



US010555980B2

(12) **United States Patent**  
**Schneider et al.**

(10) **Patent No.:** US 10,555,980 B2  
(45) **Date of Patent:** \*Feb. 11, 2020

(54) **TREATMENT OF CLOSTRIDIUM DIFFICILE INFECTION**

(71) Applicant: **Vedanta Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Jessica Schneider**, Cambridge, MA (US); **Yun-Gi Kim**, Watertown, MA (US); **Bernat Olle**, Cambridge, MA (US); **Shilpa Reddy**, Watertown, MA (US); **Jason Norman**, North Weymouth, MA (US); **Juan Patarroyo**, Lexington, MA (US)

(73) Assignee: **Vedanta Biosciences, Inc.**, Cambridge, MA (US)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **16/423,487**

(22) Filed: **May 28, 2019**

(65) **Prior Publication Data**

US 2019/0275090 A1 Sep. 12, 2019

**Related U.S. Application Data**

(63) Continuation of application No. 16/157,640, filed on Oct. 11, 2018, now Pat. No. 10,456,431, which is a continuation of application No. 15/993,037, filed on May 30, 2018, now Pat. No. 10,350,250, which is a continuation of application No. 15/630,088, filed as application No. PCT/US2017/037498 on Jun. 14, 2017, now Pat. No. 9,999,641.

(60) Provisional application No. 62/349,914, filed on Jun. 14, 2016.

(51) **Int. Cl.**

<b>A61K 35/747</b>	(2015.01)
<b>A23L 5/00</b>	(2016.01)
<b>A23L 33/135</b>	(2016.01)
<b>A61K 9/00</b>	(2006.01)
<b>A61K 9/48</b>	(2006.01)
<b>A61K 35/74</b>	(2015.01)
<b>A61K 35/742</b>	(2015.01)
<b>C12N 1/20</b>	(2006.01)
<b>A23K 50/30</b>	(2016.01)
<b>A23K 10/18</b>	(2016.01)

(52) **U.S. Cl.**

CPC .....	<b>A61K 35/747</b> (2013.01); <b>A23L 5/00</b> (2016.08); <b>A23L 33/135</b> (2016.08); <b>A61K 9/0053</b> (2013.01); <b>A61K 9/48</b> (2013.01); <b>A61K 35/74</b> (2013.01); <b>A61K 35/742</b> (2013.01); <b>C12N 1/20</b> (2013.01); <b>A23K 10/18</b> (2016.05); <b>A23K 50/30</b> (2016.05); <b>A61K 9/0031</b> (2013.01)
-----------	---

(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

6,635,260 B1	10/2003	Gerding
8,460,648 B2	6/2013	Borody
9,386,793 B2	7/2016	Blaser et al.
9,642,881 B2	5/2017	Honda et al.
9,649,345 B2	5/2017	Honda et al.
9,999,641 B2 *	6/2018	Schneider .....
10,064,904 B2 *	9/2018	Schneider .....
10,350,250 B2 *	7/2019	Schneider .....
2010/0074872 A1	3/2010	Blaser et al.
2013/0195804 A1	8/2013	Borody
2014/0199281 A1	7/2014	Henn et al.
2014/0328803 A1	11/2014	McKenzie et al.
2014/0341921 A1	11/2014	Honda et al.
2015/0037476 A1	2/2015	Dhingra et al.
2015/0079209 A1	3/2015	Kameyama et al.
2016/0022745 A1	1/2016	Wang
2016/0022746 A1	1/2016	Lawley et al.
2016/0040215 A1	2/2016	Henn et al.
2016/0113971 A1	4/2016	Kaplan et al.
2016/0193256 A1	7/2016	Honda et al.
2016/0193257 A1	7/2016	Honda et al.
2016/0228476 A1	8/2016	Cutcliffe et al.
2017/0143772 A1	5/2017	Mulder et al.
2017/0209502 A1	7/2017	Honda et al.
2017/0216378 A1	8/2017	Honda et al.
2017/0354697 A1	12/2017	Schneider et al.
2018/0264056 A1	9/2018	Schneider et al.
2019/0030098 A1	1/2019	Schneider et al.

**FOREIGN PATENT DOCUMENTS**

WO	WO 2002/007741 A1	1/2002
WO	WO 2006/050479 A1	5/2006
WO	WO 2011/152566 A1	12/2011
WO	WO 2013/037067 A1	3/2013
WO	WO 2013/037068 A1	3/2013
WO	WO 2013/080561 A1	6/2013
WO	WO 2013/182038 A1	12/2013
WO	WO 2014/082050 A1	5/2014
WO	WO 2014/121298 A1	8/2014
WO	WO 2014/121301 A1	8/2014
WO	WO 2014/121302 A1	8/2014
WO	WO 2014/145958 A2	9/2014
WO	WO 2014/153194 A2	9/2014

(Continued)

**OTHER PUBLICATIONS**

Apisarthanarak et al. 2002 (Adjunctive Intracolonic Vancomycin for Severe Clostridium difficile Colitis: Case Series and Review of the Literature: CID 2002:35). (Year: 2002).\*

(Continued)

*Primary Examiner — Mary Maille Lyons*

(74) *Attorney, Agent, or Firm — Wolf, Greenfield & Sacks, P.C.*

(57) **ABSTRACT**

Provided herein are compositions and methods for the treatment or prevention of pathogenic infections.

**20 Claims, 41 Drawing Sheets**

**Specification includes a Sequence Listing.**

(56)

**References Cited**

## FOREIGN PATENT DOCUMENTS

WO	WO 2015/006355	A2	1/2015
WO	WO 2015/051323	A1	4/2015
WO	WO 2015/077794	A1	5/2015
WO	WO 2015/095241	A2	6/2015
WO	WO 2015/156419	A1	10/2015
WO	WO 2015/164555	A1	10/2015
WO	WO 2015/179437	A1	11/2015
WO	WO 2016/086161	A1	6/2016
WO	WO 2016/086205	A2	6/2016
WO	WO 2016/086206	A1	6/2016
WO	WO 2016/086208	A1	6/2016
WO	WO 2016/086209	A1	6/2016
WO	WO 2016/086210	A1	6/2016
WO	WO 2016/185469	A1	11/2016
WO	WO 2016/201053	A1	12/2016
WO	WO 2016/203217	A1	12/2016
WO	WO 2016/203218	A1	12/2016
WO	WO 2016/203220	A1	12/2016
WO	WO 2016/203221	A1	12/2016
WO	WO 2016/203223	A1	12/2016
WO	WO 2016/209806	A1	12/2016
WO	WO 2017/008026	A1	1/2017
WO	WO 2017/035188	A1	3/2017
WO	WO 2017/075098	A1	5/2017
WO	WO 2017/085518	A1	5/2017
WO	WO 2017/085520	A1	5/2017
WO	WO 2017/089794	A1	6/2017
WO	WO 2017/089795	A1	6/2017
WO	WO 2017/091783	A2	6/2017
WO	WO 2017/148596	A1	9/2017

## OTHER PUBLICATIONS

- Alang et al., Weight gain after fecal microbiota transplantation. Open Forum Infect Dis. Feb. 4, 2015;2(1):ofv004. doi: 10.1093/ofid/ofv004. eCollection Jan. 2015.
- Blaser, The microbiome revolution. J Clin Invest. Oct. 2014;124(10):4162-5. doi: 10.1172/JCI78366. Epub Oct. 1, 2014.
- Borody et al., Therapeutic faecal microbiota transplantation: current status and future developments. Curr Opin Gastroenterol. Jan. 2014;30(1):97-105. doi: 10.1097/MOG.0000000000000027.
- Buffie et al., Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. Nature. Jan. 8, 2015;517(7533):205-8. doi: 10.1038/nature13828. Epub Oct. 22, 2014.
- Burns et al., Donor Recruitment and Eligibility for Fecal Microbiota Transplantation: Results From an International Public Stool Bank. Gastro. Apr. 2015;148(4):S96-S97.
- Calfee, *Clostridium difficile*: a reemerging pathogen. Geriatrics. Sep. 1, 2008;63(9):10-21.
- Cammarota et al., Randomised clinical trial: faecal microbiota transplantation by colonoscopy vs. vancomycin for the treatment of recurrent *Clostridium difficile* infection. Aliment Pharmacol Ther. May 2015;41(9):835-43. doi: 10.1111/apt.13144. Epub Mar. 1, 2015.
- Drancourt et al., 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol. Oct. 2000;38(10):3623-30.
- Eyre et al., Whole-genome sequencing demonstrates that fidaxomicin is superior to vancomycin for preventing reinfection and relapse of infection with *Clostridium difficile*. J Infect Dis. May 1, 2014;209(9):1446-51. doi:10.1093/infdis/jit598. Epub Nov. 11, 2013. Genbank Accession No. NR\_104687.1. NCBI. Sakamoto. Feb. 3, 2015.
- Hooper et al., Interactions between the microbiota and the immune system. Science. Jun. 8, 2012;336(6086):1268-73. doi:10.1126/science.1223490. Epub Jun. 6, 2012.
- Hughes et al., Immune activation in irritable bowel syndrome: can neuroimmune interactions explain symptoms? Am J Gastroenterol. Jul. 2013;108(7):1066-74. doi: 10.1038/ajg.2013.120. Epub May 7, 2013.
- Kakihana et al., Fecal microbiota transplantation for patients with steroid-resistant acute graft-versus-host disease of the gut. Blood. Oct. 20, 2016;128(16):2083-2088. doi: 10.1182/blood-2016-05-717652. Epub Jul. 26, 2016.
- Kassam et al., Fecal microbiota transplantation for *Clostridium difficile* infection: systematic review and meta-analysis. Am J Gastroenterol. Apr. 2013;108(4):500-8. doi: 10.1038/ajg.2013.59. Epub Mar. 19, 2013.
- Khoruts et al., Emergence of fecal microbiota transplantation as an approach to repair disrupted microbial gut ecology. Immunol Lett. Dec. 2014;162(2 Pt A):77-81. doi: 10.1016/j.imlet.2014.07.016. Epub Aug. 10, 2014.
- Leblanc et al., Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol. Apr. 2013;24(2):160-8. doi: 10.1016/j.copbio.2012.08.005. Epub Aug. 30, 2012.
- Lessa et al., Burden of *Clostridium difficile* infection in the United States. N. Engl J Med. Feb. 26, 2015;372(9):825-34. doi:10.1056/NEJMoa1408913.
- Louie et al., Fidaxomicin preserves the intestinal microbiome during and after treatment of *Clostridium difficile* infection (CDI) and reduces both toxin reexpression and recurrence of CDI. Clin Infect Dis. Aug. 2012;55 Suppl 2:S132-42. doi: 10.1093/cid/cis338.
- Marvola et al., Enteric polymers as binders and coating materials in multiple-unit site-specific drug delivery systems. Eur J Pharm Sci. Feb. 1999;7(3):259-67.
- Miller, Fidaxomicin (OPT-80) for the treatment of *Clostridium difficile* infection. Expert Opin Pharmacother. Jun. 2010;11(9):1569-78. doi:10.1517/14656566.2010.485614.
- Mullane, Fidaxomicin in Clostridium difficile infection: latest evidence and clinical guidance. Ther Adv Chronic Dis. 2014 Mar;5(2):69-84. doi:10.1177/2040622313511285.
- Narushima et al., Characterization of the 17 strains of regulatory T cell-inducing human-derived Clostridia. Gut Microbes. May-Jun. 2014;5(3):333-9. doi: 10.4161/gmic.28572. Epub Mar. 18, 2014.
- Paramsothy et al., Donor Recruitment for Fecal Microbiota Transplantation. Inflamm Bowel Dis. Jul. 2015;21(7):1600-6. doi: 10.1097/MIB.0000000000000405.
- Rossi-Tamisier et al., Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. Int J Syst Evol Microbiol. Jun. 2015;65(Pt 6):1929-34. doi:10.1099/ijss.0.000161. Epub Mar. 3, 2015.
- Shannon-Lowe et al., Prevention and medical management of *Clostridium difficile* infection. BMJ. Mar. 12, 2010;340:c1296. doi:10.1136/bmj.c1296.
- Surawicz, Fecal microbiota transplantation: what we know and what we need to know. Ann Intern Med. May 5, 2015;162(9):662-3. doi: 10.7326/M15-0609.
- Tannock et al., A new macrocyclic antibiotic, fidaxomicin (OPT-80), causes less alteration to the bowel microbiota of *Clostridium difficile*-infected patients than does vancomycin. Microbiology. Nov. 2010;156(Pt 11):3354-9. doi: 10.1099/mic.0.042010-0. Epub Aug. 19, 2010.
- Van Nood et al., Duodenal infusion of donor feces for recurrent *Clostridium difficile*. N. Engl J Med. Jan. 31, 2013;368(5):407-15. doi: 10.1056/NEJMoa1205037. Epub Jan. 16, 2013.
- Youngster et al., Fecal microbiota transplant for relapsing *Clostridium difficile* infection using a frozen inoculum from unrelated donors: a randomized, open-label, controlled pilot study. Clin Infect Dis. Jun. 2014;58(11):1515-22. doi: 10.1093/cid/ciu135. Epub Apr. 23, 2014.

\* cited by examiner

Figure 1

<b>Composition A</b>	<b>Composition B</b>	<b>Composition C</b>	<b>Composition D</b>
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_coccoidea (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_botteae (XIVa)	SEQ_07 - 59 - Blautia_producta/Blautia_coccoidea (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 - Anerostipes_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 - Marinibryantia_formata/gens (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<sup>+</sup>**bolding indicates strains other than Clostridium cluster XIVa**

**Figure 2**  
**Clostridium difficile Infection Mouse Models**

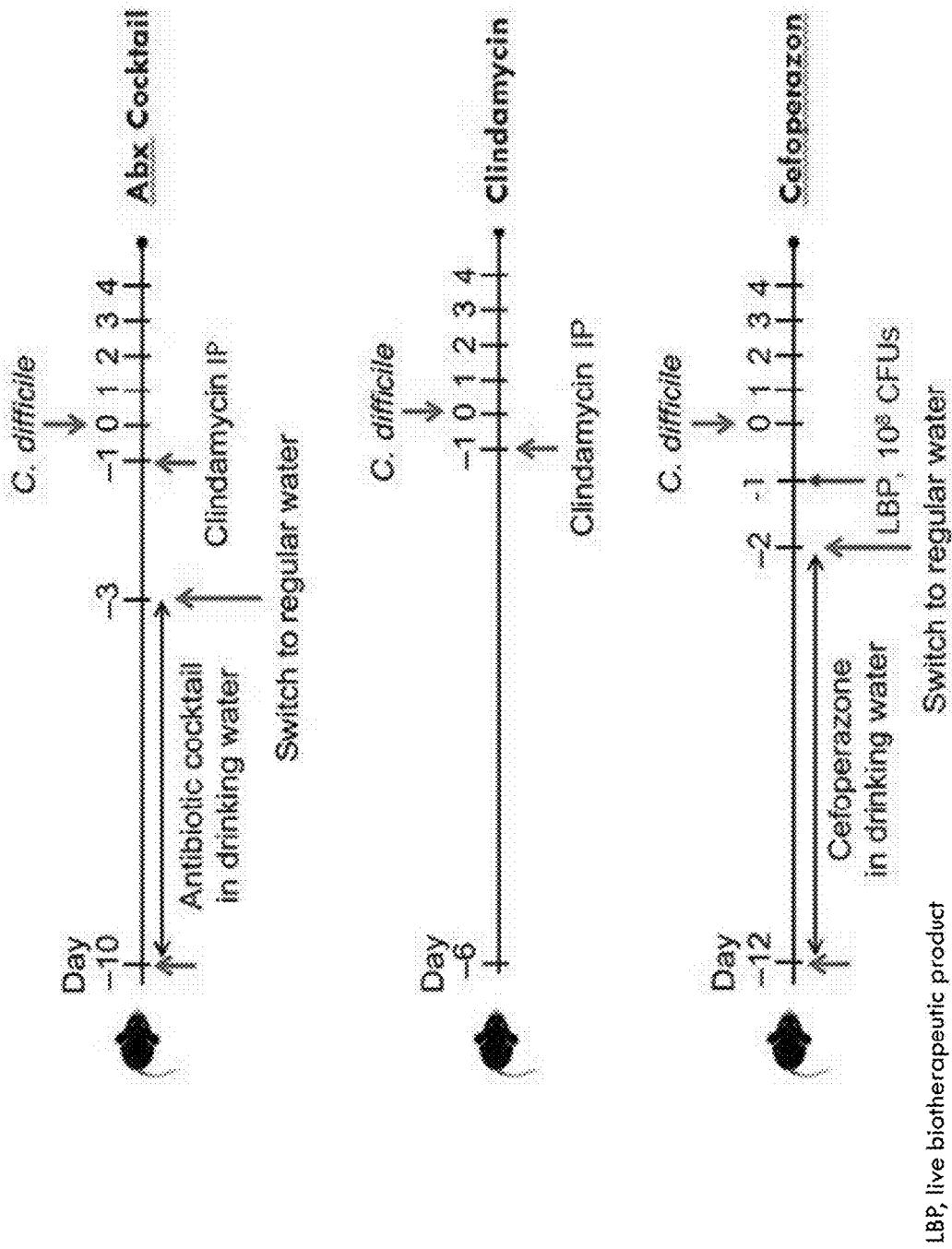


Figure 3

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

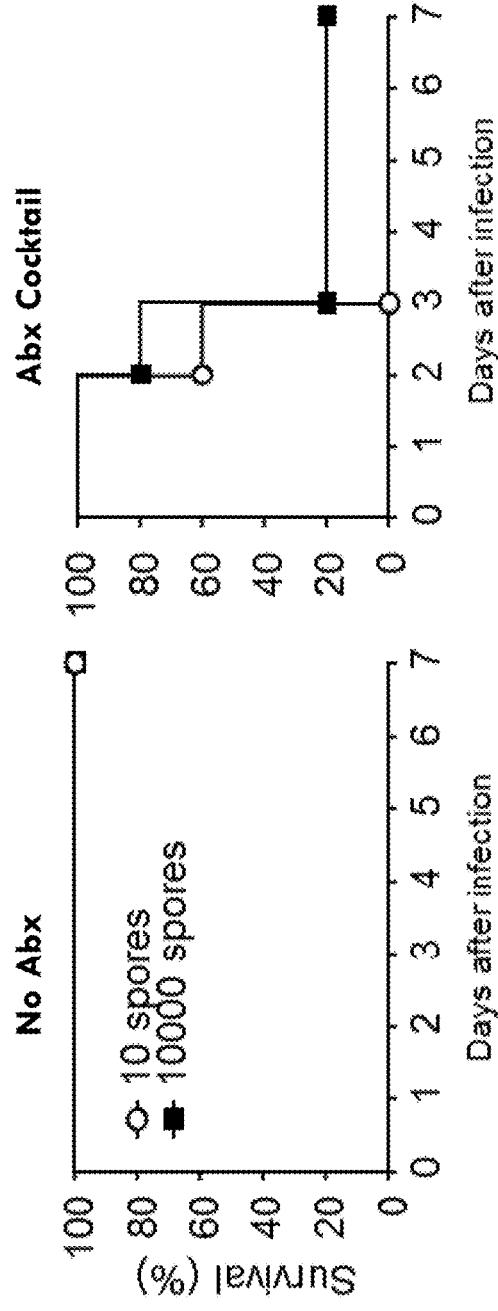


Figure 4B

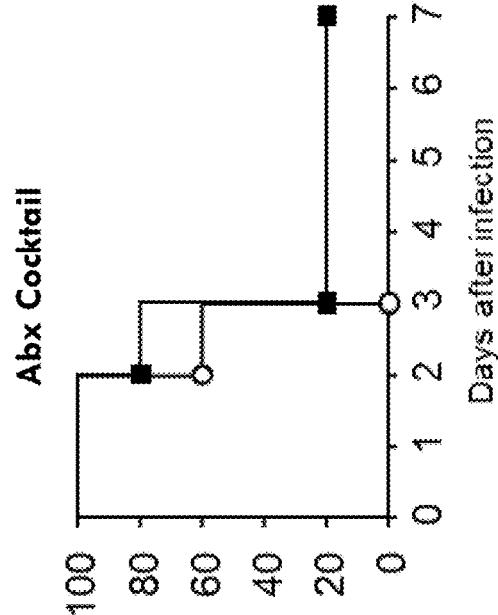


Figure 4C

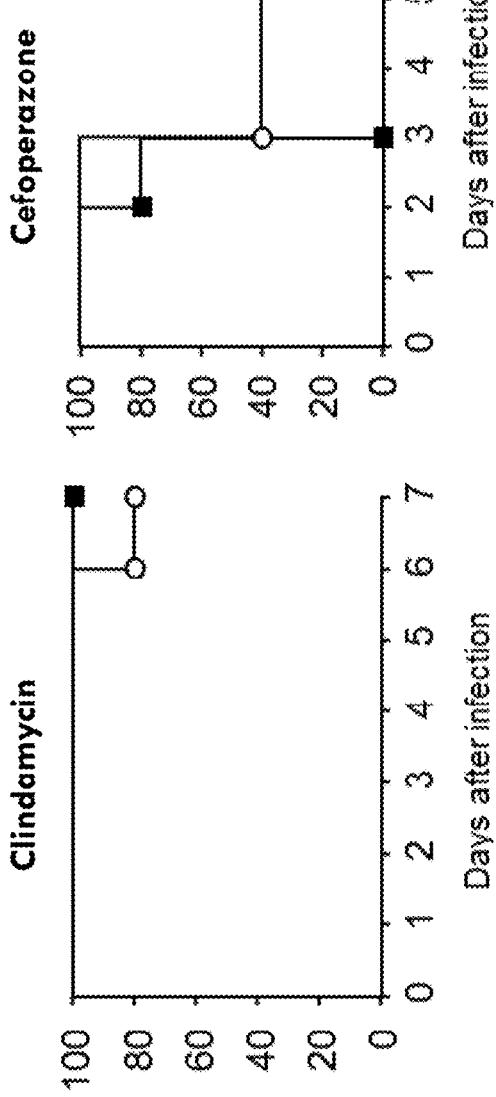


Figure 4D

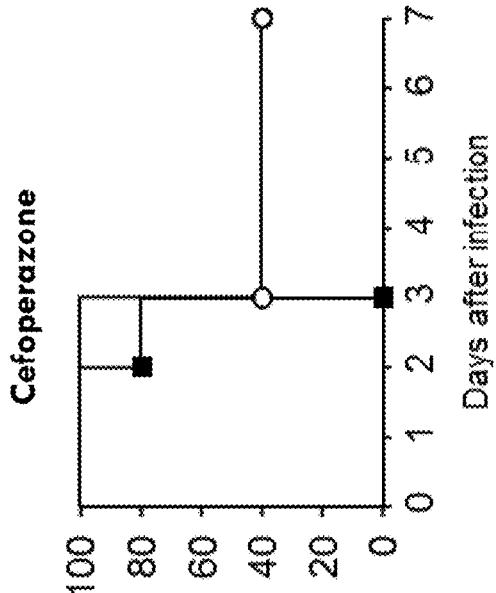


Figure 4E

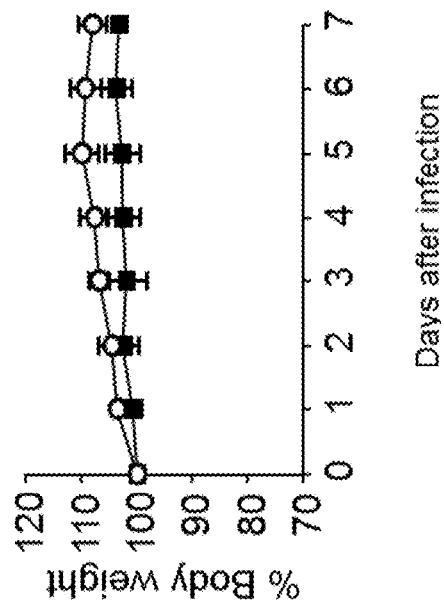
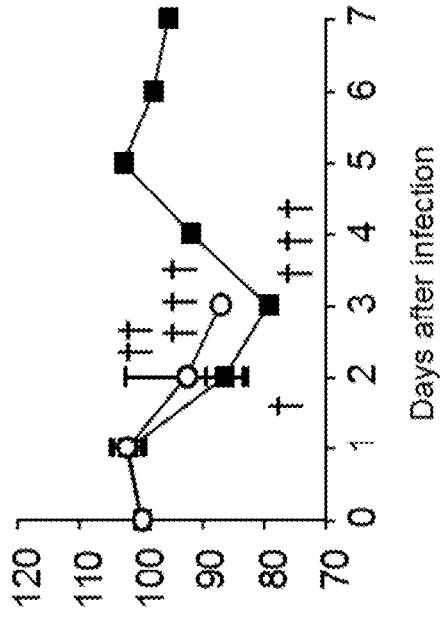
**No Abx****Abx Cocktail**

Figure 4G

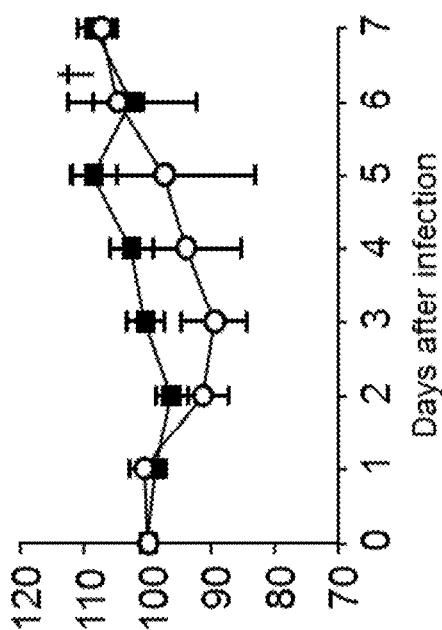
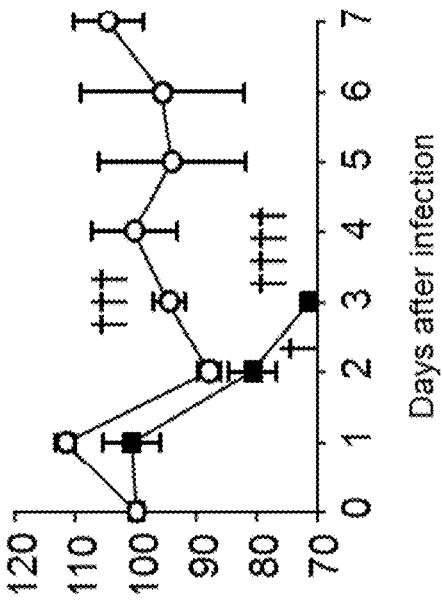
**Clindamycin**

Figure 4H

**Cefoperazone**

Days after infection

Days after infection

Days after infection

Figure 4I

No Abx

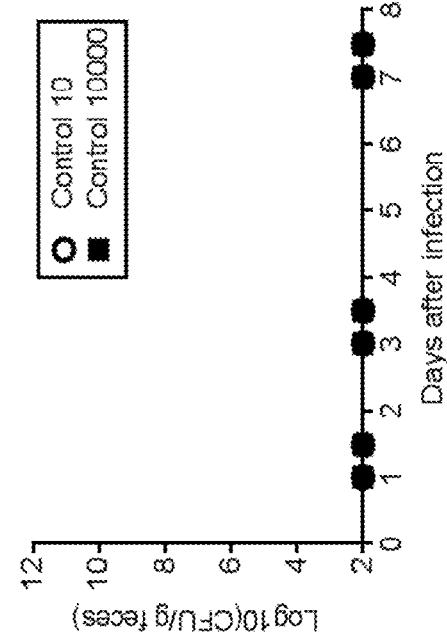


Figure 4J

Abx Cocktail

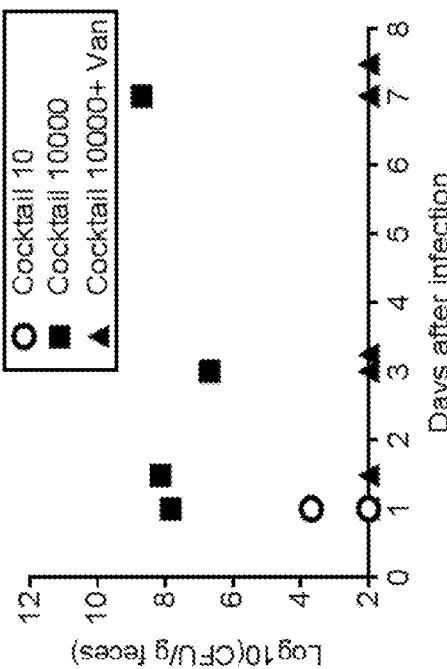


Figure 4K

Clindamycin

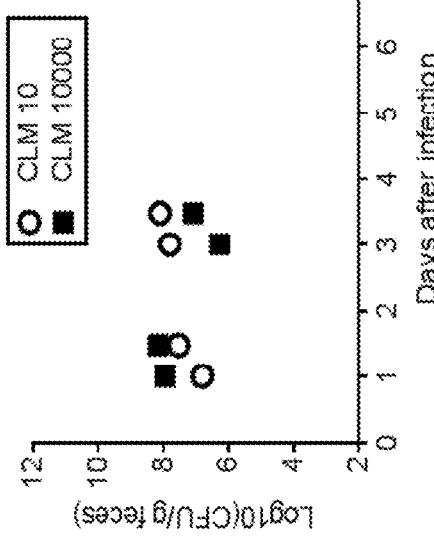


Figure 4L

Cefoperazone

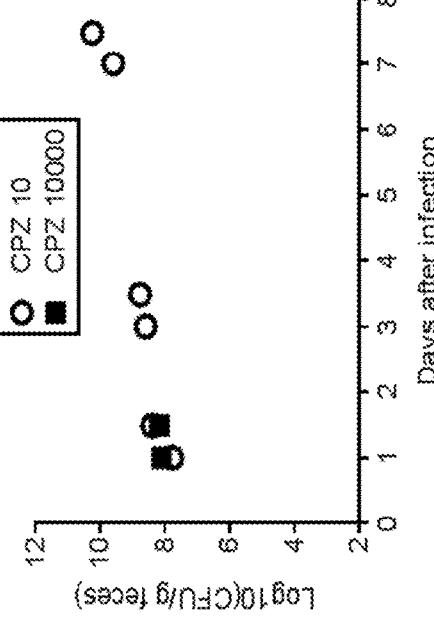


Figure 4J

Abx Cocktail

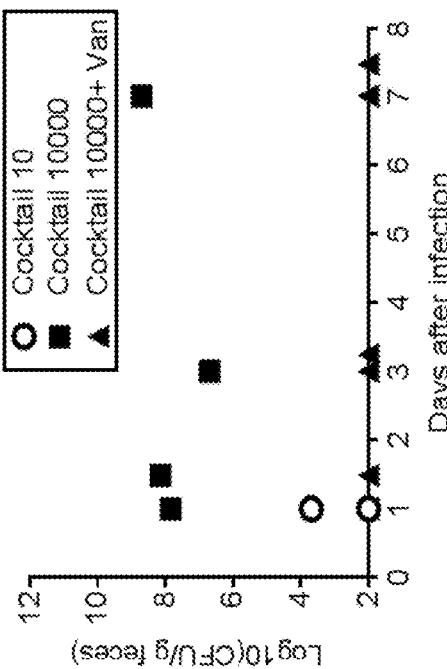


Figure 5

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	+	$10^4$

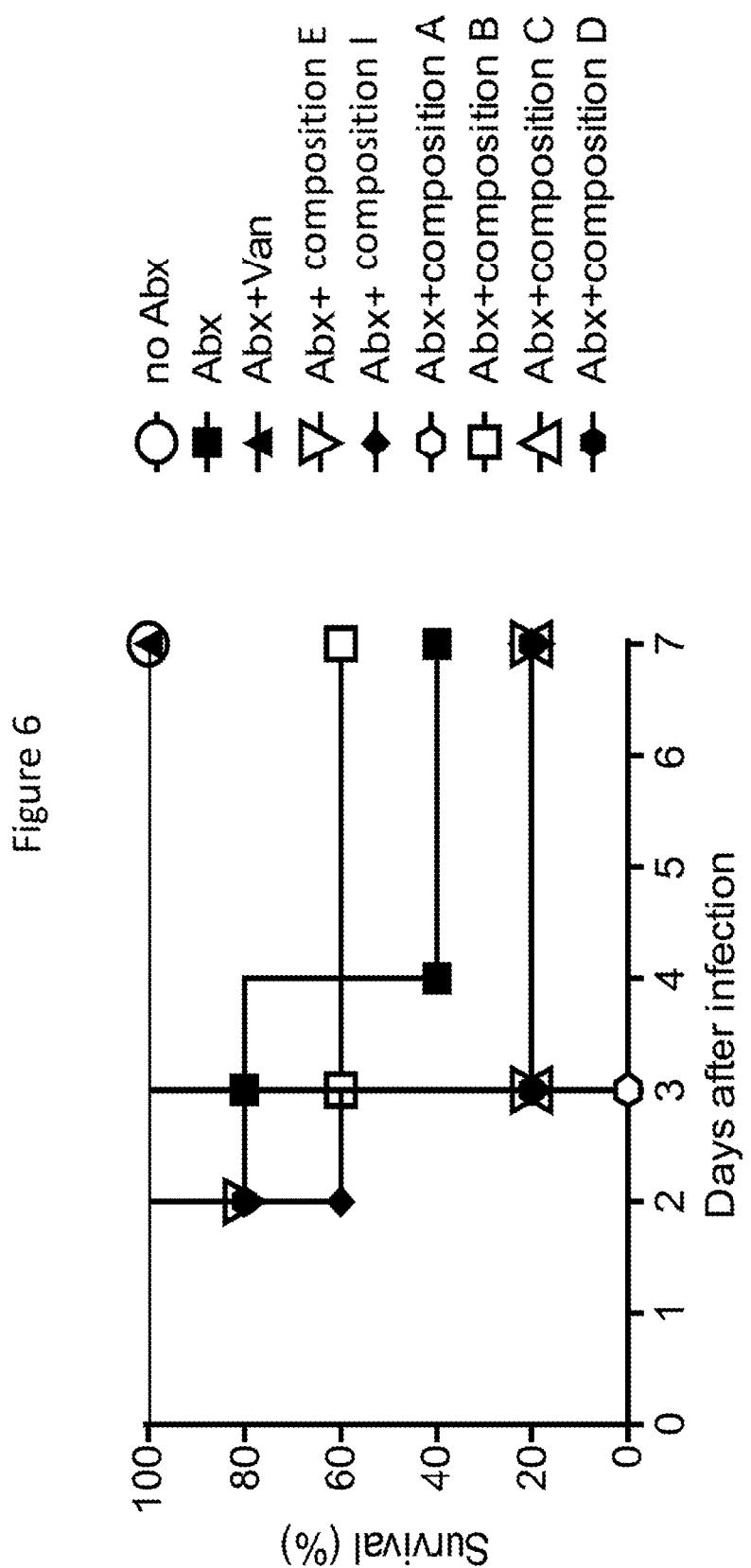


Figure 7A

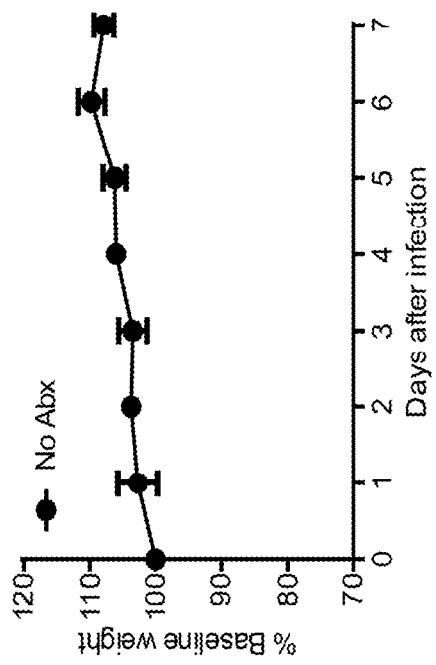


Figure 7B

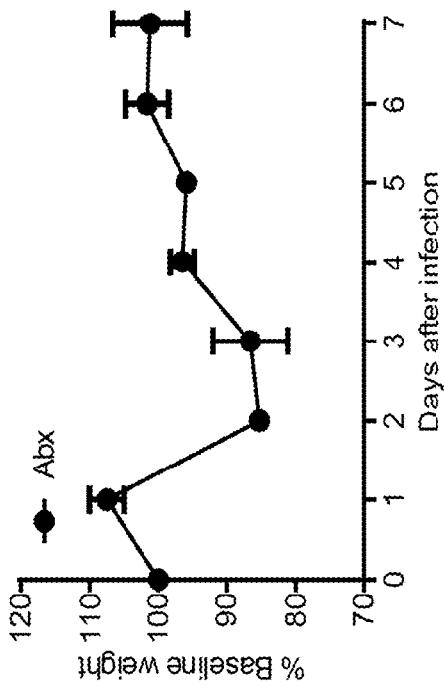


Figure 7C

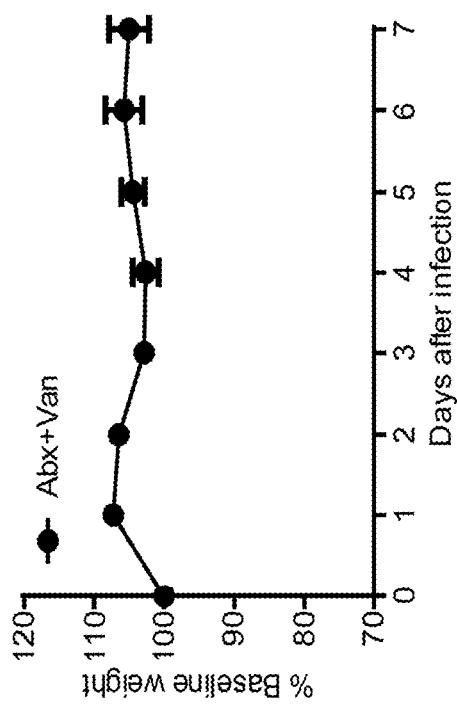


Figure 7D

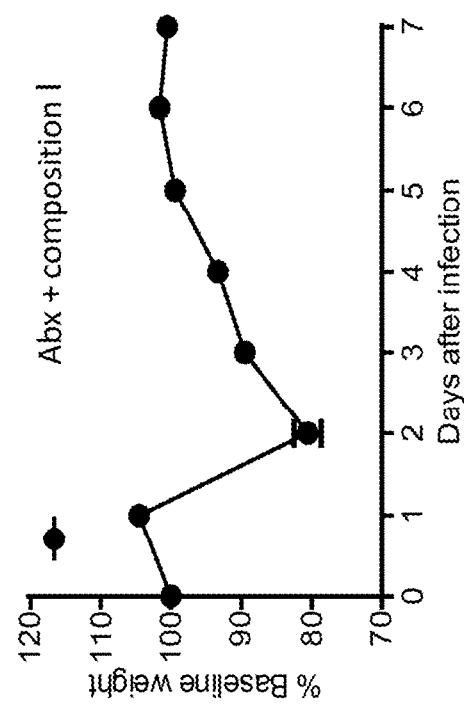


Figure 7E

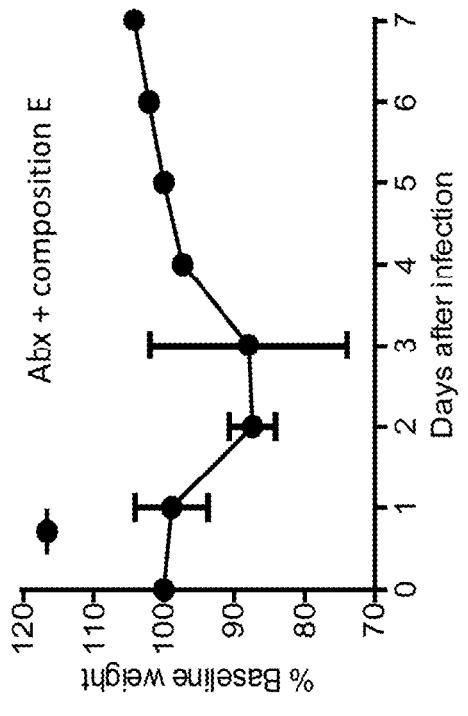


Figure 7F

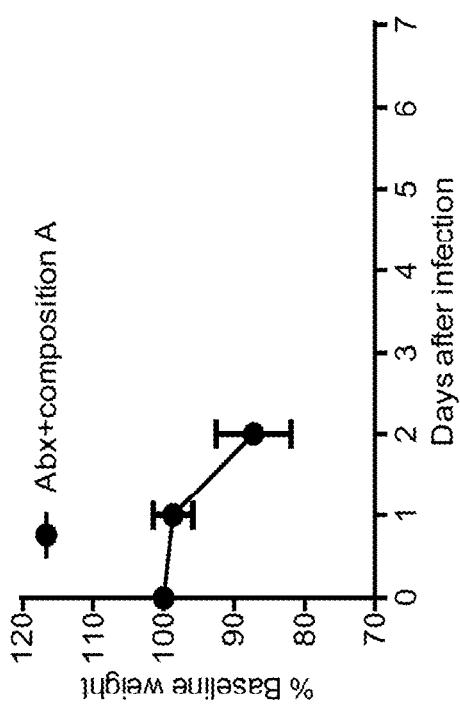


Figure 7G

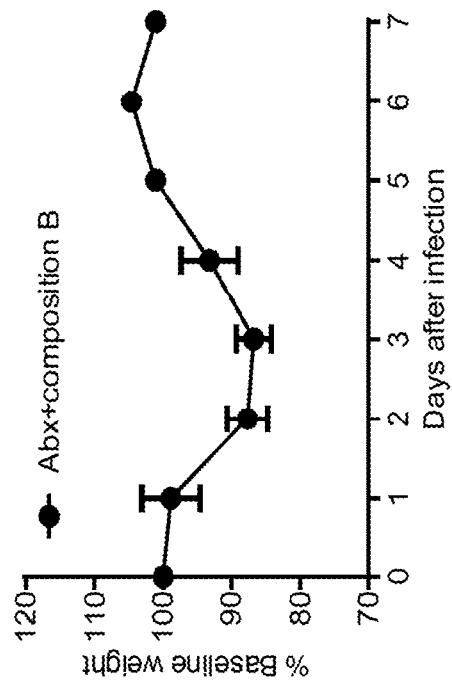


Figure 7H

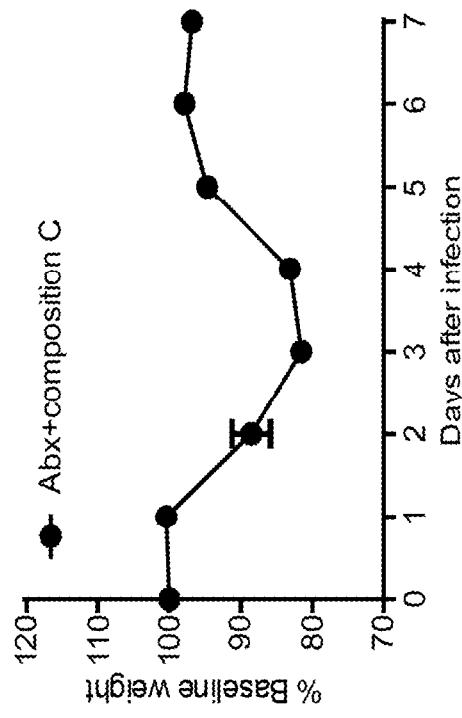


Figure 7I

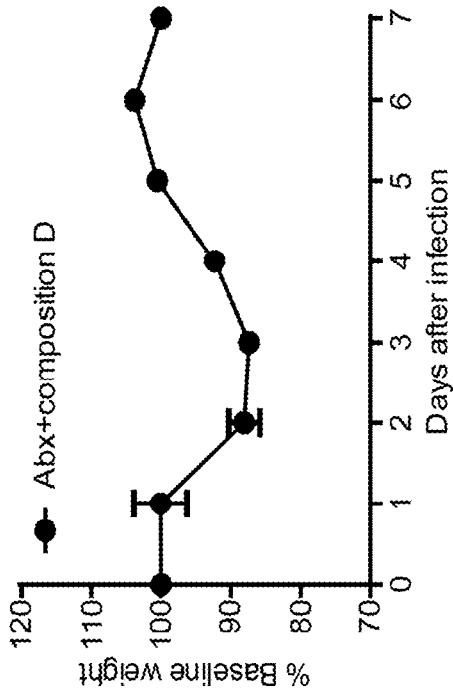


Figure 7J

Figure 8A

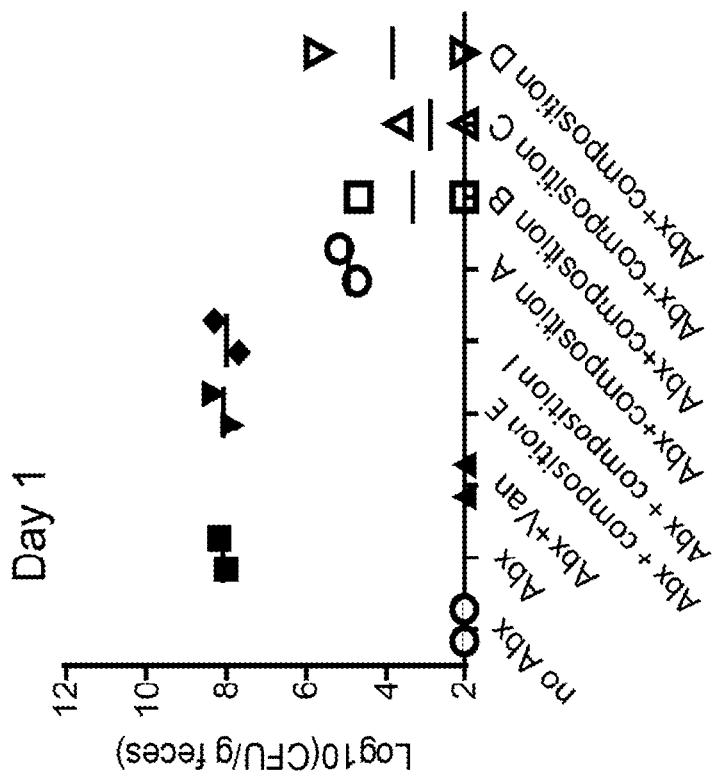
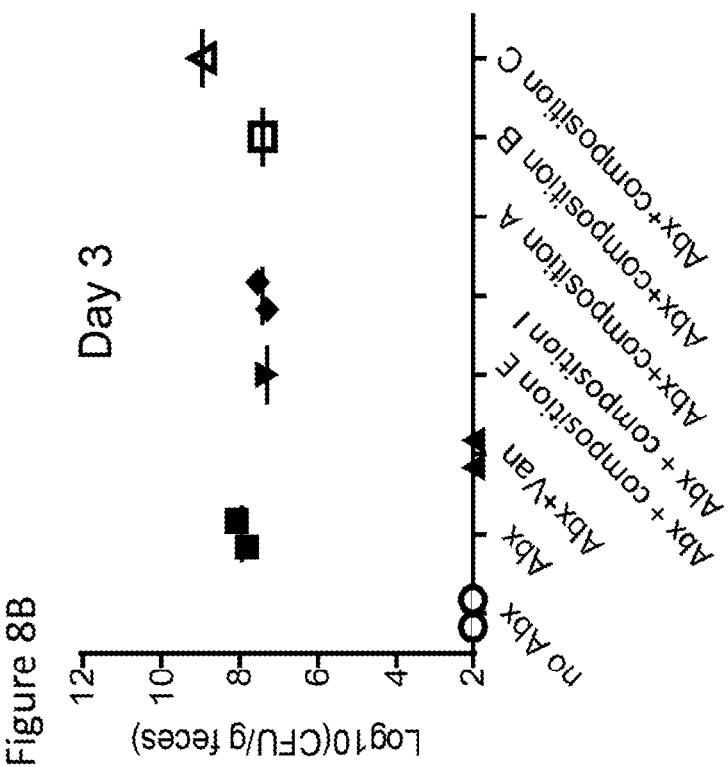


Figure 8B



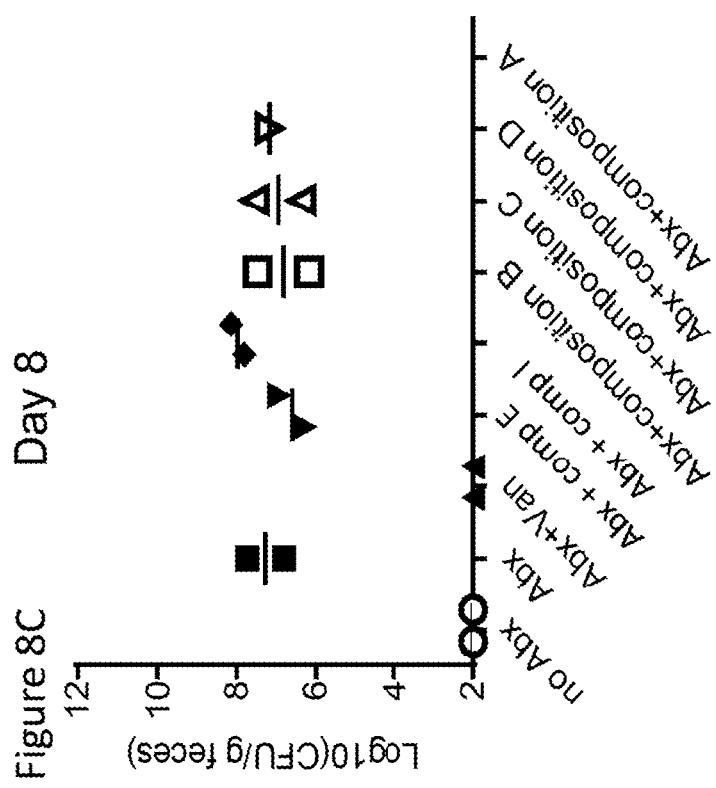
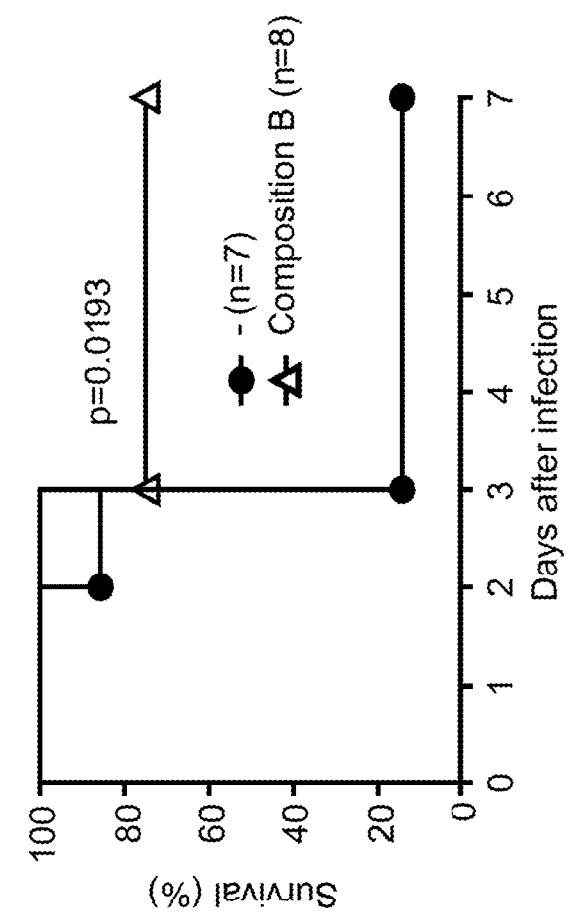


Figure 9

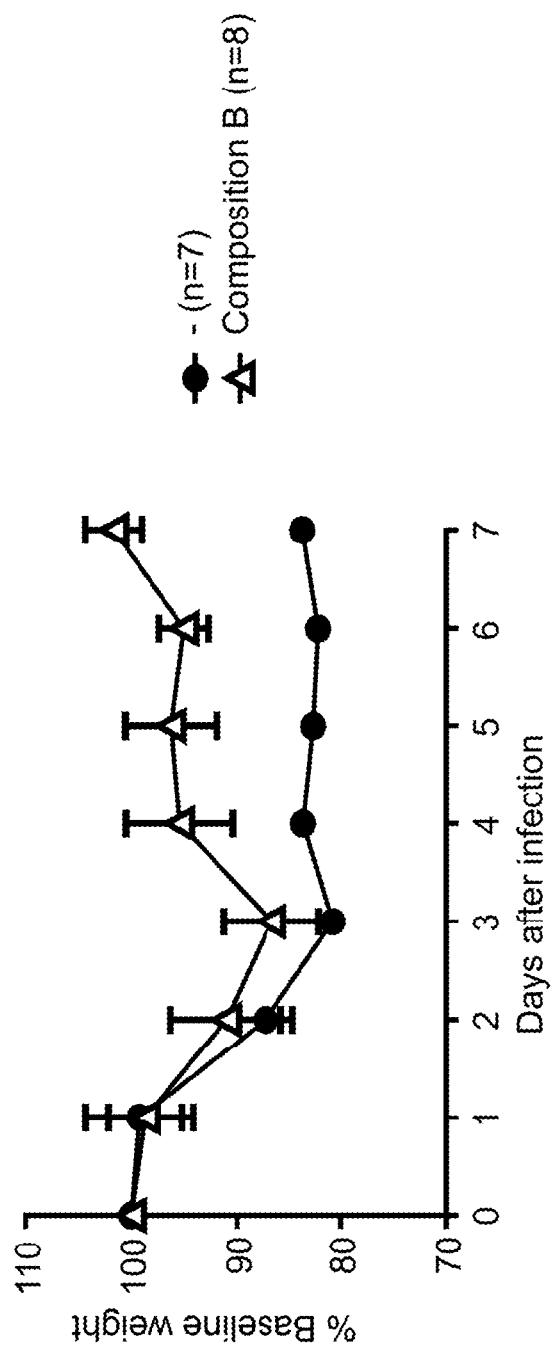
Groups	# of animals	Abx	<i>C. difficile</i> spore	CFUs LBPs
(1) Control	7	+	$10^4$	-
(2) Composition B	8	+	$10^4$	$10^8/\text{mouse}$

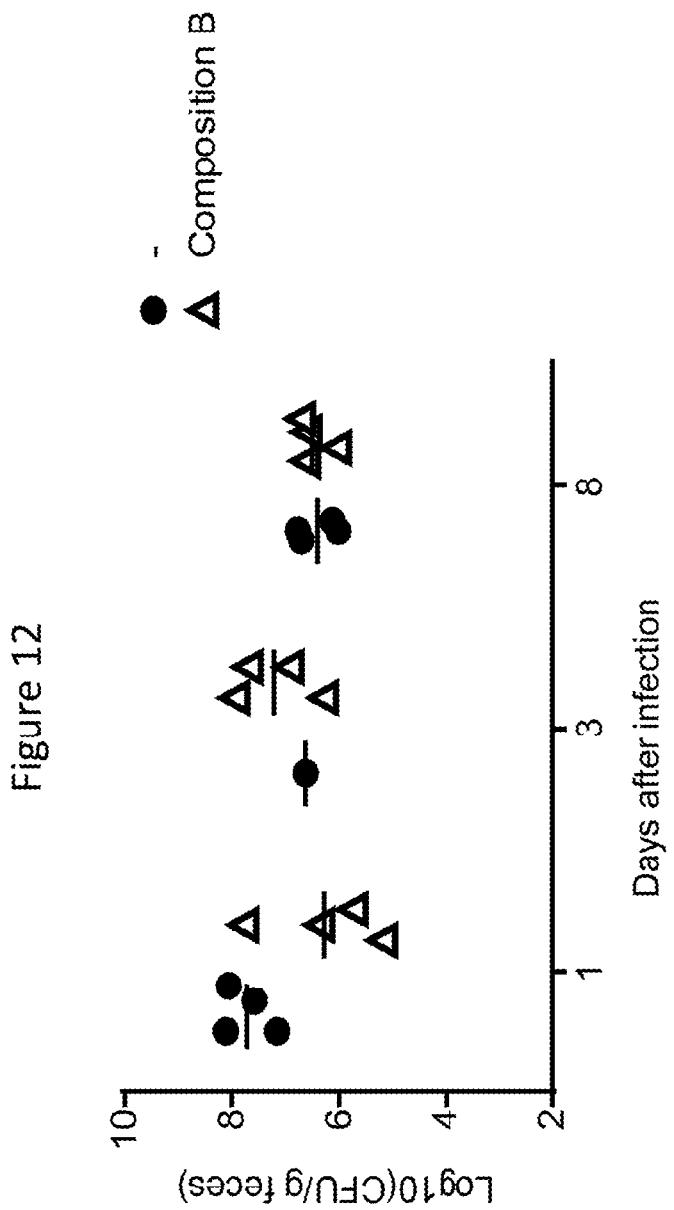
Figure 10



(Natural death + >20% weight loss)

Figure 1.1





**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_elisdenii	SEQ_54	YK54	Anaerostipes_hadrus
SEQ_27	YK149	Acidamicrococcus_fermentans /	SEQ_55	YK56	Ruminococcus_faecis
SEQ_28	YK154	Megasphaera_elisdenii	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_elisdenii
SEQ_36	YK153	Megasphaera_elisdenii	SEQ_64	YK88	Eubacterium_rectale
SEQ_37	YK163	Eubacterium_rectale	SEQ_65	YK89	Eubacterium_rectale
SEQ_38	YK191	Ruminococcus_champanellensis /	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	SCFAs
XIVa	<i>Eubacterium rectale</i> 12	A, B, L
	<i>Ruminococcus faecis</i> 8	A, I
	<i>Ruminococcus obaeum</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
IV	<i>Roseburia faecis</i> 5	A, B
	<i>Fusicatenibacter saccharivorans</i> 3	A, I
	<i>Dorea formicigenerans</i> 1	A
	<i>Dorea longicatena</i> 2	A
	<i>Flavonifractor plautii</i> 2	A, B
	<i>Ruminococcus chamaenellensis</i> 2	A
	<i>Acidaminococcus fermentans</i> 1	A, B, P
other	<i>Megasphaera elsdeni</i> 4	P
	<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	+	$10^4$	$10^8/\text{mouse}$
(3) Composition B dosed at day -2 and -1	10	+	$10^4$	$10^8/\text{mouse}$
(4) Composition B dosed at day -2, -1, 1, 2, and 3	10	+	$10^4$	$10^8/\text{mouse}$
(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 16

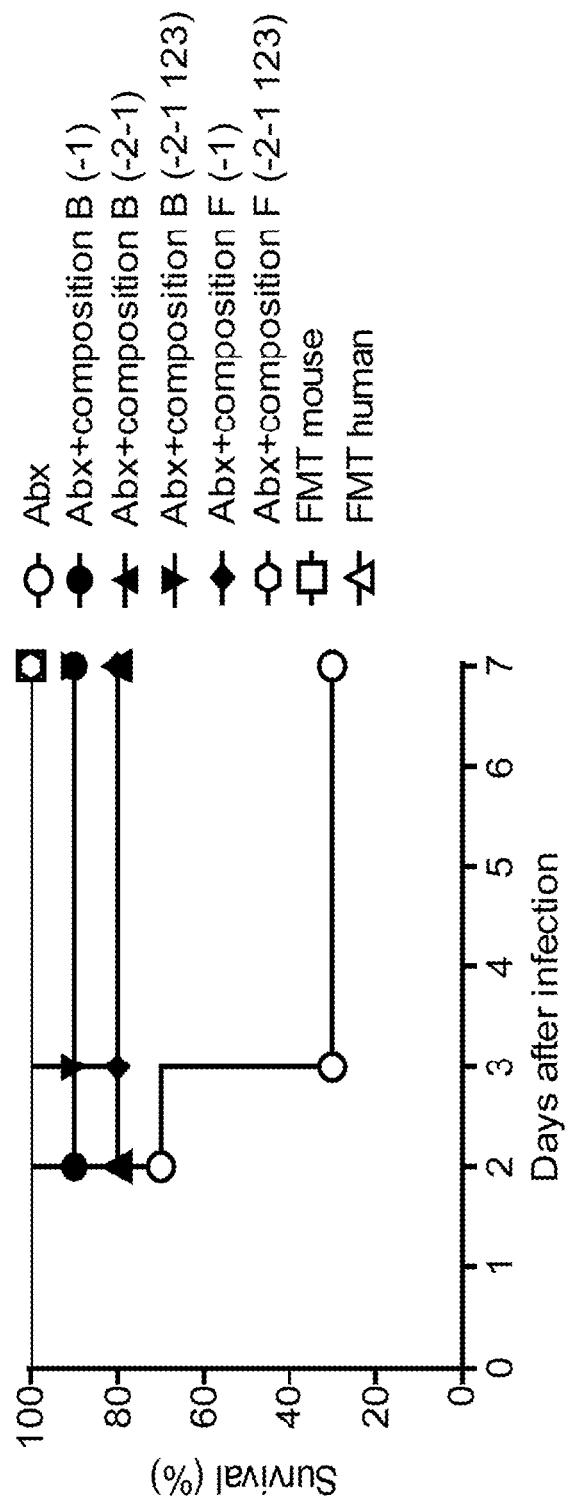


Figure 17A

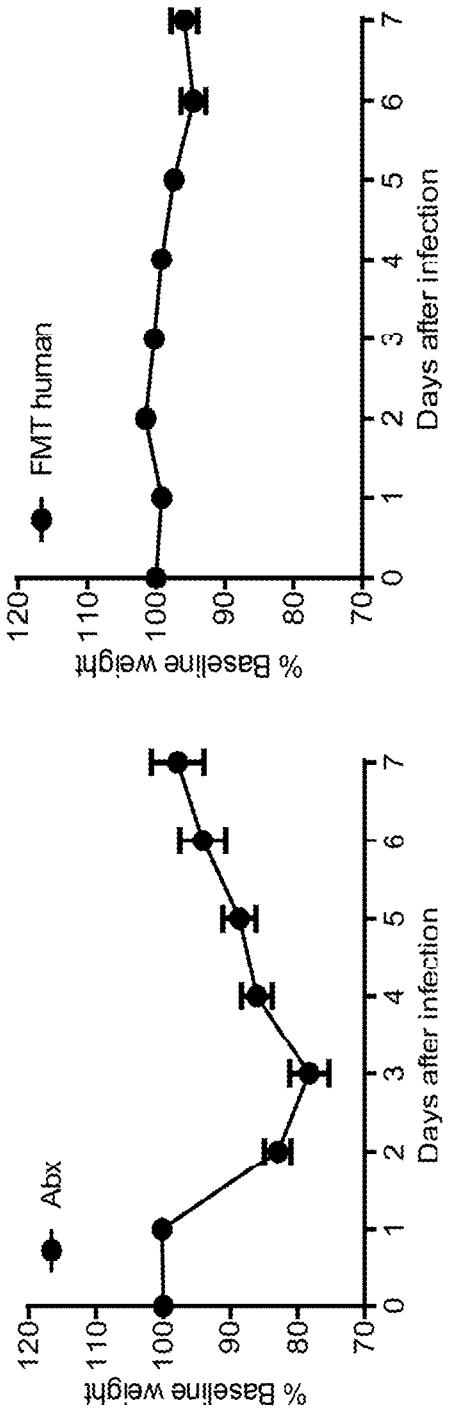


Figure 17B

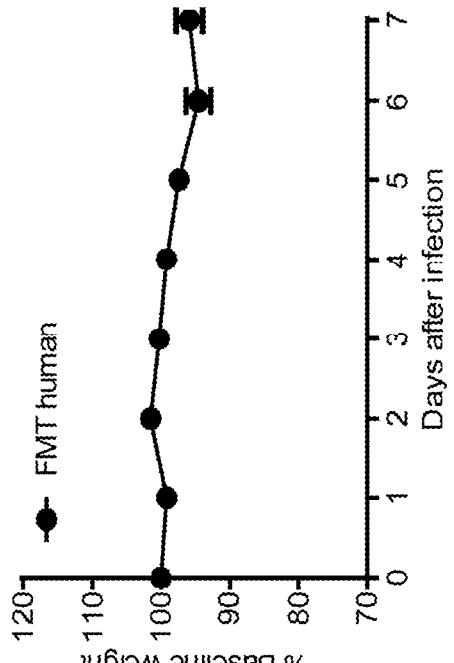


Figure 17C

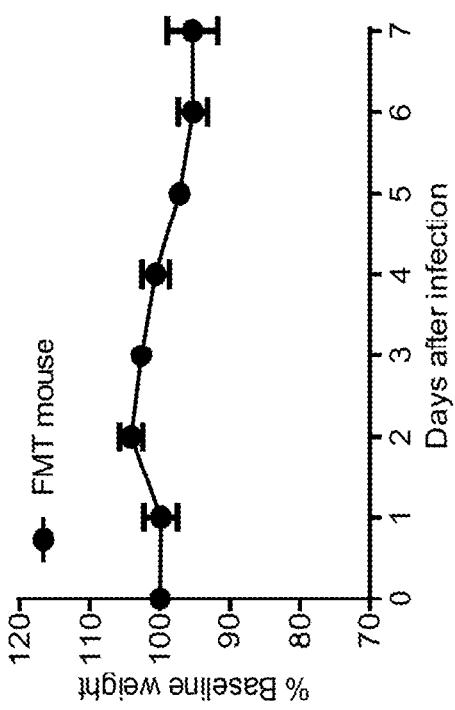


Figure 17D

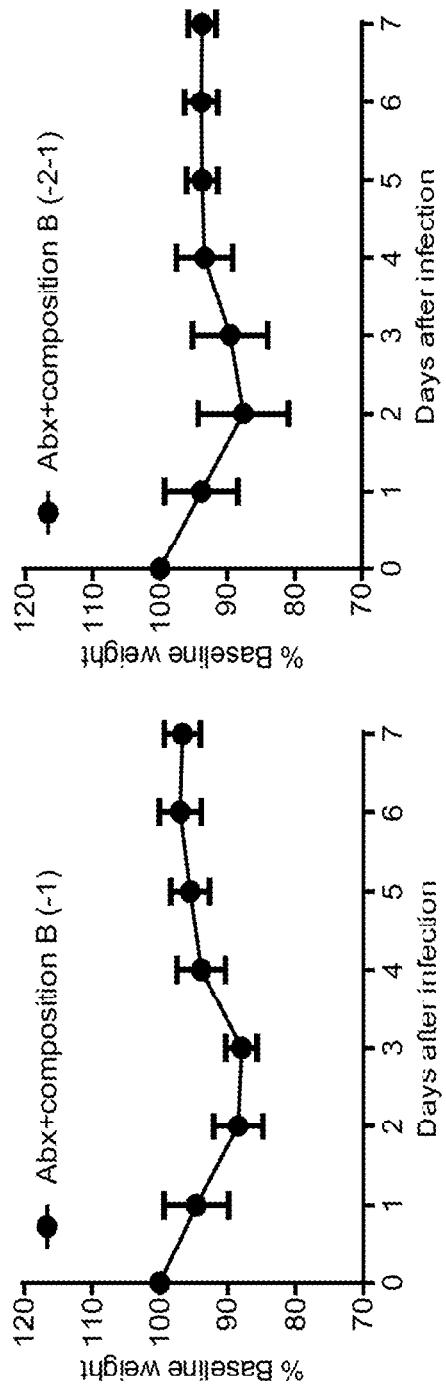


Figure 17E

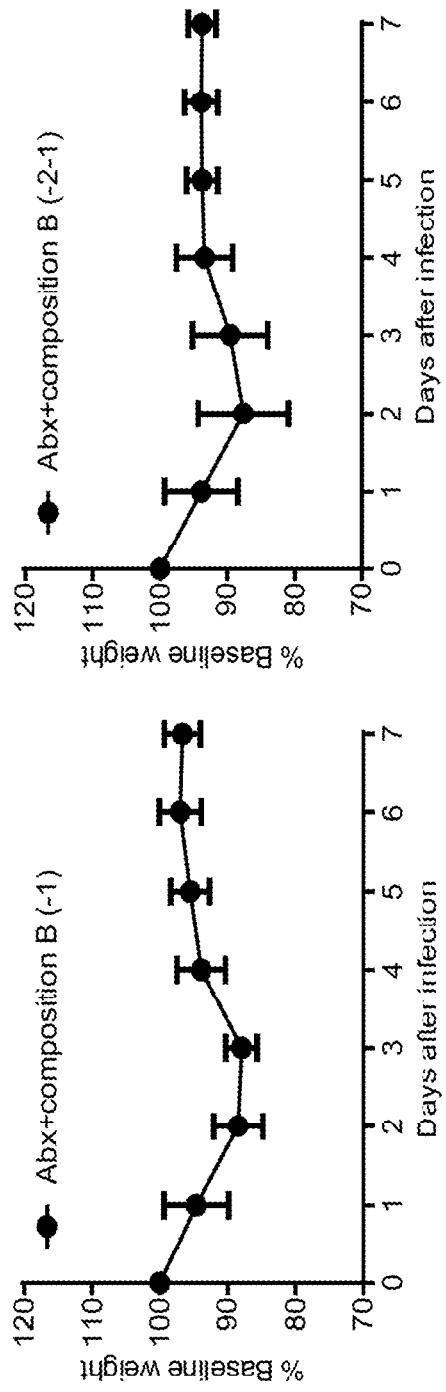


Figure 17F

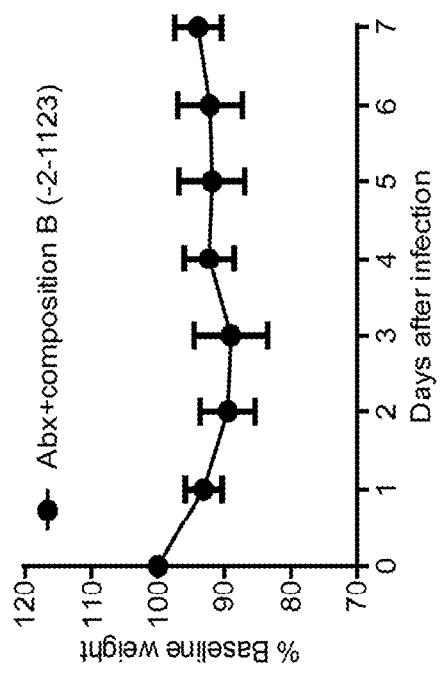
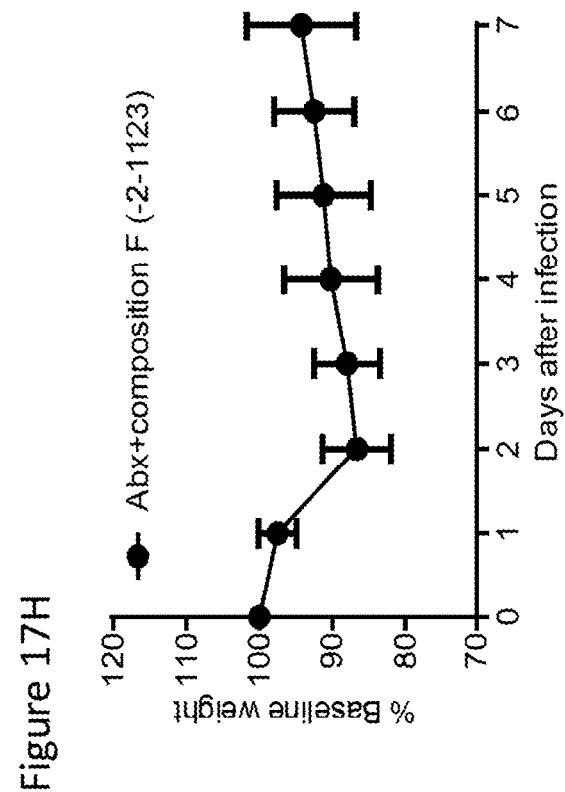
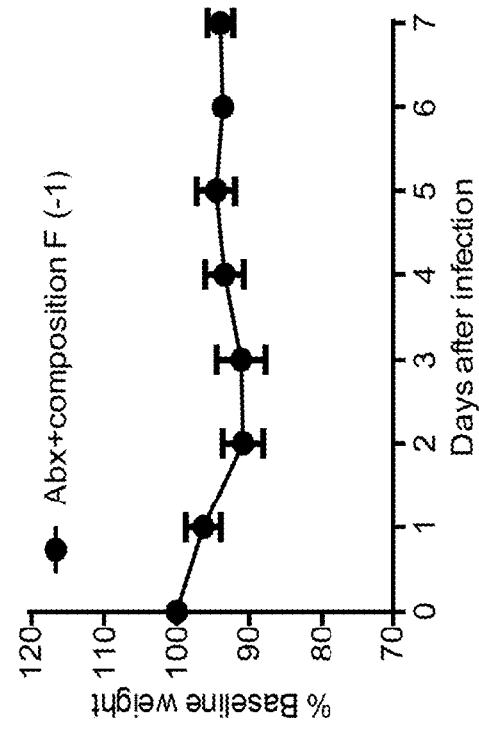
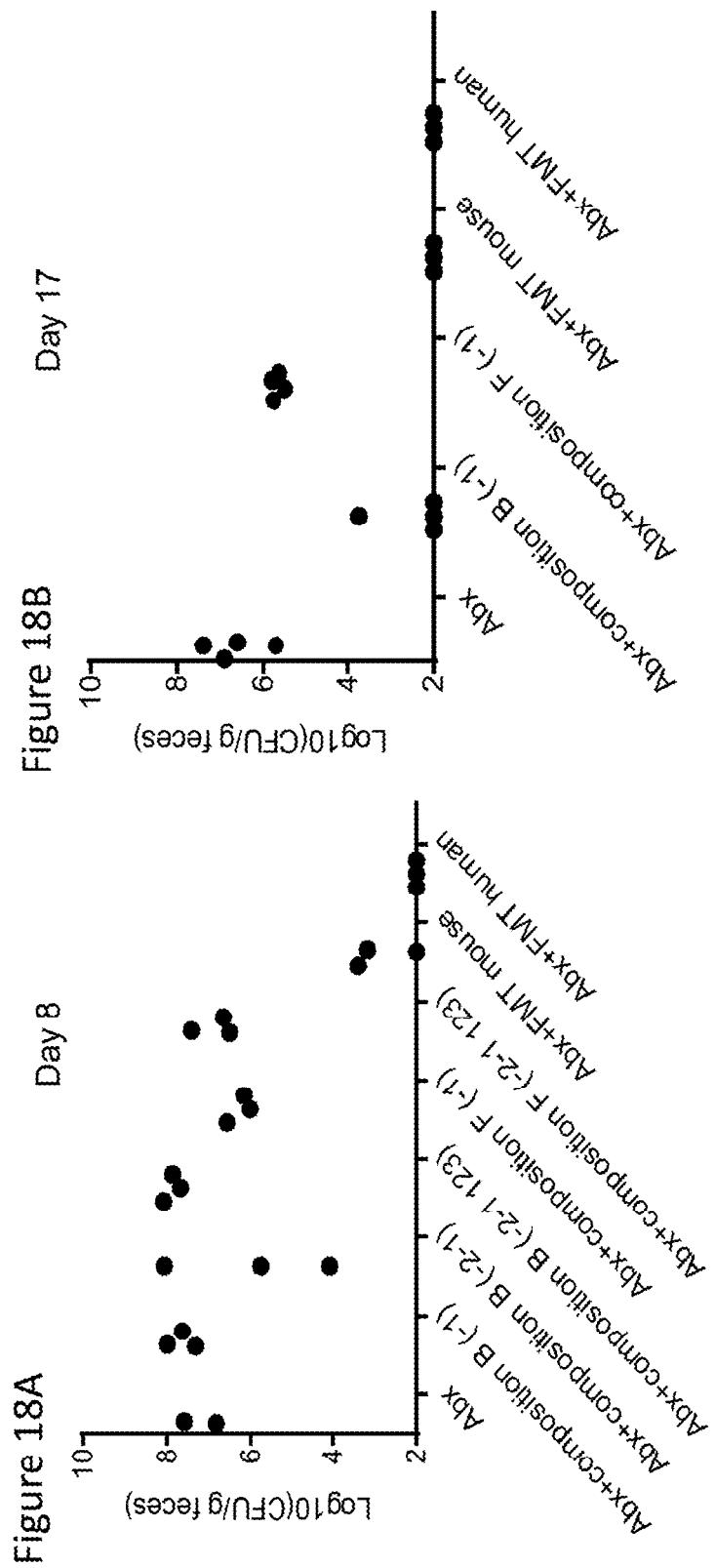


Figure 17G





## 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_lngicatena
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_eldenii
SEQ_63	YK71	Roseburia_faecis
SEQ_67	YK97	Roseburia_faecis
SEQ_40	YK99	Ruminococcus_champanellensis/Ruminococcus_albus
SEQ_38	YK191	Ruminococcus_champanellensis/Ruminococcus_albus
SEQ_47	YK27	Ruminococcus_faecis
SEQ_56	YK56	Ruminococcus_faecis
SEQ_25	YK101	Ruminococcus_oebeum
SEQ_32	YK64	Ruminococcus_oebeum

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 21

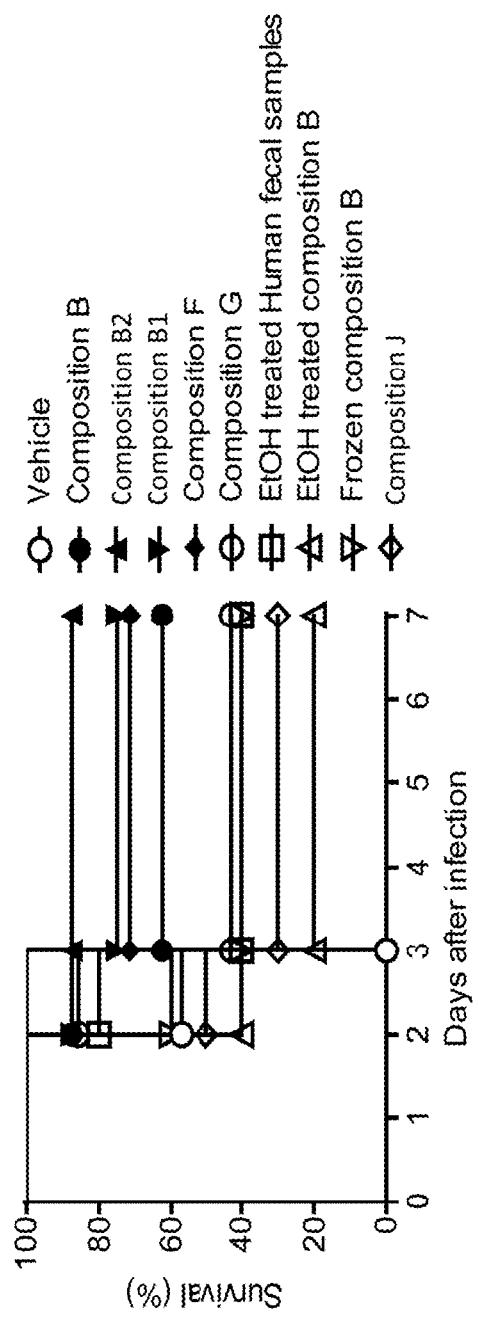


Figure 22A

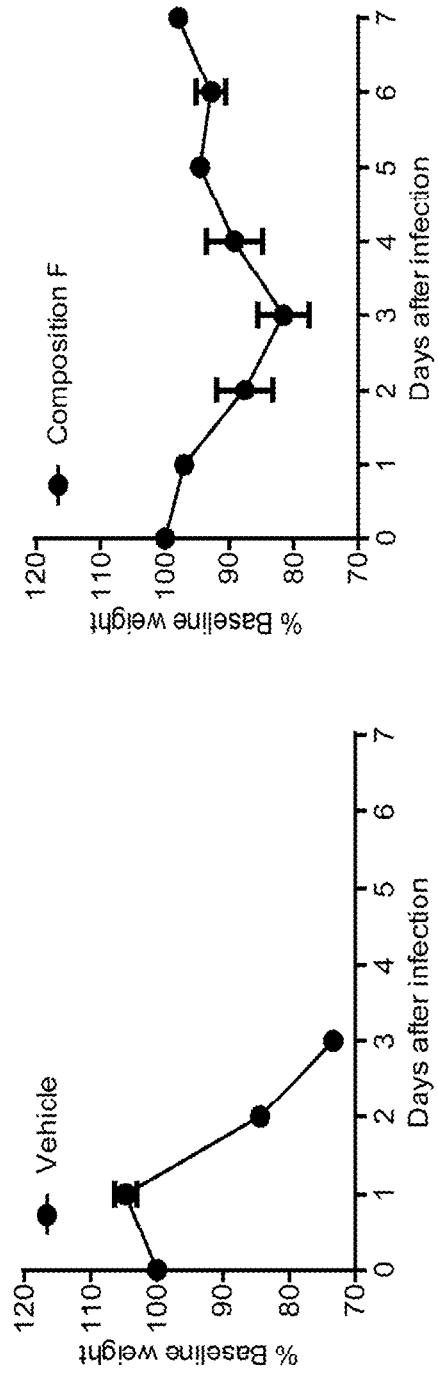


Figure 22B

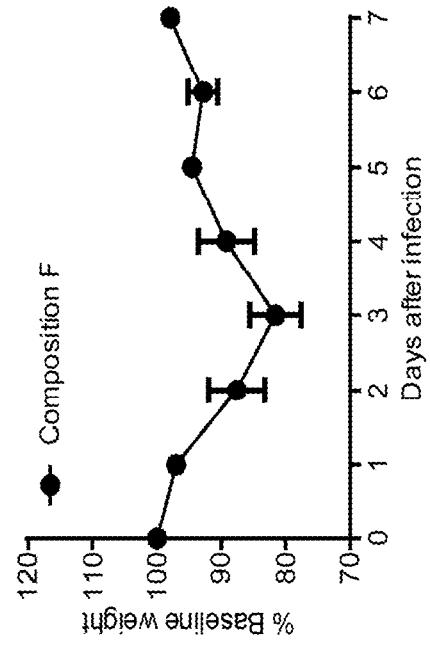


Figure 22C

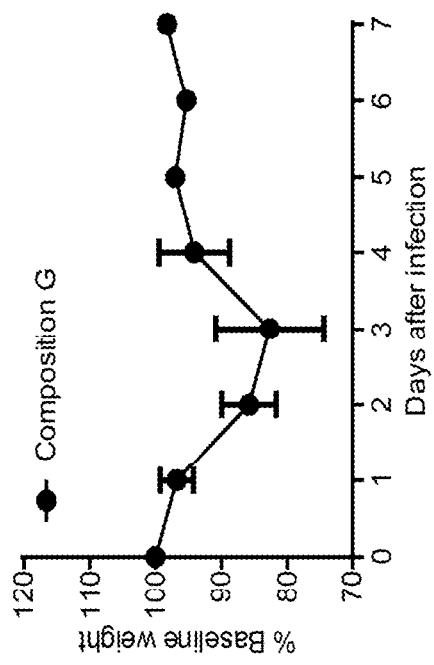


Figure 22D

Figure 22D

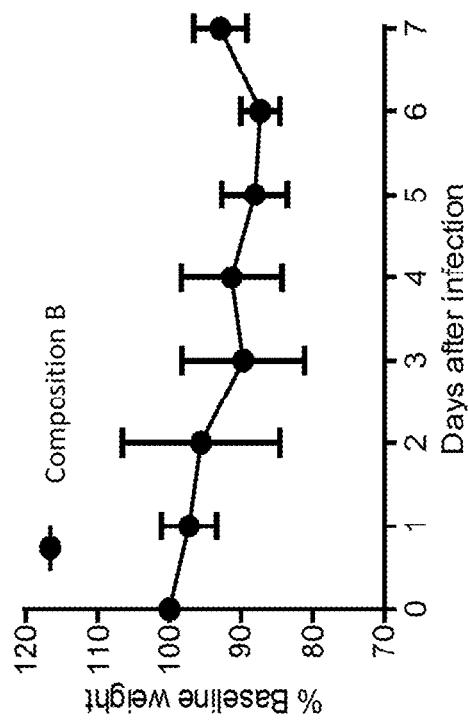


Figure 22E

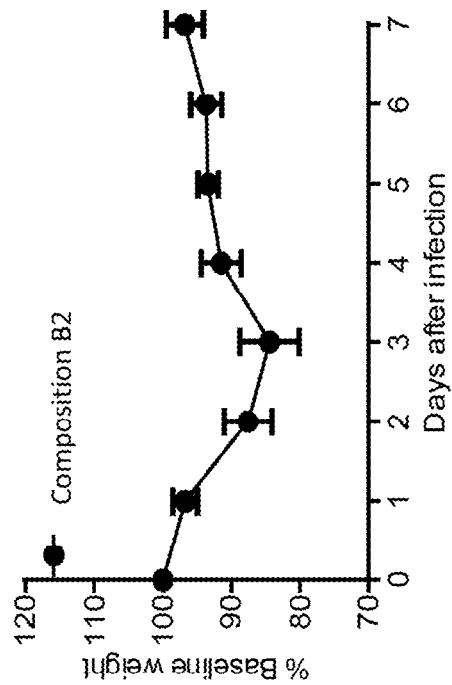


Figure 22F

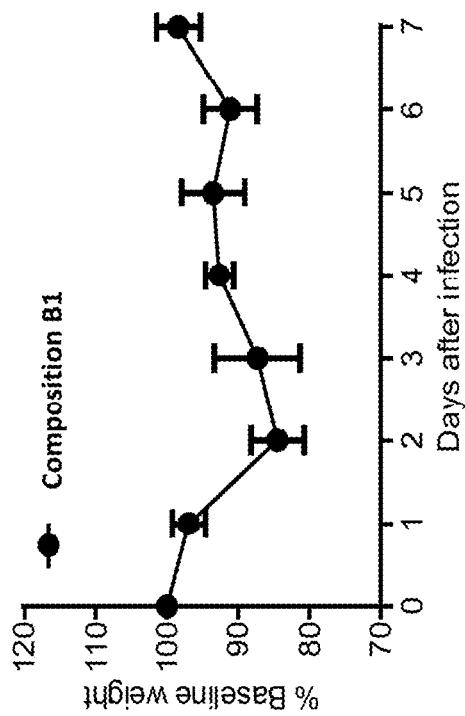


Figure 22G

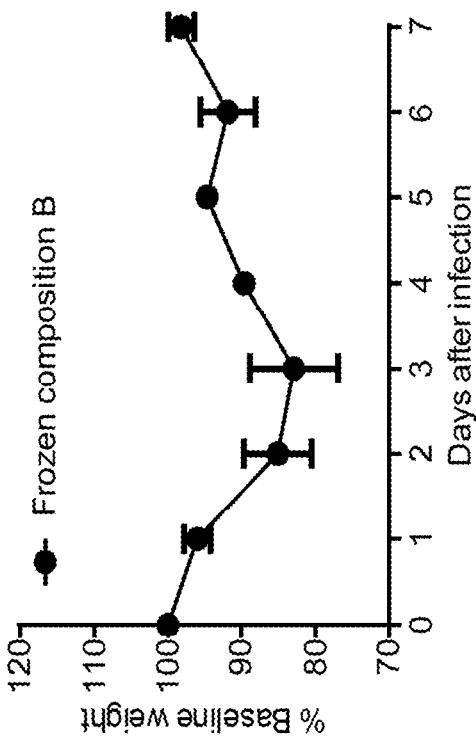


Figure 22H

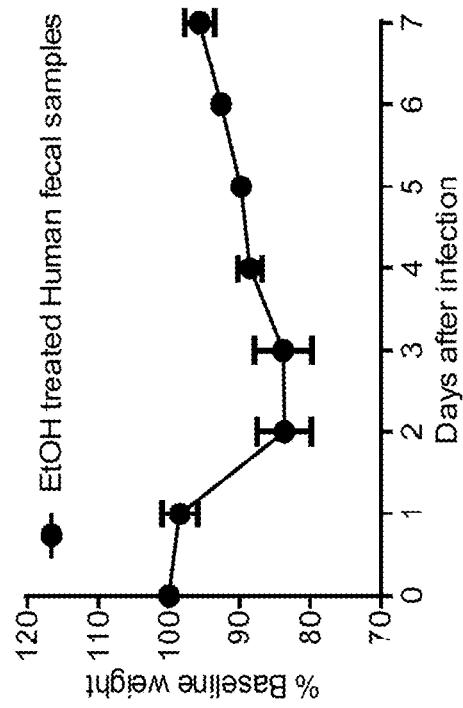


Figure 22I

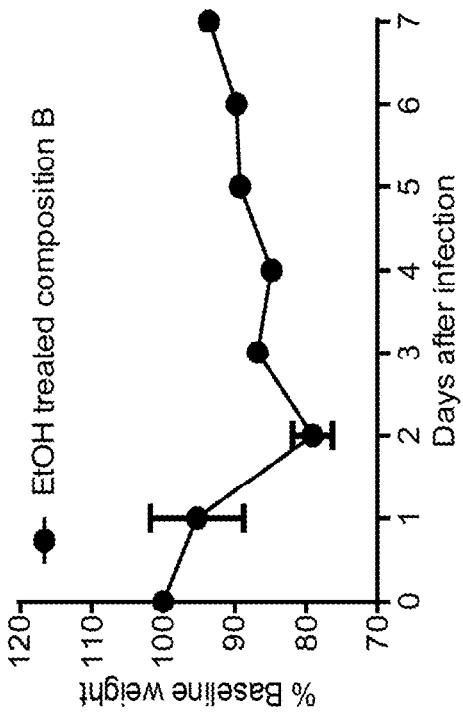
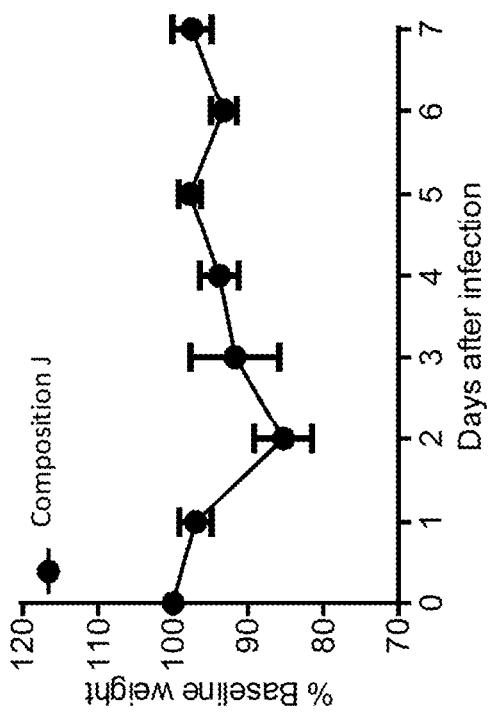


Figure 22J



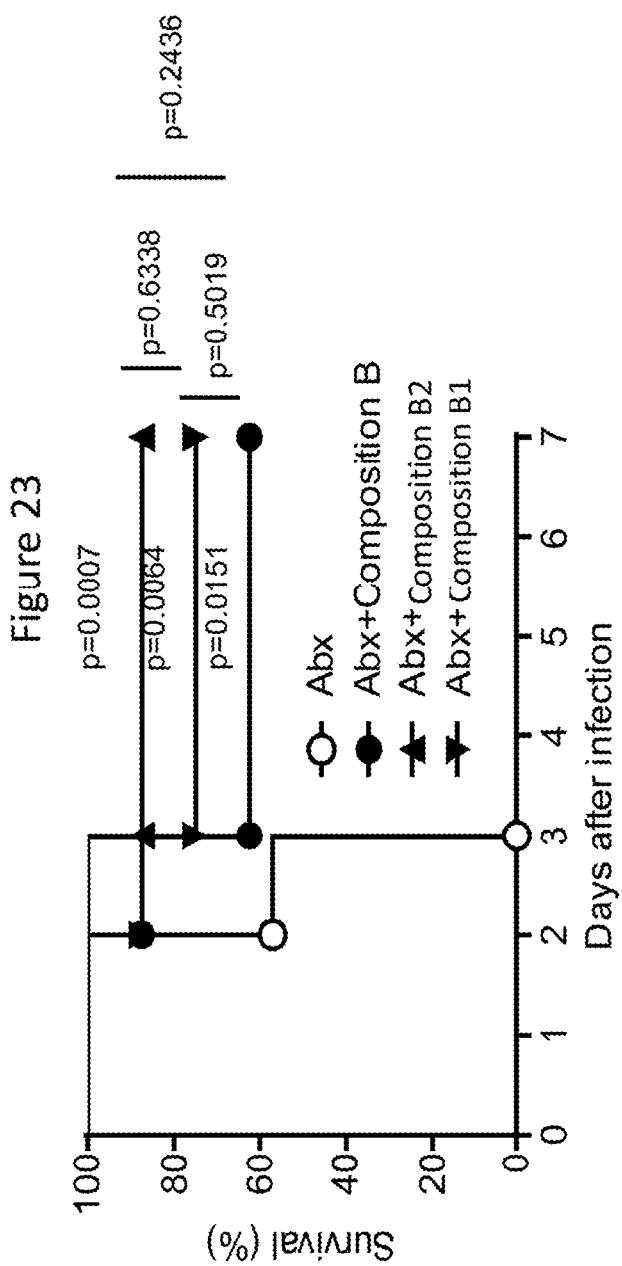


Figure 24

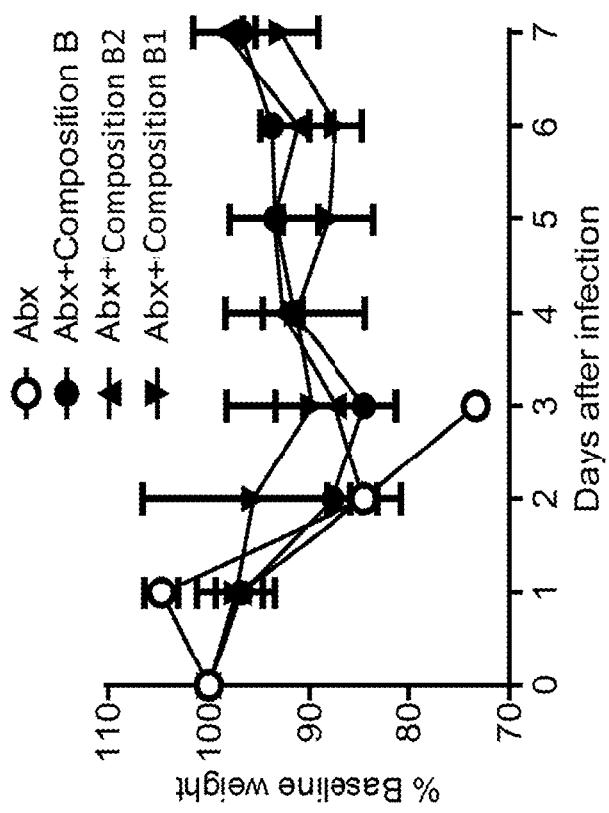


Figure 25

Groups	N	Abx	CFUs <i>C. difficile</i>	CFUs LBPs
(1) Vehicle	10	+	10 <sup>4</sup>	200ul of PBS
(2) human FMT	10	+	10 <sup>4</sup>	200ul of 10% fecal samples/mouse
(3) Composition B	10	+	10 <sup>4</sup>	10 <sup>8</sup> /mouse
(4) Composition B + 4 spores*	10	+	10 <sup>4</sup>	10 <sup>8</sup> live bacteria+spores/mouse
(5) Composition H**	10	+	10 <sup>4</sup>	10 <sup>8</sup> /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 sporificum*, and *Erysipelotoclostridium ramosum*

Figure 26  
Composition H

\*\*Composition H = The following six strains in spore form *Clostridium bolteae*,  
*Anaerotruncus colihominis*, *Clostridium symbiosum*, *Clostridium innocuum*,  
*Clostridium sporificum* 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 sporificum</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

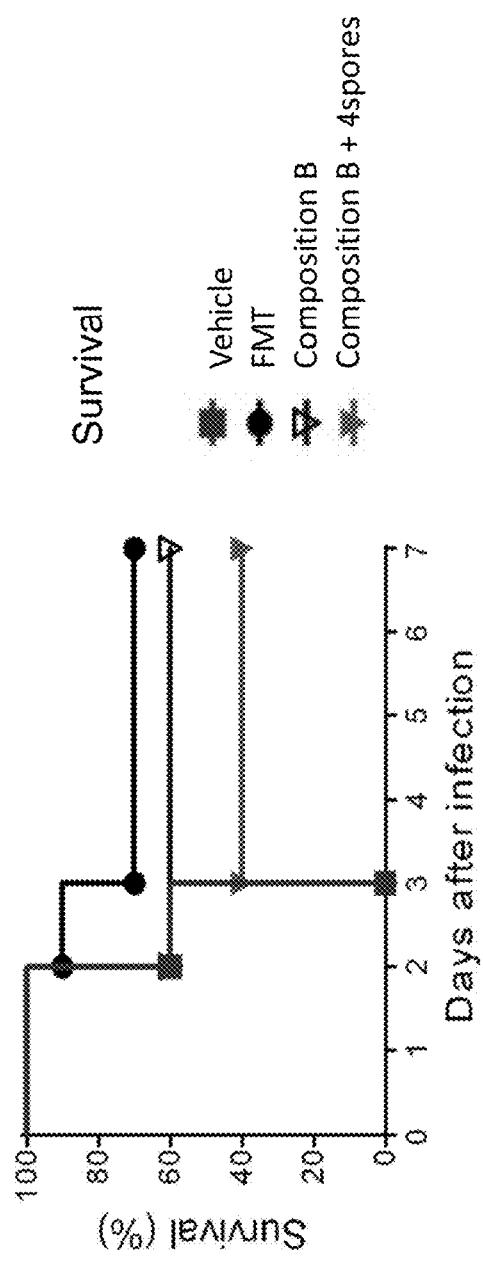


Figure 27B

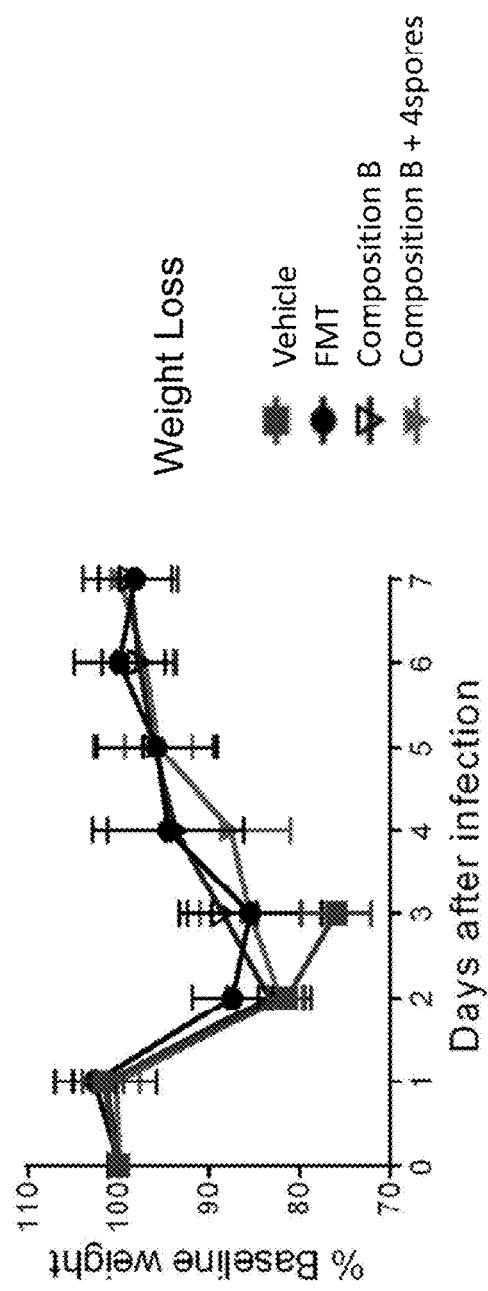


Figure 28A

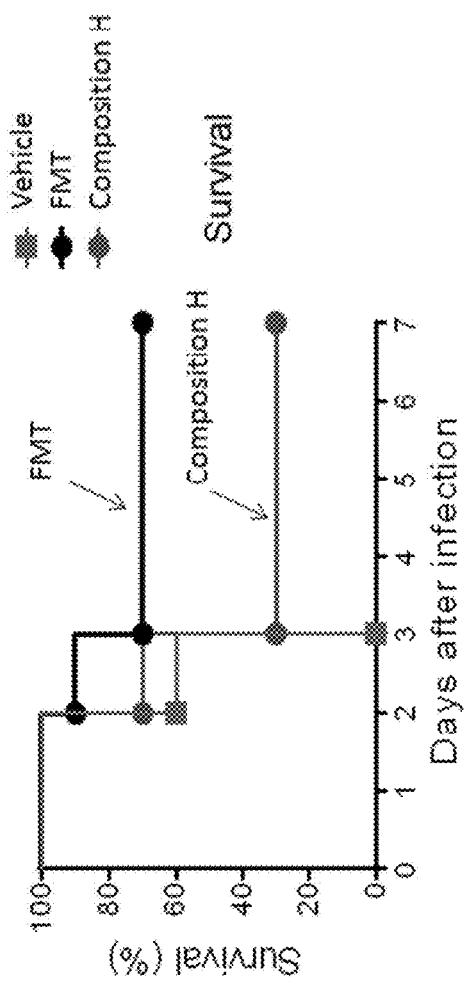


Figure 28B

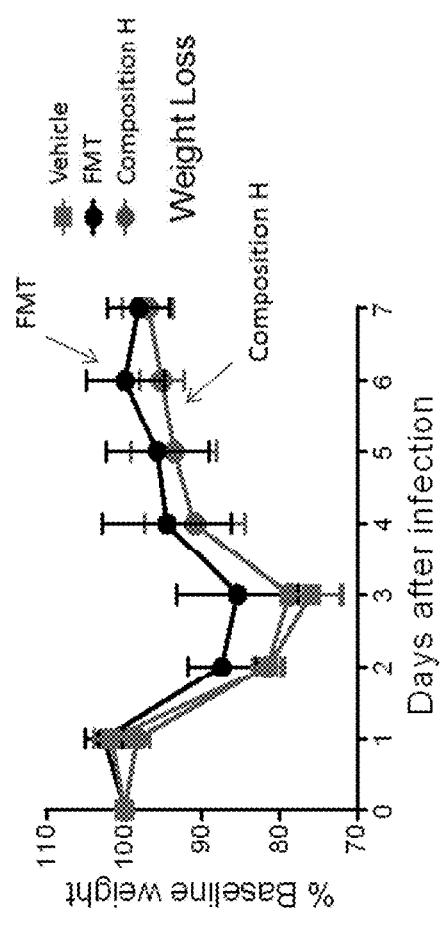


Figure 29A

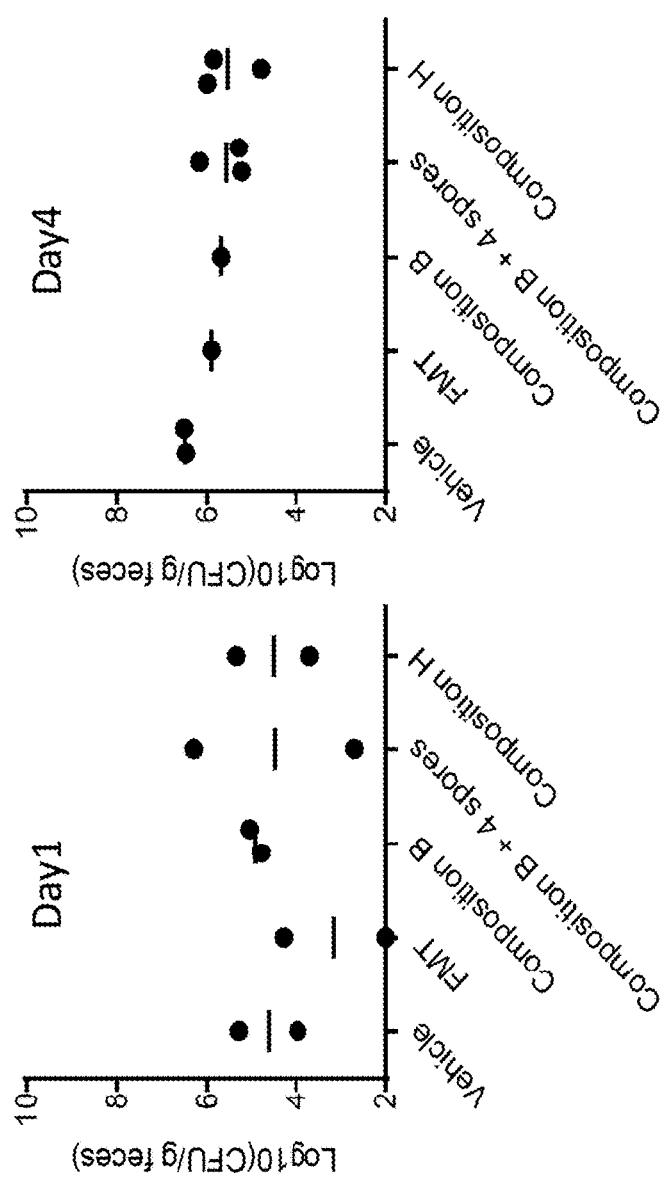


Figure 29B

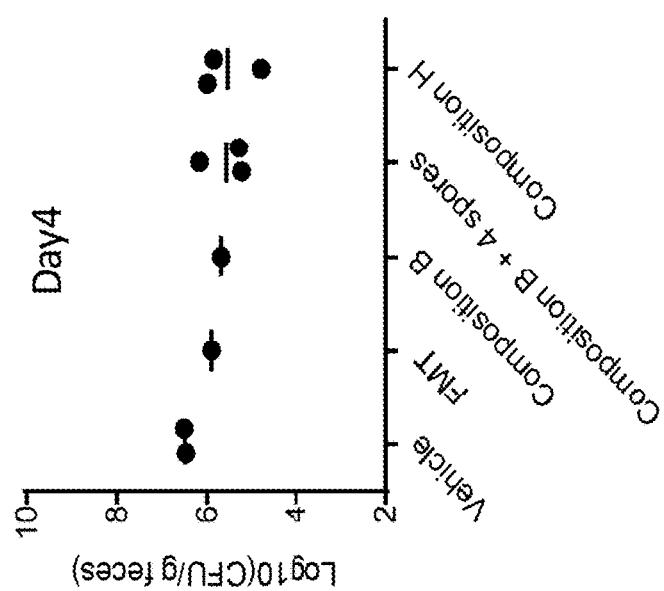


Figure 29C

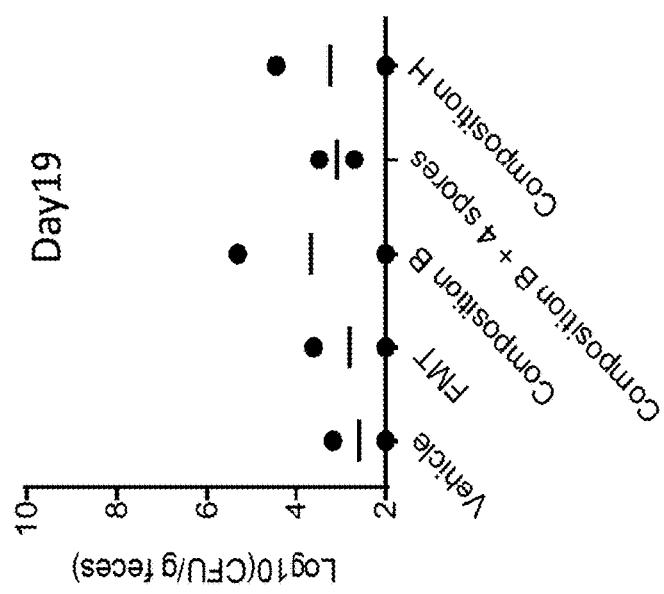


Figure 30

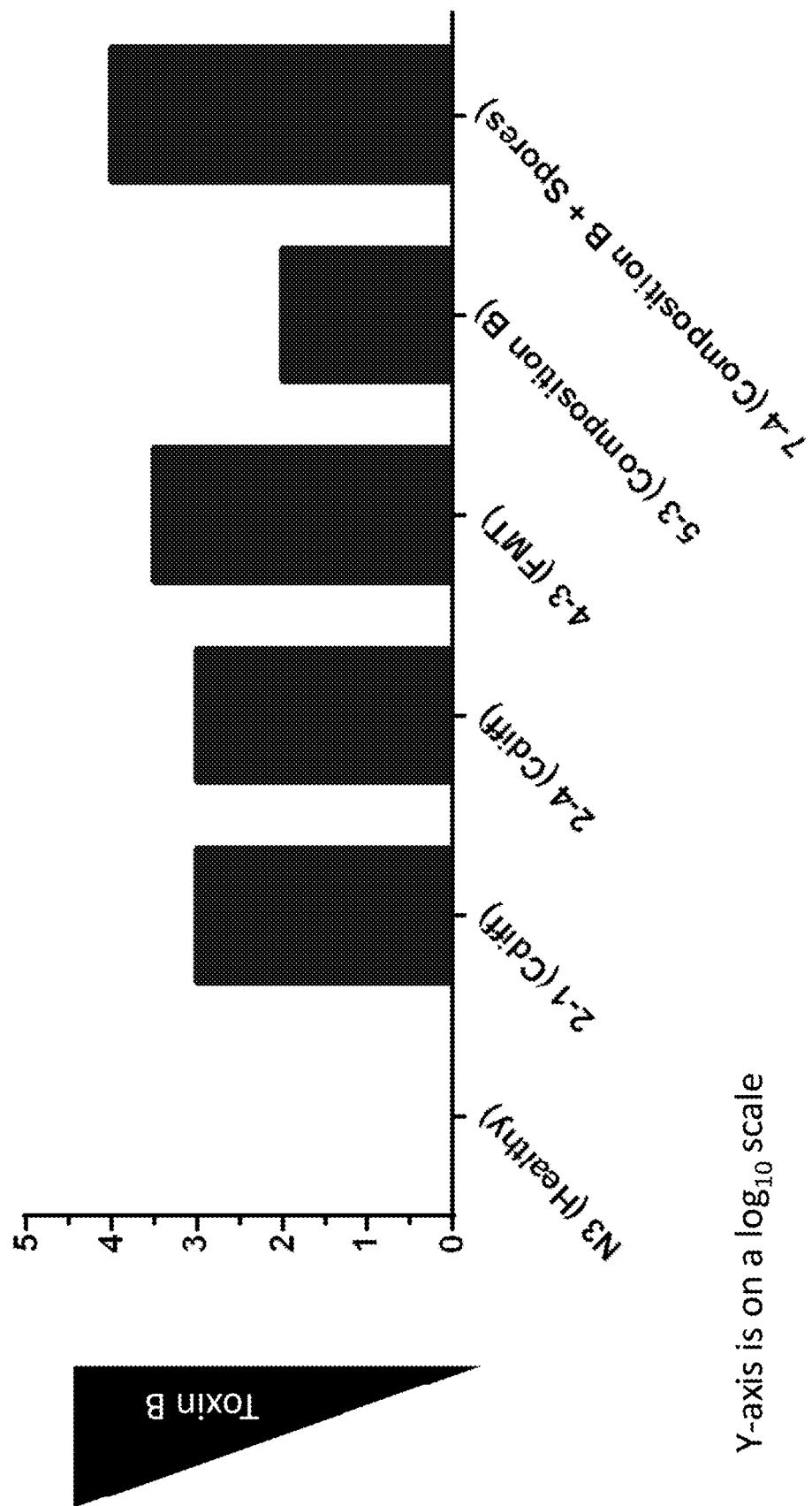


Figure 31

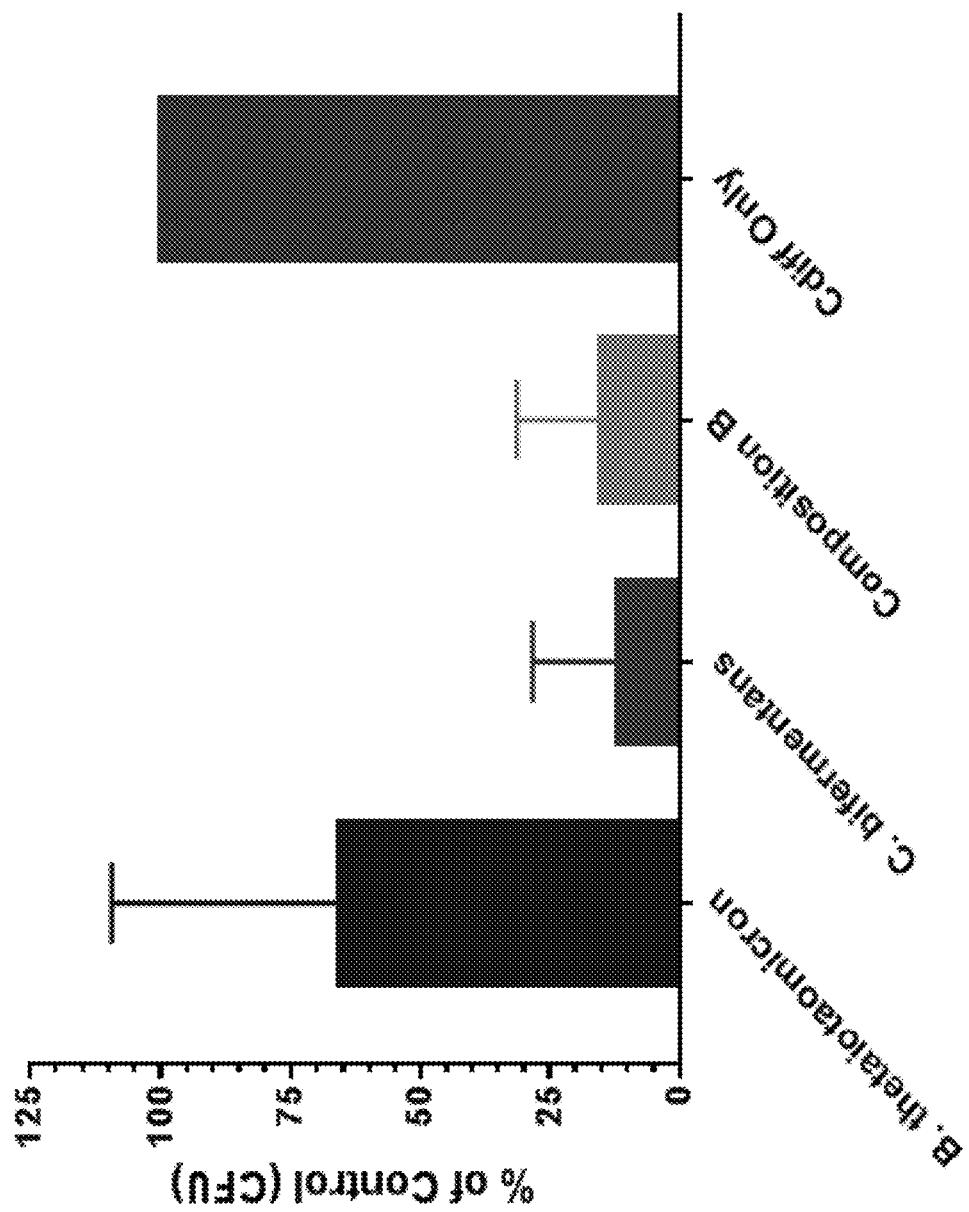
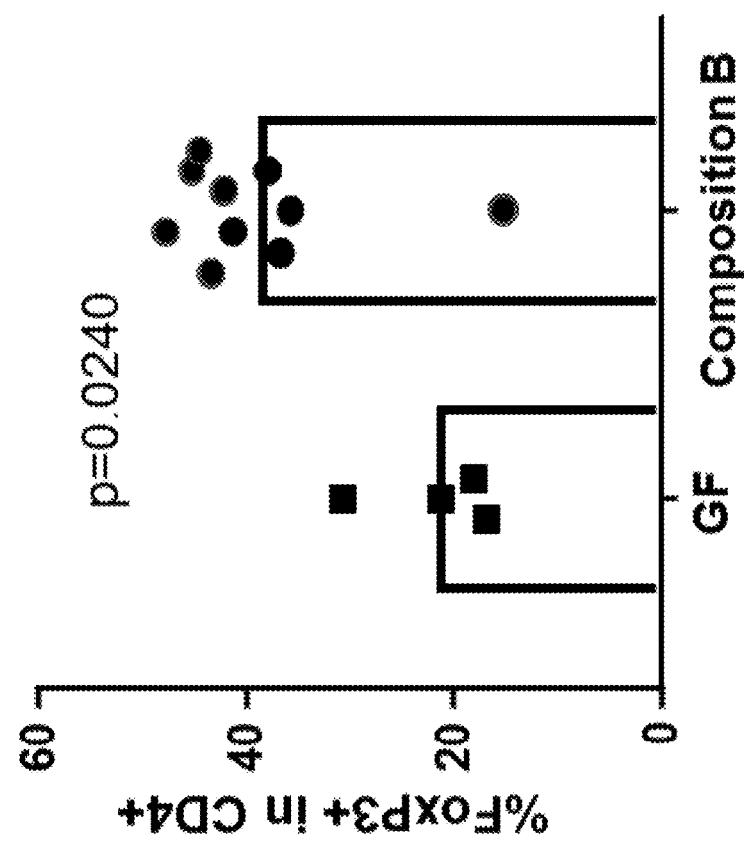


Figure 32



1

## TREATMENT OF *CLOSTRIDIUM DIFFICILE* INFECTION

### RELATED APPLICATIONS

This application is a continuation of U.S. application Ser. No. 16/157,640, filed Oct. 11, 2018, which is a continuation of U.S. application Ser. No. 15/993,037, filed May 30, 2018, which is a continuation of U.S. application Ser. No. 15/630,088, filed Jun. 22, 2017, now issued as U.S. Pat. No. 9,999,641, which is a continuation of international application number PCT/US2017/037498, filed Jun. 14, 2017, which claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application No. 62/349,914, filed Jun. 14, 2016, each of which is incorporated by reference herein in its entirety.

### FIELD OF INVENTION

The disclosure relates to compositions of purified bacterial strains, and methods for 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 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): 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.

*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-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

2

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 (Camarota 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. *Infec. 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)).

### 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 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* GD1, *Ruminococcus torques*, *Anerostipes caccae*, *Clostridium scindens*, *Marvinbryantia formatexigens*, *Eisenbergiella tayi*, *Flavimicrofractor plautii*, *Clostridium orbiscindens* 1\_3\_50AFAA, *Lachnospiraceae bacterium* 7\_1\_58FAA, *Subdoligranulum*, *Anerotruncus colihominis*, *Anerotruncus 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*, *Erysipelotrichaceae\_bacterium\_21-3*, *Blautia wexlerae*, *Clostridium disporicum*, *Erysipelotrichaceae\_bacterium\_21-3*, *Bacteroides cellulosilyticus*, *Anaerostipes hadrus*, *Eubacterium rectale*, *Ruminococcus chamaenensis*, *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 massiliensis* GD1, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryantia 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 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 scindens*, *Marvinbryantia 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 colihominis* DSM 17241, *Eubacterium fissicatena*, *Sellimonas intestinalis*, *Dracourtella massiliensis*, *Dracourtella massiliensis* 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 *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 massiliensis* 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 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*, and *Erysipelotrichaceae\_bacterium\_21-3*.

10 In some embodiments of the compositions provided herein, the composition comprises purified bacterial strains *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*, 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, *Sellimonas intestinalis* GD1, *Clostridium symbiosum* WAL-14163, *Clostridium bolteae* 90A9, *Dorea longicatena* CAG:42, *Clostridium bacterium UC5.1-1D4*, and *Erysipelotrichaceae\_bacterium\_21-3*.

20 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*, *Anaerotruncus colihominis*, *Dracourtella massiliensis*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

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

30 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*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

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

40 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*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, and *Clostridium innocuum*.

45 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*, *Anaerotruncus colihominis*, *Eubacterium fissicatena*, *Ruminococcus torques*, *Clostridium symbiosum*, *Clostridium bolteae*, *Dorea longicatena*, *Blautia producta*, *Clostridium innocuum*, *Erysipelotrichaceae\_bacterium\_21-3*, and *Bacteroides ovatus*.

50 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 *Flavimonas 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 *Flavimonas 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\_158FAA*.

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*, *Anaerostipes hadrus*, *Flavinofractor plautii*, *Eubacterium rectale*, *Ruminococcus chamanensis*, *Ruminococcus albus*, *Bifidobacterium bifidum*, *Ruminococcus faecis*, *Blautia luti*, *Roseburia faecis*, *Fusicatenibacter saccharivorans*, *Blautia faecis*, *Dorea formicigenerans*, and *Blautia hansenii*.

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 hanseenii*, *Dorea formicigenerans*, *Dorea longicatena*, *Eubacterium rectale*, *Flavinofractor plautii*, *Fusicatenibacter saccharivorans*, *Megasphaera elsdenii*, *Roseburia faecis*,

*Ruminococcus chamanellensis*, *Ruminococcus albus*, *Ruminococcus faecis*, and *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.

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.

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 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.





## 11

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 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 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 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 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.

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, 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

## 12

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.

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-alpha-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).

## 13

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 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-alpha-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

## 14

for delivery to the intestine. In some 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 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*, *Campylobacter*, Extended spectrum beta-lactamase (ESBL) producing *Enterobacteriaceae*, Multi-drug Resistant *Pseudomonas aeruginosa*, *Salmonella*, Drug resistant non-typoid *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 *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 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 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:

FIG. 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.

FIG. 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.

FIG. 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.

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

FIG. 5 shows experimental conditions evaluated in Example 2, the results for which are presented in FIGS. 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.

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

FIGS. 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. FIG. 7A shows weight of the mice that received no antibiotic treatment. FIG. 7B shows weight of the mice that received cefoperazone treatment. FIG. 7C shows weight of the mice that received cefoperazone treatment followed by vancomycin. FIG. 7D shows weight of the mice that received cefoperazone treatment followed by Composition I. FIG. 7E shows weight of

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

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

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

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

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

FIG. 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*.

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

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

FIG. 15 shows experimental conditions evaluated in Example 4, the results for which are presented in FIGS. 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.

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

FIGS. 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. FIG. 17A shows weight of the mice that received cefoperazone treatment. FIG. 17B shows weight of the mice that received cefoperazone treatment followed by FMT with fecal matter from a human. FIG. 17C shows weight of the mice that received cefoperazone treatment followed by FMT with fecal matter from a mouse. FIG. 17D shows weight of the mice that received cefoperazone treatment followed by Composition B on day -1. FIG. 17E shows weight of the mice that received cefoperazone treatment followed by Composition B on days -2 and -1. FIG. 17F shows weight of the mice that received cefoperazone treatment followed by Composition B on days -2, -1, 1, 2, and 3. FIG. 17G shows weight of the mice that received cefoperazone treatment followed by Composition F on day -1. FIG. 17H shows weight of the mice that received cefoperazone treatment followed by Composition F on days -2, -1, 1, 2, and 3.

FIGS. 18A-18B show the load of *C. difficile* in colony forming units (CFUs) in fecal pellets at various times post

infection with *C. difficile*. FIG. 18A shows *C. difficile* CFU/g feces 8 days post infection. FIG. 18B shows *C. difficile* CFU/g feces 17 days post infection.

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

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

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

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

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

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

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

FIG. 26 shows the strains in Composition H (SEQ ID NO:14—VE202-13—*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 sporadicum* (Cluster I); SEQ ID NO:81—PE5—*Clostridium bolteae* (Cluster XIVa); SEQ ID NO:80—VE202-18—*Erysipelatoclostridium ramosum* (Cluster XVIII).

FIGS. 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 FIG. 25. Mice losing >20% body weight of baseline were included in mortality numbers in survival curves. FIG. 29A shows survival/mortality of mice that received the indicated treatment prior to *C. difficile* infection. FIG. 29B shows the weight over time of mice that received the indicated treatment prior to *C. difficile* infection.

FIGS. 28A and 28B show results from the experimental conditions shown in FIG. 25. FIG. 28A shows survival/mortality of mice that received the indicated treatment prior

to *C. difficile* infection. FIG. 28B shows the weight over time of mice that received the indicated treatment prior to *C. difficile* infection.

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

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

FIG. 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. bifertamentans*, 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).

FIG. 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 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* 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* 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 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 that can be administered to subjects experiencing or having experienced a pathogenic infection to treat the infection. In some embodiments, the composi-

tions 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 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 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. 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 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 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) 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 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).

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 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 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 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 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: 1-83 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.

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.

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 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 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 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 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 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 formataxigens*, *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 cellulolyticus*, *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*.

*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 cellulolyticus*, *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, 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 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 formataxigens*, *Eisenbergiella tayi*, *Flavinofractor plautii*, *Clostridium orbiscindens* 1\_3\_50AFAA, *Lachnospiraceae bacterium* 7\_1\_58FAA, *Subdoligranulum*, *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 cellulolyticus*, *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, *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 formataxigens*, *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*, *Acidaminococ-*

*cus intestine, Ruminococcus faecis, Bacteroides cellulosilyticus, Anaerostipes hadrus, Eubacterium rectale, Ruminococcus chamanellensis, 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 masiliensis*, *Dracourtella massiliensis* GD1, *Ruminococcus torques*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryantia formaticigens*, *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., FIG. 1, Table A). As shown in FIG. 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, 5 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 10 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, 15 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 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 coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryantia formaticigens*, 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 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 coccoides*, *Eubacterium contortum*, *Eubacterium fissicatena*, *Anaerostipes caccae*, *Clostridium scindens*, *Marvinbryantia formataxigens*, and *Eisenbergiella tayi*. In some embodiments, the disclosure 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.

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 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 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: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 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 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, 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- $\Delta^4$ -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
30	SEQ_03 - 5 -	<i>Clostridium hathewayi</i> (XIVa)*
	SEQ_04 - 7 -	<i>Blautia hansenii</i> (XIVa)*
	SEQ_05 - 10 -	<i>Blautia hansenii</i> (XIVa)*
	SEQ_07 - 59 -	<i>Blautia producta/Blautia coccoides</i> (XIVa)
	SEQ_08 - 79 -	<i>Blautia hansenii</i> (XIVa)*
35	SEQ_09 - VE202-21 -	<i>Eubacterium contortum/Eubacterium fissicatena</i> (XIVa)*
	SEQ_11 - VE202-9 -	<i>Anerostipes caccae</i> (XIVa)*
	SEQ_12 - VE202-26-	<i>Clostridium scindens</i> (XIVa)*
	SEQ_13 - 136 -	<i>Marvinbryantia formatic exigens</i> (XIVa)*
	SEQ_23 - VE202-29 -	<i>Eisenbergiella tayi</i> (XIVa)*

In one aspect, the disclosure provides Composition B (See e.g., FIG. 1, Table B). As shown in FIG. 1, Composition B 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: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:









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: *Flavonifractor plautii*, *Lachnospiraceae*, *bacterium\_7\_1\_58FAA*, *Subdoligranulum Anaerotuncus colihominis*, *Eubacterium fissicatena*, *Ruminococcus torques*, *Clostridium symbiosum*, *Clostridium botteae*, *Dorea longicatena*, *Blaustia 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, Calif.) and sequences were assembled into whole genomes (Table 3). The 16S rDNA sequences were identified using Prokka and Barrnap. 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 botteae*, *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 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 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 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 *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 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: 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 N<sub>O</sub>: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 NO:22, and SEQ ID NO: 124-159. In some 10 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, 15 SEQ ID NO:18, 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 20 ID NO:18, 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:18, SEQ 25 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 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 30 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 35 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 40 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 45 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 50 NO:157.

55 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:  
60 151-159. 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  
65 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: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 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- $\Delta^4$ -cholenic 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

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% 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

40 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, 45 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

In some embodiments, the compositions disclosed herein do not include *Clostridium orbiscindens* 1\_3\_50FAA,  
65 *Flavinofractor plautii*, *Subdoligranulum* or *Lachnospiraceae bacterium* 7\_1\_58FAA. In some embodiments, the compositions disclosed herein do not include



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 - <i>Anaerotruncus colihominis</i> (IV)
SEQ_15 - VE202-14 - <i>Eubacterium fissicatena</i> (XIVa)
SEQ_16 - VE202-16 - <i>Clostridium symbiosum</i> (XIVa)
SEQ_17 - VE202-7 - <i>Clostridium bolteae</i> (XIVa)
SEQ_20 - 170 - <i>Dorea longicatena</i> (XIVa)
SEQ_19 - 16 - <i>Blaudia producta</i> (XIVa)
SEQ_21 - 189 - <i>Clostridium innocuum</i> (XVII)
SEQ_83 <i>Bacteroides ovatus</i>

In one aspect, the disclosure provides Composition C (See e.g., FIG. 1, Table C). As shown in FIG. 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 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, respectively.

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 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 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<sup>+</sup> strains and BaiCD<sup>-</sup> strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein one or more bacteria are BaiCD<sup>+</sup> strains and one or more bacteria are BaiCD<sup>-</sup> strains. In some embodiments of the one or more bacteria that are BaiCD<sup>+</sup> 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<sup>-</sup> 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<sup>+</sup> strains are selected from the bacterial species *Clostridium scindens*, *Clostridium hathewayi*, *Blautia hansenii*, *Blautia wexlerae*, *Blautia producta*, and *Blautia coccoides*. In some embodiments of the one or more bacteria that are BaiCD<sup>-</sup> 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<sup>-</sup> strains and BaiCD<sup>+</sup> 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<sup>-</sup> strains and BaiCD<sup>+</sup> 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<sup>-</sup> strains and BaiCD<sup>+</sup> strains and belong to *Clostridium* clusters IV or XIVa.

TABLE C

## Composition C

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

\*= BaiCD<sup>+</sup>

In one aspect, the disclosure provides Composition D (See e.g., FIG. 1, Table D). As shown in FIG. 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 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 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 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*, *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 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 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 having at least 97% homology with a nucleic acid sequence of SEQ ID NO:10.

15 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 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.

20 25 The strains of Composition D include both BaiCD<sub>+</sub> strains and Bai CD<sub>-</sub> strains. In some embodiments, the disclosure provides compositions comprising two or more bacteria, wherein one or more bacteria are BaiCD<sub>+</sub> strains and one or more bacteria are BaiCD<sub>-</sub> strains. In some 30 35 embodiments of the one or more bacteria that are BaiCD<sub>+</sub> 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<sub>-</sub> 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 40 45 more bacteria that are BaiCD<sub>+</sub> 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<sub>-</sub> 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<sub>-</sub> strains and BaiCD<sub>+</sub> 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<sub>-</sub> strains and BaiCD<sub>+</sub> strains and belong to *Clostridium* clusters XIVa or XVII. In some 50 55 embodiments, the disclosure provides two or more bacterial strains, wherein the bacteria are BaiCD<sub>-</sub> strains and BaiCD<sub>+</sub> strains and belong to *Clostridium* clusters IV or XIVa.

Composition D includes the non-*Clostridium* strains 60 *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. amyloyticus*, *L. amylophilus*, *L. amylotrophicus*, *L. amylovorus*, *L. animalis*, *L. antri*, *L. apodemi*, *L. aviarius*, *L. bifermentans*, *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. crustum*, *L. curvatus*, *L. delbrueckii* subsp. *bulgaricus*, *L. delbrueckii* subsp. *delbrueckii*, *L. delbrueckii* subsp. *lactis*, *L. dextrinicu*s, *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. gas-tricus*, *L. ghanensis*, *L. graminis*, *L. hampesi*, *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. kefirancaciens*, *L. kefiri*, *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. parasarraginis*, *L. parake-firi*, *L. paralimentarius*, *L. paraplanitarum*, *L. pentosus*, *L. perolens*, *L. plantarum*, *L. pontis*, *L. protectus*, *L. psittaci*, *L. rennini*, *L. reuteri*, *L. rhamnosus*, *L. rimae*, *L. rogosae*, *L. rossiae*, *L. ruminis*, *L. saerimneri*, *L. sakei*, *L. salivarius*, *L. sanfranciscensis*, *L. satsumensis*, *L. secaliphilus*, *L. sharp-eae*, *L. siliginis*, *L. spicheri*, *L. suebicus*, *L. thailandensis*, *L. ultunensis*, *L. vaccinostercus*, *L. vaginalis*, *L. versmoldensis*, *L. vini*, *L. vitulinus*, *L. zeae*, 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
20	SEQ_12 - VE202-26 - <i>Clostridium_scindens</i> (XIVa)*
	SEQ_03 - 5 - <i>Clostridium_hathewayi</i> (XIVa)*
	SEQ_05 - 10 - <i>Blautia_hansenii</i> (XIVa)*
	SEQ_01 - 71 - <i>Blautia_wexlerae</i> (XIVa)*
	SEQ_14 - VE202- 13 - <i>Anaerotruncus_colihominis</i> (IV)
25	SEQ_18 - 148 - <i>Dorea_longicatena</i> (XIVa)
	SEQ_21 - 189 - <i>Clostridium_innocuum</i> (XVII)
	SEQ_10 - 211 - <i>Flavonifractor_plautii</i> (IV)
	SEQ_02 - 102 - <i>Turicibacter_sanguinis</i> (non- <i>Clostridium</i> )
	SEQ_06 - 40 - <i>Lactobacillus_mucosae</i> (non- <i>Clostridium</i> )

\*= BaiCD+

30 In one aspect, the disclosure provides Composition F (See e.g., FIGS. 13 and 14, and Tables F1 and F2). As shown in FIG. 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 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.

45 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.

50 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*, *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*, *Dorea longicatena*, *Roseburia faecis*, *Blautia luti*, *Fusicatenibacter saccharivorans*, *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*, *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*, *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*, *Ruminococcus faecis*, *Ruminococcus faecis*, *Dorea longicatena*, *Roseburia faecis*, *Blautia luti*, *Fusicatenibacter saccharivorans*, *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 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.

position 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

TABLE F1

Composition F					
SEQ_NO	StrainID	Genus_species	SEQ_NO	StrainID	Genus_species
SEQ_24	YK96	<i>Dorea longicatena</i>	SEQ_52	YK51	<i>Eubacterium rectale</i>
SEQ_25	YK101	<i>Ruminococcus obeum</i>	SEQ_53	YK52	<i>Eubacterium rectale</i>
SEQ_26	YK110	<i>Megasphaera elsdenii</i>	SEQ_54	YK54	<i>Anaerostipes hadrus</i>
SEQ_27	YK149	<i>Acidaminococcus fermentans/ Acidaminococcus intestini</i>	SEQ_55	YK56	<i>Ruminococcus faecis</i>
SEQ_28	YK154	<i>Megasphaera elsdenii</i>	SEQ_56	YK57	<i>Ruminococcus faecis</i>
SEQ_29	YK36	<i>Ruminococcus faecis</i>	SEQ_57	YK58	<i>Dorea longicatena</i>
SEQ_30	YK95	<i>Bacteroides cellullosilyticus</i>	SEQ_58	YK65	<i>Roseburia faecis</i>
SEQ_31	YK32	<i>Anaerostipes hadrus</i>	SEQ_59	YK67	<i>Blautia luti</i>
SEQ_32	YK64	<i>Ruminococcus obeum</i>	SEQ_60	YK69	<i>Fusicatenibacter saccharivorans</i>
SEQ_33	YK73	<i>Flavonifractor plautii</i>	SEQ_61	YK70	<i>Fusicatenibacter saccharivorans</i>
SEQ_34	YK87	<i>Eubacterium rectale</i>	SEQ_62	YK71	<i>Roseburia faecis</i>
SEQ_35	YK105	<i>Flavonifractor plautii</i>	SEQ_63	YK74	<i>Megasphaera elsdenii</i>
SEQ_36	YK153	<i>Megasphaera elsdenii</i>	SEQ_64	YK88	<i>Eubacterium rectale</i>
SEQ_37	YK163	<i>Eubacterium rectale</i>	SEQ_65	YK89	<i>Eubacterium rectale</i>
SEQ_38	YK191	<i>Ruminococcus chamanellensis/ Ruminococcus albus</i>	SEQ_66	YK97	<i>Roseburia faecis</i>
SEQ_39	YK99	<i>Ruminococcus chamanellensis</i>	SEQ_67	YK98	<i>Blautia faecis</i>
SEQ_40	YK55	<i>Ruminococcus faecis</i>	SEQ_68	YK139	<i>Fusicatenibacter saccharivorans</i>
SEQ_41	YK75	<i>Bifidobacterium bifidum</i>	SEQ_69	YK141	<i>Dorea formicigenerans</i>
SEQ_42	YK90	<i>Anaerostipes hadrus</i>	SEQ_70	YK142	<i>Ruminococcus faecis</i>
SEQ_43	YK30	<i>Anaerostipes hadrus</i>	SEQ_71	YK152	<i>Blautia hansenii</i>
SEQ_44	YK31	<i>Anaerostipes hadrus</i>	SEQ_72	YK155	<i>Blautia hansenii</i>
SEQ_45	YK12	<i>Eubacterium rectale</i>	SEQ_73	YK157	<i>Eubacterium rectale</i>
SEQ_46	YK27	<i>Ruminococcus faecis</i>	SEQ_74	YK160	<i>Roseburia faecis</i>
SEQ_47	YK28	<i>Blautia luti</i>	SEQ_75	YK166	<i>Eubacterium rectale</i>
SEQ_48	YK29	<i>Ruminococcus faecis</i>	SEQ_76	YK168	<i>Eubacterium rectale</i>
SEQ_49	YK33	<i>Anaerostipes hadrus</i>	SEQ_77	YK169	<i>Eubacterium rectale</i>
SEQ_50	YK34	<i>Anaerostipes hadrus</i>	SEQ_78	YK171	<i>Eubacterium rectale</i>
SEQ_51	YK35	<i>Ruminococcus faecis</i>	SEQ_79	YK192	<i>Roseburia faecis</i>

TABLE F2

Composition F, strain groupings		
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
IV	<i>Dorea longicatena</i> 2	A
	<i>Flavonifractor plautii</i> 2	A, B
	<i>Ruminococcus chamanellensis</i> 2	A
IX	<i>Acidaminococcus fermentans</i> 1	A, B, P
	<i>Megasphaera elsdenii</i> 4	P
other	<i>Bacteroides cellullosilyticus</i> 1	A, S
	<i>Bifidobacterium 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., FIG. 19; Table G). As shown in FIG. 19, Com-

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

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

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

In one aspect, the disclosure provides Composition H (See e.g., FIG. 26, Table H). As shown in FIG. 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.

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 *Erysipeloctostridium 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 *Erysipeloctostridium 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
55			
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
60			
SEQ ID NO: 80	VE202-18	<i>Erysipeloctostridium 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 *Collinsella*. 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, Wis.), 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, FIGS. 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., 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 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 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.

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).

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., U.S. Pat. No. 6,635,260). The compositions disclosed herein can be distinguished 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 $\alpha$ / $\beta$ -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 $\alpha$  dihydroxylation is mediated by a stereo-specific NAD (H)-dependent 3-oxo- $\Delta^4$ -cholenic 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- $\Delta^4$ -cholenic acid oxidoreductase activity of BaiCD protein. The NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenic acid oxidoreductase activity of a bacterial strain may be assessed by methods such as measuring the amount of 7 $\alpha$ -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 *Pseudoflavinofractor 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 *Acetovibrio ethanolignens*. In some embodiments, the compositions provided herein do not 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. 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 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 *Pseudobutyribrio*. 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 compositions do not comprise *E. coli*. In some embodiments, the compositions do not comprise fungi, such as *Monilia* 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 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, 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 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 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 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 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 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 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.

5 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 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.

60 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 compo-

sitions 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, cefoxintin, fructose agar plates), in which taurocholate helps the spores to germinate. In addition, spores can be quantified using the dipicolinic assay (DPA assay). 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 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.

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.

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 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.

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.

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 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 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 a microbiome associated with a disease state). In 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 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, 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*; *Escherichia coli*; *Pseudomonas aeruginosa*, such as Multidrug Resistant *Pseudomonas aeruginosa*; Vancomycin Resistant Enterococci (VRE); Carbapenem Resistant Enterobacteriaceae (CRE); *Neisseria gonorrhoeae*; *Acinetobacter*; Multi-drug Resistant *Acinetobacter*; *Campylobacter*; Multi-drug resistant *Campylobacter*; *Candida*; Fluconazole-resistant *Candida*; Extended spectrum beta-lactamase (ESBL) producing Enterobacteriaceae; *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. 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 epidermidis*, *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 cholerasuis*, *Salmonella enterica*, *Salmonella enteriditis*, *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 tuberculosis*, *Mycobacterium leprae*, *Chlamydia psittaci*, *Chlamydia trachomatis*, *Chlamydia pneumoniae*, *Rickettsia rickettsii*, *Rickettsia akari*, *Rickettsia prowazekii*, *Brucella abortus*, *Brucella melitensis*, *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<sup>4</sup>-fold, 10<sup>5</sup>-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., U.S. Pat. Nos. 3,261,761; 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 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 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 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 U.S. Pat. No. 6,368,586) and compositions into which a protease inhibitor is incorporated for 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 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 10 that includes a coating that can be removed by an enzyme 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 20 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 25 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 30 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, 35 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 40 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 45 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 50 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 55 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



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  $10^9$  and  $10^9$ , 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  $10^8$  and  $10^8$ , between  $10^9$  and  $10^8$ , 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  $10^7$  and  $10^7$ , between  $10^8$  and  $10^7$ , between  $10^9$  and  $10^7$ , between  $10^{10}$  and  $10^7$ , 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  $10^6$  and  $10^6$ , between  $10^7$  and  $10^6$ , between  $10^8$  and  $10^6$ , between  $10^9$  and  $10^6$ , between  $10^{10}$  and  $10^6$ , between  $10^2$  and  $10^5$ , between  $10^3$  and  $10^5$ , between  $10^4$  and  $10^5$ , between  $10^5$  and  $10^5$ , between  $10^6$  and  $10^5$ , between  $10^7$  and  $10^5$ , between  $10^8$  and  $10^5$ , between  $10^9$  and  $10^5$ , between  $10^{10}$  and  $10^5$ , between  $10^2$  and  $10^4$ , between  $10^3$  and  $10^4$ , between  $10^4$  and  $10^4$ , between  $10^5$  and  $10^4$ , between  $10^6$  and  $10^4$ , between  $10^7$  and  $10^4$ , between  $10^8$  and  $10^4$ , between  $10^9$  and  $10^4$ , between  $10^{10}$  and  $10^4$ , between  $10^2$  and  $10^3$ , between  $10^3$  and  $10^3$ , between  $10^4$  and  $10^3$ , between  $10^5$  and  $10^3$ , between  $10^6$  and  $10^3$ , between  $10^7$  and  $10^3$ , between  $10^8$  and  $10^3$ , between  $10^9$  and  $10^3$ , between  $10^{10}$  and  $10^3$ , between  $10^2$  and  $10^2$ , between  $10^3$  and  $10^2$ , between  $10^4$  and  $10^2$ , between  $10^5$  and  $10^2$ , between  $10^6$  and  $10^2$ , between  $10^7$  and  $10^2$ , between  $10^8$  and  $10^2$ , between  $10^9$  and  $10^2$ , between  $10^{10}$  and  $10^2$ , between  $10^2$  and  $10^1$ , between  $10^3$  and  $10^1$ , between  $10^4$  and  $10^1$ , between  $10^5$  and  $10^1$ , between  $10^6$  and  $10^1$ , between  $10^7$  and  $10^1$ , between  $10^8$  and  $10^1$ , between  $10^9$  and  $10^1$ , between  $10^{10}$  and  $10^1$ , between  $10^2$  and  $10^0$ , between  $10^3$  and  $10^0$ , between  $10^4$  and  $10^0$ , between  $10^5$  and  $10^0$ , between  $10^6$  and  $10^0$ , between  $10^7$  and  $10^0$ , between  $10^8$  and  $10^0$ , between  $10^9$  and  $10^0$ , between  $10^{10}$  and  $10^0$ , between  $10^2$  and  $10^{-1}$ , between  $10^3$  and  $10^{-1}$ , between  $10^4$  and  $10^{-1}$ , between  $10^5$  and  $10^{-1}$ , between  $10^6$  and  $10^{-1}$ , between  $10^7$  and  $10^{-1}$ , between  $10^8$  and  $10^{-1}$ , between  $10^9$  and  $10^{-1}$ , between  $10^{10}$  and  $10^{-1}$ , between  $10^2$  and  $10^{-2}$ , between  $10^3$  and  $10^{-2}$ , between  $10^4$  and  $10^{-2}$ , between  $10^5$  and  $10^{-2}$ , between  $10^6$  and  $10^{-2}$ , between  $10^7$  and  $10^{-2}$ , between  $10^8$  and  $10^{-2}$ , between  $10^9$  and  $10^{-2}$ , between  $10^{10}$  and  $10^{-2}$ , between  $10^2$  and  $10^{-3}$ , between  $10^3$  and  $10^{-3}$ , between  $10^4$  and  $10^{-3}$ , between  $10^5$  and  $10^{-3}$ , between  $10^6$  and  $10^{-3}$ , between  $10^7$  and  $10^{-3}$ , between  $10^8$  and  $10^{-3}$ , between  $10^9$  and  $10^{-3}$ , between  $10^{10}$  and  $10^{-3}$ , between  $10^2$  and  $10^{-4}$ , between  $10^3$  and  $10^{-4}$ , between  $10^4$  and  $10^{-4}$ , between  $10^5$  and  $10^{-4}$ , between  $10^6$  and  $10^{-4}$ , between  $10^7$  and  $10^{-4}$ , between  $10^8$  and  $10^{-4}$ , between  $10^9$  and  $10^{-4}$ , between  $10^{10}$  and  $10^{-4}$ , between  $10^2$  and  $10^{-5}$ , between  $10^3$  and  $10^{-5}$ , between  $10^4$  and  $10^{-5}$ , between  $10^5$  and  $10^{-5}$ , between  $10^6$  and  $10^{-5}$ , between  $10^7$  and  $10^{-5}$ , between  $10^8$  and  $10^{-5}$ , between  $10^9$  and  $10^{-5}$ , between  $10^{10}$  and  $10^{-5}$ , 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  $10^{-2}$  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  $10^{-3}$  and  $10^{-3}$ , between  $10^{-2}$  and  $10^{-3}$ , between  $10^{-7}$  and  $10^{-4}$ , between  $10^{-6}$  and  $10^{-4}$ , between  $10^{-5}$  and  $10^{-4}$ , between  $10^{-4}$  and  $10^{-4}$ , between  $10^{-3}$  and  $10^{-4}$ , between  $10^{-2}$  and  $10^{-4}$ , between  $10^{-7}$  and  $10^{-5}$ , between  $10^{-6}$  and  $10^{-5}$ , between  $10^{-5}$  and  $10^{-5}$ , between  $10^{-4}$  and  $10^{-5}$ , between  $10^{-3}$  and  $10^{-5}$ , between  $10^{-2}$  and  $10^{-5}$ , or between  $10$  and  $10^2$  total CFUs per dosage amount.

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  $10^{-2}$  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  $10^{-3}$  and  $10^{-3}$ , between  $10^{-2}$  and  $10^{-3}$ , between  $10^{-7}$  and  $10^{-4}$ , between  $10^{-6}$  and  $10^{-4}$ , between  $10^{-5}$  and  $10^{-4}$ , between  $10^{-4}$  and  $10^{-4}$ , between  $10^{-3}$  and  $10^{-4}$ , between  $10^{-2}$  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 and 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	<i>Blautia wexlerae</i>	SEQ_94	96.62	207	NR_044054
SEQ ID NO: 02	102	<i>Turicibacter_sanguinis</i>	SEQ_91	97.81	183	NR_028816
SEQ ID NO: 03	5	<i>Clostridium hathewayi</i>	SEQ_105	92.42	198	NR_036928
SEQ ID NO: 04	7	<i>Blautia_hansenii</i>	SEQ_99	96.62	207	NR_104687
SEQ ID NO: 05	10	<i>Blautia_hansenii</i>	SEQ_99	98.06	206	NR_104687
SEQ ID NO: 06	40	<i>Lactobacillus_mucosae</i>	SEQ_90	87.57	185	NR_024994
SEQ ID NO: 07	59	<i>Blautia_producta</i>	SEQ_106	98.54	206	NR_113270
SEQ ID NO: 07	59	<i>Blautia_coccooides</i>	SEQ_103	98.54	206	NR_104700
SEQ ID NO: 08	79	<i>Blautia_hansenii</i>	SEQ_99	100	194	NR_104687
SEQ ID NO: 09	VE202-21	<i>Eubacterium_contortum</i>	SEQ_109	94.59	296	NR_117147
SEQ ID NO: 09	VE202-21	<i>Eubacterium_fissicatena</i>	SEQ_108	94.59	296	NR_117142
SEQ ID NO: 10	211	<i>Flavonifractor_plautii</i>	SEQ_93	98.49	199	NR_043142
SEQ ID NO: 11	VE202-9	<i>Anaerostipes_caccae</i>	SEQ_88	99.5	399	NR_028915
SEQ ID NO: 12	VE202-26	<i>Clostridium_scindens</i>	SEQ_87	95.76	354	NR_028785
SEQ ID NO: 13	136	<i>Marvinbryantia_formataxigena</i>	SEQ_89	94.66	131	NR_042152
SEQ ID NO: 14	VE202-13	<i>Anaerotruncus_colihominis</i>	SEQ_95	99.34	1365	NR_027558
SEQ ID NO: 15	VE202-14	<i>Eubacterium_fissicatena</i>	SEQ_102	93.33	1530	NR_117563
SEQ ID NO: 16	VE202-16	<i>Clostridium_symbiosum</i>	SEQ_122	98.43	1469	NR_118730
SEQ ID NO: 17	VE202-7	<i>Clostridium_bolteae</i>	SEQ_110	99.86	1390	NR_113410
SEQ ID NO: 18	148	<i>Dorea_longicatena</i>	SEQ_97	99.7	1318	NR_028883
SEQ ID NO: 19	16	<i>Blautia_producta</i>	SEQ_106	98.33	1493	NR_113270
SEQ ID NO: 20	170	<i>Dorea_longicatena</i>	SEQ_97	99.7	1318	NR_028883
SEQ ID NO: 21	189	<i>Clostridium_innocuum</i>	SEQ_98	98.64	1476	NR_029164
SEQ ID NO: 22	169	<i>Dorea_longicatena</i>	SEQ_97	99.58	475	NR_028883
SEQ ID NO: 23	VE202-29	<i>Eisenbergiella_tayi</i>	SEQ_121	100	354	NR_118643
SEQ ID NO: 24	YK96	<i>Dorea_longicatena</i>	SEQ_97	99.48	191	NR_028883
SEQ ID NO: 25	YK101	<i>Ruminococcus_obeum</i>	SEQ_85	96.81	188	NR_118692
SEQ ID NO: 26	YK110	<i>Megasphaera_elsdenii</i>	SEQ_119	96.62	207	NR_102980
SEQ ID NO: 27	YK149	<i>Acidaminococcus_fermentans</i>	SEQ_115	99.48	192	NR_074928
SEQ ID NO: 27	YK149	<i>Acidaminococcus_intestini</i>	SEQ_112	99.48	192	NR_074306
SEQ ID NO: 28	YK154	<i>Megasphaera_elsdenii</i>	SEQ_119	96.12	206	NR_102980
SEQ ID NO: 29	YK36	<i>Ruminococcus_faecis</i>	SEQ_96	99.29	425	NR_116747
SEQ ID NO: 30	YK95	<i>Bacteroides_cellulosilyticus</i>	SEQ_100	99.54	437	NR_112933
SEQ ID NO: 31	YK32	<i>Anaerostipes_hadrus</i>	SEQ_107	98.8	415	NR_104799
SEQ ID NO: 32	YK64	<i>Ruminococcus_obeum</i>	SEQ_84	99.04	415	NR_119185
SEQ ID NO: 33	YK73	<i>Flavonifractor_plautii</i>	SEQ_93	98.56	418	NR_043142
SEQ ID NO: 34	YK87	<i>Eubacterium_rectale</i>	SEQ_114	99.52	416	NR_074634
SEQ ID NO: 35	YK105	<i>Flavonifractor_plautii</i>	SEQ_93	99.26	407	NR_043142
SEQ ID NO: 36	YK153	<i>Megasphaera_elsdenii</i>	SEQ_119	96.04	429	NR_102980
SEQ ID NO: 37	YK163	<i>Eubacterium_rectale</i>	SEQ_114	99.76	415	NR_074634
SEQ ID NO: 38	YK191	<i>Ruminococcus_champanellensis</i>	SEQ_117	94.47	416	NR_102884
SEQ ID NO: 38	YK191	<i>Ruminococcus_albus</i>	SEQ_113	94.47	416	NR_074399
SEQ ID NO: 39	YK99	<i>Ruminococcus_champanellensis</i>	SEQ_117	97.28	184	NR_102884
SEQ ID NO: 40	YK55	<i>Ruminococcus_faecis</i>	SEQ_96	99.02	408	NR_116747
SEQ ID NO: 41	YK75	<i>Bifidobacterium_bifidum</i>	SEQ_118	99.45	183	NR_102971
SEQ ID NO: 42	YK90	<i>Anaerostipes_hadrus</i>	SEQ_107	98.97	194	NR_104799
SEQ ID NO: 43	YK30	<i>Anaerostipes_hadrus</i>	SEQ_107	99.48	191	NR_104799
SEQ ID NO: 44	YK31	<i>Anaerostipes_hadrus</i>	SEQ_107	98.97	194	NR_104799

TABLE 1-continued

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: 45	YK12	<i>Eubacterium_rectale</i>	SEQ_114	99.27	412	NR_074634
SEQ ID NO: 46	YK27	<i>Ruminococcus_faecis</i>	SEQ_96	99.51	412	NR_116747
SEQ ID NO: 47	YK28	<i>Blautia_luti</i>	SEQ_111	99.5	400	NR_041960
SEQ ID NO: 48	YK29	<i>Ruminococcus_faecis</i>	SEQ_96	99.03	413	NR_116747
SEQ ID NO: 49	YK33	<i>Anaerostipes_hadrus</i>	SEQ_107	99.27	413	NR_104799
SEQ ID NO: 50	YK34	<i>Anaerostipes_hadrus</i>	SEQ_107	99.51	410	NR_104799
SEQ ID NO: 51	YK35	<i>Ruminococcus_faecis</i>	SEQ_96	99.51	409	NR_116747
SEQ ID NO: 52	YK51	<i>Eubacterium_rectale</i>	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 53	YK52	<i>Eubacterium_rectale</i>	SEQ_114	99.03	413	NR_074634
SEQ ID NO: 54	YK54	<i>Anaerostipes_hadrus</i>	SEQ_107	85.82	409	NR_104799
SEQ ID NO: 55	YK56	<i>Ruminococcus_faecis</i>	SEQ_96	99.03	413	NR_116747
SEQ ID NO: 56	YK57	<i>Ruminococcus_faecis</i>	SEQ_96	98.79	413	NR_116747
SEQ ID NO: 57	YK58	<i>Dorea_longicatena</i>	SEQ_97	98.8	417	NR_028883
SEQ ID NO: 58	YK65	<i>Roseburia_faecis</i>	SEQ_92	99.27	413	NR_042832
SEQ ID NO: 59	YK67	<i>Blautia_luti</i>	SEQ_111	98.57	419	NR_041960
SEQ ID NO: 60	YK69	<i>Fusicatenibacter_saccharivorans</i>	SEQ_116	99.27	413	NR_114326
SEQ ID NO: 61	YK70	<i>Fusicatenibacter_saccharivorans</i>	SEQ_116	98.79	414	NR_114326
SEQ ID NO: 62	YK71	<i>Roseburia_faecis</i>	SEQ_92	99.28	414	NR_042832
SEQ ID NO: 63	YK74	<i>Megasphaera_elisdenii</i>	SEQ_119	96.06	431	NR_102980
SEQ ID NO: 64	YK88	<i>Eubacterium_rectale</i>	SEQ_114	99.28	415	NR_074634
SEQ ID NO: 65	YK89	<i>Eubacterium_rectale</i>	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 66	YK97	<i>Roseburia_faecis</i>	SEQ_92	99.28	414	NR_042832
SEQ ID NO: 67	YK98	<i>Blautia_faecis</i>	SEQ_104	98.02	405	NR_109014
SEQ ID NO: 68	YK139	<i>Fusicatenibacter_saccharivorans</i>	SEQ_116	99.03	412	NR_114326
SEQ ID NO: 69	YK141	<i>Dorea_formicifex</i>	SEQ_120	98.51	402	NR_044645
SEQ ID NO: 70	YK142	<i>Ruminococcus_faecis</i>	SEQ_96	98.79	413	NR_116747
SEQ ID NO: 71	YK152	<i>Blautia_hansenii</i>	SEQ_99	99.5	401	NR_104687
SEQ ID NO: 72	YK155	<i>Blautia_hansenii</i>	SEQ_99	98.79	413	NR_104687
SEQ ID NO: 73	YK157	<i>Eubacterium_rectale</i>	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 74	YK160	<i>Roseburia_faecis</i>	SEQ_92	99.03	414	NR_042832
SEQ ID NO: 75	YK166	<i>Eubacterium_rectale</i>	SEQ_114	99.27	409	NR_074634
SEQ ID NO: 76	YK168	<i>Eubacterium_rectale</i>	SEQ_114	99.27	413	NR_074634
SEQ ID NO: 77	YK169	<i>Eubacterium_rectale</i>	SEQ_114	99.28	416	NR_074634
SEQ ID NO: 78	YK171	<i>Eubacterium_rectale</i>	SEQ_114	97.87	188	NR_074634
SEQ ID NO: 79	YK192	<i>Roseburia_faecis</i>	SEQ_92	99.03	414	NR_042832
SEQ ID NO: 80	VE202-18	<i>Erysipelotrichobacter_ramosum</i>	SEQ_123	100	1485	NR_113243
SEQ ID NO: 81	PE5	<i>Clostridium_bolteae</i>	SEQ_110	100	1385	NR_113410
SEQ ID NO: 82	PE9	<i>Clostridium_disporicum</i>	SEQ_86	99.21	382	NR_026491
SEQ ID NO: 83	211-B	<i>Bacteroides_ovatus</i>	SEQ_101	95.64	436	NR_112940

40

TABLE 2

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

Strain	Whole genome homology
SEQ_10-211	<i>Lachnospiraceae bacterium 7_1_58FAA</i>
	<i>Subdoligranulum</i>
	<i>Flavinifractor plautii</i>
SEQ_14-VE202-13	<i>Anaerotruncus colihominis</i>
SEQ_15-VE202-14	<i>Eubacterium_fissicatena</i>
SEQ_16-VE202-16	<i>Ruminococcus torques</i>
	<i>Clostridium symbiosum</i>

TABLE 2-continued

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

Strain	Whole genome homology
45 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>
50 SEQ_21-189	<i>Clostridium innocuum</i>
	<i>Erysipelotrichaceae bacterium_21_3</i>

TABLE 3

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

Compo- sition B strain number	Strain iden- tifier	SEQ ID # Closest of 16S region as deter- mined by sequencing Sanger se- quencing region	Closest species based on Sanger deter- mined by sequencing region as deter- mined by WGS^	SEQ ID # of 16S regions as deter- mined by WGS	*Consensus SEQ ID # of 16S regions as deter- mined by WGS	Closest species based on WGS database	Closest species based on WGS database	Additional closely related sequences	Clos- tridium cluster
1	VE202-7	17	<i>Clostridium bolteae</i>	124, 125, 126, 127, 128	124	<i>Clostridium bolteae</i>	<i>Clostridium bolteae</i> 90A9		XIVa

TABLE 3-continued

Bacterial species with highest degree of homology based on whole genome analysis										
Composition B strain number	Strain identifier	SEQ ID # of 16S region as deter- mined by sequencing Sanger se- quencing	Closest species based on Sanger sequencing region	SEQ ID # of 16S regions as determined by WGS <sup>a</sup>	*Consensus SEQ ID # of 16S region as determined by WGS	Closest species based on Concensus SEQ ID # of 16S region as compared with 16S database	Closest species based on WGS compared versus WG databases	Additional closely related sequences	Clos- tridium cluster	
2	VE202-13	14	<i>Anaerotruncus colihominis</i>	129,130, 131	129	<i>Anaerotruncus colihominis</i>	<i>Anaerotruncus colihominis</i> DSM 17241		IV	
3	VE202-14	15	<i>Eubacterium fissicatena</i>	132, 133, 134, 135, 136	132	<i>Drancourtella massiliensis</i>	<i>Drancourtella massiliensis</i> GD1	<i>Ruminococcus torques</i> ; <i>Sellimonas intestinalis</i>	XIVa	
4	VE202-16	16	<i>Clostridium symbiosum</i>	137, 138, 139, 140	137	<i>Clostridium symbiosum</i>	<i>Clostridium symbiosum</i> WAL-14163		XIVa	
5	strain #16	19	<i>Blautia producta</i>	141, 142, 143, 144, 145	141	<i>Blautia producta</i>	<i>Clostridium bacterium</i> UC5.1-1D4	<i>Blautia product</i> ATCC 27340	XIVa	
6	strain #170	20	<i>Dorea longicatena</i>	146, 147, 148, 149, 150, 151	146	<i>Dorea longicatena</i>	<i>Dorea longicatena</i> CAG:42		XIVa	
7	strain #189	21	<i>Clostridium innocuum</i>	152, 153, 154, 155, 156	152	<i>Clostridium innocuum</i>	<i>Erysipelotrichaceae bacterium</i> 21_3		XVII	
8	strain #211	10	<i>Flavonifractor plautii</i>	157, 158, 159	157	<i>Flavonifractor plautii</i>	<i>Clostridium orbiscindens</i> 1_3_50AFAA	<i>Subdolinogranulum</i>	IV	

<sup>a</sup>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 intestinalis*, *Drancourtella massiliensis* or *Drancourtella massiliensis* GP1. 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, 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, 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, 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, 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.

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, 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*, *Anaerotruncus 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*, *Anaerotruncus 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 |  
GCCCGGAGCAGTTGATGTGAAGGATGGGTACCTGTGGACTGCATTGAACTGTCTGAGTGCCGGAGGGTAA  
GCGGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGCTTACTGGACGGT  
AACTGACGTTGAGGCTCGAACAGCGTGGGAGCAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 02 | 102 |  
CTAACCGTGGAGGTCAATTGGAAACTGGTCAACTTGAGTGCGAGAAGAGGGAAAGTGGAAATTCCATGTGTAGCGGTGAAA  
TGGCTAGAGATATGGAGGAACACCACTGGCGAAGGCGCTCCCTGGTCTGTAACGTGACACTGAGGCAGCAAAGCGT  
GGGGCAAACAGGATTAGATCCCCGGTAA

-continued

SEQ ID NO: 03|5|

ATGAAAGCCGGGGCTCAACCCCGTACTGCTTGAAACTGTTGACTTGAGTGCTTGAGAGGTAAGTGAATTCT  
 ATGTGTAGCGGGAAATGTTAGATATTAGGAGGACACCAGTGGCGAAGGCCTTACTGGAAGTGAACGTGACGTTG  
 GCTCGATTGTGGGAGCAAACAGGATTATATCCCCTGGTAA

SEQ ID NO: 04|7|

CGGAAGGTCTGATGTGAAGGTTGGGCTTACCCCGACTGCATTGAAACTGTTTCTAGAGTGCCGAGAGGTA  
 CGCGAATTCTAGTGTAGCGGTGAAATGCTTAGATATTAGGAGGACACCAGTGGCGAAGGCCTTACTGGACG  
 TAAGTGAACGTTGAGGCTCGAACAGTGGGAGCAAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 05|10|

CGATGTCTGAGTGAAGGCTGGGCTTACCCCGACTGCATTGAAACTGTTTCTAGAGTGCCGAGAGGTAAGCG  
 GAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGACACCAGTGGCGAAGGCCTTACTGGACG  
 TAACGTTGAGGCTCGAACAGTGGGAGCAAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 06|40|

TTAACCAAGAAGTGCATTGAACTGTCAGACTGGGGAAAAAAAGACAGTGCACCTCATGTTAGCGGTGAAATG  
 CTCCATATATATGGAAGAACACCAGTGGCGAAGGCCTGTCGGCTGCAACTGACGCTGAGGCTCGAATT  
 GTAAGAAAAGTATTAGTCCCTTGAA

SEQ ID NO: 07|59|

ACCCGCTTGGCTGAGGTGAGGCTGGGCTTAACCCAGGACTGCATTGAAACTGTTTAGAGTGCCTGAGGAG  
 GTAAAGCGGAATTCTAGTGTAGCGTAGATATTAGGAGGACACCAGTGGCGAAGGCCTTACTGGACGGTA  
 CGGTAACTGACGTTGAGGCTCGAACAGTGGGAGCAAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 08|79|

TAGGCTGGGCTTAACCCAGGACTGCATTGAAACTGTTTCTAGAGTGCCTGAGGAGGTAAGCGGAATTCTAGT  
 GTAGCGGTGAAATGCGTAGATATTAGGAGGACACCAGTGGCGAAGGCCTTACTGGACGGTA  
 CTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 09

|VE202-21|  
 TTGCATTGGACACTATGTCAGCTGAGTGTGGAGAGGTAAGTGAATTCTAGTGTAGCGGTGAAATGCGTAGAT  
 TAGGAGGAAACACCAGTGGCGAAGGGCTTACTGCACGTTCTGACGTTGAGGCTCGAAATCGTGGGAGCAA  
 AAAATAGATAACCTGGTAGTCCACGCCAACGATGCATACTAGGTGCGGCTGCAAGGCCATTGGTGC  
 CAGCAAACCGCAATAAGTATGCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAATAATTGACG  
 GA

SEQ ID NO: 10|211|

CCCGTCGTAGATGTGAACTGGGGCTCACCTCCAGCCTGCATTGAAACTGTTAGTCTGAGTGC  
 CGGAATTCCGTGTAGCGGTGAAATGCGTAGATATACGGAGGAAACCCAGTGGCGAAGGCCTTACTGG  
 ACTGACGCTGAGGCGCGAACAGTGGGAGCAAACAGGATTAGATAACCTCATAA

SEQ ID NO: 11

|VE202-9|  
 ACCTGATGCAGCGACGCCGCGTGAAGTAAAGTATTCGGTATGTAAGCTCTACGAGGGAAAGAAAAAGACG  
 GTACCTGACTAAGAAGCCCCGGTAACGTGCCAGCAGCCCGTAATACGTAGGGGCAAGCGTTATCCGAAT  
 TACTGGGTGAAAGGGTGCAGGTGGCATGGTAAGTCAGAAGTGAAGGCTTAACCCGGACTGCTTT  
 GAAAATGTCATGCTGGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGG  
 ACACCAAGTGGCGAAGGGCTTACTGGACTGTCACTGACACTGATGCACGAAAGCGTGGGAGCAAACAGGATTAGA  
 TACCCCTGGAAGTCCAT

SEQ ID NO: 12

|VE202-26|  
 ATGGGAGCGTAGATGGCAGTGGCCATATGTGACAGCCCTGGTCTAACCCCTAACGATTGAACTGAGTGG  
 CTGGAGTGTGGAGAGGCAGGCCGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAAACCCAGTGGCG  
 AAGGCCTGCTGGACGATGACTGACGTTGAGGCTCGAACAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGT

-continued

CCACGCCGTAACGATGACTACTAGGTGTCGGGTGCAAGGACATTGGTGCAGCAAACGCAATAAGTAGTCCA  
CCTGGGGAGTACGTTCGAAGAATGAAACTCAAAGGAAATTGACGGA

SEQ ID NO: 13 | 136 |  
CGCAGCGGAGTGATCCTAGGCTCACCTGGCTGTTGAACTGGTTTCTAGATCGTAGAGGGGGAGATTCTG  
GTGTAGCGTAAATCGTAGATATCGGAGAACACCAGTGGCGAAGGCGCCTCTGGACGGCAACTGACGTTGAG  
GCTCGAAAGTGTGGGAGCAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 14  
| VE202-13 |  
TCAAAGAGTTGATCCTGGCTCAGGACGAACGCTGGCGCGCCTAACACATGCAAGTCGAACGGAGCTTACGTT  
TGAAGTTTCCGATGGACGAATGTAAGCTTAGTGGCGACGGGTGAGTAACACGTGAGCAACCTGCCCTTCAGAGGG  
GGATAACAGCGGAAACGGCTGCTAACACCGCATGATGTTGCGGGGGCACATGCCCTGCAACCAAAGGAGCAATCC  
GCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGGCGACGATCGGTAGCCGGACT  
GAGAGGTTGAACGGCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGATATTGACA  
ATGGCGAAAGCCTGATGCAGCGACGCCGCTGAGGGAAAGACGGTCTTCGATTGTAACACCTGTTGGGAAG  
AAAAATGACGGTACCCAAAGAGGAAGCTCCGCTAACACTACGCCAGCAGCCGGTAATACGTAGGGAGCAAGCGTT  
GTCCGGAATTACTGGGTGAAAGGGAGCGTAGGGGGATGGCAAGTAGATAATGTTAACATCGGCTAACCGGTGG  
CTGCGTTCTAAACTGCCGTTCTGAGTGAAGTAGAGGCAGGGGAATTCTAGTGTAGCGGTGAAATGCGTAGATAT  
TAGGAGGAACACCACTGGCGAAGGGGGCTGCTGGGTTAACAGCAGTGGCTTACGCTGAGGCTCGAAAGCGTGGGGAGCAA  
GGATTAGATAACCTGGTAGTCCACGCCGTAACACGATGATTACTAGGTGTGGGGACTGACCCCTCCGTGCCAG  
TTAACACAATAAGTAATCCACCTGGGAGTACGCCGCAAGGTTAACAGGTTAACAGGAAATTGACGGGGGCCACAA  
GCAGTGGAGTATGTGGTTAACAGCAGGTTAACAGGCTTACGCTGAGGCTTGACATCGGATGCATAGCCTAGAG  
ATAGGTGAAGCCCTCGGGCATCCAGACAGGTGGTGATGGTTGCTGAGGCTTGACATCGGATGCATAGCCTAGAG  
GTCCCGCAACCGCGAACCCCTATTAGTTGCTACGCAAGGAGCAACTTAATGAGACTGCCGTTGACAAAACGGA  
GGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGCTACACAGCTACAAATGGCACTAAAACAG  
AGGGCGCGAACCGCGAGGTGAAGCGAATCCGAAAAAGTGTCTCAGTTGAGATTGCAAGGCTGCAACCCGCTGCA  
TGAAGTCGAAATTGCTAGTAATCGGGATCAGCATGCCGGTGAATACGTTCCGGCTTGTACACACCGCCGT  
CACACCATGGAGTCGGTAACACCGAAGCCAGTAGCCTAACCGAAGGGGGCGCTGCAAGGTGGATTGATGA  
CTGGGGTGAAGTCGTAACAAGGTAGCCGTACCGAAGGTGCGGCTGGATCACCTCTT

SEQ ID NO: 15  
| VE202-14 |  
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAAGCGAAGCGCTGTTT  
CAGAATCTCGGAGGAAGAGGACAGTGACTGAGCGCGGAGGGTAGTAACCGCTGGCAACCTGCCCTACAGG  
GGGATAACAGTTAGAAATGACTGCTAACACCGCATAGCGCACAGGACCGCATGGTAGTGTGAAAAACTCCGGTG  
GTATGAGATGGACCCCGCTGATTAGTAGTTGGGGTAAAGGCCTACCAAGCCGACGATCAGTACGCCGACCTG  
AGAGGGTGACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGACCAA  
TGGGGGAAACCTGATGCGACGCCGCTGAGGAAGAAGTATTCGGTAGTGTAAACTCTATCAGCAGGGAGA  
AGATGACGGTACCTGAGTAAGAACGACCGCTAACACGTGCCAGCAGCCGGTAATACGTATGGTCAAGCGTTA  
TCCGGATTACTGGGTGAAAGGGAGCGTAGACGGTAGGCAAGTCTGGAGTAAACCCAGGGCTAACCTGGGA  
CTGCTTGGAAACTGCAGATCTGGAGTGCAGGAGAGGTAAAGCGAATTCTAGTGTAGCGGTGAAATGCGTAGATAT  
TAGGAGGAACACCACTGGCGAAGGGGGCTTACTGGACGGTAGCTGACGTTGAGGCTCGAAAGCGTGGGGAGCAA  
GGATTAGATAACCTGGTAGTCCACGCCGTAACACGATGACTACTAGGTGTGGTAGTCAAGCAGCATCGGTGCCAG  
CAAACGCAATAAGTAGTCCACCTGGGAGTAGTGTGCAAGAATGAAACTCAAAGGAAATTGACGGGGACCCGACAA  
CGGGTGGAGCATGTGGTTAACAGCAACGCGAAGAACCTTACCTGGTCTTGACATCGGATGACGGCGAGTA

-continued

ATGTCGCCGTCCCTCGGGCATCCGAGACAGGTGGTGCATGGTTGTCGTCACTCGTGTGAGATGTTGGTTA  
 ATGCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCATATAAGGTGGCACTCTGGAGAGACTGCCAGGGAGA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCTATGCCAGGGCTACACCGTGTACAATGGGTA  
 AACAAAGGAAAGCGAGAGGGTGACCTGGAGCGAATCCAAAATAACGTCTCAGTCGGATTGAGTCTGCAACTCG  
 ACTACATGAAGCTGGAATCGCTAGTAATCGCGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC  
 GCCCCTCACACCATGGGAGTCAGTAACGCCGAAGCCAGTGACCCAACTTAGAGGAGGGAGCTGCGAAGGGGGAGA  
 CGGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 16

|VE202-16|  
 ATGAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTAAACG  
 GAAGTTTCGGATGGAAGTTGAATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGTAACCTGCCTTGACTGGG  
 GGACAAACAGTTAGAAATGACTGTAATACCGCATAAGCGCACAGTATCGATGATAACAGTGTGAAAAACTCCGGTGG  
 TACAAGATGGACCCCGCTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCAGCATCAGTAGCCGACCTGA  
 GAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCAACAAAT  
 GGGCAGAAAGCCTGATGCAAGCGACGCCGTGAGTGAAGAAGTATTGGTATGTAAGCTTACAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAACGCCCCGCTAACTACGTGCCAGCGCGTAAATACGTAGGGGAAAGCGTTAT  
 CGGATTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAGGAAAGCCCGCGCTCACTGGGAC  
 TGCTTGGAAACTGTTAACTGGAGTGTGGAGAGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTGGCGAAGCGACTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAACAG  
 GATTAGATACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTTGGGAGCAAAGCTTCCGGTGC  
 AAACGCAGTAAGTATTCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGGTTAAATCGAAGCAAGCGAAGAACCTTACAGGTCTGACATCGATCGACGGGGAGTAA  
 CGTCCCCCTCCCTCGGGCGGAGAAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGAACAGCGCAACCCATTCTAAGTAGCCAGCGTTGCCCGGAACTCTGGGAGACTGCCAGGGATAA  
 CCTGGAGGAAGGTGGGATGACGTCGAAATCATGCCCTTATGATCTGGCTACACACGTGCTACAATGGCGTAA  
 ACAAAAGAAGAACGCCGAGGTGGAGCAAATCTAAACGTCTCAGTGGACTGCAGGCTGCAACTCGC  
 CTGCACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTGTACACACCG  
 CCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGCAAGGGAGCTGCCGAAGCGGGACC  
 GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 17

|VE202-7|  
 ATGAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAAT  
 GAAGTTTCGGATGGAAGTTGATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGGG  
 GGATAACAGTTAGAAATGACTGTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCCGGTGG  
 TGAGGAGATGGATCCCGCTCTGATTAGCCAGTTGGGGGTAACGGCCACAAAGCGACGATCAGTAGCCGACCTGA  
 GAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCAACAAAT  
 GGGCAGAAAGCCTGATGCAAGCGACGCCGTGAGTGAAGAAGTATTGGTATGTAAGCTTACAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAACGCCCCGCTAACTACGTGCCAGCGCGTAAATACGTAGGGGAAAGCGTTAT  
 CGGATTACTGGGTGAAAGGGAGCGTAGACGGCGAAGCAAGTCAGTGAAGTGAAGGAAACCCAGGGCTCAACCTGGGAC  
 TGCTTGGAAACTGTTGCTAGAGTGTGGAGAGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTGGCGAAGCGCTTAACGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAACAG  
 GATTAGATACCTGGTAGTCCACGCCGAAACGATGAATGCTAGGTGTTGGGGGAAAGCCCTCGGTGCCGTG  
 AAACGCAGTAAGCATTCCACCTGGGAGTCGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG

-continued

CGGTGGAGCATGTGGTTAATTCAAGCAACCGAAGAACCTTACCAAGTCTTGACATCCTCTGACCAGCGTAA  
 CGCGCCTCCCTCGGGCAAGAGAGACAGGTGGTCATGGTTGCTCAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGCAACGAGCGAACCCCTATCCTTAGTAGCCAGCAGGTAAAGCTGGCACTCTAGGGAGACTGCCAGGGATA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGATTGGGCTACACAGTGCTACAATGGCGTA  
 AACAAAGGAAAGCAAGACAGTGATGTGGAGCAAATCCAAAATAACGTCCCAGTCGGACTGTAGTCTGCAACCCG  
 ACTACACGAAGCTGGATCGCTAGTAAATCGCAATCAGAATGTCGCGGTAAATACGTTCCGGTCTTGTACACACC  
 GCCCGTCACACCATGGGAGTCAGCAACGCCGAAGTCAGTACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCCG  
 GCAGGTAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGTGGATCACCTCCTT

SEQ ID NO: 18 | 148 |

AACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTTAACACATGCAAGTCGAGCGAAGCACTTTGGA  
 AGATTCTCGGATGAAAGACTTTGTGACTGAGCGCGGACGGGTGAGTAACGGTGGGTAACCTGCCTCATACAGGG  
 GGATAACAGTTAGAAATGACTGCTAAATACCGCATAAGACCACGGTACCGCATGGTACAGTGGAAAAACTCCGGTGG  
 TAGAGATGGACCCCGCTCTGATTAGTAGTTGGTGGGTAACGGCTACCAAGCGCAGCATCAGTAGCCGACCTGA  
 GAGGGTACCCGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGCAAAAT  
 GGAGGAAACTCTGATGCGAGCGACGCCGCTGAAGGATGAAGTATTGGTATGTAACCTCTATCAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAACGCCGGTAACACGTGCCAGCGCGGTAAATACGTAGGGGCAAGCGTTAT  
 CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCGGGCTCAACCCGGGAC  
 TGCATTTGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTAGGGCAAGGGCGCTTGCTGGACGATGACTGACGTTGAGGCTCGAAGCGTGGGAGCAACAG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGACTGCTAGGTGTCGGGGCAAAGCATTGGTCCGCGAC  
 TAACGCAATAAGCAGTCCACCTGGGAGTACGTCGAAGAAATGAAACTCAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGTGGTTAATTCAAGCAACCGAAGAACCTTACCTGATCTGACATCCGATGACCCTTCTAA  
 TGGAAAGCTTTCTCGGAAACATCGGTGACAGGTGGTCATGGTTGCTCAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCAGGTAAAGCTGGCACTCTGGAGAGACTGCCAGGGATA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGACCAGGGTACACAGTGCTACAATGGCGTA  
 AACAAAGAGAACGAACTCGGAGGTAAAGCAAATCTAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG  
 ACTACATGAAGCTGGATCGCTAGTAAATCGCAGATCAGAATGTCGCGGTAAATACGTTCCGGTCTTGTACACACC  
 GCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTACCCAAACGTAAGGAGGGAGCTGCCGAAGGTGGAC  
 CGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGTGGATCACCTCCTT

SEQ ID NO: 19 | 16 |

ATCAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTTAACACATGCAAGTCGAGCGAAGCACTTAAG  
 TGGATCTCTCGGATTGAAACTATTGACTGAGCGCGGACGGGTGAGTAACGGTGGGTAACCTGCCTCATACAG  
 GGGATAACAGTTAGAAATGGCTGCTAAACCGCATAAGGCCACAGGACCGCATGGTCTGGTGTGAAAAACTCCGGT  
 GGTATGAGATGGACCCCGCTCTGATTAGCTAGTTGGAGGGTAACGGCCACCAAGGCAGCATCAGTAGCCGGCT  
 GAGAGGGTAGACGGCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGACA  
 ATGGGGAAACCTGATGCGAGCGACGCCGCTGAAGGAAGAGTATCTCGGTATGTAACCTCTATCAGCAGGGAAAG  
 AAAATGACGGTACCTGACTAAGAACGCCGGTAACACGTGCCAGCAGCGCGGTAAATACGTAGGGGCAAGCGTT  
 ATCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGAAGGAGCAAGTGTGAAAGGCTGGGCTTAACCCCGAG  
 ACTGCATTGGAAACTGTTCTAGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATA  
 TTAGGAGGAACACCGTAGGGCAAGGGCGCTACTGGACCGTAACGTGAGGCTCGAAGCGTGGGAGCAAC  
 AGGATTAGATAACCTGGTAGTCCACGCCGAAACGATGAAATACGTAGGTGTCGGGTGGCAAAGCATTGGTCCGCA  
 GCAAACGCAATAAGTATTCCACCTGGGAGTACGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA

-continued

AGCGGTGGAGCATGTGTTAATCGAACGCGAAGAACCTTACCAAGTCTTGACATCCCTCTGACC GGCGT  
 AACGGGGCCTCCCTCGGGGCAGAGGAGACAGGTGGTCATGGTGTGTCAGCTCGTGTGAGATGTTGGTT  
 AAGTCCCACGAGGCCAACCCCTATCCTAGTAGGCCAGGGTAAGCTGGCACTCTAGGGAGACTGCCGGGA  
 TAACCCGGAGGAAGGGCGGGGACGACGTCAAATCATCATGCCCTTATGATTGGCTACACACGTGCTACAATGGCG  
 TAAACAAAGGAAGCGAGACAGCGATGTTGAGCAAATCCCCAAAATAACGTCCCAGTCGGACTGCAGTCTGCAACT  
 CGACTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTAAACGTTCCCGGGTCTGTACACA  
 CCGCCCGTCAACCATGGGAGTCAGTAACGCCAGTCAGTGAACCAACCTTACAGGAGGGAGCTGCCAGGGCG  
 GACCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 20 | 170 |

AACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCACTTGGAA  
 AGATTCTCGGATGATTCCCTTGTGACTGAGCGCGGACGGGTAGTAACCGTGGTAACCTGCCCTACACAGGG  
 GGATAACAGTTAGAAATGACTGCTAACCGCATAAGACCACGGTACCGCATGGTACAGTGGAAAAACTCCGGTGG  
 TAGAGATGGACCCCGCTCTGATTAGTAGTTGGTGGGTAACGGCTACCAAGCGACGATCAGTAGCCGACCTGA  
 GAGGGTACCCGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGCAAAAT  
 GGAGGAAACTCTGATGCGAGCGACGCCGCTGAAGGATGAAGTATTGGTATGTAACCTCTATCAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAACGACGCCGCTAACACGTGCCAGCGCGGTAAACGTAGGGGCAAGCGTTAT  
 CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCGGGCTCAACCCGGGAC  
 TGCATTTGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTAGGGCGAGGGCTTGCTGGACGATGACTGACGTTGAGGCTCGAAGCGTGGGAGCAACAG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGACTGCTAGGTGTCGGGCAAAGCATTGGTCCCGCAGC  
 TAACGCAATAACGAGTCCACCTGGGAGTACGTCGAAGAAATGAAACCTAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGTGTTAATCGAACGCAACGCGAAGAACCTTACCTGATCTGACATCCGATGACCCTCCGAA  
 TGGAAAGCTTTCTCGGAAACATCGGTGACAGGTGGTCATGGTGTGTCAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCGAGGTAAGCTGGCACTCTGGAGAGACTGCCAGGGATA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCTTATGACCGAGGCTACACAGTGTACAATGGCGTA  
 AACAAAGAGAACGCAACTCGGAGGTAAGCAAATCTAAAAAAACGTCTAGTCGGATTGTAGTCTGCAACTCG  
 ACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAATGCTCGGTAAACGTTCCCGGTCTGTACACACC  
 GCCCGTACACCATGGGAGTCAGTAACGCCAGTCAGTGAACCAACCGTAAGGAGGGAGCTGCCGAAGGTGGAC  
 CGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 21 | 189 |

ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCATAACATGCAAGTCGAGCGAAGTTCGAGGA  
 AGCTTCTCCAAAGAGACTTAGTGGCGAACGGGTAGTAACACGTAGGTAACCTGCCATGTGTCGGATAACTG  
 CTGGAAACGGTAGCTAAACCGGATAGGTACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGGGGTGAAAC  
 ATGGATGGACCTCGCGCGCATTAGTAGTTGGTGAGGTAACGCCACCAAGCGATGATGCGTAGCCGGCTGAGA  
 GGGTAAACGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTAGGGAATTCGTCAATGG  
 GGGAAACCCCTGAAACGAGCAATGCCGCGTAGTGAAGAAGGTCTCGGATGTAAGCTCTGTTGTAAGTAAGAACG  
 GCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTACCGAAAGCCACGGCTAACTACGTGCCAGCGCGGT  
 AATAACGTAGGGTGGCAAGCGTTACCGGAATCATTGGCGTAAAGGGTGCCTAGGTGGCTACTAAGTGTAGTAA  
 AGGCAATGGCTCAACCATTGTAAGCTATGGAAACTGGTATGCTGGAGTGACAGAAGAGGGCGATGGAATTCCATGTG  
 AGCGGTTAAATGCGTAGATATGGAGGAACCGAGTGGCGAAGCGGTGCGCTGGTGTGACTGACACTGAGGCA  
 CGAAAGCGTGGGAGCAAATAGGATTAGATAACCTAGTAGTCACGCCGTAACACGTGAGAAACTAAGTGTGAGGA  
 ATTCACTGCTGCGAGTTAACGCAATAAGTCTCCGCTGGGAGTATGACGCAAGTGTGAAACTCAAAGGAATTGAC

-continued

GGGGGCCCGACAAGCGGTGGAGTATGTTAATTCAAGCAGCGAAGAACCTTACCAAGGCCTTGACATGGAA  
 ACAAAATACCCTAGAGATAGGGGATAATTATGGATCACACAGGTGGTCATGGTGTGTCAGCTCGTGTGAGA  
 TGGTGGTTAACCGCAACGAGCGAACCCCTGTCGATGTTACCAGCATCAAGTGGGACTCATGCGAGACTG  
 CCGGTGACAAACCGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGGCCTGGCTACACAGTACTAC  
 AATGGCGGCCACAAAGCAGCGACACAGTGTGAAGCGAATCTCATAAAGGTCGTCAGTCGGATTGAAGTC  
 TGCAACTCGACTTCATGAAGTCGGAATCGTAGTAATCGCAGATCAGCATGCTGCGGTGAATACGTTCTCGGCCTT  
 GTACACACCGCCCGTCAAACCATGGGAGTCAGTAATACCGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTCGA  
 AGTAGGACCGATGACTGGGTTAACCGTAAAGTCAAGGTATCCCTACGGGAACGTGGGATGGATCACCTCCTT

SEQ ID NO: 22 | 169 |  
 AGTAACCGTGGTAAACCTGCCTCATACAGGGGATAACAGTTAGAAATGACTGCTAACCGCATAAGACCAACGGT  
 ACCGCATGGTACAGTGGTAAAAACTCCGGTGGTATGAGATGGACCCCGCTGATTAGGTAGTTGGTGGGTAACGG  
 CCTACCAAGCCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGACTGAGACACGGCCAGACTCCT  
 ACGGGAGGCAGCAGTGGGAAATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCTGAAGGATGAAGTATT  
 TCGGTATGTAACCTCTATCAGCAGGGAAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGC  
 AGCCCGGTAATACGTAGGGGCAAGCGTTACCGGATTTACTGGGTGAAAGGGAGCGTAGACGGCACGGCAAGCC  
 AGATGTGAAAGCC

SEQ ID NO: 23  
 | VE202-29 |  
 CAGGCTGGAGTCAGGAGAGGTAAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCGT  
 GGCAGAGGCGGTTACTGGACTGTAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGG  
 TAGTCCACCGGTAACGATGATTGCTAGGTGAGGTGATGGACCCATCGGTGCCAGCTAACGCAATAAGCA  
 ATCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGCACAGCGTGGAGCATGTG  
 GTTAAATTGCAAGCAACCGAAGAACCTTACCAAGTCTTGACATCC

SEQ ID NO: 24  
 | YK96 |  
 CGGGGCTCACCCCGGACTGCATTGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGAAATTCTAGTGTAGC  
 GGTGAAATGCGTAGATATTAGGAGGAACACCGAGTGGCGAAGGGCGCTTGCTGGAGCATGACTGACGTTGAGGCTCGA  
 AAGCGTGGGAGCAAACAGGATTAGATAACCTGGTA

SEQ ID NO: 25  
 | YK101 |  
 AGGGTCAACCCCTGGACTGCATTGAAACTGTCAGGCTGGAGTGCCGGAGAGGTAAAGCGGAATTCTAGTGTAGCG  
 TGAAATGCGTAGATATTAGGAGGAACACCGAGTGGCGAAGGGCGCTTACTGGACGGTAACTGACGTTGATGCTCGAAA  
 GCGTGGGAGCAAACAGGATTAGATAACCTGGTAAA

SEQ ID NO: 26  
 | YK110 |  
 GGGAAAGTCGGCTTAAGTGCAGGGCTTAACCCCGTGGGGACCGAAACTGTGAAGCTCGAGTGTGGAGAGGAAAG  
 CGGAATTCTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCGAGTGGCGAAAGCGGCTTCTGGACGACA  
 ACTGACGCTGAGGCGCGAAAGCCAGGGAGCAAACGGGATTAGATAACCCAGTAA

SEQ ID NO: 27  
 | YK149 |  
 TAGTCTGAGTGTGCGGGCTTAACCCCGTATGGCGTTGGACTGGAAGTCTTGAGTGCAGGAGAGGAAAGGGAA  
 TTCCCGAGTGTAGCGGTGAAATGCGTAGATATTGGAGGAACACCGAGTGGCGAAGGGCCTTCTGGACTGTGTCTGA  
 CGCTGAGATGCGAAAGCCAGGGTAGCAAACGGGATTAGATAACACCGTA

## US 10,555,980 B2

99

100

-continued

SEQ ID NO: 28

| YK154 |  
 GATAGTCGGTCTTAAGTGCAGGGCTTACCCGTGAGGGGACCGAAACTGTGAAGCTCGAGTGTGAGAGGAAAGCG  
 GAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCAAAGCGCTTCTGGACGACAAC  
 TGACGCTGAGGCCGAAAGCCAGGGAGCAGGGATTAGATACCAACGGTAA

SEQ ID NO: 29

| YK36 |  
 CGTTTGCTCCACGCTTCGAGCCTCACGTCAAGTTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCCTAATA  
 TCTACGCATTCACCGTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCC  
 CGGGGTTGAGCCCCGGGTTTCACATCAGACTGCCACTCCGTACGCTCCCTTACACCCAGTAAATCCGATAA  
 CGCTTGACCATACGTATTACCGCGTGTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCTTTCT  
 TCCCTGCTGATAGAGCTTACATACCGAAATACCTCATCGCTCACGCCGCGTCGCTGCATCAGGTTCCCCATTG  
 TGCAATATTCCCCACTGCTGCCTCCGTAGGAGTTGGA

SEQ ID NO: 30

| YK95 |  
 TGTCAACACTTCGAGCATCAGCGTCAAGTTACGTCAGTAAGCTGCCATTCCAACTGGAGTTCTCGTGTATCTAA  
 GCATTCACCGTACACCACGAATTCCGCCTACCTCTACTGCACTCAAGACGACAGTATCACTGCAATTACGG  
 TTGAGCCGCAAACTTCACAGCTGACTTAATAGTCGCCTACGCTCCCTTAAACCAATAATCCGATAACGCTT  
 GGATCCTCCGTATTACCGCGTGTGGCACGGAGTTAGCCGATCCTTATTCTGTATGGTACATACAAAAAGCCACAC  
 GTGGCTACTTTATCCCATAAAAAGTTACACCCATAGGGCAGTCATCCTCACGCTACTTGGCTGGTTCA  
 GACTCTCGTCATTGACCAATTCCCTCACTGCTGCCTCCGTAGGTAGTTGGAA

SEQ ID NO: 31

| YK32 |  
 CGGTTGTCACGCTTCGTGCTCAGTGTCAAGTTCAAGCTGCCACTGATGTTCTCTTAATATC  
 TACGCATTCACCGTACACTAGGAATTCCGCTTACCTCTCCGTCACTCCAGTCTGACAGTTCAAAAGCAGTCCA  
 GAGTTAACGCCCTGGGTTTCACCTCTGACTTGCCATACCACCTACGCACCCCTTACACCCAGTAACTCCGATAACG  
 CCTGGCCCCCTACGTATTACCGCGGCTGTGGCACGTAGTTAGCCGGGCTTCTTAGTCAGGTACCGTCAATTCTTC  
 CCTGCTGATAGAGCTTACATACCGAAATACCTCTCACTCACGCCGCGTCGCTGCATCAGGTTCCCCCATTGT  
 CAATATTCCCCACTGCTGCCTCCGTGGAGTTGGA

SEQ ID NO: 32

| YK64 |  
 CGGAATGTCACGCATTGAGCCTCACGTCAAGTTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCTTAATAT  
 CTACGCATTCACCGTACACTAGGAATTCCGCTTACCTCTCCGGACTCAAGACTAACAGTTCCAATGCAGTCCA  
 GGGGTTGAGCCCCGCCTTCACATCAGACTTGCCAGTCCGTCTACGCTCCCTTACACCCAGTAAATCCGATAAC  
 CCTTGCCCCCTACGTATTACCGCGGCTGTGGCACGTAGTTAGCCGGGCTTCTTAGTCAGGTACCGTCACTATCTT  
 CCTGCTGATAGAGTTACATACCGAGATACTTCTCCTCACGCCGCGTCGCTGCATCAGGTTCCCCCATTGT  
 GCAATATTCCCCACTGCTGCCTCCGTGGAGTTGGA

SEQ ID NO: 33

| YK73 |  
 TGCTCACGCTTCGCGTCAGCGTCAAGTTACTGTCCAGCAATCCGCTTCGCCACTGGTGTCTCTGTATATCTAC  
 GCATTCACCGTACACACGGAAATTCCGATTGCCCTCCAGCACTCAAGAAACTACAGTTCAAATGCAGGCTGGAGG  
 TTGAGCCCCAGTTTCACATCTGACTTGCAATCCGCTACACGCCCTTACACCCAGTAAATCCGATAACGCTT  
 GCCACCTACGTATTACCGCGGCTGTGGCACGTAGTTAGCCGGCTTCTTAGTCAGGTACCGTCAATTGTTCTG  
 CCTGACAAAAGAAGTTACACCCGAAAGCTTCTCCTCACGCCGCGTTGCTGGTCAGGCTTGCGCCATTGCC  
 CAATATTCCCCACTGCTGCCTCCGTGGAGTTGGA

## US 10,555,980 B2

101

-continued

102

SEQ ID NO: 34

| YK87 |  
TGTCCACGCTTCGAGCTAGCGTCAGTTATCGTCCAGTAAGCCCTCGCCACTGGTCTCCTAAATATCTAC  
GCATTTCACCGCTACACTAGGAATTCCGCTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCACACTACGGGG  
TTGAGCCCCGGGCTTACATCAGACTTGCGCACCGCTCGCTCCCTTACACCCAGTAAATCCGGATAACGCTT  
GCACCATACTGATTACCGCGCTGCTGGCACGTATTAGCCGGTCTTCTAGTCAGGTACCGTCAATTATCTCCCT  
GCTGATAGAGCTTACACCGAAACTTCTCGCTACGCCGCTCGCTCATCAGGCTTCGCCATTGTGCAA  
TATTCCCCACTGCTGACTCCCGTAGGAGTTGGAA

SEQ ID NO: 35

| YK105 |  
CGTTTCTCCACGCTTCGCGCTAGCGTCAGTTACTGTCCAGCAATCCGCTTCGCCACTGGTCTCCTCCGTATATC  
TACGCATTCACCGCTACACACCGAATTCCGATTGCCCTCCAGCACTCAAGAACTACAGTTCAAATGCAGGCTGG  
AGGTTGAGCCCCAGTTTACATCTGACTTCACTGCCATTCCGCTACACGCCCTTACACCCAGTAAATCCGGATAACG  
CTTGCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGTGGCTTACGTCAAGGTACCGTCAATTGTTTC  
GTCCCCGACAAAAGAAGTTAACACCGAACGCCCTCTCCTCACGCCGGTTGCTGGTCAGGCTTGCGCCATT  
GCCCAATATTCCCCACTGCTGCCCTGGGAAGTTGGAA

SEQ ID NO: 36

| YK153 |  
ATGTCCTGACTTCGCGCTCAGCGTCAGTTGCGTCCAGAAAGCCGCTTCGCCACTGGTCTCCTCTAATATCTA  
CGCATTTCACCGCTACACTAGGAATTCCGCTTCCTCTCCGACACTCGAGCTTACAGTTCCGGTCCCTCACGGGG  
TTAAGCCCCGACTTTAACGCGACTTGGATGCCGCTGCCCTTACGCCAATAATTCCGGACAACGCTT  
GCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGTGGCTTACGTCAAGGTACCGTCAAGGATAACGGG  
TATTGACCGCTATCCTGTTCCATATAACAGAACTTACAAACCGAACGCCGTACGTTACGCCGGTTGCT  
CCGTCAGACTTCGTCATTGCCGAAGATTCCCCACTGCTGCCCTGGGAAGTTGGAA

SEQ ID NO: 37

| YK163 |  
GTTTGCCTACGCCCTCGAGCTAGCGTCAGTTATCGTCCAGTAAGCCGCTTCGCCACTGGTCTCCTCTAATATC  
TACGCATTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCG  
GGGTTGAGCCCCGGGCTTCACATCAGACTTGGCACCGCTGCCCTTACACCCAGTAAATCCGGATAACG  
CTTGACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTCTTCTAGTCAGGTACCGTCAATTCTTC  
CCTGCTGATAGAGCTTACACCGAAATCTCTCGCTACGCCGCTCGCTCATCAGGCTTCGCCATTGTG  
CAATATTCCCCACTGCTGCCCTCCGAGGAGTTGGAA

SEQ ID NO: 38

| YK191 |  
CGTTGCTCACGCAATTGAGCCTCAGCGTCAGTTAACCGAGTAAGCCGCTTCGCCACTGATGTTCTCCTAAATATC  
TACGCATTCACCGCTACACTAGGAATTCCGCTTACCTCTACTCAAGAACACAGTTCAAATGCAGTTAT  
GGGTTAACCCATAGTTTACATCTGACTTGGGATCCCGCTACGCTCCCTTACACCCAGTAAATCCGGACAACG  
CTCGCTCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGAGCTTCCCTCTCAGGTACCGTCTTTTCTG  
CCCTGAAGACAGAGGTTACAACTCTAAACCCCTTCCCTCACGCCGCTCGCTCATCAGGCTTCGCCATTGT  
GCAATATTCCCCACTGCTGCCCTCCGAGGAGTTGGAA

SEQ ID NO: 39

| YK99 |  
TGGGCTTACCCATAACTGCATTGAAACTGTGGTTCTTGAGTGAAGTAGAGGTAAGCGGAATTCTAGTGTAGCGG  
TGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCCGCTTACTGGGTTAACTGACGCTGAGGCTCGAAA  
GCGTGGGAGCAAACAGGATTAGATAACCCAAAGTAA

-continued

SEQ ID NO: 40

| YK55 |  
 GTCAGCATCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCACTGGTGTCCCTCTAATATCTACGCATT  
 TCAACCGCTACACTAGGAATTCGCTTACCTCTCCGTACTCTAGATTGACAGTTCCAATGCAGTCCCAGGGTTGAG  
 CCCGGGTTTACATCAGACTGCCACTCGTCTACGCTCCCTTACACCCAGTAAATCCGATAACGCTTGCA  
 ATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTTCTCCCTGCTGA  
 TAGAGCTTTACATACCGAAATACCTCATCGCTACCGCGCTGCTGCATCAGGGTTCCCCATTGTGCAATATT  
 CCCACTGCTGCCCTCCGAGGGAGTTGGA

SEQ ID NO: 41

| YK75 |  
 TCATCGTTACGGTGGATCTGCGCCGGTACGGCGGGCTGGAGTGCAGTAGGGAGACTGGAATTCCGGTGTAA  
 GGTGGAAATGTGTAGATATCGGAAAGAACACCGATGGCGAAGGCAGGTCTCTGGCCGTCACTGACGCTGAGGAGCGA  
 AAGCGTGGGAGCGAACAGGATTAGATAACCGTAA

SEQ ID NO: 42

| YK90 |  
 TGAACCCAGGGCTTAACCTCTGGACTGCTTGAACGTCAAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCTAGT  
 GTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCAGGTTACTGGACTGAAACTGACACTGAGG  
 CACGAAAGCGTGGGAGCAAACAGGATTAGATAACCATGGTAA

SEQ ID NO: 43

| YK30 |  
 ACCAGGGCTTAACCTCTGGACTGCTTGAACGTCAAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAG  
 CGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCAGGTTACTGGACTGAAACTGACACTGAGGCACG  
 AAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAA

SEQ ID NO: 44

| YK31 |  
 GAACCCAGGGCTTAACCTCTGGACTGCTTGAACGTCAAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCTAGT  
 TAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCAGGTTACTGGACTGAAACTGACACTGAGG  
 ACGAAAGCGTGGGAGCAAACAGGATTAGATAACCCGGTAA

SEQ ID NO: 45

| YK12 |  
 GAGTCAGCTTCGAGCTCACGTCAAGTTACGTCCAGTAAGCCGCCACTGGTGTCCCTCTAATATCTACG  
 CATTTCACCGCTACACTAGGAATTCCGTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGGT  
 TGAGCCCCGGGTTTCACATCAGACTTGCACCCGCTGCGCTCCCTTACACCCAGTAAATCCGATAACGCTTG  
 CACCATACGTATTACCGCGCTGCGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTATCTCCCTG  
 CTGATAGAGCTTACATACCGAAATCTTCTCGTCACCGCGCTGCTGCATCAGGCTTCGCCATTGTGCA  
 ATTCCCCACTGCTGCCCTCCGAGGGAGTTGGA

SEQ ID NO: 46

| YK27 |  
 TGTCAGCTTCGAGCTCACGTCAAGTTACCGTCCAGTAAGCCGCCACTGGTGTCCCTCTAATATCTACGCA  
 TTTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCCCAGGGTTG  
 AGCCCCGGGTTTCACATCAGACTTGCACCCGCTACGCTCCCTTACACCCAGTAAATCCGATAACGCTTGCA  
 CCATACGTATTACCGCGCTGCGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTCTCCCTGCT  
 GATAGAGCTTACATACCGAAATCTTACGTCCAGTCACCGCGCTGCTGCATCAGGCTTCGCCATTGTGCA  
 ATTCCCCACTGCTGCCCTCCGAGGGAGTTGGA

SEQ ID NO: 47

| YK28 |  
 CACGTCAGTTACCGTCCAGTAAGCCGCCACTGGTGTCCCTCTAATATCTACGCAATTACCGCTACACTA  
 GGAATTCCGCTACCTCTCCGGCACTCAAGACGGGAGTTCCAATGCAGTCCCAGGGTTGAGCCCCAGCCTTCAC  
 ATCAGACTTGTCCATCCGTACGCTCCCTTACACCCAGTAAATCCGATAACGCTTGCCCCCTACGTATTACCGC

**105****106**

-continued

GGCTGCTGGCACGTAGTTAGCCGGGCTTCTTAGTCAGGTACCGTCAGTACCTCTCCCTGCTGATAGAAGTTACATA  
 CCGAGATACTTCTCCTCACGCCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAATATCCCCACTGCTGCCTC  
 CGTAGGAGTTGGG

SEQ ID NO: 48

| YK29 |  
 GTCAGCTTCGAGCTACGTCAGTTACCGTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCCTAATATCTACGCAT  
 TTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCCCAGGGTTGA  
 GCCCCGGGTTTCACATCAGACTGCCACTCCGTACGCTCCCTTACACCCAGTAAATCCGATAACGCTTGAC  
 CATACGTATTACCGCGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCAGTACCTCTCCCTGCTG  
 ATAGAGCTTACATACCGAAATACTCATCGCTCACGCCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAATATT  
 CCCCACGTGCCTCCCGTGGGAGTTGGA

SEQ ID NO: 49

| YK33 |  
 GATGCTCAGCTTCGTCAGTGTAGTTCAAGCCAGTAAGCCGCTTCGCCACTGATGTTCCCTAATATCTA  
 CGCATTTCACCGCTACACTAGGAATTCCGTTACCTCTCCGTACGTCAGTCTGACAGTTCAAAGCAGTCCCAGA  
 GTTAAGCCCTGGGTTTCACTTCTGACTTGCCATACCACCTACGCACCCCTTACACCCAGTAAATCCGATAACGCT  
 TGCCCCCTACGTATTACCGCGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTCAGTACCTCTCC  
 TGCTGATAGAGCTTACATACCGAGATACTTCTCACTCACGCCGCGTCGCTGCATCAGGGTTCCCCATTGTGCA  
 ATATTCCCCACTGCTGCCTCCCGAAGGAAGTTGGA

SEQ ID NO: 50

| YK34 | 70A\_009\_YK34\_A1\_A02  
 GTGTCAGCTTCGTCAGTGTAGTTCAAGCCAGTAAGCCGCTTCGCCACTGATGTTCCCTAATATCTACGC  
 ATTTACCGCTACACTAGGAATTCCGTTACCTCTCCGTACGTCAGTCTGACAGTTCAAAGCAGTCCCAGAGTT  
 AAGCCCTGGGTTTCACTTCTGACTTGCCATACCACCTACGCACCCCTTACACCCAGTAAATCCGATAACGCTTG  
 CCCCTACGTATTACCGCGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTCAGTACCTCTCCCTG  
 TGATAGAGCTTACATACCGAGATACTTCTCACTCACGCCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAATA  
 TTCCCCACTGCTGCCTCCCGTACGGAGTTGGA

SEQ ID NO: 51

| YK35 |  
 GTCAGCTTCGAGCCTACGTCAGTTACCGTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCCTAATATCTACGCAT  
 TTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCCCAGGGTTGA  
 GCCCCGGGTTTCACATCAGACTGCCACTCCGTACGCTCCCTTACACCCAGTAAATCCGATAACGCTTGAC  
 CATACGTATTACCGCGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCAGTACCTCTCCCTG  
 ATAGAGCTTACATACCGAAATACTCATCGCTCACGCCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAATATT  
 CCCCACGTGCCTCGCGTAGGGAGTTGGA

SEQ ID NO: 52

| YK51 |  
 TGTCAAGCTTCGAGCTACGCGTCAGTTACGTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCCTAATATCTACGC  
 ATTTACCGCTACACTAGGAATTCCGTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGTT  
 GAGCCCCGGGCTTCACATCAGACTGCCACCGCCTGCGCTCCCTTACACCCAGTAAATCCGATAACGCTTG  
 ACCATACGTATTACCGCGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTATCTCCCTG  
 TGATAGAGCTTACATACCGAAATACTTCTCGCTCACGCCGCGTCGCTGCATCAGGGTTCCGCCATTGTGCAATA  
 TTCCCCACTGCTGCCTCCCGAAGGGAGTTGGA

SEQ ID NO: 53

| YK52 |  
 TTCAAGCTTCGAGCTACGCGTCAGTTACGTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCCTAATATCTACGC  
 TTTCACCGCTACACTAGGAATTCCGTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGTTG

-continued

AGCCCCGGGCTTCACATCAGACTTGCACGCCGCTCGCTCCCTTACACCCAGTAATCCGATAACGCTTGCA  
 CCATACGTATTACCGCGCTGCTGGCACGTATTTAGCCGGTCTTAGTCAGGTACCGTCATTATCTCCCTGCT  
 GATAGAGCTTACATACCGAAATACTCTCGCTCACGCCGCTCGCATCAGGCTTCGCCATTGTGCAATAT  
 TCCCCACTGCTGCCCTCCGAGGGGAGTTGG

SEQ ID NO: 54

| YK54 |  
 TTCGGCTGCTTCCCTTCTCGCGCTCAGTGTAGTTCTGTCTAGTAAGCCGCTTCGCCACTGATGTTCCCTCC  
 TAATATCTACGCACTCACCGCTCCACAATGAATTCCGCTTACCCCTCCCGCCTCTAGTCAGTTAAAAAA  
 ACTCCCCGAGAGAAACCTGGGTTTTCTCTGACATGCCATCCCACCCCCACCCCTTATACACCCAAAAATCG  
 GATAAAAGGTGCGACCTACGTATTACCGCTGCTGGCGTAGATAGCCGGGGTTCTTATACAGGGACCGTCAT  
 TTTCTTCCCGCTGATACAGCTTACATACCGAAATACTTCTCACGCCGCTCGCATCAGGTTCCCC  
 ATTGTGCAATATTCCCCACTGCTGCCCTCCGAGGGAGTTGGGAAA

SEQ ID NO: 55

| YK56 |  
 GTTCAGCTTCGAGCCTCACGTCAAGTACCGCTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCTAATATCTACGC  
 ATTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCCGGGGTT  
 GAGCCCCGGGTTTCACATCAGACTGCCACTCCGTCTACGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
 ACCATACGTATTACCGCGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTTCTCCCTGC  
 TGATAGAGCTTACATACCGAAATACTTCATCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATA  
 TTCCCCACTGCTGCCCTCCGAGGGAGTTGG

SEQ ID NO: 56

| YK57 |  
 GTCAGCTTCGAGCTCACGTCAAGTACCGCTCAGTAAGCCGCTTCGCCACTGGTGTTCCTCTAATATCTACGCAT  
 TTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGTACTCTAGATTGACAGTTCCAATGCAGTCCGGGGTTGA  
 GCCCCGGGTTTCACATCAGACTGCCACTCCGTCTACGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
 CATACGTATTACCGCGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTTCTCCCTGCTG  
 ATAGAGCTTACATACCGAAATACTTCATCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATA  
 CCCCCACTGCTGCCCTCCGAGGGAGTTGG

SEQ ID NO: 57

| YK58 |  
 TCTCACGCTTCGAGCTCACGTCAAGTACCGCTCAGCAAGCCGCTTCGCCACTGGTGTTCCTCTAATATCTACGC  
 ATTCACCGCTACACTAGGAATTCCACTTGCCCTCCGACACTCTAGCTCACGAGTTCCAATGCAGTCCGGGGTT  
 GAGCCCCGGGTTTCACATCGGCTGCCGCTACGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
 CCCCTACGTATTACCGCGCTGGCACGTATTAGCCGGGCTCTTAGTCAGGTACCGTCATTTCTCCCTG  
 TGATAGAGTTACATACCGAAATACTTCATCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATA  
 TTCCCCACTGCTGCCCTCCGAGGGAGTTGG

SEQ ID NO: 58

| YK65 |  
 GTCAGCTTCGAGCTCACGTCAAGTACCGCTCAGTATCGTCAAGTAAGCCGCTTCGCCACTGGTGTTCCTCTAATATCTACGC  
 TTTCACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGGGTT  
 AGCCCCGGGTTTCACATCAGACTGCCACTCCGCTCGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
 CCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTCTCTCCCTG  
 TGATAGAGCTTACATACCGAAATACTTCATCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATA  
 TTCCCCACTGCTGCCCTCCGAGGGAGTTGG

109

-continued

110

SEQ ID NO: 59

| YK67 |  
AGCCCCGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTACG  
CATTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCAAGACGGCAGTTCCAATGCAGTCCGGGGT  
TGAGCCCCAGCCTTCACATCAGACTTGTCCATCCGTCAAGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
CCCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGGCTTCTAGTCAGGTACCGTCATTCTCCCTG  
CTGATAGAGTTTACATACCGAGATACTTCTTCCTCACGCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAAT  
ATTCCCCACTGCTGCCTCCGAAGGAAGTTGGA

SEQ ID NO: 60

| YK69 |  
TGCCTCAGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTACG  
CATTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCGAGCCAGACAGTTCCAATGCAGTCCAGGGT  
TAAGCCCTGGGTTTCACATCAGACTTGCCTGCGTCAAGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
CCCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGGCTTCTAGTCAGGTACCGTCATTATCTCCCTG  
CTGATAGAGTTTACATACCGAAATACCTCTCGCTCACGCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAAT  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGGA

SEQ ID NO: 61

| YK70 |  
GTTGCTCAGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTAC  
GCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCGAGCCAGACAGTTCCAATGCAGTCCAGGG  
TTAACGCCCCTGGGTTTCACATCAGACTTGCCTGCGTCAAGCTCCCTTACACCCAGTAATCCGATAACGCTT  
GCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGGCTTCTAGTCAGGTACCGTCATTATCTCCCT  
GCTGATAGAGCTTACATACCGAAATACCTCTCGCTCACGCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAA  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGGA

SEQ ID NO: 62

| YK71 |  
TGCCTCAGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTACG  
CATTTCACCGCTACACTAGGAATTCCCACTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACGGGGT  
TGAGCCCCGGCTTCACATCAGACTTGCCTGCGCACCGCTCCCTTACACCCAGTAATCCGATAACGCTTG  
CACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTCTTCTAGTCAGGTACCGTCATTCTCCCT  
GCTGATAGAGCTTACATACCGAAATACCTCTCGCTCACGCGCGTCGCTGCATCAGGGTTCCCCATTGTGCAA  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGGA

SEQ ID NO: 63

| YK74 |  
GATGCCCTGGCTTCGCGCTCACGTCAAGTACCGTCCAGTTGTCTGGCAGAAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTA  
CGCATTTCACCGCTACACTAGGAATTCCGCTTCCTCTCCGACACTCGAGCTTACAGTTGGTCCCTCACGGGG  
TTAACGCCCCGACTTTAAGACCGACTTGCATGCCCTGCGGCCCTTACGCCAATAATTCCGGACAACGCTT  
GCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGTCTTCTAGTCAGGTACCGTCAGGGATAACGGG  
TATTGACCGCTATCCTGTTGCTCCATATAACAGAACTTTACAACCCGAAGGCCGTACGTTACGGGGCT  
CCGTCAGACTTCGTCATTGCGGAAGATTCCCCACTGCTGCCTCCGGGGGAGTTGGA

SEQ ID NO: 64

| YK88 |  
GTCCCCGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCTTCGCCACTGGTGTCTCCTAATATCTACG  
CATTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACGGGGT  
TGAGCCCCGGCTTCACATCAGACTTGCCTGCGCACCGCTGCCCTTACACCCAGTAATCCGATAACGCTTG  
CACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTCTTCTAGTCAGGTACCGTCATTATCTCCCTG

-continued

CTGATAGAGCTTACATACCGAAATACTTCGCTCACGCCGCTCGCATCAGGCTTCGCCATTGTGCAAT  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 65

| YK89 |  
TGCAGCTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCTCGCCACTGGTGTCCCTCTAATATCTACGC  
ATTCACCGCTACACTAGGAATTCCGTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGTT  
GAGCCCCGGGCTTCACATCAGACTTGCGCACCGCTGCGCTCCCTTACACCCAGTAAATCCGATAACGCTG  
ACCATACGTATTACCGGGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTATTATCTCCCTGC  
TGATAGAGCTTACATACCGAAATACTTCGCTCACGCCGCTCGCATCAGGCTTCGCCATTGTGCAATA  
TTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 66

| YK97 |  
TGCAGCTTCGAGCTCAGCGTCAGTTATCGTCCAGTAAGCCGCTCGCCACTGGTGTCCCTCTAATATCTACG  
CATTACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGTT  
TGAGCCCCGGGCTTCACATCAGACTTGCGCACCGCTGCGCTCCCTTACACCCAGTAAATCCGATAACGCTG  
CACCATACGTATTACCGGGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTATTATCTCCCT  
GCTGATAGAGCTTACATACCGAAATACTTCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAAA  
TATTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 67

| YK98 |  
ATTCAGCTTCGAGCTCACGTCACTGGCAGTAAGCCGCTTCGCCACTGGTGTCCCTCTAATATCTACGCA  
TTACCGCTACACTAGGAATTCCGTTACCCCTCCGCACTCAAGCATAACAGTTCCAATGCAGTCCAGGGTTA  
AGCCCTGCCTTCACATCAGACTTGATAACGCCGCTACGCTCCCTTACACCCAGTAAATCCGATAACGCTG  
CCCTACGTATTACCGGGCTGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTATTATCTCCCTG  
GATAGAAGTTACATACCGAGATACTTCCTTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATAT  
TCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 68

| YK139 |  
GTGTCAGCTTCGAGCTCACGTCACTGGCAGTAAGCCGCTTCGCCACTGGTGTCCCTCTAATATCTACGC  
ATTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGACTCGAGCCAGACAGTTCCAATGCAGTCCAGGGTT  
AGCCCTGGGTTTCACATCAGACTTGCGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTATTATCTCC  
CCCTACGTATTACCGGGCTGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTATTATCTCCCTG  
TGATAGAGCTTACATACCGAAATACTTCGCTCACGCCGCTCGCATCAGGTTCCCCATTGTGCAATA  
TTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 69

| YK141 |  
GCCAGCTTCGAGCTCACGTCACTGGCAGTAAGCCGCTTCGCCACTGGTGTCCCTCTAATATCTACGCAT  
TTACCGCTACACTAGGAATTCCACTTACCTCTCCGACACTCTAGCTGCACAGTTCCAAGCAGTCCACAGGGTGA  
GCCCATGCCTTCACCTCAGACTTGACAGCCGCTACGCTCCCTTACACCCAGTAAATCCGATAACGCTG  
CCTACGTATTACCGGGCTGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTATTATCTCCCTG  
ATAGAAGTTACATACCGAAATACTTCATCCTTCACGCCGCTCGCATCAGGTTCCGCCATTGTGCAATATT  
CCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 70

| YK142 |  
TGATCAGCTTCGAGCTCACGTCACTGGCAGTAAGCCGCTTCGCCACTGGTGTCCCTCTAATATCTACGC  
ATTCACCGCTACACTAGGAATTCCGTTACCTCTCCGGACTCTAGATTGACAGTTCCAATGCAGTCCGGGTT  
GAGCCCCGGGTTTCACATCAGACTTGCACTCCGCTACGCTCCCTTACACCCAGTAAATCCGATAACGCTG

-continued

ACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGTGCTCTTAGTCAGGTACCGTCATTTCTTCCCTGC  
TGATAGAGCTTACATACCGAAATACTTCATCGCTACGCCGCGCTGCATCAGGGTTCCCCATTGTGCAATA  
TTCCCCACTGCTGCCTCCGGGGAGTTGGA

SEQ ID NO: 71

| YK152 |  
GTGATCAGCTTCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTAC  
GCATTTACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCTAGAAAAACAGTTCCAATGCAGTCCTGGGG  
TTAAGCCCAGCCTTCACATCAGACTTGCTCTCCGTCTACGCTCCCTTACACCCAGTAAATCCGGATAACGCTT  
GCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTCATTTCTCCCT  
GCTGATAGAAGTTACATACCGAGATACTTCTTCACGCCGCGCTGCATCAGGGTTCCCCATTGTGCAA  
TATTCCCCACTGCTGCCTCCGGGGAGTTGGA

SEQ ID NO: 72

| YK155 |  
TTGATCAGCTTCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTAC  
CATTTCACCGCTACACTAGGAATTCCGCTTACCTCTCCGGCACTCTAGAAAAACAGTTCCAATGCAGTCCTGGGG  
TAAGCCCAGCCTTCACATCAGACTTGCTCTCCGTCTACGCTCCCTTACACCCAGTAAATCCGGATAACGCTTG  
CCCCCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGGGCTCTTAGTCAGGTACCGTCATTTCTCCCTG  
CTGATAGAAGTTACATACCGAGATACTTCTTCACGCCGCGCTGCATCAGGGTTCCCCATTGTGCAAAT  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 73

| YK157 |  
GTATTCAGCTTCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTA  
CGCATTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGG  
GTTGAGCCCCGGGCTTCACATCAGACTTGCCGCACCGCCTGCGCTCCCTTACACCCAGTAAATCCGGATAACGCT  
TGCACCACATCGTATTACCGCGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTATCTCC  
TGCTGATAGAGCTTACATACCGAAATACTTCTCGCTACGCCGCGCTGCATCAGGGTTCCCCATTGTGCA  
ATATTCCCCACTGCTGCCTCCGAAGGGAGTTGGA

SEQ ID NO: 74

| YK160 |  
GTCAGCTTCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTACGC  
ATTCACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGGTT  
GAGCCCCGGGCTTCACATCAGACTTGCCGCACCGCCTGCGCTCCCTTACACCCAGTAAATCCGGATAACGCTTG  
ACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTCTCCCTG  
CTGATAGAGCTTACATACCGAAATACTTCTCGCTACGCCGCGCTGCATCAGGGTTCCCCATTGTGCAAAT  
ATTCCCCACTGCTGCCTCCGAAGGGAGTTGG

SEQ ID NO: 75

| YK166 |  
TTTCAGCTTCGAGCCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTACGC  
ATTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGGTT  
GAGCCCCGGGCTTCACATCAGACTTGCCGCACCGCCTGCGCTCCCTTACACCCAGTAAATCCGGATAACGCTTG  
ACCATACGTATTACCGCGCTGCTGGCACGTATTAGCCGGTGTCTTAGTCAGGTACCGTCATTATCTCCCTG  
TGATAGAGCTTACATACCGAAATACTTCTCGCTACGCCGCGCTGCATCAGGGTTCCCCATTGTGCA  
TTCCCCACTGCTAGCTCCGAAGGGAGTTGGA

SEQ ID NO: 76

| YK168 |  
AGCTCAGCTTCGAGCTCACGTCAAGTACCGTCCAGTAAGCCGCCTCGCCACTGGTGTTCCTCCTAATATCTACG  
CATTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGGTT

-continued

TGAGCCCCGGGTTTACATCAGACTTGCGCACCCCTGCCTTACACCCAGTAAATCCGGATAACGCTTG  
 CACCATACTATTACCGGGCTGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTATCTCCCTG  
 CTGATAGAGCTTACATACCGAAATACTTCTCGCTCACGCCGCTCGCATCAGGCTTCGCCATTGTGCAAT  
 ATTCCCCACTGCTGCCCTCCGAAGGGAGTTGGA

SEQ ID NO: 77

| YK169 |  
 GTCCAGCTTCGAGCCTCAGCGTCAGTTATCGCCAGTAAGCCGCTTCGCCACTGGTGTCTCTTAATATCTACG  
 CATTTCACCGCTACACTAGGAATTCCGCTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCGGGT  
 TGAGCCCCGGGTTTACATCAGACTTGCGCACCCCTGCCTTACACCCAGTAAATCCGGATAACGCTTG  
 CACCATACTATTACCGGGCTGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTATCTCCCTG  
 CTGATAGAGCTTACATACCGAAATACTTCTCGCTCACGCCGCTCGCATCAGGCTTCGCCATTGTGCAAT  
 ATTCCCCACTGCTGCCCTCCGAAGGGAGTTGGA

SEQ ID NO: 78

| YK171 |  
 TGAGCCGGCTACCCCGGTACTGCATTGAACTGTCGACTAGAGTGTGGAGGGTAAGCGGAATTCTCTAGTGTA  
 GCGGTGAAATGCGTAGATATTAGGAGGAACACCACTGGCGAAGGGCGTTACTGGACGATAACTGACGCTGAGGCTC  
 GAAAGCGTGGGAGCAAACAGGATTAGATAACCGGTAA

SEQ ID NO: 79

| YK192 |  
 CACGATGTCAGCTTCGAGCTAGCGTCAGTTATCGCCAGTAAGCCGCTTCGCCACTGGTGTCTCTTAATATC  
 TACGCATTCACCGCTACACTAGGAATTCCACTTACCCCTCCGACACTCTAGTACGACAGTTCCAATGCAGTACCG  
 GGTTGAGCCCCGGGTTTACATCAGACTGGCGCACCGCCTGCCTCCCTTACACCCAGTAAATCCGGATAACG  
 CTTGCACCATACGTATTACCGGGCTGCTGGCACGTATTAGCCGGTCTTAGTCAGGTACCGTCATTCTCTT  
 CCCTGCTGATAGAGCTTACATACCGAAATACTTCTCGCTCACGCCGCTCGCATCAGGGTTCCCCATTG  
 GCAATATTCCCCACTGCTGCCCTCCGAAGGGAGTTGGA

SEQ ID NO: 80

| VE202-18 |  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAATACATGCAAGTCAACGGGAGGACTTGTG  
 CTCGAGTGGCGAACGGGTGAGTAATACATAAGTAACCTGCCCTAGACAGGGGATAACTATTGGAAACGATAGCTAA  
 GACCGCATAGGTACGGACACTGCATGGTGACCGTATTAAAGTGCTCAAAGCACTGGTAGAGGGATGGACTTATGG  
 CGCATTAGCTGGTGGCGGGTAAACGGCCCACCAAGGCAGATCGTAGCCGACCTGAGAGGGTACCGGACAC  
 TGGGACTGAGACACGGCCCAGACTCCTACGGAGGCAGCAGTAGGAAATTTCGGCAATGGGGAAACCTGACCGA  
 GCAACCGCGCTGAAGGAAGAAGGTTTCGGATTGTTAAACTCTGTATAAGGAAGAACGGCGCTACAGGAAT  
 GGTAGCCGAGTGACGGTACTTTATTAGAAAGCCACGGCTAAACTACGTGCCAGCGCGGTAAACGTAGGTGGCA  
 AGCGTTATCCGAAATTGGCGTAAAGAGGGAGCAGGCGGAGCAAGGGTCTGTGGTAAAGCCTGAAGCTTAAC  
 TTCAGTAAGCCATAGAAACCAGGCAGCTAGAGTGCAGGAGAGGATCGTGGATTCCATGTGAGCGGTAAATGCC  
 TAGATATATGGAGAACACCAGTGGCGAACGGCAGATCTGGCCTGCACACTGACGCTCAGTCCGAAAGCGTGGGA  
 GCAAATAGGATTAGATAACCTAGTAGTCCACGCCGTAAACGATGAGTAACAGTGTGGATGTCAAAGTTCAGTG  
 GCAGTTAACGCAATAAGTAACCGTCCGCTGAGTAGTCAGTGCAGAATGAAACTCAAAGGAATTGACGGGGCC  
 CACAAGCGGTGGAGCATGTGGTTAATCGAAGCAACGCGAAGAACCTTACCGGTCTGACATACTCATAAAGGCT  
 CCAGAGATGGAGAGATAGCTATATGAGATAACAGGTGGTGATGGTGTGCTCAGTCGTGAGATGTTGGGTT  
 AAGTCCCGAACGAGCGAACCCCTATCGTTAGTTACCATCATTAAGTTGGGACTCTAGCGAGACTGCCAGTGAC  
 AAGCTGGAGGAAGCGAACGCCGAGGTGAAGCAAAACCCATAAAACCATCTCAGTCGGATTGTAGTCGCAACTC  
 GTGCAGAGGGAAAGCGAACGCCGAGGTGAAGCAAAACCCATAAAACCATCTCAGTCGGATTGTAGTCGCAACTC  
 GACTACATGAAGTTGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGTGAATCAGTTCTGGCCATTGTACACA

-continued

CCGCCCGTCACACCACGAGAGTTGATAACACCGAAGCCGTGGCTAACCGCAAGGAAGGAGCTGCTAAGGTGG  
ATTGATGATTGGGTGAAGTCGTAACAAGGTATCCCTACGGAACGTGGGATGGATCACCTCTT

SEQ ID NO: 81

| PE5 |  
ATGAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTGAAGG  
AAGTTTCGGATGGAATTGATTGACTGAGTGGCGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGGGG  
ATAACAGTTAGAAATGACTGCTAATACCGATAAGCGCACAGTACCGCATGGTACAGTGTGAAAAACTCCGGTGGT  
GTGAGATGGATCCCGTCTGATTAGCCAGTTGGCGGGTAACGGCCCACCAAAGCGACGATCAGTAGCCGACCTGAG  
AGGGTGACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAAATTGACACAAATG  
GGCGAAAGCCTGATGCAGCGACGCCGTGAGTGAGAAGTATTGCGTATGAAAGCTCTACAGCAGGGAAAGAA  
AATGACGGTACCTGACTAAGAAGCCCCGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTTAT  
CCGGATTACTGGGTGAAAGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAGGGCTAACCTGGAC  
TGCTTTGGAAACTGTTGCTAGAGTGTGGAGGGTAAGTGTGAAATTCTAGTGTAGCGGTGAAATGCGTAGATAT  
TAGGAGGAACACCACTGGCGAAGGCGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAA  
GGATTAGATAACCTGGTAGTCACCGCTAACCGATGAATGCTAGGTGTTGGGGCAAAGCCTTCGGTCCGTC  
CAAACCGCAGTAAGCATTCCACCTGGGAGTAGCTCGCAAGAATGAAACTCAAAGGAAATTGACGGGACCCGACA  
AGCGGTGGAGCATGTGTTAATTGAAAGCAACCGAAGAACCTTACCAAGTCTTGACATCCTCTTGACCGCGTGT  
AACGGCGCCTTCCTCGGGGAAAGAGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGTT  
AAGTCCCGAACAGCGCAACCCATTCTTAGTAGCCAGCAGGTAAAGCTGGCACTCTAGGGAGACTGCCAGGG  
ATAACCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGATTGGCTACACACGTGCTAACATGGC  
GTAAACAAAGGAAGCAAGACAGTGATGTGGAGCAAATCCAAAATAACGTCAGTGTGAACTGCAAC  
CCGACTACACGAAGCTGAATCGCTAGTAATCGCAATCAGAATGTCGCGTGAATACGTTCCGGTCTGTACA  
CACCGCCCGTCAACACCATGGGAGTCAGCAACGCCGAAGTCAGTGACCCAACCGCAAGAGAGGGAGCTGCCAGG  
CGGGGCAGGTAACGGGTGAAGTCGTAACAAGGTAGCCGTACGGAGGTGCGGCTGGATCACCTCTT

SEQ ID NO: 82

| PE9 |  
AATTGACGTTGTCGGATTACTGGCGTAAAGGGAGCGTAGGTGGACTTTAAGTGAGATGTGAAATACCCGGCT  
CAACTTGGGTGCTGCATTCAAACCTGAAAGTCTAGAGTGCAGGAGAGGAGAATGAAATTCTAGTGTAGCGGTGAA  
TGCCTAGAGATTAGGAAGAACCCAGTGGCGAAGGCAGATTCTCTGACTGTAACGTACGCTGAGGCTCGAAAGCGTG  
GGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTAGGGTTGTGACCT  
CTGTGCCCGCTAACGCTTAAGTATTCCCGCTGGGAGTAGCGTGCAGGATTAAAGGAAATTGACCGGA

SEQ ID NO: 83 | 211-B |

ACGAGCGTATCGGATTATTGGGTTAAGGGAGCGTAGGTGGATTGTTAAGTCAGTGTGAAAGTTGCGGCTAAC  
GTAAAATTGCAAGTGGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGCGGAATTGCGTGTAGCGGTGAAATGCTT  
AGATATCACGAAGAAACTCCGATTGCGAAGGCAGCTCACTAGACTGCACTGATGCTCGAAAGTGTGGTAT  
CAAACAGGATTAGATAACCTGGTAGTCCACACAGTAAACGATGAATACTCGCTGTTGCGATACAGTAAGCGGCC  
AAGCGAAAGCATTAAGTATTCCACCTGGGAGTAGCCGGCACCGTAAACGAAATTGACGGAGGCC  
CAGGGGGAAAAACATGGGTTAGTGGATGATAACGGGAGGAACCTC

SEQ ID NO: 84

| NR\_119185.1 | *Ruminococcus obeum* 16S ribosomal RNA gene, complete  
sequence  
GGCGGGCGTGCTTAACACATGCAAGTCGAACGGGAAACCTTCATTGAAGCTTCGGCAGATTGGNNTGTTCTAGTG  
GGGACGGGTGAGTAACCGCAGGGTAAACCTGCCCTATACAGGGGATAACACCGAAATGGTGCTAACCCGAT  
AAGCGCACAGGACCGCATGGCTGGTGTAAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCTAGTTGG  
CAGGGTAACGGCCTACCAAGGGCAGCATCCAGCCGGCTGAGAGGGTGAACGGCCACATTGGACTGAGACACGG

## US 10,555,980 B2

119

120

-continued

CCCAAGACTCCTACGGGAGGCAGCAGTGGGAATATTGCACAATGGGGAAACCTGATGCAGCGACGCCGCGTGAAG  
 GAAGAAGTATCTCGGTATGTAACCTCTATCAGCAGGGAAAGATACTGACGGTACCTGACTAAAGAACCCCCGGCTAAC  
 TACGTGCCAGCAGCCCGGTAACTAGTAGGGGCAAGCGTTACCGGATTACTGGGTAAAGGGAGCGTAGACGG  
 ACTGGCAAGTCTGTGAAAGGGGGCTCAACCCCTGGACTGCATTGGAAACTGTTAGTCAGTGAGTGCAGGAGA  
 GGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCAGCTACTGG  
 ACGGTAACGTGAGCTGGGGCTGAAAGCGTGGGGAGCAAACAGGATTAGATAACCGTAGTCCACGCCGAAACGA  
 TGATTACTAGGTGTTGGGGAGCAAAGCTCTCGGTGCCGCCAAACGCATTAAGTATTCCACCTGGGAGTACGTT  
 CGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGACAACGGTGGAGCATGTTAATTGAAAGCAACCG  
 AAGAACCTTACCAAGTCTTGACATCCCTCTGACCGNCCCTAACCGATCTTCCTCGGACAGGGAGACAGGTG  
 GTGCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCAGAGCGCAACCCCTATCCCAGTAG  
 CCAGCAGTCCGGCTGGGACTCTGAGGAGACTGCCAGGGATAACCTGGAGGAGGGGGATGACGTCAAATCATCA  
 TGCCTTATGATTGGGCTACACAGTGCTACAATGGGTAAACAAAGGGAGCAAGCCTGCGAAGGTAAGCAAAT  
 CCCAAAATAACGTCCCAGTCGACTGCAGTCGAACTCGACTGCACGAGCTGGAACTGCTAGTAATCGGGAT  
 CAGAATGCCCGGGTGAATACGTTCCGGGTCTGTACACACCGCCCGTACACCATGGAGTCAGTAACGCCGAAG  
 TGAGTACCTAACTGCAAAGAAGGAGCTGCCGAAGGGGGACCGATGACTGGGTGAAGTCGTAACAAGGT

SEQ ID NO: 85

|NR\_118692.1|Ruminococcus obeum strain ATCC 29174 16S ribosomal RNA  
 gene, complete sequence  
 GGCAGCTTAACACATGCAAGTCGAACGGAAACTTTCATTGAAGCTCGGAGATTGGCTGTTCTAGGGCG  
 GACGGGTGAGTAACGCGTGGGTAACCTGCCTTATACAGGGGATAACAAACAGAAATGGTTGTAATACCGCATAAG  
 CGCACAGGACCGCATGGTCTGGTGTAAAAACTCCGGTGTATAAGATGGACCCGCGTGGATTAGCTAGTGGCAG  
 GGTAACGGCCTACCAAGCGACGATCCATGCCGCGTGAAGGGGTGAACGGCCACATTGGACTGAGACACGCC  
 AGACTCTCGGAGGCAGCAGTGGGAATTGCAAAATGGGGAAACCTGATGCAGCGACGCCGCGTAGAGGAAG  
 AAGTATCTCGGTATGTAACCTCTATCAGCAGGGAAAGATACTGACGGTACCTGACTAAAGAACCCCGKCTAACTAC  
 TGCCAGCAGCCCGGTAAACGTAGGGGCAAGCGTTACCGGATTACTGGGTGAAGGGAGCGTAGACGGACTG  
 GCAAGTCTGATGTGAAAGGGGGGCTAACCCCTGGACTGCATTGGAAACTGTTAGTCAGTGAGTGCAGGAGAGGTA  
 AGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGGGGTTACTGGACGG  
 TAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCACGCCGAAACGATGAA  
 TACTAGGTGTTGGGGAGCAAAGCTTCGGTGCCGCCAACGCATTAAGTATTCCACCTGGGAGTACGTTGCA  
 AGAATGAAAATCAAAGGAATTGACGGGGACCCGACAAGCGGTGGAGCATGTTAATTGAAAGCAACGCCAAGA  
 ACCTTACCAAGTCTTGACATCCCTGACCGTCCCTAACCGATTTCTCGGACAGGGAGACAGGTGGTGC  
 ATGGTTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCAGAGCGCAACCCCTATCCCAGTAGGCCAG  
 CAGTNCGGCTGGGACTCTGAGGAGACTGCCAGGGATAACCTGGAGGAAGGGGGATGACGTCAAATCATG  
 CCTTATGATTGGGCTACACAGTGCTACAATGGGTAAACAAAGGGAGCNAGCTKGAGGTAAAGCAAATCCA  
 NAAATAACGTCCCAGTCGGACTGCAGTCGAACTCGACTGCACGAGCTGGAACTGCTAGTAATCGCGGATCAGA  
 ATGCCCGGGTGAATACGTTCCGGGTCTGTACACACCGCCCGTACACCATGGAGTCAGTAACGCCGAAGTCAG  
 TGACCTAACTGC

SEQ ID NO: 86

|NR\_026491.1|Clostridium disporicum strain DS1 16S ribosomal RNA gene,  
 partial sequence  
 GCTCAGGACGAAACGCTGGCGCGTGCCTAACACATGCAAGTCGAGCGAGTTGATTCTTCGGAGATGAAGCTAGCG  
 GCGGACGGGTGAGTAACACGTGGGACACCTGCCCTAGAGGGGAATAGCCTCCGAAAGGGAGATTAAACCGCAT  
 AAGATTGTAGCTTCGCTGAGTAGCAATTAAAGGAGCAATCCGCTATGAGATGGGCCCGCGCATTAGCTAGTT  
 GGTGAGGTAAACGGCTACCAAGCGACGATGCCAGCGTACCGCACCTGAGAGGGTGAACGGGACATTGGACTGAGACAC

-continued

```

GGCCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGACAATGGGGAAACCTGATGCAGCAACGCCGTGA
GTGATGACGGCTTCGGTTGTAAGCTCTGCTTCAGGGACGATAATGACGGTACCTGAGGAGGAAGCCACGGCTA
ACTACGTGCCAGCGCCGTTAACGTAGGTGGCGAGGCTTGCTCGGATTACTGGCTAAAGGGAGCGTAGGC
GGACTTTAAGTGGAGATGTGAAATACCGGGCTCAACTTGGTGCTGCATTCAAACGTGGAGTCTAGAGTGCAGGA
GAGGAGAATGAAATTCTAGTGTAGCGGTGAAATCGTAGAGATTAGGAAGAACACCAGTGGCGAAGGCATTCT
GGACTGTAACGTGAGCTGAGGCTGAAAGCGTGGGAGCAGAACAGGATTAGATAACCGTGGTAGTCCACGCCGTAAAC
GATGAATACTAGGTGTAGGGTTGTCATGACCTCTGCGCCGCTAACGCTTAAGTATTCCGCTGGGAGTACG
GTCGCAAGATAAAACTCAAAGGAATTGACGGGGCCCGACAAGCAGCGGAGCATGTGGTTAATTGAAAGCAACG
CGAAGAACCTTACCTAGACTGACATCTCCTGAATTACCGTAACGGGAGCCACTTGGGTGGCAGGAAGACAGG
TGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTATTGTTAGT
TGCTTACCATTTAGTGTAGGACTGACCTGAGACTGCGGGTTAACCGGAGGAAGGTGGGATGACGTCAAATCATC
ATGCCCTTATGCTAGGGCTACACACGTCTAACATGGCAAGTCAAAGAGAACGACCGCGAGGTGGAGCAAA
ACTCAAAAACCTGTCTCAGTTCGGATTGTAGGCTAACACCGCCGCTCACCATGAGAGTTGGTAGTAACTCGCAAT
CAGCATGTCGGTGAATACGTTCCGGCCTGTACACACCGCCGCTCACCATGAGAGTTGGCAATACCAACG
TACGTGATCTAACCGCAAGGGAGGAAGCGCTAACGGTAGGCTAGCGATTGGGTGAAGTCGTAACAAGGTAGCC
GTAGGAGAA

```

SEQ ID NO: 87

```

|NR_028785.1|Clostridium scindens strain ATCC 35704 16S ribosomal RNA
gene, complete sequence
GAGAGTTGATCCTGGCTCAGGATGAAACGCTGGCGCGTGCCTAACACATGCAAGTCGAACGAAGCGCCTGGGGCCG
ACTTCTCGAACGAGGAGCCTTGCAGTGGGGACGGGTAGTAACCGTGGCAACCTGCCTGCACTGGG
GGATAACAGCCAGAAATGGCTGCTAACCGCATAAGACCGAAGCGCCGATGGCGGGGCAAAGCCCCGGCG
TGCAAGATGGCCCGCTGATTAGTAGTTGGGGTAACGGCCACCAAGCCGACGATCAGTAGCCGACCTGA
GAGGGTACCGGCCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCTGGGAATATTGACAAT
GGGGAAACCTGATGCAAGCGACGCCGCTGAAGGATGAAGTATTGCTATGAAACTTCTATCAGCAGGAAGAA
GATGACGGTACCTGACTAAGAACCCCCGTAACACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTTAT
CCGGATTTACTGGGTAAAGGGAGCGTAGACGGCGATGCAAGCCAGATGTGAAAGCCGGGCTAACCCGGGAC
TGCATTTGGAACTGCGTGGAGGTGTCGGAGAGGCAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATT
AGGAGGAACCCAGTGGCGAAGGCGCTGCTGGACGATGACTGACGTTGAGGCTGAAAGCGTGGGAGCAACAG
GATTAGATAACCTGGTAGTCCACGCGTAAACGATGACTACTAGGTGTCGGTGGCAAGGCCATTGGTGCAGCAGC
AACCGAATAAGTAGTCACCTGGGAGTAAGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG
CCGTGGAGCATGGTTAACCGAAGCGAACCTTACCTGATCTGACATCCGATGCCAAAGCGCGTA
ACCGCCTTCTCGAACATCGGTGACAGGTGGCATGGTGTGCTAGCTGTGAGATGTTGGTTAA
GTCGGCAACGCGAACCCCTATCTCAGTAGCCAGCATTTGGATGGCACTCTGGAGAGACTGCCAGGGAGAA
CCTGGAGGAAGGTGGGATGACGTCAAATCATGACCGCCCTATGACCAGGGCTACACACGTGCTACAATGGCGTAA
ACAAAGGGAGCGAACCCGCGAGGGGGAAATCCAAAAAAACGTCTCAGTTGGATTGAGTCTGCAACTCGA
CTACATGAAGTTGAATCGCTAGTAATCGCAATCAGAATGTCGGTGAATACGTTCCGGTCTGTACACACCG
CCCGTCACACCATGGGAGTCAGTAACGCCGAAGCCGGTACCCAACCGTAAGGGAGGGAGCGTCGAAGGTGGGA
CCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTTC

```

SEQ ID NO: 88

```

|NR_028915.1|Anaerostipes caccae strain L1-92 16S ribosomal RNA gene,
partial sequence
GCGCTTAATACATGTCAGTCGAACGAAGCATTAGGATTGAAGTTGGATGGATTCTATGACTGAGTGGC
GGACGGGTGAGTAACCGTGGGAACCTGCCCTATACAGGGGATAACAGCTGGAAACGGCTGCTAACCGCATAA

```

-continued

CGCGCACAGAATCGCATGATTCACTGTGAAAAGCCCTGGCAGTATAGGATGGTCCCGCTCTGATTAGCTGGTTGGTG  
 AGGTAACGGCTACCAAGGCAGATCAGTAGCCGGCTTGAGAGAGTGAACGCCACATTGGACTGAGACACGGCC  
 CAAACTCCTACGGGAGGCAGCAGTGGGAATATTGACAAATGGGGTAAACCTGATGCAGCGACGCCGCTGAGTG  
 AAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAAGAAAACAGACGGTACCTGACTAAGAACGCCGCTAAC  
 TACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTACCGGAATTACTGGGTGAAAGGGTGCCTAGGTGG  
 CATGGTAAGTCAGAAGTGAAGGCCGGGCTTAACCCGGACTGCTTTGAAACTGTATGCTGGAGTGCAGGAGA  
 GGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGAACACCAGTGGCGAAGGCCGTTACTGG  
 ACTGTCACTGACACTGATGCACGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAACGA  
 TGAATACTAGGTGTCGGGCGTAGAGGCTTCGGTGCAGCAAACGAGTAAGTATTCCACCTGGGAGTACGTT  
 CGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTGAAGCAACGCG  
 AAGAACCTTACCTGGCTTGACATCCAATGACCGAACCTAACCGTTTTCTTCGAGACATTGGAGACAGGTG  
 GTGCATGGTTGTCGTCACTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAACCCCTATTTAGTAG  
 CCAGCATTAAAGGTGGGACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAGTGGGACGACGTCAAATCATCA  
 TGCCCCCTATGCCAGGGCTACACAGTGCTACAATGGCTAAACAAAGGAAGCGAAGTCGTGAGGCGAAGCAAAT  
 CCCAGAAATAACGTCTCAGTCGGATTGAGCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGTGAAT  
 CAGAATGTCACGGTGAATACGTTCCGGGCTTGTACACACCGCCCGTACACCATGGGAGTCAGTAACGCCGAAG  
 TCAGTGACCCAACCGCAAGGAGGGAGCTGCCGAAGGTGGACCGATAACTGGGTGAAGTCGTAAACAAGG

SEQ ID NO: 89

|NR\_042152.1|*Marvinbryantia formatexigens* strain 1-52 16S ribosomal RNA  
 gene, partial sequence >i|636558750|ref|NR\_114807.1|*Marvinbryantia*  
*formatexigens* strain I-52 16S ribosomal RNA gene, complete sequence  
 TGGCGGGGTGCTTAACACATGCAAGTCGAGCGAACGATTTAAATGAAGTTTCGGACGGAAATTAAATGACTGAG  
 CGCGGGACGGGTGAGTAACCGTGGATAACCTGCCTTATACAGGGGATAACAGCCAGAAATGGCTGCTAATACCGC  
 ATAAGCGCACGGTACCGCATGGTACAGTGTAAAAACTCCGGTGGTATAAGATGGTCCCGCTGGATTAGGCAGTT  
 GGCAGGGTAAAGGCCACCAACCGACGATCCATAGCCGGCTGAGAGGGTGGACGGCCACATTGGACTGAGACAC  
 GGCCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGACAAATGGGGAAACCTGATGCAGCGACGCCGCTGG  
 GTGAAGAAGTATTTCGGTATGTAAAGCCCTATCAGCAGGGAAAGAAAATGACGGTACCTGACCAAGAACGCCGCTA  
 ACTACGTGCCAGCAGCCGGTAATACGTAGGGCAAGCGTTACCGGATTACTGGGTGAAAGGGAGCGTAGAC  
 GGCCATGCAAGTCTGGTGTAAAGGGGGCTCAACCCCGACTGCATTGAAACTGTATGGCTTGAGTGCCTGG  
 GAGGTAAAGCGGAATTCTGGTGTAGCGTGAAGATCGTAGATATCAGGAGGAACACCAGTGGCGAAGGCCGTTACT  
 GGACGGTAACTGACGTTGAGGCTGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAC  
 GATGAATACCAAGGTGCGGGGACACGGCTTCGGTGCAGCAAACGCACTAAGTATTCCACCTGGGAGTACG  
 TTGCGAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTGAAGCAACG  
 CGAAGAACCTTACCAAGGTCTGACATCCGGACCCGACAGTAACGTGTCCCTCCCTGGGGCGTGGAGACAGG  
 TGGTGCATGGTTGTCGTCACTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAACCCCTGTTCCAGT  
 AGCCAGCATTCAAGGATGGGACTCTGGGAGACTGCCAGGGATAACCTGGAGGAAGGCCGGATGACGTCAAATCAT  
 CATGCCCTTATGATCTGGCTACACAGTGCTACAATGGCGTAACAGAGGGAAAGCGAACCCCGCAGGGGAGCAA  
 ATCCCGAGAAATAACGTCCCAGTTCGGATTGAGCTGCAACCCGGTACATGAAGCTGGAATCGCTAGTAATCGCG  
 ATCAGGATGCCGGGTAAATACGTTCCGGGCTTGTACACACCGCCCGTACACCATGGGAGTCGGAAATGCCGA  
 AGTCAGTGACCCAACCGGAAGGAGGGAGCTGCCGAAGGCCGGTAACGGGTGAAGTCGTAAACAA

-continued

SEQ ID NO: 90

|NR\_024994.1| *Lactobacillus mucosae* strain S32 16S ribosomal RNA gene, complete sequence  
 AGAGTTGATCTGGCTCAGGATGAACGCCGGTGTGCCTAATACATGCAAGTCGAACCGTGGCCAACTGAT  
 TGAACGTGCTTGACGGACTTGACGTTGGTTACAGCGAGTGGCGACGGGTGAGTAACACGTAGGTAACCTGCC  
 CAAAGCGGGGATAACATTGGAAACAGATGCTAATACCGCATAAACAATTGAATCGCATGATTCAAATTAAAAGA  
 TGGCTTCGGCTATCACTTGGATGGACCTGCCGGCATTAGCTTGGTAGGGTAACGCCAACCAAGGTGTGA  
 TCGTAGCCGAGTTGAGAGACTGATCGGCCAACATGGAACGTGAGACACGGTCCACTCCTACGGGAGGCAGCAGTA  
 GGGAACTTCCACAATGGCGCAAGCCTGATGGAGCAACACCGCGTGAGTGAGAAGAAGGTTTGGCTCGTAAAGCTC  
 TGGTGTAGAGAAGAACGTGCGTGAGAGCAACTGTTACGCGAGTGGCTACCGAGAAAGTCACGGCTAACTA  
 CGTGCAGCGCCGCGTAATACGTAGGTGGCAAGCGTTACCGGATTATTGGCGTAAAGCGAGCGCAGGCGTT  
 TGATAAGTCTGTGAAAGCCTTGGCTAACCAAAGAAGTCATCGGAAACTGTCAGACTTGAGTGAGAAGAG  
 ACAGTGGAACTCCATGTGAGCGGTGAATGCGTAGATATATGGAAGAACACCAGTGGCAAGGCGCTGCTGGTC  
 TGCAACTGACGCTGAGGCTCGAAAGCATGGTAGCGAACAGGATTAGATAACCTGGTAGTCCATGCCGTAACGATG  
 ATGCTAGGTGTGGAGGGTTCCGCCCTCAGTGGCGAGCTAACGCTAACGACTCCGCTGGGAGTACGACC  
 GCAAGGGTAAACTCAAAGGATTGACGGGGCCCGACAAGCGGTGGAGCATGGTTAACTGAAGCTACGCGA  
 AGAACCTTACCAAGGTCTGACATCTTGCCTAACCCCTAGAGATAGGGCTTCCCTCGGAACGCAATGACAGGTGG  
 TGCATGGTCGTCGTCAGCTGCTGTGAGATGGTTAACTCCGCAACCGTAGAGATAGGGCTTACTAGTTGC  
 CAGCATTCAAGTGGGACTCTAGTGAGACTGCCGGTACAAACCGGAGGAAGGTGGGAGCACGTCAGATCATG  
 CCCCTTATGACCTGGGACTCACACCGTCTAACATGGACGGTACACGAGTCGCGAACCTCGCAGGGCAAGCTAATCT  
 CTTAAACCGTCTCAGTCCGACTGCAGGCTGCAACTGCCGTGACGAAGTCGAATCGTAGTAATCGCGATCA  
 GCATGCCGCGGTAAACGTTCCGGGCTGTACACACCCTGCAACCATGAGAGTTGCAACACCCAAAGTC  
 GTGGGTAACCTCGGGAGCTAGCCGCTAACGGTGGGAGATGATTAGGGTAAGTCGTAACAAGGTAGCCG  
 AGGAGAACCTGCCGCTGGATCACCTCCT

SEQ ID NO: 91

|NR\_028816.1| *Turicibacter sanguinis* strain MOL361 16S ribosomal RNA gene, complete sequence  
 AGAGTTGATCATGGCTCAGGATGAACGCTGGCGCGTGCCTAATACATGCAAGTCGAAGCGAACACTCGGTGGTG  
 AGCGCGAACGGGTGAGTAACACGTAGGTTATCTGCCCATCAGACGGGACAACGATTGGAAACGATCGCTAAC  
 GGATAGGACGAAAGTTAAAGGTGCTCGGCACCACTGATGGATGAGCCTGCCGCATTAGCTAGTTGGTAGGGTA  
 AAGGCCTACCAAGGCACGATCGTAGCCGACCTGAGAGGGTAACGCCACACTGGACTGAGACACGCCAGAC  
 CCTACGGGAGGCAGCAGTAGGGAATCTCGGAATGGGAAAGCCTGACCGAGCAACGCCGTGAATGATGAAG  
 GCCTCGGGTTGAAAATTCTTTAAGGAAGAATGGCTAGTAGGAAATGGCTAGAGTGTGACGGTACCTTAT  
 GAGAAAGCCACGGCTAACTACGTGCCAGCGCCGGTAACAGTAGGTGGCAGCGTTACCGGAATTGGCG  
 TAAAGAGGCCAGGTGGTTGATTAAGTCTGTGAAAGCCCACGGTTAACCGTGGAGGGCATTGGAAACTGGT  
 CAACTTGAGTCGAGAAGAGGGAAAGTGGAACTCCATGTGAGCGGTAAACGCTAGAGATATGGAGGAACACCAGTG  
 GCGAAGGCCAGGTGGCTCTGGTCTGTAACGTGACACTGAGGCGCAGCGTGGGAGCAACAGGATTAGATA  
 AGTCCACGCCGTAACGATGAGTGCTAAGTGTGGGGTCGAACCTCAGTGTGAGTTAACGGCTAAC  
 CCTGGGAGTACGGTCGCAAGACTGAAACTCAAAGGAAATTGACGGGACCGCACAAGCGGTGGAGCATGTGGTTA  
 ATTGAGAAGCAACGCCAGAAGAACCTTACCAAGGTCTTGACATACCGTGCCTAGAGATAGGGATTACCCCTGG  
 GACAATGGATAACGGTGGTCATGGTGTGCTCAGCTCGTGTGAGATGGTTAACGCGCAACGAGCG  
 ACCCCTGCTGTTAGTGGCCAGCATTGAGTGGGACTCTAACGAGACTGCCAGTGCACAAACTGGAGGAAGGTGG  
 TGACGTCAAATCATGCCCCCTATGACCTGGCTACACACGTGCTACATGGTTGATCAAAGAGAAGCGAAGCG  
 GTGACGTGGAGCAACCTCATAAGCCAATCTCAGTGGATTGTAGGCTGCAACTCGCCTACATGAAGTGGAAATC

-continued

```
GCTAGTAATCGGAATCAGCATGTCGGTGAATACTGTTCCGGTCTTGACACACCGCCGTACACCACGGAGAG  
TTAACACACCGAAGTCAGTGGCTAACCGCAAGGAGGGACTGCCTAAGGTGGGTAGATGATTGGGTAAAGTC  
GTAACAAGGTATCCCTACCGGAAGGTGGGTGGATCACCTCCTT
```

SEQ ID NO: 92

|NR\_042832.1| Roseburia faecis strain M72/1 16S ribosomal RNA gene, partial sequence  
GATGAACGCTGGCGGGCTGCTTAACACATGCAAGTCAACGAAGCAAGCACTCTATTGATTTCCTCGGAAATGAAGATT  
TTGTGACTGAGTGGCGGACGGGTGAGTAACCGCTGGTAACCTGCCTACACAGGGGATAACAGTTGGAAACGACT  
GCTAATACCGATAAGCGCACAGGATCGCATGATCCGGTGTGAAAAACTCCGGTGGTATGAGATGGACCCCGCTG  
ATTAGCCAGTGGCAGGGTAACGGCTACCAAAGCGACGATCAGTAGCCGACCTGAGAGGGTACCGGCCACATTGG  
GACTGAGACACGGCCAAACTCTACGGGAGGCAGCAGTGGGAATATTGCACAATGGGGAAACCTGATGCAGCG  
ACGCCGCGTGAAGCGAAGAAGTATTCGGTATGTAAGCTCTACGCAAGGAAAGAAGATGACGGTACCTGACTAAG  
AAGCACCGCTAAATACGTGCCAGCAGCGCGTAATACGTATGGTCAAGCGTTATCCGATTTACTGGGTGAAA  
GGGAGCGCAGGGGGTGCAGGCAAGTGTGAAAGCCGGGCTCAACCCGGTACTGCATTGAAACTGTCGTAC  
TAGAGTGTGGAGGGGTAAGTGGATTCTAGTGTAGCGTGAATGCGTAGATATTAGGAGGAACACCAAGTGGCGA  
AGGCGCTTACTGGACGATAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAACAGGATTAGATAACCCGGTAGTC  
CACGCCGTAAACGATGAATACTAGGTGTCGGGGAGCATTGCTTCGGTGCAGCAACCGCAATAAGTATTCAC  
CTGGGGAGTACGTTCGAAGAATGAAACACTCAAAGGAATTGACGGGACCCGACAAGCGTGGAGCATGTGGTTAA  
TTCGAAGCAACCGAAGAACCTTACCAAGTCTTGACATCCGATGACAGAGTATGTAATGTACYTTCTTCGGAGC  
ATCGGTGACAGGTGGTCATGGTTGCTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCACAGCGCAACCG  
CCTGTCCTTAGTAGCCAGCGTTGGCCGGGACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGATG  
ACGTCAAATCATCATGCCCTTATGACTTGGCTACACACGTCTACAATGGCGTAAACAGGGAGCAGCG  
GAGGCGAGCAAATCTAAAAAAACGTCTAGTCCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCG  
TAGTAATCGCAGATCAGAATGTCGGTGAATACGTTCCGGGTCTGTACACACCGCCGTACACCATGGAGTT  
GGAAATGCCGAAGTCAGTGACCAACCGCAAGGAGGGAGCTGCCAAGGCAGGTTGATAACTGGGTG

SEQ ID NO: 93

|NR\_043142.1| Flavonifractor plautii strain Prevot S1 16S ribosomal RNA gene, partial sequence  
CGCTGGCGCGTGTCTAACACATGCAAGTCAACGGGTGCTCATGACGGAGGATTGCTCAATGGATTGAGTTACC  
TAGTGGCGACGGGTGAGTAACCGCTGAGGAACCTGCCTGGAGGGGAATAACACTCCGAAAGGAGTGCTAAATAC  
CGCATGAAGCAGTTGGTCGATGGCTCTGACTGCCAAAGATTATCGCTCTGAGATGGCTCGCTGATTAGCT  
AGTAGGCGGGTAAACGGCCACCTAGCGACGATCAGTAGCCGACTGAGAGGGTACCGGACATTGGACTGAG  
ACACGGCCAGACTCCTACGGAGGCAGCAGTGGGAATTGGCAATGGCGCAAGCCTGACCCAGCAACGCC  
GTGAAGGAAGAAGGTTGGGGTAAACTCTTTGTCGGGGACGAAACAAATGACGGTACCCGACGAATAAGCC  
ACGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGTGCCAGCGTTATCCGGATTTACTGGGTGAAAGGGCG  
TGTAGGCGGGATTGCAAGTCAGATGAAACTGGGGCTCAACCTCCAGCGCTGATTGAAACTGTAGTTCTTGAG  
TGCTGGAGAGGCAATCGGAATTCCGTGTAGCGGTGAATGCGTAGATACGGAGGAACACCAAGTGGCGAAGGCG  
GATTGCTGGACAGTAACTGACGCTGAGGCGCAAAGCGTGGGGAGCAACAGGATTAGATAACCCGGTAGTCCACGC  
CGTAAACGATGGATACTAGGTGTTGGGGCTGACCCCCCTCCGTGCCAGTTAACACAATAAGTATCCACCTGG  
GAGTACGATCGCAAGGGTGAACACTAAAGGAATTGACGGGGCCGACAAGCGTGGAGATGTGGTTAAITCGA  
AGCAACCGAAGAACCTTACCAAGGGCTTGACATCCACTAACGAGGCAGAGATGCGTTAGGTGCCCTCGGGAAAG  
TGGAGACAGGTGGTCATGGTTGCTGAGCTAGCGAGACTGCCGTTGACAAACAGGAGGAAGGTGGGACGACGTCAA  
TATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGACAAACAGGAGGAAGGTGGGACGACGTCAA  
TCATCATGCCCTTATGTCCTGGGCCACACACGTACTACAATGGGTAAACAGAGGGAGGCAATACCGCGAGGTGG

-continued

AGCAAATCCCTAAAAGCCATCCCAGTTGGATTGCAGGCTGAAACCCGCCGTATGAAGTTGAATCGCTAGTAATC  
CGGGATCAGCATGCCGGTGAATACTTCCCCGGCCTTGACACACCAGCCGTACACCATGAGAGTCGGAACAC  
CGAAGTCGCTAGCCTAACCGCAAGGAGGGCGGCCGAAGGTGGGTCATAATTGGGTGAAGTCGTAACAAGGT  
AG

SEQ ID NO: 94

|NR\_044054.1| *Blautia wexlerae* strain DSM 19850 16S ribosomal RNA gene,  
partial sequence  
CAAGTCGACAGGGATTANTTATTGAAACTTCGGTCGATTTAATTAACTTAGTGGCGACGGGTGAGTAACGCG  
TGGGTAACCTGCCTTATACAGGGGATAACAGTCAGAAATGGCTGCTAATACCGATAAGCGCACAGAGCTGCATGG  
CTCAGTGTGAAAACCTCCGGTGTATAAGATGGACCCGCGTGGATTAGCTTGGTGGGTAAACGGCCACCAAG  
GGCACGATCCATAGCCGGCCTGAGAGGGTGAACGCCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGC  
AGCAGTGGGAATATTGACAATGGGGAAACCCCTGATGCAGCGACGCCGTGAAGGAAGAAGTATCTCGGTATGT  
AAACTTCTATCAGCAGGGAAAGATAGTGACGGTACCTGACTAAAGAAGCCCGGTAACACTACGTGCCAGCGCGGT  
AAATACGTAGGGGCAAGCGTTACCGGATTACTGGGTGAAAGGGAGCGTAGACGGTGTGGCAAGTGTGAA  
AGGCATGGGCTAACCTGTGGACTGCAATTGGAAACTGTCAACTTGAGTGCCGGAGGGTAAGCGGAATTCTAGTG  
TAGCGGTGAAATCGCTAGATATTAGAGGAACACCGTAGCGCAAGGGCTTACTGGACGGTAACGTGAGGC  
TCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATAACTAGGTGTCGGT  
GGCAAAGCCATTGGTGCCTCGCAACCGCAGTAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAA  
GGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTAATCGAACGCGAAGAACCTTACCAAGTCTT  
GACATCCGCCTGACCGATCCTAACCGGATCTTCCTCGGGACAGCGAGACAGGTGGTGCATGGTTGCGTCAGC  
TCGTGTCGTGAGATGTTGGTTAAGTCCCGAACGAGCGAACCCCTATCCTCAGTAGCCAGCATTAAGGTGGCA  
CTCTGGGAGACTGCCAGGGATAACCTGGAGGAAGGGGGATGACGTCAAATCATCATGCCCTTATGATTGGC  
TACACACGTGCTACAATGGCTAAACAAAGGGAAAGCGAGATTGTGAGATGGAGAACGAAATCCAAAATAACGTC  
TTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAAATCGCTAGTAATCGCGGATCAGAATGCCGGTGAATA  
CGTCCCGGGCTTGTACACACCGCCGTACACCATGGAGTCAGTAACGCCGAAGTCAGTGACCTAACTGCAA  
GAAGGAGCTGCCGAAGCGGGACCGATGACTGGGTGAAGTCGTAACAAGGT

SEQ ID NO: 95

|NR\_027558.1| *Anaerotruncus colihominis* strain WAL 14565 16S ribosomal  
RNA gene, partial sequence  
AACGGAGCTTACGTTTGAAGTTTCGGATGGATGAATGTAAGCTTAGTGGCGACGGGTGAGTAACAGTGAGCAA  
CTGCCTTCAGAGGGGATAACAGCCGAAACGGCTGCTAATACCGCATGATGTTGCGGGGCACATGCCCTGCA  
ACAAAGGAGCAATCGCTGAAAGATGGCTCGCTCGATTAGCCAGTTGGCGGGGTAAAGGCCACCAAAGCGAC  
GATCGTAGCCGGACTGAGAGGGTGAACGCCACATTGGACTGAGACACGCCAGACTCCTACGGGAGGCAGCAG  
TGGGGATATTGACAATGGCGAAAGCCTGATGCGAGCGACGCCGTGAGGGAGACAGGTCTCGGATTGTAAC  
TCGTCTTGGGAAGAAAATGACGGTACCCAAAGAGGAAGCTCCGGTAACACTACGTGCCAGCGCGGTAATAC  
GTAGGGAGCAAGCGTTGCCGAAATCTGGGTGAAAGGGAGCGTAGGCGGGATGGCAAGTGAATGTTAAATCCA  
TCGGCTAACCGGTGGCTCGTTCTAAACTGCCGTTCTGAGTGAAGTAGAGGCAGGCAGGAAATTCTAGTGAGCG  
TGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCCGTGCTGGGTTAACTGACGCTGAGGCTCGAAA  
GCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTGGGGGACTGAC  
CCCTCCGTGCCGAGTTAACACAATAAGTAATCCACCTGGGAGTACGCCGCAAGGTTGAAACTCAAAGGAATTG  
ACGGGGGCCGACAAGCAGTGGAGTATGTGGTTAATCGAAGCAACGCAAGAACCTTACCGTCTGACATCG  
GCGTAATAGCCTAGAGAGTAGGTGAAGCCCTCGGGCATCCAGACAGGTGGTGCATGGTTGCGTCAGCTCGTGC  
GTGAGATGTTGGTTAAGTCCCGAACGAGCGAACCCCTTATTAGTTGCTACGCAAGAGCACTCTAATGAGACT  
GCCGTTGACAAAACGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCTGGCTACACACGTACTA

-continued

```
CAATGGCACTAAAACAGAGGGCGGCACACCGCGAGGTGAAGCGAATCCCAGAAAAGTGTCTCAGTTAGATTGCA
GGCTGCAACCCGCTGCATGAAGTCGAAATTGCTAGTAATCGCGATCAGCATGCCCGGTAAACGTTCCCGGC
CTGTACACACCACCGGGTACACCATGGGAGTCGGGTAAACACCGAAGCCAGTAG
```

SEQ ID NO: 96

```
|NR_116747.1|Ruminococcus faecis strain Eg2 16S ribosomal RNA gene,
partial sequence
ATGCAAGTCGAACGAAGCACCTGATTGATTCCTCGATGAAGATCTTGGTACTGAGTGGCGACGGGTGAGTAA
CGCGTGGGTAACCTGCCTCATACAGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAGACCACAGCACCGC
ATGGTGCAGGGTAAAAACTCCGGTGGTATGAGATGGACCCCGTCTGATTAGGTAGTTGGGGGTAACGCCCTAC
CAAGCCGACGATCAGTAGGCCACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGCCAAACTCCTACGGG
AGGCAGCAGTGGGAATATTGACAATGGGGAAACCTGATGCAGCGACGCCCGTGGCGATGAAGTATTTCGGT
ATGTAAGCTCTATCAGCAGGGAAAGAAAATGACGGTACCTGACTAAGAACGACCGCTAAATACGTGCCAGCAGCCG
CGGTAATACGTATGGTCAAGCGTTATCCGATTTACTGGGTGAAAGGGAGCGTAGACGGAGTGGCAAGTCTGATG
TGAAAACCGGGGCTCAACCCCCGGACTGCATGGAAACTGTCATCTAGAGTACCGGAGAGGTAAAGCGGAATTCT
AGTGTAGCGGTGAAATGGTAGATATTAGGAGGAACACCACTGGCGAAGGGCGCTTACTGGACGGTAACGTGTTG
AGGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCCCTGGTAGTCCACGCCGAAACGATGACTACTAGGTGTCG
GGCAGCAAAGCTGTCGGTCCCGCAGCAAACGCAATAAGTAGTCCACCTGGGAGTACGTTCGAAGAATGAAACTC
AAAGGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCTGCT
CTTGACATCTCCCTGACCGCAAGTAATGTTGCCCTTCCTCGGACAGGGATGACAGGTGGTGCATGGTTGCGT
AGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGAACCCCTATCTTAGTAGCCAGCGGTTGGCGG
GCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTTATGAGCAG
GGCTACACACGTGCTACAATGGCGAAACAAAGGGAGCGAACCGCGAGGTGAGCAAATCCAAAATAACGTCT
CAGTTGGATTGAGATGTTCTGCAACTCGACTACATGAAGCTGGATCGCTAGTAATCGCAATCAGAATGTCGGGTGA
ATACGTTCCCGGGCTTGTACACACCACCGCCGTACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGT
AAGGAGGAGCTGCCGAAG
```

SEQ ID NO: 97

```
|NR_028883.1|Dorea longicatena strain 111-35 16S ribosomal RNA gene,
partial sequence
TAACCGCGTGGGTAACCTGCCTCATACAGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAGACCACGTACC
GCATGGTACAGTGGAAAAACTCCGGTGGTATGAGATGGACCCCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCT
ACCAAGCCGACGATCAGTAGGCCACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGCCAGACTCCTACG
GGAGGCAGCAGTGGGAATATTGACAATGGGGAAACTCTGATGCAGCGACGCCCGTGAAGGGATGAAGTATTTCG
GTATGTAACCTCTATCAGCAGGGAAAGAAAATGACGGTACCTGACTAAGAACGCCCGCTAACTACGTGCCAGCAGC
CGCGGTAATACGTAGGGGCAAGCGTTATCCGATTTACTGGGTGAAAGGGAGCGTAGACGGCACGGCAAGCCAGA
TGTGAAAAGCCCCGGGCTCAACCCCGGACTGCATTGGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGATT
CCTAGTGTAGCGGTGAAATGCTGAGATATTAGGAGGAACACCAGTGGCGAAGGGCGCTTGTGGACGATGACTGACG
TTGAGGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGCCGAAACGATGACTGCTAGGTG
TCGGGGCAAGGCCATTCCGGTCCCGCAGCTAACGCAATAAGCAGTCCACCTGGGAGTACGTTCGAAGAATGAAA
CTCAAAGGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCCAAGAACCTTACCT
GATCTTGACATCCGATGCCGCTTGTGAAATGGAAGTTTCTCGGAACATCGGTGACAGGTGGTGCATGGTTGTC
GTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCAGGTTAAG
CTGGGCACTCGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTTATGA
CCAGGGCTACACACGTGCTACAATGGCGAAACAAAGAGAACGCAACTCGCGAGGGTAAGCAAATCTCAAAAATAAC
GTCTCAGTTGGATTGAGATGTTCTGCAACTCGACTACATGAAGCTGGATCGCTAGTAATCGCAAGTCAGAATGCTGCG
```

-continued

GTGAATACGTTCCGGGTCTTGATACACACCAGGCCGTCACACCATGGAGTCATAACGCCGAAGTCAGTGACCCAAC  
CGTAAGG

SEQ ID NO: 98

|NR\_029164.1| *Clostridium innocuum* strain B-3 16S ribosomal RNA gene,  
partial sequence  
ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCTAATACATGCAAGTCGAAGAACAGTCTTCAGGA  
AGCTTGCTTCAAAAAGACTTACTGGCGAACGGGTGAGTAACACGTAGGTAAACCTGCCATGTGTCGGATAACTG  
CTGGAAACGGTAGCTAAACCGATAGGTATACGGAGCGCATGGCTCTGTATATTAAAGGCCCTCAAGCGTGAAC  
ATGGATGGACCTGCGACGCATTAGCTAGTTGGTGAGGTAAACGCCAACCAAGGCGATGATGCCGTAGCCGGCTGAGA  
GGGTAAACGCCACATTGGACTGAGACACGCCAAACTCCTACGGGAGGCAGCAGTAGGGAAATTTCGTCAATGG  
GGGAAACCTGAACGAGCAATGCCCGTGAGTGAAGAAGGTCTCGGATCGTAAAGCTCTGTGTAAGTGAAGAACG  
GCTCATAGAGGAAATGCTATGGGAGTGACGGTAGTACAGAAAGCCACGGCTAACTACGTGCCAGYAGCCGGT  
AATACGTAGGTGGCAAGCGTTATCGGAATCATTGGCGTAAAGGGTGCCTAGGTGGCGTACTAAGTCTGTAGTAAA  
AGGCAATGGCTAACCAATTGTAAGTATGGAAACTGGTATGCTGGAGTGCAGAAAGAGGGCGATGGAATTCCATGTGT  
AGCGTAAATGGCTAGATATGGAGGAACACCAGTGGCGAACGGCGTGCCTGGTCTGTAACTGACACTGAGGCA  
CGAAAGCGTGGGAGCAAATAGGATTAGATACCCCTAGTAGTCCACGCCGTAACAGATGAGAAACTAAGTGTGGAGAA  
ATTCAGTGCTGCAGTTAACGCAATAAGTTCTCCGCTGGGAGTATGCACGCAAGTTNGAAACTCAAAGGAATTGAC  
GGGGGCCGACAAGCGNTGGAGTATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCGCCCTTGACATGGAA  
ACAAATACCCCTAGAGATAGGGGATAATTATGGATCACACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGA  
TGTGGGTTAACGCGCAACGAGCGAACCCCTTGCGCATGTTACAGCATCAAGTGGGACTCATGCGAGACTG  
CCGGTGACAAACGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGGCTGGCTACACAGTACTAC  
AATGGCGACCACAAAGCAGCGACTTGGTACAAGAAGCGAATCTCATAAAGATCGTCTAGTCGGATTGAAGTC  
TGCAACTCGACTTCATGAAGTCGAATCGTAGTAATCGCAGATCAGCATGCTGCCGTGAAATACGTTCTGGCCTT  
GTACACACGCCGTCAAACCATGGGAGTCAGTAATACCGAACGCCGGTGGCATAACCGTAAGGAGTGACCGCTCGA  
AGGTAGGACCGA

SEQ ID NO: 99

|NR\_104687.1| *Blautia hansenii* strain JCM 14655 16S ribosomal RNA gene,  
partial sequence  
AGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCACTTATCATTGA  
CTCTCGGAAGATTGATATTGACTGAGCGGGGACGGGTGAGTAACCGTGGGTAACCTGCCATACAGGGAA  
TAACAGTTAGAAATGGCTGCTATGCCGATAAGCGCACAGGACCGCATGGTCTGGTGTAAAAACTGAGGTGGTAT  
GAGATGGACCCCGCTGATTAGGTAGTTGGGGTAACGCCCTACCAAGCCGACGATCAGTAGCCGGCTGAGAG  
GGTGAACGGCCACATTGGACTGAGACACGCCAGACTCCTACGGAGGCAGCAGTAGGGAAATTGACAATGG  
GGAAACCTGATGCAGCAGCCCGTGAAGGAAGAAGTATCTCGGTATGAAACTTCTACAGCAGGGAAAGAAAAT  
GACGGTACCTGACTAAGAACCCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGAAAGCGTTATCCG  
GATTTACTGGGTAAAGGGAGCGTAGACGGAAGGCAAGTCGATGTGAAAGGTGGGCTTAACCCAGGACTGC  
ATTGGAAACTGTTTCTAGAGTGCGGAGAGGTAAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGG  
AGGAACACCAGTGGCGAAGGCCGCTACTGGACCGTAACTGACGTTGAGGCTCGAAGCGTGGGAGCAAACAGGAT  
TAGATAACCTGGTAGTCCACGCCGAAACGATGAATAACTAGGTGTCGGGTGCAAGCAGTCCGTGCCGAGCAA  
CGCAATAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCG  
TGGAGCATGTGGTTAATCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCTGCCGTACCGTCTTAACCG  
GAGCTTCCCTGGGACAGGCAAGCAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGGTTAAC  
CCGCAACGAGCGAACCCCTATCCTTAGTAGCCAGCAGTCCGGCTGGGCACTCTAGGGAGACTGCCGGGATAACCC  
GGAGGAAGGCGGGGACGACGTCAAATCATGCCCCTATGATTGGCTACACACGTGCTACAATGGCGTAAACA

-continued

```
AAGGGAAAGCGAACGGTGACGCTTAGCAAATCTCAAAAATAACGTCCAGTCGACTGCAGTCTGCAACTCGACTG
CAGCAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGGCTTGTACACACCGGCC
GTCACACCAGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCTATGGAGGGAGCTGCCGAAGGCCGGACCGAT
AACTGGGGTGAAGTCGTAACAAGGTAAACC
```

SEQ ID NO: 100

|NR\_112933.1| *Bacteroides cellulosilyticus* strain JCM 15632 16S ribosomal RNA gene, partial sequence  

```
AGAGTTTGATCCTGGCTCAGGATGAAACGCTAGCTACAGGCTTAACACATGCAAGTCGAGGGCAGCATGACCTAGCA
ATAGGTTGATGGCGACCGGCGCACGGGTGAGTAACACGTATCCAACCTACCGTTATTCCGGATAGCCTTCGAAA
GAAAGATTAACCGGATAGTATAACGAGAAGGCATCTTTGTTATTAAGAATTCGATAACCGATGGGATGCG
TTCCATTAGTTGTTGGCGGGTAACGCCAACCAAGACATCGATGGATAGGGTTCTGAGAGGAAGGTCCCCACA
TTGGAACGTGAGACACGGTCAAACCTCTACGGGAGGCAGCAGTGAGGAATATTGGTCAATGGACGAGAGTCTGAACC
AGCCAAGTAGCGTAAGGATGACTGCCCTATGGTTGAAACTCTTTATATGGAATAAAAGTGAGCCACGTGTGG
CTTTTGATGTACCATACGAAATAAGGATCGGCTAACCGTGTGAAAGTTGGCTAACCGTAA
ATTGCGATTATTGGGTTAAAGGGAGCGTAGGGGACTATTAAAGTCAGCTGTGAAAGTTGGCTAACCGTAA
ATCACGAAGAACCTCGATTGCGAAGGCAGCTACTGGACTGTAACGTGCTGATGCTCGAAAGTGTGGGTATCAA
CAGGATTAGATAACCGTGGTAGTCCACACAGTAAACGATGAATACTCGTGTGCTGATACAGCAAGCGGCCAAC
GAAAGCATTAAAGTATTCCACCTGGGAGTACGCCGCAACGGTAAACAGGATATTGACGGGGCCCGACAAG
CGGAGGAACATGGTTAATCGATGACCGAGGAAACCTACCGGGCTAAATTGACATCTGAATAATTGGAA
ACAGATTAGCCGAAGGCAGATGTGAAGGTGCTGCATGGTTGTCGTAGCTCGTCCGTGAGGTGTCGGCTTAAGTG
CCATAACGAGCGAACCTTATCTTAGTTAAACAGGTATGCTGAGGACTCTAGAGAGACTGCCGTGTAAGAT
GTGAGGAAGGTGGGATGACGTCGAAATCAGCACGGCCCTACGTGGGCTACACAGTGTACAATGGGGTAC
AGAAGGCAGCTACACAGCGATGTGCTAATCCAAAAGCCTCTCAGTCGGATTGGAGTCTGCAACCCACTC
CATGAAGCTGGATTGCTAGTAATCGCGATCAGCACGGCGGTGAATACGTTCCGGCTTGTACACACCGCC
CGTCAAGCCATGAAAGCCGGGGTACCTGAAGTCCGTAACCGCAAGGAGCGGCCAGGGTAAACTGGAATTGGG
CTAAGTCGTA
```

SEQ ID NO: 101

|NR\_112940.1| *Bacteroides ovatus* strain JCM 5824 16S ribosomal RNA gene, partial sequence  

```
GGCTCAGGATGAAACGCTAGCTACAGGCTTAACACATGCAAGTCGAGGGCAGCATTAGTTGCTGCAAACGTAA
GATGGCGACCGGCGCACGGGTGAGTAACACGTATCCAACCTGGCGATAACTCCGAATAGCCTTCGAAAGAAAGAT
TAATACCGGATAGCATAACGAAATATCGCATGATATTAAAGAATTTCGTTATCGATGGGATGCGTTCCATT
AGTTTGTGGCGGGTAACGCCAACCAAGACTACGATGGATAGGGTTCTGAGAGGAAGGTCCCCACATTGGAAC
TGAGACACGGTCAAACCTCTACGGGAGGCAGCAGTGAGGAATATTGGTCAATGGCGAGAGCCTGAACCAGCCAG
TAGCGTGAAGGATGAAAGGCTCTATGGTCGTAACCTCTTTATGGGAATAAAAGTTCCACGTGTGGAAATTG
TATGTACCATATGAATAAGGATCGGCTAACCTCGTGCAGCAGCGCGGTAAACGGAGGATCCGAGCGTTATCCGG
ATTATTGGTTAAAGGGAGCGTAGGTGATTGTTAAGTCAGTTGTAAGGTTGCGGCTCAACCGTAAAGTGCA
GTTGAAACTGGCAGTCTGAGTACAGTAGAGGTGGCGGAATCGTGGTGTAGCGGTGAAATGCTTAGATATCAGCA
AGAAACTCCGATTGCGAAGGCAGCTCACTAGACTGTTACTGACACTGATGCTGCAAAGTGTGGGTATCAAACAGGATT
AGATAACCTGGTAGTCCACACAGTAAACGATGAATACTCGTGTGCGATATACAGTAAGCGGCCAGCGAAAGCA
TTAAGTATTCCACCTGGGAGTACGCCGCAACGGTAAACAGGAAATTGACGGGGCCCGACAAGCGGAGGA
ACATGTGGTTAATCGATGACCGAGGAAACCTTACCCGGCTTAAATTGCAACAGAATATTGGAAACAGTAT
AGCCGTAAGGCTGTTGTGAAGGTGCTGCATGGTTGTCGTAGCTCGTGCCTGAGGTGTCGGCTTAAGTGCCATAAC
```

-continued

GAGCGCAACCCTTATCTTAGTTACTAACAGGTTATGCTGAGGACTCTAGAGAGACTGCCGTGTAAGATGTGAGGA  
 AGGTGGGATGACGTCAAATCAGCACGGCCCTACGTCCGGGCTACACACGTGTTACAATGGGGGTACAGAAGGC  
 AGCTACCTGGCACAGGATGTAATCCAAAACCTCTCTCAGTTCGGATCGAAGTCTGCAACCCGACTCGTGAAG  
 CTGGATTGCTAGTAATCGCGCATCAGCCATGGCGGGTGAATACTGTTCCCGGGCTTGTACACACCGCCCGTCAAG  
 CCATGAAAGCGGGGGTACCTGAAGTACGTAACCGCAAGGAGCGTCTAGGGTAAAATGGTAATTGGGCTA

SEQ ID NO: 102

| NR\_117563.1 | *Eubacterium fissicatena* 16S ribosomal RNA gene, partial sequence  
 TAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCGCTTACTTAG  
 ATTTCTCGGATTGAAAGAGTTTGCAGTGAACGGCGGACGGGTGAGTAACCGTGGGTAACCTGCCTCATACAGGG  
 GGATAACAGTTAGAAATGACTGTAATACCGCATAAGACCACAGTACCGCATGGTACAGTGGGAAAAACTCCGGTGG  
 TATGAGATGGACCCCGCTCTGATTAGCTAGTGGTAAGGTAACGGTTACCAAGGCAGCAGTCAGTAGCCGACCTGA  
 GAGGGTACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCAAAAT  
 GGGGAAACCTGATGCGAGCGCGTGAAGGATGAAGTATTGGTATGTAACCTCTATCAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAACCCCCGGTAACACGTGCCAGGCCCGGTAATACGTAGGGGCAAGCGTTAT  
 CCGGATTACTGGGTGAAAGGGAGCGTAGACGGTTATGTAAGTGTGATGAAAACCCGGGCTCAACCCGGGAC  
 TGCATTGGAAACTATGTAACTAGAGTGTGGAGAGGTAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCACTGGCGAAGGCGCTTACTGGACGATCACTGACGTTGAGGCTCGAAGACGGTGGGAGCAAACAG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTCGGGTGGCAAAGCATTGGTGCCGAGC  
 AAACGCAATAAGTATTCCACCTGGGAGTACGTCGAAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGTGGTTAATTGAAAGCAACCGAAGAACCTTACCTGCTTGTACATCCCAGTACCGCGTGTAA  
 TGGCGCCTCCCTCGGGCAGTGGAGACAGGTGGTGATGGTGTGCTCAGCTGTGAGATGTTGGGTTAA  
 GTCCCGCAACCGCGAACCTTATCTTAGTAGCCAGCGTTGGCCGGCACTCTAGAGAGACTGCCAGGGATAA  
 CCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGAGCAGGGTACACACGTGCTACAATGGCGTAA  
 ACAAAAGGGAGGCAATACCGCAGGGTGAGCAAATCCAAAATAACGTCAGTGGATTGAGTCTGCAACTCGA  
 CTACATGAAGGCTGAAATCGTAGTAATCGCGAATCAGAAATGTCGCGGTGAAATACGTTCCGGGTCTTGTACACACCG  
 CCCGTACACCATGGGAGTTGGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGCGGGATC  
 GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 103

| NR\_104700.1 | *Blautia coccoides* strain JCM 1395 16S ribosomal RNA gene, partial sequence  
 AGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCGCTAAGACAGAT  
 TTCTCGGATTGAAAGTCTTGTGACTGAGCGGGACGGGTGAGTAACCGTGGGTAACCTGCCTCATACAGGGGA  
 TAACAGTTAGAAATGACTGTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACCTCCGGTGTAT  
 GAGATGGACCCCGCTGATTAGCTAGTGGAGGGTAACGCCAACCAAGGCAGCAGTCAGTAGCCGGCTGAGAG  
 GGTGAACGGCCACATGGGACTGAGACACGCCAACACTCCTACGGGAGGCAGCAGTGGGAAATTGCAACATGGG  
 GGAAACCTGATGCGAGCGACGCCGTGAAGGAAGAAGTATCGGTATGTAACCTCTATCAGCAGGGAAAGAAAT  
 GACGGTACCTGACTAAGAACGCCGCTAACTACGTGCCAGCGCGGTAATACGTAGGGGCAAGCGTTATCCG  
 GATTACTGGGTGAAAGGGAGCGTAGACGGAAGAGCAAGTCGATGTGAAAGGCTGGGCTAACCCAGGACTGC  
 ATTGGAAACTGTGTTCTAGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGG  
 AGGAACACCACTGGCGAAGGCGCTTACTGGACGTTACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAGGAT  
 TAGATAACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTCGGGTGGCAAAGCATTGGTGCCGAGCAA  
 CGCAATAAGTATTCCACCTGGGAGTACGTCGAAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCGG  
 TGGAGCATGTGGTTAATTGAAAGCAACCGAAGAACCTTACCAAGTCTTGACATCCCCTGTACCGTCCGTAACGG

## US 10,555,980 B2

139

140

-continued

GGGCTTCCCTCGGGCAGAGGAGACAGGTGGTCATGGTTGTCAGTCGTGAGATGTTGGTTAAC  
 CGCAACGAGCGCAACCCTATCCTTAGTAGCCAGCACATGATGGTGGGCACTCTAGGGAGACTGCCGGGATAACC  
 CGGAGGAAGGGGGACGACGTCAAATCATCATGCCCTTATGATTGGCTACACACGTCTACAATGGCTAAAC  
 AAAGGAAAGCGAGACAGCGATGTTGAGCGAATCCAAAAAATACGCCCAGTCGGACTCGCAGTCGAACTCGACT  
 GCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGGTGAATACGTTCCGGTCTTGTACACACCGCC  
 CGTCACACCATGGGAGTCAGTAACGCCGAAAGTCAGTGACCTAACCGAAAGGAAGGAGCTGCCGAAGGCAGGACCGA  
 TAACTGGGGTGAAGTCGTAACAAGGTAACC

SEQ ID NO: 104

| NR\_109014.1 | *Blautia faecis* strain M25 16S ribosomal RNA gene, partial  
 sequence  
 ATAACAGCCAGAAAATGACTGCTAAACCGATAAGCGCACAGAACCGATGGTCGGTGTAAAAACTCCGGTGGTA  
 TAAGATGGACCCCGCTGGATTAGCTAGTTGGCAGGGCAGCGGCCTACCAAGGCAGCAGTCGGCATAGCCGGCTGAGA  
 GGGTGAACGCCACATTGGGACTGAGACACGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGCACAATGG  
 GGAAACCCCTATGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAACCTCTATCAGCAGGGAAAGATAA  
 TGACGGTACCTGACTAAGAACCCCCGCTAACTACGTGCCAGGCCGGTAAACGTAGGGGCAAGCGTTATCC  
 GGATTACTGGGTGAAAGGGAGCGTAGACGGCGCAGCAAGTCTGATGTGAAAGGCAGGGCTTAACCCCTGGACTG  
 CATTGGAAACTGCTGTGCTTGAGTGCGGAGGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAG  
 GAGGAACACCACTGGCGAACGGGGCTACTGGACGTTAAGCAGTCTGAGGCTGAAAGCGTGGGAGCAAACAGGA  
 TTAGATAACCTGGTAGTCCACGCCGAAACGATGAAACTAGGTGTCAGGGAGCACAGCTTTGGTGCCGCCCAA  
 ACGCATTAAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCG  
 GTGGAGCATGTGGTTAACGCAACCGAACGGACTTACCAATCTTGACATCCCTGACCGGACTTAACC  
 GTCCCTTCTCGGGACAGGGAGACAGGTGGCATGGTGTGCTAGCTCGTGTGAGATGTTGGTTAAAGT  
 CCCGCAACGAGCGCAACCCCTACCTAGTAGGCCAGCAGCAGTCAGTGGGACTCTGAGGAGACTGCCAGGGATAAC  
 CTGGAGGAAGGGGGATGACGCTAAACATCATGCCCTATGATTGGCTACACAGTGTACAATGGCGTAA  
 CAAAGGGAAGCGAACCGCGAGGGGGAAATCTAAAAATAACGTCGGACTGCACTGCAACTCGAC  
 TGCACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGGTGAATACGTTCCGGTCTTGTACACACCGC  
 CCGTCACACCATGGGAGTCAGTAACGCCG

SEQ ID NO: 105

| NR\_036928.1 | *Clostridium hathewayi* strain 1313 16S ribosomal RNA gene,  
 partial sequence  
 CTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAACGGTTCAATGAAAGTTTCGGATGGATT  
 TGAAATTGACTTAGCGGGGAGGGTGAGTAAACGCGTGGTAACCTGCTTACACTGGGGATAACAGTTAGAAATG  
 ACTGCTAATACCGATAAGCGCACAGGCCGATGGNCTGGTGTAAAAACTCCGGNGGTGTAAGATGGACCCCGT  
 CTGATTAGGTAGTGGGGGGTAAACGGCCACCAAGCCGACGATCAGTAGCCGACCTGAGAGGGTACCGGCCACAT  
 TGGGACTGAGACACGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGGACAATGGCGAACGGCTGATCCA  
 GCACGCCGCGTGAAGAAGTATTCGGTATGTAAGCTCTACAGCAGGGAAAGAAATGACGGTACCTGACTA  
 AGAAGCCCCGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTACCGGATTACTGGGTGA  
 AAGGGAGCGTAGACGGTTAGCGAAGTGAAGTGAAGAAGCCGGCTAACCCCGTACTGCTTGGAAACTGTTAG  
 ACTTGAGTGCAGGAGAGGTAAGTGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACAGTGGC  
 GAAGGGCGCTACTGGACTGTAACGTGAGCGTGGAGGCTGAAAGCGTGGGGAGCAAACAGGATTAGATACCGTGT  
 TCCACGCCGTAACGATGAAACTAGGTGTCGGGGCAAAGCCCTCGGTGCCGCAAACGCAATAAGTATTCC  
 ACCTGGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCGTGGAGCATGTT  
 AATTGCAAGCAACGCGAAGAACCTTACCAAGTCTGACATCCCACTGAAAACACNTTAACCGTGTACCCCTTCGGA  
 GCAGTGGAGACAGGTGGTCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCAA

-continued

```
CCCTTATCCTTAGTAGCCAGCGAGTAGAGTCGGGCACTCTGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGG
ATGACGTCAAATCATCATGCCCTTATGATTTGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAGGCAAAGG
AGCGATCTGGAGCAAACCCCCAAAATAACGCTCTCAGTTCGGATTGCAGGCTGCAACTCGCTGCATGAAGCTGGAA
CGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCGGTCTTGTACACACCGCCCGTCACACCAGGGAA
GTTGGTAACGCCGAAGTCAGTGACCCAACCGAAAGGAGGGAGCT
```

SEQ ID NO: 106

|NR\_113270.1| *Blautia producta* strain JCM 1471 16S ribosomal RNA gene,  
partial sequence  
AGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCAGTAAGACGGAT  
TTCTTCGATTGAAAGCTTTGTACTGAGCGGGGACGGGTGAGTAACCGTGGGTAACCTGCCCTACAGGGGG  
TAACAGTTAGAAATGACTGCTAATACCGATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAAACTCCGGTGGTAT  
GAGATGGACCCCGCTCTGATTAGCTAGTTGGGGGTAACGGCCCACCAAGGCACGATCAGTAGCCGGCTGAGAG  
GGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCACGAGTGGGAATTGCACAATGGG  
GGAAACCCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGAAACTTCTATCAGCAGGGAAAGAAAAT  
GACGGTACCTGACTAAGAAGCCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTATCCG  
GATTTACTGGGTGAAAGGGAGCGTAGACGGAAGGCAAGTCGATGTGAAAGGCTGGGCTTAACCCAGGACTGC  
ATTGGAAACTGTTGTTCTAGAGTGCGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGG  
AGGAACACCAGTGGCGAAGGGCGCTACTGGACCGTAACTGACGTGAGGCTGAAAGCGTGGGAGCAAACAGGAT  
TAGATAACCTGGTAGTCCACGCCGTAACGATGAATACTAGGTGTGGTGGCAAAGCCATTGGTGCAGCAGCAA  
CGCAATAAGTATTCCACCTGGGAGTAGCTTGCAGAAATGACGGGAGCCGCAACAGCG  
TGGAGCATGTGGTTAATCGAAGCAACGCAAGAACCTTACCAAGTCTTGACATCCCTGTACCGTCCCGTAACGG  
GGACTTCCCTCGGGGAGAGGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGTTAAGTC  
CCGCAACGAGGCCAACCTTACCTTAGTAGCAGCACATGATGGTGGGACTCTAGGGAGACTGCCGGGATAACC  
CGGAGGAAGGGGGGAGACGCTCAAATCATCATGCCCTTATGATTTGGCTACACACGTGCTACAATGGCGTAAAC  
AAAGGGAGCGAGACAGCGATGTTGAGCGAATCCCCAAAATAACGTCCCAGTCGGACTGCAGTCTGCAACTCGACT  
GCACAGAGCTGGAATCGCTAGTAATCGCGATCAGAATGCCCGGTGAATACGTTCCGGTCTTGTACACACCGCC  
CGTCACACCAGGGAGTCAGTAACGCCGAAGTCAGTGACCTAACGAAAGGAAGGAGCTGCCAGGGAGCG  
TAACTGGGGTGAAGTCGTAACAAGGTAACC

SEQ ID NO: 107

|NR\_104799.1| *Anaerostipes hadrus* strain DSM 3319 16S ribosomal RNA  
gene, partial sequence  
TGGCTCAGGATAACGCTGGCGCGTCTAACACATGCAAGTCGAAAGCGAGCTGCTTAACCTGATCTTCTCGGAAT  
TGACGTTTGTAGACTGAGTGGCGACGGGTGAGTAACCGTGGCAACCTGCCCTGTACAGGGGATAACAGTCAG  
AAATGACTGCTAATACCGATAAGACCACAGCACCGCATGGTGCAGGGTAAAAACTCGGTGGTACAGGATGGACC  
CGCGTCTGATTAGCTGGTGGTGGAGGTAACGGCTACCAAGGCACGATCAGTAGCCGGTTGAGAGAGTGAACGGC  
CACATTGGACTGAGACACGGCCCAAACCTCTACGGGAGGCAGCAGTGGGAATTGCAACATGGGGAAACCCCTG  
ATGCAGCGACGCCCGTGAAGTATCTGGTATGTAAGCTCTACAGCAGGGAAAGAAAATGACGGTACCT  
GACTAAGAAGCCCCGGCTAACTACGTGCCAGCCGGTAATACGTAGGGGCAAGCGTTATCGGAATTACTGG  
GTGTAAAGGGTGCCTAGGTGGTATGGCAAGTCAGAAGTGAACCGGGCTTAACCTGGGACTGCTTTGAAACT  
GTCAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGGAGGAACATCA  
GTGGCGAAGGGGGCTACTGGACTGAAACTGACACTGAGGCACGAAAGCGTGGGAGCAAACAGGATTAGATAACCT  
GGTAGTCCACGCCGTAACGATGAATACTAGGTGTGGGCGCTAGAGGCTTGGTGCAGCAGCGAAACGCAAGTAAAGT  
ATTCCACCTGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGAGCCGACAAGCGGTGGAGCATGT  
GGTTAATTGCAAGCAACGCGAAGAACCTTACCTGGTCTTGTACATCCTCTGACCGGTCTTAACCGGACCTTCCT

-continued

TCGGGACAGGAGAGACAGGTGGTGCATGGTTGTCGTCAGTCGTCGTGAGATGTTGGGTAAGTCCCACACGAG  
 CGCAACCCCTATCTTAGTAGCAGCATTCAAGTGGCACTCTAGAGAGACTGCGCAGGGATAACCTGGAGGAAGGT  
 GGGGACGACGTCAAATCATCATGCCCTATGACCAGGGTACACACGTCTACAGTGGTAAACAGAGGGAAAGCA  
 GCCTCGTGGAGAGTGAGCAAATCCAAAAAATACGTCTCAGTCGATTGAGTCTGCACACTCGACTACATGAAGCTG  
 GAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTTGTACACACCGCCGTACACCAT  
 GGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGCCGGACCGATAACTGGGTG  
 AAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTTC

SEQ ID NO: 108

|NR\_117142.1| *Eubacterium fissicatena* strain DSM 3598 16S ribosomal RNA  
 gene, partial sequence  
 GTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGTAAACACATGCAAGTCGAGCGAAGCGCTTACTTAGATT  
 CTCGGATTGAAAGAGTTTGCAGTGGCGACGGTGGAGTAACCGTGGTACAGTGGGAAACTCCGGTGGTATG  
 AACAGTTAGAAATGACTGCTAAACCGATAAGACCACAGTACCGCATGGTACAGTGGGAAACTCCGGTGGTATG  
 AGATGGACCCCGCTGTGATTAGCTAGTGGTAAGGTAACGGCTTACCAAGGCAACGATCAGTAGCCGACCTGAGGG  
 GTGACCGGCCACATTGGGACTGAGAACCGGCCAAACTCTACGGGAGGCAGCAGTGGGAAATTGCAAAATGGG  
 GAAAACCTGTGACGCGACGCCGCTGAAGGATGAAGTATTCGGTATGTAACCTCTATCAGCAGGGAAAGAAAATG  
 ACGGTACCTGACTAAGAACGCCCGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTTATCGG  
 ATTTACTGGGTGAAAGGGAGCGTAGACGGTTATGTAAGTCTGATGTGAAACCCGGGCTCAACCCGGGACTGCA  
 TTGGAAACTATGTAACTAGAGTGTGGAGAGGTAAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGA  
 GGAACACCACTGGCGAGGCGCTTACTGGACGATCACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGG  
 AGATAACCTGGTAGTCCACGCCGAAACGATGAAATACTAGGTGTGGGTGGCAAAGCCATTGGTGCAGCAAAC  
 GCAATAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAAATTGACGGGACCCGACAAGCGGT  
 GGAGCATGTGGTTAATTGAAAGCAACGCGAAGAACCTTACCTGCTCTTGACATCCACTGACCGCGTGTAAATGGC  
 GCCTCCCTCGGGGAGTGGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAGTCC  
 CGCAACGAGCGCAACCCCTATCTTAGTAGCCAGCGTTGGCCGGGACTCTAGAGAGACTGCCAGGGATAACCTG  
 GAGGAAGGTGGGAGTACGTCAGTAAACGCAATCAGTGGGAGGCTGCGGAAGGCCGGATCGATA  
 ATGAAGCTGGAAATCGCTAGTAATCGCAATCAGAATGTCGCGGTAAACGTTCCGGTCTTGTACACACCGCCG  
 TCACACCAGGGAGTTGGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGCCGGATCGATA  
 ACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTTC

SEQ ID NO: 109

|NR\_117147.1| *Eubacterium contortum* strain DSM 3982 16S ribosomal RNA  
 gene, partial sequence  
 TTGATCCTGGCTCAGGATGAACGCTGGCGACGTGTGTAAACACATGCAAGTCGAGCGAAGCAGCTTACTTTGATT  
 TCGGAATGAAAGGTTTGTGACTGAGCGGGGAGGGTGAGTAACCGTGGTAAACCTGCCCTACAGGGGATA  
 ACAGTTAGAAATGACTGCTAAACCGATAAGACCACAGTACCGCATGGTACAGTGGGAAACTCCGGTGGTATGA  
 GATGGACCCCGCTGTGATTAGCTAGTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGCCGACCTGAGAGGG  
 TGACCGGCCACATTGGGACTGAGAACCGGCCAAACTCTACGGGAGGCAGCAGTGGGAAATTGCAAAATGGGG  
 AAACCCGTGACGCGACGCCGCTGAAGGATGAAGTATTCGGTATGTAACCTCTATCAGCAGGGAAAGAAAATGA  
 CGGTACCTGACTAAGAACGCCCGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTTATCCGG  
 TTTACTGGGTGAAAGGGAGCGTAGACGGTTATGTAAGTCTGATGTGAAACCCGGGCTCAACCCGGGACTGCAT  
 TGGAAACTATGTAACTAGAGTGTGGAGAGGTAAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAG  
 GAACACCACTGGCGAAGGCCGGCTTACTGGACGATGACTGACGTTGAGGCTGAAAGCGTGGGGAGCAAACAGGATTA  
 GATACCCCTGGTAGTCCACGCCGAAACGATGAAATACTAGGTGTGGGTGGCAAAGCCATTGGTGCAGCAAAC

-continued

```

CAATAAGTATTCCACCTGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTG
GAGCATGTGGTTAACCGAAGCACCGAAGAACCTTACCTGCTCTTGACATCCCCCTGACCGCGTGTAAATGGTG
CTTTCTTCGGGACAGGGAGACAGGTGGTGCATGGTTGCGTCAGCTCGTGTGAGATGTTGGGTAAGTCCC
GCAACAGCGCAACCCTATCTTAGTAGCCAGCGTTGGCCGGCACTCTAGAGAGACTGCCAGGGATAACCTGG
AGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGAGCAGGGCTACACACGTGCTACAATGGCGTAAACAAA
GGGAGCGAAGCGTGGAGGCAAATCCAAAAAAACGCTCAGTTGGATTGAGTCGAACTCGACTACA
TGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAAATACGTTCCGGTCTTGTACACACCGCCGT
CACACCATGGGAGTTGGTAACGCCGAAGTCAGTGACCCAACCGCAAGGAGGGAGCTGCCAGGGTGGGACCGATAA
CTGGGTGAAGTCGTAACAAGGTAGCGTATCGGAAGGTGCGGCTGGATCACCTCCTTCT

```

SEQ ID NO: 110

```

|NR_113410.1|Clostridium bolteae strain JCM 12243 16S ribosomal RNA
gene, partial sequence
TTTAATTGACTGAGTGGCGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGGGGATAACAGTTAGAAAT
GACTGCTAATACCGATAAGCGCACAGTACCGCATGGTACAGTGTGAAAAACTCCGGTGGTGTGAGATGGATCCCG
TCGATTAGCCAGTTGGCGGGTAACGCCAACCAAGCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACA
TTGGGACTGAGACACGCCAACCTCTACGGGAGGCAGCAGTGGGAATATTGACAATGGCGAAAGCCTGATGC
AGCGACGCCCGGTGAGTGAAGAAGTATTCGGTATGTAAGCCTATCAGCAGGGAAAGAAATGACGGTACCTGACT
AAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAAACGTAGGGGCAAGCGTTATCCGATTTACTGGGT
AAAGGGAGCGTAGACGGCGAACGAAAGTCTGAAGTGAAAACCCAGGGCTAACCTGGACTGCTTGGAAACTGTT
TGCTAGAGTGTGGAGGGTAAGTGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGGAGGAACACCAGTGG
CGAAGGGCGCTACTGGACGATAACTGACGTTGAGGCTCGAACAGCGTGGGAGCAAACAGGATTAGATAACCTGGTA
GTCCACCGCTAACGATGAATGCTAGGTGTTGGGGCAAAGCCCTCGGTGCGTCGAAACCGAGTAAGCATTG
CACCTGGGACTACGTCGAAGAATGAAACTCAAAGGATTGACGGGACCCGACAGCGGTGGAGCATGTGTT
TAATTGCAAGCAACCGAACGAAACCTTACCAAGTCTGACATCCTTGTGACCGCGTGTAAACGGCCTTCCCTGG
GGCAAGAGAGACAGGTGGTGCATGGTGTGTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCGCAACGAGCGA
ACCCTTACCTTAGGCCAGCAGGTAAGCTGGGACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGTGGG
GATGACGTCAAATCATGCCCTATGATTGGCTACACAGTGTACAATGGCGTAAACAAAGGAAGCAAGA
CAGTGTGTGGAGCAATCCAAAAAAACGCTCCAGTTCGGACTGTAGTCGCAACCCGACTACACGAAGCTGGAA
TCGCTAGTAATCGCAATCAGAATGTCGCGGTGAAATACGTTCCGGTCTTGTACACACCGCCGTACACCATGG
AGTCAGCAACGCCGAAGTCAGTGACCCAACCGCAAGAGAGGGAGCTGCCGAAGGCGGGCAGGTAACGGGTGA
AGTC

```

SEQ ID NO: 111

```

|NR_041960.1|Blautia luti strain BInIX 16S ribosomal RNA gene, complete
sequence
GTGGGTAAACCTGCCTTATACAGGGGATAACAGTCAGAAATGACTGCTAATACCGATAAGCGCACAGAGCTGCATG
GCTCCGGTGTGAAAAACTCCGGTGTATAAGATGGACCCGCGTTGGATTAGCTAGTTGGTGGAGGTAAACGCCACCA
AGGCAGCATCCATGCCGGCTGAGAGGGTGAACGGCACATTGGACTGAGACACGGCCAGACTCCTACGGAG
GCAGCAGTGGGAATATTGCAACATGGGGAAACCCGATGCGACGCCGCGTGAAGGAAGAAGTATCTCGGT
GTAAACTCTATCAGCAGGGAAAGAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCG
GTAATACGTAGGGGCAAGCGTTATCGGATTACTGGGTGTAAGGGAGCGTAGACGGCATGGACAAGTCTGATGT
GAAAGGCTGGGCTAACCCGGACTGCATTGAAACTGCCGTCTGAGTGCAGGAGGGTAAGCGGAATTCTA
GTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCGAGTGGCGAAGGCGGTTACTGGACGGTAACGACGTTGA
GGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTCACCGGTAAACGATGAATCCTAGGTGTGG
GGAGCAAANNNTTCGGTGCCCGCAAACGCATTAAGCATCCACCTGGGAGTACGTTCGCAAGAATGAAACTCA

```

-continued

```
AAGGAATTGACGGGGACCGACAAGCGGTGGAGCATGTGGTTAACCGAAGCAACCGAAGAACCTTACCAAGTC
TTGACATCCCTGACCGAGTATGTATGGTACTTCTCGGGAGAGAGAGAGACAGGTGGCATGGTTGCGT
CAGCTCGTGTGTGGAGATGTGGTTAAGTCCCGAACCGCGAACCCCTATCCCAGTAGCCAGCGGTCGGCG
GCCACTCTGAGGAGACTGCCAGGGATAACCTGGAGGAAGGGGGATGACGTCAAATCATCATGCCCTTATGATT
GGGCTACACACGTGCTACAATGGCTAAACAAAGGAAGCAAGCCTGCGAGGGTGGCAAATCCAAAAATAACGTC
CCAGTTCGGACTGTAGTCTGCACCCGACTACACGAAGCTGAAATCGCTAGTAATCGGGATCAGAACGCGCGTG
AATACGTTCCGGGTCTGTACACACCGCCGTCACACCAGGGAGTCAGTAACGCCGAAGTCAGTGACCTAACT
```

SEQ ID NO: 112

```
|NR_074306.1| Acidaminococcus intestini RyC-MR95 strain RyC-MR95 16S
ribosomal RNA, complete sequence
CTGGCGCGTCTAACATGCAAGTCGAACGGAGAACTTATTCGTAAGTCTTAGTGGCAACGGGTGAGTAA
CGCGTGGGCAACCTGCCCTCAGTTGGGACAACATTCCGAAAGGGATGCTAAATCCGAATGTCCTCCCTCCGC
ATGGAGGAGGGAGGAAAGATGGCCTCTGCTTGAAGCTATCGCTGGAAGATGGGCCGCGTCTGATTAGCTAGTTGG
TGGGTAACCGCTCACCAAGGGCATGATCAGTAGCGGTCTGAGGGATGACGCCACATTGGACTGAGACACGG
CCAAACTCCTACGGGAGGCAGCAGTGGGAATCTCCGAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGA
GATGAAGGTCTCGGATTGTAAGGACTCTGTTAGGGACGAAAGCACCGTGTGCAACAGGTATGGTGTGACGG
TACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGTGGCAAGCGTTGCGGAATT
ATTGGCGTAAAGAGCATGTAAGCGGGCTTTAACGTTAACGCTGAAATGCGGGCTTAACCCGTATGGCGTTGGA
TACTGGAAGTCTTGAGTCAGGAGAGGAAAGGGGAATTCCAGTGTAGCGGTGAAATGCGTAGATATTGGGAGGAAC
ACCACTGGCGAAGGCCTGGGACTACGATGCAAGATTCAAAGGAATTGACGGGGCCGACAAGCGGTGGAG
CCCCGGTAGTCCTGGCGTAAACGATGGATACTAGGTGTAGGAGGTATGACCCCTCTGTGCCGGAGTTAACGAA
TAAGTATCCCGCTGGGACTACGATGCAAGATTCAAAGGAATTGACGGGGCCGACAAGCGGTGGAG
TATGTGGTTAACGCAACCGAAGAACCTTACCAAGGCTTACGATTGAGTGGAAAGACCTAGAGATAGGTCCC
TCCCTCGGGGACACGAAAACAGGTGGCATGGCTGTCAGCTCGTGTGAGATGTTGGGTTAACCGCA
ACGAGCGCAACCCATCCTATGTTACCAAGCGCTAACGGGGACTCATAGGAGACTGCCAGGATAACTGGAG
GAAGGGGGGATGACGTCAAGTCATCATGCCCTATGTTGGGCTACACAGTACTAACATGGCGAACAAAG
GGCAGCGAAACCGCGAGGTGGAGCAAATCCAGAAACCCGACCCAGTCCGATCGTAGGCTGCAACCCGCTACGT
GAAGTTGGAAATGCTAGTAATCGCAGGTCACTCGGGTAACAGTCCGGGCTTGTACACACCGCCGTC
ACACCAAGGAAAGTTGTTAACACCGAAGCCGGTGAGATAACCTTTAGGAGTCAGCTGCTAACGGGGCGATGA
TTGGGTGAAGTCGTAACAAGGTAGC
```

SEQ ID NO: 113

```
|NR_074399.1| Ruminococcus albus strain 7 16S ribosomal RNA gene,
complete sequence
AGAGTTTGATCTGGCTCAGGACGAACGCTGGCGCACGCTAACACATGCAAGTCGAACGAGCGAAAGAGTGTG
CACTCTCTAGCTAGTGGCGGAGGGTGAGTAACACGTGAGCAATGCCCTTCCGAGAGGGATACCAATTGGAAACG
ATTGTTAACCTCATAACATAACGAGCCGATGACTTTGTTACAAATGAATTGCGCAAAGATGAGCTCGCGT
CTGATTAGGTAGTTGGTGAGGTAACGGCCACCAAGCCGACGATCAGTAGCCGACTGAGAGGTTGAACGGCACAT
TGGGACTGAGACACGGCCGAGACTCCTACGGGAGGAGCAGTGGGAATTGCAACATGGCGAAAGCCTGATGCA
GCGATGCCGCGTGGAGGAAGAGTTAGGATTGTAACCTCTGTCTTGGGACGATAATGCGGTACCCAAGGA
GGAAGCTCCGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGAGCGAGCGTTGCGGAATTACTGGGTGTA
AAGGGAGCGTAGGGGGATTGCAAGTCAGGTGTGAAATTAGGGCTTAACCCCTGAACTGCACTTGAAACTGTAGT
TCTTGAGTGAAGTAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGC
GAAGGGGGCTTACTGGCTTAACGCGTGAGGCTGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAG
TCCACGCCGTAAACGATGATTACTAGGTGTGGGGGACTGACCCCTCCGTGCCGAGTTAACACAATAAGTAATCC
```

## US 10,555,980 B2

149

150

-continued

ACCTGGGGAGTACGGCCGAAGGCTGAAACTCAAAGGAATTGACGGGACCCGCACAAGCAGTGGAGTATGTGGTT  
 AATTCAAGCAACGCGAAGAACCTTACCAAGGCTTGACATCGCATAGCATAGAGATATGTGAAATCCCTCGG  
 CGACGTATAGACAGGTGGTCATGGTGTCTCAGCTCGTCTGAGATGTTGGGTTAAGTCCCGAACGAGCG  
 AACCTTACTGTAGTTAGTGTCTACGCAAGAGCACTCTAGCAGGACTGCCGTTGACAAAACGGAGGAAGGTGGGATGAC  
 GTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTACTACAATGGCTTAAACAGAGGGAAAGCAAAACAGTG  
 ATGTGGAGCAAACCTAAAGCAGTCTTAGTCGGATTGAGGTGCAACCCGCTACATGAAGTCGAATTGCTA  
 GTAATCGCGGATCAGCATGCCGGTGAATACGTTCCCGGGCTGTACACACCGCCGTACGCCATGGAGTCGG  
 TAACACCGAAGGCTGTGTTCTAACCGCAAGGAGGAAGCAGTCGAAGGTGGGATTGATGACTGGGTGAAGTCGTAA  
 CAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT

SEQ ID NO: 114

|NR\_074634.1| *Eubacterium rectale* strain ATCC 33656 16S ribosomal RNA  
 gene, complete sequence  
 AGAGTTGATCCTGGCTCAGGATGAAACGCTGGCGCGTGCTTAACACATGCAAGTCGAACGAAGCACTTTATTGAT  
 TTCTCGGGACTGATTATTTGTACTGAGTGGCGGACGGGTGAGTAACGCGTGGTAACTGCCCTGTACAGGG  
 GATAACAGTTGAAACGGCTGCTAACACCGATAAGCGCACGGCATCGCATGATGCGAGTGTGAAAAACTCCGGTGG  
 ATAAGATGGACCCGCGTGGATTAGCTAGTGGTGGAGTAACGGCCACCAAGGGCAGCATCCATAGCCGACCTGAG  
 AGGGTGACCGGCCACATTGGACTGAGACACGGCCAAACTCTACGGGAGGCAGCAGTGGGAATATTGACAATG  
 GGCAGAACGCGTGTGAGCAGCGACGCCGCGTGGAGCAAGAAGTATTCGGTATGTAAGCTCTACAGCAGGGAAAGATA  
 ATGACGGTACCTGACTAACAGAACCGCTAACACCGCTAACACCGCTAACACCGCTAACACCGCTAACACCGCTAAC  
 CCGATTTACTGGGTGAAAGGGAGCGCAGGCCGCGTGGCAAGTGTGAGATGTTGGGACTTCCGGTGGGACTTCC  
 GCATTGGAAACTGCGTACTAGAGTGTGGAGGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTA  
 GGAGGAACACCGTGGCGAAGCGCCGTTACTGGACGATAACTGACGCTGAGGCTGAAAGCGTGGGAGCAACAGG  
 ATTAGATAACCTGGTAGTCCACGCCAACGATGAAATACTAGGTGTTGGAGCATTGCTTCGGTGGCGTGC  
 AACGCAGTAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGC  
 GGTGGAGCATGTGGTTAACCGCAACCGCTAACACCGCTAACACCGCTAACACCGCTAACACCGCTAACACCG  
 CTGGACGCTCTCGGAGCAGGAGTGCAGGGTGCATGGTGTGCTCAGCTCGTGTGAGATGTTGGGTTAAG  
 TCCCGAACGAGCGAACCCCTATCTTAGTGGCAGCGGTTGGCCGGCACTCTAGAGAGACTGCCAGGGATAAC  
 CTGGAGGAAGGGGGGGATGACGCTAACACCGCTAACACCGCTAACACCGCTAACACCGCTAACACCG  
 CAAAGGGAAAGCTGTGAAGCGAGCAAATCTCAAAATAACGCTCAGTTGGACTGTAGTCTGCAACCCGAC  
 TACACGAAGCTGGAATCGCTAGTAATCGCAGATCAGAATGCTCGGTGAATACGTTCCGGTCTGTACACACCG  
 CGTCACACCAGGGAGTTGGGAATGCCGAAGCCAGTGACCTAACCGAAAGGAAGGAGCTGTCGAAGGCAGGCTCG  
 ATAACCTGGGTGAAGTCGTAACAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT

SEQ ID NO: 115

|NR\_074928.1| *Acidaminococcus fermentans* strain DSM 20731 16S  
 ribosomal RNA gene, complete sequence  
 AGAGTTGATCCTGGCTCAGGACGAAACGCTGGCGCGTGCTTAACACATGCAAGTCGAACGGAGAACTTTCTCGA  
 ATGTTCTTAGTGGCGAACGGGTGAGTAACGCGTAGGCAACCTGCCCTCTGGTGGGACAACATTGAAAGGGATG  
 CTAATACCGAATGAGATCCTCTTCCGATGGAGAGAGGTGAAAGATGGCTCTACTGTGAGCTATGCCAGAAG  
 ATGGGCTGCGTCTGATTAGCTAGTAGGTGAGGTGAGGCTCACCTAGGCATGATCAGTAGCCGGTCTGAGAGGAT  
 GAACGGCACATGGACTGAGACACGGCCAAACTCTACGGGAGGCAGCAGTGGGAATCTCCGCAATGGACGA  
 AAGTCTGACGGAGCAACGCCGCGTGAAGTGAAGGCCTTCGGGTTGAAAACCTGTGTCAGGGACGAAAGCACC  
 GATCTATAATACATTGGTGTGACGGTACCTGACGAGGAAGCCACGGCTAACACCGCTAACACCGCTAACACCG  
 ACGTAGGTGGCAAGCGTTGTCGGAATTATGGCGTAAAGAGCATGTAGGCGGGCTTTAAGTCCGACGTGAAAAT  
 CGGGGCTTAACCCGTATGGCTGGATACTGGAAAGTCTGAGTGCAGGAGAGGAAGGGAAATTCCCAGTGTAGC

-continued

GGTGAATGCGTAGATATTGGGAGGAACACCAGTGGCGAAGGCCTTCTGGACTGTGCTGACGCTGAGATGCA  
 AAGCCAGGGTAGCAAACGGGATTAGATAACCCGGTAGTCCTGGCGTAAACGATGGGTACTAGGTGTAGGAGGTATC  
 GACCCCTCTGTGCCGGAGTTAACGCAATAAGTACCCGGCTGGGACTACGATCGCAAGATTGAAACTCAAAGGA  
 TTGACGGGGGCCGACAAGCGGTGGAGTATGTGTTAATTCGACGCAACCGAAGAACCTTACCAAGGCTGACA  
 TTGAGTGAAAGACCCAGAGATGGGTCCCCTCTCGGAAGCACGAAAACAGGTGGTGCATGGCTGTCGTCAGCTCGT  
 GTCGTGAGATGTGGTTAAGTCCCGAACAGCGCAACCCCTATCCTATGTTACCAAGCACGTAATGGTGGGACTC  
 ATAGGAGACTGCCAGGGATAACCTGGAGGAAGGCAGGGATGACGTCAAGTCATCATGCCCTATGTTGGGCTAC  
 ACACGTACTACAATGGTCGGCAACAAAGGGCAGCGAAGCCGGAGGGACCAATCCCAGAAACCCGACCCAGTT  
 CGGATCGCAGGCTGCAACCCGCTCGGTGAAGTTGAATCGCTAGTAATCGCAGGTCAGCATACTGCGGTGAATACG  
 TCCCCGGGCTTGTACACACCAGCGTACACACCAGAAAGTTGTAACACCCGAAGCCGGTGAAGATAACCTTTAGG  
 AGTCAGCTGTCTAAGTGGGGCGATGATTGGGTGAAGTCGTAACAAGGTAGCCGTTGAGAACGAGCGCTGGAT  
 CACCT

SEQ ID NO: 116

|NR\_114326.1| *Fusicatenibacter saccharivorans* strain HT03-11 16S ribosomal RNA gene, partial sequence  
 TGGCTCAGGATGAAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCAGTTAAGAAGATTTCGGATGAT  
 TCTTGACTGACTGAGCGCGGACGGGTGAGTAACCGTGGTGAACCTGCCATACCGGGGATAACAGCTGGAAAC  
 GGCTGCTAATACCGATAAGCGCACAGAGCTGCATGGCTCGGTGAAAAACTCCGGTGGTATGGATGGGCCGCG  
 TCTGATTAGGCAGTTGGCGGGTAACGGCCACCAAACCGACGATCAGTAGCCGGCTGAGAGGGCGACCGGCCACA  
 TTGGGACTGAGACACGGCCAAACTCTACGGGAGGCAGCAGTGGGAATATGCAAAATGGGGAAACCTGATGC  
 AGCGACGCCGCGTGAGCGAAGAAGTATTCGGTATGTAAGCTCTACAGCAGGGAAAGATAATGACGGTACCTGACT  
 AAAAGAGCCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTATCCGATTTACTGGGT  
 AAAGGGAGCGTAGACCGCAAGGCAAGTCTGATGTAAAACCCAGGGCTTAACCCCTGGACTGCATTGGAAACTGTCT  
 GGCTCGAGTGGCGAGGGTAAAGCGGAATTCTAGTGTAGCGGTAAATCGTAGATATTAGGAAGAACACCAAGTGG  
 CGAAGGGCGCTACTGGACGTTAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAACAGGATTAGATAACCTGGTA  
 GTCCACGCCGTTAACCGATGAAATGCTAGGTGGGAGCAAGCTCTCGGTGCCGCAAACGATTAAGCATTC  
 CACCTGGGGAGTACGTTGCAAGAATGAAACTCAAAGAATTGACGGGACCCGACAAGCGTGGAGCATGTGGTT  
 TAATTGCAAGCAACGCGAAGAACCTTACCGAGTCTGACATCCGATGACCGCCGTAACGGGCTTCTCTCGG  
 AGCATTGGAGACAGGTGGTGCATGGTGTGTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCGCAACGAGCGA  
 ACCCTTATCCTCAGTAGCCAGCAGGTAAGCTGGGACTCTGTGGAGACTGCCAGGGATAACCTGGAGGAAGGTGG  
 GATGACGTCAAATCATGCCCCATTGATCTGGCTACACAGCTGCTAACATGGCTAACAAAGGGAGGCAAAG  
 CGCGAGGTGGAGCAAATCCCCAAAATAACGCTCTAGTCGACTGCAACTCGACTGCACGAAGCTGGAA  
 TCGCTAGTAATCGCAATCAGAATGTCGCGGTAAACGTTCCGGTCTTGTACACACCAGCGTCAACCATGG  
 AGTTGGTAACGCCGAAGTCAGTGACCCAAACCTTTA

SEQ ID NO: 117

|NR\_102884.1| *Ruminococcus Champanellensis* strain 18P13 16S ribosomal RNA gene, complete sequence  
 AGAGTTTGATCTGGCTCAGGACGAACGCTGGCGCACGCCAACACATGCAAGTCGAAACGGAGATAAGACTTCGG  
 TTTTATCTTAGTGGGGACGGGTGAGTAACACGTGAGCAACCTGCCCTGAGAGAGGGATAGCTCTGGAAACGGA  
 TGGTAATACCTCATAACATAGCGGTACCGCATGATACTGCTATCAAAGATTATGCTCAGAGATGGCTCGCGTCT  
 GATTAGCTAGATGGTGGAGTAACGGCTCAGTAGCCGACTGAGAGGGTGAACGGCCACATTG  
 GGACTGAGACACGGCCCAAGACTCTACGGGAGGCAGCAGTGGGAATATTGCAAAATGGCGCAAGCCTGATGCAGC  
 GATGCCGCGTGGAGGAAGAGTTTCGGATTGTAACACTCTGTCTTAAGGGACGATAATGACGGTACCTTAGGAGG  
 AAGCTCCGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGAGCGAGCGTTGTCGGAATTACTGGGTAAA

-continued

GGGAGCGTAGGCGGGATTGCAAGTCAGATGTGAAAATATGGCTTAACCCATAGACTGCATTGAAACTGTAGTTC  
 TTGAGTGAAGTAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATATTAGGAGGAACATCGTGGCGA  
 AGGCGCTTACTGGCTTTACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTC  
 CACGCTGAAACGATGATTACTAGGTGTGGGGACTGACCCCTCCGTGCCGAGTTAACACAATAAGTAATCCAC  
 CTGGGGAGTACGGCGCAAGGTTGAAACTCAAAGGAATTGACGGGGCCCGACAAGCAGTGGAGTATGTGGTTAA  
 TTCGAAGCAACCGAAAAACCTTACCAAGGTCTTGACATCGAGTGAATGATCTAGAGATAGATCAGTCCTGGGACA  
 CAAAGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTGGGTTAAGTCCCACAGAGCGAACCG  
 TACCTTAGTTGCTACGCAAGAGCACTTAGAGGGACTGCCGTTGACAAAACGGAGGAAGGTGGGATGACGTAAA  
 TCATCATGCCCTTATGACCTGGCTACACACGTACTACAATGGCAATGAACAGAGGGAAAGCAATACAGTGTG  
 AGCAAATCCCCAAAATTGTCAGTTGAGGTGCAACTCGCCTACATGAAGTCGAAATTGCTAGTAATC  
 GCAGATCAGCATGCTGCCGTGATACTGTTCCGGCCTTGACACACCGCCGTACACCATGGAGTCGTAACAC  
 CGAAGCCAGTAGCCTAACCGCAAGGAGGGCGTGTGAAAGGTGGATTGATGACTGGGTGAAGTCGTAACAAAGGT  
 AGCCGTATCGAAGGTGCGCTGGATCACCT

SEQ ID NO: 118

| NR\_102971.1 | *Bifidobacterium bifidum* S17 strain S17 16S ribosomal RNA,  
 complete sequence  
 TTTTGTGGAGGGTTGATTCTGGCTCAGGATGAACGCTGGCGCGTGTAAACACATGCAAGTCGAACGGGATCCA  
 TCGGGCTTGCCTGGTGGTGGAGGTGGCGAACGGGTGAGTAATGCGTGGCGACCGACCTGCCCATGCTCCGGAATAGCT  
 CCTGGAAACGGGTGGAATGCCGATGTTCCACATGATCGCATGTGATTGTTGGGAAAGATTCTATCGCGTGGGATG  
 GGGTCGCGTCTATCAGCTTGTGGTGGAGGTACGGCTCACCAAGGCTTCGACGGGTAGCCGCGTGGAGGGCGAC  
 CGGCCACATTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGCAATGGCGCAAG  
 CCTGATCGCAGCGACGCCGCGTGGGGATGGAGGCCTGGGTGTAAACCTCTTGTGGAGCAAGCCTCGGG  
 TGAGTGTACCTTCGAAATAAGGCCGTAACACTACGTGCCAGCCGCGTAATACGTAGGGCGAAGCGTTATCC  
 GGATTTATTGGCGTAAAGGGCTCGTAGGCGCTCGCGTCCGGTGTGAAAGTCCATCGCTTAACGGTGGATCTG  
 CGCCGGGTACGGCGGGCTGGAGTGCCTAGGGAGACTGGAATTCCGGTGTAAACGGTGGAAATGTGTAGATATCGG  
 GAAGAACACCGATGGCGAAGGCAGGTCTCGGCCGTCAGCTGACGCTGAGGGCAGCGAAAGCGTGGGAGCGAACAGGA  
 TTAGATAACCTGGTAGTCCACCGTAAACGGTGGACGCTGGATGTGGGCACGTTCCACGTGTTCCGTGTCGGAGC  
 TAACCGCTTAAGCGTCCCGCTGGGGAGTACGGCGCAAGGCTAAACAGGCTAACAGGAAATTGACGGGGCCGACAAG  
 CGCGGAGCATGCGGATTAATTGATGCAACCGAAGAACCTTACCTGGCTTGACATGTTCCGACGACGCCAGAG  
 ATGGCGTTCCCTCGGGCGGGTACAGGTGGTGCATGGCGTCACTGCGTGTGAGATGTTGGGTTAAG  
 TCCCAGCAACGAGCGAACCCCTCGCCCGTGTGGCCAGCACGTTATGGTGGGACTCACGGGGACCGCCGGGTTAA  
 CTCGGAGGAAGGTGGGATGACGTAGATCATCATGCCCTACGTCAGGGCTCACGCTGCTACAATGGCGGT  
 ACAGCGGGATGCGACATGGCGACATGGAGCGGATCCCTGAAACCGGTCTCAGTCGGATGGAGCGCTGCAACCCGG  
 CTCCGTGAAGGGGGAGTCGCTAGTAATCGGGATCAGCAACGCCGCGTGAATGCGTCCGGCTTGTACACACC  
 GCCCGTCAAGTCATGAAAGTGGCAGCACCCGAAGCCGGTGGCTAACCCCTTGTGGGATGGAGCGTCAAGGTGA  
 GGCTCGTGTGATTGGGACTAAGTCGAAACAGGTAGCCGTACCGAAGGTGCGGCTGGATCACCTCCTTCT

SEQ ID NO: 119

| NR\_102980.1 | *Megasphaera elsdenii* strain DSM 20460 16S ribosomal RNA  
 gene, complete sequence  
 AGAGTTTGTCTGGCTCAGGACGAACGCTGGCGCGTGTAAACACATGCAAGTCGAACAGAGAGAGATGAGAAGC  
 TTGCTTCTTATCAATTGAGTGGCAAACGGGTGAGTAACCGTAAGCAACCTGCCCTCAGATGGGACAACAGCTG  
 GAAACGGCTGCTAATACCGAATACGTTCTTTGTGCGATGGCAGAGGGAAAGAAAGGGAGGCTCTCGGAGCTTCTG  
 CTGAAGGAGGGCTTGCCTGATTAGCTAGTGGAGGGTAACGGCCCACCAAGGCGACGATCAGTAGCCGCTG  
 AGAGGATGAACGCCACATTGGGACTGAGACACGCCAGACTCCTACGGGAGGCAGCAGTGGGAATCTCCGCAA

-continued

TGGACGAAAGTCTGACGGAGCAACGCCGCGTAACCGATGACGCCCTCGGTTGTAAAGTTCTGTTATACGGGACGA  
 ATGGCGTAGCGGTCAATACCCGTTACGAGTACCGTAAGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCG  
 CGGTAATACGTAGGTGGCAACGGCTTACGGGATTATTGGCGTAAGGGCGCAGGCGCGTCTGTAAGTCGGTCT  
 TAAAAGTGCAGGGCTTAACCCGTTAGGGGACCGAACACTCGCATGCTAGAGTATCGGAGAGGAAAGCGGAATCCTA  
 GTGTAGCGGTAAATCGTAGATATTAGGAGAACACCAGTGGCAAAGCGGTTCTGGACGACAACGTGCGTGA  
 GGCGCGAAGCCAGGGAGCAAACGGGATTAGATAACCCGTTAGTCTGGCGTAAACGATGGATACTAGGTGAG  
 AGGTATCGACCCCTCTGTGCCGGAGTTAACGCAATAAGTATCCCGCTGGGAGTACGCCGCAAGGCTGAAACTC  
 AAAGGAATTGACGGGGCCCGACAAGCGGTGAGATGTGGTTAATTCGACGCAACGCGAAGAACCTTACCAAGC  
 CTTGACATTGATTGCTATGGATAGAGATATCCAGTCTCTCGGAGGACAAGAAAACAGGTGGTGCACGGCTGTCG  
 TCAGCTCGTGTGAGATGTTGGTTAAGTCCCACGAGCGAACCCCTATCTCTGTTACCAGCGGTCGGCC  
 GGGGACTCAGGAGAGACTGCCACGACAATCGGAGGAAGGGGGATGACGTCAGTCATGCCCTATGGCT  
 TGGGCTACACAGTACTACAATGGCTTAATAGAGGAAGCGAAGGAGCGATCCGGAGCAAACCCAAAAACAGAG  
 TCCCAGTTCGGATTGCAGGCTGCAACTCGCTGCATGAAGCAGGAATCGTAGTAATCGCAGGTCACTGC  
 TGAAATACGTTCCCGGCTTGACACACCGCCCGTCACACCAGAAAGTCATTCAACCCGAAGCGGTGAGGTAAC  
 CTTTGAGCCAGCCGTCGAAGGTGGGGCGATGATTGGGTGAAGTGTAAAGGTAGCCGTATCGGAAGGTGCG  
 GCTGGATCACCT

SEQ ID NO: 120

|NR\_044645.2| *Dorea formicigenerans* strain ATCC 27755 16S ribosomal  
 RNA gene, complete sequence  
 TTAACACGAGAGTTGATCCTGGCTCAGGATGAAACGCTGGCGCGTCTAACACATGCAAGTCAGCGAAGCACATA  
 AGTTTGATTCTCGGATGAAGACTTTGTGACTGAGCGGCGACNNNGAGTAACCGTGGTAACCTGCCCTACAC  
 AGGGGATAACAGYTAGAAATGGCTGCTAACACCGCTAACACCAGAAAGTCATTCAACCCGAAGCGGTGAGGTAAC  
 GTGGTATGAGATGGACCCGCTGATTAGTAGTGGTGGAGGTAACGGCCACCNAGCCGACGATCAGTAGCCGAC  
 CTGAGAGGGTGACCCGACATTGGACTGAGACACGCCNNACTCCTACGGGAGGCGAGCAGTGGGAATTATTGCA  
 CAATGGCGAAAGCCTGATGAGCGACGCCGTGAAGGATGAAAGTATTCGGTATGTAACCTATCAGCAGGGA  
 AGAAAATGACGGTACCTGACTAAGAACGCCGCTAACACTGAGCGACGCCGNTAACAGTGGTAAACGGTGGNNAAGCG  
 TTATCCGGATTACTGGGTGAAAGGGAGCGTAGACGGCTGTGCAAGTCTGAAGTGAAGGATGGCTAACCTGT  
 GGAAGTCTGGAAACTGTGAGCTAGAGTGTGGAGAGGTAAGTGGAAATTCTTAGTGTAGCGGTGAAATGCGTAGA  
 TATTAGGAGGAACACCAGTGGCGAAGGCCNTACTGGACGATGACTGACGGTGTGGTAGCAAAGCTATTGGTGC  
 ACAGGATTAGATAACCTGGTAGTCCACGCCGTAACAGATGACTGCTAGGTGTGGTAGCAAAGCTATTGGTGC  
 CAGCTAACGCAATAACGAGTCACCGTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGNCNGCA  
 CAAGCGGTGGAGCATGTTAAATCGAANNAACCGAAGAACCTTACCTGATCTTGACATCCGATGACCGCTTC  
 GTAATGGAAGYTTTCTCGGAAACATCGGTGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGG  
 TTAAGTCCGCAACGAGCGAACCTTATCTCAGTAGCCAGATTAGGATGGGACTCTGGAGAGACTGCCAGGG  
 ATAACCTGGAGGAAGGTGGGATGACGTNNAACTCATGCCCCCTATGACCAGGGCTACACAGTGTACAATGGC  
 GTAAACAGAGGGAGGAGCGAGGCCGAGGCCGAGCAAATCTAAAAATAACGTCAGTCGGATTGAGTCTGCAAC  
 TCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAATGCTGCCGTGAATACGTTCCGGTCTTGTACAC  
 ACCGCCGTCACACCATGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGAAAGGAGGGAGCTGCCGAAGGTGG  
 GACCGATAACTGGGGT

SEQ ID NO: 121

|NR\_118643.1| *Eisenbergiella tayi* strain B086562 16S ribosomal RNA gene,  
 partial sequence  
 GGTATAACTTAGTGGCGGACGGGTGAGTAACCGTGGAAACCTGCCCTGTACCGGGGATAACACTTAAAGG  
 TGCTAATACCGCATAAGCGCACGGAACCGCATGGTCCGTGAAAAACTCCGGTGGTACAGGATGGTCCCGCT

-continued

GATTAGCCAGTTGGCAGGGTAACGGCCTACCAAGCGACGATCAGTAGCCGGCTGAGAGGGTGAAACGCCACATTG  
 GGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGACAATGGGGAAACCCCTGATGCAGC  
 GACGCCGCGTGAAGTATTCCGTATGAAAGCTCTACGAGGGAAAGAAAATGACGGTACCTGACTAAG  
 AAGCCCCGGCTAACTACGTGCCAGCAGCCGCGTAATACTAGTAGGGGCAAGCGTTATCGGATTTACTGGGTAAA  
 GGGAGCGTAGACGGCATGGCAAGCCAGATGTGAAACCCAGGGCTCAACCTGGGATTGCACTTGGAACTGCCAGGC  
 TGGAGTGCAGGGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGAACACCAGTGGCGA  
 AGCGGCTTACTGGACTGTAACGTGAGCTGAGGCTCGAAAGCGTGGGAGCAAACAGGATTAGATAACCTGGTAGTC  
 CACCGGTTAACCGATGATTGCTAGGTGAGGGTATGGACCCATCGGTGCCAGCTAACGCAATAAGCAATCCA  
 CCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGCGGTGGAGCATGTGGTTA  
 ATTGAGAAGCAACCGCAAGAACCTTACCAAGTCTTGACATCCCAATGACGCACCTGTAAAGAGGTGTTCCCTCGGG  
 CATTGGAGACAGGTGGTGCATGGTTGCTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGAACGAGCGAAC  
 CCTTATTCTTAGTAGCCAGCAGGTAAAGCTGGCACTCTAAGGAGACTGCCGGGATAACCCGGAGGAAGGCGGGGA  
 TGACGTCAAATCATCATGCCCTTATGATTGGTACACACGTCTACAGTGGTAAACAAAGGAAGCGAGACA  
 GTGATGTGGAGCAAATCYCAGAAAATACGTCTCAGTTGGATTGAGTCTGCAACTCGACTACATGAAGCTGGAATC  
 GCTAGTAATCGCGAATCAGCATGTCGGTGAATAGTCCCGGGCTTGTACACACCGCCGTACACCATGGGAG  
 TTGGAAATGCCGAAGTCTGTGACCTAACCGAAAGGGAGGAGCAGCGAAGGCAGGTCTGATAACTGGGTAAAGTC  
 GTAA

SEQ ID NO: 122

|NR\_118730.1| *Clostridium symbiosum* strain ATCC 14940 16S ribosomal  
 RNA gene, partial sequence  
 AAACATGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCCTAACACATGCAAGTCGAACGAAGCGATT  
 AACCGAAGTTTCGGATGGAAGTTGAAATTGACTTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCCTGTAC  
 TGGGGACAACAGTTAGAAATGACTGCTAATACCGATAAGCGCACAGTATTGATGACAGTGTGAAAAACTCCG  
 GTGGTACAAGATGGACCCCGCTCTGATTAGCTAGTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGCCGAC  
 CTGAGAGGGTACCGGCCACATTGGACTGAGACACGGCCNNAACTCCTACGGGAGGCAGCAGTGGGAATTGCA  
 CAATGGCGAAAGCCTGATGCGACGCCGCGTGAGTGAAGAAGTATTGCGTATGAAAGCTCTATCAGCAGGG  
 AGAAAATGACGGTACCTGACTAAGAAGCCCGCTAACACTGCGCAGCAGCCGGTAATACGTAGGGNNAGCG  
 TTATCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAAGTGAAGAACCCCGCNCTCAACTGCG  
 GNNTGCTTGGAAACTGTTAACTGGAGTGTGGAGAGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTADA  
 TATTAGGAGGAACACNAGTGGCGAAGGCAGACTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAA  
 ACAGGATTAGATAACCTGGTAGTCCACGCCGTAACCGATGAATACTAGGTGTTGGGAGCAAAGCTCTCGGTGCC  
 TCGCAAACGCGAGTAAGTATTCCACCTGGGAGTACGTTGCGAAGAATGAAACTCAAAGGAATTGACGGGACNGCA  
 CAAGCGGTGGAGCATGTTAATTCGAANNAACCGGAAGAACCTTACCGGTCTTGACATCGACTCGACGGGG  
 GTAACGTCCNNNTNCCTCGGGCGGAGAAGACAGGTGGTGCATGGTTGCGTCAGCTCGTGTGAGATGTTGG  
 TTNAGTCCCAGAGCGAACCCCTATTCTAAGTAGCCAGCGGTTGGCGGAACTCTGGAGACTGCCAGGG  
 ATAACCTGGAGGAAGGTGGGATGACGTCNAATCATGCCCCCTTATGATCTGGCTACACACGTGCTACAATGGC  
 GTAAACANAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTCAAAATAACGTCTCAGTTGGACTGCAGGCTGCAAC  
 TCGCCTGCACGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGCGGTGAATACGTCCGGTCTTGTACAC  
 ACCGNNCGTCACACCATGGAGTCAGTAACGCCGAAGTCAGTGACCCACCGCAAGGAGGGAGCTGCCGAAGGG  
 GACCGANAACNNGGG

US 10,555,980 B2

159

-continued

160

SEQ ID NO: 123

|NR\_113243.1| *Erysipelatoclostridium ramosum* strain JCM 1298 16S  
ribosomal RNA gene, partial sequence  
AGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCCTAATACATGCAAGTCGAACCGCAGCACGGACTTGTGCTCG  
AGTGGCGAACGGGTGAGTAATACTAACCTGCCCTAGACAGGGGATAACTATTGAAACGATAGCTAACGACC  
GCATAGGTACGGACACTGCATGGTAGCGTATTAAAGTGCTCAAAGCACTGGTAGAGGGATGGACTTATGGCGCAT  
TAGCTGGTTGGGGGGTAACGGCCCACCAAGGCAGCGATGCCCTAGCCGACCTGAGAGGGTGACCGGCCACACTGGGA  
CTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTCGGCAATGGGGAAACCGCTGACCGAGCAAC  
GCCGCGTGAAGGAAGAAGGTTTCGGATTGTAACCTCTGTTATAAGGAAGAACGGCGTACAGGAATGGTAGC  
CGAGTGACGGTACTTTATTAGAAAGCCACGGCTAACACTACGTGCCAGCAGCGCGTAATACGTAGGTGCAAGCGTT  
ATCCGGAATTATTGGCGTAAAGAGGGAGCAGCGGAGCAAGGGTCTGTTGAAAGCCTGAAGCTTAACCTCAGT  
AAGCCATAGAAACCAGGAGCTAGGTGCAAGGAGAGGATCGTGAATTCCATGTTAGCGGTGAAATGCGTAGATAT  
ATGGAGGAACACCAGTGGCGAACGGCAGCATGGCCTGCAACTGACGCTCAGTCCGAAAGCGTGGGAGCAAATA  
GGATTAGATAACCTAGTAGTCCACGCCAACAGTGAAGTACTAAGTGGATGTCAGGTTCAAAAGTCAAGTGCAGTT  
ACCGAATAAGTACTCCGCTGAGTAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGGGCCGACAAGCG  
GTGGAGCATGTGGTTAATTGCAAGCAACCGCAAGAACCTTACCGGTCTGACATACTCATAAAAGGCTCAGAGAT  
GGAGAGATAGCTATATGAGATAACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGGTGGTTAAGTCCG  
CAACGAGCGAACCTTATCGTAGTTACCATCATTAAGTGGGACTCTAGCGAGACTGCCAGTGACAAGCTGGAG  
GAAGGCGGGGATGACGTCAAATCATGCCCCTTATGACCTGGCTACACACGTGCTACAATGGATGGTGCAGAGG  
GAAGCGAAGCCGCGAGGTGAAGCAAAACCCATAAAACCATTCAGTGGATTGAGATCTGCAACTCGTACATG  
AAAGTTGAATCGCTAGTAATCGCAATCAGCATGTCGCGGTGAATACGTTCTGGCCTTGTACACACCGCCGTCA  
CACCACGAGAGTTGATAAACCCGAAGCCGTTGGCTAACCGCAAGGAAGGAGCTGCTAAGGTGGATTGATT  
GGGGTGAAGTCTAACAGGTAAAC

SEQ ID NO: 124

|PROKKA\_00507 16S ribosomal RNA gene | VE202-7  
ATGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAAGG  
AAAGTTTGGATGGAATTGATTGACTGAGTGCGGACGGGTGAGTAACGCGTGGATAACCTGCCTCACACTGGGG  
ATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCCGGTGGT  
TGAGATGGATCCCGCTGATTAGCAGTTGGGGGTAACGGCCACCAAGCGACGATCAGTAGCCGACCTGAGA  
GGGTGACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGACAATGG  
GCGAAAGCCTGATGCGACGCCGCGTGAGTGAAGAAGTATTGCGTATGTAAGCTCTACAGCAGGGAAAGAAA  
TGACGGTACCTGACTAAGAACCCCGCTAACACTACGTGCCAGCAGCGCGTAATACGTAGGGCAAGCGTTATCC  
GGATTCACTGGGTAAAGGGAGCGTAGACGGCGAACGAAAGTCTGAAGTAAAAACCCAGGGCTCAACCTGGGACTG  
CTTGAAACTGTTGCTAGAGTGTGGAGAGGTAAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATTAG  
GAGGAACACCAGTGGCGAACGGGGTTACTGGACGATAACTGACGTTGAGGCTGAAAGCGTGGGAGCAAACAGGA  
TTAGATAACCTGGTAGTCCACCCGTAACAGTGAATGCTAGGTGTTGGGGCAAGCCCTCGGTGCCGTGCAA  
ACCGAGTAAGCATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGACAAGCG  
GTGGAGCATGTGGTTAATTGCAAGCAACCGCAAGAACCTTACCAAGTCTGACATCCTTGTGACCGCGTGTAA  
GCCCTCCCTCGGGCAAGAGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGGTGGTTAAGT  
CCCGCAACGAGCGAACCTTACCTTAGTACGCCAGCAGGTAGAGCTGGCACTCTAGGGAGACTGCCAGGGATAAC  
CTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTTATGATTGGCTACACACGTGCTACAATGGCTAA  
CAAAGGGAAAGCAAGACAGTGTGGAGCAAATCCAAAAATAACGTCCAGTTGGACTGTAGTCTGCAACCCGAC  
TACACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGGTCTGTACACACCGC

-continued

CCGTACACCATGGGAGTCAGCAACGCCAAGTCAGTGACCAACTCGAACAGAGAGGGAGCTGCCGAAGGCCGGC  
AGGTAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGAACAGGTGCGCTGGATCACCTCCTT

SEQ ID NO: 125

| PROKKA\_00709 16S ribosomal RNA gene | VE202-7  
ATGAGAGTTGATCCTGGCTCAGGTAGAACGCTGGCGCTGCCAACACATGCAAGTCGAACGAAGCAATTAAAAT  
GAAGTTTCGGATGGATTTGATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGG  
GGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTAAAAACTCCGGTGG  
TGTGGATGGATCCCGCTCTGATTAGCCAGTTGGCGGGTAACGGCCCACCAAAGCGACGATCAGTAGCCGACCTGA  
GAGGGTACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGACA  
GGGCAGAACGCTGATGCAGCGACGCCGCGTGAAGAAGTATTCTGGTATGAAAGCTCTACAGCAGGGAAAGAA  
AATGACGGTACCTGACTAAGAACCCCCGCTAACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTT  
CCGGATTACTGGGTGAAAGGAGCGTAGACGGCGAACAGCTGAAGTGTGAAACTCCAGGGCTCAACCTGG  
TGCTTTGGAAACTGTTGCTAGAGTGTGGAGGGTAAGTGGATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
AGGAGGAACACCAGTGGCGAAGGCGCTTACTGGACGATACTGACGTTGAGGCTCGAAAGCGTGGGAGC  
GATTAGATACCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTGGGGCAAGCCTTCGGTGC  
AAACGCAGTAAGCATTCCACCTGGGAGTACGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA  
CCGGAGCGATGGTTAATTGAAAGCACGCCAACCTTACCAAGTCTGACATCCTTGACCGCGTGTAA  
CGGCGCCTCCCTCGGGCAAGAGAGACAGGTGGTGCATGGTGTGCTCAGCTGTGCTGAGATGTTGG  
GTCCCGCAACGAGCGAACCCCTATCCTAGTAGCCAGCAGGTAAGCTGGCCTCTAGGGAGACTGCCAGGG  
ACCTGGAGGAAGGTGGGATGACGTCAGTCAGTCAGTCCAGGTAACGTTGGGACTCTAGTGTAGCTG  
AACAAAGGGAAAGCAAGACAGTGTGGAGCAAATCCAAAATAACGTCCAGTGGACTGTAGCTG  
ACTACACGAAGCTGGATCGCTAGTAATCGCGAATCAGAATGTCGCGGTAACTACGTTCCGGTCTTGAC  
GCCCGTCACACCATGGAGTCAGCAACGCCAAGTCAGTACGCCAACCGCAAGAGAGGGAGCTGCC  
GCAGGTAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGAACAGGTGCGCTGGATCACCTCCTT

SEQ ID NO: 126

| PROKKA\_01766 16S ribosomal RNA gene | VE202-7  
ATGAGAGTTGATCCTGGCTCAGGTAGAACGCTGGCGCTGCCAACACATGCAAGTCGAACGAAGCAATTAAAAT  
GAAGTTTCGGATGGATTTGATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGG  
GGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTAAAAACTCCGGTGG  
TGTGAGATGGATCCCGCTCTGATTAGCCAGTTGGCGGGTAACGGCCCACCAAAGCGACGATCAGTAGCCGACCTGA  
GAGGGTACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGACA  
GGGCAGAACGCTGATGCAGCGACGCCGCGTGAAGAAGTATTCTGGTATGAAAGCTCTACAGCAGGGAAAGAA  
ATGACGGTACCTGACTAAGAACCCCCGCTAACACTGTCGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTT  
CGGATTACTGGGTGAAAGGAGCGTAGACGGCGAACGAAGTCTGAAGTAAAAACCCAGGGCTAACCTGG  
GCTTGGAAACTGTTGCTAGAGTGTGGAGGGTAAGTGGAACTCTAGTGTAGCGGTGAAATGCGTAGATATT  
GGAGGAACACCAGTGGCGAAGGCGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGC  
ATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAATGCTAGGTGTTGGGGCAAGCCTTCGGTGC  
ACCGAGTAAGCATTCCACCTGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA  
GTGGAGCATGGTTAATTGAAAGCAACGCGAAGAACCTTACCAAGTCTGACATCCTTGACCGCGTGT  
GCCCTCCCTCGGGCAAGAGAGACAGGTGGCATGGTGTGCTCAGCTGTGCTGAGATGTTGG  
CCCGCAACGAGCGAACCCCTATCCTAGTAGCCAGCAGGTAAGCTGGCCTCTAGGGAGACTGCCAGGG  
CTGGAGGAAGGTGGGATGACGTCAAATCATGCCCTTATGATTGGCTACACAGTGTACA  
CAAAGGGAAAGCAAGACAGTGTGGAGCAAATCCAAAATAACGTCCAGTTGGACTGTAGTCTG  
CAACCCGAC

-continued

TACACGAAGCTGGAATCGCTAGTAATCGCGAACATCGAATTCGCGGTGAATACGTTCCGGGTCTTGTACACACCGC  
 CGCTCACACCATTGGGAGTCAGCAACGCCGAAGTCAGTGACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCCGGC  
 AGGTAACTGGGTGAAGTCGTAACAAGGTACCCGTATCGGAAGGTGCCGTGGATCACCTCTT

SEQ ID NO: 127

| PROKKA\_01779 16S ribosomal RNA gene | VE202-7  
 ATGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCTGCCTAACACATGCAAGTCGAACGAAGCAATTAAAAT  
 GAAGTTTCGGATGGATTTGATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGG  
 GGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCCGGTGG  
 TGTGAGATGGATCCCGCTCTGATTAGCCAGTGGCGGGTAAACGGCCCACCAAAGCGACGATCAGTAGCCGACCTGA  
 GAGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCAACAAAT  
 GGGCAGAACGCTGATGCAGCGACGCCGCGTGAAGTAAGAAGTATTCGGTATGTAAGCTCTACAGCAGGGAAAGAA  
 ATGACGGTACCTGACTAAGAGCCCGGTAACACTACGTGCCAGCAGCGCGTGAATACGTTAGGGCAAGCGTT  
 CCGGATTACTGGGTGAAAGGGAGCGTAGACGGCGAAGCAAGTCAGTGAAGTGAAGAACCCAGGGCTAACCTGGGAC  
 TGCTTGGAAACTGTTGCTAGAGTGTGGAGAGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCAAGTGGCGAAGGCGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAACAG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATGCTAGGTGTTGGGCAAAGCCCTCGGTGCCGTGCA  
 AACCGCAGTAAGCATTCCACCTGGGAGTACGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGC  
 GTGGAGCATGTGGTTAATTGCAAGCAACCGAAGAACCTTACCAAGTCTGACATCCTCTGACCGCGTGTAAAC  
 GCGCCTCCCTCGGGCAGGAGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGGTTAAG  
 TCCCGCAACGAGCGAACCCCTATCCTTAGTAGCCAGCAGGTAGCTGGGACTCTAGGGAGACTGCCAGGGATAA  
 CCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTTATGATTGGCTACACACGTGCTACAATGGCGTAA  
 ACAAAAGGAAGCAAGACAGTGATGTGGAGCAAATCCAAAATAACGTCCCAGTTGGACTGTAGTCTGCAACCCGAC  
 TACACGAAGCTGGAATCGCTAGTAATCGCGAACATCGAATGTCGCGGTGAATACGTTCCGGTCTTGTACACACCGC  
 CGTCACACCATTGGGAGTCAGCAACGCCGAAGTCAGTGACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCCGGC  
 AGGTAACTGGGTGAAGTCGTAACAAGGTACCCGTATCGGAAGGTGCCGTGGATCACCTCTT

SEQ ID NO: 128

| PROKKA\_05926 16S ribosomal RNA gene | VE202-7  
 ATGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCTGCCTAACACATGCAAGTCGAACGAAGCAATTAAAAT  
 GAAGTTTCGGATGGATTTAATTGACTGAGTGGCGGACGGGTGAGTAACCGTGGATAACCTGCCTCACACTGG  
 GATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTACGGTGTGAAAAACTCCGGTGG  
 GTGAGATGGATCCCGCTCTGATTAGCCAGTGGCGGGTAAACGGCCCACCAAAGCGACGATCAGTAGCCGACCTGAG  
 AGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCAACATG  
 GCGAAGCCGTATGCAGCGACGCCGCGTGAAGTAAGAAGTATTCGGTATGTAAGCTCTACAGCAGGGAAAGAA  
 ATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTATC  
 CGGATTACTGGGTGAAAGGGAGCGTAGACGCCAAGCAAGTCTGAAGTGAAGAACCCAGGGCTAACCTGGGACT  
 GCTTGGAAACTGTTGCTAGAGTGTGGAGAGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATTA  
 GGAGGAACACCAAGTGGCGAAGGGCGCTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAACAGG  
 ATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATGCTAGGTGTTGGGGCAAAGCCCTCGGTGCCGTGCA  
 AACCGCAGTAAGCATTCCACCTGGGAGTACGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAGC  
 GGTGGAGCATGTGGTTAATTGCAAGCAACGCCAAGAACCTTACCAAGTCTGACATCCTCTGACCGCGTGTAAAC  
 GCGCCTCCCTCGGGCAGGAGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGGTTAAG  
 TCCCGCAACGAGCGAACCCCTATCCTTAGTAGCCAGCAGGTAAGCTGGCACTCTAGGGAGACTGCCAGGGATAA  
 CCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTTATGATTGGCTACACACGTGCTACAATGGCGTAA

-continued

ACAAAGGGAAAGCAAGACAGTGTGAGCAAATCCAAAATAACGTCCCAGTCGGACTGTAGTCGAACCCGAC  
TACACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACCGC  
CCGTACACCATGGGAGTCAGAACGCCGAAAGTCAGTGACCCAACTCGCAAGAGAGGGAGCTGCCGAAGGCCGGC  
AGGTAACTGGGTGAAGTCGTAAACAAGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 129

| PROKKA\_01784 16S ribosomal RNA gene  
TCAAAGAGTTGATCCTGGCTCAGGACGAACGCTGGCGGCCCTAACACATGCAAGTCGAACGGAGCTTACGTT  
TGAAGTTCGGATGGATGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTCAGAGGG  
GGATAACAGCGGAAACGGCTGCTAAACCGCATGATGTTGGGGGGCACATGCCCTGCAACCAAAGGAGCAATCC  
GCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCGACGATCGGTAGCCGGACT  
GAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAGCAGTGGGGATATTGCACA  
ATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAAGGAAAGACGGTCTCGGATTGTAACACCTCTGCTTGGGAAG  
AAAATGACGGTACCCAAAGAGGAAGCTCGCTAAACTACGTGCCAGCAGCGCGTAATACGTAGGGAGCAAGCGTT  
GTCCGGAATTACTGGGTGAAAGGGAGCGTAGGCAGGATGGCAAGTAGAATGTTAAATCCATCGGCTAACCGGTGG  
CTGCGTTCTAAACTGCCGTTCTGAGTGAAGTAGAGGCAAGGGGAATTCCCTAGTGTAGCGGTGAAATGCGTAGATAT  
TAGGAGGAACACCAGTGGCGAAGGCGGCTGCTGGCTTAAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAACAA  
GGATTAGATAACCTGGTAGTCCACGCCGTAACCGATGATTACTAGGTGTGGGGACTGACCCCTCCGTGCCAG  
TTAACACAATAAGTAATCCACCTGGGAGTACGCCGCAAGGTTAAACTCAAAGGAATTGACGGGGGCCGCACAA  
GCAGTGGAGTATGTGGTTAATCGAAGCAACCGCAAGAACCTTACCGGTCTGACATCGGATGCATAGCCTAGAG  
ATAGGTGAAGCCCTCGGGCATCCAGACAGGGTGCATGGTTGTCGTCAGCTGTGAGATGTTGGGTTAA  
GTCCCGCAACGAGCGAACCCCTATTAGTTGCTACGCAAGAGCACTCTAAATGAGACTGCCGTTGACAAAACGGA  
GGAAGGTGGGGATGACGTCAAATCATGCCCTTACAGCTGGCTACACAGCTACAAATGGCAACTAAACAG  
AGGGCGCGACACCGCGAGGTGAAGCGAATCCGAAAAGTGTCTCAGTTGACATTGCAAGGCTGCAACCCGCTGCA  
TGAAGTCGAATTGCTAGTAATCGGGATCAGCATGCCGCTGAAATCGTCCGGCTTGTACACACCGCCGT  
CACACCATGGAGTCGTAACACCGAAGCCAGTAGCCTAACCGCAAGGGGGCGCTGCGAAGGTGGATTGATGA  
CTGGGTGAAGTCGTAAACAAGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 130

| PROKKA\_01864 16S ribosomal RNA gene  
TCAAAGAGTTGATCCTGGCTCAGGACGAACGCTGGCGGCCCTAACACATGCAAGTCGAACGGAGCTTACGTT  
TGAAGTTCGGATGGACGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTCAGAGGG  
GATAACAGCGGAAACGGCTGCTAAACCGCATGATGTTGGGGGGCACATGCCCTGCAACCAAAGGAGCAATCC  
CTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGTAACGGCCCACCAAAGCGACGATCGGTAGCCGGACTG  
AGAGGTTGAACGCCACATTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAGCAGTGGGGATATTGCACAA  
TGGCGAAAGCCTGATGCAGCGACGCCGCGTGAAGGAAAGACGGTCTCGGATTGTAACACCTCTGCTTGGGAAGA  
AAATGACGGTACCCAAAGAGGAAGCTCGGCTAAACTACGTGCCAGCAGCGCGGTAAACGTAGGGAGCAAGCGTT  
TCCGGAAATTACTGGGTGAAAGGGAGCGTAGGGGGATGGCAAGTAGAATGTTAAATCCATCGGCTAACCGGTGG  
TGCCTCTAAACTGCCGTTCTGAGTGAAGTAGAGGCAAGGGGAATTCCCTAGTGTAGCGGTGAAATGCGTAGATATT  
AGGAGGAACACCAGTGGCGAAGGCGCTGCTGGCTTAAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAACAG  
GATTAGATAACCTGGTAGTCCACGCCGTAACCGATGATTACTAGGTGTGGGGACTGACCCCTCCGTGCCAGT  
TAACACAATAAGTAATCCACCTGGGAGTACGCCGCAAGGTTAAACTCAAAGGAATTGACGGGGGCCGCACAA  
CAGTGGAGTATGTGGTTAATCGAAGCAACCGCAAGAACCTTACCGGTCTGACATCGGATGCATAGCCTAGAG  
TAGGTGAAGCCCTCGGGCATCCAGACAGGTGGTCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAG  
TCCCGCAACGAGCGAACCCCTATTAGTTGCTACGCAAGAGCACTCTAAATGAGACTGCCGTTGACAAAACGGAG

-continued

```

GAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCTGGCTACACACGTACTACAATGGCACTAAAACAGA
GGCGGGCAGACCCGCGAGGTGAAGCGAATCCCAAAAAGTGTCTCAGTCAGATTGCAGGCTGCAACCCGCTGCAT
GAAGTCGAATTGCTAGTAATCGCGATCAGCATGCCCGTGAAATACGTCAGGCTGCAACACCGCCCGTC
ACACCATGGGAGTCGTAACACCGAAGCCAGTAGCCTAACCGCAAGGGGGCGCTGTCGAAGGTGGATTGATGAC
TGGGGTGAAGTCGTAACAAGGTAGCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 131

```

| PROKKA_02671 16S ribosomal RNA gene
TCAAAGAGTTGATCCTGGCTCAGGACGAACGCTGGCGCGCCTAACACATGCAAGTCGAACGGAGCTTACGTT
TGAAGTTTCGGATGGATGAATGTAAGCTTAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTCAGAGGG
GGATAACAGCGGAAACGGCTGCTAACACCGCATGATGTTGGGGGGCACATGCCCTGCAACCAAAGGAGCAATCC
GCTGAAAGATGGGCTCGCGTCCGATTAGCCAGTTGGCGGGTAACGGCCACCAAAGCGACGATCGGTAGCCGACT
GAGAGGTGAACGGCACATTGGGACTGAGACACGCCAGACTCCTACGGAGGAGCAGCTGGGATATTGACAA
ATGGGCAGAGCTGATGCAGCGACGCCGCGTGGAGGAAGACGGTCTCGGATTGTAACCTCTGTCTTGGGAAG
AAAATGACGGTACCCAAAGAGGAAGCTCGGCTAACACTACGTGCCAGCAGCGCGGTAATACGTAGGGAGCAAGCGTT
GTCGGAATTACTGGGTGTAAGGGAGCGTAGGGGGATGCAAGTAGATAATGTTAACATCGGCTAACCGGTGG
CTCGTTCTAAACTGCCGTTCTGAGTGAAGTAGAGGCAGGGGAATTCTTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACAGTGGCGAAGGGGGCTGCTGGCTTAACGTGACGCTGAGGCTCGAAAGCGTGGGAGCAACAA
GGATTAGATAACCTGGTAGTCCACGCCGTAACCGATGATTACTAGGTGTGGGGACTGACCCCTCCGTGCCGAG
TTAACACAATAAGTAATCCACCTGGGAGTACGGCGCAAGGGTAAACAGGAAATTGACGGGGGCCGACAA
GCACTGGAGTATGTTAATTCGAAGCAACCGAAGAACCTTACCGAGTGTGACATCGGATGCAAGCTAGAG
ATAGGTGAAGCCCTCGGGCATCCAGACAGGTGGCATGGTGTGTCGTCAGCTCGTGTGAGATGTTGGTTAA
GTCGGCAACAGCGCAACCCATTATTAGTGTACGCAAGAGCACTCTAACGAGACTGCCGTTGACAAAAGGA
GGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCTGGCTACACACGTACTACAATGGCACTAAAACAG
AGGGCGCGCACCGCGAGGTGAAGCGAATCCGAAAAGTGTCTCAGTCAGATTGCAAGCTGCAACCCGCTGCA
TGAAGTCGAATTGCTAGTAATCGGGATCAGCATGCCGGTGAAATACGTTCCCGGCTTGTACACACCGCCG
CACACCATGGGAGTCGTAACACCGAAGCCAGTAGCCTAACCGCAAGGGGGCGCTGTCGAAGGTGGATTGATGA
CTGGGTGAAGTCGTAACAAGGTAGCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 132

```

| PROKKA_00690 16S ribosomal RNA gene
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTTT
CAGAATCTCGGAGGAAGAGGACAGTGACTGAGCGCGGAGCGGTGAGTAACCGCTGGCACACCTGCCTACAGG
GGATAACAGTTAGAAATGACTGCTAACCGCATAAGCGCACAGGACCGCATGGTAGTGTGAAAAACTCCGGTG
GTATGAGATGGACCCCGCTCTGATTAGGTGTGGGGAAAGGCCTACCAAGCCGACGATCAGTAGCCGACCTG
AGAGGGTACCGGCCACATTGGGACTGAGACACGCCAACCTCTACGGAGGCCAGCAGTGGGAATTGACAA
TGGGGGAAACCTGATGCAGCGCCGCGTGAAGGAAGAAGTATTCGGTATGTAACCTCTACAGCAGGGAGA
AAATGACGGTACCTGAGTAAGAAGCACCGCTAACACGTGCCAGCAGCGCGGTAATACGTATGGTCAAGCGTTA
TCCGGATTTACTGGGTAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGAAACCCAGGGCTAACCTGGGA
CTGCTTGGAAACTGCAGATCTGGAGTGCAGGAGAGGTAAAGCGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACAGTGGCGAAGGGGGCTTACTGGACGGTACTGACGTTGAGGCTCGAAAGCGTGGGAGCAACAA
GGATTAGATAACCTGGTAGTCCACGCCGTAACCGATGACTACTAGGTGTGGTGTGCAAAGCACATCGGTGCCGAG
CAAACCGAATAAGTAGTCCACCTGGGAGTACGTTGCAAGAATGAAACCTAACAGGAAATTGACGGGGACCCGACAA
GCCGTGGAGCATGTGGTTAATCGAAGCAACCGAAGAACCTTACCTGGTCTTGACATCGGATGACGGGGAGTA
ATGTCGCCGTCCCTCGGGCGTCCGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATGTTGGGTTA

```

-continued

```

AGTCCCGCAACGAGCGCAACCCTATCTCAGTAGCCAGCATATAAGGTGGGACTCTGGAGAGACTGCCAGGGAGA
ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGCCAGGGTACACACGTGCTACAATGGCGTA
ACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAACCTCCAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG
ACTACATGAAGCTGGAATCGCTAGTAATCGCGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC
GCCCGTCACACCATGGGAGTCAGTAACGCCGAAGCCAGTGACCCAACCTTAGAGGAGGGAGCTGTCGAAGGCCGG
CGGATAACTGGGGTAAGTCGTAACAAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 133

```

| PROKKA_00991 16S ribosomal RNA gene
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCCTAACACATGCAAGTCAGCGAAGCGCTGTTTT
CAGAATCTCGAGGAAGAGGACAGTGACTGAGCGGGACGGGTGAGTAACCGTGGCAACCTGCCTCATACAGG
GGGATAACAGTTAGAAATGACTGCTAACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTGAAAAACTCCGGTG
GTATGAGATGGACCCCGCTCTGATTAGGTAGTTGGGGTAAAGGCCTAACAGCGACGATCAGTAGCCGACCTG
AGAGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCACAA
TGGGGAAACCTGATGCAGCGACGCCGTGAAGGAAGAAGTATTCGGTATGTAACCTATCAGCAGGAAGA
AGATGACGGTACCTGAGTAAGAACGACCGCTAACACGTGCAAGCGCAGCGCGGTAAACGTATGGTCAAGCGTTA
TCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGGAAACCCAGGGCTAACCTGGGA
CTGCTTGGAAACTGAGATCTGGAGTGCAGCGGGAGGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACCGAGTGGCGAAGGGCGCTACTGGACGGTGAACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAA
GGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGTGCAAAGCACATCGTGCAGCAG
CAAACGCAATAAGTAGTCCACCTGGGAGTACGTTGCAAGAACCTTACCTGGCTTGACATCCGGATGACGGGAGTA
GCGGTGGAGCATGTGGTTAACGCAACGCCAACGGTACCGGTTAACGCGTAAACGATGACTACTAGGTGTCGGTGTG
ATGTCGGTCCCTCGGGCATCCGAGACAGGTGGTGTGCAAGGAGATGTGGGTAACTGGCTCGTGTGAGATGTGGGT
AGTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCATATAAGGTGGGACTCTGGAGAGACTGCCAGGGAGA
ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGCCAGGGTACACACGTGCTACAATGGCGTA
ACAAAGGGAAGCGAGAGGGTACCTGAAGCGAACCTCCAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG
ACTACATGAAGCTGGAATCGCTAGTAATCGCGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC
GCCCGTCACACCATGGGAGTCAGTAACGCCGAAGCCAGTGACCCAACCTTAGAGGAGGGAGCTGTCGAAGGCCGG
CGGATAACTGGGGTAAGTCGTAACAAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 134

```

| PROKKA_01948 16S ribosomal RNA gene
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCCTAACACATGCAAGTCAGCGAAGCGCTGTTTT
CAGAATCTCGAGGAAGAGGACAGTGACTGAGCGGGACGGGTGAGTAACCGTGGCAACCTGCCTCATACAGG
GGGATAACAGTTAGAAATGACTGCTAACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTGAAAAACTCCGGTG
GTATGAGATGGACCCCGCTCTGATTAGGTAGTTGGGGTAAAGGCCTAACAGCGACGATCAGTAGCCGACCTG
AGAGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGCACAA
TGGGGAAACCTGATGCAGCGACGCCGTGAAGGAAGAAGTATTCGGTATGTAACCTATCAGCAGGAAGA
AGATGACGGTACCTGAGTAAGAACGACCGCTAACACGTGCAAGCGCAGCGCGGTAAACGTATGGTCAAGCGTTA
TCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGATAGGCAAGTCTGGAGTGGAAACCCAGGGCTAACCTGGGA
CTGCTTGGAAACTGAGATCTGGAGTGCAGCGGGAGGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACCGAGTGGCGAAGGCCTTACTGGACGGTACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAA
GGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTCGGTGTGCAAAGCACATCGTGCAGCAG
CAAACGCAATAAGTAGTCCACCTGGGAGTACGTTGCAAGAACCTAACCTGGCTTGACATCCGGATGACGGGAGTA
GCGGTGGAGCATGTGGTTAACGCAACGCCAACCTTACCTGGCTTGACATCCGGATGACGGGAGTA

```

-continued

```

ATGTCGCCGTCCCTCGGGCATCCGAGACAGGTGGTGCATGGTTGTCGTCACTCGTGTGAGATGTTGGTTA
AGTCCCGCAACGAGCGAACCCATTCTCAGTAGCCAGCATATAAGGTGGCACTCTGGAGAGACTGCCAGGGAGA
ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCATTGGCCAGGGCTACACCGTGTACAATGGCGTA
AACAAAGGAAAGCGAGAGGGTGACCTGGAGCGAATCCAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG
ACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC
GCCCGTACACCATGGGAGTCAGTAACGCCGAAAGCCAGTGACCCAAACCTTAGAGGAGGGAGCTGCGAAGGGGGAGA
CGGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 135

```

| PROKKA_02310 16S ribosomal RNA gene
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTTT
CAGAATCTCGGAGGAAGAGGACAGTGACTGAGCGCGGACGGGTGAGTAACCGTGGCAACCTGCCCTACACAGG
GGGATAACAGTTAGAAATGACTGCTAACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTGAAAAACTCCGGTG
GTATGAGATGGACCCCGCTCTGATTAGGTAGTGGTGGGTAAGGCCCTACCAAGCGCACGATCAGTAGCCGACCTG
AGAGGGTGACCGGCCACATTGGGACTGAGACACGCCAACCTACGGGAGGCAGCAGTGGGAATATTGCACAA
TGGGGAAACCTGATGCGACGCCGCTGAAGGAAGAAGTATTCGGTATGTAACCTCTATCAGCAGGGAAAGA
AGATGACGGTACCTGAGTAAGAAGCACCGCTAAATACGTGCCAGCAGCGCGGTAAACGTATGGTCAAGCGTTA
TCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAGGCAAGTCTGGAGTGGAAACCCAGGGCTAACCTGGGA
CTGCTTGGAAACTGCAGATCTGGAGTGCCTGGAGGTAAGCGGAACTCCTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACCGTAGGCGAACGGCGCTACTGGACGGTAGCTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACA
GGATTAGATAACCTGGTAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA
CAAACGCAATAAGTAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA
GCCGTGGAGCATGTTGTTAACCGCAACCGCAAGAACCTTACCTGGTCTGACATCGGATGACGGGAGTAGA
ATGTCGCCGTCCCTCGGGCATCCGAGACAGGTGGTGCATGGTTGTCGTCACTCGTGTGAGATGTTGGTTA
AGTCCCGCAACGAGCGAACCCATTCTCAGTAGCCAGCATATAAGGTGGCACTCTGGAGAGACTGCCAGGGAGA
ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCATTGGCCAGGGCTACACCGTGTACAATGCCGTAA
AACAAAGGAAAGCGAGAGGGTGACCTGAAGCGAATCCAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG
ACTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC
GCCCGTACACCATGGGAGTCAGTAACGCCGAAAGCCAGTGACCCAAACCTTAGAGGAGGGAGCTGCGAAGGGGGAGA
CGGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 136

```

| PROKKA_02993 16S ribosomal RNA gene
TACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCTGCCTAACACATGCAAGTCGAGCGAAGCGCTGTTT
CAGAATCTCGGAGGAAGAGGACAGTGACTGAGCGCGGACGGGTGAGTAACCGTGGCAACCTGCCCTACACAGG
GGGATAACAGTTAGAAATGACTGCTAACCGCATAAGCGCACAGGACCGCATGGTGTAGTGTGAAAAACTCCGGTG
GTATGAGATGGACCCCGCTCTGATTAGGTAGTGGTGGGTAAGGCCCTACCAAGCGCACGATCAGTAGCCGACCTG
AGAGGGTGACCGGCCACATTGGGACTGAGACACGCCAACCTACGGGAGGCAGCAGTGGGAATATTGCACAA
TGGGGAAACCTGATGCGACGCCGCTGAAGGAAGAAGTATTCGGTATGTAACCTCTATCAGCAGGGAAAGA
AGATGACGGTACCTGAGTAAGAAGCACCGCTAAATACGTGCCAGCAGCGCGGTAAACGTATGGTCAAGCGTTA
TCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAGGCAAGTCTGGAGTGGAAACCCAGGGCTAACCTGGGA
CTGCTTGGAAACTGCAGATCTGGAGTGCCTGGAGGAGGTAAAGCGGAACTCCTAGTGTAGCGGTGAAATGCGTAGATAT
TAGGAGGAACACCGTAGGCGAACGGCGCTACTGGACGGTAGCTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACA
GGATTAGATAACCTGGTAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA
CAAACGCAATAAGTAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA

```

-continued

GCGGTGGAGCATGTGTTAATTGAAGCAACCGAAGAACCTTACCTGGTCTTGACATCCGGATGACGGGGAGTA  
 ATGTCGCCGTCCCTCGGGCATCCGAGACAGGTGGTCATGGTGTGTCAGCTCGTGTGAGATGTTGGTTA  
 AGTCCCGAACAGCGCAACCCCTATCTCAGTAGCCAGCATATAAGGTGGCACTCTGGAGAGACTGCCAGGGAGA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGCCAGGGTACACAGTGCTACAATGGCGTA  
 AACAAAGGGAAGCGAGAGGGTGACCTGGAGCGAATCCAAAATAACGTCTCAGTCGGATTGTAGTCTGCAACTCG  
 ACTACATGAAGCTGGATCGCTAGTAACTCGGGATCAGCATGCCCGGTGAATACGTTCCCGGTCTGTACACACC  
 GCCCGTCACACCATGGGAGTCAGTAACGCCGAAGCCAGTGACCCAACCTTAGAGGAGGGAGCTGCGAAGGGGGAG  
 CGGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCTTT

SEQ ID NO: 137

| PROKKA\_00436 16S ribosomal RNA gene  
 ATGAGAGTTGATCCTAGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAACGAAGCAATTAAACG  
 GAAGTTTCGGATGGAAGTTGAATTGACTGAGTGGCGACGGGTGAGTAACCGTGGTAACCTGCTTGTACTGGG  
 GGACAAACAGTTAGAAATGACTGCTAAACCGCATAAGCGCACAGTATCGCATGATAACAGTGTGAAAAACTCCGGTGG  
 TACAAGATGGACCCCGTCTGATTAGCTAGTTGGTAAGGTACGGCTTACCAAGGCGACGATCAGTAGCCGACCTGA  
 GAGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGTGGGAATATTGACAAT  
 GGGCGAAAGCCTGATGCAGCGACGCCGTGAGTGAAGAAGTATTTCGGTATGTAAGGCTTACAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAAGCCCCGCTAACTACGTGCCAGCCGGTAATACGTAGGGGGCAAGCGTTAT  
 CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAGGCTCAACTGCGGAC  
 TGCTTGGAAACTGTTAACTGGAGTGTGGAGGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTGGCAAGGGACTTACGGACGATAACTAGGTGTTGGGAGCAAAGCTTCCGGTGCCTCG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTTGGGAGCAAAGCTTCCGGTGCCTCG  
 AAACGCAGTAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGACAAG  
 CGGTGGAGCATGTGTTAATTGAAGCAACCGAAGAACCTTACCAAGGTCTGACATCGATCCGACGGGGAGTAA  
 CGTCCCCCTCCCTCGGGCGAGAGACAGGTGGTCATGGTTGTCGTACGCTGTGAGATGTTGGTTAA  
 GTCCCCCAACGCGAACCCCTATTCTAAGTAGCCAGGGTTGCCCGGAACTCTGGAGACTGCCAGGGATAA  
 CCTGGAGGAAGGTGGGATGACGTAAACATCATGCCCTTATGATCTGGCTACACACGTGCTACAATGGCGTAA  
 ACAAAAGAGAAGCAAGACCGCAGGTGGAGCAAATCTCAAAAATAACGTCTCAGTCGGACTGCAGGCTGCAACTCG  
 CTGCACAGCTGGAACTGCTAGTAATCGCAATCAGAATGTCGGTGAATACGTTCCGGTCTGTACACACCG  
 CCCGTACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCAACCGCAAGGAGGGAGCTGCCGAAGGGGGACC  
 GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCTTT

SEQ ID NO: 138

| PROKKA\_00685 16S ribosomal RNA gene  
 ATGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTGCCTAACACATGCAAGTCGAACGAAGCGATTAACG  
 GAAATTTTCGGATGGAAGTTGAATTGACTGAGTGGGGACGGGTGAGTAACCGTGGTAACCTGCTTGTACTGGG  
 GGACAAACAGTTAGAAATGACTGCTAAACCGCATAAGCGCACAGTATCGCATGATAACAGTGTGAAAAACTCCGGTGG  
 TACAAGATGGACCCCGTCTGATTAGCTAGTTGGTAAGGTACGGCTTACCAAGGCGACGATCAGTAGCCGACCTGA  
 GAGGGTACCGGCCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATATTGACAAT  
 GGGCGAAAGCCTGATGCAGCGACGCCGTGAGTGAAGAAGTATTTCGGTATGTAAGGCTTACAGCAGGGAAAGAA  
 AATGACGGTACCTGACTAAGAAGCCCCGCTAACTACGTGCCAGCCGGTAATACGTAGGGGGCAAGCGTTAT  
 CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCTGAAGTGAAGGCTCAACTGCGGAC  
 TGCTTGGAAACTGTTAACTGGAGTGTGGAGGGTAAGTGGAAATTCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCGTGGCAAGGGACTTACGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAG  
 GATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTTGGGAGCAAAGCTTCCGGTGCCTCG

-continued

```

AAACGCAGTAAGTATTCCACCTGGGAGTACGTTCCAAGAACACTCAAAGGAATTGACGGGACCGCACAAG
CGGTGGAGCATGTGGTTAATTCAAGCAACCGAAGAACCTTACAGGTCTGACATCGATCGACGGGGAGTAA
CGTCCCCCTCCCTCGGGCGAGAGACAGGTGGTGCATGGTCGTACAGCTGTGAGATGTTGGTTAA
GTCGGCAACGAGCGAACCCCTATTCTAAGTAGCGCAGCGTTGCCCGGAACTCTGGAGACTGCCAGGGATAA
CCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGATCTGGCTACACACGTGCTACAATGGCGTAA
ACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTAAAAAAACGTCTCAGTCAGGACTGCCAGGCTGCAACTCGC
CTGCACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTTGTACACACCG
CCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGCAAGGGAGGAGCTGCCGAAGGCAGGACC
GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 139

```

| PROKKA_01171 16S ribosomal RNA gene
ATGAGAGTTGATCCCTGGCTCAGGATGAACGCTGGCGGTGCCAACACATGCAAGTCGAACGAAGCGATTAAACG
GAAGTTTCGGATGGAAGTTGAAATTGACTGAGTAGGCGACGGGTGAGTAACCGTGGTAACCTGCCTTGACTGGG
GGACAAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGATGATAACAGTGTGAAAAACTCCGGTGG
TACAAGATGGACCCCGCTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCAGCAGTACAGTCAGGCCACCTGA
GAGGGTGACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGCAACAAT
GGCGAAAGCCTGATGCAGCGACGCCGTGAGTGAAGAAGTATTCCGGTATGTAAGCTCTATCAGCAGGGAAAGAA
AATGACGGTACCTGACTAAGAAGCCCCGCTAAACTACGTGCCAGCGCGTAAACGTAGGGGCAAGCGTTAT
CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCAGTGAAGTGAAGGCCCGCGCTCAACTCCGGGAC
TGCTTGGAAACTGTTAACTGGAGTAGCAGGTTAGGCTGAAAGCTTCCGGTCAACTGCCGAC
AGGAGGAACACCACTGGCGAAGGGGACTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAG
GATTAGATAACCTGGTAGTCCACGCCGAAACGATGAATACTAGGTGTTGGGAGCAAAGCTTCCGGTCCGTGCG
AAACGCAGTAAGTATTCCACCTGGGAGTACGTTCGCAAGAACACTCAAAGGAATTGACGGGACCGCACAAG
CGGTGGAGCATGTGGTTAATTCAAGCAACGCGAACCTTACAGGTCTGACATCGATCGACGGGGAGTAA
CGTCCCCCTCCCTCGGGCGAGAGACACGGTGGCATGGTGTGTCAGCTCGTGTGAGATGTTGGTTAA
GTCGGCAACGAGCGAACCCCTATTCTAAGTAGCGCAGCGTTGCCCGGAACTCTGGAGACTGCCAGGGATAA
CCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGATCTGGCTACACACGTGCTACAATGGCGTAA
ACAAAGAGAAGCAAGACCGCGAGGTGGAGCAAATCTAAAAAAACGTCTCAGTCAGGACTGCCAGGCTGCAACTCGC
CTGCACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTTGTACACACCG
CCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGCAAGGGAGGAGCTGCCGAAGGCAGGACC
GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 140

```

| PROKKA_05969 16S ribosomal RNA gene
ATGAGAGTTGATCCCTGGCTCAGGATGAACGCTGGCGGTGCCAACACATGCAAGTCGAACGAAGCGATTAAACG
GAAGTTTCGGATGGAAGTTGGATTGACTGAGTAGGCGACGGGTGAGTAACCGTGGTAACCTGCCTTGACTGGG
GGACAAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTATCGATGATAACAGTGTGAAAAACTCCGGTGG
TACAAGATGGACCCCGCTCTGATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCAGCAGTACAGTCAGGCCACCTGA
GAGGGTGACCGGCCACATTGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTGGGAATTGCAACAAT
GGCGAAAGCCTGATGCAGCGACGCCGTGAGTGAAGAAGTATTCCGGTATGTAAGCTCTATCAGCAGGGAAAGAA
AATGACGGTACCTGACTAAGAAGCCCCGCTAAACTACGTGCCAGCGCGGTAAACGTAGGGGCAAGCGTTAT
CCGGATTTACTGGGTGAAAGGGAGCGTAGACGGTAAAGCAAGTCAGTGAAGTGAAGGCCCGCGCTCAACTCCGGGAC
TGCTTGGAAACTGTTAACTGGAGTAGTCGAGAGGTAAAGTGAATTCTAGTGTAGCGGTGAAATGCGTAGATATT
AGGAGGAACACCACTGGCGAAGGGGACTTACTGGACGATAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAG

```

-continued

```

GATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAGCAAAGCTTCGGTGCCTCGC
AAACGCAGTAAGTATTCCACCTGGGGAGTACGTTCGAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG
CGGTGGAGCATGTGGTTAATCGAACGCAAGCGAACCTTACCAAGGTCTTGACATCGATCCGACGGGGAGTAA
CGTCCCCCTCCCTCGGGCGGAGAACAGGTGGTGATGGTTGCTCAGCTCGTGTGAGATGTTGGTTAA
GTCCCGCAACGAGCGAACCTTATTCTAAGTAGCCAGCGGTCGGCCGGAACTCTGGGAGACTGCCAGGGATAA
CCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGATCTGGCTACACACGTGCTACAATGGCTAA
ACAAAGAGAAGCAAGACCGCAGGGTGGAGCAAATCTCAAAAATAACGTCTCAGTTGGACTGCAGGCTGCAACTCGC
CTGCACAGGCTGGAATCGTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACCG
CCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCACCGCAAGGAGGGAGCTGCCGAAGGCGGGACC
GATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 141

```

| PROKKA_00279 16S ribosomal RNA gene
ATCAGAGAGTTGATCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCAGCTTAAG
TGGATCTCTCGGATTGAAACTATTGACTGAGCGCGGACGGGTGAGTAACGCGTGGTAACCTGCCTCATACAG
GGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGCTGGTGTGAAAAACTCCGGT
GGTATGAGATGGACCCCGCTGATTAGCTAGTTGGAGGGTAACGGCCACCAAGGCAGCAGTCAGTAGCCGGCT
GAGGGGTGAACGGCACATTGGGACTGAGACACGGCCAGACTCTACGGAGGAGCAGTGGGAATTGGCACA
ATGGGGAAACCTGTGAGCAGCGACGCCGCTGAAGGAAGAAGTATCTCGGTATGTAACCTCTACAGCAGGGAAAG
AAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTT
ATCGGATTACTGGGTAAAGGGAGCGTAGCGGAAGCAGAAGCAGTCGAGTGTGAAAGGCTGGGCTAACCCAGG
ACTGCATTGGAAACTGTTCTAGAGTGCGGAGAGGTAAAGCGGAATTCTAGTGAGCGTGAATGCGTAGATA
TTAGGAGGAACCCAGTGGCGAAGGCGCTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAAC
AGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGTGGCAAAGCATTGGTCCGCA
GCAAACGCAATAAGTATTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA
ACCGTGGAGCATGTGGTTAATCGAACGCAACCGAAGAACCTTACCAAGTCTTGACATCCCTGACCGGCCCCT
ACGGGGCCTCCCTCGGGAGAGGAGACAGGTGGTGATGGTGTGCTCAGCTCGTGTGAGATGTTGGTT
AAGTCCCGAACGAGCGAACCCCTACCTTAGTGAGCCAGCAGGTGAAGCTGGCACTCTAGGGAGACTGCCGGGA
TAACCCGGAGGAAGGCGGGACGACGTCAAATCATGCCCCCTATGATTGGGCTACACACGTGCTACAATGGCG
TAAACAAAGGGAGCGAGACAGCGATGTTGAGCAAATCCAAAATAACGTCCAGTCGGACTGCAGTCAGCTGCAACT
CGACTGCACGAAGCTGGAATCGTAGTAATCGCAATCAGAATGTCGCGGTAAACGTTCCGGGTCTTGTACACA
CCGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCACCTTTAGGAGGGAGCTGCCGAAGGCGG
GACCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 142

```

| PROKKA_01221 16S ribosomal RNA gene
ATCAGAGAGTTGATCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAGCGAAGCAGCTTAAG
TGGATCTCTCGGATTGAAACTATTGACTGAGCGCGGACGGGTGAGTAACGCGTGGTAACCTGCCTCATACAG
GGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGCTGGTGTGAAAAACTCCGGT
GGTATGAGATGGACCCCGCTGATTAGCTAGTTGGAGGGTAACGGCCACCAAGGCAGCAGTCAGTAGCCGGCT
GAGGGGTGAACGGCACATTGGGACTGAGACACGGCCAGACTCTACGGAGGAGCAGCAGTGGGAATTGGCACA
ATGGGGAAACCTGTGAGCAGCGACGCCGCTGAAGGAAGAAGTATCTCGGTATGTAACCTCTACAGCAGGGAAAG
AAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTT
ATCGGATTACTGGGTAAAGGGAGCGTAGCGGAAGAGCAAGTCAGTGAAAGGCTGGGCTAACCCAGG
ACTGCATTGGAAACTGTTCTAGAGTGCGGAGAGGTAAAGTGGATTCTAGTGAGCGGTAAATGCGTAGATA

```

-continued

TTAGGAGGAACACCAGTGGCGAAGGCAGCTACTGGACCGTAACGTGAGGCTGAAAGCGTGGGAGCAAAC  
 AGGATTAGATAACCCCTGGTAGTCACGCCGTAAACGATGAATACAGGTGTCGGTGCAAGGCCATTGGTGGC  
 GCAAAACGCAATAAGTATTCCACCTGGGACTACGTCGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA  
 AGCGGTGGAGCATGTGGTTAACCGAAGCACGCAAGAACCTTACCAAGTCTTGACATCCCTGACCGGGCG  
 AACGGGGCCTTCCCTCGGGCAGAGGAGACAGGTGGTGATGGTGTGTCAGCTCGTGTGAGATGTTGGTT  
 AAGTCCCACGAGCGCAACCCCTATCCTAGTAGCCAGCAGGTGAAGCTGGCACTCTAGGGAGACTGCCGGGA  
 TAACCCGGAGGAAGCGGGACAGCGATGTTGAGCAAATCCAAAAAACGTCACGGTCACTGGACTCGAGTCTGCA  
 TAAACAAAGGAAGCGGGACAGCGATGTTGAGCAAATCCAAAAAACGTCACGGTCACTGGACTCGAGTCTGCA  
 CGACTGCACGAAGCTGGAATCGTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTGTACACA  
 CGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGC  
 GACCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCCGTGGATCACCTCCTT

SEQ ID NO: 143

| PROKKA\_02318 16S ribosomal RNA gene  
 ATCAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTAAACACATGCAAGTCGAGCGAAGCCTTAAG  
 TGGATCTCTGGATTGAGCTTATTTGACTGAGCGGGACGGGTGAGTAACCGTGCGGTAACCTGCCTCATACAG  
 GGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACCTCGGT  
 GGTATGAGATGGACCCCGCTGATGTTAGCTAGTGGAGGGTAACGGCCACCAAGCGACGATCAGTAGCCGGCT  
 GAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGAGGCAGCAGTGGGAATTGCA  
 ATGGGGAAACCCCTGATGCAGCGACGCCCGTGAAGGAAGAAGTATCTGGTATGTAACCTCTATCAGCAGGGAAAG  
 AAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAACGTAGGGGCAAGCGT  
 ATCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCTGGGTTAACCCAGG  
 ACTGCATTGGAAACTGTTCTAGAGTGCAGGAGAGGTAAGCGGAATTCTAGTGTAGCGGTGAAATGCGTAGATA  
 TTAGGAGGAACACCAGTGGCGAAGGCAGCTACTGGACGGTAACGTGAGGCTGAAAGCGTGGGAGCAAAC  
 AGGATTAGATAACCCCTGGTAGTCACGCCGTAAACGATGAATACAGGTGTCGGTGCAAGGCCATTGGTGGC  
 GCAAAACCAATAAGTATTCCACCTGGGAGTACGTTCCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACA  
 AGCGGTGGAGCATGTGGTTAACCGAAGCACGCAAGAACCTTACCAAGTCTTGACATCCCTGACCGGGCG  
 AACGGGGCCTTCCCTCGGGCAGAGGAGACAGGTGGTGATGGTGTGTCAGCTCGTGTGAGATGTTGGTT  
 AAGTCCCACGAGCGCAACCCCTATCCTAGTAGCCAGCAGGTGAAGCTGGCACTCTAGGGAGACTGCCGGGA  
 TAACCCGGAGGAAGCGGGACAGCGATGTTGAGCAAATCCAAAAAACGTCACGGTCACTGGACTCGAGTCTGCA  
 TAAACAAAGGAAGCGAGACAGCGATGTTGAGCAAATCCAAAAAACGTCACGGTCACTGGACTCGAGTCTGCA  
 CGACTGCACGAAGCTGGAATCGTAGTAATCGCAATCAGAATGTCGCGGTGAATACGTTCCGGTCTGTACACA  
 CGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCTTACAGGAGGGAGCTGCCGAAGGC  
 GACCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCCGTGGATCACCTCCTT

SEQ ID NO: 144

| PROKKA\_02336 16S ribosomal RNA gene  
 ATCAGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTAAACACATGCAAGTCGAGCGAAGCCTTAAG  
 CGGATCTCTGGATTGAAACTTATTTGACTGAGCGGGACGGGTGAGTAACCGTGCGGTAACCTGCCTCATACAG  
 GGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGCGCACAGGACCGCATGGTCTGGTGTGAAAACCTCGGT  
 GGTATGAGATGGACCCCGCTGATGTTAGCTAGTGGAGGGTAACGGCCACCAAGCGACGATCAGTAGCCGGCT  
 GAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGAGGCAGCAGTGGGAATTGCA  
 ATGGGGAAACCCCTGATGCAGCGACGCCCGTGAAGGAAGAAGTATCTGGTATGTAACCTCTATCAGCAGGGAAAG  
 AAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAACGTAGGGGCAAGCGT  
 ATCCGGATTTACTGGGTGAAAGGGAGCGTAGACGGAAGAGCAAGTCTGATGTGAAAGGCTGGGTTAACCCAGG

-continued

```

ACTGCATTGGAAACTGTTTCTAGAGTGCGGAGAGGTAAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATA
TTAGGAGGAACACCAGTGGCGAAGGCGGCTACTGGACGGTAACTGACGTTGAGGCTGAAAGCGTGGGAGCAAAC
AGGATTAGATAACCCCTGGTAGTCACGCCGTAACAGATGAATACTAGGTGTGGTGGCAAAGCCATTGGTGGCACA
GCAAACGCAATAAGTATTCCACCTGGGAGTACGTTGCAAGAAATGAAACTCAAAGGAATTGACGGGACCCGACA
AGCGGTGGAGCATGTGGTTAATTGAAGCAACGCAAGAACCTACCAAGTCTGACATCCCTGTGACCGGGCGT
AACGGGCGCTTCCCTCGGGGAGAGGAGCACGGTGGCATGGTGTGAGCTCGTGTGAGATGTTGGTT
AAGTCCCACGAGCGCAACCCCTACCTTAGTAGCCAGCAGGTGAAGCTGGCACTCTAGGGAGACTGCCGGGA
TAACCCGGAGGAAGGGGGAGACGTCAAATCATCATGCCCTTATGATTTGGCTACACACGTGCTACAATGGCG
TAAACAAAGGAAGCGAGACAGCGATGTTGAGCAAATCCAAAATAACGTCCCAGTCGGACTGCAGTCTGCAACT
CGACTGCACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTAAATACGTTCCGGTCTGTACACA
CGCCCGTCACACCATGGGAGTCAGTAACGCCAGTCAGTGAACCAACCTACAGGAGGGAGCTGCCAGGGCG
GACCGATAACTGGGTGAAGTCGAAACAGGTAGCCGTATCGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 145

```

| PROKKA_04947 16S ribosomal RNA gene
ATCAGAGAGTTGATCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCACTTAAG
TGGATCTCTGGATTGAAACTATTGACTGAGCGGGACGGTGAGTAACCGTGGTAACCTGCCTACATACAG
GGGATAACAGTTAGAAATGGCTGCTAATACCGATAAGCGCACAGGACCCATGGTCTGGTGTGAAAAACTCCGGT
GGTATGAGATGGACCCCGCTGCTGATTAGCTAGTTGGAGGGTAACGGCCACCAAGGCAGCAGTCAGTAGCCGGCT
GAGAGGGTGAACGGCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGACACA
ATGGGGAAACCTGATGCAGCGACGCCGTGAAGGAAGAAGTATCTCGGTATGTAACCTTATCAGCAGGGAAAG
AAAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTTAGGGGCAAGCGTT
ATCCGGATTACTGGGTAAAGGGAGCGTAGACGGAAGGCAAGTCTGATGTGAAAGGTGGGTTAACCCAGG
ACTGCATTGGAAACTGTTTCTAGAGTGCGGAGAGGTAAAGCGGAATTCTAGTGTAGCGGTGAAATCGTAGATA
TTAGGAGGAACATCAGTGGCGAAGGCGGCTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAAC
AGGATTAGATAACCCCTGGTAGTCACGCCGTAACAGATGAATACTAGGTGTGGTGGCAAAGCCATTGGTGGCACA
GCAAACGCAATAAGTATTCCACCTGGGAGTACGTTGCAAGAAATGAAACTCAAAGGAATTGACGGGACCCGACA
AGCGGTGGAGCATGTGGTTAATTGAAGCAACGCAAGAACCTACCAAGTCTGACATCCCTGTGACCGGGCGT
AACGGGCGCTTCCCTCGGGGAGAGGAGCACGGTGGCATGGTGTGAGCTCGTGTGAGATGTTGGTT
AAGTCCCACGAGCGCAACCCCTACCTTAGTAGCCAGCAGGTGAAGCTGGCACTCTAGGGAGACTGCCGGGA
TAACCCGGAGGAAGGGGGAGACGTCAAATCATCATGCCCTTATGATTTGGCTACACACGTGCTACAATGGCG
TAAACAAAGGAAGCGAGACAGCGATGTTGAGCAAATCCAAAATAACGTCTAGTCGGACTGCAGTCTGCAACT
CGACTGCACGAAGCTGGAATCGCTAGTAATCGCAATCAGAATGTCGCGGTAAATACGTTCCGGTCTGTACACA
CGCCCGTCACACCATGGGAGTCAGTAACGCCAGTCAGTGAACCAACCTACAGGAGGGAGCTGCCAGGGCG
GACCGATAACTGGGTGAAGTCGAAACAGGTAGCCGTATCGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 146

```

| PROKKA_00208 16S ribosomal RNA gene
AACGAGAGTTGATCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCACTTAAGT
TTGATTCTCGGATGAAGACTTTGTGACTGAGCGGGACGGTGAGTAACCGTGGTAACCTGCCTACATACAGG
GGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCAAGGTACGGCATGGTACAGTGGTAAAAACTCCGGT
GTATGAGATGGACCCCGCTGATTAGTAGTTGGTGGGTAACGGCTACCAAGCCGACGATCAGTAGCCGACCTG
AGAGGGTACGGGCCACATTGGGAGTCAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGACACA
TGGAGGAAACTCTGATGCAGCGACGCCGTGAAGGTGAAGTATTCGGTATGTAACCTTATCAGCAGGGAAAGA
AAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTTAGGGGCAAGCGTTA

```

-continued

TCCGGATTTACTGGGTAAAGGGAGCGTAGACGGCACGGCAAGCAGATGTGAAAGCCGGGCTCAACCCGGGA  
 CTGCATTGGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGATTCCCTAGTGTAGCGGTGAAATGCGTAGATAT  
 TAGGAGGAACACCAGTGGCGAAGGGCGTTCTGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACA  
 GGATTAGATAACCTGGTAGTCCACGCCAACGATGACTGCTAGGTGTGGCAAAGCATTGGTGCAG  
 CTAACGCAATAAGCAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA  
 GCGGTGGAGCATGTGTTAATCGAAGCAACGCGAAGAACCTTACCTGATCTTGACATCCCAGTACCGCTTCGTA  
 ATGGAAGTTTCTCGAACATCGTGACAGGTGGTCATGGTTGTCAGCTCGTGTGAGATGTTGGGTTA  
 AGTCCCGCAACGAGCGAACCCCTATCTTCAGTAGCCAGCAGGTTAACGCTGGCACTCTGGAGAGACTGCCAGGGAT  
 AACCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCAGGGCTACACACGTGCTACAATGGCGT  
 AAACAAAGAGAAGCGAACCTCGGAGGGTAAGCAAATCTAAAAAAACGTCTCAGTTGGATTGAGTCTGCAACTC  
 GACTACATGAAGCTGGAATCGTAGTAATCGCAGATCAGAAATGCTCGGTAAATCGTCCGGTCTTGACACAC  
 CGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGA  
 CCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTACGGAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 147

| PROKKA\_00340 16S ribosomal RNA gene  
 AACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCACTTTGGA  
 AAGATTCTCGGATGATTCCTTGTGACTGAGCGCGGACGGGTGAGTAACCGCTGGTAAACCTGCCCTCATACAGG  
 GGGATAACAGTTAGAAATGACTGCTAATACCGATAAGACCACCGTACCGTACAGTGGTAAAAACTCCGGTG  
 GTATGAGATGGACCCCGCTCTGATTAGTAGTTGGTGGGTAACGCCCTACCAAGCCGACGATCAGTAGCCGACCTG  
 AGAGGGTACCGGCCACATTGGACTGAGCACCGCCAGACTCTACGGGAGGCAGCAGTGGGAATTGCACAA  
 TGGAGGAAACTCTGATGCAGCAGCCGCGTGAAGGATGAAGTATTCGGTATGTAACCTATCAGCAGGGAAAGA  
 AAATGACGGTACCTGACTAAGAAGCCCGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGCAAGCGTTA  
 TCCGGATTTACTGGGTAAAGGGAGCGTAGACGGCACGGCAAGCAGATGTGAAAGCCGGCTCAACCCGGAC  
 TGCATTGGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGGATTCCCTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCAGTGGCGAAGGGCGTTGCTGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACAG  
 GATTAGATACCTGGTAGTCCACGCGTAAACGATGACTGCTAGGTGTCGGTGGCAAAGCATTGGTGCAGCAGC  
 TAACGCAATAAGCAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGTGGTTAATCGAAGCAACGCGAAGAACCTACCTGATCTGACATCCCAGTACCGCTTCGTA  
 TGGAGCTTTCTCGGAACATCGGTACAGGTGGTCATGGTTGTCGTAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGAACAGCGAACCCCTATCTCAGTAGCCAGCAGGTTAACGCTGGGACTCTGGAGAGACTGCCAGGGATA  
 ACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCTTATGACCAGGGCTACACACGTGCTACAATGGCGTA  
 AACAAAGAGAAGCGAACCTCGCGAGGGTAAGCAAATCTAAAAAAACGTCTCAGTTGGATTGAGTCTGCAACTCG  
 ACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAGAAATGCTCGGTAAATACGTTCCGGTCTTGACACACC  
 GCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGAC  
 CGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTACGGAGGTGCGGCTGGATCACCTCCTT

SEQ ID NO: 148

| PROKKA\_01031 16S ribosomal RNA gene  
 AACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGGTGCTTAACACATGCAAGTCGAGCGAAGCACTTAAGT  
 TTGATTCTCGGATGAGACTTTGTGACTGAGCGCGGACGGGTGAGTAACCGCTGGTAAACCTGCCCTCATACAGG  
 GGGATAACAGTTAGAAATGACTGCTAATACCGATAAGACCACCGTACCGTACGGTACAGTGGTAAAAACTCCGGTG  
 GTATGAGATGGACCCCGCTCTGATTAGTAGTTGGTGGGTAACGCCCTACCAAGCCGACGATCAGTAGCCGACCTG  
 AGAGGGTACCGGCCACATTGGGACTGAGCACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGCACAA  
 TGGAGGAAACTCTGATGCAGCAGCCGCGTGAAGGATGAAGTATTCGGTATGTAACCTATCAGCAGGGAAAGA

-continued

```

AAATGACGGTACCTGACTAAGAAGCCCCGGTAACACTACGTGCCAGCAGCCGGTAATACGTAGGGGGCAAGCGTTA
TCCGGATTACTGGGTAAAGGGAGCGTAGACGGCACGGCAAGGCCAGATGTGAAAGCCGGGCTCAACCCGGGA
CTGCATTTGGAACTGCTGAGCTAGCTGGAGGGCAAGTGGAAATTCTAGTGACTGCGGTAAATGCGTAGATAT
TAGGAGGAACACCAGTGGCGAAGGGCGCTTGTGACGATGACTGACGTTGAGGTCGAAAGCGTGGGGAGCAAACA
GGATTAGATAACCTGGTAGTCCACGCGTAAACGATGACTGCTAGGTGTCGGGTGGCAAAGCATTGGTGCAGCAG
CTAACGCAATAAGCAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA
GCGGTGGAGCATGTGGTTAATTGAAGCAACGCGAAGAACCTTACCTGATCTGACATCCCAGTACCGCTCGTA
ATGGAAGCTTTCTCGAACATCGGTACAGGTGGTAGTGGCTAGCTGGTGTGCTAGCTCGTGTGAGATGTTGGGTTA
AGTCCCAGAGCGCAACCCCTATCTTCAGTAGCCAGCAGGTTAACGCTGGGACTCTGGAGACTGCCAGGGAT
AACCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCAGGGTACACACGTGCTACAATGGCGT
AAACAAAGAGAAGCGAACCTGGGAGTACGTTGCAAGAACCTTACCTGATCTGACATCCCAGTACCGCTCGTA
GACTACATGAAGCTGGAATCGTAGTAATCGCAGATCAGAATGCTGCGGTGAAATGTTCCGGGCTTGTACACAC
CGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGA
CCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 149

```

| PROKKA_01840 16S ribosomal RNA gene
| AACGAGAGTTGATCTGGCTCAGGATGAAACGCTGGCGCGTGTCTTAACACATGCAAGTCGAGCGAAGCAGCTTGGA
| AAGATTCTCGGATGATTTCTTGTGACTGAGCGCGGAGCGGTGAGTAACGCGTGGGTAACCTGCCTCATACAGG
| GGGATAACAGTTAGAAATGACTGCTAAACCGATAAGACCACCGTACCGCATGGTACAGTGGTAAAAACTCCGGTG
| GTATGAGATGGACCCCGCTCTGATTAGTAGTGGTGGGTAACGCCCTACCAAGCGCAGCATCGTAGCCGACCTG
| AGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGCACAA
| TTGAGGAAACTCTGATGCGACGCGCCTGAGGGATGAAGTATTTCGGTATGAACTCTATCAGCAGGGAGA
| AAATGACGGTACCTGACTAAGAAGCCCCGGTAACACTACGTGCCAGCAGCCGGTAATACGTAGGGGGCAAGCGTTA
| TCCGGATTACTGGGTAAAGGGAGCGTAGACGGCACGGCAAGGCCAGATGTGAAAGCCGGGCTCAACCCGGGA
| CTGCATTTGGAACTGCTGAGCTAGAGTGTGGAGAGGCCAGTGGAAATTCTAGTGACTGCGGTGAAATGCCAGATAT
| TAGGAGGAACACCAGTGGCGAAGGGCGCTTGTGACGATGACTGACGTTGAGGTCGAAAGCGTGGGGAGCAAACA
| GGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGTCGGGTGGCAAAGCATTGGTGCAGCAG
| CTAACGCAATAAGCAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA
| GCGGTGGAGCATGTGGTTAATTGAAGCAACGCGAAGAACCTTACCTGATCTGACATCCCAGTACCGCTCGTA
| ATGGAAGCTTTCTCGAACATCGGTACAGGTGGTAGTGGCTAGCTGGTGTGCTAGCTCGTGTGAGATGTTGGGTTA
| AGTCCCAGAGCGCAACCCCTATCTTCAGTAGCCAGGTTAACGCTGGGACTCTGGAGAGACTGCCAGGGAT
| AACCTGGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGACCAGGGTACACACGTGCTACAATGGCGT
| AAACAAAGAGAAGCGAACCTGGGAGGTAAGCAAATCTAAAAAAACGTCTCAGTGGATGTTGAGCTCGAACCTC
| GACTACATGAAGCTGGAATCGTAGTAATCGCAGATCAGAATGCTGCGGTGAAATGTTCCGGGCTTGTACACAC
| CGCCCGTCACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGTGGGA
| CCGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT

```

SEQ ID NO: 150

```

| PROKKA_02944 16S ribosomal RNA gene
| AACGAGAGTTGATCTGGCTCAGGATGAAACGCTGGCGCGTGTCTTAACACATGCAAGTCGAGCGAAGCGCTTAAGT
| TTGATTCTCGGATGAAGACTTTGTGACTGAGCGCGGAGGGTGAGTAACGCGTGGGTAACCTGCCTCATACAGG
| GGGATAACAGTTAGAAATGACTGCTAAACCGATAAGACCACCGTACCGCATGGTACAGTGGTAAAAACTCCGGTG
| GTATGAGATGGACCCCGCTCTGATTAGTAGTGGTGGGTAACGCCCTACCAAGCGCAGCATCGTAGCCGACCTG
| AGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATTGCACAA

```

-continued

TGGAGGAAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCGGTATGTAAACTCTATCAGCAGGGAAAGA  
 AAATGACGGTACCTGACTAAGAACGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTA  
 TCCGGATTACTGGGTGAAAGGGACGCTAGACCGCACGGCAAGCAGATGTGAAAGGCCGGCTCAACCCGGAC  
 TGCATTGGAACTGCTGAGCTAGAGTGTGGAGAGGCAAGTGAATTCTTAGTGTAGCGGTGAAATGCGTAGATATT  
 AGGAGGAACACCAGTGGCGAAGGCGCTTGCTGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAG  
 GATTAGATAACCTGGTAGTCCACGCGTAAACGATGACTGCTAGGTGTCGGTGGCAAAGCATTGGTGCCGAGC  
 TAACGCAATAAGCAGTCCACCTGGGAGTACGTCGAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAAG  
 CGGTGGAGCATGTGGTTAATCGAAGCAACCGAAGAACCTACCTGATCTGACATCCGATGACCGTTCGTAA  
 TGGAAAGTTTCTCGGAACATCGTGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTGGTTAA  
 GTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCAGGTTAACGCTGGGACTCTGGAGAGACTGCCAGGGATA  
 ACTCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGACCAGGGTACACCGTGTACATGGCGTA  
 AACAAAGAGAGGCAAACTCGCAGGGTAAGCAAATCTCAAAAATACGTCTAGTCGGATTGAGTCTGCAACTCG  
 ACTACATGAAGCTGGAATCGTAGTAACTCGCAGATCAGAATGCTGCGGTGAATACGTTCCGGTCTTGTACACACC  
 GCGCGTACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACGTAAGGAGGGAGCTGCCGAAGGTGGGAC  
 CGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCTTT

SEQ ID NO: 151

| PROKKA\_04036 16S ribosomal RNA gene  
 ACGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGTGTAAACACATGCAAGTCGAGCGAAGCGCTTGGG  
 AAGATTCTCGGATGATTCTCTTGACTGAGCGCGGACGGGTGAGTAACCGTGGTAACTGCTCATACAGG  
 GGATAACAGTTAGAAATGACTGCTAATACCGATAAGACCAACCGTACCGTACGGTACAGTGGTAAACTCCGGTG  
 GTATGAGATGGACCCCGCTTGATTAGTAGTGTGGGGTAACGGCTACCAAGCCGACGATCAGTAGCCGACCTG  
 AGAGGGTACCGGCCACATGGGACTGAGACACGGCCAGACTCTACGGGAGGCAGCAGTGGGAATATTGACACAA  
 TGGAGGAAACTCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCGGTATGTAAACTCTATCAGCAGGGAAAGA  
 AAATGACGGTACCTGACTAAGAACGCCCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGCAAGCGTTA  
 TCCGGATTACTGGGTGAAAGGGACGCTAGACGGCACGGCAAGCCAGATGTGAAAGCCCCGGCTCAACCCGGGA  
 CTGCATTGGAACCTGGTAGTGTGAGCTAGGTGCGAGGGCAAGTGAATTCTTAGTGTAGCGGTGAAATGCGTAGATAT  
 TAGGAGGAACACCAAGTGGCGAAGGCGCTTGCTGGACGATGACTGACGTTGAGGCTCGAAAGCGTGGGAGCAAACA  
 GGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGACTGCTAGGTGCGGCAAAGCATTGGTGCAG  
 CTAACGCAATAAGCAGTCCACCTGGGAGTACGTTGCAAGAATGAAACTCAAAGGAATTGACGGGACCCGACAA  
 GCGGTGGAGCATGTGGTTAATCGAAGCAACCGAAGAACCTTACCTGATCTGACATCCGATGACTGCTCGTA  
 ATGGAAAGTTTCTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTGGTTA  
 AGTCCCGCAACGAGCGAACCCCTATCTCAGTAGCCAGCAGGTTAACGCTGGCAGTGGGACTCTGGAGAGACTGCCAGGGAT  
 AACCTGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGACCAGGGTACACACGTGCTACATGGGT  
 AAACAAAGAGAAGCGAACTCGCAGGGTAAGCAAATCTCAAAAATACGTCTAGTCGGATTGAGTCTGCAACTC  
 GACTACATGAAGCTGGAATCGTAGTAACTCGCAGATCAGAATGCTGCGGTGAATACGTTCCGGTCTTGTACACAC  
 CGCCCGTACACCATGGGAGTCAGTAACGCCGAAGTCAGTGACCCAACGTAAGGAGGGAGCTGCCGAAGGTGGG  
 CGATAACTGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCTTT

SEQ ID NO: 152

| PROKKA\_00437 16S ribosomal RNA gene  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCATAACATGCAAGTCGAACGAAGTTGGAGGA  
 AGCTTGCTTCCAAAGAGAGACTTAGTGGCGAACGGGTGAGTAACACGTAGGTAACCTGCCATGTGTCGGGATAACTG  
 CTGGAAACGGTAGCTAAACCGGATAGGTATAACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGCGTGAAC  
 ATGGATGGACCTGCGCGCATTAGCTAGTTGGTGAGGTAAACGGCTACCAAGGCGATGATGCGTAGCCGGCTGAGA

-continued

GGTAAACGCCACATTGGGACTGAGACACGCCAAACTCTACGGGAGGCAGCAGTAGGAAATTCGTCAATGG  
 GGAAACCTGAACGAGCAATGCCCGTGAGTGAAGAAGGCTTCGGATCGTAAAGCTCTGTGTAAGTGAAGAACG  
 GCTCATAGAGGAATGCTATGGAGTGACGTACCCAGAAAAGCCACGCCAACTACGTGCCAGCCGGT  
 AATACGTAGGGCAAGCGTTACCGGAATCATGGCGTAAAGGTGCGTAGGTGGCGTACTAAGTCTGTAGTAAA  
 AGGCAATGGCTCAACCATTGTAAGCTATGGAACTGGTATGCTGGAGTGAGAAGAGGGCGATGGAATTCCATGTGT  
 AGCGTAAAATGCGTAGATATGGAGGAACCCAGTGGCGAAGGGCGTCGCCCTGCTGTAACGTGAGGCA  
 CGAAAGCGTGGGAGCAAATAGGATTAGATAACCTAGTAGTCACGCCGTAACCGATGAGAACTAAGTGTGAGGA  
 ATTCAGTGCTGCAAGTAAACGCAATAAGTCTCCGCTGGGAGTATGCAACGCGAAGAACCTTACCAAGGCTTGACATGGAA  
 GGGGCCCGACAAGCGGTGGAGTATGGTTAATTGAGCAACGCGAAGAACCTTACCAAGGCTTGACATGGAA  
 ACAAAATACCTAGAGATAAGGGGATAATTAGGATCACACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGA  
 TGTGGGTTAACCGCAAGCGGTGGAGTACAGGATCACACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGA  
 CCGGTGACAAACCGGAGGAAGGTGGGATGACGTAAACATCATGCCCCTATGGCCTGGCTACACACGTACTAC  
 AATGGCGGCCACAAAGAGCAGCAGCACAGTGTGAAGCGAATCTCATAAAGGCGTCTAGTTGGATTGAAAGTC  
 TGCAACTCGACTTCATGAGTCGAAGTCGCTAGTAATCGCAGATCAGCATGCTGCCGTGAATACGTTCTGGCCTT  
 GTACACACCGCCCGTCAAACCATGGAGTCAGTAATACCGAAGCCGGTGGCATAACCGTAAGGAGTGAGCGTCGA  
 AGTAGGACCGATGACTGGGTTAAGTCGTAACAAGGTATCCCTACGGGAACTGGGGATGGATCACCTCCTT

SEQ ID NO: 153

| PROKKA\_00896 16S ribosomal RNA gene  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCAATACATGCAAGTCGAACGAAGTTGGAGGA  
 AGCTTGCTTCAAAGAGACTTAGTGGCGAAGGGTAGTAACACGTAGGTAACCTGCCATGTGTCGGATAACTG  
 CTGGAAACGGTAGCTAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGCGTGAAC  
 ATGGATGGACCTGGCGCATTAGTAGTTGGTGAGGTAACGCCAACCAAGGCGATGATGCGTAGCCGGCTGAGA  
 GGGTAAACGCCACATTGGGACTGAGACACGCCAAACTCTACGGGAGGCAGCAGTAGGAAATTCGTCAATGG  
 GGAAACCTGAACGAGCAATGCCCGTGAGTGAAGAAGGCTTCGGATCGTAAAGCTCTGTGTAAGTGAAGAACG  
 GCTCATAGAGGAATGCTATGGAGTGACGGTAGCTTACCAAGAACCCACGCCAACTACGTGCCAGCCGGT  
 AATACGTAGGTGGCAAGCGTTACCGGAATCATGGCGTAAAGGTGCGTAGGTGGCGTACTAAGTGTAGTAAA  
 AGGCAATGGCTCAACCATTGTAAGCTATGGAACTGGTATGCTGGAGTGAGAAGAGGGCGATGGAATTCCATGTGT  
 AGCGTAAAATGCGTAGATATGGAGGAACCCAGTGGCGAAGGGCGTCGCCCTGCTGTAACGTGAGGCA  
 CGAAAGCGTGGGAGCAAATAGGATTAGATAACCTAGTAGTCACGCCGTAACCGATGAGAACTAAGTGTGAGGA  
 ATTCAGTGCTGCAAGTAAACGCAATAAGTCTCCGCTGGGAGTATGCAACGCGAAGAACCTTACCAAGGCTTGACATGGAA  
 GGGGCCCGACAAGCGGTGGAGTATGGTTAATTGAGCAACGCGAAGAACCTTACCAAGGCTTGACATGGAA  
 ACAAAATCCCTAGAGATAAGGGGATAATTAGGATCACACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGA  
 TGTGGGTTAACCGCAACGAGCGAACCCCTGCGATGTTACCAAGCATCAAGTGGGACTCATGCCAGACTG  
 CCGGTGACAAACCGGAGGAAGGTGGGATGACGTAAACATCATGCCCCTATGGCCTGGCTACACACGTACTAC  
 AATGGCGGCCACAAAGAGCAGCAGCACAGTGTGAAGCGAATCTCATAAAGGCGTCTAGTTGGATTGAAAGTC  
 TGCAACTCGACTTCATGAGTCGAAGTCGCTAGTAATCGCAGATCAGCATGCTGCCGTGAATACGTTCTGGCCTT  
 GTACACACCGCCCGTCAAACCATGGAGTCAGTAATACCGAAGCCGGTGGCATAACCGTAAGGAGTGAGCGTCGA  
 AGTAGGACCGATGACTGGGTTAAGTCGTAACAAGGTATCCCTACGGGAACTGGGGATGGATCACCTCCTT

SEQ ID NO: 154

| PROKKA\_02845 16S ribosomal RNA gene  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCAATACATGCAAGTCGAACGAAGTTGGAGGA  
 AGCTTGCTTCAAAGAGACTTAGTGGCGAAGGGTAGTAACACGTAGGTAACCTGCCATGTGTCGGATAACTG  
 CTGGAAACGGTAGCTAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGCGTGAAC

-continued

ATGGATGGACCTGCGCGCATTAGCTAGTTGGTGAGGTAAACGGCTACCAAGGCATGATGCGTAGCCGCCGTGAGA  
 GGGTAAACGGCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTAGGAAATTCGTCAATGG  
 GGGAAACCCCTGAACGAGCAATGCCGCGTAGTGAAGAAGGCTTCGGATCTAAAGCTCTGTGTAAGTGAAGAACG  
 GCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACAGAAAAGCCACGGCTAAACTACGTGCCAGCAGCCGGT  
 AATAACGTAGGTGGCAAGCGTTATCGGAATCATTGGCGTAAAGGTGCGTAGGTGGCGTACTAAGTCTGTAGTAAA  
 AGGCAATGGCTCAACCATTGTAAGCTATGGAAACTGGTATGCTGGAGTGCAGAAAGAGGGCATGGAATTCCATGTGT  
 AGCGGTTAAATGCGTAGATATGGAGGAACACCAGTGGCAAGGGCGTCGCCCTGGTCTGTAACTGACACTGAGGCA  
 CGAAAGCGTGGGAGGCAAATAGGATTAGATACCCCTAGTAGTCCACGCCGTAACAGATGAGAAACTAAGTGTGGAGGA  
 ATTCAGTGCTGCAGTTAACGCAATAAGTCTCCGCTGGGAGTATGCAAGCAAGCGAAGAACCTTACAGGCCCTGACATGGAA  
 GGGGGCCCGACAAGCGGTGGAGTATGTTAAATCGAAGCAAGCGAAGAACCTTACAGGCCCTGACATGGAA  
 ACAAAACCCCTAGAGATAGGGGATAATTATGGATCACACAGGTGGTGCATGGTCTGAGTGTGAAACTCAAAGGAATTGAC  
 TGTTGGGTTAAGTCCCGCAACGAGCGAACCCCTGTCGATGTTACAGCATCAAGTGGGACTCATGCGAGACTG  
 CCGGTGACAAACCGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGGCCTGGCTACACACGTACTAC  
 AATGGCGGCCACAAAGAGCAGCAGCACAGTGTGAGCGAATCTCATAAAGGTGCTCAGTTGGATTGAAAGTC  
 TGCAACTCGACTTCATGAAGTCGGATCGTAGTAAATCGCAGATCAGCATGCTGCCGTGAATACGTTCTGGCCTT  
 GTACACACCGCCCCGTCAAACCATGGGAGTCAGTAATACCGAAGCCGGTGGCATAACCGTAAGGAGTGAGCCGTGCA  
 AGGTAGGACCGATGACTGGGTTAAGTCGTAACAAGGTATCCCTACGGGAACGTGGGATGGATCACCTCCTT

SEQ ID NO: 155

| PROKKA\_04164 16S ribosomal RNA gene  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCAATACATGCAAGTCGAACGAAGTTGGAGGA  
 AGCTTGCTTCCAAAGAGACTTACTGGGAACGGGTGAGTAACACGTAGGTAACCTGCCATGTGTCGGATAACTG  
 CTGGAAACGGTAGCTAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGCGTGAAC  
 ATGGATGGACCTGCGCGCATTAGCTAGTTGGTGAGGTAAACGCCACCAAGGCATGATGCGTAGCCGCCGTGAGA  
 GGGTAAACGGCACATTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGCAGTAGGAAATTCGTCAATGG  
 GGGAAACCCCTGAACGAGCAATGCCGCGTAGTGAAGAAGGTCTCCGATCGTAAAGCTCTGTGTAAGTGAAGAACG  
 GCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACAGAAAAGCCACGGCTAAACTACGTGCCAGCAGCCGGT  
 AATAACGTAGGTGGCAAGCGTTATCGGAATCATTGGCGTAAAGGTGCGTAGGTGGCGTACTAAGTCTGTAGTAAA  
 AGGCAATGGCTCAACCATTGTAAGCTATGGAAACTGGTATGCTGGAGTGCAGAAAGAGGGCATGGAATTCCATGTGT  
 AGCGGTTAAATGCGTAGATATGGAGGAACACCAGTGGCAAGGGCGTCGCCCTGGTCTGTAACTGACACTGAGGCA  
 CGAAAGCGTGGGAGGCAAATAGGATTAGATACCCCTAGTAGTCCACGCCGTAACAGATGAGAAACTAAGTGTGGAGGA  
 ATTCAGTGCTGCAGTTAACGCAATAAGTCTCCGCTGGGAGTATGCAACGCAAGTGTGAAACTCAAAGGAATTGAC  
 GGGGGCCCGACAAGCGGTGGAGTATGTTAAATCGAAGCAAGCGAAGAACCTTACAGGCCCTGACATGGAA  
 ACAAAACCCCTAGAGATAGGGGATAATTATGGATCACACAGGTGGCATGGTGTGCTCAGCTCGTGTGAGA  
 TGTTGGGTTAAGTCCCGCAACGAGCGAACCCCTGTCGATGTTACAGCATCAAGTGGGACTCATGCGAGACTG  
 CCGGTGACAAACCGGAGGAAGGTGGGATGACGTCAAATCATGCCCCTATGGCCTGGCTACACACGTACTAC  
 AATGGCGACCACAAAGAGCAGCAGCACAGTGTGAGCGAATCTCATAAAGGTGCTCAGTTGGATTGAAAGTC  
 TGCAACTCGACTTCATGAAGTCGGATCGTAGTAAATCGCAGATCAGCATGCTGCCGTGAATACGTTCTGGCCTT  
 GTACACACCGCCCCGTCAAACCATGGGAGTCAGTAATACCGAAGCCGGTGGCATAACCGCAAGGAGTGAGCCGTGCA  
 AGGTAGGACCGATGACTGGGTTAAGTCGTAACAAGGTATCCCTACGGGAACGTGGGATGGATCACCTCCTT

SEQ ID NO: 156

| PROKKA\_04921 16S ribosomal RNA gene  
 ATGGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCATGCCAATACATGCAAGTCGAACGAAGTTGGAGGA  
 AGCTTGCTTCCAAAGAGACTTACTGGGAACGGGTGAGTAACACGTAGGTAACCTGCCATGTGTCGGGATAACTG

-continued

CTGGAAACGGTAGCTAAAACCGGATAGGTATACAGAGCGCATGCTCAGTATATTAAAGGCCATCAAGCGTGAAAC  
 ATGGATGGACCTCGCGCGATTAGCTAGTTGGTAGGTAACGGCCACCAAGGCATGATGCGTAGCCGGCTGAGA  
 GGGTAAACGCCACATTGGGACTGAGACACGGCCAAACTCTACGGGAGGCAGCAGTAGGGAAATTTCGTCAATGG  
 GGAAACCCCTGAACGAGCAATGCCGCGTAGTGAAGAAGGTCTCGGATCGTAAAGCTCTGTGTAAGTGAAGAACG  
 GCTCATAGAGGAAATGCTATGGGAGTGACGGTAGCTTACAGAAAGCCACGGCTAAGTGCAGCAGCCCGGT  
 AATACGTAGGTGGCAAGCGTTACCGGAATCATTGGCGTAAAGGGTGCCTAGGTGGCTACTAAGTCTGTAGTAAA  
 AGGCAATGGCTAACCAATTGTAAGTATGGAACTGGTATGCTGGAGTGCAGAAGAGGGCATGGAATTCCATGTGT  
 ACCGGTAAAATGGTAGATATGGAGGAACCCAGTGGGAAGGCCGTCGCTGTAACTGACACTGAGGCA  
 CGAAAGCGTGGGAGCAAATAGGATTAGATACCCCTAGTAGTCCACGCCGTAACAGTGAAGACTAAGTGTGGAGGA  
 ATTCACTGCTGCAGTTAACGCAATAAGTCTCCGCTGGGAGTATGCACGCAAGTGTGAAACTCAAAGGAATTGAC  
 GGGGGCCCGCACAGCGTGGAGTATGTGGTTAATTCAAGCAACGCCGAAAGACCTTACCAAGGCTTGACATGGAA  
 ACAAAACCCCTAGAGATAGGGGATAATTATGGATCACACAGGTGGTCATGGTTGTCGTCAAGCTGTGCGTGA  
 TGTGGGTTAAGTCCCACAGAGCGAACCCCTGTGGCATGTTACCAGCATCAAGTGGGACTCATGGAGACTG  
 CCGGTGACAAACCGAGGAAGGTGGGATGACGTCAAATCATCATGCCCTTATGGCCTGGCTACACACGTACTAC  
 AATGGCGGCCACAAAGAGCAGCAGCACAGTGTGAAGCGAATCTCATAAAGGTCGTCAAGTGGATTGAAGTC  
 TGCAACTCGACTTCATGAAGTCGAATCGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACTCGTCTCGGCCTT  
 GTACACACCGCCGTCAAACCATGGAGTCAGTAATACCCGAAGCGGTGGCATAACCGTAAGGAGTGAGCGTCGA  
 AGGTAGGACCGATGACTGGGTTAAGTCGAAACAGGTATCCCTACGGGAACGTGGGATGGATCACCTCCTT

SEQ ID NO: 157

| PROKKA\_00199 16S ribosomal RNA gene  
 TATTGAGAGTTGATCCTGGCTCAGGATGAAACGCTGGCGCGTGCTTAACACATGCAAGTCGAACGGGTGCTCATG  
 ACGGAGGATTCTCCAACGGATTGAGTTACCTAGTGGCGGACGGGTGAGTAACCGTGAGGAACCTGCCCTGGAGAG  
 GGGATAACACTCCGAAAGGAGTGCTAATACCGCATGATGCGAGTTGGGTGCGATGGCTCTGACTGCCAAAGATTAT  
 CGCTCTGAGATGCCCTCGCTGATTAGCTAGTAGGGGGTAACGCCAACCTAGGGACGATCAGTAGCCGAC  
 TGAGAGGTTGACCGGCCACATTGGACTGAGACACGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGC  
 AATGGCGCAAGCCTGACCCAGCAACGCCGCGTAGAGGAAGAAGGTTTGGGTGAAACTCTTTGTCGGGAC  
 GAAACAAATGACGGTACCGACGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGTGGCAAG  
 CGTTATCCGATTACTGGGTAAAGGGCGTAGGGGGATTGCAAGTCAGATGTGAAACTGGGGCTCAACCT  
 CCAGCGCTGCATTGAAACTGTAGTTCTGAGTGTGGAGAGGCAATCGGAATTCCGTGTAGGGTGAATGCGTA  
 GATATACGGAGAACACCGTAGGCCGAAGGGGATTGCAAGTCAGTAACGTGACGCTGAGGCGCAAAGCGTGGGAGC  
 AACACAGGATTAGATAACCTGGTAGTCCACGCCGTAACAGATGGATACTAGGTGTGGGCTGACCCCTCCGTGCC  
 GCAGTTAACACAATAAGTATCCACCTGGGAGTACGATGCAAGGTTGAAACTCAAAGGAATTGACGGGGCCCGC  
 ACAAGCGGTGGAGTATGTGGTTAATTGCAAGCAACGCGAAGAACCTTACCAAGGGCTGACATCCACTAACGAGC  
 AGAGATGCATTAGGTGCCCTCGGGAAAGTGGAGACAGGTGGCATGGTTGTCGTCAAGCTGTGAGATGT  
 TGGGTTAAGTCCGCAACGAGCGAACCCCTATTGTTAGTTGCTACGCAAGAGCAGTCAAGCTGAGACTGCCGTTGAC  
 AAAACGGAGGAAGGTGGGACGAGTCGAAATCATCATGCCCTTATGTCCTGGCACACACGTACTACAATGGTGG  
 TTAACAGAGGGAGGCAATACCGGAGGTGGAGCAATCCCTAAAGCCATCCAGTCGGATTGCAAGGCTGAAACCC  
 GCCTGTATGAAAGTTGAAATCGTAGTAATCGGGATCAGCATGCCGCGTGAATGCTCCGGGCTTGTACACAC  
 CGCCCGTCACACCATGAGAGTCGGGACACCCGAAGTCCGTAGCCTAACCGCAAGGAGGGCGCGGCCGAAGGTGG  
 TCGATAATTGGGTTGAAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGGCTGGATCACCTCCTT

-continued

SEQ ID NO: 158

| PROKKA\_00208 16S ribosomal RNA gene  
TATTGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAACGGGTGCTCATG  
ACGGAGGATTCTGCCAACGGATTGAGTTACCTAGTGGCGACGGGTGAGTAACGCGTGAGGAACCTGCCCTGGAGAG  
GGGAATAACACTCCGAAAGGAGTGCTAATACCGCATGATGCAGTTGGGTGCGATGGCTCTGACTGCCAAAGATTAT  
CGCTCTGAGATGGCCTCGCGTCTGATTAGCTAGTAGGCGGGTAACGGCCCACCTAGGCGACGATCAGTAGCCGGAC  
TGAGAGGTTGACCGGCCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGC  
ATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTCGGGTTGAAACTCTTTGTCGGGAC  
GAAAACAATGACGGTACCCGACGAATAAGCACCGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGTGGCAAG  
CGTTATCCGATTACTGGGTGAAAGGCGTGTAGGCGGATTGCAAGTCAGATGTGAAAACGGGCTCAACCT  
CCAGCCTGCATTGAAACTGTAGTTCTGAGTGCTGGAGAGGCAATCGAATCCGTGTAGCGGTGAAATGCGTA  
GATATA CGGAGAACACCAGTGGCGAAGGCGATTGCTGGACAGTAACGCGTGAGGCGAAAGCGTGGGAGC  
AACAGGATTAGATAACCTGGTAGTCCACGCCGAAACGATGGATACTAGGTGTTGGGCTGACCCCCCTCCGTGC  
CCGAGTTAACACAATAAGTATCCCACCTGGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGCCG  
ACAAGCGGTGGAGTGTGTTAATCGAACGCAACGCCAGAACCTTACCGGCTTGACATCCCACGAAAG  
CAGAGATGCATTAGGTGCCCTCGGGAAAGTGGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATG  
TTGGGTTAAGTCCGCAACGAGCGAACCCATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGA  
CAAAACGGAGGAAGGTGGGAGCAGTCAAATCATCATGCCCTTATGCTGGGACACACGTACTACAATGGTG  
GTTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAATCCCTAAAGCCATCCAGTCGGATTGCAAGGCTGAAACC  
CGCCTGTATGAAGTTGAACTCGCTAGTAAATCGGGATCAGCATGCCCGGTGAATACGTTCCGGGCTTGACACA  
CCGCCCGTACACCATGAGAGTCGGAACACCCGAAGTCCTAGCCTAACCGCAAGGAGGGCGCCGAAGGTGG  
TTCGATAATTGGGTGAAGTCGAACAAGGTAGCCGTATCGGAAGGTGCGCTGGATCACCTCTT

SEQ ID NO: 159

| PROKKA\_04460 16S ribosomal RNA gene  
TATTGAGAGTTGATCCTGGCTCAGGATGAACGCTGGCGCGTCTAACACATGCAAGTCGAACGGGTGCTCATG  
ACGGAGGATTCTGCCAACGGATTGAGTTACCTAGTGGCGACGGGTGAGTAACCGCGTGAGGAACCTGCCCTGGAGAG  
GGGAATAACACTCCGAAAGGAGTGCTAATACCGCATATGCAGTTGGGTGCGATGGCTCTGACTGCCAAAGATTAT  
CGCTCTGAGATGGCCTCGCGTCTGATTAGCTAGTAGGCGGGTAACGGCCCACCTAGGCGACGATCAGTAGCCGGAC  
TGAGAGGTTGACCGGCCACATTGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGC  
ATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTCGGGTTGAAACTCTTTGTCAGGGAC  
GAAAACAATGACGGTACCTGACGAATAAGCACGCCGTAACCTACGTGCCAGCAGCCGGTAATACGTAGGTGGCAAG  
CGTTATCCGATTACTGGGTGAAAGGCGTGTAGGCGGATTGCAAGTCAGATGTGAAAACGGGCTCAACCT  
CCAGCCTGCATTGAAACTGTAGTTCTGAGTGCTGGAGAGGCAATCGAATCCGTGTAGCGGTGAAATGCGTA  
GATATA CGGAGAACACCAGTGGCGAAGGCGATTGCTGGACAGTAACGCGTGAGGCGCAAGCGTGGGAGC  
AACAGGATTAGATAACCTGGTAGTCCACGCCGAAACGATGGATACTAGGTGTTGGGCTGACCCCCCTCCGTGCC  
GCAGTTAACACAATAAGTATCCCACCTGGGAGTACGATCGAACGGTTGAAACTCAAAGGAATTGACGGGGCCG  
ACAAGCGGTGGAGTGTGTTAATCGAACGCAACGCCAGAACCTTACCGGCTTGACATCCCACGAAAG  
AGAGATGCATTAGGTGCCCTCGGGAAAGTGGAGACAGGTGGTGCATGGTGTGTCAGCTCGTGTGAGATG  
TGGGTTAAGTCCGCAACGAGCGAACCCATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGAC  
AAAACGGAGGAAGGTGGGAGCAGTCAAATCATCATGCCCTTATGCTGGGACACACGTACTACAATGGTGG  
TTAACAGAGGGAGGCAATACCGCGAGGTGGAGCAAATCCCTAAAGCCATCCAGTCGGATTGCAAGGCTGAAACCC  
GCCTGTATGAAGTTGGAATCGCTAGTAATCGGGATCAGCATGCCCGGTGAATACGTTCCGGGCTTGTACACAC

-continued

```
CGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTCCGTAGCTAACCGCAAGGAGGGCGGCCGAAGGTGGGT
TCGATAATTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT
```

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," "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* Infection

#### 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 10-10<sup>4</sup> *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.

#### 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 (OD<sub>600</sub>) of the cultures was assessed and live biotherapeutic product (LBP) cocktails were prepared inside an anaerobic chamber adjusting inputs based upon OD<sub>600</sub> for equal CFU ratio cocktails in PBS (sterile, pre-treated).

#### *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 (FIGS. 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 10<sup>4</sup> *C. difficile* spores by oral gavage on day 0 (FIGS. 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 (FIG. 4J; black triangles).

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

The groups of mice that received cefoperazone treatment had a significant change in weight (FIG. 4H) and substantial *C. difficile* bacterial load in the fecal pellets (FIG. 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 (FIG. 4I) and all mice survived (FIG. 4A) without significant change in body weight (FIG. 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,

199

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* (FIG. 5).

In general, LBP cocktails were mixed in PYG media, and each mouse was administered a dose by oral gavage in 250  $\mu\text{L}$  pre-reduced PBS (media-free). For composition E, bacteria were mixed in equal volumes (not equal ratios/CFUs) and administered in a 250  $\mu\text{L}$  dose. Each LBP of Compositions A-D contained  $10^8$  CFUs total in a 250  $\mu\text{L}$  dose, comprised of  $10^7$  CFU of each of the bacterial strains (FIG. 1), for a total of  $10^8$  CFU administered to each animal. Composition I contained a total of  $10^6$  CFUs in a 250  $\mu\text{L}$  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 (FIG. 5). Mice evaluated for survival/mortality (FIG. 6), weight (FIGS. 7A-7I), and *C. difficile* CFUs (FIGS. 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 (FIG. 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 (FIG. 10), weight (FIG. 11), and *C. difficile* burden (CFUs per gram feces; FIG. 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

FIG. 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. FIG. 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 (FIG. 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  $\mu\text{L}$  of a 10% fecal sample s per mouse). Mice were then evaluated for survival/mortality (FIG. 16), weight (FIGS. 17A-17H), and *C. difficile* burden (CFU/gram feces) on days 1, 3, 8 and 17 after infection (FIGS. 18A and 18B). The data demonstrate that Composition B, Composition F, and FMT protect against and/or treat *C. difficile* infection.

200

*ficide* burden (CFU/gram feces) on days 1, 3, 8 and 17 after infection (FIGS. 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

FIG. 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 FIG. 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 (FIGS. 21 and 23) and change in weight (FIGS. 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 (FIG. 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* (FIG. 26).

Mice were then challenged with *C. difficile* infection with  $10^4$  *C. difficile* VPI 10463 spores and monitored for survival/mortality (FIGS. 27A and 28A), weight (FIGS. 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 (FIGS. 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

## 201

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<sup>-6</sup> dilutions of clarified pellet supernatant. Vero cell cultures were exposed to the diluted samples for approximately 18 hours, then visualized by phase contrast 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 (FIG. 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 bif fermentans*, 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 (OD<sub>600</sub>) 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 (TC-CFA) 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.

## 202

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, Calif.), 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 OD<sub>600</sub>, to reach a final consortium OD<sub>600</sub> of 0.1. *C. bif fermentans* and *B. thetaiotaomicron* were setup to compete with Cdiff individually at an OD of 0.1. The OD<sub>600</sub> 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. bif fermentans* 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 1×10<sup>-6</sup>. Plates for CFU enumeration were prepared by spreading 100 µL of each of the 1×10<sup>-4</sup> through 1×10<sup>-6</sup> 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. bif fermentans* (positive control). The results are shown in Table 5 and FIG. 31.

203

TABLE 5

Summary Results for In Vitro Competition				
Experiment Number	No Competing Strain(s)	Competition with <i>B. thetaiotaomicron</i>	Competition with <i>C. bif fermentans</i>	Competition with Composition B
n = 1	100	9.90	0.1	33.8
n = 2	100	115	39.5	0.5
n = 3	100	41.3	0.7	33.1
n = 4	100	105	14.1	0.7
n = 5	100	57.4	4.1	20.9
n = 6	100	65.6	11.7	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 vitro competition, Composition B inhibited Cdiff growth to  $15.1 \pm 16.2\%$  of control (absence of competing strain(s)). This result is consistent with the inhibition observed by the positive control, *C. bif fermentans*, of  $11.7 \pm 16.5\%$  of control. *B. thetaiotaomicron*, a negative control, yielded a negligible effect on Cdiff growth at  $65.6 \pm 43.8\%$  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. bif fermentans*. 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

204

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, Calif.).

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  $37^\circ\text{C}$ . in the anaerobic chamber. Unless otherwise noted, when the optical density (OD) was  $\geq 0.2$ , samples were collected for CFU enumeration and filtration. Inside an anaerobic chamber, a 100  $\mu\text{L}$  sample of turbid culture was collected and serially diluted 1:10 to reach a final dilution of  $1 \times 10^{-6}$ . Plates for CFU enumeration were prepared by spreading 100  $\mu\text{L}$ /dilution for the  $1 \times 10^{-4}$  through  $1 \times 10^{-6}$  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 (bio-Mérieux 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  $\mu\text{m}$  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  $\mu\text{m}$  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 ( $\mu\text{g}/\text{Log}(\text{CFU/mL})^*\text{mL}$ )							
	Acetate	Propionate	Iso-	2-Methyl-	Iso-	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

**205**

Seven strains of Composition B were found to produce significant quantities (>1 µg/Log(CFU/mL)\*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 µg/Log (CFU/mL)\*mL) of other SCFAs were also produced by the Composition B strains.

Example 10: Composition B Induces Regulatory T Cells (Tress)

Each of the bacterial strains of Composition B were grown to log phase, combined to a total dose of ~10<sup>8</sup> cfu per

**206**

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 FIG. 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.

## SEQUENCE LISTING

```

<160> NUMBER OF SEQ ID NOS: 159

<210> SEQ ID NO 1
<211> LENGTH: 210
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 1
gccccggagca gttgatgtga aggatgggtc acctgtggac tgcattggaa ctgtcatact      60
tgagtgcggg agggtaagcg gaattcctag tgttagcggtg aaatgcgtat atattaggag     120
gaacaccagt ggcaaggccg gcttactgga cggtaactga cgttgaggct cgaaacgcgtg     180
gggagcaaac aggattagat accctggtaa                                         210

<210> SEQ ID NO 2
<211> LENGTH: 184
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 2
ctaaccgtgg aggtcattgg aaactggtca acttgagtgca agaagaggaa agtggaaattc      60
catgtgttagc ggtgaaatgc gttagagatat ggaggaacac cagtggcgaa ggcggcttcc    120
tggctgttaa ctgacactga ggccgaaag cgtggggggc aaacaggatt agatccccccg     180
gtaa                                         184

<210> SEQ ID NO 3
<211> LENGTH: 196
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 3
atgaaagccg gggctcaacc ccggtaactgc tttggaaact gtttgacttg agtgctttag      60
aggttaagtgg aattccttagt gttagcggaa atgttttagat attaggagga caccagtggc    120
gaaggccggct tactggactg taactgacgt tgtggctcga ttgtggggga gcaaacaggaa    180
ttatatcccc tggtaa                                         196

<210> SEQ ID NO 4
<211> LENGTH: 211
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

```

-continued

<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 4

```
cggaaggctt gatgtgaagg ttggggctta ccccgactg cattggaaac tgttttctta      60
gagtgcggca gaggtaagcg gaattcctag tgttagcggtg aaatgttta gatattagga     120
ggaacaccag tggcgaaggc ggcttactgg acggtaactg acgttgaggc tcgaaacgct     180
ggggagcaaa caggattaga tacccctggta a                                         211
```

<210> SEQ ID NO 5  
<211> LENGTH: 207  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 5

```
cgtatgtctga gtgaaggctg gggcttaccc caggactgca ttggaaactgt ttttcttagag      60
tgccggagag gtaagcgaa ttcctagtgt agcggtgaaa tgcgtagata ttaggaggaa     120
caccagtggc gaaggcggt tactggacgg taactgacgt tgaggctcga aacgcgtggg     180
agcaaaacagg attagatacc ctggtaa                                         207
```

<210> SEQ ID NO 6  
<211> LENGTH: 179  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 6

```
ttaaccaaga agtgcattgg aactgtcaga ctgggggaa aaaaagacag tgcaactcca      60
tgtgtagcggtt ggaatgctc catatatatg gaagaacacc agtggcgaag gcggctgtct    120
ggtctgcaac tgacgctgag gctcgaattc atggtaaga aagtattagt cccttgtaa     179
```

<210> SEQ ID NO 7  
<211> LENGTH: 214  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 7

```
accgcgttgg tctgaggtga ggctggggct taaccccagg actgcattgg aaactgttgt      60
tcttagagtgc cggagaggta agcggaaattc ctatgttagc ggtgaaatgc gtatattaa    120
ggaggaacac cagtggcgaa ggcggcttac tggacggtaa ctgacgttga ggctcgaaag    180
cgtggggagc aaacaggatt agataccctg gttaa                                         214
```

<210> SEQ ID NO 8  
<211> LENGTH: 196  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 8

```
taggctgggg cttaaacccca ggactgcatt ggaaactgtt tttcttagagt gccggagagg      60
taagcgaaat tcctagtgtt gcggtgaaat gcgttagatat taggaggaac accagtggcg    120
```

## US 10,555,980 B2

**209****210**

-continued

aaggccggtt actggacggt aactgacgtt gaggctcgaa agcgtgggga gcaaacagga 180

ttagataccc tggtaa 196

&lt;210&gt; SEQ ID NO 9

&lt;211&gt; LENGTH: 297

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 9

ttgcattgga cactatgtca gctgagtgtc ggagaggtaa gtggaaattcc tagtgttagcg 60

gtgaaatgcg tagatattag gaggaacacc agtggcgaag gcggcttact gcacgtttc 120

tgacgttgag gctcgaaatc gtggggagca aacaaaaata gataccctgg tagtccacgc 180

cgtaaacgat gcatactagg tgtcgggtgg caaagccatt cggtgccgca gcaaacgcaa 240

taagtatgcc acctggggag tacgttcgca agaatgaaac tcaaataaat tgacgga 297

&lt;210&gt; SEQ ID NO 10

&lt;211&gt; LENGTH: 209

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 10

cccgctcgtag atgtgaactg ggggctcacc tccagcctgc atttgaact gtagttcttg 60

agtgcgtggag aggcaatcgg aattccgtgt gtagcggta aatgcgtaga tatacggagg 120

aacaccagtg gcgaaggcgg attgctggac agtaactgac gctgaggcgc gaaagcgtgg 180

ggagcaaaaca ggattagata ccctcataaa 209

&lt;210&gt; SEQ ID NO 11

&lt;211&gt; LENGTH: 401

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 11

acctgatgca gcgacgcccgc gtgagtgaag aagtatttcg gtatgtaaag ctctatcagc 60

agggaaagaaa aaagacggta cctgactaag aagccccggc taactacgtg ccagcagccg 120

cggtataacg tagggggccaa gcgttatccg gaattactgg gtgtaaaggg tgcgttaggtg 180

geatggtaag tcagaagtga aagcccgcccc cttAACCCGGG ggactgcttt tgaaactgtc 240

atgctggagt gcaggagagg taagcggaat ttctagtgtta gcccgtaaat gcgttagatat 300

taggaggaac accagtgccg aaggccgctt actggactgt cactgacact gatgcaccaa 360

agcgtgggaa gcaaacagga ttagataccc tggaaagtccaa 401

&lt;210&gt; SEQ ID NO 12

&lt;211&gt; LENGTH: 355

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 12

atgggagcgt agatggcgac tggccatat gtgacagccc tggctcaac cccttaactg 60

catttggaaac tgagtggctg gagtgctgaa gaggcaggcg gaattccttag tgcgtgggt 120

-continued

---

aatgcgtag atattaggag gaacaccagt ggccaaggcg gcctgctgga cgatgactga	180
cgttggggct cgaaagcgtg gggagcaac aggattagat accctggtag tccacgcgt	240
aaacgatgac tactaggtgt cgggtggcaa ggacattcgg tgccgcagca aacgcaataa	300
gtagtcacc tggggagta gttcgcaaga atgaaactca aaggaaatg acgga	355
 <b>&lt;210&gt; SEQ ID NO 13</b>	
<b>&lt;211&gt; LENGTH: 197</b>	
<b>&lt;212&gt; TYPE: DNA</b>	
<b>&lt;213&gt; ORGANISM: Artificial Sequence</b>	
<b>&lt;220&gt; FEATURE:</b>	
<b>&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide</b>	
 <b>&lt;400&gt; SEQUENCE: 13</b>	
cgcagcggag tgtatcctag gctcacctgg ctgcttcga actggtttc tagatcggt	60
agagggggag attcctggtg tagcgtgaaa tgcgtagata tctggagggaa caccagtggc	120
gaaggcggcc tcctggacgg caactgacgt tgaggctcgaa aagtgtgggg agcaaacagg	180
attagataacc ctggtaa	197
 <b>&lt;210&gt; SEQ ID NO 14</b>	
<b>&lt;211&gt; LENGTH: 1522</b>	
<b>&lt;212&gt; TYPE: DNA</b>	
<b>&lt;213&gt; ORGANISM: Artificial Sequence</b>	
<b>&lt;220&gt; FEATURE:</b>	
<b>&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide</b>	
 <b>&lt;400&gt; SEQUENCE: 14</b>	
tcaaagagtt tgatcctggc tcaggacgaa cgctggcggc gcgcctaaca catgcaagtc	60
gaacggagct tacgttttga agtttcgga tggacgaatg taagcttagt ggccggacggg	120
ttagtaaacac gtgagcaacc tgcctttcag agggggataa cagccggaaa cggctgtcaa	180
taccgcataa tggcgggggc gcacatgccc ctgcaaccaa aggagaatc cgctgaaaga	240
tgggctcgcg tccgatttagc cagttggcg ggtAACGCCG caccAAAGCG acgatcggt	300
gccggactga gaggttgaac ggccacattt ggactgagac acggcccaga ctcctacgg	360
aggcagcgtt gggggatatt gcacaatggg cggaaacgtt atgcagcgtt gcccgttgag	420
ggaagacgtt ttccggattt taaacctctg tctttggggaa agaaaaatgac ggtacccaaa	480
gaggaagctc cggctaaacta cgtgccagca gcccggtaa tacgttaggg gcaagcggt	540
tccggaatta ctgggtgtaa agggagcgtt ggcggatgg caagtagaat gttaaatcca	600
tccggtcaac cgggtggctgc gttctaaact gcccgtttagt agtgaagttag aggcaggcgg	660
aattcctagt gtagegggtt aatgcgtttaga tattaggagg aacaccagtgc gcgaaaggcgg	720
cctgctgggc tttaactgac gctggggctc gaaacgttgg gggcaaaaca ggattagata	780
ccctggtagt ccacgcccgtt aacgttggatt actagggttg gggggactga ccccttcgt	840
gcccggatgtt acacaataag taatccaccc tggggatgtt gcccggatgtt tgaaactcaa	900
aggaaatttgcg gggggccccgc acaaggcgtt ggtatgttg tttaatttgcg agcaacgcgtt	960
agaacccatc cagggttttgc catggatgc ataggcttgcg atagggttgg gcccgg	1020
gcatccagac aggtgggttgc tgggtgttgcgt cagctgtgtt cgttggatgtt tgggtttagt	1080
cccgcaacgtt ggcgttccctt tattttttgtt tgctacgttgc gggactctt atggactgcgtt	1140
cgttgcgtttt acggggatgtt gttggggatgtt cgttgcgtttt tttttttttt tatggatgtt	1200
gtttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1260
atccccaaaaa agtgcgttgc ttcagattgc aggctgcgtt cccgttgcgtt gaagtcgttgcgtt	1320

-continued

---

ttgctagtaa tcgcggatca gcatgccgctgtgaatacgt tcccgcccttgtacacacc	1380
gccccgtcaca ccatgggagt cggtAACCCAGT agcctaACCG caaggGGGGC	1440
gtctgtcgaa gtgggattga tgactgggtt gaagtcgtaa caaggttagcc gtatcgaa	1500
gtgcggctgg atcacccctt tt	1522
<210> SEQ ID NO 15	
<211> LENGTH: 1529	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic Polynucleotide	
<400> SEQUENCE: 15	
tcaggagatt tgatcctggc tcaggatcaa cgctggccgc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggccggacgg	120
gtgagtaacg cgtggcaac ctgcctcata cagggggata acagtttagaa atgactgcta	180
ataccgcata agcgcacagg accgcattgtt gtatgtgaa aaactccgtt ggtatgagat	240
ggaccgcgtt ctgatttaggt agttgggtgg gtaaaggccct accaagccga cgatcagtag	300
ccgacctgag agggtgaccg gccacattgg gactgagaca cggcccaaac tcctacggga	360
ggcagcagtg gggaaatattt cacaatgggg gaaaccctga tgcagcgcacg ccgcgtgaa	420
gaagaagtat ttccgttatgt aaacttctat cagcaggaa gaagatgacg gtacctgagt	480
aagaaggcacc ggctaaatac gtgcgcacgc cccgcgttaat acgtatggtg caagcgttat	540
ccggattttac tgggtgtaaa gggagcgttag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc ttggaaact gcagatctgg agtgccggag aggtaaacgg	660
aattcctagt gtacgggtga aatgcgtaga tattaggagg aacaccatgt gcgaaggcgg	720
cttactggac ggtactgac gttggggctc gaaagcgtgg ggagcaaaaca ggattagata	780
ccctggtagt ccacgcgtt aacgtatgc actagggtgc ggtgtgaaaa gcacatcggt	840
gccgcagcaa acgcaataag tagtccaccc gggaggtacg ttgcagaa tgaaactcaa	900
aggaattgac ggggacccgc acaagcggtg gacatgtgg tttaattcga agcaacgcga	960
agaaccttac ctggtcttga catccggatg acgggcgcgtt aatgtcgccg tcccttcggg	1020
gcatccgaga cagggtgtgc atgggtgtcg tcaagctcgat tcgtgagatg ttgggttaag	1080
tcccgcaacg agcgcacacc ttatcttcag tagccagcat ataagggtgg cactctggag	1140
agactgcacgg ggagaacactg gaggaaagggtt gggatgacgt caaatcatca tgccccttat	1200
ggccagggtt acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgg	1260
agcgaatccc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctacatgaag	1320
cttggaaatcgc tagataatcgc ggtacagcat gccgcgggtga atacgttccc gggtcttgc	1380
cacacccccc gtcacaccat gggagtcgtt aacgccccaa ggcagtgacc caacctttaga	1440
ggaggggact gtcgaaggcg ggacggataa ctggggtgaa gtcgtaaaca ggtagccgtt	1500
tcggaaagggtt cggctggatc acctcccttt	1529

<210> SEQ ID NO 16	
<211> LENGTH: 1527	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic Polynucleotide	

-continued

&lt;400&gt; SEQUENCE: 16

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaaacac atgcaagtgc	60
aacgaagcga tttaacggaa gtttcggat ggaagttgaa ttgactgagt ggccggacggg	120
ttagtacacgc gtggtaacc tgcctgtac tgggggacaa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta tcgcatgata cagtgtgaaa aactccggtg gtacaagatg	240
gaccccgctc tgattagcta gttggtaagg taacggctta ccaaggcgac gatcagtgc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaaact cctacggag	360
gcagcagtgg ggaatattgc acaatggcg aaaggctgat gcagcgcacgc cgctgagtg	420
aagaagtatt tcggtatgta aagctctatc agcagggaa aaaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcgc cgccgtataa cgtagggggc aagcgttatc	540
cggatttact ggggtaaag ggagcgtaga cggtaaagca agtctgaagt gaaagccgc	600
ggctcaactg cgggactgct ttggaaactg tttaactgga gtgtcggaga ggttaagtgg	660
atcccttagtg tagcggtgaa atgcgtagat attaggagga acaccagtgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcgtaa acgtgataa cttagtggg gggagcaaa ctcttcggtg	840
ccgtcgcaaa cgcagtaagt attccacctg gggagtagct tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcggtg agcatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga cggggagtaga acgtccccctt cccttcgggg	1020
cggagaagac aggtggtgca tgggtgtcg cagctcggt cgtgagatgt tgggttaagt	1080
cccgcaacga ggcacccct tattctaagt agccagcggt tcggccggaa actcttggga	1140
gactgccagg gataacctgg aggaagggtgg ggatgacgct aaatcatcat gccccttatg	1200
atctggctca cacacgtgct acaatggcg aaacaaagag aagcaagacc gcgaggtgg	1260
gcaaatactca aaaataacgt ctcaagttcg actgcaggct gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg tcgctgtgaa tacgttcccg ggtcttgc	1380
acaccggccg tcacaccatg ggagtctgta acgcccgaag tcagtgaccc aaccgcaagg	1440
agggagctgc cgaaggcgaa accgataact ggggtgaagt cgttaacaagg tagccgtatc	1500
ggaagggtgcg gctggatcac ctccctt	1527

&lt;210&gt; SEQ ID NO 17

&lt;211&gt; LENGTH: 1530

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 17

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaaacac atgcaagtgc	60
aacgaagcaa tttaaatgaa gtttcggat ggatttttgta ttgactgagt ggccggacggg	120
ttagtacacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa	180
taccgcataa gcgcacagta ccgcacggta cgggtgtgaaa aactccggtg gtgtgagatg	240
gatcccgctc tgattagcca gttggcgaaa taacggccca ccaaggcgac gatcagtgc	300
cgacctgaga ggggtgaccgg ccacattggg actgagacac ggcccaaact cctacggag	360
gcagcagtgg ggaatattgc acaatggcg aaaggctgat gcagcgcacgc cgctgagtg	420
aagaagtatt tcggtatgta aagctctatc agcagggaa aaaaatgacgg tacctgacta	480

-continued

agaagccccg gctaactacg tgccagcago cgcggtaata cgtaggggc aagcgttatc	540
cggatttaact gggtgtaaag ggagcgtaga cggcgaagca agtctgaagt gaaaaccag	600
ggctcaaccc tggactgct ttggaaactg ttttgctaga gtgtcggaga ggttaagtgg	660
attccttagtg tagcggtgaa atgcgttagat attaggagga acaccagtgg cgaaggccgc	720
ttactggacg ataaactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcccgtaa acgtgaatg cttagtggc gggggcaaaag cccttcggtg	840
cgtcgcaaa cgccgtaaacg attccacctg gggagtagtgc tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcggtgg agcatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc aagtcttgac atcccttgc cccggcgtgtc acggcgcctt cccttcgggg	1020
caagagagac aggtggtgca tggttgtcg cagctcgatc cgtgagatgt tgggttaagt	1080
cccgcaacga ggcacccct tatacgtcg agccagcagg taaagctggg cactctagg	1140
agactgcacg ggataaacctg gaggaagggtg gggatgacgt caaatcatca tgccccttat	1200
gatttggctt acacacgtgc tacaatggcg taaacaaagg gaagcaagac agtgtatgtgg	1260
agcaaatccc aaaaataacg tcccgatcg gactgttagtc tgcaacccga ctacacgaa	1320
ctggaatcgc tagtaatcgc gaatcagaat gtccgggtga atacgttccc gggttttgt	1380
cacacccccc gtcacccat gggagtcacg aacccccgaa gtcagtgacc caactcgaa	1440
gagagggagc tgccgaaggc gggggcaggta actgggtgg tgcgtaaaca aggtagccgt	1500
atcggaaggc gcggctggat cacccctttt	1530

<210> SEQ ID NO 18  
 <211> LENGTH: 1528  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 18

aacgagagtt tgatcctggc tcaggatgaa cgctggggc gtgtttaaca catgcgtac	60
gagcgaagca ctggaaaga ttcttcggat gaagactttt gtgactgagc ggccggacggg	120
ttagtacgcg gtggtaacc tgcctcatac agggggataa cagtttggaaa tgactgtaa	180
taccgcataa gaccacggta ccgcgttgcgta cagtttggaaa aactccggtg gtatgatg	240
gacccgcgtc tgattaggta gttgggggg taacggccta ccaagccgac gatcgttagc	300
cgacctgaga ggggtacccgg ccacattgggg actgagacac ggcccagact cctacgggag	360
gcagcgtgg ggaatattgc acaatggagg aaactctgtat gcagcgtacgc cgcgtgaagg	420
atgaagtatt tcggatgtt aacttctatc agcaggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcago cgcggtaata cgtaggggc aagcgttatc	540
cggatttaact gggtgtaaag ggagcgtaga cggcgtggc acggcgtatgt gaaagccgg	600
ggctcaaccc cggactgca ttggaaactg ctgagctaga gtgtcggaga ggcaagtgg	660
attccttagtg tagcggtgaa atgcgttagat attaggagga acaccagtgg cgaaggccgc	720
ttgctggacg atgcgtacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcccgtaa acgtgtactg cttagtggc ggtggcaaaag ccatttcggtg	840
ccgcagctaa cgcaataacg agtccacctg gggagtagtgc tcgcaagaat gaaactcaaa	900
ggaattgacg gggacccgca caagcggtgg agcatgtgg ttaattcgaa gcaacgcgaa	960

## US 10,555,980 B2

219

220

-continued

---

gaaccttacc tgatcttgac atcccgatga ccgcgtcgta atggaagctt ttcttcggaa	1020
catcggtgac aggtggtgca tggttgtcg cagctcggt cgtgagatgt tgggttaagt	1080
cccgcaacga gcgcaccccc tatcttcagt agccagcagg ttaagctggg cactctggag	1140
agactgccag ggataacctg gaggaaggtg gggatgacgt caaatcatca tgccccttat	1200
gaccagggtt acacacgtgc tacaatggcg taaacaaaga gaagcgaact cgcgaggta	1260
agcaaatacgc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctatcatgaa	1320
ctggaaatcgc tagtaatcgc agatcagaat gctgcgggtga atacgttccc gggtcttgta	1380
cacacccgcccc gtcacaccat gggaggtcagt aacgccccaa gtcagtgacc caaccgtaa	1440
gagggagctg ccgaaggtgg gaccgataac tggggtgaag tcgtaacaag gtggccgtat	1500
cggaaggtgc ggctggatca cctcccttt	1528

&lt;210&gt; SEQ\_ID NO 19

&lt;211&gt; LENGTH: 1531

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 19

atcagagagt ttgatcctgg ctcaggatga acgctggcg cgtgcttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctccttcgg attgaaactt atttgcgtga gcggcggacg	120
ggtgagtaac gcgtggtaaa cctgcctcat acagggggat aacagttaga aatggctgct	180
aataccgcat aagcgcacag gaccgcatgg tctgggtgtga aaaactccgg tggtatgaga	240
tggaccgcgc tctgattagc tagttggagg ggttaacggcc caccaaggcg acgatcgta	300
gcgcgcctga gagggtaaac gcgcacattt ggactgagac acggcccgaga ctcctacggg	360
aggcagcagt ggggatatatt gcacaatggg gaaaccctg atgcagcgcac gccgcgtgaa	420
ggaagaagta tctcggtatg taaacttcta tcagcaggaa agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gccgcgtta tacgttaggg gcaagcgta	540
tccggattta ctgggtgtaa aggaggcgtt gacggaaagag caagtctgtat gtgaaaggct	600
ggggcttaac cccaggactg cattggaaac tggggatctt gagtgcggg gaggtaaagcg	660
gaattccatg ttagcggtt aaatgcgtat atattaggag gaacaccagt ggcgaaggcg	720
gcttaactgga cggtaactga cggttgggtt cggaaagcggt gggagcaaac aggattagat	780
accctggtag tccacccgtt aaacgatgaa tactagggtt cgggtggcaa agccattcg	840
tgcgcgcagca aacgcataaa gtattccacc tggggagtac gttcgcaaga atgaaactca	900
aaggaatttga cggggccccg cacaagcggt ggagcatgtt gtttaattcg aagcaacgcg	960
aagaaccccta ccaagtcttgc acatccctctt gaccggcccg taacggggcc ttcccttcgg	1020
ggcagaggag acaggtgggtt catgggtgtc gtcagctgtt gtcgttagat gttgggttaa	1080
gtcccgcaac gagcgcacacc cctatcccttta gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgcac gtcacatcat catgcctt	1200
atgatttggg ctacacacgt gctacaatgg cgtaaacaaa gggaaagcgag acagcgatgt	1260
tgagcaatc caaaaataaa cgtcccgat cggactgcag tctgcaactc gactgcacga	1320
agctggaaatc gcttagatac gcgaatcaga atgtcgccgtt gaatacggttcc cgggttttgc	1380
tacacaccgc cccgtcacacc atgggagtca gtaacgcccc aagtcaatgtt cccaaacccca	1440
caggagggag ctggcgaaagg cgggaccgtt aactgggggtt aagtcaatgtt aaggttagccg	1500

-continued

tatcggagg tgccgttga tcacccctt t 1531

<210> SEQ ID NO 20  
<211> LENGTH: 1528  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 20

aacgagagg ttatcctggc tcaggatgaa cgctggggc gtgttaaca catgcaagtc	60
gagcgaagca ctttggaga ttcttcggat gatttcctt gtgactgagc ggccggacggg	120
ttagtaacgc gtggtaacc tgcctcatac agggggataa cagttagaaa tgactgctaa	180
taccgcataa gaccacggta ccgcattggta cagtggtaaa aactccggtg gtatgagatg	240
gaccgcgtc tgattaggtt gttgggtgggg taacggccta ccaagccgac gatcagttagc	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggccccagact cctacgggag	360
gcagcagtgg ggaatattgc acaatggagg aaactctgat gcagcgcacgc cgctgtgaagg	420
atgaagtatt tcggatgtt aacttctatc agcagggaa gaaaatgacgg tacctgacta	480
agaageccccg gctactaacg tgccagcagc cgccgtataa cgtagggggc aagcgttatc	540
cggatttact ggggtgaaag ggagcgtaga cggcacggc agccagatgt gaaagcccg	600
ggctcaaccc cgggactgca tttggaaactg ctgagctaga gtgtcggaga ggcaagtgg	660
attccttagt tagcggtgaa atgcgttagat attaggagga acaccagtgg cgaaggccgc	720
ttgctggacg atgactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacggcgtaa acgatgactg cttagtgtcg ggtggcaaa ccattcggtg	840
ccgcagctaa cgcaataaagc agtccacctg gggagtagtgc tcgcaagaat gaaactcaaa	900
ggaaattgacg gggaccgc caagcggtgg agcatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc tgatcttgac atcccgatga ccgcattgtt atggaaagctt ttcttcggaa	1020
catcggtgac aggtggtgca tggttgtcg cagctcgatc cgtgagatgt tgggttaagt	1080
cccgcaacga ggcacccccc tatcttcagt agccagcagg ttaagctggg cactctggag	1140
agactgccag ggataaacctg gaggaaagggtg gggatgacgt caaatcatca tgcccccttat	1200
gaccagggtt acacacgtgc tacaatggcg taaacaaaga gaagcgaact cgccgggtt	1260
agcaaatctc aaaaataacg tctcgttgc gattgttagtc tgcaactcga ctatcgaa	1320
ctggaaatcgc tagtaatcgc agatcagaat gctgcgggtga atacgttccc gggctttgtt	1380
cacacccgccc gtcacaccat gggagtcgt aacgccccgaa gtcagtgacc caaccgttaag	1440
gagggagctg ccgaagggtgg gaccgataac tgggggtgaag tcgttaacaag gtagccgtat	1500
cggaagggtgc ggctggatca cctccctt 1528	

<210> SEQ ID NO 21  
<211> LENGTH: 1537  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 21

atggagagg ttatcctggc tcaggatgaa cgctggggc atgcctataa catgcaagtc	60
gaacgaagg ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120

-continued

---

cgttaggtaac ctgccatgt gtccggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga gcgcatgtc agtatattaa agcgcccata aaggcgtgaa catggatgga	240
cctgcggcgc attagctgt tggtgaggta acggcccacc aaggcgatga tgcgtagccg	300
gcctgagagg gtaaaacggcc acattggac tgagacacgg cccaaactcc tacgggaggc	360
agcagttaggg aatttcgtc aatggggaa accctgaacg agcaatgcgcg cgtgaatgaa	420
gaagggtttc ggatcgtaaa gctctgtgt aagtgaagaa cggctatag aggaaatgct	480
atgggagtga cggtagctt ccagaaagcc acggctaact acgtgccagc agccgcgtt	540
atacgttagt ggcaagcgat atccggatc attggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctaa ccattgttaa ctatggaaac tggatgtctg	660
gagtgcgaaa gagggcgatg gaattccatg tggatcggtt aatgcgttag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctgtt ctgttaactga cactgaggca cgaaagcggt	780
gggagcaaat aggattagat accctagtag tccacgcgtt aaacgtatggaa aactaagtgt	840
tggaggaatt cagtgctgca gttaaacgaa taagtttcc gcctggggag tatgcacgca	900
agtgtgaaac tcaaaggaaat tgacgggggc cccgacaaggc ggtggatgtat gtggtttaat	960
tcaagcaac gcaagaacc ttaccaggcc ttgacatggaa aacaaataacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atgggtgtcg tcagctcgat tcgtgagatg	1080
ttgggttaag tcccccaacg agcgcaaccc ttgtcgatg ttaccagcat caagttgggg	1140
actcatgcga gactgcgggt gacaaaccgg aggaagggttgg ggtgcgttcaaatcatcat	1200
gccccctatg gcctgggcta cacacgtact acaatggcgcc ccacaaagag cagcgacaca	1260
gtgatgtgaa gcaaatctca taaagggtcgat ctcagttcgat attgaatgtt gcaactcgac	1320
tccatgaagt cggaaatcgat agtaatcgca gatcagatcgat ctgcgggttacatgtctcg	1380
ggcccttgcac acaccgcccc tcaaaccatg ggagtcagta atacccgaag ccgggtggcat	1440
aaccgttaagg agtggccgtt cggaaaggtagg accgtatgtt ggggttaatgt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccctt	1537

&lt;210&gt; SEQ ID NO 22

&lt;211&gt; LENGTH: 476

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 22

agtaacgcgt gggtaacctg cctcatacag gggataaca gttagaaatg actgctaata	60
ccgcataaga ccacgggtacc gcatggtaca gtggtaaaaa ctccgggtgt atgagatgga	120
cccgctgtcg attaggtatgt tggtgggtt acggcctacc aagccgacga tcagtagccg	180
acctgagagg gtgacggcc acattggac tgagacacgg cccagactcc tacgggaggc	240
agcagtgggg aatattgcac aatggaggaa actctgtatgc agcgacgcgc cgtgaaggat	300
gaagtatttc ggtatgttaaa cttctatcg cagggaaatggaa aatgcgttgc cctgactaag	360
aagccccggc taactacgtt ccaggcggcc cggtaatacg tagggggcaaa gcgttatccg	420
gatttactgg gtgttaaggaa agcgtatgttgc acggcaag ccagatgttgc aagcccc	476

&lt;210&gt; SEQ ID NO 23

&lt;211&gt; LENGTH: 354

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

-continued

<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 23

```
caggctggag tgcaggagag gtaagcggaa ttccctagtgt agcggtgaaa tgcgtagata      60
ttaggaggaa caccagtggc gaaggcggct tactggactg taactgacgt tgaggctcga     120
aagcgtgggg agcaaacagg attagatacc ctggtagtcc acgcggtaaa cgatgattgc     180
taggtgttagg tgggtatgga cccatcggtg ccgcagctaa cgcaataagc aatccacctg    240
gggagttacgt tcgcaagaat gaaactcaaa ggaatttgacg gggaccgcgca caagcggtgg   300
agcatgtggt ttaattcgaa gcaacgcgaa gaaccttacc aagtcttgc acatcc          354
```

<210> SEQ ID NO 24  
<211> LENGTH: 190  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 24

```
cggggggtca cccgggact gcatttggaa ctgctgagct agagtgtcgg agaggcaagt      60
gaaatttcta gtgttagcggt gaaatgcgtta gatattagga ggaacaccag tggcgaaggc    120
ggcttgcgtgg acgatgactg acgttggggc tcgaaagcgt ggggagcaaa caggattaga    180
taccctggta                                         190
```

<210> SEQ ID NO 25  
<211> LENGTH: 190  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 25

```
agggtcaacc cctggactgc attggaaact gtcaggctgg agtgccggag aggttaagcgg      60
aattcctagt gtageggta aatgcgtaga tattaggagg aacaccagtg gcgaaggcgg     120
cttactggac ggtaactgac gttgatgctc gaaagcgtgg ggagcaaaca ggattagata    180
acctggtaaa                                         190
```

<210> SEQ ID NO 26  
<211> LENGTH: 209  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 26

```
gggaagtccg tcttaagtgc ggggcttaac cccgtgagggg gaccgaaact gtgaagctcg      60
agtgtcggag agggaaagcgg aattcctagt gtacgggtga aatgcgtaga tattaggagg    120
aacaccagtg gcgaaagcgg ctttctggac gacaactgac gctgaggcgc gaaagccagg    180
ggagccaaacg ggattagata ccccagtaa                                         209
```

<210> SEQ ID NO 27  
<211> LENGTH: 203  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

## US 10,555,980 B2

227

228

-continued

&lt;400&gt; SEQUENCE: 27

tagtctgagt gatgeggggc ttaacccgt atggcggtgg atactggaaag tcttgaggtc	60
aggagaggaa aggggaattc ccagtgtgc ggtgaaatgc gtagatattg ggaggaacac	120
cagtggcga ggcgccttc tggactgtgt ctgacgctga gatgcgaaag ccagggtgc	180
aacgggatt agataaccacg gta	203

&lt;210&gt; SEQ ID NO 28

&lt;211&gt; LENGTH: 207

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 28

gatagtcggt cttaaatgtcg gggcttaccc cgtgaggggc ccgaaaactgt gaagtcgag	60
tgtcgagag gaaagcggaa ttccatgtgt agcgggtaaa tgcgtagata ttaggaggaa	120
caccagtggc gaaagcggct ttctggacga caactgacgc tgaggcgcga aagccaggg	180
agcaaacggg attagatacc acggtaa	207

&lt;210&gt; SEQ ID NO 29

&lt;211&gt; LENGTH: 424

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 29

cgtttgtcc acgctttcga gcctcacgtc agttaccgtc cagtaagccg ccttcgcccac	60
tggtgttctt cctaatatct acgcatttca ccgcataact aggaattccg cttaccttc	120
cggtactcta gattgacagt ttccaatgcg gtccccggg tgagccccgg gttttcacat	180
cagacttgcc actccgtcta cgctccctt acacccagta aatccggata acgcttgcac	240
catacgtatt accggggctg ctggcacgtt tttagccggt gtttcttagt caggtaccgt	300
catttttttc cctgctgata gagcttaca taccgaaata cttcatcgct cacggggcgt	360
cgctgcatca gggttttccc cattgtgcaa tattccccac tgctgcctcc cgttagagtt	420
tgga	424

&lt;210&gt; SEQ ID NO 30

&lt;211&gt; LENGTH: 441

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 30

tgtcacactt tcgagcatca gctgtcgtt cagtcgttca agtcgccttc gcaatcgagg	60
ttcttcgtga tatctaagca ttccatcgat acaccacgaa ttccgcctac ctctactgca	120
ctcaagacga ccagtatcaa ctgcattttt acgggtgagc cgccaaacttt cacagctgac	180
ttaatagtcc gcctacgctc cctttaacc caataaatcc ggataacgct tggatccctcc	240
gtattaccgc ggctgtggc acggagttag ccgatcctta ttctgtatggt acataaaaaa	300
agccacacgt ggctcaactt atcccatat aaaagaagt tacaacccat agggcagtc	360
tccttcacgc tacttggctg gttcagactc tcgtccattt accaatatcc ctcactgctg	420
cctcccgtag gtagtttggaa a	441

-continued

<210> SEQ ID NO 31  
<211> LENGTH: 422  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 31

```

ccgttgcac gcttcgtgc tcagtgtcag tttcagtcca gtaagccgc ttgcgcactg      60
atgttctcc taatatctac gcatttcacc gctacactag gaattccgct tacctctct      120
gcactccagt ctgacagttt caaaaaggcgtt cccagagttt agccctgggt tttcaactt      180
gacttgccat accacccatcg caccctttac acccagtaat tccggataac gcttgcggcc      240
tagtattttac cgccggctgtt ggcacgttgt tagccggggc ttcttagtca ggtaccgtca      300
ttttcttccc tgctgtataga gctttacata ccgaaataact tcttcactca cgccggcgctg      360
ctgcatcagg gttcccccca ttgtgcaata ttccccactt ctgcctcccg tggaaagttt      420
ga                                         422

```

<210> SEQ ID NO 32  
<211> LENGTH: 424  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 32

```

gcaaatgtca cgcatcgag cctcacgtca gttaccgtcc agtaagccgc ctgcgcact      60
gggtttcctc ctaatatctt cgcatttcac cgctacacta ggaattccgc ttacctctcc      120
ggcactcaag actaacagtt tccaaatgcag tccaggggtt gagcccccgc cttcacatc      180
agacttgcca gtccgtctac gtccttta caccctgtttt atccggataaa cgcttgcggcc      240
ctacgttta cgcggctgtc tggcacgttag ttagccgggg cttcttagtca aggtaccgtc      300
actatcttcc ctgctgtatag aagtttacat accgagatac ttcttccttc acgcggcgctc      360
gctgcatcagg ggttcccccc attgtgcaat attccccactt gctgcctcccg gtgggagttt      420
ggaa                                         424

```

<210> SEQ ID NO 33  
<211> LENGTH: 422  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 33

```

tgctcacgtt ttcgcgtca gcgtcagttt ctgtccagca atccgccttc gccactgggt      60
ttccctccgtt tatctacgca tttcaccgtt acacacggaa ttccgatttc ctctccagca      120
ctcaagaactt acagtttcaa atgcaggctg gaggttgagc cccctttt cacatctgac      180
ttgcaatccc gcctacacgc cctttacacc cagtaaatcc ggataaacgt tgccacctac      240
gtattaccgc ggctgtggc acgttagttt ccgtggctt ttcgtcagg accgtcattt      300
gttctcgccc tgacaaaaga agtttacaac ccgaaaggct tcttccttc acgcggcgctt      360
ctgggtcagg cttgcgcggcc ttggccaaata ttccccactt ctgcctcccg tggtagttt      420
ga                                         422

```

-continued

<210> SEQ ID NO 34  
<211> LENGTH: 419  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 34

tgtccacgtt	ttcgagctca	gcgtcagtt	tctgtccagta	agccgccttc	gccactggtg	60
ttccctctaa	tatctacgca	tttcaccgtt	acactaggaa	ttccgtttac	ccctccgaca	120
ctcttagtacg	acagtttcca	atgcagttacc	gggggttgagc	ccggggcttt	cacatcagac	180
ttgcccgcacc	gcctgcgcctc	cctttacacc	cagtaaatcc	ggataaacgtt	tgcaccatac	240
gtattaccgc	ggctgtgtgc	acgtatttag	ccgggtgtttc	ttagtcaggt	accgtcatta	300
tcttccctgc	tgatagagct	ttacataccg	aaataacttct	tgcgtcacgc	ggcgtcgctg	360
catcaggcatt	tcgcccattt	tgcaatattt	ccccactgtt	actcccgtag	gagtttggaa	419

<210> SEQ ID NO 35  
<211> LENGTH: 424  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 35

cgtttctcca	cgcttcgcgc	tcagcgtcag	ttactgtcca	gcaatccggcc	ttcgccactg	60
gtgttctcc	gtatatctac	gcatttcacc	gtacacacag	gaattccgat	tgcctctcca	120
gcactcaaga	actacagttt	caaatgcagg	ctggagggtt	agccccca	tttacatct	180
gacttgaat	cccgccatcaca	cgccctttac	accaggtaaa	tccggataac	gtttggcacc	240
tacgttattac	cgccggctgct	ggcacgtagt	tagccgtggc	ttatccgtca	ggtaccgtca	300
tttggggcgt	ccccgacaaa	agaagtttac	aacccgaaag	ccttcttcct	tcacgcggcg	360
ttgctgggtc	aggcttcgcgc	ccattgcccc	atattcccc	ctgctgcctc	cctgggaagt	420
ttgg						424

<210> SEQ ID NO 36  
<211> LENGTH: 445  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 36

atgtcctgac	ttcgtcgctc	agcgtcagtt	gtcggtccaga	aagccgcctt	cgccactgg	60
gttcctctta	atatctacgc	atttcaccgc	tacacttagga	attccgtttt	cctctccgac	120
actcgagttt	cacagtttgc	gtcccttcac	gggggttaagc	cccgcaactt	taagaccgac	180
ttgcgtatgcc	gccttcgcgc	cctttacgcc	caataattcc	ggacaacgtt	tgccacctac	240
gtattaccgc	ggctgtgtgc	acgtatttag	ccgtggcttt	ctcttacgg	accgtcagg	300
ataacgggta	ttgaccgcta	tcctgttgcgt	cccatataac	agaactttac	aacccgaaagg	360
ccgtcatcg	tcacgcggcg	ttgctccgtc	agactttcg	ccattgcgg	agattcccc	420
ctgctgcctc	ctggggaaagt	ttggaa				445

<210> SEQ ID NO 37  
<211> LENGTH: 421

-continued

<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 37

gtttgtcac gctttcgagc tcagcgtca	60
ttatcgtcca gtaagccgcc ttgcac	
gtgttctcc taatatctac gcatttcacc	120
gctacactag gaattccgct taccctccg	
acactctagt acgacagttt ccaatgcagt	180
accggggttg agccccgggc tttcacatca	
gacttgccgc accgcctgcg ctcccttac	240
accaggataaa tccggataac gcttgacca	
tacgttattac cgccgtgtct ggcacgtatt	300
tagccgggtc ttcttagtca ggtaccgtca	
ttatcttccc tgctgataga gctttacata	360
ccgaaaataact tcttcgtctca cgccgcgtcg	
ctgcatcagg ottcgccc ttgtgcaata ttccccactg	420
ctgcctcccg taggagttt	
g	421

<210> SEQ ID NO 38  
<211> LENGTH: 424  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 38

cgttgtcac gcattcgagc ctcagcgtca	60
gttaagccca gtaagccgcc ttgcac	
atgttctcc taatatctac gcatttcacc	120
gctacactag gaattccgct tacctctact	
tcactcaaga accacagttt caaatgcagt	180
ttatgggtta agcccatagt tttcacatct	
gacttgccgt cccgcctacg ctcccttac	240
accaggataaa tccggacaac gctgcgtccc	
tacgttattac cgccgtgtct ggcacgtagt	300
tagccggagc ttctctctca ggtaccgtct	
tttttgtcc ctgaagacag aggttacaa tcctaaaacc	360
ttcttccctc acgcggcattc	
gctgcatcag agtttctcc attgtgcaat attccccact	420
gctgcctccc gttaggagttt	
ggaa	424

<210> SEQ ID NO 39  
<211> LENGTH: 189  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 39

tgggcttacc cataaactgc atttgcattt	60
gtggttttt agtgaagtag aggtaaagg	
aattcctagt gtagcggta aatgcgtaga tattaggagg	120
aacatcagtgc gcgaaaggcg	
cttactgggc tttaactgac gctgaggctc	180
gaaaggcgtgg ggagcaaaca ggattagata	
cccaagtaa	189

<210> SEQ ID NO 40  
<211> LENGTH: 414  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 40

gtcagcatcg agtcacgtc agttaccgtc	60
cagtaagccg ctttcgcac tggttctt	

-continued

---

cctaataatct acgcatttca ccgtacact aggaattccg cttaccttc cggtactcta	120
gattgacagt ttccaatgca gtcgggggt tgagccccgg gtttcatat cagacttgcc	180
actccgtcta cgcccccctt acaccagta aatccggata acgcttgcac catacgatt	240
accgcggctg ctggcacgta tttagccggt gtttcttagt caggtaccgt catttttttc	300
cctgctgata gagcttatac taccgaaata cttcatecgct cacgcggcgt cgctgcatca	360
gggtttcccc cattgtgcaa tattccccac tgctgcetcc cgagggagtt tgga	414

<210> SEQ ID NO 41  
<211> LENGTH: 191  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 41

tcatcgctta cgggtggatct gcgcgggta cgggcgggct ggagtgcggt aggggagact	60
ggaaattcccg gtgtaacgggt ggaatgtgta gatatcgga agaacaccga tggcgaaggc	120
aggtctctgg gccgtcaactg acgctgagga gcgaaagcgt ggggagcga caggattaga	180
tacaacggta a	191

<210> SEQ ID NO 42  
<211> LENGTH: 196  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 42

tgaacccagg gcttaactct gggactgctt ttgaactgtc agactggagt gcaggagagg	60
taagcggaaat tcctagtgtt gcggtgaaat gcgttagatat taggaggaac atcagtggcg	120
aaggcggctt actggactga aactgacact gaggcacgaa agcgtgggaa gcaaacagga	180
ttagataccca tggtaa	196

<210> SEQ ID NO 43  
<211> LENGTH: 192  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 43

accaggcgtt aactctggga ctgttttga actgtcagac tggagtgcag gagaggttaag	60
cggaaattcct agtgttagcggt tgaaaatgcgt agatattagg aggaacatca gtggcgaagg	120
cggcttaactg gactgaaact gacactgagg cacgaaagcg tggggagcaa acaggattag	180
ataccctggta aa	192

<210> SEQ ID NO 44  
<211> LENGTH: 195  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 44

gaacccaggc cttaactctg ggactgctt tgaactgtca gactggagt caggaggg	60
aagcggaaatt cctagtgttag cggtgaaatg cgttagatatt aggaggaaca tcagtggcga	120

-continued

aggcggctta ctggactgaa actgacactg aggcacgaaa gcgtggggag caaacaggat	180
tagataaccc ggtaa	195

<210> SEQ ID NO 45  
<211> LENGTH: 418  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 45

gagtca	60
ttccat	120
tccat	180
tttccat	240
tttccat	300
tttccat	360
tttccat	418

<210> SEQ ID NO 46  
<211> LENGTH: 416  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 46

tgtca	60
tttccat	120
tttccat	180
tttccat	240
tttccat	300
tttccat	360
tttccat	416

<210> SEQ ID NO 47  
<211> LENGTH: 400  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 47

cacgtca	60
tttccat	120
tttccat	180
tttccat	240
tttccat	300
tttccat	360
tttccat	400

<210> SEQ ID NO 48

## US 10,555,980 B2

**239****240**

-continued

```

<211> LENGTH: 416
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 48

gtcagtttc gagtacgt cagttaccgt ccagtaagcc gccttcgcca ctggtgtcc      60
tcctaataatc tacgcatttc accgcgtacac taggaattcc gcttacacctt ccggtaactct    120
agattgacag tttccaatgc agtccccggg ttgagccccg ggttttcaca tcagacttgc      180
cactccgtct acgctccctt tacacccagt aaatccggat aacgcgttgc ccatacgat      240
taccgcggct gctggcacgt atttagccgg tgcttcttag tcaggtaccc tcattttctt      300
ccctgctgtat agagctttac ataccgaaat acttcatcgc tcacgcggcg tcgctgcac      360
agggttccc ccattgtgca atattccca ctgctgcctc ccgtggggag tttgga      416

<210> SEQ ID NO 49
<211> LENGTH: 421
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 49

gatgctcagc ttctcggtc agtgtcagtt tcagttccagt aagccgcctt cgccactgat      60
gttcctccta atatctacgc atttcaccgc tacactagga attccgctta cctctccctgc      120
actccagtct gacagttca aaagcagtcc cagagttaa ccctgggttt tcacttctga      180
cttgccatac cacctacgca ccctttacac ccagtaattc cgataacgc ttgcccccta      240
cgttattacccg cggctgtgg cacgtatc gccggggctt cttagtccagg taccgtcatt      300
ttcttcctg ctgatagagc ttacatacc gagataacttc ttcaactcagc cggcgtcgct      360
gcatacagggt ttccccatt gtcaatatt cccactgt gcctcccgaa ggaagttgg      420
a                                         421

<210> SEQ ID NO 50
<211> LENGTH: 418
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 50

gtgtcagtt cgtgttcgt gtcagttca gtcagttcaag ccgccttcgc cactgtat      60
cctcctaata tctacgcatt tcaccgttac actaggaatt ccgttacact ctctgtcact      120
ccagttctgac agttcaaaa gcagtccag agttaagccc tgggtttca cttctgactt      180
gcataccac ctacgcaccc ttacaccca gtaattccgg ataacgcttgc cccctacgt      240
attacccggg ctgctggcac gtatgttcc ggggttttcc agtcaggatc cgtcatttc      300
ttccctgtct atagagctt acataccgag atacttctt actcacgcgg cgtcgtcga      360
tcagggttcc cccattgtt caatattccc cactgctgcc tccctgttcc agtttgga      418

<210> SEQ ID NO 51
<211> LENGTH: 415
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

```

-continued

&lt;400&gt; SEQUENCE: 51

gtcagctcg agcctcacgt cagttaccgt ccagtaagcc gccttcgcac ctgggttcc	60
tccataatatac tacgcatttc accgctacac taggaattcc gettacctct ccggactct	120
agattgacag tttcaatgc agtccccgggg tttagccccgg gggtttcaca tcagacttgc	180
cactccgtct acgctccctt tacacccagt aaatccggat aacgcttgca ccatacgat	240
taccggggct gctggcacgt atttagccgg tgcttcttag tcaggtaccg tcattttctt	300
ccctgctgtat agagcttac ataccgaaat acttcategc tcacgccccg tcgctgcatc	360
agggtttccc ccatttgca atattccca ctgctgcctc gcttaggagt ttgga	415

&lt;210&gt; SEQ ID NO 52

&lt;211&gt; LENGTH: 418

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 52

tgtcagcttt cgagctcage gtcagttatc gtccagtaag cgccttcgc cactgggttt	60
cctcctaata tctacgcatt tcaccgctac actaggaatt cgcttaccc ctccgacact	120
ctagtagcgc agtttcaat gcagttccgg ggtttagccccggggcttca catcagactt	180
gcgcaccgc ctgcgtccc tttacaccca gtaaatccgg ataacgcttgc caccatacg	240
attaceggcg ctgctggcac gtathtagcc ggtgcttctt agtcaggtac cgtcattatc	300
ttccctgctg atagagctt acataccgaa atacttcttc gtcacgccccg cgctgctgca	360
tcaggcttc gcccattgtg caatattccc cactgctgc tcccgaaggg agtttggaa	418

&lt;210&gt; SEQ ID NO 53

&lt;211&gt; LENGTH: 416

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 53

ttcagcttc gagctcageg tcagttatcg tccagtaagc cgccttcgc actgggttcc	60
cctcctaataat ctacgcattt caccgctaca ctaggaattt cgcttacccc tccgacactc	120
tagtagcaca gtttcaatg cagttccggg gttttagccccggggcttca atcagacttgc	180
ccgcaccgc tgcgtccc ttacacccag taaatccggta taacgcttgc accatacgta	240
ttaccggcgc tgctggcacg tatttagccg gtgcttcttgc gtcaggtacc gtcattatct	300
tccctgctgat tagagcttac cataccgaaa tacttcttc gtcacgccccg gtcgtgcata	360
caggcttcg cccattgtgc aatattcccc actgctgcctt cccgagggga gtttggaa	416

&lt;210&gt; SEQ ID NO 54

&lt;211&gt; LENGTH: 435

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 54

ttcggcttc tttcccttc tcgcgcctca gtgtcagttt ctgtcttagta agccgccttc	60
gcacactgatg ttccctctaa tatctacgca cttcaccgctt ccacaatgaa ttccgcttac	120

## US 10,555,980 B2

243

244

-continued

---

ccctcccgcg ctctagtctg acagtttaa aaaaactccc cgagagaac cctgggttt	180
ttcttctgac atgcgatatac ccaccccac ccttataaca cccaaaaatc ggataaaagg	240
tgcgacctac gtattatacc ggctgctggg gcttagatag cgggggttc ttatacaggg	300
accgtcattt tctttccgc tgatacaget ttacataccg aaatacttct ttctcacgcg	360
gcgtcgctgc atcagggttt cccccattgt gcaatattcc caactgctgc ctccgaagg	420
ggaagttggg ggaaa	435

<210> SEQ ID NO 55  
<211> LENGTH: 418  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 55

gttcagctt cgagcctcac gtcagttacc gtccagtaag cgccttcgc cactgggtt	60
cctcctaata tctacgcatt tcacccgtac actaggaatt ccgccttacct ctccggta	120
ctagattgac agtttcaat gcagtcccg ggttgagccc cgggtttca catcagactt	180
gccactccgt ctacgctccc ttacaccca gtaaatccgg ataacgctg caccatacgt	240
attaccgcgg ctgctggcac gtathtagcc ggtgcttctt agtcaggta cgtcatttc	300
ttccctgctg atagagctt acataccgaa atacttcatc gtcacgcgg cgtcgctgca	360
tcagggtttc cccattgtg caaatattccc cactgctgcc tcccgagggg agtttgg	418

<210> SEQ ID NO 56  
<211> LENGTH: 416  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 56

gtcagtttc gagctcacgt cagttaccgt ccagtaagcc gccttcgcca ctgggttcc	60
tcttaatatac tacgcatttc accgcatacac taggaattcc gettacctct ccggta	120
agattgacag tttcaatgc agtcccgaggg ttgagccccc ggtttcaca tcagacttgc	180
cactccgtct acgctccctt tacacccagt aaatccggat aacgcttgca ccatacgat	240
taccgcggct gctggcacgt atttagccgg tgcttcttag tcaggtaccg tcattttctt	300
ccctgctgat agagcttac ataccgaaat acttcatcgc tcacgcggcg tcgctgcac	360
agggtttccc ccattgtca atattccca ctgctgcctc ccgagggggag tttgg	416

<210> SEQ ID NO 57  
<211> LENGTH: 417  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 57

tctcacgcct tcgagctcac gtcagtcata gtcagcaag cgccttcgc cactgggtt	60
cctcctaata tctacgcatt tcacccgtac actaggaatt ccacttgctt ctccgacact	120
ctagctcagc agttccaaat gcagtcccg ggttgagccc cgggtttca catctggctt	180
gcgcgtgcgt ctacgctccc ttacaccca gtaaatccgg ataacgctg cccctacgt	240
attaccgcgg ctgctggcac gtathtagcc ggggtttctt agtcaggta cgtcatttc	300

-continued

ttccctgctg atagaagttt acataccgaa atacttcatc cttcacgcgg cgtcgctgca	360
tcagagttc ctccattgtg caaatattccc cactgctgcc tcccgtaggg agtttgg	417

<210> SEQ ID NO 58  
<211> LENGTH: 417  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 58

gtcagcttc gagctcagcg tcagttatcg tccagtaagc cgccttcgcc actgggttgc	60
ctcctaataat ctacgcattt caccgctaca ctaggaattc cacttacccc tccgacactc	120
tagtacgaca gtttccaaatg cagttccggg gttgagcccc gggctttcac atcagacttg	180
cgcacccggcc tgccgtccct ttacacccag taaatccgga taacgcttgc accatacgt	240
ttaccggcgc tgctggcacg tatttagccg gtgcatttta gtcaggtaacc gtcattttc	300
ttccctgctg atagagctt acataccgaa atacttcttcc gtcacgcgg cgtcgctgca	360
tcagggttc cccattgtg caaatattccc cactgctgcc tcccgaggga gtttgg	417

<210> SEQ ID NO 59  
<211> LENGTH: 419  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 59

agcccccgctt tcgagcctca cgtcagttac cgtccagtaa gcccgccttcg ccactgggt	60
tccctctaattt atctacgcattt ttcaccgcta cactaggaat tccgcttacc tctccggcac	120
tcaagacggg cagttccaa tgcagtcccg gggttgagcc ccagccttc acatcagact	180
tgtccatccg tctacgctcc ctttacaccc agtaaatccg gataacgctt gcccctacg	240
tattaccgcg gctgtggca cgtagttacg cggggcttct tagtcaggta ccgtcattt	300
cttccctgct gatagaagtt tacataccga gataacttctt cttcacgcg gctcgctgc	360
atcagggttt ccccccattgt gcaatattcc ccactgtgc ctcccgaaagg aagtttgg	419

<210> SEQ ID NO 60  
<211> LENGTH: 419  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 60

tgctcagctt tcgagcctca cgtcagttac cgtccagtaa gcccgccttcg ccactgggt	60
tccctctaattt atctacgcattt ttcaccgcta cactaggaat tccgcttacc tctccggcac	120
tcgagccaga cagttccaa tgcagtccca gggtaagcc ctgggttttc acatcagact	180
tgccttgcgc tctacgctcc ctttacaccc agtaaatccg gataacgctt gcccctacg	240
tattaccgcg gctgtggca cgtagttacg cggggcttct tagtcaggta ccgtcattat	300
cttccctgct gatagagctt tacataccga aataacttctt cgctcacgcg gctcgctgc	360
atcagggttt ccccccattgt gcaatattcc ccactgtgc ctcccgaaagg gagtttgg	419

<210> SEQ ID NO 61

-continued

<211> LENGTH: 421  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 61

gttgctcagc ttgcgagctc acgtcagttt ccgtccagta agccgccttc gcccactggtg	60
ttcttcctaa tatctacgca tttcacccgtt acactaggaa ttccgcgttac ctctccggca	120
ctcgagccag acagtttcca atgcagtcaccc agggtaaagc cctgggtttt cacatcagac	180
ttgccttgcc gtctacgctc cctttacacc cagtaaatcc ggataacgct tgccccctac	240
gtattaccgc ggctgctggc acgttagttt ccggggcttc ttagtcaggt accgtcatta	300
tcttcctgc tgatagagct ttacataccg aaataacttct tcgctcacgc ggcgtcgctg	360
catcagggtt tccccattt tgcaatattt cccactgctg cctccgaag gaaagtttgg	420
a	421

<210> SEQ ID NO 62  
<211> LENGTH: 420  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 62

tgctcagctt tcgagcttag cgtagttat cgtagttaa gcccgccttcg ccactgggtgt	60
tcctcctaat atctacgcat ttcaccgcta cacttagaaat tccacttacc cctccgacac	120
tctagtagcga cagttccaa tgtagtaccg gggttgagcc ccggggcttc acatcagact	180
tgcgcgcacccg cctgcgcgtcc cttagacacc agtaaatccg gataacgctt gcaccatacg	240
tattaccgcg gctgctggca cgtatttaccc cggtgcttcc tagtcaggta ccgtcattct	300
tcttcctgc tgatagagct ttacataccg aaataacttct tcgctcacgc ggcgtcgctg	360
catcagggtt tccccattt tgcaatattt cccactgctg cctccgaag ggagtttgg	420

<210> SEQ ID NO 63  
<211> LENGTH: 446  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 63

gatgcctgg cttcgcgttc agcgtcagtt gtcgtccaga aagccgcctt cggccactgg	60
gttcctccta atatctacgc atttcaccgc tacacttagga attccgcctt cctctccgac	120
actcgagctt cacagtttcg gtccgcgttc ggggttaagc cccgcactt taagaccgac	180
ttgcgtatgcc gcctgcgcgc ccttacgc caataattcc ggacaacgct tgccacctac	240
gtattaccgc ggctgctggc acgttagttt ccgtggctt ctcttacggt accgtcagg	300
ataacgggtt ttgaccgcta tcctgttcgtt cccatataac agaactttac aacccgaagg	360
ccgtcatcgt tcacgcggcg ttgctccgtc agactttcgtt ccattgcggaa agattcccc	420
ctgctgcctc ccggggggag tttggaa	446

<210> SEQ ID NO 64  
<211> LENGTH: 419  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence

-continued

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 64

gtccccgttt	cggccctcag	cgtcagttat	cgtccagtaa	gcccgcattcg	ccactggtgt	60
tccctctaata	atctacgcat	ttcacccgcta	cacttaggaat	tccgcattacc	cctccgacac	120
tctagtagcga	cagttccaa	tgcagtaccg	gggttgagcc	ccgggcttc	acatcagact	180
tgcgcacccg	cctgcgcctcc	ctttacaccc	agtaaatccg	gataacgcct	gcaccatacg	240
tattaccgcg	gctgcgtggca	cgtatccatc	cggtgcattct	tagtcaggta	ccgtcattat	300
cttccctgt	gatagagctt	tacataccga	aataacttctt	cgctcacgcg	gcgtcgctgc	360
atcaggcttt	cggccattgt	gcaatattcc	ccactgctgc	ctcccgaaagg	gaagtttgg	419

&lt;210&gt; SEQ ID NO 65

&lt;211&gt; LENGTH: 418

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 65

tgtcagcttt	cggccatcgc	gtcagttatc	gtccagtaag	ccgcgcattcg	cactgggttt	60
cctcctaata	tctacgcatt	tcacccgctac	actaggaatt	ccgcttaccc	ctccgacact	120
ctagtagcgc	agttccaaat	gcagttccgg	gggttgagcc	ccgggcttc	catcagactt	180
gcccgcacccg	ctgcgcctcc	tttacaccca	gtaaatccgg	ataacgcctt	caccatacgt	240
attacegcgg	ctgcgtggcac	gtatccatc	ggtgcttctt	agtcaggta	cgtcattatc	300
ttccctgtct	atagagcttt	acataccgaa	ataacttctt	gctcacgcgg	cgtcgtgc	360
tcaggcttcc	gcccatgttgc	caatattccc	cactgctgcc	tcccgaaagg	agtttgg	418

&lt;210&gt; SEQ ID NO 66

&lt;211&gt; LENGTH: 420

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 66

tgctcagttt	tcgagctcag	cgtcagttat	cgtccagtaa	gcccgcattcg	ccactggtgt	60
tccctctaata	atctacgcat	ttcacccgcta	cacttaggaat	tccacttacc	cctccgacac	120
tctagtagcga	cagttccaa	tgcagtaccg	gggttgagcc	ccgggcttc	acatcagact	180
tgcgcacccg	cctgcgcctcc	ctttacaccc	agtaaatccg	gataacgcct	gcaccatacg	240
tattaccgcg	gctgcgtggca	cgtatccatc	cggtgcattct	tagtcaggta	ccgtcattatc	300
tctccctgtc	tgatagagct	ttacataccg	aaataacttct	cgctcacgc	ggcgctgc	360
catcagggtt	tcccccatttgc	caatattttc	ccactgctgc	ctcccgaaag	ggagtttgg	420

&lt;210&gt; SEQ ID NO 67

&lt;211&gt; LENGTH: 417

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 67

atcaggcttt	cggccatcgc	tcagttccgg	tccagtaagc	ccgcgcattcg	actgggtttc	60
------------	------------	------------	------------	-------------	------------	----

-continued

---

c tc c t a a t a t	c t a c g c a t t	c a c c g c t a c a	c t a g g a a t t c	c g c t t a c c c c	t c c g g c a c t c	120
a a g c a t a c c a	g t t t c c a a t g	c a g t c c a g g g	g t t a a g c c c c	t g c c t t c a c	a t c a g a c t t g	180
a t a c g c e g t c	t a c g e c t c c t	t t a c a c c c a g	t a a a t c c g g a	t a a c g c t c g c	c c c c t a c g t a	240
t t a c c g c g g c	t g c t g g c a c g	t a g t t a g c c g	g g g c t t c t t a	g t c a g g t a c c	g t c a t t a t c t	300
t c c c t g c t g a	t a g a a g t t t a	c a t a c c g a g a	t a c t t c t t c c	t t c a c g c g g c	g t c g c t g c a t	360
c a g g g t t c c	c c c a t t g t c	a a t a t t c c c c	a c t g c t g c c t	c c c g a g g g a a	g t t t g g a	417

<210> SEQ ID NO 68  
<211> LENGTH: 416  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 68

g t g t c a g c t t	t c g a g c t c a c	g t c a g g t t a c c	g t c c a g t a a g	c c g c t t c g c	c a c t g g t g t t	60
c t c c t a a t a	t c t a c g c a t t	t c a c c g c t a c	a c t a g g a a t t	c c g c t t a c c t	c t c c g g c a c t	120
c g a g c c a g a c	a g t t t c c a a t	g c a g t c c c a g	g g t t a a g c c c	t g g g t t t c a	c a t c a g a c t t	180
g c c t t g c c g t	c t a c g c t c c c	t t t a c a c c c a	g t t a a t c c g g	a t a a c g c t t g	c c c c t a c g t	240
a t t a c c g c g g	c t g c t g g c a c	g t a g t t a g c c	g g g g c t t c t t	a g t c a g g t a c	c g t c a t t a t c	300
t t c c c t g c t g	a t a g a g c t t t	a c a t a c c g a a	a t a c t t c t c	g o t c a c g c g g	c g t c g c t g c a	360
t c a g g g t t c c	c c c a t t g t g	c a a t a t t c c c	c a c t g c t g c c	t c c c g a g g g a	g t t t g g	416

<210> SEQ ID NO 69  
<211> LENGTH: 416  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 69

g c c a g c t t c g	a g c c t c a c g t	c a g t c a t c g t	c c a g t a a g c c	g c t t c g c c a	c t g g t g t t c c	60
t c t c t a a t a t c	t a c g c a t t t c	a c c g c t a c a c	t a g g a a t t c c	a c t t a c c t t c	c c g a c a c t t c t	120
a g c t g c a c a g	t t t c c a a a g c	a g t c c a c a g g	t t g a g c c c a t	g c t t t c a c t	t c a g a c t t g c	180
a c a g c c g t c t	a c g c t c c c t t	t a c a c c a g t	a a a t c c g g a t	a a c g c t t g c c	c c c t a c g t a t	240
t a c c g e g g c t	g c t g g c a c g t	a g t t a g c c g g	g g o t t c t t a g	t c a g g t a c c g	t c a t t t t c t t	300
c c c t g c t g a t	a g a a g t t t a c	a t a c c g a a a t	a c t t c a t c t	t c a c g e g g c g	t c g c t g c a t c	360
a g g c t t c g c	c c a t t g t g c a	a t a t t c c c c a	c t g c t g c c t c	c c g a g g g a a g	t t t g g a	416

<210> SEQ ID NO 70  
<211> LENGTH: 418  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 70

t g a t c a g c t t	t c g a g c t c a c	g t c a g g t t a c c	g t c c a g t a a g	c c g c t t c g c	c a c t g g t g t t	60
c c t c t a a t a	t c t a c g c a t t	t c a c c g c t a c	a c t a g g a a t t	c c g c t t a c c t	c t c c g g t a c t	120
c t a g a t t g a c	a g t t t c c a a t	g c a g t c c c g g	g g t t g a g c c c	c g g g t t t c a	c a t c a g a c t t	180
g c c a c t c c g t	c t a c g c t c c c	t t t a c a c c a	g t t a a t c c g g	a t a a c g c t t g	c a c c t a c g t	240
a t t a c c g c g g	c t g c t g g c a c	g t a t t a g c c	g g t g c t t c t t	a g t c a g g t a c	c g t c a t t t c	300

## US 10,555,980 B2

**253****254**

-continued

ttccctgctg atagagctt acataccgaa atacttcatc gtcacgcgg cgtcgctgca	360
tcagggttc cccatttgt caaatattcc cactgctgcc tccccggggg agtttgaa	418

<210> SEQ ID NO 71  
<211> LENGTH: 420  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 71

gatgatcagc ttgcagctc acgtcagtta cggccagta agccgcctc gccactggtg	60
ttccctctaa tatctacgca ttccaccgt acactaggaa ttccgcttac ctctccggca	120
ctctagaaaa acagttcca atgcagtcct ggggtaaggcc cccagcctt cacatcagac	180
ttgctttcc gtctacgctc ccttacacc cagtaaatcc ggataacgct tgccccctac	240
gtattaccgc ggctgtggc acgttagttac cggggcttc ttagtcaggt accgtcattt	300
tcttcctgc tgatagaagt ttacataccg agataacttct tccatcacgc ggctcgctg	360
catcagggtt tccccattt tgcaatattc ccactgctg cttccgggg gaagtttgaa	420

<210> SEQ ID NO 72  
<211> LENGTH: 418  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 72

ttgatcagct ttgcagctca cgtcagttac cgtccagtaa gcccgcctcg ccactgggt	60
tcctcctaattatctacgcat ttccaccgtta cactaggaat tccgcttacc tctccggcac	120
tcttagaaaa cagttccaa tgcagtcctg gggtaagcc ccagccttc acatcagact	180
tgctcttcgttctacgctcc ttccacacc agtaaatccg gataacgctt gcccctacg	240
tattaccgcg gctgtggca cgttagttac cggggcttc tagtcaggtt ccgtcatttt	300
cttccctgtt gatagaagtt tacataccgaa gataacttctt cttcacgcg gcgtcgctg	360
atcagggtttt ccccccattgt gcaatattcc ccactgctg cttccgaagg gagtttgaa	418

<210> SEQ ID NO 73  
<211> LENGTH: 421  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 73

gtatccatc ttgcagctc agcgtcagtt atcgtccagt aagccgcctt cgccactgg	60
gttcctccata atatctacgc atttcaccc tacactagga attccgctta cccctccgac	120
actcttagtac gacagttcc aatgcagttac cggggtttag ccccccgtt tcacatcaga	180
cgtccgcac cgcctgcgtt cccttacac ccagtaatcc ggataacgc ttgcaccata	240
cgttattaccgc cggctgtggc cacgtattta gcccgtgtt cttagtcagg taccgtcatt	300
atcttcctgtt ctgatagagc ttacatacc gaaataacttc ttgcgtcacg cggcgctcg	360
gcacaggctt ttcgcccattt gtcaatattt cccactgctt gctcccgaa gggagtttgaa	420

## US 10,555,980 B2

**255****256**

-continued

```

<210> SEQ ID NO 74
<211> LENGTH: 419
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 74
gctcagctt cggatccgc gtcagttatc gtccagtaag ccgccttcgc cactgggtt    60
cctcctaata tctacgcatt tcaccgctac actaggaatt ccacttaccc ctccgacact    120
ctagtacgc agtttcaat gcagtgccgg gggtgagccc cgggcttca catcagactt    180
gccgcaccgc ctgcgtccc ttacaccca gtaaatccgg ataacgctt caccatacg    240
attaccgcgg ctgctggcac gtattttagcc ggtgcttctt agtcaggta cgtcatttt    300
cttccctgct gatagagctt tacataccga aataacttctt cgctcacgcg gcgtcgctc    360
atcagggttt cccccattgt gcaatattcc ccactgctgc ctcccgaggg gagtttgg    419

<210> SEQ ID NO 75
<211> LENGTH: 417
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 75
tttcagctt gggatccgc gtcagttatc gtccagtaag ccgccttcgc cactgggtt    60
cctcctaata tctacgcatt tcaccgctac actaggaatt ccgcttaccc ctccgacact    120
ctagtacgc agtttcaat gcagtgccgg gggtgagccc cgggcttca catcagactt    180
gccgcaccgc ctgcgtccc ttacaccca gtaaatccgg ataacgctt caccatacg    240
attaccgcgg ctgctggcac gtattttagcc ggtgcttctt agtcaggta cgtcatttt    300
ttccctgctg atagagctt acataccga aataacttctt gtcacgcgg cgtcgtcga    360
tcagggtttt gccccattgt gcaatattcc ccactgctgc tcccgaaagg gtttgg    417

<210> SEQ ID NO 76
<211> LENGTH: 419
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 76
agctcagtt tcgagatccgc cgtcagttat cgtccagtaa gccccttcgc ccactgggtt    60
tcctcctaata atctacgcatt tcaccgctac cacttagaat tccgcttacc cctccgacac    120
tctagtacga cagttccaa tgcaatccgg gggttggccc cgggcttca acatcagact    180
tgccgcaccgc ctgcgtccc ctccatccgg agtaaatccgg gataacgctt gcaccatacg    240
tattaccgcgg gtcgtggca cgtatccgg cgggtgttcc tagtcaggta ccgtcattat    300
cttccctgct gatagagctt tacataccga aataacttctt cgctcacgcg gcgtcgctc    360
atcagggtttt cccccattgt gcaatattcc ccactgctgc ctcccgaaagg gagtttgg    419

<210> SEQ ID NO 77
<211> LENGTH: 418
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

```

-continued

&lt;400&gt; SEQUENCE: 77

gtccagctt	cgagcctcag	cgtcagttat	cgtccagtaa	gcccgcctcg	ccactggtgt	60	
tccctcta	atctacgc	ttcacccgcta	cacttaggaat	tccgcttacc	cctccgacac	120	
tctagtcga	cagttccaa	tgcagtaccg	gggttgagcc	ccgggcttcc	acatcagact	180	
tgcgcaccc	cctgcgc	cttacaccc	agtaaatccg	gataacgc	tt gcaccatacg	240	
tattaccgc	gctgcgtggca	cgtat	tttgc	tagtcaggta	ccgtcattat	300	
cttccctgt	gatagagctt	tacataccga	aatacttctt	cgctc	acgcg ggcgc	360	
atcaggctt	cgc	ccattgt	gcaatattcc	ccactg	tcgc	ctcccgaggg	418

&lt;210&gt; SEQ ID NO 78

&lt;211&gt; LENGTH: 193

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 78

tgagccgggc	tcaccccggt	actgcattgg	aactgtcgta	ctagagtgtc	ggaggggtaa	60
gggaaattcc	tagttagcg	gtgaaatcg	tagatattag	gaggaacacc	agtggcgaa	120
ggggcttact	ggacgataac	tgacgctgag	gctcgaaagc	gtggggagca	aacaggatta	180
gatacaccgg	taa					193

&lt;210&gt; SEQ ID NO 79

&lt;211&gt; LENGTH: 424

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 79

cacgatgtca	gtttcgagc	tcagcgtc	ttatcgcca	gtaagccgc	ttcgccactg	60
gtgttctcc	taatatctac	gcatttacc	gtcacactag	gaatccact	tacccctccg	120
acactctagt	acgacagttt	ccaatgc	accgggttg	agccccgggc	tttcacatca	180
gacttgccgc	accgcgtgc	ctcccttac	accaggtaaa	tccggataac	gcttgcacca	240
tacgttattac	cgccgctgct	ggcacgtatt	tagccgtgc	ttcttagtca	ggtaccgtca	300
ttcttcttcc	ctgctgatag	agctttacat	accgaaatac	ttcttcgctc	acgcggcg	360
gttgcatcag	ggtttccccc	attgtgcaat	attccccact	gctgcctccc	gaggggagtt	420
tgga						

&lt;210&gt; SEQ ID NO 80

&lt;211&gt; LENGTH: 1524

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 80

atggagagtt	tgatcctggc	tcaggatgaa	cgctggccgc	gtgcctaata	catgcaagtc	60
gaacgcgc	acttgtgctc	gagtggcgaa	cgggtgagta	atacataagt	aacctgc	120
agacaggggg	ataactattg	gaaacgatag	ctaagaccgc	ataggta	ccg acactgc	180
gtgaccgtat	taaaagtgcc	tcaaagcact	ggttagaggat	ggacttatgg	cgccattagct	240

## US 10,555,980 B2

**259****260**

-continued

---

ggttggcgcc accaaggcga cgatgcgtag ccgacctgag agggtaccg	300
gccacactgg gactgagaca cggcccgac tcctacgggaa ggcagcagta gggatttc	360
ggcaatgggg gaaaccctga ccgagcaacg ccgcgtgaag gaagaagggtt ttccggatgt	420
aaacttctgt tataaaggaa gaacggcgcc tacaggaaat ggtagecgag tgacggtaact	480
ttattagaaa gccacggcta actacgtcc agcagccgcg gtaatacgta ggtggcaagc	540
gttateccgga attattgggc gtaaaagggg agcagggggc agcaagggtc tgggtgaaa	600
gcctgaagct taacttcagt aagccataga aaccaggcag ctagagtgca ggagaggatc	660
gtggaattcc atgtgtagcg gtgaaatcg tagatataatg gaggaacacc agtggcgaag	720
gcgacgatct ggcctgcaac tgacgctcag tcccggaaac gtggggagca aataggatta	780
gataccctag tagtccacgc cgtaaacatgat gagtactaaatg tgggtgatgt caaagttcag	840
tgctgcagtt aacgcataaa gtactccgcc ttagatgtac gttcgcaaga atgaaactca	900
aaggaattga cgggggcccc cacaaggcggt ggagcatgtg gtttaattcg aagcaacgcg	960
aagaacctta ccaggcttgc acataactcat aaaggctcca gagatggaga gatagctata	1020
ttagatcacg gtgggtcgat gttgtcgta gctcgatcg ttagatgtg gtttaagtcc	1080
cgcacgcgc gcaaccctta tcgtagtta ccatcattaa gttggggact ctgcgcgac	1140
tgcgcgtgc aagctggagg aaggcgaaaa tgacgtcaaa tcatcatgcc ccttatgacc	1200
tgggctacac acgtgttaca atggatgggtg cagagggaaatg cgaagccgcg aggtgaagca	1260
aaaccctataa aaccattctc agttcggtt gtagtctgca actcgactac atgaagtgg	1320
aatcgctagt aatcgcaat cagcatgtcg cggtaataac gttctcggtt cttgtacaca	1380
ccggccgtca caccacgaga gttgataaca cccgaagccg gtggcctaacc gcacaggaa	1440
gagctgtcta aggtgggatt gatgatgggg gtgaagtcgt aacaaggatccctacggga	1500
acgtggggat ggatcacctc cttt	1524

&lt;210&gt; SEQ ID NO 81

&lt;211&gt; LENGTH: 1528

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 81

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaacac atgcaagtgc	60
aacgaagcaa ttgaaggaag ttttcggatg gaattcgatt gactgatgg cggacgggtg	120
agtaacgcgt ggataacctg cctcacactg gggataaca gtttagaaatg actgctaata	180
ccgcataacgc gcacagttacc gcatggtaca gtgtgaaaaa ctccgggtgt gtgagatgg	240
tccgcgtctg attagccagt tggcgaaaaa acggcccacc aaagcgacga tcagtagccg	300
acctgagagg gtgacggcc acattggacat tggatggggc acgtggggcc acataactcc tacgggaggc	360
agcagtgggg aatattgcac aatgggcgaa agcctgtatgc agcgacccgcg cgtgaatggaa	420
gaagtatttc ggtatgtaaa gctctatcg cagggaaatg aatgacggta cctgactaag	480
aagccccggc taactacgtg ccagcagccg cggtaatacg tagggggcaaa gcgttatccg	540
gatttactgg gtgtaaagggg agcgttagacg gcgaagcaag tctgaatgtaa aacccagg	600
ctcaaccctg ggactgtttt ggaaactgttt ttgtctatgt gtccggaggg taatgtggaaat	660
tccttagtgc gcggtgaaat gcgtagatat tagggggaaac accagtggcg aaggccgtt	720
actggacgat aactgacgtt gaggctcgaa agcgtggggc gcaaaccaggaa ttagataccc	780

-continued

tggtagtcca cgccgtaaac gatgaatgct aggtgttggg gggcaaagcc cttcggtgcc	840
gtcgcaaacg cagaatgcattccacctggg gagtacgttc gcaagaatga aactcaaagg	900
aattgacggg gacccgcaca agcggtgag catgtggttt aattcgaagc aacgcgaaga	960
accttaccaa gtcttacat cctcttgacc ggcgtgttaac ggccgcctcc cttcgccc	1020
agagagacag gtgggtcatg gttgtcgtca gtcgtgtcg tgagatgtg ggttaagtcc	1080
cgcacacgac gcaaccctta tccttagtag ccagcaggtt aagctggca ctctaggag	1140
actgcacagg ataaccttgg agaagggtggg gatgacgtca aatcatcatg ccccttatga	1200
tttgggtac acacgtgcta caatggcgta aacaaaggaa agcaagacag tggatgtggag	1260
caaatcccaa aaataacgtc ccagttcgga ctgttagtctg caacccgact acacgaagct	1320
ggaatcgcta gtaatcgca atcagaatgt cgccggtaat acgttccccg gtcttgcata	1380
caccgcccgt cacaccatgg gagtcagcaa cgcccgaaat cagtgaccctt actcgcaaga	1440
gagggagctg cccaaaggccgg ggcaggttac ttgggtgaag tggtaacaag gtggccgtat	1500
cggaagggtgc ggctggatca cctccctt	1528

&lt;210&gt; SEQ ID NO 82

&lt;211&gt; LENGTH: 385

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 82

aattcgacgt tgcggatt actggcgta aaggagcggt aggccgactt ttaagtgaga	60
tgtgaaatac cccggctcaa ctgggtgct gcatttcaaa ctggaaatgtt agatgcagg	120
agaggagaat ggaattccta gtgtacgtt gaaatgcgtt gagatttagg agaacaccag	180
tggcgaaggc gattctctgg actgtactg acgctggcc tcgaaacgtt ggggagccaa	240
caggattaga taccctggta gtccacgcgg taaacgttta atacttaggt taggggttgt	300
catgacctct gtgcggccgc taacgcattt agtattccgc ctggggagttt cggtcgcag	360
ataaaaactc aaagaaattt acggtt	385

&lt;210&gt; SEQ ID NO 83

&lt;211&gt; LENGTH: 434

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 83

acgagcgtat cggattattt ggttaagg agcgttaggtt gattgttaag tcagttgtta	60
aatgtttcggtt ctcaaccgtt aaattgcgtt tgaaactggc agtcttgcgtt acgttagagg	120
tgggcggaaat tcgtgggtta gcggtgaaat gcttagatattt cacgttgcgtt tccgatttgc	180
aaggcagctt actagactgtt cactgacactt gatgttcgtt aatgtgggtt tcaaacaggaa	240
ttagataacc ttgttagtcca cacagttttt gatgtttttt gatatacgtt	300
aaggcccaa gcaaaaggcat taagtattcc acctggggag tacggccggca acggtaaac	360
tcaaaatggaaat tgacggaaac ccggccagggg gggaaaaaca ttgggtttttt ttggatgtata	420
cgggggaggaa cctc	434

&lt;210&gt; SEQ ID NO 84

-continued

```

<211> LENGTH: 1457
<212> TYPE: DNA
<213> ORGANISM: Ruminococcus obaeum
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (65)..(66)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (960)..(960)
<223> OTHER INFORMATION: n is a, c, g, t or u

<400> SEQUENCE: 84

ggcggcgtgc ttaacacatg caagtcgaac gggaaacctt tcattgaagc ttccggcagat    60
ttgggnntgtt tcttagtggcg gacgggtgag taacgcgtgg gtaacctgcc ttatacaggg    120
ggataacaac cagaaaatggt tgctaataacc gcataagcgc acaggaccgc atggtctgg    180
gtgaaaaact ccgggtggat aagatggacc cgcgttggat tagctagttg gcagggtaac    240
ggcctaccaa ggcgacgatc catagccggc ctgagagggt gaacggccac attgggactg    300
agacacggcc cagactccta cgggaggcag cagtggggaa tattgcacaa tgggggaaac    360
cctgatgcgc cgacgcgcgc tgaaggaaga agtatctcg tatgtaaact tctatcagca    420
ggaagatag tgacggtacc tgactaagaa gccccggcta actacgtgcc agcagccgcg    480
gtatacgtt gggggcaagc gttatccgga tttactgggt gttaaggggag cgttagacgga    540
ctggcaagtc ttagtgtggaa ggcgggggct caacccctgg actgcattgg aaactgttag    600
tctttagtgc cggagggta agcggaaattc cttagttagc ggtgaaatgc gtatgttta    660
ggaggaacac cagtggcgaa ggcggcttac tggacggtaa ctgacgttga ggctcgaaag    720
cgtggggagc aaacaggatt agataccctg gtatcccacg ccgtaaacga tgattactag    780
gtgttggggaa gcaaagctt tcggtgccgc cgcaaacgc ttaagtatcc cacctggga    840
gtacgttcgc aagaatgaaa ctcaaaggaa ttgacggggaa cccgcacaag cggtggagca    900
tgtggtttaa ttcaagcaa cgcaagaacac cttaccaagt ttgcacatcc ctctgaccgn    960
cccttaaccg gatctttcct tcgggacagg ggagacaggt ggtgcattgt tgcgtcagc    1020
tcgtgtcgta agatgttggg ttaagtcccg caacgagcgc aacccctatc cccagtagcc    1080
agcagtccgg ctgggcactc tgaggagact gccaggata acctggagga aggccgggat    1140
gacgtcaaat catcatgccc cttatgatt gggctacaca cgtgctacaa tggcgtaaac    1200
aaagggaagc aagcctgcga aggttaagcaa atcccaaaaa taacgtccca gttcggactg    1260
cagtctgcaa ctgcactgca cgaagcttga atcgcttaga atcgccgatc agaatgcgc    1320
ggtgaatacg ttccgggtc ttgtacacac cgcggcgtcac accatggggag tcagtaacgc    1380
ccgaagttagc tgacctaact gcaaaagaagg agtcggccaa ggccggaccg atgactgggg    1440
tgaagtgcgtt acaaggt                                         1457

```

```

<210> SEQ ID NO 85
<211> LENGTH: 1398
<212> TYPE: DNA
<213> ORGANISM: Ruminococcus obaeum
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1083)..(1083)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1207)..(1207)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature

```

## US 10,555,980 B2

**265****266**

-continued

<222> LOCATION: (1233)..(1233)  
<223> OTHER INFORMATION: n is a, c, g, t or u

&lt;400&gt; SEQUENCE: 85

```

ggcgtgctta acacatgcaa gtcgaacggg aaactttca ttgaagcttc ggcagattg      60
gtctgttct agtggggac gggtgagtaa cgcggtggta acctgcctta tacaggggga     120
taacaaccag aatgggtgc taataccgca taagcgcaca ggaccgcatt gtctggtgtg     180
aaaaactccg gtggataag atggccccgc gttggattag ctagttggca gggtaacggc     240
ctaccaaggc gacgatccat agccggcctg agagggtgaa cggccacatt gggactgaga    300
cacgccccag actcctcggtt aggcagcagt gggaaatatt gcacaatggg ggaaaccctg   360
atgcagcgcac gccgegtgaa ggaagaagta tctcggtatg taaacttcta tcagcaggaa  420
agatagtgac ggtacactgac taagaagccc cgkctaacta cgtgccagca gccgeggtaa  480
tacgttagggg gcaagcgtta tccggattta ctgggtgtaa agggagcgtt gacggactgg  540
caagtctgtat gtgaaaggcg ggggctcaac ccctggactg cattggaaac tgtagtctt  600
gagtgccgga gaggtaagcg gaattcttag tgtagcggtg aatgcgttag atattaggag  660
gaacaccagt ggcgaaggcg gcttactgga cggtaactgaa cgttgaggct cgaaagcgtg  720
gggagcaaac aggattagat accctggtag tccacgcgcg aaacgatgaa tactaggtgt  780
tggggagcaa agcttctcggtt tgccggccca aacgcattaa gtattccacc tggggagttac 840
gttcgcagaatgaaactca aaggaaattga cggggaccccg cacaagcgtt ggagcatgtg  900
gttaattcg aagcaacgcg aagaacctta ccaagtctt acatccctt gaccgtccct  960
taaccggatc ttcccttcgg gacagggag acaggtggtg catggttgtc gtcagctgt 1020
gtcgtgagat gttgggttaa gtcggcaac gagcgcacc cctatccca gtggccagca 1080
gtncggctgg gcactctgag gagactgcca gggataacct ggaggaaggc gggatgacg 1140
tcaaatacatc atgcccctta tgattgggc tacacacgtt ctacaatggc gtaaacaag 1200
ggaagcnagc ctkcraggt aagcaaattcc canaaataac gtcccgatc ggactgcgt 1260
ctgcaactcg actgcacgaa gctggaatcg ctagtaatcg cggatcggaa tgccgcgtt 1320
aatacgttcc cgggtcttgc acacaccgcg cgtcacacca tgggagtcgtt taacgcccga 1380
agtcaagtgcg ctaactgc                                         1398

```

<210> SEQ ID NO 86  
<211> LENGTH: 1472  
<212> TYPE: DNA  
<213> ORGANISM: Clostridium disporicum

&lt;400&gt; SEQUENCE: 86

```

gttcaggacg aacgctggcg gcgtgcctaa cacatgcaag tcgagcgagt tgattcttt  60
cggagatgaa gctagcgccg gacgggtgag taacacgtgg gcaacctgcc tcataagggg 120
gaatagccctc cccaaaggaa gattaatacc gcataagatt gtagcttcgc atgaagttagc 180
aattaaagga gcaatccgtt atgagatggg cccgcggccg attagctgtt tggtgaggta 240
acggctacc aaggcgacga tgctgttgcg acctgagagg gtgtatcgcc acattggac 300
tgagacacgg cccagactcc tacgggaggc agcagtgggg aatattgcac aatggggaa 360
accctgtatgc agcaacgcgg cgtgtgtat gacggccttc ggggtgtaaa gctctgttt 420
cagggacgtt aatgacggta cctgaggagg aagccacggc taactacgtt ccagcaggcg 480
cggttaatacg taggtggcga gctgtgtccg gatttactgg gctgtaaaggc agcgttaggc 540

```

## US 10,555,980 B2

267

268

-continued

---

gacttttaag tgagatgtga aatacccgaa	ctcaacttgg gtgctgcatt tcaaactgg	600
agtcttagtgc gaggaggagg agaatggaa	tccatgtgtc ggggtgaaat gcgttagagat	660
taggaagaac accagtggcg aaggcgattc	tctggactgt aactgacgct gaggttcgaa	720
agcgtgggg gcaaaccaggat ttatatacc	ttgttagtcca cgccgtaaac gatgaatact	780
aggtagtggg gttgtcatga cctctgtgcc	gcccctaactg catatgttat tccgtctggg	840
gagtaeggc gcaagattaa aactcaaagg aattgacggg	ggcccgacaca agcagcggag	900
catgtggttt aattcgaagc aacgcgaaga	accttaccta gacttgacat ctccgtatt	960
accgcgtact gggggagcca ctccgtggc	aggaagacag gtgggtcatg gttgtcgta	1020
gtctcggtcg tgagatgtt ggttaagtcc	cgcaacgacg gcaaccctta ttgttagtt	1080
ctaccatata gttgacact ctacgtggac	tgccgggtt aaccgggagg aaggtgggg	1140
tgacgtcaaa tcatcatgcc ctttatgtct	agggctacac acgtgttaca atggcaagta	1200
caaagagaag caagaccgcg aggtggagca	aaactcaaaa acttgtctca gttcggattt	1260
taggctgaaa ctgcgttaca tgaagcttgg	gttgcttagta atcgcgttac agcatgtcg	1320
gggttaatacg ttccccggcc ttgttacacac	cgcccggtcac accatgagag ttggcaatac	1380
ccaaacgttacg tgatctaacc cgcaagggag	gaagcgttcc aaggtgggtt cagcgatgg	1440
gggtgaagtgc taacaaggta gcccgttggaa	aa	1472

&lt;210&gt; SEQ ID NO 87

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Clostridium scindens

&lt;400&gt; SEQUENCE: 87

gagagtttga ttctggctca ggttgaacgc tggcggcg	cctaacacat gcaagtgc	60
cgaaggcgct ggccccgact tcttcggaaac	gaggagcctt gcgactgagt ggcggacgg	120
ttagttaacgcg gtggcaacc	tgccttgcac tggggataaa cagccagaaa tggctgtt	180
taccgcataa gaccgaagcg ccgcattggcg	cggccggccaa agcccccggcg	240
ggcccggttc tgattaggta gttggcg	ggggtaacccca ccaagccgac gatcgttac	300
cgacctgaga ggggtgaccgg ccacattttgg	acttgagacac ggcccagact cctacggag	360
gcagcgttgg ggaatatgtc acaatggggg	aaaccctgtt gcaagcgttgc cgcgttgg	420
atgaagtatt tcggatgtt aacttctatc	agcaggaaag aagatgttgc tacgttact	480
agaaggcccg gcttaacttacg tgcgcgttgc	cgccgttataa cgttgggggc aagcgatgtt	540
cggtttaactt ggggtttaaag ggagcgttgc	cggttgcgttca agccatgtt gaaagccgg	600
ggctcaacc cgggacttgc tttggacttgc	cgttggcttgc gtgttgggaa ggcaggccgg	660
attccctatgt tagcgttgc	atgcgttgc attaggaggaa acaccatgtt cgaaggccgg	720
ctgcgttgc	atgcgttgc ttgggttgc aaagcggtt gggcaacac gatgttac	780
cttggatgtc caccgttgc	acgttgcacta ctaggttgc ggttggcaagg ccatttgc	840
cccgccgttgc	cgcaataatgt agtccatgtt ggggttgcgtt tgcgttgc gaaactcaaa	900
ggaaatttgcg	gggaccggca caagcggtt gcatgttgc ttatccgtt gcaacgcgaa	960
gaacccatcc	tgtatccgttgc atcccgatgc caaaggcggtt aacgcgttgc ttcttgc	1020
catcggttgc	agggttgcgtt cttgttgcgtt cgttgcgtt ggttgcgtt gtttgcgtt	1080
cccgcaacgc	gcgttgcgtt tatccgttgcgtt agccatgttgc ttggatgggc actcttgcgaa	1140
gacttgcgtt	gagaacccttgc aggaagggtt ggttgcgtt gtttgcgtt gtttgcgtt	1200

## US 10,555,980 B2

**269****270**

-continued

accagggcta cacacgtgct acaatggcgt aaacaaaggg aggcgaaccc gcgagggtgg	1260
gcaaataccca aaaataacgt ctcagttcg attgtagtct gcaactcgac tacatgaagt	1320
tggaatcgct agtaatcgcg aatcagaatg tcgcggtgaa tacgttcccg ggtttgtac	1380
acaccggccg tcacaccatg ggagtctcgt aacgcccgaag ccggtgaccc aacccgtaa	1440
ggagggagcc gtcgaagggtg ggaccgataa ctggggtgaa gtcgtaacaa ggttagccgt	1500
tcggaagggtg cggctggatc acctcccttc	1529

<210> SEQ ID NO 88  
<211> LENGTH: 1456  
<212> TYPE: DNA  
<213> ORGANISM: Anaerostipes caccae

&lt;400&gt; SEQUENCE: 88

gcgcctaata catgtcaagt cgaacgaagc atttaggatt gaagtttcg gatggattc	60
ctatatgact gagtggcgga cgggtgagta acgcgtgggg aacctgcct atacaggggg	120
ataacagctg gaaacggctg ctaataccgc ataagcgcac agaatcgcat gattcagtgt	180
gaaaagccct ggcagtatag gatggtcccg cgtctgatta gtgggttgtt gaggtaacgg	240
ctcaccaagg cgacgatcag tagccggctt gagagagtga acggccacat tgggactgag	300
acacggccca aactcttacg ggaggcagca gtggggata ttgcacaatg ggggtaaacc	360
ctgatgcagc gacgcccgcgt gagtgaagaa gtatttcggt atgtaaagct ctatcagcag	420
ggaagaaaac agacggtacc tgactaagaa gccccggcta actacgtgcc agcagccgcg	480
gtaatacgtt gggggcaage gttatccgga attactgggt gttaaagggtg ctaggtggc	540
atggtaagtc agaagtgaaa gcccgggct taacccgggg actgcgtttt aaactgtcat	600
gctggagtgc aggagaggtt agcggaaattc ctagtgttagc ggtgaaatgc gtatgttatt	660
ggaggaacac cagtggcgaa ggcggcttac tggactgtca ctgacactga tgcaegaaag	720
cgtggggagc aaacaggatt agataccctg gtagtccacg ccgtaaacgta tgaataactag	780
gtgtcggggc cgttagggct tcgggtccgc agcaaacgca gtaagtattt cacctgggg	840
gtacgttcgc aagaatgaaa ctcaaaggaa ttgacggggc cccgcacaag cggtggagca	900
tgtggtttaa ttcaagacaa cgcgaagaac cttacctggt cttgacatcc caatgacgca	960
accttaaccg gtttttctt tcgagacatt ggagacagg ggtgcattt tgcgtcagc	1020
tctgtcggtg agatgttggg ttaagtcccg caacgagcgc aacccttatac ttttagtagcc	1080
agcatttaag gtgggcactc tagagagact gccaggata acctggagga aggtggggac	1140
gacgtcaaat catcatgccc cttatggcca gggctacaca cgtgctacaa tggcgtaaac	1200
aaaggaaagc gaagtctgtt ggcgaagcaa atcccaagaaa taacgtctca gttcgattt	1260
tagtctgcata ctcgactaca tgaagctggta atcgcttagta atcgtgaatc agaatgtcac	1320
ggtgaataacg ttcccggttcc ttgtacacac cggccgtcac accatggggag tcagtaacgc	1380
ccgaagtcag tgacccaaacc gcaaggaggg agctgcccga ggtggggaccc ataactgggg	1440
tgaagtcgtt acaagg	1456

<210> SEQ ID NO 89  
<211> LENGTH: 1456  
<212> TYPE: DNA  
<213> ORGANISM: Marvinbryantia formatexigens

&lt;400&gt; SEQUENCE: 89

## US 10,555,980 B2

271

272

-continued

---

tggccgcgtg	cttaacacat	gcaagtgcag	cgaagcattt	taaatgaagt	tttcggacgg	60
aatttaaaat	gactgagcgg	cggacgggtg	agtaacgcgt	ggataacctg	ccttatacag	120
ggggataaca	gccagaaatg	gctgctaata	ccgcataagc	gcacggtacc	gcatggtaca	180
gtgtaaaaaa	ctccggtggt	ataagatggg	tccgcgttgg	attaggcagt	tggcgggta	240
aaggcccacc	aaaccgacga	tccatagccg	gcctgagagg	gtggacggcc	acattgggac	300
tgagacacgg	cccagactcc	tacgggaggc	agcagtgggg	aatattgcac	aatggggaa	360
accctgatgc	agcgacgccc	cgtgggtgaa	gaagtatttc	ggtatgtaaa	gccctatcag	420
cagggaaagaa	aatgacggta	cctgaccaag	aagccccggc	taactacgtg	ccagcagccg	480
cggtaatacag	tagggggcaa	gcgttatccg	gatttactgg	gtgtaaaggg	agcgttagacg	540
gccccatgcaag	tctggtgtga	aaggcggggg	ctcaaccccc	ggactgcatt	ggaaactgta	600
tggcttgagt	gccggagagg	taagcggaat	tcctggtgta	gcgggtaaaat	gcgttagat	660
caggaggaac	accagtggcg	aaggccgctt	actggacggt	aactgacgtt	gaggctcgaa	720
agcgtggggaa	gcaaacagga	ttagatacc	tggtagtc	cgccgtaaac	gatgaatacc	780
aggtgtcggg	ggacacggtc	cttcgggtgcc	gcagcaaaacg	cactaagtat	tccacctgg	840
gagtaacgttc	gcaagaatga	aactcaaagg	aattgacggg	gaccgcaca	agcggtgag	900
catgtggttt	aattcgaagc	aacgcgaaga	accttaccag	gtcttgacat	ccggacgacc	960
ggacagtaac	gtgtccccc	cttcggggcg	tccgagacag	gtgggtcatg	gttgcgtca	1020
gctcgtgtcg	tgagatgtt	ggtaagtcc	cgcaacgagc	gcaacccctg	ttcccagtag	1080
ccagcattca	ggatgggcac	tctggggaga	ctgccaggga	taacctggag	gaaggcgggg	1140
atgacgtcaa	atcatcatgc	cccttatgt	ctggctaca	cacgtgctac	aatggcgtga	1200
acagaggaa	gccaacccgc	gagggggago	aaatcccaga	aataacgtcc	cagttcggat	1260
tgtagtcgc	aaccggcta	catgaagctg	gaatcgctag	taatcgccg	tcagcatgcc	1320
gcgggtgaata	cgccccggg	tcttgtacac	accgcccgtc	acaccatggg	agtccgaaat	1380
gccccgaagtc	agtgacccaa	ccggaaggag	ggagctgccc	aaggcggggc	cggttaactgg	1440
ggtgaagtcg	taacaa					1456

&lt;210&gt; SEQ ID NO 90

&lt;211&gt; LENGTH: 1568

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Lactobacillus mucosae

&lt;400&gt; SEQUENCE: 90

agagtttgat	cctggctcag	gatgaacgcc	ggcggtgtgc	ctaatacatg	caagtgcAAC	60
cggttggccc	aactgattga	acgtgcttgc	acggacttga	cgttggtttta	ccagcgagt	120
gccccacgggt	gagtaacacg	taggtaacct	gccccaaacg	gggggataac	atttggaaac	180
agatgctaat	accgcataac	aatttgaatc	gcatgattca	aatttaaaag	atggcttcgg	240
ctatcactt	gggatggacc	tgccgcgc	tagttgtt	gtagggtaac	ggcctacaa	300
ggctgtgtat	cgtagccgag	ttgagagact	gatccgcac	aatggaaactg	agacacggtc	360
catactccct	cgggaggcag	cagtagggaa	tcttccacaa	tgggcgc	cctgtatgg	420
caacaccgcg	tgagtgaaga	agggttcgg	ctcgtaaagc	tctgttggta	gagaagaacg	480
tgcgtgagag	caactgttca	cgcagtgcac	gtatctaacc	agaaagtca	ggcttaactac	540
gtgccacgcag	ccgcggtaat	acgttaggtgg	caagcgttat	ccggatttat	tggcgtaaa	600
gcgagcgcag	gcggtttgc	aagtctgtat	tgaaagcctt	tggcttaacc	aaagaagtgc	660

-continued

atcgaaaact gtcagacttg agtgcagaag aggacagtgg aactccatgt gtacgggtgg	720
aatgcgtaga tataatggaag aacaccaggc gcgaaggcgg ctgtctggc tgcaactgac	780
gctgaggctc gaaagcatgg gtagcaca ggattagata ccctggtagt ccatgccgt	840
aacgatgagt gcttaggtttt ggagggtttc cgcccttcag tgccgcagct aacgcattaa	900
gcactccgcc tggggagtagc gaccgcaagg ttgaaactca aaggaattga cggggccccg	960
cacaaggcgt ggagcatgtg gtttaattcg aagctacgcg aagaacctt ccaggcttg	1020
acatcttgcg ccaaccctag agataggcg tttccttcgg gaacgcata gacagggtgt	1080
catggtcgtc gtcagctcg gtcgtgagat gttgggttaa gtcccgcac ggcgcacacc	1140
cttggtaacta gttgecagca ttca gttttggg cactctagtg agactgccgg tgacaaaccg	1200
gaggaaggtg gggacgacgt cagatcatca tgccccttt gacacggctt acacacgtgc	1260
tacaatggac ggtacaacga gtcgcgaact cgccggggca agctaatttc ttaaaaccgt	1320
tctcagttcg gactgcaggc tgcaactcgc ctgcacgaa tgccatcgc tagtaatcgc	1380
ggatcagcat gccgggggtga atacgttccc gggccttgcg cacaccggcc gtcacaccat	1440
gagagttgc aacacccaaa gtcgggtggg taacccttcg gggagctagc cgccataaggt	1500
ggggcagatg attagggtga agtgcgtaca aggttagccgt aggagaacct gcggctggat	1560
cacctctt	1568

<210> SEQ ID NO 91  
 <211> LENGTH: 1508  
 <212> TYPE: DNA  
 <213> ORGANISM: Turicibacter sanguinis

<400> SEQUENCE: 91	
agagtttgcatggctcag gatgcgtac ggcggcgtgc ctaatacatg caagtcgagc	60
gaaccacttc ggtggtgagc ggccaacggg ttagttaacac ttaggttata tgcccatcag	120
acggggacaa cgatggaaa cgatcgctaa taccggatag gacgaaagg taaagggtgc	180
tccggaccac ttagggatga gctgcggcg cattagtagt ttggtagggt aaaggctac	240
caagggcgcg atgcgtagcc gacgtgagag ggtgaacggc cacactgggca ctgagacacg	300
gcccagactc ctacgggagg cagcgttgcgaaatccgg caatggcga aagcctgacc	360
gagcaacgccc gcgtaatga tgaaggcatt cgggttgtaa aattctgtta taagggaaga	420
atggctctag taggaaatgg cttagtgcgtg acggtaacctt atgagaaagc cacggctaac	480
tacgtgccag cagccgcgt aatacgtagg tggcgagcgt tatccgaaat tattggcgt	540
aaagagcgcg caggtgggg attaagtctg atgtgaaagc ccacggctta accgtggagg	600
gtcattggaa actggtaaac tttagtgcgtg aagaggaaatgg tggaaatccatgtgttagcgg	660
tgaaatgcgt agagatatgg aggaacacca gttggcgtgg cggcttcctg gtctgtact	720
gacactgagg cgcgtaaagcg tggggagcaa acaggattag ataccctggt agtccacgcc	780
gtaaacgatg agtgcgtaaatgg tgggggggttc gaaacctcgat gctgtggatca acgcattaa	840
cactccgcctt gggggatcgt gtcgcgtac gtcgtac gtaaaactcaaa aggaatttgc gggggccgc	900
acaaggcgtg gagcatgtgg tttaattcgaa agcaacgcgtt agaacccttac cagggtctgt	960
cataccaggctt accgtccatg agataggatt ttcccttcgg ggacaatggc tacagggtgt	1020
gcatgggtgtt cgtcgtcgtcgt gtcgtgtgtt aatcccgcaac cgacgcgtac	1080
ccctgtcggtt agttggccagc attcgtttgg ggactctaaac gagactgcgtt gtgacaaactt	1140

-continued

---

ggaggaaggt ggggatgacg tcaaatacgc atgcccctta tgacctggc tacacacgtg	1200
ctacaatggt tggtacaaag agaagcgaag cggtaactg gagcaaacct cataaaaggca	1260
atctcagttc ggattttagg ctgcaactcg cctacatgaa gttgaaatcg ctagtaatcg	1320
cgaatcagca tgtcgccccg aatacgttcc cgggtcttg acacaccgcc cgtcacaccca	1380
cggaggtta caacacccga agtcagtggc ctaaccgaa ggagggagct gcctaagggt	1440
gggttagatga ttgggtgaa gtcgtaacaa ggtatcccta ccggaaagggtg gggttggatc	1500
acctcctt	1508

&lt;210&gt; SEQ ID NO 92

&lt;211&gt; LENGTH: 1456

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Roseburia faecis

&lt;400&gt; SEQUENCE: 92

gatgaacgct ggccggcgtgc ttaacacatg caagtcgaac gaagcactct atttgattt	60
cttcggaaat gaagatttt tgactgagtg gccggacgggt gagtaacgcg tgggtaacct	120
gcctcataca gggggataac agttggaaac gactgctaatt accgcataag cgccacaggat	180
cgcacatgatcc ggtgtgaaaa actccgggtgg tatgagatgg acccgcgtct gattagccag	240
ttggcagggt aacggcctac caaagcgcacg atcagtagcc gacctgagag ggtgaccggc	300
cacattggga ctgagacacg gcccaaactc ctacgggagg cagcagtggg gaatattgca	360
caatggggaa aaccctgtatc cagcgcacgcg gcgtgagcga agaagtattt cggtatgtaa	420
agctctatca gcagggaaaga agaatgacgg tacctgacta agaagcaccc gctaaatacg	480
tgcgcagcgcgc cgccgtataa cgtatgggtgc aagcgttatc cggttactt ggggttaaag	540
ggagcgcagg cgggtggca agtctgatgt gaaagccccgg ggctcaaccc cggtactgca	600
ttggaaactg tcgtactaga gtgtcgagg ggttaagtggg attcctagtg tagcggtgaa	660
atgcgttagat attaggagga acaccagtgg cgaaggcggc ttactggacg ataactgacg	720
ctgaggctcg aaagcgtggg gagcaaacag gattagatac cctggtagtc cacgcgtaa	780
acgatgaata ctatgttcg gggagcattt ctcttcgggtt ccgcagcaaa cgcaataagt	840
attccacctg gggagtagt tcgcaagaat gaaactcaaa ggaattgacg gggacccgca	900
caagcgggtgg agcatgtggg ttaattcgaa gcaacgcgaa gaaccttacc aagtcttgc	960
atcccgatga cagagtatgt aatgtacytt ctcttcggag catcggtgac aggtggtgca	1020
tgggttgtcg cagctcgatgt cgttagatgt tgggttaagt cccgcaacga ggcacacccc	1080
tgtccttagt agccagcggt tcggccggc actctaggaa gactgcccagg gataacctgg	1140
aggaaggcggg ggatgacgatc aaatcatcat gccccttatg acttgggtca cacacgtgt	1200
acaatggcgt aaacaaagg aagcggagcc gtgaggccga gcaaatctca aaaataacgt	1260
ctcagttcg actctgtatc gcaacccgac tacacgaacg tggaaatcgatc agtaatcgca	1320
gatcagaatg ctgcgtgaa tacgttcccg ggtcttgatc acaccgccc tcacaccatg	1380
ggagttggaa atgcccgaag tcagtgaccc aaccgcaagg agggagctgc cgaaggcagg	1440
ttcgataact ggggtg	1456

&lt;210&gt; SEQ ID NO 93

&lt;211&gt; LENGTH: 1465

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Flavonifractor plautii

&lt;400&gt; SEQUENCE: 93

## US 10,555,980 B2

277

278

-continued

cgctggggc gtgcttaaca catgcaggc gaacgggtg ctcatgacgg aggattcg	60
caatggatg agttacctag tggcgacgg gtgagtaacg cgtgaggaac ctgccttgg	120
gagggaaata acactccgaa aggagtgcta ataccgatg aagcagttgg gtcgcattgc	180
tctgactgcc aaagatttat cgctctgaga tggcctcgcc tctgatttagc tagtagggcg	240
ggtaacggcc cacctaggcg acgatcagta gccggactga gaggttggacc ggccacattg	300
ggactgagac acggcccgaa ctccctacggg aggcaagcgat gggaaatattt gggcaatggg	360
cgcacgcctg accccagcaac gcccgtgaa ggaagaaggc ttccgggttg taaaactttt	420
ttgtcggggaa cgaaacaaat gacggatccc gacgataag ccacggctaa ctacgtgcca	480
geagcccgccg taatacgttag gtggcaagcg ttatccggat ttactgggtg taaaggcg	540
gtaggcgaaaa ttgcaagtca gatgtggaaa ctgggggatc aaccccttgc ctgcatttg	600
aactgttagtt cttagtgct ggagaggcaa tcggaaattcc gtgtgttagcg gtgaaatgcg	660
tagatatacgc gaggaaacacc agtggcgaaag gccggattgtc ggacagtaac tgacgctgag	720
gcccggaaacgt gtggggagca aacaggatta gataccctgg tagtccacgc cgtaaacgat	780
ggataactagg tgggggggtt ctgacccctt ccgtggccgaa gttaacacaa taagtatccc	840
acctggggag tacgatcgca aggttggaaac tcaaaggaaat tgacgggggc cccacaaggc	900
gggtggatgt gtggtttaat tcggaaacac gccaagaacc ttaccaggc ttgacatccc	960
actaacgagg cagagatgcg ttaggtgccc ttccggggaa gtggagacag gtgggtcatg	1020
tttgtgtca gtcgtgtcg tgatgttg ggttaagtcc cgcaacgagc gcaaccctta	1080
ttgttagttt ctacgcaaga gcactctagc gagactgccc ttgacaaaac ggaggaaagg	1140
ggggacgacg tcaaatacgtt atccccctt tgcctggcc cacacacgta ctacaatgg	1200
gggttaacaga gggaggcaat accgcgaggt ggagcaaatc cctaaaagcc atcccgatc	1260
ggattgcagg ctgaaaccccg cctgtatgaa gttggaaatcg ctatgtatcg cggatcagca	1320
tgccgcgtgtt aatacggtcc cggggcttgc acacaccggc cgtcacacca tgagatcg	1380
gaacaccccg agtccgttagc ctaaccgaa ggaggcgcc gcccggatgg ggttcgataa	1440
ttggggtgaa gtcgtacaa ggtag	1465

&lt;210&gt; SEQ ID NO 94

&lt;211&gt; LENGTH: 1438

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Blautia wexlerae

&lt;220&gt; FEATURE:

&lt;221&gt; NAME/KEY: misc\_feature

&lt;222&gt; LOCATION: (19)..(19)

&lt;223&gt; OTHER INFORMATION: n is a, c, g, t or u

&lt;400&gt; SEQUENCE: 94

caagtcgaac gggaaattttt ttattgttac ttccggatgtat ttaattttat tcttagtggcg	60
gacgggtgag taacgcgtgg gtaaccccttgc ttatacagggg ggataacagt cagaaatggc	120
tgcataatacc gcataaggcgc acagagctgc atggctcagt gtggaaaact ccgggtggat	180
aagatggacc cgcgttggat tagttgttg gtggggtaac ggcccaaccaa ggcgcacgtc	240
catagccggc ctgagagggtt gaaacggccac attgggactg agacacggcc cagactctta	300
cgggaggcg cagtggggaa tattgcacaa tggggggaaac cctgtatgcag cgacgcccgc	360
tgaaggaaaga agtatctcggtatgtaaact tctatcagca gggaaatgtatcg tgcgttacc	420
tgactaagaa gccccggcta actacgtgcc agcagccgcg gtaatacgtt gggggcaagc	480

## US 10,555,980 B2

279

280

-continued

---

gttatccgga	tttactgggt	gtaaaggag	cgttagacgg	gtggcaagt	tgtatgtgaaa	540
ggcatgggc	taaacctgtgg	actgcattgg	aaactgtcat	acttgagtgc	cgagggggta	600
agcggaaatc	ctagtgttagc	ggtgaaatgc	gtagatatta	ggaggaacac	cagtggcgaa	660
ggcggtttac	tggacggtaa	ctgacggttga	ggctcgaaag	cgtggggagc	aaacaggatt	720
agataccctg	gtagtcacg	ccgtaaaacga	tgaataacta	ggtgtcgggt	ggcaaaggcca	780
ttcgggtccg	tcgcaaacgc	agtaagtatt	ccacctgggg	agtacggtcg	caagaatgaa	840
actcaaagga	attgacgggg	acccgcacaa	gcccgtggagc	atgtggtttta	attcgaagca	900
acgcgaagaa	ccttaccaag	tcttgacatc	cgcctgacc	atcctaacc	ggatcttcc	960
ttcgggacag	gcgagacagg	tggtgcatgg	ttgtcgtagc	ctcgtagtgc	gagatgttgg	1020
gttaagtccc	gcaacgagcg	caacccctat	cctcagtagc	cagcatttaa	ggtgggcact	1080
ctggggagac	tgccagggat	aacctggagg	aaggcgggg	tgacgtaaa	tcatcatgcc	1140
ccttatgatt	tggctacac	acgtgctaca	atggcgtaaa	caaaggaaag	cgagattgt	1200
agatgggca	aatccaaaaa	ataacgtccc	agttcggact	gtagtcgca	acccgactac	1260
acgaagctgg	aatcgctagt	aatcgccgat	cagaatgccc	cggtgaatac	gttcccgggt	1320
cttgtacaca	ccgccccgtca	caccatggga	gtcagtaacg	cccgaaagtca	gtgacctaac	1380
tgcaaagaag	gagctgccga	aggcgggacc	gatgactggg	gtgaagtctgt	aacaaggt	1438

&lt;210&gt; SEQ ID NO 95

&lt;211&gt; LENGTH: 1365

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Anaerotruncus colihominis

&lt;400&gt; SEQUENCE: 95

aacggagctt	acgttttgaa	gttttccgat	ggatgaatgt	aagcttagtg	gcccggcgggt	60
gagtaaacacg	ttagcaaccc	gcctttcaga	ggggataac	agccggaaac	ggctgctaat	120
accgcacat	gttgcgggg	cacatgcccc	tgcaacccaa	ggagcaatcc	gctgaaagat	180
gggctcgcgt	ccgattagcc	agttggcggg	gtaacggccc	accaaagcga	cgatcggtag	240
ccggactgag	aggttgaacg	gccacattgg	gactgagaca	cgccccagac	tcctacggga	300
ggcagcagt	ggggatattg	cacaatgggc	gaaagcctga	tgcagcgaacg	ccgcgtgagg	360
gaagacggtc	ttcggattgt	aaacctctgt	cttggggaa	gaaaatgacg	gtacccaaag	420
aggaagctcc	ggctaactac	gtgcacggcag	ccgcggtaat	acgttagggag	caagcggt	480
ccggaattac	tgggtgtaaa	gggagcgtag	gccccatggc	aagtagaaatg	ttaatccat	540
cggctcaacc	ggtggctgcg	ttctaaactg	cggttcttga	gtgaagtaga	ggcaggcgg	600
attcctagtg	tagcggtgaa	atgcgttagat	attaggagga	acaccatgg	cgaaggcggc	660
ctgctggct	ttaactgacg	ctgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	720
cctggtagtc	cacgcgtaa	acgtatgtt	ctaggtgtgg	ggggactgac	cccttccgt	780
ccgcagttaa	cacaataagt	aatccaccc	gggagtacgg	ccgcaagg	tt gaaactcaaa	840
ggaatttacg	ggggcccgca	caagcgtgg	agtatgtgt	ttaattcgaa	gcaacgcgaa	900
gAACCTTAC	AGGTCTTGAC	ATCGGCGTAA	TAGCCTAGAG	AGTAGGTGAA	GCCCTTCGGG	960
GCATCCAGAC	AGGTGGTGCA	TGGTTGTCGT	CAGTCGTGT	CGTGAGATGT	TGGGTAACT	1020
CCCGCAACGA	GCAGCAACCC	TATTATTAGT	TGCTACGCAA	GAGCACTCTA	ATGAGACTGC	1080
CGTTGACAAA	ACGGAGGAAG	GTGGGGATGA	CGTCAAATCA	TCAATGCCCT	TATGACCTGG	1140
GCTACACACG	TACTACAATG	GCACTAAAAC	AGAGGGCGGC	GACACCGCGA	GGTGAAGCGA	1200

-continued

```

atcccgaaaa aagtgtctca gttcagattt caggctgcaa cccgcctgca tgaagtcgga    1260
attgcttagta atcgccggatc agcatgccgc ggtgaatacg ttccccggcc ttgtacacac    1320
cccccgtaac accatgggag tcgggtaac acccgaaagcc agtag                                1365

```

<210> SEQ\_ID NO 96  
<211> LENGTH: 1404  
<212> TYPE: DNA  
<213> ORGANISM: Ruminococcus faecis

<400> SEQUENCE: 96

```

atgcaagtcg aacgaagcac cttgatttga ttcttcggat gaagatcttgc tgacttgagt    60
ggcgacgggg ttagtaacgc gtgggttaacc tgcctcatac agggggataa cagttagaaa    120
tgactgttaa taccgcataa gaccacagca cccatgggtc caggggtaaa aactccggtg    180
gtatgagatg gacccgcgtc tgatttagtta gttgggtgggg taacggccta ccaagccgac    240
gatcagtagc cgacacctgaga ggggtgaccgg ccacattgggg actgagacac ggcccaaact    300
cctacggggag gcagcagtgg ggaatattgc acaatgggggg aaaccctgtat gcagcgcacgc    360
cgcgtagcgt atgaagtattt tcgggtatgtaa aagctctatc agcaggaaag aaaatgacgg    420
tacactgacta agaagcaccg gctaaatacg tgccagcgc cgccgtataa cgtatggtgc    480
aagcgttatac cggattttact ggggtgtaaag ggagcgttaga cggagtggca agtctgtatgt    540
gaaaacccgg ggctcaaccc cgggactgca ttggaaactg tcaatctaga gtacggaga    600
ggtaaggcggg attcctatgt tageggtaaa atgcgtatgtt attaggagga acaccagtgg    660
cgaaggcggc ttactggacg gtaactgacg ttgaggctcg aaagcgtggg gagcaaacag    720
gattagatac cctggtagtc cacggcgtaa acgtacta ctgggtgtcg ggcagcaag    780
ctgttcggtg ccgcagcaaa cgcaataagt agtccacctg gggagttacgt tcgcaagaat    840
gaaactcaaa ggaattgacg gggacccgca caagcgggtg agcatgtgg ttaattcgaa    900
gcaacgcgaa gaacattacc tgctcttgc atctccctga cccgcaagta atgttgctt    960
tccttcggga cagggtatgc aggttgtca tgggtgtcg cagctgtgt cgtgatgt    1020
tgggttaagt cccgcaacga ggcacccccc tatctttatc agccagcggt ttggccgggc    1080
actcttagaga gactgccagg gataacctgg aggaagggtgg ggatgacgtc aaatcatcat    1140
gccccttatg agcagggtca cacacgtgtcacaatggcgtaa aacaaagg aggccagaacc    1200
gcgagggtcgca gcaaatccccca aaaataacgt ctcagttcggtt atttgtatct gcaactcgac    1260
tacatgaagc tggaaatcgct agtaatcgac aatcagaatg tcgggtgaa tacgttcccg    1320
ggtcttgc acaccggcccg tcacaccatg ggagttacgt aacggccgaag tcagtgcaccc    1380
aaccgtaagg aggagctgcc gaag                                1404

```

<210> SEQ\_ID NO 97  
<211> LENGTH: 1316  
<212> TYPE: DNA  
<213> ORGANISM: Dorea longicatena

<400> SEQUENCE: 97

```

taacgcgtgg gtaacacctgca tcatacagggg ggataacagt tagaaatgac tgctaatacc    60
gcataagacc acgtaccgcata tggtagtgcgt gtaaaaaactc cgggtggatgt agatggaccc    120
gcgtctgttggt aggttagttgg tgggttaacgc gcctaccaag ccgacgtatca gtagccgacc    180
tgagagggtg accggccaca ttgggactga gacacggccc agactctac gggaggcagc    240

```

-continued

---

agtgggaaat attgcacaat ggaggaaaact ctgatgcagc gacgcgcgt gaaggatgaa	300
gtatttcggt atgtaaacctt ctatcagcag ggaagaaaat gacggtaacct gactaagaag	360
ccccggctaa ctacgtgcc a gcagccgegg taatacgtag gggcaagcg ttatccggat	420
ttactgggtg taaagggagc gtacacggca cggcaagcca gatgtaaaa gcccggggct	480
caacccccgg actgcattt gaaactgctga gcttagagtgt cggagaggca agtggattc	540
ctagtgtacg ggtgaaatgc gtatgatatta ggaggaacac cagtgccgaa ggccgctgc	600
tggacgatga ctgacgttga ggctcgaaag cgtggggagc aaacaggatt agataccctg	660
gtatgtccacg ccgtaaacga tgactgctag gtgtcgggtg gcaaagccat tcggtgccgc	720
agctaaccgcataa agcagtc cacctggggta gtacgttgc aagaatgaaa ctcaaaaggaa	780
ttgacggggc cccgcacaag cgggtggagca tgggtttaa ttcaagc a cggcaagaaac	840
cttacctgat ctgacatcc cgtacccgc ttctgtatgg aagttttct tcggaaatc	900
ggtgacaggt ggtgcatgg tggcgttgc tgggttgttgg agatgttggg ttaagtcccg	960
caacgagcgc accccctatc ttcaatgc accaggttaa gctgggact ctggagagac	1020
tgcctgggat aacctggagg aagggtggggta tgacgtcaaa tcatcatgcc ccttatgacc	1080
agggtacac acgtgttaca atggcgtaaa caaagagaag cgaactcgcc agggtaagca	1140
aatctcaaaa ataacgtctc agttcggatt gtatgtcgca actcgactac atgaagctgg	1200
aatcgctagt aatcgatc cagaatgtcg cggtaatac gttccgggtt cttgtacaca	1260
ccggccgtca caccatggga gtcataacgc ccgaagtca g tgacccaaacc gtaagg	1316

<210> SEQ\_ID NO 98  
<211> LENGTH: 1475  
<212> TYPE: DNA  
<213> ORGANISM: Clostridium innocuum  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (905)..(905)  
<223> OTHER INFORMATION: n is a, c, g, t or u  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (942)..(942)  
<223> OTHER INFORMATION: n is a, c, g, t or u

<400> SEQUENCE: 98

atggagagtt tgatcctggc tcaggatgaa cgctggccgc atgcctaata catgcaagtc	60
gaacgaagtc ttcaaggaaatc ttgttccaa aaagacttag tggcgaacgg gtgagtaaca	120
cgttaggtac ctgcctatgt gtccggata actgtggaa acggtagcta aaaccggata	180
ggtatatacgga ggcgtatgtc tggatattaa agcgcccttc aaggcgtgaa catggatgga	240
cctgcgcacgc attatgttgc tggtaggttgc acggccacc aaggcgtatgc tgcgttagcc	300
gcctgagagg gtaaacggcc acattggac tgagacacgg cccaaactcc tacggaggg	360
agcgttaggg aattttgc aatggggaa accctgaacg agcaatgcgg cgtgtgttgg	420
gaagggttttgc ggatgttgc aagtgttgc aaggctcatag aggaaatgttgc	480
atggggatgttgc cggtagctt ccagaaagcc acggctaact acgtgtccagy agccgcgttgc	540
atacgttaggttgc ggcaagcgatc atccggatc atggggatgttgc aagggtgttgc	600
ctaagtctgttgc agtaaaaggc aatggcttgc ccattgttgc aatggatgttgc	660
gagtgccatggc gagggtgttgc aatggatgttgc aatggatgttgc	720
gaacaccatggc ggcgtatgttgc aatggatgttgc aatggatgttgc	780
ggggcaaat aggattatgttgc accctgtatgttgc aacgtatgttgc aactaagtgttgc	840

## US 10,555,980 B2

**285****286**

-continued

tggagaaatt cagtgtgca	gttaacgcaa	taagttctcc	gcctggggag	tatgcacgca	900	
agttnngaac tcaaaggaat	tgacggggc	ccgcacaagg	gntggagtat	gtgggtaat	960	
tcgaagcaac	gcgaagaacc	ttaccaggcc	ttgacatgga	aacaaatacc	ctagagatag	1020
ggggataatt	atggatcaca	caggtggtgc	atggttgtcg	tcagctcg	tcgtgagatg	1080
ttgggttaag	tcccgcacg	agcgcaccc	ttgtcgcatg	ttaccagcat	caagttggg	1140
actcatgcga	gactgcccgt	gacaaccgg	aggaagggtgg	ggatgacg	tc aaatcatcat	1200
gcccttatg	gcctgggcta	cacacgtact	acaatggcga	ccacaaagag	cagcgactt	1260
gtgacaagaa	gcgaatctca	taaagatcgt	ctcagttcgg	attgaagtct	gcaactcgac	1320
ttcatgaagt	cggaatcgct	agtaatcgca	gatcagcatg	ctgcggtgaa	tacgttctcg	1380
ggccttgtac	acaccggcccg	tcaaaccatg	ggagtcagta	atacccgaag	ccggtggcat	1440
aaccgttaagg	agtgagccgt	cgaaggtagg	accga			1475

&lt;210&gt; SEQ ID NO 99

&lt;211&gt; LENGTH: 1492

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Blautia hansenii

&lt;400&gt; SEQUENCE: 99

agagtttgat	cttggctcag	gatgaacgct	ggcgccgtgc	ttaacacatg	caagtcgagc	60
gaagcactta	tcattgactc	ttcggaaagat	ttgatatttgc	actgagcggc	ggacgggtga	120
gttaacgcgtg	ggtaacctgc	ctcatacagg	ggaataaacag	tttagaaatgg	ctgctaatgc	180
cgccataagcg	cacaggaccg	catggtctgg	tgtaaaaac	tgaggtggta	tgagatggac	240
ccgcgtctga	ttaggtagtt	ggtggggtaa	cggcctaccac	agccgacgat	cagtagccgg	300
cctgagaggg	tgaacggcca	cattgggact	gagacacggc	ccagactcct	acgggaggca	360
gcagtgggga	atattgcaca	atggggaaa	ccctgatgca	gcgacgcccgc	gtgaaggaag	420
aagtatctcg	gtatgtaaac	ttctatcgc	agggaaagaaa	atgacggtac	ctgactaaga	480
agccccggct	aactacgtgc	cagcagccgc	ggtaatacgt	agggggcaag	cgttatccgg	540
atttactggg	tgtaaaggga	gcgtagacgg	aagagcaagt	ctgatgtgaa	aggctggggc	600
ttaaccccg	gactgcatttgc	gaaactgttt	ttctagagtgc	ccggagaggt	aagcggaatt	660
cctagtgtag	cggtgaaatgc	cgtagatatt	aggaggaaca	ccagtgccgc	aggccggctta	720
ctggacggta	actgacgttg	aggctcgaaa	gcgtggggag	caaacaggat	tagataccct	780
ggtagtccac	gccgttaaacg	atgaataacta	ggtgtcgffff	tgcaaagcag	ttcggtgccg	840
cagcaaacgc	aataagtatt	ccacctgggg	agtacgttgc	caagaatgaa	actcaaagga	900
atgtgacgggg	acccgcacaa	gcggtggcgc	atgtgggtta	attcgaagca	acgcaagaa	960
ccttaccaag	tcttgacatc	tgcctgaccgc	ttccttaacc	ggagcttcc	ttcgggacag	1020
gcaagacagg	ttgtgcattgg	ttgtcgatgg	ctcgtgtcg	gagatgttgg	gttaagtccc	1080
gcaacgagcg	caacccttat	ccttagtagc	cagcagtcgg	gctgggcact	ctagggagac	1140
tgccggggat	acccggagg	aaggcgggga	cgacgtcaaa	tcatcatgcc	ccttatgtatt	1200
tgggctacac	acgtgttaca	atggcgtaaa	caaaggaaag	cgaagccgtg	acgcttagca	1260
aatctcaaaa	ataacgtccc	agttcgact	gcagtcgtca	actcgactgc	acgaagctgg	1320
aatcgctagt	aatcgcaat	cagaatgtcg	cggtgaatac	gttccgggt	cttgcacaca	1380
ccggccgtca	caccatggga	gtcagtaacg	cccgaaagtca	gtgacccaac	cttatggagg	1440

-continued

gagctgccga aggccggacc gataactggg gtgaagtgcgt aacaaggtaa cc	1492
--	------

<210> SEQ ID NO 100  
<211> LENGTH: 1473  
<212> TYPE: DNA  
<213> ORGANISM: Bacteroides cellulosilyticus

&lt;400&gt; SEQUENCE: 100

agagtttgc cctggctcgat gatgaacgct agctacaggc ttaacacatg caagtcgagg	60
ggcagcatga cctagcaata gggtgatggc gacccggcga cgggtgagta acacgtatcc	120
aacctaccgg ttattccggg atagccttc gaaagaaaaga ttaataccgg atagtataac	180
gagaaggcat cttttgtta ttaagaatt tcgataaccg atggggatgc gttccatag	240
tttgttggcg gggtaacggc ccaccaagac atcgatggat aggggttctg agaggaaagg	300
cccccacatt ggaactgaga cacggtccaa actcctacgg gaggcagcag tgaggaatat	360
tggtaatgg acgagagtct gaaccagcca agtagcgtga aggtgactg ccctatgggt	420
tgttaaacttc ttttatatgg gaataaaatgt agccacgtgt ggcttttgt atgtaccata	480
cgaataagga tcggctaact ccgtgccagc agccgcggta atacggagga tccgagcggt	540
atccggattt attgggttta aagggagcgt aggccgacta ttaagtgcgt tggtaaagg	600
tgcggctcaa ccgtaaaatt gcagttgata ctggtcgtct tgagtgcagt agaggtaggc	660
ggaattcgtg gtgtacgtt gaaatgctta gatatcaca agaactccga ttgcgaaggc	720
agcttactgg actgttaactg acgctgtatgc tcgaaagtgt gggtatcaaa caggattaga	780
taccctggta gtccacacag taaacgatga atactcgctg ttgcgatatacagcaagcg	840
gccaagcga agcattaagt attccacctg gggagtacgc cggcaacggt gaaactcaaa	900
ggaatttgcg gggcccgca caagccggagg aacatgtggt ttaattcgat gatacgcgag	960
gaaccttacc cgggcttaaa ttgcattctga ataatttggaa aacagatttag ccgcaaggca	1020
gatgtgaagg tgctgcattgg ttgcgtcgat ctcgtgcgtt gagggtgtcggtttaagtgc	1080
ataacgagcg caacccttat cttagttac taacaggtaa tgctgaggac tctagagaga	1140
ctggccgtcgtaa aagatgtgag gaagggtgggg atgacgtcaa atcagcacgg cccttacgtc	1200
cggggctaca cactgtttac aatgggggggt acagaaggca gtcacacagc gatgtatgc	1260
taatcccaaa agcctctctc agttcggatt ggagtctgcataccgactcc atgaagctgg	1320
atccgtatgt aatcgccat cagccacggc gccgtgaata cggtccggg cttgtacac	1380
accggccgtc aagccatgaa agccgggggt acctgaagtc cgtaaccgca aggagcggcc	1440
tagggtaaaa ctggtaattt gggctaaatgc gta	1473

<210> SEQ ID NO 101  
<211> LENGTH: 1459  
<212> TYPE: DNA  
<213> ORGANISM: Bacteroides ovatus

&lt;400&gt; SEQUENCE: 101

ggctcaggat gaacgcgtac tacaggctta acacatgcaa gtcgaggggc agcattttag	60
tttgcttgcataactgaagat ggcgaccggc gcacgggtga gtaacacgtta tccaacctgc	120
cgataactcc ggaatagcct ttcgaaagaa agattaatac cggatagcat acgaatatacg	180
catgatattt ttatataaga attcggat ttcgtgggg tgcgttccat tagttgttgc	240
gcggggtaac gcccaccaa gactacgtg gatagggtt ctgagaggaa ggtccccac	300
attggaaactg agacacggc caaactccta cgggaggcag cagtggagaa tattggtaaa	360

## US 10,555,980 B2

**289****290**

-continued

tgggcgagag	cctgaaccag	ccaagttagcg	tgaaggatga	aggctctatg	ggtcgtaaac	420
ttctttata	tggaaataaa	gttccacg	tgtggattt	tgtatgtacc	atatgaataa	480
ggatcgctta	actccgtgcc	agcagcccg	gtaatacgga	ggatccgagc	gttatccgga	540
tttattgggt	ttaaaggggag	cgttaggtga	ttgttaagtc	agttgtgaaa	gtttgcggct	600
caaccgtaaa	attgcagttg	aaactggcag	tcttgagtac	agtagaggtg	ggcggaaattc	660
gtgggtgtacg	ggtgaaatgc	ttagatatca	cgaagaactc	cgattgcga	ggcagctcac	720
tagactgtta	ctgacactga	tgctcgaaag	tgtgggtatc	aaacaggatt	agataacctg	780
gtagtccaca	cagtaaacga	tgaatactcg	ctgtttgcga	tatacgtaa	gcggccaagc	840
gaaagcatta	agtattccac	ctggggagta	cgcggcaac	ggtgaaactc	aaaggaattt	900
acggggggccc	gcacaagcgg	aggaacatgt	ggtttaattc	gtgatacgc	gaggaacctt	960
acccgggctt	aaattgcaac	agaatatatt	ggaaacagta	tagccgtaa	gctgttgtga	1020
aggtgctgca	tgggttgt	cagctcgta	cgtgagggt	cggcttaagt	gccataacga	1080
gogcaaccct	tatcttagt	tactaacagg	ttatgctgag	gactctagag	agactgcgt	1140
cgtaagatgt	gaggaagggt	gggatgacgt	caaattcagca	cggcccttac	gtccgggct	1200
acacaacgtgt	tacaatgggg	ggtacagaag	gcagctacct	ggcgacagga	tgctaattcc	1260
aaaaacccct	ctcagttcg	atcgaagtct	gcaaccccg	tccgtga	tgattcgct	1320
agtaatcg	catcagccat	ggcgccgt	atacgttccc	gggcctt	gtacagggt	1380
gtcaagccat	gaaagccggg	ggtacctgaa	gtacgttaacc	gcaaggagcg	tcctagggt	1440
aaactggtaa	ttggggcta					1459

&lt;210&gt; SEQ ID NO 102

&lt;211&gt; LENGTH: 1526

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Eubacterium fissicatena

&lt;400&gt; SEQUENCE: 102

tagagttga	tcctggctca	ggatgaacgc	tggcggegtg	cttaacacat	gcaagtgcag	60
cgaagegctt	tacttagatt	tcttcggatt	gaagagttt	gctactgagc	ggcggacggg	120
ttagttaacgc	gtggtaacc	tgcctcatac	agggggataa	cagttagaaa	tgactgctaa	180
taccgcataa	gaccacagta	ccgcatggta	cagtggaaa	aactccggtg	gtatgagatg	240
gacccgcgtc	tgattagcta	gttggtaagg	taacggctta	ccaaggcgcac	gatcgttagc	300
cgacctgaga	gggtgacccg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360
gcagcgtgg	ggaatattgc	acaatgggg	aaaccctgt	gcagcgcacgc	cgcgtgaagg	420
atgaagtatt	tcggtatgta	aacttctatc	agcagggaa	aaaatgacgg	tacctgacta	480
agaagccccg	gctaactacg	tgccagcgc	cgcgtataa	cgtagggggc	aagcgttatc	540
cggattttact	gggtgtaaag	ggagcgtaga	cggttatgta	agtctgtatgt	gaaaaccgg	600
ggctcaaccc	cgggactgca	ttggaaacta	tgttaactaga	gtgtcgaga	ggtaagtgg	660
attcctagtg	tagcggtgaa	atcgcttagat	attaggagga	acaccagtgg	cgaaggccgc	720
ttactggacg	atcaactgacg	ttgaggctcg	aaagcgtgg	gagcaaacag	gattagatac	780
cctggtagtc	cacgcgtaa	acgtgaata	ctaggtgtcg	ggtggcaaag	ccattcggt	840
ccgcagcaaa	cgcaataagt	attccacctg	gggagtacgt	tcgcaagaat	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcgtgg	agcatgtgg	ttaattcgaa	gcaacgcgaa	960

## US 10,555,980 B2

291

292

-continued

---

gaaccttacc tgctttgac atcccactga ccggcgtgta atggcgccctt cccttcgggg	1020
cagtggagac aggtggtgca tggttgtcg cagctcggt cgtgagatgt tgggttaagt	1080
cccgcaacga ggcacccct tatcttagt agccageggg ttggccggc actctagaga	1140
gactgecagg gataaacctgg aggaagggtgg ggatgacgta aatcatcat gcccattatg	1200
agcagggcta cacacgtgct acaatggcgta aaacaaaggagg aggcaataacc gcgagggtgta	1260
gcaaatccca aaaataacgt ctcagttcg attgtatgtc gcaactcgac tacatgaagc	1320
tggaatcgct agtaatcgct aatcagaatg tcgcggtgaa tacgttcccg ggtttgtac	1380
acaccggccg tcacaccatg ggagttggta acgcccgaag tcagtgaccc aaccgtaagg	1440
agggagctgc cgaaggcggg atcgataact ggggtgaagt cgtaacaagg tagccgtatc	1500
ggaagggtgc gctggatcac ctcccc	1526

&lt;210&gt; SEQ\_ID NO 103

&lt;211&gt; LENGTH: 1493

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Blautia coccoides

&lt;400&gt; SEQUENCE: 103

agagtttgc cctggcttag gatgaacgct ggccggcgtgc ttaacacatg caagtcgagc	60
gaagcgctaa gacagattc ttccggattga agtctttgtg actgagcggc ggacgggtga	120
gtaacgcgtg ggtaacctgc ctcatacagg gggataacag ttagaaatga ctgctaatac	180
cgcataagcg cacaggaccg catggctctgg tgtgaaaaac tccgggtgta tgagatggac	240
ccgcgtctga ttagctagtt ggaggggtaa cggcccacca aggccgacgt cagtagccgg	300
cctgagaggg tgaacggcca cattggact gagacacggc ccagactcct acgggaggca	360
gcagtggggaa atattgcaca atgggggaaa ccctgatgca ggcacgcgc gtgaaggaag	420
aagtatctcg gtatgtaaac ttcttatcgc agggaaagaa atgacggta ctgactaaga	480
agccccggct aactacgtgc cagcagccgc ggtataatcgta agggggcaag cgttatccgg	540
atttactggg tgtaaaggga gcgttagacgg aagagcaagt ctgatgtgaa aggctggggc	600
ttaaccccgact gactgcattt gaaactgttg ttcttagatgc cccggagatgt aagcgaaatt	660
cctagtgtatcg cggtaaatcg cgtatgtatt aggaggaaca ccagtggcga aggccggctta	720
ctggacggta actgacgttg aggctcgaaa gcgtggggag caaacaggat tagataccct	780
ggtagtccac gccgtaaacg atgaataacta ggtgtcggtt ggcaaaacca ttccgtgccg	840
cagcaaaacgc aataagtatt ccacccgggg agtacgttcg caagaatgaa actcaaaggaa	900
attgacgggg accccgcacaa gcggtggagc atgtggtttta attcgaagca acgcgaagaa	960
ccttaccaag tcttgacatc cctctgaccc tcccgtaacg ggggcttccc ttccggggcag	1020
aggagacagg tggtgcatgg ttgtcgctcg ctcgtgtcgagatgttgg gttaaatccc	1080
gcaacgcgcg caacccttat ccttagatgc cagcacatga tgggtggcac tcttagggaga	1140
ctggccgggaa taacccggag gaaggcgggg acgacgtcaa atcatcatgc cccttatgtat	1200
ttgggctaca cacgtgtac aatggcgtaa acaaaggaa ggcacgcgc gatgttgc	1260
gaatcccaaa aataacgtcc cagttcggac tgcagtcgtc aactcgactg cacgaagctg	1320
gaatcgctag taatcgccgta tcagaatgcc gcggtgaata cggtccggg tcttgcac	1380
accggccgc acaccatggg agtcaactac gcccgaagtc agtgcaccaa ccgaaaggaa	1440
ggagctgccg aaggcgggac cgataactgg ggtgaagtcg taacaaggta acc	1493

-continued

<210> SEQ ID NO 104  
<211> LENGTH: 1262  
<212> TYPE: DNA  
<213> ORGANISM: Blautia faecis  
<400> SEQUENCE: 104

ataacagcca gaaatgactg ctaataccgc ataagcgcac agaaccgcacat ggttcggtgt	60
gaaaaactcc ggtggtataa gatggacccg cgttggatta gctagttggc agggcagcgg	120
cctaccaagg cgacgatcca tagccggct gagagggtga acggccacat tgggactgag	180
acacggccca gactctacg ggaggcagca gtgggaaata ttgcacaatg gggaaaccc	240
tgtatgcagcg acgcccgcgtg aaggaagaag tatctcggtt tgtaaaacttc tatcagcagg	300
gaagataatg acggtaacctg actaagaagc cccggctaact tacgtgccag cagccgcgt	360
aatacgttagg gggcaagcgt tatccggatt tactgggtgt aaagggagcg tagacggcgc	420
agcaagtctg atgtgaaagg caggggctta acccctggac tgcattggaa actgctgtgc	480
ttgagtgccg gaggggtaag cgaaattccct agtgtagcgg tgaaatgcgt agatattagg	540
aggaacacca gtggcgaagg cggcttactg gacggtaact gacggttggagg ctgcggcg	600
tggggagcaa acaggattag ataccctggt agtccacgcg gttaacgcgt aatacttagt	660
gtcaggggagc acagcttctt ggtgccgcgc caaacgcatt aagtattcca octggggagt	720
acgttgcacaa gaatgaaact caaaggaaatt gacggggacc cgcacaagcg gtggagcatg	780
tggtttaatt cgaagcaacg cgaagaacctt taccaaatct tgacatccct ctgaccggga	840
cttaaccgc cttttccctt gggacagggg agacagggtt tgcatgggt tcgtcagctc	900
gtgtcgtgag atgttgggtt aagtcccgca acgagcgcac cccctatcct tagtagccag	960
cacgcgttgg tgggactctt gaggagactg ccagggtataa cctggggaaa ggccggatg	1020
acgtcaaattt atcatgcaccc ttatgattt ggctacacac gtgttacaat ggctaaaca	1080
aagggaagcg aacccgcgag ggtggcaaa tctcaaaaat aacgtcccgatc ttccggactgc	1140
agtctgcaac tcgactgcac gaagctggaa tcgctagtaa tcgcggatca gaatgcgcgc	1200
gtgaatacgt tccccgggtct tgcacacacc gcccgtcaca ccatgggagt cagtaacgccc	1260
cg	1262

<210> SEQ ID NO 105  
<211> LENGTH: 1431  
<212> TYPE: DNA  
<213> ORGANISM: Clostridium hathewayi  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (191)..(191)  
<223> OTHER INFORMATION: n is a, c, g, t or u  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (211)..(211)  
<223> OTHER INFORMATION: n is a, c, g, t or u  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (248)..(248)  
<223> OTHER INFORMATION: n is a, c, g, t or u  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (979)..(979)  
<223> OTHER INFORMATION: n is a, c, g, t or u  
<400> SEQUENCE: 105

ctcaggatga acgctggcgg cgtgtttaac acatgcaagt cgagcgaagc ggtttcaatg	60
aagttttcgg atggatttga aattgactta gcggcggacg ggtgagtaac gcgtggtaa	120

-continued

---

cctgccttac actggggat aacagttaga aatgactgct aataccgcat aagcgcacag	180
ggccgcatgg nctgggtgtga aaaactccgg nggtgtaaga tggaccgcg tctgattagg	240
tagttggngg ggttaacggcc caccaagccg acgatcagta gccgacctga gagggtgacc	300
ggcccacattg ggactgagac acggccaaa ctcctacggg aggcagcagt gggaaatatt	360
ggacaatggg cgaaaggctg atccagcgac gccgcgtgag tgaagaagta ttccggat	420
ttaaagctcta tcagcaggga agaaaatgac ggtacctgac taagaagccc cggctaacta	480
cgtgccagca gccgcggtaa tacgttaggg gcaagcgtta tccggattta ctgggtgtaa	540
agggagcgtg acgggtttag caagtctgaa gtgaaagccc ggggctcaac cccggactg	600
ctttggaaac tgtagactt gagtgcagga gaggttaagtg gaattcttag ttagcggtg	660
aaatgcgtat atattaggag gaacaccgt ggcgaaggcg ctgtactgaa	720
cgttgaggct cgaaaggctg gggagcaac aggattagat accctggtag tccacgcgt	780
aaacgatgaa tactaggtgt cggggggcaa agcccttcgg tgccgcgc aacgcaataa	840
gtattecacc tggggagtagc gttcgcaaga atgaaactca aaggattga cggggacccg	900
cacaaggcgtt ggagcatgtg gtttaattcg aagcaacgcg aagaacctta ccaagtctt	960
acatcccact gaaaacacnt taaccgtat ccctcttcgg agcagtggag acagggtgt	1020
catgggtgtc gtcagctcg gtcgttagat gttgggttaa gtcccgcac gagcgcaacc	1080
cttatacctta gtagccagcg agtagagtcg ggcactctgg ggagactgcc agggataacc	1140
tggaggaagg tggggatgac gtcaaattcat catgcccctt atgatttggg ctacacacgt	1200
gttacaatgg cgtaaacaaa gggaggcaaa ggagcgatct ggagcaacc cccaaaataa	1260
cgtctcagtt cggattgcag gctgcaactc gcctgcatga agctggatc gctagtaatc	1320
gccaatcaga atgtcgccgt gaatacgttc cccgggtt tacacaccgc cccgtcacacc	1380
atgggagttg gtaacgcccc aagtcaagtga cccaaaccgaa aggagggagc t	1431

&lt;210&gt; SEQ ID NO 106

&lt;211&gt; LENGTH: 1493

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Blautia producta

&lt;400&gt; SEQUENCE: 106

agagttttagt cctggctcag gatgaacgct ggccgcgtgc ttaacacatg caagtcgagc	60
gaagcactaa gacggatttc ttccgattga agtctttgtg actgagcggc ggacgggtga	120
gttaacgcgtg ggtaacctgc ctcatacagg gggataacag ttagaaatga ctgctaatac	180
cgcataagcg cacaggaccg catggctctgg tgtgaaaaac tccgggtgta tgagatggac	240
ccgcgtctga ttagctagtt ggagggtaa cggccacca aggccgacat cagtagccgg	300
cctgagaggg tgaacggcca cattggact gagacacggc ccagactcct acggaggca	360
gcagtgaaaa atattgcaca atggggaaa ccctgatgca ggcgcgcgc gtgaaggaag	420
aagtatctcg gtatgtaaac ttctatcagc agggaaagaaa atgacggtagt ctgactaaga	480
agccccggct aactacgtgc cagcagccgc ggtataacgt agggggcaag cgttatccgg	540
atttactggg tgtaaaggga gcgttagacgg aagagcaagt ctgatgtgaa aggctgggc	600
ttaacccag gactgcattt gaaactgttg ttcttagatgt ccggagaggt aagcggatt	660
cctagtgttag cggtaaatg cgtagatatt aggaggaaca ccagtgccga aggccgctta	720
ctggacggta actgacgttg aggctcgaaa gcgtgggag caaacaggat tagataccct	780
ggtagtccac gccgtaaacg atgaatacta ggtgtcggtt ggcaaaagcca ttccggccg	840

-continued

cagcaaacgc aataagtatt ccacctgggg agtacgttcg caagaatgaa actcaaagga	900
attgacgggg acccgacaaa gcggtggcgc atgtgggta attcgaagca acgcaagaa	960
ccttaccaag tcttgcacatc cctctgaccc tcccgttaacg gggacttccc ttggggcag	1020
aggagacagg tggtgcatgg ttgtcgtagc ctgcgtgtcg gagatgttgg gttaagtccc	1080
gcaacgagcg caacccttat ccttagtagc cagcacatga tggtggcac tctaggaga	1140
ctgcccgggta aacccggag gaaggcgggg acgacgtcaa atcatcatgc cccttatgat	1200
ttgggctaca cacgtgctac aatggcgtaa acaaaggaa gcgagacagc gatgttgagc	1260
gaatccaaaa aataacgtcc cagttcgac tgcagtctgc aactcgactg cacgaagctg	1320
gaatcgctag taatcgccga tcagaatgcc gcgggtgaata cggtccggg tcttgtacac	1380
accggccgtc acaccatggg agtcaactac gccccgatc agtgacctaacc cggaaaggaa	1440
ggagctgccc aaggcgggac cgataactgg ggtgaagtcg taacaaggta acc	1493

&lt;210&gt; SEQ ID NO 107

&lt;211&gt; LENGTH: 1515

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Anaerostipes hadrus

&lt;400&gt; SEQUENCE: 107

tggctcagga tgaacgctgg cggtcgctt aacacatgca agtcaacga agtgcctaa	60
ctgatcttct tcggaattga cggtttgttag actgagtggc ggacgggtga gtaacgcgtg	120
ggcaacactgc cctgtacagg gggataacag tcagaaatgatc ctgtaatac cgcataagac	180
cacageacccg catggtgca gggtaaaaac tccgggtgta caggatggac ccgcgtctga	240
ttagctgggtt ggtgaggtaa cggctcacca agggcagcat cagtagccgg cttgagagag	300
tgaacggcca cattggact gagacacggc ccaaactct acgggaggca gcagtgggaa	360
atattgcaca atggggaaa ccctgtatca ggcacgcgc gtgagtgaag aagtatctcg	420
gtatgtaaag ctctatcagc agggaaagaa atgacggtagt ctgactaaga agcccccgc	480
aactacgtgc cagcagccgc ggtataacgt agggggcaag cggttatccgg aattactggg	540
tgtaaagggt gcgttaggtgg tatggcaagt cagaagtgaa aacccaggcc ttaactctgg	600
gactgtttt gaaactgtca gactggagtg caggagaggta aagcggaaatt ccttagtgg	660
cggtgaaatcg cgtagatatt aggaggaaca tcagtgccgaa aggccggctta ctggactgaa	720
actgacactg aggcacgaaa gcgtggggag caaacaggat tagataccct ggttagtcac	780
gccgtaaacg atgaataacta ggtgtcgcccc ccgttagaggc ttccgtgcgc cagccaaacgc	840
agtaagtatt ccacctgggg agtacgttcg caagaatgaa actcaaaggaa attgacgggg	900
acccgcacaa gcgggtggagc atgtgggtaa attcgaagca acgcaagaa ccttacctgg	960
tcttgcacatc cttctgaccc gtccttaacc ggaccttcc ttccggacag gagagacagg	1020
tggtgcatgg ttgtcgtagc ctgcgtgtcg gagatgttgg gttaagtccc gcaacgagcg	1080
caacccttat cttagtagc cagcattca ggtggcact ctagagagac tgccaggat	1140
acccctggagg aagggtgggaa cgacgtcaaa tcatcatgcc cottatgacc agggctacac	1200
acgtgtaca atggcgtaaa cagaggaaag cagcgtcg agagtggca aatccaaaa	1260
ataacacgtctc atggcgatt gtgtctgc actcgactac atgaagctgg aatcgctagt	1320
aatcgcaat cagaatgtcg cggtgaatac gttccgggtt ctgtacaca ccggccgtca	1380
caccatggga gtcagtaacg cccgaagtcg gtgacccaaac cgtaaggagg gagctgcga	1440

## US 10,555,980 B2

**299****300**

-continued

aggcgggacc gataactggg gtgaagtcgt aacaaggtag ccgtatcgga aggtgcggct	1500
ggatcacctc ctttc	1515
<210> SEQ ID NO 108	
<211> LENGTH: 1523	
<212> TYPE: DNA	
<213> ORGANISM: Eubacterium fissicatena	
<400> SEQUENCE: 108	
gttgatcct ggctcaggat gaacgctggc ggcgtgctta acacatgcaa gtcgagcgaa	60
gcccgttact tagattctt cggattgaag agttttgcga ctgagcggcg gacgggtgag	120
taacgcgtgg gtaacctgcc tcatacaggg ggataacagt tagaaatgac tgctaatacc	180
gcataagacc acagtaccgc atggtagcgt gggaaaaact ccgggtggat gagatggacc	240
cgcgtctgat tagctagttg gtaaggtaac ggcttaccaa ggcaacgatc agtagccgac	300
ctgagagggt gaccggccac attgggactg agacacggcc caaactccctc cgaggaggcag	360
cagtggggaa tattgcacaa tgggggaaac cctgatgcag cgacgcgcg tgaaggatga	420
agtatcccg tatgttaact tctatcagca gggaaaatgg tgcgggtacc tgactaagaa	480
ccccccggcta actacgtgcc agcagccgcg gtaatacgtt gggggcaagc gttatccgga	540
tttactgggt gtaaaaggag cgtagacggt tatgttaagtc tgatgtgaaa accccgggct	600
caaccccccggg actgcattgg aaactatgtt actagagtgt cggagaggta agtggaaattc	660
ctagtgttagc ggtgaaatgc gtatgttggat gggggccac cagtgccgaa ggccgttac	720
tggacgatca ctgacgttga ggctcgaaag cgtggggagc aaacaggatt agataccctg	780
gtatccacg ccgtaaacga tgaataacttag gtgtcggttgc gcaaagccat tcgggtccgc	840
agcaaacgcgataaatttcc cacctggggat gtacgatcgcc aagaatgaaa ctcaaaaggaa	900
ttgacggggaa cccgcacaaag cgggtggagca tgtggttttaa ttgcgaaacgc cgcaagaac	960
cttacctgtt ctgcacatcc cactgaccgg cgtgtatgg cgccttcct tcggggcagt	1020
ggagacagggt ggtgcgttgt tgctgtcagc tcgtgtcgat agatgttggg ttaagtcccg	1080
caacgcgcgca aacccttatac tttatgttgc acgggtttgg cggggactc tagagagact	1140
gcggggata acctggagga aggtggggat gacgtcaaat catcatgccc cttatgagca	1200
gggctacaca cgtgtacaaat tggcgtaaacaa aaaggggaggc aataccgcga ggttggacaa	1260
atcccaaaaa taacgtctca gttcgatttgc tagtctgcaat ctcgactaca tgaagcttga	1320
atcgctatgtt atcgcaatc agaatgtcgcc ggtgtatcgttcccgatc ttgttacacac	1380
cgcgggtcac accatggggat ttggtaacgc ccgaagtcag tgacccaacc gtaaggagg	1440
agctgcggaa ggccggatcg ataactgggg tgaagtcgtt acaaggtagc cgtatcgaa	1500
ggtgccggctg gatcacctcc ttt	1523

<210> SEQ ID NO 109	
<211> LENGTH: 1524	
<212> TYPE: DNA	
<213> ORGANISM: Eubacterium contortum	

<400> SEQUENCE: 109	
tttgatcctg gctcaggatg aacgctggcg acgtgtttaa cacatgcaag tcgagcgaa	60
cactttaactt tgatttcttc ggaatgaaag gttttgtgac tgagcggcg acgggtgagt	120
aacgcgtggg taacctgcct catacagggg gataacagt tagaaatgact gctaatacc	180
cataagacca cagtagccca tggtagcgtt gggaaaaactc cgggtggatg agatggaccc	240

-continued

gcgtctgatt agctagttgg taaggtaacg gcttaccaag gcgacgatca gtagccgacc	300
tgagagggtg accggccaca ttgggactga gacacggccc aaactcctac gggaggcagc	360
agtggaaat attgcacaat ggggaaacc ctgatgcage gacgcccgt gaaggatgaa	420
gtatccgtt atgtaaactt ctatcagcag ggaagaaaat gacggtacct gactaagaag	480
ccccggctaa ctacgtgcc a cagccgccc taatacgtag gggcaagcg ttatccggat	540
ttactgggtg taaaggagc gtacggttt atgtaagtct gatgtgaaaa cccggggctc	600
aacccggga ctgcattgga aactatgtaa ctagactgtc ggagaggtaa gtggaatcc	660
tagttagcg gtgaaatgcg tagatattag gagaaacacc agtggcgaag gcggcttact	720
ggacgatgac tgacggtttag gctcgaaagc gtggggagca aacaggattt gataccctgg	780
tagtccacgc cgttaaacgtt gaataactagg tgcgggtgg caaaggccat cgggtccgca	840
geaaacgcaa taagtattcc acctggggag tacgttcgca agaatgaaac tcaaaggaaat	900
tgacggggac ccgcacaagc ggtggagcat gtgggttaat tgcggcaac gcgaaacc	960
ttacctgctc ttgacatccc cctgaccggc gtgtaatggt gctttccctt cgggacagg	1020
gagacaggtg gtgcattgtt gtcgtcagct cgtgtcgta gatgtgggt taatccgc	1080
aacgagcgca acccttatct ttagtagcca ggggttggc cgggcactct agagagactg	1140
ccagggataa cctggggagaa ggtggggatg acgtcaaatc atcatgcccc ttatggcag	1200
ggctacacac gtgctacaat ggcgtaaaca aaggaggcg aagccgtgag gtggagcaaa	1260
tccaaaaat aacgtctcag ttcggattgt agtctgcaac tgcactacat gaagetggaa	1320
tgcgttagaa tcgcgaatca gaatgtcgat gtgataatgt tccgggtct tgtacacacc	1380
gcccgtcaca ccatggggatg tggtaacgcc cgaagtcagct gacccaaaccg caaggaggaa	1440
gtgcggagg gtgggaccga taactgggtt gaagtcgtaa caaggtagcc gtatcgaaag	1500
gtgcggctgg atcacccctt ttct	1524

<210> SEQ ID NO 110  
 <211> LENGTH: 1390  
 <212> TYPE: DNA  
 <213> ORGANISM: Clostridium bolteae

&lt;400&gt; SEQUENCE: 110

tttaattga ctgagtgccg gacgggttag taacgcgtgg ataacctgcc tcacactgg	60
ggataacagt tagaaatgac tgctaatacc qcataagcgc acagtaccgc atggtacagt	120
gtaaaaact ccgggttgtt gagatggatc cgcgtctgat tagccagttg gccccgttaac	180
ggcccaccaa agcgacgatc agtagccgac ctgagagggt gacccggccac attggactg	240
agacacggcc caaactccta cgggaggcg cagtggggaa tattgcacaa tgggcgaaag	300
cctgatgcag cgacgcgcg tgagtgaaga agtatttcgg tatgtaaagc tctatcagca	360
gggaagaaaa tgacggtacc tgactaagaa gccccggcta actacgtgcc agcagccgc	420
gtaatacgtt gggggcaagc gttatccgga tttactgggt gttaaggagc cgtagacggc	480
gaagcaagtc tgaagtggaaa acccagggtt caaccctggg actgctttgg aaactgtttt	540
gctagagtgt cggagaggta agtggaaattc ctatgttagc ggtgaaatgc gtagatatta	600
ggagggacac cagtggcgaa ggcggcttac tggacgataa ctgacggttga ggctcgaaag	660
cgtggggaggc aaacaggatt agataccctg gtatccacg ccgtaaacga tgaatgttag	720
gtgttgggggg gcaaaggccct tcgggtccgt cgcaaacgca gtaagcattc cacctgggg	780

-continued

---

gtacgttcgc aagaatgaaa ctcaaaggaa ttgacgggaa cccgcacaag cggtggagca	840
tgtggtttaa ttcaagca cgcgaagaac cttaccaagt cttgacatcc tcttgacccg	900
cgtgtaacgg cgcctccct tcggggcaag agagacaggt ggtgcattgt tgctgtcagc	960
tctgtgtcgta agatgttggg ttaagtcccg caacgagcgc aacccttatac ctttagtagcc	1020
agcaggtaaa gctgggact ctagggagac tgccaggat aacctggagg aagttgggaa	1080
tgacgtcaaa tcatacatgcc ccttatgatt tgggctacac acgtgttaca atgggttaaa	1140
caaaggaaag caagacagt atgtggagca aatccaaaa ataacgtccc agttggact	1200
gtagtctgca acccgactac acgaagctgg aatcgctagt aatcgcaat cagaatgtcg	1260
cggtaatac gttcccggt cttgtacaca ccgcggcgtca caccatggga gtcagcaacg	1320
cccgaaagtca gtgacccaac tcgcaagaga gggagctgcc gaaggcgaaa caggtactg	1380
gggtgaagtc	1390

<210> SEQ\_ID NO 111  
<211> LENGTH: 1308  
<212> TYPE: DNA  
<213> ORGANISM: Blautia luti  
<220> FEATURE:  
<221> NAME/KEY: misc\_feature  
<222> LOCATION: (702)..(705)  
<223> OTHER INFORMATION: n is a, c, g, t or u

<400> SEQUENCE: 111

gtggtaacc tgcctatac agggggataa cagtcagaaa tgactgctaa taccgcataa	60
gcgcacagag ctgcattggct ccgggtgtgaa aaactccgggt ggtataagat ggacccgcgt	120
tggattagct agttgggtgag gtaacggccc accaaggcga cgatccatag ccggccttag	180
agggtgaacg gccacattgg gactgagaca cggccacagac tcctacggga ggcagcgt	240
ggaatatttgc cacaatgggg gaaaccctga tgcagcgtac ccgcgtgaag gaagaagtat	300
ctcggatgtt aaacttctat cagcaggaa gaaaatgacg gtacctgact aagaagcccc	360
ggctaactac gtgcacgcg ccgcgttaat acgtgggggg caagcggtat ccggatttac	420
tgggtgtaaa gggagcgtat acggcatggaa caagtctgtat gtgaaaggct ggggtcaac	480
cccggactg cattggaaac tgccgtt gatgtccggaa gaggttaacgc gaattcttag	540
tgtacgggtg aaatgcgtat atattaggag gaacaccatggc ggcgaaggcg gcttactggaa	600
cggtaactga ctttggggct cggaaacgcgtg gggagcaaac aggattagat accctggtag	660
tccacgcgggtt aaacgcgtat tccttaggtgt cggggagcaaa annnnnttcgg tgccggccca	720
aacgcattaa gcattccacc tggggagttac gttcgcaaga atgaaactca aaggaattga	780
cggggacccg cacaaggcgtt ggacatgtt gtttaattcg aagcaacgcg aagaacccat	840
ccaaagtcttgc acatcccttgc gaccggatgtat gtatggactt tttcccttgg gagagagagg	900
agacagggtgg tgcattgggtt tgcattgttgc gtgtcgtagt atgttgggtt aagtcccgca	960
acggagcgttcc cccctatccc cagtagccag cgggttcggcc gggcactctg aggagactgc	1020
cagggataac ctggaggaag gggggatgtat cgtcaaatca tcattccct tatgatttgg	1080
gtacacacacg tgctacaatgc gctaaacaa agggaaagcaaa gcttcggagg gtggggcaat	1140
cccaaaaata acgtcccaatgc tcggactgtat gtctgcaacc cgtacacacg aagctggaaat	1200
cgttagataat cgcggatgtat aatggccgggg tgcatacgatcc cccgggtt gtcacacccg	1260
cccgtaacac catggggatgtc agtaacgcgtt gaaatcgatgc acctaact	1308

## US 10,555,980 B2

**305****306**

-continued

<210> SEQ ID NO 112  
<211> LENGTH: 1489  
<212> TYPE: DNA  
<213> ORGANISM: Acidaminococcus intestini

<400> SEQUENCE: 112

ctggcggcgt	gcttaacaca	tgcaagtcga	acggagaact	tatttcggtt	agtcttagt	60
ggcgaacggg	tgagtaacgc	gtggcaacc	tgcctccag	ttggggacaa	cattccgaaa	120
gggatgtcaa	taccgaatgt	cctccctct	ccgcattggag	gaggggagaa	agatggcctc	180
tgcttgcag	ctatcgctgg	aagatggcc	cgcgtctgt	tagctagttg	gtggggtaac	240
ggctcaccaa	ggcgatgtac	agtagccggt	ctgagaggat	gaacggccac	attgggactg	300
agacacggcc	caaactccta	cgggaggcag	cagtgggaa	tcttcggcaa	tggacgaaag	360
tctgacggag	caacgcgcg	tgagtgtatga	aggcttccgg	attgtaaaac	tctgttgtta	420
gggacgaaag	caccgtgttc	gaacagggtca	tggtggtgac	ggtacctaac	gaggaagcca	480
cggcttaacta	ogtgcacagca	gccgcggtaa	tacgttaggtg	gcaagcgttg	tccggatta	540
ttggggctaa	agagcatgtt	ggcgggcttt	taagtctgac	gtgaaaatgc	ggggcttaac	600
cccgatatggc	gttggatact	ggaagtcttg	agtgcaggag	aggaaagggg	aattcccagt	660
gtagcggtaga	aatcgctaga	tattgggagg	aacaccagt	gogaaggcgc	ctttctggac	720
tgtgtctgac	gctgagatgc	gaaagccagg	gtagcaaacc	ggattagata	ccccggtagt	780
cctggccgta	aacgatggat	actaggtgtt	ggaggtatcg	accccttctg	tgccggagtt	840
aacgcataaa	gtatcccgcc	tgggactac	gatcgcaaga	ttgaaactca	aaggaattga	900
cgggggcccg	cacaagcggt	ggagtatgt	gtttaattcg	acgcaacgcg	aagaacctta	960
ccaaggcttg	acattgagtg	aaagacctag	agataggc	ctcccttcgg	ggacacgaaa	1020
acagggtggtg	atggctgtc	gtcagctgt	gtcgtgagat	gttgggttaa	gtcccgcaac	1080
gagcgcacc	cctatcctat	gttaccageg	cgtaaaggcg	ggactcata	ggagactgcc	1140
agggataact	tggaggaagg	cggggatgac	gtcaagtcat	catgccccctt	atgtcttggg	1200
ctacacacgt	actacaatgg	tccgcaacaa	agggcagcga	aaccgcggagg	tggagcaat	1260
cccagaaaacc	cgaccccaagt	tccgatcgta	ggctgcaacc	cgctcacgt	aagttggaaat	1320
cgcgtatata	cgcaggtag	catactgegg	tgaatacggt	cccgccctt	gtacacaccg	1380
cccgtcacac	cacgaaagtt	ggtaacaccc	gaagccggtg	agataacctt	ttaggagtca	1440
gctgtctaa	gtggggccga	tgattgggt	gaagtcgtaa	caaggtac		1489

<210> SEQ ID NO 113  
<211> LENGTH: 1500  
<212> TYPE: DNA  
<213> ORGANISM: Ruminococcus albus

<400> SEQUENCE: 113

agagtttgt	cctggcttag	gacgaacgt	ggcgccacgc	ttaacacatg	caagtgcgaac	60
gagcgaaaaga	gtgcttgcac	tctctagcta	gtggcgacg	ggtgagtaac	acgtgagcaa	120
tctgccttc	ggagagggat	accatttga	aacgattgtt	aatacctcat	aacataacga	180
agccgcattga	cattttatc	aatgaattt	cggcggaaaga	tgagctcg	tctgattagg	240
tagtttgta	ggttaacggcc	caccaagccg	acgtatgt	gccggactga	gagggtgaac	300
ggccacattg	ggactgagac	acggcccgaga	ctcctacggg	aggcagcagt	ggggaaatatt	360
gcacaaatggg	cgaaaggctg	atgcagcgt	gccgcgtgag	ggaagaaggt	ttttaggattg	420

## US 10,555,980 B2

307

308

-continued

---

taaacctctg tctttgggga cgataatgac ggtacccaag gaggaagctc cggctaacta	480
cgtgccagca gccgeggtaa tacgttaggg a cgagcgttc tccgaaatta ctgggtgtaa	540
agggagcgta ggccggattt caagtcaggt gtgaaattt a gggcttaac ccctgaactg	600
cacttgaac tgttagttt gaggtaagta gaggttaagcg gaattctcg ttagcggtg	660
aatatgcgtat atattaggag gaacatcagt ggcgaaggcg gcttactggg cttaactga	720
cgctgaggct cgaaaacgtg gggagcaac aggattagat accctggtag tccacgcgt	780
aaacatgtat tactaggtgt gggggactg accccttccg tgccgcaggta aacacaataa	840
gtatccacc tggggagtac ggccgcaagg ctgaaactca aaggaattga cggggacccg	900
cacaaggcgt ggagtatgtg gtttattcg aagcaacgcg aagaacctta ccaggcttg	960
acatcgtagt catagcatag agatatgtga aatcccttcg gggacgtata gacaggttgt	1020
gcatgggtgt cgtcagctcg tgcgtgaga tgggggtt aagtcccgc acgagcgaa	1080
cccttactgt tagttgtac gcaagagcac tctagcaggta ctggcggttga caaaacggag	1140
gaagggtggg atgacgtcaa atcatcatgc cccttatgac ctgggttaca cacgtactac	1200
aatggctgtt aacagaggga agcaaaacag tggatgtggag caaaacctta aaagcgtct	1260
tagttcgat tgcgttgc aaccggccta catgaagtgc gaattgtctat taatcgccga	1320
tca gcatgcc ggggtgataa cgttccggg ccttgcac accggccgtc acggcatggg	1380
atcggttaac acccgaaagcc tgggttctaa ccgcaaggag gaagcgtcg aaggtgggat	1440
tgatgactgg ggtgaaagtgc taacaaggta gccgtatcg aaggtgcggc tggatcacct	1500

&lt;210&gt; SEQ\_ID NO 114

&lt;211&gt; LENGTH: 1521

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Eubacterium rectale

&lt;400&gt; SEQUENCE: 114

agagtttgcctggctcgt gatgaacgct ggcggcgtgc ttaacacatg caagtcgaac	60
gaagcacttt atttgatttc ctccggact gattatttt tgactgagtg gcggacgggt	120
gagtaacgcg tgggttaccc gccttgcata gggggataac agttggaaac ggctgtctat	180
accgcataag cgcacggcat cgcacgtatgc agtgtgaaaa actccggtgg tataagatgg	240
accgcgttg gattagctat ttgggtgaggt aacggccccac caaggcgcacg atccatagcc	300
gacctgagag ggtgaccggc cacattggga ctgagacacg gcccacactc ctacggagg	360
cagcagtggg gaatattgca caatggcgaa aagcctgtatgc cagcgcaccc gcgtgagcga	420
agaagtattt cggatgtttaa agctctatca gcaggaaaga taatgacggt acctgactaa	480
gaagcaccgg ctaaataacgt gccagcagcc gcggttaatac gtatggtgc acgttatcc	540
ggatattactg ggtgtaaagg gagcgcaggc ggtgcggcaa gtctgtatgt aaagcccccgg	600
gttcaccccc ggtactgcat tggaaactgt cgtactagag tgcgtggagg gtaagcggaa	660
ttccttagtgc agcgggtaaa tgcgtatata ttaggggaa caccagtggc gaaggcggct	720
tactggacgc taactgacgc tgaggctcgaa aagcgtgggg agcaaaacagg attagatacc	780
ctggtagtcc acgcccgtaaa cgtatatac taggtgttgg gaagcattgc ttctcggtgc	840
cgtcgaaac gcagtaagta ttccacctgg ggagtagcttgc cgcacaaatg aaactcaaag	900
gaattgacgg ggacccgcac aagcggtgg a cgtgtggtaat ttcgaag caacgcgaag	960
aaccttacca agtcttgaca tccttctgac cggtaacttac ccgtacccctc tcttcggagc	1020
aggagtgaca ggtggtgcat ggtgtgcgc agtcgtgtc gtgagatgtt gggtaagtc	1080

-continued

ccgcaacgag	cgcaaccctt	atcttagta	gccagcggtt	cggccgggca	ctctagagag	1140
actgcccagg	ataacctgga	ggaaggcggg	gatgacgtca	aatcatcatg	ccccttatga	1200
cttggctac	acacgtgcta	caatggcgta	aacaaaggga	agcaaagctg	tgaagccgag	1260
caaatctcaa	aaataacgta	tcagttcgga	ctgtagtctg	caacccgact	acacgaagct	1320
ggaatcgcta	gtaatcgac	atcagaatgc	tgccgtgaat	acgttcccg	gtcttgtaca	1380
caccgccccgt	cacaccatgg	gagttggaa	tgcccgaa	cagtgcacct	accgaaagga	1440
aggagctgtc	gaaggcaggc	tcgataactg	gggtgaagtc	gtaacaaggt	agccgtatcg	1500
gaaggtgcgg	ctggatcacc	t				1521

&lt;210&gt; SEQ ID NO 115

&lt;211&gt; LENGTH: 1545

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Acidaminococcus fermentans

&lt;400&gt; SEQUENCE: 115

agagtttgat	octggctcag	gacgaacgt	ggccgcgtgc	ttaacacatg	caagtgcac	60
ggagaacttt	tttcggaaatg	ttcttagtgg	cgaacgggtt	agtaacgcgt	aggcaacctg	120
ccctctgggt	ggggacaaca	ttccgaaagg	gatgctaata	cogaatgaga	tcctttcc	180
gcatggagag	aggatgaaag	atggcctcta	cttgtaagct	atgcgcagaa	gatggcctg	240
cgtctgatta	gctagtaggt	gaggtaacgg	ctcacctagg	cgatgatcag	tagccggct	300
gagaggatga	acggccacat	tggactgag	acacggccca	aactcctacg	ggaggcagca	360
gtggggaaatc	ttccgcaatg	gacgaaagtc	tgacggagca	acgcccgcgt	agtgtatgaa	420
gccttcgggt	tgtaaaactc	tgttgtcagg	gacgaaagca	ccgatctata	atacatttg	480
gtgttgacgg	tacctgacga	ggaagccacg	gctaactacg	tgccagcagc	cgccgtata	540
cgttaggtggc	aagcgtgtc	cggattatt	ggcgtaaaag	agcatgtagg	cgggcttta	600
agtccgacgt	gaaaatcgcc	ggcttaaccc	cgtatggcgt	tggatactgg	aagtcttgag	660
tcaggaggag	gaaaggggaa	ttcccagtgt	agcggtaaaa	tgcgtagata	ttgggaggaa	720
caccagtggc	gaaggcgcct	ttctggactg	tgtctgacgc	tgagatgcga	aagccaggg	780
agcaaacggg	attagatacc	ccggtagtcc	tggccgtaaa	cgatgggtac	taggtgtagg	840
aggtatcgac	cccttctgt	ccggagttaa	cgcaataagt	accccgccct	gggactacga	900
tcgcgaagatt	gaaactcaaa	ggaattgacg	ggggccccca	caagcgggtt	agtatgttgt	960
ttaattcgac	gcaacgcgaa	gaaccttacc	aaggcttgcac	attgagtgaa	agacccagag	1020
atgggtcccc	ttcttcggaa	gcacgaaaac	aggtggtgca	tggctgtcgt	cagctcggt	1080
cgtgagatgt	tgggttaagt	cccgcaacga	gwgcaaccct	tatcctatgt	taccagcagc	1140
taatggtggg	gactcatagg	agactgccag	ggataacctg	gaggaaggcg	gggatgacgt	1200
caagtcatca	tgccccctt	gtcttggct	acacacgtac	tacaatggc	ggcaacaaag	1260
ggcagegaa	ccgcgaggcg	gagccaatcc	cagaaacccg	accccagttc	ggatcgccagg	1320
ctgcaaccgg	octgcgtgaa	gttggaaatcg	cttagtaatcg	caggtcagca	tactgcgggt	1380
aatacgttcc	cgggccttgt	acacaccgccc	cgtcacaccca	cgaaagttgg	taacacccga	1440
agccggtag	ataacctttt	aggagtcagc	tgtctaagg	ggggccgatg	attgggtga	1500
agtcgtacaca	aggtagccgt	tcgagaacga	gcccgtggat	cacct		1545

&lt;210&gt; SEQ ID NO 116

-continued

<211> LENGTH: 1423  
<212> TYPE: DNA  
<213> ORGANISM: Fusicatenibacter saccharivorans  
<400> SEQUENCE: 116

tggctcagga tgaacgctgg	60
cggcgctgtt aacacatgca	
agtgcagcga agcagttaag	
aagatttttc gatatgtttt ctgtactgt	120
agccggcgac gggtgatcaa cgcgtgggt	
acctgcggca taccggggga taacagctgg	180
aaacggctgc taataccgca taagegcaca	
gagctgcatt gctcggtgtg aaaaactccg	240
gtggtatggg atgggcccgc gtctgattag	
gcagttggcg gggtaacggc ccaccaaacc	300
gacgatcagt agccggccctg agagggcgac	
cgggccacatt gggactgaga cacggccaa actcctacgg	360
gaggcagcag tggggaaatat	
tgcacaatgg gggaaaccct gatgcagcga	420
cgccgcgtga gcgaagaagt atttcggtat	
gtaaagctct atcagcaggg aagataatga	480
cggtacctga ctaagaagcc cggtctaact	
acgtgccagc agccgcggta atacgttaggg	540
ggcaagcggt atccggattt actgggtgt	
aaggggagcgt agacggcaag gcaagtctga	600
tgtgaaaacc cagggttaa ccctgggact	
gcattggaaa ctgtctggct cgagtgcggg	660
agaggttaagc ggaattccta gtgtagcggt	
gaaatgcgt aatatttagga agaacaccag	720
tggcgaaggc ggcttactgg acggtaactg	
acggttgggc tcgaaagcgt ggggagcaaa	780
caggattaga taccctggta gtccacgccc	
taaacgcgtga atgcttaggtt ttggggagca	840
aagctcttcg gtgcccgc aaacgcatta	
agcatccac ctggggagta cgatcgcaag	900
aatgaaaactc aaaggaattt acggggaccc	
gcacaaggcg tggagcatgt ggttaattt	960
gaagcaacgc gaagaacctt accaggtctt	
gacatcccgta tgaccggccc gtaacggggc	1020
cttctttcg gagcattgga gacaggttgt	
gcatgggtgt cgtcagctcg tgcgtgaga	1080
tgttgggtta agtcccccaa cgagcgaac	
ccttattcctc agtagccagc aggttaagct	1140
gggcactctg tggagactgc caggataac	
ctggagggatgtcgtcaatca tcatgccc	1200
tatgtatctgg gctacacac	
tgctacaatg gcgtaaacaa agggaggcaa	1260
agccgcgagg tggagcaaat cccaaaaata	
acgtctcaatg tcggactgca	1320
gtctgcact cgactgcacg aagctggaa	
cgctagtaat	
ccgcaatcag aatgtcgccg tgaatacgtt	1380
cccggttctt gtacacaccc cccgtcacac	
catgggatgtt ggttaacgccc gaagtcagtg	1423
acccaacctt tta	

<210> SEQ ID NO 117  
<211> LENGTH: 1494  
<212> TYPE: DNA  
<213> ORGANISM: Ruminococcus chamanellensis  
<400> SEQUENCE: 117

agagtttgat octggctcag gacgaacgt	60
ggcgccacgc ctaacacatg caagtgcac	
ggagataaag acttcgggtt ttatcttagt	120
ggcgacggg ttagtaacac gtgagcaacc	
tgcctctgag agaggatag ctctggaa	180
cgatggtaa tacctcataa catagcgta	
ccgcatgata ctgctatcaa agatttatcg	240
ctcagagatg ggctcgctc tgattagcta	
gatggtgagg taacggctca ccatggcgcac	300
gatcagtgc cggactgaga ggttgaacgg	
ccacattggg actgagacac ggcccagact	360
cctacgggag gcagcagtgg ggaatattgc	
acaatgggcg caagectgat gcagcgtac	420
cgctggagg aagaagggtt tcggattgt	
aactcctgtc ttaagggacg ataatgacgg	480
taccttagga ggaagctccg gctaactacg	
tgccagcagc cgccgtataa cgtagggacg	540
gagcgttgc cggaattact gggtgtaaag	

-continued

ggagcgtagg	cgggattgca	agtcaagatgt	gaaaactatg	ggcttaaccc	atagactgca	600
tttgcggaaatcg	tagttcttga	gtgaagtaga	ggtaaggcgg	attccttagtg	tagcggtgaa	660
atgcgttagat	attaggagga	acatcggtgg	cgaaggcggc	ttactgggct	tttactgacg	720
ctgaggcgtcg	aaagcggtgg	gagcaaacag	gattagatac	cctggtagtc	cacgctgtaa	780
acgatgatta	ctaggtgtgg	ggggactgac	cccttccgtg	ccgcagttaa	cacaataagt	840
aatccacctg	gggagtagcgg	ccgcaagggtt	gaaaactcaaa	ggaattgacg	ggggcccgca	900
caagcagtgg	agtatgtgg	ttaattcgaa	gcaacgcgaa	aaaccttacc	aggcttgac	960
atcgagtgaa	tgtatctagag	atagatcagt	cttcggggac	acaaagacag	gtgggtgcgt	1020
gttgtcgtca	gctcgtgtcg	tgagatgttg	ggttaagtcc	cgcaacgagc	gcaaccctta	1080
ccttagttg	ctacgcaga	gcactctaga	gggactgccc	ttgacaaaac	ggaggaaggt	1140
ggggatgacg	tcaaatcatc	atgccccctta	tgacacctggc	tacacacgta	ctacaatggc	1200
aatgaacaga	gggaagcaat	acagtgtatgt	ggagcaaatc	ccaaaaatt	gtcccaagttc	1260
agattgttagg	ctgcaactcg	cctacatgaa	gtcgaaattg	ctagtaatcg	cagatcagca	1320
tgctgcggtg	aatacgttcc	cgggccttgt	acacaccgc	cgtcacacca	ttgggatgtcg	1380
taacacccgaa	agccagtagc	ctaaccgcaa	ggagggcgct	gtcgaagggt	ggattgtga	1440
ctgggggtgaa	gtcgtaacaa	ggtagccgta	tccggaaagggt	cggtctggatc	acct	1494

&lt;210&gt; SEQ ID NO 118

&lt;211&gt; LENGTH: 1533

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Bifidobacterium bifidum

&lt;400&gt; SEQUENCE: 118

ttttgtgga	gggttcgatt	ctggctcagg	atgaacgcgt	gccccgtgt	taacacatgc	60
aagtgcgaa	ggatccatcg	ggcttgcgtt	ggtggtgaga	gtggcgaacg	ggtgagtaat	120
gcgtgaccga	cctgcggccat	gctccggaa	agctccgttg	aacgggttgt	aatgcggat	180
gttccacatg	atcgcgtgt	attgtggaa	agattctatc	ggcgtggat	ggggtcgcgt	240
cctatcagct	tgttgtgag	gtaacggctc	accaaggctt	cgacgggtag	ccggcctgag	300
aggggcaccg	gccacattgg	gactgagata	cggcccgac	tccatacggtt	ggcagcgtg	360
gggaatattg	cacaatggc	gcaacgcgt	tgcagcgtac	ccgcgtgagg	gtggaggcc	420
ttcgggttgt	aaacctcttt	tgttgtggag	caagcctcg	ggtgatgtta	ccttcgaat	480
aagcgcggc	taactacgt	ccagcagccg	cgttaatacg	tagggcgcaa	gcgttatccg	540
gatttattgg	gcgtaaaggg	ctcgtaggcg	gctcgctcg	tccggtgtga	aagtccatcg	600
cttaacgggt	gatctgcgcc	gggtacgggc	gggctggagt	gcggtagggg	agactggaaat	660
tcccggtgt	acgggtgaaat	gtgttagatat	cgggaagaac	accgatggcg	aaggcaggtc	720
tctggggccgt	cactgacgct	gaggagcga	agcgtgggg	gcgaacagga	ttagataccc	780
tggtagtcca	cgcgtaaac	ggtgacgct	ggatgtgggg	cacgttccac	gtgtccgt	840
tccggatcaa	cgcgtaa	gggatgtacgg	ccgcaaggct	aaaactcaaa		900
gaaattgacg	ggggcccgca	caagcggcgg	agcatgcgg	ttaattcgat	gcaacgcgaa	960
gaaccttacc	tgggcttgcac	atgttcccg	cgacgcaga	gtggcggtt	cccttcgggg	1020
cgggttcaca	ggtgggtgcat	ggtcgtcg	agtcgtgtc	gtgagatgtt	gggttaagtc	1080
ccgcaacgag	cgcaaccctc	gccccgtgtt	gccagcacgt	tatggtgaaa	actcacgggg	1140

-continued

---

gaccgcgggg	gttaactcg	aggaagggtgg	ggatgacg	tc	agatcatcat	gcccttacg	1200
tccagggc	t	cacgc	atg	gtc	atgcgacat	g	1260
g	cgatcc	aaa	acc	cggt	tc	ca	1320
cg	gatc	gt	aa	atcg	cg	cc	1380
ca	cac	cc	gt	atcg	ca	cc	1440
cc	ggat	gg	aa	atcg	cc	tt	1500
tt	cc	gg	aa	atcg	cc	tt	1533

&lt;210&gt; SEQ ID NO 119

&lt;211&gt; LENGTH: 1552

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Megasphaera elsdenii*

&lt;400&gt; SEQUENCE: 119

agagtttgat	cctggctcag	gacgaacg	ct	ggcgcgtgc	ttaacacat	g	caagt	cgAAC	60
gagaagagat	gagaag	ctt	tttat	ca	ttc	att	cgagt	gg	120
a	gcaac	cc	ttc	ca	gg	ggaca	aa	ac	180
t	ttttt	gtc	cat	ggc	ggc	aa	ag	aa	240
gg	ctt	gcgt	tgat	ggc	ggc	gg	gg	gtc	300
cc	gtc	tgcgt	tttgc	at	ggc	at	ggc	at	360
gg	cag	tcgt	tttc	cg	aat	gg	at	gg	420
g	gac	cgat	tttc	ca	ttt	gg	at	gg	480
cc	gtt	acgt	tttc	cg	at	gg	at	gg	540
gg	ata	atcg	tttc	cg	at	gg	at	gg	600
cg	tgt	atcg	tttc	cg	at	gg	at	gg	660
gt	tag	atcg	tttc	cg	at	gg	at	gg	720
gg	agg	atcg	tttc	cg	at	gg	at	gg	780
cc	agg	atcg	tttc	cg	at	gg	at	gg	840
gt	tag	atcg	tttc	cg	at	gg	at	gg	900
at	tg	atcg	tttc	cg	at	gg	at	gg	960
at	gt	atcg	tttc	cg	at	gg	at	gg	1020
gat	ag	atcg	tttc	cg	at	gg	at	gg	1080
ag	tc	atcg	tttc	cg	at	gg	at	gg	1140
acc	age	atcg	tttc	cg	at	gg	at	gg	1200
gat	ac	atcg	tttc	cg	at	gg	at	gg	1260
ta	at	atcg	tttc	cg	at	gg	at	gg	1320
at	tg	atcg	tttc	cg	at	gg	at	gg	1380
ct	cg	atcg	tttc	cg	at	gg	at	gg	1440
ac	acc	atcg	tttc	cg	at	gg	at	gg	1500
gg	gg	atcg	tttc	cg	at	gg	at	gg	1552

&lt;210&gt; SEQ ID NO 120

&lt;211&gt; LENGTH: 1479

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Dorea formicigenerans*

-continued

---

```

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (123)..(125)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (220)..(222)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (287)..(287)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (349)..(350)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (517)..(517)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (534)..(535)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (725)..(725)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (918)..(918)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (921)..(921)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (955)..(956)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1184)..(1185)
<223> OTHER INFORMATION: n is a, c, g, t or u

<400> SEQUENCE: 120
ttaaacgaga gtttgatcct ggctcaggat gaacgctggc ggcgtgctta acacatgcaa      60
gtcgagcgaa gcacataagt ttgattcttc ggatgaagac ttttgtact gagcggcgga     120
cgnnnngagta acgcgtgggt aacctgcctc atacaggggg ataacagyta gaaatggctg     180
ctaataccgc ataagaccac agtactgcat ggtacagtgn nnaaaactcc ggtggtatga     240
gatggacccg cgtctgatta ggttagttgt gaggtAACCG cccaccnacg cgacgatcag     300
tagccgacct gagaggggtga ccggccacat tgggactgag acacggccnn gactcctacg     360
ggaggcagca gtgggaaata ttgcacaatg ggcgaaagcc tgatgcagcg acgcccgtg     420
aaggatgaag tatttcggta tgtaaacttc tatcagcagg gaagaaaatg acggtaactg     480
actaagaagc cccggctaac tacgtgccag cagccgnggt aatacgtagg gggnnagcgt     540
tatccggatt tactgggtgt aaaggagcgc tagacggctg tgcaagtctg aagtgaaagg     600
catgggctca acctgtggac tgctttggaa actgtgcagc tagagtgtcg gagaggtaa     660
tggaaattcct agtgttagcgg tgaaaatgcgt agatattagg aggaacacca gtggcgaagg     720
cgccntactg gacgatgact gacggttgagg ctgcgaaagcg tggggagcaa acaggattag     780
ataccctggt agtccacgccc gtaaacgtatc actgcttaggt gtcgggtacg aaagctattc     840
ggtgccgcag ctaacgcaat aagcagtcca cctggggagt acgttcgcaa gaatgaaact     900
caaaggaatt gacgggncc ngcacaagcg gtggagcatg tggtttaatt cgaannaacg     960

```

## US 10,555,980 B2

319

320

-continued

---

cgaagaacct tacctgatct tgacatcccg atgaccgctt cgtaatggaa gytttcctc 1020  
 ggaacatcg tgacaggtgg tgcatggtt tcgtcagctc gtgtcgtag atgttgggtt 1080  
 aagtcccgca acgagcgcaa cccttatctt cagtagccag catttaggat gggcactctg 1140  
 gagagactgc cagggataac ctggaggaag gtggggatga cgtnnaatca tcatgccct 1200  
 tatgaccagg gctcacacag tgctacaatg gcttaaacag agggaggcag agccgcgagg 1260  
 ccgagcaaat ctcaaaaata acgtctcagt tggattgtt gtctgcaact cgactacatg 1320  
 aagctggaat cgctagtaat cgccagatcg aatgctcgg tgaatacgtt cccgggtctt 1380  
 gtacacaccc cccgtcacac catggagtc agtaacgcc gaagtcagtg acccaaccga 1440  
 aaggagggag ctgcccgaagg tgggaccgt aactgggg 1479

&lt;210&gt; SEQ ID NO 121

&lt;211&gt; LENGTH: 1390

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Eisenbergiella tayi

&lt;400&gt; SEQUENCE: 121

ggtataactt agtggcggac gggtgagtaa cgcggtggaa acctgcctc taccgggg 60  
 taacacttag aaataggtgc taataccgca taagcgcacg gaaccgcatt gttccgtgt 120  
 aaaaactccg gtggtacagg atggtcccgc gtctgattag ccagttggca gggtaacggc 180  
 ctaccaaaggc gacgatcagt agccggcctg agagggtgaa cggccacatt gggactgaga 240  
 cacggcccaa actcctacgg gaggcagcag tggggatata tgcacaatgg gggaaaccct 300  
 gatgcagcga cgccgcgtga gtgaagaagt atttcggat gtaaagctt atcagcagg 360  
 aagaaaaatga cggtacctga ctaagaagcc ccggctaact acgtgccagc agccgcgtt 420  
 atacgtggg ggcaagcggtt atccggattt actgggtgtt aaggggagcgt agacggcatg 480  
 gcaaggcaga tggaaaacc caggcgttcc cttggattt gcatttggaa ctgccaggct 540  
 ggagtgcagg agaggttggc ggaattccca gtgtagcggt gaaatgcgtt gatattagga 600  
 ggaacaccag tggcgaaggc ggcttactgg actgttaactg acgttggggc tcgaaagcgt 660  
 ggggagcaaa caggattaga taccctggta gtccacgcgg taaacgttgcgtt ttgcttaggt 720  
 taggtgggtt tggaccatc ggtggccgcag ctaacgcattt aagcaatcca cctggggagt 780  
 acgttcgcaaa gaatgaaact caaaggattt gacggggacc cgcacaagcg gtggggcatg 840  
 tgggttaattt cgaagcaacg cgaagaacctt taccatgtt tgacatccca atgacgcacc 900  
 tgtaaagagg tggcccttc ggggcattgg agacagggtgg tgcatgggtt tcgtcagctc 960  
 gtgtcgtag atgttgggtt aagtcccgca acgagcgcaa cccttattct tagtagccag 1020  
 caggttaaaggc tggcactctt aaggagactg ccggggatata cccggaggaa ggcggggatg 1080  
 acgtcaaattt atcatgcccc ttatgattt ggcttacacac gtgttacaat ggcgttac 1140  
 aagggaagcg agacagtgtt gtggggatata tccytagaaat aacgttcttgcgtt ttcggatgtt 1200  
 agtctcaac tggactacat gaagctggaa tcgttacatca tcgttacatca gcatgtcg 1260  
 gtgttacatgtt tccgggtt tggatccacc cccgttacatcc ccatggggatg tggaaatgcc 1320  
 cgaagtctgtt gacctaaccg aaaggggaggaa gcaaggcgtt gcaagggttgcgtt taactgggg 1380  
 gaagtcgtt 1390

&lt;210&gt; SEQ ID NO 122

&lt;211&gt; LENGTH: 1478

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Clostridium symbiosum

-continued

---

```

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (349)..(350)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (534)..(535)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (606)..(606)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (618)..(619)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (709)..(709)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (921)..(921)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (955)..(956)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1012)..(1013)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1015)..(1015)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1081)..(1081)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1185)..(1185)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1240)..(1240)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1391)..(1392)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1470)..(1470)
<223> OTHER INFORMATION: n is a, c, g, t or u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1474)..(1475)
<223> OTHER INFORMATION: n is a, c, g, t or u

<400> SEQUENCE: 122

aaacatgaga gtttgatcct ggctcaggat gaacgctggc ggcgtgccta acacatgcaa      60
gtcgaacgaa gcgatttaac ggaagtttc ggatggaagt tgaattgact gagtggcgga     120
cgggtgagta acgcgtgggt aacctgcctt gtactgggg acaacagtta gaaatgactg     180
ctaataaccgc ataagcgcac agtattgcat gatacagtgt gaaaaactcc ggtggtacaa     240
gtatggaccgg cgtctgatta gctagtttgt aaggtaacgg cttaccaagg cgacgatcag    300
tagccgacct gagagggtga ccggccacat tgggactgag acacggccnn aactcctacg    360
ggaggcagca gtgggaaata ttgcacaatg ggcgaaagcc tgatgcagcg acgccgcgtg    420
agtgaagaag tatttcggta tgtaaagctc tatcagcagg gaagaaaatg acggtacctg    480

```

-continued

---

actaagaagc cccggctaac tacgtgccag cagccgcggt aatacgtagg gggnnnacgt	540
tatccggatt tactgggtgt aaaggggagcg tagacggtaa agcaagtctg aagtgaaagc	600
ccgcgnctca actgegggnnc tgctttgaa actgttaac tggagtgtcg gagaggttaag	660
tggaattcct agtgttagcgg tgaaatgegt agatattagg aggaacacna gtggcgaagg	720
cgacttactg gacgataact gacgttgagg ctgcggaaacg tggggagcaa acaggattag	780
atacccttgt agtccacgccc gtaaacgtat aatactagggt gttggggagc aaagtcctc	840
ggtgccgtcg caaacgcagt aagtattcca cctggggagt acgttcgcaaa gaatgaaact	900
caaaggaaatt gacggggacc ngcacaagcg gtggagcatg tggttaatt cgaannaacg	960
cgaagaacct taccaggctc tgacatcgac tcgacgggggg agtaacgtcc cnntncctc	1020
ggggcggaga agacagggtgg tgcatgggtc tgcgtcagctc gtgtcgtag atgttgggtt	1080
nagtcccgca acgagcgcaaa cccttattct aagtagccag cgggtcggcc gggactctt	1140
ggggagactgc cagggataaac ctggaggaag gtggggatga cgtcnaatca tcatgcccct	1200
tatgatctgg gctcacacgc tgctacaatg gcgtaaacan agagaagcaa gaccgcgagg	1260
tggagcaaat ctcaaaaata acgtctcagt tcggactgcg ggctgcaact cgcctgcacg	1320
aagctggaat cgctagtaat cgcaaatcgt aatgtcgccg tgaatacgtt cccgggtctt	1380
gtacacacccg nncgtcacac catggggatc agtaacgccc gaagtcagtg acccaaccgc	1440
aaggagggag ctgccgaagg cgggaccgan aacnnggg	1478

&lt;210&gt; SEQ\_ID NO 123

&lt;211&gt; LENGTH: 1488

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Erysipelatoclostridium ramosum

&lt;400&gt; SEQUENCE: 123

agagtttgat cctggctcag gatgaacgc ggcggcgtgc ctaatacatg caagtcgaac	60
gcgagcaccc ttgtcgaggt ggcgaacggg tggatataac ataagtaacc tgccctagac	120
agggggataa ctattggaaa cgtatagctaa gaccgcatacgt acggacacac tgcatggta	180
ccgttattaa aatgtgcctaa agcactggta gaggatggac ttatggcgca ttagctggtt	240
ggcggggtaa cggcccacca aggccgacgt gcgtagccgaa cctgagaggg tgaccggcca	300
cactgggact gagacacccgc ccacactctt acggggaggca gcagtagggaa attttcggca	360
atggggaaa ccctgaccga gcaacgcgcgtc gtgaaggaaag aaggtttcg gattgtaaac	420
ttctgttata aaggaagaac ggcggctaca gggaaatggta ggcgagtgac ggtactttat	480
tagaaagcca cggcttaacta cgtgccagca gcccggtaa tacgttaggtc gcaagcgta	540
tccggatta ttggcgtaa agggggagca ggcggcagca aggggtctgtc gtgaaaggcct	600
gaagcttaac ttcaactaacc catagaaacc aggcagctag agtgcaggag aggatcggtt	660
aattccatgt gtatcggtt aatgcgttataatggagg aacaccgttgcgaaaggcga	720
cgatctggcc tgcaactgac gctcagtc gaaagcgtgg ggagcaataa ggattagata	780
cccttagtgc ccacgcgtt aacgtgttgc actaagtgtt ggtatgtcaaa gttcagtgct	840
gcagttaacg caataagtac tccgcctgag tagtacgttc gcaagaatga aactcaaagg	900
aattgacggg ggcccgacca agcgggtggag catgtggttt aattcgaagc aacgcgaaga	960
accttaccatg gtcttgcacat actcataaaag gtcggagaga tggagagata gctatatgag	1020
atacagggtgg tgcatgggtc tgcgtcagctc gtgtcgtag atgttgggtt aagtcccgca	1080
acgagcgcaaa cccttatacgtagtaccat cattaagttt gggactcttag cgagactgccc	1140

-continued

```

agtgacaagg tggaggaagg cggggatgac gtcaaattcat catgcccctt atgacctggg    1200
ctacacacgt gctacaatgg atggtcaga gggaaagcgaa gccgcgaggt gaagcaaaac    1260
ccataaaaacc attctcagtt cggattttagt tctgcaactc gactacatga agttggaaatc   1320
gttagtaatc gcgaatcagc atgtcgccgt gaataacgttc tcgggecttg tacacaccgc   1380
ccgtcacacc acgagagttg ataacacccg aagccggtgg cctaaccgca aggaaggagc   1440
tgtctaagggt gggattttagt attggggatc agtcgttaaca aggttaacc                      1488

```

<210> SEQ ID NO 124  
<211> LENGTH: 1528  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 124

```

atgagagttt gatcctggct caggatgaac gctggcggcg tgcctaacac atgcaagtgc    60
aacgaagcaa ttaaaggaag ttttcggatg gaatttgatt gactgagtgg cggacgggtg    120
agtaacgcgt ggataacactg cctcacactg gggataaca gttagaaatg actgctaata   180
ccgcataaggc gcacagtacc gcattggtacg gtgtgaaaaa ctccgggtgt gtgagatgg    240
tcggcgtctg attagccgt tgccgggtta acggcccacc aaagcgacga tcagtagccg   300
acctgagagg gtgaccggcc acattgggac tgagacacgg cccaaactcc tacgggaggc   360
agcagtgggg aatattgcac aatgggcgaa agcctgtatgc agcgacgcgg cgtgagtgaa  420
gaagtatttc ggtatgtaaa gctctatcg cagggaaagaa aatgacggta cctgactaag  480
aagccccggc taactacgtg ccagcagccg cggtataacg tagggggcaaa gcgttatccg  540
gattcaactgg gtgtaaaggg agcgttagacg gcgaagcaag tctgaagtga aaaccaggg  600
ctcaaccctg ggactgcttt ggaaactgtt ttgcttagatgt gtccggagagg taagtggaa 660
tccttagtgcg gcggtgaaat gcttagatcat taggagggac accagttggcg aaggggctt 720
actggacgat aactgacgat gaggctcgaa agcgtgggg gcaaacagga tttagatccc  780
tggtagtcca cggcgtaaac gatgaatgtt aggtgttggg gggcaaaaggcc ctccgggtcc 840
gtcgcaaaacg ctagtaagcat tccacctggg gagtacgttc gcaagaatga aactcaaagg 900
aattgacggg gacccgcaca agcgggtggag catgtggttt aattcgaagc aacgcaaga  960
accttaccaaa gtcttgcacat cctcttgacc ggcgtgtaaac ggccgccttc ctccgggca 1020
agagagacag gtgggtgcgt gttgtcgta gctcgtgtcg tgagatgttg ggttaagtcc 1080
cgcaacgcgc gcaaccctta tccttagtag ccagcaggta gagctggca ctcttagggag 1140
actgccaggataaactggg ggaagggtggg gatgacgtca aatcatcatg ccccttatga 1200
tttgggtctac acacgtgtca caatggcgta aacaaaggaa agcaagacag ttagtggag 1260
caaatacccaa aaataacgtc ccagttcgaa ctgttagtctg caacccgact acacgaagct 1320
ggaatcgcta gtaatcgca atcagaatgt cgccggtaat acgttccgg gtcttgcata 1380
caccgcgggtt cacaccatgg gagtcagcaa cgcccgaaatg cagtgcacca actcgcaaga 1440
gagggagctg ccgaaggccg ggcaggtaac tgggtgaag tcgttaacaag gtagccgtat 1500
cggaagggtgc ggctggatca cctccctt 1528

```

<210> SEQ ID NO 125  
<211> LENGTH: 1530  
<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 125

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaacac atgcaagtgc	60
aacgaagcaa ttaaaatgaa gtttcggat ggattttga ttgactgagt ggccggacggg	120
ttagtaacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa	180
taccgcataa ggcacacgta ccgcacatggta cggtgtgaaa aactccggtg gtgtggatg	240
gatccgcgtc tgattagcca gttggcgcccc taacggccca ccaaagcgcac gatcagtgc	300
cgcacctgaga ggggtgacccgg ccacattggg actgagacac ggcccaaact cctacggag	360
gcagcagtgg ggaatattgc acaatggcg aaaggctgtat gcagcgcacgc cgcgtgatg	420
aagaagtatt tcggtatgtt aagctctatc agcagggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcgc cgcgtataa cgtagggggc aagcgttatc	540
cggatttact ggggttaaaag ggagcgtaga cggcgaagca agtctgtatg gaaaaccagg	600
ggctcaaccc tgggactgct ttggaaactg ttttgcataa gtgtcgaaaa ggttaagtgg	660
attccttagtg tagcgtgtt atgcgtatgat attaggagga acaccagtgg cgaaggccgc	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcgttaa acgtgtatc cttaggtgtt gggggcaaa cccttcgggt	840
ccgtcgaaac cgcgtaaacg attccacctg gggagttacgt tcgcaagaat gaaactcaaa	900
ggaattgacg gggaccggca caagcgggtgg agcatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc aagtcttgac atcccttgc cggcgtgtt acggcgcctt cccttcgggg	1020
caagagagac aggtgttgca tgggtgtcg cagctcgatc cgtgagatgt tgggttaagt	1080
cccgcaacca ggcacccct tatccttagt agccagcagg taaagctggg cactctaggg	1140
agactgcccgg ggtataacccgt gaggaaagggtg gggatgacgt caaatcatca tgccccttat	1200
gatttggctt acacacgtgc tacaatggcg taaacaaagg gaagcaagac agtgtatgtt	1260
agcaaatccc aaaaataacg tcccgatcg gactgtatgc tgcaacccga ctacacgaag	1320
ctggaatcgc tagtaatcgc gaatcagaat gtcgcgtgtt atacgttccc gggctttgtt	1380
cacacccccc gtcacccat gggagtcgc aacgccccgaa gtcagtgacc caactcgaa	1440
gagagggagc tgccgaaggc gggggcaggta actgggggtga agtcgtaaaca aggttagccgt	1500
atcggaaagggtt gcggtggat cacctccctt	1530

<210> SEQ ID NO 126  
<211> LENGTH: 1528  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 126

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaacac atgcaagtgc	60
aacgaagcaa ttaaaatgaa gtttcggat ggattttga ttgactgagt ggccggacggg	120
ttagtaacgc gtggataacc tgcctcacac tgggggataa cagttagaaa tgactgctaa	180
taccgcataa ggcacacgta ccgcacatggta cggtgtgaaa aactccggtg gtgtggatg	240
gatccgcgtc tgattagcca gttggcgcccc taacggccca ccaaagcgcac gatcagtgc	300
cgcacctgaga ggggtgacccgg ccacattggg actgagacac ggcccaaact cctacggag	360

## US 10,555,980 B2

**329****330**

-continued

gcagcagtgg ggaatattgc acaatggcg aaagcctgat gcagcgacgc cgctgtgagtg	420
aagaagtatt tcggtatgtt aagctctatc agcagggaaag aaatgacggt acctgactaa	480
gaagccccgg ctaactacgt gccagcagcc gcggtataac gtagggggca agcgttatcc	540
ggatttactg ggtgtaaagg gacgttagac ggcaagcaa gtctgaagtg aaaaccagg	600
getcaaccct gggactgctt tggaaactgt tttgctagag tgctggagag gtaagtggaa	660
tccctagtgt agcggtgaaa tgctgttagata ttaggaggaa caccagtggc gaaggcgct	720
tactggacga taactgacgt tgaggctcga aagcgtgggg agcaaacagg attagataacc	780
ctggtagtcc acgcccgtaaa cgtatgtc taggtgttgg gggcaaagcc cttecggtgcc	840
gtcgcaaacg cagtaagcat tccacctggg gagtacgttc gcaagaatga aactcaaagg	900
aattgacggg gaccggcaca agcgggtggag catgtggttt aattcgaagc aacgcegaaga	960
accttaccaa gtcttgacat cctcttgacc ggcgtgtAAC ggcgccttcc cttecggtgcc	1020
agagagacag gtgggtgcgtg gttgtcgta gctcggtcg tgagatgtg ggttaagtcc	1080
cgcacacgac gcaaccctta tccttagtag ccagcaggtt aagctggca ctcttagggag	1140
actgccaggg ataacctggg ggaaggtggg gatgacgtca aatcatcatg ccccttatga	1200
tttgggtac acacgtgcta caatggcgta aacaaaggaa agcaagacag tgatgtggag	1260
caaataccaa aaataacgtc ccagttcggta ctgtgtctg caacccgact acacgaagct	1320
ggaategcta gtaatcgca atcagaatgt cgccgtgaat acgttccgg gtcttgata	1380
caccgeccgt cacaccatgg gagtcagca cggccgaagt cagtgaccga actcgcaaga	1440
gagggagctg ccgaaggcgg ggcaggtaac tgggtgaag tcgtaacaag gtagecgat	1500
cggaagggtgc ggctggatca cctccctt	1528

&lt;210&gt; SEQ ID NO 127

&lt;211&gt; LENGTH: 1528

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 127

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaacac atgcaagtgc	60
aacgaagcaa ttaaatgaa gtttcggat ggattttgtt ttgactgagt ggcggacggg	120
ttagtacacgc gtggataacc tgcctcacac tgggggataa cagttagaaaa tgactgctaa	180
taccgcataa gcgcacagta ccgcattgtt cggtgtgaaa aactccggtg gtgtgagatg	240
gatccgcgtc tgattagcca gttggcgcccc taacggccca ccaaagcgac gatcgtac	300
cgacctgaga gggtgaccgg ccacattggg actgagacac ggcccaact cctacggag	360
gcagcagtgg ggaatattgc acaatggcg aaagcctgat gcagcgacgc cgctgtgagt	420
aagaagtatt tcggtatgtt aagctctatc agcagggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacgt tgccagcagc cgccgtataa cgtagggggc aagcgttatc	540
cgatattact ggggtgtaaag ggacgttaga cggcgaagca agtctgaagt gaaaaccagg	600
ggctcaaccc tggactgct ttggaaactg tttgtctaga gtgtcgagag ggtaaagtgg	660
atccctagtgt tagcggtgaa atgcgttagat attaggagga acaccagtgg cgaaggcgcc	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacggcgtaa acgtgaatg ctagtggttt gggccaaagc cttcggtgc	840

-continued

---

cgtcgaaac gcagtaagca ttccacactgg ggagtagtgcgtt cgcaagaatg aaactcaaag	900
gaattgacgg ggacccgcac aagegggtgga gcatgtgggtt taattcgaag caacgcgaag	960
aaccttacca agtcttgaca tcctcttgcac cggcgtgtaa cggcgccttc cttcggggc	1020
aggagagaca ggtggtgcat ggttgcgtc agtcgtgtc gtgagatgtt gggtaagtgc	1080
ccgcaacgag cgcaaccctt atccttagta gccagcagg agagctggc actctaggaa	1140
gactgccagg gataaacctgg aggaagggtgg ggatgaegtc aaatcatcat gccccttatg	1200
atttgggcta cacacgtgct acaatggcgt aaacaaaggaa aagcaagaca gtgatgtgga	1260
gcaaataccca aaataacgtc ccagttcgga ctgttagtctg caacccgact acacgaagct	1320
ggaatcgcta gtaatcgca atcagaatgt cgccgtgaat acgttcccggtt gtcttgtaca	1380
caccggccgt cacaccatgg gagtcagcaa cgcccgaaatg cagtgaccca actcgcaaga	1440
gagggagctg ccgaaggcgg ggcaggtaac tgggtgaag tcgtaacaag gtggccgtat	1500
cggaagggtgc ggctggatca cctccctt	1528

&lt;210&gt; SEQ ID NO 128

&lt;211&gt; LENGTH: 1528

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 128

atgagagttt gatcctggct caggatgaac gctggcgccg tgcctaacac atgcaagtcg	60
aacgaagcaa taaaatgaa ttttcgat ggattttaat tgactgagtg gcggacgggt	120
gagtaacgcg tggataacct gcctcacact gggggataac agttagaaat gactgctaatt	180
accgcataag cgcacagtac cgcacgttac ggtgtgaaaa actccgggtgg tggatgtgg	240
atccgcgtct gattagccag ttggcggtt aacggccac caaagcgacg atcagtagcc	300
gacctgagag ggtgacccggc cacattggga ctgagacacg gcccaaactc ctacggagg	360
cagcagtggg gaatattgca caatggcgaa aagcctgtatc cagcgcacgccc gcgtgagtga	420
agaagtattt cggatgttacgatc cggatgttacgatc gcaggaaaga aatgacggtt acctgtactaa	480
gaagccccgg ctaactacgt gccagcagcc gcggtaatac gtggggggca agcgttatcc	540
ggatttactg ggtgttacgg gacgttagac ggcaagcaaa gtctgttttttggaaaatccagg	600
gtcaaccctt gggactgtttt tggaaactgtt tttgttagatc tgccggagatc gtaatgtggaa	660
tcccttagtgtt agcggtggaaa tgcgttagata ttggggggca caccgtggc gaaggcggtt	720
tactggacga taactgacgt tgaggctcgaa aagcgtgggg agcaaacagg attagataacc	780
ctggtagtcc acggccgtaaa cgtatgttgc taggtgttgg gggcaaaatgc cttcgggtgc	840
cgtcgaaac gcagtaagca ttccacactgg ggagtagtgcgtt cgcaagaatg aaactcaaag	900
gaattgacgg ggacccgcac aagegggtgga gcatgtgggtt taattcgaag caacgcgaag	960
aaccttacca agtcttgaca tcctcttgcac cggcgtgtaa cggcgccttc cttcggggc	1020
agagagaca ggtggtgcat ggttgcgtc agtcgtgtc gtgagatgtt gggtaagtgc	1080
ccgcaacgag cgcaaccctt atccttagta gccagcagg agtggccgtt acctctaggaa	1140
gactgccagg gataaacctgg aggaagggtgg ggatgtggcgtt aaatcatcat gccccttatg	1200
atttgggcta cacacgtgct acaatggcgt aaacaaaggaa aagcaagaca gtgatgtgga	1260
gcaaataccca aaataacgtc ccagttcgga ctgttagtctg caacccgact acacgaagct	1320
ggaatcgcta gtaatcgca atcagaatgtt cgccgtgaat acgttcccggtt gtcttgtaca	1380

-continued

```

caccgcccgt cacaccatgg gagtcagcaa cgcccgaaagt cagtgaccca actcgcaaga 1440
gagggagctg ccgaaggcgg ggcaggtaac tggggtaaag tcgtaaacaag gtagccgtat 1500
cggaagggtgc ggctggatca cctccctt 1528

<210> SEQ ID NO 129
<211> LENGTH: 1522
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 129

tcaaagagtt tgatcctggc tcaggacgaa cgctggcggc ggcctaaca catgcaagtc 60
gaacggagct tacgtttga agtttcgga tggatgaatg taagcttagt ggccggacggg 120
tgagtaacac gtgagcaacc tgccttcag agggggataa cagccggaaa cggctgctaa 180
taccgcatga ttttgcgggg gcacatgccc ctgcaaccaa aggagcaatc cgctgaaaga 240
tgggctcgcg tccgatttgc cagttggcgg ggttaacggcc caccaaagcg acgatcgta 300
gcggactga gagggttgaac ggccacattt ggactgagac acggcccaga ctcctacggg 360
aggcagcagt gggggatatt gcacaatggg cgaaagctg atgcagcgac gccgcgtgag 420
ggaagacggt ttccggattt taaacctctg tctttggggg agaaaaatgac ggtacccaaa 480
gaggaagctc cggctaacta cgtgccagca gcccggtaa tacgtaggaa gcaagcggt 540
tccggaatta ctgggtgtaa agggagcgtt ggcgggatgg caagtagaat gttaaatcca 600
tcggctcaac cggtggtgc gttctaaact gccgttctt agtgaagtag aggcaggcgg 660
aattcctagt gtatcggtga aatgcgtaga tattaggagg aacaccagtgc gcaaggcgg 720
cctgctgggc tttaactgac gctgaggctc gaaagcgtgg ggagcaaaaca ggattagata 780
ccctggtagt ccacgcgtt aacgtgatt actagggtgtt gggggactga cccctccgt 840
gccgcgttta acacaataag taatccaccc tgggagatcg gccgcaaggt tgaaactcaa 900
aggaatttgc gggggccgc acaaggatgtt gatgtatgtt tttaatttgc agcaacgcga 960
agaacccatc cagggttttgc catcgatgc atagcctaga gatagggtt gcccgggg 1020
gcatccagac aggtgggtca tgggttgtgt cagctcggt cgtgagatgt tgggttaagt 1080
cccgcaacgc ggcacccctt tattattgt tgctacgaa gggactcta atgagactgc 1140
cggtgacaaa acggaggaag gtggggatgtt cgtcaatca tcatggccct tatgacctgg 1200
gtacacacgt tactacaatg gcactaaac agagggcggc gacaccgcga ggtgaagcga 1260
atccccaaaa agtgcgttgc ttccggatttgc aggctgcaac cccctgcgtt gaagtcggaa 1320
ttgcttagaa tcgggatca gcatgcccgg gttttttttt tttttttttt tttttttttt 1380
gcccgtcaca ccatggggatgtt cggtaacacc cgaaggccagt agcctaaccg caagggggc 1440
gtgtcgaaatgtt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 1500
gtgcggctgg atcacccctt 1522

```

```

<210> SEQ ID NO 130
<211> LENGTH: 1521
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 130

```

-continued

---

tcaaagagtt tgatcctggc tcaggacgaa cgctggcgcc ggcctaaca catgcaagtc	60
gaacggagct tacgtttga agtttcgga tggacgaatg taagcttagt ggccggacggg	120
ttagtaaacac gtgagcaacc tgcccttcag aggggataac agccggaaac ggctgctaat	180
accgcattgtat gttgggggg cacatcccc tcacaacaaa ggagcaatcc gctgaaagat	240
gggctcgctt cccattagcc agttggcgcc gtaacggccc accaaagcga cgatcggtag	300
ccggacttag aggttgaacg gccacattgg gactgagaca cggcccgac tcctacggg	360
ggcagcagtggggatattt cacaatggc gaaagctga tgcagcgacg ccgcgtgagg	420
gaagacggc ttccggattgt aaacctctgt ctggggggaa gaaaatgacg gtacccaaag	480
aggaagctcc ggcttaactac gtgccagcag ccgcgttaat acgttagggag caagcggtgt	540
ccggattttac tgggtgtaaa gggagcgtag gcggggatggc aagttagatg ttAAATCCAT	600
ccggctcaacc ggtggctgcg ttctaaactg ccgttcttgta gtgaagtaga ggcaggcgg	660
attccctatgt tagcggtgaa atgcgttagat attaggagga acaccagtgg cgaaggcggc	720
ctggctgggct ttaactgacg ctggggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcgttaa acgtatgatta cttagtgtgg gggactgac cccttcgt	840
ccgcgtttaa cacaataagt aatccacctg gggagttacgg ccgcgtttaa gaaactcaaa	900
ggaaattgacg gggcccgca caagcgtgg agtatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc aggttcttgac atcggatgca tagcctagag ataggtgaag cccttcgggg	1020
catccagaca ggtgggtcat ggttgtcgat agctcgatgtc gtgagatgtt gggtaatgc	1080
ccgcacacgag cgcaaccctt attattatgt gctacgcaag agcactctaa tgagactgcc	1140
gttgacaaaa cggaggaagg tggggatgac gtcaaattatcat catgccccctt atgacctggg	1200
ctacacacgt actacaatgg cactaaaaca gaggcgccg acaccgcgag gtgaagcgaa	1260
tccccaaaaa gtgtctcagt tcagattgca ggctgcaacc cgctgtcatg aagtcggaaat	1320
tgcgtatataat cgcggatcag catggcgccg tgaatacgtt cccggccctt gtacacaccc	1380
cccggtcacac catggagtc ggtaacaccc gaagccagta gcctaaccgc aagggggcg	1440
ctgtcgaagg tggattgtat gactgggggtg aagtcgtaac aaggttagccg tatcggaagg	1500
tgccggctgga tcacccctt t	1521

&lt;210&gt; SEQ ID NO 131

&lt;211&gt; LENGTH: 1522

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 131

tcaaagagtt tgatcctggc tcaggacgaa cgctggcgcc ggcctaaca catgcaagtc	60
gaacggagct tacgtttga agtttcgga tggatgaatg taagcttagt ggccggacggg	120
ttagtaaacac gtgagcaacc tgcccttcag aggggataaa cagccggaaa cggtgtctaa	180
taccgcattgtat gttgggggg gcacatgccc ctgcaaccaa aggagcaatc cgctgaaaga	240
tgggctcgcc tccgatttagc cagttggcgcc ggtaacggcc caccaaaagcg acgatcgta	300
ggccggactga gaggttgaac ggccacattt ggactgagac acggcccaga ctcctacggg	360
aggcagcagt gggggatattt gcacaatggg cggaaacgtt atgcagcgac ggcgtgttag	420
ggaagacggt ttccggattt taaacctctg tctttggggaa agaaaatgac ggtacccaaa	480
gaggaagctc cggcttaacta cgtgccagca gccgggtaa tacgtaggga gcaagcggt	540

-continued

tccggaaatta	ctgggtgtaa	agggagcgtta	ggcgggatgg	caagtagaat	gttaaatcca	600
tccggctaac	cgggtggctgc	gttctaaact	gccgttcttg	agtgaagttag	aggcaggcgg	660
aattcctagt	gtagcggtga	aatgcgtaga	tattaggagg	aacaccagtg	gcgaaggcgg	720
cctgctgggc	ttaactgac	gctgaggctc	gaaagcgtgg	ggagcaaaca	ggattagata	780
ccctggtagt	ccacgcccgt	aacgatgatt	actaggtgtg	gggggactga	ccccttcgt	840
gcccgcgtta	acacaataag	taatccacct	ggggagtagc	gcccgaaggt	tgaaaactcaa	900
aggaattgac	ggggggccgc	acaaggcgtg	gagtagtgtg	ttaattcga	agcaacgcga	960
agaacacctac	cagggtttga	catcggtgc	atagcctaga	gatagggtgaa	gcccttcggg	1020
gcatccagac	agggtggtgca	tggttgtcgt	cagctcgtgt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcccgaacccct	tattattagt	tgctacgcaa	gagcactcta	atgagactgc	1140
cgttgacaaa	acggaggaag	gtggggatga	cgtcaaatca	tcatgccccct	tatgacctgg	1200
gctacacacg	tactacaatg	gcactaaaac	agagggcggc	gacaccgcga	ggtgaaggcga	1260
atccccaaaa	agtgtctcag	ttcagattgc	aggctgcaac	cggcctgcatt	gaagtccgaa	1320
ttgctagtaa	tcgcggatca	gcatgcccgg	gtgaatacgt	tcccgccct	tgtacacacc	1380
gcccgtcaca	ccatgggagt	cgttaacacc	cgaagccagt	agcctaaccg	caaggggggc	1440
gctgtcgaag	gtgggattga	tgactgggt	gaagtcgtaa	caaggtagcc	gtateggaaag	1500
gtgcggctgg	atcacctctt	tt				1522

&lt;210&gt; SEQ ID NO 132

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 132

tacgagagtt	tgatcctggc	tcaggatgaa	cgtggcggc	gtgcctaaca	catgcgaatc	60
gagcgaagcg	ctgttttcag	aatcttcgga	ggaagaggac	agtgaactgag	cgccggacgg	120
gtgagtaacg	cgtggcaac	ctgcctcata	cagggggata	acagttagaa	atgactgcta	180
ataccgcata	agcgcacagg	accgcattgt	gtagtgtgaa	aaactccggt	ggtatgagat	240
ggacccgcgt	ctgatttagt	agttggtggg	gtaaaggcct	accaagccga	cgatcagtag	300
ccgaccttag	agggtgaccg	gccacattgg	gactgagaca	cggcccaaacc	tcctacggg	360
ggcagcagtg	ggaaatattg	cacaatgggg	gaaaccctga	tgcagcgtac	ccgcgtgaag	420
gaagaagtat	ttcggtatgt	aaacttctat	cagcagggaa	gaaaatgacg	gtacctgagt	480
aagaagcacc	ggctaaatac	gtgccagcag	ccgcggtaat	acgtatggt	caagcgttat	540
ccggattac	tgggtgtaaa	gggagcgtag	acggataggc	aagtctggag	tgaaaaccca	600
gggctcaacc	ctgggactgc	tttggaaact	gcagatctgg	agtgcggag	aggtaaggcgg	660
aattcctagt	gtagcggtga	aatgcgtaga	tattaggagg	aacaccagtg	gcgaaggcgg	720
cttactggac	ggtactgac	gttgaggctc	gaaagcgtgg	ggagcaaaca	ggattagata	780
ccctggtagt	ccacgcccgt	aacgatgact	actaggtgtc	ggtgtgcata	gcacatcggt	840
gcccgcgtta	acgcaataag	tagtccacct	ggggagtagc	ttcgcaagaa	tgaaaactcaa	900
aggaattgac	ggggacccgc	acaaggcgtg	gagcatgtgg	ttaattcga	agcaacgcga	960
agaacacctac	ctggcttga	catccggatg	acggcgcagt	aatgtcgccg	tcccttcggg	1020

## US 10,555,980 B2

339

340

-continued

---

ggtccgaga cagggttgtc atggttgtcg tcagctcgta tcgtgagatg ttgggttaag	1080
tccgcacacg agcgcacacc ttatcttcag tagccagcat ataagggtgg cactctggag	1140
agactgccag ggagaacctg gaggaaggtg gggatgacgt caaatcatca tgccccttat	1200
ggccaggcgt acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgg	1260
agcgaatccc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctacatgaag	1320
ctggaatcgc tagtaatcgc ggtcagcat gcccgggtga atacgttccc gggtcttgta	1380
cacacccccc gtcacaccat gggagtcagt aacgccccaa gccagtgacc caaccttaga	1440
ggagggagct gtcgaaggcg ggacggataa ctggggtgaa gtcgtaacaa ggtagccgta	1500
tcgaaaggta cggctggatc acctcccttt	1529

&lt;210&gt; SEQ\_ID NO 133

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 133

tacgagagtt tgatcctggc tcaggatgaa cgctggcgcc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgttttcag aatcttcgga ggaagaggac agtgactgag cggcggacgg	120
gtgagtaacg cgtggcaac ctgcctcata cagggggata acagtttagaa atgactgcta	180
ataccgcata agcgcacagg accgcattgtgt gtatgtgaa aaactccgtt ggtatgagat	240
ggaccgcgt ctgatttagt agttgggtgg gtaaaggccct accaagccga cgatcgttag	300
ccgaccttagt agggtgaccg gccacattgg gactgagaca cggcccaaact tcctacggg	360
ggcagcgtg gggatattt cacaatgggg gaaaccctgat tgcaagcgac ccgcgtgaag	420
gaagaagtat ttccgtatgt aaacttctat cagcaggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgcagcagcc cgcggtaat acgtatggtg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctggactgc ttggaaact gcagatctgg agtgccggag aggtaaaggcg	660
aattcctagt gtagcggtaa aatgcgtaga tattaggagg aacaccagtgc gcaaggccgg	720
cttactggac ggtgactgac gttggggctc gaaagcgtgg ggagcaaaaca ggattagata	780
ccctggtagt ccacgcccgtt aacgatgact actagggtgc ggtgtgaaaa gcacatcggt	840
gcccggccaa acgcaataag tagtcccacct ggggagtagc ttgcgaagaa tgaaaactcaa	900
aggaatttgc ggggacccgc acaagcggtg gacatgtgg tttaattcga agcaacgcga	960
agaacccattac ctgggttttgc catccggatc acggggcgtt aatgtcggcc tcccttcggg	1020
gcatccgaga cagggttgtc atggttgtcg tcagctcgta tcgtgagatg ttgggttaag	1080
tccgcacacg agcgcacacc ttatcttcag tagccagcat ataagggtgg cactctggag	1140
agactgccag ggagaacctg gaggaaggtg gggatgacgt caaatcatca tgccccttat	1200
ggccaggcgt acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgaa	1260
agcgaatccc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctacatgaag	1320
ctggaatcgc tagtaatcgc ggtcagcat gcccgggtga atacgttccc gggtcttgta	1380
cacacccccc gtcacaccat gggagtcagt aacgccccaa gccagtgacc caaccttaga	1440
ggagggagct gtcgaaggcg ggacggataa ctggggtgaa gtcgtaacaa ggtagccgta	1500
tcgaaaggta cggctggatc acctcccttt	1529

-continued

<210> SEQ\_ID NO 134  
<211> LENGTH: 1529  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 134

tacgagagtt tgatcctggc tcaggatcaa cgctggcgcc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgtttttag aatcttcgaa ggaagaggac agtactgag cggcgacgg	120
gtgagtaacg cgtggcaac ctgcctcata cagggggata acagtttagaa atgactgcta	180
ataccgcata agcgcacagg accgcattgt gtagtgtgaa aaactccgtt ggtatgagat	240
ggaccggcgt ctgatttagt agttgggtgg gtaaaggcc accaagccga cgatcagtag	300
ccgacctgag agggtgaccg gccacattgg gactgagaca cggcccaaac tcctacggga	360
ggcagcagtg gggaaatattt cacaatgggg gaaaccctga tgcagcgcg ccgcgtgaag	420
gaagaagtat ttccgttatgt aaacttctat cagcaggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgcacagcg ccgcggtaat acgtatggtg caagegttat	540
cggattttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc ttggaaact gcagatctgg agtgcggag aggtaaaggcg	660
aattccatgt gtageggtaa aatgcgtaga tattaggagg aacaccagtgcg aaggccgg	720
cttactggac ggtactgac gttgaggctc gaaagcgtgg ggagcaaca ggattagata	780
ccctggtagt ccacgccgt aacgatgact actagggtgc ggtgtgcaaa gcacatcggt	840
gcccgcacaa acgcaataag tagtccaccc gggaggtacg ttgcacagaa tgaaactcaa	900
aggaatttgc ggggacccgc acaagcgtg gacatgtgg tttaattcga agcaacgcga	960
agaacccattc ctggcttgc catccggat acggcgagt aatgtcgccg tccctcggg	1020
gcatccgaga caggtggcgtc atgggtgtcg tcagctcgat tcgtgagatg ttgggttaag	1080
tcccgcaacc acgcacacc ttatcttcag tagccagcat ataagggtgg cactctggag	1140
agactgccag ggagaacctg gaggaagggtg gggatgacgt caaatcatca tgcccccttat	1200
ggccaggcgt acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacccgt	1260
agcgaatccc aaaaataacg tctcgttgc gattgtatgc tgcaactcga ctacatgaag	1320
ctggaaatcgc tagataatcgc ggatcagcat gcccgggtga atacgttccc gggcttgc	1380
cacacccgccc gtcacaccat gggagtcgt aacgccccaa gccagtgacc caacccat	1440
ggaggggagct gtcgaaggcg ggacggataa ctgggggtgaa gtcgtaccaa ggtagccgt	1500
tcggaaagggtc cggctggatc acctccctt	1529

<210> SEQ\_ID NO 135  
<211> LENGTH: 1529  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 135

tacgagagtt tgatcctggc tcaggatcaa cgctggcgcc gtgcctaaca catgcaagtc	60
gagcgaagcg ctgtttttag aatcttcgaa ggaagaggac agtactgag cggcgacgg	120
gtgagtaacg cgtggcaac ctgcctcata cagggggata acagtttagaa atgactgcta	180

-continued

---

ataccgcata agcgcacagg accgcattgt gtatgtgaa aaactccgggt ggtatgagat	240
ggaccccgct ctgatttaggt agttgggtggg gtaaaggccat accaagccga cgatcagtag	300
ccgacacctgag agggtgaccg gccacattgg gactgagaca cggcccaaacc tcctacggga	360
ggcagcgtg ggaaatattt cacaatgggg gaaaccctgt tgcaegcagc ccgcgtgaag	420
gaagaagtat ttccgttatgt aaacttctat cagcaggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgcacagcag ccgcgttaat acgtatgggt caagegttat	540
ccggattttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc ttggaaact gcagatctgg agtgcggag aggtaagcgg	660
aattcctagt gtagcgggtga aatgcgtaga tattaggagg aacaccagtgc gcaaggcgg	720
cttactggac ggtactgac gttgggttc gaaagcgtgg ggagcaaaaca ggattagata	780
ccctggtagt ccacgcccgt aacgatgact actagggttc ggtgtgcaaa gcacatcggt	840
gccgcagcaa acgcaataag tagtccaccc ggggagtagc ttgcgaagaa tgaaactcaa	900
agaatttgcg ggggaccgc acaagcgtgc gagcatgtgg tttaattcga agcaacgcga	960
agaacccattac ctgggttttgc catccggatg acggggcgt aatgtcgccg tcccttcggg	1020
gcattccgaga cagggtggtc atgggtgtcg tcaagctcg tcaatggatg ttgggttaag	1080
tcccccaacg agcgcacacc ttatcttcag tagccagcat ataaggtggg cactctggag	1140
agactgcccgg ggagaacactg gaggaaagggt gggatgacgt caaatcatca tgcccttat	1200
ggccaggcgt acacacgtgc tacaatggcg taaacaaagg gaagcggaggg ggtgaccgtga	1260
agcgaatccc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctacatgaag	1320
cttggatcgc tagtaatcgc ggatcagcat gcccgggtga atacgttccc gggtcttgc	1380
cacacccccc gtcacacccat gggagtcgt aacgccccgaa gccagtggacc caacccatgt	1440
ggagggagct gtcgaaggcg ggacggataa ctgggtgaa gtcgtaccaa ggtagccgt	1500
tcggaaagggtc cggctggatc acctccctt	1529

&lt;210&gt; SEQ\_ID NO 136

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 136

tacgagatgt tgatcctggc tcaggatgaa cgctggcgcc gtgcctaaca catgcgttc	60
gagcgttttcgt aatcttcggg ggaagaggac agtgcgttc gggcggtcc	120
gttgcgttttcgt aatcttcggg ggaagaggac agtgcgttc gggcggtcc	180
ataccgcata agcgcacagg accgcattgt gtatgtgaa aaactccgggt ggtatgagat	240
ggaccccgct ctgatttaggt agttgggtggg gtaaaggccat accaagccga cgatcagtag	300
ccgacacctgag agggtgaccg gccacattgg gactgagaca cggcccaaacc tcctacggga	360
ggcagcgtg ggaaatattt cacaatgggg gaaaccctgt tgcaegcagc ccgcgtgaag	420
gaagaagtat ttccgttatgt aaacttctat cagcaggaa gaagatgacg gtacctgagt	480
aagaagcacc ggctaaatac gtgcacagcag ccgcgttaat acgtatgggt caagegttat	540
ccggattttac tgggtgtaaa gggagcgtag acggataggc aagtctggag tgaaaaccca	600
gggctcaacc ctgggactgc ttggaaact gcagatctgg agtgcggag aggtaagcgg	660
aattcctagt gtagcgggtga aatgcgtaga tattaggagg aacaccagtgc gcaaggcgg	720

-continued

cttactggac ggtgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccctggtagt ccacgcgtta aacgatgact actaggtgtc ggtgtgcaaa gcacatcggt	840
gccgcagcaa acgcaataag tagtccacct ggggagtagc ttgcgaagaa tgaaactcaa	900
aggaattgac ggggaccgc acaagcggtg gagcatgtgg ttaattcga agcaacgcga	960
agaaccttac ctggtcttga catccggatg acggggcgagt aatgtcgccg tcccttcgg	1020
gcatccgaga cagggttgtc atgggtgtcg tcagctcggt cgtagatg ttgggttaag	1080
tcccgcAACG agcgcaaccc ttatctttag tagccagcat ataaggtggg cactctggag	1140
agactgcccgg gggaaacctg gaggaagggt gggatgacgt caaatcatca tgcccttat	1200
ggccagggct acacacgtgc tacaatggcg taaacaaagg gaagcgagag ggtgacctgg	1260
agcgaatccc aaaaataacg tctcagttcg gattgttagtc tgcaactcga ctacatgaag	1320
ctggaaatcgc tagtaatcgc ggatcagcat gcccgggtga atacgttccc gggtcttcta	1380
cacacccggcc gtcacaccat gggagtcagt aacgccccgaa gccagtgacc caaccttaga	1440
ggagggagct gtcgaaggcg ggacggataa ctgggtgaa gtgcataacaa ggtagccgt	1500
tcgaaagggtc cggctggatc acctcccttt	1529

&lt;210&gt; SEQ ID NO 137

&lt;211&gt; LENGTH: 1527

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 137

atgagagttt gatcctagct caggatgaac gctggccggcg tgcctaacac atgcaagtcg	60
aacgaagcaa tttaacggaa gtttcggat ggaagttgaa ttgactgagt ggccggacggg	120
ttagttaacgc gtgggttaacc tgccttgcgtc tgggggacaa cagttagaaa tgactgctaa	180
taccgcataa ggcacacgtt tcgcatacgata cagttgtgaaa aactccgggt gtacaagatg	240
gaccccggtc tgattagcta gttggtaagg taacggctta ccaaggcgac gatcagtgc	300
cgacctgaga ggggtgacccgg ccacattggg actgagacac ggcccaaact cctacgggag	360
gcagcgtgg ggaattttgc acaatggcg aaaggctgtat gcagcgtacgc cgcgtgatg	420
aagaagtatt tcggatgtt aagctctatc agcaggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaaatcgt tgccacgcgc cgggtataa cgtagggggc aagcgttatc	540
cggattttact ggggttaaaag ggagcgtaga cggtaaaagca agtctgaagt gaaagccgc	600
ggctcaactg cgggactgtt ttggaaactg ttaactggaa gtgtcgagaa ggtaaatggaa	660
atcccttagt tagcgggtgaa atcgtagat attaggagga acaccatgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcgttaa acgatgataa cttaggtgtt gggagcaaaat ctcttcgggt	840
ccgtcgcaaa cgcgtaaatg attccacccgt gggagtagt tcgcaagaat gaaactcaaa	900
ggaaatggacg gggacccgca caagcgggtgg agcatgtgg ttaattggaa gcaacgcga	960
gaaccttacc aggtcttgcgatcgtccga cggggggatg acgtccccctt cccttcgggg	1020
cgaggaaagac aggtgggtgca tgggtgtcgat cagctcggtt cgtggatgtt tgggttaagt	1080
cccgcaacga ggcacaccccttattcttaatg agccagcggt tcggccggaa actcttggaa	1140
gactgccagg gataacctgg aggaagggtgg ggtgcgtc aaatcatcat gccccttatg	1200

-continued

---

atctgggcta cacacgtgct acaatggcgt	aaacaaagag aagcaagacc gcgagggtgga	1260
gcaaatacgt ctcagttcg actgcaggct	gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg	tcgcggtgaa tacgttcccg ggtcttgac	1380
acaccggccg tcacaccatg ggagttagta	acgcccgaag tcagtgaccc aaccgcaagg	1440
agggagctgc cgaaggcgaaa accgataact	gggggtgaagt cgtaacaagg tagccgtatc	1500
ggaagggtcg	gctggatcac ctccctt	1527

&lt;210&gt; SEQ ID NO 138

&lt;211&gt; LENGTH: 1527

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 138

atgagagttt gatcctggct caggatgaac	gctggcggcg tgcctaacac atgcaagtcg	60
aacgaagcga tttaacggaa attttcggat	ggaaggttcaa ttgacttgagt ggcggacggg	120
ttagtaacgc gtgggttaacc tgccttgcac	tgggggacaa cagtttagaaa tgactgctaa	180
taccgcataa ggcacacatg tcgcatgata	cagtgtgaaa aactccggtg gtacaagatg	240
gacccgcgtc tgattagcta gttggtaagg	taacggctta ccaaggcgac gatcgttagc	300
cgacctgaga gggtgaccgg ccacattggg	actgagacac ggcccaact cctacggag	360
gcagcgtgg ggaattattgc acaatggcg	aaaggctgat gcagcgacgc cgcgtgagtg	420
aagaagtatt tcggatgtt aagctctatc	agcagggaaag aaaatgacgg tacctgacta	480
agaagccccg gctaactacg tgccagcagc	cgcgttaata cgtagggggc aagcgttatc	540
cggattttact ggggttaaaag ggagcgtaga	cggtaaaagc agtctgaatg gaaagccgc	600
ggctcaactg cgggactgct ttggaaactg	tttaactgaa gtgtcgagaa ggttaatgttga	660
attccttagt tagcgttgaa atgcgttagat	attaggagaa acaccagtgg cgaaggcgac	720
ttactggacg ataaactgacg ttgaggctcg	aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcgttaa acgtgaata	ctagggttgtt gggagcaaa ctcttcgggt	840
cgcgtcaaaa cgcgttaagt attccacctg	gggagtagtgcg tgcgttgcgaaat gaaactcaaa	900
ggaatttgcg gggacccgca caagcgggtgg	agcatgtgg ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga	cgggggagta acgtccccctt cccttcgggg	1020
cggagaagac aggtgttgca tgggtgtcg	cagctcggt cgtgagatgt tgggttaatg	1080
cccgcaacga ggcacccct tattctaagt	agccagcggt tcggccggaa actcttggga	1140
gactgccagg gataacctgg aggaagggtgg	ggatgacgtc aaatcatcat gccccttatg	1200
atctgggcta cacacgtgct acaatggcgt	aaacaaagag aagcaagacc gcgagggtgga	1260
gcaaatacgt ctcagttcg actgcaggct	gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgcg aatcagaatg	tcgcggtgaa tacgttcccg ggtcttgac	1380
acaccggccg tcacaccatg ggagttagta	acgcccgaag tcagtgaccc aaccgcaagg	1440
agggagctgc cgaaggcgaaa accgataact	gggggtgaagt cgtaacaagg tagccgtatc	1500
ggaagggtcg	gctggatcac ctccctt	1527

&lt;210&gt; SEQ ID NO 139

&lt;211&gt; LENGTH: 1527

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

## US 10,555,980 B2

**349****350**

-continued

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 139

atgagagttt gatcctggct caggatgaac	gctggcgccg	tgcctaaacac	atgcaagt	60		
aacgaagcga tttaacggaa	gttttcggat	ggaagttgaa	ttgactgagt	ggcggacgg	120	
ttagtaacgc gtggtaacc	tgccttgtac	tgggggacaa	cagttagaaa	tgactgctaa	180	
taccgcataa gcgcacagta	tcgcatgata	cagtgtgaaa	aactccggtg	gtacaagatg	240	
gacccgcgtc tgattagcta	gttggtaagg	taacggctta	ccaaggcgac	gatcagt	300	
cgacctgaga gggtgaccgg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360	
gcagcagtgg ggaatattgc	acaatggcg	aaagcctgat	gcagcgacgc	cgcgtgagt	420	
aagaagtatt tcggtatgta	aagctctatc	agcagggaa	aaaatgacgg	tacctgacta	480	
agaagccccg gctaactacg	tgccagcagc	cgcgttaata	cgtagggggc	aagcgttatc	540	
cggatttact gggtgtaaag	ggagcgtaga	cggtaaaagca	agtctgaagt	gaaagcccgc	600	
ggctcaactg ogggactgct	ttggaaactg	tttaactgga	gtgtcgagaa	ggtaagtgg	660	
attcctagtg tagcggtgaa	atgcgttagat	attaggagga	acaccagtgg	cgaaggcgac	720	
ttactggacg ataaactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780	
cctggtagtc cacgcgtaa	acgatgaata	ctagtggtt	gggagcaaag	ctcttcggtg	840	
cgcgtcgc当地 cgcgtc当地	atccacctg	gggagtagtgc当地	tcgcaagaat	gaaactcaaa	900	
ggaattgacg gggaccgc当地	caagcgtgg	agcatgtgg	ttaattcgaa	gcaacgc当地	960	
gaaccttacc aggtcttgac	atcgatccg	cgggggagta	acgtcccctt	cccttcgggg	1020	
cggagaagac aggtggtgca	tggttgtcg	cagctcggt	cgtgagatgt	tgggttaagt	1080	
cccgcaacga ggc当地	acccctt	tattctaagt	agccagcggt	tcggccggga	actcttggga	1140
gactgccagg gataacctgg	aggaagggtgg	ggatgacg	aatcatcat	gccc当地	1200	
atctggc当地	caca	actgtgtc	acaatggcg	aaacaaagag	aagcaagacc	1260
gcaaatctca aaaataacgt	ctcagttcg	actgcaggct	gcaactcgcc	tgcaegaagc	1320	
tggaaatcgct agtaatcg	cg	aatcagaatg	tcgc当地	tacgttcccg	ggtcttgc当地	1380
acaccgc当地	tcacaccatg	ggagtc当地	acgc当地	tc当地	aaccgc当地	1440
agggagctgc cgaaggc当地	accgataact	ggggtgaagt	cgt	taacaagg	tagcgtatc	1500
ggaagggtgc当地	gctggatcac	ctc当地				1527

&lt;210&gt; SEQ ID NO 140

&lt;211&gt; LENGTH: 1527

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 140

atgagagttt gatcctggct caggatgaac	gctggcgccg	tgcctaaacac	atgcaagt	60	
aacgaagcga tttaacggaa	gttttcggat	ggaagttgaa	ttgactgagt	ggcggacgg	120
ttagtaacgc gtggtaacc	tgccttgtac	tgggggacaa	cagttagaaa	tgactgctaa	180
taccgcataa gcgcacagta	tcgcatgata	cagtgtgaaa	aactccggtg	gtacaagatg	240
gacccgcgtc tgattagcta	gttggtaagg	taacggctta	ccaaggcgac	gatcagt	300
cgacctgaga gggtgaccgg	ccacattggg	actgagacac	ggcccaaact	cctacgggag	360

-continued

---

gcagcagtgg ggaatattgc acaatggcg aaaggctgat gcagcgcacgc cgctgtagtg	420
aagaagtatt tcggtatgtt aagctctatc agcagggaaag aaaatgacgg tacctgacta	480
agaageccccg gctaactacg tgccagcagc cgctgtataa cgtagggggc aagcgttatc	540
cgatttact gggtgtaaag ggagcgtaga cggtaaagca agtctgaagt gaaagccgc	600
ggctcaactg cgggactgct ttggaaactg ttaactgga gtgtcgaga ggttaagtgg	660
attccttagt tagcggtgaa atcgtagat attaggagga acaccagtgg cgaaggcgac	720
ttactggacg ataactgacg ttgaggctcg aaagcgtggg gagcaaacag gattagatac	780
cctggtagtc cacgcccgtaa acgtgaata cttagtggg gggagcaaa ctcttcgtg	840
ccgtcgcaaa cgcgtaaatg attccacctg gggagtagt tcgcaagaat gaaactcaaa	900
gaaattgacg gggaccgcga caagcggtgg agcatgtggg ttaattcgaa gcaacgcgaa	960
gaaccttacc aggtcttgac atcgatccga cggggagta acgtccccctt cccttcggg	1020
cggagaagac aggtggtgca tggttgttgt cagctcggt cgtgagatgt tgggttaagt	1080
cccgcaacgca ggcgcaacccttattcttaatg agccagcggt tcggccggg actcttcggg	1140
gactgccagg gataacctgg aggaagggtgg ggatgacgtc aaatcatcat gccccttatg	1200
atctgggcta cacacgtgct acaatggcgt aaacaaagag aagcaagacc gcgaggtgg	1260
gcaaatactca aaaataacgt ctcaagttcg actgcaggct gcaactcgcc tgcacgaagc	1320
tggaatcgct agtaatcgct aatcagaatg tcgcgggtgaa tacgttcccg ggtcttgatc	1380
acaccggcccg tcacaccatg ggagtctgat acgcccgaag tcagtgaccc aaccgcaagg	1440
agggagctgc cgaaggcgaaa accgataact ggggtgaagt cgtacaagg tagccgtatc	1500
ggaagggtgcg gctggatcac ctccctt	1527

&lt;210&gt; SEQ ID NO 141

&lt;211&gt; LENGTH: 1531

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 141

atcagagagt ttgatcctgg ctcaggatga acgctggcg cgtgttaac acatgcaagt	60
cgagcgaagc acttaagtgg atctcttcgg attgaaactt atttgcgtga gcggccggacg	120
ggtgagtaac gcgtggtaa cctgcctat acagggggat aacagttaga aatggctgt	180
aataccgcat aagcgcacag gaccgcattgg tctgggtgtgaaaactccgg tggatgaga	240
tggaccgcgc tctgatttgc tagttggagg ggtaacggcc caccaaggcg acgatcgat	300
geccgcctga gagggtgaac ggccacatttggactgacacgc acggcccaga ctcctacgg	360
aggcagcagt gggaaatatttgcacaatgggg gaaacccttgcacgcacgc gcccgtgaa	420
ggaagaagta tctcggtatg taaacttcta tcagcaggaa agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccacgc gcccggtaa tacgtagggg gcaagcgat	540
tccggattta ctgggtgtaa agggagcgta gacggaaagag caagtctgtatgtaaaggct	600
ggggcttaac cccaggactg cattggaaac tggatgttctatc gagtgcggg gaggtaagcg	660
gaattccatg ttagcggtg aaatgcgtatc atattaggag gaacaccatg ggcgaaggcg	720
gtttactgaa cggtaactgatc cgttggggat cggaaagcggt gggagcaaa aggatttagat	780
accctggtag tccacgcgtt aaacgatgaa tacttaggtt cgggtggcaa agccattcg	840
tgcgcgcagca aacgcataaa gtattccacc tggggagttac gttcgcaaga atgaaactca	900

## US 10,555,980 B2

353

354

-continued

aaggaattga	cggggacccg	cacaagcggt	ggagcatgtg	gtttaattcg	aagcaacgcg	960
aagaacctta	ccaagtcttg	acatccctct	gaccggcccg	taacggggcc	ttcccttcgg	1020
ggcagaggag	acaggtggtg	catggttgtc	gtcagctcg	gtcgtgagat	gttgggttaa	1080
gtcccgcAAC	gagcgcAAC	cctatcctta	gtagccagca	ggtgaagctg	ggcactctAG	1140
ggagactgcc	ggggataacc	cggaggaagg	cggggacgac	gtcaaattcat	catgcccTT	1200
atgatttggg	ctacacacgt	gttacaatgg	cgtaaaacaaa	gggaagcgag	acagcgatgt	1260
ttagccaatc	ccaaaaataa	cgtccccagtt	cggaactgcag	tctgcaactc	gactgcacga	1320
agcttggaaatc	gttagtataatc	gcgaatcaga	atgtcgccgt	gaatacgttc	ccgggtcttg	1380
tacacaccgc	ccgtcacacc	atgggagtca	gtaacgcccgg	aagtcaagtga	cccaaccTTT	1440
taggaggagg	ctgcccgaagg	cggggaccat	aactggggtg	aagtctgtaac	aaggtagccg	1500
tatcggaaagg	tgcggctgga	tcacccctt	t			1531

&lt;210&gt; SEQ ID NO 142

&lt;211&gt; LENGTH: 1531

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 142

atcagagagt	ttgatcctgg	ctcaggatga	acgctggcgg	cgtgcttaac	acatgcaagt	60
cgagcgaagg	acttaagtgg	atctcttcgg	attgaaactt	atttgactga	gcggggacg	120
ggtgagtaac	gcgtggtaa	cctgcctcat	acagggggat	aacagttaga	aatggctgct	180
aataccgcat	aagcgcacag	gaccgcattgg	tctggtgtga	aaaactccgg	ttgttatgaga	240
tggaccgcgc	tctgatttgc	tagttggagg	ggtaacggcc	caccaaggcg	acgatca	300
gcgggcctga	gagggtgaac	ggccacatttgc	ggactgagac	acggccca	ctcctacggg	360
aggcagcagt	ggggatatt	gcacaatggg	ggaaaccctg	atgcagcgcac	gccgcgtgaa	420
ggaagaagta	tctcggtatg	taaacttcta	tcagcaggaa	agaaaaatgac	ggtaacctgac	480
taagaagccc	cggtactacta	cgtgccagca	gcccggtaa	tacgttaggg	gcaagcgta	540
tccggattta	ctgggtgtaa	agggagcgta	gacggaagag	caagtctgtat	gtgaaaggct	600
ggggcttaac	cccaggactg	cattggaaac	tgtttttcta	gagtgcgg	gaggtaagt	660
gaattcctag	tgtagcggt	aaatgcgtat	atattaggag	gaacaccagt	ggcgaaggcg	720
gtttacttgg	cgttacttga	cgttggaggct	cgaaagcggt	gggagcaaac	aggattagat	780
accctggtag	tccacccgt	aaacgtgaa	tacttaggtt	cggttggcaa	agccattcg	840
tgcgcgcgca	aacgcataaa	gtattccacc	tggggagttac	gttcgcaga	atgaaactca	900
aaggaattga	cggggacccg	cacaagcggt	ggagcatgtg	gtttaattcg	aagcaacgcg	960
aagaacctta	ccaagtcttg	acatccctct	gaccggcccg	taacggggcc	ttcccttcgg	1020
ggcagaggag	acaggtggtg	catggttgtc	gtcagctcg	gtcgtgagat	gttgggttaa	1080
gtcccgcAAC	gagcgcAAC	cctatcctta	gtagccagca	ggtgaagctg	ggcactctAG	1140
ggagactgcc	ggggataacc	cggaggaagg	cggggacgac	gtcaaattcat	catgcccTT	1200
atgatttggg	ctacacacgt	gttacaatgg	cgtaaaacaaa	gggaagcggg	acagcgatgt	1260
ttagccaatc	ccaaaaataa	cgtccccagtt	cggaactgcag	tctgcaactc	gactgcacga	1320
agcttggaaatc	gttagtataatc	gcgaatcaga	atgtcgccgt	gaatacgttc	ccgggtcttg	1380

## US 10,555,980 B2

**355****356**

-continued

---

tacacacccgc ccgtcacacc atgggagtca gtaacgcccg aagtcatgta cccaacctta	1440
caggagggag ctgccaagg cgggaccat aactggggtg aagtcaac aaggttagccg	1500
tatcggaaagg tgccgttga tcacccctt t	1531

<210> SEQ ID NO 143  
<211> LENGTH: 1531  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 143

atcagagagt ttgatcctgg ctcaggatga acgctggcg cgtgcttaac acatgcaagt	60
cgagcgaagg acttaagtgg atctcttcgg attgaagctt atttgactga gcggcgacg	120
ggtgagtaac gcgtggtaa cctgcctcat acagggggat aacagttaga aatggctgt	180
aataccgcatt aagcgcacag gaccgcattt gctgggtgtt aaaaactccgg tggtatgaga	240
tggaccccgcg tctgatttgc tagttggagg ggttaacggcc caccaaggcg acgatca	300
gccggcctga gagggtaac ggccacattt ggactgagac acggcccaaga ctcctacggg	360
aggcagcagt gggaatattt gcacaatggg gaaaccctt atgcagcgcac gccgcgtgaa	420
ggaagaagta tctcggtatg taaaacttcta tcagcaggaa agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gcccggtaa tacgttaggg gcaagcgta	540
tccggattta ctgggtgtt aaggagcgta gacggaaagag caagtctgtat gtgaaaggct	600
ggggcttaac cccaggactt cattggaaac ttttttcta gagtgcggg gaggtaaagcg	660
gaattccatg ttagcggtt aaatgcgtt atattaggag gaacaccagt ggcgaaggcg	720
gtttactgga cggtaacttga cgttgaggct cggaaagcgat gggagcaac aggatttagat	780
accctggtag tccacgcgtt aaacgtttaa tacttaggtt cgggtggcaa agccattcg	840
tgcgcgcgca aacgcataa gtattccacc tggggagtac gttcgcaaga atgaaactca	900
aaggaaatttga cggggaccccg cacaagcggt ggagcatgtt gtttattcg aagcaacgcg	960
aagaacctta ccaagtcttgc acatcccttgc gaccggcccg taacggggcc ttcccttcgg	1020
ggcagaggag acaggtgggtt catgggtgtc gtcagctgt gtcgttagat gttgggttaa	1080
gtcccccaac gagcgcacacc cctatcccttgc ttagccagca ggtgaagctg ggcactctag	1140
ggagactgcc gggataacc cggaggaagg cggggacgcac gtcaaatcat catgcccctt	1200
atgattttggg ttacacacgt gctacaatgg cgtaaacaaa gggaaagcgag acagcgatgt	1260
tgagcaaattt caaaaataa cgtccccatgg cggactgcag tctgcaactc gactgcacga	1320
agcttggatc gcttagatac gcaatcaga atgtcggtt gaataacgttcc cgggtcttgc	1380
tacacacccgc ccgtcacacc atgggagtca gtaacgcccg aagtcatgta cccaacctta	1440
caggagggag ctgccaagg cgggaccat aactggggtg aagtcaac aaggttagccg	1500
tatcggaaagg tgccgttga tcacccctt t	1531

<210> SEQ ID NO 144  
<211> LENGTH: 1531  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 144

atcagagagt ttgatcctgg ctcaggatga acgctggcg cgtgcttaac acatgcaagt	60
--	----

## US 10,555,980 B2

**357****358**

-continued

cgagcgaagc acttaagcgg atctcttcgg attgaaactt atttgactga gcggcggacg	120
ggtgagtaac gcgtggtaa cctgcctcat acaggggat aacagttaga aatggctgt	180
aataccgcat aagcgcacag gaccgcattg tctggtgtga aaaactccgg tggtatgaga	240
tggaccgcg tctgatttagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gagggtgaac ggccacattt ggactgagac acggcccaga ctcctacgg	360
aggcagcagt gggaaatattt gcacaatggg ggaaccctgt atgcagcgcac gccgcgtgaa	420
ggaagaagta tctcggtatg taaaacttcta tcagcaggaa agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gcccggtaa tacgttaggg gcaagcgta	540
tccggattta ctgggtgtaa agggagcgta gacggaaagag caagtctgtat gtgaaaggct	600
ggggcttaac occaggactg cattggaaac tggggcttaa gagtgcggga gaggtaagcg	660
gaattccctag ttagcggtg aaatgcgtat atattaggag gaacaccagt ggcgaaggcg	720
gtttactgga cggtaactga cgttgggggtt cggaaacgcgt gggagcaaac aggattagat	780
accctggtag tccacgcgtt aaacgatgaa tactagggtt cgggtggcaa agccattcgg	840
tgcgcgcgca aacgcataa gtattccacc tgggggttac gttcgcaaga atgaaactca	900
aaggaattga cggggacccg cacaagcggtt ggagcatgtt gtttaattcg aagcaacgcg	960
aagaacctta ccaagtcttgc acatcccttgc gaccggcccg taacggggcc ttcccttcgg	1020
ggcagaggag acagggtgtt catgggttgc gtcagctgtt gtcgtgagat gttgggttaa	1080
gtcccccaac gagcgcacacc cctatcccttgc gtagccagca ggtgaagctg ggcactctag	1140
ggagactgcc ggggataacc cggaggaagg cggggacgac gtcaaatcat catgcctt	1200
atgattttggg ctacacacgt gctacaatgg cgtaaaacaaa gggaaacgcgtt acagegtgt	1260
ttagcaaatc caaaaataa cgtccctgtt cggactgcgt tctgcaactc gactgcacga	1320
agcttggaaatc gcttagatac gcgaatcaga atgtcgccgtt gaatacgttc cgggtcttgc	1380
tacacaccgc cctgcacacc atggggatca gtaacgcggc aagtcaatgg cccaaaccttta	1440
caggaggagg ctggcgaaagg cggggccgtt aactgggggtt aagtcgttac aaggttagccg	1500
tatcgaaagg tgcggcttgc tcacccctt t	1531

&lt;210&gt; SEQ ID NO 145

&lt;211&gt; LENGTH: 1531

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 145

atcagagagt ttgatccctgg ctcaggatga acgctggcg cgtgcttaac acatgcgtt	60
cgagcgaagc acttaagtggtt atctcttcgg attgaaactt atttgactga gcggcggacg	120
ggtgagtaac gcgtggtaa cctgcctcat acaggggat aacagttaga aatggctgt	180
aataccgcat aagcgcacag gaccgcattt tctggtgtga aaaactccgg tggtatgaga	240
tggaccgcg tctgatttagc tagttggagg ggtaacggcc caccaaggcg acgatcagta	300
gccggcctga gagggtgaac ggccacattt ggactgagac acggcccaga ctcctacgg	360
aggcagcagt gggaaatattt gcacaatggg ggaaccctgtt atgcagcgcac gccgcgtgaa	420
ggaagaagta tctcggtatg taaaacttcta tcagcaggaa agaaaatgac ggtacctgac	480
taagaagccc cggctaacta cgtgccagca gcccggtaa tacgttaggg gcaagcgta	540

## US 10,555,980 B2

359

360

-continued

---

tccggattta ctgggtgtaa	agggagcgta	gacggaagag	caagtctgat	gtgaaaggct	600
ggggcctaac cccaggactg	cattggaaac	tgaaaaatac	gagtgccgga	gaggtaaagcg	660
gaatccctag tgcgttgcgt	aaatgcgttag	atattaggag	gaacatcagt	ggcgaaggcg	720
gcttaactgaa cggtaactga	cgttgaggct	cggaaagegtg	gggagcaaac	aggatttagat	780
accctggtag tccacgcccgt	aaacgatgaa	tactaggatgt	cggggtggcaa	agccattcgg	840
tgcgcgcaga aacgcaataa	gtattccacc	tggggagttac	gttcgcaga	atgaaactca	900
aaggaattga cggggacccg	cacaagcggt	ggagcatgtg	gtttaattcg	aagcaacgcg	960
aagaacctta ccaagtcttgc	acatccctct	gaccggcccg	taacggggcc	ttcccttcgg	1020
ggcagaggag acaggtggtg	catggttgtc	gtcagctcg	gtcgtgagat	gttgggttaa	1080
gtcccccaac gagcgcaccc	cctatcctta	gtagccagca	ggtgaagctg	ggcactctag	1140
ggagactgcc ggggataacc	cggaggaagg	cggggacgac	gtcaaattcat	catgcccctt	1200
atgatttggg ctacacacgt	gctacaatgg	cgtaaacaaa	gggaagcgag	acagcgatgt	1260
tgagcaatac caaaaataa	cgtccctagtt	cggactgcag	tctgcaactc	gactgcacga	1320
agcttggaaatc gcttagatac	gcgaatcaga	atgtcgccgt	gaatacgttc	ccgggtcttgc	1380
tacacaccgc cccgtcacacc	atgggagtca	gtaacgcccgg	aagtcaagtga	cccaacctta	1440
caggagggag ctgccgaagg	cgggaccgat	aactggggtg	aagtctgtaac	aaggttagccg	1500
tatcggaagg tgccgttgc	tcacccctt	t			1531

&lt;210&gt; SEQ ID NO 146

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 146

aacgagaggat tgcgttgcgc	tcaggatgaa	cgctggccgc	gtgcttaaca	catgcaagtc	60
gagcgaagca cttaagtttgc	atttttcgg	tgaagacttt	tgtgactgag	cgccggacgg	120
gtgagtaacg cgtgggttaac	ctgcctcata	cagggggata	acagttagaa	atgactgcta	180
ataccgcata agaccacgg	accgcattgtt	acagtggtaa	aaactccgg	ggtatgagat	240
ggaccgcgt ctgatttagt	atgtgggtgg	gtaacggcc	accaagccga	cgatcagtag	300
ccgacctgag agggtgaccg	gccacattgg	gactgagaca	cggcccgac	tcctacggga	360
ggcagcgtg gggaaatattg	cacaatggag	gaaactctga	tgcagcgcac	ccgcgtgaag	420
gatgaagtat ttccgttatgt	aaacttctat	cagcaggaa	gaaaatgacg	gtacctgact	480
aagaagcccc ggcttaactac	gtgccagcag	ccgcggtaat	acgttaggggg	caagcgttat	540
ccggatttac tgggtgtaaa	gggagcgtag	acggcacggc	aagccagatg	tgaaagcccc	600
gggctcaacc cccggactgc	attttggaaact	gtctgagctg	atgtgtcgag	aggcaagtgg	660
attccctagt gtacgggtga	aatgcgtaga	tattaggagg	aacaccagtg	gcgaaggcg	720
cttgcgttgc	gtgactgac	gttggggctc	gaaagcggtg	ggagcaaacaa	780
ccctggtagt ccacgcccgt	aacgatgact	gtctggatgt	gggtggcaaa	gccattcggt	840
gcccgagcta acgcaataag	cagtccacat	ggggagtgac	tgcgttgc	tgaaactcaa	900
aggaatttgc	ggggacccgc	acaagcggtg	gagcatgtgg	tttattcga	960
agaacccatc ctgatcttgc	catcccgat	accgcttcgt	aatggaaattt	tttcttcgg	1020
acatcggtga cagggtggc	atgggtgtcg	tcagctcg	tcgtgagatg	ttgggttaag	1080

-continued

tcccgcaacg	agcgcaaccc	ctatttcag	tagccagcag	gttaagctgg	gcactcttga	1140
gagactgcca	gggataacct	ggagaaagg	ggggatgacg	tcaaatacatc	atgcccccta	1200
tgaccagggc	tacacacgtg	ctacaatggc	gtaaacaag	agaagcgaac	tcgcgagggt	1260
aagcaaatact	aaaaataac	gtctcagttc	ggattgtagt	ctgcaactcg	actacatgaa	1320
gttggaatcg	ctagtaatcg	cagatcagaa	tgctgcggtg	aatacgttcc	cgggtcttgt	1380
acacaccgcc	cgtcacacca	tgggagtcag	taacgcccga	agtcagtgac	ccaaaccgtaa	1440
ggagggagct	gccgaagggt	ggaccgataa	ctggggtgaa	gtcgtaacaa	ggtagccgt	1500
tcggaaagggt	cggctggatc	acccctttt				1529

&lt;210&gt; SEQ ID NO 147

&lt;211&gt; LENGTH: 1528

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 147

aacgagagtt	tgatcctggc	tcaggatgaa	cgctggcgcc	gtgcttaaca	catgcaagtc	60
gagcgaagca	ttttggaaag	attttcggg	tgatttcctt	tgtgactgag	cggcgacgg	120
gtgagtaacg	cgtgggtaac	ctgcctcata	cagggggata	acagtttagaa	atgactgcta	180
ataccgcata	agaccacggt	accgcattgt	acagtggtaa	aaactccgg	gttatgagat	240
ggaccecggt	ctgatttagt	agttgggtgg	gtaacggcc	accaagccga	cgatcagtag	300
ccgacctgag	agggtgaccg	gccacattgg	gactgagaca	cgccccagac	tcctacggga	360
ggcagcgtg	ggaaatattt	cacaatggag	gaaactctga	tgcaagcagc	ccgcgtgaag	420
gatgaagtat	ttcggtatgt	aaacttctat	cagcaggaa	gaaaatgacg	gtacctgact	480
aagaagcccc	ggctaactac	gtgccagcag	ccgcggtaa	acgttaggggg	caagegttat	540
ccggattttac	tgggtgtaaa	ggggcgttag	acggcacggc	aagccagatg	tgaaggcccg	600
gggctcaacc	ccggactgca	tttggaaactg	ctgagctaga	gtgtcgggaga	ggcaagtgg	660
attcctagtg	tagcggtgaa	atgcgtagat	attaggagga	acaccagtgg	cgaaggccgc	720
ttgctggacg	atgactgacg	ttgaggctcg	aaagcgtggg	gagcaaacag	gattagatac	780
cctggtagtc	cacgcgtaa	acgatgactg	ctaggtgtcg	ggtgccaaag	ccattcggtg	840
ccgcagctaa	cgcaataagc	agtccacctg	gggagtagct	tgcagaagaa	gaaactcaaa	900
ggaattgacg	gggacccgca	caagcggtgg	agcatgtgg	ttaattcgaa	gcaacgcgaa	960
gaaccttacc	tgatcttgc	atcccgatga	ccgcctcgta	atggaaagctt	ttcttcggaa	1020
catcggtac	aggcggtgca	tggttgtcg	cagctcggt	cgtgagatgt	tgggttaagt	1080
cccgcaacga	gcgcacaccc	tatcttcagt	agccagcagg	ttaagctggg	cactctggag	1140
agactgcccag	ggataaacctg	gaggaagggt	gggatgacgt	caaatacatca	tgccccttat	1200
gaccagggt	acacacgtgc	tacaatggcg	taaacaaga	gaagcgaact	cgcgagggt	1260
agcaaatactc	aaaaataacg	tctcagttcg	gattgttagtc	tgcaactcga	ctacatgaag	1320
cttggaaatcgc	tagtaatcgc	agatcagaat	gctgcgggt	atacgttccc	gggtcttgta	1380
cacacccgccc	gtcacacccat	ggggatcagt	aacgcccga	gtcagtgacc	caaccgtaa	1440
gagggagctg	ccgaagggtgg	gaccgataac	ttggggtgaa	tcgttaacaa	gtagccgtat	1500
cggaagggtgc	ggctggatca	cctccctt				1528

-continued

<210> SEQ ID NO 148  
<211> LENGTH: 1529  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 148

aacgagagtt	tgatcctggc	tcaggatgaa	cgctggggc	gtgcttaaca	catgcaagtc	60
gagcgaagca	cttaagttt	attcttcgga	tgaagacttt	tgtgactgag	cggcggacgg	120
gtgagtaacg	cgtggtaac	ctgcctcata	cagggggata	acagtttagaa	atgactgcta	180
ataccgcata	agaccacggt	accgcattgt	acagtggtaa	aaactccggt	ggtatgagat	240
ggacccgcgt	ctgatttagt	agttggtggg	gtaacggcc	accaagccga	cgatcagtag	300
ccgacctgag	agggtgaccg	gccacattgg	gactgagaca	cggcccagac	tcctacggga	360
ggcagcagtg	ggaaatattt	cacaatggag	gaaactctga	tgcagcgcacg	ccgcgtgaag	420
gtatgaagtat	ttcggtatgt	aaacttctat	cagcaggaa	aaaaatgacg	gtacctgact	480
aagaagcccc	ggcttaactac	gtgccagcag	ccgcggtaat	acgttaggggg	caagcgttat	540
ccggattttac	tgggtgtaaa	gggagcgtag	acggcacggc	aagccagatg	tgaaagcccg	600
gggctcaacc	ccgggactgc	atttggaaact	gctgagctag	agtgtcggag	aggcaagtgg	660
aattcctagt	gtagcggtga	aatgcgtaga	tattaggagg	aacaccatgt	gcgaaggcgg	720
cttgctggac	gatgactgac	gttggggctc	gaaagcgtgg	ggagcaaaca	ggattagata	780
ccctggtagt	ccacgcccgt	aacgatgact	gctaggtgtc	gggtggcaaa	gccattcggt	840
gccgcagcta	acgcataaa	cagtcacact	ggggagtagc	ttcgcaagaa	tgaaactcaa	900
aggaatttgc	ggggaccgc	acaagcggtg	gagcatgtgg	tttaattcga	agcaacgcga	960
agaaccttac	ctgatcttga	catcccgatg	accgcattcg	aatggaaat	tttcttcgga	1020
acatcggtga	cagggttgtc	atgggtgtcg	tcaagctcg	tgcgtgatg	ttgggttaag	1080
tcccgcaacg	agcgcacccc	ctatcttgc	tagccagcag	gttaagctgg	gcactctgga	1140
gagactgcca	gggataaccc	ggagaaatgt	ggggatgcag	tcaaatcatc	atgcccccta	1200
tgaccagggc	tacacacgtg	ctacaatggc	gtaaacaaag	agaagcgaac	tgcgcagggt	1260
aagcaaatct	caaaaataac	gtctcagttc	ggattgttagt	tgcactcg	actacatgaa	1320
gctggaatcg	ctagtaatcg	cagatcagaa	tgctgcggtg	aatacgttcc	cgggtcttgt	1380
acacaccgcc	cgtcacacca	tggggatgcag	taacgcccga	agtcagtgcac	ccaaccgtaa	1440
ggagggagct	gccgaagggt	ggaccgataa	ctggggtgaa	gtcgtaacaa	ggtagccgt	1500
tccgaagggt	cggtggatc	acccctttt				1529

<210> SEQ ID NO 149  
<211> LENGTH: 1529  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Polynucleotide

<400> SEQUENCE: 149

aacgagagtt	tgatcctggc	tcaggatgaa	cgctggggc	gtgcttaaca	catgcaagtc	60
gagcgaagca	ctttggaaag	attcttcgga	tgatttcgtt	tgtgactgag	cggcggacgg	120
gtgagtaacg	cgtggtaac	ctgcctcata	cagggggata	acagtttagaa	atgactgcta	180
ataccgcata	agaccacggt	accgcattgt	acagtggtaa	aaactccggt	ggtatgagat	240

-continued

ggacccgcgt ctgatttagt agttgggtgg gtaacggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga	360
ggcagcagtg gggaaatattt cacaatggag gaaactctga tgcagcgacg ccgcgtgaag	420
gatgaagtat ttccgttatgt aaacttctat cagcaggaa gaaaatgacg gtacctgact	480
aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccg	600
gggcctaacc ccgggactgc atttggact gctgagctag agtgtcgagg aggcaagtgg	660
aattcctagt gtagegggtga aatgcgtaga tattaggagg aacaccatg gccaaggcgg	720
cttgctggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaca ggattagata	780
ccttggtagt ccacgcgtta aacgtgact gctaggtgtc ggggtggcaaa gccattcggt	840
gecccgacta acgcaataag cagtcacccct ggggagtagc ttgcgaagaa tgaaactcaa	900
aggaatttac gggggacccgc acaagcgttg gaggatgtgg tttaatttca agcaacgcga	960
agaaccttac ctgatcttga catcccgatg accgcttcgt aatggaaatct ttttttcgga	1020
acatcggtga cagggtgtc atgggtgtcg tcagctcggt tcgtgagatg ttgggttaag	1080
tcccgcaacg agcgcaaccc ctatttcag tagccagcag gttaagctgg gcactcttga	1140
gagactgcca gggataacct ggagaaatggt ggggatgacg tcaaatcatc atgccttta	1200
tgaccaggc tacacacgtg ctacaatggc gtaaacaatgg agaagegaac tcgaggggt	1260
aagcaaatct caaaaataac gtctcagttc ggattgttagt ctgcaactcg actacatgaa	1320
gcttggatcg ctatgtatcg cagatcagaa tgctgcgttg aatacgttcc cgggtcttgt	1380
acacaccggc cgtcacacca tggggatcg taacgcccggc agtcaatgtac ccaaccgtaa	1440
ggagggagct gcccggatgg gggccatgg ctgggggtgaa gtcgtaccaa ggttagccgt	1500
tccggatgtt cggctggatc acctcccttt	1529

&lt;210&gt; SEQ ID NO 150

&lt;211&gt; LENGTH: 1528

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 150

aacggagatgt tgatcctggc tcaggatgaa cgctggggc gtgtttaaca catgcaagtc	60
gagcgaaatcg cttaagttt atttttcgga tgaagacttt tgtgactgag cggccggacgg	120
gtgagtaacg cgtgggtaac ctgcctcata cagggggata acagtttagaa atgactgtca	180
ataccgcata agaccacggc accgcattgtt acagtggtaa aaactccgtt ggtatgagat	240
ggacccgcgt ctgatttagt agttgggtgg gtaacggcct accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccagac tcctacggga	360
ggcagcagtg gggaaatattt cacaatggag gaaactctga tgcagcgacg ccgcgtgaag	420
gatgaagtat ttccgttatgt aaacttctat cagcaggaa gaaaatgacg gtacctgact	480
aagaagcccc ggctaactac gtgccagcag ccgcggtaat acgtaggggg caagcgttat	540
ccggatttac tgggtgtaaa gggagcgtag acggcacggc aagccagatg tgaaagcccg	600
ggcctaacc ccgggactgc ttggaaactg ctgagctaga gtgtcgagg ggcgttgg	660
atccctatgt tagcggtgaa atgcgtatgtt attaggaggaa acaccatgtgg cgaaggcggc	720

## US 10,555,980 B2

367

368

-continued

---

ttgctggacg atgactgacg ttgaggctcg aaagcgtgg gагcaaасаг gattagatac	780
cctggtagtc cacggcgtaa acgatgactg ctaggtgtcg ggtggcaaag ccattcggtg	840
ccgcagctaa cgcaataaагc agtccacctg gggagtagctg tcgcaagaat gaaactcaaа	900
ggaaattgacg gggaccggca caagcggtgg agcatgtgg ttaattcgaa gcaacgcaaa	960
gaaccttacc tgatcttgac atcccgatga cegcttcgtatggaaattttgcgaa	1020
catcggtgac aggtggtgca tggttgttgt cagctcggt cgtgagatgt tgggttaagt	1080
cccgcaacga ggcgcaacccc tatcttcagt agccagcagg ttaagctggg cactctggag	1140
agactgcccgg ggataaacctg gaggaaagggt gggatgacgt caaatcatca tgccccttat	1200
gaccaggggctt acacacgtgc tacaatggcg taaacaaaga gaggcaaaact cgccgggtta	1260
agcaaatctc aaaaataaагc tctcgttgc gattgttagtgc tgcaactcga ctacatgaag	1320
cttggaaatcgc tagtaatcgc agatcagaat gctgcgggtga atacgttccc gggctttgtat	1380
cacacccgcccc gtcacaccat gggagtagt aacgccccggaa gtcagtgacc caaccgtaaаг	1440
gaggggagctg ccgaagggtgg gaccgataac tgggggtgaag tcgtaaacaag gtagccgtat	1500
cggaagggtgc ggctggatca cctccctt	1528

&lt;210&gt; SEQ\_ID NO 151

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 151

aacgagagtt tgatcctggc tcaggatgaa cgctggcgcc gtgcttaaca catgcaagtc	60
gagcgaagcg ctttggaaag attttcgga tggatttcctt tggactgag cggcgacgg	120
gtgagtaacg cgtggtaac ctgcctcata cagggggata acagttagaa atgactgcta	180
ataccgcata agaccacggt accgcattgtt acagtggtaa aaactccgggt ggtatgagat	240
ggaccggcgt ttgatttagt agttgggtgg gtaacggctt accaagccga cgatcagtag	300
ccgacctgag aggggtgaccg gccacattgg gactgagaca cggcccgac tcctacggga	360
ggcagcagtg gggaaatattt cacaatggag gaaactctga tgcagcggac cccgggtaaаг	420
gatgaagtat ttccgttatgt aaacttctat cagcaggaa gaaaatgacg gtacctgact	480
aagaagcccc ggcttaactac gtgcggcggc ccgcggtaat acgttaggggg caaccgttat	540
ccggattttac tgggtgtaaa gggagcgttag acggcacggc aagccagatg tgaaagcccg	600
gggctcaacc ccgggactgc atttggaaact gctgagctag agtgcggag aggcaagtgg	660
aatttcctagt gtacgggtga aatgcgtaga tattaggagg aacaccgtg gcgaaaggcgg	720
cttgcgtggac gatgactgac gttgaggctc gaaagcgtgg ggagcaaaaca ggattagata	780
ccctggtagt ccacgcgtta aacgtgact gctaggtgtc ggggtggcaaa gccattcggt	840
ccgcgcgtta acgcaataaагc cagtcacact gggagtagt ttcgcaagaa tgaaactcaa	900
aggaattgac gggggacccgc acaagcgggt gaggatgtt ttaattcga agcaacgcga	960
agaacccctac ctgatcttga catccccgtt actgcattgtt aatggaaatgtt tttcttcggaa	1020
acatcggtga cagggtgtgc atgggtgtcg ttagctcggt tcgtgagatg tgggttaag	1080
tcccgcaacg agcgcacccctt ctatcttcgt tagccggcgt gtttagtgg gcactctggaa	1140
gagactgcca gggataacct ggagggaaatgtt ggggtggacg tcaaatcatc atgccccctta	1200
tgaccaggggc tacacacgtg ctacaatggc gtaaacaагaagcgaac tcggcggaggt	1260

-continued

aagcaaatac caaaaataac gtctcagttc ggattgttagt ctgcaactcg actacatgaa	1320
getggaatcg ctagtaatcg cagatcagaa tgctgcggtg aatacgttcc cgggtcttgt	1380
acacaccgcc cgtcacacca tgggagtcag taacgcccga agtcagtgac ccaaccgtaa	1440
ggagggagct gccgaagggtg ggaccgataa ctgggggtgaa gtcgtaacaa ggtagccgt	1500
tcggaagggtg cggctggatc acctccctt	1529

&lt;210&gt; SEQ ID NO 152

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 152

atggagagtt tgatcctggc tcaggatgaa cgctggcgcc atgcctaata catgcaagtc	60
gaacgaagg ttgaggaaatc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgttagtaac ctgcctatgt gtccggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga ggcgcgtc agtatattaa agcgccccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtaggta acggctcacc aaggcgatga tgcgtagccg	300
gcctgagagg gttaaacggcc acattggac tgagacacgg cccaaactcc tacggagggc	360
agcagtaggg aatttcgta aatggggaa accctgaacg agcaatgcgc cgtgagtgaa	420
gaagggtcttc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atggggagtga cggtagctt ccagaaagcc acggctact acgtgcgcagc agccgcgtt	540
atacgttagt ggcaagcggtt atccggatc attggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctcaa ccattgttaag ctatggaaac tggatgtcg	660
gagtgcgaa gagggcgatg gaattccatg tgtagcggtt aatgcgttag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctggt ctgttaactga cactgaggca cgaaagcggt	780
ggggcaaat aggattagat accctagtag tccacgcgtt aaacgtatgg aactaagtgt	840
tggaggaatt cagtgcgtca gttaacgc aaatgttcc gcctggggag tatgcacgc	900
agtgtgaaac tcaaaggaaat tgaegggggc cccgcacaagg ggtggagtat gtggtttaat	960
tcaagcaac gcgaagaacc ttaccagggc ttgacatgaa aacaaatacc ctagagatag	1020
ggggataatt atggatcaca cagggtgtgc atgggtgtcg tcagctcgat tcgtgagatg	1080
ttgggttaag tcccgcaacg agcgcaaccc ttgtcgcatg ttaccagcat caagttgggg	1140
actcatgcga gactgcgggt gacaaaccgg aggaagggtgg ggatgacgtc aaatcatcat	1200
gccccttatg gcctgggcta cacacgtact acaatggcgcc cccacaaggag cagcgacaca	1260
gtgtatgtgaa gcgaatctca taaagggtcg ttcagttcg attgaagtct gcaactcgac	1320
ttcatgaagt cggaaatcgat agtaatcgca gatcagcatg ctgcgggtgaa tacgttctcg	1380
ggcccttgc acaccgcccc tcaaaccatg ggagtcagta atacccgaag ccgggtggcat	1440
aaccgttaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccctt	1537

&lt;210&gt; SEQ ID NO 153

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

-continued

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 153

atggagagtt tgatcctggc tcaggatcaa cgctggggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgttagtaac ctgcccattgt gtccggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga gcgcatgctc agtatattaa agcgccccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtgaggta acggcccacc aaggcgtgaa tgcgtagccg	300
gcctgagagg gtaaaacggcc acattggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aatttcgtc aatggggaa accctgaacg agcaatgcgcg cgtgagtgaa	420
gaaggcttc ggatcgtaaa gctctgtgt aagtgaagaa cggctcatag agggaaatgct	480
atgggagtga cggtagctt ccagaaagcc acggctaact acgtgccagc agccgcgtt	540
atacgttagt ggcaagcggtt atccggatc attggcgta aagggtgcgt aggtggcgta	600
ctaagtctgt agtaaaaggc aatggctaa ccattgtaa ctagggaaac tggatgtcg	660
gagtgccagaa gagggcgatg gaattccatg tgtagcggta aatgcgttag atatatggag	720
gaacaccagt ggccaggcgtt gtcgcctgtt ctgttaactgaa cactgaggca cgaaacgcgt	780
gggagcaaat aggatttagt acccttagtag tccacgcccgtt aaacgatgag aactaagtgt	840
tggaggaatt cagtgcgtca gttAACGCAA taagttctcc gcctggggag tatgcacgc	900
agtgtgaaac tcaaaggaaat tgacggggcc ccgcacaacgg ggtggagttat gtggtttaat	960
tgcgaagcaac gcgaaaggacc ttaccaggcc ttgacatgaa aacaaatacc ctagagatag	1020
ggggataatt atggatcaca caggtggtgc atgggtgtcg tcaagtcgtc tcgtgagatg	1080
ttgggttaag tcccccaacg agcgcacacc ttgtcgatgtt acccgatcat caagttgggg	1140
actcatgcga gactgccgtt gacaaacccgg aggaagggtgg ggatgcgtc aaatcatcat	1200
gcccccttatg gcctgggcta cacacgtact acaatggcgcc ccacaaagag cagcgacaca	1260
gtgatgtgaa gcaatctca taaaggctgt ctcagttcggtt attgaagtct gcaactcgac	1320
ttcatgaagt cggaaatcgct agtaatcgca gatcagcatg ctgcgggtaa tacgttctcg	1380
ggccttgtac acaccggcccg tcaaaccatg ggagtcgatg atacccgaag ccggtgccat	1440
aaccgttaagg agtgagccgtt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccctt	1537

&lt;210&gt; SEQ ID NO 154

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 154

atggagagtt tgatcctggc tcaggatcaa cgctggggc atgcctaata catgcaagtc	60
gaacgaagtt tcgaggaagc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgttagtaac ctgcccattgt gtccggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga gcgcatgctc agtatattaa agcgccccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtgaggta acggctacc aaggcgtgaa tgcgtagccg	300
gcctgagagg gtaaaacggcc acattggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aatttcgtc aatggggaa accctgaacg agcaatgcgcg cgtgagtgaa	420

-continued

gaaggcttc ggatcgtaaa gctctgttgc aagtgaagaa cggctcatag aggaaatgct	480
atggggagtga cggtagctt ccagaaaagcc acggctaact acgtgccagc agccgcggtt	540
atacgttaggt ggcaagcggtt atccggaaatc attggggcgta aagggtgcgtt aggtggcgta	600
ctaagtctgtt agtaaaaggc aatggctcaa ccattgttaag ctatggaaac tggatgtctg	660
gagtgccagaa gagggcgatg gaattccatg tgttagcgta aaatgcgttag atatatggag	720
gaacaccagt ggccaaggcg gtcgcctgggt ctgttaactga cactgaggca cgaaagcggt	780
gggagcaaat aggatttagat accctagtag tccacgcccgtt aaacgtatgaa aactaagtgt	840
tggggaaatt cagtgcgtca gttaacgc当地 taagtttccgccttggggag tatgcacgc当地	900
agtgtgaaac tcaaaggaaat tgacgggggc ccgcacaaggc ggtggagttat gtggtttaat	960
togaagcaac gcgaagaacc ttaccaggcc ttgacatggaa aacaaatacc ctatggatag	1020
ggggataatt atggatcaca caggtgggtgc atgggtgtcg tcagctcgatc tcgtggatgt	1080
ttggggtaag tcccccaacg agcgc当地 accccggcc ttgtcgatc ttaccaggcat caagtgggg	1140
actcatgc当地 gactgc当地 gacaaaccgg aggaagggtgg ggtatgc当地 aaatcatcat	1200
gc当地 cttatg gc当地 tgggtggctt cacacgtact acaatggccgg ccacaaagag cagcgacaca	1260
gtgtatgtgaa gcaatctca taaaaggctgtt ctcagttcgatc attgaagttt gcaactcgac	1320
ttcatgaagt cggaaatcgctt agtatacgca gatcagcatg ctgc当地 gaa tacgttctcg	1380
ggc当地 ttgtac acacccggcc tcaaaccatg ggagtc当地 atacccgaag cc当地 tggcat	1440
aaccgttaagg agtgagccgtt cgaaggtagg accgtatgactt ggggttaatg cgttacaagg	1500
tatccctacg ggaacgttggg gatggatcac ctccctt	1537

&lt;210&gt; SEQ ID NO 155

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 155

atggagagtt tgatcctggc tcaggatgaa cgctggggc atgc当地ata catgc当地	60
gaacgaagtt tcgagaaatc ttgttccaa agagacttag tggcgaaacgg gttagttaaca	120
cgttagttaac ctgc当地atgtt gtccggata actgctggaa acggtagctt aaaccggata	180
gttatacaga gcgatgttc agtataattaa agcgc当地atc aaggcgatgaa catggatgaa	240
cctgc当地ggc当地ttagtgc当地tggatgaa acggccc当地 aaggcgatgaa tgc当地tgc当地	300
gc当地tggagg gtaaacggcc acatgggac tgagacacgg cccaaactcc tacggaggc	360
agcagtaggg aatttcgcc aatggggaa accctgaaacg agcaatgc当地 cgttagtgc当地	420
gaaggcttc ggatcgtaaa gctctgttgc aagtgaagaa cggctcatag aggaaatgct	480
atggggagtga cggtagctt ccagaaaagcc acggctaact acgtgccagc agccgcggtt	540
atacgttaggt ggcaagcggtt atccggaaatc attggggcgta aagggtgcgtt aggtggcgta	600
ctaagtctgtt agtaaaaggc aatggctcaa ccattgttaag ctatggaaac tggatgtctg	660
gagtgccagaa gagggcgatg gaattccatg tgttagcgta aaatgcgttag atatatggag	720
gaacaccagt ggccaaggcg gtcgcctgggt ctgttaactga cactgaggca cgaaagcggt	780
gggagcaaat aggatttagat accctagtag tccacgcccgtt aaacgtatgaa aactaagtgt	840
tggggaaatt cagtgcgtca gttaacgc当地 taagtttccgc当地 tggggag tatgcacgc当地	900

## US 10,555,980 B2

375

376

-continued

---

agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagg ggtggagtat gtggttaat	960
togaagcaac gcgaagaacc ttaccaggcc ttgacatggaa aacaaaatacc cttagatag	1020
ggggataatt atggatcaca caggtggtgc atgggtgtcg tcagctcg tcgtgagatg	1080
ttgggttaag tcccgcacg agcgcaccc ttgtcgcatg ttaccagcat caagttggg	1140
actcatgcga gactgccgt gacaaaccgg aggaagggtgg ggatgcgtc aaatcatcat	1200
gccccttatg gcctggctca cacacgtact acaatggcga ccacaaagag cagcgacaca	1260
gtgatgtgaa gcgaatctca taaaggctgt ctcagttcg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcgggtgaa tacgttctcg	1380
ggccttgtac acaccgcccc tcaaaccatg ggagtcagta atacccgaag ccggtgccat	1440
aaccgcagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg	1500
tatccctacg ggaacgtggg gatggatcac ctccctt	1537

&lt;210&gt; SEQ\_ID NO 156

&lt;211&gt; LENGTH: 1537

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 156

atggagagtt tgatcctggc tcaggatgaa cgctggccgc atgcctaata catgcaagtc	60
gaacgaagg ttgagggatc ttgcttccaa agagacttag tggcgaacgg gtgagtaaca	120
cgttagtaac ctgcccattgt gtccggata actgctggaa acggtagcta aaaccggata	180
ggtatacaga ggcgcgtc agtatattaa agcgcacccatc aaggcgtgaa catggatgga	240
cctgcggcgc attagctagt tggtaggtt acggccacc aaggcgtgaa tgcgtagccg	300
gcctgagagg gttaaacggcc acatggggac tgagacacgg cccaaactcc tacgggaggc	360
agcagtaggg aattttcgtc aatgggggaa accctgaacg agcaatgcgg cgtgatgtgaa	420
gaaggcttc ggatcgtaaa gctctgttgt aagtgaagaa cggctcatag aggaaatgct	480
atgggagtga cggtagctt ccagaaagcc acggcttaact acgtgcgcgc agccgcggta	540
atacgttagt ggcaagcggtt atccggatc attggcgta aagggtgcgt aggtggcgta	600
ctaagtcgtt agtaaaaggc aatggctcaa ccattgttaag ctatggaaac tggatgtctg	660
gagtgccagaa gagggcgatg gaattccatg tgtagcgta aatgcgttag atatatggag	720
gaacaccagt ggcgaaggcg gtcgcctggc ctgttaactgaa cactggggca cgaaagcggt	780
gggagcaaat aggatttagat acccttagtag tccacgcgtt aaacgtgag aactaagtgt	840
tggaggaatt cagtgcgtca gttAACGCAA taagttctcc gcctggggag tatgcacgca	900
agtgtgaaac tcaaaggaat tgacgggggc ccgcacaagg ggtggagtat gtggttaat	960
togaagcaac gcgaagaacc ttaccaggcc ttgacatggaa aacaaaatacc cttagatag	1020
ggggataatt atggatcaca caggtggtgc atgggtgtcg tcagctcg tcgtgagatg	1080
ttgggttaag tcccgcacg agcgcaccc ttgtcgcatg ttaccagcat caagttggg	1140
actcatgcga gactgccgt gacaaaccgg aggaagggtgg ggatgcgtc aaatcatcat	1200
gccccttatg gcctggctca cacacgtact acaatggcgg ccacaaagag cagcgacaca	1260
gtgatgtgaa gcgaatctca taaaggctgt ctcagttcg attgaagtct gcaactcgac	1320
ttcatgaagt cggaatcgct agtaatcgca gatcagcatg ctgcgggtgaa tacgttctcg	1380
ggccttgtac acaccgcccc tcaaaccatg ggagtcagta atacccgaag ccggtgccat	1440

-continued

aaccgttaagg agtgagccgt cgaaggtagg accgatgact ggggttaagt cgtaacaagg 1500  
 tatccctacg ggaacgtggg gatggatcac ctccctt 1537

<210> SEQ ID NO 157  
 <211> LENGTH: 1529  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 157

tattgagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt 60  
 cgaacggggt gctcatgacg gaggattcgt ccaacggatt gagttaccta gtggggacg 120  
 ggtgagtaac gcgtgaggaa cctgccttgg agagggaaat aacactccga aaggagtgct 180  
 aataccgcat gatgcagttg ggtcgcatgg ctctgactgc caaagattta tcgctctgag 240  
 atggcctcgc gtctgattag cttagggcgg gggtaacggc ccaccttaggc gacgatcagt 300  
 agccggactg agaggttgac cggccacatt gggactgaga cacggcccaag actcctacgg 360  
 gaggcagcag tggggatat tgggcaatgg gcgcaagccat gaccaggcaa cgccgcgtga 420  
 aggaagaagg otttcggtt gtaaaacttct tttgtcgaaa acgaaacaaa tgacggtacc 480  
 cgacgaataa gccacggcta actacgtgcc agcagccgcg gtaatacgtt ggtggcaagc 540  
 gttatccgga tttaactgggt gtaaaggcgg ttttagggcgg attgcaagtc agatgtgaaa 600  
 actggggctt caacctccag cctgcattt aaactgttagt tcttgagtgc tggagaggca 660  
 atcggaatc cgtgtgtage ggtgaaatgc gtagatatac ggaggaacac cagtggcga 720  
 ggccggattgc tggacagtaa ctgacgctga ggcgcgaaag cgtggggagc aaacaggatt 780  
 agataccctg gtagtccacg ccgtaaacga tggatactag gtgtgggggt ctgaccctt 840  
 ccgtgcccga gttaacacaa taagtatccc acctggggag tacgatcgcg aggttgaac 900  
 tcaaaggaaat tgacgggggc ccgcacaaacg ggtggagttat tcgaagcaac 960  
 gccaagaacc ttaccaggggc ttgacatccc actaacgaaag cagagatgca ttaggtgcc 1020  
 ttccggggaaa gtggagacag gtggcgtatg gttgtcgatc gctcgatcg tgagatgtt 1080  
 ggtaagtcc cgcaacgcg ccaaccctta ttgttagttt ctacgcaaga gcactctagc 1140  
 gagactgcgc ttgacaaaac ggaggaagg gggacgcacg tcaaatcatc atgccccctt 1200  
 tgcctgggc cacacacgtt ctacaatggt ggtaacaga gggaggcaat accgcgagg 1260  
 ggagcaaatc cctaaaagcc atcccagttc ggattgcagg ctgaaaccccg cctgtatgaa 1320  
 gttggatcg ctagtaatcg cggatcagca tggccgcgggtt aatacgttcc cggcccttgt 1380  
 acacaccgcc cgtcacacca tgagatcgaa gaacacccga agtccgtacg ctaaccgaa 1440  
 ggagggcgcg gccgaagggtt ggttcgatata ttggggtgaa gtcgtaccaa ggtagccgt 1500  
 tcggaaagggtt cggctggatc acctccctt 1529

<210> SEQ ID NO 158  
 <211> LENGTH: 1530  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 158

tattgagagt ttgatcctgg ctcaggatga acgctggcgg cgtgcttaac acatgcaagt 60

## US 10,555,980 B2

379

380

-continued

---

cgaacggggt gctcatgacg gaggattcgt ccaacggatt gagttaccta gtggggacg	120
ggtgagtaac gcgtgaggaa cctgccttgg agagggaaat aacactccga aaggagtgc	180
aataccgcat gatgcagttg ggtcgcatgg ctctgactgc caaagattta tcgctcttag	240
atggcctcgc gtctgattag cttagtaggcg gggtaacggc ccaccttaggc gacgatcagt	300
agccggactg agaggttgcg cgccacatt gggactgaga cacggcccg actcctacgg	360
gaggcagcag tggggatat tgggcaatgg ggcgaagct gaccacgaa cgccgcgtga	420
aggaagaagg ctttcgggtt gtaaacttct tttgtcgaaa acgaaacaaa tgacggtacc	480
cgacgataaa gccacggcta actacgtgcc agcagccgcg gtaatacgta ggtggcaagc	540
gttatccgga ttactgggt gtaaaggcg gttaggcg attgcaagtc agatgtgaaa	600
actggggctt caacctccag cctgcattt aaactgttagt tcttgagtgc tggagaggca	660
atcggatttc cgtgtgtagc ggtgaaatgc gtagatatac ggaggaacac cagtggcgaa	720
ggcggattgc tggacagtaa ctgacgctga ggccgcggaa cgtggggagc aaacaggatt	780
agataccctg ttagtccacg ccgttaaacga tggatactag gtgtgggggg tctgaccccc	840
tccgtgcgcg agttaacaca ataagtatcc cacctgggg gtagatgcg aagggtgaaa	900
ctcaaaggaa ttgacggggg cccgcacaag cggtggagta tgggtttaa ttgcagacaa	960
cgcgaagaac cttaccaggg cttgacatcc cactaacgaa gcagagatgc attaggtgcc	1020
tttcggggaa agtggagaca ggtggatgc ggttgcgcg agctcggtc gtgagatgtt	1080
gggttaagtc ccgcaacgcg cgcaaccctt attgttagt gctacgcaag agcactctag	1140
c gagactgcc gttgacaaaa cggaggaagg tggggacgac gtcaaatcat catgcccctt	1200
atgtcctggg ccacacacgt actacaatgg tggtaacag agggaggcaa taccgcgagg	1260
tggagcaaat ccctaaagc catcccgattt cggattgcg gctgaaaccc gcctgtatga	1320
agttgaatc gctagtaatc gcggatcgc atgcccggt gaatacggtc ccggcccttgc	1380
tacacaccgc ccgtcacacc atgagagtcg ggaacacccg aagtccgtag cctaaaccgca	1440
aggagggcgc ggccgaaggt gggttcgata attgggtga agtcgtaaaca aggttagccgt	1500
atcggaaaggc gcggtggat cacctccctt	1530

&lt;210&gt; SEQ ID NO 159

&lt;211&gt; LENGTH: 1529

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Synthetic Polynucleotide

&lt;400&gt; SEQUENCE: 159

tattgagagt ttgatccctgg ctcaggatga acgctgggg cgtgcttaac acatgcaagt	60
cgaacggggt gctcatgacg gaggattcgt ccaacggatt gagttaccta gtggggacg	120
ggtgagtaac gcgtgaggaa cctgccttgg agagggaaat aacactccga aaggagtgc	180
aataccgcat aatgcagttg ggtcgcatgg ctctgactgc caaagattta tcgctcttag	240
atggcctcgc gtctgattag cttagtaggcg gggtaacggc ccaccttaggc gacgatcagt	300
agccggactg agaggttgcg cgccacatt gggactgaga cacggcccg actcctacgg	360
gaggcagcag tggggatat tgggcaatgg ggcgaagct gaccacgaa cgccgcgtga	420
aggaagaagg ctttcgggtt gtaaacttct tttgtcgagg acgaaacaaa tgacggtacc	480
tgacgataaa gccacggcta actacgtgcc agcagccgcg gtaatacgta ggtggcaagc	540
gttatccgga ttactgggt gtaaaggcg gttaggcg attgcaagtc agatgtgaaa	600

-continued

actgggggct caacccctcag cctgcattt	aaactgttagt tcttgagtgc tggagaggca	660
atcggaattc cgtgtgttagc ggtgaaatgc	gtagatacggagggacac cagtggcgaa	720
ggcggtttgc tggacagtaa ctgacgctga	ggcgcgaaag cgtggggagc aaacaggatt	780
agataccctg gtatccacg ccgtaaacga	tggatactag gtgtgggggg tctgaccct	840
cctgtccgcgtttaa acctggggag tacgatcgca	agggttggaaac	900
tcaaaggaaat tgacgggggc ccgcacaagc	ggtggagtagt gtgggttaat tcgaagcaac	960
gccaagaacc ttaccagggc ttgacatccc	actaacgaag cagagatgca ttaggtgcc	1020
ttcggggaaa gtggagacag gtgggtcgatg	gttgcgtca gtcgtgtcg tgagatgtt	1080
ggttaagtcc cgcaacgagc gcaaccctta	ttgttagttt ctacgcaaga gcactctagc	1140
gagactgccc ttgacaaaac ggaggaaggt	ggggacgacg tcaaatacatc atgcccccta	1200
tgtcctgggc cacacacgta ctacaatggt	gtttaacaga gggaggcaat accgcgaggt	1260
ggagcaaatac cctaaaagcc atcccagttc	ggattgcagg ctgaaaccccg cctgtatgaa	1320
gttggaatcg ctagtaatcg cggatcagca	tgccgcgggtg aatacgttcc cgggccttgt	1380
acacaccgccc cgtcacacca tgagatcg	gaacaccccg agtccgtac ctaaccgcaa	1440
ggagggcgcg gcccgaagggtg	ggttcgataa ttgggggtgaa gtcgtaccaa ggtagccgt	1500
tcggaaagggtg cggctggatc acctccctt		1529

What is claimed is:

1. A method to suppress an abnormal or excessive immune response comprising administering a pharmaceutical composition comprising a purified bacterial mixture consisting of 7 or more bacterial strains comprising 16S rDNA sequences of at least 97% sequence identity to 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, to the subject in an amount sufficient to suppress the abnormal or excessive immune response.

2. The method of claim 1, wherein the abnormal or excessive immune response is suppressed by inducing proliferation and/or accumulation of regulatory T cells.

3. The method of claim 1, wherein the subject is administered one or more doses of an antibiotic prior to the pharmaceutical composition.

4. The method of claim 3, wherein the antibiotic is vancomycin, kanamycin, gentamicin, colistin, metronidazole, clindamycin, fidaxomicin, or cefoperazone.

5. The method of claim 3, wherein the antibiotic is vancomycin.

6. The method of claim 1, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

7. The method of claim 1, wherein the pharmaceutical composition is formulated for oral administration.

8. The method of claim 1, wherein the pharmaceutical composition is in the form of a capsule.

9. The method of claim 1, wherein the pharmaceutical composition is formulated for delivery to the colon.

10. The method of claim 1, wherein the pharmaceutical composition comprises between  $1 \times 10^7$  and  $1 \times 10^{10}$  colony forming units (CFUs) per bacterial strain.

30 11. A method to suppress an abnormal or excessive immune response comprising administering a pharmaceutical composition comprising a purified bacterial mixture consisting of 8 bacterial strains comprising 16S rDNA sequences of at least 97% sequence identity 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, to the subject in an amount sufficient to suppress the abnormal or excessive immune response.

35 12. The method of claim 11, wherein the abnormal or excessive immune response is suppressed by inducing proliferation and/or accumulation of regulatory T cells.

13. The method of claim 11, wherein the subject is administered one or more doses of an antibiotic prior to the pharmaceutical composition.

14. The method of claim 13, wherein the antibiotic is vancomycin, kanamycin, gentamicin, colistin, metronidazole, clindamycin, fidaxomicin, or cefoperazone.

15. The method of claim 13, wherein the antibiotic is vancomycin.

16. The method of claim 11, wherein the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

17. The method of claim 11, wherein the pharmaceutical composition is formulated for oral administration.

18. The method of claim 11, wherein the pharmaceutical composition is in the form of a capsule.

19. The method of claim 11, wherein the pharmaceutical composition is formulated for delivery to the colon.

20. The method of claim 11, wherein the pharmaceutical composition comprises between  $1 \times 10^7$  and  $1 \times 10^{10}$  colony forming units (CFUs) per bacterial strain.

\* \* \* \* \*