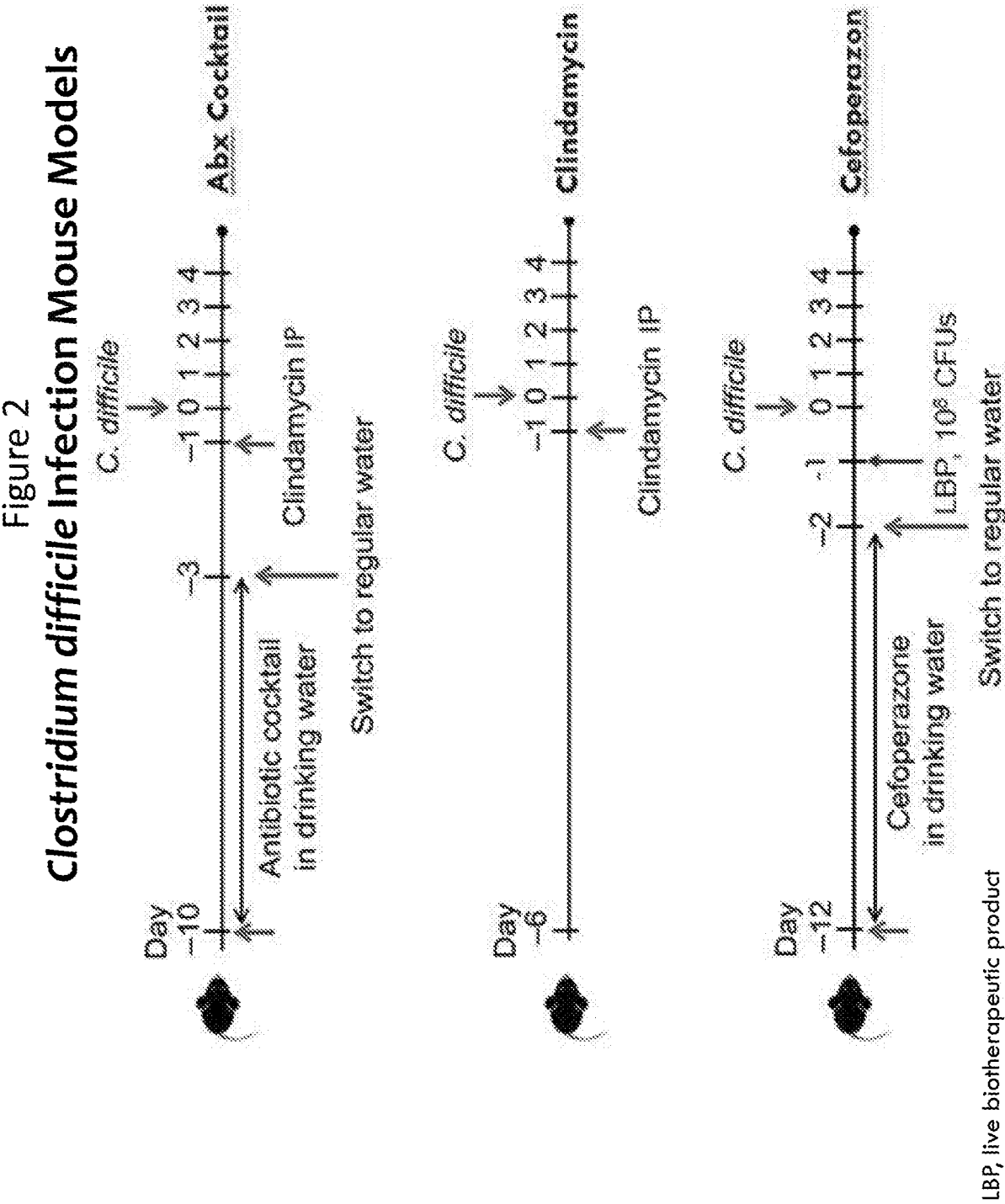


Figure 3

Groups	# of animals	Abx	<i>C. difficile</i> Spores
(1) Control	5	-	10 ¹
(2) Control	5	-	10 ⁴
(3) Abx cocktail	5	+	10 ¹
(4) Abx cocktail	5	+	10 ⁴
(5) Clindamycin	5	+	10 ¹
(6) Clindamycin	5	+	10 ⁴
(7) Cefoperazone	5	+	10 ¹
(8) Cefoperazone	5	+	10 ⁴

Figure 4A

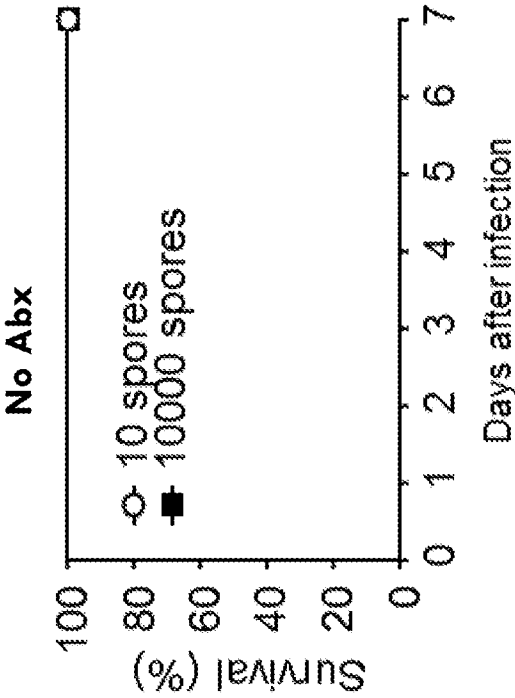


Figure 4B

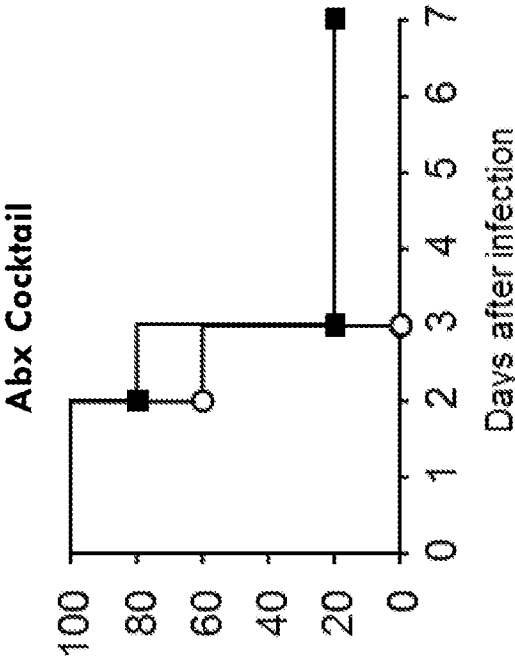


Figure 4C

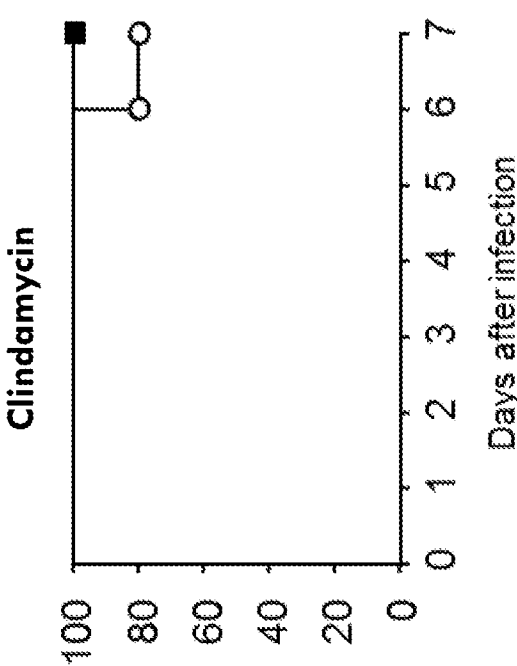


Figure 4D

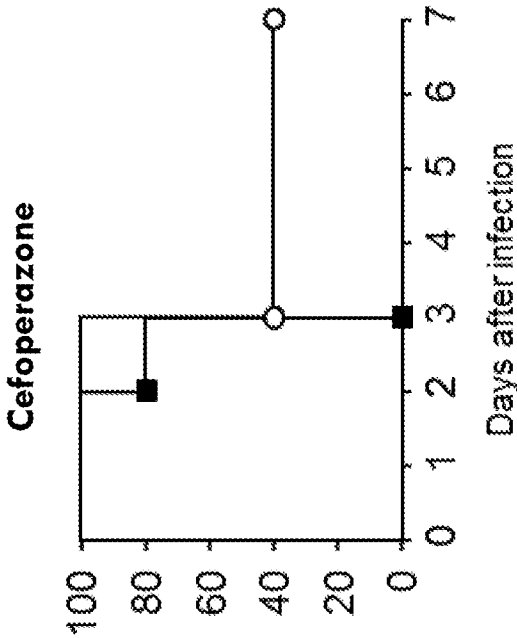


Figure 4E

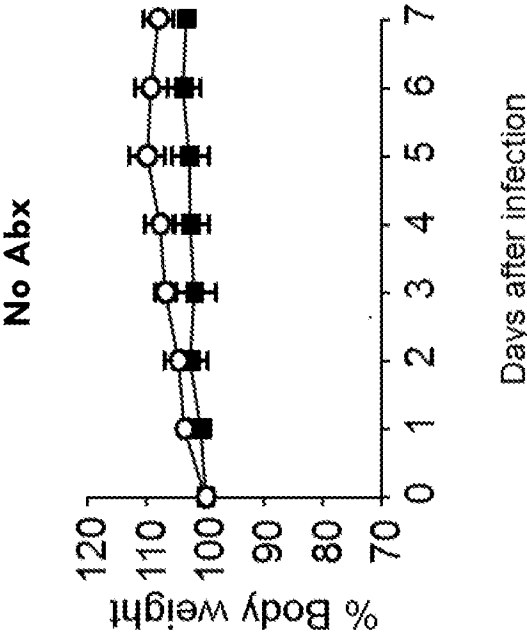


Figure 4F

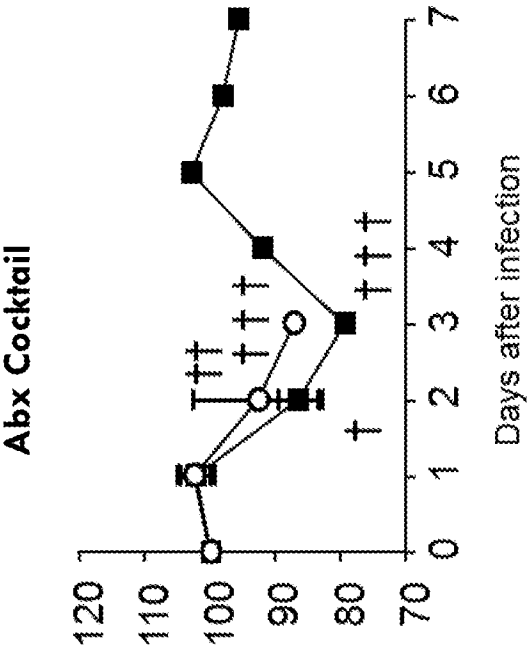


Figure 4G

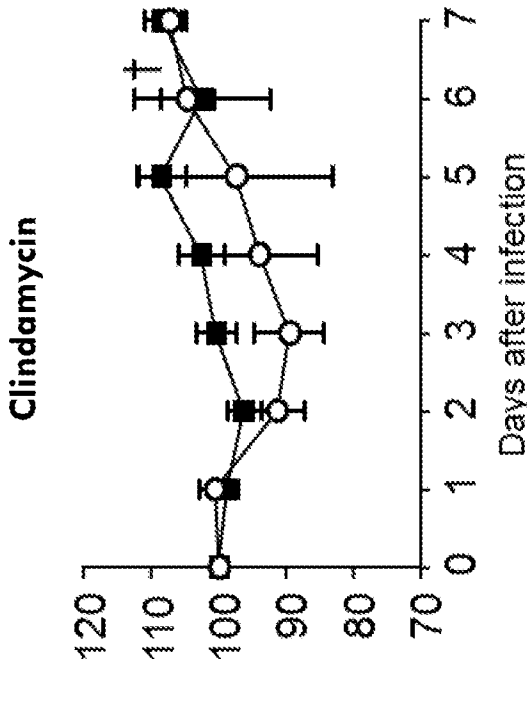
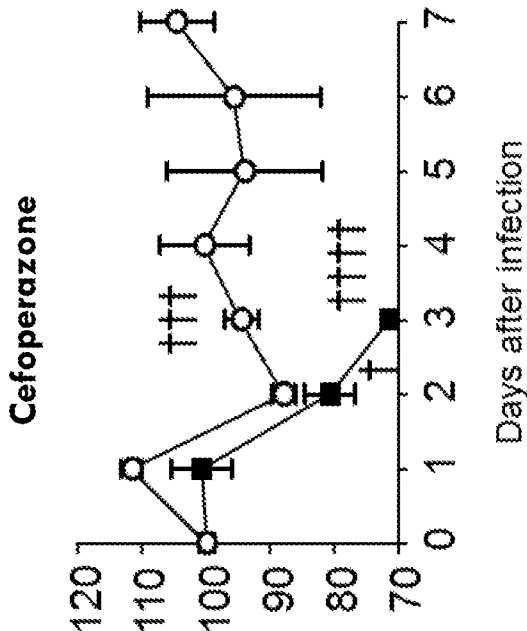


Figure 4H



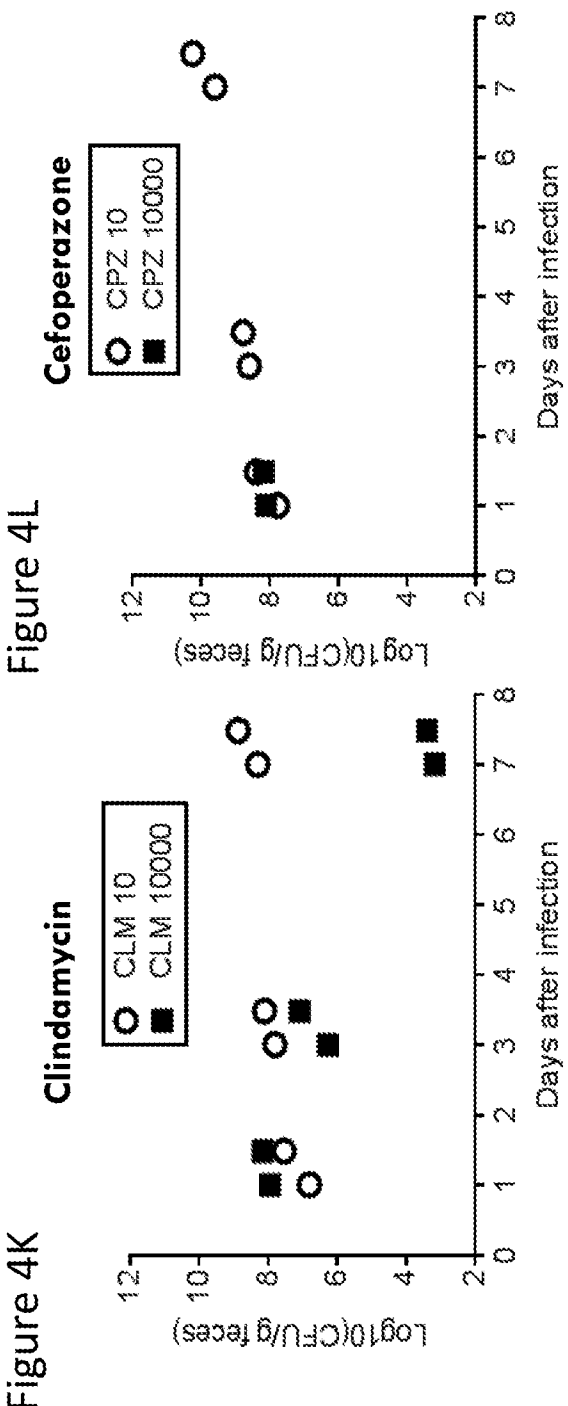
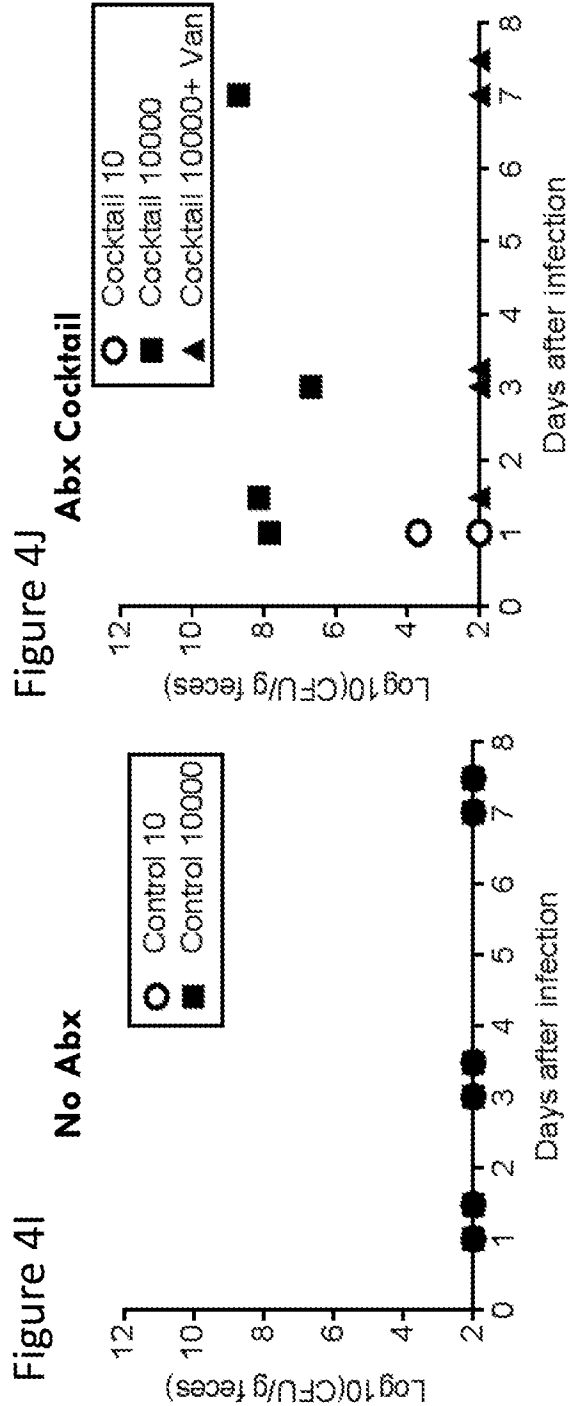
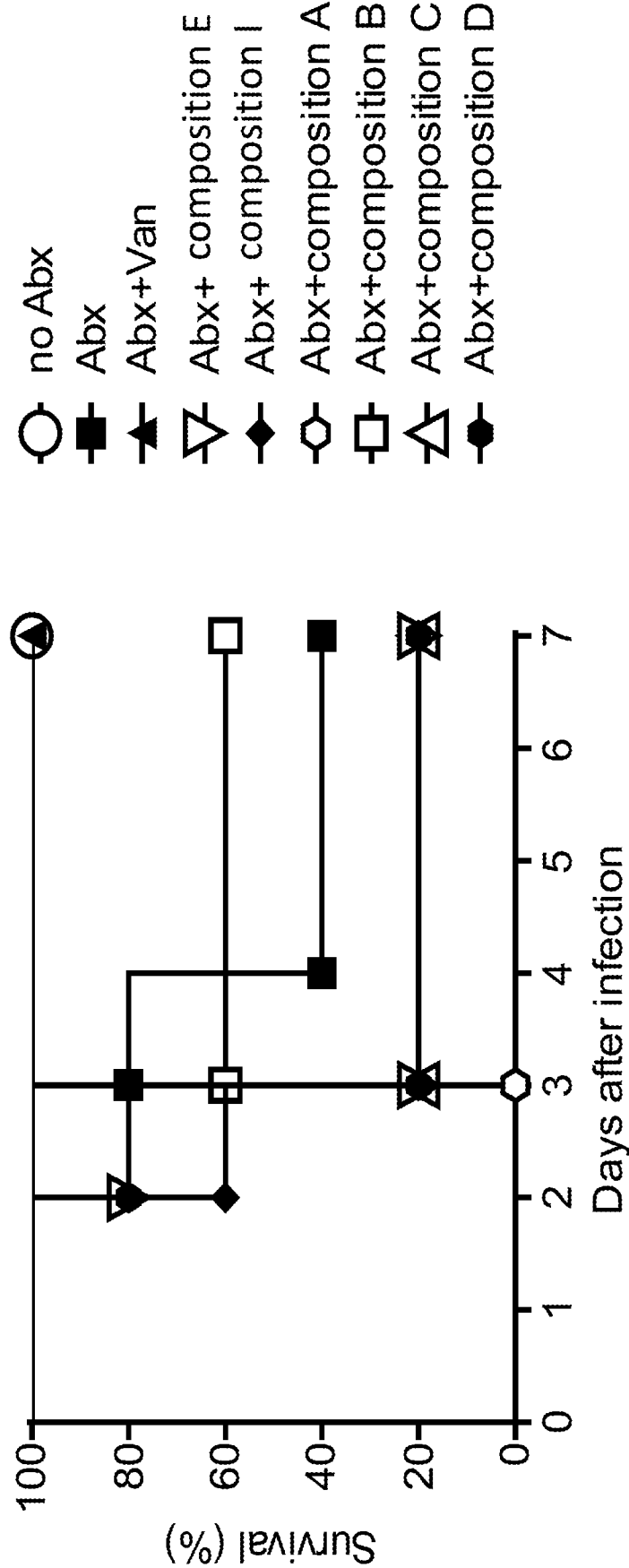
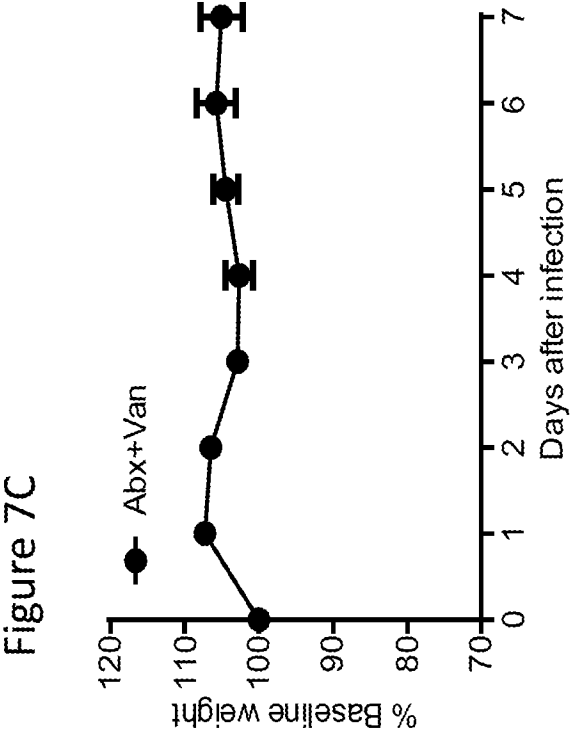
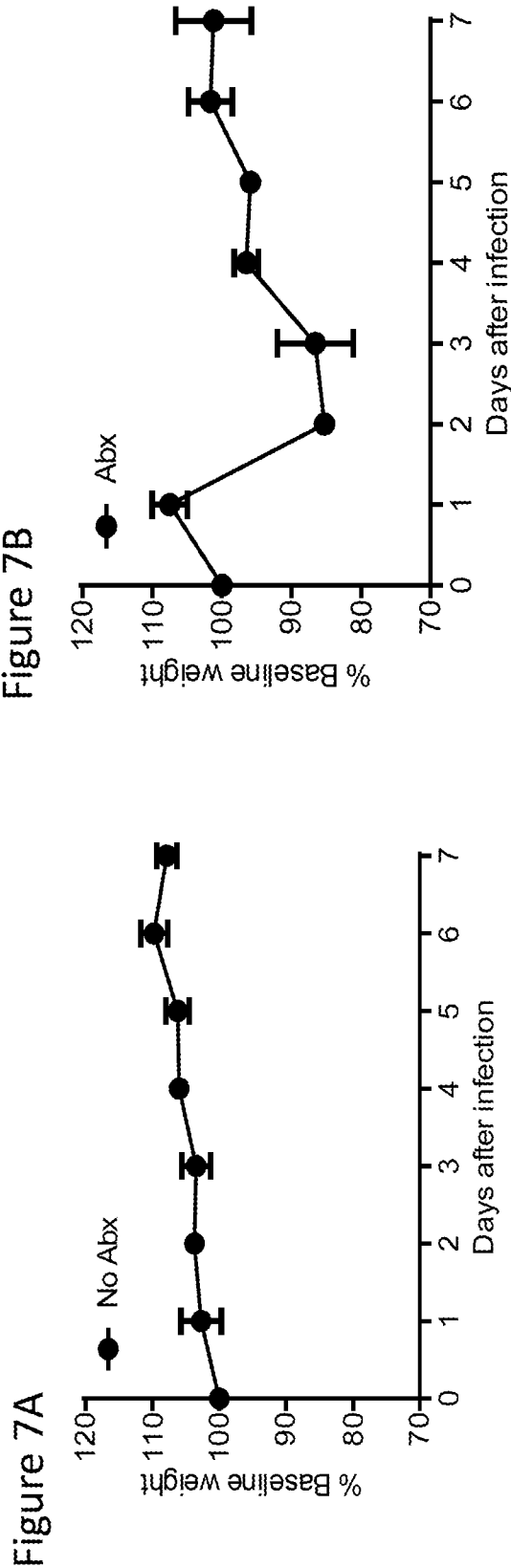


Figure 5

Groups	# of animals	Abx	CFUs (Spores)
(1) Control-	5	-	10 ⁴
(2) Control+	5	+	10 ⁴
(3) Van	5	+	10 ⁴
(4) Composition E	5	+	10 ⁴
(5) Composition I	5	+	10 ⁴
(6) Composition A	5	+	10 ⁴
(7) Composition B	5	+	10 ⁴
(8) Composition C	5	+	10 ⁴
(9) Composition D	5	+	10 ⁴

Figure 6





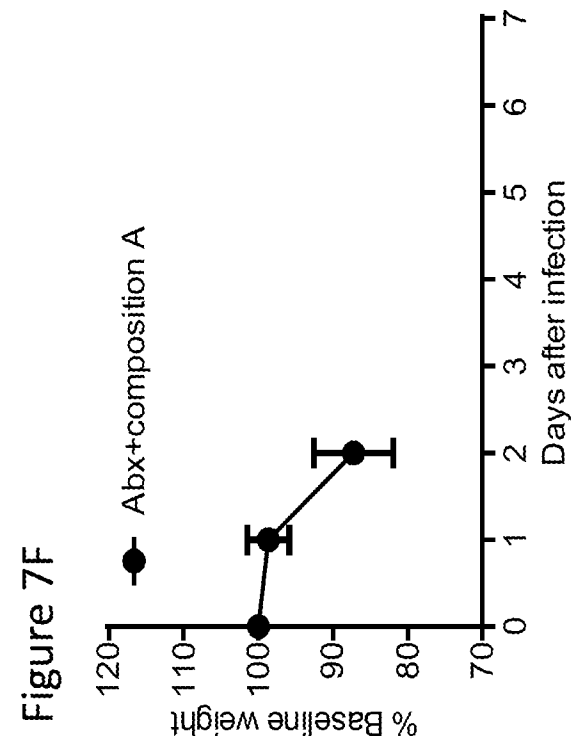
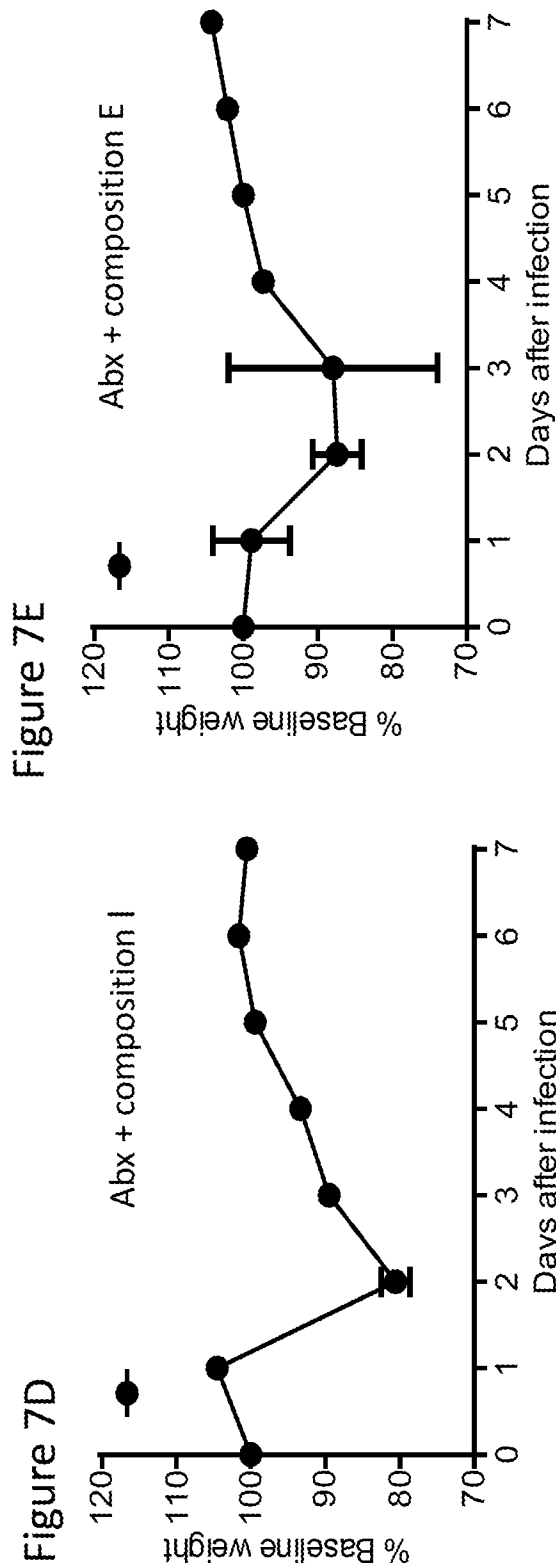


Figure 7H

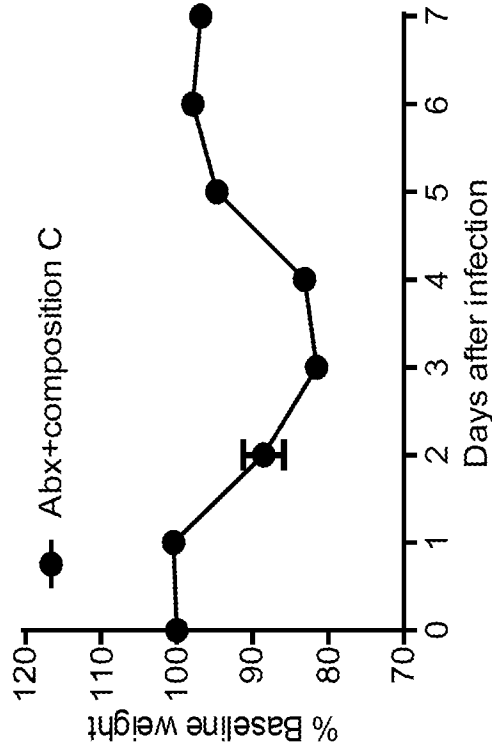


Figure 7G

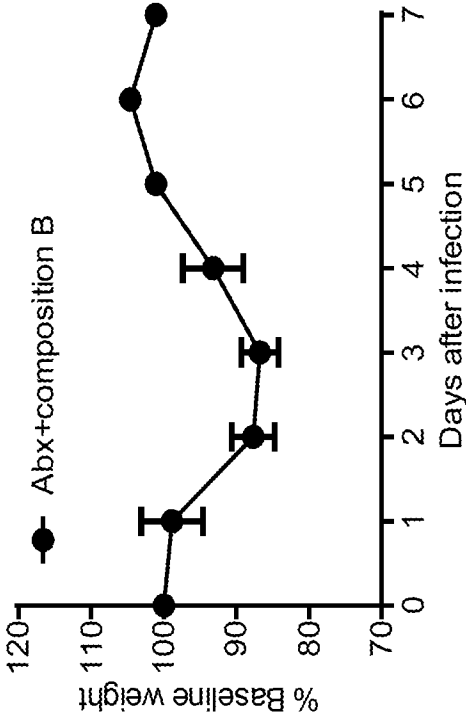
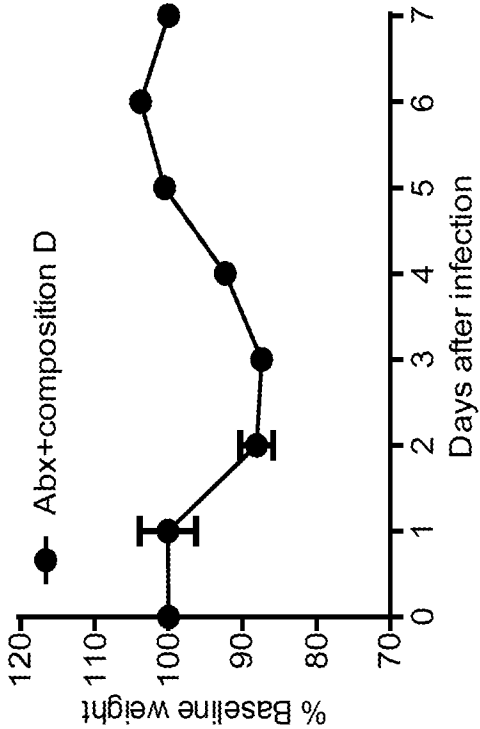
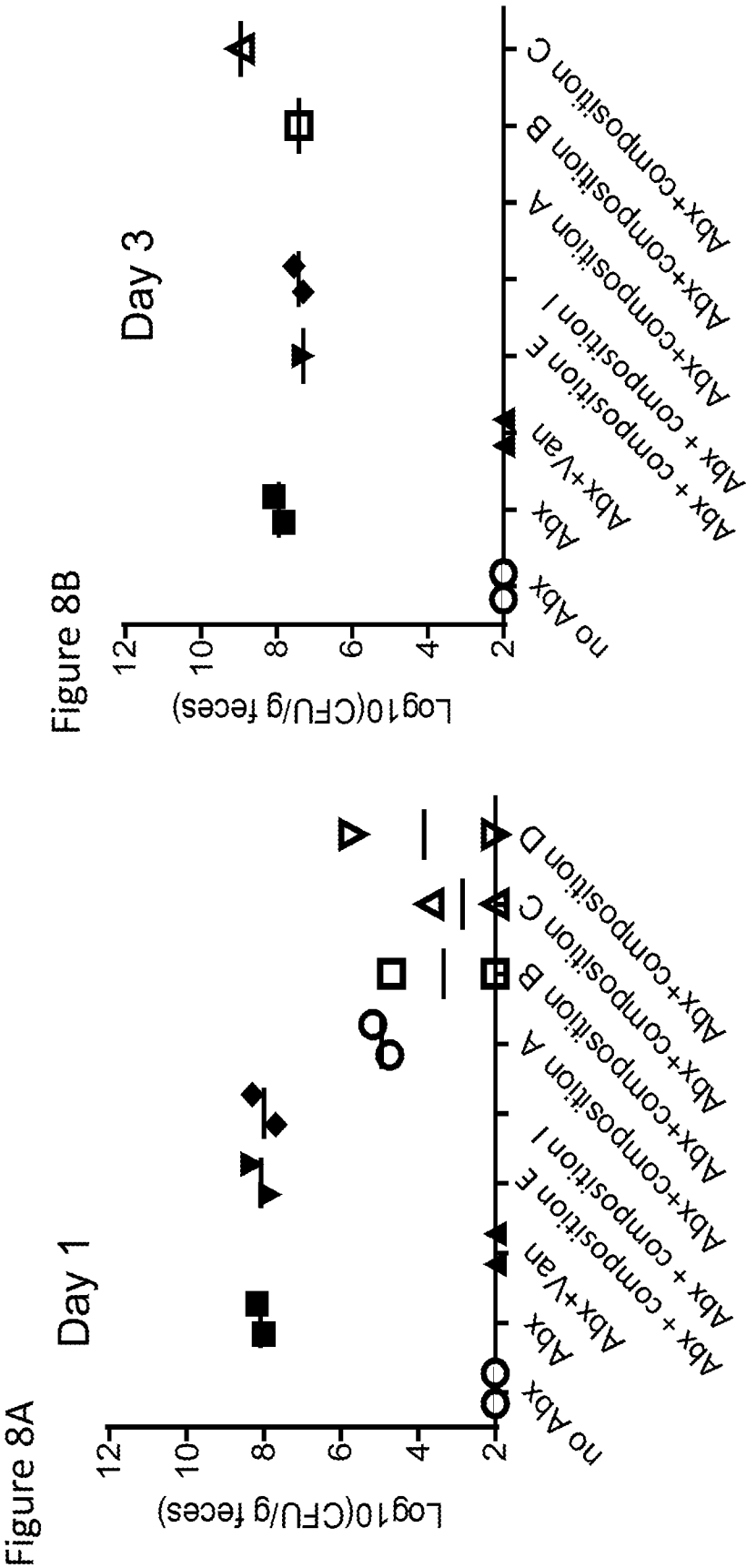


Figure 7I





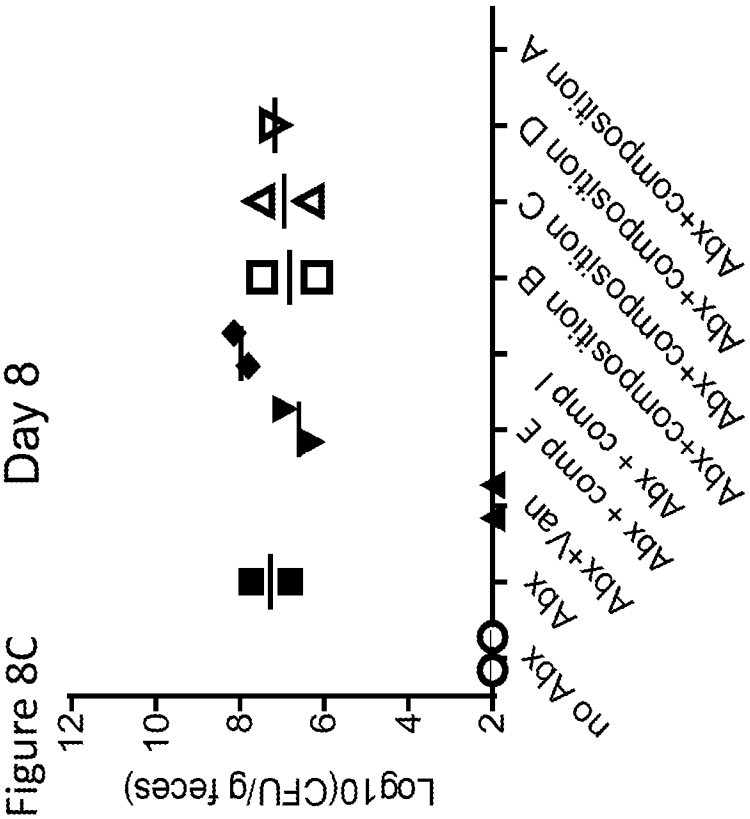
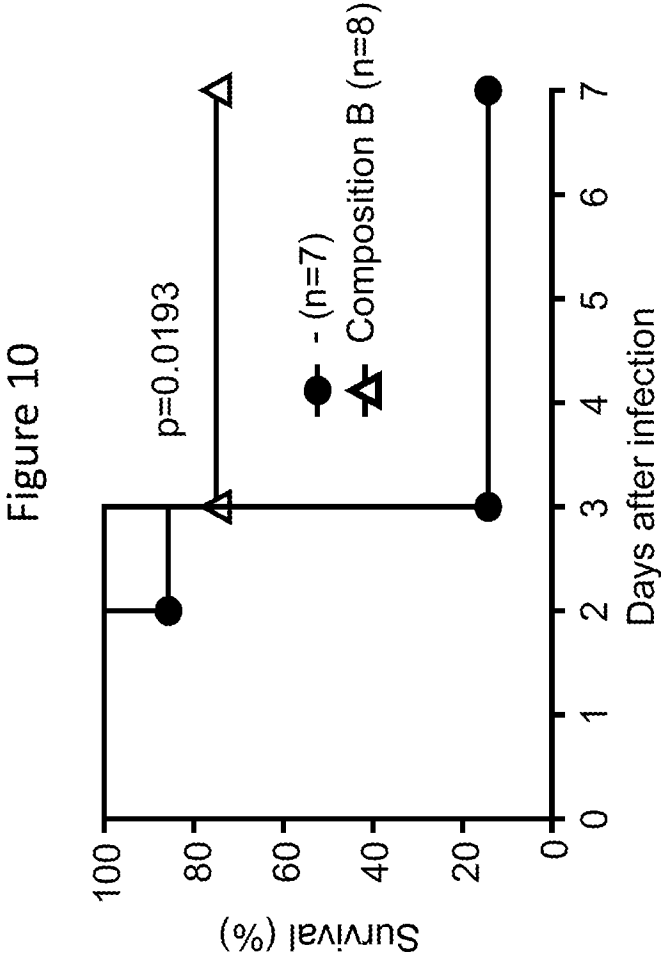


Figure 9

Groups	# of animals	Abx	<i>C. difficile</i> spore	CFUs LBP's
(1) Control	7	+	10 ⁴	-
(2) Composition B	8	+	10 ⁴	10 ⁸ /mouse



(Natural death + >20% weight loss)

Figure 11

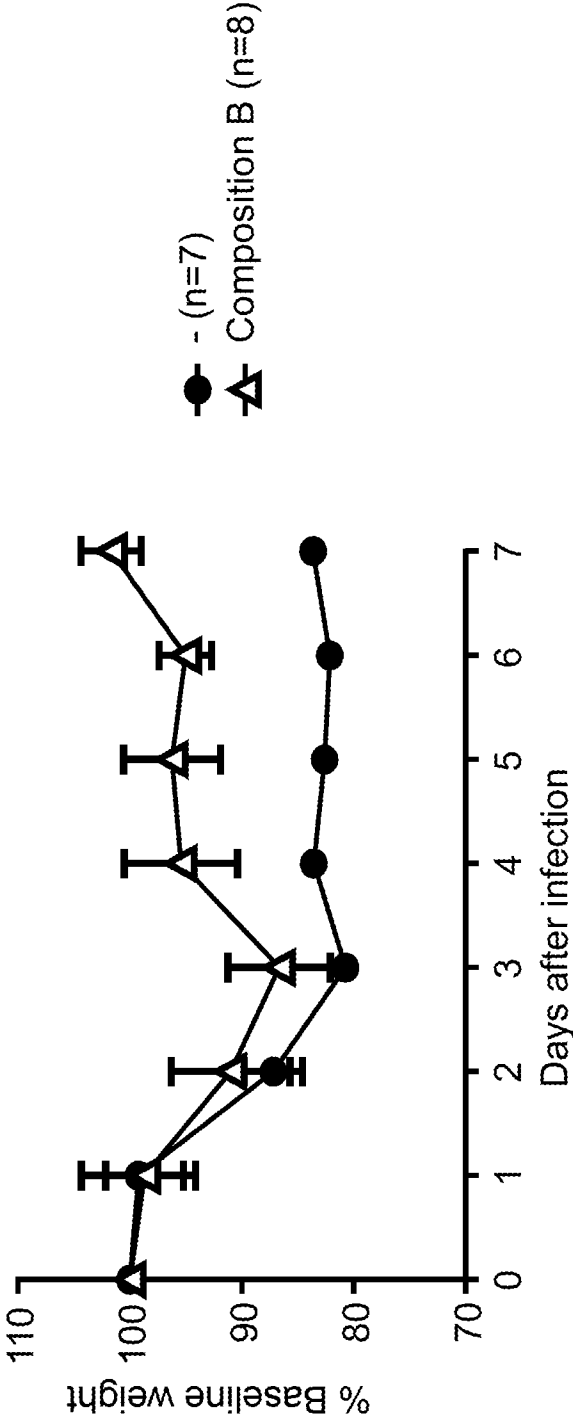


Figure 12

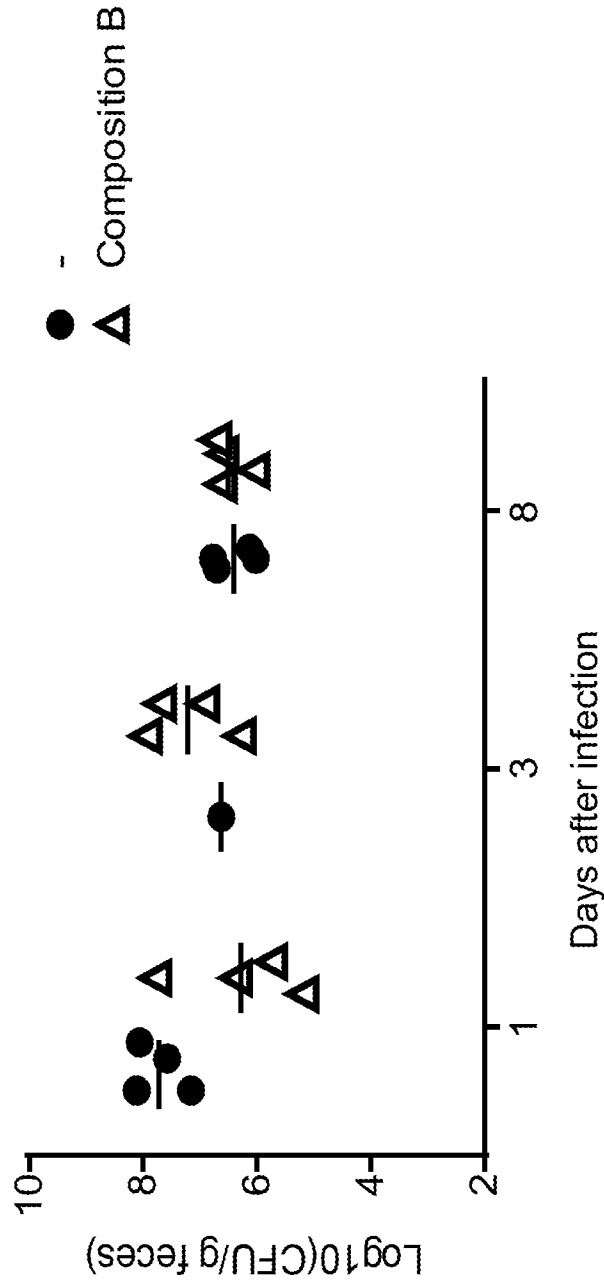


Figure 13

Composition F

SEQ_NO	StrainID	Genus_species	SEQ_NO	StrainID	Genus_species
SEQ_24	YK96	Dorea_longicatena	SEQ_52	YK51	Eubacterium_rectale
SEQ_25	YK101	Ruminococcus_obeum	SEQ_53	YK52	Eubacterium_rectale
SEQ_26	YK110	Megasphaera_elsdenii	SEQ_54	YK54	Anaerostipes_hadrus
		Acidaminococcus_fermentans /			
SEQ_27	YK149	Acidaminococcus_intestini	SEQ_55	YK56	Ruminococcus_faecis
SEQ_28	YK154	Megasphaera_elsdenii	SEQ_56	YK57	Ruminococcus_faecis
SEQ_29	YK36	Ruminococcus_faecis	SEQ_57	YK58	Dorea_longicatena
SEQ_30	YK95	Bacteroides_cellulosilyticus	SEQ_58	YK65	Roseburia_faecis
SEQ_31	YK32	Anaerostipes_hadrus	SEQ_59	YK67	Blautia_luti
SEQ_32	YK64	Ruminococcus_obeum	SEQ_60	YK69	Fusicatenibacter_saccharivorans
SEQ_33	YK73	Flavonifractor_plautii	SEQ_61	YK70	Fusicatenibacter_saccharivorans
SEQ_34	YK87	Eubacterium_rectale	SEQ_62	YK71	Roseburia_faecis
SEQ_35	YK105	Flavonifractor_plautii	SEQ_63	YK74	Megasphaera_elsdenii
SEQ_36	YK153	Megasphaera_elsdenii	SEQ_64	YK88	Eubacterium_rectale
SEQ_37	YK163	Eubacterium_rectale	SEQ_65	YK89	Eubacterium_rectale
		Ruminococcus_champanellensis /			
SEQ_38	YK191	Ruminococcus_albus	SEQ_66	YK97	Roseburia_faecis
SEQ_39	YK99	Ruminococcus_champanellensis	SEQ_67	YK98	Blautia_faecis
SEQ_40	YK55	Ruminococcus_faecis	SEQ_68	YK139	Fusicatenibacter_saccharivorans
SEQ_41	YK75	Bifidobacterium_bifidum	SEQ_69	YK141	Dorea_formicigenerans
SEQ_42	YK90	Anaerostipes_hadrus	SEQ_70	YK142	Ruminococcus_faecis
SEQ_43	YK30	Anaerostipes_hadrus	SEQ_71	YK152	Blautia_hansenii
SEQ_44	YK31	Anaerostipes_hadrus	SEQ_72	YK155	Blautia_hansenii
SEQ_45	YK12	Eubacterium_rectale	SEQ_73	YK157	Eubacterium_rectale
SEQ_46	YK27	Ruminococcus_faecis	SEQ_74	YK160	Roseburia_faecis
SEQ_47	YK28	Blautia_luti	SEQ_75	YK166	Eubacterium_rectale
SEQ_48	YK29	Ruminococcus_faecis	SEQ_76	YK168	Eubacterium_rectale
SEQ_49	YK33	Anaerostipes_hadrus	SEQ_77	YK169	Eubacterium_rectale
SEQ_50	YK34	Anaerostipes_hadrus	SEQ_78	YK171	Eubacterium_rectale
SEQ_51	YK35	Ruminococcus_faecis	SEQ_79	YK192	Roseburia_faecis

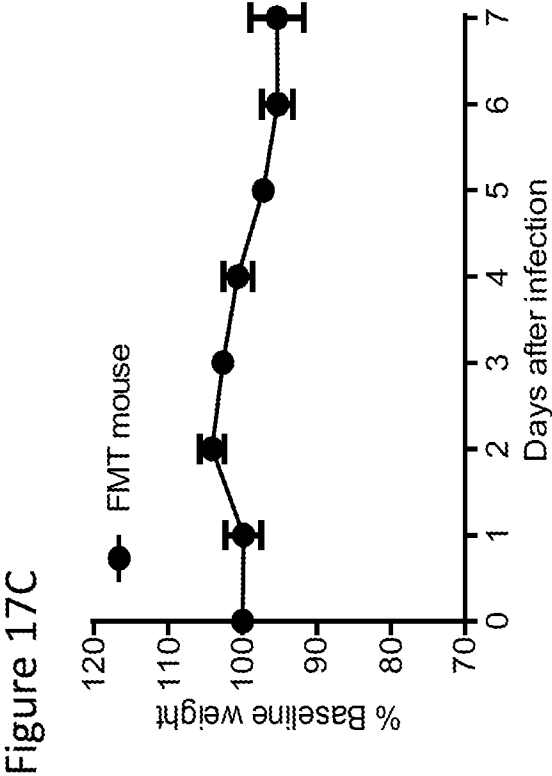
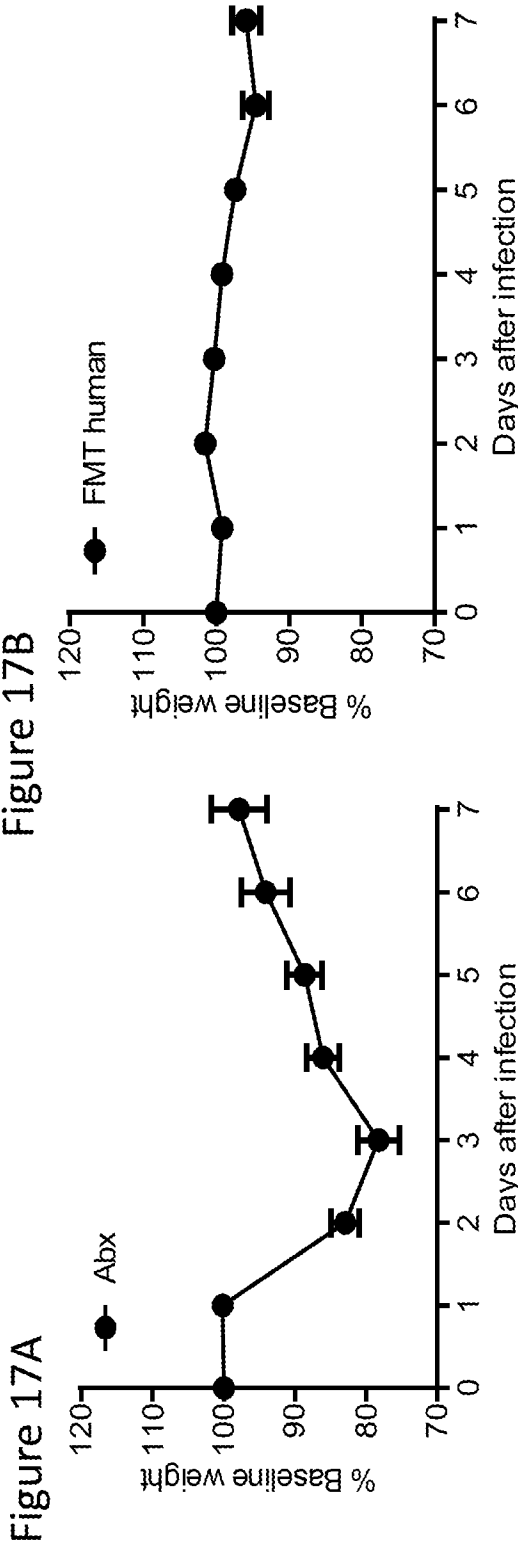
Figure 14

Cluster	Composition F	SCFA s
XIVa	<i>Eubacterium rectale</i> 12	A, B, L
	<i>Ruminococcus faecis</i> 8	A, L
	<i>Ruminococcus obeum</i> 2	A, L
	<i>Blautia faecis</i> 1	A, L
	<i>Blautia hansenii</i> 2	A, L
	<i>Blautia luti</i> 2	A, L
	<i>Anaerostipes hadrus</i> 7	B
	<i>Roseburia faecis</i> 5	A, B
	<i>Fusicatenibacter saccharivorans</i> 3	A, L
	<i>Dorea formicigenerans</i> 1	A
IV	<i>Dorea longicatena</i> 2	A
	<i>Flavonifractor plautii</i> 2	A, B
IX	<i>Ruminococcus champanellensis</i> 2	A
	<i>Acidaminococcus fermentans</i> 1	A, B, P
	<i>Megasphaera elsdeni</i> 4	P
other	<i>Bacteroides cellulosilyticus</i> 1	A, S
	<i>Bifidobacterium Bifidum</i>	L, A

A, acetate;
B, Butyrate;
L, lactate;
P, propionate;
S, succinate

Figure 15

Groups	# of animals	Abx	<i>C. difficile</i> spore	CFUs LBPs
(1) Control	10	+	10 ⁴	-
(2) Composition B dosed at day -1	10	+	10 ⁴	10 ⁸ /mouse
(3) Composition B dosed at day -2 and -1	10	+	10 ⁴	10 ⁸ /mouse
(4) Composition B dosed at day -2, -1, 1, 2, and 3	10	+	10 ⁴	10 ⁸ /mouse
(5) Composition F dosed at day -1	5	+	10 ⁴	OD Normalized
(6) Composition F dosed at day -2, -1, 1, 2, and 3	5	+	10 ⁴	OD Normalized
(7) FMT mouse	5	+	10 ⁴	200ul of 10% fecal samples/mouse
(8) FMT human	5	+	10 ⁴	200ul of 10% fecal samples/mouse



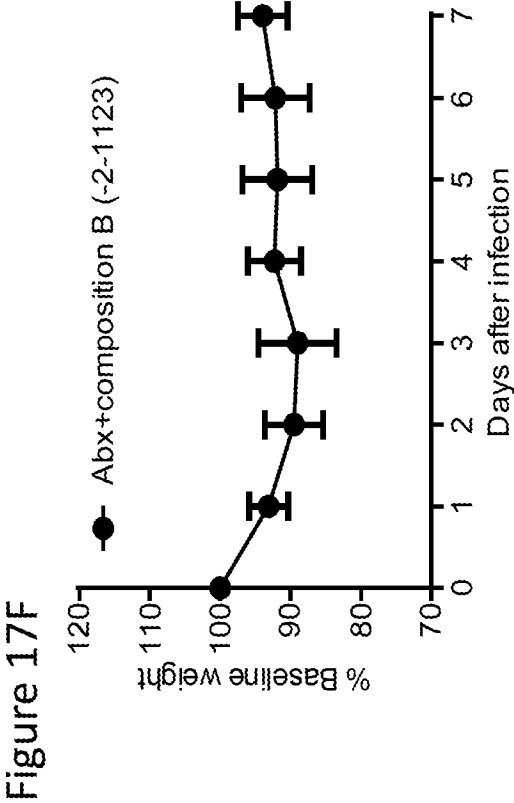
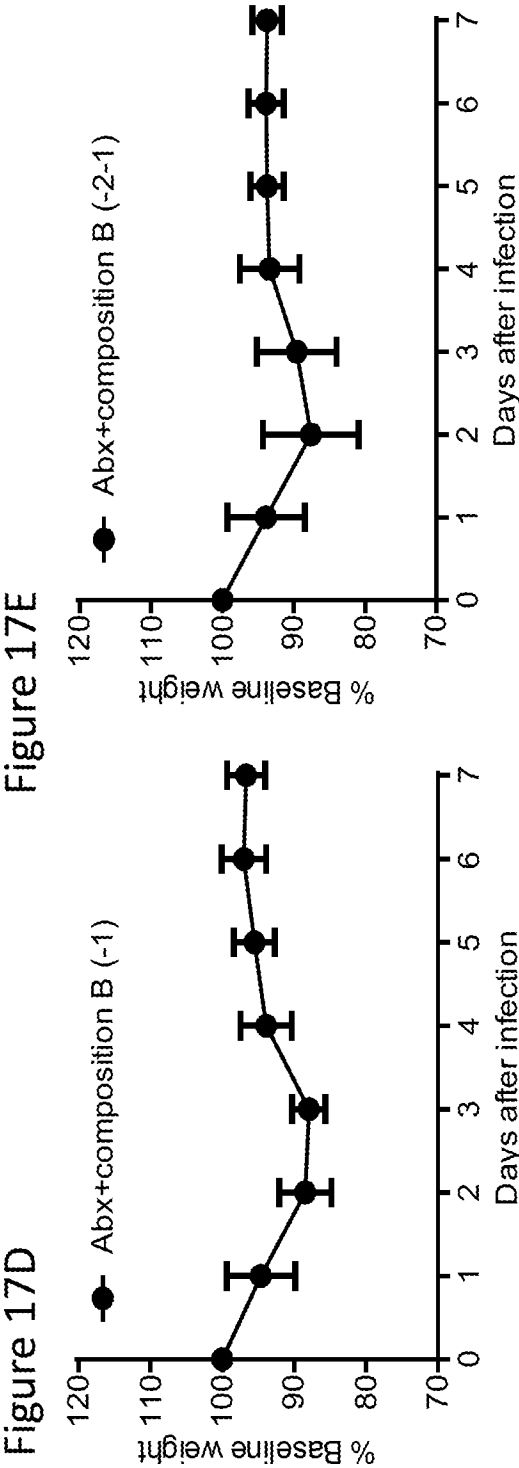


Figure 17G

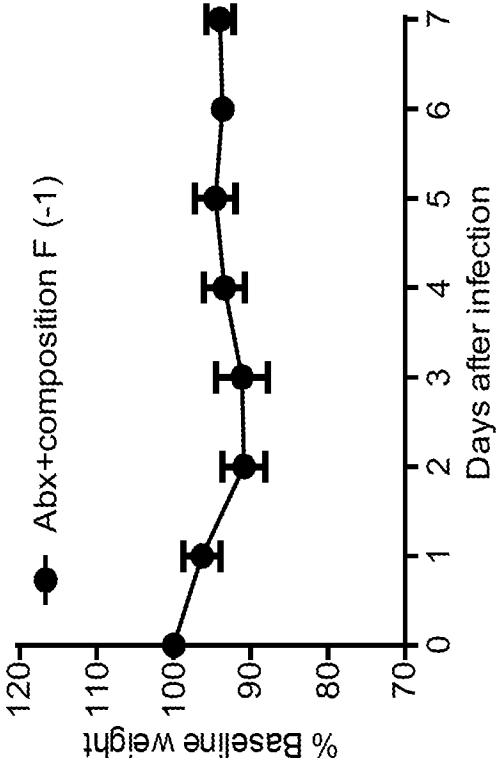
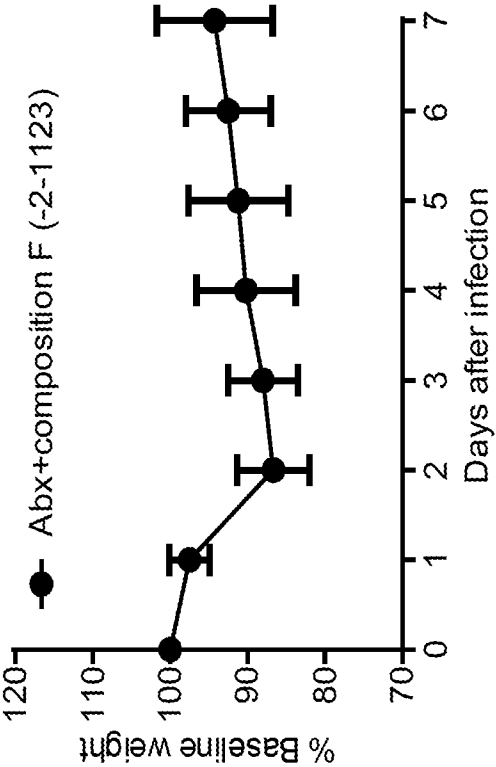


Figure 17H



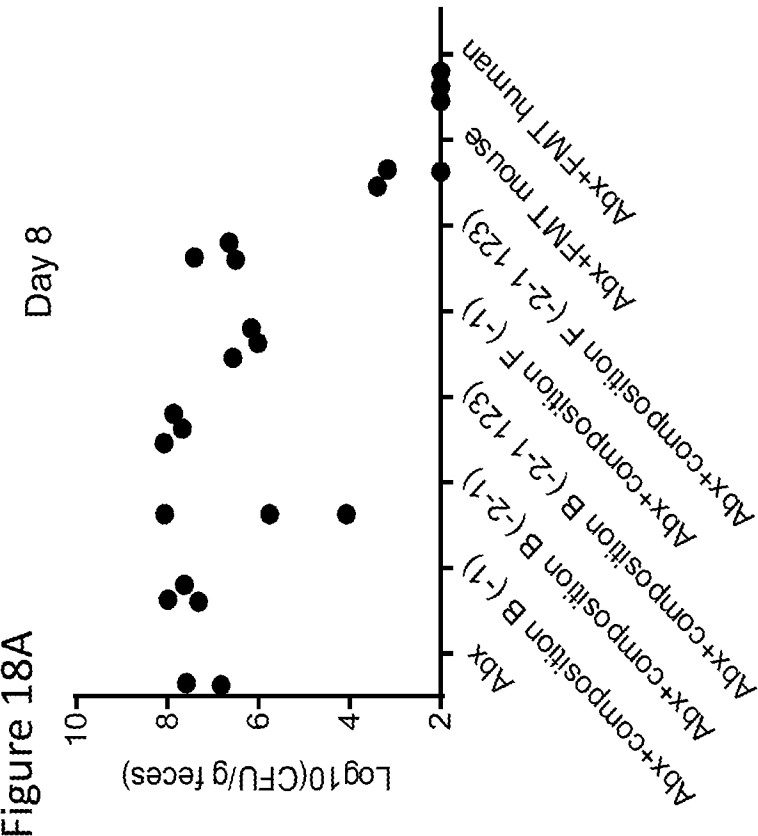
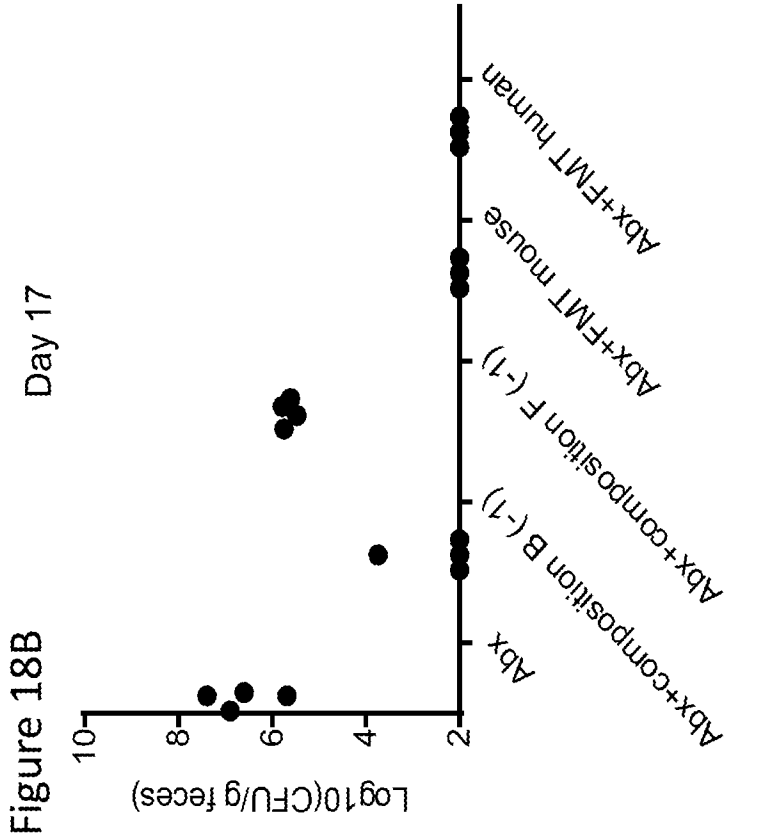


Figure 19
Composition G

SEQ_27	YK149	Acidaminococcus_fermentans/Acidaminococcus_intesti
SEQ_43	YK90	Anaerostipes_hadrus
SEQ_44	YK30	Anaerostipes_hadrus
SEQ_51	YK34	Anaerostipes_hadrus
SEQ_55	YK54	Anaerostipes_hadrus
SEQ_68	YK98	Blautia_faecis
SEQ_72	YK152	Blautia_hansenii
SEQ_70	YK141	Dorea_formicigenerans
SEQ_24	YK96	Dorea_longicatena
SEQ_34	YK87	Eubacterium_rectale
SEQ_37	YK163	Eubacterium_rectale
SEQ_46	YK12	Eubacterium_rectale
SEQ_76	YK166	Eubacterium_rectale
SEQ_77	YK168	Eubacterium_rectale
SEQ_35	YK105	Flavonifractor_plautii
SEQ_62	YK70	Fusicatenibacter_saccharivorans
SEQ_26	YK110	Megasphaera_elsdenii
SEQ_63	YK71	Roseburia_faecis
SEQ_67	YK97	Roseburia_faecis
SEQ_40	YK99	Ruminococcus_champanellensis
SEQ_38	YK191	Ruminococcus_champanellensis/Ruminococcus_albus
SEQ_47	YK27	Ruminococcus_faecis
SEQ_56	YK56	Ruminococcus_faecis
SEQ_25	YK101	Ruminococcus_obeum
SEQ_32	YK64	Ruminococcus_obeum

Figure 20

Groups	N	Abx	CFUs <i>C. difficile</i>	CFUs LBPs
(1) Vehicle	7	+	10 ⁴	200ul of PBS
(2) Composition B	8	+	10 ⁴	10 ⁸ /mouse
(3) Composition B1	8	+	10 ⁴	10 ⁸ /mouse
(4) Composition B2	8	+	10 ⁴	10 ⁸ /mouse
(5) Composition F	7	+	10 ⁴	OD Normalized
(6) Composition G	7	+	10 ⁴	OD Normalized
(7) EtOH treated Human fecal samples	5	+	10 ⁴	200ul of 10% fecal samples/mouse
(8) EtOH treated Composition B	5	+	10 ⁴	10 ⁸ /mouse
(9) Frozen Composition B	5	+	10 ⁴	10 ⁸ /mouse
(10) EtOH treated Composition J	5	+	10 ⁴	colony scrapes

Figure 22A

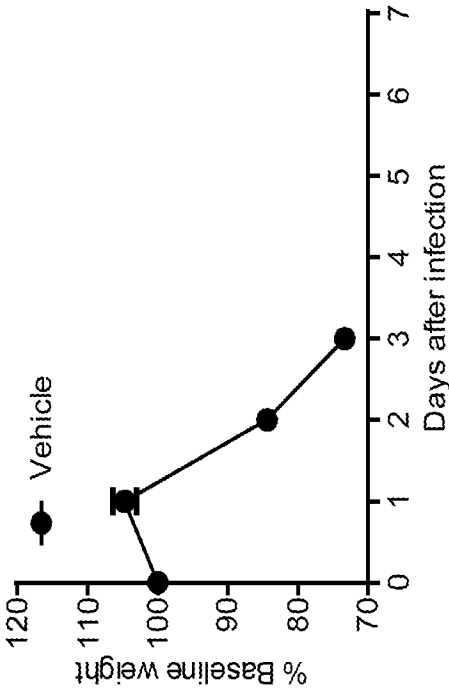


Figure 22B

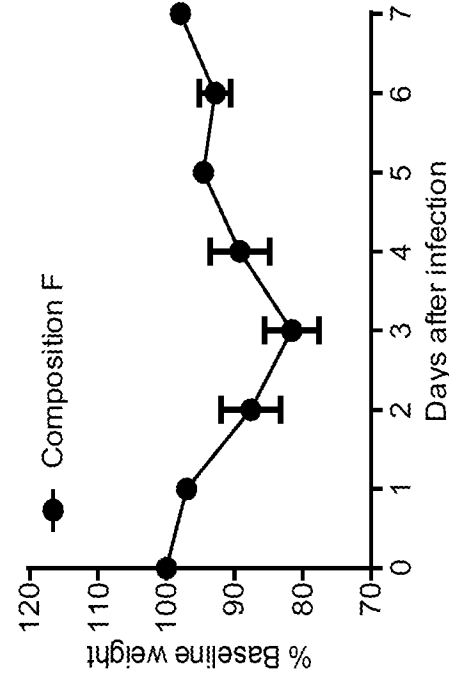


Figure 22C

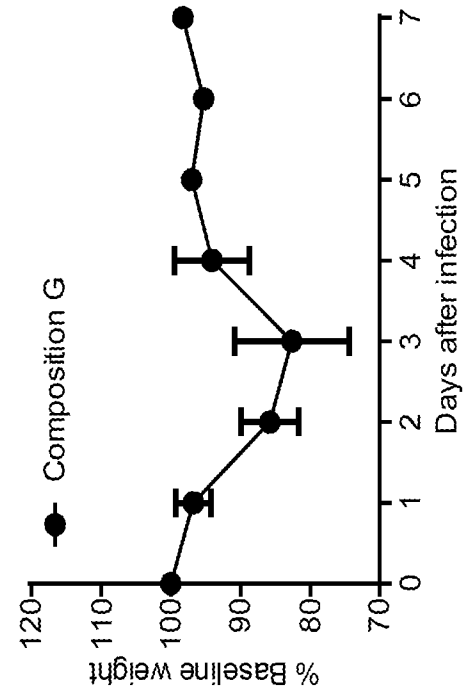


Figure 22E

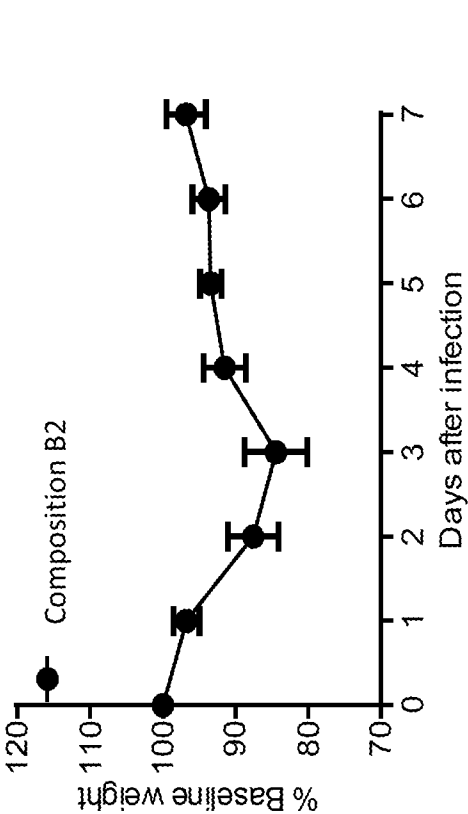


Figure 22G

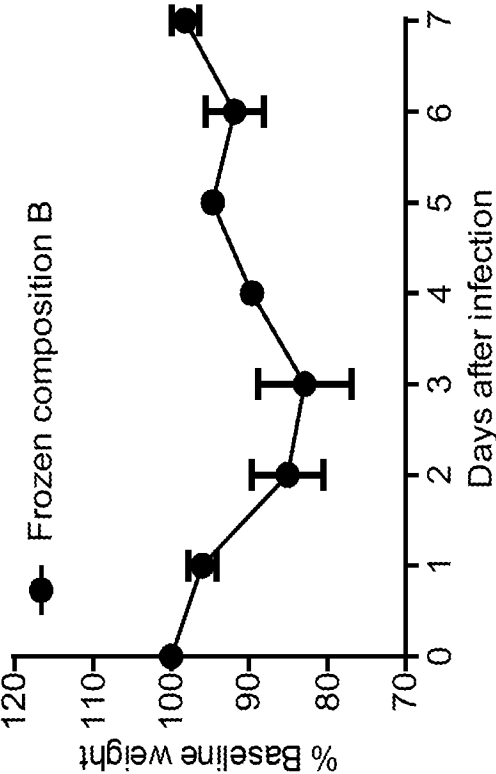


Figure 22D

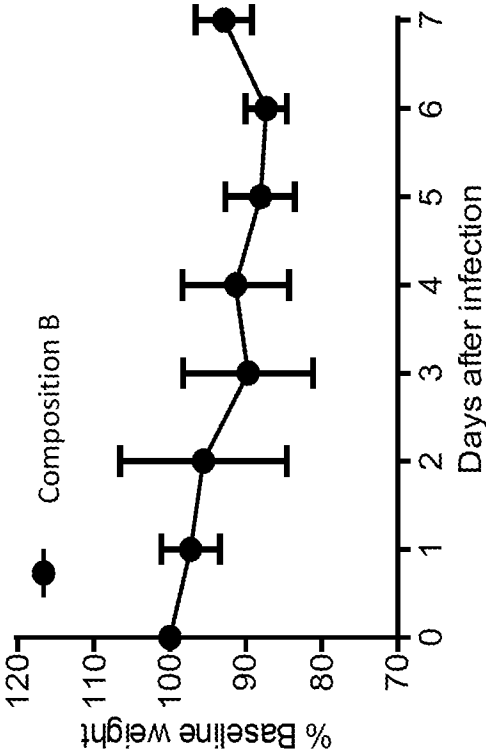


Figure 22F

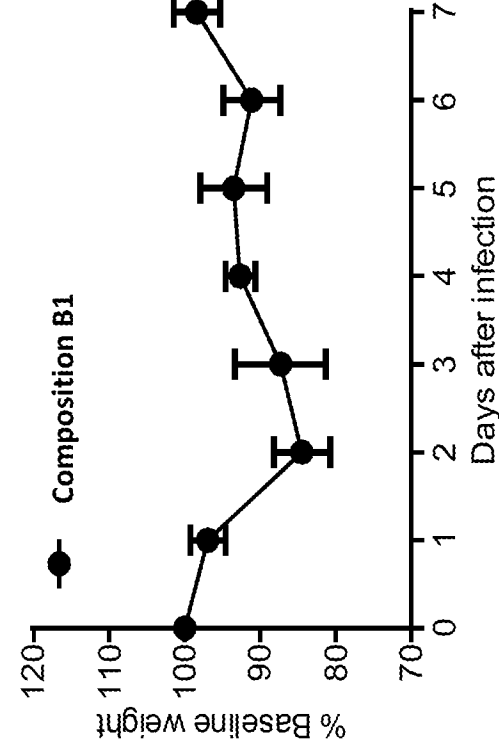


Figure 22I

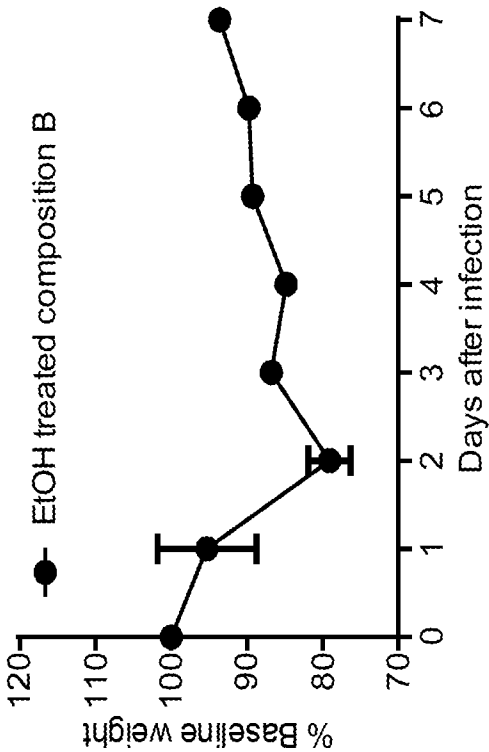


Figure 22H

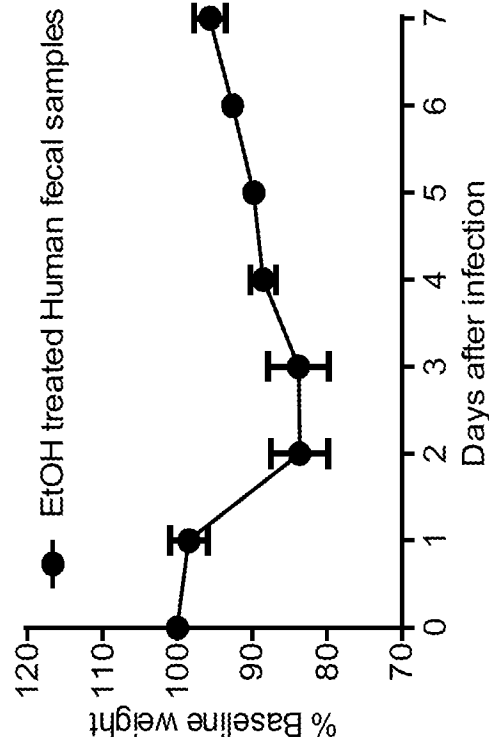
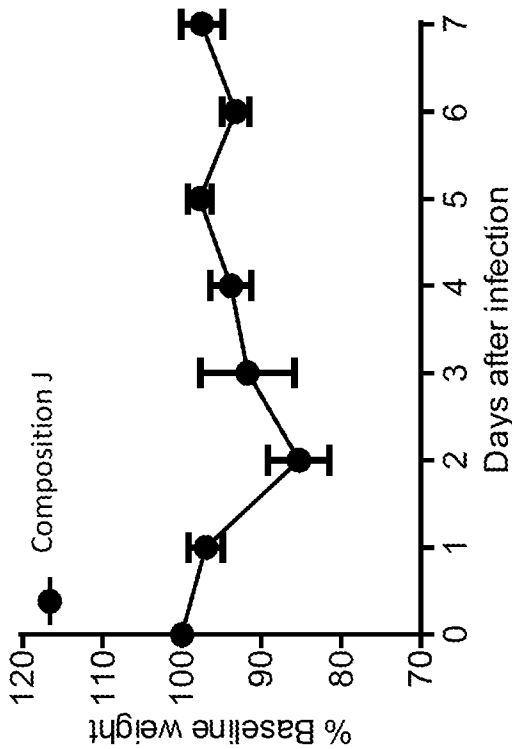


Figure 22J



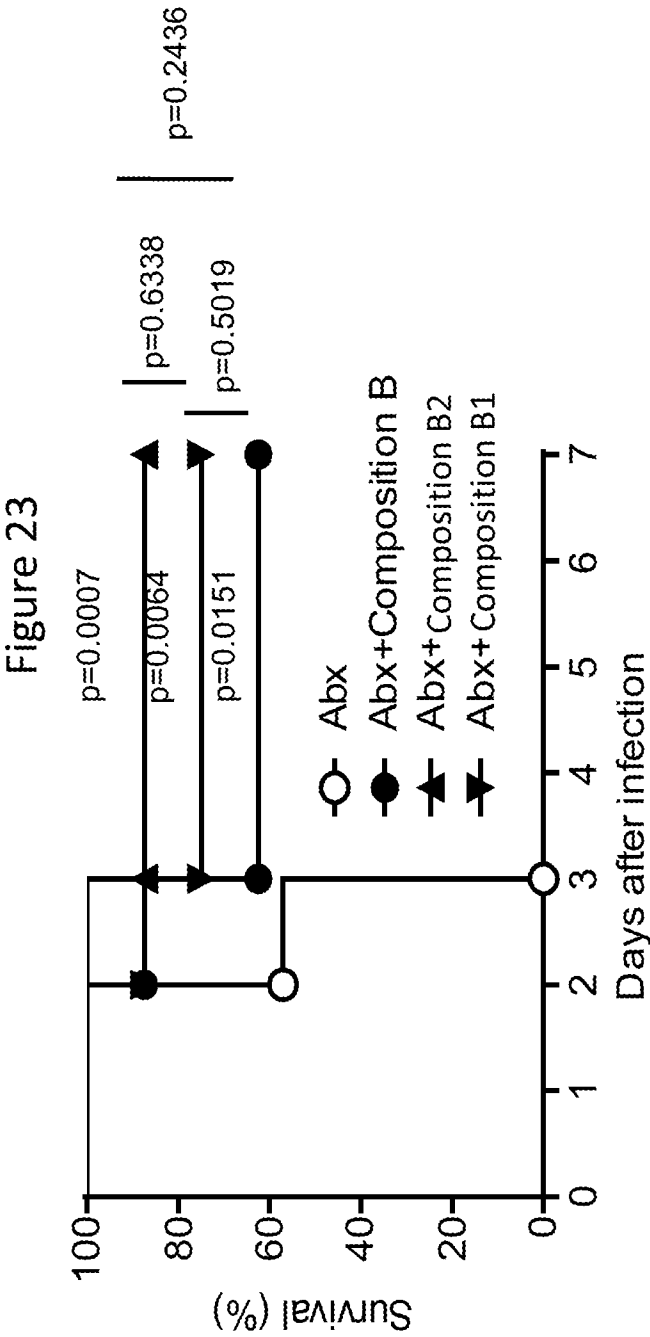


Figure 24

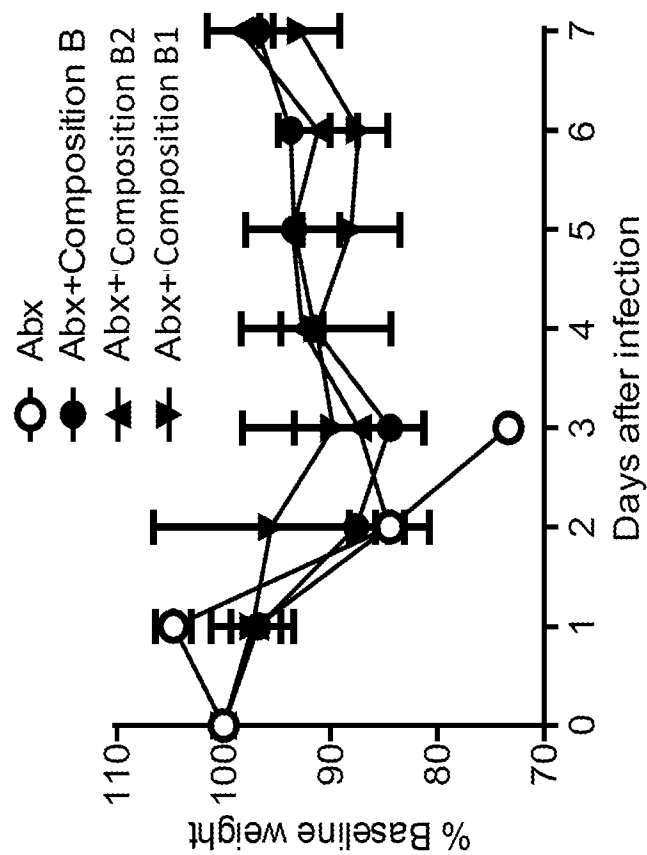


Figure 25

Groups	N	Abx	CFUs <i>C. difficile</i>	CFUs LBPs
(1) Vehicle	10	+	10 ⁴	200ul of PBS
(2) human FMT	10	+	10 ⁴	200ul of 10% fecal samples/mouse
(3) Composition B	10	+	10 ⁴	10 ⁸ /mouse
(4) Composition B + 4 spores*	10	+	10 ⁴	10 ⁸ live bacteria+spores/mouse
(5) Composition H**	10	+	10 ⁴	10 ⁸ /mouse

*Composition B + 4 spores = The strains of Composition B plus the following four strains in spore form: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum*, and *Clostridium innocuum*

**Composition H contains the following six strains in spore form: *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum*, *Clostridium innocuum*, *Clostridium disporicum*, and *Erysipelatoclostridium ramosum*

Figure 26
Composition H

**Composition H = The following six strains in spore form *Clostridium bolteae*, *Anaerotruncus colihominis*, *Clostridium symbiosum*, *Clostridium innocuum*, *Clostridium disporicum* and *Erysipelatoclostridium ramosum*

Composition H sequence info:

SEQ ID NO: 14 - VE202-13 – <i>Anaerotruncus colihominis</i>	Cluster IV
SEQ ID NO: 16 - VE202-16 – <i>Clostridium symbiosum</i>	Cluster XIVa
SEQ ID NO: 21 - 189 – <i>Clostridium innocuum</i>	Cluster XVII
SEQ ID NO: 82 - PE9 – <i>Clostridium disporicum</i>	Cluster I
SEQ ID NO: 81 – PE5 – <i>Clostridium bolteae</i>	Cluster XIVa
SEQ ID NO: 80 – VE202-18 – <i>Erysipelatoclostridium ramosum</i>	Cluster XVIII

Figure 27A

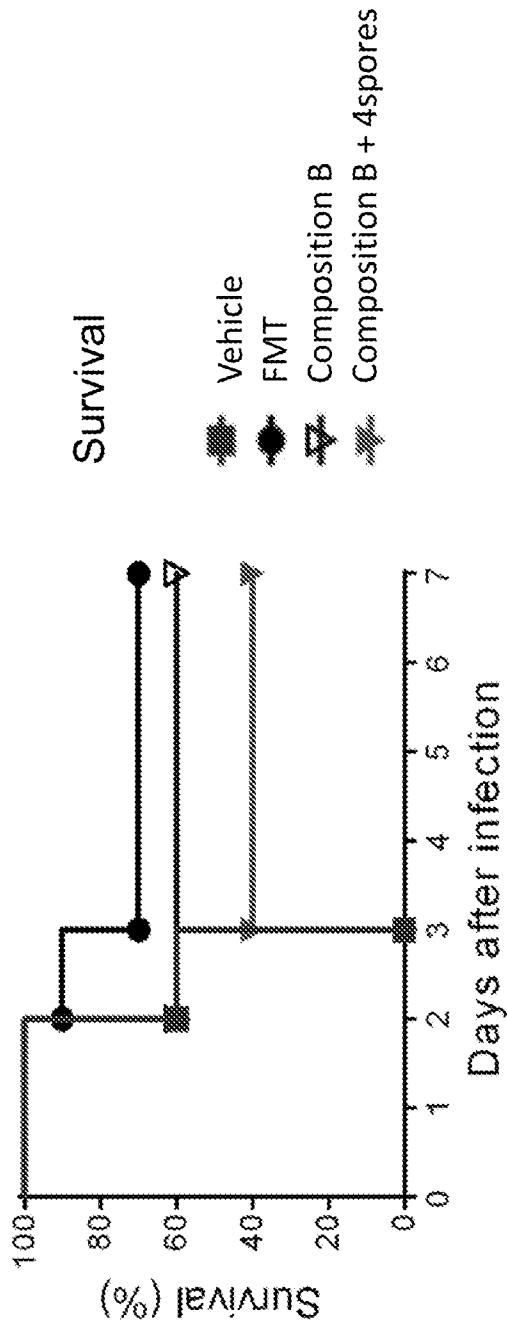
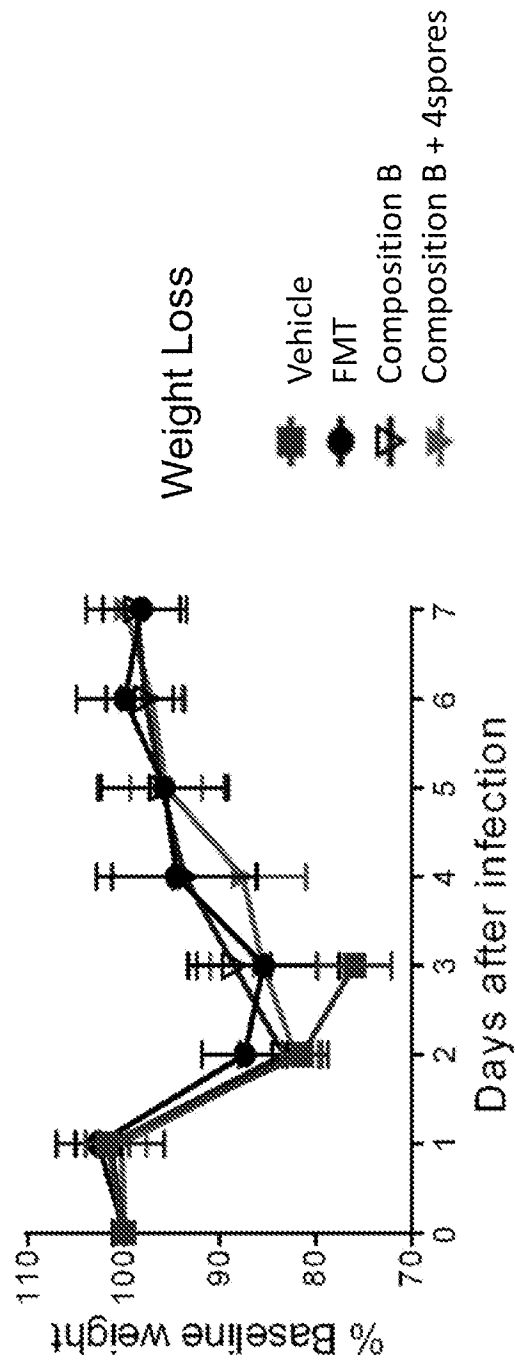
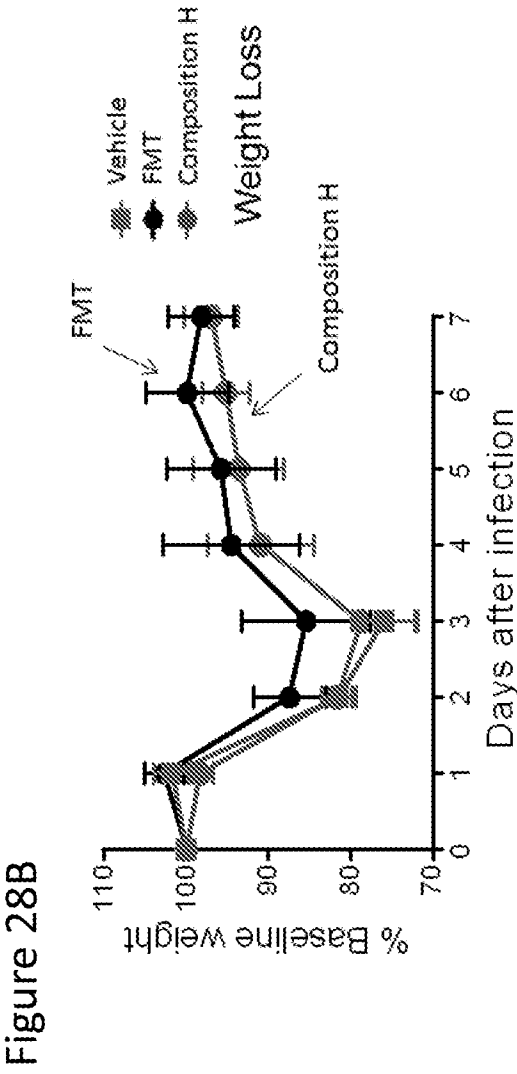
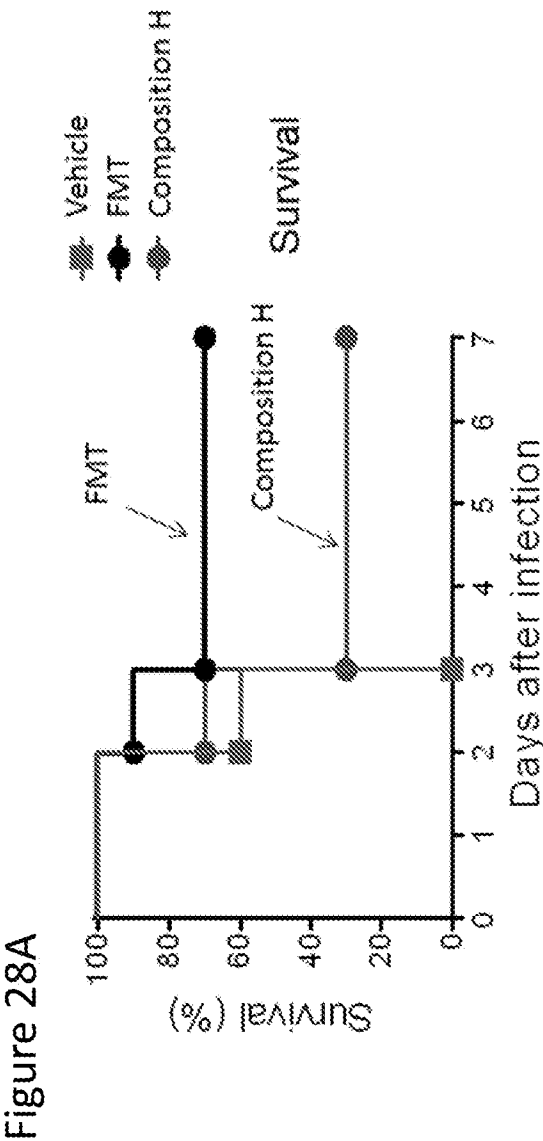
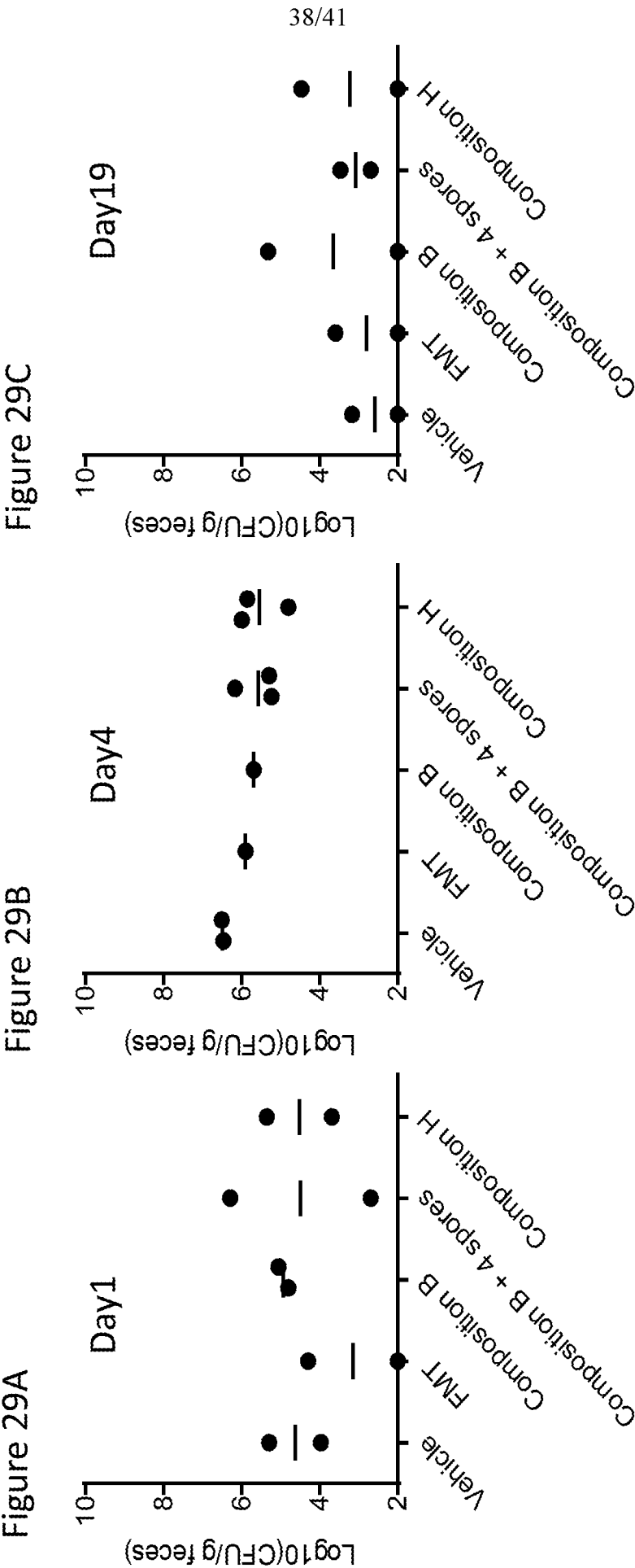


Figure 27B







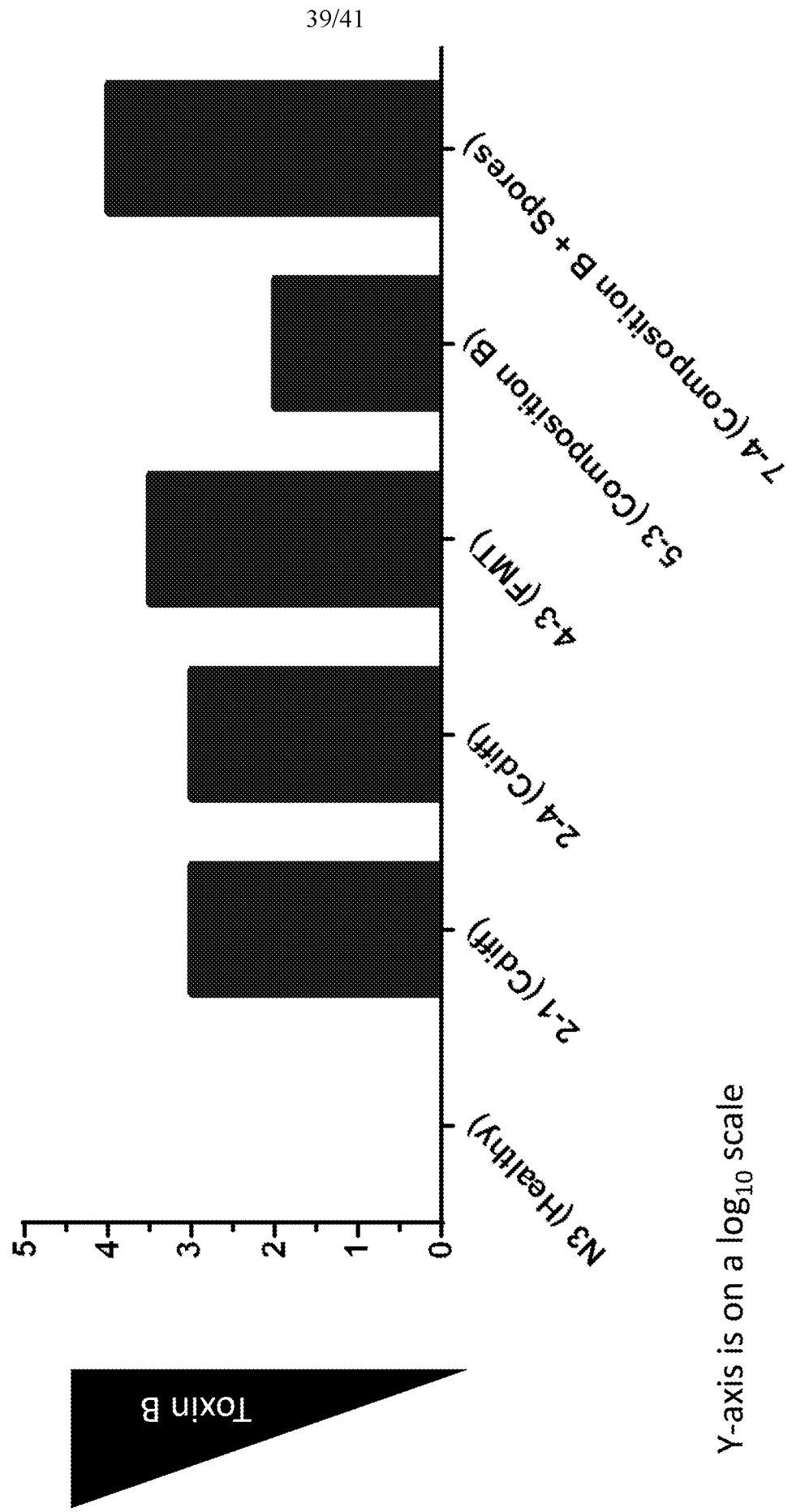


Figure 30

Figure 31

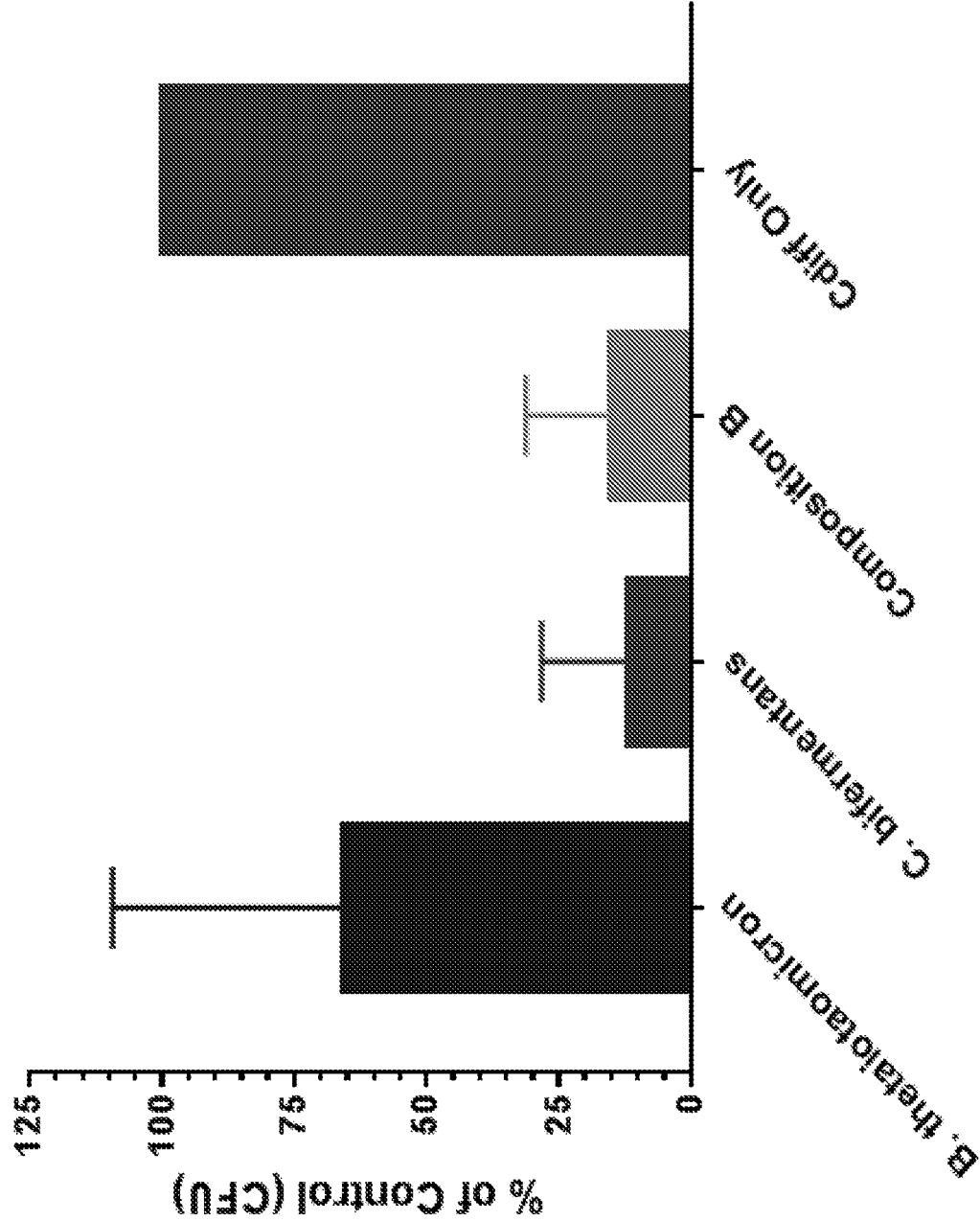
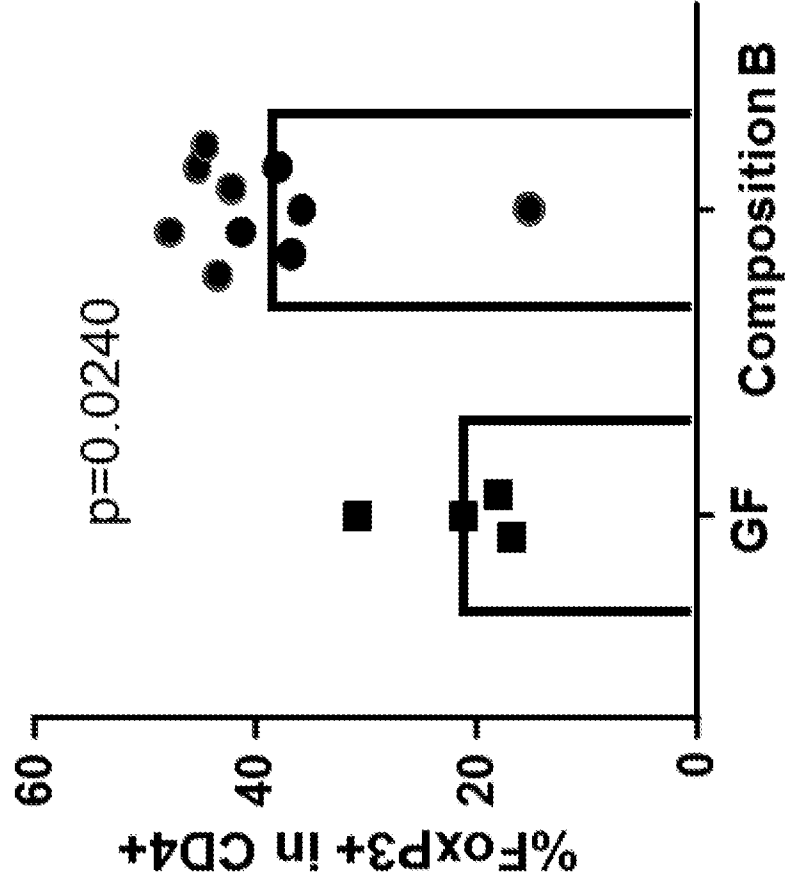


Figure 32



P074570006W000-SEQLISTING-CEW.txt
SEQUENCE LISTING

<110> Vedanta Biosciences, Inc.

<120> TREATMENT OF CLOSTRIDIUM DIFFICILE INFECTION

<130> P0745.70006W000

<140> Not Yet Assigned

<141> 2017-06-14

<150> US 62/349,914

<151> 2016-06-14

<160> 159

<170> PatentIn version 3.5

<210> 1

<211> 210

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 1

gcccggagca gttgatgtga aggatgggtc acctgtggac tgcattggaa ctgtcatact 60

tgagtgccgg agggtaagcg gaattcctag tgtagcgggtg aaatgcgtag atattaggag 120

gaacaccagt ggcgaaggcg gcttactgga cggttaactga cgttgaggct cgaaagcgtg 180

gggagcaaac aggattagat accctggtaa 210

<210> 2

<211> 184

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 2

ctaaccgtgg aggtcattgg aaactgggtca acttgagtgc agaagaggga agtggaattc 60

catgtgtagc ggtgaaatgc gtagagatat ggaggaacac cagtggcgaa ggcggcttcc 120

tggtctgtaa ctgacactga ggcgcgaaag cgtggggggc aaacaggatt agatcccccg 180

gtaa 184

<210> 3

<211> 196

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 3

atgaaagccg gggctcaacc ccggtactgc tttggaaact gtttgacttg agtgcttgag 60

aggtaagtgg aattcctagt gtagcgggaa atgttttagat attaggagga caccagtggc 120

gaaggcggct tactggactg taactgacgt tgtggctcga tttgtgggga gcaaacagga 180

ttatatcccc tggtaa 196

<210> 4
 <211> 211
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 4
 cggaaggctc gatgtgaagg ttggggctta ccccgactg cattggaaac tgtttttcta 60
 gagtgcccga gaggtgaagcg gaattcctag tgtagcgggtg aaatgcttta gatattagga 120
 ggaacaccag tggcgaaggc ggcttactgg acggtaactg acgttgaggc tcgaaagcgt 180
 ggggagcaaa caggattaga taccctggta a 211

<210> 5
 <211> 207
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 5
 cgatgtctga gtgaaggctg gggcttacc caggactgca ttggaactgt ttttctagag 60
 tgccggagag gtaagcggaa ttcctagtgt agcggtgaaa tgcgtagata ttaggaggaa 120
 caccagtggc gaaggcggct tactggacgg taactgacgt tgaggctcga aagcgtgggg 180
 agcaaacagg attagatacc ctggtaa 207

<210> 6
 <211> 179
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 6
 ttaaccaaga agtgcattgg aactgtcaga cttgggggaa aaaaagacag tgcaactcca 60
 tgtgtagcgg tggaatgctc catatatatg gaagaacacc agtggcgaag gcggctgtct 120
 ggtctgcaac tgacgctgag gctcgaattc atgggtaaga aagtattagt cccttgtaa 179

<210> 7
 <211> 214
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 7
 acccgcttgg tctgaggtga ggctggggct taaccccagg actgcattgg aaactgttgt 60
 tctagagtgc cggagaggta agcgggaattc ctagtgtagc ggtgaaatgc gtagatatta 120

P074570006W000-SEQLISTING-CEW.txt

ggaggaacac cagtggcgaa ggcggcttac tggacggtaa ctgacgttga ggctcgaaag 180
cgtgggggagc aaacaggatt agataccctg gtaa 214

<210> 8
<211> 196
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 8
taggctgggg cttaacccca ggactgcatt ggaaactgtt tttctagagt gccggagagg 60
taagcgggaat tcctagtgtg gcggtgaaat gcgtagatat taggaggaac accagtggcg 120
aaggcggctt actggacggt aactgacgtt gaggctcgaa agcgtgggga gcaaacagga 180
ttagataccc tggtaa 196

<210> 9
<211> 297
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 9
ttgcattgga cactatgtca gctgagtgtc ggagaggtaa gtggaattcc tagtgtagcg 60
gtgaaatgcg tagatattag gaggaacacc agtggcggaag gcggttact gcacgttttc 120
tgacgttgag gctcgaaatc gtggggagca aacaaaaata gataccctgg tagtccacgc 180
cgtaaacgat gcatactagg tgtcgggtgg caaagccatt cggtgccgca gcaaacgcaa 240
taagtatgcc acctggggag tacgttcgca agaataaac tcaaataaat tgacgga 297

<210> 10
<211> 209
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 10
cccgtcgtag atgtgaactg ggggctcacc tccagcctgc atttgaaact gtagttcttg 60
agtgctggag aggcaatcgg aattccgtgt gtagcggtag aatgcgtaga tatacggagg 120
aacaccagtg gcgaaggcgg attgctggac agtaactgac gctgaggcgc gaaagcgtgg 180
ggagcaaaca ggattagata ccctcataa 209

<210> 11
<211> 401
<212> DNA
<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 11
 acctgatgca gcgacgccgc gtgagtgaag aagtatttcg gtatgtaaag ctctatcagc 60
 agggaagaaa aaagacggta cctgactaag aagccccggc taactacgtg ccagcagccg 120
 cggtaatacg tagggggcaa gcgttatccg gaattactgg gtgtaaaggg tgcgtaggtg 180
 gcatggtaag tcagaagtga aagcccgggg cttaaccccc ggactgcttt tgaaactgtc 240
 atgctggagt gcaggagagg taagcggaat tcctagtgtg gcggtgaaat gcgtagatat 300
 taggaggaac accagtggcg aaggcggctt actggactgt cactgacact gatgcacgaa 360
 agcgtgggga gcaaacagga ttagataccc tggaagtcca t 401

<210> 12
 <211> 355
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 12
 atgggagcgt agatggcgac tgggccatat gtgacagccc tgggtctcaac cccttaactg 60
 catttggaac tgagtggctg gagtgtcgga gaggcaggcg gaattcctag tgtagcgggtg 120
 aaatgcgtag atattaggag gaacaccagt ggcgaaggcg gcctgctgga cgatgactga 180
 cgttgaggct cgaaagcgtg gggagcaaac aggattagat accctggtag tccacgccgt 240
 aaacgatgac tactaggtgt cgggtggcaa ggacattcgg tgccgcagca aacgcaataa 300
 gtagtccacc tggggagtac gttcgcaaga atgaaactca aaggaaattg acgga 355

<210> 13
 <211> 197
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 13
 cgcagcggag tgtatcctag gctcacctgg ctgctttcga actggttttc tagatcgtgt 60
 agagggggag attcctgggtg tagcgtgaaa tgcgtagata tctggaggaa caccagtggc 120
 gaaggcggcc tcctggacgg caactgacgt tgaggctcga aagtgtgggg agcaaacagg 180
 attagatacc ctggtaa 197

<210> 14
 <211> 1522
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 14
 tcaaagagtt tgatcctggc tcaggacgaa cgctggcggc ggcctaaca catgcaagtc 60

P074570006W000-SEQLISTING-CEW.txt

gaacggagct	tacgttttga	agtttttcgga	tggacgaatg	taagcttagt	ggcggacggg	120
tgagtaacac	gtgagcaacc	tgccttttcag	aggggggataa	cagccggaaa	cggctgctaa	180
taccgcatga	tgttgcgggg	gcacatgccc	ctgcaaccaa	aggagcaatc	cgctgaaaga	240
tgggctcgcg	tccgattagc	cagttggcgg	ggtaacggcc	caccaaagcg	acgatcggta	300
gccggactga	gaggttgaac	ggccacattg	ggactgagac	acggcccaga	ctcctacggg	360
aggcagcagt	gggggatatt	gcacaatggg	cgaagcctg	atgcagcgac	gccgcgtgag	420
ggaagacggg	cttcggattg	taaacctctg	tctttgggga	agaaaatgac	ggtacccaaa	480
gaggaagctc	cggctaacta	cgtgccagca	gccgcggtaa	tacgtaggga	gcaagcgttg	540
tccggaatta	ctgggtgtaa	agggagcgta	ggcgggatgg	caagtagaat	gttaaattcca	600
tcggctcaac	cgggtggctgc	gttctaaact	gccgttcttg	agtgaagtag	aggcaggcgg	660
aattcctagt	gtagcggtag	aatgcgtaga	tattaggagg	aacaccagtg	gcgaaggcgg	720
cctgctgggc	tttaactgac	gctgaggctc	gaaagcgtgg	ggagcaaaca	ggattagata	780
ccctggtagt	ccacgccgta	aacgatgatt	actaggtgtg	gggggactga	ccccttcctg	840
gccgcagtta	acacaataag	taatccacct	ggggagtagc	gccgcaaggt	tgaaactcaa	900
aggaattgac	gggggcccgc	acaagcagtg	gagtatgtgg	tttaattcga	agcaacgcga	960
agaaccttac	caggtcttga	catcggatgc	atagcctaga	gataggtgaa	gcccttcggg	1020
gcatccagac	aggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcaaccct	tattattagt	tgctacgcaa	gagcactcta	atgagactgc	1140
cggtgacaaa	acggaggaag	gtggggatga	cgtcaaatca	tcatgccctt	tatgacctgg	1200
gctacacacg	tactacaatg	gcactaaaac	agagggcggc	gacaccgcga	ggtgaagcga	1260
atccccgaaa	agtgtctcag	ttcagattgc	aggctgcaac	ccgcctgcat	gaagtcggaa	1320
ttgctagtaa	tcgcggatca	gcatgccgcg	gtgaatacgt	tcccgggcct	tgtacacacc	1380
gcccgtcaca	ccatgggagt	cggtaacacc	cgaagccagt	agcctaaccg	caaggggggc	1440
gctgtcgaag	gtgggattga	tgactggggg	gaagtcgtaa	caaggtagcc	gtatcgggaag	1500
gtgcggctgg	atcacctcct	tt				1522

<210> 15
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 15						
tacgagagtt	tgatcctggc	tcaggatgaa	cgctggcggc	gtgcctaaca	catgcaagtc	60
gagcgaagcg	ctgttttcag	aatcttcgga	ggaagaggac	agtgactgag	cggcggacgg	120
gtgagtaacg	cgtgggcaac	ctgcctcata	cagggggata	acagttagaa	atgactgcta	180
ataccgcata	agcgcacagg	accgcatggg	gtagtgtgaa	aaactccggg	ggtatgagat	240

P074570006W000-SEQLISTING-CEW.txt

ggacccgcgt	ctgattaggt	agttggtggg	gtaaaggcct	accaagccga	cgatcagtag	300
ccgacctgag	aggggtgaccg	gccacattgg	gactgagaca	cggcccaaac	tcctacggga	360
ggcagcagtg	gggaatattg	cacaatgggg	gaaaccctga	tgcagcgacg	ccgcgtgaag	420
gaagaagtat	ttcggtatgt	aaacttctat	cagcagggaa	gaagatgacg	gtacctgagt	480
aagaagcacc	ggctaaatac	gtgccagcag	ccgcggtaat	acgtatggtg	caagcgttat	540
ccggatttac	tgggtgtaaa	gggagcgtag	acggataggc	aagtctggag	tgaaaaccca	600
gggctcaacc	ctgggactgc	tttgaaact	gcagatctgg	agtgccggag	aggtaagcgg	660
aattcctagt	gtagcggtag	aatgcgtaga	tattaggagg	aacaccagtg	gcgaaggcgg	720
cttactggac	ggtgactgac	gttgaggctc	gaaagcgtgg	ggagcaaaca	ggattagata	780
ccctggtagt	ccacgccgta	aacgatgact	actaggtgtc	ggtgtgcaaa	gcacatcggt	840
gccgcagcaa	acgcaataag	tagtccacct	ggggagtacg	ttcgcaagaa	tgaaactcaa	900
aggaattgac	ggggacccgc	acaagcggtg	gagcatgtgg	tttaattcga	agcaacgcga	960
agaaccttac	ctggctctga	catccggatg	acgggcgagt	aatgtcgccg	tccttcggg	1020
gcatccgaga	caggtggtgc	atggttgctg	tcagctcgtg	tcgtgagatg	ttgggttaag	1080
tcccgcacg	agcgaaccc	ttatcttcag	tagccagcat	ataagggtgg	cactctggag	1140
agactgccag	ggagaacctg	gaggaagggtg	gggatgacgt	caaatcatca	tgccccttat	1200
ggccagggct	acacacgtgc	tacaatggcg	taaacaaagg	gaagcgagag	ggtgacctgg	1260
agcgaatccc	aaaaataacg	tctcagttcg	gattgtagtc	tgcaactcga	ctacatgaag	1320
ctggaatcgc	tagtaatcgc	ggatcagcat	gccgcggtag	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagt	aacgcccga	gccagtgacc	caaccttaga	1440
ggagggagct	gtcgaaggcg	ggacggataa	ctgggggtgaa	gtcgtaacaa	ggtagccgta	1500
tcggaagggtg	cggctggatc	acctccttt				1529

<210> 16
 <211> 1527
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 16						
atgagagttt	gatcctggct	caggatgaac	gctggcggcg	tgccaaacac	atgcaagtcg	60
aacgaagcga	tttaacggaa	gttttcggat	ggaagttgaa	ttgactgagt	ggcggacggg	120
tgagtaacgc	gtgggtaacc	tgccctgtac	tgggggacaa	cagttagaaa	tgactgctaa	180
taccgcataa	gcgcacagta	tcgcatgata	cagtgtgaaa	aactccggtg	gtacaagatg	240
gacccgcgtc	tgattagcta	gttggttaag	taacggctta	ccaaggcgac	gatcagtagc	300
cgacctgaga	gggtgaccgg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360
gcagcagtg	ggaatattgc	acaatggg	aaagcctgat	gcagcgacgc	cgcgtagagt	420

P074570006W000-SEQLISTING-CEW.txt

aagaagtatt	tcggtatgta	aagctctatc	agcaggggaag	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc	540
cggatttact	gggtgtaaag	ggagcgtaga	cggtaaagca	agtctgaagt	gaaagcccgc	600
ggctcaactg	cgggactgct	ttggaaactg	tttaactgga	gtgtcggaga	ggtaagtgga	660
attcctagtg	tagcgggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcgac	720
ttactggacg	ataactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780
cctggtagtc	cacgccgtaa	acgatgaata	ctaggtgttg	gggagcaaag	ctcttcgggtg	840
ccgtcgcaaa	cgcagtaagt	attccacctg	gggagtacgt	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgggtg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	aggtcttgac	atcgatccga	cgggggagta	acgtcccctt	cccttcgggg	1020
cggagaagac	aggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcaaccct	tattctaagt	agccagcggg	tcggccggga	actcttggga	1140
gactgccagg	gataacctgg	aggaagggtg	ggatgacgtc	aaatcatcat	gccccttatg	1200
atctgggcta	cacacgtgct	acaatggcgt	aaacaaagag	aagcaagacc	gcgaggtgga	1260
gcaaattctca	aaaataacgt	ctcagttcgg	actgcaggct	gcaactcgcc	tgcacgaagc	1320
tggaatcgct	agtaatcgcg	aatcagaatg	tcgcggtgaa	tacgttcccg	ggtcttgtac	1380
acaccgcccg	tcacaccatg	ggagtcagta	acgcccgaa	tcagtgaccc	aaccgcaagg	1440
agggagctgc	cgaaggcggg	accgataact	gggggtgaagt	cgtaacaagg	tagccgtatc	1500
ggaagggtgcg	gctggatcac	ctccttt				1527

<210> 17
 <211> 1530
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 17						
atgagagttt	gatcctggct	caggatgaac	gctggcggcg	tgcctaacac	atgcaagtcg	60
aacgaagcaa	ttaaaatgaa	gttttcggat	ggattttttga	ttgactgagt	ggcggacggg	120
tgagtaacgc	gtggataacc	tgcctcacac	tgggggataa	cagttagaaa	tgactgctaa	180
taccgcataa	gcgcacagta	ccgcatggta	cgggtgtgaaa	aactccgggtg	gtgtgagatg	240
gatccgcgtc	tgattagcca	gttggcgggg	taacggccca	ccaaagcgac	gatcagtagc	300
cgacctgaga	gggtgaccgg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360
gcagcagtgg	ggaatattgc	acaatgggcg	aaagcctgat	gcagcgacgc	cgcgtagtg	420
aagaagtatt	tcggtatgta	aagctctatc	agcaggggaag	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc	540
cggatttact	gggtgtaaag	ggagcgtaga	cggcgaagca	agtctgaagt	gaaaaccag	600

P074570006W000-SEQLISTING-CEW.txt

ggctcaaccc	tgggactgct	ttggaaactg	ttttgctaga	gtgtcggaga	ggtaagtgga	660
attcctagt	tagcgggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc	720
ttactggacg	ataactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780
cctggtagtc	cacgccgtaa	acgatgaatg	ctaggtgttg	gggggcaaag	cccttcggtg	840
ccgtcgcaaa	cgcagtaagc	attccacctg	gggagtacgt	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgggtg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	aagtcttgac	atcctcttga	ccggcgtgta	acggcgcctt	cccttcgggg	1020
caagagagac	aggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcaaccct	tatccttagt	agccagcagg	taaagctggg	cactctaggg	1140
agactgccag	ggataacctg	gaggaagggt	gggatgacgt	caaatcatca	tgccccttat	1200
gatttgggct	acacacgtgc	tacaatggcg	taaacaaagg	gaagcaagac	agtgatgtgg	1260
agcaaatccc	aaaaataacg	tcccagttcg	gactgtagtc	tgcaaccgga	ctacacgaag	1320
ctggaatcgc	tagtaatcgc	gaatcagaat	gtcgcggtga	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagc	aacgcccga	gtcagtgacc	caactcgcaa	1440
gagagggagc	tgccgaaggc	ggggcaggta	actgggggtga	agtcgtaaca	aggtagccgt	1500
atcggaaggt	gcggctggat	cacctccttt				1530

<210> 18
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 18						
aacgagagtt	tgatcctggc	tcaggatgaa	cgctggcggc	gtgcttaaca	catgcaagtc	60
gagcgaagca	ctttggaaga	ttcttcggat	gaagactttt	gtgactgagc	ggcggacggg	120
tgagtaacgc	gtgggtaacc	tgccatcatc	agggggataa	cagttagaaa	tgactgctaa	180
taccgcataa	gaccacggta	ccgcatggta	cagtggtaaa	aactccggtg	gtatgagatg	240
gacccgcgtc	tgattaggtg	gttggtgggg	taacggccta	ccaagccgac	gatcagtagc	300
cgacctgaga	gggtgaccgg	ccacattggg	actgagacac	ggcccagact	cctacgggag	360
gcagcagtg	ggaatattgc	acaatggagg	aaactctgat	gcagcgacgc	cgcgtaagg	420
atgaagtatt	tcggtatgta	aacttctatc	agcagggaa	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc	540
cggatttact	gggtgtaaag	ggagcgtaga	cgccacggca	agccagatgt	gaaagccccg	600
ggctcaaccc	cgggactgca	tttggaactg	ctgagctaga	gtgtcggaga	ggcaagtgga	660
attcctagt	tagcgggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc	720
ttgctggacg	atgactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780

P074570006W000-SEQLISTING-CEW.txt

cctggtagtc	cacgccgtaa	acgatgactg	ctaggtgtcg	ggtggcaaag	ccattcgggtg	840
ccgcagctaa	cgcaataagc	agtccacctg	gggagtagct	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgggtg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	tgatcttgac	atccccgatga	ccgcttcgta	atggaagctt	ttcttcggaa	1020
catcggtgac	aggtgggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcaacccc	tatcttcagt	agccagcagg	ttaagctggg	cactctggag	1140
agactgccag	ggataacctg	gaggaagggtg	gggatgacgt	caaatcatca	tgccccttat	1200
gaccagggct	acacacgtgc	tacaatggcg	taaacaaaga	gaagcgaact	cgcgagggtta	1260
agcaaatctc	aaaaataacg	tctcagttcg	gattgtagtc	tgcaactcga	ctacatgaag	1320
ctggaatcgc	tagtaatcgc	agatcagaat	gctgcgggtga	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagt	aacgcccga	gtcagtgacc	caaccgtaag	1440
gagggagctg	ccgaagggtg	gaccgataac	tgggggtgaag	tcgtaacaag	gtagccgtat	1500
cggaagggtgc	ggctggatca	cctccttt				1528

<210> 19
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 19						
atcagagagt	ttgatcctgg	ctcaggatga	acgctggcgg	cgtgcttaac	acatgcaagt	60
cgagcgaagc	acttaagtgg	atctcttcgg	attgaaactt	atttgactga	gcggcggacg	120
ggtgagtaac	gcgtgggtaa	cctgcctcat	acagggggat	aacagttaga	aatggctgct	180
aataccgcat	aagcgcacag	gaccgcatgg	tctggtgtga	aaaactccgg	tggtatgaga	240
tggacccgcg	tctgattagc	tagttggagg	ggtaacggcc	caccaaggcg	acgatcagta	300
gccggcctga	gaggggtgaac	ggccacattg	ggactgagac	acggcccaga	ctcctacggg	360
aggcagcagt	ggggaatatt	gcacaatggg	ggaaaccctg	atgcagcgac	gccgcgtgaa	420
ggaagaagta	tctcggtatg	taaacttcta	tcagcagga	agaaaatgac	ggtacctgac	480
taagaagccc	cggctaacta	cgtgccagca	gccgcggtaa	tacgtagggg	gcaagcgtta	540
tccggattta	ctgggtgtaa	agggagcgta	gacggaagag	caagtctgat	gtgaaaggct	600
ggggcttaac	cccaggactg	cattggaaac	tgtttttcta	gagtgccgga	gaggtaagcg	660
gaattcctag	tgtagcgggtg	aatgcgtag	atattaggag	gaacaccagt	ggcgaaggcg	720
gcttactgga	cggttaactga	cgttgaggct	cgaaagcgtg	gggagcaaac	aggattagat	780
accctggtag	tccacgccgt	aaacgatgaa	tactaggtgt	cgggtggcaa	agccattcgg	840
tgccgcagca	aacgcaataa	gtattccacc	tggggagtag	gttcgcaaga	atgaaactca	900
aaggaattga	cggggacccg	cacaagcggg	ggagcatgtg	gtttaattcg	aagcaacgcg	960

P074570006W000-SEQLISTING-CEW.txt

aagaacctta ccaagtcttg acatccctct gaccggcccg taacggggcc ttcccttcgg	1020
ggcagaggag acaggtggtg catggttgct gtcagctcgt gtcgtgagat gttgggttaa	1080
gtcccgcaac gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgac gtcaaatacat catgcccctt	1200
atgatttggg ctacacacgt gctacaatgg cgtaaataaa gggaagcgag acagcgatgt	1260
tgagcaaatc ccaaaaataa cgtcccagtt cggactgcag tctgcaactc gactgcacga	1320
agctggaatc gctagtaatc gcgaatcaga atgtcgcggt gaatacgttc ccgggtcttg	1380
tacacaccgc ccgtcacacc atgggagtca gtaacgcccg aagtcagtga cccaacctta	1440
caggagggag ctgccgaagg cgggaccgat aactggggtg aagtcgtaac aaggtagccg	1500
tatcggaagg tgcggctgga tcacctcctt t	1531

<210> 20
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 20	
aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc	60
gagcgaagca ctttgaaga ttcttcggat gatttccttt gtgactgagc ggcggacggg	120
tgagtaacgc gtgggtaacc tgcctcatac agggggataa cagttagaaa tgactgctaa	180
taccgcataa gaccacggta ccgcatggta cagtggtaaa aactccggtg gtatgagatg	240
gaccgcgctc tgattaggtg gttggtgggg taacggccta ccaagccgac gatcagtagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccagact cctacgggag	360
gcagcagtg ggaatattgc acaatggagg aaactctgat gcagcgacgc cgcgtgaagg	420
atgaagtatt tcggtatgta aacttctatc agcagggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcagc cgcggttaata cgtagggggc aagcgttatc	540
cggatttact ggggtgtaaag ggagcgtaga cggcacggca agccagatgt gaaagccccg	600
ggctcaacc cgggactgca tttggaactg ctgagctaga gtgtcggaga ggcaagtgga	660
attcctagt tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcggc	720
ttgctggacg atgactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgccgtaa acgatgactg ctaggtgtcg ggtggcaaag ccattcggtg	840
ccgcagctaa cgcaataagc agtccacctg gggagtacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc tgatcttgac atccccgatga ccgcttcgta atggaagctt ttcttcggaa	1020
catcggtgac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaacccc tatcttcagt agccagcagg ttaagctggg cactctggag	1140

P074570006W000-SEQLISTING-CEW.txt

agactgccag	ggataacctg	gaggaaggtg	gggatgacgt	caaatcatca	tgccccttat	1200
gaccagggct	acacacgtgc	tacaatggcg	taaacaaaga	gaagcgaact	cgcgagggta	1260
agcaaatctc	aaaaataacg	tctcagttcg	gattgtagtc	tgcaactcga	ctacatgaag	1320
ctggaatcgc	tagtaatcgc	agatcagaat	gctgcggtga	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagt	aacgcccga	gtcagtgacc	caaccgtaag	1440
gagggagctg	ccgaaggtgg	gaccgataac	tgggggtgaag	tcgtaacaag	gtagccgtat	1500
cggaaggtgc	ggctggatca	cctccttt				1528

<210> 21
 <211> 1537
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 21	
atggagagtt	tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc 60
gaacgaagtt	tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca 120
cgtaggtaac	ctgcccattgt gtccgggata actgctggaa acggtagcta aaaccggata 180
ggtatacaga	gcgcatgctc agtatattaa agcgcccatc aaggcgtgaa catggatgga 240
cctgcggcgc	attagctagt tggtaggta acggcccacc aaggcgatga tgcgtagccg 300
gcctgagagg	gtaaacggcc acattgggac tgagacacgg cccaaactcc tacgggagggc 360
agcagtaggg	aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagtgaa 420
gaaggtcttc	ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct 480
atgggagtga	cggtagctta ccagaaagcc acggctaact acgtgccagc agccgcggta 540
atacgtaggt	ggcaagcgtt atccggaatc attgggcgta aagggtgcgt aggtggcgta 600
ctaagtctgt	agtaaaaggc aatggctcaa ccattgtaag ctatggaaac tggtagctg 660
gagtgcagaa	gagggcgatg gaattccatg tgtagcggta aaatgcgtag atatatggag 720
gaacaccagt	ggcgaaggcg gtcgcctggc ctgtaactga cactgaggca cgaaagcgtg 780
gggagcaaat	aggattagat accctagtag tccacgccgt aaacgatgag aactaagtgt 840
tggaggaatt	cagtgtgca gttaacgcaa taagttctcc gcctggggag tatgcacgca 900
agtgtgaaac	tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat 960
tcgaagcaac	gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag 1020
ggggataatt	atggatcaca caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg 1080
ttgggttaag	tcccgcacg agcgcaaccc ttgtcgcatt ttaccagcat caagttgggg 1140
actcatgcca	gactgccggt gacaaaccgg aggaaggtgg ggatgacgtc aaatcatcat 1200
gccccttatg	gcctgggcta cacacgtact acaatggcgg ccacaaagag cagcgacaca 1260
gtgatgtgaa	gcgaatctca taaaggtcgt ctcagttcgg attgaagtct gcaactcgac 1320

P074570006W000-SEQLISTING-CEW.txt

ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcggtgaa tacgttctcg	1380
ggccttgtac acaccgcccc tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440
aaccgtaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccttt	1537

<210> 22
 <211> 476
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 22 agtaacgcgt gggtaacctg cctcatacag ggggataaca gttagaaatg actgctaata	60
ccgcataaga ccacggtacc gcatggtaca gtggtaaaaa ctccggtggt atgagatgga	120
cccgcgtctg attaggtagt tgggtgggta acggcctacc aagccgacga tcagtagccg	180
acctgagagg gtgaccggcc acattgggac tgagacacgg ccagactcc tacgggaggc	240
agcagtgggg aatattgcac aatggaggaa actctgatgc agcgacgccg cgtgaaggat	300
gaagtatttc ggtatgtaaa cttctatcag cagggaagaa aatgacggta cctgactaag	360
aagccccggc taactacgtg ccagcagccg cggtaatacg tagggggcaa gcgttatccg	420
gatttactgg gtgtaaaggg agcgtagacg gcacggcaag ccagatgtga aagccc	476

<210> 23
 <211> 354
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 23 caggctggag tgcaggagag gtaagcggaa ttcctagtgt agcggtgaaa tgcgtagata	60
ttaggaggaa caccagtggc gaaggcggct tactggactg taactgacgt tgaggctcga	120
aagcgtgggg agcaaacagg attagatacc ctggtagtcc acgcggtaaa cgatgattgc	180
taggtgtagg tgggtatgga cccatcgggt cgcagctaa cgcaataagc aatccacctg	240
gggagtacgt tcgcaagaat gaaactcaaa ggaattgacg gggacccgca caagcgggtg	300
agcatgtggt ttaattcgaa gcaacgcgaa gaaccttacc aagtcttgac atcc	354

<210> 24
 <211> 190
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 24 ccggggctca ccccgggact gcatttggaa ctgctgagct agagtgtcgg agaggcaagt	60

ggaattccta gtgtagcggg gaaatgcgta gatattagga ggaacaccag tggcgaaggc 120
 ggcttgctgg acgatgactg acgttgaggc tcgaaagcgt ggggagcaaa caggattaga 180
 taccctggta 190

<210> 25
 <211> 190
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 25
 aggggtcaacc cctggactgc attggaaact gtcaggctgg agtgccggag aggtaagcgg 60
 aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg 120
 cttactggac ggtaactgac gttgatgctc gaaagcgtgg ggagcaaaca ggattagata 180
 acctggtaaa 190

<210> 26
 <211> 209
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 26
 gggaagtctgg tcttaagtgc ggggcttaac cccgtgaggg gaccgaaact gtgaagctcg 60
 agtgtcggag aggaaagcgg aattcctagt gtagcggtag aatgcgtaga tattaggagg 120
 aacaccagtg gcgaaagcgg ctttctggac gacaactgac gctgaggcgc gaaagccagg 180
 ggagcaaacg ggattagata cccagtaa 209

<210> 27
 <211> 203
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 27
 tagtctgagt gatgcggggc ttaaccccgt atggcggttg atactggaag tcttgagtgc 60
 aggagaggaa aggggaattc ccagtgtagc ggtgaaatgc gtagatattg ggaggaacac 120
 cagtggcgaa ggcgcccttc tggactgtgt ctgacgctga gatgcgaaag ccagggtagc 180
 aaacgggatt agataccacg gta 203

<210> 28
 <211> 207
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 28
 gatagtcggt ctttaagtgcg gggcttacct cgtgagggga ccgaaactgt gaagctcgag 60
 tgtcggagag gaaagcggaa ttcctagtgt agcggtgaaa tgcgtagata ttaggaggaa 120
 caccagtggc gaaagcggct ttctggacga caactgacgc tgaggcgcgga aagccagggg 180
 agcaaacggg attagatacc acggtaa 207

<210> 29
 <211> 424
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 29
 cgtttgctcc acgctttcga gcctcacgtc agttaccgtc cagtaagccg ccttcgccac 60
 tgggtgttcct cctaataatct acgcatttca ccgctacact aggaattccg cttacctctc 120
 cggctactcta gattgacagt ttccaatgca gtcccggggg tgagccccgg gttttcacat 180
 cagacttgcc actccgtcta cgctcccttt acaccagta aatccggata acgcttgac 240
 catacgtatt accgcggctg ctggcacgta tttagccggg gcttcttagt cagggtaccgt 300
 catTTtcttc cctgctgata gagctttaca taccgaaata cttcatcgct cacgcggcgt 360
 cgctgcatca gggtttcccc cattgtgcaa tattccccac tgctgcctcc cgtaggagtt 420
 tgga 424

<210> 30
 <211> 441
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 30
 tgtcacactt tcgagcatca gcgtcagtta cagtccagta agctgccttc gcaatcggag 60
 ttcttcgtga tatctaagca tttcaccgct acaccacgaa ttccgcctac ctctactgca 120
 ctcaagacga ccagtatcaa ctgcaatttt acggttgagc cgcaaacttt cacagctgac 180
 ttaatagtcc gcctacgctc cctttaaacc caataaatcc ggataacgct tggatcctcc 240
 gtattaccgc ggctgctggc acggagttag ccgaccta ttcgtatggg acatacaaaa 300
 agccacacgt ggctcacttt attcccatat aaaagaagtt tacaacccat agggcagtc 360
 tccttcacgc tacttggtg gttcagactc tcgtccattg accaatattc ctactgctg 420
 cctcccgtag gtagtttgga a 441

<210> 31
 <211> 422
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 31

```

ccgttgtcac gctttcgtgc tcagtgtcag tttcagtcca gtaagccgcc ttcgccactg      60
atgttcctcc taatatctac gcatttcacc gctacactag gaattccgct tacctctcct      120
gcactccagt ctgacagttt caaaagcagt cccagagtta agccctgggt tttcactttct      180
gacttgccat accacctacg caccctttac acccagtaat tccggataac gcttgccccc      240
tacgtattac cgcggtctgt ggcacgtagt tagccggggc ttcttagtca ggtaccgtca      300
ttttcttccc tgctgataga gctttacata ccgaaatact tcttcactca cgcgggcgtcg      360
ctgcatcagg gttcccccca ttgtgcaata ttccccactg ctgcctcccg tggaagtttg      420
ga                                                                           422

```

<210> 32

<211> 424

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 32

```

gcgaatgtca cgcattcgag cctcacgtca gttaccgtcc agtaagccgc cttcgccact      60
gggtgttcctc ctaatatcta cgcatttcac cgctacacta ggaattccgc ttacctctcc      120
ggcactcaag actaacagtt tccaatgcag tccagggggt gagccccgc ctttcacatc      180
agacttgcca gtccgtctac gctcccttta caccagtaa atccggataa cgcttgcccc      240
ctacgtatta ccgcggtctg tggcacgtag ttagccgggg cttcttagtc aggtaccgtc      300
actatcttcc ctgctgatag aagtttacat accgagatac ttcttccttc acgcggcgtc      360
gctgcatcag ggtttcccc attgtgcaat attccccact gctgcctccc gtgggagttt      420
ggaa                                                                           424

```

<210> 33

<211> 422

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 33

```

tgctcacgct ttcgcgctca gcgtcagtta ctgtccagca atccgccttc gccactgggtg      60
ttcctccgta tatctacgca tttcaccgct acacacggaa ttccgattgc ctctccagca      120
ctcaagaact acagtttcaa atgcaggctg gaggttgagc cccagtttt cacatctgac      180
ttgcaatccc gcctacacgc cctttacacc cagtaaatac ggataacgct tgccacctac      240
gtattaccgc ggctgctggc acgtagttag ccgtggctta ttcgtcaggt accgtcattt      300
gtttcgtccc tgacaaaaga agtttacaac ccgaaagcct tcttccttca cgcgggcgttg      360

```

ctgggtcagg cttgcgcca ttgccaata ttccccactg ctgcctcccg tggtagtttg 420
ga 422

<210> 34
<211> 419
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 34
tgtccacgct ttcgagctca gcgtcagtta tcgtccagta agccgccttc gccactgggtg 60
ttcctcctaa tatctacgca tttcaccgct aactagga ttccgcttac ccctccgaca 120
ctctagtacg acagtttcca atgcagtacc ggggttgagc cccgggcttt cacatcagac 180
ttgccgcacc gcctgcgctc cctttacacc cagtaaattc ggataacgct tgcaccatac 240
gtattaccgc ggctgctggc acgtatttag ccggtgcttc ttagtcaggt accgtcatta 300
tcttcctgc tgatagagct ttacataccg aaatacttct tcgctcacgc ggcgtcgctg 360
catcaggctt tcgcccattg tgcaatattc cccactgctg actcccgtag gagtttgga 419

<210> 35
<211> 424
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 35
cgtttctcca cgcttcgctc tcagcgtcag ttactgtcca gcaatccgcc ttcgccactg 60
gtgttcctcc gtatatctac gcatttcacc gctacacacg gaattccgat tgcctctcca 120
gcactcaaga actacagttt caaatgcagg ctggagggtg agccccagt tttcacatct 180
gacttgcaat cccgcctaca cgccctttac acccagtaaa tccggataac gcttgccacc 240
tacgtattac cgcggtgctt ggcacgtagt tagccgtggc ttattcgtca ggtaccgtca 300
tttgtttcgt ccccgacaaa agaagtttac aaccggaaag ccttcttcct tcacgcggcg 360
ttgctgggtc aggcttgctc ccattgcca atattcccca ctgctgcctc cctgggaagt 420
ttgg 424

<210> 36
<211> 445
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 36
atgtcctgac ttcgcgcctc agcgtcagtt gtcgtccaga aagccgcttt cgccactggt 60
gttcctccta atatctacgc atttcaccgc tacactagga attccgcttt cctctccgac 120

P074570006W000-SEQLISTING-CEW.txt

actcgagctt cacagtttcg gtcccctcac ggggttaagc cccgcacttt taagaccgac	180
ttgcgatgcc gcctgcgcgc cctttacgcc caataattcc ggacaacgct tgccacctac	240
gtattaccgc ggctgctggc acgtagttag ccgtggcttt ctcttacggg accgtcaggg	300
ataacgggta ttgaccgcta tcctgttcgt cccatataac agaactttac aaccggaagg	360
ccgtcatcgt tcacgcggcg ttgctccgtc agactttcgt ccattgcgga agattcccca	420
ctgctgcctc cctgggaagt ttgga	445

<210> 37
 <211> 421
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 37	
gtttgctcac gctttcgagc tcagcgtcag ttatcgtcca gtaagccgcc ttcgccactg	60
gtgttcctcc taatatctac gcatttcacc gctacactag gaattccgct taccctccg	120
acactctagt acgacagttt ccaatgcagt accgggggtg agccccgggc tttcacatca	180
gacttgccgc accgcctgcg ctccctttac acccagtaaa tccggataac gcttgcacca	240
tacgtattac cgcggtgct ggcacgtatt tagccggtgc ttcttagtca ggtaccgtca	300
ttatcttccc tgctgataga gctttacata ccgaaatact tcttcgctca cgcggcgtcg	360
ctgcatcagg ctttcgcca ttgtgcaata ttccccactg ctgcctcccg taggagtttg	420
g	421

<210> 38
 <211> 424
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 38	
cgttgctcac gcattcgagc ctcagcgtca gttaagccca gtaagccgcc ttcgccactg	60
atgttcctcc taatatctac gcatttcacc gctacactag gaattccgct tacctctact	120
tcactcaaga accacagttt caaatgcagt ttatgggtta agcccatagt tttcacatct	180
gacttgcatg cccgcctacg ctccctttac acccagtaat tccggacaac gctcgctccc	240
tacgtattac cgcggtgct ggcacgtagt tagccggagc ttcctcctca ggtaccgtct	300
tttttcgtcc ctgaagacag aggtttacaa tcctaaaacc ttcttcctc acgcggcatc	360
gctgcatcag agtttcctcc attgtgcaat attccccact gctgcctccc gtaggagttt	420
ggaa	424

<210> 39
 <211> 189

<212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 39
 tgggcttacc cataaactgc atttgaaact gtggttcttg agtgaagtag aggtaagcgg 60
 aattcctagt gtagcggtag aatgcgtaga tattaggagg aacatcagtg gcgaaggcgg 120
 cttactgggc tttaactgac gctgaggctc gaaagcgtgg ggagcaaaca ggattagata 180
 cccaagtaa 189

<210> 40
 <211> 414
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 40
 gtcagcatcg agctcacgtc agttaccgtc cagtaagccg ccttcgccac tgggtgttcct 60
 cctaatatct acgcatttca ccgctacact aggaattccg cttacctctc cggtagtcta 120
 gattgacagt ttccaatgca gtcccggggg tgagccccgg gttttcacat cagacttgcc 180
 actccgtcta cgctcccttt acaccagta aatccggata acgcttgacac catacgtatt 240
 accgcggctg ctggcacgta tttagccggg gcttcttagt caggtaccgt cattttcttc 300
 cctgctgata gagctttaca taccgaaata cttcatcgct cacgcggcgt cgctgcatca 360
 gggtttcccc cattgtgcaa tattccccac tgctgcctcc cgaggaggtt tgga 414

<210> 41
 <211> 191
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 41
 tcatcgctta cgggtgatct gcgccgggta cgggcgggct ggagtgcggt aggggagact 60
 ggaattcccg gtgtaacggt ggaatgtgta gatatcgga agaacaccga tggcgaaggc 120
 aggtctctgg gccgtcactg acgctgagga gcgaaagcgt ggggagcgaa caggattaga 180
 tacaacggta a 191

<210> 42
 <211> 196
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 42
 tgaaccagg gcttaactct gggactgctt ttgaactgtc agactggagt gcaggagagg 60

taagcggaaat tcctagtgtgta gcggtgaaat gcgtagatat taggaggaac atcagtggcg 120
aaggcggcctt actggactga aactgacact gaggcacgaa agcgtgggga gcaaacagga 180
ttagatacca tggtaa 196

<210> 43
<211> 192
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 43
accagggcctt aactctggga ctgcttttga actgtcagac tggagtgcag gagaggtaag 60
cggaattcct agtgtagcgg tgaaatgcgt agatattagg aggaacatca gtggcgaagg 120
cggcttactg gactgaaact gacactgagg cacgaaagcg tggggagcaa acaggattag 180
ataccctggt aa 192

<210> 44
<211> 195
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 44
gaacccaggg cttaactctg ggactgcttt tgaactgtca gactggagtg caggagaggt 60
aagcggaaatt cctagtgtag cggtgaaatg cgtagatatt aggaggaaca tcagtggcga 120
aggcggccta ctggactgaa actgacactg aggcacgaaa gcgtggggag caaacaggat 180
tagatacccc ggtaa 195

<210> 45
<211> 418
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 45
gagtcagctt tcgagctcag cgtcagttat cgtccagtaa gccgccttcg cactggtgt 60
tcctcctaata atctacgcat ttcaccgcta cactaggaat tccgcttacc cctccgacac 120
tctagtacga cagtttccaa tgcagtaccg gggttgagcc ccgggctttc acatcagact 180
tgccgcaccg cctgcgctcc ctttacaccc agtaaattccg gataacgctt gcaccatacg 240
tattaccgcg gctgctggca cgtatttagc cgggtgcttct tagtcaggta ccgtcattat 300
cttccctgct gatagagctt tacataccga aatacttctt cgctcacgcg gcgtcgctgc 360
atcaggcttt cgcccattgt gcaatattcc cactgctgc ctcccaggag agtttggga 418

<210> 46
 <211> 416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 46
 tgtcagcttt cgagctcacg tcagttaccg tccagtaagc cgccttcgcc actggtgttc 60
 ctctaataat ctacgcattt caccgctaca ctaggaattc cgcttacctc tccggtactc 120
 tagattgaca gtttccaatg cagtcccggg gttgagcccc gggttttcac atcagacttg 180
 ccactccgtc tacgctccct ttacaccag taaatccgga taacgcttgc accatacgta 240
 ttaccgcggc tgctggcacg tatntagccg gtgcttctta gtcaggtacc gtcattttct 300
 tccctgctga tagagcttta cataccgaaa tacttcatcg ctcacgcggc gtcgctgcat 360
 cagggtttcc cccattgtgc aatattcccc actgctgcct cccgtaggag tttgga 416

<210> 47
 <211> 400
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 47
 cacgtcagtt accgtccagt aagccgcctt cgccactggt gttcctccta atatctacgc 60
 atttcaccgc tacactagga attccgctta cctctccggc actcaagacg ggcagtttcc 120
 aatgcagtcc cgggggttag cccagcctt tcacatcaga cttgtccatc cgtctacgct 180
 ccctttacac ccagtaaate cggataacgc ttgcccccta cgtattaccg cggctgctgg 240
 cacgtagtta gccggggctt cttagtcagg taccgtcatt ttcttccctg ctgatagaag 300
 ttacataacc gagatacttc ttccttcacg cggcgctcgt gcatcagggt ttccccatt 360
 gtgcaatatt cccactgct gcctcccgta ggagtttggg 400

<210> 48
 <211> 416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 48
 gtcagctttc gagctcacgt cagttaccgt ccagtaagcc gccttcgcca ctggtgttcc 60
 tcctaataatc tacgcatttc accgctacac taggaattcc gcttacctct ccggtactct 120
 agattgacag tttccaatgc agtcccgggg ttgagccccg ggttttcaca tcagacttgc 180
 cactccgtct acgctccctt tacaccaggt aaatccggat aacgcttgca ccatacgat 240
 taccgcggct gctggcacgt atttagccgg tgcttcttag tcaggtaccg tcattttctt 300
 ccctgctgat agagctttac ataccgaaat acttcatcgc tcacgcggcg tcgctgcatc 360

agggtttccc ccattgtgca atattcccca ctgctgcctc ccgtggggag tttgga 416

<210> 49
 <211> 421
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 49
 gatgctcagc tttcgtgctc agtgtcagtt tcagtccagt aagccgcctt cgccactgat 60
 gttcctccta atatctacgc atttcaccgc tacactagga attccgctta cctctcctgc 120
 actccagtct gacagtttca aaagcagtcc cagagttaag ccctggggtt tcacttctga 180
 cttgccatac cacctacgca ccctttacac ccagtaattc cggataacgc ttgcccccta 240
 cgtattaccg cggctgctgg cacgtagtta gccggggcct cttagtcagg taccgtcatt 300
 ttcttccctg ctgatagagc ttacatacc gagatacttc ttcactcacg cggcgctcgt 360
 gcatcagggt ttccccatt gtgcaatatt cccactgct gcctcccgaa ggaagtttgg 420
 a 421

<210> 50
 <211> 418
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 50
 gtgtcagctt cgtgctcagt gtcagtttca gtccagtaag ccgccttcgc cactgatgtt 60
 cctcctaata tctacgcatt tcaccgctac actaggaatt ccgcttacct ctctgcact 120
 ccagtctgac agtttcaaaa gcagtcccag agttaagccc tgggttttca cttctgactt 180
 gccataccac ctacgcaccc ttacaccca gtaattccgg ataacgcttg cccctacgt 240
 attaccgagg ctgctggcac gtagttagcc ggggcttctt agtcaggtag cgtcattttc 300
 ttccctgctg atagagcttt acataccgag atacttcttc actcacgagg cgtcgctgca 360
 tcagggtttc cccattgtg caatattccc cactgctgcc tcccgtaggg agtttgga 418

<210> 51
 <211> 415
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 51
 gtcagcttcg agcctcacgt cagttaccgt ccagtaagcc gccttcgcca ctggtgttcc 60
 tcctaataatc tacgcatttc accgctacac taggaattcc gcttacctct ccggtactct 120
 agattgacag tttccaatgc agtcccgggg ttgagccccg ggttttcaca tcagacttgc 180

P074570006W000-SEQLISTING-CEW.txt

cactccgtct acgctccctt tacacccagt aaatccggat aacgcttgca ccatacgtat	240
taccgcggtt gctggcacgt atttagccgg tgcttcttag tcaggtaccg tcattttctt	300
ccctgctgat agagctttac ataccgaaat acttcatcgc tcacgcggcg tcgctgcac	360
agggtttccc ccattgtgca atattcccca ctgctgcctc gcgtaggagt ttgga	415

<210> 52
 <211> 418
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 52	
tgtagcttt cgagctcagc gtcagttatc gtccagtaag ccgccttcgc cactggtggt	60
cctcctaata tctacgcatt tcaccgctac actaggaatt ccgcttacc ctcgacact	120
ctagtacgac agtttccaat gcagtaccgg ggttgagccc cgggctttca catcagactt	180
gccgcaccgc ctgcgctccc ttacaccca gtaaatccgg ataacgcttg caccatacgt	240
attaccgagg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattatc	300
ttccctgctg atagagcttt acataccgaa atacttcttc gtcacgcgg cgctgctgca	360
tcaggctttc gccattgtg caatattccc cactgctgcc tcccgaagg agtttgga	418

<210> 53
 <211> 416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 53	
ttcagctttc gagctcagc tcagttatcg tccagtaagc cgccttcgcc actggtgttc	60
ctcctaatat ctacgcattt caccgctaca ctaggaattc cgcttacc ctcgacactc	120
tagtacgaca gtttccaatg cagtaccggg gttgagcccc gggctttcac atcagacttg	180
ccgcaccgcc tgcgctccct ttacaccag taaatccgga taacgcttgc accatacgt	240
ttaccgcggc tgctggcacg tatttagccg gtgcttctta gtcaggtag gtcattatct	300
tccctgctga tagagcttta cataccgaaa tacttcttcg ctcacgcggc gtcgctgcat	360
caggctttcg ccattgtgc aatattcccc actgctgcct cccgagggga gtttgg	416

<210> 54
 <211> 435
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 54	
ttcggctctgc tttcccttc tcgcgcctca gtgtcagttt ctgtctagta agccgccttc	60

P074570006W000-SEQLISTING-CEW.txt

gccactgatg ttcctcctaa tatctacgca cttcaccgct ccacaatgaa ttccgcttac	120
ccctccccgc ctctagtctg acagttttta aaaaactccc cgagagaaac cctggggttt	180
ttcttctgac atgcgatatc ccacccccac cttttataca cccaaaaatc ggataaaagg	240
tgcgacctac gtattatacc ggctgctggg gcgtagatag ccggggggttc ttatacaggg	300
accgtcattt tctttccgc tgatacagct ttacataccg aaatacttct ttctcacgcg	360
gcgtcgctgc atcagggttt ccccatgtgt gcaatattcc ccactgctgc ctcccgaagg	420
ggaagttggg ggaaa	435

<210> 55
 <211> 418
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 55	
gttcagcttt cgagcctcac gtcagttacc gtccagtaag ccgccttcgc cactggtggt	60
cctcctaata tctacgcatt tcaccgctac actaggaatt ccgcttacct ctccggtact	120
ctagattgac agtttccaat gcagtcccgg ggttgagccc cgggttttca catcagactt	180
gccactccgt ctacgctccc ttacaccca gtaaatccgg ataacgcttg caccatacgt	240
attaccgcgg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattttc	300
ttccctgctg atagagcttt acataccgaa atacttcacg gtcacgcgg cgtcgctgca	360
tcagggtttc ccccatgtgt caatattccc cactgctgcc tcccagggg agtttgga	418

<210> 56
 <211> 416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 56	
gtcagctttc gagctcacgt cagttaccgt ccagtaagcc gccttcgcca ctggtgttcc	60
tcctaataatc tacgcatttc accgctacac taggaattcc gcttacctct ccggtactct	120
agattgacag tttccaatgc agtcccggg ttgagccccg ggttttcaca tcagacttgc	180
cactccgtct acgctccctt tacaccaggt aaatccggat aacgcttgca ccatacgtat	240
taccgcggct gctggcacgt atttagccgg tgcttcttag tcaggtagcg tcattttctt	300
ccctgctgat agagctttac ataccgaaat acttcacg gcacgcggcg tcgctgcatc	360
aggggtttccc ccattgtgca atattcccca ctgctgcctc ccgaggggag tttgga	416

<210> 57
 <211> 417
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 57

tctcacgctt	tcgagctcac	gtcagtcac	gtccagcaag	ccgccttcgc	cactggtgtt	60
cctcctaata	tctacgcatt	tcaccgctac	actaggaatt	ccacttgcc	ctccgacact	120
ctagctcagc	agttccaaat	gcagtcccgc	gggtgagccc	cgggctttca	catctggctt	180
gccgtgccgt	ctacgtcccc	tttacaccca	gtaaatccgc	ataacgcttg	ccccctacgt	240
attaccgcgc	ctgctggcac	gtagttagcc	ggggcttctt	agtcaggtag	cgtcattttc	300
ttccctgctg	atagaagttt	acataccgaa	atacttcac	cttcacgcgc	cgtcgctgca	360
tcagagtttc	ctccattgtg	caatattccc	cactgctgcc	tcccgtaggg	agtttgg	417

<210> 58

<211> 417

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 58

gtcagctttc	gagctcagcg	tcagttatcg	tccagtaagc	cgccttcgcc	actggtgttc	60
ctcctaatat	ctacgcattt	caccgctaca	ctaggaattc	cacttacccc	tccgacactc	120
tagtacgaca	gtttccaatg	cagtaccggg	gttgagcccc	gggctttcac	atcagacttg	180
ccgcaccgcc	tgcgtccct	ttacaccag	taaatccgga	taacgcttgc	accatacgta	240
ttaccgcggc	tgctggcacg	tatttagccg	gtgcttctta	gtcaggtagc	gtcattcttc	300
ttccctgctg	atagagcttt	acataccgaa	atacttcttc	gttcacgcgc	cgtcgctgca	360
tcagggtttc	ccccattgtg	caatattccc	cactgctgcc	tcccagggga	gtttgga	417

<210> 59

<211> 419

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 59

agccccgctt	tcgagcctca	cgtcagttac	cgtccagtaa	gccgccttcg	ccactggtgt	60
tcctcctaata	atctacgcat	ttcaccgcta	cactaggaat	tccgcttacc	tctccggcac	120
tcaagacggg	cagtttccaa	tgagtcctcg	gggttgagcc	ccagcctttc	acatcagact	180
tgtccatccg	tctacgtcc	ctttacacc	agtaaaccg	gataacgctt	gccccctacg	240
tattaccgcg	gctgctggca	cgtagttagc	cggggcttct	tagtcaggta	ccgtcatttt	300
cttccctgct	gatagaagtt	tacataccga	gatacttctt	ccttcacgcg	gcgtcgctgc	360
atcagggttt	ccccattgt	gcaatattcc	ccactgctgc	ctcccgaagg	aagtttgga	419

<210> 60
 <211> 419
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 60
 tgctcagctt tcgagcctca cgtcagttac cgtccagtaa gccgccttcg ccactggtgt 60
 tcttcctaata atctacgcat ttcaccgcta cactaggaat tccgcttacc tctccggcac 120
 tcgagccaga cagtttccaa tgcagtccca ggggttaagcc ctggggtttc acatcagact 180
 tgccttgccg tctacgctcc ctttacaccc agtaaattccg gataacgctt gccccctacg 240
 tattaccgcg gctgctggca cgtagttagc cggggcttct tagtcaggta ccgtcattat 300
 cttccctgct gatagagctt tacataccga aatacttctt cgctcacgcg gcgtcgctgc 360
 atcagggttt cccccattgt gcaatattcc ccactgctgc ctcccgaagg gagtttgga 419

<210> 61
 <211> 421
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 61
 gttgctcagc tttcgagctc acgtcagtta ccgtccagta agccgccttc gccactggtg 60
 ttcttcctaa tatctacgca tttcaccgct aactaggaat ttccgcttac ctctccggca 120
 ctcgagccag acagtttcca atgcagtccc aggggttaagc cctggggtttt cacatcagac 180
 ttgccttgcc gtctacgctc cttttacacc cagtaaattcc ggataacgct tgccccctac 240
 gtattaccgc ggctgctggc acgtagttag ccggggcttc ttagtcagggt accgtcatta 300
 tcttccctgc tgatagagct ttacataccg aaatacttct tcgctcacgc ggcgtcgctg 360
 catcagggtt tccccattg tgcaatattc cccactgctg cctcccgaag gaaagtttgg 420
 a 421

<210> 62
 <211> 420
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 62
 tgctcagctt tcgagctcag cgtcagttat cgtccagtaa gccgccttcg ccactggtgt 60
 tccttcctaata atctacgcat ttcaccgcta cactaggaat tccacttacc cctccgacac 120
 tctagtacga cagtttccaa tgcagtaccg ggggttagacc ccgggctttc acatcagact 180
 tgccgcaccg cctgcgctcc ctttacaccc agtaaattccg gataacgctt gcaccatacg 240
 tattaccgcg gctgctggca cgtatttagc cgggtgcttct tagtcaggta ccgtcattct 300

P074570006W000-SEQLISTING-CEW.txt

tcttccctgc tgatagagct ttacataccg aaatacttct tcgctcacgc ggcgtcgctg	360
catcaggggtt tccccattg tgcaatattc cccactgctg cctcccgaag ggagtttggg	420

<210> 63
 <211> 446
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 63	
gatgccctgg cttcgcgctc agcgtcagtt gtcgtccaga aagccgcttt cgccactggt	60
gttcctccta atatctacgc atttcaccgc tacactagga attccgcttt cctctccgac	120
actcgagctt cacagtttcg gtcccctcac ggggttaagc cccgcacttt taagaccgac	180
ttgcgatgcc gcctgcgctc cctttacgcc caataattcc ggacaacgct tgccacctac	240
gtattaccgc ggctgctggc acgtagttag ccgtggcttt ctcttacggt accgtcaggg	300
ataacgggta ttgaccgcta tcctgttcgt cccatataac agaactttac aaccggaagg	360
ccgtcatcgt tcacgcggcg ttgctccgct agactttcgt ccattgcgga agattcccca	420
ctgctgcctc ccgggggggag tttgga	446

<210> 64
 <211> 419
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 64	
gtcccgcttt cgagcctcag cgtcagttat cgtccagtaa gccgccttcg ccaactggtgt	60
tcctcctaata atctacgcat ttcaccgcta cactaggaat tccgcttacc cctccgacac	120
tctagtacga cagtttccaa tgcaagtacc ggggttagacc ccgggctttc acatcagact	180
tgccgcaccg cctgcgctcc ctttacaccc agtaaattcc gataacgctt gcaccatacg	240
tattaccgcg gctgctggca cgtatttagc cgggtgcttct tagtcaggta ccgtcattat	300
cttccctgct gatagagctt tacataccga aatacttctt cgctcacgcg gcgtcgctgc	360
atcaggcttt cgccattgt gcaatattcc ccaactgctgc ctcccgaagg gaagtttgg	419

<210> 65
 <211> 418
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 65	
gtgcagcttt cgagctcagc gtcagttatc gtccagtaag ccgccttcgc cactggtgtt	60
cctcctaata tctacgatt tcaccgctac actaggaatt ccgcttacc ctccgacact	120

P074570006W000-SEQLISTING-CEW.txt

ctagtagcgac agtttccaat gcagtagccg ggttgagccc cgggctttca catcagactt	180
gccgcaccgc ctgctgtccc ttacaccca gtaaattccg ataacgcttg caccatacgt	240
attaccgcgg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattatc	300
ttccctgctg atagagcttt acataccgaa atactttctt gtcacgcgg cgtcgtgca	360
tcaggctttc gccattgtg caatattccc cactgctgcc tcccgaagg agtttgga	418

<210> 66
 <211> 420
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 66 tgctcagctt tcgagctcag cgtcagttat cgtccagtaa gccgccttcg ccaactggtgt	60
tcctcctaata atctacgcat ttcaccgcta cactaggaat tccacttacc cctccgacac	120
tctagtagca cagtttcaa tgtagtagcc ggttgagccc cgggctttc acatcagact	180
tgccgcaccg cctgctgcc ctttacccc agtaaattcc gataacgctt gcaccatacg	240
tattaccgcg gctgctggca cgtatttagc cggtagcttct tagtcaggta ccgtcattct	300
tcttccctgc tgatagagct ttacataccg aaatacttct tcgctcacgc ggcgtcgtg	360
catcaggggt tccccattg tgcaatatc cccactgctg cctcccgaag ggagtttgga	420

<210> 67
 <211> 417
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 67 attcagcttt cgagctcacg tcagttaccg tccagtaagc cgccttcgcc actggtgttc	60
ctcctaataat ctacgcattt caccgctaca ctaggaattc cgcttaccac tccggcactc	120
aagcatacca gtttccaatg cagtccaggg gttaagcccc tgcctttcac atcagacttg	180
atacgccgtc tacgctccct ttacaccag taaatccgga taacgctcgc cccctacgta	240
ttaccgcggc tgctggcacg tagttagccg gggcttctta gtcaggtagc gtcattatct	300
tccctgctga tagaagttta cataccgaga tacttcttcc ttcacgcggc gtcgtgcat	360
cagggtttcc cccattgtgc aatattcccc actgctgcct cccgagggaa gtttgga	417

<210> 68
 <211> 416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

P074570006W000-SEQLISTING-CEW.txt

<400> 68
gtgtcagctt tcgagctcac gtcagttacc gtccagtaag cgccttcgc cactggtgtt 60
cttcctaata tctacgcatt tcaccgctac actaggaatt cgccttacct ctccggcact 120
cgagccagac agtttccaat gcagtcccag ggttaagccc tgggttttca catcagactt 180
gccttgccgt ctacgtccc tttacacca gtaaatecgg ataacgcttg cccctacgt 240
attaccgagg ctgctggcac gtagttagcc ggggttctt agtcaggtag cgtcattatc 300
ttccctgctg atagagcttt acataccgaa atacttcttc gtcacgcgg cgtcgctgca 360
tcagggtttc cccattgtg caatattccc cactgctgcc tcccaggga gtttgg 416

<210> 69
<211> 416
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 69
gccagcttcg agcctcacgt cagtcacgt ccagtaagcc gccttcgcca ctggtgttcc 60
tcctaataatc tacgcatttc accgctacac taggaattcc attacctct ccgacactct 120
agctgcacag tttccaaagc agtccacagg ttgagcccat gcctttcact tcagacttgc 180
acagccgtct acgctccctt tacaccagc aaatccggat aacgcttgcc ccctacgtat 240
taccgaggct gctggcacgt agttagccgg ggcttcttag tcaggtagcg tcatcttctt 300
ccctgctgat agaagtttac ataccgaaat acttcacct tcacggggcg tcgctgcac 360
aggctttcgc ccattgtgca atattccca ctgctgcctc ccgagggaag tttgga 416

<210> 70
<211> 418
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 70
tgatcagctt tcgagctcac gtcagttacc gtccagtaag cgccttcgc cactggtgtt 60
cctcctaata tctacgcatt tcaccgctac actaggaatt cgccttacct ctccggtact 120
ctagattgac agtttccaat gcagtcccgg ggttgagccc cgggttttca catcagactt 180
gccactccgt ctacgtccc tttacacca gtaaatecgg ataacgcttg caccatacgt 240
attaccgagg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattttc 300
ttccctgctg atagagcttt acataccgaa atacttcac gtcacgcgg cgtcgctgca 360
tcagggtttc cccattgtg caatattccc cactgctgcc tcccggggg agtttggga 418

<210> 71
<211> 420
<212> DNA
<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 71

gatgatcagc tttcgagctc acgtcagtta ccgtccagta agccgccttc gccactggtg	60
ttcctcctaa tatctacgca tttcaccgct aactagga ttccgcttac ctctccggca	120
ctctagaaaa acagtttcca atgcagtcct ggggttaagc cccagccttt cacatcagac	180
ttgctcttcc gtctacgctc cctttacacc cagtaaattc ggataacgct tgccccctac	240
gtattaccgc ggctgctggc acgtagttag ccggggcttc ttagtcaggt accgtcattt	300
tcttccctgc tgatagaagt ttacataccg agatacttct tccttcacgc ggcgtcgtg	360
catcagggtt tccccattg tgcaatattc cccactgctg cctcccgggg gaagtttgga	420

<210> 72

<211> 418

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 72

ttgatcagct ttcgagctca cgtcagttac cgtccagtaa gccgccttcg cactggtgt	60
tcctccta atctacgcat ttcaccgcta cactaggaat tccgcttacc tctccggcac	120
tctagaaaaa cagtttccaa tgcagtcctg ggggttaagc ccagcctttc acatcagact	180
tgctcttccg tctacgctcc ctttacaccc agtaaattcc gataacgctt gccccctacg	240
tattaccgcg gctgctggca cgtagttagc cggggcttct tagtcaggta ccgtcatttt	300
cttccctgct gatagaagtt tacataccga gatacttctt ccttcacgcg gcgtcgtgc	360
atcagggttt cccccattgt gcaatattcc cactgctgc ctcccgaagg gagtttgg	418

<210> 73

<211> 421

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 73

gtatttcagc tttcgagctc agcgtcagtt atcgtccagt aagccgcctt cgccactggt	60
gttcctccta atatctacgc atttcaccgc tacactagga attccgctta cccctccgac	120
actctagtag gacagtttcc aatgcagtag cgggggttag ccccggtt tcacatcaga	180
cttgccgcac cgctgctgct ccctttacac ccagtaaattc cggataacgc ttgcaccata	240
cgtattaccg cggctgctgg cacgtattta gccggtgctt cttagtcagg taccgtcatt	300
atcttccctg ctgatagagc tttacatacc gaaatacttc ttcgctcacg cggcgtcgt	360
gcatcaggct ttcgcccatt gtgcaatatt cccactgct gcctcccga gggagtttgg	420

a

421

<210> 74
 <211> 419
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 74
 gctcagcttt cgagctcagc gtcagttatc gtccagtaag ccgccttcgc cactggtggt 60
 cctcctaata tctacgcatt tcaccgctac actaggaatt ccacttacc ctccgacact 120
 ctagtacgac agtttccaat gcagtaccgg ggttgagccc cgggctttca catcagactt 180
 gccgcaccgc ctgcgctccc ttacaccca gtaaattccgg ataacgcttg caccatacgt 240
 attaccgagg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattctt 300
 cttccctgct gatagagctt tacataccga aatacttctt cgctcacgag gcgtcgctgc 360
 atcagggttt cccccattgt gcaatattcc ccactgctgc ctcccaggag gagtttgga 419

<210> 75
 <211> 417
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 75
 ttccagcttc gagcctcagc gtcagttatc gtccagtaag ccgccttcgc cactggtggt 60
 cctcctaata tctacgcatt tcaccgctac actaggaatt ccgcttacc ctccgacact 120
 ctagtacgac agtttccaat gcagtaccgg ggttgagccc cgggctttca catcagactt 180
 gccgcaccgc ctgcgctccc ttacaccca gtaaattccgg ataacgcttg caccatacgt 240
 attaccgagg ctgctggcac gtatttagcc ggtgcttctt agtcaggtag cgtcattatc 300
 ttccctgctg atagagcttt acataccgaa atacttcttc gctcacgagg cgtcgctgca 360
 tcaggctttc gccattgtg caatattccc cactgctagc tcccgaagga gtttgga 417

<210> 76
 <211> 419
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 76
 agtcagctt tcgagctcag cgtcagttat cgtccagtaa gccgccttcg ccactggtgt 60
 tcctcctaata atctacgcat ttaccgcta cactaggaat tccgcttacc cctccgacac 120
 tctagtacga cagtttccaa tgcagtaccg ggggttgagcc ccgggctttc acatcagact 180
 tgccgcaccg cctgcgctcc ctttacacc agtaaattccg gataacgctt gcaccatacg 240
 tattaccgag gctgctggca cgtatttagc cgggtgcttct tagtcaggta ccgtcattat 300

P074570006W000-SEQLISTING-CEW.txt

cttccctgct gatagagctt tacataaccga aatacttctt cgctcacgcg gcgtcgctgc	360
atcaggcttt cgcccattgt gcaatattcc ccactgctgc ctcccgaagg gagtttgga	419

<210> 77
 <211> 418
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 77 gtccagcttt cgagcctcag cgtcagttat cgtccagtaa gccgccttcg ccactggtgt	60
tcctcctaata atctacgcat ttcaccgcta cactaggaat tccgcttacc cctccgacac	120
tctagtacga cagtttcaa tgcagtaccg gggttgagcc ccgggctttc acatcagact	180
tgccgcaccg cctgcgctcc ctttacaccc agtaaattccg gataacgctt gcaccatacg	240
tattaccgcg gctgctggca cgtatttagc cggtgcttct tagtcaggta ccgtcattat	300
cttccctgct gatagagctt tacataaccga aatacttctt cgctcacgcg gcgtcgctgc	360
atcaggcttt cgcccattgt gcaatattcc ccactgctgc ctcccgaagg agtttgga	418

<210> 78
 <211> 193
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 78 tgagccgggc tcaccccggt actgcattgg aactgtcgta ctagagtgtc ggaggggtaa	60
gcggaattcc tagtgtagcg gtgaaatgcg tagatattag gaggaacacc agtggcgaag	120
gcggcttact ggacgataac tgacgctgag gctcgaaagc gtggggagca aacaggatta	180
gatacaccgg taa	193

<210> 79
 <211> 424
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 79 cacgatgtca gctttcgagc tcagcgtcag ttatcgtcca gtaagccgcc ttcgccactg	60
gtgttcctcc taatatctac gcatttcacc gctacactag gaattccact taccctccg	120
acactctagt acgacagttt ccaatgcagt accgggggtg agccccgggc tttcacatca	180
gacttgccgc accgcctgcg ctccctttac acccagtaaa tccggataac gcttgacca	240
tacgtattac cgcggtgct ggcacgtatt tagccggtgc ttcttagtca ggtaccgtca	300
ttcttcttcc ctgctgatag agctttacat accgaaatac ttcttcgctc acgcggcgctc	360

P074570006W000-SEQLISTING-CEW.txt

gctgcatcag ggtttcccc attgtgcaat attccccact gctgcctccc gaggggagtt 420
tgga 424

<210> 80
<211> 1524
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 80
atggagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcctaata catgcaagtc 60
gaacgcgagc acttgtgctc gaggggcgaa cgggtgagta atacataagt aacctgccct 120
agacaggggg ataactattg gaaacgatag ctaagaccgc ataggtacgg aactgcatg 180
gtgaccgtat taaaagtgcc tcaaagcact ggtagaggat ggacttatgg cgcattagct 240
ggttggcggg gtaacggccc accaaggcga cgatgcgtag ccgacctgag aggggtgaccg 300
gccacactgg gactgagaca cggcccagac tcctacggga ggcagcagta gggaattttc 360
ggcaatgggg gaaaccctga ccgagcaacg ccgcgtgaag gaagaagggtt ttcggattgt 420
aaacttctgt tataaaggaa gaacggcggc tacaggaaat ggtagccgag tgacgggtact 480
ttattagaaa gccacggcta actacgtgcc agcagccgcg gtaatacgta ggtggcaagc 540
gttatccgga attattgggc gtaaagaggg agcaggcggc agcaagggtc tgtggtgaaa 600
gcctgaagct taacttcagt aagccataga aaccaggcag ctagagtgcg ggagaggatc 660
gtggaattcc atgtgtagcg gtgaaatgcg tagatatatg gaggaacacc agtggcgaag 720
gcgacgatct ggctgcaac tgacgtcag tcccgaaagc gtggggagca aataggatta 780
gataccctag tagtccacgc cgtaaacgat gagtactaag tgttggtatgt caaagttcag 840
tgctgcagtt aacgcaataa gtactccgcc tgagtagtac gttcgcaaga atgaaactca 900
aaggaattga cggggggccc cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg 960
aagaacctta ccaggctctg acatactcat aaaggctcca gagatggaga gatagctata 1020
tgagatacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg ggttaagtcc 1080
cgcaacgagc gcaaccctta tcgttagtta ccatcattaa gttggggact ctagcgagac 1140
tgccagtgcg aagctggagg aaggcgggga tgacgtcaaa tcatcatgcc ctttatgacc 1200
tgggctacac acgtgctaca atggatggtg cagaggggaag cgaagccgcg aggtgaagca 1260
aaaccataa aaccattctc agttcggatt gtagtctgca actcgactac atgaagttgg 1320
aatcgctagt aatcgcaat cagcatgtcg cggatgaatac gttctcgggc cttgtacaca 1380
ccgcccgtca caccacgaga gttgataaca cccgaagccg gtggcctaac cgcaaggaag 1440
gagctgtcta aggtgggatt gatgattggg gtgaagtcgt aacaagggtat ccctacggga 1500
acgtggggat ggatcacctc cttt 1524

<210> 81
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 81
 atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg 60
 aacgaagcaa ttgaaggaag ttttcggatg gaattcgatt gactgagtgg cggacgggtg 120
 agtaacgcgt ggataacctg cctcacactg ggggataaca gttagaaatg actgctaata 180
 ccgcataagc gcacagtacc gcatggtaca gtgtgaaaaa ctccggtggt gtgagatgga 240
 tccgcgtctg attagccagt tggcggggta acggcccacc aaagcgacga tcagtagccg 300
 acctgagagg gtgaccggcc acattgggac tgagacacgg cccaaactcc tacgggagggc 360
 agcagtgggg aatattgcac aatgggcgaa agcctgatgc agcgacgccg cgtgagtгаа 420
 gaagtatttc ggtatgtaaa gctctatcag cagggaagaa aatgacggta cctgactaag 480
 aagccccggc taactacgtg ccagcagccg cggtaatagc tagggggcaa gcgttatccg 540
 gatttactgg gtgtaaaggg agcgtagacg gcgaagcaag tctgaagtga aaaccagggg 600
 ctcaaccctg ggactgcttt ggaaactgtt ttgctagagt gtcggagagg taagtggaat 660
 tcctagtgtg gcggtgaaat gcgtagatat taggaggaac accagtggcg aaggcggctt 720
 actggacgat aactgacgtt gaggctcgaa agcgtgggga gcaaacagga ttagataccc 780
 tggtagtcca cgccgtaaac gatgaatgct aggtgttggg gggcaaagcc cttcgggtgcc 840
 gtcgcaaacg cagtaagcat tccacctggg gactacgttc gcaagaatga aactcaaagg 900
 aattgacggg gacccgcaca agcgggtggag catgtgtgtt aattcgaagc aacgcgaaga 960
 accttaccaa gtcttgacat cctcttgacc ggcgtgtaac ggcgccttcc cttcggggca 1020
 agagagacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg ggttaagtcc 1080
 cgcaacgagc gcaaccctta tccttagtag ccagcaggta aagctgggca ctctagggag 1140
 actgccaggg ataacctgga ggaagggtgg gatgacgtca aatcatcatg ccccttatga 1200
 tttgggctac acacgtgcta caatggcgta aacaaagggg agcaagacag tgatgtggag 1260
 caaatcccaa aaataacgtc ccagttcgga ctgtagtctg caaccgcact acacgaagct 1320
 ggaatcgcta gtaatcgca atcagaatgt cgcggtgaat acgttcccgg gtcttgtaca 1380
 caccgcccgt cacaccatgg gagtcagcaa cgcccgaagt cagtgacca actcgcaaga 1440
 gagggagctg ccgaaggcgg ggcaggtaac tggggtgaag tcgtaacaag gtagccgtat 1500
 cggaaggtgc ggctggatca cctccttt 1528

<210> 82
 <211> 385
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 82
aattcgacgt tgtccggatt actgggcgta aagggagcgt aggcggactt ttaagtgaga 60
tgtgaaatac ccgggctcaa cttgggtgct gcatttcaaa ctggaagtct agagtgcagg 120
agaggagaat ggaattccta gtgtagcggg gaaatgcgta gagattagga agaacaccag 180
tggcgaaggc gattctctgg actgtaactg acgctgaggc tcgaaagcgt ggggagcaaa 240
caggattaga taccctggta gtccacgccg taaacgatga atactaggtg taggggttgt 300
catgacctct gtgccgccgc taacgcatta agtattccgc ctggggagta cggtcgcaag 360
attaaaactc aaagaaattg acgga 385

<210> 83

<211> 434

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 83
acgagcgtat cggattattg ggtttaaggg agcgtaggtg gattgttaag tcagttgtga 60
aagtttgagg ctcaaccgta aaattgcagt tgaaactggc agtcttgagt acagtagagg 120
tgggcggaat tcgtggtgta gcggtgaaat gcttagatat cacgaagaac tccgattgag 180
aaggcagctc actagactgt cactgacact gatgctcgaa agtgtgggta tcaaacagga 240
ttagataccc tggtagtcca cacagtaaac gatgaatact cgctgtttgc gatatacagt 300
aagcggccaa gcgaaagcat taagtattcc acctggggag tacgccggca acggtgaaac 360
tcaaagaaat tgacggaagc ccgcccaggg gggaaaaaca tgggggttag ttggatgata 420
cggggaggaa cctc 434

<210> 84

<211> 1457

<212> DNA

<213> Ruminococcus obeum

<220>

<221> misc_feature

<222> (65)..(66)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (960)..(960)

<223> n is a, c, g, t or u

<400> 84
ggcggcgtgc ttaacacatg caagtcgaac gggaaacctt tcattgaagc ttcggcagat 60
ttggnnrtgtt tctagtggcg gacgggtgag taacgcgtgg gtaacctgcc ttatacaggg 120
ggataacaac cagaaatggg tgctaatacc gcataagcgc acaggaccgc atgggtctggt 180
gtgaaaaact ccggtgggtat aagatggacc cgcgttggat tagctagttg gcagggtaac 240

P074570006W000-SEQLISTING-CEW.txt

ggcctaccaa ggcgacgac catagccggc ctgagagggg gaacggccac attgggactg	300
agacacggcc cagactccta cgggaggcag cagtggggaa tattgcacaa tgggggaaac	360
cctgatgcag cgacgccgcg tgaaggaaga agtatctcgg tatgtaaact tctatcagca	420
gggaagatag tgacggtacc tgactaagaa gccccggcta actacgtgcc agcagccgcg	480
gtaatacgta gggggcaagc gttatccgga tttactgggt gtaaagggag cgtagacgga	540
ctggcaagtc tgatgtgaaa ggcgggggct caaccctgg actgcattgg aaactgttag	600
tcttgagtgc cggagaggta agcgggaattc ctagtgtagc ggtgaaatgc gtagatatta	660
ggaggaacac cagtggcgaa ggcggcttac tggacggtaa ctgacgttga ggctcgaaag	720
cgtggggagc aaacaggatt agataccctg gtagtcacg ccgtaaacga tgattactag	780
gtgttgggga gcaaagctct tcggtgccgc cgcaaacgca ttaagtattc cacctgggga	840
gtacgttcgc aagaatgaaa ctcaaaggaa ttgacgggga cccgcacaag cggaggagca	900
tgtggtttaa ttcgaagcaa cgcaagaac cttaccaagt cttgacatcc ctctgaccgn	960
cccttaaccg gatctttcct tcgggacagg ggagacaggt ggtgcatggt tgctcgtcagc	1020
tcgtgtcgtg agatgttggg ttaagtcccg caacgagcgc aaccctatc cccagtagcc	1080
agcagtccgg ctgggcactc tgaggagact gccagggata acctggagga aggcggggat	1140
gacgtcaaat catcatgcc cttatgattt gggctacaca cgtgctacaa tggcgtaaac	1200
aaagggaagc aagcctgcga aggtaagcaa atccccaaaa taacgtcca gttcggactg	1260
cagtctgcaa ctcgactgca cgaagctgga atcgctagta atcgcggatc agaatgccgc	1320
ggtgaatacg ttcccgggtc ttgtacacac cgcccgtcac accatgggag tcagtaacgc	1380
ccgaagtcag tgacctaaact gcaaagaagg agctgccgaa ggcgggaccg atgactgggg	1440
tgaagtcgta acaagg	1457

<210> 85
 <211> 1398
 <212> DNA
 <213> Ruminococcus obeum

<220>
 <221> misc_feature
 <222> (1083)..(1083)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1207)..(1207)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1233)..(1233)
 <223> n is a, c, g, t or u

<400> 85	
ggcgtgctta acacatgcaa gtcgaacggg aaacttttca ttgaagcttc ggcagatttg	60

P074570006W000-SEQLISTING-CEW.txt

gtctgtttct agtggcggac ggggtgagtaa cgcgtgggta acctgcctta tacaggggga	120
taacaaccag aaatggttgc taataccgca taagcgcaca ggaccgcatg gtctggtgtg	180
aaaaactccg gtggtataag atggacccgc gttggattag ctagttggca gggtaacggc	240
ctaccaaggc gacgatccat agccggcctg agagggtgaa cggccacatt gggactgaga	300
cacgccccag actcctcggg aggcagcagt ggggaatatt gcacaatggg ggaaaccctg	360
atgcagcgac gccgcgtgaa ggaagaagta tctcggatg taaacttcta tcagcagggg	420
agatagtgac ggtacctgac taagaagccc cgkctaacta cgtgccagca gccgcggtaa	480
tacgtagggg gcaagcgta tccggattta ctgggtgtaa agggagcgta gacggactgg	540
caagtctgat gtgaaaggcg ggggctcaac ccctggactg cattggaaac tgtagtctt	600
gagtgccgga gaggtaagcg gaattcctag tgtagcgggtg aaatgcgtag atattaggag	660
gaacaccagt ggcgaaggcg gcttactgga cggtaactga cgttgaggct cgaaagcggtg	720
gggagcaaac aggattagat accctggtag tccacgccgc aaacgatgaa tactaggtgt	780
tggggagcaa agctcttcgg tgccgccgca aacgcattaa gtattccacc tggggagtac	840
gttcgcaaga atgaaactca aaggaattga cggggacccg cacaagcggg ggagcatgtg	900
gtttaattcg aagcaacgag aagaacctta ccaagtcttg acatccctct gaccgtccct	960
taaccggatc tttccttcgg gacaggggag acaggtgggtg catggttgtc gtcagctcgt	1020
gtcgtgagat gttgggttaa gtcccgcaac gagcgcaacc cctatcccca gtagccagca	1080
gtncggctgg gcactctgag gagactgcca gggataacct ggaggaaggc ggggatgacg	1140
tcaaatcatc atgcccctta tgatttgggc tacacacgtg ctacaatggc gtaaacaaag	1200
ggaagcnagc ctkcgraggt aagcaaattc canaaataac gtcccagttc ggactgcagt	1260
ctgcaactcg actgcacgaa gctggaatcg ctagtaatcg cggatcagaa tgccgcgggtg	1320
aatacgttcc cgggtcttgt acacaccgcc cgtcacacca tgggagtcag taacgcccga	1380
agtcagtgac ctaactgc	1398

<210> 86
 <211> 1472
 <212> DNA
 <213> Clostridium disporicum

<400> 86	
gctcaggacg aacgctggcg gcgtgcctaa cacatgcaag tcgagcgagt tgattctctt	60
cggagatgaa gctagcggcg gacgggtgag taacacgtgg gcaacctgcc tcatagaggg	120
gaatagcctc ccgaaaggga gattaatacc gcataagatt gtagcttcgc atgaagtagc	180
aattaaagga gcaatccgct atgagatggg cccgcggcgc attagctagt tggtgaggta	240
acggctcacc aaggcgacga tgcgtagccg acctgagagg gtgatcggcc acattgggac	300
tgagacacgg ccagactcc tacgggaggc agcagtgggg aatattgcac aatgggggaa	360
accctgatgc agcaacgccg cgtgagtgat gacggccttc gggttgtaaa gctctgtctt	420
cagggacgat aatgacggta cctgaggagg aagccacggc taactacgtg ccagcagccg	480

P074570006W000-SEQLISTING-CEW.txt

cggtaatac	taggtggcg	gcgttgctccg	gattttactgg	gcgtaaagg	agcgtaggcg	540
gacttttaag	tgagatgtga	aatacccggg	ctcaacttgg	gtgctgcatt	tcaaactgga	600
agtctagagt	gcaggagagg	agaatggaat	tcctagtgtg	gcggtgaaat	gcgtagagat	660
taggaagaac	accagtggcg	aaggcgattc	tctggactgt	aactgacgct	gaggctcgaa	720
agcgtgggga	gcaaacagga	ttagataccc	tggtagtcca	cgccgtaaac	gatgaatact	780
aggtgtaggg	gttgtcatga	cctctgtgcc	gccgctaacg	cattaagtat	tccgcctggg	840
gagtacggtc	gcaagattaa	aactcaaagg	aattgacggg	ggcccgccaca	agcagcggag	900
catgtggttt	aattcgaagc	aacgcgaaga	accttaccta	gacttgacat	ctcctgaatt	960
acccgtaact	ggggaagcca	cttcgggtggc	aggaagacag	gtggtgcatg	gttgtcgtca	1020
gctcgtgtcg	tgagatgttg	ggttaagtcc	cgcaacgagc	gcaaccctta	ttgttagttg	1080
ctaccattta	gttgagcact	ctagcgagac	tgcccggtt	aaccgggagg	aagggtgggga	1140
tgacgtcaaa	tcatcatgcc	ccttatgtct	agggctacac	acgtgctaca	atggcaagta	1200
caaagagaag	caagaccg	aggtggagca	aaactcaaaa	acttgctctca	gttcggattg	1260
taggctgaaa	ctcgcctaca	tgaagctgga	gttgctagta	atcgcgaaatc	agcatgtcgc	1320
ggtgaatac	ttcccggg	ttgtacacac	cgcccgtcac	accatgagag	ttggcaatac	1380
ccaacgtac	tgatctaacc	cgcaagggag	gaagcgtcct	aaggtagggg	cagcgattgg	1440
ggtgaagtc	taacaaggta	gccgtaggag	aa			1472

```
<210> 87
<211> 1529
<212> DNA
<213> Clostridium scindens
```

<400>	87							
gagagttttga	tcctgggtca	ggatgaacgc	tggcggcgctg	cctaacacat	gcaagtcgaa		60	
cgaagcgcct	ggccccgact	tcttcggaac	gaggagcctt	gcgactgagt	ggcggacggg		120	
tgagtaacgc	gtgggcaacc	tgctttgcac	tgggggataa	cagccagaaa	tggctgctaa		180	
taccgcataa	gaccgaagcg	ccgcatggcg	cggcggccaa	agccccggcg	gtgcaagatg		240	
ggcccgcgtc	tgattaggta	gttggcgggg	taacggccca	ccaagccgac	gatcagtagc		300	
cgacctgaga	gggtgaccgg	ccacattggg	actgagacac	ggcccagact	cctacggggag		360	
gcagcagtg	ggaatattgc	acaatggggg	aaaccctgat	gcagcgacgc	cgcgtaaggg		420	
atgaagtatt	tcggtatgta	aacttctatc	agcaggggaag	aagatgacgg	tacctgacta		480	
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc		540	
cggatttact	gggtgtaaag	ggagcgtaga	cggcgatgca	agccagatgt	gaaagcccgg		600	
ggctcaaccc	cgggactgca	tttggaactg	cgtggctgga	gtgtcggaga	ggcaggcgga		660	
attcctagt	tagcggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc		720	
ctgctggacg	atgactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac		780	

P074570006W000-SEQLISTING-CEW.txt

cctggtagtc	cacgccgtaa	acgatgacta	ctaggtgtcg	ggtggcaagg	ccattcggtg	840
ccgcagcaaa	cgcaataagt	agtccacctg	gggagtagct	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgggtg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	tgatcttgac	atcccgatgc	caaagcgcgt	aacgcgctct	ttcttcggaa	1020
catcggtagc	aggtgggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcaacccc	tatcttcagt	agccagcatt	ttggatgggc	actctggaga	1140
gactgccagg	gagaacctgg	aggaagggtg	ggatgacgtc	aaatcatcat	gccccctatg	1200
accagggcta	cacacgtgct	acaatggcgt	aaacaaaggg	aggcgaaccc	gcgaggggtg	1260
gcaaatccca	aaaataacgt	ctcagttcgg	attgtagtct	gcaactcgac	tacatgaagt	1320
tggaatcgct	agtaatcgcg	aatcagaatg	tcgcggtgaa	tacgttcccg	ggtcttgtac	1380
acaccgcccg	tcacaccatg	ggagtcagta	acgcccgaag	ccggtgaccc	aaccgcgaag	1440
ggagggagcc	gtcgaagggt	ggaccgataa	ctggggtgaa	gtcgtaacaa	ggtagccgta	1500
tcggaagggt	cggctggatc	acctccttc				1529

<210> 88
 <211> 1456
 <212> DNA
 <213> Anaerostipes caccae

<400> 88	
gcgcttaata	catgtcaagt
cgaacgaagc	atttaggatt
gaagttttcg	gatggatttc
	60
ctatatgact	gagtggcgga
cgggtgagta	acgcgtgggg
aacctgccct	atacaggggg
	120
ataacagctg	gaaacggctg
ctaataccgc	ataagcgcac
agaatcgcat	gattcagtgt
	180
gaaaagccct	ggcagtatag
gatggtcccg	cgtctgatta
gctggttggt	gaggtaacgg
	240
ctcaccaagg	cgacgatcag
tagccggctt	gagagagtga
acggccacat	tgggactgag
	300
acacggccca	aactcctacg
ggaggcagca	gtgggggaata
ttgcacaatg	ggggtaaacc
	360
ctgatgcagc	gacgccgcgt
gagtgaagaa	gtatttcggt
atgtaaagct	ctatcagcag
	420
ggaagaaaac	agacggtacc
tgactaagaa	gccccggcta
actacgtgcc	agcagccgcg
	480
gtaatacgta	gggggcaagc
gttatccgga	attactgggt
gtaaaggggtg	cgtaggtggc
	540
atggtaagtc	agaagtgaag
gccccgggct	taaccccggg
actgcttttg	aaactgtcat
	600
gctggagtgc	aggagaggta
agcgggaattc	ctagtgtagc
ggtgaaatgc	gtagatatta
	660
ggaggaacac	cagtggcgaa
ggcggcttac	tggactgtca
ctgacactga	tgcacgaaag
	720
cgtggggagc	aaacaggatt
agataccctg	gtagtccacg
ccgtaaacga	tgaataactag
	780
gtgtcggggc	cgtagaggct
tcggtgccgc	agcaaacgca
gtaagtattc	cacctgggga
	840
gtacgttcgc	aagaatgaaa
ctcaaaggaa	ttgacgggga
ccgcacaaag	cggtaggagca
	900
tgtggtttaa	ttcgaagcaa
cgcgaagaac	cttacctggt
cttgacatcc	caatgaccga
	960
accttaaccg	gttttttctt
tcgagacatt	ggagacaggt
ggtgcatggt	tgctcgtcagc
	1020
tcgtgtcgtg	agatgttggg
ttaagtcccg	caacgagcgc
aaccctatc	tttagtagcc
	1080

P074570006W000-SEQLISTING-CEW.txt

agcatttaag	gtgggcactc	tagagagact	gccagggata	acctggagga	aggtggggac	1140
gacgtcaa	atcatgccc	cttatggcca	gggctacaca	cgtgctacaa	tggcgtaa	1200
aaagggaagc	gaagtcgtga	ggcgaagcaa	atcccagaaa	taacgtctca	gttcggattg	1260
tagtctgcaa	ctcgactaca	tgaagctgga	atcgctagta	atcgtgaatc	agaatgtcac	1320
ggtgaatacg	ttccccgggtc	ttgtacacac	cgcccgtcac	accatgggag	tcagtaacgc	1380
ccgaagtcag	tgaccaacc	gcaaggaggg	agctgccgaa	ggtgggaccg	ataactgggg	1440
tgaagtcgta	acaagg					1456

<210> 89
 <211> 1456
 <212> DNA
 <213> Marvinbryantia formatexigens

<400> 89	
tggcggcgtg	cttaacacat gcaagtcgag cgaagcattt taaatgaagt tttcggacgg 60
aatttaaaat	gactgagcgg cggacgggtg agtaacgcgt ggataacctg cttatacag 120
ggggataaca	gccagaaatg gctgctaata ccgcataagc gcacgggtacc gcatggtaca 180
gtgtgaaaaa	ctccgggtgtg ataagatggg tccgcgttgg attaggcagt tggcggggta 240
aaggcccacc	aaaccgacga tccatagccg gcctgagagg gtggacggcc acattgggac 300
tgagacacgg	cccagactcc tacgggaggg agcagtgggg aatattgcac aatgggggaa 360
accctgatgc	agcgacgccg cgtgggtgaa gaagtatttc ggtatgtaaa gccctatcag 420
caggaagaa	aatgacggta cctgaccaag aagccccggc taactacgtg ccagcagccg 480
cggtaatacg	taggggggcaa gcgttatccg gatttactgg gtgtaaaggg agcgtagacg 540
gccatgcaag	tctggtgtga aaggcggggg ctcaaccccc ggactgcatt ggaaactgta 600
tggcttgagt	gccggagagg taagcggaat tcctggtgta gcggtgaaat gcgtagatat 660
caggaggaac	accagtggcg aaggcggtt actggacggt aactgacgtt gaggctcgaa 720
agcgtgggga	gcaaacagga ttagataccc tggtagtcca cgccgtaa
aggtgtcggg	ggacacggtc cttcgggtgcc gcagcaa
gagtacgttc	gcaagaatga aactcaaagg aattgacggg gacccgcaca agcgggtggag 900
catgtggttt	aattcgaagc aacgcgaaga accttaccag gtcttgacat ccggacgacc 960
ggacagtaac	gtgtccttcc cttcggggcg tccgagacag gtggtgcatg gttgtcgtca 1020
gctcgtgtcg	tgagatgttg ggttaagtcc cgcaacgagc gcaaccctg ttcccagtag 1080
ccagcattca	ggatgggcac tctggggaga ctgccagga taacctggag gaaggcgggg 1140
atgacgtcaa	atcatcatgc cccttatgat ctgggctaca cacgtgctac aatggcgtga 1200
acagagggaa	gcgaaccgcg gagggggagc aaatcccaga aataacgtcc cagttcggat 1260
tgtagtctgc	aaccgggcta catgaagctg gaatcgctag taatcgcgga tcagcatgcc 1320
gcggtgaata	cgttcccggg tcttgtacac accgcccgtc acaccatggg agtcggaaat 1380

gcccgaagtc agtgacccaa ccggaaggag ggagctgccg aaggcggggc cggtaactgg 1440
 ggtgaagtcg taacaa 1456

<210> 90
 <211> 1568
 <212> DNA
 <213> Lactobacillus mucosae

<400> 90
 agagtttgat cctggctcag gatgaacgcc ggcggtgtgc ctaatacatg caagtcgaac 60
 gcgttgccc aactgattga acgtgcttgc acggacttga cgttggttta ccagcgagtg 120
 gcggacgggt gagtaacacg taggtaacct gcccacaaagc gggggataac atttggaaac 180
 agatgctaata accgcataac aatttgaatc gcatgattca aattttaaag atggcttcgg 240
 ctatcacttt gggatggacc tgcggcgcat tagcttggtg gtagggtaac ggcctaccaa 300
 ggctgtgatg cgtagccgag ttgagagact gatcgccac aatggaactg agacacggtc 360
 catactccta cgggaggcag cagtagggaa tcttcacaa tgggcgcaag cctgatggag 420
 caacaccgag tgagtgaaga agggtttcgg ctcgtaaagc tctgttggtta gagaagaacg 480
 tgcgtgagag caactgttca cgcagtgcg gtatctaacc agaaagtcac ggctaactac 540
 gtgccagcag ccgcggtaat acgtaggtgg caagcgttat ccggatttat tgggcgtaaa 600
 gcgagcgag gcggtttgat aagtctgatg tgaaagcctt tggcttaacc aaagaagtgc 660
 atcggaact gtcagacttg agtcagaag aggacagtgg aactccatgt gtagcggtgg 720
 aatgcgtaga tatatggaag aacaccagtg gcgaaggcgg ctgtctggtc tgcaactgac 780
 gctgaggctc gaaagcatgg gtagcgaaca ggattagata ccctggtagt ccatgccgta 840
 aacgatgagt gctaggtggt ggagggtttc cgcccttcag tgccgcagct aacgcattaa 900
 gcactccgcc tggggagtag gaccgcaagg ttgaaactca aaggaattga cgggggccc 960
 cacaagcggg ggagcatgtg gtttaattcg aagctacgcg aagaacctta ccaggctctg 1020
 acatcttgcg ccaaccctag agatagggcg tttccttcgg gaacgcaatg acagggtggg 1080
 catggtcgtc gtcagctcgt gtcgtgagat gttgggttaa gtcccgcaac gagcgcaacc 1140
 cttgttacta gttgccagca ttcagttggg cactctagtg agactgccgg tgacaaaccg 1200
 gaggaagggt gggacgacgt cagatcatca tgccccttat gacctgggct acacacgtgc 1260
 tacaatggac ggtacaacga gtcgcgaact cgcgagggca agctaatttc ttaaaaccgt 1320
 tctcagttcg gactgcaggc tgcaactcgc ctgcacgaag tcggaatcgc tagtaatcgc 1380
 ggatcagcat gccgcggtga atacgttccc gggccttgta cacaccgccc gtcacacat 1440
 gagagtttgc aacacccaaa gtcggtggg taacccttcg gggagctagc cgcctaagg 1500
 ggggcagatg attaggtgta agtcgtaaca aggtagccgt aggagaacct gcggctggat 1560
 cacctcct 1568

<210> 91
 <211> 1508

<212> DNA

<213> *Turicibacter sanguinis*

<400> 91

```

agagtttgat catggctcag gatgaacgct ggcggcgtgc ctaatacatg caagtcgagc      60
gaaccacttc ggtggtgagc ggcgaacggg tgagtaacac gtaggttata tgcccatcag      120
acggggacaa cgattgaaa cgatcgctaa taccggatag gacgaaagtt taaaggtgct      180
tcggcaccac tgatggatga gcctgcggcg cattagctag ttggtagggt aaaggcctac      240
caaggcgacg atgcgtagcc gacctgagag ggtgaacggc cacactggga ctgagacacg      300
gccagactc ctacgggagg cagcagtagg gaatcttcgg caatgggcga aagcctgacc      360
gagcaacgcc gcgtgaatga tgaaggcctt cgggttgtaa aattctgtta taagggaaga      420
atggctctag taggaaatgg ctagagtgtg acggtacctt atgagaaagc cacggctaac      480
tacgtgccag cagccgcggt aatacgtagg tggcgagcgt tatccggaat tattgggcgt      540
aaagagcgcg caggtggttg attaagtctg atgtgaaagc ccacggctta accgtggagg      600
gtcattggaa actggtcaac ttgagtgcag aagagggaa ggaattcca tgtgtagcgg      660
tgaaatgcgt agagatatgg aggaacacca gtggcgaagg cggcttcctg gtctgtaact      720
gacctgagg cgcgaaagcg tggggagcaa acaggattag ataccctggg agtccacgcc      780
gtaaacgatg agtgctaagt gttgggggtc gaacctcagt gctgaagtta acgcattaag      840
cactccgcct ggggagtagc gtcgcaagac tgaaactcaa aggaattgac ggggaccgcg      900
acaagcggtg gagcatgtgg ttttaattcga agcaacgcga agaaccttac caggtcttga      960
cataccagtg accgtcctag agataggatt ttcccttcgg ggacaatgga tacaggtggt     1020
gcatggttgt cgtcagctcg tgtcgtgaga tgttgggtta agtcccgcaa cgagcgcaac     1080
ccctgtcgtt agttgccagc attcagttgg ggactctaac gagactgcca gtgacaaact     1140
ggaggaaggt ggggatgacg tcaaatacat atgcccctta tgacctgggc tacacacgtg     1200
ctacaatggg ttgtacaaag agaagcgaag cggtgacgtg gagcaaacct cataaagcca     1260
atctcagttc ggattgtagg ctgcaactcg cctacatgaa gttggaatcg ctagtaatcg     1320
cgaatcagca tgtcgcggtg aatacgttcc cgggtcttgt acacaccgcc cgtcacacca     1380
cgagagttta caacacccga agtcagtggc ctaaccgcaa ggagggagct gcctaagggtg     1440
gggtagatga ttgggggtgaa gtcgtaacaa ggtatcccta ccggaagggtg gggttggatc     1500
acctcctt                                     1508

```

<210> 92

<211> 1456

<212> DNA

<213> *Roseburia faecis*

<400> 92

```

gatgaacgct ggcggcgtgc ttaacacatg caagtcgaac gaagcactct atttgatttt      60
cttcggaaat gaagattttg tgactgagtg gcggacgggt gagtaacgcg tgggtaacct      120
gcctcataca gggggataac agttggaaac gactgctaata accgcataag cgcacaggat      180

```

P074570006W000-SEQLISTING-CEW.txt

cgcatgatcc ggtgtgaaaa actccggtgg tatgagatgg acccgcgctct gattagccag	240
ttggcaggggt aacggcctac caaagcgacg atcagtagcc gacctgagag ggtgaccggc	300
cacattggga ctgagacacg gcccaaactc ctacgggagg cagcagtggg gaatattgca	360
caatggggga aaccctgatg cagcgacgcc gcgtgagcga agaagtattt cggtatgtaa	420
agctctatca gcagggaaga agaatgacgg tacctgacta agaagcaccg gctaaatacg	480
tgccagcagc cgcggttaata cgtatggtgc aagcgttatc cggatttact gggtgtaaag	540
ggagcgcagg cgggtcggca agtctgatgt gaaagcccgg ggctcaacc cggactgca	600
ttggaaactg tcgtactaga gtgtcggagg ggtaagtgga attcctagtg tagcgggtgaa	660
atgcgtagat attaggagga acaccagtgg cgaaggcggc ttactggacg ataactgacg	720
ctgaggctcg aaagcgtggg gagcaaacag gattagatac cctggtagtc cacgccgtaa	780
acgatgaata ctaggtgtcg gggagcattg ctcttcgggtg ccgcagcaaa cgcaataagt	840
attccacctg gggagtacgt tcgcaagaat gaaactcaaa ggaattgacg gggacccgca	900
caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa gaaccttacc aagtcttgac	960
atcccgatga cagagtatgt aatgtacytt ctcttcggag catcggtgac aggtggtgca	1020
tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt cccgcaacga gcgcaacccc	1080
tgtccttagt agccagcggg tcggccgggc actctagggg gactgccagg gataacctgg	1140
aggaaggcgg ggatgacgtc aaatcatcat gcccttatg acttgggcta cacacgtgct	1200
acaatggcgt aaacaaaggg aagcggagcc gtgaggccga gcaaattctca aaaataacgt	1260
ctcagttcgg actgtagtct gcaacccgac tacacgaagc tggaatcgct agtaatcgca	1320
gatcagaatg ctgcgggtgaa tacgttcccg ggtcttgtac acaccgcccg tcacaccatg	1380
ggagtggaa atgcccgaag tcagtgacct aaccgcaagg agggagctgc cgaaggcagg	1440
ttcgataact ggggtg	1456

<210> 93
 <211> 1465
 <212> DNA
 <213> Flavonifractor plautii

<400> 93	
cgctggcggc gtgcttaaca catgcaagtc gaacgggggtg ctcattgacgg aggattcgtc	60
caatggattg agttacctag tggcggacgg gtgagtaacg cgtgaggaa ctcgcttgga	120
gaggggaata acactccgaa aggagtgtca ataccgatg aagcagttgg gtcgcatggc	180
tctgactgcc aaagatttat cgctctgaga tggcctcgcg tctgattagc tagtaggcgg	240
ggtaacggcc cacctaggcg acgatcagta gccggactga gaggttgacc ggccacattg	300
ggactgagac acggcccaga ctcctacggg aggcagcagt ggggaatatt gggcaatggg	360
cgcaagcctg acccagcaac gccgcgtgaa ggaagaaggc tttcggggtg taaacttctt	420
ttgtcgggga cgaaacaaat gacggtacct gacgaataag ccacggctaa ctacgtgcc	480

P074570006W000-SEQLISTING-CEW.txt

gcagccgcgg taatacgtag gtggcaagcg ttatccggat ttactgggtg taaagggcgt	540
gtaggcggga ttgcaagtca gatgtgaaaa ctgggggctc aacctccagc ctgcatttga	600
aactgtagtt cttgagtgtc ggagaggcaa tcggaattcc gtgtgtagcg gtgaaatgcg	660
tagatatacg gaggaacacc agtggcgaag gcggattgct ggacagtaac tgacgctgag	720
gcgcgaaagc gtggggagca aacaggatta gataccctgg tagtccacgc cgtaaaccgat	780
ggatactagg tgtgggggggt ctgaccccct ccgtgccgca gttaacacaa taagtatccc	840
acctggggag tacgatcgca aggttgaaac tcaaaggaat tgacgggggc ccgcacaagc	900
ggtggagtat gtggtttaat tcgaagcaac gcgaagaacc ttaccagggc ttgacatccc	960
actaacgagg cagagatgcg ttaggtgccc ttcggggaaa gtggagacag gtggtgcatg	1020
gttgctcgtca gctcgtgtcg tgagatgttg ggttaagtcc cgcaacgagc gcaaccctta	1080
ttgttagttg ctacgcaaga gactctagc gagactgccg ttgacaaaac ggaggaaggt	1140
ggggacgacg tcaaatcatc atgcccctta tgtcctgggc cacacacgta ctacaatggt	1200
ggtaacaga gggaggcaat accgcgaggt ggagcaaadc cctaaaagcc atcccagttc	1260
ggattgcagg ctgaaacccg cctgtatgaa gttggaatcg ctagtaatcg cggatcagca	1320
tgccgcggtg aatacgttcc cgggccttgt acacaccgcc cgtcacacca tgagagtcgg	1380
gaacacccga agtccgtagc ctaaccgcaa ggagggcgcg gccgaaggtg ggttcgataa	1440
ttgggggtgaa gtcgtaacaa ggtag	1465

<210> 94
 <211> 1438
 <212> DNA
 <213> *Blautia wexlerae*

<220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n is a, c, g, t or u

<400> 94	
caagtcgaac gggaattant ttattgaaac ttcggtcgat ttaatttaat tctagtggcg	60
gacgggtgag taacgcgtgg gtaacctgcc ttatacaggg ggataacagt cagaaatggc	120
tgctaatacc gcataagcgc acagagctgc atggctcagt gtgaaaaact ccggtggtat	180
aagatggacc cgcgttggat tagcttgttg gtggggtaac ggcccaccaa ggcgacgac	240
catagccggc ctgagagggg gaacggccac attgggactg agacacggcc cagactccta	300
cgggaggcag cagtggggaa tattgcacaa tgggggaaac cctgatgcag cgacgccgcg	360
tgaaggaaga agtatctcgg tatgtaaact tctatcagca gggaagatag tgacggtacc	420
tgactaagaa gccccggcta actacgtgcc agcagccgcg gtaatacgta gggggcaagc	480
gttatccgga ttactgggt gtaaaggag cgtagacggt gtggcaagtc tgatgtgaaa	540
ggcatgggct caacctgtgg actgcattgg aaactgtcat acttgagtgc cggaggggta	600
agcgggaattc ctagtgtagc ggtgaaatgc gtagatatta ggaggaacac cagtggcgaa	660

P074570006W000-SEQLISTING-CEW.txt

ggcggcttac	tggacggtaa	ctgacgttga	ggctcgaaag	cgtgggggagc	aaacaggatt	720
agataccctg	gtagtccacg	ccgtaaacga	tgaataacta	ggtgtcgggt	ggcaaagcca	780
ttcggtgccg	tcgcaaacgc	agtaagtatt	ccacctgggg	agtacgttcg	caagaatgaa	840
actcaaagga	attgacgggg	acccgcacaa	gcggtggagc	atgtggttta	attcgaagca	900
acgcgaagaa	ccttaccaag	tcttgacatc	cgcttgaccg	atccttaacc	ggatctttcc	960
ttcgggacag	gcgagacagg	tggtgcatgg	ttgtcgtcag	ctcgtgtcgt	gagatgttgg	1020
gttaagtccc	gcaacgagcg	caacccttat	cctcagtagc	cagcatttaa	ggtgggcact	1080
ctggggagac	tgccagggat	aacctggagg	aaggcgggga	tgacgtcaaa	tcatcatgcc	1140
ccttatgatt	tggtgtacac	acgtgctaca	atggcgtaaa	caaagggaag	cgagattgtg	1200
agatggagca	aatcccaaaa	ataacgtccc	agttcggact	gtagtctgca	acccgactac	1260
acgaagctgg	aatcgctagt	aatcgcggat	cagaatgccg	cgggtgaatac	gttcccgggt	1320
cttgtagaca	ccgcccgtca	caccatggga	gtcagtaacg	cccgaagtca	gtgacctaac	1380
tgcaaagaag	gagctgccga	aggcgggacc	gatgactggg	gtgaagtcgt	aacaaggt	1438

<210> 95
 <211> 1365
 <212> DNA
 <213> Anaerotruncus colihominis

<400> 95						
aacggagctt	acgttttgaa	gttttcggat	ggatgaatgt	aagcttagtg	gcggacgggt	60
gagtaacacg	tgagcaacct	gcctttcaga	gggggataac	agccggaaac	ggctgctaata	120
accgcatgat	gttgcggggg	cacatgcccc	tgcaaccaaa	ggagcaatcc	gctgaaagat	180
gggctcgcgt	ccgattagcc	agttggcggg	gtaacggccc	accaaagcga	cgatcggtag	240
ccggactgag	aggttgaacg	gccacattgg	gactgagaca	cggcccagac	tcctacggga	300
ggcagcagtg	ggggatattg	cacaatgggc	gaaagcctga	tgcagcgacg	ccgcgtgagg	360
gaagacggtc	ttcggattgt	aaacctctgt	ctttggggaa	gaaaatgacg	gtacccaaag	420
aggaagctcc	ggctaactac	gtgccagcag	ccgcggtaat	acgtagggag	caagcgttgt	480
ccggaattac	tggtgtgtaa	gggagcgtag	gcgggatggc	aagtagaatg	ttaaattccat	540
cggctcaacc	ggtggctgcg	ttctaaactg	ccgttcttga	gtgaagtaga	ggcaggcgga	600
attcctagtg	tagcggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc	660
ctgctgggct	ttaactgacg	ctgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	720
cctggtagtc	cacgccgtaa	acgatgatta	ctaggtgtgg	ggggactgac	cccttcctgt	780
ccgcagttaa	cacaataagt	aatccacctg	gggagtacgg	ccgcaagggt	gaaactcaaa	840
ggaattgacg	ggggcccgcg	caagcagtgg	agtatgtggt	ttaattcgaa	gcaacgcgaa	900
gaaccttacc	aggtcttgac	atcggcgtaa	tagcctagag	agtaggtgaa	gcccttcggg	960
gcatccagac	aggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tggtttaagt	1020

P074570006W000-SEQLISTING-CEW.txt

cccgcaacga gcgcaaccct tattattagt tgctacgcaa gagcactcta atgagactgc	1080
cgttgacaaa acggaggaag gtggggatga cgtcaaatca tcatgccctt tatgacctgg	1140
gctacacacg tactacaatg gcactaaaac agagggcggc gacaccgca ggtgaagcga	1200
atcccagaaa aagtgtctca gttcagattg caggctgcaa cccgcctgca tgaagtcgga	1260
attgctagta atcgcggatc agcatgccgc ggtgaatacg ttcccgggcc ttgtacacac	1320
cgcccgtcac accatgggag tccgggtaac acccgaagcc agtag	1365

<210> 96
 <211> 1404
 <212> DNA
 <213> Ruminococcus faecis

<400> 96 atgcaagtcg aacgaagcac cttgatttga ttcttcggat gaagatcttg gtgactgagt	60
ggcggacggg tgagtaacgc gtgggtaacc tgcctcatatc agggggataa cagttagaaa	120
tgactgctaa taccgcataa gaccacagca ccgcatgggtg caggggtaaa aactccgggtg	180
gtatgagatg gacccgcgtc tgattaggta gttggtgggg taacggccta ccaagccgac	240
gatcagtagc cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact	300
cctacgggag gcagcagtgg ggaatattgc acaatggggg aaaccctgat gcagcgacgc	360
cgcgtgagcg atgaagtatt tcggtatgta aagctctatc agcagggag aaaatgacgg	420
tacctgacta agaagcaccg gctaaatacg tgccagcagc cgcggttaata cgtatggtgc	480
aagcgttatc cggatttact ggggtgtaaag ggagcgtaga cggagtgga agtctgatgt	540
gaaaacccgg ggctcaacc cgggactgca ttggaaactg tcaatctaga gtaccggaga	600
ggtaagcggg attcctagt tagcggtgaa atgcgtagat attaggagga acaccagtgg	660
cgaaggcggc ttactggacg gtaactgacg ttgaggctcg aaagcgtggg gagcaaacag	720
gattagatac cctggtagtc cacgccgtaa acgatgacta ctaggtgtcg ggcagcaaag	780
ctgttcgggtg ccgcagcaaa cgcaataagt agtccacctg gggagtagct tcgcaagaat	840
gaaactcaaa ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa	900
gcaacgcgaa gaaccttacc tgctcttgac atctccctga ccggcaagta atgttgccctt	960
tccttcggga cagggatgac aggtggtgca tgggtgtcgt cagctcgtgt cgtgagatgt	1020
tgggttaagt cccgcaacga gcgcaacccc tatctttagt agccagcggg ttggccgggc	1080
actctagaga gactgccagg gataacctgg aggaagggtg ggatgacgtc aaatcatcat	1140
gccccttatg agcagggcta cacacgtgct acaatggcgt aaacaaaggg aggcagaacc	1200
gcgaggtcga gcaaatccca aaaataacgt ctgagttcgg attgtagtct gcaactcgac	1260
tacatgaagc tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttccccg	1320
ggtcttgtag acaccgcccg tcacaccatg ggagtcagta acgcccgaag tcagtgaccc	1380
aaccgtaagg aggagctgcc gaag	1404

<210> 97
 <211> 1316
 <212> DNA
 <213> Dorea longicatena

<400> 97
 taacgcgtgg gtaacctgcc tcatacaggg ggataacagt tagaaatgac tgctaataacc 60
 gcataagacc acgtaccgca tgggtacagt gtaaaaactc cgggtggtatg agatggaccc 120
 gcgtctgatt aggtagttag tggggtaacg gcctaccaag cggacgatca gtagccgacc 180
 tgagaggggtg accggccaca ttgggactga gacacggccc agactcctac gggaggcagc 240
 agtgggggaat attgcacaat ggaggaaact ctgatgcagc gacgccgcgt gaaggatgaa 300
 gtatttcggt atgtaaactt ctatcagcag ggaagaaaat gacggtacct gactaagaag 360
 ccccggtctaa ctacgtgccg gcagccgcgg taatacgtag ggggcaagcg ttatccggat 420
 ttactgggtg taaagggagc gtagacggca cggcaagcca gatgtgaaaa gcccggggct 480
 caaccccggt actgcatttg gaactgctga gctagagtgt cggagaggca agtggaaattc 540
 ctagtgtagc ggtgaaatgc gtagatatta ggaggaacac cagtggcgaa ggcggcttgc 600
 tggacgatga ctgacgttga ggctcgaaag cgtggggagc aaacaggatt agataccctg 660
 gtagtccacg ccgtaaacga tgactgctag gtgtcgggtg gcaaagccat tcggtgccgc 720
 agctaacgca ataagcagtc cacctgggga gtacgttcgc aagaatgaaa ctcaaaggaa 780
 ttgacgggga cccgcacaag cgggtggagca tgtggtttta ttcgaagcaa cgcaagaac 840
 cttacctgat cttgacatcc cgatgaccgc ttcgtaatgg aagtttttct tcggaacatc 900
 ggtgacaggt ggtgcatggt tgtcgtcagc tcgtgtcgtg agatgttggg ttaagtccc 960
 caacgagcgc aaccctatc ttcagtagcc agcaggttta gctgggcact ctggagagac 1020
 tgccagggat aacctggagg aaggtgggga tgacgtcaaa tcatcatgcc cttatgacc 1080
 agggctacac acgtgctaca atggcgtaaa caaagagaag cgaactcgcg agggtaagca 1140
 aatctcaaaa ataacgtctc agttcggatt gtagtctgca actcgactac atgaagctgg 1200
 aatcgctagt aatcgcatg cagaatgctg cgggtgaatac gttcccgggt cttgtacaca 1260
 ccgcccgtca caccatggga gtcataacgc ccgaagtcag tgaccaacc gtaagg 1316

<210> 98
 <211> 1475
 <212> DNA
 <213> Clostridium innocuum

<220>
 <221> misc_feature
 <222> (905)..(905)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (942)..(942)
 <223> n is a, c, g, t or u

<400> 98

P074570006W000-SEQLISTING-CEW.txt

atggagagtt tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc	60
gaacgaagtc ttcaggaagc ttgcttccaa aaagacttag tggcgaacgg gtgagtaaca	120
cgtaggtaac ctgcccattgt gtccgggata actgctggaa acggtagcta aaaccggata	180
ggtatacggg gcgcatgctc tgtatattaa agcgcccttc aaggcgtgaa catggatgga	240
cctgcgacgc attagctagt tggtagagga acggcccacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaaccggc acattgggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagtgaa	420
gaaggtcttc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctta ccagaaagcc acggctaact acgtgccagy agccgcggta	540
atacgtaggt ggcaagcgtt atccggaatc attgggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctcaa ccattgtaag ctatggaaac tggtagctg	660
gagtgacaga gagggcgatg gaattccatg tgtagcggta aaatgcgtag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctggc ctgtaactga cactgaggca cgaaagcgtg	780
gggagcaaatt aggattagat accctagtag tccacgccgt aaacgatgag aactaagtgt	840
tggagaaatt cagtgtctga gttaacgcaa taagtcttcc gcctggggag tatgcacgca	900
agttngaaac tcaaaggaat tgacgggggc ccgcacaagc gntggagtat gtggtttaat	960
tcgaagcaac gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg	1080
ttgggttaag tcccgaacg agcgcaacc ttgtcgcatt ttaccagcat caagttgggg	1140
actcatgcca gactgccgtt gacaaaccgg aggaagggtg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcga ccacaaagag cagcgacttg	1260
gtgacaagaa gcgaatctca taaagatcgt ctgagttcgg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgtt agtaatcgca gatcagcatg ctgcggtgaa tacgttctcg	1380
ggccttgtag acaccgccc tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440
aaccgtaagg agtgagccgt cgaaggtagg accga	1475

<210> 99
 <211> 1492
 <212> DNA
 <213> *Blautia hansenii*

<400> 99	
agagtttgat cctggctcag gatgaacgct ggcggcgtgc ttaacacatg caagtcgagc	60
gaagcactta tcattgactc ttcggaagat ttgatatttg actgagcggc ggacgggtga	120
gtaacgcgtg ggtaacctgc ctcatacagg ggaataacag ttagaaatgg ctgctaattgc	180
cgcataagcg cacaggaccg catggtctgg tgtgaaaaac tgaggtggta tgagatggac	240
ccgctctga ttaggtagtt ggtggggtaa cggcctacca agccgacgat cagtagccgg	300
cctgagaggg tgaacggcca cattgggact gagacacggc ccagactcct acgggaggga	360

P074570006W000-SEQLISTING-CEW.txt

gcagtgggga	atattgcaca	atgggggaaa	ccctgatgca	gcgacgccgc	gtgaaggaag	420
aagtatctcg	gtatgtaaac	ttctatcagc	agggaagaaa	atgacgggtac	ctgactaaga	480
agccccggct	aactacgtgc	cagcagccgc	ggtaatacgt	agggggcaag	cgttatccgg	540
atttactggg	tgtaaagga	gcgtagacgg	aagagcaagt	ctgatgtgaa	aggctggggc	600
ttaaccccag	gactgcattg	gaaactgttt	ttctagagt	ccggagaggt	aagcgggaatt	660
cctagtgtag	cggtgaaatg	cgtagatatt	aggaggaaca	ccagtggcga	aggcggctta	720
ctggacggta	actgacgttg	aggctcgaaa	gcgtggggag	caaacaggat	tagataccct	780
ggtagtccac	gccgtaaacg	atgaatacta	ggtgtcgggg	tgcaaagcag	ttcggtgccg	840
cagcaaacgc	aataagtatt	ccacctgggg	agtacgttcg	caagaatgaa	actcaaagga	900
attgacgggg	acccgcacaa	gcggtggagc	atgtggttta	attcgaagca	acgcgaagaa	960
ccttaccaag	tcttgacatc	tgcctgaccg	ttccttaacc	ggagctttcc	ttcgggacag	1020
gcaagacagg	tggtgcatgg	ttgtcgtcag	ctcgtgtcgt	gagatgttgg	gttaagtccc	1080
gcaacgagcg	caaccctat	ccttagtagc	cagcagtcgg	gctgggcact	ctagggagac	1140
tgccggggat	aaccgggagg	aaggcgggga	cgacgtcaaa	tcatcatgcc	ccttatgatt	1200
tgggctacac	acgtgctaca	atggcgtaaa	caaagggaag	cgaagcggtg	acgcttagca	1260
aatctcaaaa	ataacgtccc	agttcggact	gcagtctgca	actcgactgc	acgaagctgg	1320
aatcgctagt	aatcggaat	cagaatgtcg	cggtgaatac	gttcccgggt	cttgtagaca	1380
ccgcccgtca	caccatggga	gtcagtaacg	cccgaagtca	gtgaccaaac	cttatggagg	1440
gagctgccga	aggcgggacc	gataactggg	gtgaagtcgt	aacaaggtaa	cc	1492

<210> 100
 <211> 1473
 <212> DNA
 <213> Bacteroides cellulosilyticus

<400> 100	
agagtttgat	cctgggtcag gatgaacgct agctacaggc ttaacacatg caagtcgagg 60
ggcagcatga	cctagcaata ggttgatggc gaccggcgca cgggtgagta acacgtatcc 120
aacctaccgg	ttattccggg atagcctttc gaaagaaaga ttaataccgg atagtataac 180
gagaaggcat	ctttttgtta ttaaagaatt tcgataaccg atgggggatgc gttccattag 240
tttgttggcg	gggtaacggc ccaccaagac atcgatggat aggggttctg agaggaaggt 300
ccccacatt	ggaactgaga cacggtccaa actcctacgg gaggcagcag tgaggaatat 360
tggtcaatgg	acgagagtct gaaccagcca agtagcgtga aggatgactg ccctatgggt 420
tgtaaacttc	ttttatatgg gaataaagtg agccacgtgt ggctttttgt atgtaccata 480
cgaataagga	tcggctaact ccgtgccagc agccgcggta atacggagga tccgagcggt 540
atccggatth	attgggttta aaggagcgt aggcggacta ttaagtcagc tgtgaaagtt 600
tgcggtcaa	ccgtaaaatt gcagttgata ctgggtcgtct tgagtgcagt agaggtaggc 660

P074570006W000-SEQLISTING-CEW.txt

ggaattcgtg	gtgtagcggg	gaaatgctta	gatatcacga	agaactccga	ttgcgaaggg	720
agcttactgg	actgtaactg	acgctgatgc	tcgaaagtgt	gggtatcaaa	caggattaga	780
taccctggta	gtccacacag	taaacgatga	atactcgctg	tttgcgatat	acagcaagcg	840
gccaagcgaa	agcattaagt	attccacctg	gggagtagcg	cggcaacggg	gaaactcaaa	900
ggaattgacg	ggggcccgcg	caagcggagg	aacatgtggg	ttaattcgat	gatacgcgag	960
gaaccttacc	cgggcttaaa	ttgcatctga	ataatttgga	aacagattag	ccgcaaggca	1020
gatgtgaagg	tgctgcatgg	ttgtcgtcag	ctcgtgccgt	gaggtgtcgg	cttaagtgcc	1080
ataacgagcg	caacccttat	ctttagttac	taacagggtca	tgctgaggac	tctagagaga	1140
ctgccgtcgt	aagatgtgag	gaagggtggg	atgacgtcaa	atcagcacgg	cccttacgtc	1200
cggggctaca	cacgtgttac	aatggggggg	acagaaggca	gctacacagc	gatgtgatgc	1260
taatcccaaa	agcctctctc	agttcggatt	ggagtctgca	acccgactcc	atgaagctgg	1320
attcgctagt	aatcgcgcat	cagccacggc	gcggtgaata	cgttcccggg	ccttgtagac	1380
accgcccgtc	aagccatgaa	agccgggggg	acctgaagtc	cgtaaccgca	aggagcggcc	1440
tagggtaaaa	ctggtaattg	gggctaagtc	gta			1473

<210> 101
 <211> 1459
 <212> DNA
 <213> Bacteroides ovatus

<400> 101	
ggctcaggat	gaacgctagc
tacaggctta	acacatgcaa
gtcgaggggc	agcatttttag
60	
tttgcttgca	aactgaagat
ggcgaccggc	gcacgggtga
gtaacacgta	tccaacctgc
120	
cgataactcc	ggaatagcct
ttcgaaagaa	agattaatac
cggatagcat	acgaatatcg
180	
catgatattt	ttattaaaga
atttcggtta	tcgatgggga
tgcgttccat	tagtttgttg
240	
gcggggtaac	ggcccaccaa
gactacgatg	gataggggtt
ctgagaggaa	ggccccccac
300	
attggaactg	agacacggtc
caaactccta	cgggaggcag
cagtgaggaa	tattgggtcaa
360	
tgggcgagag	cctgaaccag
ccaagtagcg	tgaaggatga
aggctctatg	ggtcgtaaac
420	
ttcttttata	tgggaataaa
gttttccacg	tgtggaattt
tgtatgtacc	atatgaataa
480	
ggatcggcta	actccgtgcc
agcagccgcg	gtaatacgga
ggatccgagc	gttatccgga
540	
tttattgggt	ttaaaggagg
cgtagggtgga	ttgttaagtc
agttgtgaaa	gtttgcggct
600	
caaccgtaaa	attgcagttg
aaactggcag	tcttgagtac
agtagagggtg	ggcggaattc
660	
gtggtgtagc	ggtgaaatgc
ttagatatca	cgaagaactc
cgattgcgaa	ggcagctcac
720	
tagactgtta	ctgacactga
tgctcgaaag	tgtgggtatc
aaacaggatt	agataccctg
780	
gtagtccaca	cagtaaacga
tgaatactcg	ctgtttgcga
tatacagtaa	gcggccaagc
840	
gaaagcatta	agtattccac
ctggggagta	cgccggcaac
ggtgaaactc	aaaggaattg
900	
acggggggccc	gcacaagcgg
aggaacatgt	ggtttaattc
gatgatacgc	gaggaacctt
960	
acccggggctt	aaattgcaac
agaatatatt	ggaaacagta
tagccgtaag	gctgtttgtga
1020	

P074570006W000-SEQLISTING-CEW.txt

aggtgctgca tggttgtcgt cagctcgtgc cgtgaggtgt cggcttaagt gccataacga	1080
gcgcaaccct tatcttttagt tactaacagg ttatgctgag gactctagag agactgccgt	1140
cgtaagatgt gaggaaggtg gggatgacgt caaatcagca cggcccttac gtccggggct	1200
acacacgtgt tacaatgggg ggtacagaag gcagctacct ggcgacagga tgctaattccc	1260
aaaaacctct ctcaagttcgg atcgaagtct gcaacccgac ttcgtgaagc tggattcgct	1320
agtaatcgcg catcagccat ggcgcggtga atacgttccc gggccttgta cacaccgccc	1380
gtcaagccat gaaagccggg ggtacctgaa gtacgtaacc gcaaggagcg tcctagggta	1440
aaactggtaa ttggggcta	1459

<210> 102
 <211> 1526
 <212> DNA
 <213> Eubacterium fissicatena

<400> 102	
tagagtttga tcctgggtca ggatgaacgc tggcggcgtg cttaacacat gcaagtcgag	60
cgaagcgctt tacttagatt tcttcggatt gaagagtttt gcgactgagc ggcggacggg	120
tgagtaacgc gtgggtaacc tgcctcatalc agggggataa cagttagaaa tgactgctaa	180
taccgcataa gaccacagta ccgcatggta cagtgggaaa aactccggtg gtatgagatg	240
gacccgcgtc tgattagcta gttggtaagg taacggctta ccaaggcgac gatcagtagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcagtgg ggaatattgc acaatggggg aaacctgat gcagcgacgc cgcgtaagg	420
atgaagtatt tcggtatgta aacttctatc agcagggaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcagc cgcggttaata cgtagggggc aagcgttatc	540
cggatttact ggggtgtaaag ggagcgtaga cggttatgta agtctgatgt gaaaacccgg	600
ggctcaaccc cgggactgca ttggaaacta tgtaactaga gtgtcggaga ggtaagtgga	660
attcctagtg tagcgggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcggc	720
ttactggacg atcactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cagccgtaa acgatgaata ctaggtgtcg ggtggcaaag ccattcgggtg	840
ccgcagcaaa cgcaataagt attccacctg gggagtacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc tgctcttgac atcccactga ccggcgtgta atggcgctt cccttcgggg	1020
cagtggagac aggtgggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaaccct tatcttttagt agccagcggg ttggccgggc actctagaga	1140
gactgccagg gataacctgg aggaaggtgg ggatgacgtc aaatcatcat gcccttatg	1200
agcagggcta cacacgtgct acaatggcgt aaacaaaggg aggcaatacc gcgaggttga	1260
gcaaatccca aaaataacgt ctcaagttcgg attgtagtct gcaactcgac tacatgaagc	1320

P074570006W000-SEQLISTING-CEW.txt

tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttcccg ggtcttgtac	1380
acaccgccccg tcacaccatg ggagttggta acgcccgaag tcagtgacct aaccgtaagg	1440
agggagctgc cgaaggcggg atcgataact ggggtgaagt cgtaacaagg tagccgtatc	1500
ggaaggtgcg gctggatcac ctcctt	1526

<210> 103
 <211> 1493
 <212> DNA
 <213> *Blautia coccoides*

<400> 103 agagtttgat cctggctcag gatgaacgct ggcggcgtgc ttaacacatg caagtcgagc	60
gaagcgctaa gacagatttc ttcggattga agtctttgtg actgagcggc ggacgggtga	120
gtaacgcgtg ggtaacctgc ctcatcacagg gggataacag ttagaaatga ctgctaatac	180
cgcataagcg cacaggaccg catggtcttg tgtgaaaaac tccggtggta tgagatggac	240
ccgcgtctga ttagctagtt ggaggggtaa cggcccacca aggcgacgat cagtagccgg	300
cctgagaggg tgaacggcca cattgggact gagacacggc ccagactcct acgggaggca	360
gcagtgggga atattgcaca atgggggaaa ccctgatgca gcgacgccgc gtgaaggaag	420
aagtatctcg gtatgtaaac ttctatcagc aggaagaaa atgacggtag ctgactaaga	480
agccccggct aactacgtgc cagcagccgc ggtaatacgt agggggcaag cgttatccgg	540
atttactggg tgtaaaggga gcgtagacgg aagagcaagt ctgatgtgaa aggctggggc	600
ttaaccccag gactgcattg gaaactgttg ttctagagtg ccggagaggt aagcgggaatt	660
cctagtgtag cggtgaaatg cgtagatatt aggaggaaca ccagtggcga aggcggctta	720
ctggacggta actgacgttg aggtcga aaa gcgtggggag caaacaggat tagataccct	780
ggtagtccac gccgtaaacg atgaatacta ggtgtcgggt ggcaaagcca ttcggtgccg	840
cagcaaagcg aataagtatt ccacctggg agtacgttcg caagaatgaa actcaaagga	900
attgacgggg acccgcaaa gcggtggagc atgtggttta attcgaagca acgcaagaa	960
ccttaccag tcttgacatc cctctgaccg tcccgtaacg ggggcttccc ttcggggcag	1020
aggagacagg tgggtgcatg ttgtcgtcag ctctgtctgt gagatgttg gttaaagtccc	1080
gcaacgagcg caacccttat ccttagtagc cagcacatga tgggtgggcac tctagggaga	1140
ctgccgggga taaccggag gaaggcggg acgacgtcaa atcatcatgc cccttatgat	1200
ttgggctaca cacgtgctac aatggcgtaa acaaaggga gcgagacagc gatgttgagc	1260
gaatccaaa aataacgtcc cagttcggac tgcagtctgc aactcgactg cacgaagctg	1320
gaatcgctag taatcgcgga tcagaatgcc gcggtgaata cgttcccggg tcttgtagac	1380
accgcccgtc acaccatggg agtcagtaac gccgaagtc agtgacctaa ccgaaaggaa	1440
ggagctgccg aaggcgggac cgataactgg ggtgaagtcg taacaaggta acc	1493

<210> 104
 <211> 1262

<212> DNA

<213> *Blautia faecis*

<400> 104

```

ataacagcca gaaatgactg ctaataccgc ataagcgcac agaaccgcat gggttcggtgt      60
gaaaaactcc ggtggtataa gatggacccg cgttggatta gctagttggc agggcagcgg      120
cctaccaagg cgacgatcca tagccggcct gagaggggtga acggccacat tgggactgag      180
acacggccca gactcctacg ggaggcagca gtggggaata ttgcacaatg ggggaaaccc      240
tgatgcagcg acgccgctg aaggaagaag tatctcggtg tgtaaacttc tatcagcagg      300
gaagataatg acggtacctg actaagaagc cccggctaac tacgtgccag cagccgcggt      360
aatacgtagg gggcaagcgt tatccggatt tactgggtgt aaagggagcg tagacggcgc      420
agcaagtctg atgtgaaagg caggggctta acccctggac tgcattggaa actgctgtgc      480
ttgagtgcg gaggggtaag cggaattcct agtgtagcgg tgaaatgcgt agatattagg      540
aggaacacca gtggcgaagg cggcttactg gacggtaact gacgttgagg ctcgaaagcg      600
tggggagcaa acaggattag ataccctggg agtccacgcc gtaaacgatg aatactaggt      660
gtcagggagc acagctcttt ggtgccgccc caaacgcatt aagtattcca cctggggagt      720
acgttcgcaa gaatgaaact caaaggaatt gacggggacc cgcacaagcg gtggagcatg      780
tggtttaatt cgaagcaacg cgaagaacct taccaaactt tgacatccct ctgaccggga      840
cttaaccgtc cttttccttc gggacagggg agacaggtgg tgcatggttg tcgtcagctc      900
gtgtcgtgag atgttgggtt aagtcccgca acgagcgcaa cccctatcct tagtagccag      960
cacgcartgg tgggcactct gaggagactg ccagggataa cctggaggaa ggcggggatg     1020
acgtcaaatc atcatgcccc ttatgatttg ggctacacac gtgctacaat ggcgtaaaca     1080
aaggaagcg aacccgcgag ggtgggcaaa tctcaaaaat aacgtcccag ttcggactgc     1140
agtctgcaac tcgactgcac gaagctggaa tcgctagtaa tcgcggatca gaatgccgcg     1200
gtgaatacgt tcccgggtct tgtacacacc gcccgtcaca ccatgggagt cagtaacgcc     1260
cg

```

<210> 105

<211> 1431

<212> DNA

<213> *Clostridium hathewayi*

<220>

<221> misc_feature

<222> (191)..(191)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (211)..(211)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (248)..(248)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (979)..(979)

<223> n is a, c, g, t or u

<400> 105

```

ctcaggatga acgctggcgg cgtgcttaac acatgcaagt cgagcgaagc ggtttcaatg      60
aagttttcgg atggatttga aattgactta gcggcggacg ggtgagtaac gcgtgggtaa      120
cctgccttac actgggggat aacagttaga aatgactgct aataccgcat aagcgcacag      180
ggccgcatgg nctggtgtga aaaactccgg nggtgtaaga tggaccgcg tctgattagg      240
tagttgngg ggtaacggcc caccaagccg acgatcagta gccgacctga gagggtgacc      300
ggccacattg ggactgagac acggcccaa ctcctacggg aggcagcagt ggggaatatt      360
ggacaatggg cgaaagcctg atccagcgac gccgcgtgag tgaagaagta tttcggtatg      420
taaagctcta tcagcagga agaaaatgac ggtacctgac taagaagccc cggctaacta      480
cgtgccagca gccgcggtaa tacgtagggg gcaagcgtta tccggattta ctgggtgtaa      540
agggagcgta gacggtttag caagtctgaa gtgaaagccc ggggctcaac cccggtactg      600
ctttggaaac tgtagactt gagtgcagga gaggtaagtg gaattcctag tgtagcgggtg      660
aaatgcgtag atattaggag gaacaccagt ggcgaaggcg gcttactgga ctgtaactga      720
cgttgaggct cgaaagcgtg gggagcaaac aggattagat accctggtag tccacgccgt      780
aaacgatgaa tactaggtgt cggggggcaa agcccttcgg tgccgccgca aacgcaataa      840
gtattccacc tggggagtac gttcgaaga atgaaactca aaggaattga cggggacccg      900
cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg aagaacctta ccaagtcttg      960
acatcccact gaaaacacnt taaccgtgat ccctcttcgg agcagtggag acaggtggtg     1020
catggttgtc gtcagctcgt gtcgtgagat gttgggttaa gtcccgcaac gagcgcaacc     1080
cttatcctta gtagccagcg agtagagtcg ggcactctgg ggagactgcc agggataacc     1140
tggaggaagg tggggatgac gtcaaatcat catgcccctt atgatttggg ctacacacgt     1200
gctacaatgg cgtaaacaaa gggaggcaaa ggagcgatct ggagcaaacc caaaaataa     1260
cgtctcagtt cggattgcag gctgcaactc gcctgcatga agctggaatc gctagtaatc     1320
gcgaatcaga atgtcgcggg gaatacgttc ccgggtcttg tacacaccgc ccgtcacacc     1380
atgggagttg gtaacgcccg aagtcagtga cccaaccgaa aggagggagc t              1431

```

<210> 106

<211> 1493

<212> DNA

<213> Blautia producta

<400> 106

```

agagtttgat cctggctcag gatgaacgct ggcggcgtgc ttaacacatg caagtcgagc      60
gaagcactaa gacggatttc ttcggattga agtctttgtg actgagcggc ggacgggtga      120
gtaacgcgtg ggtaacctgc ctcatcacagg gggataacag ttagaaatga ctgctaatac      180

```

P074570006W000-SEQLISTING-CEW.txt

cgcataagcg	cacaggaccg	catggtctgg	tgtgaaaaac	tccggtggta	tgagatggac	240
ccgcgtctga	ttagctagtt	ggaggggtaa	cggcccacca	aggcgacgat	cagtagccgg	300
cctgagaggg	tgaacggcca	cattgggact	gagacacggc	ccagactcct	acgggaggca	360
gcagtgggga	atattgcaca	atgggggaaa	ccctgatgca	gcgacgccgc	gtgaaggaag	420
aagtatctcg	gtatgtaaac	ttctatcagc	agggaagaaa	atgacggtag	ctgactaaga	480
agccccggct	aactacgtgc	cagcagccgc	ggtaatacgt	agggggcaag	cgttatccgg	540
atttactggg	tgtaaaggga	gcgtagacgg	aagagcaagt	ctgatgtgaa	aggctggggc	600
ttaaccccag	gactgcattg	gaaactgttg	ttctagagtg	ccggagaggt	aagcgggaatt	660
cctagtgtag	cggtgaaatg	cgtagatatt	aggaggaaca	ccagtggcga	aggcggctta	720
ctggacggta	actgacgttg	aggctcgaac	gcgtggggag	caaacaggat	tagataccct	780
ggtagtccac	gccgtaaacg	atgaatacta	ggtgtcgggt	ggcaaagcca	ttcgggtgccg	840
cagcaaacgc	aataagtatt	ccacctgggg	agtacgttcg	caagaatgaa	actcaaagga	900
attgacgggg	acccgcacaa	gcggtggagc	atgtggttta	attcgaagca	acgcgaagaa	960
ccttaccaag	tcttgacatc	cctctgaccg	tcccgtaacg	gggacttccc	ttcggggcag	1020
aggagacagg	tgggtgcatg	ttgtcgtcag	ctcgtgtcgt	gagatgttgg	gttaagtccc	1080
gcaacgagcg	caacccttat	ccttagtagc	cagcacatga	tgggtgggcac	tctagggaga	1140
ctgccgggga	taaccgagag	gaaggcgggg	acgacgtcaa	atcatcatgc	cccttatgat	1200
ttgggctaca	cacgtgctac	aatggcgtaa	acaaagggaa	gcgagacagc	gatgttgagc	1260
gaatcccaaa	aataacgtcc	cagttcggac	tgcagtctgc	aactcgactg	cacgaagctg	1320
gaatcgctag	taatcgcgga	tcagaatgcc	gcggtgaata	cgttcccggg	tcttgtagac	1380
accgcccgtc	acaccatggg	agtcagtaac	gcccgaagtc	agtgacctaa	ccgaaaggaa	1440
ggagctgccg	aaggcgggac	cgataactgg	ggtgaagtcg	taacaaggta	acc	1493

<210> 107
 <211> 1515
 <212> DNA
 <213> Anaerostipes hadrus

<400> 107	tggtcagga	tgaacgctgg	cggcgtgctt	aacacatgca	agtcgaacga	agctgcttaa	60
	ctgatcttct	tcggaattga	cgttttgtag	actgagtggc	ggacgggtga	gtaacgcgtg	120
	ggcaacctgc	cctgtacagg	gggataacag	tcagaaatga	ctgctaatac	cgcataagac	180
	cacagcaccg	catggtgcag	gggtaaaaac	tccggtggta	caggatggac	ccgcgtctga	240
	ttagctgggt	ggtgaggtaa	cggctcacca	aggcgacgat	cagtagccgg	cttgagagag	300
	tgaacggcca	cattgggact	gagacacggc	ccaaactcct	acgggaggca	gcagtgggga	360
	atattgcaca	atgggggaaa	ccctgatgca	gcgacgccgc	gtgagtgaag	aagtatctcg	420
	gtatgtaaag	ctctatcagc	agggaagaaa	atgacggtag	ctgactaaga	agccccggct	480

P074570006W000-SEQLISTING-CEW.txt

aactacgtgc cagcagccgc ggtaatacgt aggggggcaag cgttatccgg aattactggg	540
tgtaaagggg gcgtaggtgg tatggcaagt cagaagtga aaccagggc ttaactctgg	600
gactgctttt gaaactgtca gactggagtg caggagaggt aagcggaatt cctagtgtag	660
cggtgaaatg cgtagatatt aggaggaaca tcagtggcga aggcggctta ctggactgaa	720
actgacactg aggcacgaaa gcgtggggag caaacaggat tagataccct ggtagtccac	780
gccgtaaacg atgaatacta ggtgtcgggg ccgtagaggc ttcggtgccg cagccaacgc	840
agtaagtatt ccacctgggg agtacgttcg caagaatga actcaaagga attgacgggg	900
acccgcacaa gcggtggagc atgtggttta attcgaagca acgcgaagaa ccttacctgg	960
tcttgacatc cttctgaccg gtccttaacc ggacctttcc ttcgggacag gagagacagg	1020
tggtgcatgg ttgtcgtcag ctcgtgtcgt gagatgttgg gttaagtccc gcaacgagcg	1080
caacccttat ctttagtagc cagcatttca ggtgggcact ctagagagac tgccagggat	1140
aacctggagg aaggtgggga cgacgtcaaa tcatcatgcc ccttatgacc agggctacac	1200
acgtgctaca atggcgtaaa cagaggaag cagcctcgtg agagtgagca aatcccaaaa	1260
ataacgtctc agttcggatt gtagtctgca actcgactac atgaagctgg aatcgctagt	1320
aatcgcgaat cagaatgtcg cggtgaatac gttcccgggt cttgtacaca ccgcccgtca	1380
caccatggga gtcagtaacg cccgaagtca gtgacccaac cgtaaggagg gagctgccga	1440
aggcgggacc gataactggg gtgaagtcgt aacaaggtag ccgtatcgga aggtgcggct	1500
ggatcacctc ctttc	1515

<210> 108
 <211> 1523
 <212> DNA
 <213> Eubacterium fissicatena

<400> 108	
gtttgatcct ggctcaggat gaacgctggc ggcgtgctta acacatgcaa gtcgagcgaa	60
gcgctttact tagatttctt cggattgaag agttttgcga ctgagcggcg gacgggtgag	120
taacgcgtgg gtaacctgcc tcatacaggg ggataacagt tagaaatgac tgctaatacc	180
gcataagacc acagtaccgc atggtacagt gggaaaaact ccggtgggtat gagatggacc	240
cgcgtctgat tagctagttg gtaaggtaac ggcttaccaa ggcaacgatc agtagccgac	300
ctgagagggg gaccggccac attgggactg agacacggcc caaactccta cgggaggcag	360
cagtggggaa tattgcacaa tgggggaaac cctgatgcag cgacgccgcg tgaaggatga	420
agtatttcgg tatgtaaact tctatcagca gggaagaaaa tgacggtacc tgactaagaa	480
gccccggcta actacgtgcc agcagccgcg gtaatacgta gggggcaagc gttatccgga	540
tttactgggt gtaaagggag cgtagacggg tatgtaaagtc tgatgtgaaa acccggggct	600
caacccggg actgcattgg aaactatgta actagagtgt cggagaggta agtggaattc	660
ctagtgtagc ggtgaaatgc gtagatatta ggaggaacac cagtggcgaa ggccggcttac	720
tggacgatca ctgacgttga ggctcgaaag cgtggggagc aaacaggatt agataccctg	780

P074570006W000-SEQLISTING-CEW.txt

gtagtccacg	ccgtaaacga	tgaatactag	gtgtcgggtg	gcaaagccat	tcggtgccgc	840
agcaaacgca	ataagtattc	cacctgggga	gtacgttcgc	aagaatgaaa	ctcaaaggaa	900
ttgacgggga	cccgcacaa	cggtggagca	tgtggtttaa	ttcgaagcaa	cgcaagaac	960
cttacctgct	cttgacatcc	cactgaccgg	cgtgtaatgg	cgctttccct	tcggggcagt	1020
ggagacaggt	ggtgcatggt	tgtcgtcagc	tcgtgtcgtg	agatgttggg	ttaagtcccc	1080
caacgagcgc	aacccttatt	tttagtagcc	agcggtttgg	ccgggcactc	tagagagact	1140
gccagggata	acctggagga	aggtggggat	gacgtcaaat	catcatgccc	cttatgagca	1200
gggctacaca	cgtgctacaa	tggcgtaaac	aaagggaggc	aataccgcga	ggttgagcaa	1260
atcccaaaaa	taacgtctca	gttcggattg	tagtctgcaa	ctcgactaca	tgaagctgga	1320
atcgctagta	atcggaatc	agaatgtcgc	ggtgaatacg	ttcccgggtc	ttgtacacac	1380
cgcccgtcac	accatgggag	ttggtaacgc	ccgaagtcag	tgaccaacc	gtaaggaggg	1440
agctgccgaa	ggcgggatcg	ataactgggg	tgaagtcgta	acaaggtagc	cgtatcggaa	1500
ggtgcggctg	gacacctcc	ttt				1523

<210> 109
 <211> 1524
 <212> DNA
 <213> Eubacterium contortum

<400> 109	
tttgatcctg	gctcaggatg aacgctggcg acgtgcttaa cacatgcaag tcgagcgaag 60
cactttactt	tgattttctt ggaatgaaag gttttgtgac tgagcggcgg acgggtgagt 120
aacgcgtggg	taacctgcct catacagggg gataacagtt agaaatgact gctaataccg 180
cataagacca	cagtaccgca tggtagctg ggaaaaactc cgggtggtatg agatggaccc 240
gcgtctgatt	agctagttag taaggtaacg gcttaccaag gcgacgatca gtagccgacc 300
tgagaggggtg	accggccaca ttgggactga gacacggccc aaactcctac gggaggcagc 360
agtggggaat	attgcacaat gggggaaacc ctgatgcagc gacgccgcgt gaaggatgaa 420
gtatttcggt	atgtaaactt ctatcagcag ggaagaaaat gacgggtacct gactaagaag 480
ccccgggctaa	ctacgtgccca gcagccgcgg taatacgtag ggggcaagcg ttatccggat 540
ttactgggtg	taaagggagc gtagacggtt atgtaagtct gatgtgaaaa cccggggctc 600
aaccccgga	ctgcattgga aactatgtaa ctagagtgtc ggagaggtaa gtggaattcc 660
tagttagcg	gtgaaatgcg tagatattag gaggaacacc agtggcgaag gcggcttact 720
ggacgatgac	tgacgttgag gctcgaaagc gtggggagca aacaggatta gataccctgg 780
tagtccacgc	cgtaaacgat gaatactagg tgtcgggtgg caaagccatt cggtgccgca 840
gcaaacgcaa	taagtattcc acctggggag tacgttcgca agaataaaac tcaaaggaat 900
tgacggggac	ccgcacaagc ggtggagcat gtggtttaat tcgaagcaac gcgaagaacc 960
ttacctgctc	ttgacatccc cctgaccggc gtgtaatggt gcctttcctt cgggacaggg 1020

P074570006W000-SEQLISTING-CEW.txt

gagacaggtg	gtgcatgggt	gtcgtcagct	cgtgtcgtga	gatgttgggt	taagtcccg	1080
aacgagcgca	acccttatct	ttagtagcca	gcggtttggc	cgggcactct	agagagactg	1140
ccagggataa	cctggaggaa	ggtggggatg	acgtcaaate	atcatgcccc	ttatgagcag	1200
ggctacacac	gtgctacaat	ggcgtaaaca	aagggaggcg	aagccgtgag	gtggagcaaa	1260
tccccaaaaa	aacgtctcag	ttcggattgt	agtctgcaac	tcgactacat	gaagctggaa	1320
tcgctagtaa	tcgcgaatca	gaatgtcgcg	gtgaatacgt	tcccgggtct	tgtacacacc	1380
gcccgtcaca	ccatgggagt	tggtaacgcc	cgaagtcagt	gacccaaccg	caaggaggga	1440
gctgccgagg	gtgggaccga	taactggggg	gaagtcgtaa	caaggtagcc	gtatcggaag	1500
gtgcggctgg	atcacctcct	ttct				1524

<210> 110
 <211> 1390
 <212> DNA
 <213> Clostridium bolteae

<400> 110						
ttttaattga	ctgagtggcg gacgggtgag taacgcgtgg ataacctgcc tcacactggg 60					
ggataacagt	tagaaatgac tgctaataacc gcataagcgc acagtaccgc atggtacagt 120					
gtgaaaaact	ccggtgggtgt gagatggatc cgcgtctgat tagccagttg gcggggtaac 180					
ggcccaccaa	agcgacgatc agtagccgac ctgagagggg gaccggccac attgggactg 240					
agacacggcc	caaactccta cgggaggcag cagtggggaa tattgcacaa tgggcgaaag 300					
cctgatgcag	cgacgcccg	tgagtgaaga	agtatttcgg	tatgtaaagc	tctatcagca	360
gggaagaaaa	tgacggtacc	tgactaagaa	gccccggcta	actacgtgcc	agcagcccg	420
gtaatacgt	gggggcaagc	gttatccgga	tttactgggt	gtaaaggagg	cgtagacggc	480
gaagcaagtc	tgaagtga	accagggt	caaccctggg	actgctttgg	aaactgtttt	540
gctagagtgt	cggagaggta	agtgggaattc	ctagtgtagc	ggtgaaatgc	gtagatatta	600
ggaggaacac	cagtggcgaa	ggcggcttac	tgacgataa	ctgacgttga	ggctcgaaag	660
cgtggggagc	aaacaggatt	agataccctg	gtagtccacg	ccgtaaacga	tgaatgctag	720
gtgttggggg	gcaaagccct	tcggtgccgt	cgcaaacgca	gtaagcattc	cacctgggga	780
gtacgttcgc	aagaatgaaa	ctcaaaggaa	ttgacgggga	ccgcacaaag	cggtggagca	840
tgtggtttaa	ttcgaagcaa	cgcgaagaac	cttaccaagt	cttgacatcc	tcttgaccgg	900
cgtgtaacgg	cgccttcctt	tcggggcaag	agagacaggt	ggtgcatggg	tgtcgtcagc	960
tcgtgtcgtg	agatgttggg	ttaagtcccg	caacgagcgc	aacccttatc	cttagtagcc	1020
agcaggtaaa	gctgggcact	ctaggagac	tgccagggat	aacctggagg	aaggtgggga	1080
tgacgtcaaa	tcatcatgcc	ccttatgatt	tggtctacac	acgtgctaca	atggcgtaaa	1140
caaagggaag	caagacagt	atgtggagca	aatcccaaaa	ataacgtccc	agttcggact	1200
gtagtctgca	acccgactac	acgaagctgg	aatcgctagt	aatcgcgaa	cagaatgtcg	1260
cgggtgaatac	gttcccgggt	cttgtacaca	ccgccgtca	caccatggga	gtcagcaacg	1320

cccgaagtca gtgaccaac tcgcaagaga gggagctgcc gaaggcgggg caggtaactg 1380
gggtgaagtc 1390

<210> 111
<211> 1308
<212> DNA
<213> *Blautia luti*

<220>
<221> misc_feature
<222> (702)..(705)
<223> n is a, c, g, t or u

<400> 111
gtgggtaacc tgccttatac agggggataa cagtcagaaa tgactgctaa taccgcataa 60
gcgcacagag ctgcatggct ccggtgtgaa aaactccggt ggtataagat ggacccgcgt 120
tggattagct agttggtgag gtaacggccc accaaggcga cgatccatag ccggcctgag 180
agggtgaacg gccacattgg gactgagaca cggcccagac tcctacggga ggcagcagtg 240
gggaatattg cacaatgggg gaaaccctga tgcagcgacg ccgctgtaag gaagaagtat 300
ctcggtatgt aaacttctat cagcaggga gaaaatgacg gtacctgact aagaagcccc 360
ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat ccggatttac 420
tgggtgtaaa gggagcgtag acggcatgga caagtctgat gtgaaaggct ggggctcaac 480
cccgggactg cattggaac tgcccgtctt gagtgccgga gaggtaagcg gaattcctag 540
tgtagcggtg aaatgcgtag atattaggag gaacaccagt ggcgaaggcg gcttactgga 600
cggtaaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat accctggtag 660
tccacgcggt aaacgatgaa tcctaggtgt cggggagcaa annnnttcgg tgccgccgca 720
aacgcattaa gcattccacc tggggagtag gttcgcaaga atgaaactca aaggaattga 780
cggggacccg cacaagcgtt ggagcatgtg gtttaattcg aagcaacgag aagaacctta 840
ccaagtcttg acatccctct gaccgagtat gtatggtact tttccttcgg gagagagagg 900
agacagggtg tgcatggttg tcgtcagctc gtgtcgtgag atgttgggtt aagtcccgca 960
acgagcgcaa cccctatccc cagtagccag cggttcggcc gggcactctg aggagactgc 1020
cagggataac ctggaggaag gcggggatga cgtcaaatca tcatgcccct tatgatttgg 1080
gctacacacg tgctacaatg gcgtaaacaa agggaagcaa gcctgagagg gtgggcaaat 1140
cccaaaaata acgtcccagt tcggactgta gtctgcaacc cgactacacg aagctggaat 1200
cgctagtaat cgcggtacag aatgccgcgg tgaatacggt cccgggtctt gtacacaccg 1260
cccgtcacac catgggagtc agtaacgcc gaagtcagtg acctaact 1308

<210> 112
<211> 1489
<212> DNA
<213> *Acidaminococcus intestini*

P074570006W000-SEQLISTING-CEW.txt

```

<400> 112
ctggcggcgt gcttaacaca tgcaagtcga acggagaact tatttcggta agttcttagt      60
ggcgaacggg tgagtaacgc gtgggcaacc tgccctccag ttggggacaa cattccgaaa      120
gggatgctaa taccgaatgt cctccctcct ccgcatggag gagggaggaa agatggcctc      180
tgcttgcaag ctatcgctgg aagatgggcc cgcgtctgat tagctagttg gtggggtaac      240
ggctcaccaa ggcatgatc agtagccggg ctgagaggat gaacggccac attgggactg      300
agacacggcc caaactccta cgggaggcag cagtggggaa tcttccgcaa tggacgaaag      360
tctgacggag caacgccgag tgagtgatga aggtcttcgg attgtaaaac tctgttggtta      420
gggacgaaag caccgtgttc gaacaggtca tgggtgttgac ggtacctaac gaggaagcca      480
cggctaacta cgtgccagca gccgcggtaa tacgtaggtg gcaagcgttg tccggaatta      540
ttgggcgtaa agagcatgta ggccggcctt taagtctgac gtgaaaatgc ggggcttaac      600
cccgtatggc gttggatact ggaagtcttg agtgcaggag aggaaagggg aattcccagt      660
gtagcgggtg aatgcgtaga tattgggagg aacaccagtg gcgaaggcgc ctttctggac      720
tgtgtctgac gctgagatgc gaaagccagg gtagcaaacg ggattagata ccccggtagt      780
cctggccgta aacgatggat actaggtgta ggaggtatcg accccttctg tgccggagtt      840
aacgcaataa gtatcccgc tggggactac gatcgcaaga ttgaaactca aaggaattga      900
cggggggccc cacaagcggg ggagtatgtg gtttaattcg acgcaacgcg aagaacctta      960
ccaaggcttg acattgagtg aaagacctag agataggtcc ctcccttcgg ggacacgaaa     1020
acaggtggtg catggctgtc gtcagctcgt gtcgtgagat gttgggttaa gtcccgaac     1080
gagcgcaacc cctatcctat gttaccagcg cgtaaaggcg gggactcata ggagactgcc     1140
agggataact tggaggaagg cggggatgac gtcaagtcac catgcccctt atgtcttggg     1200
ctacacacgt actacaatgg tcggcaacaa agggcagcga aaccgagagg tggagcaaat     1260
cccagaaacc cgaccccagt tcggatcgta ggctgcaacc cgcctacgtg aagttggaat     1320
cgctagtaat cgaggtcag catactgcgg tgaatacgtt cccgggcctt gtacacaccg     1380
cccgtcacac cacgaaagtt ggtaacacc gaagccggtg agataacctt ttaggagtca     1440
gctgtctaag gtggggccga tgattggggt gaagtcgtaa caaggtagc      1489

```

```

<210> 113
<211> 1500
<212> DNA
<213> Ruminococcus albus

```

```

<400> 113
agagtttgat cctggctcag gacgaacgct ggccggcacgc ttaacacatg caagtcgaac      60
gagcgaaaga gtgcttgac tctctagcta gtggcggacg ggtgagtaac acgtgagcaa      120
tctgcctttc ggagagggat accaattgga aacgattgtt aatacctcat aacataacga      180
agccgcatga ctttggtatc aaatgaattt cgccgaaaga tgagctcgcg tctgattagg      240
tagttggtga ggtaacggcc caccaagccg acgatcagta gccggactga gaggttgaac      300

```

P074570006W000-SEQLISTING-CEW.txt

ggccacattg	ggactgagac	acggcccaga	ctcctacggg	aggcagcagt	ggggaatatt	360
gcacaatggg	cgaagcctg	atgcagcgat	gccgcgtgag	ggaagaaggt	tttaggattg	420
taaacctctg	tctttgggga	cgataatgac	ggtacccaag	gaggaagctc	cggctaacta	480
cgtgccagca	gccgcggtaa	tacgtaggga	gcgagcgttg	tccggaatta	ctgggtgtaa	540
agggagcgtg	ggcgggattg	caagtcaggt	gtgaaattta	ggggcttaac	ccctgaactg	600
cacttgaaac	tgtagttctt	gagtgaagta	gaggtgaagc	gaatttcctag	tgtagcgggtg	660
aaatgcgtag	atattaggag	gaacatcagt	ggcgaaggcg	gcttactggg	ctttaactga	720
cgctgaggct	cgaagcgtg	gggagcaaac	aggattagat	accctggtag	tccacgccgt	780
aaacgatgat	tactaggtgt	ggggggactg	acccttccg	tgccgcagtt	aacacaataa	840
gtaatccacc	tggggagtac	ggccgcaagg	ctgaaactca	aaggaattga	cggggacccg	900
cacaagcagt	ggagtatgtg	gtttaattcg	aagcaacgcg	aagaacctta	ccaggtcttg	960
acatcgtacg	catagcatag	agatatgtga	aatcccttcg	gggacgtata	gacaggtggt	1020
gcatggttgt	cgtcagctcg	tgtcgtgaga	tgttgggggt	aagtcccgca	acgagcgcaa	1080
cccttactgt	tagttgctac	gcaagagcac	tctagcagga	ctgccgttga	caaaacggag	1140
gaagtggtgg	atgacgtcaa	atcatcatgc	cccttatgac	ctgggctaca	cacgtactac	1200
aatggctgtt	aacagaggga	agcaaaacag	tgatgtggag	caaaacccta	aaagcagtct	1260
tagttcggat	tgtaggctgc	aaccgccta	catgaagtcg	gaattgctag	taatcgcgga	1320
tcagcatgcc	gcggtgaata	cgttcccggg	ccttgtagac	accgcccgtc	acgccatggg	1380
agtcggtaac	accgaagcc	tgtgttctaa	ccgcaaggag	gaagcagtcg	aaggtgggat	1440
tgatgactgg	ggtgaagtcg	taacaaggta	gccgtatcgg	aaggtgcggc	tggatcacct	1500

<210> 114

<211> 1521

<212> DNA

<213> Eubacterium rectale

<400> 114

agagtttgat	cctggctcag	gatgaacgct	ggcggcgtgc	ttaacacatg	caagtcgaac	60
gaagcacttt	atttgatttc	cttcgggact	gattattttg	tgactgagtg	gcggacgggt	120
gagtaacgcg	tgggtaacct	gccttgtaga	gggggataac	agttggaaac	ggctgctaata	180
accgcataag	cgcacggcat	cgcatgatgc	agtgtgaaaa	actccggtgg	tataagatgg	240
acccgcgttg	gattagctag	ttggtgaggt	aacggcccac	caaggcgacg	atccatagcc	300
gacctgagag	ggtgaccggc	cacattggga	ctgagacacg	gcccaaactc	ctacgggagg	360
cagcagtggg	gaatattgca	caatgggcga	aagcctgatg	cagcgacgcc	gcgtgagcga	420
agaagtattt	cggtatgtaa	agctctatca	gcagggaaga	taatgacggg	acctgactaa	480
gaagcaccgg	ctaaatacgt	gccagcagcc	gcggtaatac	gtatggtgca	agcgttatcc	540
ggatttactg	ggtgtaaagg	gagcgcaggc	ggtgcggcaa	gtctgatgtg	aaagcccggg	600
gctcaacccc	ggtactgcat	tggaaactgt	cgtagtagag	tgtcggaggg	gtaagcggaa	660

P074570006W000-SEQLISTING-CEW.txt

ttcctagtgt agcgggtgaaa tgcgtagata ttaggaggaa caccagtggc gaaggcggct	720
tactggacga taactgacgc tgaggctcga aagcgtgggg agcaaacagg attagatacc	780
ctggtagtcc acgccgtaaa cgatgaatac taggtgttgg gaagcattgc ttctcgggtgc	840
cgtcgcaaac gcagtaagta ttccacctgg ggagtacgtt cgcaagaatg aaactcaaag	900
gaattgacgg ggacccgcac aagcgggtgga gcatgtggtt taattcgaag caacgcgaag	960
aaccttacca agtcttgaca tccttctgac cggtaactta ccgtaccttc tcttcggagc	1020
aggagtgaca ggtgggtgcat ggttgctcgc agctcgtgtc gtgagatgtt gggttaagtc	1080
ccgcaacgag cgcaaccctt atcttttagta gccagcgggt cggccgggca ctctagagag	1140
actgccaggg ataacctgga ggaaggcggg gatgacgtca aatcatcatg ccccttatga	1200
cttgggctac acacgtgcta caatggcgta aacaaaggga agcaaagctg tgaagccgag	1260
caaactctcaa aaataacgtc tcagttcgga ctgtagtctg caacccgact acacgaagct	1320
ggaatcgcta gtaatcgag atcagaatgc tgcgggtgaat acgttcccgg gtcttgtaca	1380
caccgcccgt cacaccatgg gagttgggaa tgcccgaagc cagtgacctt accgaaagga	1440
aggagctgtc gaaggcaggc tcgataactg ggggtgaagtc gtaacaaggt agccgtatcg	1500
gaagggtcgg ctggatcacc t	1521

<210> 115
 <211> 1545
 <212> DNA
 <213> Acidaminococcus fermentans

<400> 115	
agagtttgat cctgggtcag gacgaacgct ggccggcgtgc ttaacacatg caagtcgaac	60
ggagaacttt cttcggaatg ttcttagtgg cgaacgggtg agtaacgcgt aggcaacctg	120
ccctctggtt ggggacaaca ttccgaaagg gatgctaata ccgaatgaga tcctctttcc	180
gcatggagag aggatgaaag atggcctcta cttgtaagct atcgccagaa gatgggcctg	240
cgtctgatta gctagtaggt gaggtaacgg ctcacctagg cgatgatcag tagccggtct	300
gagaggatga acggccacat tgggactgag acacggccca aactcctacg ggaggcagca	360
gtggggaatc ttccgcaatg gacgaaagtc tgacggagca acgccgcgtg agtgatgaag	420
gccttcgggt tgtaaaactc tgttgtcagg gacgaaagca ccgatctata atacattttg	480
gtgttgacgg tacctgacga ggaagccacg gctaactacg tgccagcagc cgcggttaata	540
cgtaggtggc aagcgttgct cggaattatt gggcgtaaag agcatgtagg cgggctttta	600
agtccgacgt gaaaatgcgg ggcttaacct cgtatggcgt tggatactgg aagtcttgag	660
tgcaggagag gaaaggggaa ttcccagtgt agcggtgaaa tgcgtagata ttgggaggaa	720
caccagtggc gaaggcgcct ttctggactg tgtctgacgc tgagatgcga aagccagggg	780
agcaaacggg attagatacc ccggtagtcc tggccgtaaa cgatgggtac taggtgtagg	840
aggtatcgac cccttctgtg ccggagttaa cgcaataagt accccgcctg gggactacga	900

P074570006W000-SEQLISTING-CEW.txt

tcgcaagatt gaaactcaaa ggaattgacg ggggcccgca caagcgggtgg agtatgtggt	960
ttaattcgac gcaacgcgaa gaaccttacc aaggcttgac attgagtgaag agaccagag	1020
atgggtcccc ttcttcggaa gcacgaaaac aggtggtgca tggctgtcgt cagctcgtgt	1080
cgtgagatgt tgggttaagt cccgcaacga gcgcaaccct taccctatgt taccagcacg	1140
taatggtggg gactcatagg agactgccag ggataacctg gaggaaggcg gggatgacgt	1200
caagtcatca tgccccttat gtcttgggct acacacgtac tacaatggtc ggcaacaaag	1260
ggcagcgaag ccgcgaggcg gagccaatcc cagaaacccg accccagttc ggatcgcagg	1320
ctgcaacccg cctgcgtgaa gttggaatcg ctagtaatcg caggtcagca tactgcggtg	1380
aatacgttcc cgggccttgt acacaccgcc cgtcacacca cgaaagttag taacaccgca	1440
agccggtgag ataacctttt aggagtcagc tgtctaaggt ggggccgatg attgggggtga	1500
agtcgtaaca aggtagccgt tcgagaacga gcggctggat cacct	1545

<210> 116
 <211> 1423
 <212> DNA
 <213> Fusicatenibacter saccharivorans

<400> 116	
tggctcagga tgaacgctgg cggcgtgctt aacacatgca agtcgagcga agcagttaag	60
aagattytic ggatgattct tgactgactg agcggcggac gggtagagtaa cgcgtgggtg	120
acctgccccca taccggggga taacagctgg aaacggctgc taataccgca taagcgcaca	180
gagctgcatg gctcgggtgtg aaaaactccg gtggtatggg atgggcccgc gtctgattag	240
gcagttggcg gggtaacggc ccaccaaacc gacgatcagt agccggcctg agagggcgac	300
cggccacatt gggactgaga cacggcccaa actcctacgg gaggcagcag tggggaatat	360
tgcacaatgg gggaaaccct gatgcagcga cgccgcgtga gcgaagaagt atttcggtat	420
gtaaagctct atcagcaggg aagataatga cggtagctga ctaagaagcc ccggctaact	480
acgtgccagc agccgcggta atacgtaggg ggcaagcgtt atccggattt actgggtgta	540
aagggagcgt agacggcaag gcaagtctga tgtgaaaacc cagggcttaa ccctgggact	600
gcattggaaa ctgtctggct cgagtgccgg agaggtaagc ggaattccta gtgtagcggg	660
gaaatgcgta gatattagga agaaccacag tggcgaaggc ggcttactgg acggtaactg	720
acgttgaggc tcgaaagcgt ggggagcaaa caggattaga taccctggta gtccacgccg	780
taaacgatga atgctagggt ttggggagca aagctcttcg gtgccgccgc aaacgcatta	840
agcattccac ctggggagta cgttcgcaag aatgaaactc aaaggaattg acggggaccc	900
gcacaagcgg tggagcatgt ggtttaattc gaagcaacgc gaagaacctt accaggtctt	960
gacatcccga tgaccggccc gtaacggggc cttctcttcg gagcattgga gacaggtggt	1020
gcatggttgt cgtcagctcg tgtcgtgaga tgttgggtta agtcccgcaa cgagcgcaac	1080
ccttatcctc agtagccagc aggtaaagct gggcactctg tggagactgc cagggataac	1140
ctggaggaag gtgggggatga cgtcaaatca tcatgccctt tatgatctgg gctacacacg	1200

P074570006W000-SEQLISTING-CEW.txt

tgctacaatg gcgtaaacia agggaggcaa agccgagagg tggagcaa	at cccaaaaata	1260
acgtctcagt tcggactgca gtctgcaact cgactgcacg aagctggaat	cgctagtaat	1320
cgcgaaatcag aatgtcgcgg tgaatacgtt cccgggtctt gtacacaccg	cccgtcacac	1380
catgggagtt ggtaacgccc gaagtcagtg acccaacctt tta		1423

<210> 117
 <211> 1494
 <212> DNA
 <213> Ruminococcus champanellensis

<400> 117		
agagtttgat cctggctcag gacgaacgct ggcggcacgc ctaacacatg	caagtcgaac	60
ggagataaag acttcggttt ttatcttagt ggcggacggg tgagtaacac	gtgagcaacc	120
tgcctctgag agagggatag cttctggaaa cggatggtaa tacctcataa	catagcggta	180
ccgcatgata ctgctatcaa agatttatcg ctacagagatg ggctcgcgtc	tgattagcta	240
gatggtgagg taacggctca ccatggcgac gatcagtagc cggactgaga	ggttgaacgg	300
ccacattggg actgagacac ggcccagact cctacgggag gcagcagtg	gggaatattgc	360
acaatgggcg caagcctgat gcagcgatgc cgcgtggagg aagaaggttt	tcggattgta	420
aactcctgtc ttaagggacg ataatgacgg taccttagga ggaagctccg	gctaactacg	480
tgccagcagc cgcggtaata cgtagggagc gagcgttgtc cgggaattact	gggtgtaaag	540
ggagcgtagg cgggattgca agtcagatgt gaaaactatg ggcttaaccc	atagactgca	600
tttgaaactg tagttcttga gtgaagtaga ggtaagcggg attcctagt	gtagcggtgaa	660
atgcgtagat attaggagga acatcggttg cgaaggcggc ttactgggct	tttactgacg	720
ctgagggtcg aaagcgtggg gagcaaacag gattagatac cctggtagtc	cacgctgtaa	780
acgatgatta ctaggtgtgg ggggactgac cccttccgtg ccgcagttaa	cacaataagt	840
aatccacctg gggagtacgg ccgcaagggt gaaactcaaa ggaattgacg	ggggcccgcg	900
caagcagtg agtatgtggt ttaattcgaa gcaacgcgaa aaaccttacc	aggtcttgac	960
atcgagtga tgatctagag atagatcagt ccttcgggac acaaagacag	gtggtgcatg	1020
gttgctgctca gctcgtgtcg tgagatgttg ggttaagtcc cgcaacgagc	gcaaccctta	1080
cctttagttg ctacgcaaga gcactctaga gggactgccg ttgacaaaac	ggaggaaggt	1140
ggggatgacg tcaaatacat atgcccctta tgacctgggc tacacacgta	ctacaatggc	1200
aatgaacaga gggaagcaat acagtgatgt ggagcaaata cccaaaaatt	gtcccagttc	1260
agattgtagg ctgcaactcg cctacatgaa gtcggaattg ctagtaatcg	cagatcagca	1320
tgctgcggtg aatacgttcc cgggccttgt acacaccgcc cgtcacacca	tgggagtcgg	1380
taacaccgga agccagtagc ctaaccgcaa ggagggcgct gtcgaagg	tgaggattga	1440
ctgggggtgaa gtcgtaacaa ggtagccgta tcggaagggt cggctggatc	acct	1494

<210> 118

<211> 1533

<212> DNA

<213> Bifidobacterium bifidum

<400> 118

```

tttttgtgga gggttcgatt ctggctcagg atgaacgctg gcggcgtgct taacacatgc      60
aagtcgaacg ggatccatcg ggctttgctt ggtggtgaga gtggcgaacg ggtgagtaat      120
gcgtgaccga cctgccccat gctccggaat agctcctgga aacgggtggt aatgccggat      180
gttccacatg atcgcatgtg attgtgggaa agattctatc ggcgtaggat ggggtcgcgt      240
cctatcagct tgttggtgag gtaacggctc accaaggctt cgacgggtag ccggcctgag      300
agggcgaccg gccacattgg gactgagata cggcccagac tcctacggga ggcagcagtg      360
gggaatattg cacaatgggc gcaagcctga tgcagcgacg ccgcgtgagg gatggaggcc      420
ttcgggttgt aaacctcttt tgtttgggag caagccttcg ggtgagtgtg cctttcgaat      480
aagcgccggc taactacgtg ccagcagccg cggtaatagc tagggcgcaa gcgttatccg      540
gatttattgg gcgtaaaggg ctcgtaggcg gctcgtcgcg tccggtgtga aagtccatcg      600
cttaacggtg gatctgcgcc gggtaggggc gggctggagt gcggtagggg agactggaat      660
tcccggtgta acggtggaat gtgtagatat cgggaagaac accgatggcg aaggcaggctc      720
tctgggccgt cactgacgct gaggagcgaa agcgtgggga gcgaacagga ttagataccc      780
tggtagtcca cgccgtaaac ggtggacgct ggatgtgggg cacgttcac gtgttccgtg      840
tcggagctaa cgcgttaagc gtccgcctg gggagtacgg ccgcaaggct aaaactcaaa      900
gaaattgacg ggggcccgcg caagcggcgg agcatgcgga ttaattcgat gcaacgcgaa      960
gaaccttacc tgggcttgac atgttcccga cgacgccaga gatggcgttt cccttcgggg      1020
cgggttcaca ggtggtgcat ggtcgtcgtc agctcgtgtc gtgagatgtt gggttaagtc      1080
ccgcaacgag cgcaaccctc gccccgtgtt gccagcacgt tatggtggga actcacgggg      1140
gaccgccggg gttaactcgg aggaagggtg ggatgacgtc agatcatcat gcccttacg      1200
tccagggtt cacgcatgct acaatggccg gtacagcggg atgcgacatg gcgacatgga      1260
gcggatccct gaaaaccggt ctcatgtcgg atcggagcct gcaaccggc tccgtgaagg      1320
cggagtcgct agtaatcgcg gatcagcaac gccgcggtga atgcgttccc gggccttgta      1380
cacaccgccc gtcaagtcag gaaagtgggc agcaccgaa gccggtggcc taacccttg      1440
tgggatggag ccgtctaagg tgaggctcgt gattgggact aagtcgtaac aaggtagccg      1500
taccggaagg tgcggctgga tcacctcctt tct      1533

```

<210> 119

<211> 1552

<212> DNA

<213> Megasphaera elsdenii

<400> 119

```

agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgaac      60
gagaagagat gagaagcttg cttcttatca attcgagtgg caaacgggtg agtaacgcgt      120

```



```

aagcaacctg cccttcagat ggggacaaca gctggaaacg gctgctaata ccgaatacgt      180
tctttttgtc gcatggcaga gggaagaaag ggaggctctt cggagctttc gctgaaggag      240
gggcttgctg ctgattagct agttggaggg gtaacggccc accaaggcga cgatcagtag      300
ccggtctgag aggatgaacg gccacattgg gactgagaca cggcccagac tcctacggga      360
ggcagcagtg gggaatcttc cgcaatggac gaaagtctga cggagcaacg ccgctgaac      420
gatgacggcc ttcgggttgt aaagttctgt tatacgggac gaatggcgta gcggtcaata      480
cccgttacga gtgacggtac cgtaagagaa agccacggct aactacgtgc cagcagccgc      540
ggtaatacgt aggtggcaag cgttgtccgg aattattggg cgtaaagggc gcgcaggcgg      600
cgctgtaagt cgggtcttaa agtgcggggc ttaaccccg gaggggaccg aaactgcat      660
gctagagtat cggagaggaa agcgggaattc ctagtgtagc ggtgaaatgc gtagatatta      720
ggaggaacac cagtggcgaa agcggctttc tggacgacaa ctgacgctga ggcgcgaaag      780
ccaggggagc aaacgggatt agatacccg gtagtcctgg ccgtaaacga tggatactag      840
gtgtaggagg tatcgacccc ttctgtgccg gagttaaagc aataagtatc ccgcctgggg      900
agtacggccg caaggctgaa actcaaagga attgacgggg gcccgcacaa gcggtggagt      960
atgtggttta attcgacgca acgcaagaa ccttaccagg ccttgacatt gattgctatg     1020
gatagagata tccagttcct cttcgaggaa caagaaaaca ggtggtgcac ggctgtcgtc     1080
agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt atcttctgtt     1140
accagcgggt cggccgggga ctcaggagag actgccgag acaatgcgga ggaaggcggg     1200
gatgacgtca agtcatcatg ccccttatgg cttgggctac acacgtacta caatggctct     1260
taatagaggg aagcgaagga gcgatccgga gcaaacccca aaaacagagt ccagttcggg     1320
attgcaggct gcaactcgcc tgcataagc aggaatcgct agtaatcgca ggtcagcata     1380
ctgcggtgaa tacgttcccg ggccttgtag acaccgccg tcacaccag aaagtcattc     1440
acacccgaag ccggtgaggt aaccttttgg agccagccgt cgaagggtggg ggcgatgatt     1500
ggggtgaagt cgtaacaagg tagccgtatc ggaagggtgc gctggatcac ct             1552

```

```

<210> 120
<211> 1479
<212> DNA
<213> Dorea formicigenerans

```

```

<220>
<221> misc_feature
<222> (123)..(125)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (220)..(222)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (287)..(287)

```

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (349)..(350)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (517)..(517)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (534)..(535)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (725)..(725)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (918)..(918)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (921)..(921)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (955)..(956)

<223> n is a, c, g, t or u

<220>

<221> misc_feature

<222> (1184)..(1185)

<223> n is a, c, g, t or u

<400> 120

ttaaacgaga gtttgatcct ggctcaggat gaacgctggc ggcgtgctta acacatgcaa 60

gtcgagcgaa gcacataagt ttgattcttc ggatgaagac ttttgtgact gagcggcgga 120

cgnnngagta acgctgggt aacctgcctc atacaggggg ataacagyta gaaatggctg 180

ctaataccgc ataagaccac agtactgcat ggtacagtgn nnaaaactcc ggtggtatga 240

gatggacccg cgtctgatta ggtagttggt gaggtaacgg ccaccnagc cgacgatcag 300

tagccgacct gagagggtga ccggccacat tgggactgag acacggccnn gactcctacg 360

ggaggcagca gtggggaata ttgcacaatg ggcgaaagcc tgatgcagcg acgccgcgtg 420

aaggatgaag tatttcggta tgtaaaacttc tatcagcagg gaagaaaatg acggtacctg 480

actaagaagc cccggctaac tacgtgccag cagccnggt aatacgtagg gggnnagcgt 540

tatccggatt tactgggtgt aaaggagcg tagacggctg tgcaagtctg aagtgaagg 600

catgggctca acctgtggac tgctttggaa actgtgcagc tagagtgtcg gagaggtaag 660

tggaattcct agtgtagcgg tgaaatgcgt agatattagg aggaacacca gtggcgaagg 720

cggcntactg gacgatgact gacgttgagg ctcgaaagcg tggggagcaa acaggattag 780

P074570006W000-SEQLISTING-CEW.txt

ataccctggt agtccacgcc gtaaacgatg actgctaggt gtcgggtagc aaagctattc	840
ggtgccgcag ctaacgcaat aagcagtcca cctggggagt acgttcgcaa gaatgaaact	900
caaaggaatt gacgggggcc ngcacaagcg gtggagcatg tggtttaatt cgaannaacg	960
cgaagaacct tacctgatct tgacatcccg atgaccgctt cgtaatggaa gyttttcttc	1020
ggaacatcgg tgacaggtgg tgcattggtg tcgtcagctc gtgtcgtgag atgttgggtt	1080
aagtcccgca acgagcgcaa cccttatctt cagtagccag catttaggat gggcactctg	1140
gagagactgc cagggataac ctggaggaag gtggggatga cgttnaatca tcatgcccct	1200
tatgaccagg gctacacacg tgctacaatg gcgtaaacag agggaggcag agccgcgagg	1260
ccgagcaaat ctcaaaaata acgtctcagt tcggattgta gtctgcaact cgactacatg	1320
aagctggaat cgctagtaat cgagatcag aatgctgcgg tgaatacggt cccgggtctt	1380
gtacacaccg cccgtcacac catgggagtc agtaacgcc gaagtcagt acccaaccga	1440
aaggagggag ctgccgaagg tgggaccgat aactggggg	1479

<210> 121
 <211> 1390
 <212> DNA
 <213> Eisenbergiella tayi

<400> 121	
ggtataactt agtggcggac ggggtagtaa cgcgtgggaa acctgccctg taccggggga	60
taacacttag aaataggtgc taataccgca taagcgcacg gaaccgcatg gttccgtgtg	120
aaaaactccg gtggtacagg atggtcccgc gtctgattag ccagttggca gggtaacggc	180
ctaccaaagc gacgatcagt agccggcctg agaggggtgaa cggccacatt gggactgaga	240
cacggcccaa actcctacgg gaggcagcag tggggaatat tgcacaatgg gggaaaccct	300
gatgcagcga cgcgcgtga gtgaagaagt atttcggtat gtaaagctct atcagcaggg	360
aagaaaatga cggtagctga ctaagaagcc ccggctaact acgtgccagc agccgcggta	420
atacgtaggg ggcaagcgtt atccggattt actgggtgta aaggagcgt agacggcatg	480
gcaagccaga tgtgaaaacc cagggtcaa ccttgggatt gcatttgga ctgccaggct	540
ggagtgcagg agaggtgaag ggaattccta gtgtagcggg gaaatgcgta gatattagga	600
ggaacaccag tggcgaaggc ggcttactgg actgtaactg acgttgaggc tcgaaagcgt	660
ggggagcaaa caggattaga taccctggta gtccacgcgg taaacgatga ttgctaggtg	720
taggtgggta tggaccatc ggtgccgcag ctaacgcaat aagcaatcca cctggggagt	780
acgttcgcaa gaatgaaact caaaggaatt gacggggacc cgcacaagcg gtggagcatg	840
tggtttaatt cgaagcaacg cgaagaacct taccaagtct tgacatccca atgacgcacc	900
tgtaaagagg tgttcccttc ggggcattgg agacaggtgg tgcattggtg tcgtcagctc	960
gtgtcgtgag atgttgggtt aagtcccgca acgagcgcaa cccttattct tagtagccag	1020
caggtaaagc tgggcactct aaggagactg ccggggataa cccggaggaa ggcggggatg	1080

```

acgtcaaadc atcatgcccc ttatgatttg ggctacacac gtgctacaat ggcgtaaaca      1140
aaggggaagcg agacagtgat gtggagcaaa tcyagaaaat aacgtctcag ttcggattgt      1200
agtctgcaac tcgactacat gaagctggaa tcgctagtaa tcgagaaatca gcatgtcgcg      1260
gtgaatacgt tcccgggtct tgtacacacc gcccgtcaca ccatgggagt tggaaatgcc      1320
cgaagtctgt gacctaaccg aaagggagga gcagccgaag gcaggtctga taactggggt      1380
gaagtcgtaa                                                                1390

```

```

<210> 122
<211> 1478
<212> DNA
<213> Clostridium symbiosum

```

```

<220>
<221> misc_feature
<222> (349)..(350)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (534)..(535)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (606)..(606)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (618)..(619)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (709)..(709)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (921)..(921)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (955)..(956)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (1012)..(1013)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (1015)..(1015)
<223> n is a, c, g, t or u

```

```

<220>
<221> misc_feature
<222> (1081)..(1081)
<223> n is a, c, g, t or u

```

<220>
 <221> misc_feature
 <222> (1185)..(1185)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1240)..(1240)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1391)..(1392)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1470)..(1470)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (1474)..(1475)
 <223> n is a, c, g, t or u

<400> 122
 aaacatgaga gtttgatcct ggctcaggat gaacgctggc ggcgtgccta acacatgcaa 60
 gtcgaacgaa gcgatttaac ggaagttttc ggatggaagt tgaattgact gagtggcgga 120
 cgggtgagta acgcgtgggt aacctgcctt gtactggggg acaacagtta gaaatgactg 180
 ctaataccgc ataagcgac agtattgcat gatacagtgt gaaaaactcc ggtggtacaa 240
 gatggacccg cgtctgatta gctagttggg aaggtaacgg cttaccaagg cgacgatcag 300
 tagccgacct gagagggtga ccggccacat tgggactgag acacggccnn aactcctacg 360
 ggaggcagca gtggggaata ttgcacaatg ggcgaaagcc tgatgcagcg acgccgcgtg 420
 agtgaagaag tatttcggta tgtaaagctc tatcagcagg gaagaaaatg acggtacctg 480
 actaagaagc cccggctaac tacgtgccag cagccgcggt aatacgtagg gggnnagcgt 540
 tatccggatt tactgggtgt aaaggagcg tagacggtaa agcaagtctg aagtgaaagc 600
 ccgcgntca actgcggnnc tgctttggaa actgtttaac tggagtgtcg gagaggtaag 660
 tggaattcct agtgtagcgg tgaaatgctg agatattagg aggaacacna gtggcgaagg 720
 cgacttactg gacgataact gacgttgagg ctcgaaagcg tggggagcaa acaggattag 780
 ataccctggg agtccacgcc gtaaacgatg aatactaggt gttggggagc aaagctcttc 840
 ggtgccgtcg caaacgcagt aagtattcca cctggggagt acgttcgcaa gaatgaaact 900
 caaaggaatt gacggggacc ngcacaagcg gtggagcatg tggtttaatt cgaannaacg 960
 cgaagaacct taccaggtct tgacatcgac tcgacggggg agtaacgtcc cnntnccttc 1020
 ggggcggaga agacagggtg tgcatggttg tcgtcagctc gtgtcgtgag atgttggggt 1080
 nagtcccgca acgagcgcaa cccttattct aagtagccag cggttcggcc gggaactctt 1140
 gggagactgc cagggataac ctggaggaag gtggggatga cgtcnaatca tcatgccctt 1200
 tatgatctgg gctacacacg tgctacaatg gcgtaaacan agagaagcaa gaccgcgagg 1260

P074570006W000-SEQLISTING-CEW.txt

tggagcaaat	ctcaaaaata	acgtctcagt	tcggactgca	ggctgcaact	cgctgcacg	1320
aagctggaat	cgctagtaat	cgcgaaatcag	aatgtcgcgg	tgaatacgtt	cccgggtctt	1380
gtacacaccg	nncgtcacac	catgggagtc	agtaacgccc	gaagtcagtg	acccaaccgc	1440
aaggagggag	ctgccgaagg	cgggaccgan	aacnnggg			1478

<210> 123
 <211> 1488
 <212> DNA
 <213> *Erysipelatoclostridium ramosum*

<400> 123	
agagtttgat	60
cctggctcag	
gatgaacgct	
ggcggcgtgc	
ctaatacatg	
caagtcgaac	
gcgagcactt	120
gtgctcgagt	
ggcgaacggg	
tgagtaatac	
ataagtaacc	
tgccctagac	
agggggataa	180
ctattggaaa	
cgatagctaa	
gaccgcatag	
gtacggacac	
tgcatggtga	
ccgtattaaa	240
agtgcctcaa	
agcactggta	
gaggatggac	
ttatggcgca	
ttagctggtt	
ggcggggtaa	300
cggcccacca	
aggcgacgat	
gcgtagccga	
cctgagaggg	
tgaccggcca	
cactgggact	360
gagacacggc	
ccagactcct	
acgggaggca	
gcagtaggga	
atgttctggca	
atgggggaaa	420
ccctgaccga	
gcaacgccgc	
gtgaaggaag	
aagggttttcg	
gattgtaaac	
ttctgttata	480
aaggaagaac	
ggcggctaca	
ggaaatggta	
gccgagtgac	
ggtactttat	
tagaaagcca	540
cggctaacta	
cgtgccagca	
gccgcggtaa	
tacgtaggtg	
gcaagcgtaa	
tccggaatta	600
ttgggcgtaa	
agaggagca	
ggcggcagca	
agggctctgtg	
gtgaaagcct	
gaagcttaac	660
ttcagtaagc	
catagaaacc	
aggcagctag	
agtgcaggag	
aggatcgtgg	
aattccatgt	720
gtagcggtag	
aatgcgtaga	
tatatggagg	
aacaccagtg	
gcgaaggcga	
cgatctggcc	780
tgcaactgac	
gctcagtcct	
gaaagcgtgg	
ggagcaaata	
ggattagata	
ccctagtagt	840
ccacgccgta	
aacgatgagt	
actaagtgtt	
ggatgtcaaa	
gttcagtgct	
gcagttaacg	900
caataagtac	
tccgcctgag	
tagtacgttc	
gcaagaatga	
aactcaaagg	
aattgacggg	960
ggccgcaca	
agcgggtggag	
catgtggttt	
aattcgaagc	
aacgcgaaga	
accttaccag	1020
gtcttgacat	
actcataaag	
gctccagaga	
tggagagata	
gctatatgag	
atacaggtgg	1080
tgcatggttg	
tcgtcagctc	
gtgtcgtgag	
atgttgggtt	
aagtcccgc	
acgagcgcaa	1140
cccttatcgt	
tagttaccat	
cattaagtgt	
gggactctag	
cgagactgcc	
agtgacaagc	1200
tggaggaagg	
cggggatgac	
gtcaaactcat	
catgcccctt	
atgacctggg	
ctacacacgt	1260
gctacaatgg	
atggtgcaga	
gggaagcgaa	
gccgcgaggt	
gaagcaaaac	
ccataaaacc	1320
attctcagtt	
cggattgtag	
tctgcaactc	
gactacatga	
agttggaatc	
gctagtaatc	1380
gcgaatcagc	
atgtcgcggg	
gaatacgttc	
tcgggccttg	
tacacaccgc	
ccgtcacacc	1440
acgagagttg	
ataacacccg	
aagccggtgg	
cctaaccgca	
aggaaggagc	
tgtctaaggt	1488
gggattgatg	
attgggggtga	
agtcgtaaca	
aggtaacc	

<210> 124

<211> 1528

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 124

```

atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg      60
aacgaagcaa ttaaaggaag ttttcggatg gaatttgatt gactgagtgg cggacgggtg      120
agtaacgcgt ggataacctg cctcacactg ggggataaca gttagaaatg actgctaata      180
ccgcataagc gcacagtacc gcatggtagc gtgtgaaaaa ctccggtggg gtgagatgga      240
tccgcgtctg attagccagt tggcggggta acggcccacc aaagcgacga tcagtagccg      300
acctgagagg gtgaccggcc acattgggac tgagacacgg cccaaactcc tacgggaggg      360
agcagtgggg aatattgcac aatgggcgaa agcctgatgc agcgacgccg cgtgagtga      420
gaagtatttc ggtatgtaaa gctctatcag caggaagaa aatgacggta cctgactaag      480
aagccccggc taactacgtg ccagcagccg cggtaatagc tagggggcaa gcgttatccg      540
gattcactgg gtgtaaaggg agcgtagacg gcgaagcaag tctgaagtga aaaccaggg      600
ctcaaccctg ggactgcttt ggaaactgtt ttgctagagt gtcggagagg taagtggaat      660
tcctagtgtg gcggtgaaat gcgtagatat taggaggaac accagtggcg aaggcggctt      720
actggacgat aactgacgtt gaggctcgaa agcgtgggga gcaaacagga ttagataccc      780
tggtagtcca cgccgtaaac gatgaatgct aggtgttggg gggcaaagcc cttcgggtgcc      840
gtcgcaaacg cagtaagcat tccacctggg gactacgttc gcaagaatga aactcaaagg      900
aattgacggg gacccgcaca agcgggtggg catgtgtgtt aattcgaagc aacgcgaaga      960
accttaccaa gtcttgacat cctcttgacc ggcgtgtaac ggcgccttcc cttcggggca     1020
agagagacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg ggttaagtcc     1080
cgcaacgagc gcaaccctta tccttagtag ccagcaggta gagctgggca ctctagggag     1140
actgccaggg ataacctgga ggaagggtgg gatgacgtca aatcatcatg ccccttatga     1200
tttgggctac acacgtgcta caatggcgta aacaaagga agcaagacag tgatgtggag     1260
caaatcccaa aaataacgtc ccagttcgga ctgtagtctg caacccgact acacgaagct     1320
ggaatcgcta gtaatcgca atcagaatgt cgcggtgaat acgttcccgg gtcttgtaca     1380
caccgcccgt cacaccatgg gagtcagcaa cgcccgaagt cagtgacca actcgcaaga     1440
gagggagctg ccgaaggcgg ggcaggtaac tggggtgaag tcgtaacaag gtagccgtat     1500
cggaaggtgc ggctggatca cctccttt                                     1528

```

<210> 125

<211> 1530

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

P074570006W000-SEQLISTING-CEW.txt

```

<400> 125
atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg      60
aacgaagcaa ttaaaatgaa gttttcggat ggatttttga ttgactgagt ggcggacggg      120
tgagtaacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa      180
taccgcataa gcgcacagta ccgcatggta cgggtgtgaaa aactccggtg gtgtgggatg      240
gatccgcgtc tgattagcca gttggcgggg taacggccca ccaaagcgac gatcagtagc      300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag      360
gcagcagtgg ggaatattgc acaatgggcg aaagcctgat gcagcgacgc cgcgtagagt      420
aagaagtatt tcggtatgta aagctctatc agcaggaag aaaatgacgg tacctgacta      480
agaagccccg gctaactacg tgccagcagc cgcggttaata cgtagggggc aagcgttatc      540
cggatttact ggggtgtaaag ggagcgtaga cggcgaagca agtctgaagt gaaaaccag      600
ggctcaaccc tgggactgct ttggaaactg ttttgctaga gtgtcggaga ggtaagtgga      660
attcctagtg tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcggc      720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac      780
cctggtagtc cagccgtaa acgatgaatg ctaggtgttg gggggcaaag cccttcgggtg      840
ccgtcgcaaa cgagtaagc attccacctg gggagtacgt tcgcaagaat gaaactcaaa      900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa      960
gaaccttacc aagtcttgac atcctcttga ccggcgtgta acggcgcctt cccttcgggg      1020
caagagagac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt      1080
cccgcaacga gcgcaaccct tatccttagt agccagcagg taaagctggg cactctaggg      1140
agactgccag ggataacctg gaggaagggt gggatgacgt caaatcatca tgccccttat      1200
gatttgggct acacacgtgc tacaatggcg taaacaaagg gaagcaagac agtgatgtgg      1260
agcaaattccc aaaaataacg tcccagttcg gactgtagtc tgcaaccgga ctacacgaag      1320
ctggaatcgc tagtaatcgc gaatcagaat gtcgcggtga atacgttccc gggctttgta      1380
cacaccgccc gtcacacat gggagtcagc aacggccgaa gtcagtgacc caactcgcaa      1440
gagaggggagc tgccgaaggc ggggcaggta actggggtga agtcgtaaca aggtagccgt      1500
atcggaaggt gcggctggat cacctccttt      1530

```

```

<210> 126
<211> 1528
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic Polynucleotide

```

```

<400> 126
atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg      60
aacgaagcaa ttaaaatgaa gttttcggat ggatttttga ttgactgagt ggcggacggg      120

```


P074570006W000-SEQLISTING-CEW.txt

tgagtaacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta ccgcatggta cgggtgtgaaa aactccggtg gtgtgagatg	240
gatccgcgtc tgattagcca gttggcgggg taacggccca ccaaagcgac gatcagtagc	300
cgacctgaga ggggtgaccg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcagtgg ggaatattgc acaatgggcg aaagcctgat gcagcgacgc cgcgtgagtg	420
aagaagtatt tcggtatgta aagctctatc agcaggaag aaatgacggt acctgactaa	480
gaagccccgg ctaactacgt gccagcagcc gcggtaatc gtagggggca agcgttatcc	540
ggatttactg ggtgtaaagg gagcgtagac ggcgaagcaa gtctgaagtg aaaaccagg	600
gctcaaccct gggactgctt tggaactgt tttgctagag tgtcggagag gtaagtggaa	660
ttcctagtgt agcggtgaaa tgcgtagata ttaggaggaa caccagtggc gaaggcggct	720
tactggacga taactgacgt tgaggctcga aagcgtgggg agcaaacagg attagatacc	780
ctggtagtcc acgccgtaaa cgatgaatgc taggtgttgg gggcaaagcc cttcgggtgcc	840
gtcgcaaacg cagtaagcat tccacctggg gactacgttc gcaagaatga aactcaaagg	900
aattgacggg gacccgcaca agcgggtggag catgtggttt aattcgaagc aacgcgaaga	960
accttaccaa gtcttgacat cctcttgacc ggcgtgtaac ggcgccttcc cttcggggca	1020
agagagacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg ggttaagtcc	1080
cgcaacgagc gcaaccctta tccttagtag ccagcaggta aagctgggca ctctagggag	1140
actgccaggg ataacctgga ggaaggtggg gatgacgtca aatcatcatg ccccttatga	1200
tttgggctac acacgtgcta caatggcgta aacaaagga agcaagacag tgatgtggag	1260
caaatcccaa aaataacgtc ccagttcgga ctgtagtctg caaccgcact acacgaagct	1320
ggaatcgcta gtaatcgca atcagaatgt cgcgggtgaat acgttcccgg gtcttgtaca	1380
caccgcccgt cacaccatgg gagtcagcaa cgcccgaagt cagtgacca actcgcaaga	1440
gagggagctg ccgaaggcgg ggcaggtaac tggggtgaag tcgtaacaag gtagccgtat	1500
cggaaggtgc ggctggatca cctccttt	1528

<210> 127
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 127	
atgagagttt gatcctggct caggatgaac gctggcgggc tgcctaacac atgcaagtcg	60
aacgaagcaa ttaaaatgaa gttttcggat ggatttttga ttgactgagt ggcggacggg	120
tgagtaacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta ccgcatggta cgggtgtgaaa aactccggtg gtgtgagatg	240
gatccgcgtc tgattagcca gttggcgggg taacggccca ccaaagcgac gatcagtagc	300

P074570006W000-SEQLISTING-CEW.txt

cgacctgaga	gggtgaccgg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360
gcagcagtgg	ggaatattgc	acaatgggcg	aaagcctgat	gcagcgacgc	cgcgtagagt	420
aagaagtatt	tcggtatgta	aagctctatc	agcaggaag	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc	540
cggatttact	gggtgtaaag	ggagcgtaga	cggcgaagca	agtctgaagt	gaaaacccag	600
ggctcaaccc	tgggactgct	ttggaaactg	ttttgctaga	gtgtcggaga	ggtaagtgga	660
attcctagt	tagcggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc	720
ttactggacg	ataactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780
cctggtagtc	cacgccgtaa	acgatgaatg	ctaggtgttg	ggggcaaagc	ccttcggtgc	840
cgtcgcaaac	gcagtaagca	ttccacctgg	ggagtacgtt	cgcaagaatg	aaactcaaag	900
gaattgacgg	ggacccgcac	aagcggtgga	gcatgtggtt	taattcgaag	caacgcgaag	960
aaccttacca	agtcttgaca	tcctcttgac	cggcgtgtaa	cggcgccttc	ccttcggggc	1020
aggagagaca	ggtggtgcat	ggttgctcgtc	agctcgtgtc	gtgagatgtt	gggttaagtc	1080
ccgcaacgag	cgcaaccctt	atccttagta	gccagcaggt	agagctgggc	actctagggg	1140
gactgccagg	gataacctgg	aggaagggtg	ggatgacgtc	aaatcatcat	gccccttatg	1200
atttgggcta	cacacgtgct	acaatggcgt	aaacaaagg	aagcaagaca	gtgatgtgga	1260
gcaaattcca	aaataacgtc	ccagttcgga	ctgtagtctg	caacccgact	acacgaagct	1320
ggaatcgcta	gtaatcgca	atcagaatgt	cgcggtgaat	acgttcccgg	gtcttgta	1380
caccgcccgt	cacaccatgg	gagtcagcaa	cgcccgaagt	cagtgacca	actcgcaaga	1440
gagggagctg	ccgaaggcgg	ggcaggtaac	tgggggtgaag	tcgtaacaag	gtagccgtat	1500
cggaagggtgc	ggctggatca	cctccttt				1528

<210> 128
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 128	
atgagagttt	gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg 60
aacgaagcaa	ttaaaatgaa gttttcggat ggattttaat tgactgagtg gcggacgggt 120
gagtaacgcg	tggataacct gcctcacact gggggataac agttagaaat gactgctaata 180
accgcataag	cgcacagtac cgcattggtac ggtgtgaaaa actccgggtg tgtgagatgg 240
atccgcgtct	gattagccag ttggcggggg aacggcccac caaagcgacg atcagtagcc 300
gacctgagag	ggtgaccggc cacattggga ctgagacacg gcccaaactc ctacgggagg 360
cagcagtggg	gaatattgca caatgggcca aagcctgatg cagcgacgcc gcgtgagtga 420
agaagtattt	cggtatgtaa agctctatca gcagggaaga aaatgacggg acctgactaa 480

P074570006W000-SEQLISTING-CEW.txt

gaagccccgg	ctaactacgt	gccagcagcc	gcggtaat	ac	gtagggggca	agcgttatcc	540
ggatttactg	ggtgtaaagg	gagcgtagac	ggcgaagcaa	gtctgaagt	g	aaaaccagg	600
gctcaaccct	gggactgctt	tggaactgt	tttgctagag	tgtcggagag	g	taagtggaa	660
ttcctagtgt	agcggtgaaa	tgcgtagata	ttaggaggaa	caccagtggc	g	aaaggcggct	720
tactggacga	taactgacgt	tgaggctcga	aagcgtgggg	agcaaacagg	a	ttagatacc	780
ctggtagtcc	acgccgtaaa	cgatgaatgc	taggtgttgg	ggggcaaagc	c	cttcgggtgc	840
cgtcgcaaac	gcagtaagca	ttccacctgg	ggagtacgtt	cgcaagaatg	a	aactcaaag	900
gaattgacgg	ggacccgcac	aagcggtgga	gcatgtgggt	taattcgaag	c	aacgcgaag	960
aaccttacca	agtcttgaca	tcctcttgac	cggcgtgtaa	cggcgccttc	c	cttcggggc	1020
aagagagaca	ggtggtgcat	ggttgtcgtc	agctcgtgtc	gtgagatgtt	g	gggttaagtc	1080
ccgcaacgag	cgcaaccctt	atccttagta	gccagcaggt	aaagctgggc	a	ctctaggga	1140
gactgccagg	gataacctgg	aggaagggtg	ggatgacgtc	aaatcatcat	g	cccccttatg	1200
atttgggcta	cacacgtgct	acaatggcgt	aaacaaagg	aagcaagaca	g	tgatgtgga	1260
gcaaatccca	aaataacgtc	ccagttcggg	ctgtagtctg	caacccgact	a	acacgaagct	1320
ggaatcgcta	gtaatcgca	atcagaatgt	cgcggtgaat	acgttcccgg	g	tcttgtaca	1380
caccgcccgt	cacaccatgg	gagtcagcaa	cgcccgaagt	cagtgaacca	a	ctcgcaaga	1440
gagggagctg	ccgaaggcgg	ggcaggtaac	tgggggtgaag	tcgtaacaag	g	tagccgtat	1500
cggaaggtgc	ggctggatca	cctccttt					1528

<210> 129

<211> 1522

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 129

tcaaagagtt	tgatcctggc	tcaggacgaa	cgctggcggc	gcgccctaaca	catgcaagtc	60
gaacggagct	tacgttttga	agttttcgga	tggatgaatg	taagcttagt	ggcggacggg	120
tgagtaacac	gtgagcaacc	tgcctttcag	agggggataa	cagccggaaa	cggctgctaa	180
taccgcatga	tgttgcgggg	gcacatgcc	ctgcaaccaa	aggagcaatc	cgctgaaaga	240
tgggctcgcg	tccgattagc	cagttggcgg	ggtaacggcc	caccaaagcg	acgatcggta	300
gccggactga	gaggttgaac	ggccacattg	ggactgagac	acggcccaga	ctcctacggg	360
aggcagcagt	gggggatatt	gcacaatggg	cgaaagcctg	atgcagcgac	gccgcgtgag	420
ggaagacggt	cttcggattg	taaacctctg	tctttgggga	agaaaatgac	ggtacccaaa	480
gaggaagctc	cggctaacta	cgtgccagca	gccgcggtaa	tacgtaggga	gcaagcgttg	540
tccggaatta	ctgggtgtaa	agggagcgta	ggcgggatgg	caagtagaat	gttaaattcca	600
tcgggtcaac	cgggtggctgc	gttctaaact	gccgttcttg	agtgaagtag	aggcaggcgg	660

P074570006W000-SEQLISTING-CEW.txt

aattcctagt gtagcgggtga aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cctgctgggc ttttaactgac gctgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgatt actagggtgtg gggggactga ccccttcctg	840
gccgcagtta acacaataag taatccacct ggggagtagc gccgcaaggt tgaaactcaa	900
aggaattgac gggggcccg cacaagcagtg gagtatgtgg ttttaattcga agcaacgcga	960
agaaccttac caggtcttga catcggtatgc atagcctaga gatagggtgaa gcccttcggg	1020
gcatccagac aggtggtgca tgggtgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcacga gcgcaaccct tattattagt tgctacgcaa gagcactcta atgagactgc	1140
cggtgacaaa acggaggaag gtggggatga cgtcaaatca tcatgcccct tatgacctgg	1200
gctacacacg tactacaatg gcactaaaac agagggcggc gacaccgcga ggtgaagcga	1260
atccccgaaaa agtgtctcag ttcagattgc aggtctgaac ccgcctgcat gaagtcggaa	1320
ttgctagtaa tcgcggatca gcatgccgcg gtgaatacgt tcccgggcct tgtacacacc	1380
gcccgtcaca ccatgggagt cggtaacacc cgaagccagt agcctaaccg caaggggggc	1440
gctgtcgaag gtgggattga tgactggggg gaagtcgtaa caaggtagcc gtatcggaag	1500
gtgcggctgg atcacctcct tt	1522

<210> 130
 <211> 1521
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 130	
tcaaagagtt tgatcctggc tcaggacgaa cgctggcggc ggcctaaca catgcaagtc	60
gaacggagct tacgttttga agttttcgga tggacgaatg taagcttagt ggcggacggg	120
tgagtaacac gtgagcaacc tgcctttcag aggggataac agccggaaac ggctgcta	180
accgcatgat gttgcggggg cacatgcccc tgcaacaaaa ggagcaatcc gctgaaagat	240
gggctcgcgt ccgattagcc agttggcggg gtaacggccc accaaagcga cgatcggtag	300
ccggactgag aggttgaacg gccacattgg gactgagaca cggcccagac tcctacggga	360
ggcagcagtg ggggatattg cacaatgggc gaaagcctga tgcagcgacg ccgcgtgagg	420
gaagacggtc ttcggattgt aaacctctgt ctttggggaa gaaaatgacg gtacccaaag	480
aggaagctcc ggctaactac gtgccagcag ccgcggtaat acgtagggag caagcgttgt	540
ccggaattac tgggtgtaaa gggagcgtag gcgggatggc aagtagaatg ttaaattccat	600
cggctcaacc ggtggctgcg ttctaaactg ccgttcttga gtgaagtaga ggcaggcgga	660
attcctagt tagcgggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcggc	720
ctgctgggct ttaactgacg ctgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cagccgtaa acgatgatta ctagggtgtgg ggggactgac cccttcctgt	840

P074570006W000-SEQLISTING-CEW.txt

ccgcagttaa cacaataagt aatccacctg gggagtacgg ccgcaagggt gaaactcaaa	900
ggaattgacg ggggcccgc caagcagtgg agtatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcggatgca tagcctagag ataggtgaag cccttcgggg	1020
catccagaca ggtggtgcat ggttgctgctc agctcgtgctc gtgagatggt gggttaagtc	1080
ccgcaacgag cgcaaccctt attattagtt gctacgcaag agcactctaa tgagactgcc	1140
gttgacaaaa cggaggaagg tggggatgac gtcaaatcat catgcccctt atgacctggg	1200
ctacacacgt actacaatgg cactaaaaca gagggcgggc acaccgcgag gtgaagcgaa	1260
tcccgaaaaa gtgtctcagt tcagattgca ggctgcaacc cgcctgcatg aagtcggaat	1320
tgctagtaat cgcgatcag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg	1380
cccgtcacac catgggagtc ggtaacacc gaagccagta gcctaaccgc aagggggggc	1440
ctgtcgaagg tgggattgat gactggggtg aagtcgtaac aaggtagccg tatcggaagg	1500
tgcggctgga tcacctcctt t	1521

<210> 131
 <211> 1522
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 131	
tcaaagagtt tgatcctggc tcaggacgaa cgctggcggc ggcctaaca catgcaagtc	60
gaacggagct tacgttttga agttttcgga tggatgaatg taagcttagt ggcggacggg	120
tgagtaacac gtgagcaacc tgcctttcag agggggataa cagccgaaa cggctgctaa	180
taccgcatga tggtgcgggg gcacatgcc ctgcaaccaa aggagcaatc cgctgaaaga	240
tgggctcgcg tccgattagc cagttggcgg ggtaacggcc caccaaagcg acgatcggtg	300
gccggactga gaggttgaac ggccacattg ggactgagac acggcccaga ctctacggg	360
aggcagcagt gggggatatt gcacaatgg cgaaagcctg atgcagcgac gccgcgtgag	420
ggaagacggg cttcggattg taaacctctg tctttgggga agaaaatgac ggtacccaaa	480
gaggaagctc cggctaacta cgtgccagca gccgcggtaa tacgtaggga gcaagcgttg	540
tccggaatta ctgggtgtaa agggagcgta ggcgggatgg caagtagaat gttaaattcca	600
tcggctcaac cgggtggctgc gttctaaact gccgttcttg agtgaagtag aggcaggcgg	660
aattcctagt gtagcgtga aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cctgctgggc tttaactgac gctgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgatt actaggtgtg gggggactga ccccttcctg	840
gccgcagtta acacaataag taatccacct ggggagtacg gccgcaagg tgaactcaa	900
aggaattgac gggggccgc acaagcagtg gagtatgtgg tttaattcga agcaacgcga	960
agaaccttac caggtcttga catcgatgc atagcctaga gataggtgaa gcccttcggg	1020

P074570006W000-SEQLISTING-CEW.txt

gcatccagac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaaccct tattattagt tgctacgcaa gagcactcta atgagactgc	1140
cgttgacaaa acggaggaag gtggggatga cgtcaaatca tcatgccctt tatgacctgg	1200
gctacacacg tactacaatg gcactaaaac agagggcggc gacaccgcga ggtgaagcga	1260
atcccgaaaa agtgtctcag ttcagattgc aggctgcaac ccgcctgcat gaagtcggaa	1320
ttgctagtaa tcgcggatca gcatgccgcg gtgaatacgt tcccgggcct tgtacacacc	1380
gcccgtcaca ccatgggagt cggtaacacc cgaagccagt agcctaaccg caaggggggc	1440
gctgtcgaag gtgggattga tgactggggg gaagtcgtaa caaggtagcc gtatcggaag	1500
gtgcggctgg atcacctcct tt	1522

<210> 132
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 132 tacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg	120
gtgagtaacg cgtgggcaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agcgcacagg accgcatggg tagtgtgtaa aaactccggg ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaaaggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgacc gccacattgg gactgagaca cggcccaaac tcctacggga	360
ggcagcagtg gggaatattg cacaatgggg gaaaccctga tgcagcgacg ccgcgtgaag	420
gaagaagtat ttcggtatgt aaacttctat cagcaggga gaaaatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgccagcag ccgcggtaat acgtatggtg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc tttggaaact gcagatctgg agtgccggag aggtaagcgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttactggac ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact actaggtgtc ggtgtgcaaa gcacatcggt	840
gccgcagcaa acgcaataag tagtccacct ggggagtagc ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctggtcttga catccggatg acgggcgagt aatgtcgccg tcccttcggg	1020
gcgtccgaga cagggtgtgc atggttgtcg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgcacag agcgaaccc ttatcttcag tagccagcat ataaggtggg cactctggag	1140
agactgccag ggagaacctg gaggaagggt gggatgacgt caaatcatca tgccccttat	1200

P074570006W000-SEQLISTING-CEW.txt

ggccagggct	acacacgtgc	tacaatggcg	taaacaaagg	gaagcgagag	ggtgacctgg	1260
agcgaatccc	aaaaataacg	tctcagttcg	gattgtagtc	tgcaactcga	ctacatgaag	1320
ctggaatcgc	tagtaatcgc	ggatcagcat	gccgcggtga	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagt	aacgcccga	gccagtgacc	caaccttaga	1440
ggagggagct	gtcgaaggcg	ggacggataa	ctgggggtgaa	gtcgtaacaa	ggtagccgta	1500
tcggaaggtg	cggctggatc	acctccttt				1529

<210> 133
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 133	
tacgagagtt	tgatcctggc tcaggatgaa cgctggcggc gtgcctaaca catgcaagtc 60
gagcgaagcg	ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg 120
gtgagtaacg	cgtgggcaac ctgcctcata cagggggata acagttagaa atgactgcta 180
ataccgcata	agcgcacagg accgcatggg ttagtggtgaa aaactccggg ggtatgagat 240
ggacccgcgt	ctgattaggt agttggtggg gtaaaggcct accaagccga cgatcagtag 300
ccgacctgag	aggggtgacc gccacattgg gactgagaca cggcccaaac tcctacggga 360
ggcagcagtg	gggaatatgg cacaatgggg gaaaccctga tgcagcgacg ccgcgtgaag 420
gaagaagtat	ttcggtatgt aaacttctat cagcagggaa gaagatgacg gtacctgagt 480
aagaagcacc	ggctaaatac gtgccagcag ccgcggtaac acgtatgggt caagcgttat 540
ccggatttac	tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca 600
gggctcaacc	ctgggactgc tttggaaact gcagatctgg agtgccggag aggtaagcgg 660
aattcctagt	gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg 720
cttactggac	ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata 780
ccctggtagt	ccacgccgta aacgatgact actaggtgtc ggtgtgcaaa gcacatcggg 840
gccgcagcaa	acgcaataag tagtccacct ggggagtagc ttcgcaagaa tgaaactcaa 900
aggaattgac	ggggacccgc acaagcgggt gagcatgtgg tttaattcga agcaacgcga 960
agaaccttac	ctggctctga catccggatg acgggcgagt aatgtcgccg tcccttcggg 1020
gcacccgaga	caggtggtgc atggttgctg tcagctcgtg tcgtgagatg ttgggttaag 1080
tcccgcacg	agcgaaccc ttatcttcag tagccagcat ataagggtgg cactctggag 1140
agactgccag	ggagaacctg gaggaagggt gggatgacgt caaatcatca tgccccttat 1200
ggccagggct	acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctga 1260
agcgaatccc	aaaaataacg tctcagttcg gattgtagtc tgcaactcga ctacatgaag 1320
ctggaatcgc	tagtaatcgc ggatcagcat gccgcggtga atacgttccc gggctcttgta 1380

P074570006W000-SEQLISTING-CEW.txt

cacaccgccc gtcacacccat gggagtcagt aacgcccga gccagtgacc caaccttaga	1440
ggaggggagct gtcgaaggcg ggacggataa ctgggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggtg cggctggatc acctccttt	1529

<210> 134
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 134 tacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg	120
gtgagtaacg cgtgggcaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agcgcacagg accgcatggt gtagtgtgaa aaactccggt ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaaaggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgacc gccacattgg gactgagaca cggcccaaac tcctacggga	360
ggcagcagtg gggaatattg cacaatgggg gaaaccctga tgcagcgacg ccgcgtgaag	420
gaagaagtat ttcggtatgt aaacttctat cagcagggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgccagcag ccgcggtaac acgtatggtg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc tttggaaact gcagatctgg agtgccggag aggtaagcgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttactggac ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact actaggtgtc ggtgtgcaaa gcacatcggg	840
gccgcagcaa acgcaataag tagtccacct ggggagtagc ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcgggt gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctggtcttga catccggatg acgggcgagt aatgtcgccg tcccttcggg	1020
gcacccgaga cagggtggtgc atggttgctg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgcacag agcgcacacc ttatcttcag tagccagcat ataaggtggg cactctggag	1140
agactgccag ggagaacctg gaggaagggt gggatgacgt caaatcatca tgccccttat	1200
ggccagggct acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgg	1260
agcgaatccc aaaaataacg tctcagttcg gattgtagtc tgcaactcga ctacatgaag	1320
ctggaatcgc tagtaatcgc ggatcagcat gccgcggtga atacgttccc gggctcttga	1380
cacaccgccc gtcacacccat gggagtcagt aacgcccga gccagtgacc caaccttaga	1440
ggaggggagct gtcgaaggcg ggacggataa ctgggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggtg cggctggatc acctccttt	1529

P074570006W000-SEQLISTING-CEW.txt

<210> 135
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 135
 tacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcctaaca catgcaagtc 60
 gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg 120
 gtgagtaacg cgtgggcaac ctgcctcata cagggggata acagttagaa atgactgcta 180
 ataccgcata agcgcacagg accgcatggt gtagtgtgaa aaactccggt ggtatgagat 240
 ggacccgcgt ctgattaggt agttggtggg gtaaaggcct accaagccga cgatcagtag 300
 ccgacctgag aggggtgacc gccacattgg gactgagaca cggcccaaac tcctacggga 360
 ggacgacgtg gggaatattg cacaatggg gaaaccctga tgcagcgacg ccgcgtgaag 420
 gaagaagtat ttcggtatgt aaacttctat cagcagggaa gaagatgacg gtacctgagt 480
 aagaagcacc ggctaaatac gtgccagcag ccgcggtaac acgtatggtg caagcgttat 540
 ccggatttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca 600
 gggctcaacc ctgggactgc tttggaaact gcagatctgg agtgccggag aggtaagcgg 660
 aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg 720
 ctactggac ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata 780
 ccctggtagt ccacgccgta aacgatgact actaggtgtc ggtgtgcaaa gcacatcgg 840
 gccgcagcaa acgcaataag tagtccacct ggggagtagc ttcgcaagaa tgaaactcaa 900
 aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga 960
 agaaccttac ctggtcttga catccggatg acgggcgagt aatgtcgccg tcccttcggg 1020
 gcatccgaga caggtggtgc atggttgctg tcagctcgtg tcgtgagatg ttgggttaag 1080
 tcccgaacg agcgaaccc ttatcttcag tagccagcat ataaggtggg cactctggag 1140
 agactgccag ggagaacctg gaggaagggt gggatgacgt caaatcatca tgccccttat 1200
 ggccagggct acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctga 1260
 agcgaatccc aaaaataacg tctcagttcg gattgtagtc tgcaactcga ctacatgaag 1320
 ctggaatcgc tagtaatcgc ggatcagcat gccgcggtga atacgttccc gggctcttga 1380
 cacaccgccc gtcacacat gggagtcagt aacgcccga gccagtgacc caaccttaga 1440
 ggagggagct gtcgaaggcg ggacggataa ctggggtgaa gtcgtaacaa ggtagccgta 1500
 tcggaagggtg cggctggatc acctccttt 1529

<210> 136
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 136

tacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg	120
gtgagtaacg cgtgggcaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agcgcacagg accgcatggt gtagtgtgaa aaactccggt ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaaaggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccaaac tcctacggga	360
ggcagcagtg gggaatattg cacaatgggg gaaaccctga tgcagcgacg ccgcgtgaag	420
gaagaagtat ttcggtatgt aaacttctat cagcagggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgccagcag ccgcggtaat acgtatggtg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc tttggaaact gcagatctgg agtgccggag aggtaagcgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttactggac ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact actaggtgtc ggtgtgcaaa gcacatcgg	840
gccgcagcaa acgcaataag tagtccacct ggggagtagc ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctggtcttga catccggatg acgggcgagt aatgtcgccg tcccttcggg	1020
gcatccgaga caggtggtgc atggttgctg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgcacg agcgaaccc ttatcttcag tagccagcat ataaggtggg cactctggag	1140
agactgccag ggagaacctg gaggaagggt gggatgacgt caaatcatca tgccccttat	1200
ggccagggct acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgg	1260
agcgaatccc aaaaataacg tctcagttcg gattgtagtc tgcaactcga ctacatgaag	1320
ctggaatcgc tagtaatcgc ggatcagcat gccgcggtga atacgttccc gggctcttga	1380
cacaccgccc gtcacacat gggagtcagt aacgcccgaa gccagtgacc caaccttaga	1440
ggagggagct gtcgaaggcg ggacggataa ctggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggtg cggctggatc acctccttt	1529

<210> 137

<211> 1527

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 137

atgagagttt gatcctagct caggatgaac gctggcggcg tgcctaacac atgcaagtcg	60

P074570006W000-SEQLISTING-CEW.txt

aacgaagcaa tttaacggaa gttttcggat ggaagttgaa ttgactgagt ggcggacggg	120
tgagtaacgc gtgggtaacc tgccttgtagc tgggggacaa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta tcgcatgata cagtgtgaaa aactccggtg gtacaagatg	240
gacccgcgtc tgattagcta gttggttaagg taacggctta ccaaggcgac gatcagtagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcagtgg ggaatattgc acaatgggag aaagcctgat gcagcgacgc cgcgtgagt	420
aagaagtatt tcggtatgta aagctctatc agcagggag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcagc cgcggtaata cgtagggggc aagcgttatc	540
cggatttact ggggtgtaaag ggagcgtaga cggtaaagca agtctgaagt gaaagccgc	600
ggctcaactg cgggactgct ttggaaactg tttaactgga gtgtcggaga ggtaagtgga	660
attcctagtg tagcgggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cagccgtaa acgatgaata ctaggtgttg gggagcaaag ctcttcggtg	840
ccgtcgcaaa cgcagtaagt attccacctg gggagtacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga cgggggagta acgtcccctt cccttcgggg	1020
cggagaagac aggtggtgca tgggtgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaaccct tattctaagt agccagcggg tcggccggga actcttgga	1140
gactgccagg gataacctgg aggaagggtg ggatgacgtc aaatcatcat gcccttatg	1200
atctgggcta cacacgtgct acaatggcgt aaacaaagag aagcaagacc gcgaggtgga	1260
gcaaattctca aaaataacgt ctcatgtcgg actgcaggct gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttcccg ggtcttgtag	1380
acaccgcccg tcacaccatg ggagtcagta acgcccgaag tcagtgacct aaccgcaagg	1440
agggagctgc cgaaggcggg accgataact ggggtgaagt cgtaacaagg tagccgtatc	1500
ggaagggtgc gctggatcac ctctttt	1527

<210> 138
 <211> 1527
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 138	
atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg	60
aacgaagcga tttaacggaa attttcggat ggaagttgaa ttgactgagt ggcggacggg	120
tgagtaacgc gtgggtaacc tgccttgtagc tgggggacaa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta tcgcatgata cagtgtgaaa aactccggtg gtacaagatg	240

P074570006W000-SEQLISTING-CEW.txt

gacccgcgtc tgattagcta gttggtaagg taacggctta ccaaggcgac gatcagtagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcagtgg ggaatattgc acaatgggcg aaagcctgat gcagcgacgc cgcgtgagtg	420
aagaagtatt tcggtatgta aagctctatc agcaggaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcagc cgcggtaata cgtagggggc aagcgttatc	540
cggatttact ggggtgtaaag ggagcgtaga cggtaaagca agtctgaagt gaaagcccg	600
ggctcaactg cgggactgct ttggaaactg ttttaactgga gtgtcggaga ggtaagtgga	660
attcctagtg tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgccgtaa acgatgaata ctaggtgttg gggagcaaag ctcttcgggtg	840
ccgtcgcaaa cgcagtaagt attccacctg gggagtacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga cgggggagta acgtcccctt cccttcgggg	1020
cggagaagac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaaccct tattctaagt agccagcggg tcggccggga actcttgga	1140
gactgccagg gataacctgg aggaaggtgg ggatgacgtc aaatcatcat gccccttatg	1200
atctgggcta cacacgtgct acaatggcgt aaacaaagag aagcaagacc gcgaggtgga	1260
gcaaacttca aaaataacgt ctcatgtcgg actgcaggct gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttcccg ggtcttgtag	1380
acaccgcccg tcacaccatg ggagtcagta acgcccgaag tcagtgacct aaccgcaagg	1440
agggagctgc cgaaggcggg accgataact ggggtgaagt cgtaacaagg tagccgtatc	1500
ggaaggtgcg gctggatcac ctccttt	1527

<210> 139
 <211> 1527
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 139	
atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg	60
aacgaagcga tttaacggaa gttttcggat ggaagttgaa ttgactgagt ggcggacggg	120
tgagtaacgc gtgggtaacc tgccttgtagc tgggggacaa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta tcgcatgata cagtgtgaaa aactccggtg gtacaagatg	240
gacccgcgtc tgattagcta gttggtaagg taacggctta ccaaggcgac gatcagtagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcagtgg ggaatattgc acaatgggcg aaagcctgat gcagcgacgc cgcgtgagtg	420

P074570006W000-SEQLISTING-CEW.txt

aagaagtatt	tcggtatgta	aagctctatc	agcaggaag	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcagc	cgcggttaata	cgtagggggc	aagcgttatc	540
cggatttact	gggtgtaaag	ggagcgtaga	cggtaaagca	agtctgaagt	gaaagcccg	600
ggctcaactg	cgggactgct	ttggaaactg	tttaactgga	gtgtcggaga	ggtaagtgga	660
attcctagt	tagcggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcgac	720
ttactggacg	ataactgacg	ttgaggctcg	aaagcggtgg	gagcaaacag	gattagatac	780
cctggtagtc	cacgccgtaa	acgatgaata	ctaggtgttg	gggagcaaag	ctcttcgggtg	840
ccgtcgcaaa	cgagtaagt	attccacctg	gggagtacgt	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcggtgg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	aggtcttgac	atcgatccga	cgggggagta	acgtcccctt	cccttcgggg	1020
cggagaagac	aggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgaacga	gcgcaaccct	tattctaagt	agccagcggt	tcggccggga	actcttgga	1140
gactgccagg	gataacctgg	aggaaggtgg	ggatgacgtc	aatcatcat	gccccttatg	1200
atctgggcta	cacacgtgct	acaatggcgt	aaacaaagag	aagcaagacc	gcgaggtgga	1260
gcaaactca	aaaataacgt	ctcagttcgg	actgcaggct	gcaactcgcc	tgcacgaagc	1320
tggaatcgct	agtaatcgcg	aatcagaatg	tcgcggtgaa	tacgttccc	ggtcttgtag	1380
acaccgccc	tcacaccatg	ggagtcagta	acgccgaag	tcagtgacc	aaccgcaagg	1440
agggagctgc	cgaaggcggg	accgataact	ggggtgaagt	cgtaacaagg	tagccgtatc	1500
ggaaggtgcg	gctggatcac	ctccttt				1527

<210> 140
 <211> 1527
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 140	
atgagagttt	gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtcg 60
aacgaagcga	tttaacggaa gttttcggat ggaagttgga ttgactgagt ggcggacggg 120
tgagtaacgc	gtgggtaacc tgccttgtag tgggggacaa cagttagaaa tgactgctaa 180
taccgcataa	gcgcacagta tcgcatgata cagtgtgaaa aactccggtg gtacaagatg 240
gacccgcgtc	tgattagcta gttggttaagg taacggctta ccaaggcgac gatcagtagc 300
cgacctgaga	gggtgaccgg ccacattggg actgagacac ggcccaaact cctacgggag 360
gcagcagtgg	ggaatattgc acaatgggcg aaagcctgat gcagcgacgc cgcgtagtg 420
aagaagtatt	tcggtatgta aagctctatc agcaggaag aaaatgacgg tacctgacta 480
agaagccccg	gctaactacg tgccagcagc cgcggttaata cgtagggggc aagcgttatc 540
cggatttact	gggtgtaaag ggagcgtaga cggtaaagca agtctgaagt gaaagcccg 600

P074570006W000-SEQLISTING-CEW.txt

ggctcaactg cgggactgct ttggaaactg tttactgga gtgtcggaga ggtaagtgga	660
attcctagt tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgccgtaa acgatgaata ctaggtgttg gggagcaaag ctcttcggtg	840
ccgtcgcaaa cgagtaagt attccacctg gggagtacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga cgggggagta acgtcccctt cccttcgggg	1020
cggagaagac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaaccct tattctaagt agccagcggg tcggccggga actcttgga	1140
gactgccagg gataacctgg aggaagggtg ggatgacgtc aaatcatcat gccccttatg	1200
atctgggcta cacacgtgct acaatggcgt aaacaaagag aagcaagacc gcgagggtga	1260
gcaaactca aaaataacgt ctcatgtcgg actgcaggct gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttcccg ggtcttgtag	1380
acaccgcccgc tcacaccatg ggagtcagta acgcccgaag tcagtgaccc aaccgcaagg	1440
agggagctgc cgaaggcggg accgataact ggggtgaagt cgtaacaagg tagccgtatc	1500
ggaagggtgcg gctggatcac ctctttt	1527

<210> 141
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 141	
atcagagagt ttgatcctgg ctccaggatga acgctggcgg cgtgcttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctcttcgg attgaaactt atttgactga gcggcggacg	120
ggtgagtaac gcgtgggtaa cctgcctcat acagggggat aacagttaga aatggctgct	180
aataccgcat aagcgcacag gaccgcatgg tctggtgtga aaaactccgg tggtagtata	240
tggacccgcg tctgattagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gaggggtgaac ggccacattg ggactgagac acggcccaga ctcttacggg	360
aggcagcagt ggggaatatt gcacaatggg ggaaccctg atgcagcgac gccgcgtgaa	420
ggaagaagta tctcggtatg taaacttcta tcagcagggg agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gccgcggtaa tacgtagggg gcaagcggtta	540
tccggattta ctgggtgtaa agggagcgta gacggaagag caagtctgat gtgaaaggct	600
ggggcttaac cccaggactg cattggaaac tgtttttcta gagggtccgga gaggttaagcg	660
gaattcctag tgtagcgggtg aaatgcgtag atattaggag gaacaccagt ggcgaaggcg	720
gcttactgga cggtaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat	780

P074570006W000-SEQLISTING-CEW.txt

accctggtag tccacgccgt aaacgatgaa tactaggtgt cgggtggcaa agccattcgg	840
tgccgcagca aacgcaataa gtattccacc tggggagtag gttcgcaaga atgaaactca	900
aaggaattga cggggacccg cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg	960
aagaacctta ccaagtcttg acatccctct gaccggcccg taacggggcc ttcccttcgg	1020
ggcagaggag acaggtggtg catggttgtc gtcagctcgt gtcgtgagat gttgggttaa	1080
gtcccgcaac gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgac gtcaaatacat catgcccctt	1200
atgatttggg ctacacacgt gctacaatgg cgtaaacaaa ggaagcgag acagcgatgt	1260
tgagcaaatc caaaaataa cgtcccagtt cggactgcag tctgcaactc gactgcacga	1320
agctggaatc gctagtaatc gcgaatcaga atgtcgcggg gaatacgttc ccgggtcttg	1380
tacacaccgc ccgtcacacc atgggagtcg gtaacgcccg aagtcagtga cccaaccttt	1440
taggagggag ctgccgaagg cgggaccgat aactgggggtg aagtcgtaac aaggtagccg	1500
tatcggaagg tgcggctgga tcacctcctt t	1531

<210> 142
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 142	
atcagagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctcttcgg attgaaactt atttgactga gcggcggacg	120
ggtgagtaac gcgtgggtaa cctgcctcat acagggggat aacagttaga aatggctgct	180
aataccgcat aagcgcacag gaccgcatgg tctggtgtga aaaactccgg tggatatgaga	240
tggacccgcg tctgattagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gaggggtgaac ggccacattg ggactgagac acggcccaga ctcctacggg	360
aggcagcagt ggggaatatt gcacaatggg ggaaaccctg atgcagcgac gccgcgtgaa	420
ggaagaagta tctcgggtatg taaacttcta tcagcagggg agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gccgcggtaa tacgtagggg gcaagcgтта	540
tccggattta ctgggtgtaa agggagcgta gacggaagag caagtctgat gtgaaaggct	600
ggggcttaac cccaggactg cattggaaac tgtttttcta gaggtaagtg	660
gaattcctag tgtagcgggtg aaatgcgtag atattaggag gaacaccagt ggcgaaggcg	720
gcttactgga cggttaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat	780
accctggtag tccacgccgt aaacgatgaa tactaggtgt cgggtggcaa agccattcgg	840
tgccgcagca aacgcaataa gtattccacc tggggagtag gttcgcaaga atgaaactca	900
aaggaattga cggggacccg cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg	960

P074570006W000-SEQLISTING-CEW.txt

aagaacctta ccaagtcttg acatccctct gaccggcccc taacggggcc ttcccttcgg	1020
ggcagaggag acaggtggtg catggttgtc gtcagctcgt gtcgtgagat gttgggttaa	1080
gtcccgcaac gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgac gtcaaatacat catgcccctt	1200
atgatttggg ctacacacgt gctacaatgg cgtaaataaaa ggggaagcggg acagcgatgt	1260
tgagcaaatc ccaaaaataa cgtcccagtt cggactgcag tctgcaactc gactgcacga	1320
agctggaatc gctagtaatc gcgaatcaga atgtcgcggg gaatacgttc ccgggtcttg	1380
tacacaccgc ccgtcacacc atgggagtca gtaacgcccg aagtcagtga cccaacctta	1440
caggagggag ctgccgaagg cgggaccgat aactgggggtg aagtcgtaac aaggtagccg	1500
tatcggaagg tgcggctgga tcacctcctt t	1531

<210> 143
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 143	
atcagagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctcttcgg attgaagctt atttgactga gcggcggacg	120
ggtgagtaac gcgtgggtaa cctgcctcat acagggggat aacagttaga aatggctgct	180
aataccgcat aagcgcacag gaccgcatgg tctggtgtga aaaactccgg tggatatgaga	240
tggacccgcg tctgattagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gaggggtgaac ggccacattg ggactgagac acggcccaga ctcctacggg	360
aggcagcagt ggggaatatt gcacaatggg ggaaaccctg atgcagcgac gccgcgtgaa	420
ggaagaagta tctcgggtatg taaacttcta tcagcagggg agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gccgcggtaa tacgtagggg gcaagcgtaa	540
tccggattta ctgggtgtaa agggagcgta gacggaagag caagtctgat gtgaaaggct	600
ggggcttaac cccaggactg cattggaaac tgtttttcta gagtgccgga gaggtaagcg	660
gaattcctag tgtagcgggtg aaatgcgtag atattaggag gaacaccagt ggcgaaggcg	720
gcttactgga cggtaaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat	780
accctggtag tccacgccgt aaacgatgaa tactaggtgt cgggtggcaa agccattcgg	840
tgccgcagca aacgcaataa gtattccacc tggggagtac gttcgcaaga atgaaactca	900
aaggaattga cggggacccg cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg	960
aagaacctta ccaagtcttg acatccctct gaccggcccc taacggggcc ttcccttcgg	1020
ggcagaggag acaggtggtg catggttgtc gtcagctcgt gtcgtgagat gttgggttaa	1080
gtcccgcaac gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag	1140

P074570006W000-SEQLISTING-CEW.txt

ggagactgcc	ggggataacc	cggaggaagg	cggggacgac	gtcaaatacat	catgcccctt	1200
atgatttggg	ctacacacgt	gctacaatgg	cgtaaataaaa	gggaagcgag	acagcgatgt	1260
tgagcaaatac	ccaaaaataa	cgtcccagtt	cggactgcag	tctgcaactc	gactgcacga	1320
agctggaatc	gctagtaatc	gcgaatcaga	atgtcgcggg	gaatacgttc	ccgggtcttg	1380
tacacaccgc	ccgtcacacc	atgggagtca	gtaacgccc	aagtcagtga	cccaacctta	1440
caggagggag	ctgccgaagg	cgggaccgat	aactgggggtg	aagtcgtaac	aaggtagccg	1500
tatcggaagg	tgcggtctga	tcacctcctt	t			1531

<210> 144
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 144	
atcagagagt	ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt 60
cgagcgaagc	acttaagcgg atctcttcgg attgaaactt atttgactga gcggcggacg 120
ggtgagtaac	gcgtgggtaa cctgcctcat acagggggat aacagttaga aatggctgct 180
aataccgcat	aagcgcacag gaccgcatgg tctggtgtga aaaactccgg tggatatgaga 240
tggacccg	tctgattagc tagttggagg ggtaacggcc caccaaggcg acgatcagta 300
gccggcctga	gaggggtgaac ggccacattg ggactgagac acggcccaga ctcctacggg 360
aggcagcagt	ggggaatatt gcacaatggg ggaaaccctg atgcagcgac gccgcgtgaa 420
ggaagaagta	tctcggtatg taaacttcta tcagcagga agaaaatgac ggtacctgac 480
taagaagccc	cggctaacta cgtgccagca gccgcggtaa tacgtagggg gcaagcgтта 540
tccggattta	ctgggtgtaa agggagcgta gacggaagag caagtctgat gtgaaaggct 600
ggggcttaac	cccaggactg cattggaaac tgtttttcta gagtgccgga gaggtaagcg 660
gaattcctag	tgtagcgggtg aaatgcgtag atattaggag gaacaccagt ggcgaaggcg 720
gcttactgga	cggtaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat 780
accctggtag	tccacgccgt aaacgatgaa tactaggtgt cgggtggcaa agccattcgg 840
tgccgcagca	aacgcaataa gtattccacc tggggagtac gttcgcaaga atgaaactca 900
aaggaattga	cggggacccg cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg 960
aagaacctta	ccaagtcttg acatccctct gaccggccc taacggggcc ttcccttcgg 1020
ggcagaggag	acaggtgggtg catggttgtc gtcagctcgt gtcgtgagat gttgggttaa 1080
gtcccgaac	gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag 1140
ggagactgcc	ggggataacc cggaggaagg cggggacgac gtcaaatacat catgcccctt 1200
atgatttggg	ctacacacgt gctacaatgg cgtaaataaaa gggaagcgag acagcgatgt 1260
tgagcaaatac	ccaaaaataa cgtcccagtt cggactgcag tctgcaactc gactgcacga 1320

P074570006W000-SEQLISTING-CEW.txt

agctggaatc gctagtaatc gcgaatcaga atgtcgcggt gaatacgttc ccgggtcttg	1380
tacacaccgc ccgtcacacc atgggagtca gtaacgccc aagtcagtga cccaacctta	1440
caggagggag ctgccgaagg cgggaccgat aactggggtg aagtcgtaac aaggtagccg	1500
tatcggaagg tgcggctgga tcacctcctt t	1531

<210> 145
 <211> 1531
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 145	
atcagagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctcttcgg attgaaactt atttgactga gcggcggacg	120
ggtgagtaac gcgtgggtaa cctgcctcat acagggggat aacagttaga aatggctgct	180
aataccgcat aagcgcacag gaccgcatgg tctggtgtga aaaactccgg tggtagaga	240
tggacccgcg tctgattagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gaggtgtaac ggccacattg ggactgagac acggcccaga ctcctacggg	360
aggcagcagt ggggaatatt gcacaatggg ggaaaccctg atgcagcgac gccgcgtgaa	420
ggaagaagta tctcgggtatg taaacttcta tcagcagga agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gccgcggtaa tacgtagggg gcaagcgta	540
tccggattta ctgggtgtaa agggagcgta gacggaagag caagtctgat gtgaaaggct	600
ggggcttaac cccaggactg cattggaac tgttttcta gagtgccgga gaggtaagcg	660
gaattcctag tgtagcgggtg aaatgcgtag atattaggag gaacatcagt ggcgaaggcg	720
gcttactgga cggttaactga cgttgaggct cgaaagcgtg gggagcaaac aggattagat	780
accctggtag tccacgccgt aaacgatgaa tactaggtgt cgggtggcaa agccattcgg	840
tgccgcagca aacgcaataa gtattccacc tggggagtac gttcgcaaga atgaaactca	900
aaggaattga cggggacccg cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg	960
aagaacctta ccaagtcttg acatccctct gaccggccc taacggggcc ttcccttcgg	1020
ggcagaggag acaggtggtg catggttgtc gtcagctcgt gtcgtgagat gttgggttaa	1080
gtcccgaac gagcgcaacc cctatcctta gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgac gtcaaatacat catgcccctt	1200
atgatttggg ctacacacgt gctacaatgg cgtaaacaaa ggggaagcgag acagcgatgt	1260
tgagcaaata ccaaaaataa cgtcctagtt cggactgcag tctgcaactc gactgcacga	1320
agctggaatc gctagtaatc gcgaatcaga atgtcgcggt gaatacgttc ccgggtcttg	1380
tacacaccgc ccgtcacacc atgggagtca gtaacgccc aagtcagtga cccaacctta	1440
caggagggag ctgccgaagg cgggaccgat aactggggtg aagtcgtaac aaggtagccg	1500

tatcgggaagg tgcggctgga tcacctcctt t

1531

<210> 146
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 146
 aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc 60
 gagcgaagca ctttaagtttg attcttcgga tgaagacttt tgtgactgag cggcggacgg 120
 gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagttagaa atgactgcta 180
 ataccgcata agaccacggt accgcatggt acagtggtaa aaactccggt ggtatgagat 240
 ggacccgcgt ctgattaggt agttgggtgg gtaacggcct accaagccga cgatcagtag 300
 ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga 360
 ggacagcagt gggaatattg cacaatggag gaaactctga tgcagcgacg ccgctgaag 420
 gatgaagtat ttcggtatgt aaacttctat cagcagggaa gaaaatgacg gtacctgact 480
 aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat 540
 ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccc 600
 gggctcaacc ccgggactgc atttggaact gctgagctag agtgctggag aggcaagtgg 660
 aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg 720
 cttgctggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata 780
 ccctggtagt ccacgccgta aacgatgact gctaggtgtc ggggtggcaa gccattcgg 840
 gccgcagcta acgcaataag cagtccacct ggggagtacg ttcgcaagaa tgaaactcaa 900
 aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga 960
 agaaccttac ctgatcttga catcccgatg accgcttcgt aatggaagtt tttcttcgga 1020
 acatcggtag caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg ttgggttaag 1080
 tcccgaacg agcgaaccc ctatcttcag tagccagcag gttaagctgg gcaactctga 1140
 gagactgcca gggataacct ggaggaaggt ggggatgacg tcaaatacat atgccctta 1200
 tgaccagggc tacacacgtg ctacaatggc gtaaacaag agaagcgaac tcgagaggg 1260
 aagcaaatct caaaaataac gtctcagttc ggattgtagt ctgcaactcg actacatgaa 1320
 gctggaatcg ctagtaatcg cagatcagaa tgctgcggtg aatacgttcc cgggtcttgt 1380
 acacaccgcc cgtcacacca tgggagtcag taacgcccga agtcagtac ccaaccgtaa 1440
 ggagggagct gccgaagggt ggaccgataa ctggggtgaa gtcgtaacaa ggtagccgta 1500
 tcggaagggt cggctggatc acctcctt 1529

<210> 147
 <211> 1528
 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 147

```

aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc      60
gagcgaagca ctttgaaaag attcttcgga tgatttcctt tgtgactgag cggcggacgg      120
gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagttagaa atgactgcta      180
ataccgcata agaccacggt accgcatggt acagtggtaa aaactccggt ggtatgagat      240
ggacccgcgt ctgattaggt agttggtggg gtaacggcct accaagccga cgatcagtag      300
ccgacctgag aggggtgacc gccacattgg gactgagaca cggcccagac tcctacggga      360
ggcagcagtg gggaatattg cacaatggag gaaactctga tgcagcgacg ccgcgtgaag      420
gatgaagtat ttcggtatgt aaacttctat cagcagggaa gaaaatgacg gtacctgact      480
aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat      540
ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccc      600
gggctcaacc ccggactgca tttggaactg ctgagctaga gtgtcggaga ggcaagtgga      660
attcctagtg tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcggc      720
ttgctggacg atgactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac      780
cctggtagtc cacgccgtaa acgatgactg ctaggtgtcg ggtggcaaag ccattcgggtg      840
ccgcagctaa cgcaataagc agtccacctg gggagtacgt tcgcaagaat gaaactcaaa      900
ggaattgacg gggacccgca caagcgggtg agcatgtggt ttaattcgaa gcaacgcgaa      960
gaaccttacc tgatcttgac atcccgatga ccgcttcgta atggaagctt ttcttcggaa     1020
catcggtgac aggtggtgca tggttgtcgt cagctcgtgt cgtgagatgt tgggttaagt     1080
cccgcaacga gcgcaacccc tatcttcagt agccagcagg ttaagctggg cactctggag     1140
agactgccag ggataacctg gaggaaggtg gggatgacgt caaatcatca tgccccttat     1200
gaccagggct acacacgtgc tacaatggcg taaacaaaga gaagcgaact cgcgagggta     1260
agcaaattct aaaaataacg tctcagttcg gattgtagtc tgcaactcga ctacatgaag     1320
ctggaatcgc tagtaatcgc agatcagaat gctgcggtga atacgttccc gggctcttgta     1380
cacaccgccc gtcacaccat gggagtcagt aacgcccga gtcagtgacc caaccgtaag     1440
gagggagctg ccgaaggtgg gaccgataac tgggggtgaag tcgtaacaag gtagccgtat     1500
cggaaggtgc ggctggatca cctccttt                                     1528

```

<210> 148

<211> 1529

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 148

P074570006W000-SEQLISTING-CEW.txt

aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc	60
gagcgaagca ctttaagtttg attcttcgga tgaagacttt tgtgactgag cggcggacgg	120
gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agaccacggt accgcatggt acagtggtaa aaactccggt ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaacggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga	360
ggcagcagtg gggaatattg cacaatggag gaaactctga tgcagcgacg ccgcgtgaag	420
gatgaagtat ttcggtatgt aaacttctat cagcagggaa gaaaatgacg gtacctgact	480
aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccc	600
gggctcaacc ccgggactgc atttggaact gctgagctag agtgctggag aggcaagtgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttgctggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact gctaggtgtc ggggtggcaaa gccattcgg	840
gccgcagcta acgcaataag cagtccacct ggggagtagc ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctgatcttga catcccgatg accgcttcgt aatggaagct tttcttcgga	1020
acatcggtag caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgaacg agcgaaccc ctatcttcag tagccagcag gtttaagctgg gcaactctga	1140
gagactgcca gggataacct ggaggaaggt ggggatgacg tcaaatacatc atgcccctta	1200
tgaccagggc tacacacgtg ctacaatggc gtaaacaaag agaagcgaac tcgcgaggg	1260
aagcaaatct caaaaataac gtctcagttc ggattgtagt ctgcaactcg actacatgaa	1320
gctggaatcg ctagtaatcg cagatcagaa tgctgcgggtg aatacgttcc cgggtcttgt	1380
acacaccgcc cgtcacacca tgggagtcag taacgcccga agtcagtgac ccaaccgtaa	1440
ggagggagct gccgaagggt ggaccgataa ctggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggt cggctggatc acctccttt	1529

<210> 149
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 149	
aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc	60
gagcgaagca ctttggaaag attcttcgga tgatttcctt tgtgactgag cggcggacgg	120
gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagttagaa atgactgcta	180

P074570006W000-SEQLISTING-CEW.txt

ataccgcata agaccacggt accgcatggt acagtggtaa aaactccggt ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaacggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga	360
ggcagcagtg gggaatattg cacaatggag gaaactctga tgcagcgacg ccgcgtgaag	420
gatgaagtat ttcggtatgt aaacttctat cagcagggaa gaaaatgacg gtacctgact	480
aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccc	600
gggctcaacc ccgggactgc atttggaact gctgagctag agtgctggag aggcaagtgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttgctggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact gctaggtgtc ggggtggcaaa gccattcggt	840
gccgcagcta acgcaataag cagtccacct ggggagtagc ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcggtg gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctgatcttga catcccgatg accgcttcgt aatggaagct tttcttcgga	1020
acatcggtag caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgcaacg agcgcaacc ctatcttcag tagccagcag gtttaagctgg gcaactctga	1140
gagactgcca gggataacct ggaggaaggt ggggatgacg tcaaatacatc atgcccctta	1200
tgaccagggc tacacacgtg ctacaatggc gtaaacaaag agaagcgaac tcgagagggg	1260
aagcaaatct caaaaataac gtctcagttc ggattgtagt ctgcaactcg actacatgaa	1320
gctggaatcg ctagtaatcg cagatcagaa tgctgcgggtg aatacgttcc cgggtcttgt	1380
acacaccgcc cgtcacacca tgggagtcag taacgcccga agtcagttag ccaaccgtaa	1440
ggagggagct gccgaagggt ggaccgataa ctggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggt cggctggatc acctccttt	1529

<210> 150
 <211> 1528
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 150	
aacgagagtt tgatcctggc tcaggatgaa cgctggcggc gtgcttaaca catgcaagtc	60
gagcgaagcg ctttaagtttg attcttcgga tgaagacttt tgtgactgag cggcggacgg	120
gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agaccacggt accgcatggt acagtggtaa aaactccggt ggtatgagat	240
ggacccgcgt ctgattaggt agttggtggg gtaacggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga	360

P074570006W000-SEQLISTING-CEW.txt

ggcagcagtg	gggaatattg	cacaatggag	gaaactctga	tgacgagcag	ccgcgtgaag	420
gatgaagtat	ttcgggtatgt	aaacttctat	cagcagggaa	gaaaatgacg	gtacctgact	480
aagaagcccc	ggctaactac	gtgccagcag	ccgcggtaat	acgtaggggg	caagcgttat	540
ccggatttac	tgggtgtaaa	gggagcgtag	acggcacggc	aagccagatg	tgaaagcccc	600
ggctcaaccc	cgggactgca	tttggaactg	ctgagctaga	gtgtcggaga	ggcaagtgga	660
attcctagtg	tagcgggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggcggc	720
ttgctggacg	atgactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780
cctggtagtc	cacgccgtaa	acgatgactg	ctaggtgtcg	ggtaggcaaag	ccattcgggtg	840
ccgcagctaa	cgcaataagc	agtccacctg	gggagtacgt	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgggtg	agcatgtggt	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	tgatcttgac	atcccgatga	ccgcttcgta	atggaagttt	ttcttcggaa	1020
catcggtgac	aggtgggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgaacga	gcgcaacccc	tatcttcagt	agccagcagg	ttaagctggg	cactctggag	1140
agactgccag	ggataacctg	gaggaagggtg	gggatgacgt	caaatacatca	tgccccttat	1200
gaccagggct	acacacgtgc	tacaatggcg	taaacaaaga	gaggcaaact	cgcgagggtta	1260
agcaaatctc	aaaaataacg	tctcagttcg	gattgtagtc	tgcaactcga	ctacatgaag	1320
ctggaatcgc	tagtaatcgc	agatcagaat	gctgcgggtga	atacgttccc	gggtcttgta	1380
cacaccgccc	gtcacaccat	gggagtcagt	aacgcccga	gtcagtgacc	caaccgtaag	1440
gagggagctg	ccgaagggtg	gaccgataac	tgggggtgaag	tcgtaacaag	gtagccgtat	1500
cggaagggtgc	ggctggatca	cctccttt				1528

<210> 151
 <211> 1529
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 151	
aacgagagtt	tgatcctggc
tcaggatgaa	cgctggcggc
gtgcttaaca	catgcaagtc
	60
gagcgaagcg	ctttgggaag
attcttcgga	tgatttcctt
tgtgactgag	cggcggacgg
	120
gtgagtaacg	cgtgggtaac
ctgcctcata	cagggggata
acagttagaa	atgactgcta
	180
ataccgcata	agaccacggt
accgcatggt	acagtggtaa
aaactccggt	ggtatgagat
	240
ggacccgcgt	ttgattaggt
agttggtggg	gtaacggcct
accaagccga	cgatcagtag
	300
ccgacctgag	agggtgaccg
gccacattgg	gactgagaca
cggcccagac	tcctacggga
	360
ggcagcagtg	gggaatattg
cacaatggag	gaaactctga
tgacgagcag	ccgcgtgaag
	420
gatgaagtat	ttcgggtatgt
aaacttctat	cagcagggaa
gaaaatgacg	gtacctgact
	480
aagaagcccc	ggctaactac
gtgccagcag	ccgcggtaat
acgtaggggg	caagcgttat
	540

P074570006W000-SEQLISTING-CEW.txt

ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccc	600
gggctcaacc ccgggactgc atttgaact gctgagctag agtgtcggag aggcaagtgg	660
aattcctagt gtagcggtag aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg	720
cttgctggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgccgta aacgatgact gctaggtgtc ggggtggcaaa gccattcggg	840
gccgcagcta acgcaataag cagtccacct ggggagtacg ttcgcaagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcgggtg gagcatgtgg ttttaattcga agcaacgcga	960
agaaccttac ctgatcttga catcccgatg actgcttcgt aatggaagtt tttcttcgga	1020
acatcggtga cagggtggtgc atgggtgtcg tcagctcgtg tcgtgagatg ttgggttaag	1080
tcccgcacag agcgcacccc ctatcttcag tagccagcag gtttaagctgg gcaactctgga	1140
gagactgcca gggataacct ggaggaaggt ggggatgacg tcaaatacatc atgcccctta	1200
tgaccagggc tacacacgtg ctacaatggc gtaaacaag agaaagcgaac tcgcgagggg	1260
aagcaaatct caaaaataac gtctcagttc ggattgtagt ctgcaactcg actacatgaa	1320
gctggaatcg ctagtaatcg cagatcagaa tgctgcgggtg aatacgttcc cgggtcttgt	1380
acacaccgcc cgtcacacca tgggagtcag taacgcccga agtcagtgc ccaaccgtaa	1440
ggagggagct gccgaagggt ggaccgataa ctgggggtgaa gtcgtaacaa ggtagccgta	1500
tcggaagggtg cggctggatc acctccttt	1529

<210> 152

<211> 1537

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 152

atggagagtt tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgtaggtaac ctgcccattg gtccgggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga gcgcatgctc agtatattaa agcgcacatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tgggtaggta acggctcacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaaccggc acattgggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagttaa	420
gaaggctctc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctta ccagaaagcc acggctaact acgtgccagc agccgcggta	540
atacgtaggt ggcaagcgtt atccggaatc attgggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctcaa ccattgtaag ctatggaaac tggatatgctg	660
gagtcagaaa gagggcgatg gaattccatg tgtagcggta aaatgcgtag atatatggag	720

P074570006W000-SEQLISTING-CEW.txt

gaacaccagt	ggcgaaggcg	gtcgcctggg	ctgtaactga	caactgaggca	cgaaagcgtg	780
gggagcaa	aggattagat	accctagtag	tccacgccgt	aaacgatgag	aactaagtgt	840
tggaggaatt	cagtgtctga	gttaacgcaa	taagttctcc	gcctggggag	tatgcacgca	900
agtgtgaaac	tcaaaggaat	tgacgggggc	ccgcacaagc	ggtggagtat	gtggttta	960
tcgaagcaac	gcgaagaacc	ttaccaggcc	ttgacatgga	aacaaatacc	ctagagatag	1020
ggggataatt	atggatcaca	caggtggtgc	atggttgtcg	tcagctcgtg	tcgtgagatg	1080
ttgggttaag	tcccgcacg	agcgcaacc	ttgtcgcgtg	ttaccagcat	caagttgggg	1140
actcatgcga	gactgccggt	gacaaaccgg	aggaagggtg	ggatgacgtc	aaatcatcat	1200
gccccttatg	gcctgggcta	cacacgtact	acaatggcgg	ccacaaagag	cagcgacaca	1260
gtgatgtgaa	gcgaatctca	taaaggctcg	ctcagttcgg	attgaagtct	gcaactcgac	1320
ttcatgaagt	cggaatcgct	agtaatcgca	gatcagcatg	ctgcggtgaa	tacgttctcg	1380
ggccttgtac	acaccgccc	tcaaaccatg	ggagtcagta	ataccgaag	ccggtggcat	1440
aaccgtaagg	agtgagccgt	cgaaggtagg	accgatgact	ggggttaagt	cgtaacaagg	1500
tatccctacg	ggaacgtggg	gatggatcac	ctccttt			1537

<210> 153

<211> 1537

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 153

atggagagtt	tgatcctggc	tcaggatgaa	cgctggcggc	atgcctaata	catgcaagtc	60
gaacgaagtt	tcgaggaagc	ttgcttccaa	agagacttag	tggcgaacgg	gtgagtaaca	120
cgtaggtaac	ctgcccattg	gtccgggata	actgctggaa	acggtagcta	aaaccggata	180
ggtatacaga	gcgcatgctc	agtatattaa	agcgcccatc	aaggcgtgaa	catggatgga	240
cctgcggcgc	attagctagt	tggtgaggta	acggcccacc	aaggcgatga	tgcgtagccg	300
gcctgagagg	gtaaaccggc	acattgggac	tgagacacgg	cccaaactcc	tacgggaggc	360
agcagtaggg	aattttcgtc	aatgggggaa	accctgaacg	agcaatgccg	cgtgagttaa	420
gaaggctctc	ggatcgtaaa	gctctgttgt	aagtgaagaa	cggctcatag	aggaaatgct	480
atgggagtga	cggtagctta	ccagaaagcc	acggctaact	acgtgccagc	agccgcggta	540
atacgtaggt	ggcaagcggt	atccggaatc	attgggcgta	aagggtgcgt	aggtggcgta	600
ctaagtctgt	agtaaaaggc	aatggctcaa	ccattgtaag	ctatggaaac	tggtatgctg	660
gagtgcagaa	gagggcgatg	gaattccatg	tgtagcggta	aaatgcgtag	atatatggag	720
gaacaccagt	ggcgaaggcg	gtcgcctggg	ctgtaactga	caactgaggca	cgaaagcgtg	780
gggagcaa	aggattagat	accctagtag	tccacgccgt	aaacgatgag	aactaagtgt	840
tggaggaatt	cagtgtctga	gttaacgcaa	taagttctcc	gcctggggag	tatgcacgca	900

P074570006W000-SEQLISTING-CEW.txt

agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat	960
tcgaagcaac gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg	1080
ttgggttaag tcccgcaacg agcgcaaccc ttgtcgcacg ttaccagcat caagttgggg	1140
actcatgcga gactgccggt gacaaaccgg aggaagggtg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcgg ccacaaagag cagcgacaca	1260
gtgatgtgaa gcgaatctca taaaggctcg ctcagttcgg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcggtgaa tacgttctcg	1380
ggccttgtac acaccgcccg tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440
aaccgtaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccttt	1537

<210> 154
 <211> 1537
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 154	
atggagagtt tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgtaggtaac ctgcccattg gtccgggata actgctggaa acggtagcta aaaccggata	180
ggatatacaga gcgcatgctc agtatattaa agcgcccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtaggta acggctcacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaacggcc acattgggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagttaa	420
gaaggtcttc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctta ccagaaagcc acggctaact acgtgccagc agccgcggta	540
atacgtaggt ggcaagcggt atccggaatc attgggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatgggtcaa ccattgtaag ctatggaaac tggatatgctg	660
gagtgcagaa gagggcgatg gaattccatg ttagcggta aaatgcgtag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctggt ctgtaactga cactgaggca cgaaagcgtg	780
gggagcaaatt aggattagat accctagtag tccacgccgt aaacgatgag aactaagtgt	840
tggaggaatt cagtgtgca gttaacgcaa taagttctcc gcctggggag tatgcacgca	900
agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat	960
tcgaagcaac gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg	1080

P074570006W000-SEQLISTING-CEW.txt

ttgggttaag tcccgcaacg agcgcaaccc ttgtcgcatg ttaccagcat caagttgggg	1140
actcatgcga gactgccggt gacaaaccgg aggaaggtgg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcgg ccacaaagag cagcgacaca	1260
gtgatgtgaa gcgaatctca taaaggtcgt ctcagttcgg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcggtgaa tacgtttctcg	1380
ggccttgtac acaccgcccg tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440
aaccgtaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccttt	1537

<210> 155
 <211> 1537
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 155	
atggagagtt tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgtaggtaac ctgcccattg gtccgggata actgctggaa acggtagcta aaaccggata	180
ggatatacaga gcgcatgctc agtatattaa agcgcccatc aaggcgtgaa catggatgga	240
cctgcgggcg attagctagt tggtaggta acggcccacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaacggcc acattgggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagttaa	420
gaaggtcttc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctta ccagaaagcc acggctaact acgtgccagc agccgcggta	540
atacgtaggt ggcaagcgtt atccggaatc attgggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctcaa ccattgtaag ctatggaaac tggtagctg	660
gagtcagaa gagggcgatg gaattccatg ttagcggta aaatgcgtag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctggc ctgtaactga cactgaggca cgaaagcgtg	780
gggagcaaatt aggattagat accctagtag tccacgccgt aaacgatgag aactaagtgt	840
tggaggaatt cagtgtgca gttaacgaa taagttctcc gcctggggag tatgcacgca	900
agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat	960
tcgaagcaac gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atggttgtcg tcagctcgtg tcgtgagatg	1080
ttgggttaag tcccgcaacg agcgcaaccc ttgtcgcatg ttaccagcat caagttgggg	1140
actcatgcga gactgccggt gacaaaccgg aggaaggtgg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcga ccacaaagag cagcgacaca	1260

P074570006W000-SEQLISTING-CEW.txt

gtgatgtgaa gcgaatctca taaaggctcg ctcagttcgg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcggtgaa tacgttctcg	1380
ggccttgtac acaccgcccg tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440
aaccgcaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccttt	1537

<210> 156
 <211> 1537
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Polynucleotide

<400> 156	
atggagagtt tgatcctggc tcaggatgaa cgctggcggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgtaggtaac ctgcccattgt gtccgggata actgctggaa acggtagcta aaaccggata	180
ggatatacaga gcgcatgctc agtatattaa agcgcccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtagagta acggcccacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaacggcc acattgggac tgagacacgg cccaaactcc tacgggagggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgccg cgtgagttaa	420
gaaggctctc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctta ccagaaagcc acggctaact acgtgccagc agccgcggta	540
atacgtaggt ggcaagcgtt atccggaatc attgggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctcaa ccattgtaag ctatggaaac tggtagctg	660
gagtgacaga gagggcgatg gaattccatg ttagcggta aaatgcgtag atatattggag	720
gaacaccagt ggcgaaggcg gtcgcctggc ctgtaactga cactgaggca cgaaagcgtg	780
gggagcaaatt aggttagat accctagtag tccacgccgt aaacgatgag aactaagtgt	840
tggaggaatt cagtgtgca gttacgcaa taagttctcc gcctggggag tatgcacgca	900
agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat	960
tcgaagcaac gcgaagaacc ttaccaggcc ttgacatgga aacaaatacc ctagagatag	1020
ggggataatt atggatcaca cagggtggtg atggttgtcg tcagctcgtg tcgtgagatg	1080
ttgggttaag tcccgcaacg agcgcaaccc ttgtcgcatt ttaccagcat caagttgggg	1140
actcatgcga gactgccggg gacaaaccgg aggaagggtg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcgg ccacaaagag cagcgacaca	1260
gtgatgtgaa gcgaatctca taaaggctcg ctcagttcgg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcggtgaa tacgttctcg	1380
ggccttgtac acaccgcccg tcaaaccatg ggagtcagta ataccgaag ccggtggcat	1440

aaccgtaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg 1500
tatccctacg ggaacgtggg gatggatcac ctccttt 1537

<210> 157
<211> 1529
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Polynucleotide

<400> 157
tattgagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt 60
cgaacggggt gtcacatgac gaggattcgt ccaacggatt gagttacctg gtggcggacg 120
ggtgagtaac gcgtgaggaa cctgccttgg agaggggaat aacactccga aaggagtgtc 180
aataccgcat gatgcagttg ggtcgcacat ctctgactgc caaagattta tcgctctgag 240
atggcctcgc gtctgattag ctagtaggcg gggtaacggc ccacctaggc gacgatcagt 300
agccggactg agaggttgac cggccacatt gggactgaga cacggcccag actcctacgg 360
gaggcagcag tgggggaatat tgggcaatgg gcgcaagcct gaccagcaa cgccgcgtga 420
aggaagaagg ctttcggggt gtaaacttct tttgtcgggg acgaaacaaa tgacgggtacc 480
cgacgaataa gccacggcta actacgtgcc agcagccgcg gtaatacgtg ggtggcaagc 540
gttatccgga tttactgggt gtaaagggcg ttagggcggg attgcaagtc agatgtgaaa 600
actgggggct caacctccag cctgcatttg aaactgtagt tcttgagtgc tggagaggca 660
atcggaattc cgtgtgtagc ggtgaaatgc gtagatatac ggaggaacac cagtggcgaa 720
ggcggattgc tggacagtaa ctgacgtgta ggcgcgaaag cgtgggggagc aaacaggatt 780
agataccctg gtagtccacg ccgtaaacga tggatactag gtgtgggggt ctgacccct 840
ccgtgccgca gtaaacacaa taagtatccc acctggggag tacgatcgca aggttgaaac 900
tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtgggtttaat tcgaagcaac 960
gcgaagaacc ttaccagggc ttgacatccc actaacgaag cagagatgca ttaggtgccc 1020
ttcggggaaa gtggagacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg 1080
ggtaagtcc cgcaacgagc gcaaccctta ttgttagttg ctacgcaaga gcaactctagc 1140
gagactgccg ttgacaaaac ggaggaaggt ggggacgacg tcaaatacatc atgcccctta 1200
tgtcctgggc cacacacgta ctacaatggg ggttaacaga gggaggcaat accgcgaggt 1260
ggagcaaata cctaaaagcc atcccagttc ggattgcagg ctgaaacccg cctgtatgaa 1320
gttggaatcg ctagtaatcg cggatcagca tgccgcggtg aatacgttcc cgggccttgt 1380
acacaccgcc cgtcacacca tgagagtcgg gaacacccga agtccgtagc ctaaccgcaa 1440
ggagggcgcg gccgaaggtg ggttcgataa ttgggggtgaa gtcgtaacaa ggtagccgta 1500
tcggaaggtg cggctggatc acctccttt 1529

<210> 158

<211> 1530

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

<400> 158

tattgagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt	60
cgaacgggggt gctcatgacg gaggattcgt ccaacggatt gagttacctt gtggcgggacg	120
ggtgagtaac gcgtgaggaa cctgccttgg agaggggaat aacactccga aaggagtgtc	180
aataccgcat gatgcagttg ggtcgcattg ctctgactgc caaagattta tcgctctgag	240
atggcctcgc gtctgattag ctagtaggcg gggtaacggc ccacctaggc gacgatcagt	300
agccggactg agaggttgac cggccacatt gggactgaga cacggcccag actcctacgg	360
gaggcagcag tggggaatat tgggcaatgg gcgcaagcct gaccagcaa cgccgcgtga	420
aggaagaagg ctttcgggtt gtaaacttct tttgtcgggg acgaaacaaa tgacgggtacc	480
cgacgaataa gccacggcta actacgtgcc agcagccgcg gtaatacgtg ggtggcaagc	540
gttatccgga tttactgggt gtaaagggcg tgtaggcggg attgcaagtc agatgtgaaa	600
actgggggct caacctccag cctgcatttg aaactgtagt tcttgagtgc tggagaggca	660
atcggaattc cgtgtgtagc ggtgaaatgc gtagatatac ggaggaacac cagtggcgaa	720
ggcggattgc tggacagtaa ctgacgctga ggcgcgaaag cgtgggggagc aaacaggatt	780
agataccctg gtagtccacg ccgtaaacga tggatactag gtgtgggggg tctgaccccc	840
tccgtgccgc agttaacaca ataagtatcc cacctgggga gtacgatcgc aaggttgaaa	900
ctcaaaggaa ttgacggggg ccgcacaaag cggtaggagta tgtgggtttaa ttcgaagcaa	960
cgcaagaac cttaccaggg cttgacatcc cactaacgaa gcagagatgc attaggtgcc	1020
cttcggggaa agtggagaca ggtggtgcat ggttgtcgtc agctcgtgtc gtgagatgtt	1080
gggttaagtc ccgcaacgag cgcaaccctt attgttagtt gctacgcaag agcactctag	1140
cgagactgcc gttgacaaaa cggaggaagg tggggacgac gtcaaatacat catgccctt	1200
atgtcctggg ccacacacgt actacaatgg tggtaaacag agggaggcaa taccgcgagg	1260
tggagcaaat ccctaaaagc catcccagtt cggattgcag gctgaaaccc gcctgtatga	1320
agttggaatc gctagtaatc gcggatcagc atgccgcggg gaatacgttc ccgggccttg	1380
tacacaccgc ccgtcacacc atgagagtcg ggaacacccg aagtcgtag cctaaccgca	1440
aggagggcgc ggccgaaggt gggttcgata attgggggtga agtcgtaaca aggtagccgt	1500
atcggaaggt gcggctggat cacctccttt	1530

<210> 159

<211> 1529

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Polynucleotide

P074570006W000-SEQLISTING-CEW.txt

```

<400> 159
tattgagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt      60
cgaacgggggt gctcatgacg gaggattcgt ccaacggatt gagttacctt gtggcgggacg      120
ggtgagtaac gcgtgaggaa cctgccttgg agaggggaat aacactccga aaggagtgtc      180
aataccgcat aatgcagttg ggtcgcattg ctctgactgc caaagattta tcgctctgag      240
atggcctcgc gtctgattag ctagtaggcg gggtaacggc ccacctaggc gacgatcagt      300
agccggactg agaggttgac cggccacatt gggactgaga cacggcccag actcctacgg      360
gaggcagcag tggggaatat tgggcaatgg gcgcaagcct gacctagcaa cgccgcgtga      420
aggaagaagg ctttcgggtt gtaaacttct tttgtcaggg acgaaacaaa tgacgggtacc      480
tgacgaataa gccacggcta actacgtgcc agcagccgcg gtaatacgta ggtggcaagc      540
gttatccgga tttactgggt gtaaagggcg ttaggcggg attgcaagtc agatgtgaaa      600
actgggggct caacctccag cctgcatttg aaactgtagt tcttgagtgc tggagaggca      660
atcggaattc cgtgtgtagc ggtgaaatgc gtagatatac ggaggaacac cagtggcgaa      720
ggcggattgc tggacagtaa ctgacgctga ggcgcgaaag cgtgggggagc aaacaggatt      780
agataccctg gtagtccacg ccgtaaacga tggatactag gtgtgggggg tctgaccctt      840
ccgtgccgca gttaacacaa taagtatccc acctggggag tacgatcgca aggttgaaac      900
tcaaaggaat tgacgggggc ccgcacaagc ggtggagtat gtggtttaat tcgaagcaac      960
gcgaagaacc ttaccagggc ttgacatccc actaacgaag cagagatgca ttaggtgccc     1020
ttcggggaaa gtggagacag gtggtgcatg gttgtcgtca gctcgtgtcg tgagatgttg     1080
ggttaagtcc cgcaacgagc gcaaccctta ttgttagttg ctacgcaaga gcaactctagc     1140
gagactgccg ttgacaaaac ggaggaaggt ggggacgacg tcaaatacatc atgcccctta     1200
tgtcctgggc cacacacgta ctacaatggg ggttaacaga gggaggcaat accgcgaggt     1260
ggagcaaata cctaaaagcc atcccagttc ggattgcagg ctgaaaccgc cctgtatgaa     1320
gttggaatcg ctagtaatcg cggatcagca tgccgcggtg aatacgttcc cgggccttgt     1380
acacaccgcc cgtcacacca tgagagtcgg gaacaccgga agtccgtagc ctaaccgcaa     1440
ggagggcgcg gccgaagggt ggttcgataa ttgggggtgaa gtcgtaacaa ggtagccgta     1500
tcggaagggt cggtctggatc acctcctttt                                     1529

```