



- (51) **International Patent Classification:**
C12N 9/90 (2006.01) *C12N 15/10* (2006.01)
C12P 7/10 (2006.01)
- (21) **International Application Number:**
PCT/US2013/051718
- (22) **International Filing Date:**
23 July 2013 (23.07.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/675,241 24 July 2012 (24.07.2012) US
- (71) **Applicant (for all designated States except US):** **BP CORPORATION NORTH AMERICA INC.** [US/US]; 501 Westlake Park Boulevard, Houston, Texas 77079 (US).
- (72) **Inventors; and**
- (71) **Applicants (for US only):** **NUNN, David Neal Jr.** [US/US]; 150 W. Warrenville Road, Mail Code 200-1W, Naperville, Illinois 60563 (US). **LUGINBUHL, Peter** [CH/US]; 150 W. Warrenville Road, Mail Code 200-1W, Naperville, Illinois 60563 (US). **LI, Ling** [CN/US]; 150 W. Warrenville Road, Mail Code 200-1W, Naperville, Illinois 60563 (US). **BURJA, Adam Martin** [AU/US]; 150 W. Warrenville Road, Mail Code 200-1W, Naperville, Illinois 60563 (US). **BARTNEK, Jon Peter Flash** [US/US]; 150 W. Warrenville Road, Mail Code 200-1W, Naperville, Illinois 60563 (US).
- (74) **Agents:** **KOLMAN, Michael F et al.;** BP Corporation North America Inc., 150 W. Warrenville Road, Mail Code

200-1W; Customer No. 04249, Naperville, Illinois 60563 (US).

- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) **Title:** XYLOSE ISOMERASES AND THEIR USES

(57) **Abstract:** This disclosure relates to novel xylose isomerases and their uses, particularly in fermentation processes that employ xylose-containing media.



WO 2014/018552 A1

XYLOSE ISOMERASES AND THEIR USES

1. BACKGROUND

[0001] The efficient, commercial production of biofuels from plant material, such as sugarcane, requires the fermentation of pentoses, such as xylose. Xylose in plant material typically comes from lignocellulose, which is a matrix composed of cellulose, hemicelluloses, and lignin. Lignocellulose is broken down either by acid hydrolysis or enzymatic reaction, yielding xylose in addition to other monosaccharides, such as glucose (Maki *et al.*, 2009, Int. J. Biol. Sci. 5:500-516).

[0002] Fungi, especially *Saccharomyces cerevisiae*, are commercially relevant microorganisms that ferment sugars into biofuels such as ethanol. However, *S. cerevisiae* does not endogenously metabolize xylose, requiring genetic modifications that allow it to convert xylose into xylulose. Other organisms, whose usefulness in ethanol production is limited, are able to metabolize xylose (Nevigot, 2008, Microbiol. Mol. Biol. Rev. 72:379-412).

[0003] Two pathways have been identified for the metabolism of xylose to xylulose in microorganisms: the xylose reductase (XR, EC 1.1.1.307) / xylitol dehydrogenase (XDH, EC 1.1.1.9, 1.1.1.10 and 1.1.1.B19) pathway and the xylose isomerase (XI, EC 5.3.1.5) pathway. Use of the XR/XDH pathway for xylose metabolism creates an imbalance of cofactors (excess NADH and NADP+) limiting the potential output of this pathway for the production of ethanol. The XI pathway, on the otherhand, converts xylose to xylulose in a single step and does not create a cofactor imbalance (Young *et al.*, 2010, Biotechnol. Biofuels 3:24-36).

[0004] Because *S. cerevisiae* does not possess a native XI, it has been desirable to search for an XI in another organism to insert into *S. cerevisiae* for the purpose of biofuels production. Several XI genes have been discovered, although little or no enzymatic activity upon expression in *S. cerevisiae* has been a common problem. The XI from *Piromyces* sp. E2 was the first heterologously expressed XI in *S. cerevisiae* whose enzymatic activity could be observed (WO 03/062430).

2. SUMMARY

[0005] Due to the physiology of *S. cerevisiae* and the process of commercial biofuel production, there are other characteristics besides activity that are valuable in a commercially useful XI. During fermentation, the pH of the yeast cell and its environment can become more acidic (Rosa and Sa-Correia, 1991, Appl. Environ. Microbiol. 57:830-835). The ability of the XI to function in an acidic environment is therefore highly desirable. Therefore, there is still a need in the art for XI enzymes with enhanced activity to convert xylose to xylulose for biofuels production under a broader range of commercially relevant conditions.

[0006] The present disclosure relates to novel xylose isomerases. The xylose isomerases have desirable characteristics for xylose fermentation, such as high activity, tolerance to acidic conditions (*i.e.*, pH levels below 7, *e.g.*, pH 6.5 or pH 6), or both.

[0007] The present disclosure has multiple aspects. In one aspect, the disclosure is directed to XI polypeptides. The polypeptides of the disclosure typically comprise amino acid sequences having at

least 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 98%, 99% or 100% sequence identity to any of the XI polypeptides of Table 1, or the catalytic domain or dimerization domain thereof, or are encoded by nucleic acid sequences comprising nucleotide sequences having at least 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 98%, 99% or 100% sequence identity to any of the nucleic acids of Table 1:

TABLE 1					
SEQ ID NO:	Clone No.	Organism Classification	Type of Sequence	Catalytic Domain	Dimerization Domain
1	1754MI2_001	<i>Bacteroidales</i>	DNA		
2	1754MI2_001	<i>Bacteroidales</i>	Amino Acid	2-376	377-437
3	5586MI6_004	<i>Bacteroidales</i>	DNA		
4	5586MI6_004	<i>Bacteroidales</i>	Amino Acid	2-376	377-437
5	5749MI1_003	<i>Bacteroidales</i>	DNA		
6	5749MI1_003	<i>Bacteroidales</i>	Amino Acid	2-381	382-442
7	5750MI1_003	<i>Bacteroidales</i>	DNA		
8	5750MI1_003	<i>Bacteroidales</i>	Amino Acid	2-381	382-442
9	5750MI2_003	<i>Bacteroidales</i>	DNA		
10	5750MI2_003	<i>Bacteroidales</i>	Amino Acid	2-381	382-442
11	5586MI5_004	<i>Bacteroides</i>	DNA		
12	5586MI5_004	<i>Bacteroides</i>	Amino Acid	2-375	376-435
13	5586MI202_004	<i>Bacteroides</i>	DNA		
14	5586MI202_004	<i>Bacteroides</i>	Amino Acid	2-377	378-438
15	5586MI211_003	<i>Bacteroides</i>	DNA		
16	5586MI211_003	<i>Bacteroides</i>	Amino Acid	2-376	377-437
17	5606MI1_005	<i>Bacteroides</i>	DNA		
18	5606MI1_005	<i>Bacteroides</i>	Amino Acid	2-377	378-438
19	5606MI2_003	<i>Bacteroides</i>	DNA		
20	5606MI2_003	<i>Bacteroides</i>	Amino Acid	2-378	379-439
21	5610MI3_003	<i>Bacteroides</i>	DNA		
22	5610MI3_003	<i>Bacteroides</i>	Amino Acid	2-377	378-439
23	5749MI2_004	<i>Bacteroides</i>	DNA		
24	5749MI2_004	<i>Bacteroides</i>	Amino Acid	2-377	378-438
25	5750MI3_003	<i>Bacteroides</i>	DNA		
26	5750MI3_003	<i>Bacteroides</i>	Amino Acid	2-377	378-438
27	5750MI4_003	<i>Bacteroides</i>	DNA		
28	5750MI4_003	<i>Bacteroides</i>	Amino Acid	2-377	378-438
29	5751MI4_002	<i>Bacteroides</i>	DNA		
30	5751MI4_002	<i>Bacteroides</i>	Amino Acid	2-376	377-437
31	5751MI5_003	<i>Bacteroides</i>	DNA		
32	5751MI5_003	<i>Bacteroides</i>	Amino Acid	2-377	378-438
33	5751MI6_004	<i>Bacteroides</i>	DNA		
34	5751MI6_004	<i>Bacteroides</i>	Amino Acid	2-377	378-438
35	5586MI22_003	<i>Clostridiales</i>	DNA		
36	5586MI22_003	<i>Clostridiales</i>	Amino Acid	2-375	376-439
37	1753MI4_001	<i>Firmicutes</i>	DNA		
38	1753MI4_001	<i>Firmicutes</i>	Amino Acid	2-374	375-440
39	1753MI6_001	<i>Firmicutes</i>	DNA		

TABLE 1

SEQ ID NO:	Clone No.	Organism Classification	Type of Sequence	Catalytic Domain	Dimerization Domain
40	1753MI6_001	<i>Firmicutes</i>	Amino Acid	2-374	375-440
41	1753MI35_004	<i>Firmicutes</i>	DNA		
42	1753MI35_004	<i>Firmicutes</i>	Amino Acid	2-375	376-441
43	1754MI9_004	<i>Firmicutes</i>	DNA		
44	1754MI9_004	<i>Firmicutes</i>	Amino Acid	2-375	376-440
45	1754MI22_004	<i>Firmicutes</i>	DNA		
46	1754MI22_004	<i>Firmicutes</i>	Amino Acid	2-375	376-440
47	727MI1_002	<i>Firmicutes</i>	DNA		
48	727MI1_002	<i>Firmicutes</i>	Amino Acid	2-372	373-436
49	727MI9_005	<i>Firmicutes</i>	DNA		
50	727MI9_005	<i>Firmicutes</i>	Amino Acid	2-374	375-438
51	727MI27_002	<i>Firmicutes</i>	DNA		
52	727MI27_002	<i>Firmicutes</i>	Amino Acid	2-374	375-439
53	1753MI2_006	<i>Neocallimastigales</i>	DNA		
54	1753MI2_006	<i>Neocallimastigales</i>	Amino Acid	2-376	377-437
55	5586MI3_005	<i>Neocallimastigales</i>	DNA		
56	5586MI3_005	<i>Neocallimastigales</i>	Amino Acid	2-376	377-437
57	5586MI91_002	<i>Neocallimastigales</i>	DNA		
58	5586MI91_002	<i>Neocallimastigales</i>	Amino Acid	2-376	377-437
59	5586MI194_003	<i>Neocallimastigales</i>	DNA		
60	5586MI194_003	<i>Neocallimastigales</i>	Amino Acid	2-376	377-438
61	5586MI198_003	<i>Neocallimastigales</i>	DNA		
62	5586MI198_003	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
63	5586MI201_003	<i>Neocallimastigales</i>	DNA		
64	5586MI201_003	<i>Neocallimastigales</i>	Amino Acid	2-376	377-438
65	5586MI204_002	<i>Neocallimastigales</i>	DNA		
66	5586MI204_002	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
67	5586MI207_002	<i>Neocallimastigales</i>	DNA		
68	5586MI207_002	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
69	5586MI209_003	<i>Neocallimastigales</i>	DNA		
70	5586MI209_003	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
71	5586MI214_002	<i>Neocallimastigales</i>	DNA		
72	5586MI214_002	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
73	5751MI3_001	<i>Neocallimastigales</i>	DNA		
74	5751MI3_001	<i>Neocallimastigales</i>	Amino Acid	2-375	376-437
75	5753MI3_002	<i>Prevotella</i>	DNA		
76	5753MI3_002	<i>Prevotella</i>	Amino Acid	2-376	377-439
77	1754MI1_001	<i>Prevotella</i>	DNA		
78	1754MI1_001	<i>Prevotella</i>	Amino Acid	2-377	378-439
79	1754MI3_007	<i>Prevotella</i>	DNA		
80	1754MI3_007	<i>Prevotella</i>	Amino Acid	2-377	378-439
81	1754MI5_009	<i>Prevotella</i>	DNA		
82	1754MI5_009	<i>Prevotella</i>	Amino Acid	2-375	376-437
83	5586MI1_003	<i>Prevotella</i>	DNA		

TABLE 1

SEQ ID NO:	Clone No.	Organism Classification	Type of Sequence	Catalytic Domain	Dimerization Domain
84	5586MI1_003	<i>Prevotella</i>	Amino Acid	2-377	378-439
85	5586MI2_006	<i>Prevotella</i>	DNA		
86	5586MI2_006	<i>Prevotella</i>	Amino Acid	2-377	378-439
87	5586MI8_003	<i>Prevotella</i>	DNA		
88	5586MI8_003	<i>Prevotella</i>	Amino Acid	2-377	378-439
89	5586MI14_003	<i>Prevotella</i>	DNA		
90	5586MI14_003	<i>Prevotella</i>	Amino Acid	2-377	378-439
91	5586MI26_003	<i>Prevotella</i>	DNA		
92	5586MI26_003	<i>Prevotella</i>	Amino Acid	2-377	378-439
93	5586MI86_001	<i>Prevotella</i>	DNA		
94	5586MI86_001	<i>Prevotella</i>	Amino Acid	2-376	377-438
95	5586MI108_002	<i>Prevotella</i>	DNA		
96	5586MI108_002	<i>Prevotella</i>	Amino Acid	2-377	378-439
97	5586MI182_004	<i>Prevotella</i>	DNA		
98	5586MI182_004	<i>Prevotella</i>	Amino Acid	2-377	378-439
99	5586MI193_004	<i>Prevotella</i>	DNA		
100	5586MI193_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
101	5586MI195_003	<i>Prevotella</i>	DNA		
102	5586MI195_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
103	5586MI216_003	<i>Prevotella</i>	DNA		
104	5586MI216_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
105	5586MI197_003	<i>Prevotella</i>	DNA		
106	5586MI197_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
107	5586MI199_003	<i>Prevotella</i>	DNA		
108	5586MI199_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
109	5586MI200_003	<i>Prevotella</i>	DNA		
110	5586MI200_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
111	5586MI203_003	<i>Prevotella</i>	DNA		
112	5586MI203_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
113	5586MI205_004	<i>Prevotella</i>	DNA		
114	5586MI205_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
115	5586MI206_004	<i>Prevotella</i>	DNA		
116	5586MI206_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
117	5586MI208_003	<i>Prevotella</i>	DNA		
118	5586MI208_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
119	5586MI210_002	<i>Prevotella</i>	DNA		
120	5586MI210_002	<i>Prevotella</i>	Amino Acid	2-374	375-437
121	5586MI212_002	<i>Prevotella</i>	DNA		
122	5586MI212_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
123	5586MI213_003	<i>Prevotella</i>	DNA		
124	5586MI213_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
125	5586MI215_003	<i>Prevotella</i>	DNA		
126	5586MI215_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
127	5607MI1_003	<i>Prevotella</i>	DNA		

TABLE 1

SEQ ID NO:	Clone No.	Organism Classification	Type of Sequence	Catalytic Domain	Dimerization Domain
128	5607MI1_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
129	5607MI2_003	<i>Prevotella</i>	DNA		
130	5607MI2_003	<i>Prevotella</i>	Amino Acid	2-376	377-442
131	5607MI3_003	<i>Prevotella</i>	DNA		
132	5607MI3_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
133	5607MI4_005	<i>Prevotella</i>	DNA		
134	5607MI4_005	<i>Prevotella</i>	Amino Acid	2-376	377-438
135	5607MI5_002	<i>Prevotella</i>	DNA		
136	5607MI5_002	<i>Prevotella</i>	Amino Acid	2-376	377-439
137	5607MI6_002	<i>Prevotella</i>	DNA		
138	5607MI6_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
139	5607MI7_002	<i>Prevotella</i>	DNA		
140	5607MI7_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
141	5608MI1_004	<i>Prevotella</i>	DNA		
142	5608MI1_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
143	5608MI2_002	<i>Prevotella</i>	DNA		
144	5608MI2_002	<i>Prevotella</i>	Amino Acid	2-375	376-437
145	5608MI3_004	<i>Prevotella</i>	DNA		
146	5608MI3_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
147	5609MI1_005	<i>Prevotella</i>	DNA		
148	5609MI1_005	<i>Prevotella</i>	Amino Acid	2-376	377-438
149	5610MI1_003	<i>Prevotella</i>	DNA		
150	5610MI1_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
151	5610MI2_004	<i>Prevotella</i>	DNA		
152	5610MI2_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
153	5751MI1_003	<i>Prevotella</i>	DNA		
154	5751MI1_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
155	5751MI2_003	<i>Prevotella</i>	DNA		
156	5751MI2_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
157	5752MI1_003	<i>Prevotella</i>	DNA		
158	5752MI1_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
159	5752MI2_003	<i>Prevotella</i>	DNA		
160	5752MI2_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
161	5752MI3_002	<i>Prevotella</i>	DNA		
162	5752MI3_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
163	5752MI5_003	<i>Prevotella</i>	DNA		
164	5752MI5_003	<i>Prevotella</i>	Amino Acid	2-376	377-438
165	5752MI6_004	<i>Prevotella</i>	DNA		
166	5752MI6_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
167	5753MI1_002	<i>Prevotella</i>	DNA		
168	5753MI1_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
169	5753MI2_002	<i>Prevotella</i>	DNA		
170	5753MI2_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
171	5753MI4_002	<i>Prevotella</i>	DNA		

TABLE 1					
SEQ ID NO:	Clone No.	Organism Classification	Type of Sequence	Catalytic Domain	Dimerization Domain
172	5753MI4_002	<i>Prevotella</i>	Amino Acid	2-376	377-438
173	5752MI4_004	<i>Prevotella</i>	DNA		
174	5752MI4_004	<i>Prevotella</i>	Amino Acid	2-376	377-438
175	727MI4_006	<i>Rhizobiales</i>	DNA		
176	727MI4_006	<i>Rhizobiales</i>	Amino Acid	2-373	374-435

[0008] In specific embodiments, a polypeptide of the disclosure comprises an amino acid sequence having:

(1) (a) at least 97% or 98% sequence identity to SEQ ID NO:78 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:78) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:78 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:78) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(2) (a) at least 95%, 97% or 98% sequence identity to SEQ ID NO:96 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:96) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:96 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:96) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(3) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:38 or the catalytic domain thereof (amino acids 2-374 of SEQ ID NO:38), and optionally further comprises one, two, three, four or all five of (i) SEQ ID NO:206 or SEQ ID NO:207; (ii) SEQ ID NO:208; (iii) SEQ ID NO:209; (iv) SEQ ID NO:210; and (v) SEQ ID NO:211;

(4) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:2 or the catalytic domain thereof (amino acids 2-374 of SEQ ID NO:2);

(5) at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:58 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:58),

(6) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:42 or the catalytic domain thereof (amino acids 2-375 of SEQ ID NO:42), and optionally further comprises one, two or all three of (i) SEQ ID NO:206 or SEQ ID NO:207; (ii) SEQ ID NO:210; and (iii) SEQ ID NO:211;

(7) (a) at least 97% or 98% sequence identity to SEQ ID NO:84 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:84), and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:84 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:84) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(8) (a) at least 97% or 98% sequence identity to SEQ ID NO:80 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:80) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:80 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:80) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(9) at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:54 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:54);

(10) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:46 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:46), and optionally further comprises SEQ ID NO:206 or SEQ ID NO:207;

(11) at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:16 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:16);

(12) at least 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:82 or the catalytic domain thereof (amino acids 2-375 of SEQ ID NO:82); and/or

(13) at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:32 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:32).

[0009] The XIs of the disclosure can be characterized in terms of their activity. In some embodiments, a XI of the disclosure has at least 1.3 times the activity of the *Orpinomyces sp.* XI assigned Genbank Accession No. 169733248 ("Op-XI") at pH 7.5, for example using the assay described in any of Examples 4, 6 and 7. In certain specific embodiments, a XI of the disclosure has an activity ranging from 1.25 to 3.0 times, from 1.5 to 3 times, from 1.5 to 2.25 times, or from 1.75 to 3 times the activity of Op-XI at pH 7.5.

[0010] The XIs of the disclosure can also be characterized in terms of their tolerance to acidic environments (e.g., at a pH of 6.5 or 6). In some embodiments, a XI of the disclosure has at least 1.9 times the activity of the Op-XI at pH 6, for example using the assay described in Example 7. In certain specific embodiments, a XI of the disclosure has an activity ranging from 1.9 to 4.1 times, from 2.4 to 4.1 times, from 2.4 to 3.9 times, or 2.4 to 4.1 times the activity of Op-XI at pH 6.

[0011] Tolerance to acidic environments can also be characterized as a ratio of activity at pH 6 to activity at pH 7.5 ("a pH 6 to pH 7.5 activity ratio"), for example as measured using the assay of Example 7. In some embodiments, the pH 6 to pH 7.5 activity ratio is at least 0.5 or at least 0.6. In various embodiments, the pH 6 to pH 7.5 activity ratio is 0.5-0.9 or 0.6-0.9.

[0012] In another aspect, the disclosure is directed to a nucleic acid which encodes a XI polypeptide of the disclosure. In various embodiments, the nucleic acid comprises a nucleotide sequence with at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 98%, 99% or 100% sequence identity to the nucleotide sequence of any one of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, and 175, or the portion of any of the foregoing sequences encoding a XI catalytic domain or dimerization domain.

[0013] The nucleic acids of the disclosure can be codon optimized, e.g., for expression in eukaryotic organisms such as yeast or filamentous fungi. Exemplary codon optimized open reading frames for

expression in *S. cerevisiae* are SEQ ID NO:238 (encoding a XI of SEQ ID NO:54), SEQ ID NO:239 (encoding a XI of SEQ ID NO:58), SEQ ID NO:244 (encoding a XI of SEQ ID NO:78), SEQ ID NO:245 (encoding a XI of SEQ ID NO:96), SEQ ID NO:246 (encoding a XI of SEQ ID NO:38), SEQ ID NO:247 (encoding a XI of SEQ ID NO:78), SEQ ID NO:248 (encoding a XI of SEQ ID NO:96), and SEQ ID NO:249 (encoding a XI of SEQ ID NO:38). In various embodiments, the disclosure provides nucleic acids comprising nucleotide sequences having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity, or having 100% sequence identity, to the nucleotide sequence of any one of SEQ ID NOs:238, 239, 244, 245, 246, 247, 248 and 249, or the portion of any of the foregoing sequences encoding a XI catalytic domain or dimerization domain.

[0014] In other aspects, the disclosure is directed to a vector comprising a XI-encoding nucleotide sequence, for example a vector having an origin of replication and/or a promoter sequence operably linked to the XI-encoding nucleotide sequence. The promoter sequence can be one that is operable in a eukaryotic cell, for example in a fungal cell. In some embodiments, the promoter is operable in yeast (*e.g.*, *S. cerevisiae*) or filamentous fungi.

[0015] In yet another aspect, the disclosure is directed to a recombinant cell comprising a nucleic acid that encodes a XI polypeptide. Particularly, the cell is engineered to express any of the XI polypeptides described herein. The recombinant cell may be of any species, and is preferably a eukaryotic cell, for example a yeast cell. Suitable genera of yeast include *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Klockera*, *Schwanniomyces*, *Issatchenkia* and *Yarrowia*. In specific embodiments, the recombinant cell is a *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *I. orientalis*, *K. marxianus* or *K. fragilis*. Suitable genera of filamentous fungi include *Aspergillus*, *Penicillium*, *Rhizopus*, *Chrysosporium*, *Myceliophthora*, *Trichoderma*, *Humicola*, *Acremonium* and *Fusarium*. In specific embodiments, the recombinant cell is an *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Penicillium chrysogenum*, *Myceliophthora thermophila*, or *Rhizopus oryzae*.

[0016] The recombinant cell may also be mutagenized or engineered to include modifications other than the recombinant expression of XI, particularly those that make the cell more suited to utilize xylose in a fermentation pathway. Exemplary additional modifications create one, two, three, four, five or even more of the following phenotypes: (a) increase in xylose transport into the cell; (b) increase in aerobic growth rate on xylose; (c) increase in xylulose kinase activity; (d) increase in flux through the pentose phosphate pathway into glycolysis, (e) decrease in aldose reductase activity, (f) decrease in sensitivity to catabolite repression, (g) increase in tolerance to biofuels, *e.g.*, ethanol, (h) increase tolerance to intermediate production (*e.g.*, xylitol), (i) increase in temperature tolerance, (j) osmolarity of organic acids, and (k) a reduced production of byproducts.

[0017] Increases in activity can be achieved by increased expression levels, for example expression of a hexose or pentose (*e.g.*, xylose) transporter, a xylulose kinase, a glycolytic enzyme, or an

ethanologenic enzyme is increased. The increased expression levels are achieved by overexpressing an endogenous protein or by expressing a heterologous protein.

[0018] Other modifications to the recombinant cell that are part of the disclosure are modifications that decrease the activity of genes or pathways in the recombinant cell. Preferably, the expression levels of one, two, three or more of the genes for hexose kinase, MIG-1, MIG-2, XR, aldose reductase, and XDH are reduced. Reducing gene activity can be achieved by a targeted deletion or disruption of the gene (and optionally reintroducing the gene under the control of a different promoter that drives lower levels of expression or inducible expression).

[0019] In yet other aspects, the disclosure is directed to methods of producing fermentation products, for example one or more of ethanol, butanol, diesel, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, itaconic acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β -lactam antibiotic and a cephalosporin. Typically, a cell that recombinantly expresses a XI of the disclosure is cultured in a xylose-containing medium, for example a medium supplemented with a lignocellulosic hydrolysate. The media may also contain glucose, arabinose, or other sugars, particularly those derived from lignocellulose. The media may be of any pH, particularly a pH between 3.0 and 9.0, preferably between 4.0 and 8.0, more preferably between 5.0 and 8.0, even more preferably between 6.0 and 7.5. The culture may occur in any media where the culture is under anaerobic or aerobic conditions, preferably under anaerobic conditions for production of compounds mentioned above and aerobically for biomass / cellular production. Optionally, the methods further comprise recovering the fermentation product produced by the recombinant cell.

3. BRIEF DESCRIPTION OF THE FIGURES

[0020] **FIGS. 1A-1B** are maps for the vector pMEV- Δ xylA (MEV3 xylA del) and PCR-BluntII-TOPO-xylA, respectively, used in the activity-based screen for XIs.

[0021] **FIG. 2** illustrates the experimental strategy for the two-step marker exchange approach.

[0022] **FIG. 3** is a map of the vector p426PGK1 for expressing XI in yeast strain, *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108).

[0023] **FIG. 4** shows the growth rates on xylose containing media of selected clones expressed in yeast strain, *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108).

[0024] **FIGS. 5A-5D** are maps for the vectors pYDAB-006, pYDURA01, pYDPt-005 and pYDAB-0006, respectively, all used in creating strains of industrial *S. cerevisiae* strain yBPA130 with a single genomic copy of select XI clones.

[0025] **FIG. 6** is a map of vector YDAB008-rDNA for multiple XI integration into *S. cerevisiae* strain yBPB007 and yBPB008.

[0026] **FIGS. 7A-7D** show monosaccharide (including xylose) utilization and ethanol production by strains of industrial *S. cerevisiae* with multiple copies of XI clones integrated into ribosomal DNA loci.

[0027] FIG. 8: Production of ethanol from glycolytic and pentose phosphate (“PPP”) pathways. Not all steps are shown. For example, glyceraldehyde-3-phosphate is converted to pyruvate via a series of glycolytic steps: (1) glyceraldehyde-3-phosphate to 3-phospho-D-glycerol-phosphate catalyzed by glyceraldehyde-3-phosphate dehydrogenase (TDH1-3); (2) 3-phospho-D-glycerol-phosphate to 3-phosphoglycerate catalyzed by 3-phosphoglycerate kinase (PGK1); (3) 3-phosphoglycerate to 2-phosphoglycerate catalyzed by phosphoglycerate mutase (GPM1); (4) 2-phosphoglycerate to phosphoenolpyruvate catalyzed by enolase (ENO1; ENO2); and (5) phosphoenolpyruvate to pyruvate catalyzed by pyruvate kinase (PYK2; CDC19). Other abbreviations: DHAP = dihydroxy-acetone-phosphate; GPD = Glycerol-3-phosphate dehydrogenase; RHR2/HOR2 = DL-glycerol-3-phosphatase; XI = xylose isomerase; GRE = xylose reductase / aldose reductase; XYL = xylitol dehydrogenase; XKS = xylulokinase; PDC = pyruvate decarboxylase; ADH = alcohol dehydrogenase; ALD = aldehyde dehydrogenase; HXK = hexokinase; PGI = phosphoglucose isomerase; PFK = phosphofructokinase; FBA = aldolase; TPI = triosephosphate isomerase; ZWF = glucose-6 phosphate dehydrogenase; SOL = 6-phosphogluconolactonase; GND = 6-phosphogluconate dehydrogenase; RPE = D-ribulose-5-Phosphate 3-epimerase; RKI = ribose-5-phosphate ketol-isomerase; TKL = transketolase; TAL = transaldolase. Heavy dashed arrows indicate reactions and corresponding enzymes that can be reduced or eliminated to increase xylose utilization, particularly in the production of ethanol, and heavy solid arrows indicate reactions and corresponding enzymes that can be increased to increase xylose utilization, particularly in the production of ethanol. The enzymes shown in FIG. 8 are encoded by *S. cerevisiae* genes. The *S. cerevisiae* genes are used for exemplification purposes. Analogous enzymes and modifications in other organisms are within the scope of the present disclosure.

4. DETAILED DESCRIPTION

4.1 Xylose Isomerase Polypeptides

[0028] A “xylose isomerase” or “XI” is an enzyme that catalyzes the direct isomerisation of D-xylose into D-xylulose and/or vice versa. This class of enzymes is also known as D-xylose ketoisomerases. A xylose isomerase herein may also be capable of catalyzing the conversion between D-glucose and D-fructose (and accordingly may therefore be referred to as a glucose isomerase).

[0029] A “XI polypeptide of the disclosure” or a “XI of the disclosure” is a xylose isomerase having an amino acid sequence that is related to any one of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, or 176. In some embodiments, the xylose isomerase of the disclosure has an amino acid sequence that is at least about 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity thereto, or to a catalytic or dimerization domain

thereof. The xylose isomerase of the disclosure can also have 100% sequence identity to one of the foregoing sequences.

[0030] The disclosure provides isolated, synthetic or recombinant XI polypeptides comprising an amino acid sequence having at least about 80%, *e.g.*, at least about 81 %, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete (100%) sequence identity to a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, or 176, over a region of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 residues, or over the full length of the polypeptide, over the length of catalytic domain, or over the length of the dimerization domain.

[0031] The XI polypeptides of the disclosure can be encoded by a nucleic acid sequence having at least about 80%, about 85%, about 86%, about 87%, about 88%, about 89%, or about 90% sequence identity to 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, or 175, or by a nucleic acid sequence capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, or 175, or to a fragment thereof. Exemplary nucleic acids of the disclosure are described in Section 4.2 below.

[0032] In specific embodiments, a polypeptide of the disclosure comprises an amino acid sequence having:

(1) (a) at least 97% or 98% sequence identity to SEQ ID NO:78 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:78) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:78 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:78) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(2) (a) at least 95%, 97% or 98% sequence identity to SEQ ID NO:96 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:96) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:96 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:96) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(3) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:38 or the catalytic domain thereof (amino acids 2-374 of SEQ ID NO:38), and optionally further

comprises one, two, three, four or all five of (i) SEQ ID NO:206 or SEQ ID NO:207; (ii) SEQ ID NO:208; (iii) SEQ ID NO:209; (iv) SEQ ID NO:210; and (iv) SEQ ID NO:211;

(4) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:2 or the catalytic domain thereof (amino acids 2-374 of SEQ ID NO:2);

(5) at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:58 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:58),

(6) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:42 or the catalytic domain thereof (amino acids 2-375 of SEQ ID NO:42), and optionally further comprises one, two or all three of (i) SEQ ID NO:206 or SEQ ID NO:207; (ii) SEQ ID NO:210; and (iii) SEQ ID NO:211;

(7) (a) at least 97% or 98% sequence identity to SEQ ID NO:84 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:84), and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:84 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:84) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(8) (a) at least 97% or 98% sequence identity to SEQ ID NO:80 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:80) and/or (b) at least 80%, 85%, 90%, 93% or 95% sequence identity to SEQ ID NO:80 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:80) and further comprises (i) SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) SEQ ID NO:214;

(9) at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:54 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:54);

(10) at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:46 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:46) , and optionally further comprises SEQ ID NO:206 or SEQ ID NO:207;

(11) at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:16 or the catalytic domain thereof (amino acids 2-376 of SEQ ID NO:16);

(12) at least 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:82 or the catalytic domain thereof (amino acids 2-375 of SEQ ID NO:82) ; and/or

(13) at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:32 or the catalytic domain thereof (amino acids 2-377 of SEQ ID NO:32).

[0033] An example of an algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, 1990, J. Mol. Biol. 215:403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. These initial neighborhood word hits act as starting points to find longer HSPs containing them. The word hits are expanded in both directions along each of the two sequences being compared for as far as the

cumulative alignment score can be increased. Extension of the word hits is stopped when: the cumulative alignment score falls off by the quantity X from a maximum achieved value; the cumulative score goes to zero or below; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff & Henikoff, 1992, Proc. Nat'l. Acad. Sci. USA 89:10915-10919) alignments (B) of 50, expectation (E) of 10, M⁵, N⁻⁴, and a comparison of both strands.

[0034] Any of the amino acid sequences described herein can be produced together or in conjunction with at least 1, *e.g.*, at least (or up to) 2, 3, 5, 10, or 20 heterologous amino acids flanking each of the C- and/or N-terminal ends of the specified amino acid sequence, and or deletions of at least 1, *e.g.*, at least (or up to) 2, 3, 5, 10, or 20 amino acids from the C- and/or N-terminal ends of a XI of the disclosure.

[0035] The XIs of the disclosure can be characterized in terms of their activity. In some embodiments, a XI of the disclosure has at least 1.3 times the activity of the *Orpinomyces sp.* XI assigned Genbank Accession No. 169733248 ("Op-XI") at pH 7.5, for example using the assay described in any of Examples 4, 6 and 7. In certain specific embodiments, a XI of the disclosure has an activity ranging from 1.25 to 3.0 times, from 1.5 to 3 times, from 1.5 to 2.25 times, or from 1.75 to 3 times the activity of Op-XI at pH 7.5.

[0036] The XIs of the disclosure can also be characterized in terms of their tolerance to acidic environments (*e.g.*, at a pH of 6.5 or 6). In some embodiments, a XI of the disclosure has at least 1.9 times the activity of the Op-XI at pH 6, for example using the assay described in Example 7. In certain specific embodiments, a XI of the disclosure has an activity ranging from 1.9 to 4.1 times, from 2.4 to 4.1 times, from 2.4 to 3.9 times, or 2.4 to 4.1 times the activity of Op-XI at pH6.

[0037] Tolerance to acidic environments can also be characterized as a ratio of activity at pH 6 to activity at pH 7.5 ("a pH 6 to pH 7.5 activity ratio"), for example as measured using the assay of Example 7. In some embodiments, the pH 6 to pH 7.5 activity ratio is at least 0.5 or at least 0.6. In various embodiments, the pH 6 to pH 7.5 activity ratio is 0.5-0.9 or 0.6-0.9.

[0038] The xylose isomerases of the disclosure can have one or more (*e.g.*, up to 2, 3, 5, 10, or 20) conservative amino acid substitutions relative to the polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, or 176 or to the portion thereof of discussed above. The conservative substitutions can be chosen from among a group having a similar side chain to the reference amino acid. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and

glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Accordingly, exemplary conservative substitutions for each of the naturally occurring amino acids are as follows: ala to ser; arg to lys; asn to gln or his; asp to glu; cys to ser or ala; gln to asn; glu to asp; gly to pro; his to asn or gln; ile to leu or val; leu to ile or val; lys to arg; gln or glu; met to leu or ile; phe to met, leu or tyr; ser to thr; thr to ser; trp to tyr; tyr to trp or phe; and, val to ile or leu.

[0039] The present disclosure also provides a fusion protein that includes at least a portion (*e.g.*, a fragment or domain) of a XI polypeptide of the disclosure attached to one or more fusion segments, which are typically heterologous to the XI polypeptide. Suitable fusion segments include, without limitation, segments that can provide other desirable biological activity or facilitate purification of the XI polypeptide (*e.g.*, by affinity chromatography). Fusion segments can be joined to the amino or carboxy terminus of a XI polypeptide. The fusion segments can be susceptible to cleavage.

4.2 Xylose Isomerase Nucleic Acids

[0040] A “XI nucleic acid of the disclosure” is a nucleic acid encoding a xylose isomerase of the disclosure. In certain embodiments, the xylose isomerase nucleic acid of the disclosure is encoded by a nucleotide sequence of any one of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, or 175, or a sequence having at least about 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity thereto. The xylose isomerase nucleic acid of the disclosure can also have 100% sequence identity to one of the foregoing sequences.

[0041] The present disclosure provides nucleic acids encoding a polypeptide of the disclosure, for example one described in Section 4.1 above. The disclosure provides isolated, synthetic or recombinant nucleic acids comprising a nucleic acid sequence having at least about 70%, *e.g.*, at least about 71 %, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81 %, 82%, 83%, 84%, 85%, 86%, 87%, 88%; 89%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, or complete (100%) sequence identity to a nucleic acid of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, or 175, over a region of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, or 2000 nucleotides.

[0042] Nucleic acids of the disclosure also include isolated, synthetic or recombinant nucleic acids encoding a XI polypeptide having the sequence of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, or 176, and subsequences thereof (e.g., a conserved domain or a catalytic domain), and variants thereof.

[0043] To increase the likelihood that a XI polypeptide is recombinantly expressed, a XI nucleic acid may be adapted to optimize its codon usage to that of the chosen cell. Several methods for codon optimization are known in the art. For expression in yeast, an exemplary method to optimize codon usage of the nucleotide sequences to that of the yeast is a codon pair optimization technology as disclosed in WO 2006/077258 and/or WO 2008/000632. WO2008/ 000632 addresses codon-pair optimization. Codon-pair optimization is a method wherein the nucleotide sequences encoding a polypeptide are modified with respect to their codon-usage, in particular the codon-pairs that are used, to obtain improved expression of the nucleotide sequence encoding the polypeptide and/or improved production of the encoded polypeptide. Codon pairs are defined as a set of two subsequent triplets (codons) in a coding sequence. Boles codon optimization (see Table 2 of Wiedemann and Boles, 2008, Appl. Environ. Microbiol. 74:2043-2050) can also be used to optimize expression and activity of XIs in yeast. Alternatively, the XI sequence can be optimized using commercially available software, such as Gene Designer (DNA2.0). Preferably, codon optimized sequences avoid nucleotide repeats and restriction sites that are utilized in cloning the XI nucleic acids, by adjusting the settings in commercial software or by manually altering the sequences to substitute codons that introduce undesired sequences, for example with highly utilized codons in the organism of interest. Exemplary codon optimized open reading frames for expression in *S. cerevisiae* are SEQ ID NO:238 (encoding a XI of SEQ ID NO:54), SEQ ID NO:239 (encoding a XI of SEQ ID NO:58), SEQ ID NO:244 (encoding a XI of SEQ ID NO:78), SEQ ID NO:245 (encoding a XI of SEQ ID NO:96), SEQ ID NO:246 (encoding a XI of SEQ ID NO:38), SEQ ID NO:247 (encoding a XI of SEQ ID NO:78), SEQ ID NO:248 (encoding a XI of SEQ ID NO:96), and SEQ ID NO:249 (encoding a XI of SEQ ID NO:38). In various embodiments, the disclosure provides nucleic acids comprising nucleotide sequences having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity, or having 100% sequence identity, to the nucleotide sequence of any one of SEQ ID NOs:238, 239, 244, 245, 246, 247, 248 and 249, or the portion of any of the foregoing sequences encoding a XI catalytic domain or dimerization domain.

4.3 Host Cells and Recombinant Expression

[0044] The disclosure also provides host cells transformed with a XI nucleic acid and recombinant host cells engineered to express XI polypeptides. The XI nucleic acid construct may be

extrachromosomal, on a plasmid, which can be a low copy plasmid or a high copy plasmid. The nucleic acid construct may be maintained episomally and thus comprise a sequence for autonomous replication, such as an autosomal replication sequence. Alternatively, a XI nucleic acid may be integrated in one or more copies into the genome of the cell. Integration into the cell's genome may occur at random by non-homologous recombination but preferably, the nucleic acid construct may be integrated into the cell's genome by homologous recombination as is well known in the art. In certain embodiments, the host cell is bacterial or fungal (e.g., a yeast or a filamentous fungus).

[0045] Suitable host cells of the bacterial genera include, but are not limited to, cells of *Escherichia*, *Bacillus*, *Lactobacillus*, *Pseudomonas*, and *Streptomyces*. Suitable cells of bacterial species include, but are not limited to, cells of *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Lactobacillus brevis*, *Pseudomonas aeruginosa*, and *Streptomyces lividans*.

[0046] Suitable host cells of the genera of yeast include, but are not limited to, cells of *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Klockera*, *Schwanniomyces*, *Phaffia*, *Issatchenkia* and *Yarrowia*. In specific embodiments, the recombinant cell is a *S. cerevisiae*, *C. albicans*, *S. pombe*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *H. polymorpha*, *K. lactis*, *I. orientalis*, *K. marxianus*, *K. fragilis*, *P. pastoris*, *P. canadensis*, *K. marxianus* or *P. rhodozyma*. Exemplary yeast strains that are suitable for recombinant XI expression include, but are not limited to, Lallemand LYCC 6391, Lallemand LYCC 6939, Lallemand LYCC 6469, (all from Lallemand, Inc., Montreal, Canada); NRRL YB-1952 (ARS (NRRL) Collection, U.S. Department of Agriculture); and BY4741.

[0047] Suitable host cells of filamentous fungi include all filamentous forms of the subdivision Eumycotina. Suitable cells of filamentous fungal genera include, but are not limited to, cells of *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Corynascus*, *Chaetomium*, *Cryptococcus*, *Filobasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Hypocrea*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Scytalidium*, *Schizophyllum*, *Sporotrichum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, and *Trichoderma*. In certain aspects, the recombinant cell is a *Trichoderma* sp. (e.g., *Trichoderma reesei*), *Penicillium* sp., *Humicola* sp. (e.g., *Humicola insolens*); *Aspergillus* sp. (e.g., *Aspergillus niger*), *Chrysosporium* sp., *Fusarium* sp., or *Hypocrea* sp. Suitable cells can also include cells of various anamorph and teleomorph forms of these filamentous fungal genera.

[0048] Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*,

Fusarium sulphureum, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[0049] Typically, for recombinant expression, the XI nucleic acid will be operably linked to one or more nucleic acid sequences capable of providing for or aiding the transcription and/or translation of the XI sequence, for example a promoter operable in the organism in which the XI is to be expressed. The promoters can be homologous or heterologous, and constitutive or inducible.

[0050] Preferably, the XI polypeptide is expressed in the cytosol and therefore lacks a mitochondrial or peroxisomal targeting signal.

[0051] Where recombinant expression in a filamentous fungal host is desired, the promoter can be a fungal promoter (including but not limited to a filamentous fungal promoter), a promoter operable in plant cells, a promoter operable in mammalian cells.

[0052] As described in U.S. provisional application no. 61/553,901, filed October 31, 2011, the contents of which are hereby incorporated in their entireties, promoters that are constitutively active in mammalian cells (which can derived from a mammalian genome or the genome of a mammalian virus) are capable of eliciting high expression levels in filamentous fungi such as *Trichoderma reesei*. An exemplary promoter is the cytomegalovirus ("CMV") promoter.

[0053] As described in U.S. provisional application no. 61/553,897, filed October 31, 2011, the contents of which are hereby incorporated in their entireties, promoters that are constitutively active in plant cells (which can derived from a plant genome or the genome of a plant virus) are capable of eliciting high expression levels in filamentous fungi such as *Trichoderma reesei*. Exemplary promoters are the cauliflower mosaic virus ("CaMV") 35S promoter or the Commelina yellow mottle virus ("CoYMV") promoter.

[0054] Mammalian, mammalian viral, plant and plant viral promoters can drive particularly high expression when the associated 5' UTR sequence (i.e., the sequence which begins at the transcription start site and ends one nucleotide (nt) before the start codon), normally associated with the mammalian or mammalian viral promoter is replaced by a fungal 5' UTR sequence.

[0055] The source of the 5' UTR can vary provided it is operable in the filamentous fungal cell. In various embodiments, the 5' UTR can be derived from a yeast gene or a filamentous fungal gene. The 5' UTR can be from the same species, one other component in the expression cassette (e.g., the promoter or the XI coding sequence), or from a different species. The 5' UTR can be from the same

species as the filamentous fungal cell that the expression construct is intended to operate in. In an exemplary embodiment, the 5' UTR comprises a sequence corresponding to a fragment of a 5' UTR from a *T. reesei* glyceraldehyde-3-phosphate dehydrogenase (gpd). In a specific embodiment, the 5' UTR is not naturally associated with the CMV promoter

[0056] Examples of other promoters that can be used include, but are not limited to, a cellulase promoter, a xylanase promoter, the 1818 promoter (previously identified as a highly expressed protein by EST mapping *Trichoderma*). For example, the promoter can suitably be a cellobiohydrolase, endoglucanase, or β -glucosidase promoter. A particularly suitable promoter can be, for example, a *T. reesei* cellobiohydrolase, endoglucanase, or β -glucosidase promoter. Non-limiting examples of promoters include a *cbhl*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pki1*, *gpd1*, *xyn1*, or *xyn2* promoter.

[0057] For recombinant expression in yeast, suitable promoters for *S. cerevisiae* include the MF α 1 promoter, galactose inducible promoters such as the GAL1, GAL7 and GAL10 promoters, glycolytic enzyme promoters including the TPI and PGK promoters, the TDH3 promoter, the TEF1 promoter, the TRP1 promoter, the CYC1 promoter, the CUP1 promoter, the PHO5 promoter, the ADH1 promoter, and the HSP promoter. Promoters that are active at different stage of growth or production (e.g., idiophase or trophophase) can also be used (see, e.g., Puig *et al.*, 1996, Biotechnology Letters 18(8):887-892; Puig and Pérez-Ortín, 2000, Systematic and Applied Microbiology 23(2):300-303; Simon *et al.*, 2001, Cell 106:697-708; Wittenberg and Reed, 2005, Oncogene 24:2746-2755). A suitable promoter in the genus *Pichia* sp. is the AOX1 (methanol utilization) promoter.

[0058] The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the nucleic acid sequence encoding the XI polypeptide. Culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art. As noted, many references are available for the culture and production of many cells, including cells of bacterial and fungal origin. Cell culture media in general are set forth in Atlas and Parks (eds.), 1993, The Handbook of Microbiological Media, CRC Press, Boca Raton, FL, which is incorporated herein by reference. For recombinant expression in filamentous fungal cells, the cells are cultured in a standard medium containing physiological salts and nutrients, such as described in Pourquie *et al.*, 1988, Biochemistry and Genetics of Cellulose Degradation, eds. Aubert, *et al.*, Academic Press, pp. 71-86; and Ilmen *et al.*, 1997, Appl. Environ. Microbiol. 63:1298-1306. Culture conditions are also standard, e.g., cultures are incubated at 30°C in shaker cultures or fermenters until desired levels of XI expression are achieved. Preferred culture conditions for a given filamentous fungus may be found in the scientific literature and/or from the source of the fungi such as the American Type Culture Collection (ATCC). After fungal growth has been established, the cells are exposed to conditions effective to cause or permit the expression of a XI.

[0059] In cases where a XI coding sequence is under the control of an inducible promoter, the inducing agent, *e.g.*, a sugar, metal salt or antibiotics, is added to the medium at a concentration effective to induce XI expression.

[0060] In addition to recombinant expression of a XI polypeptide, a host cell of the disclosure may further include one or more genetic modifications that increase the cell's ability to utilize xylose as a substrate in a fermentation process. Exemplary additional modifications create one, two, three, four, five or even more of the following phenotypes: (a) increase in xylose transport into the cell; (b) increase in aerobic growth rate on xylose; (c) increase in xylulose kinase activity; (d) increase in flux through the pentose phosphate pathway into glycolysis, (e) modulating in aldose reductase activity, (f) decrease in sensitivity to catabolite repression, (g) increase in tolerance to biofuels, *e.g.*, ethanol, (h) increase tolerance to intermediate production (for example xylitol), (i) increase in temperature tolerance, (j) osmolarity of organic acids, and (k) a reduced production of byproducts.

[0061] As illustrated below, a modification that results in one or more of the foregoing phenotypes can be a result of increasing or decreasing expression of an endogenous protein (*e.g.*, by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20) or a result of introducing expression of a heterologous polypeptide. For avoidance of doubt, “decreasing” or “reducing” gene expression encompasses eliminating expression. Decreasing (or reducing) the expression of an endogenous protein can be accomplished by inactivating one or more (or all) endogenous copies of a gene in a cell. A gene can be inactivated by deletion of at least part of the gene or by disruption of the gene. This can be achieved by deleting the some or all of a gene coding sequence or regulatory sequence whose deletion results in a reduction of gene expression in the cell. Examples of modifications that increase xylose utilization or yield of fermentation product are described below.

[0062] **Increasing xylose transport:** Xylose transport can be increased directly or indirectly. For example, a recombinant cell may include one or more genetic modifications that result in expression of a xylose transporter. Exemplary transporters include, but are not limited to GXF1, SUT1 and At6g59250 from *Candida intermedia*, *Pichia stipitis* (now renamed *Scheffersomyces stipitis*; the terms are used interchangeably herein) and *Arabidopsis thaliana*, respectively (Runquist *et al.*, 2010, Biotechnol. Biofuels 3:5), as well as HXT4, HXT5, HXT7, GAL2, AGT1, and GXF2 (see, *e.g.*, Matsushika *et al.*, 2009, Appl. Microbiol. Biotechnol. 84:37-53). Other transporters include PsAraT, SUT2-4 and XUT1-5 from *P. stipitis*; GXS1 from *Candida intermedia*; XylHP and DEHA0D02167 from *Debaryomyces hansenii*; and YALI0C06424 from *Yarrowia lipolytica* (see, *e.g.*, Young *et al.*, 2011, Appl. Environ. Microbiol. 77:3311-3319). Xylose transport can also be increased by (over-) expression of low-affinity hexose transporters, which are capable of non-selectively transporting sugars, including xylose, into the cell once glucose levels are low (*e.g.*, 0.2-1.0 g/l); and includes CgHXT1 – CgHXT5 from *Colletotrichum graminicola*. The foregoing modifications can be made singly or in combinations of two, three or more modifications.

[0063] Increasing xylulose kinase activity: Xylulose kinase activity can be increased by overexpression of a xylulose kinase, *e.g.*, xylulose kinase (XKS1; *Saccharomyces* genome database (“SGD”) accession no. YGR194C) of *S. cerevisiae*, particularly where the recombinant cell is a yeast cell. In one embodiment, a *S. cerevisiae* cell is engineered to include at least 2 additional copies of xylulose kinase under the control of a strong constitutive promoter such as TDH3, TEF1 or PGK1. In another embodiment, overexpression of an endogenous xylulose kinase was engineered. This xylulose kinase having improved kinetic activities through the use of protein engineering techniques known by those skilled in the art.

[0064] Increasing flux through the pentose phosphate pathway: This can be achieved by increasing expression of one or more genes in the pentose phosphate pathway, for example *S. cerevisiae* transaldolase TAL1 (SGD accession no. YLR354C), transketolase TKL1 (SGD accession no. YPR074C), ribulose 5-phosphate epimerase RPE1 (SGD accession no. YJL121C) and ribose-5-phosphate ketoisomerase RKI1 (SGD accession no. YOR095C) and/or one or more genes to increase glycolytic flux, for example *S. cerevisiae* pyruvate kinase PYK1/CDC19 (SGD accession no. YAL038W), pyruvate decarboxylase PDC1 (SGD accession no. YLR044C), pyruvate decarboxylase PDC5 (SGD accession no. YLR134W), pyruvate decarboxylase PDC6 (SGD accession no. YGR087C), the alcohol dehydrogenases ADH1-5 (SGD accession nos. YOL086C, YMR303C, YMR083W, YGL256W, and YBR145W, respectively), and hexose kinase HXK1-2 (SGD accession nos. YFR053C and YGL253W, respectively). In one embodiment, the yeast cell has one additional copy each of TAL1, TKL1, RPE1 and RKI1 from *S. cerevisiae* under the control of strong constitutive promoters (*e.g.*, PGK1, TDH3, TEF1); and may also include improvements to glycolytic flux (*e.g.*, increased copies of genes such as PYK1, PDC1, PDC5, PDC6, ADH1-5) and glucose-6-phosphate and hexokinase. The foregoing modifications can be made singly or in combinations of two, three or more modifications.

[0065] Modulating aldose reductase activity: A recombinant cell can include one or more genetic modifications that increase or reduce (unspecific) aldose reductase (sometimes called ald-keto reductase) activity. Aldose reductase activity can be reduced by one or more genetic modifications that reduce the expression of or inactivate a gene encoding an aldose reductase, for example *S. cerevisiae* GRE3 (SGD accession no. YHR104W).

[0066] In certain embodiments, GRE3 expression is reduced. In one aspect, the recombinant cell is a yeast cell in which the GRE3 gene is deleted. Deletion of GRE3 decreased xylitol yield by 49% and biomass production by 31%, but increased ethanol yield by 19% (Traff-Bjerre *et al.*, 2004, Yeast 21:141-150). In another aspect, the recombinant cell is a yeast cell which has a reduction in expression of GRE3. Reducing GRE3 expression has been shown to result in a two-fold decrease in by-product (*i.e.*, xylitol) formation and an associated improvement in ethanol yield (Traff *et al.*, 2001, Appl. Environ. Microbiol. 67:5668-5674).

[0067] In another embodiment, the recombinant cell is a cell (optionally but not necessarily a yeast cell) in which GRE3 is overexpressed. In a study analyzing the effect of GRE3 overexpression in *S. cerevisiae* to investigate the effect on xylose utilization, an increase of about 30% in xylose consumption and about 120% in ethanol production was noted (Traff-Bjerre *et al.*, 2004, Yeast 21:141-150).

[0068] **Decreasing xylose reductase activity:** A recombinant cell may include one or more genetic modifications that reduce xylose reductase activity. Xylose reductase activity can be reduced by one or more genetic modifications that reduce the expression of or inactivate a gene encoding a xylose reductase.

[0069] **Decreasing sensitivity to catabolite repression:** Glucose and other sugars, such as galactose or maltose, are able to cause carbon catabolite repression in Crabtree-positive yeast, such as *S. cerevisiae*. In one study, xylose was found to decrease the derepression of various enzymes of an engineered *S. cerevisiae* strain capable of xylose utilization by at least 10-fold in the presence of ethanol. Xylose also impaired the derepression of galactokinase and invertase (Belinchon & Gancedo, 2003, Arch. Microbiol. 180:293-297). In certain embodiments, in order to reduce catabolite sensitivity, yeast can include one or more genetic modifications that reduce expression of one or more of GRR1 (SGD accession no. YJR090C), the gene assigned SGD accession no. YLR042C, GAT1 (SGD accession no. YKR067W) and/or one or more genetic modifications that decrease expression of one or more of SNF1 (SGD accession no. YDR477W), SNF4 (SGD accession no. YGL115W), MIG1 (SGD accession no. YGL035C) and CRE1 (SGD accession no. YJL127C). In further embodiments, yeast can include one or more genetic modifications that result in overexpression of the pentose phosphate pathway enzymes. In yet further embodiments, yeast can include one or more genetic modifications that reduce expression of hexo-/glucokinase. In yet a further embodiment, yeast can include one or more genetic modifications that modulate the activity of one or more GATA factors, for example GAT1, DAL80 (SGD accession no. YKR034W), GZF3 (SGD accession no. YJL110C) and GLN3 (SGD accession no. YER040W). The foregoing modifications can be made singly or in combinations of two, three or more modifications.

[0070] **Increasing tolerance to biofuels (e.g., ethanol), pathway intermediates (e.g., xylitol), organic acids and temperature:** For efficient bioethanol production from lignocellulosic biomass, it is useful to improve cellular tolerance to toxic compounds released during the pretreatment of biomass. In one study, the gene encoding PHO13 (SGD accession no. YDL236W), a protein with alkaline phosphatase activity, was disrupted. This resulted in improved ethanol production from xylose in the presence of three major inhibitors (*i.e.*, acetic acid, formic acid and furfural). Further, the specific ethanol productivity of the mutant in the presence of 90 mM furfural was four fold higher (Fujitomi *et al.*, 2012, Biores. Tech., 111:161-166). Thus, in one embodiment, yeast has one or more genetic modifications that reduce PHO13 expression. In other embodiments, yeast, bacterial and fungal cells are evolved under selective conditions to identify strains that can withstand higher

temperatures, higher levels of intermediates, higher levels of organic acids and/or higher levels of biofuels (e.g., ethanol). In yet other embodiments, yeast are engineered to reduce expression of FPS1 (SGD accession no. YLL043W); overexpress unsaturated lipid and ergosterol biosynthetic pathways; reduce expression of PHO13 and/or SSK2 (SGD accession no. YNR031C); modulate global transcription factor cAMP receptor protein, through increasing or decreasing expression; increase expression of MSN2 (SGD accession no. YMR037C), RCN1 (SGD accession no. YKL159C), RSA3 (SGD accession no. YLR221C), CDC19 and/or ADH1; or increase expression of Rice ASR1. The foregoing modifications can be made singly or in combinations of two, three or more modifications.

[0071] Reducing production of byproducts: Glycerol is one of the main byproducts in C6 ethanol production. Reducing glycerol is desirable for increasing xylose utilization by yeast. Production of glycerol can be reduced by deleting the gene encoding the FPS1 channel protein, which mediates glycerol export, and GPD2 (SGD accession no. YOL059W), which encodes glycerol-3-phosphate dehydrogenase; optionally along with overexpression of GLT1 (SGD accession no. YDL171C) and GLN1 (SGD accession no. YPR035W). In one study, FPS1 and GPD2 were knocked-out in one *S. cerevisiae* strain, and in another were replaced by overexpression of GLT1 and GLN1, which encode glutamate synthase and glutamine synthetase, respectively. When grown under microaerobic conditions, these strains showed ethanol yield improvements of 13.17% and 6.66%, respectively. Conversely, glycerol, acetic acid and pyruvic acid were found to all decrease, with glycerol down 37.4% and 41.7%, respectively (Zhang and Chen, 2008, Chinese J. Chem. Eng. 16:620-625).

[0072] Production of glycerol can also be reduced by deleting the NADH-dependent glycerol-3-phosphate dehydrogenase 1 (GPD1; SGD accession no. YDL022W) and/or the NADPH-dependent glutamate dehydrogenase 1 (GDH1; SGD accession no. YOR375C). Sole deletion of GPD1 or GDH1 reduces glycerol production, and double deletion results in a 46.4% reduction of glycerol production as compared to wild-type *S. cerevisiae* (Kim *et al.*, 2012, Bioproc. Biosys. Eng. 35:49-54). Deleting FPS1 can decrease production of glycerol for osmoregulatory reasons.

[0073] Reducing production of acetate can also increase xylose utilization. Deleting ALD6 (SGD accession no. YPL061W) can decrease production of acetate.

[0074] ADH2 can also be deleted to reduce or eliminate acetaldehyde formation from ethanol and thereby increase ethanol yield.

[0075] The foregoing modifications to reduce byproduct formation can be made singly or in combinations of two, three or more modifications.

[0076] In addition to ethanol production, a recombinant XI-expressing cell of the disclosure can be suitable for the production of non-ethanolic fermentation products. Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible by a eukaryotic microorganism such as a yeast or a filamentous fungus. Such fermentation products may be, for example, butanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, itaconic acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β -

lactam antibiotic or a cephalosporin. A preferred modified host cell of the disclosure for production of non-ethanolic fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

[0077] Cells expressing the XI polypeptides of the disclosure can be grown under batch, fed-batch or continuous fermentations conditions. Classical batch fermentation is a closed system, wherein the compositions of the medium is set at the beginning of the fermentation and is not subject to artificial alternations during the fermentation. A variation of the batch system is a fed-batch fermentation in which the substrate is added in increments as the fermentation progresses. Fed-batch systems are useful when catabolite repression is likely to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the medium. Batch and fed-batch fermentations are common and well known in the art. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned medium is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth. Continuous fermentation systems strive to maintain steady state growth conditions. Methods for modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology.

4.4 Fermentation Methods

[0078] A further aspect the disclosure relates to fermentation processes in which the recombinant XI-expressing cells are used for the fermentation of carbon source comprising a source of xylose. Thus, in certain embodiments, the disclosure provides a process for producing a fermentation product by (a) fermenting a medium containing a source of xylose with a recombinant XI-expressing cell as defined herein above, under conditions in which the cell ferments xylose to the fermentation product, and optionally, (b) recovery of the fermentation product. In some embodiments, the fermentation product is an alcohol (*e.g.*, ethanol, butanol, *etc.*), a fatty alcohol (*e.g.*, a C8-C20 fatty alcohol), a fatty acid (*e.g.*, a C8-C20 fatty acid), lactic acid, 3-hydroxypropionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, an amino acid, 1,3 -propanediol, itaconic acid, ethylene, glycerol, and a β -lactam antibiotic such as Penicillin G or Penicillin V and fermentative derivatives thereof and cephalosporins. The fermentation process may be an aerobic or an anaerobic fermentation process.

[0079] In addition to a source of xylose the carbon source in the fermentation medium may also comprise a source of glucose. The source of xylose or glucose may be xylose or glucose as such or may be any carbohydrate oligo- or polymer comprising xylose or glucose units, such as *e.g.*, lignocellulose, xylans, cellulose, starch and the like. Most microorganisms possess carbon catabolite repression that results in sequential consumption of mixed sugars derived from the lignocellulose, reducing the efficacy of the overall process. To increase the efficiency of fermentation, microorganisms that are capable of simultaneous consumption of mixed sugars (*e.g.*, glucose and

xylose) have been developed, for example by rendering them less sensitive to glucose repression (see, *e.g.*, Kim *et al.*, 2010, Appl. Microbiol. Biotechnol. 88:1077-85 and Ho *et al.*, 1999, Adv. Biochem. Eng. Biotechnol. 65:163-92). Such cells can be used for recombinant XI expression and in the fermentation methods of the disclosure.

[0080] The fermentation process is preferably run at a temperature that is optimal for the recombinant XI-expressing cells. Thus, for most yeasts or fungal host cells, the fermentation process is performed at a temperature which is less than 38°C, unless temperature tolerant mutant strains are used, in which case the temperature may be higher. For most yeast or filamentous fungal host cells, the fermentation process is suitably performed at a temperature which is lower than 35°C, 33°C, 30°C or 28°C. Optionally, the temperature is higher than 20°C, 22°C, or 25°C.

[0081] An exemplary process is a process for the production of ethanol, whereby the process comprises the steps of: (a) fermenting a medium containing a source of xylose with a transformed host cell as defined above, whereby the host cell ferments xylose to ethanol; and optionally, (b) recovery of the ethanol. The fermentation medium can also comprise a source of glucose that is also fermented to ethanol. The source of xylose can be sugars produced from biomass or agricultural wastes. Many processes for the production of monomeric sugars such as glucose generated from lignocellulose are well known, and are suitable for use herein. In brief, the cellulolytic material may be enzymatically, chemically, and/or physically hydrolyzed to a glucose and xylose containing fraction. Alternatively, the recombinant XI-expressing cells of the disclosure can be further transformed with one or more genes encoding for enzymes effective for hydrolysis of complex substrates such as lignocellulose, and include but are not limited to cellulases, hemicellulases, peroxidases, laccases, chitinases, proteases, and pectinases. The recombinant cells of the disclosure can then be fermented under anaerobic in the presence of glucose and xylose. Where the recombinant cell is a yeast cell, the fermentation techniques and conditions described for example, by Wyman (1994, Biores. Technol. 50:3-16) and Olsson and Hahn-Hagerdal (1996, Enzyme Microb. Technol. 18:312-331) can be used. After completion of the fermentation, the ethanol may be recovered and optionally purified or distilled. Solid residue containing lignin may be discarded or burned as a fuel.

[0082] The fermentation process may be run under aerobic and anaerobic conditions. In some embodiments, the process is carried out under microaerobic or oxygen limited conditions. Fermentation can be carried out in a batch, fed-batch, or continuous configuration within (bio)reactors.

5. EXAMPLES

5.1 Materials and Methods

5.1.1 Yeast Culture

[0083] Unless stated otherwise for a particular example, yeast transformants were grown in SC-ura media with about 2% glucose at 30°C for about 24 hours. The media contains approx. 20 g agar, approx. 134 g BD Difco™ Yeast Nitrogen Base without amino acids (BD, Franklin Lakes, New

Jersey), and approx. 2 g SC amino-acid mix containing about 85 mg of the following amino acids unless noted (quantity listed in parentheses): L-Adenine (21.0), L-Alanine, L-Arginine, L-Asparagine, L-Aspartic Acid, L-Cysteine, Glutamine, L-Glutamic Acid, Glycine, L-Histidine, Myo-Inositol, L-Isoleucine, L-Leucine (173.4), L-Lysine, L-Methionine, p-Aminobenzoic Acid (8.6), L-Phenylalanine, L-Proline, L-Serine, L-Threonine, L-Tryptophan, L-Tyrosine, L-Valine).

5.1.2 Xylose Isomerase Activity

[0084] XI activity in cell lysates was determined using a method based on that of Kersters-Hilderson *et al.*, 1986, *Enzyme Microb. Technol.* 9:145-148, in which enzymatic conversion of xylose to xylulose by the XI is coupled with the enzymatic conversion of the product (xylulose) to xylitol via the enzyme sorbitol dehydrogenase (SDH). SDH activity requires the oxidation of NADH to NAD⁺. The rate of oxidation of NADH is directly proportional to the rate of SDH conversion of D-xylulose to D-xylitol and is measured by the decrease in absorbance at 340 nm. One unit of enzyme activity as measured by this assay is a decrease of 1 μ mole of NADH per minute under assay conditions. All reactions, solutions, plates, and spectrophotometer were equilibrated to about 35°C prior to use. Assays were performed either on fresh lysates immediately after preparation or lysates that had been frozen at -20°C immediately after preparation. Assays were performed using a BioTek Model: Synergy H1 Hybrid Reader spectrophotometer and 96-well plates (Corning, Model #Costar® #3598). All spectrophotometric readings were performed at 340 nm. A standard curve of NADH was generated with each assay with concentrations ranging from 0 to about 0.6 mM.

[0085] The reaction buffer used for experiments at pH 7.5 was about 100 mM Tris-HCl (pH 7.5). The assay mix was prepared as follows: reaction buffer to which was added about 10 mM MgCl₂, 0.15 mM NADH and 0.05 mg/ml SDH (Roche, catalog #50-720-3313). For experiments where activity was also measured at pH 6, the buffer was changed to about 100 mM sodium phosphate, pH 6. The assay mix for the entire experiment was then prepared as follows: about 10 mM MgCl₂, 1.2 mM NADH and 0.02 mg/ml SDH.

[0086] Any sample dilutions were performed using the reaction buffer as diluent. Reactions were set up by aliquotting about 90 μ l of assay mix into each well of the plates. About 10 μ l of each XI sample was added to the wells. The reactions were started by the addition of about 100 μ l substrate solution (about 1 M D-xylose). Reactions were mixed and read immediately using kinetic assay mode for about 10 minutes. Volumetric activity (VA) units are in milli-absorbance (mA) units per minute per ml of lysate added to the reactions (mA/min/ml). Background VA rates of negative control wells (no enzyme added) were subtracted from VA of samples. Determination of fold improvement over positive control (FIOPC) was obtained by dividing the VA of the XI-samples by the VA observed for a control (*Orpinomyces* xylose isomerase, NCBI:169733248 (Op-XI)) expressed using the same host and expression vector. In some characterizations, the slope of an NADH standard curve was used to convert VA (mA/min) to μ mole-NADH/min (or Units). If protein quantitation was performed, specific activities (SA) were calculated where the units for SA are (μ mole NADH⁺/min/mg, or U/mg

lysate protein). All activities listed (VA or SA) account for any dilutions, volumes of lysate added, and protein concentrations for the lysates assayed.

5.2 Example 2: Activity-Based Discovery Screen For Xylose Isomerases

[0087] Libraries used for the activity-based discovery (“ABD”) screen were in the format of excised phagemids. These libraries were constructed as described in U.S. Patent 6,280,926. Sources for these libraries were environmental rumen samples collected from the foregut of deceased herbivores.

[0088] An *Escherichia coli* screening strain was constructed to identify genes from the environmental libraries encoding xylose isomerase activity. Specifically, *E. coli* strain SEL700, a MG1655 derivative that is *recA*⁻, phage lambda resistant and contains an F' plasmid, was complemented with plasmid pJC859, a derivative of pBR322 containing the *E. coli recA* gene (Kokjohn *et al.*, 1987, J. Bacteriol. 169:1499-1508) to generate a wild-type *recA* phenotype.

[0089] A two-step marker exchange procedure was then used to delete the entire coding sequence of the endogenous *xyIA* xylose isomerase gene. Briefly, pMEV3, a plasmid with a *pir*-dependent replicon (*ori6RK*) encoding kanamycin-resistance and the *sacB* levansucrase, was used as a vector for construction of the *xyIA* deletion plasmid. A fragment of DNA containing the flanking regions of the *xyIA* gene (0.7 kb of sequence 5' and 0.9 kb of sequence 3' of *xyIA*) and containing *BsaI* restriction sites was generated by overlap extension PCR using primers, ligated to pMEV3 digested with *BbsI*, and transformed into *E. coli* by electroporation. Clones were confirmed by sequencing, resulting in plasmid pMEV3-Δ*xyIA* (FIG. 1A).

[0090] The pMEV3-Δ*xyIA* plasmid was then transformed into strain *E. coli* strain SEL700 (MG1655 Δ^r, Δ(*recA-srl*)306,*srl*-301::Tn10-84(Tets), [F' *proAB*, *lacI*^r, *ZAM15*, Tn10 (Tet^r)] pJC859). Single-crossover events were selected for by plating on LB agar plates containing kanamycin (final concentration, about 50 μg/ml). After confirmation of integration of pMEV3- Δ*xyIA* on the chromosome, a second crossover event was selected for by growth on LB agar media containing sucrose (FIG. 2). Colonies displaying resistance to kanamycin and the ability to grow on sucrose were screened both by PCR characterization with primers flanking the *xyIA* gene to confirm gene deletion and by growth on a modified MacConkey media (ABD media), comprised of: MacConkey Agar Base (Difco™ #281810) (approximate formula per liter: Pancreatic Digest of Gelatin (17.0 g) Peptones (meat and casein) (3.0 g), Bile Salts No. 3 (1.5 g), Sodium Chloride (5.0 g), Agar (13.5 g), Neutral Red (0.03 g), Crystal Violet (1.0 g, Xylose (30.0 g) and Kanamycin (50 mg). The ABD media contained neutral red, a pH indicator that turns red at a pH <6.8. Colonies of mutants lacking *xyIA* appeared white on this media while colonies with restored xylose metabolism ability appeared red in color due to the fermentation of xylose to xylulose, which lowered the pH of the media surrounding those colonies.

[0091] Following the successful deletion of *xyIA*, the resulting strain was cured of pJC859 by the following method: The *xyIA* deletion strain was grown for about 24 hours in LB media containing tetracycline at a final concentration, about 20 μg /ml, at around 37°C. The next day the cells were sub-

cultured (1:100 dilution) into LB tetracycline (at the same concentration) media and incubated at about three different temperatures (30, 37, and 42°C). Cells were passaged the same way as above for about two more days. Dilutions of the resulting cultures were plated on LB plates to isolate single colonies. Colonies were replica plated onto LB agar plates with and without Carbenicillin (at about 100 µg /ml, final concentration). Carbenicillin resistant colonies were deemed to still contain vector pJC859 whereas carbenicillin sensitive colonies were cured of pJC859, restoring the *recA* genotype of strain SEL700. This strain, SEL700 $\Delta xyIA$, was used for the ABD screening.

[0092] The ABD screening method was verified by creating a positive control strain by PCR amplification of the *xyIA* gene from *E. coli* K12 and cloning into the PCR-BluntII TOPO vector (Invitrogen, Carlsbad, California) using standard procedures. This vector (PCR-BluntII-TOPO-*xyIA*, FIG. 1B) was then transformed into the screening strain (SEL700 $\Delta xyIA$). Complementation of the xylose phenotype was verified by growth of transformants on ABD media and appearance of red halos indicating xylose utilization.

[0093] The libraries were screened for XI activity by infecting strain SEL700 $\Delta xyIA$ with the excised phagemid libraries. Infected cells were plated onto ABD media and only colonies with red “halos” (indicating xylose fermentation), were carried forward. Positives were purified to single colonies, and regrown on ABD media to confirm phenotype.

5.3 Example 2: Sequence-Based Discovery For Xylose Isomerases

[0094] Libraries used for sequence-based discovery (“SBD”) were in the format of genomic DNA (gDNA) extractions. These libraries were constructed as described in U.S. Patent No. 6,280,926. Sources for these libraries were samples collected from the guts of deceased herbivores.

[0095] XI genes often exist in conserved gene clusters (Dodd *et al.*, 2011, Molecular Microbiol. 79:292-304). In order to obtain full length XI gene sequences from metagenomic samples, primers were designed to both upstream and downstream conserved DNA sequences found in several *Bacteroides* species, typically xylulose kinase and xylose permease, respectively. These flanking DNA sequences were obtained from public databases. Sample genomic DNA was extracted from eleven different animal rumen samples. Left flanking consensus primer has the sequence 5'-GCIGCICARGARGGNATYGTVTT-3' (SEQ ID NO:177) (this primer codes for the amino acid motif AAQEGIV(F) (SEQ ID NO:178)). Right flanking consensus primer has the sequence 5'-GCDATYTCNGCRATRTACATSGG-3' (SEQ ID NO:179) (this primer codes for the amino acid motif PMYIAEIA (SEQ ID NO:180)). PCR reactions were carried out using touchdown cycling conditions, and hot start Platinum® Taq DNA polymerase (Invitrogen, Carlsbad, California). PCR products of expected size were purified and subcloned into pCR4-TOPO vector system (Invitrogen, Carlsbad, California). Positive colonies from the TOPO-based PCR libraries were transformed into TOP10 (Invitrogen, Carlsbad, California) and the transformants grown on LB agar plates with kanamycin (about 25 µg/ml final concentration). Resistant colonies were picked and inoculated into 2 columns each of a 96-deep well plate in about 1.2 ml LB kanamycin (25 µg/ml final concentration)

media per well. Cultures were grown overnight at about 30°C. The next day plasmids were purified and inserts sequenced. Sequence analysis revealed multiple full length XI genes. Identification of putative ORFs was done by identifying start and stop codons for the longest protein coding region, and subsequent manual curation based on homology to published xylose isomerase DNA sequences.

5.4 Example 3: XI Sequence Analysis

[0096] Plasmids from both ABD and SBD screens were purified and vector inserts were sequenced using an ABI 3730xl DNA Analyzer and ABI BigDye® v3.1 cycle sequencing chemistry. Identification of putative ORFs was done by identifying start and stop codons for the longest protein coding region, and subsequent manual curation based on homology to published xylose isomerase DNA sequences. The XI ORF identified are set forth in Table 2 below, which indicates the sequences and source organism classification for each XI determined from either the ABD or SBD libraries as well as their assigned sequence identifiers. The putative catalytic domains (based on sequence alignments with other XIs) are underlined.

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
1754MI2_001	<i>Bacteroidales</i>	DNA	1	ATGGCAGTTAAAGAATATTTCCCGAGATAGGCAAGATCGCCTTTGAAGAAAGGAGTCC AAGAACCCATGGCATTCACACTACTACAAATCCAGAGCAGGTAGTAGCCGGAAGAAAATG AAAGATTGGTTCAAGTTCGCTATGGCATGGTGCACACCCCTCTGCGTGAAGGTGGCAG CAGTTCGGTCCGTACCAAGAAATCCCTTGGAAACACAGTGCACACTGCACCTCGAAAGA GCAAGAACAAAATGGACGCGAGTTTCGAGATCATGAGCAAGCTCGGTATCGAGTATTTC TGCTTCCACGATGTTGACCTTATCGACGAGGCTGACACTGTTGAAGAGTACGAGGCTAAC ATGAAGGCTATCACAGCTTACGCAAGAGGAGAAAATGGCGTACTGGCATCAAACTCCTC TGGGAAACAGCCAATGTATTTCGCAACAAGAGATATATGACGCGCTTCTACCAACCTT GACTTCAAGTGGCTGCACGCGTATGCTCCAGATCAAGAACGCTATCGACGCAACTATC GCTCTCGGTGGTACTGCTATGTATTTCGGGCGGCGGTGAGGGTTACATGAGCCTTCTC AACACCGATATGAAGAGAGAGAAAGACATGGCTACCATGCTTACCAATGGCAGCGGAC TATGCTCGTTCTAAGGCTTCAAGGTTACCTTCTTATCGAGCCTAAGCCAATGGAGCGG ATGAAGCAGCAGTACAGTCTGATACTGAGACTGTGAGTGTCTGAGTGTCTCCGCGCCCATGGT CTTGACAGGACTTCAAGGTAAACATCGAGGTAAACACCGTACTCTCGAGGCGCACAC TTCGAGCAGAGCTCCAGTGGCGGTTGACGAGGATGCTCGGAGGATCGACGCGCAAC CGTGTGACTACCAAGAGGCTGGGATACCGACAGTTCCTATCGACCTCTATGAGCTC GTACAGGCTATGATGGTTATCATCAAGGGCGGGTCTCGTGGCGGTACCAACTTCGAC GCCAAGACCGGTCGTAACCAACAGACCTCGAGGATATCTCATCGCTCATGTATCCGGC ATGGATGTCAATGCGACCGCTCTCTCATCGCTGACCTTCTCGAGAAATCTCTGAT CCTGCAATGGTCAAGGAGCTTACGCTTCTTACGACTCAGGATGGCAAGGACTTCGAG AAGGCAAGCTTACTCTCGAGCAGTGTGATTCGATTCGCAAGAAAGAACGGCGAGCCTAAG AGCACCAGCGGAAAGCAGGAGCTACGAGCTATCGTCAATCTCTACATCTAA
1754MI2_001	<i>Bacteroidales</i>	Amino Acid	2	MAVKEYFEI GKIAFEKGESKNPMAPHYYNPEQVAVGKMKDWFKFAWAWHTLCAEGGD QFGPGTKRFPWNTGATALERAKNMDAGFEIMSKLIGIEYFCFHDVLDLIDEADTVEEYEAN MKAITAYAKEKMAATGILKLLWGTANVFGNKRYMNGASTNPDENVAARAMLQIKNAIDATI ALGGDCYVFWGGREGYMSLLNTDMKREKEHMTMLTMARDYARSKGFGKFTLIEPRMEP MKHQYDVIDTETVVGFLEAHGLDKDFKVNIEVNHATLAGHTFEHELQCAVDAGMLGSDAN RGDIQNGWDIDQFPIDLYELVQANMVIKGGGLVGGTNFDKIRNRNSTDLEDIFTAHSV MDVMARALLIAADLLEKSPIPAMVKERYASYDSGMGKDFENGKLTLEQVVDFARKNGEPK STSGKQELYESIVNLYI
5586MI6_004	<i>Bacteroidales</i>	DNA	3	ATGGCAACAAAGAGTACTTCCCGAGATCGGGAATCAAAATCGAAGGCAAGGATTCC AAGAACCCGCTTGCATTCCATTATACAAATCCTGAGCAGGTCTGTCGGCAAGCCGATG AAGGACTGGCTCAAGTTCGCTATGGCATGGTGGCACACCCCTCTGCGCAGAGGTAGCGAC CAGTTCGGCGGACCCCAAGTCAATCCCTTGGAAACAAAGCTTCGGATCCCATCGCAAG GCCAAGCAGAAAAGTCCAGCGCGGTTTCGAGATCATGACAGAGCTCGGTATCGGATACAT TGCTTCCACGATGTAGACCTCATCGACGAGCGCCGCCACCATCGAGGAGTATGAGGCGGAT CTCAAGGAGATCGTCGCTTACCTCAAGGAGAGCAGGCGCCAGACCGGATCAAGTCTCTT TGGGACCGCCCAACGCTTTCGGTCAACAAGCGGTACATGAACGGCGCCTCCACCAACCTT GATTTCGAGCTCGACGCGCGGCTTGGTCCAGATCAAGAACGCCATGGACGCCACCATC GAGCTCGGCGGCGAGTGTATGCTTCTGGGCGGCGGCGGAGGCTACATGAGCCTCTC AACACCGACATGAAGCGTGAGAAAGCAGCATATGGCCACCATGCTCGGATGGCCCGGAC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586M16_004	<i>Bacteroidales</i>	Amino Acid	4	TATGACGGCGCAAGGCTTCAAGGCGACCTTCTCATCGAGCCCCAAGCCCATGGAGCCG ACCAAGCACCAGTATGACGTGACACCGAGACCGTTCATCGGTTTCTCCGTGCCAAGCGT CTTGACAAGGACTTCAAGGTCAACATCGAGTCAATCAGCCACCCCTCCCGGCCACACC TTGAGCATGAGCTCCAGTGGCGCCGATGCCGTCTCTCGGATCCATCGACGCCAAC CGCGCGACTATCAGAACGGCTGGATACCGACAGTCCCGATCGACTCTATGAGCTC ACCCAGGCATGATGCTCATCTCAAGAAATGGCGCCCTCGCGGGTACCAACTCGAC GCCAAGACCGTCCGAACCTCCAGCACCTGGACGACATCATCATGCCACCGTCAAGCGT ATGACATCATGGCACCGGCTCTCTCGTGGTCCGACGCTCTCACCAAGTCCGAGCTT CCCAAGATGCTCAAGGAGGCTTACGCTTCTCGACTCGGCAAGGGCAAGGAGTTCGAA GAGGCAAGCTCACTCTCGAGAGGCTCGTAGTACGCCAAGACCAAGGGCGAGGCCAAG GCCACCGGGCAAGCAGGAGCTCTACGAGACCATCTCATCATGTACATCTAA MANKEYFPEIGKIKFEGKDSKNPLAFHYYNPEQVVGKPKMDWLKFAMAWHHTLCAEGSD QFGGPTKSPFNKASDPIAKAKQKVADAGFEIMQKLGIGYYCFHVDVLLIDEPATIEEYDAD LKEIVAYLKEKQAOQTGILKLVGTANVFGHKRYMNGASTNPEDVAARAVQIKNMDATI ELGECYVFWGGREGYMSLINTDMKREKQHMATMLGMARDYARGKFGKGTFLIEPKMEP TKHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELQCAADAGLLGSDAN RGDYQNGWDTQFPIDLYELTQAMVILKNGGLVGGTNFDAKTRRNSTDLDIIIAHVS MDIMARALIVAADVLTKSELPKMLKERYASFDGKGEFEFGKLTLEQVVEYAKTKGEPK ATSGKQELYETIVNMYI
5749M11_003	<i>Bacteroidales</i>	DNA	5	ATGAATTTTATAAGCGGAAAAAGAAATTTTCCCCGGAATAGGAAAAGATCAGTTTGAA GGACCGGAGTCAAAGAACCCGATGGCGTTTCATATATACGAAAAAAGGTTGGTATG GGTAAACACATGAAGGATCATCTCGTTTIGCAATGGCTTACGGCATACGCTTTGTGCC GAAGGGGCGACCAATTTGGCGGTGGTACGAAAAACATTCCTCGAATGCTGCTGCCGAC CCGATCAGCCGTGCCAAATATAAGATGGATGAGCGGTTCCGAGTTTATGACAAAATGCAGC ATCCCTTATTACTGTTTCCATGATGTGACGTGGTGGAGGAGCTCCACGCTGGCTCAG TTTGAAAAGACCTTCATACGATGGTAGGCCATGCCAAGGCTTCAGCAGGCAACCGGA AAAAAAGCTTATGGTCTACTGCCAACGTTTCAGCAACAAACGCTATATGAACGGGCT GCCACTAATCCTGACTTCGCGCGTGGCTTGCGCGGTACGCGAGATCAAGAATGCGATC GATGCTGTATCGCGCTGGAACGTAATAAGTAACTATGTGTCTGGGCGGACCTGAAGATAT ATGGCTTTCTCAATACCGATATGAACGCGGAAAAAGACCACTGGCCCATGATGCTGACG ATGGCACGGACTATGGCGCAAGAACGGTTTCAAGAGTACTTCTGATCGAGCGGAAA CCGATGGAACCGACCAAGCATCAATATGATGTGACTCGGAACTGTAACTGGCTTCCTA CGTATATAGGCTGGATAAGACTTCGCTTGAATATCGAAGTAAATCATGCAACCTG GCCGACATACGTTGACGACGAATTCGAGGTGCTGCTGATCGCGGTATGCTGTGAGT ATCGATGCCAACCGTGTGACTACAGAAATGGCTGGATACCGACCAATCCCAGTGGAC ATCTACGAAGTACTCAGGCTTGGTGGTCAITCTGCAAGTGGTGGTGTGACACCGGC GGAACGAAGTTCGATGCCAAGACCGCGCAACTGACCGACCTGGACGATATCTCTCTG GCTCATATAGGTGGTATGGATGCGTTGCCCGTGCCCTGATCAGGCTGTGCTGCCATCT GAAACTCCGATTACACGAAGATGCGTGGCGACGTTACACGAGCTTCGATGGTGGCGAA GGCAAGCGTTTGAAGACGGTAACTTCTCTGGAAGACCTGCGTACGATCGCTCTCCG GACGGAGAACCGAAGATGGTACGGGCAACAGGAATTAATGAGATGATCTCAATTTA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5749MI1_003	<i>Bacteroidales</i>	Amino Acid	6	TACATATAA MNYFKEKEFFPGIGKIEFEGRESKNPMAFHYIDENKVVMGKTLKDLHLPFAMAYWHTLCA EGGDQFGGKTKFPWNAADPI SRAKYKMDAAFEFMTKCNIPYCFHDVDVDEAPTIAE FEKDLHTMVGHAKGLQOATGKLLWSTANVFSNKRYMNGAATNPYFPVACAGTQIKNAI DACIALDGENYVFWGREGYMLNNTDMKREKHLAMMLTMARDYGRKNFGKGTFLVEPK PMEPTKHQYDVTETVIGFLRHYGLDKDFAINIEVNHATLAGHTFEHELQAAADAGMLCS IDANRGDYQNGWDTQFPMDIYELTQAWLVILEAGGLTTGGTGNFDAKTRNSTDLDDIFL AHIGGMDAFARALITAAAILLENSDYTKMRAERYTSFDGGGKAFEDGKLSLEDLRLTALR DGEPKMVSQKELYEMIINQYI
5750MI1_003	<i>Bacteroidales</i>	DNA	7	ATGAATTACTTTAAAGGTGAGAAAGAGTTCTTCCCGGGAATCGGGAAATAGAGTTTGAA GGACGTGAATCGAAGAAATCCGATGGCTTTTCACTACTATGACGAGAAACAAGGTTGTCTATG GGGAAGACCTTGAAGGACCATCTCGCTTTTGGATGGCTTATGGCATACGCTGTGTGG GAAGCGCGGACCATCTCGGCGGCGGACGAGGCAATTCCTTGAATACCGGGCGGAT CGTATTTCCCGTGCCAAAGTATAAGATGGATGCTGCTTTTGAAGTTATGACGAAATGTAAAC ATCCCGTACTATTTTCCATGATGTGGAATGGTGAAGTCCGACACATCGGCGGAA TTTGAAGAAAGACCTTGATACGATGAGTGAATATGCCAAGCAGCATCAGGAGGCAACCGGG AAAAAATCTGTGTGCTACCGCAATGTGTCAGCAATAAACGTTATATGAACGGGCT GCCACAAATCCGATTTCCCTGCTGCTGCTGCGGGTACGCGAGATCAAGAATGCTATC GACGCTTGATTCGCTGCGGCGGAAATATGTTCTGGGCGGCTGCTGAAGGTAT ATGAGCTTGTGAACACCAATATGAACCGGAAAGGAACTCTCGCAATGATGTTGACG ATGGCTCGGATATGCGCGTAAGAACGGCTTCAAAGGTACTTCTCGTAGAGCCATAA CCGATGGAACCGACCAACATCAGTATGATGAGACACAGAAACTGTTATCGGCTTCCIG CGTATTACGGCTTGAACAGGACTTGGCATCAACATCGAAGTGAATCATGCTACATG GCTGGACATACATTCGAACATGAGCTTCAGGCGCTGCCGATGCCGATGCTGTGCAGC ATCGACGCAACCGCGGATACAGAAATGGTGGACACGAGATCAGTCCCGGTGAC ATCTACGAATGACACAGCGCTGGTGGTTATCTCTGAAGCGGTGGCTGACTACCGGT GGTACGAATTCGACGCAAGACGCGCGCACTCGACTGACCTGGACGATATCTTCTG GCACACATCGGTGGTATGATTCGTTGCGCGTCTTGTGAGGCGGCTGCCGATATATG GAACACTCCGATTACAAAAGATGCGTGCCGAAAGTATGCCAGCTTCGATCAAGCGAC GGCAAGAACTTCGAAGATGTTAACTCTTCTCGAGGACCTCCGACCAATCGCTCTTGGC TCCGCGAACCAGCAAAATCAGCGGGAACAGGAATGTATGAAATGATTAATCAACAG TACATTTAA
5750MI1_003	<i>Bacteroidales</i>	Amino Acid	8	MNYFKEKEFFPGIGKIEFEGRESKNPMAFHYIDENKVVMGKTLKDLHLPFAMAYWHTLCA EGADQFGGKTKAFPWNTGADRI SRAKYKMDAAFEFMTKCNIPYCFHDVDVDEAPTIAE FEKDLHTMVGHAKGLQOATGKLLWSTANVFSNKRYMNGAATNPYFPVACAGTQIKNAI DACIALDGENYVFWGREGYMLNNTDMKREKHLAMMLTMARDYARKNFGKGTFLVEPK PMEPTKHQYDVTETVIGFLRHYGLDKDFAINIEVNHATLAGHTFEHELQAAADAGMLCS IDANRGDYQNGWDTQFPMDIYELTQAWLVILEAGGLTTGGTGNFDAKTRNSTDLDDIFL AHIGGMDAFARALITAAAILLENSDYTKMRAERYASFDQDGKGFEDGKLSLEDLRLTALA SGEPKQISGKELYEMIINQYI

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5750M12_003	<i>Bacteroidales</i>	DNA	9	ATGAATTTATTTAAAGTGAAGGAGAGTCTTCCCTGGAATCGGGAAAAATAGAGTTTGAA GGACGTGAGTCGAAGATCCGATGGCTTTTCATATATTAATGATGAAGAAAAAGGTCGTAATG GGCAAGACCTTGAAAGATCACCTCGCTTGGCAATGGCTTACTGGCATACGTTGTGGCGG GAAGGGCAGACCAAGTTGGCGGTGGCACAATAATCATCCCTGGAAATACCGCAGCGGAT CGTATTTCCCGCGCTAAATATAAATGATGCTGCTTTCCAGTTTATGACCAAGTGCAGT ATCCGTACTATTTTCCATGATGTGACGTGGTGGAGGCTCCGGCACTGGCGGAA TTTGAAAAGGACCTGCATACGATGGTGGGATTCGCCAAACACACCAAGGAAAGCAACGGA AAGAACTGTTGTGTTACAGCAATGATTCGGGCATAACGTTATATGAACGAGCG GCTACCAATCTTATTTCCCGCTGTCGTTGTCGGTACGCGATCAAGAAATGCAATC GACGCTGTATCGAGCTGGTGGAGAACTATGTATTTCTGGGGCGGCGGAGGCTAC ATGAGCCTGCTGAACACCAATATGAACGTAAGGATCAITTTGGCCATGATGCTGACA ATGGCAGCGATATGCGCGCAAGATGTTTCAAGGGTACTTTCCTGGTGAATCTAAG CCGATGGAACCGACCAACATCAGTATGACGAGATACGAAACCGTATCGGCTTCCTG CGCCATATGGCTCGACAAGGATTCGCTATCAACATTCAGTGAACCATGCTACATIG GCGGCCATACATCGAACATGAACCTCAGGCTGCTGCGGATCGCGGTATGCTGTGCGAGC ATCGATGCAAAATAGGCGGACTATCAGAAIGGTTGGGATACGATACGATTCCTCCGTAAGC ATTTACGAACTGACACAGGCTGGTGGTTATCTCTGGAAGGGGGGACTGACAACCGGA GGTACGAACTTCGATGCGAAGACCGTCGTAACCTGACTGACCTCGACGATATCTTCTCTG GCCATATCGGCGGTATGGATTCTGTCACGTCGTTGATGGCAGCTGCCGATATCTCTG GAACATTCGATTAAGAAGATGCGTCCGACGTTACGCCAGCTTCGACACCGGCGAC GGCAAGAGTTCGAAGACGGCAACTCTTCTCGAAGACCTCGCGCAATTCGCTGCGC GCGACGAACCGAGCAGATCAGCGGCAAGCAGGATTTGATGAGATGATATCAATCAG TATATTTAA
5750M12_003	<i>Bacteroidales</i>	Amino Acid	10	MNYFKGEKFFPGIGKTEFEGRSKNPMAFHYVDENKVMGKTLKDHLPFAMAYWHFLCA EGADQFGGCTKSPWNATAADRIIRAQYKMDAAFEFMTKCSIPYCFHDVDVDEAPALAE FEKDLHTMVGFPAKHQEQATGKLLWSTANVFGHKRYMNGAATNPYFPVACAGTQIKNAI DACIELGGENYVFWGGREGYMSLNTNMKREKDHLMMLTWARDYARKNGFKGTFLEVSK PMPETHQYDADTETVIGFLRHYGLDKDFAINIEVNHAITLAGHTFEHELQAAADAGMLCS IDANRGDYQNGWDTDQFPVDIYELTQAWLVILEAGGLTTGGTNFDTAKTRNSTDLDDIFL AHIGGMDSFARALMAAADILEHSDYKMKRAERYASFDQDGKKEFDGKLLEDLRLTALA GDEPKQISGKQELYEMINQYI
5586M15_004	<i>Bacteroides</i>	DNA	11	ATGAAACAGTATTTCCGAAACATCTCCGCCATCAAGTTTGAGGGCGTCGAGAGCAAGAT CCCCTGGCTTACCGCTACTACGACCGGACCGCGTCTGTCATGGGTAAAGATGAGCGAA TGGTTTAAGTTCGCTATGCTGGTGGCACACCTCTCGCCGAGGGCTCCGATCAGTTC GGTCCCGGCAAAAGACCTTCCCTGGAAACCGCCGCCACCCCGTGGAGGCTGCCAAG GACAAGGCCGACGCTGGCTTCGAGATCATGCAGAAACTCGGCATCGAGTACTACTGCTTC CACGAGTTGACCTCGTGGCCGAGCTCCCGAGTGGAGGCTACGAGGAAGAACCTCAAG GAGATCGTGGCTTATCTCAAGCAGAAACAGGCTGAGACGGGCTCAAGTGTCTCTGGGGC ACTGCCAACGCTCTTCGACACAAGCGCTACATGAACGGAGCTCCACGAAACCCGACTTC GATGTCGTGGCAGCGCTATCGTGAGATCAAGAACGCCATCGATGCTACCATCGAGCTG GCGGGACCAACTACGCTCTCTGGGGCGGTCCGGAAGGCTACATGAGCCTGCTCAACACC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				GATATGAAGCGCGAGAGGAGCAGATGGCTACGATGTTGACGATGGCAGCGACTATGCC CGTTCTAAGGGATTCAAGGGCACGTTCTCATCGAACCCAAACCCATGGAACCCACGAAG CATAGTACGATCGGACACGAGACGGTTCATCGGATTCCTCGTTCATGGTCTCGAC AAGGATTTCAAGGTCAACATCGAGTCAACACGCCACGGCTGGCCGGACACACGTTGAG CATGAGCTGGCCTGGCGGTAGACCCGATATGCTCGGACGATCGATGCCAATCGCGGC GACTATCAGAACGGATGGACACGACGAGTCCCATCGACCTACCAACTCAGCAG GCTATGCTCAGATCATCCGCAACGAGGTTCAAGGACGGTGGACCAATTTTGACGCT AAGACGGCGGCAACAGCACCGCTCGAGGATATCTTCATCGCTCAGTAGCAGCCATG GACGCCATGGCCACGCCCTGTTCGGCTCGGATATCATCGAGAAGTCGCCCATCTGC ACGATGTTCAAGGAGCGTTACGCCAGCTTCGATGCCGGAAGGCAAGCGCTTCGAAGAA GGCAAGATGACCTCGAGGAAGCTACGAGTATGGCAAGAGTCCGGGAGCCCCAAGCAG ACCAGCGGAAGCAGGAGCTCTACGAAGCCATTGTCAATATGTATTGA
5586MI5_004	<i>Bacteroides</i>	Amino Acid	12	MKQYFPNLSAIKFEVESKNPLAYRYDRDRVVMGKMSWFKFAMCWHTLCAEGSDQF GPGTKTFWNAAADPVQAADKADAGFEIMQKLGIEYYCFHDVDLVAEPDVETYEKNLK EIVAYLKQKQAEETGIKLLWGTANVFGHKRYMNGASTNPDFDVARAIVQIKNAIDATIEL GGTNYVFWGREGYMSLLNTDMKREKHEMATMLTMDYARSKGFKGFLIEPKMEPTK HQYDADTEIVIGFLRAHGLDKDFKVNIEVNHAITLAGHTFEHELACAVDADMLGSI DANRG DYQNGWDTQFPIDHYELTQAMLIIRNGGFKDGTNFDKTRNSTDLEDFIAHVAAM DAMAHALLSAADIIIEKSPICITMVKERYASFDAGEGRFEEGKMTLEEAYEYKKGVEPKQ TSGKQELYEAIVNMY
5586MI202_004	<i>Bacteroides</i>	DNA	13	ATGGCAACAAAAGAGTATTTTCCCGAATAGAAAGATTAAATTCGAAGGTAAGAGAGT ATGAACCCGATGGCATCGTTACTACGATGCTGAGAAGGTAATCATGGGTAAGAAGAIG AAAGATTGGTTGAAGTTGGCTATGGCTGGTGGCACACTCTCTGCGCAAGGTGGTGAC CAATTCCGTGGCGGAACGAAACAATCCCTTGGAAATGGTACTCTGACGCTTTGCAAGCA GCTAAAAATAAATTGGATGCAGGTTTGAATTCATGCAAGATGGGTATCGAATACTAT TGCTTCCAGATGTAGACTGATTTCTGAAGGTGCAAGCATCGAAGAAATACGAAGCTAAC TTGAAGCTATCGTAGCTATGCAAAAAGAAAACAGGCTGAACTGGTATCAAGCTGTIG TGGGTACTGTCTAACGTAATCGGTATGACGATGATGACGCTGCTACCAATCCT GATTCGACGTTGTAGCAGCGCTGCTGTTCAGATCAAGACGCTATTGACGCTACTATC GAAGTGGTGGTTCAAACTAIGTATCTGGGCGGTGCGCAAGGTTACATGCTTTGCTG AACACTGACAGAAACGTGAAAAGAACACCTTGCAAGATGTTGACTATCGTCTGTGAC TATGCACTGCTCGTGGCTTCAAAGGTACTTCTGATTAGCCGGAACCGATGGAAACCG ACAAAACATCAGTATGATGATAGTACTGAAACAGTATCGGCTTCTGAAAGCTCACGGT TTGGATAAGGATTTCAAAGTAAACATCGAGTTAAATCAGCAACTTTGGCTGGCCATACT TTCGAACACGAACTGGGTAGCTGTTGACACGGCATGTAGGTTCTATCGACGCTAAC CGTGGTGACTACCAGAACGGTGGGATACAGCAATTCCTATCGATACTACGAACTG ACTCAAGCTATGATGAGATCATCCGCAACGGTGGTGGTAAATGGCGGTACTAACTTC GACGCTAAGACCCGTCTAACTTACCGACCTTGGAAAGTATCTTCATCGCTCACATGCA GGATGGATGCTATGGCAGCTGCTTGGAAAGTGCAGCTAAATTAATGGAAGAAATCTCT TATAAGAAAATGTTGGCTGATCGTTACGATCATTCGACGGTGGCAAGGTAAGGAATTC GAAGAAAGCAAAATGTTCTTTGGAAGATGTTGATGCTTATGCGAAAGCTAACGGCGAACCG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				AAGGAGTGGCTCAAGTTTGGCATGGCATGGTGGCACACACATCTGCGCAGAGGGTGGCGAC CAGTTTGGCGGGTACGAAGAAATTTCCCTGGAATGGAGATCCCGATCCGGTACAGGCC GCAAAAACAAAGTAGACCGCGGCTTGAATTCATGACCAAGATGGGATAGAGTATTTTC TGTTCCACGACGTGGATCTCGTCAGCGAGGAGCAACCATCGAGGAGTATGAGGCCAAC CTGAAGGAAGTGGTGGCTACATCAAGGAAAACGAGCCGAGACGGGGATCAAAAACCTC TGGGCACTGCCAACGTGTTGAGCCACGCGCTACATGAACGAGCGCCACCAACCCC GACTTCGATGTAGTGGCCGCGCGCTGCGATCAAGATCAAGAAATGCTATCGACGCCACGATA GCCTTAGTGGCACCACTACGTGTTCTGGGTGGCCGTGAAGTTTACATGAGCCCTGCTC AACACCGACGAGAGCGGAGAGGAGCATCTGGCAATGATGCTCGCATGGCCGCGAC TATGCGGTGCAAAAGGCTTACCGGCACCTTCTTATCGAGCCCAAGCCGATGGAGCCC ACCAAGCACCATGATGTAGACACCGAGACTGTGATAGGCTTCTCCGTGCCACGGC CTCGCAAGGACTTCAAGGTCAACATAGAGGTGAACCAACCCACCTGGCCGCGCATACC TTCGAGCTGAGTGGCAGTGGCGTGGACAAGCGTATGCTCGGAGCATCGACGCCAAC CGCGTGACTACCAAGACGGCTGGGATACCGACAGTCCCCATCGACAACTACGAGCTG ACCCAGGCGCATGATGAGATAATACGCAACGCGGCTTCGGCAACGGCGATGCAACTC GACGCCAAGACACGCCCAACTCCACCGACCTGGAGGATATCTTATAGCCCCACATAGCA GGCATGGACGCCATGGCCCGCCCTGCTCAGGCGACGAGAAGTGTGGAGAAAATCGCCC TACAGGAAGATGCTCGCGAGCGGTAGCACCGTCTTGTATCGCGCCAGGGAAGGCAATTT GAAGAGGCGCAATGTGCTCACCGACCTTGTGAGTATGCCAAGAGAGCATGGCGAGCCC ACACAGACTTCGGCAAGCAGGAACCTCTATGAGGCAATCGTCAATATGTATTCGCTAA MATKEYFPFGTKIKFEGVNSYNPLAYRYDAERIVLGKPMKEWLKFAWAWHHTLCAPGGD QFGGGTKNFPWNGDPPDVQAAKNKVDAGFEFMKMGIEYFCFHDVLDLVEAATIEEYEAN LKEVGYIKEKQAEETGKNLWGTANVFSHARYMNGAATNPDPDVARAAVQIKNAIDATI ALGGTNYVFWGGREGYMSLINTDQKREKEHLAMMLRMARDYARAKGFTGTFLIEPKMEP TKHQYDVTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSLDAN RGDYQNGWDTQFPIDNYELTQAMMQIRNGGFGNGGCFNDAKTRRNSTDLEDIFIAHIA GMDAMARALLSAAEVLEKSPYRKMFAERYAPFDAGQGKAFEEGAMSLTDLVEYAKEHGE TQTSKGQELYEAIVNNYC
5606M11_005	<i>Bacteroides</i>	Amino Acid	18	
5606M12_003	<i>Bacteroides</i>	DNA	19	ATGGCAACAAGGAATATTTCCCATATAGGAAGATCCAGTTCAAAGCCACGGAATCG TACGATCCGATGTCGTATCGTTACTATGACCGCGAGCGGTAGTTCGGGCAAGCCCAIG AAGGAATGGCTGAATTCGCCATGGCATGGTGGCACACATGTGCGCGAGGGCGGAC CAGTTCGGCGGGGAAGCAAGAGTCCCCTGGAACGAGGCGAGGAGCCATGACCATC GCCAAGCAAGAGGTGACCGCGGCTCGAGATCATGACGAGCTCGGCAICGAGTATTC TGCTTCCACGACATCGACTGATCGCGACCTGGCGACGACATCGAGACTATGAGAAC CGTATGCACGAAATCACCGCACCTGAAGGAAGATGGCGCCACGGGCATCAAGAAC CTGTGGGCGACTGCCAACGTGTTGGCCACGCGCTATATGAACGGCGCGCCCAAC CCCGACTTCGACGTGTGGCACGCGCATGTGCGATCAAGAACCGCATCGACGCCACC ATCGTCTAGGCGGTACAAACTATGTATTTCTGGGCGCGCGAGGGCTACATGAGCTG CTGAACCGCACCAAGCGCGAGAGACCTTGGCTACCATGCTGACCATGGCAGG GACTATGCGCGCGCAATGGCTTCACCGGAACGTTCTGTATCGAGCCCAACCCATGGAG CCGACGACGATCAGTATGATGTGGATACCGAGACCGTAACTCGGCTTCCTGAAGGCCAC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5606M12_003	<i>Bacteroides</i>	Amino Acid	20	AACCTGGACAAGGACCTCAAGGTGAACATCGAGGTGAACCATGCCACTCTGGCCGGCCAC ACATTCGAGCATGAGCTGGCAGTAGCCGTGGACAACGGCATGCTGGGAGCATCGAGCC AACCGGGGCACTATCAGAACGGCTGGGACACCGACCATGCTCCCATCGACAACATATGAG CTGACCCAGGCCATGATGCAGATAATCCGCAACGGTGGCTCGGCAACGGCGGTACCAAC TTTCAGCCCAAGACAGCTCGCAACTCCACCGACCTGGACACATCTTCATCGCTCATATC GCCGTATGGACGCTATGCCCGCGCTCCGTCAGCGCAGCCGACGCTTGAGAACTCG CCTTACAAGAAGATGCTGGCCGACCGCTACGCTTCATTCCAGACGGCGGAGGCAAGAA TTTCAGGAGGCAAGATGACTCTGAGGATGCTGGCTAGCTGGCTACGCCAAGAAATCCCGAA CCCGCTCAGACCAAGCGGCAAGCAGGAACCTCAGAGGCCATCATCAACATGTACGCTGA MATKEYFPHIGKIQFKGTESYDPMYSYRYDAERVVLGKPMKEWLKFAAWHHTLCAEGGD QFGGTTKFPWNEGEDAMTIAKQKADAGFEIMQKLGIEYFCFHDIDLIGLDGDDIEDYEN RMHEITAHLEKMAATGINKLWGTANVFGHARYMNGAATNPDFDVVARACVQIKNAIDAT IALGGTNVFWGGREGYMSLLNTDQREKEHLATMLTMARDYARANGFTGTFLIEPKPME PSKHQYDVTETVIGFLKAHNLKDFKVNIEVNHATLAGTFFHEHELAVDNGLMSIDA NRGDYQNGWDTDQFPIDNYELTQAMMQIIRNGGLNGGTNFDKTRRNSTDLDIDFI AGMDAMARAPLSAADVLEKSPYKRLADRYASFPDSGEGKFEKGMTLEDVVAYAKRNP PAQTSKQELYEAIINMYA
5610M13_003	<i>Bacteroides</i>	DNA	21	ATGGCAACAAAAGAAATTTTCCCGAGATTGGTAAATCAAGTTTGAGGGCCCGGAAGC CGCAATCCCTCGCATTCGCTACTACGGCCCGGAGAAAGTCGTTCTTGGCAAGAAAGT AAAGACTGGTCAAGTTTGGATGGCTTGGTGGACACACTGTGCGCCAGGGCAGCGAC CAGTTTGGTGGGACACCAAGCAGTTTCCGTGGAACTGCCAGTGAACCCATCGAGGCC GCCAAGGATAAGGTGGATGCCGATTTGAATTCATGACCAAGATGGGATGAGTACTTC TGCTTCCAGATGTGGATCTCGTCGCGAGGCGGCCACTGTCAGAGGATGAGGCTAAC CTCAAGACCATCGTCGCTACATCAAAAGAGAAACAAGCCGAGACCGGCATCAAGAACCTG TGGGCACAGCCCAACGTAATCGGACACAAACGCTACATGAACGGTGCAGCCACCAACCC GACTTGTATGTCGTGGCAGCGCCATCGTGCAATCAAGAACGCCATCGACGCCACATC GAGTTGGCGGCGACGAGTTACGTCCTTTGGGGCGGCGCGAGGGCCCATGAGCCTGCTC AACACCGACAGAGCGGAGAGGAGCACCTTGACGCAATGCTGACATGGCAGCGGAC TATGCCCGCGCAGTGGTTCAACGGCACCTTCCTCATCGAGCCCAAGCCCATGGAGCCG ACCAAGCACCATAATGATGTGGACACCGACCGCTCATCGGTTTCTGGGTGCCATGGT CTGGACAAGGACTTCAAGGTCAACATCGAGTGAACCAACGCTACACTGGCCGAGCACACC TTTCAGCGGAACTGGCAGTGGCGTCGACAAACGGTCTACTCGGCTCAATCGACGCCAAC CGTGGTGAATCAGATGGTGGGACACCGATCAGTCCCATCGACCACTATGATG GTTTCAGGGCATTTGAGATTAATCGCAATGGTGGTTTCCAGCAGCGTGCCACCACTC GATGCCAAGACCCCGCAACTCGACCGACCTCGAGGACATCTTCATCGCCCAATCGCC GCGATGGATGCCATGGCTCATCGCTGGAGAGTGTGCTCCATCATCGAGGAGTGGCC TACTGCCAGATGTCAGGATCGCTATGCTCATTTGATCCGGCATCGGCAAGGACTTT GAGGACGGAAGTTGACACTGGAACAAGCCTACGAGTACGATCAATCAATATGATTCATTA AAGCAGACGAGTGGCAAGCAAGAACTGTACGATCAATCAATATGATTCATTA MATKEFFPEIGKIFEGRESRNPLAFRYYGPEKVVGLGKMKDFKFAAWHHTLCAQGT QFGGDTKQFPWNTASDPMQAAKDKVDAGFEFMTKMGIEYFCFHDVLDLVAEAAATVEEYEA
5610M13_003	<i>Bacteroides</i>	Amino Acid	22	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				LKTI VAYTKEQAE TGIKNLWGTANVFGHKRYMNGAATNPDPDVVARAIVQIKNAIDATI ELGGTSYVFWGGREGHMSLINTDQREKEHLARMLTMARDYARGFNGTFLIEPKMEP TKHQYDVTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFERELAVADNGLLGSDAN RGDYQNGWDTQFPIDHYELVQMLQIIRNGGTDGGTDFDAKTRRNSTDLEDI FIAHIA AMDAMAHAALESAA SIIIEESPYCQMVKD RYASFDSGIGKDFEDGKLTLEQAYEYKQVGE P KQTSKGQELYESIINMY S I
5749MI2_004	<i>Bacteroides</i>	DNA	23	ATGGCAACAAAAAGAGTATTTCTCTGGTATAGGAAAGATTAAATTTGAAGGTAAAGAGAGT AAGAATCCGATGGCATTCGGCTATATGATGCCAATAAAGTAATCATGGGCAAGAGATG AGCGAGTGGCTGAAGTTGCCATGGCTGGTGGCACACATTTGCGCCGAAGGTGGTGAC CAGTTTGGTGGTGGAACAAAAGACTTCCCGTGAACACGATTCGGACAAACGCCGTAGAACGA GCCAACATAAAGTAGATGCCGTTTGAATTTATGCAGAAAATGGGCATCGAATACTAT TGCTTCCATGATGAGACCTCTGCACCTGAAGCTGTACCATTAAGAAATATGAAGCCAAAT CTGAAGGAAATAGTAGCCTATCCGAAACAGAAACAGGCTGAAACAGGTATCAAACTCTIG TGGGTACGGCAATGTATTTGGTACAAACCGTATATGAATGGTGTCTACCAATCCG GATTTGATGTAGTGGCTGCTGCTGTACAGATTAAAGATCGGATAGACGCTACAAT GAACTCGGTGGTAGCACTACGTTCTGGGGGGCCGCGAAGGTTATATGAGCTTGCTC AATACAGACCAGAAACGTGAGAAAGAGCATTTGGCACAAATGTTGACCAATGGCTCGTGAC TATGCTCGTCCAAAGGATCAAGGTTACCTTCTCTGGTGAACCCAAACCGATGGAACCA ACTAAACACAGTATGATGATACGGAACCTGTAATCGGCTTCTCAAGGCTCTATAAT TTGATAAGGATTTCAAGTAAATATTGAATAACCATGCTACATTTGGCCGGTCTACT TTTGAACAGAAATGGCTGTTGCCGTAGACACGATATGCTTGGCTCTATCGATGCCAAC CGCGGTGATATCAGAACGGTGGGATACAGCAGTCCAGTCCCATGACAACTTCGAGCTT ATCCAAGCATGATGCAGATTTTCGGGGTGGTGGCTTCAAGATGGTGTACAACTTC GACGCTAAGACTCGTCTGTAACCTCTACCGACCTGGAAGATATTTTTCATTCACACATCGCT GGTATGATGCTATGGCACGCTGCTTGGAAAGTGCAGCCAGTGTCTTGAGGAATCTCT TATAAGAAAATGTTGGCTGACCGCTATGCTCGTTCGATAGTGGCAAGGTAAGGATTT GAAGAAGGCAAGCTGACATGGAAAGACGTTGTAGTTTATGCCAAGCAGAAATGGCGAGCT AAACAGACACGCGGTAAAGCAGGAATTTGATGAGGCAATTTGAAATATGATGCTGTA MATKEYFPGIGIKIFEGKESKNPWAFRYDANKVIMGKMSWLKFAMAWHHTLCAEGGD QFGGGKTTFPWNDSDNAVEAANHKVDAGFEFMQKMGIEYYCFHDVLDLCEAATIEEYEA LKEI VAYPKQKQAE TGIKNLWGTANVFGHKRYMNGAATNPDPDVVARAIVQIKNAIDATI ELGGSYVFWGGREGYMSLINTDQREKEHLAQM LTMARDYARAKFGKGTFLVEPKMEP TKHQYDVTETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFERELAVADNMLGSDAN RGDYQNGWDTQFPIDNFELIQAMQIIRGGGPKDGGTDFDAKTRRNSTDLEDI FIAHIA GMDAMARALE SAAKLLIEESPYKMLADRYASFDSGKGKEFEEGKLTLEDVVVYAKQNGEP KQTSKGQELYESIINMY A
5749MI2_004	<i>Bacteroides</i>	Amino Acid	24	
5750MI3_003	<i>Bacteroides</i>	DNA	25	ATGGCAACAAAAAGAGTATTTCTCTGGAATAGGAAAGATTAAATTTGAAGGAAAAGAGAGT AAGAATCCGATGGCATTCGGTGTGCTACGATGCAGAAAAAGTTATCATGGGTAAAGAAATG AAAGATTTGGTGAAGTTTGAATGGCGTGGTGGCATACACITTTGTGCAGAAAGCGGTGAC CAATTCGGTGGCGGTACAAAGAGTTTCCCCGGAACGATATATGATAAAAATTCAGGCT GCTAAAAACAAGATGGATGCCGGTCTTTGAGTTTATGCAGAAAGATGGGATCGAATACTAT

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5750M13_003	<i>Bacteroides</i>	Amino Acid	26	TGTTTTACGATGTAGACCTCTGCACGGAAGCTGATACCAATTGAAGAAACGAAAGCTAAT TTGAAAGAAATCGTAGTTTACGCAAGCAAAACAGGAGTAGAAACAGGTATCAAAATTTATIG TGGGTACTGCCAATGTATTTCGTTCATGAACGCTATATGTAATGGTGGCTACCAACCCA GATTTGATGTGTAGCCCGTCTGTTCAGATTAAAGATGCAATTAAGTCTACCAAT GAACTAGGTGGCTTAAACTATGTCTCTGGGTGGACCGAAGGTTATATGTCTTTCTG AACACTGATCAGAAACGTGAGAAAGCAATCTTGACAAATGCTGACCATGCCCTGAC TATGCCGTGCCGTGGCTTCAAAGTACATTTCTGGTTGAACCCGAAACCGATGGAACCA ACCAACATCAATATGACGTAGATACAGAAACAGTTATCGGTTTTTTGAAAGCTCATGCT TTGGATAAAGACTTTAAAGTAAATATGAAGTAAATCATGCAACATTAGCCGGTCAACACA TTTGAACAGAACTGGCAGTGGCTTCGACACGGTATGCTGGGTTCTATTGACGCTAAT CGTGGTATGTCAAACGGTTGGGATACAGACCAATTTCCCATTTGATAACTATGAACGTG ACTCAAGCCATGATGAGATTAATTCGTAACGGTGGTTTGGCAATGGGTACGAATTTT GACGTAACAACTCGCCGTAATTTCTACTGATCTTGAGAGATATCTTTCTGCTCACATCGCA GGTATGGATGCTATGGCACGTGCATTTGAAAGTGGCGCAAGTTGTTGGAAGAAATCTCCC TATAAGAGATGCTGGCAGAACGTTATGCATCCTTTGACACGGTAAAGGTAAAGAGTTT GAAAGGGTAAAGTTGACCTTGGAGGATCTTGTGCTTTAIGCAAAAGTCAATGGCGAACCG AAACAAATCAGTGGTAAACAAGAAATGTATGAGGCAATTTGAAATATGTATTTGCTAA MATKEYFPGIGIKIFEGKESKNPMAFRCYDAEKVIMGKRMKDWLKFAMAWHHTLCAPFGD QFGGGTKSPRNDYTDKIQAANKRMWDAGFEFMQMGIEYYCFHVDVLCLEADTIEEYAN LKEIVVYAKQKQVETGKLLWGNTANVGHRYMNGAATNPDPDVVARAAVQIKNAIDATI ELGGLNVYFWGGREGYMSLINTDQKREKEHIAQMLTIARDYARARFKGTFLVEPKPMEP TKHQYDVTETVIGFLKAHALDKDFKVNIEVNHATLAGHTFEHELAVDNGMLGSLDAN RGDCQNGWDTQFPIDNYELTQAMMQUIRNGGLNGGNTFNDAKTRRNSTDLDGIFIAHIA GMDAMARALESAAKLLEESPYKMLAERYASFDSGKGKEFEGKLTLEDIVAYKVNQEP KQISGKQELYEAIVNMYC
5750M14_003	<i>Bacteroides</i>	DNA	27	ATGGCAACAAAGAGTATTTCCCGGAATAGAAAGATTAAATTCGAAGGTAAGAGAGC AAGAACCAGATGGCATTCGGTTATACGATGCCGATAAAGTAATCATGGGTAAAGAAATG AGCGAATGGCTGAAGTTCGCCATGGCATGGTGACACACCTTTGCGCAGAAAGGTGGTGAC CAGTTCGGTGGCGGAACAAAGAAATCCCTCGAACGGTGAGGCTGACAAGGTTCAAGCT GCCAAGAACAAATGGACCGCGCTTTGAATTCATGCAGAAATGGGTATCGAATACTAC TGCTTCCAGATGTAGACCTCIGGAAAGAGCGAGACCATTAAGAAATACGAAGCCAAAC TTGAAGGAAATCGTAGCGTATGCCAAGCAGAAACAAGAACCCGGCATCAAGCTGTG TGGGTACTGCCAACGTATTTCGGCCATGCCCGTACATGAATGGTGACGCCACCAACCCC GATTCGATGTGTGGACCGTGCAGCCGTCCTCAATCAAAAGCGCCATCGACGCTACTATC GAGCTGGAGGTTTCGAACTATGTCTTGGGCGGTTCGCAAGGCTACATGTCTATTGCTG AATACAGACCAGAGCTGAGAAAGAGCACCTTCGCACAGATGTTGACCATCGCCCGGAC TATGCCCGTCCCGTGGCTTCAAAGGTACCTTCTGATTGAACCGCAACCGATGGAACCT ACAAACACACAGTATGATGTAGACCCGAAACCGTTATCGGCTCTTGAAGGCCCAAT CTGGACAAAGATTTCAAGGTAAACATCGAAGTGAACCAACGCTACTTTGGCGGGCCACAC TTCGAGCAGCAACTCGCAGTAGCCGTAGACACGGTATGCTCGGCTCCATCGATCGCAAC CGTGGTGACTACCAAGCGCTGGGATACAGACAGTTCCCATTTGACAACTTCGAACGTG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5750MI4_003	<i>Bacteroides</i>	Amino Acid	28	ACCCAGGCAATGATGCAAAATCATCCGTAAACGGCGGCTTTGGCAATGGCGGTACAAATCTC GATGCCAAGACCCGTCGTAACCTCCACCGACCTGGAAGACATCTTCAATGCCCCACATCGCC GGTATGGACGCTGATGGCAGCTGCTGAAAGTGCAGCCAAATTTCTTGAAGAGTCTCCT TACAAGAAATGCTTGGCGACCGCTATGCTTCTTCCAGCTGGTAAAGCAAGGAATTC GAAGACGGCAAGCTGACACTGGAGGATTTGGCAGCTTACGCAAAAGCCCAACGGTGAGCCG AAACAGACCGCGGCAAGCAGGAGTTGATGAGGCAATCGTAAATATGTAATGCTGCTGA MATKEYFPGIGKIKFEKESKNPMFRYYDADKIVIMGKMWSEWLKFAMAWHHTLCAEGGD QFGGGTKKFPWNGEADKVAANKMDAGFEFMQMGIEYYCFHDVLDLCEEAETIEEYAN LKEIVAYAKQQAETGTLKLGWGTANVFGHARYMNGAATNPDPDVVARAAVQIKSAIDATI ELGGSNYVFWGGREGYMSLLNTDQREKEHLAQMLTIARDYARARGFEGTFLIEPKMEP TKHQYDVTETVIGFLKAHNLKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSDAN RGDYQNGWDTQFPIDNFELTQMMQIIRNGFGNGGTFNDAKTRRNSTDLEDIFIAHIA GMDVMARALESAAKLLEESPYKKMLADRYASFDSGKGKEFEDGKLTLEDLAAYKANGEP KQISGKQGLYEAIIVNMYC
5751MI4_002	<i>Bacteroides</i>	DNA	29	ATGACAAAAGAGTATTTTCCAAACCATTTGGTAAATTCAGTTTGAAGGTAAAGAGAGTAAG AATCCATTAGCATATCGTTATTAGCATGCTAACAAAGTAATAATGGTAAAAAGATGAGC GAATGGCTCAAGTTTGAATGGCATGGTGGCACATTTGTGTGCTGAGGTAGCGACGAG TTTGGTCTGGCACCAAGTCATTTCCATGGAACGCATCAACCGACCGTATGCAGGCTGCA AAAGATAAGGCTGACGAGGCTTCGAAATCATGCAAAAATGGGCAATCGAATACTACTGT TTCCATGATGTGACCTCATCGACCCAGCAGCATATTTCCAACATACGAAAAGAAATCTC AAGGAAATCGTTGCATACCTCAAGCAAAAACAGCGCCGAGCAGGTATCAAAATGCTATGG GGTACAGCTAACGTTATGGCCACAAGCGTTATATGACCGGTGATACCAATCCTGAC TTTGACGTTGTGACAGAGCTATCGTGCAATCAAGAAATGCTATCGATGCAACAATCGAA CTGGCGGCACGAACCTAGTATTTCTGGGGTGGTCGGAAGGTTACATGTCACGTCTCAAC ACCGACCAAAAGCGGAGAAAGAGCACATGGTACCATGTAGGAATGGCACCGTACTAT GCACGTTCTAAAGGCTTACTGGTACTCTCTTATCGAGCCAAAGCCTATGGAACCAACT AAGCATCAATACGACGTCGATACAGAAACITGTTATGGTTCTTCCATCAACAAATACGAGCTGTT GACAAGGACTTCAAGTAAATATCGAAGTGAACACACGCTACATTTGGCTGGCCATACCTTC GAACATGAATTAGCATGTGCTGTTGATGCAAGTATGCTTGGTCCATCGATGCTAACCGT GGTGATATCAGAAATGGCTGGGATACAGATCAGTTCCTTCCATCAACAAATACGAGCTGTT CAGGCCATGATGAGATTAATCGCAATGGTGGTTCGGTAACGGTGGTACAAACTTCGAC GCTAAGACAGCTCGTAATCAACCGATTTGGAAGACATCATCATTTGCTACAGCTTTCAGCT ATGGATGCTATGGCAGCTGCTCTTGAATGTCGTCAGACATCTTCAAACTCACCTTAT CCACAGATGGTGGCCAAACCGTTATGCAAGTTTTCAGCAAGGTATAGGTAAGATTTGAA GACGGCAAGCTCACCTCGAGCAAGTATACGAATATGTAAGACCGTGGCGCAACCGCT ATTACAAGCGGCAACAGGAGCTCTACGAAGCTATCGTTAAATATGTAATGCTGTA
5751MI4_002	<i>Bacteroides</i>	Amino Acid	30	MTKEYFPTGKIQFEGESKNPLAYRYDANKVIMGKMWSEWLKFAMAWHHTLCAEGSDQ FPGTKSFPWNASTDRMQAAKDKADAGFEIMQKLGIEYYCFHDVLDLDPADDDIPTYEKNL KEIVAYLKQQAETGTLKLGWGTANVFGHKRYMNGASTNPDPDVVARAIVQIKNAIDATIE LGGTNYVFWGGREGYMSLLNTDQREKEHMTMLGMARDYARSKGFTGTLIEPKMEPT KHQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDANR

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5751MI5_003	<i>Bacteroides</i>	DNA	31	GDMQNGWDTDQFPINNVELVQAMQIIRNGFGNGGTNFAKTRRNSTDLEIIIAHVSA MDAMARALECAADILQNSPIPOVMVANRYASFQKIGKDFEDGKLTLEQVVEYGKTVGEPA ITSGKQELYEAIVNMYC ATGGCTAACAAAGAAATTTTCCCGGTATTGGTAAAAATCAAAATCGAAGGTAAGAGAGC AAGAACCCCATGGCATATCGTTACTACGATGCTGAGAAAGTAGTCCTTTGGCAAGAAATATG AAAGACTGGTTCAAGTTTGGATGGCTTGGTGGCACACATGTGCGCGAGGGTAGCGAC CAGTTTGGTCCCGGCACTAAGICTTTCCCTTGAACACCGAGAGTGGCCCATGCGAGCA GCTAAGGACAAAGTTGAGCGTGGCTCGAGTTTCATGACCAAGATGGGTATTGAATACTTC TGCTTCCAGATGTAGACCTCGTTGCCGAGGCGACACTGTTGAGGAGTACGAGGCTCGC ATGAAGGAAATCGTTGCTTACATCAAGGAGAGGTGGCGAGACTGGCATCAAGAACTG TGGGTACAGCTAACGTATTGGCAACAAGCGCTACATGAACGGTGTCTACTAACCC GACTTTGACGTTGTGGCTCGGCTATCGTTCAAAATCAAGAACGCTATCGACGCTACTATC GAGCTCGGTGGTACGTCATACGTTCTGGGCGGCGGAGGGTTACATGAGCCTCTTG AACCCGACAGAGCGTGAGAAAGACACCTTGGCTACTATGCTCATAATGGCACCGGAC TAGCTCGCGCTAAGGTTTCAAGGTACATTCCTCATCGAGCCCAAGCCCATGGAGCCC ACAAAGCACAGTACGATGTTGACATGAGACTGTAATCGGCTTCTTAAGGCACACAAC CTTGACAGGACTTCAAGGTTAAGTTGAGTTAAGTTAACCCAGCAACTCTCGCTGGTCAACA TTTGAGCACGAGCTCGCTGTGCTGTGACGCTGGCATGCTTGGCAGCATCGACGCTAAC CGCGTGACTACCAAGAGCGCTGGATACAGCAACTCCCATCGCAACTTCGACCTC ACTCAAGCTATGCTCGAGATCATCGGCAACGATGGTTTCAAGGATGGTGTACAACTTC GACGCTAAGACTCGCGCAACAGACCGACCTCGAGGATATCTTCATCGCACACATCGCT GCTATGAGCGCTATGGCAGCTGCTCGAGAGCGCTGCTGAGTACTCGAGGAGTCACT CTGCCCCAAATGAAGAGGACCGCTATGCATCGTTTCGAGCTGGCATGGGTAAAGGACTTC GAGGACGGCAAGCTCACCTGGAGCAAGTTTACGAGTATGTAAGAGGTGGCGGAGCGCC AAGCAGACTAGCGGCAAGAGCTGTAAGGCTATCGCTCAACATGACGTATATAA
5751MI5_003	<i>Bacteroides</i>	Amino Acid	32	MANKEFFPGIGIKIFEGKESKNPWYRYDYDAEKVVLGKNMKDWKFAMAWHHTLCAEGSD QFGPGTKSPWNTAECPMQAADKVDAGFEFMKMGIEYFCFHDVLDLVAEADTVEEYEAR MKEIVAYLKEKVAETGLKNLWGTANVFGNKRVMNGAATNPDPDVARAIVQIKNAIDATI ELGGTSYVFWGGREGYMSLLNTDQREKEHLATMLTMARDYARAKGFKGTFLIEPRMEP TKHQYDVTETVIGFLKAHNLDDFKVNIENVNATLAGHTFEHELACAVDAGMLGSLDAN RGDYQNGWDTDQFPIDNFDLTQAMLEIIRNDGFKDGGTNDKTRRNSTDLEDIIFAHIA AMDAMARALESAVALEESALPQMKKDRYASFDAGMGKDFEDGKLTLEQVVEYGKKVGEF KQTSKGQELYEAIVNMYC
5751MI6_004	<i>Bacteroides</i>	DNA	33	ATGGCTAACAAAGAAATTTTCCCGGTATTGGTAAAAATCAAAATCGAAGGCAAGAAAGC AAGAACCCCATGGCATATCGTCATACGATGCCGAGAAGTAGTCCTTTGGTAAAGATG AAGGACTGGTTCAAGTTTGGATGGCTTGGTGGCACACACTCTGTGCGCGAGGGTAGCGAC CAGTTTGGTCCCGGCACTAAGICTTTCCCTTGAACACCGAGAGTGGCCCATGCGAGCT GCTAAGGACAAAGTTGAGCGCGGCTCGAGTTTCATGACCAAGATGGGTATTGAATACTTC TGTTTCCAGATGTAGACCTCGTTGCCGAGGCGACACCGTTGAGGAGTACGAAGCTCGC ATGAAGGAAATCGTTGGCTTACATCAAGGAGAGATGGCGAGCGGATCAAGAACTG TGGGTACAGCAACGTTATTCGGCAACAAGCGCTACATGAACGGTGTCTACTAACCC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5751M16_004	<i>Bacteroides</i>	Amino Acid	34	GACTTTGACGTTGTGGCTCGCGCAATCGTTTACATCAAGAACGCCATCGACGCTACTATC GAGCTCGGCGGTACCTTACGTGTTCTGGGGCGGCGGAGGGTTACATGACTCTCTTG AACACCGACAGAGCGGAGAGGAGCACCCTGGCTACCATGCTCACCATGGCTCGCGAC TATGCTCGGCTAAGGCTTCAAGGTACATTCCTTATCGAGCCCAAGCCATGGAGCCC ACCAAGCACAGTATGACGTGATACCGAGACCGTTATCGGCTTCTCAAGGCTCAGGCG CTGGACAAGGACTTCAAGGTGAACATCGAGTTAACCATGCAACTCTCGCGGCCACACA TTCGAGCACGAACTCGCTGCGCTGTGACGCTGGCATGTGGCGAGCATCGACGCTAAC CGGGCGACTACAGAACGCTGGATACCGACCATTCCTCCATCGAACACTTCGACCTC ACTCAGGCTAIGCTCGAGATCATCGCAACGGTGGTTTCAAGACGGTGTACAAACTTC GACGTAAGACCCGTCGAAACAGCACCGATCTTGAGGACATCTTTCATCGCTCACATCGCT GCTATGGACGCAATGGCACGCGCTCGAGAGCGCTGCCGCTGTGCTCGAGCAGAGCCCC CTTCCCCAGATGAAGAAAGACCGCTACGATCGTTCGATCGCGCATGGGCAAGGACTC GAGGACGCAAGCTCACTCTGGACAGGTTTACGAGTATGTAAGAGTAGGCGGAGCCCC AAGCAGACGACGCGGCAAGCAGGAAGTGTACGAGGCTATCCTCAACATGTATGTATAA MANKEFFPGTGKIKFEGKESKNPMAYRHYDAERVLGKMKDWFKFAMAWHHTLCAFGSD QFGPVTKSPFNQAECPMQAAKDVDAAGFEFMKMGIEYFCFHDVLDLVAEADTVEEYEAR MKEIVAYLKEKMAETGILKNLWGTANVFGNRYMNGAATNPEDVVARAIQIKNAIDATI ELGGTSYVFWGGREGYMTLINTDQKREKEHLATMLTWARDYARAKGFKGTFLIEPKMEP TKHQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSLDAN RGDYQNGWDTQFPIDNFDLTQAMLEIIRNGGFKDGGTNPDAKTRRNSDLEDFIAHIA AMDAMARALESAAAVLEQSPLPQMKKDRYASFDAGMKGDFEDGKLTLEQVVEYEGKKVGEF KQISGKQELYEAILLNMYV
5586M122_003	<i>Clostridiales</i>	DNA	35	ATGAAAGAAATATTTCTCTATGACAAAAAAGTTGAATATGAGGGCGGACCATCTAAAAAT CCATTTGGGTTTAAATACTATGATCCGAAAGAAATATAGCAGGCAAGCCTATGAAAGAA CATCTTAAATTTGCTATGAGTTGGTGCAATACATTTTGGGGCGGTGCAGACCCATTT GGCACAACTATGGACAGAACTACGCGCGGACTTACCCACCCATGGAAATTTGCAAAG GCAAAAGTAGATGCAGGCTTTGAGTTTATGCAAAAACCTCGGTATAGAGTATTTTGTITT CACGATGGGATATTGCACCGGAAGGAGCAGTTTGTGAAACCAAGAAAAAAGCTTTGG GAAATAGTAGATTATACAGCAAAAGATGAATGAAACAGGCATAAAGTTGCTTTGGGGT ACTGCAAACTGCTTTAATGCTCCAGTTATATGACGGTGCAGGAACATCAIGCAATGGG CACAGTTTGCAATGACGCGCGCACAGATAAAAAATGCAATGAAAGTACCGTTAAACIG GGTGAAAAGGCTATGTTTCTGGGCGGAAGAGAGGTTATGAAACACTTCTCAATACG GATATGGCACTTGAACATGACAAATGGCAAGACTATGCAATATGGCAGTGTATATGGC AGAACGATTTGTTTGGCGGTGATTTTATATGAAACCAAGCAAGCAAGCAAGCAAGCA CATCAATATGACTTTGACTCGGCAACTGTTTGGGATTTTGGAGAAAGTACGGTTTAGAT AAGGATTTTAACTTAATATAGAGCAAAATCATCGACACTGCAGGTCAATATGAA CATGAATGACTGTAGCGCGTATAAACGGTTCATTTGGCAGCATAGTCAAAATAGCGGC GATCCCAATCTTGGCTGGGATACCGCAATTTCCCAACAGATGTTTATTCGGCAACCTT TGATGCTTGAAGTAAAGCAGCGGCTTACAAAACGAGGCTTCTAAATTTTGTATGCA AAGGTCAGAAAGAGGCTCATTTACGTTTGTATGACATGTTTATGCATATATCAGCGTATG GACACTTTTGGCTGGGTTTATATAAGGCATATGAAATAATGAGGACGCGCAGATAGAT

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI22_003	<i>Clostridiales</i>	Amino Acid	36	GAAATTTGTAAGAAAGACATACGCAAGCTATAATACAGGCATAGGCAAGATATATATAGAT GGAAAGGCAAGCCCTTGAAAGTTTGAAAGAAATATATCTTTCAAAATGATAATGTTGTAATG CAAAGCGGCAGACAGGAATATCTGAAACAGTTTGAATATATATATTTGTTTAAAGCATAA MKEYFPMTKKKVEYEGASKNPFAPKYIDAERIAGKPMKEHLKFAMSWHILCAGGADPF GTTTMDRTYGGGLTDPMELAKAKVDAGFEFMQKLGIEYFCHDADIAPEGSSFVETKKNFW EIVDYIQQKMEITGIKLLWGATANCNAPRYMHGAGTSCNASHSFAYAAAQIKNAIEATVKL GGKGYVFWGGREGYETLLNTDMALELDNMRMLMHMAVDYGRSIGFDGDFYIEPKPEPTK HQYDFSATVLGFLRYGLDKDFKLNIEANHATLAGHTFHELTVARINGAFGSI DANSG DNLGWDTPFPPTDVSATLCMLEVIRAGGFTNGGLNFDKVRRGSTFFDDIVVAYITSGM DTFALGFTKAYEIIEDGRIDEFVKERYASYNTIGIKDII DGKASLESLEEYILSNDNVVM QSGRQYLETVLNNILFKA
1753MI4_001	<i>Firmicutes</i>	DNA	37	ATGAAAGAAATTTCCCAAAATATCTCTGAGATTAAATTCGAAGGAAAAAGACAGCAAAAAT CCITTTGGTTTCCATTACTACAAACCCAGACCCTAAATCATCTTAGGCAAAACCAATGAAGAA CACCTCCCATTCGCTATGGCTTGGTGGCAAAATCTTGGTGAACAGGTGTTGATATGTTT GGCGTGGCCCGAGCTGATAAGAGTTTCGGTCTAAAGTTGGCACAAATGGAAACACGCTAAG GCCAAAGTCGATGCCGTTTCGAATTCATGAAGAACTCGGTATCAGATATTTCTGCTTC CATGATGTTGACTTAGTCCAGAAATGTCAGATATCAAGATACAAACAAAGAAATAGAT GAAATCAGTGACTACATCTTAGAAAAGATGAAGGCACAGATATTAAGTGTATTATGGGGC ACCGCCAAATATGTTCTTAACCCAGCTTCTGCAATGGTGGGTTCCACAAACAGTSCG GATGCTTCGCTTTCGCGCTGCTCAAGTTAAGAAAGCCITAGATATCACCGTTAAATTA GGTGTAGGGTTACGCTTCTCGGGTGGTCTGTAAGGTTACGAAACATTACTCAATACA GACGTTAAATTCGAACAAAGAAACATGCTCGTTAATGAAGATGGCTGTGAATATGGC CGTCCATCGGTTTCAAAGCGGATTCATATCGAACCAAAACCAAAACCAATGA CACCAATATGACTTCGACCGCGTACAGCTATTGGCTTTAAGAGCCACGCTTAGAC AAAGACTTCAAGTTGAACATCGAAGCTAACCCGCTACATTAGCGGTCATACATTCCAA CAGGATTTAAGAAATCTCGCCATTAATGGTATGTTAGGTTCTATCGATCGTAACCAAGC GATATGCTCTTAGGTTGGATACAGACGAATCCCATTTGATGCTACAGTGGACACAA TGATGTACGAAGTCTTAAAGAAATGGTGGCTTACAGGTTGTTCAACTTIGACTCCAAA ACAGCTCGTCCATCCTACAAATGGAAGATATGTTCTTAGCTATATCTTAGGTATGGAT ACATTCGCTTAGGTTTAAATCAAAGCTGCTCAAATCATCGAAGATGGCGGTATTTGATCAA TTCATCGAAAAGAAATATCTTCCTTCGTTGAACAGAAATCGGTCAAAAGATCTTAAAC AACAGACAAGCTTAAAGAAATTAACGATTACGTTAGCTTGAAGATGGTGGTCTCCAGAACTT CCAGGTAGTGGTCTGTCGAAGAAATGCTCGAAGCCATCGTTACGATGTTCTTATTCGGCAAG TAA
1753MI4_001	<i>Firmicutes</i>	Amino Acid	38	MKEIFPNIPEIKFEGKDSKNPFATHYVNPDQIILGKPMKEHLPFAMAWHNLGATGVDMF GAGPADKSEFAGKVGTMWEHAKAKVDAGFEFMKLGIRYFCHDVLVPECADIKDTNKELD EISDYIILEKMGKTDIKLWGATANPFSNPRFCNGAGSTNSADVAFAAQVKKALDITVKL GGRYVFWGGREGYETLLNTDVKEQENIARLMKMAVEYGRSIGFGDFYIEPKPEPMK HQYDFDAATAIGFLRAHGLDKDFKLNIEANHATLAGHTFQHDLRISAINGLGSI DANQG DMLLGWDTPFPFDVYSATQCMYEVILKNGGLTGTFNFDSTRRPSTFMDMFLAYITLGM TFALGLIKAAQIIEDGRIDQFIEKKYSSFRETEIGQKILNNKTSLSKLSLDYACKMGAPEL

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
1753MI6_001	<i>Firmicutes</i>	DNA	39	PGSGRQEMLEAIVNDVLFQK ATGAAAGAAATTTCCCAATAATTCCTGAGATTAAATTCGAGGAGAAACACAGCAAAAT CCTTTGGCTTTCCATTACTACACACAGACCAATCATCTTAGGTAAACCAATGAAAGAA CACCTCCCATTCGCTATGGCTTGGTGGCACAACTCTGGTGAACAGGCTTIGATATGTT GGCGTGGCCAGCTGATAAGAGTTTCGGTGAAGTTGGACAAATGGAACACAGCTAAG GCCAAGTCGATGCCGCTTCGAATTCATGAAGAACTTGGTATCAGATATTTCTGCTTC CATGATGTTGACTTAGTTCAGAAATGTCAGATATCAAGATACAAACAAAGAAATTAGAT GAAATCAGTGACTACATCTTAGAAAGATGAAGGCACAGATATCAAGTGTATGGGGC ACGCCAATATGTTCTTAACCCACGTTTCTGCAATGGTGGGGTCCACAAACAGTGGC GATGCTTGGCTTTCGGCGCTGCTCAAGTTAAGAAAGCCTTAGATATCACCGTTAAATTA GGTGGTAGGGTTACGTTCTGGGGTGGTCTGTTAATGAAGATGGCTTGAATATGGC GACGTTAAATTCGAACAAAGAAACATTCGCTGTTAATGAAGATGGCTTGAATATGGC CGTCCATCGGTTTCAAGGCGATTCATATCGAACCAACCAAGAAACCAATGAAA CACCAATATGACTTCGAGCGGCTACAGCTATGGCTTCCTAAGAGCCACGGCTTAGAC AAAGACTTCAAGTTGAACATCGAAGCTAACACGCTACATAGCGGTCATACATTCCAA CACGATTAAAGAAATTCGGCCATTAATGATGTTAGTTCATGATGCTAACCAAGGC GATATGCTTAGTTGGATACAGACGAATTCCTATTCGATGCTACAGTGGACACAA TGATGTACGAAGCTTAAAGAAATGGTCTTACAGGTGGTTTCAACTTGACTCCAAA ACAGTGTGCTTACATACACAAATGAAGATATGTTCTAGCTATATCTAGGTATGGAT ACATTCGCTTAGTTTAAATCAAGCTGCTCAATCATCGAAGATGGCGTATGATCAA TTCAATCGAAAGAAATATCTTCCCTCCGTGAACAGAAATCGTCAAGAGATCTTAAAC AACAGACAAAGCTTAAAGAAATATCCGATTACGCTTGAAGATGGGTCCAGAACTT CCAGGTAGTGGTCTCAAGAAATGCTCGAAGCCATCGTTAAGCATGTTCTATTCGGCAAG TAA MKEIFPNIPEIKFEGKSKNPFAPHYINPDQILILGKPMKEHLPFAMAWHNLGATGVDMF GAGPADKSFQAKVGTMEHAKAKVDAGFEFMKKLIRYFCFHDVDLVPECADIKDTNKELD EISDYILLEKMGTDIKLWGTANMFNPRFCNGAGSTNSADVFAFAAQVKKALDITVKL GGRGYFVGREGYETLLNTDVKEQENIARLMKMAVEYGRSIGFGDFYIEPKPEPMK HQYDFDAATAIGFLRAHGLDKFLNIEANHATLAGHTFQHLRISAINMGLSISANQOG DMLLGWDTDFPFVDVYSAIQCMYEVILKNGGLTGFNFDSKTRRPSYTMEDMFLAYILGMD TFALGLIKAAQIIEIDGRIDQFIERKYSFRETETIGQILNNKNTSLKELSDYACKMGAPEL PGSGRQEMLEAIVNDVLFQK ATGGAATATTTCCCTTTTCGTCAATCGGTCCATACAGAGGACCAACCTCAACTGAACCA TTCCGTTTCAAGTACTACGATGCCAACCGTGTGTTCTTGGAAACCAATGAAAGATGG ATGCCATTCGCTATGGCTTGGTGGCAACCTCGCGCTGCCGATCCGACATGTTCCGGC GGCAACACCATGGACAGTCTTGGGAGTCGATAAGAAAGACCCCAATGGGCTATGCC AAAGCCAAAGTTGATGCCGCTTCAAGATTCATCGAAGATGGGATCGAATACCTACTGC TTCCAGATGTCGACCTCGTCCAGATCGGACGACATCACCGTTATGTACAGAGATC GATGAGATCGGTGATTTACCTTCTCAAGAAACAGAGGAAACCGGTATCAAGCTTCTTGG TCAACCGCAATGCCTCGGACACCGCGTTTTCATGAACGCTGTTCCAGCAACTCC GCCGAAGTCTATTGCTTCGCCCGCCGCCAGATCAAGAAAGCTCTTGAGCTCTGCGTCAAA
1753MI6_001	<i>Firmicutes</i>	Amino Acid	40	
1753MI35_004	<i>Firmicutes</i>	DNA	41	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
1754MI9_004	<i>Firmicutes</i>	Amino Acid	44	GGCTTCATTAAAGAAAAGTATTCAAGCTACGAAAGTGGAATTGGTAAGAAGATCCGCGAC AAGCAGACAACTTTGGAAGAGCTTCTGCCCCGCGCAGAAATGAAAAGCCATCTGAT CCAGGTTCAGGCCCGGAGGAAATATCTGGAAGAGTGTGTTAACAATATCCTCTTTTCGCGGA TAA MSEFFKNIPEIKFEGKDSKNPWAFFKYNNPELITMGKKMSEHLFFAMAWHHNLGANGVDMF GSGTADKSFQIPGTMHAKVDAKAGIEFMKKLIGIKYYCWHVDVLVPEDPNDINVTNKRL DEISDYILEKTKGTDIKCLWGIANFSPNFRMNGAGSTNSADVCFCAAQVKKALEITVK LGGRYVFWGGREGYETLLNTDVKLEQENIANLHMAVDYGRSIFGKGFYIEPKPEPM SHQYDFDAATAIGFLROYGLDKDFKWNIEANHASLANHTFQHELISRINGMLGSVDANQ GNPILGWDIDNFPWNVYDATLAMYEVLLKAGGLTGGNFDSKNRRPSNTFEDMFHAYIMG DTFALGLIKAAEIIEDGRIDGFIKEKYSSESGIGKKIRDKQTLEELAAARAAEMKKPSD PGSGREYLEGVVNNILFRG ATGAGCGAGTTTTTTAAGAATAITCTCAAAATAAAATACGAAGGAAAAAGATAGCAAAAAT CCCTGGGCATTCAGATATTACAATCTGAATTGACAATCATGGGTAAAAAGATGAGCGAA CATCTCCATTCGCAATGGCATGGTGCAATACTTGGCGCAACGGCGTIGATATCTTT GGTCAGGGAACAGCAGACAAAGICTTTCGACAGATTCTTGAACTATGGAGCATGCAAAAG GCTAAGGTTGATGCTGGTATAGATTATGAAGAAGCTCGGAATCAATATTACTGCTGG CAGCATGTTGACCTTGTCTGAGGATCCAAACGATATCAACGTAACTAACAAACGCTIG GACGAAATTTAGATTACATCTTGAAGAGCAAAAGGACAGACATTAAGTGTCTCTGG GGAACGCAAAACATGTTGGTAACTCGCTTATGAACCGTGACGGCTACACAACTCT GCTGACGTTTACTGTTTCTGCGCTCAGGTAAAGGCTCTTGAGATTACTGTAAG CTTGGTGGCGGAGGTTATGTTTTCTGGGIGGCGGAGGTACGAACTCTTCTCAAT ACAGACGTAAAACTTGAACAGGAAATATCGCAACCTCATGCATATGGCTGTTGATTAT GGCCGCTCAATCGGTTCAAGGGAGACTTCTACATCGAGCTAAGCCAAAGGAGCCCAATG AGCCATCAGTATGATTTGATGTCGAACAGCAATCGGCTCTCCGCGAGTATGGCTC GACAAAGATTTAAGATGAACATCGAAGCTAACATGCCCTACCTTGCAATCACACTTC CAGCAGAGCTTTGATCAGCCGCAATAACGGAATGCTTGGTCTGTAGATGCAAACTCAG GGAATCCAATCTTGGCTGGGATACAGATAACTCCCAGGAATGTTACGATGCAACT CTGGCAATGTACGAAGTTCTCAAGCTGGCGGTCAACAGGTGGCTTCACTTTGACTCA AAGAACCGTCGCCATCAATACCTTTGAAGATATGTTCCAGCTTATATCATGGGTAG GATACTTTTGCCCTTATTAAGGCTGAGAAATATTGAAGCGCAGAAATGAC GGCTTCATCAAGAAAAGTATTCAAGCTTTGAAGTGAATGGTAAAGAATTCGTGAC AAGCAGACAAGTTTGGAAAGCTTCAGCTCGTGGCTGAAATGAAAAGCCATCTGAT CCAGGTTCAGGCCCGGAGGAAATACCTCGAAGGAGTGTGTTAACAACATCCTCTTTTCGCGGA TAA
1754MI22_004	<i>Firmicutes</i>	DNA	45	ATGAGCGAGTTTTTTAAGAATAITCTCAAAATAAAATACGAAGGAAAAAGATAGCAAAAAT CCCTGGGCATTCAGATATTACAATCTGAATTGACAATCATGGGTAAAAAGATGAGCGAA CATCTCCATTCGCAATGGCATGGTGCAATACTTGGCGCAACGGCGTIGATATCTTT GGTCAGGGAACAGCAGACAAAGICTTTCGACAGATTCTTGAACTATGGAGCATGCAAAAG GCTAAGGTTGATGCTGGTATAGATTATGAAGAAGCTCGGAATCAATATTACTGCTGG CAGCATGTTGACCTTGTCTGAGGATCCAAACGATATCAACGTAACTAACAAACGCTIG GACGAAATTTAGATTACATCTTGAAGAGCAAAAGGACAGACATTAAGTGTCTCTGG GGAACGCAAAACATGTTGGTAACTCGCTTATGAACCGTGACGGCTACACAACTCT GCTGACGTTTACTGTTTCTGCGCTCAGGTAAAGGCTCTTGAGATTACTGTAAG CTTGGTGGCGGAGGTTATGTTTTCTGGGIGGCGGAGGTACGAACTCTTCTCAAT ACAGACGTAAAACTTGAACAGGAAATATCGCAACCTCATGCATATGGCTGTTGATTAT GGCCGCTCAATCGGTTCAAGGGAGACTTCTACATCGAGCTAAGCCAAAGGAGCCCAATG AGCCATCAGTATGATTTGATGTCGAACAGCAATCGGCTCTCCGCGAGTATGGCTC GACAAAGATTTAAGATGAACATCGAAGCTAACATGCCCTACCTTGCAATCACACTTC CAGCAGAGCTTTGATCAGCCGCAATAACGGAATGCTTGGTCTGTAGATGCAAACTCAG GGAATCCAATCTTGGCTGGGATACAGATAACTCCCAGGAATGTTACGATGCAACT CTGGCAATGTACGAAGTTCTCAAGCTGGCGGTCAACAGGTGGCTTCACTTTGACTCA AAGAACCGTCGCCATCAATACCTTTGAAGATATGTTCCAGCTTATATCATGGGTAG GATACTTTTGCCCTTATTAAGGCTGAGAAATATTGAAGCGCAGAAATGAC GGCTTCATCAAGAAAAGTATTCAAGCTTTGAAGTGAATGGTAAAGAATTCGTGAC AAGCAGACAAGTTTGGAAAGCTTCAGCTCGTGGCTGAAATGAAAAGCCATCTGAT CCAGGTTCAGGCCCGGAGGAAATACCTCGAAGGAGTGTGTTAACAACATCCTCTTTTCGCGGA TAA
1754MI22_004	<i>Firmicutes</i>	Amino Acid	46	MSEFFKNIPEIKFEGKDSKNPWAFFKYNNPELITMGKKMSEHLFFAMAWHHNLGANGVDMF GSGTADKSFQIPGTMHAKVDAKAGIEFMKKLIGIKYYCWHVDVLVPEDPNDINVTNKRL DEISDYILEKTKGTDIKCLWGIANFSPNFRMNGAGSTNSADVCFCAAQVKKALEITVK LGGRYVFWGGREGYETLLNTDVKLEQENIANLHMAVDYGRSIFGKGFYIEPKPEPM SHQYDFDAATAIGFLROYGLDKDFKWNIEANHASLANHTFQHELISRINGMLGSVDANQ GNPILGWDIDNFPWNVYDATLAMYEVLLKAGGLTGGNFDSKNRRPSNTFEDMFHAYIMG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
727MI1_002	<i>Firmicutes</i>	DNA	47	DTFALGLTKAAEIIEDGRIDGFIKEKYSFESGIGKIRDKQTSLEELAAARAAEMKKPSPD PGSGREEYLEGVVNNILFRG ATGATATTTGAAAAATATTCGCGCAATTCCTTATGAGGGTCCGAAGAGACAAAATCCGCTG GCGTTTAAATCTATGATCCGACAAAGATCGTATGGAAGCCCATGAAGGAGCATCTIG CCCTTTGCAATGGCTGGTGGCACAACCTTGGCGGGCCGGAACCGATATGTTCCGGCGC GATACCGCGACAAATCCTTCGGTGGTAAAGGACAAATGGAGCATCCCAAAGCGAAA GTCGATGCGGCTTTGAGTTCATGACAGAGCTGGGATCCGCTATTTCTGCTTCCATGAT GTGGATCTTGTTCGGAGGGGATGATATAAAGGAGACCAACCGCGCTTGGACGAGATC AGCGATTACATCCTTGAAAAGATGAAGGGCACCGATATCAAGTCCCTTTGGGGCACGGCC AAATGTTCTCAATCCGCGCTTATGAACGGCGCAGGCTCTCCAAATTCGCCGATGTA TTCGCTTTGGCGACACAGGCCAAGAGGCTTGGATCTGACCTCAAACTCGGCGGG CGCGCTATGTTCTTGGGGGAGCTGAGGGCTATGAGACACTTCTCAATACCGACATG AAGTTCGACGAGGAGATATCGGAAGCTCATGCATATGGCTGCTGATACGGCGCAGC ATAGGCTTTACCGGTGATTTCTATATCGAGCCCAACCCGAAAGAGCCGATGAACACAG TATGATTCGATGCAGCCACTGCGATAGGCTTCTCCGCGAGTACGACTCGATAAGGAC TTCAGCTCAACATCGAGGCAACACCCACACTGGCAGGTACACTTCCAGACGAT CTGCTGTTTCCGCAATAACGGAATGCTGGGAGCATTGACGCCAACACAGGGCGATATG CTCCTCGGCTGGATACCGACGAGTCCCGTTCATATGATATGATGCGACCATGTGCATG TATGAGGTCTCAAGTCAGACGGCTCACCGGGCTTTAACTTCGACTCCAAATCAGC CGCCGAGCTATACGTCGAGGATATGTTTACAAGCTATATCTCGGATGGACATTTT GCCCCGCTTCTGAAAGCGCGGAGCTTATCGAAGACGGAAGGCTTGACGCTTCTGTC AAAGAACGCTATTCAGCTATGAGCGGCGATCGGCGAAGATCCGAGCGGAGAAACC GATTTGAAGGAATTGGCGGAATATCGGACTCCCTCGAGCCCCGAACTTCGGGCGAGC GGAAACAGGAACAGCTCGAGAGCATAGTAAATCAGATACTTTCCGGATAA MI FENI PALPYEGPKSNPLAFKFYDPDKIVMGKPMKEHLPFAMAWHNLGAAAGTDMFGR DTADKSFQAVKGTMEHAKAKVDAGFEFMQKLGTRYFCFHDVDLVPEADDIKETNRLDEI SDYILLEKMGTDIKCLWGTANMFSPRFMNGAGSSNSADVFAPAAQAQAKALDLTVKLGG RGYVFWGGREGYETLLNTDMKFEQENIAKLHMAVDYGRSIGFTGDFYIEPKPEPMKHQ YDFDAATATGFLRQYGLDKFLNIEANHATLAGHTTFQHLRVSAINGLGSDINQGD LLGWDTEPFPNVYDATCMYEVLSKSDGLTGGNFNFDKSRPSYTVEDMFTSYILGMDTF ALGLLKAELIIEGRDAFVKERYSSYESGIGAKIRSGETDLKELAEYADSLGAPELPGS GKQEQLESIVNQILFG ATGAGCGAGTTTTCGCGAGCATTCCTCAAAATTCCTTTGAAGGCAAGACAGCGGCAAT CCCCGGGCTTCAATATACTACGACCGCACAGGATGATCTGGGCAAGCCCATGAAGGAG CACTTCCCTTCGCCATGGCTGGTGGCAACAACCTGTCCGCGCGGGCACCGATATGTTT GGCCGGGACACCGCCGACAAGTCTTCGCGCAGGTCAAGGCGACCATGGAACACGCGCAAG GCCAAGGTGGACCGGGCTTTGAGTTCATGAAGAAGCTGGCATCCGCTACTTCTGCTTC CACGACGTGGACATCGTGCCCGAAGCCGACGACATCAAGGAACCAACCGCGCTCTGGAC GAGATCTCGGACTATCTCTGGAGAAATGAAGGACCCGACATCCAGTGGCTGTGGGGC ACCGCCACATGTTCCGCAACCCCGCTATATGAACGGCGGGCGAGCTCCAACCTCGCC GACGTACTGCTTCGCGCGGGCCAGATCAAAAAGGCCCTTGGACATCACCGTGAAGCTG
727MI1_002	<i>Firmicutes</i>	Amino Acid	48	
727MI9_005	<i>Firmicutes</i>	DNA	49	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				GGCGGCAAGGCTACGTGTTCTGGGCGGCCGAGAGGCTACGAGACCTGCTGAACACC GATATGAAGTTCGAGCAGGAGAACATCGCCCGCTGATGACATGGCGTGGACTACGGC CGCAGCATCGGCTTACCGCGGATTTCTACATGAGCCCAAGCCCAAGGAGCCCATGAAG CACCAGTACGACTTCGAGCGCCGCCACCGCATAGGCTTTTTCGCGCAGTACGGCCTGGAC AAGGATTTCAAGCTGAACATCGAGTCCAACCCAGCCACCTCGCGGGCCATACCTTCCAG CAGACCTCGCGGCTTCCGCCATCAACGGCATGCTGGGCTCCATCGACCAACCCAGGGC GACTACCTGCTGGGCTGGGATACCGACGAGTTCCCTACAGCTATACGAGACCAACATG TGATGTAGAGGTGCTCAAGCCCGGAGTCTACCCGGGCTTCAATTCGACGCCAAG AACCGCGTCCAGCTACACCCCGGAGATATGTTCCAGCTACATCCTTGGATGGAC AGCTTCGCCCTGGCCTGATCAAGCCCGGAGTCTATCGAGACGGTGGCCTGGACGCC TTCGTCCGGACCGTACCGAGCTGGGAGACCGGCATCGCGGATAGATCCGCAAGGGC GAGACCACTGGCGGAGTGGCGAGTACCGCGCCGATGGCGCGCCGCGCTGCC GGCAGCGCGCCAGGAATACCTGGAGGCGGTGGTCAACATATCCTGTTCAATATA MSEFFASLPKIPFEGKDSANPLAFKYDDADRMILGKPMKEHLFPAMAWHNLCAAGTDMF GRDTADKSFQVKGTMFHAKAKVDAGFEFMKKLIGRYFCHFDVDIVPEADDIKETNRRLD EISDYILEMKMGTDIOCLWGTANMFGNPRYMNAGSSNADVYCFAAQIKKALDITVKL GGKGYVFWGGREGYETLLNTDMKEFEQENIARLMHMAVDYGRSIGFTGFYIEPKPEPMK HQYDFDAATAIGFLROYGLDKFKLNIESNHATLAGHTFQHDLRVSAINGMLGSI DANQG DYLLGWDTDEFPPYSVYETIMCMYEVKAGLTGGFNFDKNRRPSYTPEDMFHAYILGMD SFALGLIKAAELIEDRLDAFVRDRIQSWETGIGDKIRKETTIAELAEYARMGAPALP GSGRQYLEGVWNILFK
727M19_005	<i>Firmicutes</i>	Amino Acid	50	
727M127_002	<i>Firmicutes</i>	DNA	51	ATGAAGACCTATTTCAAAAAATCCCGTGTATCCCTACGAGGACCGAAGTCCCAGAAAT CCGCTGTGTTCAAATCTATGACGGGACCGCATCGTTCTCGGCAAGCCCATGAAGGAG CATCTGCCCTTCGCCATGGCTGGTGCAATCTGGTGTCTCCGGACCGACATGTTT GGCCGGATACCGCGACAAGTCTTCCGAGCGGAGAGGACCCATGGAGCATGCCAAG GCCAAGTGGACGCTGGCTTCGAGTTTATGAAGGTGGCATCCGGTATTTCTGCTTC CATGACGTGGATCTGGTCCCGAAGCGGACGACATCAAGGAGACCAACCGCGCTCTGAT GAGATCAGCGACTACATCCTCAAGAAATGAGGCGACGGATATCAAGTCCCTCTGGGC ACCGCAACATGTTCCGCAATCCCGGTTTATGAACGGCGGGGAGCTCCAACAGCGCG GACGTGTTCTGTTTGGCGCGCCAGGTGAAGAGGCCITGGACATCACCGTCAAGCTG GGCGCGGGGCTATGTTCTGGGCGCGGTGAGGGTATGAGTCCCTGCTGAACACG GACGTGAAGTTTGACAGGAGAACATCGCAAGTCTATGACGCCAGCCCAAGGAGCCCATGAGC CGCAGATCGGCTTACCGGCGATTTCTACATCGAGCCAGCCCAAGGAGCCCATGAAG CACCATGCGACTTCGATCCGCGACCGCATCGGCTTCTCAGGAGTACGGCGCTCGAT AAGGACTTCAAGATGAACATGAAGCCAAACCGGACCCCTGGCGGCGCACACCTTCAG CAGACCTCAGGATCAGCGCCATCAACGGGATGCTGGCTCCATCGACGCCAACCCAGGGC GACCTCTCTGCTGGATGGACACCGACGAATTCCTCTCAACGCTCTATGAGGCCACCATG TGATGTACGAGGTCTCAAGCGCGGCGCTCACCGGCGCTTCAACTTCGACTCAAG AACCGCGCTCCCTCTACACCATGAGGATATGTTCCACGCTACATCCTGGGATGAC ACCTTCGCGCTGGTCTCTCAAGCGCGGAGTCTCAGGACCGGTGGATCGACAAA TTCGTGGAGGAGCGCTACGCCAGCTACAAGACCGGCATCGGGCCCAAGATCCGTTCGGC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
727MI27_002	<i>Firmicutes</i>	Amino Acid	52	GAGACCAGCTTCAGGAGCTGGCCGCTATGCCGACAAAGTTGGGGCGGCTGCCCTCCCGGAGGAGGAGTACCTGAGAGAGCATCGTCAACAGGTCCTTCGGGATGTGAMKTYFKTPIVPIYEGPKSQNPISFKFYDADRIVLGPKMKHLPFAMAWHNLGAAGTDMFGRDTADKSPGAEKGTMEHAKAKVDAGFEFMKKVGIYFCFHDVDLVPEADDIKETNRRLDELSDYILKKMGTDIKLWGATANFNPRFMNGAGSSNADVFCFAAQVKKALDITVKLGGRYVFWGGREGYESLINTDVKEQENIAKIMHLAVDYGRSIGFTGDFYIEPKPEPMKHQYDFDAATAIGFLROYGLDKDFKMNIEANHATLAGHTFQHDLRISAINMGLGSI DANQGDLILGWDDDEFFPNVYEATMCMYEVILKAGLITGGFNFDSKNRPPSYTMEDMFHAYILGMDTFALGLLKAELIEDGRIDKFVEERYASYKTGTGAKIRSGETTILQELAAYADKLGAPALPGSGRQYLESIVNQVLFMG
1753MI2_006	<i>Neocallimastigales</i>	DNA	53	ATGGCTAAAGAGTATTTCCAGAGATTTGGCAAAATCAAGTTTGAAGGCAAGGACAGCAAA AACCCAAATGGCTTTCACACTACTATGACCCCGAGAAGGTGATCATGGGCAAGCCTATGAAA GACTGGCTCCGCTTCGCTATGGCATGGTGGCACACCCCTCTCGCAGAGGTGGCGACAG TTCCGTGGCGGCACTAAGAAAGTTCCTTGGAAACAACGGGCTGACGCTAGAAATCGCA AAACAGAAAGGCTGACGAGGTTTCGAAATCATCAGAAAGCTCGGCATCCCATATTTTCG TTCCACGAGCTGGACCTCGTGTCTGAGGGCGCATGTAGAGAGATGAGGCTAACCTC AAGGCTATCAGACATACCTCGCTGTGAAGATGAAGAAAACAGGCATCAAGCTCCTGTGG TCTACTGCCAACGTAATTCGGCAACGGCGCTACATGAACGGTGTCTTACCAACCCCTGAC TTCAGCGTGTGTCTCGGCTATCGTGCAGATTAAAGAACGCTATCGACGTGGTATCAAG CTCCGGCTGAGAACTAGGTGTCTGGGGGGGACGCGAAGGCTACATGAGCCTCCTCAAC ACCGACCAAGGCTGAGAAAGGAGCACATGGCCACTATGCTCACTATGGCTCGGACTAC GCTCGGCTAAGGGCTCAAGGGCACATTCCTCATCGAGCTAAGCCATGGAGCCTCT AAGCACAGTATGACGTTGACACTGAGACTGTCTCGGCTTCCTCAAGGCACACACCTC GACAAGGACTTCAAGGTGAACATCGAGGTGAACACGCAACTCTCGCTGGCCACACCTC GAGCACGAGCTCGAGTGGCAGTGGACAAACACATGCTCGGCTCTATCGACGCTAACCGT GGTGACTACAGAAATGGCTGGGATGATGACCACTCCAAATCGACCAAGTACGAACTCGT CAGGCTTGGATGGAATCATCCGTGGCGGGTCTCGGCACTGGCGGACGAACTTCGAC GCTAAGACTCGTCTGTAACCTACCGACCTCGAAGACATCTCATCGCACACATCGCAGGC ATGGACGCTATGGCACCGGCACTCGAATCAGCTGCTAAGCTCCTCGAAGAGTCTCCATAC AAGGCAATGAAGGAGCTCGCTACGCTTCAATCGACAAACGCTATCGGTAAGGACTTCGAA GATGGCAAGCTCACTCTCGAGCAGGCTTACGAATACGATACGTAAGAAAGGTTGGTGAGC TAAGACTTTCGCAAGGAGCTTACGAAGCCATCGTTGCAATGTACGCTTAA
1753MI2_006	<i>Neocallimastigales</i>	Amino Acid	54	MAKEYFPELGKIKFEKDSKNPMAPHYYDPEKVMGKPMKDWLRFAMAWHNLCAEGGDQ FGGGTTKTFPNNGADAVEIAKQKADAGFEIMQKLGI PYCFHDVDLVSEFASVEEYANL KAITDYLAVMKETGKILLWSTANVFNGRYMNAGSTNPDVDVVARAIVQIKNAIDAGIK LGAENYVFWGGREGYNSLLNTDQREKEHMAITMARDYARAKGFKGTFLIEPKPEPS KHQYDVTETVIGFLKAHNLDKDFKNIEVNHATLAGHTFEHFLAVAVNNMIGSTDANR GDYQNGWDTDQFPIDQYELVQAWNELIRGGGLGTGGTNFAKTRRNTDLEDIFIAIAG MDAMARALESAAKLLEESPYKAMKAARYASFDNGIGKDFDGGKLTLEQAYEYGGKVGEPK QTSKGQELYEAIAMYA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI3_005	<i>Neocallimastigales</i>	DNA	55	ATGGCTAAAGAAATTTTCCACAGAGATGGTAAATCAAGTTCGAAGGCAAGGATTCAAAAG AATCCAATGGCTTTCCATTAATCATGACAGAGAGGTAATCATATGGGCAAAACCCCATGAAG GACTGGCTCCGTTTCGCTATGGTGGTGGACACACTCTGTGACAGAGGCGGCGACACAG TTCCGTGGCGGTACGAAGAAATTCCTTGGAAAGAGGGTGTAAATGCTTCGAGATGGCT AAGCAGAGGGCTGACGCTGGTTTCGAAATCATGACAGAGCTTGGCATTCCTTACTTCTGC TTCCACGATGTTGACCTCGTTTCTGAAGGCGCATGTTGAGGAGTATGAGGCCCAACCTC AAGGCTATCATGCTATCTCGCGTGAAGATGAAGGAGACTGGCATTAAGCTCCTGTGG TCTACTGCCAACGTTTCGGCAATGGCCGTTACATGAATGGTCTTCACCAACCCCTGAC TTCCAGCTTGTGTCGCGCATCGTTTTCAGATTAAAGAACGCTATCGATGACGATCAAG CTGGTGTGAGAACTATGTTTCTGGGCGGTCGTGAAGTTTACATGAGCCTCCTGAAC ACAGACCAGAAAGCTGAGAGGACACATGGCTACTATGCTCACTATGGCTCGGACTAC GCTCGCAGCAAGGCTTCAAGGTAATTCCTCATCGAGCCTAAGCCATGGAGCCATCT AAGCACGACTACGCTGACACAGAGACTGTTATCGGCTTCTTGAAGCACACAACTT GACAAGGACTCAAGTGAACATCGAGGTGAACACGCAACTCGCTGGTACACCTTC GAGCAGGCTCGCTGCTGCTGTCGACAACTATGCTTCTATCGATGCTAACCGC GGTACTACCAAGATGGTGGGATACGACCACTTCCAAATGACCAAGTACGAGCTCGTT CAGGCTTGGATGGAGATCATCCGTTGGTGGCGTCTCGGCACAGGTGGTACAACTTCGAC GCTAAGACTCGTCTGTAATCTACCGACCTCGAGGACATTTTCATTGCTCACATCGCTGGT ATGGACGCTAAGGCTCGGCTCTTGAAGTACGAGCTAAGCTCCTTGAAGAGTCTCCATAC AAGAAGTGAAGGCTGCGGCTTATGCTTCTTCGACAGGCGCATGGTAAAGGACTTTGAG AACGGCAAGCTCACACTCGAACAGGTTTATGAGTATGGTAAAGAGGTAGGTGAGCCCAAG CAGACTTCGGCAAGCAGGAGCTCTTCGAGGCAATCGTGGCCATGTACGCATAA
5586MI3_005	<i>Neocallimastigales</i>	Amino Acid	56	MAKEFFPEIGKIKFEGKSKNPMAPHYYDAEKVIMGKPMKDWIRFAMAWHHTLCAEGGDQ FGGGTRKTFPNEGANAWEIAKQKADAGFEIMQKLGIPYFCHFDVLDVSEGASVEEYFANL KAITDYLAVKMKETGILKLLWSTANVFGNGRYMNGASTNPDFDVVARAIVQIKNAIDAGIK LGAENYVFWGGREGYMSLLNTDQREKEHMAATMLTMARDYARSKGFKGTFLIEPKMEPS KHQYDVTETVIGFLKAHNLDKDFKNIEVNHATLAGHTFEHELAVADNNMLGSDANR GDYQNGWDITDQFPIDQYELVQAMMEIRGGGLGTGGTNFDAKIRRNSTDLEDIFIAHAG MDAMARALESAAKLLEESPYKKMKAAARYASFDGSMGKDFENGKLTLEQVYEYKKGVEPK QTSKGQELFEAIVAMYA
5586MI91_002	<i>Neocallimastigales</i>	DNA	57	ATGGCTAAAGAGATTTTCCAGAGATGGTAAATCAAGTTCGAAGGCAAGGATTCCAAAG AATCCAATGGCATTCACATATATGATGACAGAGAAAGTATATATGGTAAAGCCTATGAAG GAGTGGCTCCGCTTTCGAATGGCATGGTGGCACACACTCTGTGACAGAGGTTGGCGACAG TTTGGTGGTGGCATAAGAAATTCCTATGGAAGAGGCGACTGACGCTGACGATGCTGCT AAGCAGAGGCTGATGACGTTTCGAAATCATGACAGAACTCGGTTCCCATATTTTTCG TTCCACGACATGACCTCGTTTCGAAAGGCAACAGCATTGAAGAGTATGAGGCTAACCTC CAGGCAATCATGATATCTGAAGTGAAGATGGAAGAGACAGGCAATCAAACCTCTTGTGG TCAACTGCCAACGTTATTCGGCAATGGTCCGTACATGAATGAGTGTTCACAAACCCGAC TTTGACGTGGTGGCTCGTCCCATCGTTTCAGATTGAAGAACGCAATGACGCTGGTATCAA CTGGTGTGAGAACTATGTTCTGGGCGGTCGCAAGGTACATGACGCTTCTGAAC ACTGACCAGAAAGCTGAGAAAGGACACATGGCAACCATGCTCACTATGGCTCGGACTAC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI91_002	<i>Neocallimastigales</i>	Amino Acid	58	GCTCGAGCAAGGGTTTCAAGGGCACTTCTCTCATTTGAGCCAAAGCCAAATGGAGCCATCT AAGCACCATGATGACGTTGACACGAGACTGTCATCGGCTTCCTCAAGGCACACAACTC GACAAGGATTTCAAGGTGAACATCGAAGTGAACACACGCTACACTTCCAGGTTCATACCTTC GAGACGAACTTGCTGTGGCTGTGACAAIAGCATGCTCGGTCTATCGACGCTAAACGCT GGTGACTATCAGAACGGTTGGGACACTGACCACTTCCCAATCGACCACTAGCAACTCGTT CAGGCTTGATGAAATCATCCGTGGTGTGCTCGGCACAGGTGGTACTAACTTCGAT GCTAAGACTCGTAACTCAACTGACCTCGAGGACATCTCATCGCACACATCTCTGGT ATGGATGCAATGGCAGCTGCTCTCGAATCGGGCGCTAAACTTCTTGAGGAGTCTCCATAC TGGCTATGAAGAAGGCTCGTTACGCTTCTTCGACAGCGCATCGGTAAAGGACTTCGAG GACGGCAAACTCACGCTCGAGAGGCTTACGATACGGCAAGAAAGTCCGGCAACCCAAAG CAGACTTCTGGCAAGCAGGAACCTACGAGGCAATCGTTGCCATGTACGCATAA MAKEYFPEIGKIKFEGKDSKNPMAPHYYDAEKVIMGKPMKEWLRFAMAWHHTLCAEGGDQ FGGGTKKFPWNEGTDVTTIAKQKADAGFEIMQKLGFYFCFHDIDLVSSENSIEEYEANL QAITDYLKVKMEETGILKLWSTANVFGNGRYMNGASTNPNFDVVVARAIVQIKNAIDAGIK LGAENYVFWGGREGYMSLINTDQKREKHEMATMLTMARDYARSKGFGKGTFLIEPKPMEPS KHQYDVTETVIGFLKAHNLDKDFKNIEVNHATLAGHTFEHELAVAVDNGMLGSDANR GDYQNGWDIDQFPIDQYELVQAWMEILIRGGGLGTGGTNFDAKTRRNSTDLEDIFIAHISG MDAMARALESAAKLLESPYCAMKKARYASFDSGIGKDFDGLKLTLEQAYEYGVKVGEPK QTSKGQELYEAIVAMYA
5586MI194_003	<i>Neocallimastigales</i>	DNA	59	AATGGCAAAAGAGTATTTCCCTACGATCGGTAGATCGTTTATGAAGGACCGGAGTCCAAG AACCCTATCGCATTTCTATTACTATGACGACGACGCGGTAGTAGCTGGTAAAAAAATGAAA GATTGATGCGTTTCGCTATGGCATGGTGGCACACCTCTGTGCAGAAAGGTGCAGACGAG TTCGGTGGAGGCACCAACAACCTCCCGTGGAGTGAAGTCCGATGCCGTAAACCATCGCC AAGCAGAAAGCAGACGCGAGTTTGGATCATGCAGAAACTCGGCTTCCCGTATTTCTGT TTCATGACGTGGATGCTGTCAGCAAGGACGACGCTAGAAAGATACGAGGCGAACCTC GCAGCATCACCGATTATCTCAAGCAGAAAATGGACGAGTCGGGTATCAAACTCCTTTGG TCCACTGTAAACGTATTCGGTCACGCCCTTACATGAACGCTGCCAGCACCATACTCTGAC TTTGATGCTGTTGCCGTGCGATTGTGCAGATCAAGAAIAGTATCGACGAGGTATCAAA CTCGGCGCAGAGAACTACGTCTTCTGGGGCGTCTGTAAGTTATATGAGCCTGCTCAAT ACCGACCAAGAAACGCGAGAAAGAGCATACGGCAATGATGCTGCGTATGGCGCGTGACTAT GCCCGCAGCAAAAGTTTCAAAGGTACCTTCCTCATGAACCAACCAACCATGGAGCGCTCC AAGCACCATGATGACGTAGATACCGAGACGGTGTAGTTTCTCTAAAGCACACGCTTTG GAGAAAGACTTTAAGTTAAACATCGAAGTGAACACGCTACCTCGCCGGTACACACTTC GAGCAGAACTGGCAGTAGCCGTAGATTAACGGATGCTCGGTTCGATCGATGCCAACCGC GGTGACTATCAGAACGGATGGATACCGACCACTTCCCCATCGATAACTCGAACTGACCC CAAGCATGGATGCAGATCGTACGTAAACGGTGGTCTCGGCACAGGCGGACGAACITTCGAC TCCAAGACCCGTCGTAACTCCACCATCTCGAGGATATCTTCATCGCTCACATCAGTGGT ATGGACGCTGTGCCGTGCCCTATTGAATGCCGTAGAGATCATGGAGAAATCACCGATC CCTGCTATGCTCAAAGACGCTTACGCTTCTCGATAGCGGTCTGGGTAAAGATTTTCGAG GACGGCAAACTGACCTTGAGCAAGTCTATGATACGGTAAAGAAAGTAGCGGAACCCAAA CAAACCGCGGCAACAAAGAACTCTATGAGGCTATCGTTGCCCTCTACGCTAAATAA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI194_003	<i>Neocallimastigales</i>	Amino Acid	60	MAKEYFTTGKIVYEGPESKNPMAPHYYDAERVVAGKKMKDWMRFAMAWWHITLCAEGADQ FGGTTKHPFWSEGPDAVTIAKQKADAGFEIMQKLGFPYCFHDVDLVSEGSSVEEYEA AAITDYLKQKMDSEGIKLLWSTANVFHARYMNGASTNPDFDVARAIVQIKNAIDAGIK LGAENYVFWGGREGYMSLLNTDQREKEHTAMLMRDYARSKFGFTFLIEPKMEPSK KHQYDVTETVIGFLKAHGLEKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSDANR GDYQNGWDTDQFPIIDNFEITQAFMQIVRNGGGLGTGTGTFNFSKTRNSTDIEDIFIAHISG MDACARALLNAVEIMEKSPIPAMLKERYASFDGSLGKDFEDGKLTLEQVYVEYGGKVGEPK QTSKQELYETIVALYAK
5586MI198_003	<i>Neocallimastigales</i>	DNA	61	ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAAATTTGAAGGCCCGGAGTCCAAGAAC CCGATGGCATTTCACTACTATGACGACAGAGCGGTCTGAGCGGTAAACAAATGAAGAG TGGATGCGTTTCGCTATGGCTTGGTGGCACACCTCTCTGTGGGAAGGGCGGACCGATTCT GGAGCGGAACGAAGAAGTTCCTTGGAAACGAAGCGCTAACGCTTTGGAGATCGCAAG CACAAAGCGATGCGGATTTGAGATCATGAGAACTCGGCATCCCTATTCTGTCTTC CATGACGTGGATCTCATCGCGGAGGGGTTTCGGTAGAAGATACGAAGCAACCTCGCT GCCATACCGATTACCTCAACAGAAAATGGACGAGACTGGCATCAAACTGCTGTGGTCC ACGGCGAACGCTTTCAGCAACCCCGTTATATGAACGGCGACGACGACCCCGATTTC GATGATAGCGGTGCCATCGTCCAGATCAAGAACGCTATCGACCGCGGTATCAAACTC GGAGCAGAACTATGCTCTCTGGGTGGTCCGAGGGCTATATGAGCCTCCTCAACACT GACGAGCGGAGAGAAAGAGCATGGCTACCATCGGTATGGCGGTGACTACGCG CGTGCCAAAGGATTCAAGGGCACCTTCTCATCGAACCCAAACCATGTGAGCCGTCCAAA CATCAGTATGATGTCGATACCGAGACCGTTCATCGGTTCTCTCAAAGCGCATGGACTCGAC AAGGATTTCAAAGTCAATATCGAGTCAACCGCCACCTCGCAGGCCACACGTTTCGAA CAGAACTGGCTTGGCTGTAGTCCGGCATGCTCGGTTTCGATTGACGCCAACCCGGGT GACGCCAGAACGGATGGACACCGACCGATTCCCTATTGATAACTTCGAACTCACACAG GCTTTCATGACGATCGTCCGCAACGGCGGTTTCGGAACAGCGGTACGAACTTCGACGCC AAGACCGCGGTAACTCCACCGACTGGAGGACATCTTCATCGCCCATATCAGCGGATG GACGCTTGGCGCAGTGGCTTACTCAATGCTGCGAAATCCTCGAGAAGAGCCCGATTCCG GCGATGCTCAAAGAGCGTTATGCTTCTTCGCGCGGCTCGGAAAGGACTTCGAGGAG GGAAACTGACTTTCGAGCAGGTCTATGAGTACGGCAAGAAAGTCGGCAACCCCAACAG ACCAGCGCAACAGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG
5586MI198_003	<i>Neocallimastigales</i>	Amino Acid	62	MKEYFPEITGKIQFEGPESKNPMAPHYYDAERVVAGKTMKEMRFAWAWWHITLCAEGGQF GGGTTKFPWNEGANALEIAKHKADAGFEIMQKLGIPYCFHDVDLIAEGSSVEEYEA AITDYLKQKMDTEGILKLLWSTANVFHARYMNGASTNPDFDVARAIVQIKNAIDAGIKL GAENYVFWGGREGYMSLLNTDQREKEHTAMLMRDYARAKFGFTFLIEPKMEPSK KHQYDVTETVIGFLKAHGLEKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDANR DAQNGWDTDQFPIIDNFEITQAFMQIVRNGGFGTGTGTFNFSKTRNSTDIEDIFIAHISG MDACARALLNAVEIMEKSPIPAMLKERYASFDGSLGKDFEDGKLTLEQVYVEYGGKVGEPKQ QTSKQELYETIVALYAK
5586MI201_003	<i>Neocallimastigales</i>	DNA	63	ATGGCAAAAGAGTATTTCCCTACGATCGGTAAGATCGTTTATGAAGGACCGGAATCCAAG AACCTTATGGCATTTTATCTACTATGACGACGAGCGGTAGTAGTGGTAAAAAATAA GATTGGATGCGTTTCGCTATGGCATGGTGGCACACCTCTCTGCAAGAGGTGCAGACCAG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				TTCCGTTGGAGGCAACCAACACTTCCCGTGAATGAAGGTCCCGATGCCGTAAACCATCGCC AAGCAAGAGCAGACGCGAGTTTTCAGATCATCAGAAATCGGCTTCCCGTATTTCTGT TTCCATGACGTGGATCGTTCGGGAAGGCGAGCGGTAGAGAGTACGAGCGGAACCTC GCAGCCATCACCGATTATCTCAAGCAGAAAAATGGACGAGTCGGGTATCAAACTCCTTGG TCCACTGCTAACTGATTCGGTCACGCCGTTACATGAACGTGCCAGCAATCCTGAC TTTGATGTCGTTGCCGTCGATGTGAGATCAAGATGCTATCGACGAGGTATCAAA CTCGGCGCAGAGAACTACGTCCTTGGGCGGTCGTGAAGTTATATGAGCCCTGCTCAAC ACCGACAGAAACCGCAGAAAGCATACGGCAATGATGCTGCGTATGGCGCGTGACTAT GCCCGCAGCAAGGTTTCAAGGTACCTTCTCATCGAACCAACCCATGGAGCGTCC AAGCACCATGACGTAGATACCGAGACGGTGATAGTTTCTCTAAAGCACACGGTTTG GAGAAAGACTTAAAGTTAAACATCGAAGTGAACACGTCACCTCGCCGTCACACTTC GAGCAGCACTGGCAGTAGCGGTAGATAACGGCATGCTCGGTTCGATGCCAACC GGTGACTATCAGAACCGGATGGATACCGACCACTTCCCATCGATACTCGAATGACC CAAGCATGGATGAGATCGTACGTACGTAACGGTGGTTCGGCACAGCGGACGAATTCGAC TCCAAGACCGGTCGTAACCTCACCGATCTCGAGGATATCTTCATCGTCACATCAGTGT ATGACGCTTGTGCCGTCCTTAAAGTCCGTAGAGATCATGGAGAAATCACCGATC CCTGCTATGCTCAAGAGCGTTACGCTTCTTCGATAGCGTCTGGTAAAGATTTCGAG GACGGCAAACTGACCTTGAAGCACTATGAGTACGGTAAAGAAAGTAGCGGAACCAAA CAAAACGCGGCAACCAAGACTCTATGAGGCTATCGTTCGCTTACGCTAAATAA
5586MI201_003	<i>Neocallimastigales</i>	Amino Acid	64	MAKEYFPTIGKIYEGESKNPMAPHYYDAERVVAGKKMDMRFAMAWHHTLCAEGADQ FGGTRKHPWNEGPDATVIAKQADAGFEIMQKLGFPYFCFHDVLDLVGSSVEEYFANL AAITDYLRQMDSEGLKLLWSTANVFHARYMNGASTNPDPDFVARAIVQIKNAIDAGIK LGAENYVFWGGREGYMSLLNTDQREKEHTAMLMARDYARSKGFKGTFLEPKMEPS KHQYDVTETVIGFLKAHGLEKDFKNIEVNHATLAGHTFEHELAVAVDNGMLGSDANR GDYQNGWDITDQFPIDNFELTQAMNQIVRNGGLGTGGTNFDSKIRRNSTDLIEDIFIAHISG MDACARALLNAVEIMEKSPIPAMLKERYASFDSGLGKDFEDGKLTLEQVYEGKKVGEPK QTSQKQELYEAI VALYAK
5586MI204_002	<i>Neocallimastigales</i>	DNA	65	ATGAAAGAGTATTTCCTGAGGTCGGTAAGATCCAATTGAAGGCCCGGAGTCTAAGAAC CCGATGGCATTTCACTACTATAGCGCAGAGCGGTCGTAGCCGGTAAACAATGAAGAG TGGATGCGTTTCGCTATGGCTTGGTGGCACACCTCTGTGAGAGAGCGCGGACCATTC GGAGCGGAACGAAGCATTTCCCGTGAATGAAGCGCTAACGCTTTGGAGATCGCAAA CACAAAGCCGATCGGGATTTCGAGATCATGCAAACTCGGCATCCCTATTTCTGTTTC CATGACGTGGATCTCATCGCGGAGGCGGTTCCGTAGAAGATACGAACCAACCTCGCT GCTATCACCGACTACCTCAAGCAGAAAAATGGAGAGACCGGCATCAAACTGTGGTCC ACGGCGAACGCTGTTCAGCAACCCCGTTATATGAACGGCGGAGCACGAACCCGATTC GATGTAGTAGCGGTGCCATCGTGCAGATCAAGAATGCCATCGACCGCGGCATCAACTG GGCGAGAGAACTATGCTTCTGGGCGGTCGGAGGGGTACATGAGCTGCTCAACACC GACGAGCGCGGAGAAAGAGCATATGGCTACTATGCTCGTATGGCGGTGACTAGCGG CGTGCCAAAGGATTCAAGGGCACCTTCTCATCGAACCCAAACCGTGTGAGCCGTCCTCAA CATCATGATGATGTCGATACCGAGACCGTTCATCGGTTTCTCAAAGCGCATGGACTCGAC AAGGATTTCAAGGTTAATATCGAGGTCAACCAACGCCACCTTCGCGAGGCCACACGTTTCGAA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586M1204_002	<i>Neocallimastigales</i>	Amino Acid	66	CACGAACGGCTTGGCTGTAGATGCCGGCATGCTCGTTTCGATTGACGCCAACCCGGGT GACGCCCAACGGATGGGACACCGACACAGTTCCTTATGATAACTTCGAACTCACACAG GCTTTCATGCAGATCGTCGCAACGGCGGTTTCGGAACAGCGGTACGAACTTCGACGCC AAGACACGCCGTAATCCACCGACTTGGAGGACATCTTCAATCGCCCATATCAGCGGCAATG GACGCTTGGCACGTCGCTTCTCAACGCCATCGAAATCCTCGAGAAAGCCCGATCCCG GCTATGCTCAAGACCGTTATGCTTCTTGTGCGGCGATCGGAAAGGACTTTTGAGGAG GGCAAACTGACTTTCGAGCAGGTCATGAGTACGGCAAGAGGTTCGGAACCCCAACACAG ACCAGCGGCAACAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG MKEYFPEVGKIQFEGPESKNPMATHYDAERVVAGTKMKFWMRFAMAWHTLCAEGGDQF GGGTTKFPWNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGGSVEEYETNLA AITDYLKQKMDETGIKLLWSTANVFNPRYMNGASTNPFDDVVARAIVQIKNAIDAGIKL GAENYVFWGGREGYMSLLNTDQRRKEHMAIMLRMARDYARAKGFKGTFLIEPKPCPSK HQYDVTETVIGFLKAHGLDKDFKYNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWTDIDQFPIDNFEITQAFMQIVRNGGFGTGTNFDATKRNSTDLEDIFIAHISGM DACARALLNAIEILERSPI PAMLKDRYASFDGGIGKDFEEGKLTTFEQVYVEYKKGVEPKQ TSGQELYETIVALLYAK
5586M1207_002	<i>Neocallimastigales</i>	DNA	67	ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAAATTTGAAGGCCCGGAGTCCAAGAAC CCGATGGCGTTTCACTACTATGACGCTGAGCGGTCGTAGCCGGTAAACAATGAAAGAG TGATGCGTTTCGCTATGGCTTGGTGCACACCTCTGTGGGAAGGCGCGACCGCTTC GGAGAGGAAACGAAGAAATTCCTTGGAAACGAAGGGGCAACGCTTTGGAGATCGCAAG CACAAAGCCGATCGGGATTCGAGATCATGCAGAAATCGGCATCCCTATTCTGTTC CATGACGTGGATCTCATCGCGAGGCGAATCGGTAGAAGATACGAGCCAAACCTCGCT GCCATCACCGATTACCTCAACAGAAAATGGACGAGACCGGCATCAAACTGTGTGTCTC ACGGCAACGCTGTTACGCAACCCCGTTATATGAACGGCGCAGCACGAAACCCGATTTC GAIGTAGTGGCACGCGTATCGTACAAATCAAGAACGCTATCGACGCGGTATCAAACTC GGAGCAGAACTATGTCTTCTGGGCGGTCCGAGGGCTATATGTCTCGTCTCTCAACACC GACCAGCGCGAGAGAAAGACATATGGCTACTATGCTCCGTATGGCGGTGACTACGCG CGTCCAAAGGATTCAGGGCACCTTCCTCATCGAACCCAAACCGTGTGAGCCGTCCTCAA CATAGTACGATGTGGACACAGAGCCGTATCGGTTTCCTTAAAGCGCATGGATCGAC AAGGATTTCAAAGTCAATATCGAGGTCAACCGCCACCTTCGAGGCCACACGTTTCGAA CACGAACCTGGCTTGGCTGTAGTCCGGCATGCTCGTTTCGATTGACGCCAACCCGGGT GACGCCAGAACGGATGGGACACCGACCAATTCCTTATGATAAATTCGAACTTCACTAG GCTTTCATGCAGATCGTCCGCAACGGCGTTTCGGAACAGCGGTACGAACTTCGACGCC AAGACACGCGGTAACCTCACCGACTGGAGGACATCTTCATCGCCCATATCAGCGGATG GACGCTTGGCTCGTGGCTTGTCAATGTCTCGAAATCCTCGAGAAAGCCCGATCCCG GCTATGCTCAAGAGCGTTATGCTTCTTTCGCGCGGCTCGGAAAGGACTTTGAGGAG GGCAAACTGACTTTCGAGCAGGTCATGAGTACGGCAAGAGGTTCGGAACCCCAACACAG ACCAGCGGCAACAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATGA
5586M1207_002	<i>Neocallimastigales</i>	Amino Acid	68	MKEYFPEIGKIQFEGPESKNPMATHYDAERVVAGTKMKFWMRFAMAWHTLCAEGGDQF GGGTTKFPWNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGGSVEEYENLA AITDYLKQKMDETGIKLLWSTANVFNPRYMNGASTNPFDDVVARAIVQIKNAIDAGIKL

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI209_003	<i>Neocallimastigales</i>	DNA	69	GAENYVFWGREGYMSLLNTDQRREKEHMAIMLRMARDYARSKGFKGTFLEPKPCEPSK HQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWDTDQFPIDNFEITQAFMQIVRNGGFGTGNFDKTRNSTDIEDIFIAHISGM DACARALLNAVEIILEKSPPI PAMLKERYASFDGGIGKDFEKGKLTFEQVVEYGGKVGEPKQ TSGKQELYETIVLYAK ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAAATTTGAAGGCCCGGAGTCCAAAGAAC CCGATGGGCTTTCACCTACTATGACGACAGAGCGGTAGTAGCGGGTAAACCAATGAAAGAA TGGATCGCTTTCGCCATGGCATGGTGGCACACCTCTGTGCAGAAAGGGCGGACCCAGTTC GGAGGAGGAACGAAGCATTTCCCGTGAATGAAGGCGCTAACGCTTGGAGATCGCCAAA CACAAAGCCGATCGGGATTCGAGATCATGAGAAACTCGGCATCCCTATTCTGTTC CATGACGTGGATCTCATCGCCGAGGGCGATTCCGTGGAGGATACGAGCTAACCCGCT GCCATACCGATACCTCAACAGAAAAATGGACGAGACCGGCATCAAACTGCTGTGTCC ACGGGAACGCTTTCAGCAACCCCGTTACATGAACGGTGGAGACGAAACCCGGATTTC GATGATGGCACGGCTATCGTACAAATCAAGAACGCTATCGACGCCGATCAAACTC GGAGCAGAACTATGTCTTCTGGGCGGTCCGAGGGCTATATGTCGCTCCTCAACACC GACGAGCTCGGAGAAAGAGCATATGGCTACTATGCTCCGTATGGCGGTGACTACGCG CGTGCCAAAGGATTCAAGGGCACCTTCTCATCGAACCCAAACCATGTGAGCCGTCCTAAA CATCAGTACGATGTGACACAGAGACTGTATCGGTTCTCAAAGCGCATGGACTCGAC AAGGATTTCAAAGTCAACATCGAGTCAACCGCCACCTCGCAGGTCACACACGTTGAA CAGAACTGGCTTGGCTGTAGTCCGGCATGCTCGTTGATTAATTCGAACTCACACAG GACGCCAGAACGGATGGACACTGACCATTCCTTATGATAACTTCGAACTCACACAG GCTTTCATGACAGATCGTCGCAACGGCGGTTTCGGAACAGGGGTACGAACTTCGACGCC AAGACAGCGGTAACTCCACCGATTGGAGGACATCTTATCGCCCATATCAGCGGCATG GACGCTTGTCCGTGGCTGCTCAACGCCATCGAAATCTCGAGAGAGCCGATCCCG GCTATGCTCAAGAGCGGTACGCTTCTTTCGCGGCGCATCGGAAAGGACTTTGAGGAT GGTAACTGACTTTCGAGCAGGCTATGATGATCGGCAAGAGGTCGGAGAACCCCAACAG ACCAGCGCAACAGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAGTAA MKEYFPEIGKIQFEGESKNPMFHYDAERVVAGKTMKEMRFAWAWHILCAEGGQDF GGGTHFFWNEGANALEIAKHKADAGFEIMQKLGIPYCFHDVDLIAEGDSVEEYEANPA AITDYLKOKMDETGIKLLWSTANVFSPNRYMNGASTNPDEFVARAIVQIKNAIDAGIKL GAENYVFWGREGYMSLLNTDQRREKEHMAIMLRMARDYARAKGFKGTFLEPKPCEPSK HQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWDTDQFPIDNFEITQAFMQIVRNGGFGTGNFDKTRNSTDIEDIFIAHISGM DACVRALLNAIEIILEKSPPI PAMLKERYASFDGGIGKDFEKGKLTFEQVVEYGGKVGEPKQ TSGKQELYETIVLYAK ATGAAAGAGTATTTCCCTGAGATCGGAAAGATCCAAATTCGAAGGCCCGGAGTCCAAAGAT CCTATGGCATTTACCTACTATGACGACAGAGCGGTAGTAGCGGGTAAACCAATGAAAGAG TGGATCGCTTTCGCCATGGCATGGTGGCACACCTCTCGCAGAAAGGGCGGACCCAGTTC GGAGGCGGACGAAGCATTTCCCTTGAATGAAGGTGCAACGCTTGGAGATCGCCAAAG CACAAAGCCGATGAGGCTTCGAGATCATGCAAGAACTCGGCATCCCTATTCTGTTC CATGACGTGGATCTGATCGCCGAGGGCGGTTCCGTAGAAGAGTATGAAGCTAATTTAACG
5586MI209_003	<i>Neocallimastigales</i>	DNA	70	GAENYVFWGREGYMSLLNTDQRREKEHMAIMLRMARDYARSKGFKGTFLEPKPCEPSK HQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWDTDQFPIDNFEITQAFMQIVRNGGFGTGNFDKTRNSTDIEDIFIAHISGM DACARALLNAVEIILEKSPPI PAMLKERYASFDGGIGKDFEKGKLTFEQVVEYGGKVGEPKQ TSGKQELYETIVLYAK ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAAATTTGAAGGCCCGGAGTCCAAAGAAC CCGATGGGCTTTCACCTACTATGACGACAGAGCGGTAGTAGCGGGTAAACCAATGAAAGAA TGGATCGCTTTCGCCATGGCATGGTGGCACACCTCTGTGCAGAAAGGGCGGACCCAGTTC GGAGGAGGAACGAAGCATTTCCCGTGAATGAAGGCGCTAACGCTTGGAGATCGCCAAA CACAAAGCCGATCGGGATTCGAGATCATGAGAAACTCGGCATCCCTATTCTGTTC CATGACGTGGATCTCATCGCCGAGGGCGATTCCGTGGAGGATACGAGCTAACCCGCT GCCATACCGATACCTCAACAGAAAAATGGACGAGACCGGCATCAAACTGCTGTGTCC ACGGGAACGCTTTCAGCAACCCCGTTACATGAACGGTGGAGACGAAACCCGGATTTC GATGATGGCACGGCTATCGTACAAATCAAGAACGCTATCGACGCCGATCAAACTC GGAGCAGAACTATGTCTTCTGGGCGGTCCGAGGGCTATATGTCGCTCCTCAACACC GACGAGCTCGGAGAAAGAGCATATGGCTACTATGCTCCGTATGGCGGTGACTACGCG CGTGCCAAAGGATTCAAGGGCACCTTCTCATCGAACCCAAACCATGTGAGCCGTCCTAAA CATCAGTACGATGTGACACAGAGACTGTATCGGTTCTCAAAGCGCATGGACTCGAC AAGGATTTCAAAGTCAACATCGAGTCAACCGCCACCTCGCAGGTCACACACGTTGAA CAGAACTGGCTTGGCTGTAGTCCGGCATGCTCGTTGATTAATTCGAACTCACACAG GACGCCAGAACGGATGGACACTGACCATTCCTTATGATAACTTCGAACTCACACAG GCTTTCATGACAGATCGTCGCAACGGCGGTTTCGGAACAGGGGTACGAACTTCGACGCC AAGACAGCGGTAACTCCACCGATTGGAGGACATCTTATCGCCCATATCAGCGGCATG GACGCTTGTCCGTGGCTGCTCAACGCCATCGAAATCTCGAGAGAGCCGATCCCG GCTATGCTCAAGAGCGGTACGCTTCTTTCGCGGCGCATCGGAAAGGACTTTGAGGAT GGTAACTGACTTTCGAGCAGGCTATGATGATCGGCAAGAGGTCGGAGAACCCCAACAG ACCAGCGCAACAGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAGTAA MKEYFPEIGKIQFEGESKNPMFHYDAERVVAGKTMKEMRFAWAWHILCAEGGQDF GGGTHFFWNEGANALEIAKHKADAGFEIMQKLGIPYCFHDVDLIAEGDSVEEYEANPA AITDYLKOKMDETGIKLLWSTANVFSPNRYMNGASTNPDEFVARAIVQIKNAIDAGIKL GAENYVFWGREGYMSLLNTDQRREKEHMAIMLRMARDYARAKGFKGTFLEPKPCEPSK HQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWDTDQFPIDNFEITQAFMQIVRNGGFGTGNFDKTRNSTDIEDIFIAHISGM DACVRALLNAIEIILEKSPPI PAMLKERYASFDGGIGKDFEKGKLTFEQVVEYGGKVGEPKQ TSGKQELYETIVLYAK ATGAAAGAGTATTTCCCTGAGATCGGAAAGATCCAAATTCGAAGGCCCGGAGTCCAAAGAT CCTATGGCATTTACCTACTATGACGACAGAGCGGTAGTAGCGGGTAAACCAATGAAAGAG TGGATCGCTTTCGCCATGGCATGGTGGCACACCTCTCGCAGAAAGGGCGGACCCAGTTC GGAGGCGGACGAAGCATTTCCCTTGAATGAAGGTGCAACGCTTGGAGATCGCCAAAG CACAAAGCCGATGAGGCTTCGAGATCATGCAAGAACTCGGCATCCCTATTCTGTTC CATGACGTGGATCTGATCGCCGAGGGCGGTTCCGTAGAAGAGTATGAAGCTAATTTAACG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI214_002	<i>Neocallimastigales</i>	Amino Acid	72	GCTATCACCGATTACCTCAAAACAGAAAATGGACGAGACCGGCATCAAACTGCTGTGGTCC ACTGCGAAGCTGTTCCGTTAACGCACGTTATATGAACGGCGGAGACGAAACCCCGATTTC GATGTAGTGGCAGCGCTATCGTGAGATCAAGAACGCTATCGACGCCGATCAAACTG GGCGAGAGAACTACGTTCTTGGGGCGGTCCGAGGGAATATATGTGCTCCCTGAACACC GACCAGAACCGTGAGAAAGACATATGGCTACCATGCTCCGTATGGCGGTACTACGCG CGTCCAAAGGATTCAAAGGTACGTTCTCATCGAGCCAAACCGGTGAGCCGTCCTCAAA CATAGTACGACGTGGACACTGAGACCGTCACTGGTTCCTCAAAAGCCCATGGTCTCGGC AAGGATTTCAAAGTGAACATCGAGTGAATCAGCCACCCCTCGCAGGGACACGTTCTGAA CAGGAACGGCTTGGCGGTAGATGCCGATGCTCGGTTCGATCGACGCCAACCGCGGT GACGACAAACGGATGGACACCGACAGTTCCTTATTAATAATTCGAACTACCCAG GCATTCATGACGATCGTCCGCAACGGCGTTTCGGAACAGCGGTACGAACTTCGACGCC AAGACACCGCTAATCCACCGACTGGAGGACATCTTCATCGCCATATCAGCGGCATG GACGCTTGCGCGTGGCTTGGTCAATGCTGTGAAATCCTTGAAGAGCCCGATCCCG GGATGCTCAAAGAGCGTTACGCCCTCTTTCAGACGGTATGGTAAAGACTTTGAGGAG GGCAAGCTGACCTTCGACGAGTCTATGATACGGCAACAGGTCGGGAAACCCCAACAG ACCAGCGCAAGCAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG
5586MI214_002	<i>Neocallimastigales</i>	Amino Acid	72	MKEYFPEIGKIQFEGEPESKNPMATHYYDAERVVAGTKMKEMRFALAWHHLCAEGGDQF GGGTHFPWNEGANALETAKHKADAGFEIMQKLGIPYCFHVDVLLIAEGGSVEEYEANLT AITDYLKQKMDETGIKLLWSTANFVGNARYMNGASTNDFDVVARAI VQIKNAIDAGIKL GAENYVFWGREGYMSLLNTDQKREKEHMAIMLRMARDYARSKGFKGTFLIEPKPEPSK HQYDVTETVIGFLKAHGLGKDFKVNIEVNHATLAGTTFEHELACAVDAGMLGSI DANRG DAQNGWDTDQFPIDNFELTQAFMQIVRNGGFGTGTNFDKTRNSTDLEDFIAHISGM DACARALLNAVEILEKSPIPAMLKERYASFDSGMGKDFEEGKLTTFEQVYEGKVGEPPKQ TSGKQELYETIVALYAK
5751MI3_001	<i>Neocallimastigales</i>	DNA	73	ATGAAAGAGTATTTCCACAAATCGGCAAGATCCCATTTGAGGGACCCAGAGTCAAGAAC CCAATGGCATCCACTACTATGACGACAGAGCGGTAGTTCGCGGTAAGACAATGAGGAA TGGATGCGTTTCGCTATGGCTGGTGGCACACITCTCTGTGTGAGGTAGCGATCAGTTC GGCCCTGGTACAAAGAGTTCCTTGAACGAGGCGGAGACAGCCCTTACGCGCGCTAAG CACAGGCAGATGCTGGCTTCGAGGTATGCAAGCTCGGCATCCCATATTTCTGCTTC CAGCATGACACCTTATCGACGAGGCTGCTAACGTGGCTGAGTATGAGGCAACCTCGCT GCTATCACTGACTACCTGAAGGAGAGATGAGGAGACTGGCTAAAGCTCCTCTGGTCT ACAGCCAACTGCTCGGTAACTGCTGCTATATGAACGCTTCTACAAATCCTGACTTC GACGTTGTGGCTCGTGCCATCGTACAGATTAGAACGCTATCGACGCTGGTATCAAGCTT GGTGTGAGAACTACGTGTTCTGGGCGCGCGAGGCTACATGAGCTTCTGAACT GACCAGAGCGGAGAGGACATGGCAACTATGCTCGGCATGGCTCGGACTATGCCC CGCGCTAAGGATTACCGGTACCTTCTCATGAGCCAAAGCCAAATGGAGCCCAACAAAG CATCAGTATGATGTTGACACAGAGACCGTTATCGGTTTCTCAAGGCTACGGTCTGGAC AAGGACTTCAAGGTGAACATCGAGGTGAACCGCTACTCTCGCCGGTACACACTTCGAG CACGAGCTCGCTTGGCTGTGACGCTGGTATGCTCGGTCTATCGACGCTAACCCGGGT GACGCTCAGAACGGATGGATACGACCAAGTCCCAATCGACAACTTCGAGCTGACACAG GCTTGGATGCAGATTGTTCCCAATGGCGGTCTTGGCACAGGTGGTACCAACTTCGACGCA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5751M13_001	<i>Neocallimastix</i>	Amino Acid	74	AAGACCCGTCGTAACTTACCGACCTCGAGGACATCTTCACTCGTCACTCTCCGGTATG GACGCTTGTCACGCGCTCTCTCAACGCAGTAGAGATATCGAGAACTCTCCAATCCCA ACAAATGCTGAAGACCGCTATGCAAGCTTCGACTCAGGTATGGTAAAGACTTCGAGGAC GGCAAGCTCACACTTGAGCAGGTTATGAGTATGGTAAGAAAGTTCGACGAGCCAAAGCAG ACCTCTGGTAAGCAGGAACTCTATGAGACCATCGTTGCTCTCTATGCAAAATAA MKEYFPQIGKI PFEGPSKNPMATHYDAERVVAGTKMKWMMRFAMAWHHTLCAEGSDQF GPGTKFPWNEGETALERAKHKADAGFEVMQKLGI PYFCFHDVDLIDEGANVAEYEAFLA AITDYLKEKMEETGVKLLWSTANVFNARYMNGASTNDFDVVARAIVQIKNAIDAGIKL GAENYVFWGGREGYMSLLINTDQKREKHEMATMLGMARDYARAKFTGTFLIEPKPMEPTK HQYDVTETVIGFLKAHGLDKDEKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG DAQNGWDDQFPIDNFELTQAWQIVRNGLGTGGTNFDTAKTRNSTDIEDIFIAHISGM DACARALLNAVEILENSPIPTMLKDRYASFDSGMGKDFEDGKLTLEQVYVEYGGKVDPEKQ TSGKQELYETIVALLYAK
5753M13_002	<i>Prevotella</i>	DNA	75	ATGGCTAAAGAAATACCTCCCTCCATCGGCAAAATCCCTTTGAAGGAGCGGACAGCAAA AATCCCTCGCTTTCCATTAATATGACGCGGACGCGTGGTTATGGGCAAGCCCCATGAAG GAATGGCTTAAATTCGCATGGCTGGTGGCACACGCTGGCGCAGGCTCCGGAGACCCC TTCGGCGGCAGACCCGAGCTACGAATGGACAAGGGCGAATGCCCTACTGCCGCGCC AAAGCCAAAGCCGACGCGGTTTGAATCATGCAAAAGCTGGGTATCGAATACATTCTGC TTCACGATGTGGACCTATCGAGATTGCGATGACATGCGGAATACGAAGCCCGCAIG AAGGACATCAGGACTACCTGCTGGAAGAAATGAAGGAGACCGGCATCAAGAACCCTGG GGCACGCCAATGTCTCGGCCACAAGCGCTACATGAACGCGCCGCGCACCAATCCGACG TTCGATGTGGTGGCGCTGCGCGCTCCAGATCAAGAAGCGCTGGACGCCACCATCAAG CTGGGCGGCTCCAATATGTGTTTGGGGGGCGCGGAAGGTATTAACCCCTCCTCAAC ACCCAGATCGAGCGGGAAGAACACCTGGCCAAATTTGTCACGCGCCCGCGGACTAT GCCGCGCAAGGGCTCAAGGGCACCTTCCTCATTGAGCCCAACCCATGGAACCCACC AAGCACCTAGACGCTGGATACGAGACGGTCACTGGCTTCCTCCGTGCCAACGGCTG GACAAGGACTCAAGGTGAACATCGAGGTGAACACGCCACCTGGCGGCGCCACACCTTC GAGCATGAGCTCACCGTGGCGCGGAGAACGGTTCTGGGTCCATCGGTGCCAACCGC GGGACGCCAGAACGGCTGGGACGGACCACTTCCCTGTGACCCCGTACGATCTTACC CAGGCCATGATGAGGTGCTGTGAACGGCGGTTCCGCAACGGCGGCGCACCACTTCGAC GCCAAATCCGCGCTCTCCACGACCTGAGGACATCTCATGCCATATTTCCGCC ATGGATGCCATGGCCACGCTTGTGTTAAGCAGCTGCGGTGCTGATGGAAGAGCCCCCTG TGCCAGATGCTCAAGGAGGCTTATGCCAGCTTCAGCGGCGGCTCGGCAACAGTTCGAG GAAGCAAGGCTACCTGGAAGCCTGTACGAATACGCAAGTCCAGGTGACCGCTT GTCGCTCCGGCAAGCAGGAGCTTACGAGACTCTCTGAACTGTATGCCGTCAAGTAA
5753M13_002	<i>Prevotella</i>	Amino Acid	76	MAKEYFPSIGKI PFEGGDSKNPLAFHYIDAGRVVMGKPMKEWLKFAMAWHHTLGQASGDP FGGQTRSYEWDKGECPYCRAKAKADAGFEIMQKLGIEYCFHDVDLIEDCDDIAEYFARM KDITDYLLEKMKETGINKNLWGTANVFGHKRYMNGAGTNPOFVVAAVAQIKNALDARIK LGGSNYVFWGGREGYITLLNTQMQREKDLAKLLTAARDYARAKFGKGTFLIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKNIEVNHATLAGHTFEHETVARENGFLGSGANR GDAQNGWDDQFPVDPYDLTQAMMQVLLNGGFGNGGTNFDKLRNSTDPEDIFIAHISA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
1754MI1_001	<i>Prevotella</i>	DNA	77	MDAMAHALLNAAAALLESPLCQMVKERYASFDGGLGKQFEKGKATLEDLYEYAKVQGEPPV VASGKQELYEYLLNLVAVK ATGGCAAAAGAGATATTTCCCTTTACCGGTAAAGATTCCTTCGAAAGGAAAGACAGTAAG AAATGTAATGGCTTCCACTACTACGAGCCTGAGAAGGTGCTGATGGGAAAGAAATGAAG GACTGGCTGAAGTTCCGCTATGGCTTGGTGGCATACACTGGGTGGCGCTTCGTGCTGACAG TTTGGTGGTCAGACTCGTTTCATACGAGTGGCAAGGCTGCTGACGCTTTCAGCGCGCT AAGGATAAGATGGAGCTGGCTTCGAGATCATGGACAAGCTGGGCATCGAGTACTTCTGC TTCCACGATGTTGACCTCGTTGAAGAGGGTGACACCATCGAGGAGTATGAGGCTCGCATG AAGGCCATCACCGACTACGCTCAGGAGAAAGATGAAGCAGTCCCCAACATCAAGCTGCTC TGGGTACCGCAACGTAATCGGTAAACAAGCGTATGCTAACGGTGTCTTACCAACCCC GACTTCGACGTAGTGGCTCGGCCATCGTTTCAGATCAAGAAGCTATTGATGCTACCATC AAGCTGGTGGTACCAACTATGTTCTGGGTGGTCTGAGGGCTATATGAGTCTGCTG AACACCGACGAGAAGCTGAGAAGAGCACATGGCTACTATGCTGACCATGGCTGGCGAC TATGCTCGGGCCAAAGGATTCAGGGTACATTCCTCATTCAGCCGAAGCCCATGGAGCCC AGCAAGCACAGTATGATGTGATACAGAGACCGTTATCGGCTTCCTGAAGGCACACAAC CTGGACAAGGACTTCAAGGTGAACATCGAGGTGAACACCGCTACACTCGCTGGTCAATACC TTTCAGCACGAGCTGGCTTGGCTGTTGACGCTGGTATGCTTGGTTCATTCGACGCTAAC CGTGGTATGCTCAGAACGCTTGGGATACCGACAGTCCCATCGACAACCTACGAGCTG ACACAGGCTATGCTCGAGATCATCGCAATGGTGGTCTGGCAATGGTGTACCAACTTC GATGTAAGATCCGTCGTAACAGCACCGACCTCAGGATCTCTTCATCGCTCACATCAGT GGTATGGATGCTATGGCACGGCTCTGATGAACGCTGCTGACATCCTTGAGAACTCTGAG CTGCCCGCAATGAAGAAAGGCTCGCTACGCAAGCTTCAGCAGGGTGTGGTAAGGACTTC GAAGATGGCAAGCTGACCCCTTGACGAGGTTTACGAGTATGTAAGAGGTGGGTGAGCCC AAGCAGACTTCTGTTAAGCAGGAGAAGTACGAGACCATCGTTCCTCTATGCAAAATAA MAKEYFFFTGKLPFEGKDSKNVMAPHYIYEPEKVMGKMKDWLKFAMAWWHITLGGASADQ FGQTRSYEDWKAGDAVQRAKMDAGFEIMDKLGEYFCFHDVDLVEEGDTIEYEARM KAITDYAQEKMKQFPNLIKLLWGTANVFGNKRANGASTNPDVDVVARAIVQIKNAIDATI KLGGTNYVFWGGREGYMSLLNTDQREKEHMAITMLTMDARYARAKGFGTFLIEPKMEP SKHQYDVTETVIGFLKAHNLDKDFKVNIEVNHTLAGHTFEHELACAVDAGMLGSTDAN RGDAQNGWDTQFPIDNYELTQAMLEIIRNGGLNGGTTNFDKIRRNSTDLEDLFIAHIS GMDAMARALMNAADILENSELPAKMKARYASFDQGVGKDFEDGKLTLEQVVEYKRVGEP KQTSKQEKYETIVLYAK ATGGCAAAAGAGTATTTCCGTTTACCGGTAAAGATTCCTTCGAAAGGAAAGACAGCAAG AACGTAATGGCTTTCATTAATGATGAGCCTGAAAGGTGCTGATGGGCAAGAAATGAAG GATTGGCTGAAATTCGCCATGGCTTGGTGGCACACCTCGGTGGAGCCGCGCCGACGAG TTCCGTGGACAGACCCGAGCTATGAGTGGGACAGGCCGAGGATGCCGTACAGCGTGTCT AAGGACAAGATGGACCGCGCTTCGAGATCATGGACAACCTGGGCATCGAGTATTTCTGC TTCCACGATGTCGACCTCGTCGACGAGGTGCTACCGTTGAGGAGTATGAGGCTCGCATG AAAGCCATCACCGACTATGCCCCAGGTCAAGATGAAGGAATCCCCAACATCAACATGCTC TGGGGCACCGCCCAACGTTTCGGCAACAGCGTTATGCCACGCGCTCCACCAACCCC GACTTCGACGTGGTGGCACGCGCTATCGTTTCAGATCAAGAAATGCCATCGACGCTACCATC
1754MI1_001	<i>Prevotella</i>	Amino Acid	78	
1754MI3_007	<i>Prevotella</i>	DNA	79	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				AAGCTCGCGGTGAGAACTACGTGTTCTGGGCGGACGCGAGGGCTACATGAGCCTGCTC AATACCGATCAGAAACGTGAGAAGGAACACATGCCACCATGCTCACCATGGCGCGAC TATGTCGAGCAAGGATTCAGGGCACCTTCTCATCGAACCCAAACCATGGAGCCT TCCAAGCACAGTATGATGTCGACACCGAGACGGTTCATCGGCTTCTCGCGGCCCAAC CTCGACAAGGACTTCAAGGTGAACATCGAGTCAACACCGCACGCTCGCGGCCACAC TTCGAGCAAGAACTGGCTGCGCTGACGCGGATGTCGGCAGCATCGACGCAAC CGCGGACGACAGAAAGGCTGGATACCGACAGTTCGCCATCGAACACTACGAATG ACACAGGCCATGCTGGAGATCATCCGCAATGGGGCTCGGCAATGGTGTACCACTTC GACGCCAAGATCCGTCTAACAGCAGCAGACCTCGAAGATCTTTCATCGCTCACATCAGC GGTATGGATGCCATGGCTCGCGCTGCTCAACGCCCGCCATCCTCGAGGAGAGCGAA CTGCGCGCATGAAGAAGGCCCGCTACGCTTCTTCGACGAAGGTATCGGCAAGGACTTC GAAGCGGCAAACTCACCTCGAGCAGGTTTACGAGTACGCAAGAGGTAGCGGAGGCC AAGCAGACTCCGGCAAGCAAGAGATCGAGACCATCTGGCTCTCTACAGCAATAA
1754MI3_007	<i>Prevotella</i>	Amino Acid	80	MAKEYFPFTGKIPEEGESKNVMAFYHYPEKVVVMKKMDWLKFAMAWHHTLGGASADQ FGGQTRSYEWDKAEDAVQRAKDKMDAGFEIMDKLIGIEYCFHDVLDVDEGATVEEYFARM KAITDYAQVKMKEYPNILKLLWGTAHVFGNKRYANGASTNPDVVARALVQIKNAIDATI KLGGQNYVFWGGREGYMSLLNTDQREKEHMAITMLTMARDYARSKGFKGTFLLIEPKMEP SKHQYDVTETVIGFLRAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSTDAN RGDAQNGWDQDQFPIDNYELTQAMLEIIRNGGLNGGTFNDAKIRRNSTDLEDLFTAHIS GMDAMARALLNAAAIIESEL PAMKKARYASFDEGIGKDFEDGKLTLEQVVEYEGKKVGEP KQTSKGQKEYETIVALYSK
1754MI5_009	<i>Prevotella</i>	DNA	81	ATGAAAGAGTATTTCCCGCAAAATGGAAGAATTCCTTCGAGGGACCGAGAGCAAGAGT CCATTTGGCGTCCATTATATGAGCGGATCGCATGGTGTCTCGGAAAGAGATGGAGAT TGGCTGAAATTCGCCATGGCATGGTGGCACACCTTGGCCAGGCCAGCGCGCACCATTC GGCGACACACGTGATACGAGTGGGATAGGCTGGAGATCGGATACAAAGGGCAAG GATAAGTGGACGCCGATTCGAGATCATGGAAATTTGGGTATCAAGTACTTCTGCTTC CATGATGTGGATCTCGTCGAGGAAGCTCCCAACCATCGCCGAATATGAGGAGCGTATGAG ATCATCACGACTATGCGCTCGAGAAGATGAAGCCACTGGCATCAAACTCCTTTGGGT ACAGCCAATGTTTCGGACATAAGATATATGAATGGGCGGCCCAACCCGGAGTTC GGTGTGTCGCCAGGGTGTCTCCAGATCAAGAACCGGATCGACGCCACCATCAAGCTG GGAGGAACAAACTATGTCTCTGGGTGGCGGAGGCTACATGAGCTGCTCAACACC CAGATGCAAGGAGGAAGGACCATCTCGCCATATGCTCAAGGCTGCTGCTGACTATGCT CGCGCAAGGATTCAGGGCACATTCCTCATCGAGCCGAGCCGATGGAACCTACTAAG CATCAGTACGATGTCGACACTGAGACGCTGATCGGCTTCTCGCGCAACGGTCTTGAC AAGGATTTCAAGGTCAACATCGAGTCAATCACGCCACTCTTCGGGTTCACACTTCGAG CATGAGCTCGCGTGGTGTTCGACAATGGTCTCTTGGCTCAATCGATCGGAACAGGGGA GATTATCAGAACGGTTGGGACACCGACCGAGTTCCTCTGTTGATCTCTTGTATTGACCCAG GCCATGCTCCAGATCACTCCGTAAACGAGGCCCTCGGTAAATGGTGGATCCAACTTCAGGCC AAGCTTCGCGGTAACTCCACTGATCTGAGGATATATTCATGCCCCATATTTGCGGTATG GACGCTATGGCCAGGCTCTCTTGGCGCGCGCATGATCGGAGGAGTCTCCTATCCCG GCTATGGTCAAGAGCGGTACGCTACGCTTCGACGAAGGTGAGGGCAAGAGATTCGAGGAT

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
1754MI5_009	<i>Prevotella</i>	Amino Acid	82	GGTAAGATGAGTCTGGAGGAACCTTTGATTACGCGAAGACTCACGGAGAGCCCGCCAG AAGAGTGGCAACACAGAGACTCTACGAAACCCITGTCAACATGTACATCAATAA MKEYFPQIGKIPFEGPESKSPIAHYEYEDRMVLGKRMEWLKFMAMWHTLGGASDQF GGQREYEWKAGDPIQRAKDKMDAGFEIMEKLGIKYFCFHDVLDVEEPTIAEYERMR ILTDYALEKMKATGILKLLWGTAHVGHKRYMNGAATNPEFVVARAAVQLKNAIDATIKL GGTNYVFWGREGYMSLINTQMQRKDLANMLKAARDYARAKGFGTFLIEPKMEPTK HQYDVTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELAVAVDNGLLGSI DANRG DYQNGWDDQFPVDLFDLTQAMQLIIRNGGLNGGNSFEDAKLRNSTDEPDI FIAHICGM DAMARALIAAAAI VEE SPI PAMVKERYASFDEGEGRKFEDGKMSLEELVDYAKTHGEPAQ KSGKQELYETLVNMYIK
5586MI1_003	<i>Prevotella</i>	DNA	83	ATGGCAAAAGAGTATTTCCGTTTACCGGTAAAGATTCCTTCGAGGGAAAGACAGTAAG AAATGTAATGGCGTTCCACTACTACGAGCCCGAGCGGTGGTAATGGGCAAGAAGATGAAG GAGTGGCTGAAGTTTGCCATGGCTGGTGGCACACGCTGGTGGAGCGAGTGCACCCAG TTTGGCGGACAGACCCGAGCTACGAGTGGGACAAGGTGAAGACGCGTGCAGCGTGCC AAGGACAGATGATGCGCGCTTCGAGATCATGACAAAGCTGGGCATCGAGTATTTTCG TTCATGATGTCGATCTGTTGACGAGGTGCCATGTGAGGAGTATCAGGCTCGCATG CAGCCATCACCGACTATGCGGAGGAGAAGATGAAGCAGTATCCTGCCATCAAGCTGCTG TGGGTACGGCCAATGTTTGGCAACAAGCGTATGCCAACGGTGCTTACCAATCCC GACTTCGATGTGGTGGCGCGCATCTGTCAGATTAGAAATGCCATTGATGCCACCAIC AAGCTGGCGGCGAGCAACTATGTGTTCTGGGCGGTTCGCGAGGGCTACAIGTCGCTGCTC AACACCGACAGAAAGCTGAGAAGAACACATGCCCGGATGCTGACCATGGCCCGGAC TATGCCCGCTCGAAGGCTTCAAGGGCAACTTCTGATTGAGCCCAAGCCCATGGAGCGG TCGAAGCATCAGTACGAGTGGACACCGAGACGGTTATCGGATTCCTCCGCGCATGGC CTTGACAAGGACTTCAAGGTGAACATCGAGTGAACCATGCCACGCTGCCCGGTCTATCC TTCGAGCACGAACTGGCTGCGCGTAGATGCGGCATGCTGGGCGAGCATGATGCCAAC CGCGGCGCACAGAACGGATGGACACCGACAGTCCCATCGCAACTATGAGTTG ACACAGGCCATGATGGAGATTAATCCGCAATGGCGGTCTGGGTCTTGGCGGTACCAATTTC GATGCCAAGATTCCCGTAACTCCACCGACCTGGAAGACCTTTCATGCCCCACATCAGT GGCATGGACCGCATGGCTCGTGGCTCCTTAATGCTGCCACATTCGGAGAACAGCGAA CTGCCCCCATGAAGAAAGCGCGTACGCCCTCGTTCGACAGTGGCATGGGCAAGGACTTC GAGGACGGCAAACTGACCTTGAGCAGGTTACGCAAAAGCAAAAGTTCGGCGAACCT AAGCAGACTCCGGCAAGCAGGAGAAGTACGACCATCTGCGGTCTCTATGCCAAGTAA MAKEYFPFTGKIPFEGKDSKNMAPHYIEPERVVMGKKMEWLKFMAMWHTLGGASADQ FGGQTRSFEWDKAEDAVQRAKDKMDAGFEIMDKLIGIEYCFHDVLDVDEGATVEEYFARM QAITDYAQEKMKQYPAIKLLWGTAHVGNKRYANGASTNPDVDVARAIVQIKNAIDATI KLGGSNVFWGREGYMSLINTDQREKEHMAHMLTMDARYARKGFGKGNFLIEPKMEP SKHQYDVTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDAN RGDAQNGWDDQFPIDNYELTQAMMEIIRNGGLGCGTGNFEDAKIRRNSTDLEDLFTAHIS GMDAMARALLNAADILENSEL PAMKKARYASFDGSMGKDFEDGKLTLEQVYIEYGGKVGEP KQTSKGQKERYETIVALYAK
5586MI1_003	<i>Prevotella</i>	Amino Acid	84	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI2_006	<i>Prevotella</i>	DNA	85	ATGGCAAAAGAGTATTTTCCGTTTACAGGTAAATTCCTTCGAAGGAAAGACAGTAAG AACGTATATGGCTTTCACACTACTACAGCCCGGAAAGGTCTGTCGATGGGAAAGAAATGAAA GACTGGCTGAAGTTTCGCATGGCTGGTGGCACACATGGGTGGCGCCAGCGCCGACGACG TTGGCGCCAGACACGACGCTATGAGTGGGCAAGGCTGCCGATGCCGTGACGCGGCA AAGGACAAAGATGACGCGGCTTCGAAATCATGACAAAGCTGGGCATCGAGTATTTCTGC TTCCACGCGTGGACCTCGTTGAGGAGGAGCACCACATGAGGAGTATGAGGCCCGCATG AAGGTATCACCCGATATGCCAGGAGAAATGAAACAGTATCCAGCATCAAGCTGCTC TGGGGACCGCCAATGTGTTTGGCAACAAGCTACGCCACGCGCCGACGACCAACCCC GACTTCGACGTCGTGGCCGTCGATCGTGCAGATCAAGAACGCCATCGATGCCACCATC AAGCTGGCGGCAACCACTACGTGTTCTGGCGGAGCGGAGGCTACATGAGCCTGCTC AACACCGACAGAGCGGAGAGAGCACATGGCCACCATGCTCACCATGGCCCGGAC TACGCCGCGCAAGGATCAAGGCGACCTTCTCATCGAGCCCAAGCCCATGGAGCGG TCGAAGCACAGTACGAGTGGACACCGAGACCGTCTCATCGGTTCTTCTGAAGGCCACGCT CTGGACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACGCTGGCCGGCCACAC TTCGAGCATGAGTGGCTGGCGCTCGCGCTCGACCGGTATGCTGGCGAGCATCGATGCCAAC CGCGGACGCCCCAGAACGGCTGGACACCGACCGAGTCCCATCGACAACTTCGAGCTC ACCAGGCCATGATGGAATATCCGCAACGGCGGCTCGGCAACGGCGGACCAACTTC GACGTAAGATCCGCGCAACTCCACCGACCTCGAGGACCTTCTCATCGCCACATCAGC GGCATGACGCCCATGGCCGCGCTGATGAACGCTGCCACATTCGGAAGAACGCGAG CTGCCCCCATGAAGAGGACGCTACGCTGAGTTCGACGCGGCTCGGCAAGGACTTT GAGGATGGCAAGCTCTCGCTGGAGAGGTCTACGAGTATGCAAGAGGTGGAAGAGCCC AAGCAGACGCGGCAAGCAGGAGAGTACGAGACCATCTGCGCCCTCTATGCCCCAATAA MAKEYFPFGKIPEEGKSKNMFHYIPEKVMGKKMKDWIKFAMAWHHTLGGASADQ FGGQTRSYEWDKAADAVQRAKDKMDAGFEIMDKLIGIEYFCFHDVLDVEFGATIEEYFARM KAITDYAQEKMKQYPSIKLLWGTTANVFGNKRYANGASTNPDVDVVARALVQIKNAIDATI KLGGTNYVFWGGREGYMSLINTDQREKEHMTMLTMARDYARAKFGKTFLLIEPRMEP SKHQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDAN RGDQNGWDTDQFPIDNFEITQAMMEIIRNGGLNGGTNFDKIRRNSTDLEDLFTAHIS GMDAMARALMNAADIMENSELPAMKKARYASFDAGIGKDFEDGKLSLEQVYIEYKKEVEEP KQTSKGQEKYETIVALYAK
5586MI2_006	<i>Prevotella</i>	Amino Acid	86	
5586MI8_003	<i>Prevotella</i>	DNA	87	ATGGCAAAAGAGTATTTTCGCTTTCACAGGCAAGATTCCTTCGAGGGGAAAGACAGTAAG AACGTATATGGCTTTCACACTACTACAGCCCGGAGCGTGTGTGATGGGCAAGAGATGAAG GAGTGGCTGAAGTTTCGCATGGCTGGTGGCACACATGGGTGGCGCATCGGCCGACGAG TTGGAGGCGACACACGACGCTACGAGTGGGCAAGGCTGCCGACGCGCGCGCGCGCC AAGGACAAAGATGACGCGGCTTCGAGATTATGGAACAAGCTGGGCATCGAGTACTTCTGC TTCCACGATGATAGCTCGTTGAGGAGGGTGAGACCATAGCCGAGTACGAGCGCCGCAIG AAGGAAATCACCGACTACGCACAGGAGAGATGAAGCAGTTCCTCAACATCAAGCTGCTC TGGGGACAGCCCAACGTTTCGGCAACAAGCGTACGCCAACGGCGGCTCGACCAACCCC GACTTCGAGCGTGTGGCACGCGCATCGTGCAGATCAAGAACGCCATCGACGCCACCATC AAGCTCGCGGCTCCAACTATGTGTTCTGGCGGAGCGGAGGCTATGAGCCTGCTC AACACCGACAGAGCGGAGAGAGGACATGGCCACCATGCTCACCATGGCCCGGAC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI8_003	<i>Prevotella</i>	Amino Acid	88	TATGACGCGCCAAAGGATTCAGGGCACATTCTCATCGAGCCGAAGCCCATGGAGCCC TCGAAGCACACAGTACGAGTAGACACAGAGACCGTTCATCGGCTTCTCCGTGCACACGGG CTGGACAAGGACTTCAAGGTGAACATCGAGTAAACACACGACACATGCGCGCCACAC TTTCAGACAGAGCTGGCTTGGCCGTGACCGCTGGCATGCTGGGACGATCGACGCCAAC CGTGGCGACGACAGAACGATGGACACCGACAGTTCCTCCATCGAACACTTCGAGCTC ACACAGGCCATGATGGAATCATCGCAATGGCGGACTGGCAATGGCGGACCAACTTC GACGCCAAGATCCGTCTAAACAGCACCGACCTCGAAAGACCTCTTCATCGCCCCACATCAGC GGCATGGACGCGCATGGACGCGCTGCTCAAGCTGCCACATCCTGGAGCACAGCGAG CTGCCCAGATGAAGAGGAGGCTAGCCAGCTTCGACGAGGATCGGCAAGGACTTC GAAGACGCAAGCTCACACTCGAGAGGTCTACGAGTACGGAAGAGGTCTGAAGAGCCC CGTCAGACACGCGGCAAGCAGGAGAGTACGAGACCATCGTCGCCCTCTATGCCAAGTAA MAKEYFAFTGKIPFEGKDSKNVMAFHYPERVVMGKKMKEWLKFAMAWHHTLGGASADQ FEGQTRSVEWDKAADAVQRAKDKMDAGFEIMDKLGIYFCFHDVDLVEEGETIAEYERRM KEITDYAQEKMKQFPNLIKLLWGTANVFGNKRYANGASTNPEDVVARIVQIKNAIDATI KLGGSNVFWGGREGYMSLINTDQKREKEHMTMLTMDYARAKGFKGTFLIEPKMEP SKHQYDVTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSTDAN RGDAQNGWDTDQFPIDNFELTQAMMEIIRNGGLNGGTFNPAKIRRNSTDLEDLFIAHIS GMDAMARALINAADILEHSELPMKKERYASFDAGIKDFEDGKLTLEQVVEYKKEVEEP RQTSKQEKYETIVLYAK
5586MI14_003	<i>Prevotella</i>	DNA	89	AUGGCAAAAGAGTATTTCCGTTTACTGGTAGATTCTTTCGAGGGAAAGGATAGTAAG AATGTAATGGCTTCCACTATTACGAGCCCGAGAAAGTCGTGATGGGAAAGAAGATGAAG GACTGGCTGAAGTTTCGCAATGGCTTGGTGCAACATACATGGTGGTGCTCGACACGAG TTCCGTGGAGAGACCCGACGCTACGAGTGGAGCAAGGCTGCTGATCCCGTTCAGCGGCC AAGGACAAGATGGACCGCGCTTTGAGATTATCGATAAGCTGGGCATCCAGTACTTCTGT TTCACGATATAGACCTCGTTCAGGAGGCAGATACCATTCGAGAATATCAGGAGCGCAIG AAGCAATTACCGACTATGCTCTGGAGAAGATGAAGCAGTTCCTCCCAACATCAAGTTGCTC TGGGTACCGCTAACGTAATTAGCAACAAGCGTATATGAACGGTGTCTTACCAATCCC GACTTCGACGTGGTGGCGGTGCTTCAGATCAAGAACGCTATTGAIGCAACCAIC AAACTCGGTGGTACCAACTATGATCTGGGGTGGTCTGAGGGTTACATGAGCCTATTG AATACCGACAGAACGCTGAAAGGAGCACATGGCAATGATGCTCGGTATGGCTCGCGAC TAIGCCCGCAGCAAGGATTCAGGTACGTTCTCTCAGCCGAGCGGATGAGGAGCGCC TCTAAGCATCAGTATGATGTCGATCGGAGACTGTGATGGTTTCTGAGGACACCGT CTGGACAAGGACTTCAGGTGAACATCGAGGTGAACACCGCTACACTGGCTGGTCACTACC TTCGAGCATGAGCTGGCTGGCTTTCGCGAGGTATGCTGGGCTCTATCGACGCTAAC CGCGGTGATGCCCCAGAACGGCTGGATACCGACAGTTCCTCCATCGAACATCAGAGCTG ACACAGGCTATGATGGAATCATCCGCAACGGTGGTCTGGGCAATGGTGGTACCACTTC GACGCTAAGATCCGCGTAACCTACCGACCTCGAGGATCTGTTTCATCGCTCATCAGT GGTATGGATGCTATGGCCGTGTTTGTGATGTCGCGACATCTGGAGAACTCTGAA CTGCCCCGCTATGAAGAGGCCCCGCTACGCCAGTTCGACACGGTTCGGTAAGGACTTC GAGGATGGCAAGCTGACCTTCGAGCAGGTTTACGAATATGGAAGAAAGTGAAGAGCGG AAGCAGACCTCTGGCAAGCAGGAGAAATACGAGACCATCGTGTGCTCTGTATGCTAAATAA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI14_003	<i>Prevotella</i>	Amino Acid	90	MAKEYPFTGKIPFEGKDSKNVMAFYHPEYKVMGKKMKDWLKFAMAWHHTLGGASADQ FGGETRSYEWSKAADPVQRAKMDAGFEIMDKLGIYFCHDIDLVEADTIAEYERM KAITDYALEKMQFPNIKLLWGTANVFSNKRMYNGASTNPDVDVARAIVQIKNAIDATI KLGGTNYVFWGGREGYMSLLNTDQREKEHMAMLGMARDYARSKGFKGTFLIEPKMEP SKHQYDVTETVIGFLKAHGLDKDFKVNIEVNHTLAGHTFEHELACAVDAGMLGSTDAN RGDAQNGWDTQDFPINNYELTQAMLEIIRNGGLGNGGTNFDKIRRNSTDLEDLFTAHIS GMDAMARALLNAAADILENSELPMKKARYASFNDGIGKDFEDGKLTTFEQYVEYGKKVEEP KQTSKGQEKYETIVLYAK
5586MI26_003	<i>Prevotella</i>	DNA	91	ATGGCAAAAGAGTATTTCCGTTTACCGGTAATAATTCCTTCGAGGGGAAAGGACAGTAAG AAATGTAATGGCTTTCACCTACTACGAGCCTGAGCGCGTAGTCATGGGAAAGAGATGAAG GATTGGTTGGCATTTGCAATGGCTTGGTGGCACACACTGGGTGGCGCTTCTGCCGACGAG TTTGGTGGTCAGACCCGAGTTACGAATGGACAAGGCTGCTGATGCTTTCAGCGTGCT AAGGACAAGATGGATGCCGGCTTCGAGATTATGGATAAGCTGGGAATCGAGTTCCTTCTGC TGGCACGATATCGACCTCGTTGAAGAGGGTGAGACCATTTGAAGAGTATGAGCGCCGATG AAGGCTATCACCGACTATGCTCTTGAGAGATGCAGCAGTATCCCAACATCAAGAACCTC TGGGGAACAGCCCAATGTGTTTGGCAACAAGCGTTATGCCAACGGTGCCAGCACAAACCCA GACTTTGACGTCGTTGCTCGTATCGTACAGATTAGAAITGCTATCGACGCTACTATC AAGTTGGTGGTCAGAAATATGTGTTCTGGGTGGCGGTGAGGCTACATGAGCCTGCTC AATATGACCAAGAGCGTGAGAAGGAGCACATGGCTACAAATGCTGACCATGGCACGCGAC TATGCCCGCAGCAAGGATTCAGGGTAACTTCCTCATTTAGCCCCAAGCCCATGGAGCCG TCAAAGCACAGTATGATGTTGACACCGAGACCGTATGCGGTTTCTGCGTGCCCAACAC CTTGACAGGATTTCAAGGTAATATCGAGTTAACCACTACTCTGGCTGGTCTACT TTCGAGCAGAACTGGCATCGCTGTTGACGCTGGTATGCTGTTGTTTCTATCGATGCTAAC CGTGGTATGCCAGAAATGGCTGGGATACCGACAGTCCCATCAACAACATGAACTC ACTCAGGCTATGCTTGAGATCATCCGTAATGGTGGTCTGGGCTTGGCGGCACAACTTC GATGCCAAGATTCGTCGTAACCAACAGATCTTGAGGATCTCTTCATCGCTCACATCAGT GGTATGGATGCCATGGCCGCTGCTGCTGAATGCTGCTGCTATCTTGAGGAGAGCGGAG CTGCCTAAGATGAAGAGGAGCGTTATGCTTCTTCGATGCGGTTATCGTTAAGGACTTC GAGGATGGCAAGCTTACCCTTGACGAGGCTTACGAGTATGTAAGAGGTTGAGGAGCCC AAGCAGACTTCAGGCAAGCAGGAGAGTACGAGACCATCGTGTGCTCTGTATGCAAAATAA MAKEYPFTGKIPFEGKDSKNVMAFYHPEYKVMGKKMKDWLRFAMAWHHTLGGASADQ FGGQTRSIEWDKAADAVQRAKMDAGFEIMDKLGIYFCHDIDLVEEGETIEEYERRM KAITDYALEKMQFPNIKLLWGTANVFSNKRMYNGASTNPDVDVARAIVQIKNAIDATI KLGGQNYVFWGGREGYMSLLNTDQREKEHMAMLGMARDYARSKGFKGTFLIEPKMEP SKHQYDVTETVCGFLRAHNLDKDFKVNIEVNHTLAGHTFEHELACAVDAGMLGSTDAN RGDAQNGWDTQDFPINNYELTQAMLEIIRNGGLGNGGTNFDKIRRNSTDLEDLFTAHIS GMDAMARALLNAAADILEESELPMKKARYASFNDGIGKDFEDGKLTTFEQYVEYGKKVEEP KQTSKGQEKYETIVLYAK
5586MI86_001	<i>Prevotella</i>	DNA	93	ATGAAACAGTATTTTCCCGCAGATTGGAAAGATACCCCTTCGAGGGTGTAGAGAGCAAGAT GTGATGGCTTTCACCTATATGAGCAGAAAGAGTAGTCATGGGCAAGCCTATGAAGAA TGGCTGCGCTTCGCTATGGCGTGGTGGCACACGCTGGGCGAGCGGACCCCTTC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI108_002	<i>Prevotella</i>	Amino Acid	96	TTTCGAGCACGAACCTGGCTTGGCCGCTTGATGCCGGCATGCTCGGCTCTATCGATGCCAAC CGCGCGGAGCTCAGAACGGCTGGACACCGACAGTCCCATCGAACATACGAGCTC ACTCAGGCCATGATGGAATCATCGTAAATGGGGTCTGGCAACGGCGCACGAATCTC GATGCCAAGATCCGTCTGTAACAGCACCGACCTCGAGGACCTCTTATCGCTCACATCAGC GGCATGGATGCCATGGCACCGCTCTGATGAACGCTGCTGCCATCTTCAAGAGAGCGAG CTCGCCGCATGAAGAGCGCCGCTATCTTCTCGACAGGATCGGCAAGGACTTC GAGGACGGCAAGTTGTCTATGAGCAGGTCTACGAATATGTAAGAAGGTGAGGAGCCC AAGCAGACCTCGGGCAGCAGGAGAGATCAGACCATCGTGGCCCTCTATGCCAAGTAA MAKEYFPFTGKVPFEGTESKNVMAFHYEPEKVMGKKMDWLKFAMAWHHTLGGASADQ FGGQTRSIEWDKAADAVQRAKDKMDAGFEIMDKLGIEYFCFHDVDLVEEGETVAEYEARM KVITDYALEKMQQFPNLIKLGWTANVFVGHKRYANGASTNPDVDVVARAIVQIKNAIDATI KLGGTNVFWGGREGYMSLLNTDQKREKEHMAIMLTARDYARAKGFKFTFLIEPKMEP SKHQYDVPTEIVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DAN RGDAQNGWDIDQFPIDNYELTQAMMEIIRNGGLNGGTFNEDAKIRRNSTDLEDLFIAHIS GMDAMARALMNAAILLESEL PAMKKARYASFDEGIGKDFEDGKLSLEQVVEYKKEVEEP KQTSKQEKYEITIVLYAK
5586MI182_004	<i>Prevotella</i>	DNA	97	ATGGCAAAAGAGTATTTCCGTTTGTGGTAAGATTCCTTCGAGGGAAAGGATAGTAAG AAATGTAATGGCTTTCACATATTACGAACACAGAGAAGTTCGTGATGGGAAAGAAGATGAAG GACTGGCTGAAGTTCCGCATGGCATGGTGCCACACATGGCAGCAGCCAGTCCGACCCG TTTGGAGGTACAGACCCGAGCTACGAGTGGGACAAGGTGACGATGCTGTGACGCGCGCA AAGGACAAAGATGGATGCCGATTTGAGATCATGGACAAGCTGGGCATCGAGTACTTTCG TTCCACGATGTAGACCTCGTTGAGGAGGAGCAACTGTGAGGAGTACGAGGCTCGCATG AAGCCATCACCGCATGTGATTGGAGAAGATGAAGAAGTATCCCAACATCAAGAACCTC TGGGTACAGCCAATGTATT CAGCAACAAGCGTATATGAACGGTCCAGCACCAACCCC GACTTCGACGTTGTTGCACGTGCCATCGTACAGATAAAGAACGCCATTACGCTACCATC AAGCTCGCGGT CAGAACTACGTGTTCTGGGCGGACGTGAGGGATACATGAGCCTGCTC AACACCGACAGAACCGGAGAGAGCACATGGCAACCATGCTGACCATGGCTCGCGAC TACGCTCGCAAGAAGCGTTTCAAGGCGACATTCCTCATCGAGCCTAAGCCCATGGAACCC TCAAAGCACAGTACGAGTAGACAGAGACCGTATGCGGTTTCTCCGCGCCCATGGT CTTGACAAAGGATTTCAAGGTGAACATGAGGTGAACCCAGCTACCTCGCCGGCCACAC TTTGAGCATGAACCTGGCTTGCCTCGCATCGACACGGCATGTCGGCAGCATGATGCCAAC CGCGGACGCTT CAGAACGGCTGGACACCGACAGTCCCATCGAACATACGAGCTG ACTCAGGCCATGCTCGAAATCATCCGCAACGGTGGTCTGGGCAACGGCGGTACCAACTTC GACGCCAAGATCCGTCGTAACCTACCGACCTCGAGGATGTGTTTCTCGCCCATCATCAGC GGTATGGACGCCATGGCAGCTGCTGCTCAATGCAGACGCCATATCTGGAGGAGAGCGAG CTGCTTGCATGAAGAAGGAGCGTTACGCCAGCTTCGACAGCGGCATCGGCAAGGATTC GAGGACGGCAAGCTCACATTGAGCAGGCCTATGAGTATGTAAGAAGGTGAGGAGCCA AAGCAGACCTCTGGCAGCAGGAGAGATGAGCATATAGAGCCCTCTACGCTAAGTAG
5586MI182_004	<i>Prevotella</i>	Amino Acid	98	MAKEYFPVFGKIPFEGKDSKNVMAFHYEPEKVMGKKMDWLKFAMAWHHTLGGASADP FGGQTRSIEWDKAADAVQRAKDKMDAGFEIMDKLGIEYFCFHDVDLVEEGETVAEYEARM KAITDYALEKMKKEYPNLIKNLWGTTANVFSNKRVMNGASTNPDVDVVARAIVQIKNAIDATI

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI193_004	<i>Prevotella</i>	DNA	99	KLGGQNYVFWGGREGYMSLLNTDQKREKEHMAWMLTMDYARKNGFKGTLFIEPKMEP SKHQYDVTETVCGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDNGMLGSDAN RGDVQNGWDTDOFPIDNYELTQAMLEIIIRNGGLNGGTFNDAKIRRNSTDLEDLFIAHIS GMDAMARALLNAAAIIEESELPAKKERYASFDSGIGKDFEDGKLTLEQAYEYKKEVEEP KQTSKQELYETIVALYK ATGACTAAAGAGTATTTCCCTACCATGGCAAGATTCCCTTTGAGGGACCTGAAAGCAAG AACCCGCTTGCAATTCATATAGCCCGACCGCTGCTGTCATGGCAAGATGAAA GACTGGTGGCTTTCGGCATGGCTGGTGGACACCTGGCCAGCCCTCCGGCCGACAG TTCCGGCGGACAGACCCCACTATGCTGGGATGATCCGGATGCGGATGACACGTGCC AAAGCCAGGCGGACGCGGTTTCGAAATCATGAGAAATGCGGATGAAATTCCTGTC TTCCAGCATCGACCTGGTGGATGCGATGAAATCGCCGATGACGAGCCCGGATG AAGGACATCACCGACTATCTGCTCAAGATGAAAGAGACCGGATCAAGAACCTTTGG GGAACGGCAACGATTTGGCCACAGCGTACATGAACGGCGCCGACCAACCCGAT TTCCAGCTGTGGCGGTCGGCTCCAGTCAAGAACCGCATCGACGCCACCATCAAG TTGGCGGTGAGAACTATGTGTTCTGGGGCGGCTGAAGCTACCAAGCCCTGCTCAAT ACCCAGATCAGCGCGAAGAACACATGGCGGTATGTTGGCACTGGCCCGCGACTAT GGCGTGACACGGTTTCAAGGACGTTCCATCATCGAGCCCAACCGATGGAGCCGACC AAGCACAGTACGATCAGGATACGGAAACCGTCACTGGCTTCTCGCGCGCATGGCCTC GACAAGACTTCAAGTCAACATCAGGTGAACATGACCTGGCGGCGCACACCTTC GAGACGAGTGGCTGGCGCTGACACCGGATGCTGGGAGCATCGACGCCAACCGG GGTGTGCCAGAACGGCTGGACACCGACCTCCGATCGATCGATACTAGCTGACG CTGGCATGCTCCAGATCATCCGACCGGCGCTGGACCGGCGGCTCGAACTTCGAT GGAGCTGGTCCGAACTCCAGCATCCGGAAGATATCTTCATCGGACATCAGCGCC ATGATGCCATGGCGCGGCTGCTCAATGCTGTCGCCATCTCGAGGAATCGCCCATC CCGGCATGGTCAGGAACGTTACGCTCGTTCGACACGGAAGAGGAGGGAATATGAG GAAGGAGGCTGTCTCGAAGACATCGTGGCTATGCCAAGCCCAACCGGCAACGAAA CAGATTTCCGGCAAGCAGGAACCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG MTKEYFPTIGKIPFEGPESKNPLAFHYEPDRLVMGKMKDWLRFAMAWWHITLQASGDQ FGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCHDIDLVEDADEIAEYARM KDI TDYLLVKMKETGINKLWGTANVFGHKRYMNGAATNPDPDLARAQVQIKNAIDATIK LGGQNYVFWGGREGYQTLNTQMOREKEHMRMLALARDYGRAHGFKGTLFIEPKMEPT KHQYDQDETIVGFLRRHLDKDFKVNIEVNHATLAGHTFEHELACAVDHGMLGSDANR GDAQNGWDTDOFPIDNYELTQAMLEIIIRNGGLAPGGSFNAKLRNSTDPEDIFIAHISA MDAMARALNVAIIIEESPAMPVRERYASFDGKREYEGRLSLEDIVAYAKAHGEPK QISGKQELYETIVALYK
5586MI193_004	<i>Prevotella</i>	Amino Acid	100	
5586MI195_003	<i>Prevotella</i>	DNA	101	ATGGCAAAAGAGTATTTCCCGCAGATCGAAAGATCGGCTTGAGGGTCTGCAAGCAAG AACCCGCTGGCATTCATATATGACCGCGAGCGCTGGTGGTATGGGTAAACCCATGAAA GACTGGTTAAATTCGCCCTCGCGTGGTGGCACAGCTCGCCAGGCGCTCGGCGACCGG TTCCGGCGGACAGACCGCTCCTACGAGTGGGACAGGGCGAATGCCCTACTGCCGCGCC CGCGCAAGCGGACCGGCTTCGAGATCATGAAAGCTCGGATCGGCTATTTCTGTC TTCCACGACGTCGACCTCATCGAAGACACGACGACATCGCCGAATATGAGCGCGGCTC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				AAGGACATCAGGACATACCTGCTCGAAAGAGTGCAGGAAACCGGCATCAAGAACCCTTGG GGCAGGGCAATGTCTTCGGTCAAGCGCTACATGAACGGCGCCGACCAATCCGCAG TTCGACATCGTCGCCCGCTCGCTCCAGTCAAGAACCGCTCGACGCCACCATCAAG CTCGTGGCTCGAATAGTCTTCTGGGGCGCCGGAAGTATTAACAGCTGCTCAAC ACCCAGATCAGCGCGAGAAACACCTCGCCAAGCTCCTCACCGCCCGCGGACTAT GCCCCGCAAGGGCTCCAGGACCTTCTCATGAGCCCAAGCCGATGGAGCGGACC AAGACCATACGATGTCGACACGGAGACTGTAATCGGATTCCTCCGCGCAACGGACTG GACAAGGACTTCAAGTCAACATCAGAGTCAACACCGCCCTCGCCGCCATACCTTC GAGCATGAGCTGACCGTCGCCCGGAGAACGGATTCCTCGCAGCATCGACGCCAACCGC GGTGACGCCAGAACCGCTGGACACCGACCTTCCCGTGACCGCTACGACCTACCC CAGGCCATGATCAGGTGCTCCTGAACGGGGTTTCGGCAACGGCGGCACCAATTCGAC GCCAAGTCCGTCCGAGCTCCAGCATCCGAGGACATCTCATCGCCACATCAGCGCG ATGACGGCATGGCCACCGCTGCTGAACCGCGCGCTTCTCGAGGAGAGCGCGCTG CCCGGATGTTCAAGGAGGTTACGCTCCTTCGACAGCGCTTCGGCAAGCAGTTCGAG GAGGAAAGGCCACGCTGGAGGACCTTACGACTACGCCAAGGCCCATGGCGAGCGCTC GCCGCTCCGCAAGCAGGAACTGTGAACTTACCTGAATCTGTATGCAAGTAA
5586MI195_003	<i>Prevotella</i>	Amino Acid	102	MAKEYFPQIGKIGFEGPASKNPLAFHYIDAERVVMGKPMKDFKFAAWHSLGQASGDP FGGQTRSYEWDKGECPYCRARAKADAGFEIMQKLGIGYFCFHDVDLIETDDIAEYEARL KDITDYLLERMQETGINKNLWGIANVFHKRYMNGAGTNPQDIVARAAVQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMQRKDLAKILLTAARDYARAKFGQFTFLIEPKMEPT KHQYDVTETVIGFTRANGLDKDFKNIEVNHTLAGHTFEHETVARENGFLGSTDANR GDAQNGWDIDQFPVDAYDLTQAMMQVLLNGGFGNGGTNFDALRRSSTDPEIDIFIAHISA MDAMAHALLNAAAILLESPLPAMVKERYASFDSGLGKQFEGKATLEDLYDYAKAHGEPV AASGKQELCETYLNLAK
5586MI196_003	<i>Prevotella</i>	DNA	103	ATGACAAAAGAGATATTCCTTACCATCGGCAAGATCCCCTTGAGGGACCCGAGAGCAAA AACCCCTCGCTTTTCAATTAATATGAGCCCGACCGCTGTCATGGGCAAGAAGATGAAA GACTGGCTCGCTTCGCCATGGCTGGTGGCACACCTTGGGCCAGGCTCCGGCGGACAG TTTGGCGGCAGACCCGCACTATGCTGGATGATCGGATGCGGATGCGGATGCGGATGCGG AAAGCAAGGCCGACGCGGTTTCGAAATCATGCAGAAACTGGGCAATGAATCTTCTGCG TTCCACGACATCGACCTGATCGAGGATACCGATGACATCGTCGAGTATGAGGCCCGGATG AAGGACATCACCGACTATCTGTGTTCAAGATGAAAGAGACCGGCATCAAGAATCTCTGG GGAACGCCAACGTTATTCGGGCAAGCGCTATATGAACGGGCTGCCACCAACCCGAT TTCGACGTGCTGGCCGTGCGCGCTCCAGATCAAGAACCGCATCGACGCCACCATCAAG CTGGCGGCAGAAATATGTGTTCTGGGCGGCGTGAAGTACGAGCCTGTCTCAAT ACCCAGATCGAGCGGAAAGAACATGGCGGCTATGTGGCACTAGCCCCGCGCTAT GGCCGTGCACACGGTTTCAAGGCGACGTTCCTCATCGAGCCCAACCGATGAGCGCGACC AAGCACCATGATCAGGATACGAGACCGTCTCGGTTTCTGCGCGCCATGGGCTC GACAAGGACTTCAAGTCAACATCGAGGTGAACCATGTAACCTGGCGGCGCACACCTTC GAGCAGGACTGGCTGGCCGTGACACCGGATGCTGGGAGTATTCAGCGCCACCGC GGTGACGCCGAGAACGGTGGGACACCGACCTTCCGATCGATGAGCTGAGCTGACG CTGGCCATGCTCCAGATCATCCGCAACGGCGGCTGGCACCCGCGGCTCGAACTTCGAT

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI196_003	<i>Prevotella</i>	Amino Acid	104	GCGAAGCTGCGTCCGAACTCCACCGATCCGGAAGATATCTTCATCGGCACATCAGCGCC ATGGATGCCATGGCCCGCCCTGGTCAACGCTGTGCGCATCTTGAGGAATCGCCCAT CCGGACATGGTCAAGAGCGCTACGCTTCGTTGACAGAGAAAGGAGGAGTAGGAA GAGGGAAACTTTCCTCGAGGACCTCGTGGCTATGCCAAGCCCGCAGCGGAAACGAAA CAGATTCCGGCAAGCAGAACTCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG MTKEYFPTIGKIPFEGESKNPLAFHYYPDLVVGKMKDWLRFAMAWHHTLGQASGDQ FGGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDIDLIEDTDDIVEYARM KDI TDYLLVKMKETGINKNLWGTVANVGHKRYMNGAATNPEDVLARAQIKNAIDATIK LGGQNYVFWGGREGYQSLINTQMQREKEHGMRLALARDYGRAHGFKGTFLIEPKMEPT KHQYDQDTEVTIGFLRRHGLDKDFKNVIEVNHATLAGHTFEHELACAVDHGMLGSDANR GDAQNGWDTDQFPIDNYELTLAMLQIIRNGGLAPGGSNFDAKLRRNSTDPEDIFIAHISA MDAMARALVNAVALLEESPIPDVVKERYASFDGKGREYEGKLSFEDLVAYAKAHGEPK QISGKQELYETIVALLYCK
5586MI197_003	<i>Prevotella</i>	DNA	105	ATGACAAAAGAGTATTTCCCTACCATCGGCAAGATCCCCTTTGAGGGACCCGAGAGCAAA AACCCCTCGCTTTTCATTAATGAGCCCGACCGCTGGTTCATGGGCAAGAAAGATGAAA GACTGGCTGCTTTTCGCATGGCTGGTGGCACACCTCGGCGAGGCTCGGCGACAG TTTGGCGGCAGACCCGCTATGCTGGTGGATGATCCGGATTGCCGTATGCACGTGCC AAAGCCAAAGCCGACGCGGTTTCGAAATCATGAGAACTGGGCAATGAATTCCTTCG TTCACGCATCGACCTGATCGAGATACCGATGACATCGAGTATGAGGCGCGGATG AAGACATCACCGACTATCTGTGTTCAAGTCAAGAGACCGGCATCAAGAATCTCTGG GGAACGCCAACGTTATCGGGCAAGCGCTATATGAACGCGCTGCCACCAACCCCGAT TTCGACGTGCTGGCGCTGGCGCCCGACAGATCAAGAACCGCATCGACGCCACCATCAAG CTGGCGGCAGAAATATGTTGTTGGGGGGGGTGAAGGTACAGAGCCTGTCTCAAT ACCCAGATCGAGCGCGAAAGAACACATGGCGGCTATGTTGGCACTAGCCCCGCTAT GGCGTGACACACGGTTTCAAGGGCAGCTCCTCATGAGCCCAACCGATGGAGCGGACC AAGACCATGATCAGGATACGAGACCGTCACTCGGTTCCTGGCGGCTGAGGTACAGAGCCTGTCAAT GACAAGGACTTCAAGTCAACATCGAGGTGAACCATGCTACCTGGCGGCGCCACACCTTC GAGCAGGCTGGCTGGCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT GGTGAACCGCAGGACGGTGGGACCGACCGTCCAGTCCCGATCGATACTATGAGCTGACG CTGGCATGCTCCAGATCATCCGCAACGGCGCTGGCACCGCGGCTCGAACCTTCGAT GCGAAGCTGGTTCGAACTCCAGCATCCGGAAGATATCTTCATCGGCAACATCAGCGCC ATGGATGCTATGGCGCGGCTGGTCAACGCTGTCGCTGCGCATCTTGAAGAAATCGCCCAT CCGGACATGGTCAAGGAGCGCTACGCTTCGTTGACAGCGGAAAGGAGGAGTACGAA GAGGGGAAACTTTCCTCGAGGACCTCGTGGCTATGCCAAGCCCGCAGCGGAAACGAAA CAGATTCCGGCAAGCAGAACTCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG
5586MI197_003	<i>Prevotella</i>	Amino Acid	106	MTKEYFPTIGKIPFEGESKNPLAFHYYPDLVVGKMKDWLRFAMAWHHTLGQASGDQ FGGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDIDLIEDTDDIVEYARM KDI TDYLLVKMKETGINKNLWGTVANVGHKRYMNGAATNPEDVLARAQIKNAIDATIK LGGQNYVFWGGREGYQSLINTQMQREKEHGMRLALARDYGRAHGFKGTFLIEPKMEPT KHQYDQDTEVTIGFLRRHGLDKDFKNVIEVNHATLAGHTFEHELACAVDHGMLGSDANR GDAQNGWDTDQFPIDNYELTLAMLQIIRNGGLAPGGSNFDAKLRRNSTDPEDIFIAHISA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI199_003	<i>Prevotella</i>	DNA	107	MDAMARALVNAVAIIIEESPIPDVMKERYASFDSGKGREYEEGKLSFEDLVAYAKAHGEPK QISGKQELYETIVALYCK ATGACAAAAGAGATATTCCCTACCATCGGCAAGATCCCCTTTGAGGGACCCGAGAGCAAAA AACCCCTCGCTTTTCACTACTATGAGCCCGACCGCTGCTGTCATGGCAAGAAGATGAAA GACTGGCTGGTTTCGGCATGGCTGGTGGCACACCTTGGCCAGGCTCCGGCGGACAG TTTGGCGGCAGACCCGCACTATGCTGGATGATCCGGATGCGGATGCGGATGACAGTGCC AAAGCCAGCCGAGCGCGGTTTCGAAATCATGCAAAATGCGGCAATGAAATTCCTTGC TTCCAGCATCGACCTGATCGAGATACCGATGACATCGTCGATGAGGCCCGGATG AAGGACATCACCGACTATCTGCTGCTCAAGATGAAAGAGACCGGCATCAAGAAATCTCTGG GGAAGGCCAACGTTTCGGGCAAGCGCTATATGAACGGCGCTGCCAACACCCGAT TTCCAGCTGCTGGCCGTGCCCGTCCAGATCAAGAACCGCATCGACCCACCATCAAG CTGGCGGCGCAGAAATTATGTGTTCTGGGCGGCGGTGAAGCTACCGAGCCCTGCTCAAT ACCCAGATCAGCGGAAAGGAAACATGGCGCTATGTCGCACTAGCCCGGCTGCTCAAT GGCGTGACACCGTTTCAAGGACGTTCTTATCGAGCCCAACCGATGAGCCGACCC AAGCACCTAGCATCAGGATACGAGACCGTCTATCGGTTTCTGCGCGCCATGCGCTC GACAGGACTTCAAGTCAACATCAGGTGAACCATGCTACCTGGCGGCCACACCTC GAGCAGAGCTGGCTGGCGCTGACACCGCATGCTGGCAGATTTGACGCCAACCGC GGTGACGCCAGAACGGCTGGACACCGACCTCCGATCCGATGATGAGCTGAGCTGACG CTGGCCATGCTCCAGATCATCCGCAACGGCGCTACCGCGGCTCGAACTTCGAT CGAAGCTCGTCCGAACTCCACCATCCGGAAGATGCTTTCATCGGACATCAGCGCC ATGATGCGCATGGCGCGCTGCTCAACGCTGTCGCCATCTTGAGGAATCGCCCAT CCGGACATGCTCAAGGAGCGCTACGCTTCGTTGACAGCGAAAGGAGGAGTACGAA GAGGGAAACTTCTCTCGAGGACCTCGTGGCTATGCCAAGCCACCGCGAACCGAAA CAGATTTCCGGCAAGCAGGAACCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG MTKEYFTIGKIPFEGPESKNPLAFHYEPDRLVMGKMKDWLRFAMAWHITLGOASGDQ FGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDIIDIEDDDIVEYEARM KDI TDYLLVKMKETGINKLWGTANVFGHKRYMNGAATNPDPDLARAAVQIKNAIDATIK LGGQNYFEGGREGYQSLLNTQMQREKEHMRMLALARDYGRAHGFKGFTLIEPKMEPT KHQYDQDETIVIGFLRRHGLDKFVNIEVNHATLAGHTFEHELACAVDHMLGSLDANR GDAQNGWDIDQFPI DNYELT LAMQLIIRNGGLAPGGSNFDKLRNSTDPEDFIAHISA MDAMARALVNAVAIIIEESPIPDVMKERYASFDSGKGREYEEGKLSFEDLVAYAKAHGEPK QISGKQELYETIVALYCK ATGGCAAAGAGATATTTCCCGACAACTCGAAAGATCCCCTTCGAGGGCGTTGAGAGCAAG AATCCCTTGTCTTCCATATATGACCGCGAGCGCTGCTGTCATGGCAAGCCCATGAAG GACTGTTCAAGTTCCGATGGCTGGTGGCAACCTGGCCAGGCTTCGCGGACCCG TTCCGGCGGCAGACCCGCTCTACAGTGGGCAAGGGCAGTGGCCCTACTGCGCGGCC CGCGCAAGCTGACCGCGCTTCGAGATCATCAGAACTCGGAATCGGCTACTATTGC TTCCAGCATCGACCTGGTGGAGACACCGAGACATCGCCGAATACGAGGCCCGCATG AAGGACATCACCGACTACCTCGTCAGAAAGCAGAGGACCGGCATCAAGAACCTTGG GGCACCGGAACGTTTTCGGCAACAGCGCTACATGAACCGCGCTCGATGCCACCATCAAG TTCCAGATCGTCCCGCGCGCTGAGATCAAGAACCGCATCGACCCACCATCAAG
5586MI199_003	<i>Prevotella</i>	Amino Acid	108	MTKEYFTIGKIPFEGPESKNPLAFHYEPDRLVMGKMKDWLRFAMAWHITLGOASGDQ FGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDIIDIEDDDIVEYEARM KDI TDYLLVKMKETGINKLWGTANVFGHKRYMNGAATNPDPDLARAAVQIKNAIDATIK LGGQNYFEGGREGYQSLLNTQMQREKEHMRMLALARDYGRAHGFKGFTLIEPKMEPT KHQYDQDETIVIGFLRRHGLDKFVNIEVNHATLAGHTFEHELACAVDHMLGSLDANR GDAQNGWDIDQFPI DNYELT LAMQLIIRNGGLAPGGSNFDKLRNSTDPEDFIAHISA MDAMARALVNAVAIIIEESPIPDVMKERYASFDSGKGREYEEGKLSFEDLVAYAKAHGEPK QISGKQELYETIVALYCK ATGGCAAAGAGATATTTCCCGACAACTCGAAAGATCCCCTTCGAGGGCGTTGAGAGCAAG AATCCCTTGTCTTCCATATATGACCGCGAGCGCTGCTGTCATGGCAAGCCCATGAAG GACTGTTCAAGTTCCGATGGCTGGTGGCAACCTGGCCAGGCTTCGCGGACCCG TTCCGGCGGCAGACCCGCTCTACAGTGGGCAAGGGCAGTGGCCCTACTGCGCGGCC CGCGCAAGCTGACCGCGCTTCGAGATCATCAGAACTCGGAATCGGCTACTATTGC TTCCAGCATCGACCTGGTGGAGACACCGAGACATCGCCGAATACGAGGCCCGCATG AAGGACATCACCGACTACCTCGTCAGAAAGCAGAGGACCGGCATCAAGAACCTTGG GGCACCGGAACGTTTTCGGCAACAGCGCTACATGAACCGCGCTCGATGCCACCATCAAG TTCCAGATCGTCCCGCGCGCTGAGATCAAGAACCGCATCGACCCACCATCAAG
5586MI200_003	<i>Prevotella</i>	DNA	109	ATGGCAAAGAGATATTTCCCGACAACTCGAAAGATCCCCTTCGAGGGCGTTGAGAGCAAG AATCCCTTGTCTTCCATATATGACCGCGAGCGCTGCTGTCATGGCAAGCCCATGAAG GACTGTTCAAGTTCCGATGGCTGGTGGCAACCTGGCCAGGCTTCGCGGACCCG TTCCGGCGGCAGACCCGCTCTACAGTGGGCAAGGGCAGTGGCCCTACTGCGCGGCC CGCGCAAGCTGACCGCGCTTCGAGATCATCAGAACTCGGAATCGGCTACTATTGC TTCCAGCATCGACCTGGTGGAGACACCGAGACATCGCCGAATACGAGGCCCGCATG AAGGACATCACCGACTACCTCGTCAGAAAGCAGAGGACCGGCATCAAGAACCTTGG GGCACCGGAACGTTTTCGGCAACAGCGCTACATGAACCGCGCTCGATGCCACCATCAAG TTCCAGATCGTCCCGCGCGCTGAGATCAAGAACCGCATCGACCCACCATCAAG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				CTCGGGCACCGGCTACGTGTTCTGGGGCGCGGGAAGGCTACTACACCTTGTGAAC ACCCAGATCAGCGCGAGAGGACACCTCGCCAAAGATGCTACCGCGCGCGGACTAC GCGCGCGCAACGGCTTCAAGGGACCTTCTCATCGAGCCCAAGCCGATGAGCCACCC AAGCAACATACGACGCTGGACACGAGACCGTGATCGGCTTCTCGCGCAATGGCTG GACAAGGACTTCAAGGTGAACATCAGAGTGAACACCGCCCTCGCGGCCACACCTTC GAGCAGAGCTCACCGTGGCGTTGACAAAGGCTTCTCGCAGCATCGACCCCAACCGC GGCAGCGCCAGAACGGCTGGATACCGACCACTTCCGGTGGATCCGTACGATCTCAC CAGCGATGATCCAGATCATCCGCAACGGCGGCTTCAAGACGGCGGACCAACTTCGAC GCCAGGCTCGCGGCTCTCCACGACCGGAGACATCTCATCGCCACATCAGCGG ATGACCGCATGGCCACCGCCCTGCTGAACCGCCGCGCTCATCGAGAGAGCGCGCTC TGGAGATGGTGGCCAAAGCTTACGCTTCTTCGACAGCGGCTCGGCAAAAGTTTCGAG GAAGCAAGGCCACCGCTCGAGGAATCTACGATATGCCAAGCGGAACGCTGAGGTCAAG GCCGAATCCGGCAAGCAGGAGCTTACGAGACCTTCTGAACCTTACCGGAAATAG
5586MI200_003	<i>Prevotella</i>	Amino Acid	110	MAKEYFTPIGKIPFEGVESKNPLAFHYIDAERVVMGPKMDWKFAMAWHHTLQASADP FGQTRSVEWDKGECPYCRARAKADAGFEIMQKLGIGYCFHDIDLVEDEDIAEYFARM KIDITDYLVEKQKETGIRNLWGTANVFNKRYMNGAATNPQDIVARAALQIKNAIDATIK LGGTYVFWGGREGYYTLLNTQMOREKDLAKMLTAARDYARANGFKGTFLIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHETVAVDNGFLGSTDANR GDAQNGWDIDQFPVDPYDLTQAMIQLIRNGGFKDGGTNFDRLRSSSTDPEIDIFIAHISA MDAMAHALLNAAAVIEESPLCEMVAKRYASFDSGLGKKFEGKATLEELYEYAKANGEVK AESGKELYETLLNLAK
5586MI203_003	<i>Prevotella</i>	DNA	111	ATGGCACAAAGCGTATTTCTACCATCGGAAAATCCCTTCGAGGGACCCGAAAGCAAG AATCCCTGGCATTCATATTATAGCCCGACCGCTGTCTGGGCAAGAAGATGAAG GACTGGCTCGCTTCGCCATGGCTGGTGCCACACGCTGGGCCAGGCTTCGGCGCACAG TTCGGCGGCAGACCCGCCACTACGCTGGCAGAGCCCGCACGCTTGAACGGGCC AAGCCAAAGCGGATGCCGTTTCAGATCATGACAGAACTGGGCAATCGAATTTCTTGC TTCCAGATGTGGACCTCATCGAAGAGGGCGCCACGATCGAGGAATACGAGCAGCGGATG CAGCAGATCAGGATATCTGTGTTCAAGATCAAGAGAGACCGGCATCGCAACCTTGG GGTACGGCAACGTGTTGGACACGAGCGCTACATGAACGGCGCGCCACGAACCCGAT TTCGATGTCGTGGCCCGCGCGCTGCAGATCAAGACGGCCATCGACCCACCATCAAG TTGGCGCGGAGAACTATGTTCTGGCGCGCGGAGGATATGAGCCTGTCTCAAT ACGAGATGCACCGGAGAACTGCATCTGGCAAGATGTCGCCCGCGCGCGGCTAC GGACGCGCCACGGCTCAAGGGACCTTCTCATCGAACCAAGCCGATGGAACCCACC AAGCATCAGTATGACAGGATACGAGACGCTCATCGTTTCTGCGCGCTACGGCTG GACGAAGCTTCAAGGTGAACATCGAGGTCAACACGCTACGCTGGCGGCCATACCTTC GAACACGAACTGGCCACGGCGGTGCATGCCGCGCTGTGGCAGCATCGACGCCAACCGC GGCAGCGCCAGAACCGGTGGTACCGACCGAGTCCGAGTCCGACACTGCAACTGACCC CTGGCGATGCTGAGGTCAATCCGCAACGGCGGTCTGGCCCGCGGCTCGAATTTTCGAT GCCAAGCTCGCGCGGAACTCCACGATCCGGAGACATCTTCATTTGCCACATCAGCGG ATGGATGCGATGGCGGCGCTGCTCAATGCGCGCGCTCTGCGAGACGTCCCGGAT CCGGCGATGGTCAAGGCGGTTACGCTTCGTTGACAGCGCGCGCGGCAAGGATTTTCGAA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI203_003	<i>Prevotella</i>	Amino Acid	112	GAGGGAAGGATGACGCTGGAAGACCTCGTGGCCATATGCCAGGACCCACGGCGAGCCGAAG CGGACCTCGGGCAAGCAGGAACCTATGAGACCCCTCGTGGCGCTTATTTGCAAAATAG MAQAYFPRTGKIPFEGPESKNPLAFHYYPDRVLVGLKMKDWLRFAMAWHHTLQASGDQ FGGQTRHYAWDEPATPLERAKAKADAGFEIMOKLIGIEFFCFHDVLDLIEGATIEEYEQRM QQITDYLVLVKKMETGIRNLWGTVNFVGHRYMNGAATNPEDVVARAAVQIKTAIDATIK LGGNYVFWGGREGYMSLINTQMHREKLHLGKMLAAARDYGRAHGFKGFTLIEPKMEPT KHQYDQDTETVIGFLRRYGLDEDFKVNIEVNHATLAGHTFEHELATAVDAGLLGSDANR GDAQNGWDTDQFPIDNYELTLAMLOVIRNGGLAPGGSNFDKLRNRSTDPEDIFIAHISA MDAMARALINAAALCETSPIAMVKARYASFDGAGKDFEGRMTLEDLVAYARTHGEPK RTSGKQELYETLLVALYCK
5586MI205_004	<i>Prevotella</i>	DNA	113	ATGACCAACGAGTATTTCCCGGAATCGGTGTGATTCGGTTTGAAGGACAGGAAAGCAAG AAATCCCTGGCTTTCATATATGACGCCAACCGGTAGTATGGGCAAAACCCATGAAG GAATGGTTCAAAATTTGCCATGGCTGGTGCCATACGCTGGGCGAGCATCGGCCGATCCC TTCGGGGACAGACCCGCTCCTACGATGGACAAGGGCAGTGCCTTACTGCCGTGCC CGCCAGAGGCCGACGCCGCTTTGAACATGATCAGAAAGCTGGGAATCGGCTATTTTCG TTCACGATGTGAATATCATCGAGACTCGGAGACATTCGAGTATCAGGCCCGTAIG AAGGACATCAGGACTATCTGCTGGTGAAGTGAAGAAACGGGCATCAAGAAATCTGTGG GGCACGGCCAAACGCTCTCGGCCACAAGCGCTATATGAACGGCGCCGCCAACCCGCAA TTCGACGTGGTAGCCGCTGGGTCCAGATCAAGAAAGCCCTGGACGCCACCATCAAG CTGGGGCGAGCAATTAATGTGTTCTGGGGCGCGGGAAGGCTACTACACCTTTTGAAC ACGCAGATCGAGCGGAGAAAGCACACCTGGCCAGATGCTCAAGGGCGCCCGGACTAT GCCCCGCAAGGGATTCAAGGACGCTTCCTCATTTAGCCCAAGCCATGGAGCCACC AAGCACAGTACGAGTAGATACGAGACCGTATGTTTCTGCGGCCAACAGGGCTG GACAAGGACTTCAAGGTGAATATCGAAGTGAACACGCCCTGGCCGCCATACCTTC GAGCAGGCTCACCGTGGCCGCAAAACGGCTTCTGGCAGCATCGACGCCAACCCG GGAGCGCCAGAACGGCTGGGTACAGACCACTTCCCGTGGACGCCCTTGACCTCAC CAGCCATGATGAGGTCCTGCTCAACGGCGGATTCGGCAACGGCGGCGACCAACTTCGAC GCCAACTGGCCGCTTCTCCAGGATCCGAGGACATCTCATCGCCACATCAGCGCC ATGGACGCATGGCCACGCCCTCCTGAACGCCCGCCCTATCCTGGAGAGAGCCCCATG CCGGCATGGTGAAGGAGCGCTACGCTTCTTCGACAAATGGCCTTGGCAAGAAGTTTCGAG GAAGGAAAGGCCACGCTGGAAGAGCTGTACGACTGTCCCAAGAAGAAGCGCGAGCTGTG GCCGCTTCGGAAAGAGGAACATGACGAAACGCTGTGAACCTGTACGCCCAAGTAA
5586MI205_004	<i>Prevotella</i>	Amino Acid	114	MTNEYFPGIGVIPFEGGESKNPLAFHYIDANRVVMGKPMKEFKFAMAWHHTLQASADP FGGQTRSYAWDKGECPCYRQKADAGFELMQILGIGYFCFHDVNIIECEDIAEYEFARM KDI TDYLLVKKMETGIRNLWGTVNFVGHKRYMNGAATNPQDVVARAAVQIKNALDIK LGGSNVFWGGREGYYTLLNTQMOBEKDLAQLKAARDYARGKFGKFTLIEPKMEPT KHQYDQDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHETVARENGFLGSDANR GDAQNGWDTDQFPVDFDLTQAMMQVLLNGGFGNGGTFNDAKLRNRSTDPEDIFIAHISA MDAMAHALLNAAALLESPPMPGMVKERYASFDNGLGKKFEGKATLEELYDYAKKNGEFV AASGKQELYETLLNLYAK

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI208_003	<i>Prevotella</i>	Amino Acid	118	GCACGGCCAAAGGTTTCAAGGGCAGCTTCTTCATCGAACCCAGCCGATGGAACCCACC AAGCACCATGATGACAGGATACCGAGACCGTATCGGATTCCTCCGAGCCACCGGCTG GACAAGGACTTCAAGGTCAACATCGAGGTGAACACACGCCCTCTGGCGCCATACCTTC GAGACGAATGCGCAACCGCGTGCACAAACGGCATGCTCGGAGCATCGAACGCCAACCGC GGAGACGCCAGAACGGCTGGACACCGACCTTCCCGATCGACAACTCGAGCTCAG CTTGCCATGATGAGATAATCCGCAACGGCGCTGACCGCGGCTTGAACCTCGAC GCAAAGTGGCGCCGCAATCCACCGATCCCGAGGACATCTTCATCGCCACATCAGCGCG ATGGACGCGATGGCCCGCCCTCGTCAACGGCGCGCCATCTCTCGGAGAGTCGCCCGTT CCGGCTATGGTCAAGGACCGCTATGCTTCGTTCGACTGGCAAGGGCAAGGACTTCGAA GACGGCAAACTGACTCTCGAAGACATCGTCCCTACCGCAGGAGAAATGGCGAGCCGAAA CAGATTTCCGGCAAGCAGGAACCTACGAAACATCTCTCGCTCTTACTGCAAGTAA MSTEYFPITGKI PFEGPESKNPMAPHYYEPEKLVMGKKMKDWLRFAMAWHHTLGAASGDQ FGGQTRSVAWDKGDPCPYSRARAKVDAGFEIMQKLGIEFFCFHDIDLVEDTDDIAEYARM KDI TDYLLLEKMEATGIKNLWGTANVFGHKRYMNGAATPNFVAVARAQVQIKNAIDATIK LGGENYVFWGGREGYMSLINTQMORKEHLAKMLTAARDYARAKGFKGFLIEPKMEPT KHQYDQDETIVIGFLRSHGLDKDFKNVIEVNHATLAGHTFEHELATAVDNMGMLGSDANR GDAQNGWTDQFPIDNFELTLLAMQILIRNGGLAPGGSNFDKLRNSTDPEDFIAHISA MDAMARALVNAAAII LGSPVPMVKDRYASFDCKGKDFDGLTLEDIVAYARENGEPK QISGKQELYETIVALLYCK
5586MI210_002	<i>Prevotella</i>	DNA	119	AIGTCATATTTTCCTACTATCGGTAACATCCCTTTTGAGGGTGTAGAGAGCAAGAAATCCC CTTGCCCTTCCATATATATGACGCTTCCCGGTAGTTATGGGCAAGCCCATGAAGGAGTGG CTCAAGTTTGCCATGGCTGGTGGCACACGCTGGGTGAGCATCGCCGACCCCTTTCGGC GGACAAACCGCAGTATGCTGGGACAAAGCGAGTGCCCTACTGCGGTGCCCGTGCC AAGCCCGACCGCGCTTCGAGCTCATGCAGAACTGGGCATCGAGTATTTCTGCTCCAC GACATTGACCTCATCGAGGACTGCGACGACATTCAGAGTACGAGCGCGCTCTGAAGGAC ATTACGGACTACCTCCTGGAGAAGATGAAGACCGGTATCAAGAACCTGTGGGTACG GCCAATGTGTTTCGTTAACAGCGTTACATGAACGGTGTCTGATACCAACCTCAGTTTAC GTGTGGCCCGCGTGGCTCCAGATCAAGACGCCATTCAGCTACCATCAAGCTGGC GGTCCAACTATGTGTTGGGTGGCGTGGGTACTACACGCTTCTGAACACCCAG ATGCAGCGTGAGAAGAAATCACCTGGCTGCCATGCTCAAGGTGCCCGGACTATGCCCGC GCCAAGGTTTCAAGGCACCTTCCCTATTGAGCCCAAGCCCATGAGCCACCAAGCAC CAGTACGACGTAGACACCGAGACCGTGATTTGATTCCTCCGCGCCAAAGGTCTGGAGAAG GACTTCAAGGTGAACATGAGGTGAACACCGTACTCTTGCCGGTCAACCTTCGAGCAC GAGCTCACCGTGGCCGTGAGAACGCTTCTGGGTTCATTCAGCCCAACCGCGGAGAT GCCCAGACGGCTGGGACACCGACAGTTCCCGGTAGATGCTTTCACCTCACCCAGGCC ATGATGCAGATTCCTCAACGGAGGCTCCGGCAATGGCGGTACCAACTTGACGCCAAG CTGCGCGCTTCTCCACGACCCGAGGACATCTTCATCGCGCACATCAGCGCCATGGAT GCCATGGCTACGCCCTGCTCAATGCAGCTGCCGTGGAGGAGAGCCCGCTTTGCAAG ATGGTCAAGGAGCGTTACGCTTCTTCGACAGCGGTCTTGCAAGCAGTTCGAGGAGGA AAGGCTACGCTGGAAGATCTGTATGCCATGCGTCAAGAACGCTGAGCCCGTGGTGGCT TCCGGCAGCAGGAACTGTACGAAAACCTTCTGAACCTCTATGCAAAATGGTAA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI210_002	<i>Prevotella</i>	Amino Acid	120	MSYFTTIGNIPFEGVESKNPLAFHYDASRVVMGKPMKEWLFAMAWHHTLQASADPFG GQTRSYAWDKGECPCYCRARAKADAGFELMQKLGLIEYFCSHDIDLIEDCDDIAEYEARLKD ITDYLLEKMKKTGINKLWGTANVFNKRYMNGAATNPQFDVVARAAVQIKNAIDATIKLG GSNYVFWGGREGYYTLLNTQMQRKNHLLAAMUKAARDYARANGFKGTFLEPKPMPTKH QYDVTETVIGFLRANGLEKDFKVNIEVNHATLAGHTTTEHLETVARENGFLGSDANRGD AQNGWDTDFPVDADLTQAMQIILNGSGNGGTNFDKLRSSSTDPEIDIFIAHISAMD AMAHALLNAAAVLESPLECKMVKERYASFDSGLGKQFEFGKATLEDLYAYAVKNGEPVVA SGKQELYETFLNLYAKW
5586MI212_002	<i>Prevotella</i>	DNA	121	ATGTCAACTGAGTATTTCCCTACATCGGCAAGATTCCCTTCGAGGGACCCGAGAGCAAG AACCCCATGGCCCTTCCACTACTATGAACCCGAAAGTTGGTGATGGGCAAGAAAGATGAAG GACTGGCTGGCTTTCGCAATGGCTGGTGACACCCCTTGGAGCCGCATCCGGCCGACAG TTGGCGGACAGACCCGAGTTACGCCCTGGACAAGGGGAGCTGCCCTACAGCCGCGCC CGGCCAAGGTCAGCGCGGCTTCGAGATCATGAGAAGTCGGCATAGAGTCTTCTGTC TTCCATGACATCGACCTGGTCGAGATACCGACGACATCGCGGAGTATGAAGCCCGGATG AAAGACATCAGGACTATCTGCTGGAAGAGATGAGGTTACCGGCATCAAGAACCTCTGG GGCAGGCCAATGCTTCGGTCACAAGCGTTATATGAACGATGACGCCAACAACCCGAT TTTCGAGTGGTCGAAGGGCGGCTGCAGATCAAGAACCGCATCGACGCCACCATCAAG CTGGTGGTGAGAACTATGTGTTCTGGGTGGACGCGAGGTTATATGAGCCTGCTCAAC ACCCAGATCAGAGGAGAAGAACCTTGCAGAGATGCTACCGCGCACGTGACTAT GCACGGCCAAAGTTTCAAGGACAGTTCTTCTCGAATCGAACCCGACCGATGGAACCCACC AAGACCCAGTATGACAGGATACCGAGACGTTTCTCGAATCGAACCCGACCGATGGAACCCACC GACAAGGACTCAAGGTCAACATCGAGGTGAACACCGACCCCTCGCGGCGCCATACCTTC GAGACGAACTGGCCACCGCGCTGACAAACGGCATGCTCGGAGCATCGACGCCAACCGC GGAGACCCAGAACCGCTGGACACCGACCTCCGATCCGATCGACAACTCGAGCTCAGG CTTGCCATGATGAGATAATCCGCAACGGCGCTGGCACCGGCGGTTTGAACCTTCGAC GCAAGCTGGCGCGCAATCCACGATCCGAGGACATCATATCGCCACATCAGCGCG ATGGACGCCATGGCCGCGCCCTCGTCAACGCGCGCCCATCTCGCGGAGTCGCCCGTT CCGGCTATGGTCAAGGACCGCTATGCTTCGACTCGGCAAGGCAAGGACTTCGAA GACGGCAAACTGACTCTCGAAGACATCGTCCGCTACGCCAGGAGAAATGGCGAGCCGAAA CAGATTTCCGGCAAGCAGGAACCTACGAAACATCTCGTCTCTTTACTGCAAGTAA MSTIEYFTIGKIPFEGESKNPMFHYEPEKLVGMKKMDWLRFAMAWHHTLGAASGDQ FGQTRSYAWDKGDCPYSRARAKVDAGFEIMQKLGLIEFFCFHDIDLVEDTDDIAEYEARM KDI TDYLLLEKMEVTGINKLWGTANVFNKRYMNDAAATNPDAVVARAAVQIKNAIDATIK LGGENYVFWGGREGYMSLLNTQMOREKEHLAKMLTAARDYARAKGFKGTFLEPEPMEPT KHQYDQDETETVIGFLRSHGLDKFVNIEVNHATLAGHTTTEHLETVARENGFLGSDANR GDAQNGWDIDQFPIDNFELTAMQIILRNGGLAPGGSNFDKLRNSTDPEDIIIAHISA MDAMARALVNAIAAILGESVPVAMVKDRYASFDCGKGFDEGKLTLEDIVAYARENGEPK QISGKQELYETIVALYCK
5586MI212_002	<i>Prevotella</i>	Amino Acid	122	
5586MI213_003	<i>Prevotella</i>	DNA	123	ATGACCAACGAGTATTTTCCCGGAATCGGTGTGATTCGGTTGAAGGACGAGAAAGCAAG AATCCCTGGCTTTCCTATATTATGACGCCAACCGCGTAGTGATGGGCAACCCATGAAG GAATGGTTCAAAATTTGCCATGGCCTGGTGGCATACGCTGGGCGAGGCATCGGCCGATCCC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				TTCCGGCGACAGACCCGCTCCTACGATGGGACAGGGCGGAGTGCCTTACTGCCGTGCC CGCCAGAAAGCCGACGCGGCTTTGAAGTATGATGAGAAAGTGGGAATCGGCTATTTCTGC TTCCAGGATGTGATATCATCGAGACTCGGAGACATTCGCGAGTATGAGGCCCGGTATG AAGGACATCAGGACTATCTGTGTGAAGATGAAGAAACGGGCATCAAGAAATCTGTGG GGCAGGCCAACGCTTTCGGCCACAAGCGCTATATGAACGGCCCGCCACCAACCCGCAA TTCCAGCTGTAGCCGCTGCGGTCCAGATCAAGAACGCTTGAGCCGACCATCAAG CTGGCGGACGAAATATGTGTCTGGGGCGCGGGAAGCTACTACACCTTTTGAAC ACGAGATCAGCGGGAAGACACCTGGCCAGATGCTCAAGGGCCCGCGGACTAT GCCCGGCAAGGATCAAGGGCAGCTTCTCATTTAGCCCAAGCCCATGGAGCCACCC AAGCACATACGCTAGATACGAGACCGTATGTTTCTGCGGCCAACGGGTG GACAAGGACTCAAGGTGAATATCGAAGTGAACACCGCCCTGGCCGCCATACCTTC GAGCAGGCTCACCTGGCCCGGAAAACGGCTTCTGGCAGCATCGACGCCAACCGC GGAGCGCCAGAACGGCTGGATACAGACCATTTCCCTGGACGCTTGACCTACAC CAGGCCATGATGAGGTCTCTCAACGGCGGATTCGGCAACGGCGGACCAACTTCGAC GCCAACTGCGCGCTTCTCCAGGATCCGAGGACATCTTCATCGCCACATCAGCGCC ATGACGCAATGCGCCCGCCCTCTGAAACGCGCCGCTTCTGGAAGAGAGCCCGCATG CCGGCATGGTGAAGGAGCGCTACGCTTCTTCGACAAATGGCTTGGCAAGAAGTTCGAG GAAGAAAGGCCACGCTGGAAGCTGTACGACTATGCCAAGAAAGAACGGCGAGCGTGTG GCCGCTTCGGAAAGCAGGAACTGACGAAACGCTGTGTAACCTGTACGCCAAGTAA
5586MI213_003	<i>Prevotella</i>	Amino Acid	124	MTNEYFPGTGVIPFEGESKNPLAFHYDANRVVMGPKMEWFKFAMAWHHTLGQASADP FGGQTRSYAWDKGECPCYRARKADAGFELMQKLIGYGFCHDVIIEDEDIAEYFARM KDIITYLLVKMKETGINKNLWGIANVGHKRYMNGAATNPQDVVARAAVQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMORKEKDLAQLKAARDYARGKFGKFTLIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHFLTARENGFLGSDANR GDAQNGWDIDQFPVDAFDLTQAMMQVLLNGGFGNGGTNFDKLRSSSTDPEIDIFIAHISA MDAMAHALLNAAAILLESPMPGMVKERYASFDNGLGKKFEEGKATLEELYDYAKNGEPV AASGKQELYETLLNLYAK
5586MI215_003	<i>Prevotella</i>	DNA	125	ATGGCAAAAGAGATATTCCTCCGAGATCGGAAAGATCGGCTTGAGGGTCTTGAGAGCAAG AACCCGATGGCATTCATATTTATGACGCGGAGGTGTCTGTCTCGGAAAGAAGATGAAG GACTGGCTGAAGTTCGGGATGGCTGGTGATACGCTCGGACAGGCTTCGGCGGACCCA TTCCGGCGGACAGACTCGAGCTATGATGGCAAGGAGTGCCTACTGCGGTGCTGCGTGC CGGCCAAGCGCAGCGCGCTTCGAGCTCATGCAAGAGTCCGGCATCGAGTACTTCTGTC TTCCACGACATCGACCTCATCGAGGACTCGGACGACATCGACGAGTACGAGGCCGGAIG AAGGACATCACCGACTACCTGCTGGAGAGATGAAGGAGACCGGAATCAAGAAATCTCTGG GGAACGGCCAACGCTTTCGGTCAAGCGCTACATGAACGGCGCTGACCAATCCGAG TTTGAAATCGTCGCCCGGCTGCCGTCCAGATCAAGAACGGCTCGACGCCACCATCAAG CTCGCGGCTCCAACCTACGCTTCTGGGGCGCGCGGCTATTACCGCTGCTGTAAT ACCCAGATCGAGCGGAGAGGACCATCTCGCAGGCTTCCCATCGAGCGAAGCCGATGAGCGGACCC GCGCGGCCAAGGGTTCAGGGGACCTTCCCATCGAGCGAAGCCGATGAGCGGACCC AAGCACCATGATGACGTCGACACGAGACCGTCACTGGTTCTCTCCGCAAGATGGGCTC GACAAGGACTTCAAGGTCAATATCGAGGTGAACACGCCACCTCGCGGCCATACCTTC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5586MI215_003	<i>Prevotella</i>	Amino Acid	126	GAGCAGAGCTGACCGCGCGCCGGAGAAACGGCTTCTCGCAGCATCGACGCCAACCGC GGCAGCGCCAGAACGGCTGGGACACCGACCGATTCCCGGTGGACGCTTCGATCTCAGC CGGCCATGATGAGATCCTGCTCAATGGCGTTTCGGCAACGGCGGACCACTCGAC GCCAAGTGGCGCGAGCTCCACCGATCCGAGACATCTTCATCGCCACATCAGCGG ATGACCGCATGGCCACCGCTGTGAATCGCGCGCATCTCTCGAGGAAGCCCGCTG CCGCCCTGGTCAAGCAGCGCTATCGTCTTCCACAGCGCTTCCGCAAGCAGTTCGAG GAGGTAAAGCCACGCTCGAGGACCTGTACGATACGGAAGAGCACGGCGAGCCCGT CGGCCCTCGGCAAGCAGGAGCTCTCGGAGACCTATCTCAACCTCTACGCGAAATAA MAKEYFPQIGKIGFEGLESKNPMAPHYYDAERVVLGKKMDWLKFAMAWHHTLGQASGDP FGQTRSIEWDKGECYCRARAKADAGFELMQLGIEYFCFHDIDLIEDDDIDEYEARM KDI TDYLLLEKMKETGIKNLWGTVNFHKRYMNGAATNPQFEIVARA VQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMREKDHLLRLTAARDYARAKGFKGTFPIEPKPEPT KHQYDVDTETVIGFLRQGLDKDFKVNIEVNHATLAGHTFEHELTAARENGFLGSI DANR GDAQNGWTDIDQFPVDAFDLITRAMQILLNGFGNGGTNFDKLRSSITDPEDIFIAHISA MDAMAHALINAAAIILESPLPALVKQRYASFDGLGKQFEFGKATLEDLYAYAKEHGPV AASGQELCETLYNLVAK
5607MI1_003	<i>Prevotella</i>	DNA	127	ATGAGTAAAGAGTATTTCTCTGGGATTTGGCAAAATCCCGTATGAGGGCCGAGCAAG AAATGATGGCATTCACATATTAATGATCCCGAACGCTGGTTCATGGGCAAGAAAATGAAA GACTGGTTCAAGTTTCGCTATTCCTGGTGGCATAACCTGGCGAGCGCAGTCTGACCCAG TTTGGCGGACAGACCCGTTTCTATGAATGGGCAAAAGCCGAGACCCCTTGCAGCGTGCC AAGACAAGATGGATGCCGTTTGAATCATCGAAGCTGGGCATCGAGTATTTCTGT TTCCATGATGTGACCTCATCGAGAGCGGATACCATCGAGGAATATGAAGCCCGCAIG CAGCGATTACCGATACGCGCTGGAGAAGATGAAGGAAACGGGTATCAAGTTGCTGTGG GGCACTGCCAACGTTGTCGCCACAAAGCTTACATGAACGGCGCCGCCCAATCCCGAC TTCAATGTCGTGGCAGCTGACCGTGCAGATCAAGAACGCTCGATGCTACCATCAAG TTGGCGGAAACGAGTACGTCTTCTGGGGCGTCTGTAAGCTATCAGAGCCTGCTCAAC ACCCAGATCAGCGCGAGAAAGAACCTGGCCAGATGCTCAGCGAGCCCGTGTACTAT GCCGTGCTAAGGGCTTCAAGGGACCTTCTCATTGAGCCCAAGCCGATGGAACCCACC AAGCACCATATGACAGGACACCGAGACCGTTATCGGCTTCTTGGTGCCTAATGGCCTT GACAAGGACTTAAAGTCAACATGAGGTCAACCATGCCCGTGGTGGCCACACCTT GCACATGATGTGGAGTGGCTGGATTAACGGTATGTCGGCAGCATCGAATGCTAACCGT GGTGACCAACAGACCGCTGGGATACAGACCACTTCCCATCAACAGTATGAATCAAC AATGCTATGCTGCAGATCATGCACGCGCGGCTTCAAGGACGGCGGTACCAACITTGAC GCCAAGTGGCGGCAACAGTACCGACCCCGAGACATCTTACCGCTCACATCAGTGTG ATGGACGCTTGGCCGCTGCTGTGAGTGTGCTGCCGATATCTTGAAGAGCGAGTTG CCTGAAATGCTCAAGGAACGCTATGCCAGCTTACCGCGGTGAAGGCAAGCGCTTGTGAG GATGGCCAGATGACTCTGTAGGAACCTGTTGCCATATCCCATGCCATGCCGAGCGCTGCT ACCATCAGTGGCAAGGAGAAAATATGAAGCCATCTGTGGCTTGTGACCTCAAGTAA
5607MI1_003	<i>Prevotella</i>	Amino Acid	128	MSKEYFPQIGKIPYEGAESKNVMAFHYIDPERVVMGKKMDWFKFAIWWHTLGQASADQ FGQTRFVWEDKAEDPLQRAKDKMDAGFEIMQKLGIEYFCFHDVLIIEADTIEEYARM QAITDYALEKMKMATGIKLLWGTVNFHKRYMNGAATNPDNVVARAAVQIKNALDATIK

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5607M12_003	<i>Prevotella</i>	DNA	129	LGGTSYVFWGGREGYQSLNLTQMOREKNHLAKMLTAARDYARAKFGKFTLIEPKMEPT KHQYDQDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFAHELAVAVDNGMLGSDANR GDHQNGWDTDQFPINSYELTNAMLQIMHGGGFKDGGTNFADKLRRNSTDPEDIFIAHISG MDALARALLSAAADI LEKSELPEMLKERYASFDAGEGKRFEDGQMTLEELVAYAKSHGEPA TISGKQEKYEAI VALHVK ATGAGTAAAGAGTATTATCTCCTGAGATTGGCAAAATCCCGTTTGAGGGTCCCGAGAGCAAG AATGTGATGGCGGCTTCCATTAATGAACCCGAGACGCTGCTCATGGTAAAGAAGATGAAA GACTGGCTCAAGTTTGGCATGTGCTGGTGACAGCCCTGAGTGGTCCGACGATGCCGACAG TTCCGGCGGACAGACACGTTTCTACGAGTGGGACAAAGCCGATACCCCTGACGCGTGC AAGGACAAAATGGATGCCGGATTGAAATCATGAGAAGTTGGGCATCGAGTACTTGTGC TTCCAGATGTGGACCTCATCGAGAGGCGTATACCATCGAGGAATACGAGGCCCGCATG AAGGCCATTACCGACTATGCGCTGGAGAAGATCGAGGCCACCGGCATCAAGTTGCTTGG GGCACTGCCAATGTGTGGCCACAAAGCGTACATGAACGGCGCCGCCAATCCGAT TTCAATGTCGTGGACGTCGCGCTCCAAATCAAGAAATGCCATCGATGCCACCATCAAG CTGGCGGACAGGTTACGTTCTTCTGGGTGGTGTGAGGCTATCAGAGTCTGCTCAAC ACGAGATCGAGCGGAGAAGGACCATCTGCCCGCATCTGGCGGACGCGCCGCGATAT GGCGTGCCCATGGCTTCAAGGGCACTTCTTATCGAGCCCAAAACCATGGAGCCACCC AAGCACCATGATGTGACACCGAGACCGTCTCGCTTCTCGTCCGACCGGCTG GACAAGGACTCAAGGTTAAACATCGAGTCAATCATGCTGCGGACACACACTTTC AGCCAGAACTGGCTGGCGCTGGACAAAGGATGCTGGGACAGCATCGACGCCAACCGC GGCGATTATCAGAAATGGCTGGACACCGACCTTCCCATCGACAGCTCGAGCTCAAC CAGGCCATGCTGCAGATCATGCGCGCGCGGCTTCAAGACGAGGATCAACATTCGAT GCCAAGCTGGTCCGAACAGTACCGACCTGAGGACATCTTATCGCCACATCAGCGGT ATGATGCCATGGCAGCGGCTGTGAGCGCTGCGCTATCTCGAGGATGGCGAGTTG CCGCGATGCTCAAGGCACGTTATGCCAGCTTGACCGGCGAGGTAAGCGCTTGAG GACGGGAGATGACGCTGAGGAGCTGGTGGATTATGCAAGGATTATGCCAAATCGCAC GGCGAGCCTGATGTCATCAGCGGCAAGCAGGAGAAGTTTGAACCATCTGTTGGCCCTTAC GCCAAGTAA
5607M12_003	<i>Prevotella</i>	Amino Acid	130	MSKEYYPEIKGIPFEGPESKNVMAFHYPERVVMGKKMKDWLFAMCWHSLGQASADQ FGGQTRFYEWDKADTPLQRAKDKMDAGFEIMQKLGIEYCFHDVDLIEADTIEYEARM KAITDYALEKMQATGIKLLWGTANVFGHKRYMNGAATNPENNVARAQVQIKNAIDATIK LGGTSYVFWGGREGYQSLNLTQMOREKDLARMLAARAHGFRHGFKFTLIEPKMEPT KHQYDQDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFESHELAVAVDNGMLGSDANR GDYQNGWDTDQFPIDSFELTQAMLQIMRGGGFKDGGTNFADKLRRNSTDPEDIFIAHISG MDAMARGLLSAAAILLEDGELPAMLKARYASFDQEGEGRFEDGEMTLEQLVDYAKDYAKSH GEPDVISGKQEKFFETIVALYAK ATGACCAACGAGTATTTCCCGGAATCGGTGTGATTCGGTTTGAGGACAGGAAAGCAAG AATCCCCGTGGCTTCCATTAATGACGCCAACCGCTAGTGCATGGGCAAAACCCATGAAG GAATGGTTCAAAATTTGCCATGGCTGGTGACATACCTGGGGCAGGCATCGGCCGATCCC TTCCGGCGGACAGACCGCTCCTACGATGGGACAAAGGCGGCTACTGCGCGTGC CGCCAGAGCCGACGCGGCTTGAACATGATGACAGAGCTGGGAATCGGCTATTCTGC
5607M13_003	<i>Prevotella</i>	DNA	131	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5607M13_003	<i>Prevotella</i>	Amino Acid	132	TTCCACGATGTGGATATCATCGAGACTCGAGGACATTGCCGAGTATGAGGCCCGTATG AAGACATCAGCGACTATCTGCTGGTGAAGATGAAGAAACGGGCATCAAGAAATCTGTGG GGCAGGCCAACGCTCTCGGCCACAAGCGCTATATGAACGGCGCCGCCACCAACCCGAA TTTCAGCTGGTAGCCCGCTCGGTCCAGATCAAGAACGCCCTGGAGCCACCAATCAAG CTGGCGGAGCAATATATGTGTTCTGGGGCGCGGAGGCTACTACACCTTTTGAAC ACGAGATCGAGCGGAGAAAGACACCTGGCCAGATGCTCAAGCGCGCCCGCGACTAT GCCCGGCAAGGATCAAGGACAGTTCCTCATTTGAGCCCAAGCCATGGAGGCCACCC AAGACACGACTACGCTAGATACGAGACCGTATGTTCTCTGGCGCCAAACGGGCCG GACAAGGACTCAAGGTGAATATCGAAGTGAACACGCTGCTGGCGGCCATACCTTC GAGACGAGCTCACCGTGGCCCGGAAAACGGCTTCTGGGAGCATCGACGCCAACCGC GGAGACGCCAGAACGGCTGGGATACAGACCACTTCCCCGTGGACGCTTGGACCTACCC CAGGCCATGATGAGGTCCTGCTCAACGGCGGATTCGGCAACGGCGGCACCAACTCGAC GCCAACTGGCGCCGTTCTCCAGGATCCCGAGGACATCTTCATCGCCACATCAGCGCC ATGGACGCTATGGCCACGCTTCTGAACGGCGCCCATCTTGGAGAGAGACCCCATG CCGGCATGGTGAAGAGCGCTACGCTTCTCGACAATGGCTTGGCAAGAGTTGAG GAAGAAAGGCCACGCTGGAAGAGCTGACGATATGCCAAGAACGGCGAGCGCTGTG GCCGCTTCGGAAAGCAGGAACCTACGAAACGCTGTGAACCTGTACGCCAAGTAA MTNEYFPGTGVIPFEGQESKNPLAFHYDANRVVMGKPMKEWFKFAMAWHHTLGQASADP FGQTRSAYWDKGECPYCRARQKADAGFELMQKLIGYFCFHDVDI IEDCEDIAEYARM KDI TDYLLVMKETGIKNLWGTANVGHKRYMNGAATNPQDVVARAAVQIKNALDATIC LGGSNYVFWGGREGYYTILNTQMQRKDHQAQMLKAARDYARGKFGKFLIEPKMEPT KHQYDVTETVIGFLRANGPKDFKNIEVNHATLAGHTFEHELTVAENGFLGSDANR GDAQNGWTDQFPVDAFDLTQAMQVLLNGFGNGGTNFAKLRRSSTDPEDFIAHISA MDAMAHALINAAAILESPMPGMVKERYASFDNGLGKFEFGKATLEELYDYAKKNPEPV AASGQELYETILLNLYAK
5607M14_005	<i>Prevotella</i>	DNA	133	ATGACTAAAGAGTATTTCCCTTCCGTCGGCAAGATTGCCTTGAAGGACCCGAAAGCAAG AACCCTATGGCCTTCCATATATGACGCCAATCGCTGGTAATGGGAAAGCCGATGAAA GAATGGCTTAAATTTGCCATGGCTGGTGCCACACCTGGCCAGGCCCTTCGACAGCCCC TTCGGCGTCAGACCCGCTCCTACGAGTGGGCAAGGGCGAGTGCCTACTGCCCGGCC AAGGCCAAGCCGATGCCGCTTTGAACGTGAGTGAAGAACCTGGCATCGAGTATTTTGC TTCACGATATAGACCTGGTGAAGACTCGATGATATCGCGAATACGAGGCCCGCAIG AAGACATACCGACTATCTCTGGAGAGATGAAGAAACCGGCATCAAGAACCTCTGG GGAACGCCCAACGTTTCGGCCCAAGCGCTATATGAACGGCGCCGCCAACCCCTCAG TTCGACATCGTGGCCGCTGCTCCAGATCAAGAACGCCCTGGATGCCACCATCAAG CTGGGGGCTCCAATATGTGTTCTGGGGCGCTGAGGCTACTATACCTCTCTGAAC ACCCAGATCGAGAGAGAGAACCCACCTGGCCAAAGATGCTACCCGCCCGCCGACTAT GCCCGTGCCAAGGGCTCAAGGGACCTTCTCATCAAGCAAGCCGATGGAGCCACCC AAGCACCTAGTACGCTAGATACGAGACCGTATCGGCTTCTTCGCGCCCAACGGCTG GACAAGGACTTCAAGGTGAATATGAGGTGAACACCGCCCTGGCGGCCACACCTTC GAGCAGGACTCACCGTGGCGCGAGAACGGCTTCTGGGAGCATCGACGCCCAACCGC GGAGACGCCAGAACGGCTGGGATACGGACCACTTCCCGGTGGATGCCCTCGACCTACCC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				CAGGCTATGATGAGATCCTTCTGAACGGAGGCTTCGGCAACGGCGGTACCAACTTCGAC GCCAACTGGCGGCTCCTCCAGGACCCCGAGGACATCTTCATCGCCACATCAGCGCT ATGATGCGATGCCCGCCCTGCTGAATGAGCGCCCATCTCTGGAGAAAGCCGCTT CCGAAGATGCTGAAAGAGCGTTATGCCAGCTTTGACGGCGCTCTGGCAAGAGTTGAA GAAGGCAAGGCTCTCTGGAAGAACTCTACGAGTATGCCAGAGCAACGAGAGCCCGTG GCCGCTTCGGCAAGCAGGAGCTCTCGAAACGTAACCTGACCTCTACGCTAAGTAA MTKEYFSPVKGIAFEGPESKNPMAPHYYDANRVVMGKPMKEWLKFAMAWHHTLQASADP FGGQTRSIEWDKGECPCRAKAKADAGFELMOKLGIIEYCFHDDI DLVEDDDIAEYEARM KDITDYLLEKMKETGINKLWGTVANVFGHKRYMNGAATNPQDIVARAAYQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMOREKDHAKMLTAARDYARAKFGKTFLEPKMEPT KHQYDVTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHFLTVAENGFLGSDANR GDAQNGWDIDQFPVDAFDLTQAMQILLNGGFGNGTDFDAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAAI LEE SPLPKMLKERYASFDGGLGKKEEGKASLEELYEYAKSNGEPV AASGKQELCETYLNLYAK ATGGCTAAGAAATACCTCCCTCCATCGGCAAAATCCCTTTTGAAGGAGCGGACAGCAAA AATCCCTCGCTTTTCATATATGACGCGGAGCGTGTGTATGGCAAGCCCATGAAG GAATGGCTTAAATTCGCCATGGCTGGTGGCAACAGCTGGCGAGGCTCCGGAGACCCC TTCGGCGGCAGACCCGAGCTACGAATGGCAAGGCGGAATGCCCTACTGCCGGGCC AAGGCCAAGCGCGAGCGGCTTTGAAATCATGCAAAAGCTGGGCATCGAATACTTCTGC TTCACGATGTGGACCTTATCGAGGATGGATGACATTCGCCGATACGAAGCCCGCATG AAGACATCAGGACTACCTGCTGAAAGATGAAGGAGACCGGCATCAAGAACCTCTGG GGCACCGCAATGTCTCGGCCAAGCGCTACATGAACGGCGCGGACCAATCCGAG TTCGATGTGGTGGCGCGTCCAGATCAAGAACGCCCTGGACGCCACCATCAAG CTGGCGGCTCCAATATGTCTTGGCGCGCGCGAAGCTATTACCCCTCCTCAAC ACACAGATGACGCGGAAAGACACCTGGCAAGTTGCTGACGGCGCGCGGACTAT GCCCGGCAAGGCTTCAAGGCACTTCTCATTTAGCCCAACCCATGGAACCCACC AAGCACCATACGAGTGGATACGAGAGCGGTCTCGCTTCTCTCGTCCGCAACGGCTG GACAAAGGACTCAAGGTGAACATCGAGTGAACACACCCCTGGCGGCCACACCTTC GAGCATGAGCTCACCGTGGCGCGAGAACGGTTTCTGGGCTCCATCGATGCCAACCGC GGCAGCGCCAGAACGCTGGACACGACCCAGTTCCTGTGGACCCGACGATCTTACC CAGGCCATGATGAGGTGCTGTGAACGGCGCTTCGGACGCGGCGGCACTTCCGAC GCCAACTCGCGCTCCTCCACCGACCTGAGACATCTTCATCGCCCATATTTCCGCC ATGGATGCCATGGCCCGCTTGTCTAACGAGCTGCCGTGCTGGAAGAGAGCCCGCTG TGCCAGATGGTCAAGGAGCGTTATCGCAGCTTCGACGATGGCTCGGCAACAGTTCCGAG GAAGGCAAGGCTACCTCGAAGACCTGTACGAATACGCCAAGGCCCGGCTGAACCGCTT GTCGCTCGGCAAGCAGGAGCTTACGAGACTCTCTGACCTGTATCCCGTCAAGTAA MAKEYFSPSTGKIPFEGADSKNPIAFHYDAGRVVMGKPMKEWLKFAMAWHHTLQASGDP FGGQTRSIEWDKGECPCRAKAKADAGFELMOKLGIIEYCFHDDI DLVEDDDIAEYEARM KDITDYLLEKMKETGINKLWGTVANVFGHKRYMNGAATNPQDIVARAAYQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMOREKDHAKMLTAARDYARAKFGKTFLEPKMEPT KHQYDVTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHFLTVAENGFLGSDANR
5607MI4_005	<i>Prevotella</i>	Amino Acid	134	
5607MI5_002	<i>Prevotella</i>	DNA	135	
5607MI5_002	<i>Prevotella</i>	Amino Acid	136	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5607MI6_002	<i>Prevotella</i>	DNA	137	GDAQNGWDIDQFPVDPYDLTQAMQVLLNGFGNGGTNFAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAAVLEESPLCQMVKERYASFDDGLGKQFEGKATLEDLEYAKAQGEFV VASGKQELYETLLNLYAVK ATGACCAAGAATAATTTCCCTACCTCGGAAGATCCCTTCGAGGGCCCCGAAAGCAAG AACCCATAGGGCTTCCATTAATGACCCCAACCGTCTGGTGTATGGCAAGAAATGAAA GACTGGCTGGCTTTCGCCATGGCTGGTGGCACACCTCGCCAGCGCTGGGGCCAGCAG TTCCGGCCAGACCCGAGTTATCGTGGACGAGGAGAAATGCCGTACGAGCGGCC CGTGCCAAAGGCTACCGCGGCTTCGAGATCATCGAATACTCGGATTCGAGTTCTTCTGC TTCCACGACATCGACCTGATCGAGATACCGACGACATCGCGGATATGAGGCCCGCTG AAAGACATCACGGACTATCTGCTCGAAGATGAAAGCCACTGGCATCAAAAATCTGTGG GGAACGGCAACGTGTTGGCCCAAGCGTTGCATGAACGGCGCCGACCAACCCGGAC TTCCGCGTGTGGCCCGCTGCCCTCCAGATCAAGAAGCCATCGACGCCACCATCAAG CTGGCGGCGAGAACTATGTTCTGGGTTGGCGGCAAGCTACAGAGCTGCTCAAC ACCCAGATCGAGCTGAGAAAGAGACCTGGCGGCTGCTGTCCTGGCCCGGACTAT GGCCGCGCCACGGCTCAAGGTACCTTCTGATCGAGCCCAAGCCGATGGACCGCAGC AAACACCATGACGACGAGATACGGAACCTGTCATCGTTTCTGCGCGCCACGGTCTA GACAAGGACTTCAAGTCAATATCGAGGTGAACCATGCCAGCTGGCGGCCACACCTTC GAACACGAATGGCTGGCGCTGGATCAAGGTATGCTGGCAGCATCGACGCCAACCGC GGTGACGACAGAACGGCTGGATACCGACCGTCCGATCGACAACTTCGAGCTGACCC CTTTCATGCTCCAGATCATCGCAACGGTGGCTGGCATCCCGGGGCTCGAATTTGAT GCCAAGCTGCGCGCAACTCCACGATCCGGAAGACATTTTCATCGGCACATCAGCGCC ATGGACGCAATGGCCCGCATTTGGTCAATGGCGCCGCTATCTGGAGGAGAGCGCTAT CCGAAGATGCTCAAGGAGCTTACGCTTCGTTTCGACAGCGGCAAGGCAAGGAATACGAG GAAGGCAAGCTGACGCTCGAAGACATCGTGGCTATGCCAAGCGCAACGAGAACCGAAG CAGATTTCCGGCAACAGGAACCTACGAGACGCTTGTGCACTCTATAGCAATAA MTKEYFPTVYKIPFEGPESKNPMAFYYPNRLVMGKKMDWLRFMAWHTLQASGDQ FGGQTRSAYAWDEGECPEYERARAKADAGFEIMOKLIGIEFFCFHDIIDLIEDDDIAEYEARL KDITDYLLEKMKATGINKLWGTANVGHKRCMNGAATNPFAVLARAAYQIKNAIDATIK LGGENYVFWGGREGYTSLLNTQMQRKEHLRLSLARDYGRAHGFKFTLIEPKPMGPT KHQYDQDTETVIGFLRRHGLDKDEKVNIEVNHATLAGTTEHELACAVDHGMLGSDANR GDAQNGWDIDQFPIDNFELTSLMLQIRNGGLAPGGSNFDKLRRNSTDPEDIFIAHISA MDAMARALVNAAAILLESAIPKMVKERYASFDSGKGKEYEGLKLTLEDIVAYAKANGEPK QLSGKQELYETLLVALYSK ATGACCAAGGGTATTTCCCTACCATCGGACGATTCCTTCGAGGGGAACTGAAAGCAAG AATCCCTCGCATTCATTAATGATGACCCGACCGGCTCGTACTGGCAAGAAAATGAAA GACTGGCTGGCTTTCGGGATGGCTGGTGGCACACCTGGCCAGCGCTCCGGCCGACAG TTCCGGCGCCAGACCCGAGCTATGCTGGCAAGGCGGAGTCCCTATGAGCGCGCC AAGGCCAAAGCCGACGCGGCTTCGAGATCATCGAATACTCGGATTCGAGTTCTTGT TTCCACGACATGACCTCGTTGAGATACCGACGACATCGCGGATATGAGGCCCGGATG AAGGACATTACCGACTATCTCTGTTCAAGTGAAGGAGCGGGAATCAAGAACCTCTGG GGTACGGCCAAATGCTTCGGCCACAGCGCTATATGAACGGCGCCGCCCAATCCCGAC
5607MI6_002	<i>Prevotella</i>	Amino Acid	138	
5607MI7_002	<i>Prevotella</i>	DNA	139	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5607MI7_002	<i>Prevotella</i>	Amino Acid	140	TTTCGACGTGGTGGCCCGCGCGCTCCAGATCAAGAAACGCCCTCGATGCCACCATCAAG CTGGCGGTGAAACATATGTGTTCTGGGGCGCGCGCAAGGCTATATGAGCCTGCTCAAC ACGCAGATGCAGCGTGAGAAAGAGACCTGGCGCGATGCTGGTCCGCCCGCGACTAC GCCCGCCACCGCTTCAAGGTACCTTCTCATCGAGCCAAACCGATGGAACCGAAC AAGCACCATGACGACGATACGGAACCGTATCGGCTTCTTCCGCCCGCGCGCTG GACAAGGATTTCAAGGTGAACATCGAAGTGAACACCGCTGCGCTGGCGCGCACACCTC GAGCAGAACTGGCCACCGCGCTGACTGCGGCTGCTGGCAGCATCGACGCCAATCGC GGCAGCGCTCAGAACGGTGGATACCGACCGTCCCGATCGCAACTCGAACTCAGC CTGGCCATGCTGCAGATTATCCGCAACGGGCTGGCAGCCCGCGGCTCGAACTTCGAC GCCAACTGGCGCGTAACTCCACCGATCCGGAAGATATCTTCATCGCCACATCAGTGG ATGACGGCATGGCCGCTGGCTGGTCAACCGCGCGCAATCTGGGAAGAGTCTCCCATC CCGCAGATGAAGAAAGAACGCTACGCTCGTTCGACAGCGCAAGGCAAGGAATCGAA GAGGCAAGCTCTGCTCGAAGACCTCGTGGCTATGCCAAGCGCAACGGAGAACCGAAA CAGATCTCCGGCAGGAGCAATATAGACCATCGTCCGCCCTTATGCAAAATAG MTKGYFPTIGRI PFEGTESKNPLAFHYE PDRVLVGKKMKDWLRFAMAWHHTLGQASGDQ FGQTRSAYWDKAECPYERAKAKADAGFEIMQKLGIEFFCFHDIDLVEDTDDIAEYEARM KDI TDYLLVMKMETGINKLWGTANVFHKRYMNGAATNPEDVVVRAAVQIKNALDATIK LGENYVFWGGREGYMSLINTQMQRKEHLGRMIVAARDYARAHGFKGTFLIEPKMEPT KHQYDQDETIVIGFLRRHGLDKFVNIEVNHATLAGHTFEHELATAVDCGLLSIDANR GDAQNGWTDQFPI DNFELTLMLQIIRNGGLAPGGSNFDKLRNSTDPEDIFIAHISA MDAMARALVNAAA IWEESP I PQMKERYASFDGKGEFEFGKLCLEDLVAYAKANGEPK QISGRQELYETIVALLYCK ATGACCAACGAGTATTTCCCGGAATCCGTGTGATTCGGTTGAAGGACAGGAAAGCAAG AATCCCATGGCTTTCATATTATGACGCCAACCGGTAGTATGGGCAAAACCCATGAAG GAATGGTTCAAAATTTGCCATGGCTGGTGCAATACGTGGGCGAGGATCGGCCGATCCC TTCGGCGGACAGACCCGCTCCTACGATGGGACAGGGCGAGTGCCTTACTGCCGTGCC CGCCAGAGGCCGACGCCGCTTGAACGTGATGAGAGCTGGGTATCGGCTATTTTGTG TTCACGATGTGATATCATCGAGACTGCGAAGACATGCGGAGTATGAGGCCCGGTAG AAGGACATCAGGACTATCTGCTGTGATGAAGTGAAGAAACGGGCATCAAGAACCTGTGG GGCAGGCCAACGCTCTTCGGCCCAAGCGCTATATGAACGGCGCTGCCCAACCCGACG TTCGACGTGGTGGCGCGCTGGTCCAGATCAAGAACCGCTGGAGCCACCATCAAG CTGGCGGACGCAATTAAGTGTCTGGGGCGCGCGAAGGTATTAACCTTTTGGAAAC ACGCAGATCGGGCGGAGAGGACACCTGGCCAGATGCTCAAGGACCCGCTGATAT GCCCGGCAAGGATTAAGGACGCTTCTCATTTAGCCCAAGCCATGGAGCCACCC AAGCACCATGACGATAGATACGAGACCGTATGCTTCTGGCGCAACCGGACTG GACAAGGACTTCAAGGTGAATATCGAAGTGAACACCGCCCTGGCGGCCACACCTTC GAGCACGAACTCACCGTGGCGCGGAAACCGCTTCTGGCAGCATCGACGCCACCGC GGAGACGCCAGAACCGTGGGTACAGACCATTTCCCATAGATGCCCTTGACCTACCC CAGGCCATGATGAGGTCTGCTCAACGGCGGATTCGGCAACGGCGGACCAACTTCGAC GCCAACTGGCGCGCTTCTCCAGGATCCGAGGACATCTTCATCGCCACATCGCGCC ATGGACGCCATGGCCCGCCCTCTGAAACGCCCGCCATCTCTGGAAGAGAGCCCCATG
5608MI1_004	<i>Prevotella</i>	DNA	141	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5608MI1_004	<i>Prevotella</i>	Amino Acid	142	CCGGGCAATGGTGAAGAGCGCTACGCTTCCTTCGACAAATGGCCTTGGCAAGAAGTTCGAG GAAGGAAAGGCCACGCTGGAAGAGCTGTACGACTATGCCAAGAAGAACGGCGAGCCCTGTG GCCGCTTCGGCAAGCAGGAACCTGACGAAACGCTGCTGACCTGTACGCCAAGTAA MTNEYFPGIGV I PFEQESKNPMFHY YDANRVVMGKPMKEWFKFAMAWHHTLGQASADP FGGQTRSYAWDKGECPCYRARQKADAGFELMQKLIGYGFCEHVDV D I I EDCEDIAEYEARM KDITDYLIVKMKETGTKNLWGTANVFHKRYMNGAATNPQDFWAAVQIKNALDATIK LGGSNYVFWGGREGYYTLWNTQMRREKDHAKMLKAARDYARGKFGKTFLEPKPMEPT KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSDANR GDAQNGWDTDQFPIDAFDLTQAMMQVLLNGGFGNGGTNFDKLRSSSTDPEDFIAHISAM MDAMAHALLNAAAII LEESPMPGMVRYASFNDNLGKKEEGKATLEELYDYAKKNGEVP AASGQELYETLLNLYAK
5608MI2_002	<i>Prevotella</i>	DNA	143	ATGAAAGAAATACTTCCCTACCATCGAAAAAATCCCTTTCGAGGGGCCCTCAGAGCAAGAAT CCGCTCGCATTCATTAATAGCGCAACCGCGTGTGCGCGGCAACCCATGAAGGAC TGCTCAAGTTCGCCATGGCTTGGTGGCACACCTGGGCGAGCATCGGAGACCCCTTC GGCGCCAGACCCGACGTACGAGTGGGACAAAGCCGAGTCCCTTACTGCCGTGCCGT GAAAAGGCCGACGCGGCTTCGAGATCATGAGAAACTTGAATCGAGTACTTCTGCTTC CATGACATCGACCTTGTGGAAGACTGCGAGGACATTCGCCAGTACGAGGCCCGCATGAAG GACATCAGGACTACCTCTTGAGAAAGATGAAGGCCACCGGCATCAAGAACCTGTGGGGC ACCGCAACGCTTTGGCAACAAGCGTACATCAACGGCGAGCGACCAACCCCTCACTTC GACATCGTTGCCGTGCAGTGTCCAGATCAAGAACGCAATCGACGCAAAATCAAGCTG GGCGTACCGGTACGTATCTGGGGCGCGGAGGCTACTACACCTCCTGAACACC CAGATGCGCGGAGAGGACACCTTGCCAGATGCTACCGCAGCGCGGACTACGCC CGCGCAAGGGAATCAAGGGCAATTCCTCATCGAGCCCAAGCCCATGGAGCCCAACG CACCAGTACGATGTTGACACGGAAACCGTCATCGGCTTCCTCGCGCAACCGGCTGGAC AAGGACTTCAAGGTGAACATCGAGTGAACCGCACCCCTGGCGGCGCACACCTTCGAG CAGAGTCAACGTCGCGTGGACAAAGGCTTCCTGGGAGCATCGACGCAACCGCGGC GACGCCAGAACGGCTGGACACTGACCACTTCCCTGTGGATCCTTACGACCTCACCCAG GCAATGATGAGATTATCGCAACGGCGCTTCAAGGACGGCGGCGCACCACTTCGACGCC AAACTCCGCGCAGTCCACGGACCCGAGGACATCTTATCGCCCATCATGAGGCAATG GATGCAATGGCACACGCCCTCATCAACGCTGTCAGTGTGAGGAAAGCCCTCTGTGC GAGATGGTGTCAAAGCGTACGCCAGCTTGACAGCGGCTTGCGCAAGAGTTCGAGGAA GGCAAGCCACTCTCGAGGAGATCTACGAGTATGCCAAGAGGCCCGCCGACCCGCTGCC GCCTCCGGCAAGCAGGAGCTCTACGAGACACTCTCAATCTGTACGCTAAATAA
5608MI2_002	<i>Prevotella</i>	Amino Acid	144	MKEYFPTTGKIPFEGPQSKNPLAFIHYDANRVVAGKPMKDWLKFAMAWHHTLGAAADPF GGQTRSYEMDKAECPCYRAREKADAGFEIMQKLIGYFCFHDIDLVEDEEDIAEYEARMK DITDYLLKMKATGINKLWGTANVFGNKRYMNGAATNPQDFVARAAVQIKNAIDATIKL GGTGYVFWGGREGYYTLINTQMRREKDHAKMLTAARDYARAKGKTFLEPKPMEPTK HQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSDANR GDAQNGWDTDQFPVDPVDTQAMMQIIRNGGFKDGTNFDKLRSSSTDPEDFIAHISAM DAMAHALLNAAAII LEESPMPGMVRYASFDSGLGKKEEGKATLEELYDYAKKAPVA AASGQELYETLLNLYAK

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5608MI3_004	<i>Prevotella</i>	DNA	145	ATGACCAAGAGATATTTCCCTACAAATCGAAAGATTTCCCTCGAAGGCCCGGAGAGCAAG AATCCGCTGGCATTCATTAATGAACCCGACAGAATCATCTCGGAGGAAGATGAAG GACTGGCTGGCTTCGCGTGGCTGGTGGACACCTCGCGAGCGTCCGGCGACAG TTGGAGGCCAGACCCGAACTATGCTGGAGAGCCGAAATGCCGCTAGAGCGCGG AAAGCCAGGCCGACGCGGCTTCGAGCTGATCGAAGCTGGGCATCGAGTATTTCTGC TTCCACGCTAGACCTCATAGAGGCGCGCAACCATCGAAGATATGAGGAGCGCATG GGCATCAAAACCGACTACCTGCTCGGGAAGATGAAGAGACAGGTATCAAGAACCTTGG GGCACGCCAAACGTTTCGCGCCACAAGCGTTACATGAACGAGCCGCCACCAACCCGAC TTGACGTTGGTGGCGCTGGCGCTGCAGATCAAGAACGCGCATCGACGCCACCATCAAG CTGGCGCGAGAAATACGTATTTCTGGGCGGACGCGAGGCTATGCAAGCCTGCTCAAC ACTCAGATCGAGCGGAGAAACACCTGGAGCGCATGCTGGCTGCAGCCCGGACTAT GCGCGCGCCACGATTCAGGGCATTTCTCATCGAGCCCAACCCATGGAGCGCTACC AAGCACGATACGACAGGATACGAGACCGTTATCGCTTCTCGGAGGAACGGCTC GACAAGGATTTCAAGTTAAACATCGAGGTGAACACCGCCACCTGGCGGGCCACACCTTC GACACGAACTGGCGTGGCAGTGGACAAACGCTGCTTGGCAGCATCGACGCCAACCGC GGCACGCGCAGAACCGATGGACACCGACCTTCCCATCGACAACTTCGAGCTCAC CAGCCATCTGCAGATAATCCGCAACGGCGGACTGGGACCGCGGATCGAACTTCGAC GCCAAGCTGGCGCGCAATTCACCGACCTGAGGATATCTTCATCGCCACATCAGTGG ATGGACGCAATGGACGCGCTGGCAAAACGCGCGCAATCATCGAAGAGAGCCCATC CCCCAATCTGAAGGAGGCTACGATCGTTCGACAGGCAAGGCAAGGAGTTCGAG GACGGCAACTGAGCTCGAAGAACTGAGCTACGCAAGCGCAAGCGGAGCGCGGAAAG CAGATTTCCGGCAAGCAGAACTCTACGAAACCATAGTGCCCTCTATTGCAAGTAA
5608MI3_004	<i>Prevotella</i>	Amino Acid	146	MTKEYFPTIGKIPFEGPESKNPLAFHYIYEPDRIILGRKMDWIRFAVAWHHTLGQSGDQ FGGQTRNYAWDEPECPEVERAKAKADAGFELMQKLIGIEYFCHFDVDLIEFAATIEEYERM GLITDYLLGKMKETGILKNLWGTANVGHKRYMNGAATNPDFDVVARAAVQIKNAIDATIK LGGENYVFWGGREGYASLLNTQMQRKDHILGRMLAAARDYGRAHGFKGTFLIEPKMEPT KHQYDQDTETVIAFLRRNLGDKDFKVNIEVNHATLAGHTFEHELAVADNGLLGSIDANR GDAQNGWDTDQFPIDNFELTQAMLQILRNGGLTGGSNFDKLRNSTDPEDIFIAHISA MDAMARALANAAAIIEESPIPAMLKERYASFDSGKGKEFEDGKLSLEELVAYAKANGEPK QISGKQELLYETIVALYCK
5609MI1_005	<i>Prevotella</i>	DNA	147	ATGGCACAAAGAATACTTCCCTACCATGGGAAATCCCTTCGAGGGCACTGAGAGCAAG AATCCCTTGCCTTCCATTAATGAGCCGGAGCGCATTTGTCTGGCGAAACCCATGA GAATGGCTCAAGTTTGCCATGGCTGGTGGCACACGCTGGGGCAGGCATCGGCCGATCCC TTGGCGGCGCAAAACCGCAGCTATGCTGGGTAAGGGGAAATGCCCTACTGCGCTGCC CGGCCAAGCGGACGCGGCTTCGAGATTATGCAAAAGCTGGGCATCGAGTACTTCTGC TTCCACGATATCGACCTGGTAGAAGACTGAGCATATTCGGGAATACGAAGCCCGCATG AAGGACATCAACGCTACCTCTCGAGAGATGAAGGAACCGGATCAAGAACCTCTGG GGCACCGCAATGTGTTGGTCAAGCGCTACATGAACGCGCCGCGCCACCAACCCGAG TTGACGCTAGTGGCCGCTGCGCTGTTGAGTTAAGAACGCGCATTGACGCCACCATCAAG TTGGCGGTGCCAATACGTGTTCTGGGCGCGCGGAGGCTATTACAGCCTCCTGAAC ACCAGATCGAGCGGGGAGGAGGACCATCGCCAAAGCTGCTACGGGACCGCGGACTAT

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5609MI1_005	<i>Prevotella</i>	Amino Acid	148	GCCCCGCCAACGGCTTCAAGGGAACCTTCTGATTTGAGCCCAAGCCCATGGAGCCACCC AAGCACCATACGACGTGGATACGAGAGCGGTTCCTGGCTTCCTCCGGCCCAACGGGCTG GACAAAGGACTTCAAGTGAATATCGAGGTGAACACACCGCTTGGCCGCCACACCTTT GAGCAGAGCTCACCGTGGCCCGGAGAAAGGCTTCTGGCAGCATCGACGCCAAACGGC GGCGATGCCAGAACGGCTGGATACGACCTTCCCGGTAGACGCTTATGAGCTCAC CAGGCCATGATGAGGTCTCTGAACGAGGCTTCGGCAACGGCGGACCAACTTCGAC GCCAAAGTGGCCGCTCTCCAGGACCCCGGAGACATCTTCATCGCCCATATCAGTGG ATGGATGCCATGGCCACCGCTGCTCAACGGCCGCGCTGCTGGAGAAAGCCCGCTG TGCCAGATGGTGAAGAGCGCTACCGAGCTTACAGCGCTCCGGGCAAGCAGTTCGAG GAAGAAAGGCCACCCCTGGAGACCTGTACAACTACGCCAAAGCCACCGGTGAACCGCTG GTTCCTCCGGCAAGCAGGAACCTTACGAGACCTCTGAACTCTATGCAAGTAG MAQYFPTIGKI PFEGETSKNPLAFHYPERIVCGKPKMEWLKFAMAWHHTLGQASADP FGGQTRSVAWDKGECPYCRARAKADAGFEIMQKLGIEYFCFHDIDLVEDCDDIAEYARM KDITDYLLKMKETGIKNLWGTANVFGHKRYMNGAATNPQDVVARAAVQIKNAIDATIK LGANYVFWGGREGYSLINTQMQREKDLAKLLTAARDYARANGFKGTFLIEPKMEPT KHQYDVTETVIGFLRANGDKDFKNIEVNHATLAGHTFEHELTVAENGFLGSDANR GDAQNGWTDQFPVDAYELTQAMQVLLNGFGNGGTFDAKLRRSSDTPEDIFIAHISA MDAMAHALINAAVLESPLCQMVKERYASFDSPGPKQFEFGKATLEDLYNYAKATGEPV VASGQELYETLLNLYAK
5610MI1_003	<i>Prevotella</i>	DNA	149	AAGGCACAGAAATACCTTCCCTACCATTTGGGAAAATCCCCITCGAGGGCACTGAGAGCAAG AATCCCCCTGCTTTCATTAATACTATGAGCCGGAGCGCATTTGCTCGGCAAAACCCATGAAA GAATGGCTCAAGTTTGCCATGGCTGTGGCACACGCTGGGCAGGCATCGGCCGATCCC TTCCGGCGCCAAACCCGAGCTATGCTGGGTAAGGGCAATGCCCTACTGCCGTGCC CGTGCCAAAGCGGACGCCGTTTGGATTATGCAAAAGCTGGGCATCGAGTACTTCTGC TTCCAGATATCGACCTGGTAGAAGACTTGACGATATTGGGAATACGAAGCCCGCAIG AAGGACATCAGGACTACCTCCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG GGCACCGCCAAITGTGTTGGTCACAAGCGCTACATGAACGGCCCGGCACCAATCCGCAG TTTGACGTGGTGGCCGCTGCTGCGTGCAATCAAGAACGCCATTGACGCCACCATCAAG TTGGCGGTGCCAATTACGTGTTCTGGGGCGCGCGAGGCTATTACGCCCTCCTGAAC ACCCAGATGACGGGAGAGGACACCTGGCCAAAGCTGCTCACGGACCCCGCGCTAT GCCCCGCCAACCGCTTCAAGGGAACCTTCTGATTTAGCCCAAGCCCATGGAGCCACCC AAGCACAGTACGACGTGGATACGAGACGGTCAATTTGGCTTCTCCGGCCCAACGGCTG GACAAGGACTTCAAGTGAATATCGAGGTGAACACCGCCGCTGGCCGCCACACCTTT GAGCAGAACTCACCGTGGCCCGGAGAACGGCTTCTGGCAGCATCGACGCCCAACCGC GGCGATGCCAGAACCGCTGGGATAGGACCAAGTTCGGGTAGACGCTTATGAGTACCC CAGGCCATGATGAGGTGCTCTGAACGGAGGCTTCGGCACGGCCGGCACCAACTTCGAC GCCAAAGTGGCCGCTCTCCAGGACCTGGAGACATCTTCATCGCCCATATCAGTGG ATGGATGCCATGGCCACCGCTGCTCAACGCCCGCGCTGCTGGAGAAAGCCCGCTG TGCCAGATGGTGAAGAGCGCTACCCAGCTTACAGCGGCTCCGGGCAAGCAGTTCGAG GAAGGAAAGGCCACCTGGAGGACCTGTACAACTACGCCAAAGCCAAAGGTGAACCGCTG GTTCCTCCGGCAAGCAGGAACCTTACGAGACCTCTGAACTCTATGCAAGTAG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5610MI1_003	<i>Prevotella</i>	Amino Acid	150	MAQYFPTIGKIPFEGTESKNPLAFHYPERIVCGKPKMEWLKFAMAWHHTLGQASADP FGGQTRSYAWDKGECPCYCRARAKADAGFEIMQKLGIEYFCHDIDLVEDCDDIAEYEARM KDITDYLLEKMKETGINKLWGTANVFGHKRYMNGAGTNPOFDVARAAVQIKNAIDATIK LGGANYVFWGGREGYYSLNTQMOREKDLAKLLTAARDYARANGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLDKFKVNIENVNATLAGHTEHELTVAENGFLGSDANR GDAQNGWDIDQFPVDAYDLTQAMQVLLNGGFGNGGTNFDALRRSSTDIEDIFIAHISA MDAMAHALLNAAAALLESPLCQMVKERYASFDSPGPKQFEGKATLEDLYNAYKANGEPV VASGKQELYETLLNLYAK
5610MI2_004	<i>Prevotella</i>	DNA	151	ATGGCAAAAATAATTTCCCTACCATCGCAAGATTCTTTTGAAGGAACCGACAGCAAG AGTCCCTCGCCCTTCCATTAATGACGCCCCAGCGCTTGTGATGGGAAACCCCATGAAG GAATGGCTCAAGTTGGCATGGCTGGTGACACACCTGGCCAGGCATCGGCCGACCCC TTGGGGGTGACAGCCCGCACTATGCTGGGATGAAGGGAATGCCCTACTGCCGCGCC AAAGCCAGGCCGAGCGCGCTTCGAGATCATGAGAAATGGGCATCGAGTACTTCTGC TTCCAGATGTGGACCTGGTGAAGACTGGAGACATCGCGAGTACGAAGCCCGCATG AAGGACATCAGGACTACCTGCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG GGCAGGCCAATGTTCGGCCACAAGCGTTACATGAACGGCGCCGGACCAACCCGACG TTTGACATGTGGCCCGCTGCCGTCCAGTCAAAACCGCCCTGGACGCCACCATCAAG CTGGGGGTTCCAACTACGTGTTCTGGGGCAGCGCGAAGCTACTACACCTCTCTGAAC ACCCAGATGACGCGGAGAAAGCACCTGGCCAAAGCTCTGACCGCGCCCGCGGACTAC GCCGGCCAAAGGTTCAAGGAACTTCAAGGAACTTCTCATCGAGCCAAACCCATGGAGCCAC AAGACCCAGTACGACGTGACACCGAGACCGTAATCGGCTTCTCGTCCCAACGCGCTG GACAAGGACTCAAGGTGAACATCGAGGTGAACACACCGCTTCTCGTGGCCACACCTTC GAGACGAACTCACCGTCGCCGTGAAAACGGCTTCTCGGATCGATCGACGCCAACCGC GGCAGCCCGAGAACGGCTGGACACCGACCTCCCGTAGACGCTATGACCTCAC CAGGCCATGATGAGGTGCTGTGAACGGCGTTTCGGCAATGGCGGTACCAACTTCGAC GCCAAGCTCCGCGCTCTCCAGGATCCGGAAGACATCTTCATCGCCACATCAGCGCC ATGGACGCCATGGCCACGCGCTGCTGAACGCGCGCGCTGCTGGAAGAAAGCCGCTT CCCGCATGGCGAAAGAGCGCTACGCTCTTTGACAGCGGACTTGGCAAGAAAGTTGAA GAGGAAAGGCCACCTCGAAGAGTGTACGACTATGCCAAGGCTAACGACGCCCTGTG GCGGCTCCGGCAAGCAGGAACCTTACGAAACCTTCTTGAACCTCTATGCAAAATAG MAKEYFPTIGKIPFEGTDSKSPFLAFHYIDAQRYVMGPKMEWLKFAMAWHHTLGQASADP FGGQTRHYAWDEGECPCYCRARAKADAGFEIMQKLGIEYFCHDIDLVEDCDDIAEYEARM KDITDYLLEKMKETGINKLWGTANVFGHKRYMNGAGTNPOFDVARAAVQIKNALDATIK LGGSNYVFWGSREGYYTLLNTQMOREKDLAKLLTAARDYARAKGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLDKFKVNIENVNATLAGHTEHELTVAENGFLGSDANR GDAQNGWDIDQFPVDAYDLTQAMQVLLNGGFGNGGTNFDALRRSSTDIEDIFIAHISA MDAMAHALLNAAAALLESPLPAMAKERYASFDGLGKGFEGKATLEELYDYAKANDAPV AASGKQELYETFLNLYAK
5610MI2_004	<i>Prevotella</i>	Amino Acid	152	MAQYFPTIGKIPFEGTESKNPLAFHYPERIVCGKPKMEWLKFAMAWHHTLGQASADP FGGQTRSYAWDKGECPCYCRARAKADAGFEIMQKLGIEYFCHDIDLVEDCDDIAEYEARM KDITDYLLEKMKETGINKLWGTANVFGHKRYMNGAGTNPOFDVARAAVQIKNAIDATIK LGGANYVFWGGREGYYSLNTQMOREKDLAKLLTAARDYARANGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLDKFKVNIENVNATLAGHTEHELTVAENGFLGSDANR GDAQNGWDIDQFPVDAYDLTQAMQVLLNGGFGNGGTNFDALRRSSTDIEDIFIAHISA MDAMAHALLNAAAALLESPLCQMVKERYASFDSPGPKQFEGKATLEDLYNAYKANGEPV VASGKQELYETLLNLYAK
5751MI1_003	<i>Prevotella</i>	DNA	153	ATGGCAAAAATAATTTCCCTACCATCGCAAGATTCTTTTGAAGGAACCGACAGCAAG AGTCCCTCGCCCTTCCATTAATGACGCCCCAGCGCTTGTGATGGGAAACCCCATGAAG GAATGGCTCAAGTTGGCATGGCTGGTGACACACCTGGCCAGGCATCGGCCGACCCC TTGGGGGTGACAGCCCGCACTATGCTGGGATGAAGGGAATGCCCTACTGCCGCGCC AAAGCCAGGCCGAGCGCGCTTCGAGATCATGAGAAATGGGCATCGAGTACTTCTGC TTCCAGATGTGGACCTGGTGAAGACTGGAGACATCGCGAGTACGAAGCCCGCATG AAGGACATCAGGACTACCTGCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG GGCAGGCCAATGTTCGGCCACAAGCGTTACATGAACGGCGCCGGACCAACCCGACG TTTGACATGTGGCCCGCTGCCGTCCAGTCAAAACCGCCCTGGACGCCACCATCAAG CTGGGGGTTCCAACTACGTGTTCTGGGGCAGCGCGAAGCTACTACACCTCTCTGAAC ACCCAGATGACGCGGAGAAAGCACCTGGCCAAAGCTCTGACCGCGCCCGCGGACTAC GCCGGCCAAAGGTTCAAGGAACTTCAAGGAACTTCTCATCGAGCCAAACCCATGGAGCCAC AAGACCCAGTACGACGTGACACCGAGACCGTAATCGGCTTCTCGTCCCAACGCGCTG GACAAGGACTCAAGGTGAACATCGAGGTGAACACACCGCTTCTCGTGGCCACACCTTC GAGACGAACTCACCGTCGCCGTGAAAACGGCTTCTCGGATCGATCGACGCCAACCGC GGCAGCCCGAGAACGGCTGGACACCGACCTCCCGTAGACGCTATGACCTCAC CAGGCCATGATGAGGTGCTGTGAACGGCGTTTCGGCAATGGCGGTACCAACTTCGAC GCCAAGCTCCGCGCTCTCCAGGATCCGGAAGACATCTTCATCGCCACATCAGCGCC ATGGACGCCATGGCCACGCGCTGCTGAACGCGCGCGCTGCTGGAAGAAAGCCGCTT CCCGCATGGCGAAAGAGCGCTACGCTCTTTGACAGCGGACTTGGCAAGAAAGTTGAA GAGGAAAGGCCACCTCGAAGAGTGTACGACTATGCCAAGGCTAACGACGCCCTGTG GCGGCTCCGGCAAGCAGGAACCTTACGAAACCTTCTTGAACCTCTATGCAAAATAG MAKEYFPTIGKIPFEGTDSKSPFLAFHYIDAQRYVMGPKMEWLKFAMAWHHTLGQASADP FGGQTRHYAWDEGECPCYCRARAKADAGFEIMQKLGIEYFCHDIDLVEDCDDIAEYEARM KDITDYLLEKMKETGINKLWGTANVFGHKRYMNGAGTNPOFDVARAAVQIKNALDATIK LGGSNYVFWGSREGYYTLLNTQMOREKDLAKLLTAARDYARAKGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLDKFKVNIENVNATLAGHTEHELTVAENGFLGSDANR GDAQNGWDIDQFPVDAYDLTQAMQVLLNGGFGNGGTNFDALRRSSTDIEDIFIAHISA MDAMAHALLNAAAALLESPLPAMAKERYASFDGLGKGFEGKATLEELYDYAKANDAPV AASGKQELYETFLNLYAK
5751MI1_003	<i>Prevotella</i>	DNA	153	ATGGCAAAAATAATTTCCCTACCATCGCAAGATTCTTTTGAAGGAACCGACAGCAAG AGTCCCTCGCCCTTCCATTAATGACGCCCCAGCGCTTGTGATGGGAAACCCCATGAAG GAATGGCTCAAGTTGGCATGGCTGGTGACACACCTGGCCAGGCATCGGCCGACCCC TTGGGGGTGACAGCCCGCACTATGCTGGGATGAAGGGAATGCCCTACTGCCGCGCC AAAGCCAGGCCGAGCGCGCTTCGAGATCATGAGAAATGGGCATCGAGTACTTCTGC TTCCAGATGTGGACCTGGTGAAGACTGGAGACATCGCGAGTACGAAGCCCGCATG AAGGACATCAGGACTACCTGCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG GGCAGGCCAATGTTCGGCCACAAGCGTTACATGAACGGCGCCGGACCAACCCGACG TTTGACATGTGGCCCGCTGCCGTCCAGTCAAAACCGCCCTGGACGCCACCATCAAG CTGGGGGTTCCAACTACGTGTTCTGGGGCAGCGCGAAGCTACTACACCTCTCTGAAC ACCCAGATGACGCGGAGAAAGCACCTGGCCAAAGCTCTGACCGCGCCCGCGGACTAC GCCGGCCAAAGGTTCAAGGAACTTCAAGGAACTTCTCATCGAGCCAAACCCATGGAGCCAC AAGACCCAGTACGACGTGACACCGAGACCGTAATCGGCTTCTCGTCCCAACGCGCTG GACAAGGACTCAAGGTGAACATCGAGGTGAACACACCGCTTCTCGTGGCCACACCTTC GAGACGAACTCACCGTCGCCGTGAAAACGGCTTCTCGGATCGATCGACGCCAACCGC GGCAGCCCGAGAACGGCTGGACACCGACCTCCCGTAGACGCTATGACCTCAC CAGGCCATGATGAGGTGCTGTGAACGGCGTTTCGGCAATGGCGGTACCAACTTCGAC GCCAAGCTCCGCGCTCTCCAGGATCCGGAAGACATCTTCATCGCCACATCAGCGCC ATGGACGCCATGGCCACGCGCTGCTGAACGCGCGCGCTGCTGGAAGAAAGCCGCTT CCCGCATGGCGAAAGAGCGCTACGCTCTTTGACAGCGGACTTGGCAAGAAAGTTGAA GAGGAAAGGCCACCTCGAAGAGTGTACGACTATGCCAAGGCTAACGACGCCCTGTG GCGGCTCCGGCAAGCAGGAACCTTACGAAACCTTCTTGAACCTCTATGCAAAATAG MAKEYFPTIGKIPFEGTDSKSPFLAFHYIDAQRYVMGPKMEWLKFAMAWHHTLGQASADP FGGQTRHYAWDEGECPCYCRARAKADAGFEIMQKLGIEYFCHDIDLVEDCDDIAEYEARM KDITDYLLEKMKETGINKLWGTANVFGHKRYMNGAGTNPOFDVARAAVQIKNALDATIK LGGSNYVFWGSREGYYTLLNTQMOREKDLAKLLTAARDYARAKGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLDKFKVNIENVNATLAGHTEHELTVAENGFLGSDANR GDAQNGWDIDQFPVDAYDLTQAMQVLLNGGFGNGGTNFDALRRSSTDIEDIFIAHISA MDAMAHALLNAAAALLESPLPAMAKERYASFDGLGKGFEGKATLEELYDYAKANDAPV AASGKQELYETFLNLYAK

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				TTTCGGCGCCAGACCCGACAGTACAGTGGGATCTTTCGACGACACCCCGAGCAGCGCGCA AAGACAAAGCTCGACGCGGCTTCGAAATAATGAGAAACITGGAATCAAGTATTTCTGT TTCACGATGTTGACCTTATCGAAGACAGCGAGATATTGGACATATGAGGCTCGTCTC AAGAACCTTACAGACTAGCTGCAGAGCAGATGAAGCTCCAGACATCAAGCTCCTCTGG GGTACAGCGAATGTATTCGGCAACAAGCGCTACATGAACGTCGGGTACAAACCCCTGAT TTCGATGTAGTTGCCGCGCAGCGTTTTCAGATTAAAGAACGCTATCGACGACCATCAAG CTCGGTGTACAGCTATGTATTTCTGGGCGGCTGTCGAGGATATCAGAGCTTGTCTCAAC ACTCAGATCGAGCTGAGAAGGACACCTCGCAACCATGCTTACATCGCTCGGACTAT GCTCGCAGCAAGGCTTACCGGAACCTTCTTATCGAGCTAAGCCGATGGAGCTACAC AAACCCAGTACGCTAGATACAGAGACTGTTGTCGGCTTCTCAAGCACACCGGCTG GACAAGGACTTCAAGGTAATATCGAGGTTAACACGCACTCTCGAGGCCACACCTTC GAGCAGCACTCACCGTGTGTGATTAACGGAATGCTCGGTCTATCGACGCTAACCGC GGTATGCACAGAACCGCTGGATACAGACCAAGTCCCTTAAGCGCTGAGGAGCTTACC CTCGTATGATGAGATATCCGTAATGGTGGCTTGGCAACGAGGATCCAACCTTCGAC GCAAGCTTCGCGCACTCTACGATCTGAAGACATCTTCATCGCACACATCTGCGGT ATGGATGCAATGGCACACGCTCTCCTCAAGCAGCTGCAATATCGAGGACTCTCCTATC CCTACAAATGGTTAAGGAGGTTACGCTTCTTCGACAGCGGTATGGTAAGGACTTCGAG GATGGAAGCTTACCTCGAGGATCTCTACAGCTACGCGTGAAGAACGGAGAGCGCAAG CAGACGCGCAAGCAGGAGCTCTATGAGACTCTCATGAATATCTATGCAAGTAA
5751MI1_003	<i>Prevotella</i>	Amino Acid	154	MAKQYFPQIGIKKEFTESKNPLAFHYDANRVVLGKMEEWLKFMAWHTLQASGDQ FGGQTRSYEWDLIAATPEQRAKDKLDAGFEIMEKLIGIKYCFHDVLDLIEDSDDIATYEAL KDLTDYAEQMKLHDIKLLWGIANVFNKRYMNGAATNPDFDVARAAYQIKNAIDATIK LGTSYVFWGGREGYQSLINTQOREKDLHATMLTIARDYARSKGFTGTFLEPKMEPT KHQYDVTETVVGFLKAHGLDKDFKVNIEVNHATLAGHTFEHETVAVDNMGMLGSDANR GDAQNGWDIDQFPVSABELTLAMNQIIRNGGLNGSGSNFDAKLRRNSTDPEDIFIAHICG MDAMAHALLNAAAIIEESPIPTMVKERYASFDSGMGKDFEDGKLTLEDLYSYGVKNGEPK QTSAKQELYETLMNIYCK
5751MI2_003	<i>Prevotella</i>	DNA	155	ATGGCAAAAAGAAATTTTCCACAAGTAGGCAAGATTCCAATTGAGGGTCTCGAAAGTACT AACGTACTCGCATTCCTACTACTATGATCCAGAACGCGAAGTTCTTGGTAAGAAAAATGAAA GATTGGCTGAAGTATGCTATGGCTTGGTGGCACACACTCGGTGAGGCAAGTGGCGACCAA TTCGGTGGTCAAACTCGTTCGTATGAATGGATGAAGCCGACGATGTTCTCAACGGCA AAGGATAAATGGATGCTGGTTTTGAATTGATGACCAAACTTGGCATTAATACTACTGTC TTCCATGATGTCGACCTTATGAAGAGGTGCAACAAATGAAGAAATATGAAGCTCGTATG CAAGCTATCACCGACTACGCAATTAGAAAAACAAAAAGAACCCGGCATTAAGCTCCTTTGG GGTACTGCTAATGTGTTGGTCATAAGCGTTATGAATGGTGGCGGCAACAAACCCCTGAC TTTGTATGATGCTCGGCTGCTACAAATCAAGAACGCTATCGATCAACTATCAAG CTTGGTGGTCAAACTATGATTTCTGGGTGGCGCGAAGGTATATGAGTTTGTCTCAAC ACTCAATGCAACGCGAAAAAGACCACTTGGCAAGATGCTTACCGGAGCTCGGACTAT GCTCGTGTAAAGGCTCAAGGGTACATTCCTCGTTGAACCTAAGCCATGGAACCAACT AAGCATCAATATGATACCGATACGAAAACTGTGATGTTTCTCGTGGTAAATGGTCTT GAAAAAGACTTCAAGGTGAACATTGAAGTGAACCATGCTACTCTCGCTCAGCACACTTC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5751MI2_003	<i>Prevotella</i>	Amino Acid	156	GAACACGAACTCGTGTGGCTGTGACAAATGGCAATGCGATGCTCGGTCTCTATCGACGCTAACCGT GGCGATGCTCAAAATGGCTGGGATACGACCAATTCCTCAATCGACAACTACGAACCTCACC CTCGTATGCTCCAAATCATTCGCAATGGTGTGCTTGGCAATGGCGTAGCAACCTCGAC GCTAAGATCGTCTGTAATAGCACCGACCTTGAAGACCTCTTATCGCTCACATCAGTGGT ATGGATGCTATGGCTCGTGCCTCTCAATGCTGCTGCAATCGTGAAGAGCGCAAT CCTGCTATGTTAAGCAGCGTTATGCAAGCTCTGATGAGTATGGTAAAGGACTTCGAA GAAGGAAAACCTCACTCTCGAAACACTCGTAGACTATGCTAAGGCTAACGGCGAAACCTGCT ACAGTAAGGGGCAAGCAAGAAAGTATGAAACTCTCGTGTCTCTCTACGCTAAGTAA MAKEFFPQYVKIPFEGPESTNVLAHYDPEREVLGKKMKDWLKYAMAWHHTLQASGDQ FGGQTRSIEWDEADDVLQRAKDKMDAGFELMTKLGIYYCFHDVLDLIEGATIEEYARM QAITDYALEKQKETGIKLWGTANVFGHKRYMNGAATNPEDVVARAAVQIKNAIDATIK LGGQNYVFWGGREGYMSLLNTQMREKDLAKMLTAARDYARAKGFKGTFIVEPKMEPT KHQYDTEFTVIGFLRANGLEKDFKVNIEVNHATLAQHTFEHELAVAVDNGMLGSI DANR GDAQNGWTDIDQFPI DNYELTLAMLQIIRNGGLNGGSLNDAKIRRNSTDLEDLFIAHISG MDAMARALINAAAIVKSEI PAMLKQRYASSDAGMGKDFEFGKLTLEQLVDYAKANGEP TVSGKQKEYETLVLYAK
5752MI1_003	<i>Prevotella</i>	DNA	157	ATGACTAAAGAGTATTTCCCGGGAATCGGAAGATTCGGTTTGAAGGAACCAAGACGAAG AACCCCTTGGCCTTCCATATTATTAACGCCCTCCAGGTAGCGATGGGCAAGCCCATGAAG GACTGGCTCAAGATGTCATGGCTGTGGCACACCTGGCGAGCCCTCGCAGACCCC TTTGGGGCAGACCCGCTCTACGAATGGGACAGGGCGAGTGCCTTATTCGCCGCC AAGCAGAGGCCGATGCCGCTTTGAGCTCATCGAAGCTGGGCATCGAGTACTACTGC TTCACGACGTGGACATCATCGAGACTCGAGGACATTCGCGAGTACGAGGCCCGCATG AAGGACATCAGGACTACCTGCTGGAGAGCAGAAAGACCGGCATCAAGAACCTCTGG GGCACCGCAACGTTTGGCCACAAAGCGCTACATGAACGGCGCCGCCCAACCCCTCAG TTTGACATTTGGCCCGTCCGCGTCCAGATCAAGAACGCTGTGATGCCACCATCAAG CTGGGTGTTACCACTAGTGTCTGGGTGGCGGAGGCTACTACGGTGTCTAAC ACCCAGATCGAGCGGAGAGAACACCTGGCCAGATGCTACCGCCCGCCGCGACTAC GCCCGCCAAAGGCTTCAAGGGACCTTCTCATTTAGCCCAACCCATGGAGCCACCC AAGCACCTAGTACGACGTGGACACCGAGACCGTGTGTTTCATCGCGCCCAACGGCTG GACAAGGACTTCAAGGTAAACATTGAGGTAAACACGCCCTGGCGGCCACACCTT GAGCAGAGCTCACCGTGGCCGAGAGACGGCTTCCCTGCTCCATCGACGCCCAACCGC GGAGTCCCAGAACGGTGGGTACGGACCACTTCCCTCGACCGCTGGATCTCAC CAGGCTATGATGAGGTTCCTCAACGGTGGTTCGGCAATGGCGGCCACCACTTGAC GCCAAGCTCCCGCTCTCCACGATCCGAGGACATCTTCATCGCCACATCAGCGCC ATGGATGCCATGGCACAGCCCTCTGAAACGAGCCGCTCTCTGGAGAGAGAGCCCTG CCCGCATGGTCAAGGAGCGTTACGCTTCTTCGACAGCGCTCTGGGCAAGAGTTTCGAA GAAGCAAGGCCCTCCCTGGAAGACTTACGAATATGCCAAGAAAGATGGAGAGCCCGTG GCCGCTTCGGGCAACAGGAGCTTCGGAACCTTACTTGAACCTCTATGCAAGTAG
5752MI1_003	<i>Prevotella</i>	Amino Acid	158	MTKEYFPQYVKIPFEGTFSKNPLAFHYINASQVAMGKPMKDWLKYAMAWHHTLQASADP FGGQTRSIEWDKGECPCRAKQKADAGFELMQKLGIYYCFHDVDI IEDEDIAEYARM KDITDYLEKQKETGIKNLWGTANVFGHKRYMNGAATNPQDIVARAADVQIKNALDATIK

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5752MI2_003	<i>Prevotella</i>	DNA	159	LGGTNYVFWGGREGYYTLLNTQMOREKNHLAKMLTAARDYARAKFGKGTFLIEPKMEPT KHQYDVTETVIGFIRANGLDKDFKVNIEVNHATLAGHTEHELTIVARENGFLGSDANR GDAQNGWTDQFPIDALDLTQAMQVILNGFGNGGTNFDALRRSSTDPEDIFIAHISA MDAMAHALLNAAAIIIEESPLPAMVKERYASFDSGLGKKEGKASLEELYEYAKKNGEVP AASGKQELCETYLNLYAK ATGACTAAAGAGTATTTCCCGGGGAATCGAAAGATTCCGTTTGAAGGACCAAGAGCAAG AACCCCTGGCCTTCCATATTATAACGCCTCCAGGTAGTGCATGGCAAGCCCCATGAAG GACTGGCTCAAGTATGCCATGGCTGGTGACACACCTGGGCCAGGCCTCTGCAGACCCC TTTGGCGGCAGACCCGCTCCTACGAATGGACAAGGGCGAGTCCCCCTACTGCCGCGCC AAGCAGAAGCCCGATGCCGCTTTAGCTCATGCAGAAGCTGGGCATCGAGTACTACTGC TTCCAGAGCTGGACATCATCGAGACTCGAGGACATTGCCGAGTACGAGGCCCGCATG AAGGACATCAGCGACTACCTGCTGGAGAAGCAGAAAAGAGACCGGCATCAAGAACCCTTGG GGCACGCCAACCGTGTGGCCACAAGCGCTACATGAACGGCGCCGCCAACCAACCTCAG TTTGACATTGGCCCGTGGCCGCTCCAGTCAAGAACGCCCTGGATGCCACCATCAAA CTGGTGTGTAACAACCTACGTTCTGGGTGGCGGAAGCTACTACACGCTGCTCAAC ACCCAGATCAGCGGGAAGAAGAACACCTGGCAAGATGTCACCGCGCCCGCGACTAC GCCCCGCCAAGGCTTCAAGGCACCTTCTCATTTGAGCCCAACCATGGAGGCCACCC AAGCACAGTACGACCTGGACACCGAGACCGTGTGTTTATCCGCGCCCAACGGCCTG GACAAGGACTTCAAGGTAACATTAAGTAACACACCGCCACCTGGCGGCCACACCTTT GAGACGAGCTCACCGTGGCGGAGAAACGGCTTCTGGGCTCCATCGACGCCAACCCG GGAGTGGCCAGAACCGCTGGATACGACCGCTTCCCATCGACGCCCTGATCTCAC CAGCTATGATGACGATCATCTCAACGGTGGCTTCGGAATGGCGGACCAACTTTGAC GCCAAGCTCGCGCTCTCCACGATCCGAGGACATCTTCATCGCCACATCAGCGCC ATGATGCCATGGCACACCGCTCCTGAACGCGACCGCCATCCTGGAAGAGAGCCCGCTG CCGCGCATGGTCAAGGAGCGTTACGCTTCTCGACAGCGCTTGGCAAGAAAGTTGAA GAAGCAAGCCCTCCCTGGAAGACTTACGAATATGCCAAGAAATGGAGAGCCCGTG GCCGCTTCGGCAACAGAGGACTCTCGAAACTTACTTGAACCTCTATGCAAAAGTAG MTKEYFPGLGKIPFEGTKSKNPLAFHYINASQVVMGKPMKDWLKYAMAWWHITLGOASADP FGQTRSYEWDKGEPCYRAKQKADAGFELMQLGIEYYCFHDVDIIEDCEDIAEYEARM KDI TDYLLKQKETGINKNLWGTVANVFGHKRYMNGAATNPQDIVARA VQIKNALDARIK LGGTNYVFWGGREGYYTLLNTQMOREKNHLAKMLTAARDYARAKFGKGTFLIEPKMEPT KHQYDVTETVIGFIRANGLDKDFKVNIEVNHATLAGHTEHELTIVARENGFLGSDANR GDAQNGWTDQFPIDALDLTQAMQVILNGFGNGGTNFDALRRSSTDPEDIFIAHISA MDAMAHALLNAAAIIIEESPLPAMVKERYASFDSGLGKKEGKASLEELYEYAKKNGEVP AASGKQELCETYLNLYAK ATGGCAAAAGAGTATTTCCCGACTATCGGCAAGATTCCCTTCGAGGGCGTCGAATCCAAG AACCCGATGGCATTCACCTACTATGACGGAACCGCGTCCGTGGATGGCAAGCCCCATGAAG GACTGGCTCAAGTTCGCGATGGCTGGTGACACACCTGGGACAGGCTTCGGCGGACCCG TTTGGCGGCAGACCCGCTCCTACGAGTGGGACAAGCGGAGTCCCCCTACTGCCGCGCC AAGGCCAAGCCGCGCGCTTCGAGATCATCAGAAGCTCGGTATCGAGTACTACTGTC TTCCATGACATCGACCTCGTGGAGGACACCGAGGACATCGCCGAGTACGAGGCCCGCATG
5752MI2_003	<i>Prevotella</i>	Amino Acid	160	MTKEYFPGLGKIPFEGTKSKNPLAFHYINASQVVMGKPMKDWLKYAMAWWHITLGOASADP FGQTRSYEWDKGEPCYRAKQKADAGFELMQLGIEYYCFHDVDIIEDCEDIAEYEARM KDI TDYLLKQKETGINKNLWGTVANVFGHKRYMNGAATNPQDIVARA VQIKNALDARIK LGGTNYVFWGGREGYYTLLNTQMOREKNHLAKMLTAARDYARAKFGKGTFLIEPKMEPT KHQYDVTETVIGFIRANGLDKDFKVNIEVNHATLAGHTEHELTIVARENGFLGSDANR GDAQNGWTDQFPIDALDLTQAMQVILNGFGNGGTNFDALRRSSTDPEDIFIAHISA MDAMAHALLNAAAIIIEESPLPAMVKERYASFDSGLGKKEGKASLEELYEYAKKNGEVP AASGKQELCETYLNLYAK ATGGCAAAAGAGTATTTCCCGACTATCGGCAAGATTCCCTTCGAGGGCGTCGAATCCAAG AACCCGATGGCATTCACCTACTATGACGGAACCGCGTCCGTGGATGGCAAGCCCCATGAAG GACTGGCTCAAGTTCGCGATGGCTGGTGACACACCTGGGACAGGCTTCGGCGGACCCG TTTGGCGGCAGACCCGCTCCTACGAGTGGGACAAGCGGAGTCCCCCTACTGCCGCGCC AAGGCCAAGCCGCGCGCTTCGAGATCATCAGAAGCTCGGTATCGAGTACTACTGTC TTCCATGACATCGACCTCGTGGAGGACACCGAGGACATCGCCGAGTACGAGGCCCGCATG
5752MI3_002	<i>Prevotella</i>	DNA	161	ATGGCAAAAGAGTATTTCCCGACTATCGGCAAGATTCCCTTCGAGGGCGTCGAATCCAAG AACCCGATGGCATTCACCTACTATGACGGAACCGCGTCCGTGGATGGCAAGCCCCATGAAG GACTGGCTCAAGTTCGCGATGGCTGGTGACACACCTGGGACAGGCTTCGGCGGACCCG TTTGGCGGCAGACCCGCTCCTACGAGTGGGACAAGCGGAGTCCCCCTACTGCCGCGCC AAGGCCAAGCCGCGCGCTTCGAGATCATCAGAAGCTCGGTATCGAGTACTACTGTC TTCCATGACATCGACCTCGTGGAGGACACCGAGGACATCGCCGAGTACGAGGCCCGCATG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				AAGGACATCACCGATACCTCGTCGAGAAAGCAAGAAACCGGCATCAAGAACCCTGG GGCAGGGCAACGTTTCGGCAACAAGCGCTACATGAACGGCGCCGCCAGAACCCGACG TTCGACGTCGTCGCGCGCGCTCCAGATCAAGAAGCGCATCGACCAACCATCAAG CTCGGGTACCGGTTACGTTCTGGGGCGCGTGAAGCTACTACACCTCTCTGAAC ACCCAGATCGAGCGGAGAGAACACCTCGCCAAGATGCTACCGCGCCCGCGACTAC GCGCGCGCCACGGTTCAGGGCACCTCTCATGAGCCCAAGCCCATGAGCCACCC AAGCACCATACGATCGGACACGAGACCGGTGATCGGCTCTCTGGCGCAACGGTCTG GACAAGGACTTCAAGGTCAATATCGAGGTGAACACCGCCCTCGCCGGCCACACCTTC GAGCAGGACTCACCGTGGCTGCGATAACGGCTCTCTCGGCTCCATCGACGCCAACCGC GGCAGCGCCAGAACGGCTGGACACCGACACCGAGTCCCCGTGGACCCGTACGACCTACC CAGGCCATGATGAGATCATCCGCAACGGCGGTTCAAGGACGGCGGCACCAACTTCGAC GCCAAGCTCGCGCGCTCTCCAGCACCGGAGGACATCTCATCGCCACATCAGCGCG ATGACGGCATGGCCACCGCCCTGCTGAACCGCGCGCTCATCGAGAGAGCGCGCTC TGCAAGATGGTCGAGGAGGCTACGCTCTCTTCGACAGCGGCTCGGCAAGCAGTTCGAG GAAGCAAGGCCACCTCGAGGACCTTACGAGTATGCCAAGAAGATGGCGAGCGCTC GTCGCTCGCGCAAGCAGGAGCTCTACGAGACGCTGCTGAACCTTTACGCGAAGTAG
5752MI3_002	<i>Prevotella</i>	Amino Acid	162	MAKEYFPTIGKIPFEGVESKNPMAFYHYDANRVVMGKPMKDWLKFAMAWHHTLGQASGDP FGGQTRSYEWDKGECPYRAKAKADAGFEIMQKLGIEYCFHDIDLVEDTEDIAEYFARM KDITDYLVKEQKETGINKNLWGIANVFGNKRYMNGAATNPQFDVVARAAVQIKNAIDATIK LGCTGYVFWGGREGYYTLLNTQMORKEKDLAKMLTAARDYARAHGFQGTFLIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKNIEVNHTLAGHTFEHFLTAVDNGFLGSTDANR GDAQNGWDIDQFPVDPYDLTQAMMQIIRNGGFKDGGTNFADKLRRSTDPEIDIFIAHISA MDAMAHALLNAAAVIEESPLCKMVEERYASFDSGLGKQFEGKATLEDLYEYAKNGEPV VASGKQELYETLLNLYAK
5752MI5_003	<i>Prevotella</i>	DNA	163	ATGGCAAAAGAGATATTTCCCGACAAATCGGTAAAGTCCCTTCGAGGGACCCGAGTCCAAAG AACCCGATGGCATTCACATACTATGACGGGAGCGGTGGTGTGATGGGCAAGAAGATGAAG GACTGGTTCAAGTTCGGATGGCTGGTGGCACACCTTGGGCCAGGCTTCGCGCCGACCCG TTCGGCGGCAGACCCGCTCCTACGAGTGGCAAGGGCGAAGGCCCTGCTCCCGCGCC CGGCCAAGGCTGACCGCGGTTTCGAGATCATGCAAGAACTGGGCATCGGCTACTACTGC TTCACGACATCGACCTGGTGGAGACACCGGAGACATCGCCGAGTATGAAGCCCGCATG AAGGACATCACCGATACCTCGTGAGAGAGCAGAAGGAGACCGGCATCAAGAACCCTCTGG GGCACGGCCAAACGTATTCGGCAACAAGCCCTACATGAACGGCGCCGCCAAGAACCCGAG TTCGACATCGCCCGCGCGGCTCGAGACCAAGAACGACCATCGATGCCACCATCAAG CTGGCGGCAACGGTACGTTCTGGGGCGGCTGAGGTACTACACCTCTCTGAAC ACCCAGATCGAGCGGAGAGGACACCTTGCACAGATGCTACCGCGGCTCGGACTAT GCCCGGCCACCGCTTCAAGGACCTTCTTCATCGAGCGCAACCCATGGAGCCGACCC AAGCACCATGATACGCTGGACACGAGACCGGTGATCGGCTCTCTCGCGCCAAACGGCTG GACAAGGACTTCAAGGTGAACATCGAAGTGAACACGCCACCTTCGCGGCGCACACCTTC GAGCAGGCTCACCGTGGCGTTGACAAACGGCTTCTTCGCGAGCATCGACGCCAACCCG GGAGACGCCAGACCGGTGGGTACGACCGAGTTCGCGGTGATCCGACGACCTCACCC CAGCGCATGATCCAGATCATCCGCAATGGCGGCTTCAAGGACGGCGGTACCAACTTCGAC

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5752MI5_003	<i>Prevotella</i>	Amino Acid	164	GCCAAAGTCGCGCGCTTCCACCGACCCGAGGACATCTTCATCGCCACATCAGCGCG ATGACGCGCATGGCCACGCGCTGTGAACGCGCGCGCTGCTCGAGGAGAGCCCGCTC TGGAGATGGTTGCAAGCGCTTACGCTTCCACGAGCGCTTCCGCAAGAGTTCCGAG GAAGGCAACGCCACCTCGAGGAATCTACGATACGCCAAGGCGAGGCGAGGTCGTT GCCGAATCCGCGCAAGCAGGAATCTACGAGACCTGCTGACCTCTACCGCAAGTAG MAKEYFPTIGKIPFEGESKNPMAFYDDAERVVMGKMKDFKFMAMWHTLQASADP FGGQTRSVEWDKGEPCSRARAKADAGFEIMQKLGIGYCYCFHDIDLVEDIEDAEYARM KDI TDYLVKQKETGINKNLWGTVGNKPYMNGAATNPQDIAARAALQIKNAIDATIK LGGTYVFWGGREGYTYLLNTQMREKDLAKMLTAARDYARAHGFKGTFIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKNIEVNHATLAGHTFEHGLTVAVDNGFLGSDANR GDAQNGWDTDQFPVDPYDLTQAMIQIIRNGGFKDGTNFAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAAVLEESPLCEMVAKRYASFDSGLGKKEFNATLEELYEYAKAGEVV AESGQELYETILLNLYAK
5752MI6_004	<i>Prevotella</i>	DNA	165	ATGGCAAAAGAGTATTTCCCGACAAATCGGAAGATCCCCITCGAGGGCGCTGAGAGCAAG AATCCCCITGCTTTCCACTATTATGACGCGGAGCGTGTGGTCATGGGCAAGCCCATGAAG GACTGGTTCAAGTTTCGGATGGCTGTGGCACACCTGGCGCAGGCTTCGCGCCGCCCG TTCGGCGGCAGACCCGCTCTACGAGTGGGCAAGGGCGAGTGCCTACTGCGCGCTACTGCG CGCCAGAGGCTGACGCGGTTTCGAGATCATCGAAGCTCGGCATCGGCTACTACTGCG TTCACGACATCGACCTGTCGAGCACCGGAGACATCGCGAGTACGAGGCGCGCAIG AAGGACATCACCGACTACCTCGTCGAGAAGCAGAGGAGACCGGCATCAAGAACCCTGG GGCAGGCCAACGTTTCGGCAACAGCGCTACATGAACGCGCGCCGACGACCCGACG TTCGACATCGTCGCCACGCGGCTCGAGATCAAGAAGCGGATCGGCGCACCATCAAG CTCGGCGGACCGGTTAGTGTTCGGGGCGGCTGAGGTACTACACCCCGCGCGCTAC ACCCAGATCGAGCGCGAGAGGACACCTCGCCAAAGTGTACACCGCGCGCGGCTAC GCCGCGCAACGGCTTCAAGGCGACCTTCTCATCGAGCGAAGCGGATGGAGCCACC AAGCACAGTATGACGTGGACACGAGACCGTGTGATCGGCTTCTCGCGCCACAGGCTG GACAAGGACTTCAAGGTGAACATCGAGGTGAACACCGCCCTCGCGCGCCACACCTTC GAGCACGAGCTCACCGTGGCGTTCGACAAACGGCTTCTCGGACGATCGACGCCAACCGC GGTGACGCCAGAACGGCTGGATACCGACAGTTCGCGGTGGATCCGTACGATCTACCC CAGCGATGATCCAGATCATCCGCAACGGCGGCTTCAAGGATGGCGGACCACTTCGAC GCCAAGCTCCGCGCTCTCCACGACCGGAGGACATCTCATCGCCACATCAGCGCG ATGACGCGATGGCCACGCGCTGTGAACGCGCGCGCTCATCGAGAGAGCGCGCTC TGGAGATGGTCGCCAAGCGCTACGCTTCTTCGACAGCGGCTCTCGGCAAGAAGTTTCGAG GAAGCAACGCCACCTCGAGGAATCTACGATACGCGGAGCGGAGCGGTGAGGTCAAG GCCGAATCCGCGCAAGCAGGAGCTCTACGAGACCTTCTGAACTCTACCGCAAGTAG
5752MI6_004	<i>Prevotella</i>	Amino Acid	166	MAKEYFPTIGKIPFEGESKNPMAFYDDAERVVMGKMKDFKFMAMWHTLQASADP FGGQTRSVEWDKGEPCSRARQKADAGFEIMQKLGIGYCYCFHDIDLVEDIEDAEYARM KDI TDYLVKQKETGINKNLWGTVGNKRYMNGAATNPQDIAHAALQIKNAIGATIK LGGTYVFWGGREGYTYLLNTQMREKDLAKMLTAARDYARANGFKGTFIEPKMEPT KHQYDVTETVIGFLRANGLDKDFKNIEVNHATLAGHTFEHGLTVAVDNGFLGSDANR GDAQNGWDTDQFPVDPYDLTQAMIQIIRNGGFKDGTNFAKLRRSSTDPEDIFIAHISA

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5753MI1_002	<i>Prevotella</i>	DNA	167	MDAMAHALLNAAAIVIEESPLCEMVAKRYASFDSGLGKKFEEGNATLEELYEYAKANGEVK AESGQELYETLLNLYAK ATGGCAAAAAGAGATATTTCCCCACATCGGGAAGATTCTCTTGAAGGAGTCGAGAGCAAG AACCCCTTGCAATCCATATATGACGCAACCGCATGTGTCATGGCAAGCCCCATGAAG GACTGGTTCAAGTTCCGATGGCATGGTGGCACACCTGGGACAGCCCTCCGACAGACCCG TTCCGGCGGCAGACCCGCTCTACGAATGGACAAGGGCGAATGCCCTACTGCGCGGCC AGGGCAAGGCCGATGCGGCTTCGAGATCATGAGAAATCGGATTCGAGTATTTCTGC TTCCATGACATCGACCTGGTAGAGACTGGACGATCGCCGAGTACGAGGCCCGCATG AAGGACATCAGGACTATCTCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG GGCACGCCAACGTTTCGGCAACAAGCGTTACATGAACGGCGCCGACCAATCCGACG TTCCAGTAGTGGCCCGCTGCCGTCAGATCAAGAACCGCATCGACGCCACCATCAAG CTCGGCGGTTCCAACATATGTGTTCTGGGGCGCGGTGAAGGATACACCTGCTGAAC ACCCGATCAGCGCGAGAAGGACACCTCGCAAACTGCTACCGCGCCCGCGCATAT GCCCGAAGAACGGTTCAAGGACACCTTCTTATCGAGCCCAAGCCGATGGAGCCACAC AAGCACCTAGTACGACCTGGATACGACCGGCTGATCGGCTCTCTCGCGCCAAACGGCCTG GAGAAAGACTCAAGGTGAACATCGAGGTGAACACCGCCCTGCGCGGCCCATACCTTC GAGCATGAACCTCACCTGGCGGTGGACAAAGGCTTCTTGGGATCCATCGACGCCAACCGC GGCAGCCCGAGAACGGCTGGATACGACCGACCTCCCGGTAGACCCGTACGACCTCAC CAGGCCATGATCAGATCATCCGCAACGGCGCTCGGCAACGGCGGTACCAACTTCGAC GCCAACTGGCGCTTCTCCACCATCTGAGGACATCTTATCGCCACATCAGCGCC ATGACCGCATGGCCCGCCCTGCTCAACGACCGCCGCTGCTGGAAGAAAGTCCGCTC TGTGAGATGTTCAAGGAGCGCTACGCTTCTTCGACAGCGCTCGGCAAGAAAGTTCGAA GAGGCAAGGCTACCTGGAAGAAATCTACGATATGCCAAGAGCGCGCAACCCGCTG GTCGCTTCGGCAAGCAGGAGCTACGAAACCTGCTGAACTCTACGCCAAGTAG MAKEYFPTIGKIPFEGVESKNPLAFHYIDANRMVMGKPMKDFKAMAWWHITLGOASADP FGQTRSYEWDKGEPCYRARAADAGFEIMQKLGIEYCFHDI DLVEDDDIAEYFARM KDI TDYLLKMKETGIKNLWGTANVFGNKRMYNGAGTNPOFDVVARAAVQIKNAIDATIK LGGSNYVEWGGREGYYTLLNTQMREKDLGLKLLTAARDYARKNGFKGTFLEPKMEPT KHQYDVTETVIGFLRANGLEKDFKVNIEVNHATLAGHTFEHETLVAVDNGFLGSDANR GDAQNGWDIDQFPVDPYDLTQAMQIIRNGGLNGGTFNDAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAAIVIEESPLCEMVAKRYASFDSGLGKKFEEGNATLEELYEYAKKSPEPV VASGQELYETLLNLYAK ATGGCTAAGAATACTTCCCTCCATCGGCAAAATCCCTTTTGAAGGAGGCGACAGCAAA AATCCCTCGCTTTCATATATGACCGCGGACGCTGTTATGGCAAGCCCCATGAAG GAATGGCTTAAATTCGCCATGGCTGGTGGCACACGCTGGCCAGGCCCTCGGAGACCCC TTCCGGCGGCAGACCCGAGCTACGAATGGGCAAGGGCGAATGCCCTACTGCGCGGCC AAAGCCAGGCCGACCGCGGTTTGAATCATGCAAAAGCTGGGTATCGAATACTTCTGC TTCCACGATGTGGACCTTATCGAGATTGGATGACATTCGCGAATACGAAGCCCGCATG AAGGACATCAGGACTACCTGCTGAAAGATGAAGGAGCCGGCATCAAGAACCTCTGG GGCACGCCCAATGTCTCGGCCCAAGCGCTACATCAAGAACCGCATCGACGCCACCATCAAG TTCCAGTAGTGGCCCGCTGCCGTCAGATCAAGAACCGCATCGACGCCACCATCAAG
5753MI1_002	<i>Prevotella</i>	Amino Acid	168	
5753MI2_002	<i>Prevotella</i>	DNA	169	

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
				CTCGGGGTACCAAGTATGTAATCTTGGGGGGCGCGGAGGCTACTACACCTCTCTGAAC ACCCAGATGACGGTGAAGAAAGACACCTGGCCAAAGATGCTACCGCGAGCCCGGACTAC GCCCCGCAAGGGCTCAAGGGACCTTCTCATCGAGCCCAAGCCGATGGAGCCACAC AAGCAACAGTACGACGTTGACACGGAGACCGTGTATCGGCTCCCTGGCGCAACGGCTG GACAAGGACTTCAAGGTGAACATCGAGGTGAACACCGCCCTGGCCGCCACACCTTC GAGCAGAACTACCGTGGCTGTGACAAAGGCTTCTGGGCTCCATCGACGCCAACCGC GGCAGCCCAAGACGGCTGGATACGGACCACTTCCGGTAGACCCGTACGACCTACAC CAGCCATGATGACGATATCCGCAACGGCGCTTCAAGACGGCGGACCACTTCGAT GCCAACTGGCGGCTCTCCACGATCCGGAAGACATCTCATCGCCACATCAGCGCT ATGATGCCATGGCACACGCCCTGCTCAACGCCCGCGCTGTGTAAGAGAGAGCCGCTG TGCAACATGGTCAAGGAGCGTTACCGCGCTTCGACAGCGGCTTGGCAAGAAGTTCGAG GAAGGGAAGGCAACGCTGGAGGAATCTATGACTATGCCAAGAGAGCGGCGAACCGTC GTGGCTTCGGCAAGCAGGAACCTACGAACCATCTGAACCTTATGCCCAAGTAG
5753MI2_002	<i>Prevotella</i>	Amino Acid	170	MAKEYFPSIGKIPFEGGDSKNPLAFHYIDAGRVVMGKPKMEWLKFAMAWHHTLGQASGDP FGGQTRSYEWDKGECPYCRAKAKADAGFEIMQKLGIEYFCHDVDLIEDDDIAEYFARM KDITDYLLKMKETGIRNLWGTANVFHKRYMNGAATNPQFDVVARAAYQIKNAIDATIK LGGTSYVFWGGREGYYTLLNTQMOREKDLAKMLTAARDYARAKFGKTFLEPKMEPT KHQYDVTETVIGSLRANGLDKDFKVNIEVNHATLAGHTFEHETVAVDNGFLGSTDANR GDAQNGWDIDQFPVDPYDLTQAMMQIIRNGGFKDGGTNFADLRSSSTDPEFIAHISA MDAMAHALLNAAAVLEESPLCNMVKERYAGFDSGLGKKEGKATLEEIYDYAKKSGEPV VASGKQELLYETILNLVAK
5753MI4_002	<i>Prevotella</i>	DNA	171	ATGTCAAAAGAGATATTCCTACAAATCGGACGGTCCCCCTCGAGGGACCTGAGAGCAAG AATCCGCTGGCGTTCCACTATTACAGCCGGACCGGCTCGTCTGGGAGGAAAAATGAAG GACTGGCTGGCTTCGCAATGGCTGGTGCATACGCTCGGGCAGGCTTCGGCGGACCA TTCGGCGGACAGACCTGCACATACGCTGGATGAAGGCGAGTGTCCGCTCTGCCGGCA AAGCCAAAGCTGACGCCGGCTTGAACATGATGACGAAACTGGGCAATCGGCTATTTCTGC TTCCACGAGCTGGACCTGGTCGAGGAGGCCGACACCATTGAAGAAATACGAGGAGCGGATG CGGATCATCAACGACTACCTGCTCGAGAAGATGGAAGAGACCGGCATCGCAATCTGG GGAACGCCAATGTCTTCGGACACAAGCGCTATATGAACGGCGCCGCCAATCCCGAC TTCGACGTGCTGGCCGTGCCGGTCCAGATCAAGAAATGCATCGATCCACCATCAAA CTGGGTGTGAGAACTATGTTCTGGGTGGCGGAGGCTATACGAGCCTGTCTCAAC ACGAGATGCAACGGGAAAAACACACCTCGAAATATGCTCAGGCAAGCCCGGCGCTAT GGCGTGGCCACGGTTTCAAGGGAAGCTTCTGATCGAGCCCAAGCCGATGAGCGCGAC AAGCATCAGTACGACGAGTACGAGACGCTATCGGTTCTCTGGCTGTACGGCTG GACAAGGATTTCAAGGTGAACATCGAGGTGAACACGCCCTCGTCCCGGACACACCTTC GAGCAGAACTGGCCATGCGGTGATGCCGGCTGTGGCAGCATCGATGCCAACCGC GGCAGCGCCAGAACCGGTGGATACCGACCTAGTCCCGTCCGAACTACGAACTACG CTGGCGATGCTGAGATCATCCGCAATGGCGGACTCGACCCGGCGGATCGAACTTCGAT GCCAAGTTGGCGCGCAATTCACCGATCCGGAAGACATCTTCATCGCCCACTCAGCGG ATGGACGCGATGGCCGCTGCTGCTCAATGGCGGCGCATCTGGACCGAATCGCGGAT CAGGATATGGTCAGGACCGCTATGCTTCTTCGACAGCGGCAAGGGCAGGATTCGAG

TABLE 2

Clone No.	Class of organism	Type of Sequence	SEQ ID NO:	Sequence
5753M14_002	<i>Prevotella</i>	Amino Acid	172	GAAGGCAGACTCAGTCTGGAAGACCTCGTGGCTATATGCGAAGGAGCAGCGTGAGCCGCGC CAGATCTCCGGCAGGAGAACTTATGAAACCATCGTAGCGCTTACITGACAGTTAA MSKEYFPRTGRVPFEGPESKNPLAFHYYPDRVLVGRMKDWLRFAMAWHHTLQASGDQ FGGQTCTYAWDEGECPCRAKAKADAGFELMQKLGIGYFCFHDVLDLVEADTIEEYERM RLITDYLLEKMEETGIRNLWGTVFHKRYMNGAATNPEDVVARAAVQIKNAIDATIK LGGENYVFWGGREGYTSLLNTQMREKHHLGNMIRAARDYGRAHGFKGTFLIEPKMEPT KHQYDQDTETVIGFLRCHGLDKFKNIEVNHATLAGHTFEHELATAVDAGLLGSDANR GDAQNGWDTDQFPIDNYELTLMLQIIRNGGLAPGSGNFDKLRNSTDPEDIFIAHISA MDAMARALINAAAIIWTESPIQDMVRDRYASFDGKGRFEFEGRISLEDIVAYAKEHGEPR QISGRQELYETIVALLYCR
5752M14_004	<i>Prevotella</i>	DNA	173	ATGACTAAAGAGTATTTCCCGGGAAATCGGAACGATTCGGTTTGAAGGAACCAAGACGAAG AACCCCTTGGCCTTCCATATATATAACGCCCTCCAGGTAGTATGGGCAAGCCCATGAAG GACTGGCTCAAGTATGCCATGGCTGGTGCCACACCCCTGGCCAGGCCTCTGCAGACCCC TTTGGCGGCAGACCCGCTCTACGAATGGGACAAGGGCAGTGCCTACTGCCGCGCC AAGCAGAAAGCCGATGCCGCTTTGAGCTCATGCAGAAAGCTGGGCATCGAGTACTACTGC TTCCACGACGTGGACATCATCGAGACTCGGAGACATTCGAGTACGAGGCCCGCATG AAGACATACGGACTACCTGCTGGAGAAGCAAGAGACCGGCATCAAGAACCTCTGG GGCACCGCAACCGTGTGGCCACAAGCGCTACATGAACGGCGCCGCCAACCCCTCAG TTIGACATGTGGCCGTGGCGCTCCAGATCAAGAACGCCCTGGATGCCGCTCAAAA CTGGGTGTACCAACTAGTGTCTGGGGTGGCGGGAAGGCTACTACCGCCCGCGGACTAC ACCCAGATGCAGCGGGAGAAGAACACCTGGCCAAAGATGCTCACCGCCCGCGGACTAC GCCCGGCAAGGGCTCAAGGCCCTTCTCATTTAGCCCAAAACCATGGAGCCACCC AAGCACAGTACGACGTGGACACCGAGACCGTGTGTTTTCATCGCGCCCAACGGCCTG GACAAGGACTTCAAGGTAAACATGAGGTAAACACGCCCTTGGCTCCATCGACGCCAACCGC GAGCAGGCTCACCGTGGCCGCGAGAACGGCTTCTGGGTCCATCGACGCCAACCGC GGAGTCCCAGAACGGCTGGGTACGGACCGTTCCTCCCATCGACGCCCTGGATCTCACC CAGGCTATGATGCAGGTTCCTCAACGGTGGCTTCGGCAATGGCGGCACCAACTTIGAC GCCAAGCTCCGCGCTCTCCACGATCCGAGGACATCTCATCGCCACATCAGCGCC ATGGATGCCATGGCACAGCCCTCTGTAAACGAGCCGCCATCTCTGGAGAGAGCCCCCTG CCCGCCATGGTCAAGGAGCGTTACGCTTCTTCACAGCGGTCTGGGCAAGAAGTTCGAA GAAGGCAAGGCCCTCCCTGGAAGAACTTACGAATATGCCAAGAAGAAATGGAGAGCCGTG GCCGCTTCGGCAACAGGAGCTTCGGAACCTTACTTGAACCTCTATGCAAGTAG
5752M14_004	<i>Prevotella</i>	Amino Acid	174	MTKEYFPGITIPFEGTKSNPLAFHYYNASQVMGKPMKDWLKYAMAWHHTLQASADP FGGQTRSYEWDKGECPYCRAKQKADAGFELMQKLGIEYCFHDVDIIEDECIDAEYEARM KDIIDYLLEKQKETGIRNLWGTVFHKRYMNGAATNPQDIVARAIVQIKNALDAIK LGGTNYVFWGGREGYTYLLNTQMREKHHLAKMLTAARDYARAKGFKGTFLIEPKMEPT KHQYDQDTETVIGFIRANGLDKDFKNIEVNHATLAGHTFEHETVARENGFLGSDANR GDAQNGWDTDQFPIDALDLTQAMMQVILNGGFNGGNTFNAKLRNSTDPEDIFIAHISA MDAMAHALLNAAAIIIESPLIPAMVVKERYASFDGKGGKFEFGKASLEELYEYAKNGEPV AASGKQELCETYNLYAK

5.5 Example 4: Quantification Of XI Enzyme Activity

[0097] The clones identified in the ABD and SBD screens (see Table 2) were subcloned into vector p426PGK1 (FIG. 3), a modified version of p426GPD (ATCC accession number 87361) in which the GPD promoter was replaced with the PGK1 promoter from *Saccharomyces cerevisiae* (ATCC accession number 204501) gDNA. The clones were then transformed into yeast strain MYA11008.

[0098] Cells were grown as described in the materials and methods. Cell pellets were resuspended in about 300µl of lysis buffer: approximate concentrations (50 mM NaH₂PO₄ (pH 8.0), 300 mM NaCl, 10 mM imidazole (Sigma, #I5513), to which was added about 2 µl/ml beta-mercaptoethanol (BME)), and protease inhibitor cocktail tablet (Roche, 11836170001) (1 tablet for about 10 ml cell extract). The cell suspension was added to a 2 ml screw-cap microcentrifuge tube that had been pre-aliquotted with about 0.5 ml of acid washed glass beads (425-600 µm). Cells were lysed using a FastPrep-24 (MP Biomedicals, Solon, OH) at amplitude setting of about 6 for about 3 repetitions of about 1 minute. Cells were chilled on ice for about 5 minutes between repetitions. Samples were centrifuged at about 10,000 x g for about 10 minutes at 4°C. Recovered supernatants were used in the XI enzyme activity assay. XI enzyme activity was performed as described in the materials and methods. Results are shown in Table 3.

Table 3		
XI activity at pH 7.5		
SEQ ID NO:	Volumetric Activity	FIOPC
2	-60.73	2.58
4	-21.84	0.93
6	0.86	-0.05
8	-2.14	0.12
10	-2.38	0.13
12	-12.82	0.54
14	-26.97	1.45
16	-76.50	4.12
18	-15.32	0.83
20	-5.33	0.29
22	0.48	-0.03
24	0.36	-0.02
26	0.81	-0.04
28	-6.65	0.36
30	-9.10	0.49
32	-38.10	2.05
34	-21.76	1.17
36	-13.82	0.59
38	-17.58	0.75

Table 3		
XI activity at pH 7.5		
SEQ ID NO:	Volumetric Activity	FIOPC
40	-12.34	0.52
42	-74.88	3.18
44	-37.10	1.57
46	-35.57	1.51
48	-24.69	1.05
50	-32.23	1.37
52	-26.72	1.13
54	-90.79	3.85
56	-39.89	1.69
58	-74.26	3.15
60	-11.91	0.64
62	-15.43	0.83
64	-12.98	0.70
66	-27.45	1.48
68	-29.43	1.59
70	-4.54	0.24
72	-8.93	0.48
74	-0.20	0.01
76	-0.33	0.02
78	-50.55	2.15
80	-57.13	2.42
82	-58.09	2.47
84	-46.42	1.97
86	-35.95	1.53
88	-2.16	0.09
90	-32.77	1.39
92	-30.82	1.31
94	-8.16	0.35
96	-46.18	1.96
98	-30.05	1.28
100	-8.40	0.45
102	-8.34	0.45
104	-3.80	0.20
106	-4.81	0.26
108	-12.06	0.65
110	-6.10	0.33

Table 3		
XI activity at pH 7.5		
SEQ ID NO:	Volumetric Activity	FIOPC
112	-7.71	0.42
114	-4.17	0.22
116	-7.07	0.38
118	-13.50	0.73
120	-1.15	0.06
122	0.03	0.00
124	-4.41	0.24
126	-0.85	0.05
128	-14.60	0.79
130	-17.26	0.93
132	-0.75	0.04
134	-11.55	0.62
136	-7.20	0.39
138	0.16	-0.01
140	-3.63	0.20
142	-3.63	0.20
144	-1.20	0.06
146	-16.77	0.90
148	-2.00	0.11
150	-1.40	0.08
152	-3.63	0.20
154	-7.09	0.38
156	-0.96	0.05
158	-2.79	0.15
160	-3.23	0.17
162	-10.17	0.55
164	-0.51	0.03
166	-3.43	0.19
168	-5.65	0.30
170	-2.35	0.13
172	-1.20	0.06
174	-2.29	0.12
176	-1.92	0.08
Op-XI (ABD)	-23.56	NA
Op-XI (SBD)	-18.55	NA
Vo – ctrl	-1.74	NA

5.6 Example 5: Growth of Yeast Containing XI Clones on Xylose

[0100] A subset of the XI genes from Example 3 were expressed in *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108) and assayed for ability to confer the ability to grow on xylose. This assay was carried out as follows: colonies were isolated on SC-ura + 2% glucose agar plates and inoculated into about 3 ml “pre-cultures” of both SC –ura 2% glycerol and SC –ura 2% xylose media, incubated at about 30°C, about 220 rpm, overnight. Cells were harvested by centrifugation (about 100 x g, 5 minutes), supernatant discarded and washed twice and resuspended in about 1 ml of SC –ura 2% xylose. Cells were inoculated into Biolector plates, containing SC –ura, 2% xylose, and inoculums were normalized to two different starting optical densities of about OD₆₀₀ 0.2 and 0.4. Plates were covered using gas permeable seals and incubated in a BioLector microfermentation device (m2p-labs, Model G-BL100) at about 30°C for about 4 days at 800 rpm and 90% humidity. Growth readings from the Biolector were acquired for 60-100 hours according to manufacturer’s recommendations. Results are shown in FIG. 4.

5.7 Example 6: Ethanol Production Under Anaerobic Conditions

[0101] A subset of the XI expressing yeast clones in strain *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108) were assayed for ability to ferment xylose to ethanol (EtOH). In brief, single colonies were inoculated into about 25 ml of SC –ura medium supplemented with about 0.1% glucose and about 3% xylose. Cultures were incubated under microaerobic conditions at about 30°C and about 200 rpm. Samples were harvested at about 0, 24, 48, 72 h, and ethanol concentration determined via HPLC standard assays. Ethanol productivity was calculated, and listed in units of grams of EtOH per liter per hour, and FIOPC was generated comparing productivity of the control Op-XI. Results are shown in Table 4.

Table 4						
Anaerobic EtOH Production						
	Time (h)					
SEQ ID NO:	0	24	48	72	EtOH (g/L/h)	FIOPC
6	0.28	0	0	0	-0.004	-0.5
8	0	0	0	0	0.000	0.0
10	0	0	0	0	0.000	0.0
14	0.37	0.28	0.71	1.24	0.013	1.7
16	0.33	0.275	0.72	1.06	0.011	1.4
18	0.29	0.135	0.31	0.595	0.005	0.6
20	0.33	0	0	0	-0.004	-0.5
22	0.32	0	0	0	-0.004	-0.5
24	0.28	0	0	0	-0.004	-0.5
26	0.26	0	0	0	-0.003	-0.4
28	0.23	0.385	1.015	1.54	0.019	2.5
30	0.27	0	0	0.07	-0.003	-0.3

Table 4						
Anaerobic EtOH Production						
	Time (h)					
SEQ ID NO:	0	24	48	72	EtOH (g/L/h)	FIOPC
32	0	0.165	0.48	0.815	0.012	1.5
34	0	0.125	0.33	0.615	0.009	1.1
36	0	0	0	0	0.000	0.0
46	0	0.285	0.905	1.625	0.023	3.0
60	0.45	0.35	0.87	1.39	0.014	1.8
62	0	0	0	0.065	0.001	0.1
64	0.38	0.275	0.735	1.18	0.012	1.6
66	0	0	0.12	0.22	0.003	0.4
68	0	0.05	0.275	0.5	0.007	0.9
70	0	0	0	0	0.000	0.0
72	0.119	0	0.054	0.1685	0.001	0.1
74	0.21	0.11	0.275	0.57	0.005	0.7
76	0.28	0	0	0	-0.004	-0.5
90	0	0.24	0.69	1.09	0.016	2.0
100	0.104	0.642	0.141	0.366	0.001	0.2
102	0.185	0	0	0.054	-0.002	-0.2
104	0.235	0.536	0	0	-0.005	-0.7
106	0.188	0.4835	0	0	-0.004	-0.6
108	0.19	0.5855	0.1455	0.313	0.000	0.0
110	0.3	0	0	0.05	-0.003	-0.4
112	0.19	0.5535	0.106	0.1135	-0.003	-0.4
114	0.174	0	0	0	-0.002	-0.3
116	0.15	0	0.0515	0.211	0.001	0.1
118	0.177	0.7075	0.5065	0.941	0.009	1.1
120	0.153	0	0	0	-0.002	-0.2
122	0.169	0.553	0	0.074	-0.003	-0.5
124	0.125	0	0	0	-0.002	-0.2
126	0.32	0	0	0	-0.004	-0.5
128	0	0	0	0	0.000	0.0
130	0	0	0	0	0.000	0.0
132	0.121	0	0	0	-0.002	-0.2
134	0.118	0	0	0.1105	0.000	0.0
136	0.108	0	0	0	-0.001	-0.2
138	0.172	0.513	0	0	-0.004	-0.6
140	0.17	0.542	0	0.3135	0.000	-0.1
142	0.102	0	0	0	-0.001	-0.2
144	0.28	0	0	0	-0.004	-0.5
146	0.103	0.635	0.263	0.563	0.004	0.5
150	0.27	0	0	0	-0.003	-0.4
149	0.27	0	0	0	-0.003	-0.4
152	0.17	0	0	0	-0.002	-0.3
154	0.23	0	0	0	-0.003	-0.4
156	0.23	0	0	0	-0.003	-0.4
158	0.4	0	0.105	0.23	-0.002	-0.2

Table 4						
Anaerobic EtOH Production						
	Time (h)					
SEQ ID NO:	0	24	48	72	EtOH (g/L/h)	FIOPC
160	0.38	0	0	0	-0.005	-0.6
162	0.36	0.055	0.23	0.41	0.001	0.2
164	0.32	0	0	0	-0.004	-0.5
166	0.31	0	0	0	-0.004	-0.5
168	0.32	0	0.295	0.6	0.005	0.6
170	0.164	0.4995	0	0	-0.004	-0.5
172	0.27	0	0	0	-0.003	-0.4
174	0.3	0	0.17	0.345	0.001	0.2
OP-XI (pos)	0.2385	0.5875	0.6965	0.81508	0.008	NA
Host-(neg)	0.23625	0.088125	0	0	-0.003	NA

5.8 Example 7: Impact of pH on XI Activity

[0102] Extracts from strain *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108, expressing XI gene candidates in vector p426PGK1, were prepared as described in the Materials and Methods and assayed for XI activity at pH 7.5 and pH 6.0. Percent activity listed was calculated by dividing the VA at pH 6 by the VA at pH 7.5 and multiplying by 100. Results are listed in Table 5.

Table 5				
XI activity at pH 6 and pH 7.5				
SEQ ID NO:	Organism Classification	VA, pH 6 (U/ml)	VA, pH 7.5 (U/ml)	Percent activity (pH 6)
2	<i>Bacteroidales</i>	1.92	2.59	74%
14	<i>Bacteroides</i>	0.32	0.98	32%
16	<i>Bacteroides</i>	1.16	2.40	48%
32	<i>Bacteroides</i>	1.17	2.21	53%
38	<i>Firmicutes</i>	2.46	2.77	89%
42	<i>Firmicutes</i>	1.71	2.18	79%
44	<i>Firmicutes</i>	0.19	0.25	76%
46	<i>Firmicutes</i>	1.49	1.95	76%
50	<i>Firmicutes</i>	0.81	0.95	86%
52	<i>Firmicutes</i>	0.02	0.08	26%
54	<i>Neocallimastigales</i>	1.46	2.90	51%
58	<i>Neocallimastigales</i>	1.89	3.05	62%
68	<i>Neocallimastigales</i>	1.50	1.97	76%
72	<i>Neocallimastigales</i>	0.57	1.04	55%
78	<i>Prevotella</i>	2.40	3.61	67%
80	<i>Prevotella</i>	1.52	2.29	66%
82	<i>Prevotella</i>	1.48	1.65	89%

Table 5				
XI activity at pH 6 and pH 7.5				
SEQ ID NO:	Organism Classification	VA, pH 6 (U/ml)	VA, pH 7.5 (U/ml)	Percent activity (pH 6)
84	<i>Prevotella</i>	1.79	2.96	61%
96	<i>Prevotella</i>	2.13	3.56	60%
116	<i>Prevotella</i>	0.06	0.13	47%
Host-neg		0.04	0.02	NA
Op-XI		0.61	1.25	49%

5.9 Example 8: K_m for Selected XI Clones

[0103] The K_m and V_{max} at pH 6 were determined for a subset of the XI clones, expressed on p426PGK1 vector in *Saccharomyces cerevisiae* CEN.PK2-1Ca (ATCC: MYA1108), using the XI activity assay described in the Materials and Methods and varying the concentrations of xylose from about 40-600mM. Results shown are calculated using the Hanes Plot, which rearranges the Michaelis-Menten equation ($v = V_{max}[S] / (K_m + [S])$) as: $([S] / v = K_m / V_{max} + [S] / V_{max})$, where plotting $[S] / v$ against $[S]$, resulting in a straight line and where the y intercept = K_m / V_{max} , the slope = $1 / V_{max}$, and the x intercept = $-K_m$. Results are listed in Table 6.

Table 6		
K_m determination for 3 XIs		
SEQ ID NO:	K_m	V_{max}
78	35.2	27.6
96	33.7	28.0
38	28.8	28.6

5.10 Example 9: Quantification of XI Activity Expressed From Single Genomic Integration Locus

[0104] A vector named pYDAB006 (FIG. 5A) for integration into locus YER131.5 (between YER131W and YER132C) in the *S. cerevisiae* genome was constructed using conventional cloning methods. The vector backbone with a PacI site at each end was derived from pBluescript II SK (+) (Agilent Technologies, Inc. Santa Clara, California) by standard PCR techniques, which contained only the pUC origin of replication and *bla* gene encoding ampicillin resistance protein as a selectable marker. Two 300-base pair segments named YER131.5-A and YER 131.5-B were amplified from yeast genomic DNA by standard PCR techniques and connected with a multiple cloning site (MCS 1: 5'-GGCGCGCCTCTAGAAAGCTTACGCGTGAGCTCCCTGCAGGGATATCGGTACCGCGGCCGC-3' (SEQ ID NO:181)) using the overlapping PCR technique. The PCR primers used in the overlapping PCR are shown in Table 7 below:

Table 7		
Primers Used in pYDAB006 Construction		
Primer	SEQ ID NO:	Sequence (PacI site is underlined)
131.5AF	182	cacc <u>attaattaa</u> AGCTTTGTAAATATGATGAGAGAATAATATAAATCAAACG
131.5AR	183	GGCGCGCCTCTAGAAAGCTTAATCGACAAGAACAACCTTCTATTATATAGGTATGAAA
131.5BF	184	GCAGGGATATCGGTACCCACCAGCGGCCGCTGAAGAAGGTTTATTTTCGTTTCGCTGT
131.5BR	185	cacc <u>attaattaa</u> CCCAGGTGAGACTGGATGCTCCATA
ABMCSF	186	GCCTCTAGAAAGCTTACGCGTGAGCTCCCTGCAGGGGATATCGGTACCCACCAGCGGCCGC
ABMCSR	187	CGCTGGTGGGTACCGATATCCCTGCAGGGAGCTCACGCGTAAGCTTTCTAGAGGCGCGCC

[0105] The overlapping PCR product was then ligated with the vector backbone resulting in plasmid pYDAB006.

[0106] A vector named pYDURA01 (FIG. 5B) for generating yeast selectable and recyclable marker was constructed using similar method as pYDAB006. The URA3 expression cassette was amplified from yeast genomic DNA by standard PCR techniques. The 200 base pair fingerprint sequence (named R88: TGCGTGTGCCGCGAGTCCACGTCTACTCGCGAACCGAGTGCAGGCGGGTCTTCGGCCAGGACGGCCGTGCGTGACCCGGCCGCCAGACGAAACGGACCGCGCTCGCCAGACGCTACCCAGCCCGTTCATGCCGGCCGCGAGCCGACCTGTCTCGGTCGCTTCGACGCACGCGCGGTCCCTTTCGGGTACTCGCCTAAGAC (SEQ ID NO:188)) at both sides of URA3 cassette was amplified by standard PCR techniques from the genomic DNA of yBPA317, which was a diploid strain having genotypes MATa / MATalpha; URA3/ura3; YDL074.5::P(TDH3)-CBT1-T(CYC1)-R88 YLR388.5::P(TDH3)-StBGL-T(CYC1)-R88/YLR388.5::P(TDH3)-StBGL-T(CYC1)-R88. The primers used in the amplification are described in Table 8 below:

Table 8		
Primers Used in pYDURA01 Construction		
Primer	SEQ ID NO:	Sequence (KpnI and NotI sites are underlined)
NotI-KpnI-R88-F	189	caatagcggccgcggtaccTGCGTGTGCCGCGAGTCCAC
R88-BamHI-R	190	TGTTAGGATCCGTCTTAGGCGAGTACCCGAAAGG
BamHI-ura-F	191	caataggatccAGGCATATTTATGGTGAAGAATAAGT
ura-Xho-R	192	TGTTACTCGAGAAATCATTACGACCGAGATTCCCG
XhoI-R88-F	193	caatactcgagTGCGTGTGCCGCGAGTCCAC
R88-NotI-R	194	TGTTAGCGGCCGCGTCTTAGGCGAGTACCCGAAAGG

[0107] An expression cassette was generated for the XI genes by cloning into a vector named pYDPt005 (FIG. 5C). pYDPt005 was generated using similar method as pYDAB006. It contained a TDH3 promoter and a PGK1 terminator flanking a multiple cloning site (MCS 2: 5'-ACTAGTGGATCCCTCGAGGTCGACGTTTAAAC-3' (SEQ ID NO:195), where single underline is *SpeI* site, double underline is *XhoI* site, and jagged underline is *PmeI* site). The promoter and the terminator were amplified from *S. cerevisiae* genomic DNA; an *AscI* site was added to the 5' end of the TDH3 promoter while a *KpnI* site was added to the 3' end of the PGK1 terminator during amplification. Primers used in the amplification are described in Table 9.

Table 9		
Primers Used in pYDPt005 Construction		
Primer	SEQ ID NO:	Sequence (<i>AscI</i> and <i>KpnI</i> sites are underlined)
TDH-F	196	CACCAGGCGCGCCTCTAGAAAGCTTACGCGTAGTTTATC ATTATCAATACTGCCATTTCAAAGA
overlap-TDH-R	197	AACGTCGACCTCGAGGGATCCACTAGTTCGAAACTAAGT TCTTGGTGTTTAAAACT
overlap-PGK-F	198	GTGGATCCCTCGAGGTCGACGTTTAAACATTGAATTGAA TTGAAATCGATAGATCAAT
PGK-R	199	CACCAGCGGCCGCGGTACCGATATCCCTGCAGGGAGCTC GAAATATCGAATGGGAAAAAAAAAACTGGAT

[0108] An *Orpinomyces* sp. XI gene (NCBI:169733248) was cloned in this vector between the *SpeI* and *XhoI* sites. The *Orpinomyces* sp. XI expression cassette and R88-Ura-R88 fragment were then cloned into vector pYDAB006 using *AscI*, *KpnI* and *NotI* sites; the resulting plasmid was named pYDABF006 (FIG. 5D). Subsequently, the *Orpinomyces* sp. XI gene in pYDABF006 was replaced with a subset of the XI genes of Table 2 by digestion of pYDABF006 with *SpeI* and *PmeI* and ligation to a DNA fragment encoding the appropriate XI sequence which had been amplified from p426PGK1-XI constructs. A *SpeI* site followed by a Kozak-like sequence (6 consecutive adenines) was added immediately in front of the start codon of the XI genes while a *PmeI* site was added to the 3' end of the XI genes during amplification.

[0109] XI gene integration cassettes were extracted by *PacI* digestion and used to transform yeast strain yBPA130 using standard techniques. Transformants were selected for growth on SC-Ura (Synthetic Complete, Ura dropout) agar plates. Integration position and existence of XI cassette in transformants was confirmed by PCR using the primers shown in Table 10.

Table 10		
Primers Used in Integration Verification		
Primer	SEQ ID NO:	Sequence
5' of integration	200	ACAGGGATAACAAAGTTTCTCCAGC
3' of integration	201	CATACCAAGTCATGCGTTACCAGAG
5' of R88-ura-R88	202	TTTCCCATTCGATATTTTCGAGCTCC
3' of integration	203	CATACCAAGTCATGCGTTACCAGAG

[0110] Confirmed clones were then grown about 18 hours in liquid YPD to allow looping out of the URA3 marker and were selected for growth on SC + 5-FOA agar plate. The absence of the URA3 marker was confirmed by PCR.

[0111] Strains containing the confirmed XI expression cassettes were inoculated into about 3ml of modified YP Media (YP+0.1% Glucose + 3.0% Xylose) and incubated overnight at about 30°C and about 220 rpm. These overnight cultures were subcultured into about 25 ml of the same media to about OD₆₀₀ = 0.2. Samples were incubated overnight at about 30°C and about 220 rpm. Cultures were harvested when OD₆₀₀ was between about 3 and 4. Pellets were collected by centrifugation for about 5 minutes at about 4000 rpm. The supernatant was discarded and pellets washed with about 25 ml of distilled-deionized water and centrifuged again using the same conditions. Supernatant was discarded and the pellet frozen at about -20 °C until lysis and characterization.

[0112] Cell pellets were thawed and about 200 mg of each pellet sample was weighed out into 2 ml microcentrifuge tubes. About 50 µl of Complete®, EDTA-free Protease Inhibitor cocktail (Roche Part# 11873 580 001) at 5 times the concentration stated in the manufacturer's protocol was added to each sample. To this was added about 0.5 ml of Y-PER Plus® Dialyzable Yeast Protein Extraction Reagent (Thermo Scientific Part# 78999) (YP+) to each sample. Samples were incubated at about 25°C for about 4 hours on rotating mixer. Sample supernatants were collected after centrifugation at about 10,000 x g for about 10 minutes for characterization.

[0113] Total protein concentrations of the XI sample extracts prepared above were carried out using Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad, cat# 500-0006, Hercules CA) which is a modified version of the Bradford method (Bradford).

[0114] Yeast physiological pH ranges are known to range from about pH 6 to about pH 7.5 (Pena, Ramirez *et al.*, 1995, J. Bacteriology 4:1017-1022). Ranking of XI activity at yeast physiological pH was accomplished using the assay conditions at pH 7.5 and modified for pH 6.0 as described in the materials and methods. The specific activities of 20 XIs when expressed from a single copy integrated into the yeast YER131.5 locus were evaluated. The results are listed in Table 11.

Table 11			
SA of XI Expressed in an Industrial <i>S. cerevisiae</i>			
SEQ ID NO:	Organism Classification	SA, pH6 (U/mg)	SA, pH 7.5 (U/mg)
2	<i>Bacteroidales</i>	0.86	1.08
14	<i>Bacteroides</i>	0.33	1.07
16	<i>Bacteroides</i>	0.57	1.05
32	<i>Bacteroides</i>	0.53	1.00
38	<i>Firmicutes</i>	1.00	0.94
42	<i>Firmicutes</i>	0.79	0.82
44	<i>Firmicutes</i>	0.08	0.10
46	<i>Firmicutes</i>	0.62	0.69
50	<i>Firmicutes</i>	0.35	0.41
52	<i>Firmicutes</i>	0.01	0.03
54	<i>Neocallimastigales</i>	0.64	1.17
58	<i>Neocallimastigales</i>	0.79	1.10
68	<i>Neocallimastigales</i>	0.01	0.02
72	<i>Neocallimastigales</i>	0.22	0.40
78	<i>Prevotella</i>	1.10	1.45
80	<i>Prevotella</i>	0.74	1.11
82	<i>Prevotella</i>	0.54	0.60
84	<i>Prevotella</i>	0.76	1.06
96	<i>Prevotella</i>	1.10	1.62
116	<i>Prevotella</i>	0.03	0.06
Host neg ctrl		0.00	0.02

5.11 **Example 10: Identification of Sequence Motifs In Acid Tolerant XIs**

[0115] The proposed mechanism of xylose isomerases can be summarized as follows: (i) binding of xylose to xylose isomerase, so that O3 and O4 are coordinated by metal ion I; (ii) enzyme-catalyzed ring opening (the identity of the ring-opening group remains a subject for further investigation; ring opening may be the rate limiting step in the overall isomerization process); (iii) chain extension (sugar binds in a linear extended form) in which O2 and O4 now coordinate metal ion I; (iv) O2 becomes deprotonated causing a shift of metal ion II from position 1 to an adjacent position 2 in which it coordinates O1 and O2 of the sugar together with metal ion I; (v) isomerization via an anionic transition state arises by a hydride shift promoted by electrophilic catalysis provided by both metal ions; (vi) collapse of transition state by return of metal ion II to position 1; (vii) chain contraction to a pseudo-cyclic position with ligands to metal ion I changing from O2/O4 back to O3/O4; (viii) enzyme-catalyzed ring closure; (ix) dissociation of xylulose from xylose isomerase (Lavie *et al.*, 1994, Biochemistry 33(18), 5469-5480).

[0116] Many XIs identified contained one or both of two signature sequences characteristic of XIs, [LI]EPKP.{2}P (SEQ ID NO:204) and [FL]HD[^K]D[LIV].[PD].[GDE] (SEQ ID NO:205). Additional sequence motifs present in the top performing *Firmicutes* and *Prevotella* XIs were identified. The motifs are located near the active site including residues in direct contact with the D-xylose and/or the metal ions. The motifs are shown in Table 12 below:

Table 12			
XI Sequence Motifs			
XI Source	Motif	Sequence	SEQ ID NO:
<i>Firmicutes</i>	1A	P[FY][AST][MLVI][AS][WYFL]W[HT]N[LFMG]GA	206
<i>Firmicutes</i>	1B	P[FY][AS].{2}[WYFL]W[HT][^TV].GA	207
<i>Firmicutes</i>	2	[GSN][IVA]R[YFHG][FYLV]C[FW]HD.D	208
<i>Firmicutes</i>	3	T[ASTC][NK][^L]F.[NDH][PRKAG][RVA][FY]C	209
<i>Firmicutes</i>	4	[WFY]D[TQVI]D.[FY][PF][^T].{2,4}[YFH]S[ATL]T	210
<i>Firmicutes</i>	5	GF[NH]FD[SA]KTR	211
<i>Prevotella</i>	1A	FG.QT[RK].{2}E[WYF][DNG].{2,3}[DNEG][AT]	212
<i>Prevotella</i>	1B	FG.QT[RK].{2}E[WYF][DNG].{3}[^C][^P]	213
<i>Prevotella</i>	2	[FW]HD.D[LVI].[DE]EG[^P][TSD][IV][EA]E	214

5.12 Example 11: In Vivo Evaluation of Xylose Isomerase

[0117] Haploid *S. cerevisiae* strain yBPA130 (MATa::ura3) and yBPA136 (MATalpha::ura3) were genetically modified to enhance C5 xylose utilization during fermentation. The modification includes the following: the native glucose repressible alcohol dehydrogenase II gene ADH2 was disrupted by inserting an expression cassette of the endogenous transaldolase gene TAL1 (SEQ ID NO:215) and xylulokinase gene XKS1 (SEQ ID NO:216). PHO13 encoding the native alkaline phosphatase specific for p-nitrophenyl phosphate gene was disrupted by inserting the native transketolase-1 gene TKL1 (SEQ ID NO:217). Native aldose reductase gene GRE3 was disrupted by inserting native D-ribulose-5-phosphate 3-epimerase gene RPE1 (SEQ ID NO:218) and Ribose-5-phosphate ketol-isomerase gene RKI1 (SEQ ID NO:219). Also one expression cassette of native galactose permease gene GAL2 (SEQ ID NO:220) was integrated into the *S. cerevisiae* strain, resulting in haploid strains pBPB007 (MATa::ura3) and pBPB008 (MATalpha::ura3). The genotype of pBPB007 and pBPB008 is adh2::TAL1-XKS1, pho13::TKL1-XKS1, gre3::RPE1-RKI1 and YLR388.5::GAL2. The sequences are shown in Table 13, below:

Table 13			
Sequence Name	Type of sequence	SEQ ID NO:	Sequence
<i>TAL1 (S. cerevisiae)</i>	DNA	215	ATGTCCTGAACCAGCTCAAAAGAAACAAAAGGTTGCTAACAACCTCT CTAGAACAATTGAAAGCCTCCGGCACTGTCTGTTGTTGCCGACACT GGTGATTTTCGGCTCTATTGCCAAGTTTCAACCTCAAGACTCCACA ACTAACCCATCATTGATCTTGGCTGCTGCCAAGCAACCAACTTAC GCCAAGTTGATCGATGTTGCCGTGGAATACGGTAAGAAGCATGGT AAGACCACCGAAGAACAAGTCGAAAATGCTGTGGACAGATTGTTA GTCGAATTCGGTAAGGAGATCTTAAAGATTGTTCCAGGCAGAGTC TCCACCGAAGTTGATGCTAGATTGTCTTTTGACACTCAAGCTACC ATTGAAAAGGCTAGACATATCATTAAATTGTTTGAACAAGAAGGT GTCTCCAAGGAAAGAGTCCTTATTAAAATTGCTTCCACTTGGGAA GGTATTCAAGCTGCCAAAGAATTGGAAGAAAAGGACGGTATCCAC TGTAATTTGACTCTATTATTCTCCTTCGTTCAAGCAGTTGCCTGT GCCGAGGCCCAAGTTACTTTGATTTCCTTCCCATTTGTTGGTAGAATT CTAGACTGGTACAAATCCAGCACTGGTAAAGATTACAAGGGTGAA GCCGACCCAGGTGTTATTTCCGTCAAGAAAATCTACAACCTACTAC AAGAAGTACGGTTACAAGACTATTGTTATGGGTGCTTCTTTTCA AGCACTGACGAAATCAAAAACCTGGCTGGTGTGACTATCTAACA ATTTCTCCAGCTTTATTGGACAAGTTGATGAACAGTACTGAACCT TTCCCAAGAGTTTGGACCCTGTCTCCGCTAAGAAGGAAGCCGGC GACAAGATTTCTTACATCAGCGACGAATCTAAATTGAGATTTCGAC TTGAATGAAGACGCTATGGCCACTGAAAAATTGTCCGAAGGTATC AGAAAATTCTCTGCCGATATTGTTACTCTATTCGACTTGATTGAA AAGAAAGTTACCGCTTAA
<i>XKSI (S. cerevisiae)</i>	DNA	216	ATGTTGTGTTTCAGTAATTCAGAGACAGACAAGAGAGGTTTCCAAC ACAATGTCTTTAGACTCATACTATCTTGGGTTTGATCTTTTCGACC CAACAACCTGAAATGTCTCGCCATTAACCAGGACCTAAAAATTGTC CATTCAGAAACAGTGGAATTTGAAAAGGATCTTCCGCATTATCAC ACAAAGAAGGGTGTCTATATACACGGCGACACTATCGAATGTCCC GTAGCCATGTGGTTAGAGGCTCTAGATCTGGTTCTCTCGAAATAT CGCGAGGCTAAATTTCCATTGAACAAAGTTATGGCCGTCTCAGGG TCCTGCCAGCAGCACGGGTCTGTCTACTGGTCTTCCCAAGCCGAA TCTCTGTTAGAGCAATTGAATAAGAAACCGGAAAAAGATTTATTG CACTACGTGAGCTCTGTAGCATTGCAAGGCAAACCGCCCCCAAT TGGCAAGACCAGTACTGCAAAGCAATGTCAAGAGTTTGAAGAG TGCATAGGTGGGCTGAAAAAATGGCTCAATTAACAGGGTCCAGA GCCCATTTTAGATTTACTGGTCCTCAAATTCTGAAAATTGCACAA TTAGAACCAGAAGCTTACGAAAAAACAAGACCATTCTTTTAGTG TCTAATTTTTTGAAGTTCTATCTTAGTGGGCCATCTTGTGAAATTA GAGGAGGCAGATGCCTGTGGTATGAACCTTTATGATATACGTGAA AGAAAATTCAGTGATGAGCTACTACATCTAATTGATAGTTCTTCT AAGGATAAAACTATCAGACAAAAATTAATGAGAGCACCCATGAAA AATTTGATAGCGGTACCATCTGTAAATATTTTATTGAGAAGTAC GGTTCATACAAACTGCAAGGTCTCTCCCATGACTGGGGATAAT TTAGCCACTATATGTTCTTTACCCCTGCGGAAGAATGACGTTCTC GTTTCCTTAGGAACAAGTACTACAGTTCTTCTGGTCACCGATAAG TATCACCCCTCTCCGAACATCATCTTTTC ATTCATCCAACCTCTGCCAAACATTATATGGGTATGATTTGTTAT TGTAATGGTTCTTTGGCAAGGGAGAGGATAAGAGACGAGTTAAAC AAAGAACGGGAAAAATAATTATGAGAAGACTAACGATTGGACTCTT TTTAATCAAGCTGTGCTAGATGACTCAGAAAGTAGTGAAAAATGAA

Table 13			
Sequence Name	Type of sequence	SEQ ID NO:	Sequence
			TTAGGTGTATATTTTCTCTGGGGGAGATCGTTCCTAGCGTAAAA GCCATAAAACAAAAGGGTTATCTTCAATCCAAAACGGGTATGATT GAAAGAGAGGTGGCCAAGTTCAAAGACAAGAGGCACGATGCCAAA AATATTGTAGAATCACAGGCTTTAAGTTGCAGGGTAAGAATATCT CCCCTGCTTTTCGGATTCAAACGCAAGCTCACAACAGAGACTGAAC GAAGATACAATCGTGAAGTTTGATTACGATGAATCTCCGCTGCGG GACTACCTAAATAAAAAGGCCAGAAAGGACTTTTTTTGTAGGTGGG GCTTCTAAAAACGATGCTATTGTGAAGAAGTTTGCTCAAGTCATT GGTGCTACAAAGGGTAATTTTAGGCTAGAAACACCAAACTCATGT GCCCTTGGTGGTTGTTATAAGGCCATGTGGTCATTGTTATATGAC TCTAATAAAATTGCAGTTCCTTTTGATAAATTTCTGAATGACAAT TTTCCATGGCATGTAATGGAAAGCATATCCGATGTGGATAATGAA AATTGGGATCGCTATAATTCCAAGATTGTCCCTTAAGCGAACTG GAAAAGACTCTCATCTAA
<i>TKL1 (S. cerevisiae)</i>	DNA	217	ATGACTCAATTCACTGACATTGATAAGCTAGCCGTCTCCACCATA AGAATTTTGGCTGTGGACACCGTATCCAAGGCCAATCAGGTCAC CCAGGTGCTCCATIGGGTATGGCACCAGCTGCACACGTTCTATGG AGTCAAATGCGCATGAACCCAACCAACCCAGACTGGATCAACAGA GATAGATTTGTCTTGTCTAACGGTCACGCGGTGCTTTGTTGTAT TCTATGCTACATTTGACTGGTTACGATCTGTCTATTGAAGACTTG AAACAGTTCAGACAGTTGGGTTCAGAACACCAGGTCATCCTGAA TTTGAGTTGCCAGGTGTTGAAGTTACTACCGGTCCATTAGGTCAA GGTATCTCCAACGCTGTTGGTATGGCCATGGCTCAAGCTAACCTG GCTGCCACTTACAACAAGCCGGGCTTTACCTTGTCTGACAACTAC ACCTATGTTTTCTTGGGTGACGGTTGTTTGCAAGAAGGTATTTCT TCAGAAGCTTCCTCCTTGGCTGGTCATTTGAAATTGGGTAACCTG ATTGCCATCTACGATGACAACAAGATCACTATCGATGGTGTCTACC AGTATCTCATTCGATGAAGATGTTGCTAAGAGATACGAAGCCTAC GGTTGGGAAGTTTGTACGTAGAAAATGGTAACGAAGATCTAGCC GGTATTGCCAAGGCTATTGCTCAAGCTAAGTTATCCAAGGACAAA CCAACCTTGATCAAAATGACCACAACCATTTGGTTACGGTTCCCTTG CATGCCGGCTCTCACTCTGTGCACGGTGCCCCATTGAAAGCAGAT GATGTTAAACAACATAAAGAGCAAATTCGGTTTCAACCCAGACAAG TCCTTTGTTGTTCCACAAGAAGTTTACGACCCTACCAAAAGACA ATTTTAAAGCCAGGTGTGGAAGCCAACAACAAGTGGAACAAGTTG TTCAGCGAATACCAAAAGAAATTTCCAGAATTAGGTGCTGAATTG GCTAGAAGATTGAGCGGCCAACTACCCGCA AATTGGGAATCTAAGTTGCCAACTTACACCGCCAAGGACTCTGCC GTGGCCACTAGAAAATTATCAGAAACTGTTCTTGAGGATGTTTAC AATCAATTGCCAGAGTTGATTGGTGGTTCTGCCGATTTAACACCT TCTAACTTGACCAGATGGAAGGAAGCCCTTGACTTCCAACCTCCT TCTTCCGGTTCAGGTAACCTACTCTGGTAGATACATTAGGTACGGT ATTAGAGAACACGCTATGGGTGCCATAATGAACGGTATTTACAGCT TTCGGTGCCAACCTACAAACCATAACGGTGGTACTTTCTTGAACCTC GTTTCTTATGCTGCTGGTGCCGTTAGATTGTCCGCTTTGTCTGGC CAGGAGTTATTTGGGTTGCTACACATGACTCTATCGGTGTGCGT GAAGATGGTCCAACACATCAACCTATTGAACTTTAGCACACTTC AGATCCCTACCAACATTCAAGTTTGGAGACCAGCTGATGGTAAC GAAGTTTCTGCCGCCTACAAGAACTCTTTAGAATCCAAGCATACT CCAAGTATCATTGCTTTGTCCAGACAAAACCTTGCCACAATTGGAA GGTAGCTCTATTGAAAGCGCTTCTAAGGGTGGTTACGTACTACAA

Table 13			
Sequence Name	Type of sequence	SEQ ID NO:	Sequence
			GATGTTGCTAACCCAGATATTATTTTAGTGGCTACTGGTTCCGAA GTGTCTTTGAGTGTGAAGCTGCTAAGACTTTGGCCGCAAAGAAC ATCAAGGCTCGTGTTGTTTCTCTACCAGATTCTTCACTTTTGAC AAACAACCCCTAGAATACAGACTATCAGTCTTACCAGACAACGTT CCAATCATGTCTGTGAAGTTTGGCTACCACATGTTGGGGCAAA TACGCTCATCAATCCTTCGGTATTGACAGATTGTTGGTGCCTCCGGT AAGGCACCAGAAGTCTTCAAGTTCTTCGGTTTCACCCAGAAAGGT GTTGCTGAAAGAGCTCAAAAGACCATTGCATTCTATAAGGGTGAC AAGCTAATTTCTCCTTTGAAAAAGCTTTCTAA
<i>RPE1 (S. cerevisiae)</i>	DNA	218	ATGGTCAAACCAATTATAGCTCCCAGTATCCTTGCTTCTGACTTC GCCAACTTGGGTTGCGAATGTCATAAGGTCATCAACGCCGGCGCA GATTGGTTACATATCGATGTCATGGACGGCCATTTTGTTCCAAAC ATTACTCTGGGCCAACCAATTGTTACCTCCCTACGTCGTTCTGTG CCACGCCCTGGCGATGCTAGCAACACAGAAAAGAAGCCCACTGCG TTCTTCGATTGTCACATGATGGTTGAAAATCCTGAAAAATGGGTC GACGATTTTGCTAAATGTGGTGCTGACCAATTTACGTTCCACTAC GAGGCCACACAAGACCCTTTGCATTTAGTTAAGTTGATTAAGTCT AAGGGCATCAAAGCTGCATGCGCCATCAAACCTGGTACTTCTGTT GACGTTTTATTTGAACTAGCTCCTCATTGGGATAGGCTCTTGTT ATGACTGTGGAACCTGGGTTTGGAGGCCAAAAATTCATGGAAGAC ATGATGCCAAAAGTGGAACTTTGAGAGCCAAGTTCCCCCATTG AATATCCAAGTCGATGGTGGTTTGGGCAAGGAGACCATCCCGAAA GCCGCCAAAGCCGGTGCCAACGTTATTGTCGCTGGTACCAGTGTT TTCACTGCAGCTGACCCGCACGATGTTATCTCCTTCATGAAAGAA GAAGTCTCGAAGGAATTGCGTTCTAGAGATTGCTAGATTAG
<i>RK11 (S. cerevisiae)</i>	DNA	219	ATGGCTGCCGGTGTCCTCAAAAAATTGATGCGTTAGAATCTTTGGGC AATCCTTTGGAGGATGCCAAGAGAGCTGCAGCATACAGAGCAGTT GATGAAAATTTAAAATTTGATGATCAGAAAATTTATGGAATTGGT AGTGGTAGCACAGTGGTTTATGTTGCCGAAAGAATTGGACAATAT TTGCATGACCCTAAATTTTATGAAGTAGCGTCTAAATTCATTTGC ATTCCAACAGGATTTCAATCAAGAACTTGATTTGGATAACAAG TTGCAATTAGGCTCCATTGAACAGTATCCTCGCATTTGATATAGCG TTTGACGGTGCTGATGAAGTGGATGAGAATTTACAATTAATTAAA GGTGGTGGTGCTTGTCTATTTCAAGAAAAATTGGTTAGTACTAGT GCTAAAACCTTCATTGTCGTTGCTGATTCAAGAAAAAGTCACCA AAACATTTAGGTAAGAAGTGGAGGCAAGGTGTTCCCATTTGAAAT GTACCTTCCTCATACGTGAGGGTCAAGAATGATCTATTAGAACAA TTGCATGCTGAAAAAGTTGACATCAGACAAGGAGGTTCTGCTAAA GCAGGTCCGTGTTGTAAGTGAATAATAAATTCATTATCGATGCG GATTTCCGTGAAATTTCCGATCCAAGAAAATTGCATAGAGAAATC AACTGTTAGTGGGCGTGGTGGAAACAGGTTTATTCATCGACAAC GCTTCAAAAGCCTACTTCGGTAATTCTGACGGTAGTGTTGAAGTT ACCGAAAAGTGA
<i>GAL2 (S. cerevisiae)</i>	DNA	220	ATGGCAGTTGAGGAGAACAATATGCCTGTTGTTTCACAGCAACCC CAAGCTGGTGAAGACGTGATCTCTTCACTCAGTAAAGATTCCCAT TTAAGCGCACAAATCTCAAAAGTATTCTAATGATGAATTGAAAGCC GGTGAGTCAGGGTCTGAAGGCTCCCAAGTGTTCCTATAGAGATA CCCAAGAAGCCCATGTCTGAATATGTTACCGTTTCCTTGCTTTGT TTGTGTGTTGCCTTCGGCGGCTTCATGTTTGGCTGGGATACCGGT ACTATTTCTGGGTTTGTGTCCAAACAGACTTTTTGAGAAGGTTT GGTATGAAACATAAGGATGGTACCCACTATTTGTCAAACGTCAGA

Table 13			
Sequence Name	Type of sequence	SEQ ID NO:	Sequence
			ACAGGTTTAAATCGTCGCCATTTTCAATATTGGCTGTGCCTTTGGT GGTATTATACTTTCCAAAGGTGGAGATATGTATGGCCGTAAAAAG GGTCTTTTCGATTGTCGTCTCGGTTTATATAGTTGGTATTATCATT CAAATTGCCTCTATCAACAAGTGGTACCAATATTTTCATTGGTAGA ATCATATCTGGTTTGGGTGTCGGCGGCATCGCCGTCTTATGTCCT ATGTTGATCTCTGAAATTGCTCCAAAGCACTTGAGAGGCACACTA GTTTCTTGTTATCAGCTGATGATTACTGCAGGTATCTTTTTGGGC TACTGTACTAATTACGGTACAAAGAGCTATTGCAACTCAGTTCAA TGGAGAGTTCCATTAGGGCTATGTTTCGCTTGGTCATTATTTATG ATTGGCGCTTTGACGTTAGTTCCTGAATCCCCACGTTATTTATGT GAGGTGAATAAGGTAGAAGACGCCAAGCGTTCATTGCTAAGTCT AACAAGGTGTCACCAGAGGATCCTGCCGTCCAGGCAGAGTTAGAT CTGATCATGGCCGGTATAGAAGCTGAAAAACTGGCTGGCAATGCG TCCTGGGGGAATTATTTTCCACCAAGACCAAAGTATTTCAACGT TTGTTGATGGGTGTGTTTGTTCAAATGTTT CAACAATTAAACCGGTAACAATTATTTTTTCTACTACGGTACCGTT ATTTTCAAGTCAGTTGGCCTGGATGATTCCCTTTGAAACATCCATT GTCATTGGTGTAGTCAACTTTGCCTCCACTTTCTTTAGTTTGTGG ACTGTCGAAAACTTGGGACATCGTAAATGTTTACTTTTGGGCGCT GCCACTATGATGGCTTGTATGGTCATCTACGCCTCTGTTGGTGTT ACTAGATTATATCCTCACGGTAAAAGCCAGCCATCTTCTAAAGGT GCCGGTAACTGTATGATTGTCTTTACCTGTTTTTATATTTTCTGT TATGCCACAACCTGGGCGCCAGTTGCCTGGGTGTCATCACAGCAGAA TCATTCCCACTGAGAGTCAAGTCGAAATGTATGGCGTTGGCCTCT GCTTCCAATTGGGTATGGGGTTCTTGATTGCATTTTTCACCCCA TTCATCACATCTGCCATTAACTTCTACTACGGTTATGTCTTTCATG GGCTGTTTGGTTGCCATGTTTTTTTATGTCTTTTCTTTGTTCCA GAACTAAAGGCCTATCGTTAGAAGAAATCAAGAATTATGGGAA GAAGGTGTTTACCTTGGAAATCTGAAGGCTGGATTCCCTTCATCC AGAAGAGGTAATAATTACGATTTAGAGGATTTACAACATGACGAC AAACCGTGGTACAAGGCCATGCTAGAATAA

[0118] A vector named pYDAB008 rDNA (FIG. 6) for integration xylose isomerase into ribosomal DNA loci in *S. cerevisiae* genome was constructed using conventional cloning methods. This vector can confer high copy number integration of genes and resulting in high-level expression of proteins. The vector was derived from pBluescript II SK (+) (Agilent Technologies, Inc., Santa Clara, California). The pUC origin of replication and *bla* gene encoding ampicillin resistance was amplified with specific primer sequences as a selectable marker for cloning. A 741 base-pair segment R1 region, 253 base-pair R3 region and a 874 base-pair R2 region were amplified from yeast genomic DNA by PCR amplifications. A multiple cloning site of SEQ ID NO:181 (: 5'-GGCGCGCCTCTAGAAAGCTTACGCGTGAGCTCCCTGCAGGGATATCGGTACCGCGGCCGC-3') was inserted between the R1 and R3/R2 regions by assembly using overlapping PCR. All primers

used in above reactions are shown in Table 14. Overlapping PCR products were then ligated in one reaction and result in rDNA integration plasmid named pYDAB008 rDNA (FIG. 6).

Table 14		
Primers Used in pYDAB008 rDNA vector construction		
Primer	SEQ ID NO:	Sequence (<i>Pac I</i> restriction site is underlined)
<i>Pac I</i> -rDNA(R1)-R	221	CACCATTAATTAACCCGGGGCACCTGTCACCTTTGGAA
rDNA (R1)-over-R	222	CGCGTAAGCTTTCTAGAGGCGCGCCAAGCTTTTACACTCTTG ACCAGCGCA
AB vector-MCS-R	223	CCGCTGGTGGGTACCGATATCCCTGCAGGGAGCTCACGCGTA AGCTTTCTAGAGGCG
rDNA(R3)-over-R	224	CTGCAGGGATATCGGTACCCACCAGCGGCCGAGGCCTTGG GTGCTTGCTGGCGAA
rDNA(R3)-over-R	225	ACCTCTGCATGCGAATTCTTAAGACAAATAAAATTTATAGAG ACTTGT
rDNA(R2)-over-R	226	GTCTTAAGAATTCGCATGCAGAGGTAGTTTCAAGGT
<i>Pac I</i> -rDNA(R2)-R	227	CACCATTAATTAATACGTATTTCTCGCCGAGAAAACTT

[0119] pYDABF 0015 (a plasmid comprising a Boles codon optimized nucleic acid of SEQ ID NO:244, encoding a xylose isomerase of SEQ ID NO:78) and pYDABF-0026 (a plasmid comprising a Boles codon optimized nucleic acid of SEQ ID NO:245, encoding a xylose isomerase of SEQ ID NO:SEQ ID NO:96) (both described in Example 10) were digested with Asc I and Kpn I restriction enzymes (New England Biolabs Inc., MA, USA) and the XI-coding insert ligated to pYDAB008 rDNA integration vector described above (FIG. 6). The resulting plasmids were named pYDABF-0033(SEQ ID NO:78) and pYDABF-0036 (SEQ ID NO:96). Additionally, Boles codon optimized nucleic acids encoding xylose isomerase of SEQ ID NO:54 and SEQ ID NO:58 (SEQ ID NO:238 and SEQ ID NO:239, respectively) were ordered from Genewiz (Genewiz Inc., NJ, USA) were digested with Asc I and Kpn I restriction enzymes (New England Biolabs Inc., MA, USA) and ligated to pYDAB008 rDNA integration vector (FIG. 6). The codon-optimized sequences are set forth in Table 15, below:

Table 15		
Sequence Description	SEQ ID NO:	Sequence
Codon optimized DNA encoding XI of SEQ ID NO:54	238	ATGGCTAAGGAATACTTCCCAGAAATTGGTAAGATTAAGTTCGAA GGTAAGGACTCTAAGAACCCAATGGCTTTCCACTACTACGACCCA GAAAAGGTTATTATGGGTAAGCCAATGAAGGACTGGTTGAGATTC GCTATGGCTTGGTGGCACACCTTGTGTGCTGAAGGTGGTGACCAA TTCGGTGGTGGTACTAAGAAGTTCCTATGGAACAACGGTGCTGAC GCTGTTGAAATTGCTAAGCAAAAGGCTGACGCTGGTTTCGAAATT ATGCAAAAGTTGGGTATTCCATACTTCTGTTTCCACGACGTTGAC TTGGTTTCTGAAGGTGCTTCTGTTGAAGAATACGAAGCTAACTTG AAGGCTATTACCGACTACTTGGCTGTTAAGATGAAGGAAACCGGA

Table 15		
Sequence Description	SEQ ID NO:	Sequence
		ATTAAGTTGTTGTGGTCTACCGCTAACGTTTTTCGGTAACGGTAGA TACATGAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTTGCT AGAGCTATTGTTCAAATTAAGAACGCTATTGACGCTGGTATTAAG TTGGGTGCTGAAAACCTACGTTTTCTGGGGTGGTAGAGAAGGTTAC ATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACACATG GCTACCATGTTGACCATGGCTAGAGACTACGCTAGAGCTAAGGGT TTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCATCT AAGCACCATAACGACGTTGACACCGAAAACCGTTATTGGTTTTCTTG AAGGCTCACAACCTTGGACAAGGACTTCAAGGTTAACATTGAAGTT AACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTGGCT GTTGCTGTTGACAACAACATGTTGGGTTCTATTGACGCTAACAGA GGTGACTACCAAAACGGTTGGGACACCGACCAATTCCCAATTGAC CAATACGAATTGGTTCAAGCTTGGATGGAAATTATTAGAGGTGGT GGTGGGTGACAGGTGGTACTAAGCTTCGACGCTAAGACCAGAAGA AACTCTACCGACTTGAAGACATTTTCATTGCTCACATTGCTGGT ATGGACGCTATGGCTAGAGCTTTGGAATCTGCTGCTAAGTTGTTG GAAGAATCTCCATACAAGGCTATGAAGGCTGCTAGATACGCTTCT TTCGACAACGGTATTGGTAAGGACTTCGAAGACGGTAAGTTGACC TTGGAACAAGCTTACGAATACGGTAAGAAGGTTGGTGAACCAAAG CAAACCTCTGGTAAGCAAGAATTGTACGAAGCTATTGTTGCTATG TACGCTTAA
Codon optimized DNA encoding XI of SEQ ID NO:58	239	ATGGCTAAGGAATACTTCCCAGAAATTGGTAAGATTAAAGTTCGAA GGTAAGGACTCTAAGAACCCAATGGCTTTCCACTACTACGACGCT GAAAAGGTTATTATGGGTAAGCCAATGAAGGAATGGTTGAGATTC GCTATGGCTTGGTGGCACACCTTGTGTGCTGAAGGTGGTGACCAA TTCGGTGGTGGTACTAAGAAGTTCCCATGGAACGAAGGTACTGAC GCTGTTACCATTGCTAAGCAAAAGGCTGACGCTGGTTTCGAAATT ATGCAAAAGTTGGGTTTCCCATACTTCTGTTTCCACGACTTACG TTGGTTTCTGAAGGTAAGTCTATTGAAGAATACGAAGCTAAGTTG CAAGCTATTACCGACTACTTGAAGGTAAAGATGGAAGAAACCGGA ATTAAGTTGTTGTGGTCTACCGCTAACGTTTTTCGGTAACGGTAGA TACATGAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTTGCT AGAGCTATTGTTCAAATTAAGAACGCTATTGACGCTGGTATTAAG TTGGGTGCTGAAAACCTACGTTTTCTGGGGTGGTAGAGAAGGTTAC ATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACACATG GCTACCATGTTGACCATGGCTAGAGACTACGCTAGATCTAAGGGT TTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCATCT AAGCACCATAACGACGTTGACACCGAAAACCGTTATTGGTTTTCTTG AAGGCTCACAACCTTGGACAAGGACTTCAAGGTTAACATTGAAGTT AACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTGGCT GTTGCTGTTGACAACGGTATGTTGGGTTCTATTGACGCTAACAGA GGTGACTACCAAAACGGTTGGGACACCGACCAATTCCCAATTGAC CAATACGAATTGGTTCAAGCTTGGATGGAAATTATTAGAGGTGGT GGTGGGTGACTGGTGGTACAACTTCGACGCTAAGACCAGAAGA AACTCTACCGACTTGAAGACATTTTCATTGCTCACATTTCTGGT ATGGACGCTATGGCTAGAGCTTTGGAATCTGCTGCTAAGTTGTTG GAAGAATCTCCATACTGTGCTATGAAGAAGGCTAGATACGCTTCT TTCGACTCTGGTATTGGTAAGGACTTCGAAGACGGTAAGTTGACC TTGGAACAAGCTTACGAATACGGTAAGAAGGTTGGTGAACCAAAG CAAACCTCTGGTAAGCAAGAATTGTACGAAGCTATTGTTGCTATG TACGCTTAA

Table 15		
Sequence Description	SEQ ID NO:	Sequence
Codon optimized DNA encoding XI of SEQ ID NO:78	244	ATGGCTAAGGAATATTTCCCATTCACCGGTAAGATTCCATTCGAA GGTAAGGACTCTAAGAACGTTATGGCTTTCCACTATTATGAACCA GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGGTGAC GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC TTGGTTGAAGAAGGTGACACCATTGAAGAATATGAAGCTAGAATG AAGGCTATTACCGACTATGCTCAAGAAAAGATGAAGCAATTCCCA AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTTCGGTAACAAG AGATAIGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCAT AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT TATATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACAC ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC TTGAAGGCTCACAACCTTGGACAAGGACTTCAAGGTTAACATTGAA GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCTATTGACGCTAAC AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT GACAACTATGAATTGACCCAAGCTATGTTGGAAATTATTAGAAAC GGTGGTTTGGGTAAACGGTGGAACCAACTTCGACGCTAAGATTAGA AGAACTCTACCGACTTGAAGACTTGTTCATTGCTCACATTTCT GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGACATT TTGGAATACTCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT TCTTTCGACCAAGGTGTTGGTAAGGACTTCGAAGACGGTAAGTTG ACCTTGAACAAGTTTATGAATATGGTAAGAAGGTTGGTGAACCA AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT TTGTATGCTAAGTAA
Codon optimized DNA encoding XI of SEQ ID NO:96	245	ATGGCTAAGGAATATTTCCCATTCATTGGTAAGGTTCCATTCGAA GGTACTGAATCTAAGAACGTTATGGCTTTCCACTATTATGAACCA GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGCTGAC GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC TTGGTTGAAGAAGGTGAAACCGTTGCTGAATATGAAGCTAGAATG AAGGTTATTACCGACTATGCTTTGGAAGAAGATGCAACAATTCCCA AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTTCGGTCACAAG AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCAT AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT TATATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACAC ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC TTGAGGGCTCACGGTTTGGACAAGGACTTCAAGGTTAACATTGAA GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCTATTGACGCTAAC AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT

Table 15		
Sequence Description	SEQ ID NO:	Sequence
		GACAACTATGAATTGACCCAAGCTATGATGGAAATTATTAGAAAC GGTGGTTTGGGTAACGGTGGAACCAACTTCGACGCTAAGATTAGA AGAACTCTACCGACTTGAAGACTTGTTTCATTGCTCACATTTCT GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGCTATT TTGGAAGAATCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT TCTTTCGACGAAGGTATTGGTAAGGACTTCGAAGACGGTAAGTTG TCTTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGAAGAACCA AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT TTGTATGCTAAGTAA

[0120] The resulting plasmids were named pYDABF-0033 (SEQ ID NO:78) and pYDABF-0036 (SEQ ID NO:96), pYDABF-0231(SEQ ID NO:54) and pYDABF-0232 (SEQ ID NO:58).

[0121] The rDNA integration cassette was linearized by Pac I restriction enzyme digestion (New England Biolabs Inc., MA, USA) and purified with DNA column purification kit (Zymo Research, Irvine, CA, USA). The integration cassette was transformed into modified haploid *S. cerevisiae* strain pBPB007 (MATa::ura3) and pBPB008 (MAT alpha::ura3) using the standard protocol described in previous examples. Transformants were plated on SC-xylose (SC complete +2% xylose) agar plates, about 2-3 days at about 30°C. Colonies that grew on SC-xylose agar plates were then checked by colony PCR analysis with primer sets shown in Table 16 (SEQ ID NOs:228, 229, 230, 231) to confirm the presence of xylose isomerase in the genome.

Table 16		
Primers Used in Integration Verification		
Primer	SEQ ID NO:	Sequence
N16PCR_F	228	CCCCATCGACAACTACGAGCTCACT
N16PCR_R	229	CAACTTGCCGTCCTCGAAGTCCTTG
N05PCR_F	230	CGAGCCTGAGAAGGTCGTGATGGGA
N05PCR_R	231	TACGTCGAAGTCGGGGTTGGTAGAA
N08PCR_F	240	TACTTGGCTGTAAAGATGAAG
N08PCR_R	241	ATCTAGCAGCCTTCATAGCCTT
N17PCR_F	242	CGAAGGTACTGACGCTGTTACC
N17PCR_R	243	CGAAAGAAGCGTATCTAGCCTT

[0122] Confirmed haploid strains were BD31328 (MATa), BD31336 (MATalpha), BD31526 (MATa) and BD31527 (MATalpha), BD34364 (MATa) and BD34365 (MATalpha), BD34366 (MATa) and BD34367 (MATalpha). Diploid strains BD31378 (expressing a xylose isomerase of SEQ ID NO:96), BD31365 (expressing a xylose isomerase of SEQ ID NO:78), BD34369 (expressing a xylose isomerase of

SEQ ID NO:54) and BD34377 (expressing a xylose isomerase of SEQ ID NO:58) were generated by conventional plate mating on YPXylose (YP+2% xylose) agar plates, about 2 days at about 30°C. Colony PCR with specific primers checking mating types were performed (shown in Table 17) and single colonies having MATa and MATalpha were picked as diploid strains BD 31378 (SEQ ID NO:96), BD31365 (SEQ ID NO:78), BD34369 (SEQ ID NO:54) and BD34377 (SEQ ID NO:58).

[0123] A linear fragment encoding the URA3 sequence (SEQ ID NO:237; TTAATTAAGTTAATTACCTTTTTTTCGAGGCATATTTATGGTGAAGAATAAGTTTTGACCATC AAAGAAGGTTAATGTGGCTGTGGTTTCAGGGTCCATAAAGCTTTTCAATTCATCATTTTTTTT TTATTCTTTTTTTTGATTCCGGTTTCCTTGAAATTTTTTTTGATTCCGTAATCTCCGAACAGAAG GAAGAACGAAGGAAGGAGCACAGACTTAGATTGGTATATATACGCATATGTAGTGTGAAG AACATGAAATTGCCAGTATTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAAACG AAGATAAATCATGTGCGAAAGCTACATATAAGGAACGTGCTGCTACTCATCCTAGTCCTGTTG CTGCCAAGCTATTTAATATCATGCACGAAAAGCAAACAACTTGTGTGCTTCATTGGATGTT CGTACCACCAAGGAATTACTGGAGTTAGTTGAAGCATTAGGTCCCAAAATTTGTTTACTAAA AACACATGTGGATATCTTGACTGATTTTTCCATGGAGGGCACAGTTAAGCCGCTAAAGGCAT TATCCGCCAAGTACAATTTTTTACTCTTCGAAGACAGAAAATTTGCTGACATTGGTAATACA GTCAAATTGCAGTACTCTGCGGGTGTATACAGAATAGCAGAATGGGCAGACATTACGAATG CACACGGTGTGGTGGGCCAGGTATTGTTAGCGGTTTGAAGCAGGCGGCAGAAGAAGTAAC AAAGGAACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCATGCAAGGGCTCCCTAGCTACTG GAGAATATACTAAGGGTACTGTTGACATTGCGAAGAGCGACAAAGATTTTGTTATCGGCTTT ATTGCTCAAAGAGACATGGGTGGAAGAGATGAAGGTTACGATTGGTTGATTATGACACCCG GTGTGGGTTTAGATGACAAGGGGAGACGCATTGGGTCAACAGTATAGAACCGTGGATGATGT GGTCTCTACAGGATCTGACATTATTATTGTTGGAAGAGGACTATTGCAAAGGGAAGGGATG CTAAGGTAGAGGGTGAACGTTACAGAAAAGCAGGCTGGGAAGCATATTTGAGAAGATGCGG CCAGCAAACTAAAAAACTGTATTATAAGTAAATGCATGTATACTAACTCACAAATTAGA GCTTCAATTTAATTATATCAGTTATTACCCGGGAATCTCGGTCGTAATGATTTTTATAATGAC GAAAAAAAAAAAAATTGGAAAGAAAAAGCTTCATGGCCTTTATAAAAAGGAACCATCCAATA CCTCGCCAGAACCAAGTAACAGTATTTTACGGTTAATTAA) was transformed into BD 31378 (SEQ ID NO:96), BD31365 (SEQ ID NO:78), BD34369 (SEQ ID NO:54) and BD34377 (SEQ ID NO:58) by a conventional transformation protocol, and transformants were plated on SCXylose-URA (Synthetic Complete, Uracil dropout) for selection. Colonies were checked by PCR with primers shown in Table 17, SEQ ID NO:235, SEQ ID NO:236). Confirmed strains are BD31446 (SEQ ID NO:78), BD31448 (SEQ ID NO:96), BD34373 (SEQ ID NO:54) and BD34378 (SEQ ID NO:58).

Table 17		
Primers Used in Mating Type Verification		
Primer	SEQ ID NO:	Sequence
1-mating type-R	232	AGTCACATCAAGATCGTTTAT
2-mating type alpha-F	233	GCACGGAATATGGGACTACTT
3-mating type a-F	234	ACTCCACTTCAAGTAAGAGTT
Ura fix-F	235	GAACAAAAACCTGCAGGAAACGAAGAT
Ura fix-R	236	GCTCTAATTTGTGAGTTTAGTATACATGCAT

[0124] Table 18 below shows the genotypes of the resulting yeast strains:

Table 18		
Strain Construction		
Name	Parent Strain	Description
pBPB007	yBPA130	<i>MATa, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2</i>
pBPB008	yBPA136	<i>MATalpha, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2</i>
BD31328	pBPB007	<i>MATa, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:96)</i>
BD31336	pBPB008	<i>MATalpha, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:96)</i>
BD31526	pBPB007	<i>MATa, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:78)</i>
BD31527	pBPB008	<i>MATalpha, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:78)</i>
BD34364	pBPB007	<i>MATa, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:238)</i>
BD34365	pBPB008	<i>MATalpha, ura3, adh2 :: TAL1-XKS1, pho13:: TKL1-XKS1, gre3:: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:238)</i>

Table 18		
Strain Construction		
Name	Parent Strain	Description
BD34366	pBPB007	<i>MATa, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:239)
BD34367	pBPB008	<i>MATalpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:239)
BD31378	BD31328 BD31336	<i>MATa/alpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:96)
BD31365	BD31526 BD31527	<i>MATa/alpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:78)
BD34369	BD34364 BD34365	<i>MATa/alpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:238)
BD34377	BD34366 BD34367	<i>MATa/alpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:239)
BD31448	BD31378	<i>MATa/alpha, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:96)
BD31446	BD31365	<i>MATa/alpha. adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2, rDNA::XI (SEQ ID NO:78)

5.13 Example 12: Fermentation Performance of Yeast Strain Expressing Different Xylose Isomerases

[0125] Fermentation performances of two different XI-expressing yeast strains were evaluated using the DasGip fermentation systems (Eppendorf, Inc.). DasGip fermenters allowed close control over agitation, pH, and temperature ensuring consistency of the environment during fermentation. DasGip fermenters were used to test performance of the yeast strains expressing the XI genes on hydrolysate (Hz)

(neutralized with magnesium bases) as a primary carbon source. Prior to the start of fermentation strains were subjected to propagation testing consisting of two steps as described below.

[0126] SEED 1: About 1 ml of strain glycerol stock was inoculated into about 100 ml of YP (Yeast extract, Peptone) medium containing about 2% glucose and about 1% xylose in the 250 ml bellco baffled flask (Bellco, Inc.). Strains were cultivated at about 30°C with about 200 rpm agitation for at least 18 hours until at full saturation. Optical density was assessed by measuring light absorbance at wavelength of 600 nm.

[0127] SEED 2: About 20 ml of saturated SEED 1 (see preceding paragraph) was inoculated into 3L Bioflo unit (New Brunswick, Inc.) containing about 2.1 L of basal medium at pH 6.0 (1% v/v inoculation). Cultivation was conducted at about 30°C in a fed batch mode with constant air flow of about 2L/min. Agitation ramp (rpm) was about 200-626 rpm over about 15 hours starting at about 5 hours of elapsed fermentation time (EFT). Feeding profile was about 0-4.8 ml/min over 20 hours. The basal medium contained (per 1L): about 20 % of neutralized hydrolysate (Hz); about 20 g/L sucrose (from cane juice); about 35 ml of nutrients mixture (Table 19), about 1 ml of vitamin mixture (Table 20); about 0.4 ml of antifoam 1410 (Dow Corning, Inc.) and water. Feed medium contained (per 1L): about 20% neutralized hydrolysate (Hz), about 110 g/L sucrose (from cane juice), about 35 ml of nutrient mixture; about 1 ml of vitamin mixture, about 0.4 ml of antifoam 1410 (Dow Corning, Inc.) and water.

Table 19		
Nutrients mixture		
Component	FW g/mol	Conc.
KH ₂ PO ₄ H ₂ O	154.1	99.1 g/L
Urea	60.06	65.6 g/L
MgSO ₄ -7H ₂ O	192.4	14.6 g/L
DI Water	NA	To 1.0L

Table 20	
Vitamin mixture (1000x)	
Components	mM
ZnSO ₄	100
H ₃ BO ₃	24
KI	1.8
MnSO ₄	20
CuSO ₄	10
Na ₂ MoO ₄	1.5
CoCl ₂	1.5
FeCl ₃	1.23

[0128] DasGip fermentation: Strains were tested in small scale fermentation using the DasGip system in the industrially relevant medium containing detoxified hydrolysate and sucrose. Strains were propagated as described above; DasGip inoculation was performed using the following protocol:

[0129] Cell dry weight of SEED 2 was assessed based on the final optical density. Cell dry weight and optical density (600 nm) correlation was used to estimate the volume of the SEED 2 culture needed for fermentation. Targeted inoculation level was about 7% v/v; about 1.5 g/L cell dry weight. Appropriate volume of SEED 2 culture was harvested by centrifugation (about 5000 rpm for 10 min) to pellet the cells and resuspended in about 17.5 ml of PBS. Resuspended cell solution was used to inoculate a 500 ml DasGip unit containing about 250 ml of detoxified hydrolysate and nutrient solution (about 3.5 ml/100 ml of medium). Fermentation was performed at about 32°C at pH 6.3 with about 200 rpm. The duration of fermentation was about 92 hours with regular sampling. Sampling was conducted by a 25 ml steriological pipette through the port in the head plate of the DasGip unit. About 3 ml of culture were taken out, harvested by centrifugation (about 5000 rpm for 10 min) to pellet the cells and the supernatant was submitted for analysis. Standard analytical techniques such as high-pressure liquid chromatography (HPLC) were used to determine concentration of sugars and ethanol in the medium. Fermentation performances for yeast strains BD31378 (expressing a xylose isomerase of SEQ ID NO:96) and BD31365 (expressing a xylose isomerase of SEQ ID NO:78) are presented in FIG. 7A and FIG. 7B, respectively.

[0130] Serum bottle fermentation: Fermentation performances of BD34373 (SEQ ID NO:54) and BD34378 (SEQ ID NO:58) were evaluated using the serum bottle fermentation system. New Wheaton Thin-Fling Lyp Stopper (VWR Inc., PA, USA) wrap individual 125mL Anaerobic Media bottles (VWR Inc., PA, USA) allowed close control over agitation, pH, and temperature ensuring consistency of the environment during fermentation. Serum bottle fermentations were used to test performance of the yeast strains expressing XI genes on clean sugar media (see Table 21 (below), supplemented with nutrients (Table 19, above) and vitamins (Table 20, above)).

Table 21	
Clean sugar	
Component	Conc.
Glucose	80 g/L
Xylose	80 g/L
Arabinose	5.0 g/L
Acetic acid	8.0 g/L

[0131] Yeast cells were inoculated into about 200 ml of YP (Yeast extract, Peptone) medium containing about 0.5 % glucose and about 3% xylose in the 500 ml bellco baffled flask (Bellco, Inc.). Strains were cultivated at about 30°C with about 200 rpm agitation for at least 18 hours until at full saturation. Optical

density was assessed by measuring light absorbance at wavelength of 600 nm. Targeted inoculation level was about 1.5 g/L cell dry weight. Appropriate volume of SEED culture was harvested by centrifugation (about 3500 rpm for 10 min) to pellet the cells and resuspended in about 20 ml of media. Resuspended cell solution was used to inoculate a 150 ml serum bottle containing about 90 ml of fermentation media. The autoclaved stopper was placed into serum bottles and then the serum bottles were clamped with aluminium seals (Bellco Inc., USA) by using a seal crimper (Bellco Inc., USA). The aluminium seal cap was peeled off and then the needle (Fisher, USA) inserted. Fermentation was performed at about 35°C, pH 5.5 with about 200 rpm. The duration of fermentation was about 44 hours. Fermentation performances for yeast strains BD34373 (SEQ ID NO:54) is presented in FIG. 7C and BD34378 (SEQ ID NO:58) are presented in FIG. 7D.

5.14 Example 13: Comparative Activity of XI's Encoded by Codon Optimized vs. Non-Optimized Open Reading Frames

[0132] The coding sequences for the XIs of SEQ ID NOs:38, 78 and 96 were subject to codon optimization using two approaches: the Boles codon optimization method and the DNA 2.0 Gene Designer software. The codon optimized sequences are set forth in Table 22 below:

Table 22		
Sequence Description	SEQ ID NO:	Sequence
Boles codon optimized DNA encoding XI of SEQ ID NO:78	244	ATGGCTAAGGAATATTTCCCATTCACCGGTAAGATTCCATTTCGAA GGTAAGGACTCTAAGAACGTTATGGCTTTCCACTATTATGAACCA GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGGTGAC GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT ATGGACAAGTTGGGTATTGAATATTTCTGTTTCACGACGTTGAC TTGGTTGAAGAAGGTGACACCATTGAAGAATATGAAGCTAGAATG AAGGCTATTACCGACTATGCTCAAGAAAAGATGAAGCAATTCCTCA AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTCTGGTAACAAG AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT TATATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACAC ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC TTGAAGGCTCACAACCTGGACAAGGACTTCAAGGTTAACATTGAA GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCTATTGACGCTAAC AGAGGTGACGCTCAAAAACGGTTGGGACACCGACCAATTCCCAATT GACAACTATGAATTGACCCAAGCTATGTTGGAAATTATTAGAAAC GGTGGTTTGGGTAACGGTGGAACCAACTTCGACGCTAAGATTAGA AGAACTCTACCGACTTGAAGACTTGTTCATTGCTCACATTTCT GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGACATT TTGGAAAACCTCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT TCTTTCGACCAAGGTGTTGGTAAGGACTTCGAAGACGGTAAGTTG

Table 22		
Sequence Description	SEQ ID NO:	Sequence
		ACCTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGGTGAACCA AAGCAAACCTCTGGTAAGCAAGAAAAAGTATGAAACCATTGTTGCT TTGTATGCTAAGTAA
Boles codon optimized DNA encoding XI of SEQ ID NO:96	245	ATGGCTAAGGAATATTTCCCATTCATTGGTAAGGTTCCATTGCAA GGTACTGAATCTAAGAACGTTATGGCTTTCCACTATTATGAACCA GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGCTGAC GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTTCGAAATT ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC TTGGTTGAAGAAGGTGAAACCGTTGCTGAATATGAAGCTAGAATG AAGGTTATTACCGACTATGCTTTGGAAGATGCAACAATCCCA AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTCGGTCACAAG AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT AAGTTGGGTGGTACTAACTATGTTTTCTGGGTGGTAGAGAAGGT TATATGTCTTTGTTGAACACCGACCAAAAAGAGAGAAAAGGAACAC ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC TTGAGGGCTCACGGTTTGGACAAGGACTTCAAGGTTAACATTGAA GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCTATTGACGCTAAC AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT GACAACTATGAATTGACCCAAGCTATGATGGAAATTATTAGAAAC GGTGGTTTGGGTAACGGTGGAAACCAACTTCGACGCTAAGATTAGA AGAACTCTACCGACTTGAAGACTTGTTCATTGCTCACATTTCT GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGCTATT TTGGAAGAATCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT TCTTTCGACGAAGGTATTGGTAAGGACTTCGAAGACGCTAAGTTG TCTTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGAAGAACCA AAGCAAACCTCTGGTAAGCAAGAAAAAGTATGAAACCATTGTTGCT TTGTATGCTAAGTAA
Boles codon optimized DNA encoding XI of SEQ ID NO:38	246	ATGAAGGAAATTTTCCCAAACATTCCAGAAATTAAGTTCGAAGGT AAGGACTCTAAGAACCCATTGCTTTCCACTATTATAACCCAGAC CAAATTATTTTGGGTAAGCCAATGAAGGAACACTTGCCATTGCT ATGGCTTGGTGGCACAACCTGGGTGCTACCGGTGTTGACATGTTT GGTGCTGGTCCAGCTGACAAGTCTTTCGGTGCTAAGGTTGGTACT ATGGAACACGCTAAGGCTAAGGTTGACGCTGGTTTTCGAATTCATG AAGAAGTTGGGTATTAGATATTTCTGTTTCCACGACGTTGACTTG GTCCAGAATGTGCTGACATTAAGGACACCAACAAGGAATTGGAC GAAATTTCTGACTATATTTTGGAAAAGATGAAGGGTACTGACATT AAGTGTGTTGGGGTACTGCTAACATGTTCTCTAACCCAAGATTCT TGTAACGGTGCTGGTTCTACCAACTCTGCTGACGTTTTCTGCTTTC GCTGCTGCTCAAGTTAAGAAGGCTTTGGACATTACCGTTAAGTTG GGTGGTAGAGGTTATGTTTTCTGGGGTGGTAGAGAAGGTTATGAA ACCTTGTGTAACACCGACGTTAAGTTTCGAACAAGAAAACATTGCT AGATTGATGAAGATGGCTGTTGAATATGGTAGATCTATTGGTTTC AAGGGTGACTTCTATATTGAACCAAAGCCAAAGGAACCAATGAAG CACCAATATGACTTCGACGCTGCTACCGCTATTGGTTTTCTTGAGG GCTCACGGTTTGGACAAGGACTTCAAGTTGAACATTGAAGCTAAC

Table 22		
Sequence Description	SEQ ID NO:	Sequence
		CACGCTACCTTGGCTGGTCACACCTTCCAACACGACTTGAGAATT TCTGCTATTAAACGGTATGTTGGGTCTATTGACGCTAACCAAGGT GACATGTTGTTGGGTGGGACACCGACGAATTCCCATTGACGTT TATTCTGCTACCCAATGTATGTATGAAGTTTTGAAGAACGGTGGT TTGACCGGTGGTTTCAACTTCGACTCTAAGACCAGAAGACCATCT TATACCATGGAAGACATGTTCTTGGCTTATATTTTGGGTATGGAC ACCTTCGCTTTGGGTGTTGATTAAAGGCTGCTCAAATTATTGAAGAC GGTAGAATTGACCAATTCATTGAAAAGAAGTATTCTTCTTTTCA GAAACCGAAATTGGTCAAAAAGATTTTGAACAACAAGACCTCTTTG AAGGAATTGCTGACTATGCTTGTAAGATGGGTGCTCCAGAATTG CCAGTTTCTGGTAGACAAGAAATGTTGGAAGCTATTGTTAACGAC GTTTTGTTCCGTAAGTAA
DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:78	247	ATGGCTAAGGAATACTTTCATTACCCGAAAGATACCATTTGAA GGTAAAGATTCTAAAAACGTAATGGCTTTTCATTATTACGAACCA GAAAAAGTTGTTATGGGCAAAAAGATGAAAGATTGGTTGAAATTT GCGATGGCTTGGTGGCATACTCGGGGGAGCTTCCGCTGATCAA TTTGGCGGACAAACCAGATCATAACGAATGGGATAAAGCAGGCGAT GCCGTGCAGAGAGCAAAGGATAAAATGGATGCTGGTTTCGAAATT ATGGATAAGCTAGGTATCGAATACTTCTGCTTCCATGACGTCGAT TTGGTTGAAGAGGGCGATACTATCGAGGAATACGAGGCGAGAATG AAGGCTATAACAGACTACGCCCAGGAGAAAATGAAACAATTTCCCT AACATCAAATTACTCTGGGGTACTGCCAATGTGTTTGGTAACAAA AGATACGCAAACGGGGCTTCAACTAATCCTGACTTCGATGTTGTT GCAAGAGCCATTGTTCAAATCAAAAACGCGATAGACGCTACTATT AACTAGGTGGCAGCAATTACGTCCTTTTGGGGTGAAGGGAAGGT TACATGTCCTCTGCTTAATACAGATCAGAAGAGAGAGAAGGAACAC ATGGCAACAATGCTCACTATGGCCCGTGACTACGCAAGAGCAAAA GGTTTTAAGGGCACTTTCCCTTATCGAACCAGGCTATGGAACCA TCAAAACACCAATATGATGTTGACACAGAACTGTGATCGGCTTT TTGAAAGCTCATAACTTGGACAAGGATTTCAAAGTAAACATTGAA GTTAATCATGCTACACTAGCAGGACACACATTTGAACACGAACG GCCTGTGCGGTAGATGCAGGGATGCTGGGTCTATCGACGCTAAT AGAGGGGATGCTCAAAATGGTTGGGATACCGATCAATTTCCAATC GACAATTACGAATTAACACAAGCTATGTTGGAGATTATTAGAAAT GGAGGTTTGGGTAATGGGGGTACAAACTTCGATGCTAAGATTTCGT CGAAATTCCACAGACTTAGAAGATTTGTTTCATTGCGCATATATCT GGTATGGATGCTATGGCCAGAGCATTAAATGAATGCCGCTGACATC TTAGAAAACAGTGAACCTCCAGCAATGAAAAAGGCCAGATATGCC TCTTTTCGATCAAGGTGTAGGAAAAGATTTTGAAGCAGGCAAGTTG ACTTTAGAACAAGTCTATGAATACGGTAAAAAGGTCGGCGAACCT AAGCAAACCAGCGGAAAGCAAGAGAAATACGAGACTATCGTGGCT CTTTATGCAAAATAA
DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:96	248	ATGGCCAAGGAGTACTTCCCTTTTATCGGCAAGGTCCCATTTGAA GGGACAGAATCCAAAAACGTCATGGCTTTTCACTACTATGAACCT GAGAAGGTAGTTATGGGTAAAAAGATGAAAGATTGGTTGAAGTTT GCAATGGCATGGTGGCATACTTGGGTGGGGCCTCTGCTGATCAA TTTGGAGGACAAACTAGATCATAACGAATGGGATAAAGCAGCTGAT GCCGTTCAAAGAGCCAAAGATAAAATGGATGCCGGGTTTCGAAATC ATGGACAAATTGGGTATCGAATATTTCTGCTTCCATGATGTAGAC CTTGTGAGGAGGGTGAAACCGTCGCTGAATATGAGGCGAGAATG AAGGTTATTACGGATTACGCACTAGAAAAGATGCAGCAGTTTCCA

Table 22		
Sequence Description	SEQ ID NO:	Sequence
		AACATAAACTATTGTGGGTACTGCTAATGTTTTCGGACATAAA CGTTACGCTAACGGAGCTTCCACTAATCCAGACTTTGATGTTGTC GCGAGAGCTATCGTTCAAATCAAAAATGCAATCGATGCTACAATT AAGTTAGGAGGGACAAATTACGTGTTCTGGGGTGGTAGAGAAGGT TACATGAGCCTGCTTAATACAGATCAAAAGAGAGAAAAGGAGCAC ATGGCAACAATGCTAACAATGGCTAGAGATTATGCCCAGCTAAG GGCTTCAAAGGCACTTTTCTGATAGAACCTAAACCAATGGAACCA TCTAAACACCAATACGATGTAGACACCGAAACTGTAATAGGCTTC CTTCGTGCACATGGTTTGGATAAAGATTTAAGGTGAACATTGAA GTGAATCATGCTACTTTAGCCGGTCACACTTTTGAACATGAATTA GCATGTGCTGTTGATGCGGGAATGTTGGGTTCTATCGATGCCAAC AGAGGCGACGCCCCAAAATGGTTGGGACACAGACCAGTTTCCTATT GACAATTACGAACTCACCCAAGCTATGATGGAAATTATCAGGAAT GGGGGACTGGGAAATGGTGGTACGAACTTTGATGCGAAGATAAGG AGAACTCTACTGACTTAGAAGATTTGTTTATAGCACATATTTCA GGTATGGACGCTATGGCAAGAGCTTTAATGAATGCCGAGCAATC TTGGAGGAAAGTGAACCTCCAGCTATGAAAAAGGCAAGATACGCA AGTTTTGATGAGGTATTGGCAAAGACTTCGAAGATGGTAAACTA TCTTTAGAACAAAGGTACGAGTATGGCAAAAAGGTAGAGGAACCA AAACAAACATCAGGCAAAACAAGAGAAATATGAAACAATTGTCGCT CTTTACGCGAAGTAA
DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:38	249	ATGAAGGAAATCTTCCCTAACATCCCAGAGATCAAATTGCAAGGC AAAGACTCTAAAAATCCATTTGCCTTCCACTATTACAACCCAGAC CAGATCATTTTAGGTAAACCAATGAAGGAGCACTTGCCATTTGCT ATGGCTTGGTGGCATAATCTAGGCGCCACTGGTGTGATATGTTT GGTGCAGGCCCTGCGGACAAATCTTTCGGAGCTAAAGTAGGAAT ATGGAACATGCAAAAGCGAAAGTTGATGCTGGGTTTGAGTTTCATG AAGAAATTAGGAATCAGATATTTCTGCTTTCATGATGTTGACTTG GTTCTGAGTGTGCTGACATTAAGGATACAAACAAGGAACCTGAT GAAATCTCTGACTACATTTTGGAAAAGATGAAAGGTACTGACATA AAGTGTTTGTGGGGCACGGCTAATATGTTTTCCAATCCAAGATTT TGTAACGGCGCTGGCTCAACTAATTCAGCAGATGTCTTTGCATTC GCTGCTGCACAAGTCAAGAAAGCACTTGACATTACAGTCAAACCTG GGTGGGAGAGGATACGTTTTCTGGGGTGGTAGAGAAGGCTACGAA ACATTGTTGAATACAGACGTTAAGTTTGAACAAGAGAATATTGCA AGGTTAATGAAAAATGGCAGTGGAATATGGGCGTTCTATAGGTTTT AAAGGTGATTTCTACATTGAGCCAAAACCAAAGGAACCTATGAAA CATCAATACGATTTTCGATGCCGCAACAGCAATAGGTTTTCCTTAGA CCCCACGGGTGGATAAAGACTTTAAGCTCAATATCGAAGCCAAC CACGCAACACTTGCAGGCCATACATTTCAACATGATCTTAGAATA TCTGCTATTAAACGGAATGCTCGGCTCAATTGATGCCAATCAGGGT GATATGCTACTAGGTTGGGATACTGATGAGTTTCCATTTGATGTA TACTCCGCTACACAATGCATGTATGAGGTGTTGAAAAATGGTGGT CTGACCGGTGGCTTCAACTTCGATAGTAAGACCAGACGTCCTTCA TACACTATGGAAGATATGTTTCTGGCGTATATCTTAGGTATGGAC ACATTTGCTTTAGGTCTAATCAAAGCCGCTCAAATCATTGAAGAT GGCAGAAATGACCAGTTTATAGAAAAGAAATACTCCAGTTTTCGA GAAACCGAAATCGGACAAAAGATTCTCAATAACAAAACCTTCATTG AAGGAATTATCTGATTACGCCTGTAAGATGGGTGCGCCAGAATTA CCTGGAAGCGGTAGACAAGAGATGCTTGAAGCTATCGTGAATGAT GTATTGTTTGGAAAATAA

[0133] The codon optimized DNA sequences were synthesized and incorporated into expression cassettes substantially as described in Example 12. Yeast strains were generated that included single copies of individual XI open reading frames integrated into the yeast YER131.5 locus. Strains confirmed to contain the XI expression cassettes were inoculated into about 3ml of modified YP Media (YP+0.1% Glucose + 3.0% Xylose) and incubated overnight at about 30°C and about 220 rpm. These overnight cultures were subcultured into about 25 ml of the same media to about OD₆₀₀ = 0.2. Samples were incubated overnight at about 30°C and about 220 rpm. Cultures were harvested when OD₆₀₀ was between about 3 and 4. Pellets were collected by centrifugation for about 5 minutes at about 4000 rpm. The supernatants were discarded and pellets washed with about 25 ml of distilled-deionized water and centrifuged again using the same conditions. Supernatants were discarded and the pellet frozen at about -20 °C until lysis and characterization.

[0134] Cell pellets were thawed and about 200 mg of each pellet sample was weighed out into 2 ml microcentrifuge tubes. About 50 µl of Complete®, EDTA-free Protease Inhibitor cocktail (Roche Part# 11873 580 001) at 5 times the concentration stated in the manufacturer's protocol was added to each sample. To this was added about 0.5 ml of Y-PER Plus® Dialyzable Yeast Protein Extraction Reagent (Thermo Scientific Part# 78999) (YP+) to each sample. Samples were incubated at about 25°C for about 4 hours on rotating mixer. Sample supernatants were collected after centrifugation at about 10,000 x g for about 10 minutes for characterization.

[0135] Total protein concentrations of the XI sample extracts prepared above were carried out using Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad, cat# 500-0006, Hercules CA) which is a modified version of the Bradford method (Bradford). In this assay, optical density readings were taken on a spectrophotometer set to 595 nm, and the standard curve was plotted as a linear regression line.

[0136] XI activity was determined using assay conditions at pH 7.5 as described in the Section 5.1.2. The specific activities of the codon optimized Xis are shown in Table 23.

Table 23			
Protein SEQ ID NO.	Nucleic Acid SEQ ID NO:	Codon optimization method	SA, pH 7.5 (U/mg)
78	77	Native	0.43
78	247	DNA2.0-Codon Optimization	0.14
78	244	Boles Codon Optimization	0.53
96	95	Native	0.42
96	248	DNA2.0-Codon Optimization	0.14
96	245	Boles Codon Optimization	0.40
38	37	Native	0.25
38	249	DNA2.0-Codon Optimization	0.27
38	246	Boles Codon Optimization	0.40

[0137] Because these specific activity data are determined on the basis of the total cellular protein mass, any variations in specific activity for any given XI are due to expression levels. These data demonstrate that the Boles codon optimization approach improves the expressibility of bacterial XIs in *S. cerevisiae*.

5.15 Example 14: Comparative Activity of XI's of The Disclosure vs. *Orpinomyces sp.* XI

[0138] The specific activities of exemplary XIs of the disclosure were compared to the specific activity of a known XI, *Orpinomyces sp.* XI assigned Genbank Accession No. 169733248. The XIs were incorporated into expression cassettes substantially as described in Example 12. Yeast strains were generated that included single copies of individual XI open reading frames integrated into the yeast YER131.5 locus. Activity of the individual clones was measured at pH 7.5 using a similar approach to that used in Example 12, except that the total protein concentrations were based on optical density readings at 450 nm and 595 nm, with the standard curve was plotted as a parametric fit. Results are shown in Table 24, below.

Table 24				
Protein SEQ ID NO.	Nucleic Acid SEQ ID NO:	Codon optimization method	SA, pH 7.5 (U/mg)	% activity as compared to <i>Orpinomyces</i> XI
NA	NA	Host negative control (no recombinant XI)	0.005	7%
Genbank Accession No. 169733248		<i>Orpinomyces sp.</i> XI (Native)	0.071	100%
96	95	Native	0.137	193%
54	238	Boles codon optimization	0.193	272%
58	239	Boles codon optimization	0.205	288%

[0139] While various specific embodiments have been illustrated and described, it will be appreciated that various changes can be made without departing from the spirit and scope of the invention(s).

WHAT IS CLAIMED IS:

1. A recombinant polypeptide comprising an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:78.
2. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93% or 95% sequence identity to amino acids 2-377 of SEQ ID NO:78 and which comprises the amino acid sequence of (a) the amino acid sequence of SEQ ID NO:212 or SEQ ID NO:213 and/or (b) the amino acid sequence of SEQ ID NO:214.
3. The recombinant polypeptide of claim 2, which comprises an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:78.
4. The recombinant polypeptide of any one of claims 1 to 3, which comprises an amino acid sequence having at least 97% or 98% sequence identity to SEQ ID NO:78.
5. A recombinant polypeptide comprising an amino acid sequence having at least 95%, 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:96.
6. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93% or 95% sequence identity to amino acids 2-377 of SEQ ID NO:96 and further comprises the amino acid sequence of (a) SEQ ID NO:212 or SEQ ID NO:213 and/or (b) SEQ ID NO:214.
7. The recombinant polypeptide of claim 6, which comprises an amino acid sequence having at least 95%, 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:96.
8. The recombinant polypeptide of any one of claims 5 to 7, which comprises an amino acid sequence having at least 95%, 97% or 98% sequence identity to SEQ ID NO:96.
9. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-374 of SEQ ID NO:38.
10. The recombinant polypeptide of claim 9, which comprises one, two, three, four or all five of (i) SEQ ID NO:206 or SEQ ID NO:207; (ii) SEQ ID NO:208; (iii) SEQ ID NO:209; (iv) SEQ ID NO:210; and (v) SEQ ID NO:211.
11. The recombinant polypeptide of claim 9 or claim 10, which comprises an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:38.
12. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-374 of SEQ ID NO:2.
13. The recombinant polypeptide of claim 12 which comprises an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:2.
14. A recombinant polypeptide comprising an amino acid sequence having at least 93%, 95%, 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:58.

15. The recombinant polypeptide of claim 14 which comprises an amino acid sequence having at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:58.

16. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-375 of SEQ ID NO:42.

17. The recombinant polypeptide of claim 16 which comprises one, two or all three of (a) the amino acid sequence of SEQ ID NO:206 or SEQ ID NO:207; (b) the amino acid sequence of SEQ ID NO:210; and (c) the amino acid sequence of SEQ ID NO:211.

18. A recombinant polypeptide comprising an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:84.

19. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93% or 95% sequence identity to amino acids 2-376 of SEQ ID NO:84 and which comprises (i) the amino acid sequence of SEQ ID NO:212 or SEQ ID NO:213 and/or (ii) the amino acid sequence of SEQ ID NO:214,

20. The recombinant polypeptide of claim 19, which comprises an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:84.

21. The recombinant polypeptide of any one of claims 18 to 20, which comprises an amino acid sequence having at least 97% or 98% sequence identity to SEQ ID NO:84.

22. A recombinant polypeptide comprising an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:80.

23. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93% or 95% sequence identity to amino acids 2-377 of SEQ ID NO:80 and which comprises (a) the amino acid sequence of SEQ ID NO:212 or SEQ ID NO:213 and/or (b) the amino acid sequence of SEQ ID NO:214.

24. The recombinant polypeptide of claim 22, which comprises an amino acid sequence having at least 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:80.

25. The recombinant polypeptide of any one of claims 22 to 24, which comprises an amino acid sequence having at least 97% or 98% sequence identity to SEQ ID NO:80.

26. A recombinant polypeptide comprising an amino acid sequence having at least 93%, 95%, 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:54.

27. The recombinant polypeptide of claim 26 which comprises an amino acid sequence having at least 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:54.

28. A recombinant polypeptide comprising an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:46.

29. The recombinant polypeptide of claim 28 which comprises the amino acid sequence of SEQ ID NO:206 or SEQ ID NO:207.

30. The recombinant polypeptide of claim 28 or claim 29 which comprises an amino acid sequence having at least 80%, 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:46.

31. A recombinant polypeptide comprising an amino acid sequence having at least 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-376 of SEQ ID NO:16.

32. The recombinant polypeptide of claim 31 which comprises an amino acid sequence having at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:16.

33. A recombinant polypeptide comprising an amino acid sequence having at least 85%, 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-375 of SEQ ID NO:82.

34. The recombinant polypeptide of claim 33 which comprises an amino acid sequence having at least 85%, 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:82.

35. A recombinant polypeptide comprising an amino acid sequence having at least 90%, 93%, 95%, 97% or 98% sequence identity to amino acids 2-377 of SEQ ID NO:32.

36. The recombinant polypeptide of claim 35 which comprises an amino acid sequence having at least 90%, 93%, 95%, 97% or 98% sequence identity to SEQ ID NO:32.

37. A polypeptide comprising an amino acid sequence having a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity, or having 100% sequence identity, to SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, or 176, or to a catalytic or dimerization domain thereof.

38. A nucleic acid which encodes a polypeptide according to any one of claims 1 to 37.

39. The nucleic acid of claim 38 which is codon optimized for expression in a eukaryotic cell.

40. The nucleic acid of claim 39, which is codon optimized for expression in yeast.

41. The nucleic acid of claim 40, which comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity, or having 100% sequence identity, to the nucleotide sequence of any one of any one of SEQ ID NOs:238, 239, 244, 245, 246, 247, 248 and 249, or a portion thereof encoding at least a xylose isomerase catalytic or dimerization domain.

42. The nucleic acid of claim 39, which is codon optimized for expression in a filamentous fungus.

43. A nucleic acid comprising a nucleotide sequence having at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 98%, or at least 99% sequence identity, or having 100% sequence identity, to the nucleotide sequence of any one of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, and 175, or to a portion thereof encoding a xylose isomerase catalytic or dimerization domain.

44. The nucleic acid of claim 43, which has less than 100% sequence identity to said nucleotide sequence of any one of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, and 175, or to said portion thereof encoding a xylose isomerase catalytic or dimerization domain, respectively.

45. The nucleic acid of claim 43 or claim 44 codon optimized for expression in a eukaryotic cell.

46. The nucleic acid of claim 45, which is codon optimized for expression in yeast.

47. The nucleic acid of claim 46, which comprises the nucleotide sequence of any one of SEQ ID NOS:238, 239, 244, 245, 246, 247, 248 and 249, or a portion thereof encoding at least a xylose isomerase catalytic or dimerization domain.

48. The nucleic acid of claim 45, which is codon optimized for expression in a filamentous fungus.

49. A vector comprising the nucleic acid of any one of claims 38 to 48.

50. The vector of claim 49 which further comprises an origin of replication.

51. The vector of claim 49 or claim 50 which further comprises a promoter sequence operably linked to said nucleotide sequence.

52. The vector of claim 51, wherein the promoter sequence is operable in yeast.

53. The vector of claim 51, wherein the promoter sequence is operable in filamentous fungi.

54. A recombinant cell comprising engineered to express the polypeptide of any one of claims 1 to 36.

55. The recombinant cell of claim 54 which is a eukaryotic cell.

56. The recombinant cell of claim 54 which is a yeast cell.

57. The recombinant cell of claim 56 which is a yeast cell of the genus *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Klockera*, *Schwanniomyces*, *Issatchenkia* or *Yarrowia*.

58. The recombinant cell of claim 57, wherein the yeast cell is of the species *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* or *K. fragili*, or *Issatchenkia orientalis*.

59. The recombinant cell of claim 58, which is a *S. cerevisiae* cell.

60. The recombinant cell of claim 59, comprising one or more genetic modifications resulting in at least one, any two, any three, any four or all of the following phenotypes:

- (a) an increase in transport of xylose into the cell;
- (b) an increase in xylulose kinase activity;
- (c) an increase in aerobic growth rate on xylose;
- (d) an increase in flux through the pentose phosphate pathway into glycolysis;
- (e) a decrease in aldose reductase activity;
- (f) a decrease in sensitivity to catabolite repression;
- (g) an increase in tolerance to ethanol, intermediates, osmolarity or organic acids;

and

- (h) a reduced production of byproducts.

61. The recombinant cell of claim 60, wherein one or more genetic modifications result in increased expression levels of one or more of a hexose or pentose transporter, a xylulose kinase, an enzyme from the pentose phosphate pathway, a glycolytic enzyme and an ethanologenic enzyme.

62. The recombinant cell of claim 61, wherein the increased expression levels are achieved by overexpressing an endogenous gene or expressing a heterologous gene in the recombinant cell.

63. The recombinant cell of claim 60, wherein one or more genetic modifications result in decreased expression levels of one or more of a hexose kinase gene, the MIG1 gene, and the MIG2 genes.

64. The recombinant cell of claim 60 which is engineered to express the xylose reductase (XR), xylose kinase (XK) and xylitol dehydrogenase (XD) pathway.

65. The recombinant cell of claim 56 in which the nucleic acid is operably linked to a promoter that is insensitive to catabolite repression.

66. The recombinant cell of claim 65 in which the promoter is the TDH3 promoter, the PGK1 promoter, the TEF1 promoter, or the ADH1 promoter.

67. The recombinant cell of claim 54 which is a filamentous fungal cell.

68. The recombinant cell of claim 67, wherein the filamentous fungal cell is of the genus *Aspergillus*, *Penicillium*, *Rhizopus*, *Chrysosporium*, *Myceliophthora*, *Trichoderma*, *Humicola*, *Acremonium* or *Fusarium*.

69. The recombinant cell of claim 67, wherein the filamentous fungal cell is of the species *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Penicillium chrysogenum*, *Myceliophthora thermophila*, or *Rhizopus oryzae*.

70. A host cell transformed with the vector of any one of claims 49 to 53.

71. The host cell of claim 70 which is a prokaryotic cell.

72. The host cell of claim 71 which is a bacterial cell.

73. The host cell of claim 70 which is a eukaryotic cell.

74. A method for producing a fermentation product, comprising culturing the recombinant cell of any one of claims 54 to 69 in medium containing xylose under conditions in which the fermentation product is expressed.

75. The method of claim 74, wherein the xylose in the medium is provided by lignocellulosic hydrolysate.

76. The method of claim 74 or claim 75, wherein the fermentation product is ethanol.

77. The method of claim 74 or claim 75, wherein the fermentation product is butanol, diesel, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, itaconic acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β -lactam antibiotic or a cephalosporin.

78. The method of claim 77, wherein the recombinant cell comprises a genetic modification that results in decreased alcohol dehydrogenase activity.

79. The method of claim 77 or claim 78, wherein the recombinant cell expresses one or more enzymes that confers on the cell the ability to produce said fermentation product.

80. The method of any one of claims 74 to 79, wherein the medium further comprises glucose.

81. The method of any one of claims 74 to 79, wherein the recombinant cell is cultured under anaerobic conditions.

82. The method of any one of claims 74 to 81, further comprising recovering the fermentation product.

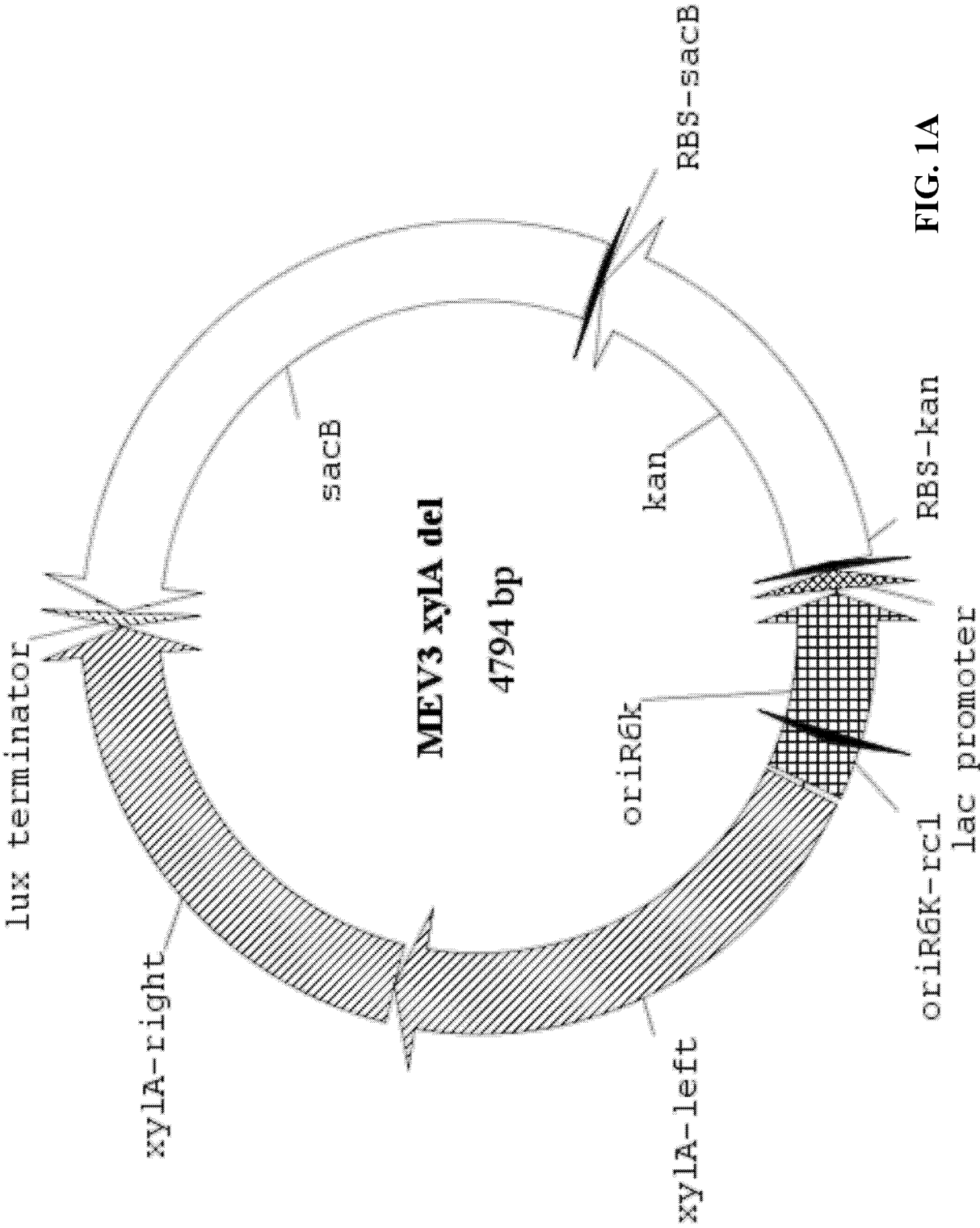
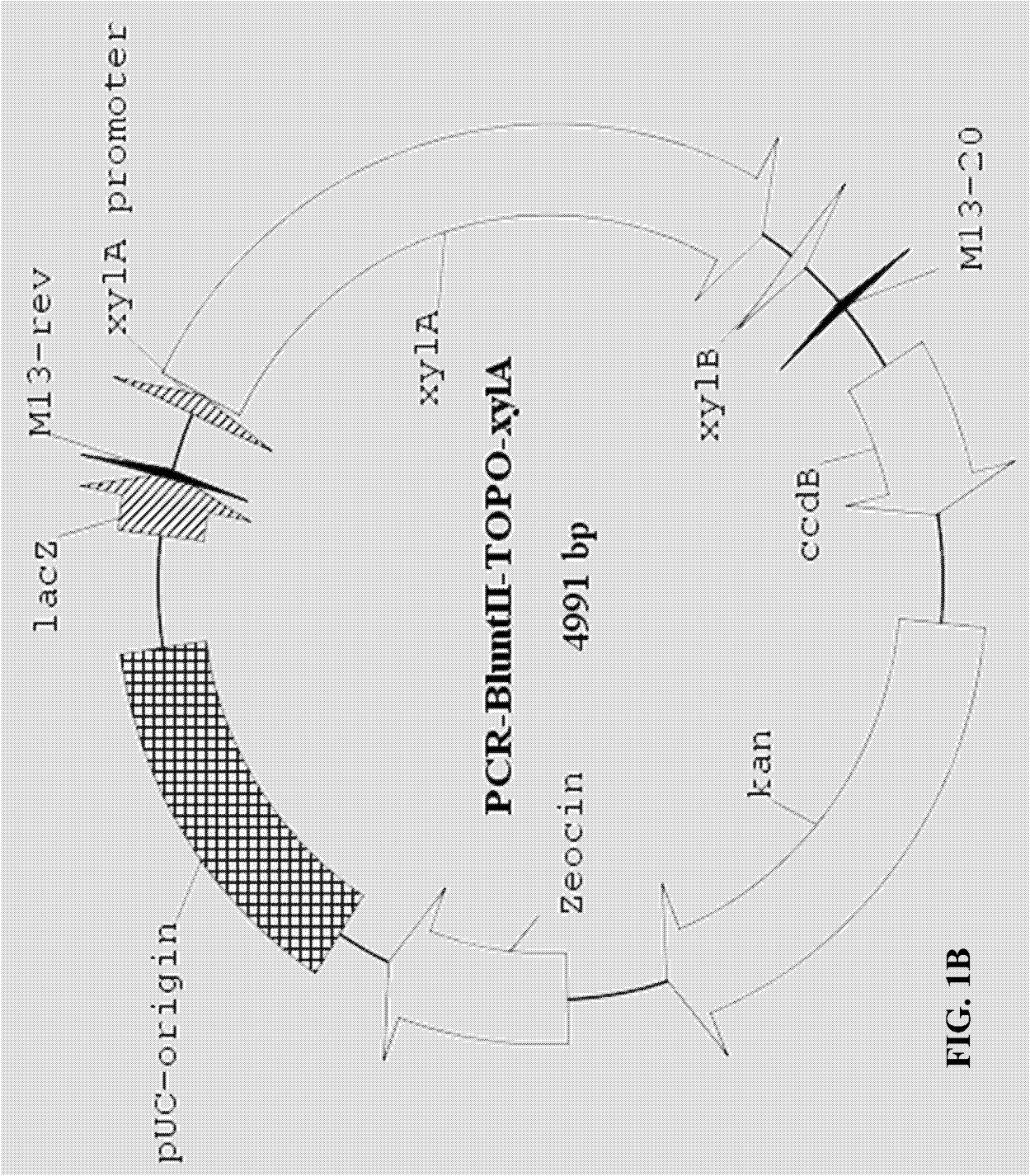


FIG. 1A



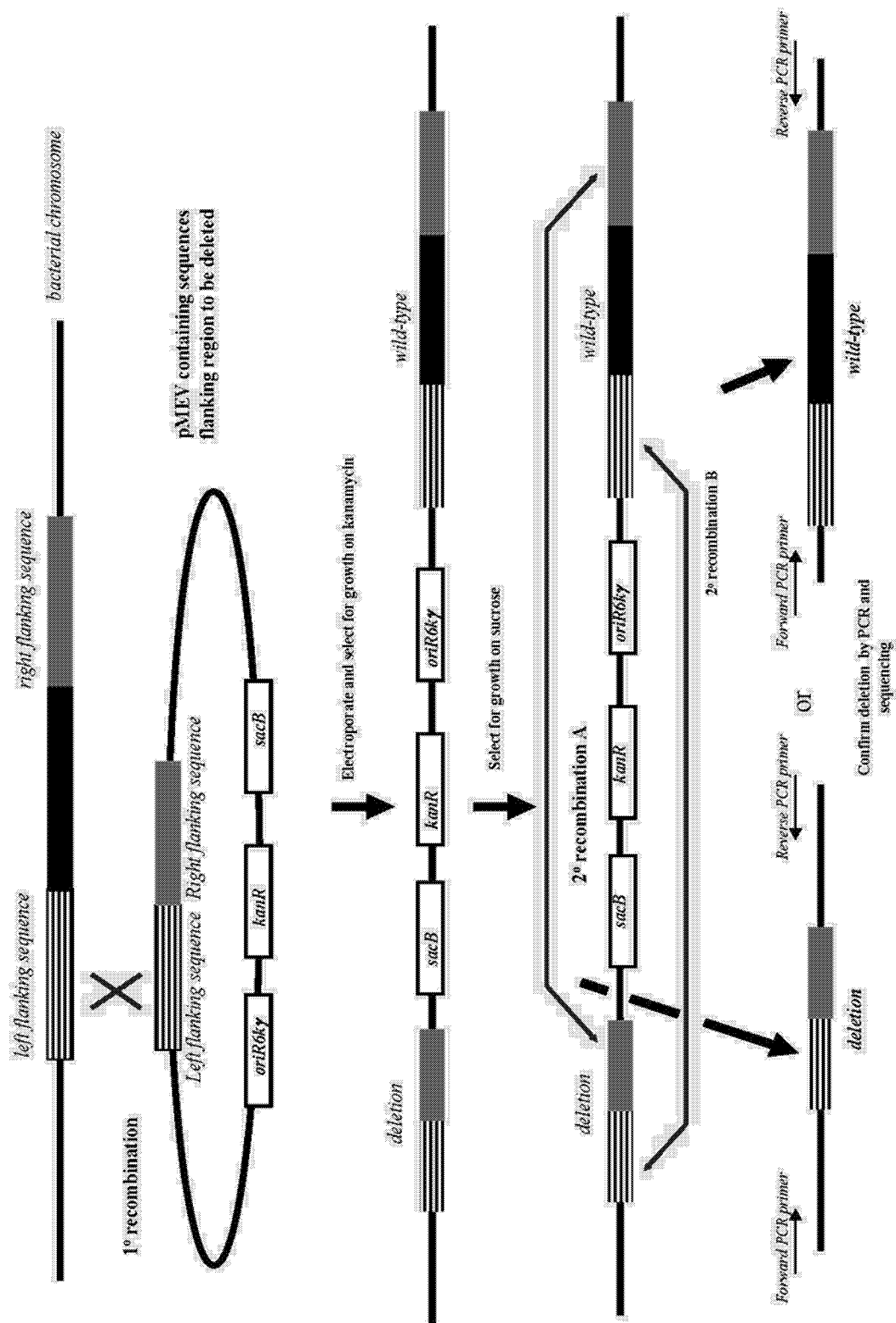


FIG. 2

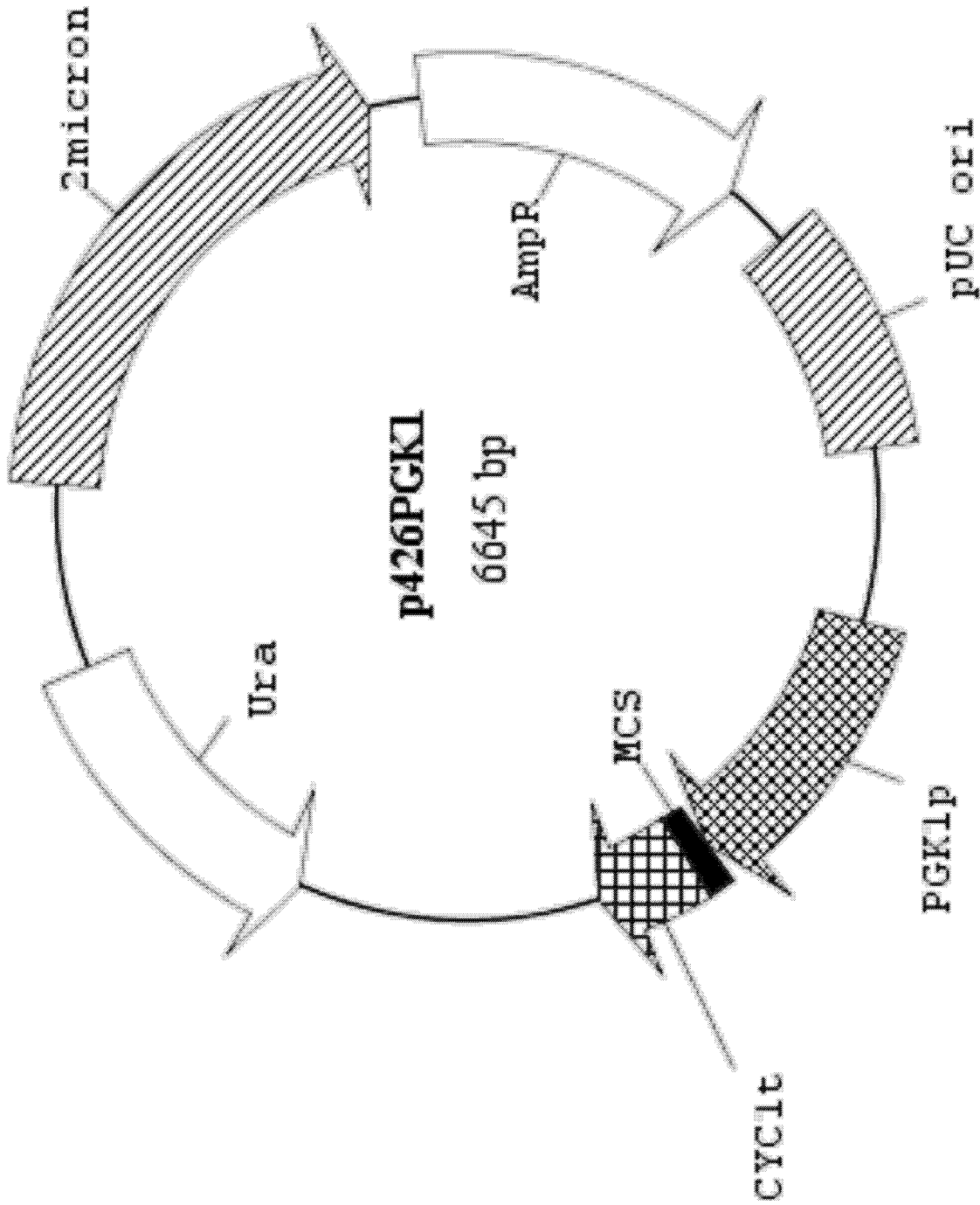


FIG. 3

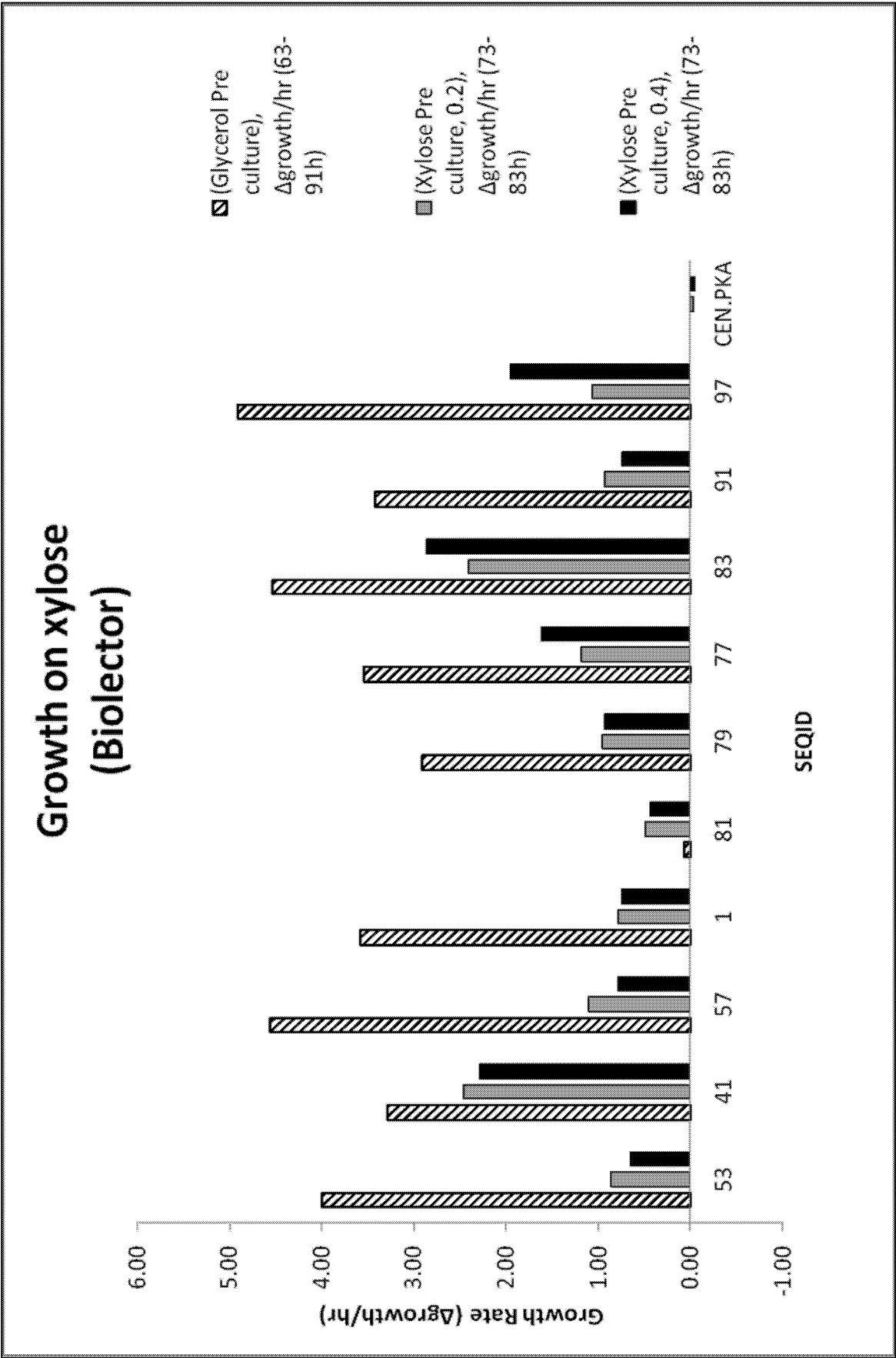


FIG. 4

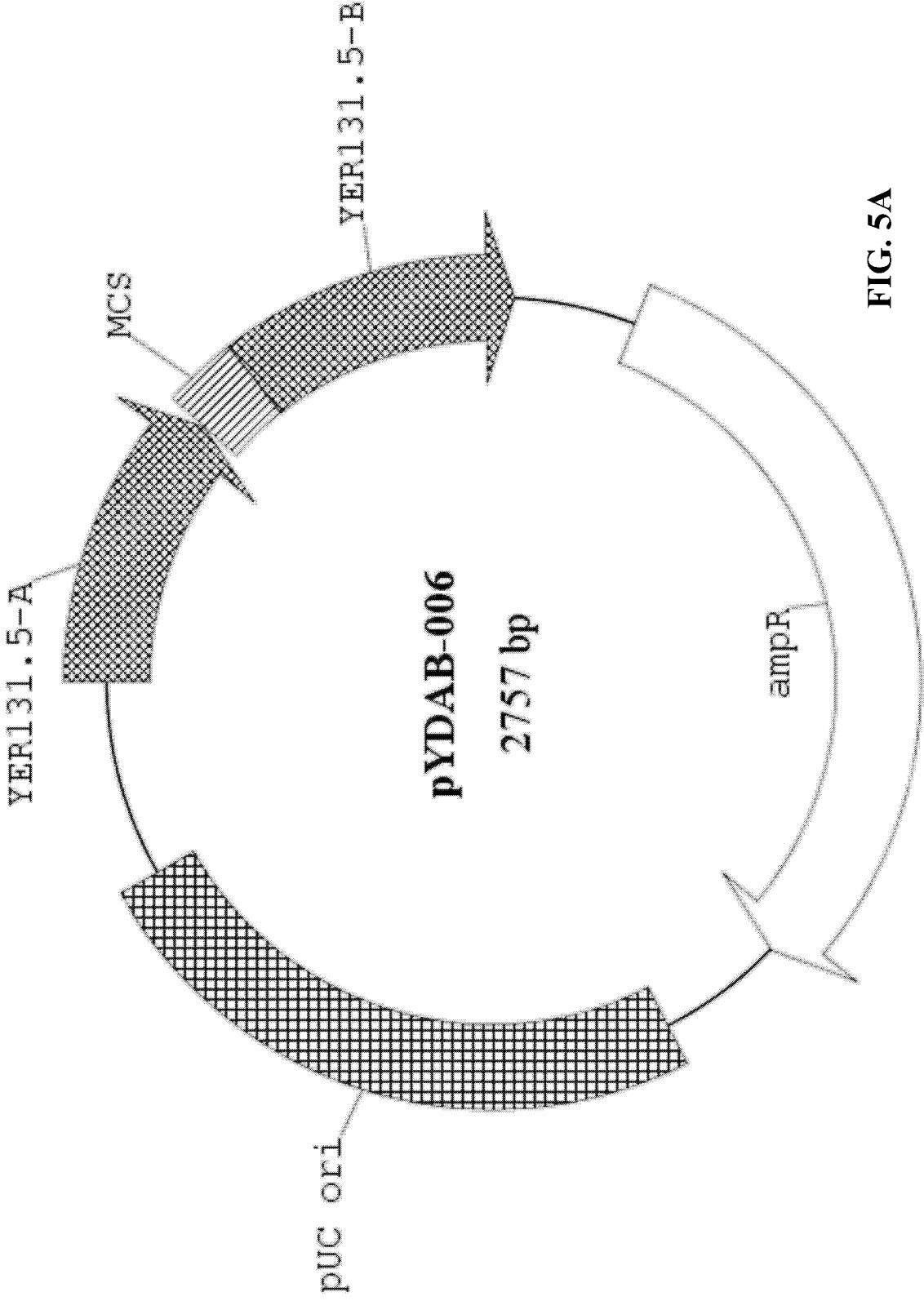


FIG. 5A

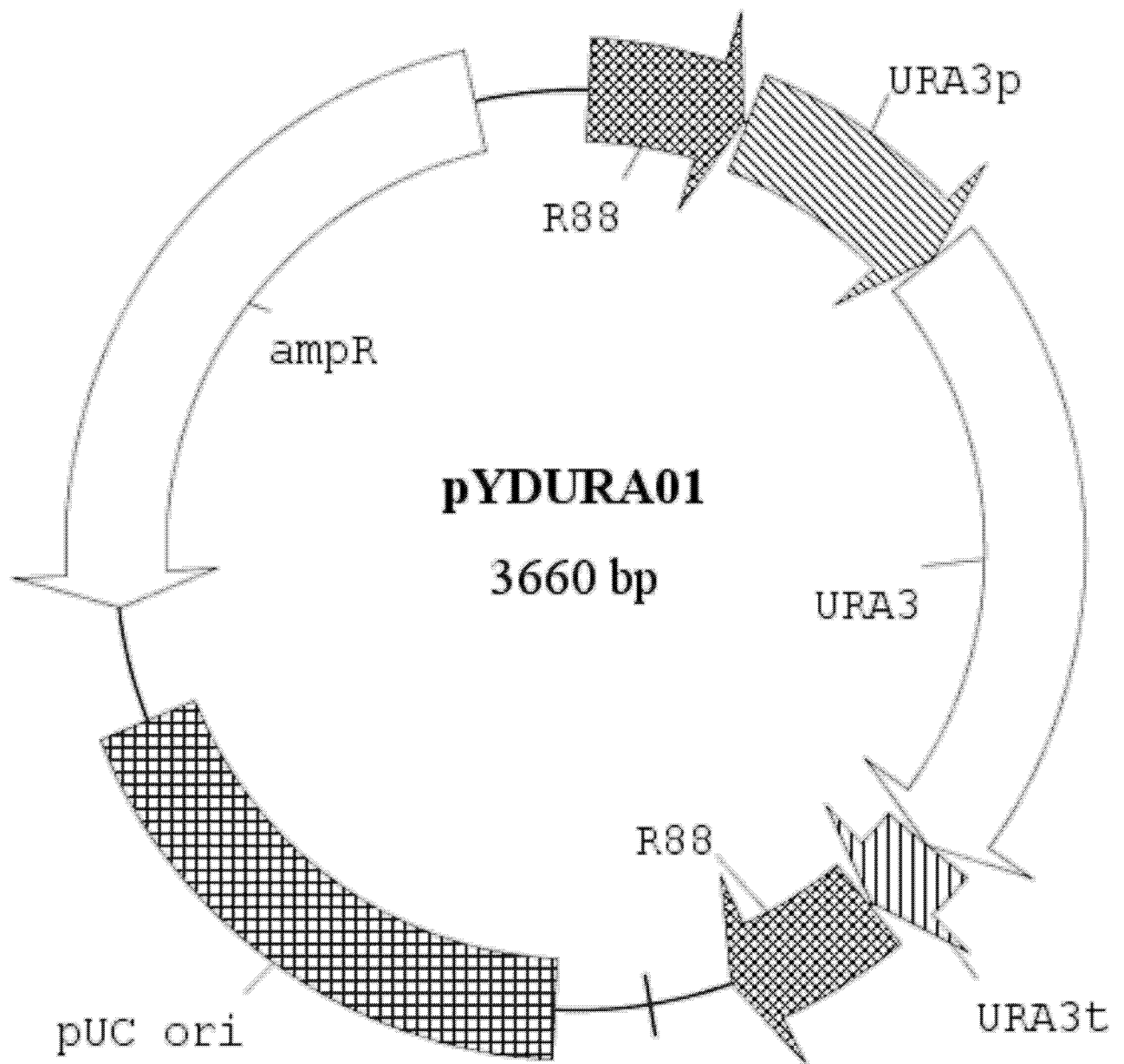


FIG. 5B

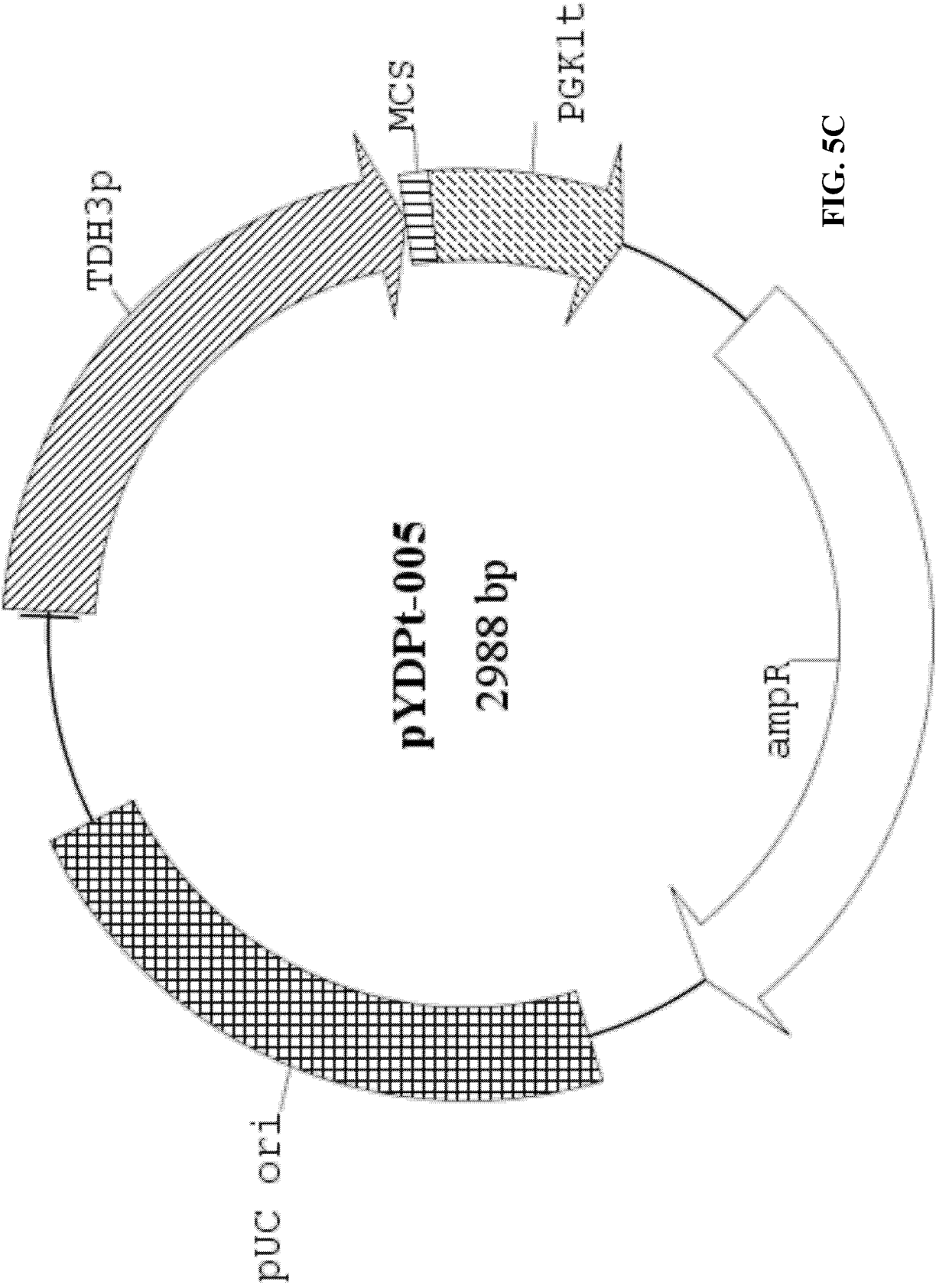


FIG. 5C

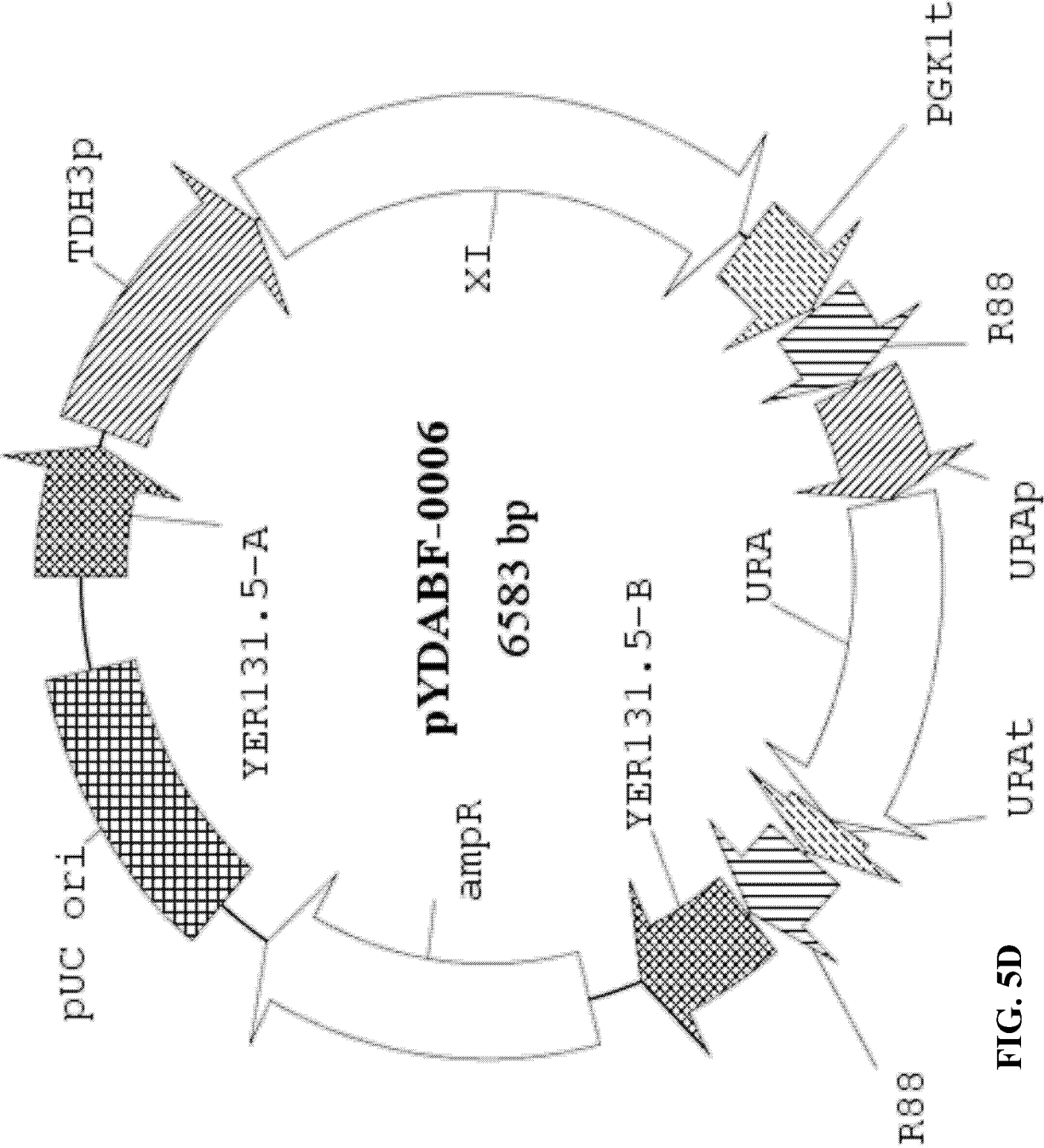
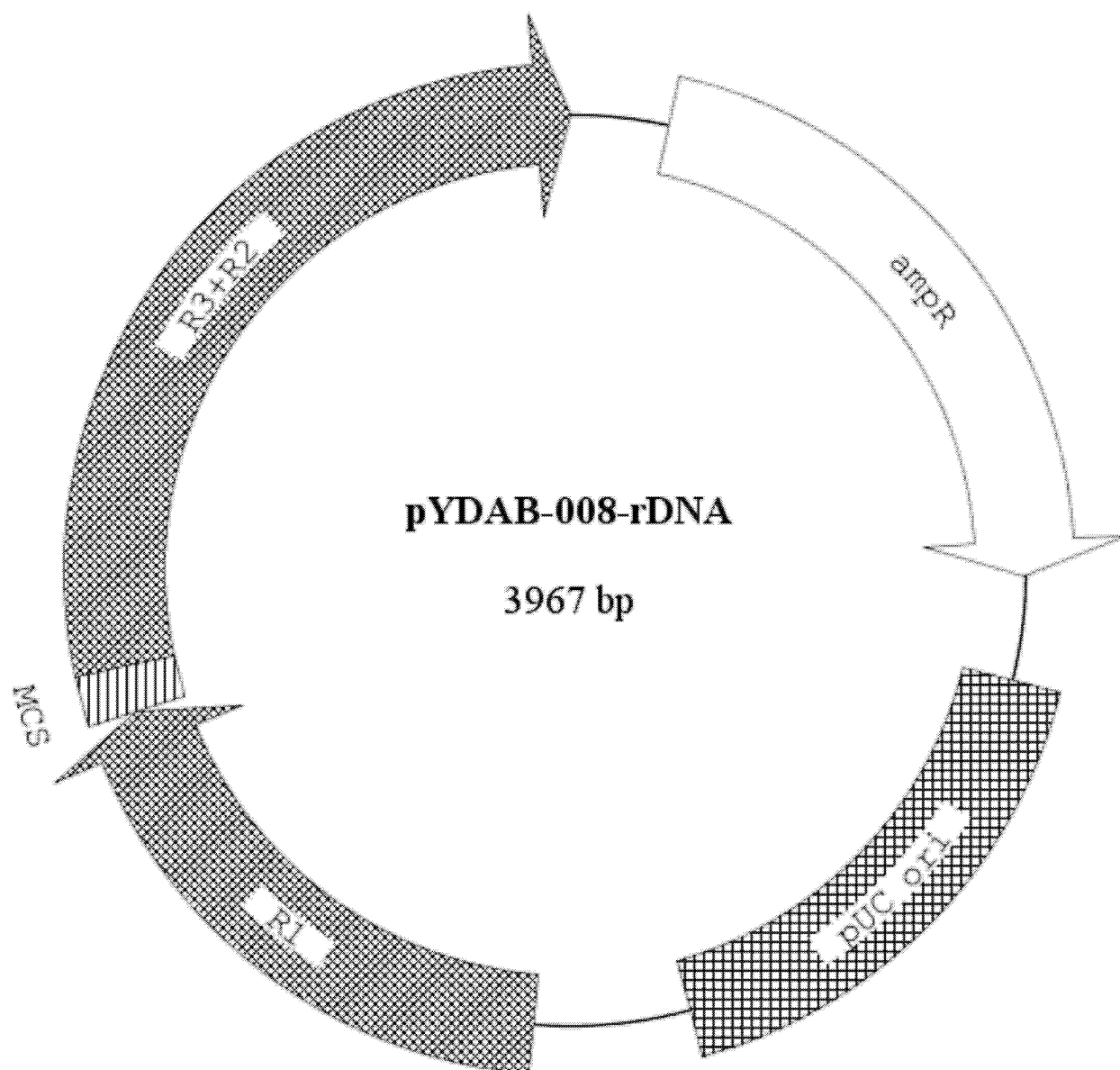


FIG. 5D

**FIG. 6**

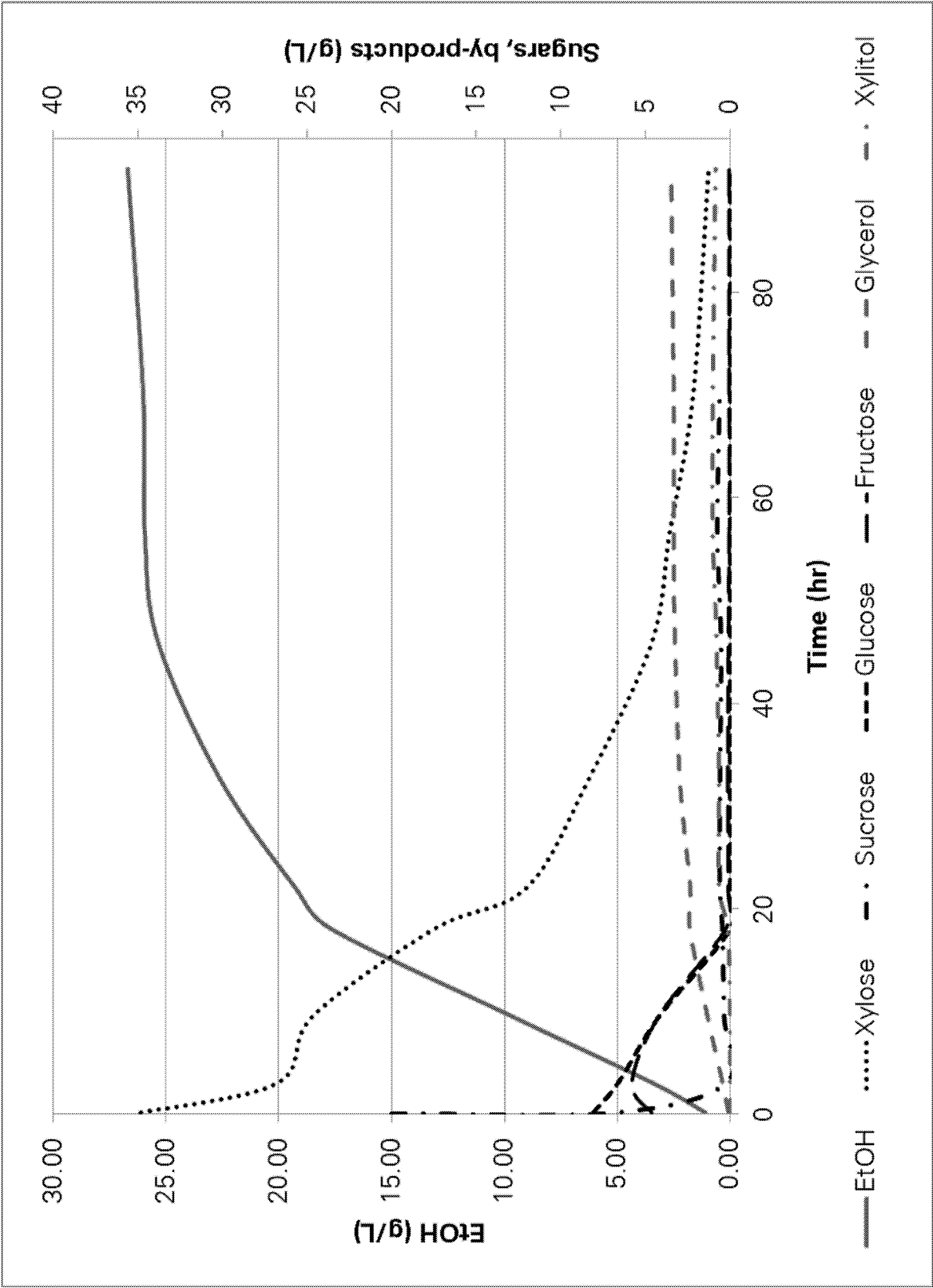


FIG. 7A

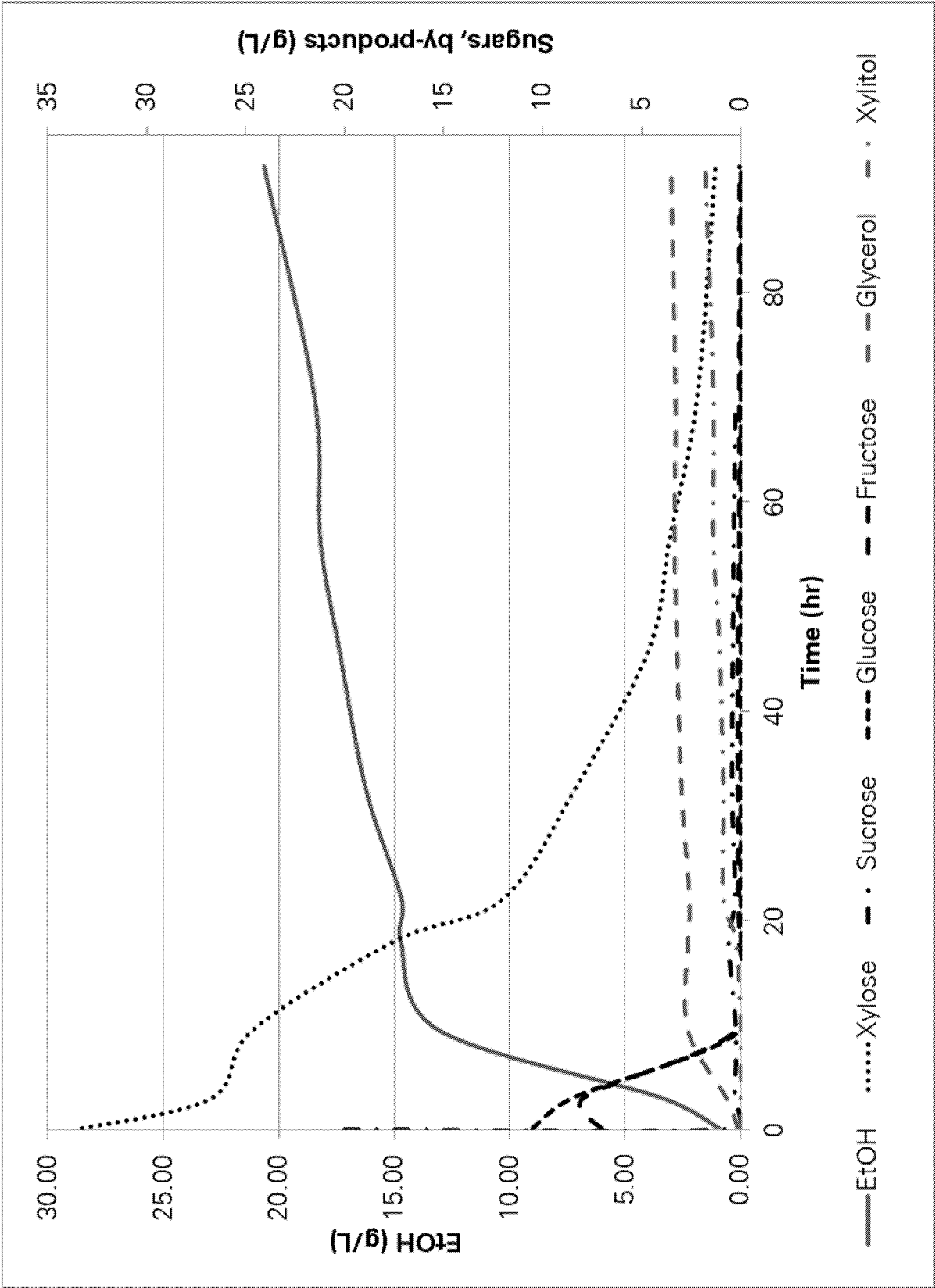


FIG. 7B

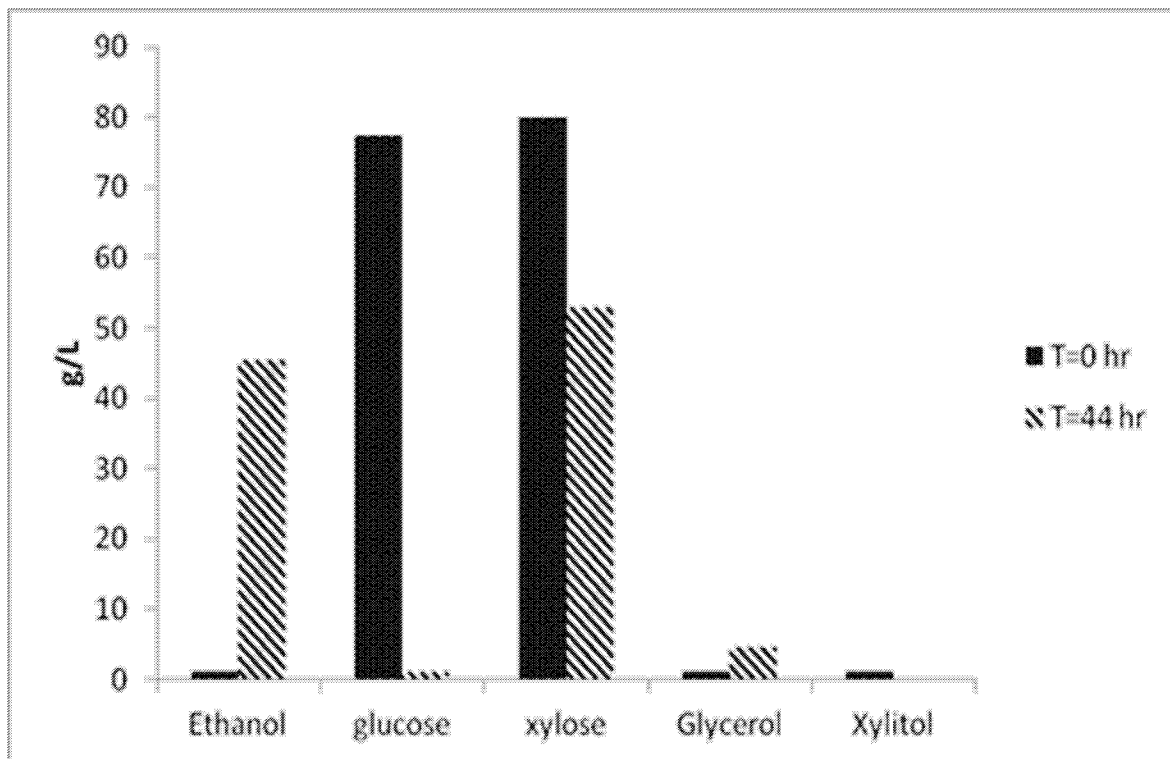


FIG. 7C

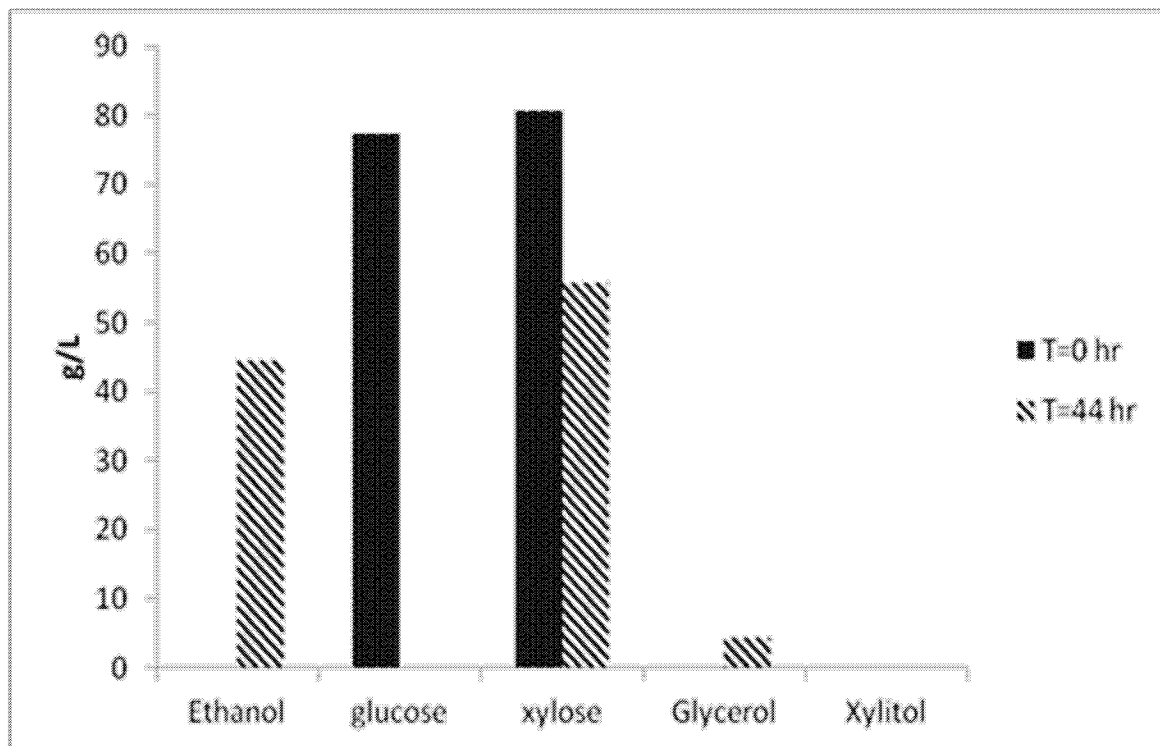


FIG. 7D

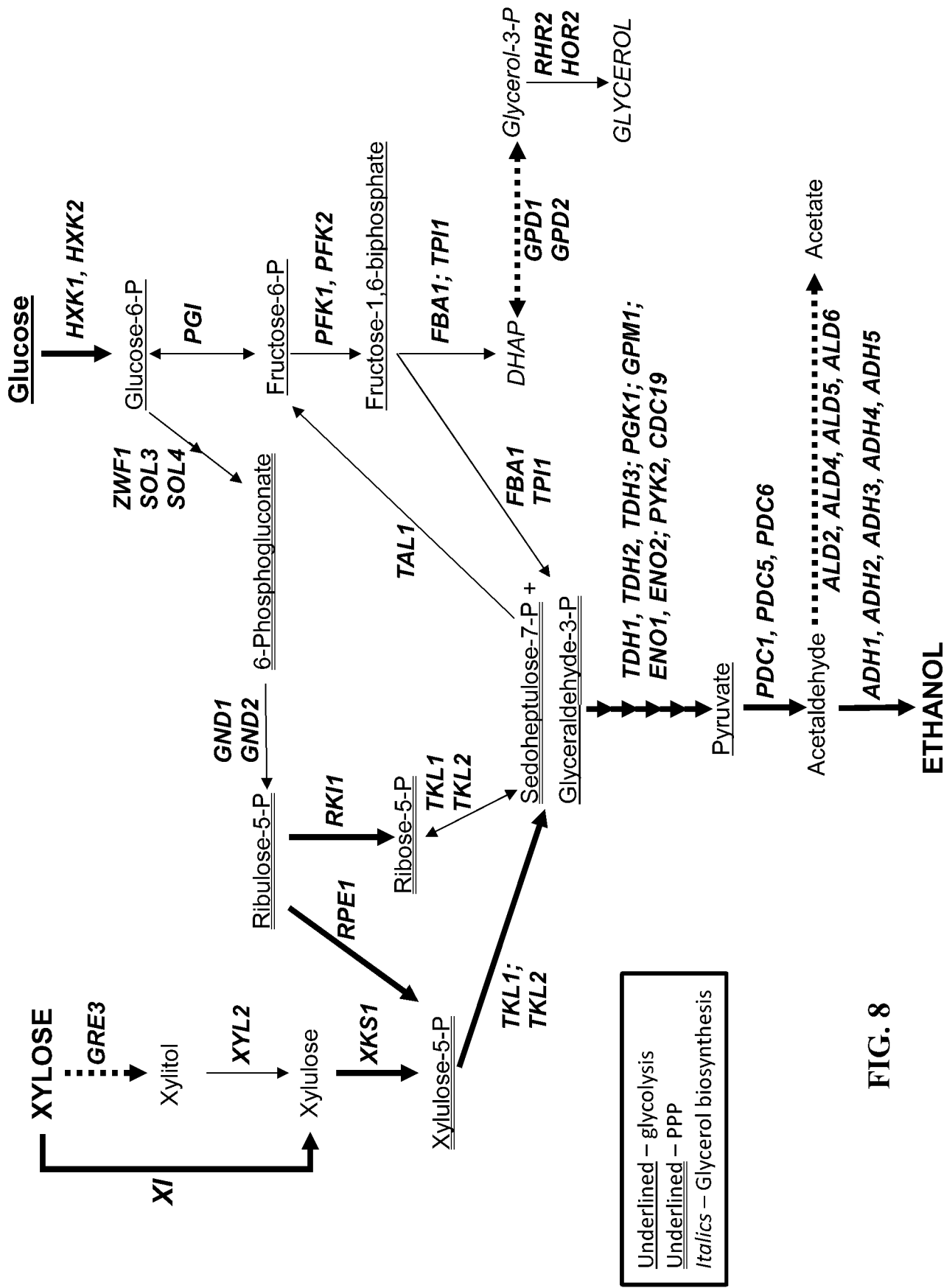


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/051718

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/90 C12P7/10 C12N15/10
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C12P

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

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

EPO-Internal, Sequence Search, WPI Data, BIOSIS, COMPENDEX, EMBASE, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102 174 549 A (UNIV SHANDONG) 7 September 2011 (2011-09-07) the whole document, in particular SEQ ID NOs: 1, 2 and claims -& DATABASE Geneseq [Online] 10 November 2011 (2011-11-10), "Ethanol production related xylose isomerase (XI), SEQ ID 2.", XP002714606, retrieved from EBI accession no. GSP:AZM80432 Database accession no. AZM80432 sequence & US 2012/225452 A1 (BAO XIAOMING [CN] ET AL) 6 September 2012 (2012-09-06) ----- -/--	1-4, 37-82



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

14 October 2013

Date of mailing of the international search report

15/01/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Bassias, Ioannis

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/051718

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/090731 A1 (MARTEK BIOSCIENCES CORP [US]; BAYNE ANNE-CECILE V [US]; LIPPMEIER JAME) 28 July 2011 (2011-07-28) the whole document, in particular SEQ ID NO: 6 (Fig. 11) and claims -----	37-40, 42-46, 48-56, 70-75, 77,79,82
X	WO 2010/074577 A1 (ROYAL NEDALCO B V [NL]; TEUNISSEN ALOYSIUS WILHELMUS RUDOLPHUS HUBERTU) 1 July 2010 (2010-07-01) the whole document -& DATABASE EMBL [Online] 3 August 2010 (2010-08-03), "Sequence 18 from Patent W02010074577.", XP002714607, retrieved from EBI accession no. EMBL:HD069893 Database accession no. HD069893 sequence -& DATABASE EMBL [Online] 3 August 2010 (2010-08-03), "Sequence 19 from Patent W02010074577.", XP002714608, retrieved from EBI accession no. EM PAT:HD069894 Database accession no. HD069894 sequence -& DATABASE EMBL [Online] 3 August 2010 (2010-08-03), "Sequence 20 from Patent W02010074577.", XP002714609, retrieved from EBI accession no. EM PAT:HD069895 Database accession no. HD069895 sequence -----	37-82
X	WO 03/062430 A1 (ROYAL NEDALCO B V [NL]; OP DEN CAMP HUBERTUS JOHANNES [NL]; HARHANGI H) 31 July 2003 (2003-07-31) cited in the application the whole document -/--	7,37-70, 72-82

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/051718

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	-& KUYPER M ET AL: "High-level functional expression of a fungal xylose isomerase: The key to efficient ethanolic fermentation of xylose by <i>Saccharomyces cerevisiae</i>", FEMS YEAST RESEARCH, WILEY-BLACKWELL PUBLISHING LTD, GB, NL, vol. 4, no. 1, 1 October 2003 (2003-10-01), pages 69-78, XP002312913, ISSN: 1567-1356, DOI: 10.1016/S1567-1356(03)00141-7 -----	
X	HARHANGI H R ET AL: "Xylose metabolism in the anaerobic fungus <i>Piromyces</i> sp. strain E2 follows the bacterial pathway", ARCHIVES OF MICROBIOLOGY, SPRINGER, DE, vol. 180, no. 2, 13 June 2003 (2003-06-13), pages 134-141, XP002400714, ISSN: 0302-8933, DOI: 10.1007/S00203-003-0565-0 the whole document -----	37,38, 43,44
X	US 2012/142067 A1 (DESFOUGERES THOMAS [FR] ET AL) 7 June 2012 (2012-06-07) the whole document -----	43,44, 49-52, 70,73
X	WO 2009/109634 A1 (DSM IP ASSETS BV [NL]; KLAASSEN PAUL [NL]; LAAN VAN DER JAN METSKE [NL]) 11 September 2009 (2009-09-11) the whole document -----	43-47, 49-52
A	DATABASE UniProt [Online] 23 March 2010 (2010-03-23), "RecName: Full=Xylose isomerase; EC=5.3.1.5;", XP002714610, retrieved from EBI accession no. UNIPROT:D3I1K1 Database accession no. D3I1K1 sequence -----	1-82
A	DATABASE UniProt [Online] 19 October 2011 (2011-10-19), "RecName: Full=Xylose isomerase; EC=5.3.1.5;", XP002714611, retrieved from EBI accession no. UNIPROT:F9D530 Database accession no. F9D530 sequence -----	1-82
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/051718

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	RONALD E HECTOR ET AL: "Growth and fermentation of D-xylose by Saccharomyces cerevisiae expressing a novel D-xylose isomerase originating from the bacterium Prevotella ruminicola TC2-24", BIOTECHNOLOGY FOR BIOFUELS, BIOMED CENTRAL LTD, GB, vol. 6, no. 1, 30 May 2013 (2013-05-30), page 84, XP021152661, ISSN: 1754-6834, DOI: 10.1186/1754-6834-6-84 -----	1-82

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/051718

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-4(completely); 37-82(partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-4(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 97% sequence identity to a amino acids 2-377 of SEQ ID NO: 78, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 78 or a nucleic acid having at least 40% identity to SEQ ID NO: 77, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

2. claims: 5-8(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 80% sequence identity to a amino acids 2-377 of SEQ ID NO: 96, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 96 or a nucleic acid having at least 40% identity to SEQ ID NO: 95, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

3. claims: 9-11(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 80% sequence identity to a amino acids 2-374 of SEQ ID NO: 38, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 38 or a nucleic acid having at least 40% identity to SEQ ID NO: 37, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

4. claims: 12, 13(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 80% sequence identity to a amino acids 2-374 of SEQ ID NO: 2, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 2 or a nucleic acid having at least 40% identity to SEQ ID NO: 1, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

containing xylose under conditions in which the fermentation product is expressed.

5. claims: 14, 15(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 93% sequence identity to a amino acids 2-376 of SEQ ID NO: 58, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 58 or a nucleic acid having at least 40% identity to SEQ ID NO: 57, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

6. claims: 16, 17(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 80% sequence identity to a amino acids 2-375 of SEQ ID NO: 42, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 42 or a nucleic acid having at least 40% identity to SEQ ID NO: 41, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

7. claims: 18-21(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 97% sequence identity to a amino acids 2-376 of SEQ ID NO: 84, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 84 or a nucleic acid having at least 40% identity to SEQ ID NO: 83, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

8. claims: 22-25(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 97% sequence identity to a amino acids 2-377 of SEQ ID NO: 80, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 80 or a nucleic acid having at least 40% identity to SEQ ID NO: 79, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

9. claims: 26, 27(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 93% sequence identity to a amino acids 2-376 of SEQ ID NO: 54, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 54 or a nucleic acid having at least 40% identity to SEQ ID NO: 53, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

10. claims: 28-30(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 80% sequence identity to a amino acids 2-376 of SEQ ID NO: 46, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 46 or a nucleic acid having at least 40% identity to SEQ ID NO: 45, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

11. claims: 31, 32(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 90% sequence identity to a amino acids 2-376 of SEQ ID NO: 16, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 16 or a nucleic acid having at least 40% identity to SEQ ID NO: 15, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

12. claims: 33, 34(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 85% sequence identity to a amino acids 2-375 of SEQ ID NO: 82, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 82 or a nucleic acid having at least 40% identity to SEQ ID NO: 81, vectors and recombinant host cells comprising

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

13. claims: 35, 36(completely); 37-82(partially)

A recombinant polypeptide (xylose isomerase) comprising an amino acid sequence with at least 90% sequence identity to a amino acids 2-377 of SEQ ID NO: 32, the nucleic acid encoding a polypeptide with at least 70% identity to SEQ ID NO: 32 or a nucleic acid having at least 40% identity to SEQ ID NO: 31, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

14-88. claims: 37-82(partially)

A nucleic encoding a polypeptide with at least 70% identity to SEQ ID NOs: 4, 6, 8, 10, 12, 14, 18, 20, 22, 24, 26, 28, 30, 34, 36, 40, 44, 48, 50, 52, 56, 60, 62, 62, 66, 68, 70, 72, 74, 76, 86, 88, 90, 92, 94, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, or a nucleic acid having at least 40% identity to SEQ ID NOs: 3, 5, 7, 9, 11, 13, 17, 19, 21, 23, 25, 27, 29, 33, 35, 39, 43, 47, 49, 51, 55, 59, 61, 63, 65, 67, 69, 71, 73, 75, 85, 87, 89, 91,, 93, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, vectors and recombinant host cells comprising said nucleic acid and a method for producing a fermentation product (e.g. ethanol) by culturing said recombinant host cells in medium containing xylose under conditions in which the fermentation product is expressed.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/051718

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CN 102174549 A	07-09-2011	CN 102174549 A	07-09-2011
		EP 2679686 A1	01-01-2014
		US 2012225452 A1	06-09-2012
		WO 2012113120 A1	30-08-2012
WO 2011090731 A1	28-07-2011	AU 2010343095 A1	09-08-2012
		CA 2785971 A1	28-07-2011
		CN 102822343 A	12-12-2012
		EA 201290567 A1	28-02-2013
		EP 2519635 A1	07-11-2012
		JP 2013515503 A	09-05-2013
		US 2011189228 A1	04-08-2011
		WO 2011090731 A1	28-07-2011
WO 2010074577 A1	01-07-2010	CA 2746818 A1	01-07-2010
		CN 102307987 A	04-01-2012
		EP 2367928 A1	28-09-2011
		JP 2012513753 A	21-06-2012
		US 2011318790 A1	29-12-2011
		WO 2010074577 A1	01-07-2010
WO 03062430 A1	31-07-2003	AT 418616 T	15-01-2009
		BR 0306740 A	28-12-2004
		CA 2474033 A1	31-07-2003
		CN 1703514 A	30-11-2005
		CY 1108905 T1	04-09-2013
		DK 1468093 T3	16-02-2009
		EP 1468093 A1	20-10-2004
		ES 2319757 T3	12-05-2009
		JP 4334352 B2	30-09-2009
		JP 5113800 B2	09-01-2013
		JP 2005514951 A	26-05-2005
		JP 2009213481 A	24-09-2009
		JP 2012231794 A	29-11-2012
		PT 1468093 E	15-04-2009
		SI 1468093 T1	30-04-2009
		US 2008014620 A1	17-01-2008
		US 2010035306 A1	11-02-2010
		US 2012064607 A1	15-03-2012
		US 2013084617 A1	04-04-2013
		WO 03062430 A1	31-07-2003
US 2012142067 A1	07-06-2012	AR 084087 A1	17-04-2013
		AU 2011334846 A1	13-06-2013
		CA 2819323 A1	07-06-2012
		EP 2646544 A1	09-10-2013
		FR 2968313 A1	08-06-2012
		US 2012142067 A1	07-06-2012
		WO 2012072793 A1	07-06-2012
WO 2009109634 A1	11-09-2009	AU 2009221105 A1	11-09-2009
		CA 2715147 A1	11-09-2009
		CN 101965400 A	02-02-2011
		EA 201001436 A1	28-02-2011
		EP 2250263 A1	17-11-2010
		JP 2011512833 A	28-04-2011
		US 2011189728 A1	04-08-2011
		WO 2009109634 A1	11-09-2009

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/051718

Patent document cited in search report	Publication date	Patent family member(s)	Publication date