



Office de la Propriété  
Intellectuelle  
du Canada

Un organisme  
d'Industrie Canada

Canadian  
Intellectual Property  
Office

An agency of  
Industry Canada

CA 2879680 A1 2014/01/30

(21) **2 879 680**

(12) **DEMANDE DE BREVET CANADIEN  
CANADIAN PATENT APPLICATION**

(13) **A1**

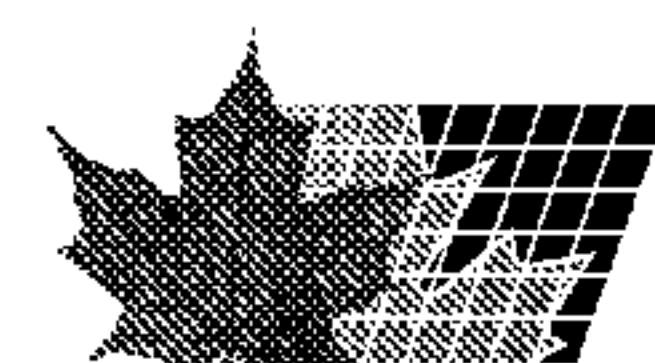
(86) Date de dépôt PCT/PCT Filing Date: 2013/07/23  
(87) Date publication PCT/PCT Publication Date: 2014/01/30  
(85) Entrée phase nationale/National Entry: 2015/01/16  
(86) N° demande PCT/PCT Application No.: US 2013/051718  
(87) N° publication PCT/PCT Publication No.: 2014/018552  
(30) Priorité/Priority: 2012/07/24 (US61/675,241)

(51) Cl.Int./Int.Cl. *C12N 9/90* (2006.01),  
*C12N 15/10* (2006.01), *C12P 7/10* (2006.01)  
(71) Demandeur/Applicant:  
BP CORPORATION NORTH AMERICA INC., US  
(72) Inventeurs/Inventors:  
NUNN, DAVID NEAL, JR., US;  
LUGINBUHL, PETER, US;  
LI, LING, US;  
BURJA, ADAM MARTIN, US;  
BARTNEK, JON PETER FLASH, US  
(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : XYLOSE ISOMERASES ET LEURS UTILISATIONS  
(54) Title: XYLOSE ISOMERASES AND THEIR USES

(57) **Abrégé/Abstract:**

This disclosure relates to novel xylose isomerases and their uses, particularly in fermentation processes that employ xylose-containing media.



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

(19) World Intellectual Property  
Organization  
International Bureau



WIPO | PCT



(10) International Publication Number

WO 2014/018552 A1

(43) International Publication Date  
30 January 2014 (30.01.2014)

## (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 a 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 (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).

**[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*, *Kllockera*, *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  $\beta$ -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- $\Delta$ 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, M<sup>5</sup>, N<sup>-4</sup>, 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 entirety, promoters that are constitutively active in mammalian cells (which can be 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 entirety, promoters that are constitutively active in plant cells (which can be 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  $\beta$ -glucosidase promoter. A particularly suitable promoter can be, for example, a *T. reesei* cellobiohydrolase, endoglucanase, or  $\beta$ -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 $\alpha$ 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 aldo-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  $\beta$ -

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  $\beta$ -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<sup>+</sup>. 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  $\mu$ 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 MgCl<sub>2</sub>, 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 MgCl<sub>2</sub>, 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  $\mu$ l of assay mix into each well of the plates. About 10  $\mu$ l of each XI sample was added to the wells. The reactions were started by the addition of about 100  $\mu$ 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  $\mu$ mole-NADH/min (or Units). If protein quantitation was performed, specific activities (SA) were calculated where the units for SA are ( $\mu$ mole NADH<sup>+</sup>/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*<sup>-</sup>, 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 *xylA* 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 *xylA* deletion plasmid. A fragment of DNA containing the flanking regions of the *xylA* gene (0.7 kb of sequence 5’ and 0.9 kb of sequence 3’ of *xylA*) 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*xylA* (FIG. 1A).

[0090] The pMEV3-Δ*xylA* plasmid was then transformed into strain *E. coli* strain SEL700 (MG1655 Δ<sup>r</sup>, Δ(*recA-srl*)306,*srl*-301::Tn10-84(Tets), [F’ *proAB*, *lacI*<sup>r</sup>, *ZAM15*, Tn10 (Tet<sup>r</sup>)] 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*xylA* 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 *xylA* 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 *xylA* 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 *xylA*, the resulting strain was cured of pJC859 by the following method: The *xylA* 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  $\Delta xylA$ , was used for the ABD screening.

[0092] The ABD screening method was verified by creating a positive control strain by PCR amplification of the *xylA* gene from *E. coli* K12 and cloning into the PCR-BluntII TOPO vector (Invitrogen, Carlsbad, California) using standard procedures. This vector (PCR-BluntII-TOPO-*xylA*, FIG. 1B) was then transformed into the screening strain (SEL700  $\Delta xylA$ ). 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  $\Delta xylA$  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          | ATGGCAGTTAAAGAATATTTCCCGGAGATAGGCAAGATCGCCTTTGAAGGAAAGGAGTCC<br>AAGAACCCTATGGCATTCACACTACTACAATCCAGAGCAGGTAGTAGCCGGAAGAAAAATG<br>AAAGATTGGTTCAAAGTTCGCTATGGCATGGTGGCACACCCCTCTGCGCTGAAGGTGGCGAC<br>CAGTTCGGTCCTGGTACCAAGAAATTCCTTTGGAACACAGGTGCAACTGCACTCGAAAGA<br>GCAAAGAACAAAAATGGACGCAGGTTTCGAGATCATGAGCAAGCTCGGTATCGAGTATTTCTC<br>TGCTTCCACGATGTTGACCTTATCGACGAGGCTGACACTGTTGAAGAGTACGAGGCTAAC<br>ATGAAGGCTATCACAGCTTACGCAAAAGGAGAAAAATGGCCGCTACTGGCATCAAACTCCTC<br>TGGGAAACAGCCAAATGTATTGGGCAACAAGAGATATATGAACGGCGCTTCTACCAACCCT<br>GACTTCAACGTGGCTGCACGGCTATGCTCCAGATCAAGAACGCTATCGACGCAACTATC<br>GCTCTCGGTGGTGACTGCTATGTATTCTGGGGCGCGTGAAGGTTACATGAGCCTTCTCTC<br>AACACCGATATGAAGAGAGAGAAAGACACATGGCTACCATGCTTACCATGGCACGCGAC<br>TATGCTCGTTCTAAGGGCTTCAAGGGTACCTTCCTTATCGAGCCTAAGCCAATGGAGCCG<br>ATGAAGCACCAAGTACGATGTCGATACTGAGACTGTGCTAGGTTTCCTCCGCGCCCATGGT<br>CTTGACAAAGGACTTCAAGGTAACACATCGAGGTTAACCACGCTACTCTCGCAGGCCACACC<br>TTCGAGCACGAGCTCCAGTGCGCCGTTGACGCAGGCATGCTCGGAAGCATCGACGCCAAC<br>CGTGGTGACTACCAAGACGGCTGGGATACCGACCAGTTCCTTATCGACCTCTATGAGCTC<br>GTACAGGCTATGATGGTTATCATCAAGGGCGCGGCTCGTCCGCGGTACCAACTTCGAC<br>GCCAAGACCCGTCGTAACTCAACAGACCTCGAGGATATCTTCATCGCTCATGTATCCGGC<br>ATGGATGTCATGGCACGGCTCTCCTCATCGCTGTGACCTTCTCGAGAAAATCTCCTATT<br>CCTGCAATGGTCAAGGAGCGTTACGCTTCTCTACGACTCAGGCATGGCAAGGACTTCGAG<br>AACGGCAAAGCTTACTCTCGAGCAGGTTGTGATTTTGCAGAAAAAGACGGCGAGCCCTAAG<br>AGCACGCGGAAAGCAGGAGCTCTACGAGTCTATCGTCAATCTCTACATCTAA |
| 1754MI2_001 | <i>Bacteroidales</i> | Amino Acid       | 2          | MAVKEYFPEIGKIAFEGKESKNPMAFHYYNPEQVVAGKKMKDWFKFAMAWWHTLCAEGGD<br>QFGPGTKKFPNWTGATALERAKNKMMDAGFEIMSKLGIEYFCFHDVDLIDEADTVEEYEAN<br>MKALITAYAKEKMAATGIKLLWGTANVFGNKRYMNGASTNPDFNVAARAMLQIKNAIDATI<br>ALGGDCYVFWGGREGYMSLNTDMKREKEHMA TMLTMARDYARSKGFKGTFLIEPKPMEP<br>MKHQYDVDTETVVGF LRAHGLDKDFKVNIEVNHATLAGHTFEHELQCAVDAGMLGSI DAN<br>RGDYQNGWDTDQFPIDLYELVQAMMVIIKGGGLVGGTNFDAKTRNSTDLEDIFIAHVSG<br>MDVMARALLIAADLLEKSPIPAMVKERYASYDSGMGKDFENGKLTLEQVVDFARKNGEPK<br>STSGKQELYESI VNL YI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 5586MI6_004 | <i>Bacteroidales</i> | DNA              | 3          | ATGGCAAAACAAAGAGTACTTCCCGGAGATCGGGAAAAATCAAATTCGAAGGCAAGGATTCC<br>AAGAACCCGCTTGCAATTCATTATTACAATCCTGAGCAGGTGTCCTGCGGCAAGCCGATG<br>AAGGACTGGCTCAAAGTTCGCTATGGCATGGTGGCACACCCCTCTGCGCAGAGGGTAGCGAC<br>CAGTTCGGCGGACCCACCAAAGTCATTCCCTTGGAACAAAAGCTTCGGATCCCATCGCAAAG<br>GCCAAGCAGAAAAGTCGACGCCGTTTCGAGATCATGCAGAAGCTCGGTATCGGATACTAT<br>TGCTTCCACGATGTAGACCTCATCGACGAGCCCGCCACCATCGAGGAGTATGAGGCCGAT<br>CTCAAGGAGATCGTCGCTTACCTCAAGGAGAAAGCAGGCCCCAGACCGGCATCAAGCTCCTT<br>TGGGGCACCGCCAAACGTCTTCGGTCACAAGCGGTACATGAACGGCGCCTCCACCAAACCTT<br>GATTTCGACGTCGCAGCCCGCGCCATGGTCCAGATCAAGAAACGCCATGGACGCCACCATC<br>GAGCTCGGCGCGAGTGCTATGTCTTCTGGGGCGCGCGAGGGCTACATGAGCCTCCTC<br>AACACCGACATGAAGCGGTGAGAAGCAGCATATGGCCACCATGCTCGGCATGGCCCCGGAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| TABLE 2     |                      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------------|----------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism    | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|             |                      |                  |            | TATGACGGGCAAGGGCTTCAAGGGCACCTTCCTCATCGAGCCCCAAGCCCATGGAGCCG<br>ACCAAGCACCATGATGACGTCGACACCGAGACCGTCATCGGTTTCCTCCGTGCCAACGGT<br>CTTGACAAAGGACTTCAAGGTCAACATCGAGGTCAATCACGCCACCCCTCGCCGGCCACACC<br>TTCGAGCATGAGCTCCAGTGGCCGCCGATGCCGGTCTCCTCGGATCCATCGACGCCAAC<br>CGCGCGACTATCAGAACGGCTGGGATACCGACCAAGTTCCTCGGATCGACCTCTATGAGCTC<br>ACCCAGGCCATGATGGTCATCCTCAAGAATGGCGCCTCGTCGGCGGTACCAACTTCGAC<br>GCCAAGACCCGTCGCAACTCCACCGACCTGGACGACATCATCGCCCCACGTCAGCGGT<br>ATGGACATCATGGACGGCACTCCTCGTCGTCGGGACGTCCTCACCAAGTCCGAGCTT<br>CCCAAGATGCTCAAGGAGCGTTACGCTTCCTTCGACTCCGGCAAGGCAAGGAGTTCGAA<br>GAGGCCAAGCTCACTCTCGAGCAGGTCGTAGAGTACGCCAAGACCAAGGGCGAGCCCCAAG<br>GCCACCAGCGGCAAGCAGGAGCTCTACGAGACCATCGTCAACATGTACATCTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 5586MI6_004 | <i>Bacteroidales</i> | Amino Acid       | 4          | MANKEYFPEIGKIKFEGKDSKNPLAFHYYNPEQVVCCKPMKDWLKFAMAWHHTLCAEGSD<br>QFEGPTKSFPPNKA <sup>SDPIAKAKQKVDAGFEIMQKLGIGYYCFHVDVLDLIDEPATIEEY</sup> EAD<br><u>LKEIVAYLKEKQAQTGIKLLWGTANVF</u> GHKRYMNGASTNP <sup>DFDVAARAMVQIKNAMDATI</sup><br><u>ELGGEYVFWGGREGYMSL</u> INTDMKREKQH <sup>MATMLGMARDYARGKFGKGTFLIEPKPMEP</sup><br>TKHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELQCAADAGLLGSIDAN<br>RGDYQNGWDTDQFPIDLYELTQAMMVILKNGGLVGGTNEDAKTRRNSTDLDIIIAHVS<br>G<br>MDIMARALLVAADVLT <sup>KSEL</sup> PKMLKERYAS <sup>FD</sup> SGKGKEFE <sup>E</sup> GKLTLEQVVEYAKTKGEPK<br>ATSGKQELYETIVNMYI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5749MI1_003 | <i>Bacteroidales</i> | DNA              | 5          | ATGAATTTTATAAAGCGGAAAAAGAATTCTCCCCGGAATAGGAAAGATTCA <sup>GT</sup> TGAA<br>GGACGCGAGTCAAAGAACCCGATGGCGTTTCATTATTATGACGAAAAACAAGGTGGTGATG<br>GGTAAAAACACTGAAGGATCATCTTCGTTTGGCAATGGCTTACTGGCATA <sup>CGCT</sup> TTGTGCC<br>GAAGGGGCGACCA <sup>GT</sup> TTGGCGGTGGTACGAAAACATTCCCCTGGAATGCTGTCGCCGAC<br>CCGATCAGCCGTGCCAAATATAAGATGGATGCAGCGTTCGAGTTTATGACAAAAATGCAGC<br>ATCCCTTATTACTGTTTCCATGATGTGGACGTGGTGACGAAGCTCCCACGCTGGCTCAG<br>TTTGAAAAAGACCTTCATACGATGGTAGGCCATGCCAAAGGGCTTCAGCAGGCCAACCGGA<br>AAAAAACTGTTATGGTCTACTGCCAACGTGTTCAGCAACAAACGCTATATGAACGGGGCT<br>GCCACTAATCCTGACTTCTGGCCGTGGCTTGTCGGGTACGCAGATCAAGAAATGCCGATC<br>GATGCCGTGATCGCGCTGGACCGTGAAAACTATGTGTTCTGGGGCGGACGTGAAGGATAT<br>ATGGGCTTGCTCAATACCGATATGAAACGCGAAAAAGACCATCTGGCCATGATGCTGACG<br>ATGGCACGCGACTATGGCCGCAAGAACGGTTTCAAAGGTACTTTCCTGATCGAGCCGAAA<br>CCGATGGAACCGACCAAGCATCAATATGATGTCGACTCGGAAACTGTAA <sup>TCGGCT</sup> TCCTA<br>CGTCATTATGGCCTGGATAAAGACTTCGCCCTGAATATCGAAGTAAATCATGCAACCCCTG<br>GCCGGACATACGTTCGAGCACGAATTGCAGGCTGCTGTCGATGCCGGTATGCTGTGCAGT<br>ATCGATGCCAACCGTGGTGACTACCAGAATGGCTGGGATACCGACCAATTCCC <sup>CGAT</sup> GGAC<br>ATCTACGAACTGACTCAGGCTTGGCTGGTCATTCTGCAAGGTGGTGGTCTGACAAACCGGC<br>GGAACGAACTTCGATGCCAAGACCCGCCGCAACTCGACCGACCTGGACGATATCTTCCTG<br>GCTCATATAGGTGGTATGGATGCGTTTGCCCGTGCCCTGATCACGGCTGCTGCCATCCTT<br>GAAAACTCCGATTACACGAAGATGCGTGCCGGAACGTTACACCAGCTTCGATGGTGGCGAA<br>GGCAAAGCGTTTGAAGACCGGTAAACTTTCCTCTGGAAGACCTGCGTACGATCGCTCTCCGC<br>GACGGAGAACCGGAAGATGGTCAGCGGCAACACAGGAATTATATGAGATGATTC <sup>CAAT</sup> TTA |

| TABLE 2     |                      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------|----------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism    | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5749MI1_003 | <i>Bacteroidales</i> | Amino Acid       | 6          | TACATATAA<br>MNYFKGEKEFFPGIGKIEFEGRESKNPMAFHYYDENKVVMGKTLKDHLLRFAMAYWHTLCA<br>EGGDQFGGGTKTFPWNAAAADPISRAKYKMDAAFEFMTKCSIPYYCFHDVDVVDEAPTLLAQ<br>FEKDLHTMVGHAKGLQQATGKKLLWSTANVFSNKRYSMNAGAATNPDFSAVACAGTQIKNAI<br>DACIALDGENYVFWGGREGYMGLLNTDMKREKDHLLAMMLTMARDYGRKNGFKGTFLLIEPK<br>PMEPTKHQYDVDSETVIGFLRHYGLDKDFALNIEVNHATLAGHTFEHELQAAVDAGMLCS<br>IDANRGDYQNGWDTDQFPMDIYELTQAWLVILQGGGLTTGGTNFDDAKTRRNSTDLLDDIFL<br>AHIGGMDAFARALITAAAILENS DYTKMRAERYTSFDGGEGKAFEDGKLSLEDLRTIALR<br>DGEPKMVSQKQELYEMILNLYI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5750MI1_003 | <i>Bacteroidales</i> | DNA              | 7          | ATGAATTACTTTAAAGGTGAGAAAAGAGTTCTTCCCGGGAATCGGGAAAAATAGAGTTTGAA<br>GGACGTGAATCGAAGAATCCGATGGCTTTTCATTACTATGACGAGAACAAAGTTGTCTATG<br>GGGAAGACCTTGAAGGACCATCTGCGTTTTGGCATGGCTTATTGGCATACGCTGTGTGCG<br>GAAGGCGCCGACCAAGTTCGGCGCGGACGAAAGGCATTTCCCTGGAATACCGGGGCGGAT<br>CGTATTTCCCGTGCCAAGTATAAGATGGATGCTGCTTTTGAGTTTATGACGAAATGTAAC<br>ATCCCGTACTATTGTTTCCATGATGTGGATGTGGTGAATGAAGCTCCGACACTGGCCGAA<br>TTTGAAAAAAGACTTGCATACGATGGTCGAATATGCCAAGCAGCATCAGGAGGCCAACCGGG<br>AAAAAACTGTTGTGCTACCCGCCAATGTGTTCAGCAATAAACGTTATATGAACGGGGCT<br>GCCACAAATCCGTATTTCCCTGCTGTCGCTTGTCGGGTACGCAGATCAAGAAATGCTATC<br>GACGCTTGATTGCCCTGGCGCGGCAAAACTATGTGTTCTGCGGCGGTCTGTGAAGGGGTAT<br>ATGAGCTTGTTGAACACCAATATGAAACCGCAAGGAAAGAACATCTCGCCATGATGTTGACG<br>ATGGCTCGCGATTATGCGCGTAAGAACGGCTTCAAAGGTACTTTCCTGGTAGAGCCATAA<br>CCGATGGAACCGACCAACACATCAGTATGATGTGGACACAGAAACTGTTATCGGCTTCCTG<br>CGTCATTACGGCCTTGACAAAGGACTTTGCCATCAACATCGAAGTGAATCATGCTACATTG<br>GCTGGACATACATTCGAACATGAGCTTCAGGCGGTGCCGATGCCGTATGCTGTGCAGC<br>ATCGACGCCAACCGCGCGGATTACCAAGATGTTGGGACACGGATCAGTTCCTCGGTCGAC<br>ATCTACGAACTGACACAGCGTGGCTGGTTATCCTCGAAGCGGGTGGCCTGACTACCGGT<br>GGTACGAACTTCGACGCCAAGACGCGCCGCAACTCGACTGACCTGGACGATATCTTCCTG<br>GCACACATCGGTGGTATGGATTGTTTGGCCGTGCTTGTATGGCGGTGCCGATATATTG<br>GAACACTCCGATTACAAAAAAGATGCGTGCCGAAACGTTATGCCAGCTTCGATCAAGGCGAC<br>GGCAAGAAAGTTCGAAGATGTTAAACTCCTTCTCGAGGACCTCCGCACCATCGCTCTTGCC<br>TCCGGCGAACCGAAGCAAAATCAGCGGGGAAACAGGAATTGTATGAAATGATTATCAACCAG<br>TACATTTAA |
| 5750MI1_003 | <i>Bacteroidales</i> | Amino Acid       | 8          | MNYFKGEKEFFPGIGKIEFEGRESKNPMAFHYYDENKVVMGKTLKDHLLRFAMAYWHTLCA<br>EGADQFGGGTKAFPWNNTGADRI SRAKYKMDAAFEFMTKCNIPYYCFHDVDVVDEAPTLLAE<br>FEKDLHTMVEYAKQHQEATGKKLLWSTANVFSNKRYSMNAGAATNPYFPAVACAGTQIKNAI<br>DACIALGGENYVFWGGREGYMSLLNTNMKREKEHLAMMLTMARDYARKNGFKGTFLLVEPK<br>PMEPTKHQYDVDTETVIGFLRHYGLDKDFAINIEVNHATLAGHTFEHELQAAADAGMLCS<br>IDANRGDYQNGWDTDQFPVDIYELTQAWLVILEAGGLTTGGTNFDDAKTRRNSTDLLDDIFL<br>AHIGGMDSFARALMAAADILEHSDYKMKRAERYASFQDGDGKKFFEDGKLLLEDLRTIALA<br>SGEPKQISGKQELYEMIINQYI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| TABLE 2     |                      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|----------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism    | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 5750MI2_003 | <i>Bacteroidales</i> | DNA              | 9          | ATGAATTATTTAAAGGTGAAAAAGAGTTTTTCCCTGGAATCGGGAAAAATAGAGTTTGAA<br>GGACGTGAGTCGAAGAATCCGATGGCTTTTCATTATTATGATGAAAAACAAGGTCGTAATG<br>GGCAAGACCTTGAAAGATCACCTCCGCTTTGCAATGGCTTACTGGCATACGTTGTGCGCG<br>GAAGCGCAGACCAAGTTTGGCGGTGGCACAAAATCATCCCTGGAATACCGCAGCGGAT<br>CGTATTTCCCGCGCTAAATATAAAAATGGATGCTGCTTTTCGAGTTTATGACCAAGTGCAGT<br>ATCCCGTACTATTGTTTCCATGATGTGGACGTGGTGACGAAGCTCCGGCACTGGCCGAA<br>TTTGAAAAGGACCTGCATACGATGGTGGGATTCGCCAAACAACACAGGAAGCAACCGGA<br>AAGAAACTGTTGTGCTACAGCCAAATGTATTCGGGCATAAACGTTATATGAACGGAGCG<br>GCTACCAATCCTTAATTTCCCGGCTGTCGCTTGTCGGGTACGCAGATCAAGAAATGCAATC<br>GACGCTGTATCGAGCTGGTGGAGAGAACTATGTATTTCTGGGCGGACGCGGAAGGCTAC<br>ATGAGCCTGCTGAACACCAATATGAACCGTGAAACGTAAGAGGATCATTTGGCCATGATGCTGACA<br>ATGGACCGGATTATGCCCCGCAAGATGGTTCAAGGGTACTTTCCTGGTGAATCTAAG<br>CCGATGGAACCGACCAAAACATCAGTATGACGCAGATACGGAAACCGTGATCGGCTTCCTG<br>CGCCACTATGGCCTCGACAAAGGATTTCGCTATCAACATTGAAGTGAACCATGCTACATTG<br>GCCGGCCATACATTCGAACATGAACCTTCAGGCTGCTGCCGATGCCGTATGCTGTGCAGC<br>ATCGATGCAAAATAGAGCGGACTATCAGAAATGGTTGGGATACGGATCAGTTCCTCCCGTAGAC<br>ATTTACGAACTGACACAGGCTGGCTGGTTATCCTGGAAGCGGCGGACTGACAAACCGGA<br>GGTACGAACTTCGATCGGAAGACCCGTCGTAACTCGACTGACCTCGACGATATCTTCCTG<br>GCCATATCGGCGGTATGGATTGTTTGACAGTGCCTTGATGGCAGCTGCCGATATCCTG<br>GAACATTCTGATTACAAGAAAGATGCGTGCCGAACGTTACGCCAGCTTCGACCAGGGCGAC<br>GGCAAGAAAGTTCGAAGACGGCAAACTCCTTCTCGAAGACCTCGGCACAATTCCTCTGCC<br>GGCAGCAACCGAAGCAGATCAGCGGCAAGCAGGAGTTGATGAGATGATTATCAATCAG<br>TATATTTAA |
| 5750MI2_003 | <i>Bacteroidales</i> | Amino Acid       | 10         | MNYFKGEKEFFPGIGKIEFEGRESKNPMAFHYYDENKVMGKTLKDHLRFAMAYWHTLCA<br>EGADQFGGGTKSFPWNNTAADRI <del>SR</del> AKYKMDAAFEFMTKCSI <del>P</del> YYCFHDVDVVDEAPALAE<br><u>FEKDLHTMVGF</u> AKQH <del>Q</del> EATGKLL <del>W</del> STANVFGHKRYMNGAATNPYFP <del>V</del> AVACAGTQIKNAI<br>DACIELGGENYVFWGGREGYMSLLNTNMKREK <del>DL</del> AMMLTMARDYARKNGFKGTF <del>LV</del> ESK<br>PMEPTKHQYDADTETVIGFLRHYGLDKDFAINIEVN <del>HA</del> TLAGHTFEHELQ <del>AA</del> ADAGMLCS<br>IDANRGDYQNGWDTDQFPVDIYELTQ <del>AW</del> LVI <del>L</del> EAGGLTTGGTNF <del>DA</del> KTRRNSTD <del>LD</del> DDIFL<br><u>AHIGG</u> MSFARALMAAADILEHSDYK <del>K</del> MRAERYASFDQGDGKKFFEDGKLLLEDLRTIALA<br>GDEPKQISGKQELYEMIINQYI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI5_004 | <i>Bacteroides</i>   | DNA              | 11         | ATGAAACAGTATTTCCCGAACATCTCCGCCATCAAGTTTGAGGGCGTCGAGAGCAAGAAT<br>CCCTGGCTTACCGCTACTACGACCGCGACCGCGTCGTCTGTTAAGTAAGAGATGAGCGAA<br>TGTTTAAAGTTCGCTATGTGCTGGTGGCACACCCCTCTGCGCCGAGGGCTCCGATCAGTTC<br>GGTCCCGGCACAAAAGACCTTCCCTGGAACCGCCCGCCGACCCCGTCGAGGCTGCCAAG<br>GACAAGGCCGACGCTGGCTTCGAGATCATGCAGAAACTCGGCATCGAGTACTGCTTC<br>CAGACGTTGACCTCGTGGCCGAGGCTCCCGACGTGGAGACCTACGAGAAACCTCAAG<br>GAGATCGTGGCTTATCTCAAGCAGAAACAGGCTGAGACGGGCATCAAGCTGCTCTGGGGC<br>ACTGCCAACGTCCTCGGACACAAAGCGCTACATGAACGGAGCCTCCACGAACCCCGACTTC<br>GATGTCGTGGCACGCGCTATCGTGCAGATCAAGAACGCCATCGATGCTACCATCGAGCTG<br>GGCGGCACCAACTACGTCCTTCTGGGGCGGTGCGGAAGGCTACATGAGCCTGCTCAACACC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| TABLE 2       |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|---------------|--------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|               |                    |                  |            | GATATGAAGCGGAGAAGGAGCACATGGCTACGATGTTGACGATGGCACGCGACTATGCCCGTTCTAAGGGATTCAAGGGCACGTTCCCTCATCGAACCCAAACCCCATGGAACCCACGAAGCATCAGTACGATGGGACACCGAGACGGTCATCGGATTCCTCCGTGCTCATGGTCTCGAC AAGGATTTCAAGGTCAACATCGAGGTCAACCAAGCCAGCTGGCCGGACACACGTTTCGAGCATGAGCTGGCCTGCGCCGTAGACGCCGATATGCTCGGCAGCATCGATGCCAATCGCGGC GACTATCAGAACCGGATGGACACCGACCAGTTCCTCCATCGACCACTACGAACTCACGCAG GCTATGCTGCAGATCATCCGCAACGGAGGTTTCAAGGACGGTGGCACCAATTTTGACGCT AAAGCGCCGCCAACAGCACCGACCTCGAGGATATCTTTCATCGCTCACGTAGCAGCCATG GACGCCATGGCCCAAGCCCTGTTGTGGCTGCCGATATCATCGAGAAAGTCGCCCATCTGC ACGATGGTCAAGGAGCGTTACGCCAGCTTCGATGCCGGCAAGGCAAGCGCTTCGAAGAA GGCAAGATGACCCCTCGAGGAAGCCTACGAGTATGGCAAGAAGGTCGGGGAGCCCCAAGCAG ACCAGCGGAAAGCAGGAGCTCTACGAAGCCATTGTCAATATGTATTGA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI5_004   | <i>Bacteroides</i> | Amino Acid       | 12         | MKQYFPNISAIKFEGVESKNPLAYRYYDRDRVVMGKKMSEWKFAMCWWHTLCAEGSDQF GP <sup>GT</sup> KT <sup>FP</sup> WNAAAADPVOAAKDKADAGFEIMQKLGIEYYCFHDVDLVAEAPDVETYEKNLK EIVAYLKQKQAE <sup>TGI</sup> KL <sup>LM</sup> GTANVFGHKRYMNGASTNP <sup>DFD</sup> VVARIVQIKNAIDATIEL GG <sup>TNY</sup> VFWGGREGYMSLLNTDMKREKEHEMATMLTMARDYARSKGFKGT <sup>FL</sup> IEPKPMEPTK HQYDADTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTEHELACAVDADMLGSIDANRG DYQNGWD <sup>TD</sup> QFPIDHYELTQAMLQIIRNGGFKDGGTNF <sup>DA</sup> KTRRNSTDLEDIFIAHVAAM DAMAHALLSAADII <sup>E</sup> KSPIC <sup>TM</sup> VKERYASF <sup>DA</sup> GEGKRFEEGKMTLEEAYEYGKKVGE <sup>PKQ</sup> TSGKQELYEAI <sup>V</sup> NMY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI202_004 | <i>Bacteroides</i> | DNA              | 13         | ATGGCAACAAAAGAGTATTTTCCCGGAATAGGAAAAGATTAAATTCTGAAGGTTAAAGAGAGT ATGAACCCGATGGCATATCGTTACTACGATGCTGAGAAAGGTAATCATGGGTAAGAAAGATG AAAGATTGGTTGAAAGTTTGCTATGGCTTGGTGGCACACTCTCTGCGCAGAAAGGTGGTGAC CAATTGCGTGGCGGAACGAAACAATTCCTTTGGAATGGTGACTCTGACGCTTTTGCAAGCA GCTAAAAATAAATTGGATGCAGGTTTCGAATTCTATGCAGAAGATGGGTATCGAATACTAT TGCTTCCACGATGTAGACCTGATTTCTGAAGGTGCAAGCATCGAAGAATACGAAGCTAAC TTGAAAGCTATCGTAGCTTATGCAAAAAGAAAACAGGCTGAAACTGGTATCAAGCTGTTG TGGGTACTGCTAACGTATTGCGTCATGCACGTTATATGAACGGTGTGCTACCAATCCT GATTTCGACGTTGTAGCACGCGCTGCTGTTCAGATCAAGAACGCTATTGACGCTACTATC GAACTGGGTGGTTCAAAC <sup>TAT</sup> GTATTTCTGGGCGGTCTCGGAAGGTTACATGTC <sup>TT</sup> TGCTG AACACTGACCAGAAACGTGAAAAAGAACACCTTGCAAGATGTTGACTATCGCTCGTGAC TATGCACGTGCTCGTGGCTTCAAAGGTACTTTCCCTGATTGAGCCGAAACCGATGGAAACCG ACAAAACATCAGTATGATGTAGATACTGAAACAGTTATCGGCTTCCTGAAAGCTCACGGT TTGGATAAGGATTTCAAAGTAAACATCGAGGTTAATCACGCAACTTTGGCTGGCCATACT TTCGAACACGAACTGGCTGTAGCTGTTGACAAACGGCATGTTAGGTTCTATCGACGCTAAC CGTGGTGACTACCAAGAACGGTTGGGATACTGACCAATTCCTTATCGATAACTACGAAC <sup>TG</sup> ACTCAAGCTATGATGCAGATCATCCGCAACGGTGGTTTGGGTAAATGGCGGTACTAACTTC GACGCTAAGACCCGTGTAAC <sup>TCT</sup> ACCGACCTGGAAGATATCTTCATCGCTCACATTGCA GGTATGGATGCTATGGCACGTGCTCTGGAAAAGTGCAGCTAAATTA <sup>CT</sup> GGAAGAATCTCCT TATAAGAAAAATGTTGGCTGATCGTTACGCATCATTCGACGGTGGCAAGGTAAGGAATTC GAAGAAGGCAAAATTGTC <sup>TT</sup> TGGAAGATGTTGTAGCTTATCGGAAAGCTAACGGCGAACCG |

| TABLE 2       |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|---------------|--------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5586MI202_004 | <i>Bacteroides</i> | Amino Acid       | 14         | AAGCAAAACGCGCAAGCAAGAATTGTATGAAGCAATCGTGAATATGTATTGCTAA<br>MATKEYFPGIGIKIKFEGKESMNP MAYRYDDAEKVIMGKKMDWLKFAMAWWHTLCAEGGD<br>QFGGGTKQFPWNGSDALQAAKNKLDAGFEFMQMGI EYYCFHVDVLI SEGASIEEYEAN<br>LKAI VAYAKEKQAE TG I KLLWGTANVFGHARYMNGAATNPDPDVVARAAVQIKNAIDATI<br>ELGGSNYVFWGGREGYMSLLNTDQKREKEHLAKMLT IARDYARARGFKGTFLIEPKPMEP<br>TKHQYD VDTETVIGFLKAHGLDKDFKVNIEVNHAT IAGHTFEHELAVAVDNGMLGSI DAN<br>RGDYQNGWDTDQFPIDNYELTQAMMQ IIRNGGLNGGTFNFDAKTRNSTDLEDIFIAHIA<br>GMDAMARALESAAKLLEESPYKKMLADRYASFDGGKGKEFEEGKLSLEDVVAYAKANGEP<br>KQTS GKQELYEAI VNMYC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5586MI211_003 | <i>Bacteroides</i> | DNA              | 15         | ATGGCAAAAGAGTATTTTCTGGCGTGAAAAAAATCCAGTTCGAGGGTAAGGACAGTAAG<br>AATCCAATGGCTTACCGTTATTATGATGCAGAGAAGGTCATCATGGGTAAGAAGATGAAG<br>GATTGGTTGAAGTTCGCTATGGCTTGGTGGCACACTTTGTGCGCTGAGGGCGCAGACCAG<br>TTCGGTGGCGGTACTAAGACTTTCCTTTGGAACGAAGGTGCAAA CGCTTTGGAAGTTGCT<br>AAGAATAAGGCTGATGCTGCTTTCGAGATTATGGAGAAAGCTTGGCATCGAGTACTACTGT<br>TCCACGATGTAGACCTCGTTGAGGAGGCTGCAACTATCGAGGAGTATGAGGCTAACATG<br>AAGCTATCGTTGCTTATCTTAAGGAGAAGCAGGCTGCTACTGGCAAGAAGCTTCTTTGG<br>GGTACTGCTAACGTATTCGGCAACAAGCGCTATATGAACGGTGCTTCTACAAACCCGTGAC<br>TTCGACGTTGTTGCTCGCGCTTGTTGTTCAGATT AAGAA CGCTATCGACGCTACTATCGAA<br>CTTGGTGGTACAAAAC TACGTATTCTGGGTGCGCGAGGGTTATATGAGCCTTCTTTAAC<br>ACAGATATGAAGCGTGAGAAAGGAGCACATGGCAACTATGCTTACTAAGGCTCGCGACTAC<br>GCTCGTTCAAAGGGCTTTACTGGTACATTCTTTATCGAGCCAAAGCCAA TGGAAACCATCA<br>AAGCATCAGTATGATGTTGATACTGAGACTGTTTGTGGTTTCTTGAGGGCTCACGGTCTT<br>GACAAGGACTTCAAGGTAAACATCGAGGTTAAACCAGCTACTTTGGCTGGTCACACATTC<br>GAGCACGAGTTGGCTGCTGCTGTTGATAACGGTATGCTTGGCTCTATCGACGCTAAACCGC<br>GGTGACTAC CAGAAACGGTTGGGATACTGACCACTCCCTATCGACAACTTCGAGCTTATT<br>CAGGCTATGATGCAGATTATCCGCAACGGTGGTCTTGGCAACGGTGGTACAAACTTCGAC<br>GCTAAGACTCGTCGTAACTCAACTGACCTTGAGGATATCTTCATCGCACACATCGCTGGT<br>ATGGATGCAATGGCTCGCGCTCTTGAGAACGCAGACACCTTTTGGAGAACTCTCCAATC<br>AAGAAGATGCTTGTGCTGAGCGTTACGCTTCATTCGACAGCGGCAAGGGTAAGGAGTTTCGAG<br>GAAGGCAAGTTGAGCCTTGGGGACATCGTTGCTTATGCTAAGCAGAAACGGTGAGCCCTAAG<br>CAGACAAGCGGTAAGCAGGAGCTTTACGAGGCTATCGTAAACATGTACTGCTAA |
| 5586MI211_003 | <i>Bacteroides</i> | Amino Acid       | 16         | MAKEYFPGVKKIQFEGKDSKNP MAYRYDDAEKVIMGKKMKDWLKFAMAWWHTLCAEGADQ<br>FGGGTKTFPWNEGANA LEVAKNKADAGFEI MEKLGIEYYCFHVDV DLVEEAA TIEEYEANM<br>KAIVAYLKEKQAA TGKLLWGTANVFGNKRYMNGASTNPDPDVVARACVQIKNAIDATIE<br>LGCTNYVFWGGREGYMSLLNTDMKREKEHMA TMLTKARDYARSKGFTGTFLIEPKPMEPS<br>KHQYD VDTETVCGFLRAHGLDKDFKVNIEVNHAT IAGHTFEHELAAA VDNMGLGSI DANR<br>GDYQNGWDTDQFPIDNFELIQAMMQ IIRNGGLNGGTFNFDAKTRNSTDLEDIFIAH IAG<br>MDAMARALENAADLLEN SP I KKMVAERYASFDSGKGKEFEEGKLSLGDIVAYAKQNGEPK<br>QTS GKQELYEAI VNMYC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5606MI1_005   | <i>Bacteroides</i> | DNA              | 17         | ATGGCGACAAAAGAATACTTTCCTCCGGAATAGGGAAAAATCAAGTTTGAGGGTGTGAATAGC<br>TATAATCCGCTGGCATACAGATAT TACGATGCCGAGCGCATAGTCCTTGGCAAGCCGATG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|--------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|             |                    |                  |            | AAGGAGTGGCTCAAGTTGCCATGGCATGGTGGCACACACTCTGCGCAGAGGGTGGCGAC<br>CAGTTTGGCGCGGTACGAAGAATTTCCTCGAATGGAGATCCCGATCCGGTACAGGCC<br>GCAAAAAACAAAGTAGACCGCGCTTCGAATTCATGACCAAGATGGGAATAGAGTATTC<br>TGTTCCACGACGTGGATCTCGTCAGCGAGGCAGCAACCATCGAGGAGTATGAGGCCAAC<br>CTGAAGGAAGTGGTGGCTACATCAAGGAAAAAGCAGGCCGAGAGCGGGATCAAAAAACCTC<br>TGGGCACTGCCAACGTGTTCAGCCACGCGCGCTACATGAACGGAGCCGCCAACCAACCCC<br>GACTTCGATGTAGTGGCCCCGCGAGCCGTGCAGATCAAGAAIGCTATCGACGCCACGATA<br>GCCTTAGTGGCACTAAGTACGTGTTCTGGGTGGCGTGAAGTTACATGAGCCTGCTC<br>AACCCGACAGAAAGCGGAGAGGAGCATCTGGCAATGATGCTCCGATGGCCCGCGAC<br>TATGCGGTGCAAAAAGGCTTCACCGGCACCTTCCTTATCGAGCCCAGCGCGATGGAGCCC<br>ACCAAGCACCCAGTATGATGTAGACACCGAGACTGTGATAGGCTTCCTCCGTGCCACGGC<br>CTCGACAAAGGACTTCAAGGTCAACATAGAGTGAACCAACGCCACCTGGCCGGCCATACC<br>TTCAGCATGAGCTGGCAGTGGCCGTGGACAAACGGTATGCTCGGCAGCATCGACGCCAAC<br>CGGGTGACTACCAAGACGGCTGGGATACCGACCACTCCCATCGACAACTACGAGCTG<br>ACCCAGGCCATGATGCAGATAATACGCAACGGCGGCTTCGGCAACGGCGGATGCAACTTC<br>GACGCCAAGACACGCCCAACTCCACCGACCTGGAGGATATCTTCATAGCCACATAGCA<br>GGCATGGACGCCATGGCCCCGCCCTGCTCAGCGCAGCAGAAGTGTGGAGAAATCGCCC<br>TACAGGAAGATGCTCGCCGAGCGCTACGCACCCGTTTGATCGCGGCCAGGAAAGGCAATT<br>GAAGAGGGCGCAATGTGCTCACCGACCTTGTGGAGTATGCCAAGGAGCATGGCGAGCCC<br>ACACAGACTTCGGCAAGCAGGAACCTCTATGAGGCAATCGTCAATATGTATTGCTAA<br>MATKEYFPGIGIKIFEGVNSYNPLAYRYDAERIVLGKPMKEWLKFAMAWHHTLCAEGGD<br>QFGGKTNFPWNGDDPVPQAAKNKVDAGFEFMTKMGIETYCFHDVDLVSEAAIEEYEAN<br>LKEVVGYIKEKQAE TG IKNLWGTANVFSHARYMNGAATNPDFDVVARAAVQIKNAIDATI<br>ALGGTNYVFWGREGYMSLLNTDQKREKEHLAMLMRMARDYARAKGFTGTFLIEPKPMEP<br>TKHQYDVDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSIDAN<br>RGDYQNGWDTDQFFIDNYELTQAMMQIIRNGGFNGGCNFDKTRRNSTDLEDIFIAHIA<br>GMDAMARALLSAAEVLEKSPYRKMLAERYAPFDAGQGKAFEEGAMSLTDLVEYAKEHGE<br>P<br>TQTS GKQELYEAI VNM YC |
| 5606MI1_005 | <i>Bacteroides</i> | Amino Acid       | 18         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5606MI2_003 | <i>Bacteroides</i> | DNA              | 19         | ATGGCAACAAAGGAATATTTCCCCATATAGGGAAGATCCAGTTCAAAGGCACGGAATCG<br>TACGATCCGATGTCGTATCGTTACTATGACGCCGAGCGCGTAGTTCTGGGCAAGCCCCATG<br>AAGGAATGGCTGAAATTCGCCATGGCATGGTGGCACACATGTGCGCCGAGGGCGGCAC<br>CAGTTCGGCGCGGAACGAAAGTTCCCTTGGAACGAGGGCGGAGCGCCATGACCATC<br>GCCAAGCAGAAGGCTGACGCCGGCTTCGAGATCATGCAAGAAGCTCGGCATCGAGTATTC<br>TGCTTCCACGACATCGACCTGATCGCGACCTGGCGACGACATCGAGGACTATGAGAAC<br>CGTATGCACGAAATCACCCACACCTGAAGGAGAAGATGGCCGCCACGGGCATCAAGAAC<br>CTGTGGGCACTGCCAACGTGTTCGGCCACGCACGCTATATGAACGGCGCCGCCAAC<br>CCGACTTCGACGTTGTGGCAGCGCATGTGCGAGATCAAGAACGCCATCGACGCCACC<br>ATCGCTCTAGCGGTACAAACTATGTATTTCTGGGCGCGCGGAGGGCTACATGAGCCTG<br>CTGAACACCCGACCAAGCGCGAGAAAGAGCACCTGGCTACCATGCTGACCATGGCACGC<br>GACTATGCCCGCGCAATGGCTTCACCGGAACGTTCCCTGATCGAGCCCAACCCCATGGAG<br>CCCAGCAAGCATCAGTATGATGTGGATACCGAGACCGGTAATCGGCTTCCTGAAGGCCAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------|--------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|             |                    |                  |            | AACCTGGACAAGGACTTCAAGGTGAACATCGAGGTGAACCATGCCACTCTGGCCGGCCAC<br>ACATTCGAGCATGAGCTGGCAGTAGCCGTGGACAACGGCATGCTGGCAGCATCGACGCC<br>AACCGCGGCGACTATCAGAAACGGCTGGGACACCGACCAGTTCCCCATCGACAACATATGAG<br>CTGACCCAGGCCATGATGCAGATAATCCGCAACGGTGGCCTCGGCAACGGCGGTACCAAC<br>TTCGACGCCAAGACACGTGCAACTCCACCGACCTGGACGACATCTTCATCGCTCACATC<br>GCCGGTATGGACCGCTATGCCCGCGCTCCGCTCAGCGAGCCGACGTGCTTGAGAAAGTCG<br>CCTTACAAAGAAGATGCTGGCCGACCGCTACGCTTCATTCGACAGCGGCGAGGCAAGAAG<br>TTCGAGGAAGGCAAGATGACTCTGGAGGATGTCGTGGCTACGCCAAGAAATCCCGAA<br>CCGCTCAGACCAGCGGCAAGCAGGAACTCTACGAGGCCATCATCAACATGTACGCCCTGA<br>MATKEYFPHIGKIQFKGTEsyDPMsYRYyDAERVVlGKPMKEWLKFAMAWWHtLCAEGGD<br>QFGGGTKKFPWNEGEDAMtIAKQKADAGFEIMQKLGIEYFCFHdIDlGLDdDIEDYEN<br>RMHEI tAHLKEKMAATGIKNLWGTANVFGHARyMNGAATNPdFVVARACVQIKNAIDAT<br>IALGGTNyVFWGGREGYMSLLNTDQKREKEHLATMLTMARDYARANGFTGTFLIEPKPME<br>PSKHQYDvDTETvIGFLKAHNLdKDFKVNIEVNhATLAGHTfEHElAVAVdNGMLGSIDA<br>NRGDYQNGWdTDQfPIDNyELtQAMMQiIRNGGLNGGTnFDAKTRRNStDLDDiFfIAHI<br>AGMDAMARAPLSAADvLEKSPYKkMLADRYASfDSGEGKkFEEGKMTLEDvVAYAKKNPE<br>PAQTSgKQELyEAIINMYA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 5606MI2_003 | <i>Bacteroides</i> | Amino Acid       | 20         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5610MI3_003 | <i>Bacteroides</i> | DNA              | 21         | ATGGCAACAAAAGAATTTTtCCCGAGATtGGTAAAAATCAAGTTTGAGGGCCGCGAAAGC<br>CGCAATCCcCTCGCATtCCGCTACTACGGCCcCGAGAAAAGtCGTtCTtTGGCAAGAAAGATG<br>AAAGACTGGTtCAAGTTTGCGATGGCTTGtTGGCACACACtGTGCGCCcCAGGGCACCCGAC<br>CAGTTTGtTGGCGCACACCAAGCAGTTTCCGTGGAACACtGCCAGTGACCCCATGCAGGCC<br>GCCAAGGATAAGGTGGATGCCGGATTTGAATTCATGACCAAGATGGGCATtGAGTACTtC<br>TGCTtCCACGATGTGGATCTCGTCGCCGAGGGCCcCAGtGTcGAGGAGTATGAGGCTAAc<br>CTCAAGACCATCGTCGCCTACATCAAAAGAGAAACAAGCCGAGACCGGCATCAAGAAACCTG<br>TGGGCACAGCCAAcGTATTCGGACACAAAACGTACATGAACGGTGCCGCcCACCAACCCc<br>GACTtTGATGTcGTGGCACCGCCCATCGTGCAATCAAGAACCCCATCGACGCCACCATC<br>GAGTTGGCGGCACGAGTTACGTCTtTTGGGGCGCGCGAGGGCCACATGAGCCtGCTC<br>AACCCGACCAGAAAGCGGAGAGGAGCACCTTGcACGCATGCTGACCATGGCACCGGAC<br>TATGCCCGCGCACGTGGTTTCAACGGCACCTTcCTCATCGAGCCCAAGCCCATGGAGCCG<br>ACCAAGCACCAATATGATGTGGACACCGAGACCGTCATCGGTtTCTGCGTGCCCATGGT<br>CTGGACAAAGGACTtCAAGGTCAACATCGAGGTGAACcCAGCTACACTGGCCGGACACACc<br>TTCGAGCGCGAACTGGCAGTGCCCGTCGACAAACGGTCTACTCGGTCAATCGACGCCCAc<br>CGTGGTGACTATCAGAATGGTTGGACACCGATCAGTTCCCCATCGACCACTATGAGTTG<br>GTTcAGGGCATGTGCAGATtATCCGCAATGGTGGTTTCACCGACGGTGGCACCAACTtC<br>GATGCCAAAGACCCCGCAACTCGACCGACCTCGAGGACATCTTCATCGCCcCACATCGCC<br>GCGATGGATGCCATGGCTCATGCGCTGGAGAGTGtGCCTCCATCATCGAGGAGTCGCCc<br>TACTGCCAGATGGTCAAGGATCGCTATGCCTCATtTGACTCCGGCATCGGCAAGGACTtT<br>GAGGACGGCAAGTTGACACTGGAACAAGCCtACGAGTACGGTAAGCAAGTGGCGGAACCC<br>AAGCAGACcCAGTGGCAAGCAAGAACtGTACGAGTCAATCATCAATATGtATtCCATtTAA<br>MATKEFFPEIGIKfEGRESRNPlAFRYyGPEKvVLGKMKDwFKFAMAWWHtLCAQGTD<br>QFGGDtKQFPWNTASDPMQAAKDKVDAGFEfMTKMGIeYFCFHdVDLVAEAAtVEEYEA |
| 5610MI3_003 | <i>Bacteroides</i> | Amino Acid       | 22         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------|--------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|             |                    |                  |            | <u>L</u> KTI VAYI KEKQAE TGIKNLWG TANVF GHKRYMNGAATNPDFDVVARAI VQIKNAIDATI<br><u>E</u> LGGTSYVFWGGREGHMSLLNTDQKREKEHLARMLTMARDYARARGFNGTFLIEPKPMEP<br><u>T</u> KHQYDVDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFERELAVAVDNLGLSIDAN<br><u>R</u> GDYQNGWDTDQFPIDHYELVQGMLQIIRNGGFTDGGTNFDAKTRRNSTDLEDIFIAHIA<br><u>A</u> MDAMAHAALESAASIIEESPYCQMVKDRYASFDSGIGKDFEDGKLTLEQA YEY GKQVGE P<br><u>K</u> QTS GKQELYESIINMYSI                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 5749MI2_004 | <i>Bacteroides</i> | DNA              | 23         | ATGGCAACAAAAAGAGTATTTTCCTGGTATAGGAAAAGATTAAATTTGAAGGTAAAGAGAGT<br>AAGAA TCCGATGGCATTCGGCTATTATGATGCCAATAAAGTAATCATGGGCAAGAAAGATG<br>AGCGAGTGGCTGAAAGTTTGGCATGGCTTGGTGGCACACATTTGCGCCGAAAGGTGGTGAC<br>CAGTTTGGTGGTGGACAAAAGACTTCCCGTGGAAACGATTTCGGACAAACGCCGTAGAAAGCA<br>GCCAACCATAAAGTAGATGCCGGTTTGAATTTATGCAGAAAATGGGCATCGAATACTAT<br>TGCTTCCATGATGTAGACCTCTGCAC TGAAGCTGTACCATTGAAGAATATGAAGCCAAAT<br>CTGAAGGAAAATAGTAGCCTATCCGAAACAGAAACAGGCTGAAACAGGTATCAAACTTCTG<br>TGGGTACGGCAAAATGTATTGGTCACAAAACGCTATATGAATGGTGTCTACCAAATCCG<br>GATTTTGATGTAGTGGCTCGTGTCTGTACAGATTAAAGATCGGATAGACGCTACAAT<br>GAACTCGGTGGTAGCAACTACGTGTTCTGGGGCGCGGTGAAGTTATATGAGCTTGCTC<br>AATACAGACCAGAAACGTGAGAAAGAGCATTTGGCACAAA GTTGACCATGGCTCGTGAC<br>TATGCTCGTGCCAAAGGATTCAAGGTACCTTCCTGGTTGAACCCAAACCGATGGAACCA<br>ACTAAACACCAGTATGATGTAGATACGGAAC TGTAAATCGGCTTCCTCAAAGGCTCATAAT<br>TTGGATAAGGATTTCAAGGTAATAATTTGAAGTAAACCATGCTACATTTGCCCGGTCACTACT<br>TTTGAACACGAAATGGCTGTTGCCGTAGACAAACGATATGCTTGGCTCTATCGATGCCAAC<br>CGCGGTGACTATCAGAACGGTTGGGATAC TGAACCATGCCCCATTGACAACTTCGAGCTT<br>ATCCAAGCCATGATGCAGATTA TTCGCGGTGGTGGCTTCAAAGATGGTGGTACAAACTTC<br>GACGCTAAGACTCGTCGTAAC TCTACCGACCTGGAAGATAATTTTCA TTGCACACATCGCT<br>GGTATGGATGCTATGGCACGTGCTTTGGAAAGTGCAGCCAAAGTTGCTTGAGGAATCTCCT<br>TATAAGAAAAATGTTGGCTGACCGCTATGCATCGTTCGATAGTGGCAAAGGTAAGGAGTTT<br>GAAGAAGGCAAGCTGACATTGGAAGACGTTGTAGTTTATGCCAAGCAGAA TGGCGAGCCT<br>AAACAGACCAGCGTAAGCAGGAATTGTATGAGGCAATTGTAAATATGTATGCCTGA |
| 5749MI2_004 | <i>Bacteroides</i> | Amino Acid       | 24         | <u>M</u> ATKEYFPGIGIKIFEGKE SKNPMAFRYYDANKVIMGKKMSEWLKFAMAWWHTLCAEGGD<br><u>Q</u> FGGGTKTFPWNDSDNAVEAAANHKVDAGFEFMQKMGIEYYCFHDV DLTCEAATIEEYEAN<br><u>L</u> KEI VAYPKQKQAE TGIKLLWGTANVF GHKRYMNGAATNPDFDVVARAAVQIKNAIDATI<br><u>E</u> LGGSNYVFWGGREGYMSLLNTDQKREKEHLAQMLTMARDYARAKGFKGTF LVEPKPMEP<br><u>T</u> KHQYDVDTETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELAVAVDNDMLGSIDAN<br><u>R</u> GDYQNGWDTDQFPIDNFELIQAMMQIIRGGGFKDGGTNFDAKTRRNSTDLEDIFIAHIA<br><u>G</u> MDAMARALESAAKLLLEESPYKKMLADRYASFDSGKGKFEFEEGKLTLEDVVVYAKQNGEP<br><u>K</u> QTS GKQELYESIINMYA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5750MI3_003 | <i>Bacteroides</i> | DNA              | 25         | ATGGCAACAAAAAGAGTATTTTCCTGGAATAGGAAAAGATTAAATTTGAAGGAAAAGAGAGT<br>AAGAACCCGATGGCATTCGGTTGCTACGATGCAGAAAAAAGTTATCATGGGTAAAGAGAA TG<br>AAAGATTGGT TGAAGTTTGCAATGGCGTGGTGGCATAACACTTGTGCAGAAAGCGGTGAC<br>CAATTCGGTGGCGGTACAAAAGAGTTTCCCCCGGAACGACTATACTGATAAAATTCAGGCT<br>GCTAAAAACAAGATGGATGCCCGTTT TGAGTTTATGCAGAAAGATGGGGATCGAATACTAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|--------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|             |                    |                  |            | TGTTTTCAGATGTAGACCTCTGCACGGGAAGCTGATACCATTGAAGAATACGGAAGCTAAT<br>TTGAAAGAAATCGTAGTTTACGCCAAAGCAAAAGCAGGTAGAAACAGGTATCAAATTATTG<br>TGGGTACTGCCAATGTATTCCGTCATGAACGCTATATGAATGGTGGGCTACCAACCCA<br>GATTTGATGTTGTAGCCCGTGTCTGTTCAAGTTAAGAAATGCAATTGATGCTACCAATT<br>GAACTAGGTGGCTTAAACTATGTGTTCTGGGTGGACGCGGAAGTTATATGTCTTTGCTG<br>AACACTGATCAGAAACGTGAGAAAGAACATCTTGACAAATGCTGACCATTGCCCGTGAC<br>TATGCCCGTGCCCGTGGCTTCAAAGGTACATTCTTGGTTGAACCGAAACCGATGGAACCA<br>ACCAAAACATCAATATGACGTAGATACAGAAACAGTTATCGGTTTTTTGAAAGCTCATGCT<br>TTGGATAAAGACTTTAAAGTAAATATTGAAGTAAATCAIGCAACATTAGCCGGTCATACA<br>TTTGAACACGAACTGGCAGTGGCTGTCGACAAACGGTATGCTGGGTTCTATTGACGCTAAT<br>CGTGGTGATTTGTCAAAACGGTTGGGATACAGACCAATTTCCCATTGATAACTATGAAC TG<br>ACTCAAGCCATGATGCAGATTATTCGTAACGGTGGTTTGGGCAATGGTGGTACGAATTTT<br>GACGTAAAACTCGCCGTAATTTCTACTGATCTTGGAGATATCTTCATTGCTCACATCGCA<br>GGTATGGATGCTATGGCACGTGCAATTGGAAAGTGGGCCAAGTTGTTGGAAGAATCTCCC<br>TATAAGAAAGATGCTGGCAGAACGTTATGCATCCTTTGACAGCGGTAAGGGTAAAGAGTTT<br>GAAGAGGGTAAGTTGACCTTGGAGGATCTTGTGCTTATGCAAAAGTCAATGGCGAACC G<br>AAACAAATCAGTGGTAAACAAGAAATTGTATGAGGCAATTGTGAATATGTATTGCTAA |
| 5750MI3_003 | <i>Bacteroides</i> | Amino Acid       | 26         | MATKEYFPGIGKIKFEGKESKNPMAFRCYDAEKVIMGRMKDWLKFAMAWWHTLCAEGGD<br>QFGGGTKSFPRNDYT <del>DKIQAAKNKMDAGFEFMQKM</del> GIEYYCFHDDV <del>DLCTEADTIEE</del> YEAN<br>LKEIVVYAKQKQVETGIKLLWGTANVFGHERYMN <del>GAATNP</del> DFDVVARAAVQIKNAIDATI<br>ELGGLNYVFWGGREGYMSLLNTDQKREKEHLAQMLTIARDYARARGFKGTFIVEPKPMEP<br>TKHQYDVDTETVIGFLKAHALDKDFKVNIEVN <del>HATLAGHTFEHELAV</del> VDNGLGSLIDAN<br>RGDCQNGWDTDQFPIDNYELTQAMMQIIIRNGGLNGGTN <del>FDAKTR</del> NSTDLGDI <del>FIAHIA</del><br>GMDAMARALESAAKLLEESP <del>YKKMLAERYAS</del> FD <del>SGKGKEFE</del> EGKLTLEDLVAYAKVNGEP<br>KQISGKQELYEALVNM <del>Y</del> C                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 5750MI4_003 | <i>Bacteroides</i> | DNA              | 27         | ATGGCAACAAAAGAGTATTTTCCCGGAATAGGAAAAGATTAAATTCGAAGGTAAGAGAGC<br>AAGAACCCGATGGCATTCGGTTATTACGATGCCGATAAAGTAATCATGGGTAAGAAAAATG<br>AGCGAATGGCTGAAGTTCGCCATGGCATGGTGGCACACTCTTTGCGCAGAAAGGTGGTGAC<br>CAGTTCGGTGGCGGAACAAAGAAATTC <del>CCCTGGAACGGTGAGGCTGACAAAGGTT</del> CAGGCT<br>GCCAAGAACAAAAATGGACCGCGGCTTTGAATTCATGCAGAAAAATGGGTATCGAATACTAC<br>TGCTTCCACGATGTAGACCTCTGCGAAGAAGCCGAGACCATTGAAGAATACGAAGCCAAC<br>TTGAAGGAAATCGTAGCGTATGCCAAGCAGAAACAAGCAGAAACCGGCATCAAGCTGTTG<br>TGGGTACTGCCAACGTATTCGGCCATGCCCGTACATGAATGGTGCAGCCACCAACCCC<br>GATTCGATGTTGTGGCACGTGCAGCCGTC <del>CAAAATCAAAAAGCGCCATCGACGCTACTATC</del><br>GAGCTGGGAGGTTCGAACTATGTGTTCTGGGCGGT <del>CGCGAAGGCTACATGT</del> CATTGCTG<br>AATACAGACCAGAAAGCGTGAGAAAGAGCACCTCGCACAGATGTTGACCATCGCCCCGCAC<br>TATGCCCGTGCCCGTGGCTTCAAAGGTACCTTCCTGATTGAACCGAAACCGATGGAACCT<br>ACAAAACACCAGTATGATGTAGACACCGAAACCGTTATCGGCTTCTTGAAAGGCCCACAAT<br>CTGGACAAAGATTCAAGGTAAACATCGAAGTGAACCCACGCTACTTTTGGCGGGCCACACC<br>TTCGAGCACGAACTCGCAGTAGCCGTAGACAAACGGTATGCTCGGCTCCATCGATGCCAAC<br>CGTGGTGACTACCAGAACGGCTGGGATACAGACCAGTTC <del>CCCCATTGACAACTTCGAACTG</del>       |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------|--------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|             |                    |                  |            | ACCCAGGCAATGATGCAAAATCATCCGTAACGGCGGCTTTGGCAATGGCGGTACAAACTTC<br>GATGCCAAGACCCGTCGTAACTCCACCGACCTGGAAGACATCTTCATTGCCCCACATCGCC<br>GGTATGGACGTGATGGCACCTGCACTGGAAAGTGACGCCAAATTGCTTGAAGAGTCTCCT<br>TACAAGAAAGATGCTTGCCGACCGCTATGCTTCCTTCGACAGTGGTAAAGGCAAGGAATTC<br>GAAGACGGCAAGCTGACACTGGAGGATTTGGCAGCTTACGCAAAAAGCCAAACGGTGAGCCG<br>AAACAGACCAAGCGCAAGCAGGGATTGTATGAGGCAATCGTAAATATGTACTGCTGA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 5750MI4_003 | <i>Bacteroides</i> | Amino Acid       | 28         | MATKEYFPGIGIKIFEGKESKNPMAFRIYDADKVI MGKKMSEWLKFAMAWWHTLCAEGGD<br>QFGGTTKKFPWNGEADKVQAAKNKMDAGFEFMQMGIEYYCFHVDVLDCEEAETIEEYEAN<br>LKEIVAYAKQKQAE TG I K L L W G T A N V F G H A R Y M N G A A T N P D F D V V A R A A V Q I K S A I D A T I<br>ELGGSNYVFWGGREGYMSLLNTDQKREKEHLAQMLT IARDYARARGFKGTFLIEPKPMEP<br>TKHQYDVDTE TV I G F L K A H N L D K D F K V N I E V N H A T L A G H T F E H E L A V A V D N G M L G S I D A N<br>RGDYQNGWDTDQFPIDNFELTQAMMQIIRNGGFNGGTFNFDAKTRNSTDLEDIFIAHIA<br>GMDVMARALESAAKLLEESPYKKMLADRYASFDSGKGKEFEDGKLTLEDLAAAYAKANGEP<br>KQISGKQGLYEAI V N M Y C                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5751MI4_002 | <i>Bacteroides</i> | DNA              | 29         | ATGACAAAAGAGTATTTTCCAACCATTTGGTAAAAATTCAGTTTGAAGGTAAGAGAGTAAG<br>AATCCATTAGCATATCGTTATTACGATGCTAACAAAGTAATAATGGGTAAAAAGATGAGC<br>GAATGGCTCAAGTTTGCAATGGCATGGTGGCACACTTTGTGTGCTGAGGGTAGCGACCCAG<br>TTTGGTCTGTGGCAACAAGTCATTCCTATGGAAACGCATCAACCGACCGTATGCAGGCTGCA<br>AAAGATAAGGCTGACGCAGGCTTCGAAATCATGCAAAAAACCTGGGCATCGAATACTACTGT<br>TTCCATGATGTTGACCTCATCGACCCAGCAGACGATATTCCAACATACGAAAAGAAATCTC<br>AAGGAAATCGTTGCATACCTCAAGCAAAAACAGGCCGAGACAGGTATCAAAATTGCTATGG<br>GGTACAGCTAACGTATTTGGCCACAAGCGTTATATGAACGGTGCACTACCAATCCTGAC<br>TTTGACGTTGTTGCACGAGCTATCGTGCAAAATCAAGAAATGCTATCGATGCAACAATCGAA<br>CTGGCGGCACGAACTACGTATTTCTGGGTGGTCGCGAAGGTTACATGTCACTGCTCAAC<br>ACCGACCAAAAAGCGGAGAAAAGAGACATGGCTACCATGTTAGGAATGGCACGTGACTAT<br>GCACGTTCTAAAGGCTTTACTGGTACTCTCCTTATCGAGCCAAAGCCTATGGAACCAACT<br>AAGCATCAATACGACGTGATACAGAAACTGTTATTGGTTTCTCAAAGCTCACGGATTTA<br>GACAAGGACTTCAAGGTAATAATCGAAGTGAACCACGCTACATTGGCTGGCCATACCTTC<br>GAACATGAATTAGCATGTGCTGTTGATGCAGGTATGCTTGGTTCCATCGATGCTAAACCGT<br>GGTGATATGCAGAAATGGCTGGGATACAGATCAGTTCCTATCAACAATTACGAGCTCGTT<br>CAGGCCATGATGCAGATTATCCGCAATGGTGGTTTCGGTAAACGGTGTACAAACTTCGAC<br>GCTAAGACACCGTCGTAATTCAACCGATTTGGAAGACATCATCATTTGCTCACGTTTCAGCT<br>ATGGATGCTATGGCAGCTGCTCTTGAATGTGCTGCAGACATTCCTCAAAAACTCACCTATT<br>CCACAGATGGTGGCCAACCGTTATGCAAGTTT TGACAAAGGTATAGGTAAGATTTCGAA<br>GACGGCAAGCTCACCCCTCGAGCAAGTATACGAATATGTTAAGACCGTCGGCGAACCAGCT<br>ATTACAAGCGGCAAAACAGGAGCTCTACGAAGCTATCGTTAAATATGTATTGCTGA |
| 5751MI4_002 | <i>Bacteroides</i> | Amino Acid       | 30         | MTKEYPPTIGIKIQFEGKESKNPLAYRIYDANKVIMGKKMSEWLKFAMAWWHTLCAEGSDQ<br>FGPGTKSFPWNASTDRMQAAKDKADAGFEIMQKLGIEYYCFHVDVLDLIDPADDDIPTYEKNL<br>KEIVAYLKQKQAE TG I K L L W G T A N V F G H K R Y M N G A S T N P D F D V V A R A I V Q I K N A I D A T I E<br>LGGTNYVFWGGREGYMSLLNTDQKREKEHMATMLGMARDYARSKGFTGTLTIEPKPMEPT<br>KHQYDVDTE TV I G F L K A H G L D K D F K V N I E V N H A T L A G H T F E H E L A C A V D A G M L G S I D A N R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| TABLE 2     |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------|--------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|             |                    |                  |            | GDMQNGWDTDQFPINNYELVQAMQIIRNGGFNGGTFNFDAKTRRNSTDLEDI IIAHVSA<br>MDAMARALECAADILQNSPIPQMVANRYASFDKIGKDFEDGKLTLEQVVEYEGKTVGEPA<br>ITSGKQELYEAI VNMYC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 5751MI5_003 | <i>Bacteroides</i> | DNA              | 31         | ATGGCTAACAAAGAATTTTCCCGGTATTGGTAAAAATCAAAATTCGAAGGTAAAGAGAGC<br>AAGAACCCCATGGCATATCGTTACTACGATGCTGAGAAAGGTAGTCCTTGGCAAGAAATATG<br>AAAGACTGGTTCAAGTTTCCGATGGCTTGGTGGCACACATTTGCGCCGAGGGTAGCGAC<br>CAGTTTGGTCCCGGCACTAAGTCTTTCCCTTGGAAACACCGCAGAGTCCCCCATGCAGGCA<br>GCTAAGGACAAGGTTGACGCTGGCTTCGAGTTTCATGACCAAGATGGGTATTGAATACTTC<br>TGCTTCCACGATGTAGACCTCGTTGCCGAGGCCGACACTGTTGAGGAGTACGAGGCTCGC<br>ATGAAGGAAATCGTTGCTTACATCAAGGAGAAAGGTGGCCGAGACTGGCATCAAGAAACCTG<br>TGGGTACAGCTAACGTATTGGCAACAAGCGCTACATGAACGGTGTCTACTACTAACCCC<br>GACTTTGACGTTGTGGTCGGCTATCGTTCAAATCAAGAACGCTATCGACGCTACTATC<br>GAGTCGGTGGTACGT CATA CGTATTCTGGGGCGCGGAGGGTTACATGAGCCTCTTG<br>AACCCGACCAGAAAGCGTGAGAAAGAGCACCTGGCTACTATGCTCACTATGGCACGCGAC<br>TACGCTCGCGCTAAGGGTTTCAAGGGTACATTCTCATCGAGCCCAGCCCATGGAGCCC<br>ACAAAGCACCGTACGATGTTGACACTGAGACTGTAAATCGGCTTCCTTAAGGCACACAAAC<br>CTTGACAAAGGACTTCAAGGTTAACATTGAGGTTAACCCACGCAACTCTCGCTGGTCACACA<br>TTTGAGCACGAGCTCGCTTGCTGTTGACGCTGGCATGCTTGGCAGCATCGACGCTAAC<br>CGCGGTGACTACCAAGAACGGCTGGGATACTGACCAATTCCCCATCGACAACTTCGACCTC<br>ACTCAAGCTATGCTCGAGATCATCCGCAACGATGGTTTCAAGGATGGTGGTACAAAACCTC<br>GACGCTAAGACTCGCCGCAACAGCACCGACCTCGAGGATATCTTCATCGCACACATCGCT<br>GCTATGGACGCTATGGCACGTGCTCTCGAGAGCGCTGCTGCAGTACTCGAGGAGTCAGCT<br>CTGCCCCAAATGAAGAAGGACCGCTATGCATCGTTCGACGCTGGCATGGTAAGGACTTC<br>GAGGACGGCAAGCTCACCCCTGGAGCAAGTTTACGAGTATGTTAAAGAGGTGGCGGAGCCC<br>AAGCAGACTAGCGGCAAGCAAGAGCTGTATGAGGCTATCCTCAACATGTACGTATAA |
| 5751MI5_003 | <i>Bacteroides</i> | Amino Acid       | 32         | MANKEFFPGIGIKIFEGKESKNPMAYRYDDAEKVVLGKNMKDWFKFAMAWHHTLCAEGSD<br>QFGPGTKSFPWNTAECPMQAAKDKVDAGFEFMTKMGI EYFCFHDVLDLVAEADTVEEYEAR<br>MKEIIVAYIKEKVAETGIKNLWGTANVFGNKRYSMNGAATNPDPDVVARAIVQIKNAIDATI<br>FLGGTSYVFWGGREGYMSLLNTDQKREKEHLATMLTMARDYARAKGFKGTFLLIEPKPMEP<br>TKHQYDVDTETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDAN<br>RGDYQNGWDTDQFPIDNFDLTQAMLEIIRNDGFKDGGTNFDAKTRRNSTDLEDIFIAHIA<br>AMDAMARALESAAAVLEESALPQMKKDRYASFDA GMGKDFEDGKLTLEQVVEYEGKKVGE<br>KQTS GKQELYEAILNMYV                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5751MI6_004 | <i>Bacteroides</i> | DNA              | 33         | ATGGCTAACAAAGAATTTTCCCGGTATTGGTAAAAATCAAAATTCGAAGGCAAGAAAGC<br>AAGAACCCCATGGCATATCGTCACTACGATGCCGAGAAAGGTAGTCCTTGGTAAGAAATG<br>AAGGACTGGTTCAAGTTTCCGATGGCTTGGTGGCACACTCTGTGCGCCGAGGGTAGCGAC<br>CAGTTCCGCCCCGTGACCAAGTCTTTCCCTTGGAAACCAAGCCGAGTCCCCCATGCAGGCT<br>GCTAAGGACAAGGTTGACGCGCGCTTCGAGTTTCATGACCAAGATGGGTATCGAATACTTC<br>TGTTTCCACGATGTAGACCTCGTTGCCGAGGCCGACACCGTIGAGGAGTACGGAAGCTCGC<br>ATGAAGGAAATCGTGGCTTACATCAAGGAGAAAGATGGCCGAGACCGGCATCAAGAACCTG<br>TGGGTACAGCCAAACGTATTTCGGCAACAAGCGCTACATGAACGGTGTCTACTACTAACCCC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| TABLE 2      |                      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--------------|----------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism    | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|              |                      |                  |            | GACTTTGACGTTGTGGCTCGGCAATCGTTCAGATCAAGAACGCCATCGACGCTACTATC<br>GAGCTCGGCGGTACCTCTTACGTGTTCTGGGGCGCGCGAGGGTTACATGACTCTCTTG<br>AACACCGACCAAGCGCGAGAAGGACACCTGGCTACCATGCTCACCATGGCTCGGAC<br>TATGCTCGGCTAAGGGCTTCAAGGTTACATTCTTTATCGAGCCCAAGCCCATGGAGCCC<br>ACCAAGCACCAAGTATGACGTGGATACCGAGACCGTTATCGGCTTCTCAAGGCTCACGGC<br>CTGGACAAGGACTTCAAGGTGAACATCGAGGTTAACCATGCAACTCTCGCGGCCACACA<br>TTCGAGCACGAACTCGCTTGGCTGTTGACGCTGGCATGCTGGGAGCATCGACGCTAAC<br>CGGGCGACTACCAAGACGGCTGGGATACCGACCACTTCCCCATCGACAACTTCGACCTC<br>ACTCAGGCTATGCTCGAGATCATCCGCAACGGTGGTTTCAAGGACGGTGATCAAACTTC<br>GACGCTAAGACCCGTGCGCAACAGCACCGATCTTGAGGACATCTTCATCGCTCACATCGCT<br>GCTATGGACGCAATGGCACGCGCGCTCGAGAGCGCTGCCGTGTGCTCGAGCAGAGCCCC<br>CTTCCCCAGATGAAGAAAGACCGCTACGCATCGTTCGATCCCGGCATGGGCAAGGACTTC<br>GAGGACGGCAAGCTCACTCTGGAGCAGGTTTACGAGTATGGTAAGAAGGTAGGCGAGCCC<br>AAGCAGACCAGCGGCAAGCAGGAACGTGTACGAGGCTATCCCTCAACATGTATGTATAA                                                                                                                                                                                                                                                                                                                                            |
| 5751MI6_004  | <i>Bacteroides</i>   | Amino Acid       | 34         | <u>MANKEFFPGIGIKIFEGKESKNPMAYRHYDAEKVVVLGKKMKDWFKFAMAWHHTLCAEGSD</u><br><u>QEGPVTKSFPNWQAECPMQAAKDKVDAGFEFMTKMGIEYFCFHDVDLVAAEADTVEEYEAR</u><br><u>MKEIVAYIKEKMAETGIKNLWGTVNVEGNKRYMNGAATNPDFVVARAIVQIKNAIDATI</u><br><u>ELGGTSYVFWGGREGYMTLLNTDQKREKEHLATMLTMARDYARAKGFKGTFLIEPKPMEP</u><br><u>TKHQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN</u><br><u>RGDYQNGWDTDQFPIDNFDLTQAMLEIIRNGGFKDGGTNFDAKTRRNSTDLEDIFIAHIA</u><br><u>AMDAMARALESAAAIVLEQSPLPQMKKDRYASFDAGMGKDFEDGKLTLEQVVEYGKKVGEF</u><br><u>KQTS GKQELYEAILNMYV</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI22_003 | <i>Clostridiales</i> | DNA              | 35         | ATGAAAGAATATTTTCTATGACAAAAAAAGTTGAATATAGGGCGCAGCATCTAAAAAT<br>CCATTTGCGTTTAAATACTATGATGCCGAAAGAATATATAGCAGGCAAGCCTATGAAAGAA<br>CATCTTAAATTTGCTATGAGTTGGTGGCATAACACTTTGTGCGGGCGGTGCAGACCCATT<br>GGCACAACAACATATGGACAGAACATACGGCGGACTTACCGACCCAAATGGAAATTGCAAAG<br>GCAAAAGTAGATGCAGGCTTGAGTTTATGCAAAAACCTCGGTATAGAGTATTTTGTTTT<br>CACGATGCGGATATTGCACCGGAAGGACAGTTTGTGTGAAACAAAGAAAAACTTTTGG<br>GAAATAGTAGATTATATACAGCAAAAAGATGAATGAAACACAGGCATAAAAGTTGCTTTGGGT<br>ACTGCAAACTGCTTTAATGCTCCACGTTATATGCACGGTGCAGGAACATCATGCAATGCG<br>CACAGTTTTCATATGCAGCCGCACAGATAAAAAATGCAATTGAAGCTACCGTTAAACTG<br>GGTGAAAAGGCTATGTTTCTGGGCGGAAGAGAGGGTTATGAAACACTTCTCAATACG<br>GATATGGCACTTGAACTTGACAATATGGCAAGACTTATGCATATGGCAGTTGATTATGGC<br>AGAAAGCATTGGTTTTCACGGTGATTTTATATCGAACCAAGCCAAAGGAACCAACAAAA<br>CATCAATATGACTTTGACTCGGCAACTGTTTGGGATTTTGGAAAAGTACGGTTTAGAT<br>AAGGATTTTAAACTTAATATAGAGGCAAAATCATGCGACACTTGCAGGTATACATTTGAA<br>CATGAATTGACTGTAGCGCGTATAAACGGTGCAATTTGGCAGCATAGATGCAAAATAGCGGC<br>GATCCCAATCTTGGCTGGGATACCGACCAATTCCCAACAGATGTTTATTGGCAACCCTT<br>TGATGCTTGAAGTGATAAGAGCAGGCGGCTTTACAAAACGGAGGCTCTTAAATTTTGATGCA<br>AAGGTCAGAAGAGGCTCATTTACGTTTGATGACATTTTATGCAATATATCAGCGGTATG<br>GACACTTTTGGCGCTGGGTTTATAAAGGCATATGAAATAATTGAGGACGGCAGAAATAGAT |

| TABLE 2      |                      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------------|----------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism    | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|              |                      |                  |            | GAATTTGTAAGAAAGATACGCAAGCTATAATAACAGGCATAGGCAAGATATTATAGAT<br>GGAAAGGCAAGCCTTGAAAGTTTGGAAGAAATATATTCTTTCAAATGATAATGTTGTAATG<br>CAAAGCGCAGACAGGAATATCTTGAAACAGTTTTGAATAATATTTTGTTTAAAGCATAA<br>MKEYFPMTKKVEYEGAASKNPF <del>AFKY</del> YDAERIIAGKPMKEHLKFAMSWWHILCAGGADPF<br>GTTTMDRITYGGLTDPEMIAKAKVDAGFEFMQKLGIEYFCFHDADIAPEGSSFVETKKNEW<br>EIVDYIQQKMNETGIKLLWGTANCFNAPRYMHGAGTSCNAHSFAYAAAQIKNAIEATVKL<br>GGKGYVFWGGREGYETLLNTDMALELDNMARLMHMAVDYGRSIGFDGDFYIEPKPKEPTK<br>HQYDFDSATVLGFLRKYGLDKDFKLNIEANHATLAGHTFEHELTVARINGAFGSIDANSG<br>DPNLGWDTDQFPPTDVYSATLCMLEVIRAGGFTNGGLNFDAKVRRGSTFFDDIVYAYISGM<br>DTFALGF <del>IKAYE</del> ILIEDGRIDEFVKERYASYNTGIGKDIIDGKASLESLEEYILSNDNVVM<br>QSGRQEYLETVLNILFKA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI22_003 | <i>Clostridiales</i> | Amino Acid       | 36         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 1753MI4_001  | <i>Firmicutes</i>    | DNA              | 37         | ATGAAAGAAAATTTCCCAAATATTCCTGAGATTAAATTCGAAGGAAAAGACAGCAAAAAAT<br>CCTTTTGCTTTCCATTACTACAACCCAGACCAATCATCTTAGGCAAACCAATGAAAGAA<br>CACCTCCCATTCGCTATGGCTTGGTGGCACAATCTTGGTGAACAGGTGTTGATATGTTT<br>GGCGCTGGCCAGCTGATAAGAGTTTCGGTGCTAAAGTTGGCACAATGGAACACGCTAAG<br>GCCAAAGTCGATGCCGGTTTCGAATTCAATGAAGAAACTCGGTATCAGATATTTCTGCTTC<br>CATGATGTTGACTTAGTTCAGAAATGTCAGATATCAAAGATACAAACAAAGAATTAGAT<br>GAAATCAGTGACTACATCTTAGAAAAAGATGAAAGGCACAGATATTAAGTGTTTATGGGCG<br>ACCGCCAATATGTTCTCTAACCCACGCTTCTGCAATGGTGCGGGTTCCACAAACAGTGGC<br>GATGCTTTCGCTTTCGCCGCTGCTCAAGTTAAGAAAAGCCTTAGATATCACCGTTAAATTA<br>GGTGGTAGGGGTACGTCTTCTGGGTGGTCGTGAAGGTTACGAAACATTACTCAATACA<br>GACGTTAAATTCGAACAAGAAAACATTGCTCGTTTAATGAAGATGGCTGTTGAATATGGC<br>CGTTCCATCGGTTTCAAAGCGGATTCTATATCGAACCAAAACCAAAAGAACCAATGAAA<br>CACCAATATGACTTCGACGCCGCTACAGCTATTGGCTTCTTAAGAGCCCCACGGCTTAGAC<br>AAAGACTTCAAGTTGAACATCGAAGCTAACCCACGCTACATTAGCGGGTCATACATTCCAA<br>CACGATTTAAGAAATCTCCGCCATTAATGGTATGTTAGGTTCTATCGATGCTAACCAAGGC<br>GATATGCTCTTAGGTTGGGATACAGACGAATTCCCATTGATGCTACAGTGGGACACAA<br>TGATGTACGAAGTCTTAAAGAATGGTGGTCTTACAGGTGGTTTCAACTTTGACTCCAAA<br>ACACGTCGTCCATCCTACACAATGGAAGATATGTTCTTAGCCTATATCTTAGGTATGGAT<br>ACATTTCGCTTAGGTTTAATCAAAGCTGCTCAAATCATCGAAGATGGCCGTATTGATCAA<br>TTCATCGAAAAGAAATATTCTTCTCCGTGAAACAGAAATCGGTCAAAAGATCTTAAAC<br>ACAAGACAAGCTTAAAGAAATTATCCGATTACGCTTGCAAGATGGGTGCTCCAGAACTT<br>CCAGGTAGTGGTCGTCAAGAAATGCTCGAAGCCATCGTTAACGATGTCTTATTCGGCAAG<br>TAA |
| 1753MI4_001  | <i>Firmicutes</i>    | Amino Acid       | 38         | MKEIFPNIPEIKFEGKDSKNPFAFHYYNPDQIILGKPMKEHLPFAMAWWHNLGATGVDME<br>GAGPADKSEFKA <del>VG</del> TMEHAKAKVDAGFEFMKKLGIRYFCFHDVDLVPECADIKDTNKELD<br>EISDYIILEKMGTDIKCLWGTANMFNPRFCNAGAGSTNSADVAF <del>FAAAQ</del> VKKALDITVKL<br>GGRGYVFWGGREGYETLLNTDVKFEQENIARLMKMAVEYGRSIGFKGDFYIEPKPKEPMK<br>HQYDFDAATAIGFLRAHGLDKDFKLNIEANHATLAGHTFQHDLRISAINGLGSI <del>DANQG</del><br>DMLLGWDTDEFPPDVYSATQCMYEVLKNGGLTGGFNFD <del>SK</del> TRRPSYTMEDMFLAYILGMD<br>TFALGLIKAAQIIEDGRIDQGFIEKKYSSFRETEIGQKILNNKTSLKELSDYACKMGAPEL                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| TABLE 2      |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------|-------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 1753MI6_001  | <i>Firmicutes</i> | DNA              | 39         | PGSGRQEMLEAIVNDVLF GK<br>ATGAAAGAAAATTTCCCAAATATTCCTGAGATTAAATTCGAAGGAAAAAGACAGCAAAAAAT<br>CCTTTTGCTTTCCATTACTACAACCCAGACCAAATCATCTTAGGTAAACCAATGAAAGAA<br>CACCTCCCATTCGCTATGGCTTGGTGGCACAAATCTTGGTGCAACAGGTGTGATATGTTT<br>GGCGCTGGCCAGCTGATAAGAGTTTCGGTGTCTAAAGTTGGCACAAATGGAACACGCTAAG<br>GCCAAAGTCGATGCCGGTTTCGAATTCAATGAAGAAACTTGGTATCAGATATTTCTGCTTC<br>CATGATGTTGACTTAGTTCAGAAATGTGCAGATATCAAAGATACAAAACAAAGAAATTAGAT<br>GAAATCAGTGACTACATCTTAGAAAAGATGAAAGGCACAGATATCAAGTGTTTATGGGGC<br>ACGCCCAATATGTTCTCTAAACCCACGTTTCTGCAATGGTCGGGTTCCACAAACAGTGCG<br>GATGCTTTCGCTTTCGCCGCTGCTCAAGTTAAGAAAAGCCTTAGATATCACCGTTAAATTA<br>GGTGGTAGGGGTACGTCTTCTGGGTGGTCGTGAAGGTTACGAAACATTACTCAATACA<br>GACGTTAAATTCGAACAAGAAAACATTTGCTCGTTTAAATGAAGATGGCTGTTGAATATGGC<br>CGTTCATCGGTTTCAAAGCGGATTTCTATATCGAACCAAAACCAAGAACCAATGAAA<br>CACCAATATGACTTCGACGCCGCTACAGCTATTGGCTTCTTAAGAGCCCAACGGCTTAGAC<br>AAAGACTTCAAGTTGAACATCGAAGCTAACCCACGTACATTAGCGGGTCATACATTCCAA<br>CACGATTTAAGAAATCTCCGCCATTAAATGGTATGTTAGGTTCTATCGATGCTAACCAAGGC<br>GATATGCTCTTAGGTTGGGATACAGACGAAATTTCCATTTGATGTTACAGTGGACACAA<br>TGATGTACGAAGTCTTAAAGAAATGGTGGTCTTACAGGTGGTTTCAACTTTGACTCCAAA<br>ACACGTGCTCCATCCTACACAAATGGAAGATATGTTCTTAGCCTATATCTTAGGTATGGAT<br>ACATTCGCTTTAGGTTTAAATCAAAGCTGCTCAAATCATCGAAGATGGCCGTATTGATCAA<br>TTCATCGAAAAGAAATATTTCTTCTTCCGTGAACACAGAAATCGGTCAAAAGATCTTAAAC<br>ACAAGACAAAGCTTAAAAGAAATATCCGATTACGCTTGCAAGATGGTGCTCCAGAACTT<br>CCAGGTAGTGGTCGTCAAGAAATGCTCGAAGCCATCGTTAACGATGTCTTATTCGGCAAG<br>TAA |
| 1753MI6_001  | <i>Firmicutes</i> | Amino Acid       | 40         | MKEIFPNIP EIKFEGKDSKNPFAFHYYNPDQIIILGKPMKEHLPPFAMAWWHNLGATGVD MF<br>GAGPADKSF GAKVGTMEHAKAKVDAGFEFMKKLIGIRYFCFHDVLDLVPECADIKDTNKELD<br>EISDYIILEKMKGTDIKCLWGTANMFSPNPRFCNGAGSTNSADVFAFAAAQVKKALDITVKL<br>GGRGYVFWGGREGYETLLNIDVKFEQENIARLMKMAVEYGRSIGFKGDFYIEPKPEPMK<br>HQYDFDAATAIGFLRAHGLDKDFKLNIEANHATLAGHTTFQHDLRISAINGMLGSI DANQG<br>DMLLGWDTDEFPPFDVYSATQCMYEVLKNGGLTGGFNFD SKTRRPSYTMEDMFLAYILGMD<br>TFALGLIKAAQI IEDGRIDQFIEKKYSSFRETEIGQKILNNKTSLKELSDYACKMGAPEL<br>PGSGRQEMLEAIVNDVLF GK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 1753MI35_004 | <i>Firmicutes</i> | DNA              | 41         | ATGGAATATTTCCCTTTCGTCAAAATCGGTCCAATACAAAGGGACCAACCTCAACTGAACCA<br>TTCGCTTTC AAGTACTACGATGCCAACCGTGTCGTTCTTGAAAAACCAATGAAAGAATGG<br>ATGCCATTCGCTATGGCTTGGTGGCACAACTTCGGCGCTGCCGGTACCGACATGTTCCGC<br>GGCAACACCATGGACAAGTCTTGGGGAGTCGATAAAGAAAAAGACCCAAATGGGCTATGCC<br>AAAGCCAAAGTTGATGCCGGCTTCGAATT CATGCAGAAAGATGGGCATCGAATACTACTGC<br>TTCACAGATGTCGACCTCGTCCCAGAGTGCACGACATCACCGTTATGTACCCAGAGACTC<br>GATGAGATCGGTGATTACCTTCTCAAGAAAA CAGAAAGGAACCCGGTATCAAGCTTCTTTGG<br>TCAACCGCCCAATGCCTTCGGACACCGCCGTTTCATGAACGGTGCTGGTTCAGCAACTCC<br>GCCGAAGTCTATTGCTTCGCCCGCCGCCAGATCAAGAAAAGCTCTTGAGCTCTGCGTCAAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| TABLE 2      |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|              |                   |                  |            | CTCGGTGGCAAAGGCTATGTCTTCTGGGTGGACGTGAAGGCTACGAAAACCCTTCTCAAC<br>ACCGACATGAAGTTCGAACAAGAGAACATCGCCAACTTATGAGATGCGCCCGTGACTAC<br>GGCCGCAAGATCGGTTTCAAAGCGACTTCTACATCGAACCAAAACCAAGAGCCACA<br>AAGCATCAGTATGACTTCGACGCCGTACGCCATCGGATTCCTCGTCAGTACGGTCTC<br>GACAAAGACTTCAAGATGAACATCGAAGCCAAACCAACGCTACCTTAGCTGGCCACACCTTC<br>GAACACGAACTCCGCGTCTCCGCCATGAACGGCATGTCGGTTCCATCGACGCCAAACGAA<br>GGGATATGCTCCTCGGATGGGATGTCGACCGTTTCCAGCCAAACGTCATAGCGCCACC<br>TTCGCCATGCTCGAAGTCATCAAAGCCGTGGACTTACCGGTGGCTTCAACTTCGACGCC<br>AAGACCCGCCGCTTCCAAACACTATGAAGATATGTTCAAGGCTTTCGTCTTGGTATG<br>GATACCTTCGCTTTAGGTCTTCTCAATGCCGAAGCCATCATCAAAGACGGCCGCATCGAC<br>AAGTTCGTCGAGGATAGATATGCCAGCTTCAAGACCCGCATCGGTGCTAAGGTCGCGGAT<br>CACTCCGCTACCTTGAGGATTTAGCTGCCACGCCCTTGAGACCAAGGTTTGCCCAAGAT<br>CCAGGCAGCGCGACGAGGAAGAACTCCAGGAAATCCTCAACCAAGTTAATGTTCCGGTAAG<br>AAATAA                                                                                                                                                                                                                                                                                                                                                                                                                |
| 1753MI35_004 | <i>Firmicutes</i> | Amino Acid       | 42         | MEYFPFVKSVQYKGPTSTPEFAFKYYDANRVVLGKPKMEWMPFAMAWWHNLGAAGTDMFG<br>GNTMDKSWGVDKEKDPMGYAKAKVDAGFEFMQKMGIEYYCFHDVLDLVPECCDDITVMYQRL<br>DEIGDYLLKKQKETGIKLLWSTANAFGHRRFMNGAGSSNSAEVYCFAAAQIKKALELCVK<br>LGKGYYFWGGREGYETLLNTDMKFEQENIANLMRCARDYGRKIGFKGDFYIEPKPKPT<br>KHQYDFDDAATAIGFLRQYGLDKDFKMNIEANHATLAGHTFEHEL RVSAMNGLGSI DANE<br>GDMLLGWDVDVDRFPANVYSATFAMLEVIKAGGLTGGFNFDAKTRRASNTYEDMFKA FVLGM<br>DTFALGLLNAEAI IKDGRIDKFVEDRYASFKTGIGAKVRDHSATLEDIAAHALET KVC PD<br>PGSGDEEELQEILNQLMFGKK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 1754MI9_004  | <i>Firmicutes</i> | DNA              | 43         | ATGAGCGAATTTTTTAAGAAATATCCAGAGATTAATAATCGAAGGAAAAGATAGTAAAAAT<br>CCATGGGCATTCAAGTATTACAATCCTGAATTGACCATTATGGGTAAAAAAATGTCTGAA<br>CATCTTCCTTTTGCAATGGCCTGGTGCAATAACCTTGCGCAAAATGGAGTTGATATGTTT<br>GGTTCGGGAACCGCCGATAAATCTTTCGGTCAGGCTCCGGGAACATGGAGCACGCAAG<br>GCTAAGGTAGATGCAGGTATCGAGTTTATGAAGAAACTCGGAATCAAGTACTACTGCTGG<br>CATGATGTAGACCTTGTTCTGAAGATCCAAACGATATCAACGTAAACAAAACAAGCGCCTT<br>GATGAGATTTCAGATTATATCCTTGAAAAAACAAGGGAAC TGACATCAAAGTGTCTCTGG<br>GGAAC TGCTAACATGTT CAGTAATCCCCGCTTATGAACGGGCGAGGCTCAACAAAAC TCT<br>GCTGACGTTTACTGCTTTCAGCTGCCCAGGTTAAAAAGGCTCTTGAGATTACCGTAAAG<br>CTTGGTGGCCCGGTTATGTATTCTGGGGTGGACGCGAAGGTTATGAAAAC TCTTCTTAAT<br>ACAGATGTAAAGCTTGAACAGGAAAAATATTGCAAACTTATGCACATGGCAGTTGATTAT<br>GGCCGTTCAATCGGTTTCAAGGGAGACTTCTACATCGAGCTAAGCCAAAGGAGCGCGATG<br>AGTCATCAGTATGATTTTGATGCCGCAACTGCAATCGGCTTCCTCCGCCAGTATGGCCTC<br>GACAAAGACTTTAAGATGAACATTGAGGCTAACCCAGCTTCTCTTGCAAA TCATACCTTC<br>CAGCATGAGCTTTATATCAGCCGCATTAAACGGAATGCTTGGTTCTGTAGATGCTAAACCAG<br>GGAAATCCAATTC TCGGCTGGGATACAGATAA CTTCCCTTGGAATGTCTACGACGCAACT<br>CTTGCAATGTACGAAGTACTCAAGGCTGGTGGACTTACAGGTGGCTTCAACTTTGACTCA<br>AAGAACCGCCGCCCATCAAAATACATTTGAAGATATGTTCCACGCTTACATCATGGGAATG<br>GACACTTTTGCTCTTGGTCTTATTAAAGGCTGCAGAAAATTATTGAAGACGGAAGAAATCGAT |

| TABLE 2      |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|--------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|              |                   |                  |            | GGCTTCATTAAAGAAAAGTATTCAAGCTACGAAAGTGGAATTGGTAAGAAAGATCCGCGAC<br>AAGCAGACAACTTTGGAAGAGCTTGCTGCCCGTGCCGCAGAAAATGAAAAAGCCATCTGAT<br>CCAGGTTCAGGCCCGCAGGAAATATCTGGAAGGAGTTGTTAACAAATATCCTCTTTCCGCGGA<br>TAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 1754MI9_004  | <i>Firmicutes</i> | Amino Acid       | 44         | MSEFFKNIPEIKFEGKDSKNPWAFKYYNPELTI MGKKMSEHLPFAMAWWHNLGANGVDMF<br>GSGTADKSFQI PGTM EHAKAKVDAGIEFMKKLGIKYCWHVDLVPEDPNDINVTNKRL<br>DEISDYILEKTKGTDIKCLWG TANMFNPRFMNGAGSTNSADVCF AAAQVKKALEITVK<br>LGGRGYVFWGGREGYETLLNTDVKLEQENIANLMHMAVDYGRSIGFKGDFYIEPKPKPEM<br>SHQYDFDAATAIGFLRQYGLDKDFKMNIEANHASLANHTFQHEL YISRINGMLGSVDANQ<br>GNPILGWDTDNFPWNVYDATLAMYEVLKAGGLTGGFNFDSKNRRRPSNTFEDMFHAYIMGM<br>DTFALGLIKA AEI IEDGRIDGFIKEKYSSYESGIGKKIRDKQTTLEELAARAAEMKKPSD<br>PGSGREEYLEGVVNNILFRG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 1754MI22_004 | <i>Firmicutes</i> | DNA              | 45         | ATGAGCGAGTTTTTTAAGAAATATTCTCAAAATAAAATACGAAGGAAAAAGATAGCAAAAAT<br>CCCTGGGCATTC AAGTATTACAATCCTGAATGACAATCATGGGTAAAAAGATGAGCGAA<br>CATCTCCATTCCGAATGGCATGGTGGCATAACCTTGCGCAAAACGGCGTTGATATGTTT<br>GGTCAGGGAACAGCAGACAAAGTCTTTCGGACAGATTCTCTGGAAC TATGGAGCATGC AAAG<br>GCTAAGGTTGATGCTGGTATAGAGTTTATGAAGAAGCTCGGAATCAAATATTACTGCTGG<br>CACGATGTTGACCTTGTTCTTGAGGATCCAAAACGATATCAACGTAACTAA CAAAACGTCTG<br>GACGAAAATTTCAGATTACATCCTTGAAAAGACAAAAGGAACAGACATTAAGTGTCTCTGG<br>GGAAC TGCAAAACATGTTCCGTAACCCCTCGCTTATGAACGGTGCAGGCTCTACAAAAC TCT<br>GCTGACGTTTACTGTTTCTGCTCCGCTCAGGTAAAAAAGGCTCTTGAGATTACTGTAAAG<br>CTTGGTGGCCGAGGTTATGTTTCTGGGGTGGCCGCGAAGGTACGAAAAC TCTTCTCAAT<br>ACAGACGTAAAACTTGAACAGGAAAAATATCGAAACCTCATGCATATGGCTGTTGATTAT<br>GGCCGCTCAATCGGTTTCAAGGGAGACTTCTACATCGAGCCTAAAGCCAAAAGGAGCCAA TG<br>AGCCATCAGTATGATTTTGATGCTGCAACAGCAATCGGCTTCTCCGCCAGTATGGCCTC<br>GACAAAGATTTTAAGATGAACATCGAAGCTAACCATGCCTCACTTGCAAA TCACACCTTC<br>CAGCACGAGCTTTGTATCAGCCGCATAAACGGAATGCTTGGTTCTGTAGATGCAAA TCAG<br>GGAAATCCAATTC TTGGCTGGGATACAGATAA CTTCCCATGGAA TGTTACGATGCAACT<br>CTGGCAATGTACGAAAGTTCTCAAGGCTGGCGGTCTAACAGGTGGCTTCAA ACTTTGACTCA<br>AAGAACCGTCGCCCATCAAATACTTTTGAAGATATGTTCCACGCTTATATCATGGGTATG<br>GATACTTTTGCCCTTGCCCTTATTAAGGCTGCAGAAAATTA TTGAAGACGGCAGAAATTGAC<br>GGCTTCATCAAAGAAAAGTATTCAAGCTTTGAAAGTGGAA TTGGTAAGAAAGATT CGTGAC<br>AAGCAGACAAAGTTTGGAAGAGCTTGCA GCTCGTGCCGCTGAAATGAAAAAGCCATCTGAT<br>CCAGGTT CAGGCCCGCAGGAAATACCTCGAAGGAGTTGTTAACAA CATCCTCTTTCCGCGGA<br>TAA |
| 1754MI22_004 | <i>Firmicutes</i> | Amino Acid       | 46         | MSEFFKNIPOIKYEGKDSKNPWAFKYYNPELTI MGKKMSEHLPFAMAWWHNLGANGVDMF<br>GQGTADKSFQI PGTM EHAKAKVDAGIEFMKKLGIKYCWHVDLVPEDPNDINVTNKRL<br>DEISDYILEKTKGTDIKCLWG TANMFNPRFMNGAGSTNSADVCF AAAQVKKALEITVK<br>LGGRGYVFWGGREGYETLLNTDVKLEQENIANLMHMAVDYGRSIGFKGDFYIEPKPKPEM<br>SHQYDFDAATAIGFLRQYGLDKDFKMNIEANHASLANHTFQHEL CI SRINGMLGSVDANQ<br>GNPILGWDTDNFPWNVYDATLAMYEVLKAGGLTGGFNFDSKNRRRPSNTFEDMFHAYIMGM                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

| TABLE 2    |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.  | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 727MI1_002 | <i>Firmicutes</i> | DNA              | 47         | DTFALGLIKAAEIIEDGRIDGFIKEKYSSFESGIGKKIRDKQTSLEELAARAAEMKKPSD<br>PGSGREEYLEGVVNNILFRG<br>ATGATATTGAAAAATATCCCGCAATTCTTTATGAGGGTCCGAAGAGCACAAAATCCGCTG<br>GCGTTTAAATTCTATGATCCGGACAAAGATCGTTATGGGAAAGCCCATGAAAGGAGCATCTG<br>CCCTTTGCAATGGCTGGTGGCACAAACCTTGGCGGGCCGGAACCGATATGTTCCGGCGC<br>GATACCGCCGACAAATCCTTCGGTGGTAAAGGCACAAATGGAGCATGCCAAAGCGAAA<br>GTCGATGCCGGCTTTGAGTTCAATGCAGAAGCTGGGGATCCGCTATTTCTGCTTCCATGAT<br>GTGGATCTTGTTCCGGAGCGGATGATATAAAGGAGACCAACCGCGTCTGGACGAGATC<br>AGCGATTACATCCTTGAAAAAGATGAAAGGCACCGATATCAAGTGCCCTTTGGGCACGGCC<br>AATATGTTCTCAAAATCCGCGCTTTATGAACGGCGCAGGCTCCTCCAATTTGCCGATGTA<br>TTCGCTTTTGGCGCAGCACAGGCCAAAGAGCCCTTGATCTGACCGTCAAACTCGGCGGG<br>CGGGCTATGCTTCTGGGGCGGACGTGAGGGCTATGAGACACTTCTCAATACCGACATG<br>AAGTTCGAGCAGGAGAATATCGCGAAGCTCATGCATATGGCTGTCGATTACGGCCGCAGC<br>ATAGGCTTTACCGTGATTTCTATATCGAGCCCAAACCGAAGAGCCGATGAAACACACAG<br>TATGATTTTCGATGCAGCCACTGCGATAGGCTTCCTCCGCCAGTACGGACTCGATAAGGAC<br>TTCAAGCTCAACATCGAGGCCAAACCCAGCCACACTGGCAGGTCACACTTCCAGCACGAT<br>CTGCGTGTTCCGCAATAAACGGAATGCTGGGCAGCATTGACGCCAACCCAGGGCGATATG<br>CTCCTCGGCTGGATACCGACGAGTTCCCGTTCAATGTATATGATCGGACCATGTGCATG<br>TATGAGGTGCTCAAGTCAGACGGGCTCACCGGGCTTTAACTTCGACTCCAATCACGC<br>CGCCGAGCTATACGGTCGAGGATATGTTTACAAGCTATATCCTCGGCATGGACACTTTT<br>GCCCTCGGCCTTCTGAAAGCGGCGGAGCTTATCGAAGACGGAAGGCTTGACGCCCTTCGTC<br>AAAGAACGCTATTCAGCTATGAGAGCGGCATCGGCGCAAAGATCCGCAGCGGAGAAAACC<br>GATTTGAAGGAATTGGCGGAATATGCGGACTCCCTCGGAGCCCCGAACTTCGGGGCAGC<br>GGAAAACAGGAACAGCTCGAGAGCATAGTAAATCAGATACTTTTCGGATAA<br>MIFENIPAIPEGPKSTNPLAFKFYDPDKIVMGKPMKEHLPFAMAWWHNLGAAGTDMFGR<br>DTADKSFGAVKGTMEHAKAKVDAGFEFMQKLGIRYFCFHDVDLVPEDDIKETNRRLDEI<br>SDYILEKMKGTDIKCLWGTMFESNPRFMNGAGSSNSADVFAFAAAQAKKALDLTVKLG<br>RGYVFWGGREGYETLLNTDMKFEQENIAKLMHMAVDYGRSIGFTGDFYIEPKPEPMKHQ<br>YDFDAATAIGFLRQYGLDKDFKLNIEANHATLAGHTFQHDLRVSAINGMLGSIDANQGD<br>LLGWDTDEFPFNVYDATMCMYEVLKSDGLTGGFNFDSKSRRPSYTVEDMFTSYILGMDTF<br>ALGLLKAAELIEDGRLDADFVKERYSSYESGIGAKIRSGETDIKELAEYADSLGAPELPGS<br>GKQEQLESIVNQILFG<br>ATGAGCGAGTTTGTGCCAGCATTCGCCAAAAATTCCTTTGAAGGCAAGGACAGCGCCAAT<br>CCCCTGGCGTTCAAATACTACGACGCCGACAGGATGATACTGGGCAAGCCCATGAAGGAG<br>CACCTTCCCTTCGCCATGGCCTGGTGCCACAACTGTGCGCCCGGGCACCCGATATGTTT<br>GGCCGGGACACCCCGACAAAGTCTTCGGCCAGGTCAAGGGCACCATGGAACACGCCCAAG<br>GCCAAGGTGGAACGGGGCTTTGAGTTCATGAAGAAGTGGGCATCCGCTACTTCTGCTTC<br>CACGACGTGGACATCGTGCCCGAAGCCGACGACATCAAGGAAACCAACCGCGTCTGGAC<br>GAGATCTCCGACTATATCTGGAGAAAATGAAAGGCACCGACATCCAGTGCCTGTGGGC<br>ACCGCCAAACATGTTCGGCAACCCCCGCTATATGAACGGCGGGCAGCTCCAACCTCCGCC<br>GACGTATACTGCTTCGCCCCGGCCCCAGATCAAAAAGGCCCTGGACATCACCGTGAAGCTG |
| 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                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|             |                   |                  |            | GGGGCAAGGGCTACGTGTTCTGGGGCGGGCGGAGGGCTACGAGACCCCTGCTGAACACC<br>GATATGAAGTTCGAGCAGGAGAACATCGCCCCGCTGATGCACATGGCCGTGGACTACGGC<br>CGACGATCGGCTTCACCGCGGATTCTACATCGAGCCCAAGCCCAAGAGCCCATGAAG<br>CACCAGTACGACTTCGACGCCGCCACCGCCATAGGCTTTTTCGCCCAGTACGGCCTGGAC<br>AAGGATTTCAAGCTGAACATCGAGTCCAACCCACGCCACCTGGCGGCCATACCTTCCAG<br>CAGACCTGCGCGTTTCCGCCATCAACGGCATGCTGGCTCCATCGACGCCAACCAAGGC<br>GACTACCTGCTGGCTGGGATACCGACGAGTCCCCTACAGCGTATACGAGACCAACCATG<br>TGATGTACGAGGTGCTCAAGGCCGGAGGTCTCACCGCGGCTTCAATTCGACGCCAAG<br>AACCGCCGCCAGCTACACCCCGAGGATATGTTCCACGCCATACATCCTTGGGATGGAC<br>AGCTTCGCCCTGGCCTGATCAAGGCCGCCGAGCTCATCGAGGACGGTCGCTGGACGCC<br>TTCGTCGGGACCGCTACCAAGACTGGGAGACCGGCATCGCGGATAAGATCCGCAAGGGC<br>GAGACCACACTGGCCGAGTGGCCGAGTACGCCGCCGATGGCGCGGCCCGCTGCC<br>GGACGGCGCCAGGAATACCTGGAGGGCTGGTCAACAATATCCTGTTCAAAATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 727MI9_005  | <i>Firmicutes</i> | Amino Acid       | 50         | <u>MSEFFASIPKIPFEGKDSANPLAFKY</u> <u>YDADRMILGKPMKEHL</u> <u>PFAMAWHNLCAAGTDMF</u><br><u>GRDTADKSFQGVKGTMEHAKAKVDAGFEFMKKLIGIRYFCFHDVDIVPEADDIKETNRRLD</u><br><u>EISDYILLEKMGTDIQCLWGTANMFNPRYMNGAGSSNSADVYCFAAAQIKKALDITVKL</u><br><u>GGKGYVFWGGREGYETLLNTDMKFEQENIARLHMVADYGRSIGFTGDFYIEPKPEPMK</u><br><u>HQYDFDAATAIGFLRQYGLDKDFKLNIESNHATLAGHTFQHDLRVSAINGMLGSIDANQG</u><br><u>DYLLGWDTDEFPYSVYETTMCMYEVLKAGGLTGGFNFD</u> <u>AKNRRRPSYTPEDMFHAYIILGMD</u><br><u>SFALGLIKAAELIEDGRLD</u> <u>AFVRDRYQSWETGIGDKIRKGETTLAELAEYAAARMGAPALP</u><br><u>GSGRQYLEGVVNNILFK</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 727MI27_002 | <i>Firmicutes</i> | DNA              | 51         | ATGAAGACCTATTTCAAAAAAATCCCCGTGATCCCCTACGAGGGACCGAAATCCCAGAAT<br>CCGTGTGCTTCAAATTCTATGACGGGACCGCATCGTTCTCGGCAAGCCCATGAAGGAG<br>CATCTGCCCTTCGCCATGGCTGGTGGCACAAATCTGGTGCTGCCGGAACGGACATGTTT<br>GGCCGGATACCCCGACAAAGTCCTTCGGAGCGGAGAAAGGCACCATGGAGCATGCCAAG<br>GCCAAGGTGGACCGCTGGCTTCGAGTTTATGAAGAAGTGGGCATCCGGTATTTCTGCTTC<br>CATGACGTGGATCTGGTCCCGGAAGGGACGACATCAAGGAGACCAACCGCGTCTCGAT<br>GAGATCAGCGACTACATCCTCAAGAAAGATGAAGGGCAGGATATCAAGTGCCTCTGGGGC<br>ACCGCCAACATGTTCCGGCAATCCCCGGTTCATGAACGGCGCGGCAGCTCCAACAGCGCG<br>GACGTGTTCTGCTTTCGCCGGCCCCAGGTGAAGAAGGCCTTGACATCACCGTCAAGCTG<br>GGCGCCGGGGCTATGTGTTCTGGGGCGCCGTGAGGGGTATGAGTCCCTGCTGAACACG<br>GACGTGAAGTTTGACAGGAGAACATCGCCAAAGCTCATGCACCTTGCCGTGGACTACGGC<br>CGACGATCGGCTTCACCGCGGATTCTACATCGAGCCCAAGCCCAAGGAGCCCATGAAG<br>CACCAGTACGACTTCGATGCCGCCACCGCCATCGGCTTCCTCAGGCAGTACGGCCTCGAT<br>AAGGACTTCAAGATGAACATTGAAGCCAACCAACGACCCCTGGCCGCCACACCTTCCAG<br>CAGACCTCAGGATCAGGCCCATCAACGGGATGCTGGGCTCCATCGACGCCAACCAAGGC<br>GACCTCCTGCTGGATGGGACACCGACGAATTCCCCTTCAACGTCTATGAGGCCACCATG<br>TGCATGTACGAGGTCTCAAGGCCGGCGGCTCACCGGGGCTTCAACTTCGACTCAAAG<br>AACCGCCGTCCCTCCTACACCATGGAGGATATGTTCCACGCCTACATCCTGGGCATGGAC<br>ACCTTCGCCCTGGTCTTCTCAAGGCCGGGAGCTCATCGAGGACGGTCGGATCGACAAA<br>TTCGTGGAGGAGCGCTACGCCAGCTACAAGACCGGCATCGGGGCCAAGATCCGTTCCGGC |

| TABLE 2     |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|---------------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 727MI27_002 | <i>Firmicutes</i>         | Amino Acid       | 52         | GAGACCAGCTTCAGGAGCTGGCCGCCTATGCCGACAAAGTTGGCGCGCCTGCCCTTCCC<br>GGCAGCGGCGTCAGGAGTACCTGGAGAGCATCGTCAACCAGGTGCTCTTCGGGATGTGA<br>MKTYFKKIPVIPYEGPKSQNPLSEKFYDADRIVLGKPMKEHLPFAMAWHNLGAAGTDMF<br>GRDTADKSFGAEKGTMEHAKAKVDAGFEFMKKVGIRYFCFHDVDLVP EADD DIKETNRRLD<br>EISDYILKKMGTDIKCLWGTANMFGNPRFMNGAGSSNSADVFCFAAAQVKKALDITVKL<br>GGRGYVFWGREGYESLINTDVKFEQENIAKLMHLAVDYGRSIGFTGDFYIEPKPEPMK<br>HQYDEDAATAIGFLRQYGLDKDFKMNIEANHATLAGHTFQHDLRISAINGMLGSIDANQG<br>DLLLLGWDTDEFPPNVYEATMCMYEVLKAGGLTGGFNFD SKNRRPSYTMEDMFHAYIILGMD<br>TFALGLLLKAAELIEDGRIDKFVEERYASYKTGTGIGAKIRSGETTTLQELAAAYADKLGAPALP<br>GSGRQEYLESIVNQVLFGM                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 1753MI2_006 | <i>Neocallimastigales</i> | DNA              | 53         | ATGGCTAAAGAGTATTTTCCAGAGATTGGCAAAATCAAGTTTGAAGGCAAGGACAGCAAA<br>AACCCAAATGGCTTTCCTACTACTATGACCCCCGAGAAGGTGATCATGGGCAAGCCTATGAAA<br>GACTGGCTCCGCTTCGCTATGGCATGGTGGCACACCCCTCTGCCAGAAGGTGGCGACCAG<br>TTCGGTGGCGGCACATAAGAAAGTTCCCTTTGGAACAACGGCGCTGACGCTGTAGAAATCGCA<br>AAACAGAAAGGCTGACGCAGGTTTCGAAATCATGCAGAAAGCTCGGCATCCCATATTTCTGC<br>TTCCACGACGTGGACCTCGTGTCTGAGGGCGCATCTGTAGAAGAGTATGAGGCTAACCTC<br>AAGGCTATCACAGACTACCTCGCTGTGAAGATGAAGGAAACAGGCATCAAGCTCCTGTGG<br>TCTACTGCCAACGTATTTCGGCAACGGCCGCTACATGAACGGTGCTTCTACCAACCCCTGAC<br>TTCGACGTCGTTGCTCGCGCTATCGTGCAGATTAAAGAACGCTATCGACGCTGGTATCAAG<br>CTCGGCGCTGAGAACTACGTGTTCTGGGCGGACGCGAAGGTACATGAGCCTCCTCAAC<br>ACCGACCAAGACCGTGAGAAAGGAGCACATGGCCACTATGCTCACTATGGCTCGCGACTAC<br>GCTCGCGCTAAGGCTTCAAGGGCACATTCTTCATCGAGCCTAAGCCAAATGGAGCCTTCT<br>AAGCACCAGTATGACGTTGACACTGAGACTGTCTATCGGCTTCTCTCAAGGCACACAACTC<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCAACTCTCGTGGCCACACCTTC<br>GAGCACGAGCTCGCAGTGGCAGTGGACAACAACATGTCGGCTCTATCGACGCTAACCGT<br>GGTGACTACCAGAAATGGCTGGGATACTGACCACTTCCCAATCGACCAGTACGAACTCGT<br>CAGGCTTGGATGGAATCATCCGTGGCGGGGTCTCGGCACTGGCGGCACGAACTTCGAC<br>GCTAAGACTCGTCGTAACCTACCGACCTCGAAGACATCTTCATCGCACACATCGCAGGC<br>ATGGACGCTATGGCACGCGCACTCGAATCAGCTGCTAAAGCTCCTCGAAGAGTCTCCATAC<br>AAGGCAATGAAGGCAGCTCGCTACGCTTCATTTCGACAACGGTATCGGTAAGGACTTCGAA<br>GATGGCAAGCTCACTCTCGAGCAGGCTTACGAATACGGTAAGAAGGTGGTGAGCCCTAAG<br>CAGACTTCTGGCAAGCAGGAGCTCTACGAAGCCATCGTTGCAATGTACGCTTAA |
| 1753MI2_006 | <i>Neocallimastigales</i> | Amino Acid       | 54         | MAKEYFPEIGKIKFEGKDSKNPMAFYHYDPEKIVIMGKPMKDWLRFAMAWHNLCAEGGDQ<br>FGGGTKKFPWNNGADAVEIAKQKADAGFEIMQKLGIPYFCFHDVDLVSEGASVEEYEANL<br>KAITDYLAVKMKETGIKILWSTANVFNGRYMNASTNPDFDVVARAIVQIKNAIDAGIK<br>LGAENYVFWGREGYMSLLNTDQKREKEHMATMLTMARDYARAKGFKGTFLIEPKPEPS<br>KHQYDVDTETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELAVAVDNNMLGSI DANR<br>GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGLGTGGTNFDAKTRNSTDLEDIFIAH IAG<br>MDAMARALESAAKLLEESPYKAMKAARYASFNDNGIGKDFEDGKLTLEQAYEYGKKVGEPK<br>QTS GKQELYEAI VAMYA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| TABLE 2      |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--------------|---------------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI3_005  | <i>Neocallimastigales</i> | DNA              | 55         | ATGGCTAAAGAAATTTTCCACAGAGATTGGTAAAAATCAAGTTCGAAGGCAAGGATTCAAAG<br>AATCCAATGGCTTTCATTACTATGATGCAGAGAAGGTAATCATGGGCAAAACCCATGAAG<br>GACTGGCTCCGTTTCGCTATGGCATGGTGGCACACACTCTGTGCAGAGGGCGGCAGCAG<br>TTCGGTGGCGGTACGAAGAAAGTTCCCTTGGAACGAGGGTGTCTAATGTGTGAGATTGCT<br>AAGCAGAAAGGCTGACGCTGGTTTCGAAATCATGCAGAAAGCTTGGCATTCCTTACTTCTGC<br>TTCACGATGTTGACCTCGTTTCTGAAGGCGCATCTGTTGAGGAGTATGAGGCCAACCTC<br>AAGCTATCACTGACTATCTCGCGGTGAAGATGAAGGAGACTGGCATTAAGCTCCTGTGG<br>TCTACTGCCAACGTGTTCCGCAATGGCCGTTACATGAATGGTGCTTCCACCAACCCGTGAC<br>TTCGACGTTGTTGCTCGGCCCATCGTTCAGATTAAAGAACGCTATCGATGCAGGTATCAAG<br>CTCGGTGCTGAGAACTATGTGTTCTGGGCGGTCTGTAAGGTTACATGAGCCTCCTGGAAC<br>ACAGACCAGAAGCGTGAGAAAGGAGCACATGGCTACTATGCTCACTATGGCTCGCGACTAC<br>GCTCGCAGCAAGGGCTTCAAGGGTACTTTCCTCATCGAGCCTAAGCCAAATGGAGCCATCT<br>AAGCACCACTACGACGTTGACACAGAGACTGTTATCGGCTTCTTGAAGGCACACAAACCTT<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGAACTCGTGGTCACACCTTC<br>GAGCACGAGCTCGCTGTGGCTGTCGACAACTATGCTTGGTTCTATCGATGCTAAACCGC<br>GGTGACTACCGAATGGTTGGGATACGGACCAAGTCCCAATTGACCAGTACGAGCTCGTT<br>CAGGCTTGGATGGAGATCATCCGTGGTGGCGGTCTCGGCACAGGTGGTACAAACTTCGAC<br>GCTAAGACTCGTCGTAACCTACCGACCTCGAGGACATTTTCATTGCTCACATCGCTGGT<br>ATGGACGCTATGGCTCGGCTCTTGAGTCAGCAGCTAAGCTCCTTGAGGAGTCTCCATAC<br>AAGAAGATGAAGGCTGCCCGTTATGCTTCTTCGACAGCGGCATGGTAAAGGACTTTGAG<br>AACGGCAAGCTCACACTCGAACAGGTTTATGAGTATGTTAAGAAAGGTAGGTGAGCCCCAAG<br>CAGACTTCTGGCAAGCAGGAGCTCTTCGAGGGCAATCGTGGCCATGTACGCAATAA |
| 5586MI3_005  | <i>Neocallimastigales</i> | Amino Acid       | 56         | <u>MAKEFFPEIGKIKFEGKDSKNPMAFHYYDAEKVIMGKPMKDWLRFAMAWHHTLCAEGGDQ</u><br><u>FGGTTKKFPWNEGANAWEIAKQKADAGFEIMQKLGIPIYFCFHDVDLVSEGASVEEYEANL</u><br><u>KAITDYLAVKMKETGIKLLWSTANVFNGNRYMNGASTNPDFDVVARAIVQIKNAIDAGIK</u><br><u>LGAENYVFWGGREGYMSLLNTDQKREKEHMA TMLTMARDYARSKGFKGTFLIEPKPMEPS</u><br><u>KHQYDVDTETVIGFLKAHNLDKDFKNIEVNHATLAGHTFEHELAVAVDNNMLGSI DANR</u><br><u>GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGLGTGGTNFDAKTRRNSTDLEDIFIAHIAG</u><br><u>MDAMARALESAAKLLEESPYKKMKAARYASFDSGMGKDFENGKLTLEQVVEYGKKVGEPK</u><br><u>QTS GKQELFEAIVAMYA</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI91_002 | <i>Neocallimastigales</i> | DNA              | 57         | ATGGCTAAAGAGTATTTTCCAGAGATTGGTAAAAATCAAGTTTGAAGGCAAGGATTCCAAG<br>AATCCAATGGCATTCCACTATTATGATGCAGAGAAAGTGATTATGGGTAAAGCCTATGAAG<br>GAGTGGCTCCGCTTTGCAATGGCATGGTGGCACACACTCTGTGCAGAGGGTGGCGAACAG<br>TTTGGTGGTGGCACTAAGAAATTCCCATGGAACGAGGGCACTGACGCTGTGACGATTGCT<br>AAGCAGAAAGGCTGATGCAGGTTTCGAAATCATGCAGAAACTCGGTTTCCCATATTTTGC<br>TTCACGACATTGACCTCGTTTCGGAAGGCAACAGCATTGAAGAGTATGAGGCTAACCTC<br>CAGGCAATCACTGATTATCTGAAAGTGAAGATGGAAGAGACAGGCATCAAACTCTTGTGG<br>TCAACTGCCAACGTTATTCGGCAATGGTCGCTACATGAATGGTGCTTCCACAAACCCAGAC<br>TTTGACGTGGTGGCTCGTGCCATCGTTCAGATTAAAGAACGCAATTGACGCTGGTATCAAA<br>CTCGGTGCTGAGAACTATGTATTCTGGGGCGGTGCGGAAGGCTACATGAGCCTTCTGAAC<br>ACTGACCAGAAGCGTGAGAAAGGAGCACATGGCAACCATGCTCACTATGGCTCGCGACTAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| TABLE 2       |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------|---------------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|               |                           |                  |            | GCTCGAGCAAGGGTTTCAAGGGCACTTTCCTCATTGAGCCAAAAGCCAAATGGAGCCATCT<br>AAGCACCAAGTATGACGTTGACACGGAGACTGTCTCGGCTTCCTCAAGGCACACAAACCTC<br>GACAAGGATTTCAAGGTGAACATCGAAGTGAACCAACGCTACACTTGCAGGTCTACTATTTC<br>GAGCACGAACTTGTGTGGCTGTTGACAAATGGCATGCTCGGTTCCTATCGACGCTAAACCGT<br>GGTGACTATCAGAAACGGTTGGGACACTGACCAAGTCCCAATCGACCAGTACGAACTCGTT<br>CAGGCTTGGATGGAATCATCCGTGGTGGTCTCGGCACACAGGTGGTACTAACTTCGAT<br>GCTAAGACTCGTCGTAACCTCAACTGACCTCGAGGACATCTTCATCGCACACATCTCTGGT<br>ATGGATGCAATGGCACGTGCTCTCGAATCGGCGGCTAAACTTCTTGAGGAGTCTCCATAC<br>TGGCTATGAAGAAAGGCTCGTTACGCTTCCTTCGACAGCGGCATCGGTAAAGACTTCGAG<br>GACGGCAAACTCACGCTCGAGCAGGCTTACGAGTACGGCAAGAAAAGTCGGCGAACCACAAG<br>CAGACTTCTGGCAAGCAGGAACTCTACGAGGCAATCGTTGCCATGTACGCATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 5586MI191_002 | <i>Neocallimastigales</i> | Amino Acid       | 58         | MAKEYFPEIGKIKFEGKDSKNPMAFHYYDAEKVIMGKPMKEWLRFAMAWHHTLCAEGGDQ<br>FGGGTKKFPWNEGTDAVTIAKQKADAGFEIMQKLGFPEYFCFHDIDLVSEGNSIEEYEANL<br>QAITDYLLKVKMEETGIKLLWSTANVFGNGRYMNGASTNPDFDVVARAIVQIKNAIDAGIK<br>LGAENYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARSKGFKGTFLIEPKPMEPS<br>KHQYDVDTETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSI DANR<br>GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGLGTGGTNEDAKTRNSTDLEDIFIAHISG<br>MDAMARALESAAKLLEESPYCAMKKARYASFDSIGDKDFEDGKLTLEQAYEYGKKVGEPK<br>QTSKGQELYEAI VAMYA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5586MI194_003 | <i>Neocallimastigales</i> | DNA              | 59         | ATGGCAAAAAGAGTATTTCCCTACGATCGGTAAGATCGTTTATGAGGACCCGGAGTCCAAG<br>AACCCTATGGCATTTCTATTACTATGACGCAGACGCGGTAGTAGCTGGTAAAAAATGAAA<br>GATTGGATCGGTTTCGCTATGGCATGGTGGCACACCCCTCTGTGCAAGAGTGCAGACCAG<br>TTCGGTGGAGGCACCAAAACACTTCCC GTGAGTGAAGGTCCC GATGCCGTAAACCATCGCC<br>AAGCAGAAAGCAGACGCAGGTTT TGAGATCATGCAGAAAAC TCGGCTTCCC GTATTCTGT<br>TTCCATGACGTGGATCTGTCAGCGAAGGCAGCAGCGTAGAAGAGTACGAGGCGAAACCTC<br>GCAGCCATCACCGATTATCTCAAGCAGAAAA TGACGAGT CCGGTATCAAACTCCTTTGG<br>TCCACTGCTAACGTATTCCGTCACGCCCGTTACATGAACGGTGCCAGCACCAATCCTTGAC<br>TTTGATGTGTTGCCGTGCGATTGTGCAGATCAAGAA TGTATCGACGCAGGTATCAAA<br>CTCGGCGCAGAGAACTACGTCTTCTGGGCGGTCTGTGAAGGTTATATGAGCCTGCTCAAT<br>ACCGACCAAGAAACGCGAGAAAGAGCATACGGCAATGATGCTGCGTATGGCGCGTGACTAT<br>GCCCCAGCAAAAGGTTTCAAAGGTACCTTCCTCATCGAACCCAAACCCATGGAGCCGTCC<br>AAGCACCAGTATGACGTAGATACCGAGACGGTGTAGGTTTCTCTCAAAGCACACGGTTTG<br>GAGAAAGACTTTAAAGGTAACATCGAAGTGAACCAACGCTACCCCTGCCGGTCACACTTTC<br>GAGCACGAACTGGCAGTAGCCGTAGATAACCGGCATGCTCGGTTCGATCGATGCCAAACCGC<br>GGTGACTATCAGAACGGATGGGATACCGACCAAGTTCGCCATCGATAACTTCGAACTGACC<br>CAAGCATGGATGCAGATCGTACGTAACGGTGGTCTCGGCACAGGCGGAACGAACTTCGAC<br>TCCAAGACCCGT CGTAACTCCACCGATCTCGAGGATATCTTCATCGCTCACATCAGTGGT<br>ATGGACGCTTGTGCCCGTGCCCTATTGAATGCCGTAGAGATCATGGAGAAATCACCGGATC<br>CCTGCTATGCTCAAAGAGCGTTACGCTTCCTTCGATAGCGGTCTGGTAAAGATTTCGAG<br>GACGGCAAACTGACCCCTTGAGCAAGTCTATGAGTACGGTAAGAAAGTAGGCGAACCCTAAA<br>CAAAACGCGGCAAAACAAGAACTCTATGAGGGCTATCGTTGCCCTCTACGCTAAATAA |

| TABLE 2       |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|---------------|---------------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 5586MI194_003 | <i>Neocallimastigales</i> | Amino Acid       | 60         | MAKEYFPTIGKIVYEGPESKNPMAFYHYDAERVVAGKKMKDWMRFAMAWHHTLCAEGADQ<br>FGGGTKHFPWSEGPDAVTIAKQKADAGFEIMQKLGFYPFCFHDVDLVSEGSSVEEYEANL<br>AAITDYLKQKMDESGIKLLWSTANVFGHARYMNGASTNPDFDVARAIVQIKNAIDAGIK<br>LGAENYVFWGGREGYMSLLNTDQKREKEHTAMMLRMARDYARSKGFKGTFLIEPKPMEPS<br>KHQYDVDTETVIGFLKAHGLEKDFKVNIEVNHATLAGHTFEHELAVAVDNGMLGSI DANR<br>GDYQNGWDTDQFPIDNFELTQAWMQIVRNGGLGTGGTNFDSKTRNSTDLEDIFIAHISG<br>MDACARALLNAVEI MEKSPIPAMLKERYASFDSGLGKDFEDGKLTLEQVVEYGKKVGEPK<br>QTSKQELYEAI VALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI198_003 | <i>Neocallimastigales</i> | DNA              | 61         | ATGAAAGAGTATTTCCTGAGATCGGTAAGATCCAATTTGAAGGCCCGGAGTCCAAAGAAC<br>CCGATGGCATTTCAC TACTATGACGCAGAGCGCGTCGTAGCCGGTAAAAACAATGAAAGAG<br>TGGATGCCGTTTCGCTATGGCTTGGTGGCACACCCCTCTGTGCGGAAGCGCGCACCACTTC<br>GGAGCGGAAACGAAAGAAGTTCCTCTGGAACGAAGGCGCTAACGCTTTGGAGATCGCCAAAG<br>CACAAAGCCGATGCGGGATTGAGATCATGCAGAAACTCGGCATCCCTTATTCTGTCTTC<br>CATGACGTGGATCTCATCGCCGAGGGCGGTTCCGTTAGAGAGTACGAAGCCAACCTCGCT<br>GCCATCACCGATTACCTCAAACAGAAAAATGGACGAGACTGGCATCAAAC TGCTGTGGTCC<br>ACGGCGAACGTCCTCAGCAACCCCCGTTATATGAACGGCGCCAGCACGAACCCCGATTTC<br>GATGTAGTAGCGGTGCCATCGTCCAGATCAAGAACGCTATCGACGCCGGTATCAAAC T C<br>GGAGCAGAGAACTATGTCTTCTGGGTGGTCGCGAGGGCTATATGAGCCTCCTCAACACT<br>GACCAGCGCCGAGAGAAAGACATATGGCTACCATGCTCCGTATGGCGGTGACTACGCG<br>CGTGCCAAAGGATTCAAGGGCACCTTCCTCATCGAAACCAAAACCATGTGAGCCGTCCAAA<br>CATCAGTATGATGTCGATACCGAGACCGTCATCGGTTTCCTCAAAGCGCATGGACTCGAC<br>AAGGATTTCAAAGTCAATATCGAGGTCAACCAACGCCACCCCTCGCAGGCCACACGTT CGAA<br>CACGAACTGGCTTGGCTGTAGATGCCGGCATGCTCGGTTCGATTGACGCCAACCGCGGT<br>GACGCCCAGAACGGATGGGACACCGACCGAGTTCCCTATTGATAAACTTCGAACTCACACAG<br>GCTTTCATGCAGATCGTCCGCAACGGCGGTTTCGGAACAGGCGGTACGAACTTCGACGCC<br>AAGACACGCCGTAACTCCACCGACTTGGAGGACATCTTCATCGCCCATATCAGCGGCATG<br>GACGCTTGCGCACGTGCGTTACTCAATGCTGTCGAAATCCTCGAGAAGAGCCCGATTCCG<br>GCGATGCTCAAAGAGCGTTATGCTTCCTTTGACGGCGGCATCGGAAAGGACTTCGAGGAG<br>GGAAAACTGACTTTCGAGCAGGTCTATGAGTACGGCAAGAAAGTCGGCGAAACCCCAACAG<br>ACCAGCGGCAAAACAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG |
| 5586MI198_003 | <i>Neocallimastigales</i> | Amino Acid       | 62         | MKEYFPEIGKIQFEGPESKNPMAFYHYDAERVVAGKTMKEWMRFAMAWHHTLCAEGGDQF<br>GGGTTKKFPWNEGANA LEIAKHKADAGFEIMQKLGI PYFCFHDVDLIAEGGSVEEYEANLA<br>AITDYLKQKMDETGIKLLWSTANVFSNPRYMNGASTNPDFDVARAIVQIKNAIDAGIKL<br>GAENYVFWGGREGYMSLLNTDQRREKEHEMATMLRMARDYARAKGFKGTFLIEPKPCEPSK<br>HQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG<br>DAQNGWDTDQFPIDNFELTQAFMQIVRNGGFGTGGTNFDAKTRNSTDLEDIFIAHISGM<br>DACARALLNAVEI LEKSPIPAMLKERYASFDDGGIGKDFEEGKLTTFEQVVEYGKKVGEPKQ<br>TSGKQELYETI VALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 5586MI201_003 | <i>Neocallimastigales</i> | DNA              | 63         | ATGGCAAAAGAGATATTCCCTACGATCGGTAAGATCGTTTATGAAGGACCGGAATCCAAG<br>AACCCTATGGCATTTCA TTTACTATGACGCAGAGCGCGTAGTAGCTGGTAAAAAATGAAA<br>GATTGGATGCGGTTTCGCTATGGCATGGTGGCACACCCCTCTGTGCAGAAGGTGCAGACCAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| TABLE 2       |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------|--------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|               |                    |                  |            | TTCCGGTGGAGGCACCAAAACACTTCCCCTGGAAATGAAGGTCCTCGATCCCCTAACCATCGCC<br>AAGCAGAAAGCAGACGCAGGTTTGGAGATCATGCAGAAACTCGGCTTCCCCTATTTCTGT<br>TTCCATGACGTGGATCTGGTCGGCGAAGCAGCAGCGTAGAAGATACGAGGCGAACCTC<br>GCAGCCATCACCGATTATCTCAAGCAGAAAAATGGACGAGTCGGGTATCAAACCTCCTTTGG<br>TCCACTGCTAACGTATTCCGTACGCCCTTACATGAACGGTGCCAGCACCAATCCTGAC<br>TTTGATGTCTGTTGCCCCGTCGATTGTGCAGATCAAGAAATGCTATCGACGCAGGTATCAAA<br>CTCGGCGCAGAGAACTACGTCTTCTGGGGCGGTCTGTGAAGGTATATGAGCCTGCTCAAC<br>ACCGACCAGAAACGGCAGAAAGAGCATACGGCAATGATGCTGCGTATGGCGGTGACTAT<br>GCCGCAGCAAAGGTTTCAAAGGTACCTTCTCATCGAACCCAAACCCATGGAGCCGTCC<br>AAGCACCAAGTATGACGTAGATACCGAGACGGTGATAGGTTTCTCAAAGCACACGGTTTG<br>GAGAAAGACTTTAAGGTAACATCGAAGTGAACCCACGCTACCTCGCCGGTCACACTTC<br>GAGCACGAACTGGCAGTAGCCGTAGATAACGGCATGCTCGGTTCGATCGATGCCAACCGC<br>GGTACTATCAGAACGGATGGGATACCGACCACTTCCCCATCGATAACTTCGAACTGACC<br>CAAGCATGGATGCAGATCGTACGTAACGGTGGTCTCGGCACAGCGGGAACGAACTTCGAC<br>TCCAAGACCCGTCGTAACCTCACCGATCTCGAGGATATCTTCATCGCTCACATCAGTGGT<br>ATGGACGCTTGTGCCCCGTCCTTATGAATGCCGTAGAGATCATGGAGAAATCACCGATC<br>CCTGCTATGCTCAAAGAGCGTTACGCTTCTTCGATAGCGGTCTGGGTAAGATTTCGAG<br>GACGCCAAACTGACCCCTTGAGCAAGTCTATGAGTACGGTAAGAAAGTAGGCCGAACCCAAA<br>CAAAACGCGGCAACAAGAACTCTATGAGGCTATCGTTGCCCTCTACGCTAAATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI201_003 | Neocallimastigales | Amino Acid       | 64         | <u>MAKEYFP</u> <u>TI</u> <u>GKI</u> <u>VE</u> <u>GP</u> <u>ES</u> <u>KN</u> <u>PM</u> <u>AF</u> <u>HY</u> <u>YD</u> <u>AE</u> <u>RV</u> <u>V</u> <u>AG</u> <u>KK</u> <u>MD</u> <u>WM</u> <u>RF</u> <u>AM</u> <u>AW</u> <u>WH</u> <u>TL</u> <u>CA</u> <u>EG</u> <u>AD</u> <u>Q</u><br><u>FG</u> <u>GG</u> <u>TK</u> <u>HF</u> <u>PN</u> <u>EG</u> <u>PD</u> <u>AV</u> <u>TI</u> <u>AK</u> <u>Q</u> <u>K</u> <u>AD</u> <u>AG</u> <u>FE</u> <u>IM</u> <u>Q</u> <u>KL</u> <u>FP</u> <u>Y</u> <u>FC</u> <u>FD</u> <u>VD</u> <u>LV</u> <u>GE</u> <u>SS</u> <u>VE</u> <u>EY</u> <u>EA</u> <u>N</u> <u>L</u><br><u>AA</u> <u>IT</u> <u>D</u> <u>Y</u> <u>L</u> <u>K</u> <u>Q</u> <u>K</u> <u>MD</u> <u>ES</u> <u>GI</u> <u>K</u> <u>LL</u> <u>W</u> <u>ST</u> <u>AN</u> <u>V</u> <u>F</u> <u>G</u> <u>H</u> <u>AR</u> <u>Y</u> <u>M</u> <u>NG</u> <u>AS</u> <u>TN</u> <u>PD</u> <u>FD</u> <u>V</u> <u>VA</u> <u>RI</u> <u>V</u> <u>Q</u> <u>I</u> <u>K</u> <u>NA</u> <u>ID</u> <u>AG</u> <u>IK</u><br><u>LG</u> <u>AE</u> <u>NY</u> <u>V</u> <u>FW</u> <u>GG</u> <u>RE</u> <u>GY</u> <u>MS</u> <u>LL</u> <u>NT</u> <u>DQ</u> <u>K</u> <u>RE</u> <u>KE</u> <u>HT</u> <u>AM</u> <u>LM</u> <u>RL</u> <u>M</u> <u>AR</u> <u>D</u> <u>Y</u> <u>AR</u> <u>SK</u> <u>G</u> <u>F</u> <u>K</u> <u>G</u> <u>T</u> <u>FL</u> <u>IE</u> <u>PK</u> <u>P</u> <u>ME</u> <u>PS</u><br><u>KH</u> <u>Q</u> <u>Y</u> <u>D</u> <u>V</u> <u>D</u> <u>T</u> <u>E</u> <u>T</u> <u>VI</u> <u>G</u> <u>FL</u> <u>KA</u> <u>H</u> <u>G</u> <u>LE</u> <u>K</u> <u>D</u> <u>FK</u> <u>V</u> <u>N</u> <u>IE</u> <u>V</u> <u>N</u> <u>H</u> <u>AT</u> <u>L</u> <u>AG</u> <u>HT</u> <u>FE</u> <u>HE</u> <u>LA</u> <u>VA</u> <u>VD</u> <u>N</u> <u>G</u> <u>ML</u> <u>GS</u> <u>I</u> <u>D</u> <u>AN</u> <u>R</u><br><u>GD</u> <u>Y</u> <u>Q</u> <u>NG</u> <u>W</u> <u>D</u> <u>T</u> <u>D</u> <u>Q</u> <u>FP</u> <u>I</u> <u>DN</u> <u>F</u> <u>EL</u> <u>T</u> <u>Q</u> <u>AW</u> <u>M</u> <u>Q</u> <u>I</u> <u>VR</u> <u>NG</u> <u>GL</u> <u>GT</u> <u>GT</u> <u>N</u> <u>FD</u> <u>SK</u> <u>TR</u> <u>NR</u> <u>ST</u> <u>D</u> <u>LE</u> <u>D</u> <u>I</u> <u>F</u> <u>I</u> <u>A</u> <u>H</u> <u>IS</u> <u>G</u><br><u>MD</u> <u>AC</u> <u>AR</u> <u>AL</u> <u>LN</u> <u>AVE</u> <u>I</u> <u>ME</u> <u>K</u> <u>SP</u> <u>I</u> <u>P</u> <u>AM</u> <u>I</u> <u>K</u> <u>ERY</u> <u>AS</u> <u>FD</u> <u>SG</u> <u>LG</u> <u>KD</u> <u>FD</u> <u>EG</u> <u>KL</u> <u>TL</u> <u>EQ</u> <u>VY</u> <u>EY</u> <u>G</u> <u>K</u> <u>K</u> <u>V</u> <u>GE</u> <u>PK</u><br><u>Q</u> <u>T</u> <u>S</u> <u>G</u> <u>K</u> <u>Q</u> <u>E</u> <u>L</u> <u>Y</u> <u>E</u> <u>A</u> <u>I</u> <u>V</u> <u>A</u> <u>L</u> <u>Y</u> <u>A</u> <u>K</u> |
| 5586MI204_002 | Neocallimastigales | DNA              | 65         | ATGAAAGAGTATTTCCCTGAGGTCGGTAAGATCCAATTTGAAGGCCCGGAGTCTAAGAAC<br>CCGATGGCATTTCACTACTATGACGCAGAGCGCGTCGTAGCCGGTAAAAACAATGAAAGAG<br>TGGATGCGTTTCGCTATGGCTTGGTGGCACACCCCTCTGTGCAGAAGGCGGACCACTTC<br>GGAGGCGGAACGAAGCATTTCCCGTGGAATGAAGGCGCTAACGCTTTGGAGATCGCCAAA<br>CAGAAAGCCGATGGGGATTCGAGATCATGCAGAAACTCGGCATCCCCTATTTCTGTTTC<br>CATGACGTGGATCTCATCGCCGAGGGCGGTTCCGGTAGAAGAGTACGAAACCAACCTCGCT<br>GCTATCACCGACTACCTCAAGCAGAAAAATGGACGAGACCGGCATCAAACCTGCTGTGTTCC<br>ACGGCGAACGTGTTCAGCAACCCCCGTTATATGAACGGCGGAGCACGAACCCCGATTTC<br>GATGTAGTAGCGGTGCCATCGTGCAGATCAAGAAATGCCATCGACGCCGGCATCAAACCTG<br>GGGCAGAGAACTATGTCTTCTGGGGCGGTCCGAGGGCTACATGAGCCTGCTCAACACC<br>GACCAGCGCCGAGAGACATATGGCTACTATGCTCCGTATGGCGGTGACTACGCG<br>CGTGCCAAAGGATTCAAGGGCACCTTTCTCATCGAAACCCAAACCGGTGTGAGCCGTCCAAA<br>CATCAGTATGATGTCGATACCGAGACCGTCATCGGTTTCTCTCAAAGCGCATGGACTCGAC<br>AAGGATTTCAAGGTTAATATCGAGGTCAACCCACGCCACCCCTCGCAGGCCACACGTTTCGAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| TABLE 2       |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------|---------------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|               |                           |                  |            | CACGAACTGGCTTGGCGTGTAGATGCCGGCATGCTCGGTTTCGATTGACGCCAACCCGCGGT<br>GACGCCAGAACCGGATGGGACACCGACCAGTTCCTATTGATAAATTTCGAACTCACACAG<br>GCTTTCATGCAGATCGTCCGCAACGGCGGTTTCGGAACAGCGGTACGAACTTCGACGCC<br>AAGACACGCCGTAACTCCACCGACTTGGAGGACATCTTCATCGCCCATATCAGCGGCATG<br>GACGCTTGGCACGTGCGTTGCTCAACGCCCATCGAAATCCTCGAGAAAGAGCCCGATCCCG<br>GCTATGCTCAAAAGACCGTTATGCCCTCCTTTGATGGCGGCATCGGAAAGGACTTTGAGGAG<br>GGCAAACTGACTTTCGAGCAGGTCTATGAGTACGGCAAGAAGGTCGGAGAACCCCAACAG<br>ACCAGCGGCAACAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 5586MI204_002 | <i>Neocallimastigales</i> | Amino Acid       | 66         | MKEYFPEVGKIQFEGPESKNPMAFHYDYDAERVVAGKTMKEWMRFAMAWHHTLCAEGGDQF<br>GGGTKHFFPNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGGSVEEYETNLA<br>AITDYLKQKMDETGIKLLWSTANVFSNPRYMNGASTNPDFDVVARAIVQIKNAIDAGIKL<br>GAENYVFWGGREGYMSLLNTDQRREKEHMATMLRMARDYARAKGFKGTFLIEPKPCEPSK<br>HQYDVTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDANRG<br>DAQNGWDTDQFPIDNFELTQAFMQIVRNGFGTGGTNFDAKTRRNSTDLEDIFIAHISGM<br>DACARALLNAIEILLEKSPIPAMLKDRYASFDGGIGIKDFEEGKLTFEQVVEYEGKKVGEPKQ<br>TSGKQELYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI207_002 | <i>Neocallimastigales</i> | DNA              | 67         | ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAATTTGAAGGCCCGGAGTCCAAGAAC<br>CCGATGGCGTTTCACTACTATGACGCTGAGCGCGTCGTAGCCGGTAAACAATGAAAGAG<br>TGGATGCGTTTCGCTATGGCTTGGTGGCACACCCCTCTGTGCGGAAGCGGCACCAAGTTC<br>GGAGGAGGAACGAAGAAATCCCCTGGAACGAAGGGGCAACCGCTTTGGAGATCGCCAAG<br>CACAAAGCCGATCGGGATTTCGAGATCATGCAGAAACTCGGCATCCCTTATTTCTGTTTC<br>CATGACGTGGATCTCATCGCCGAGGGCGAATCGGTAGAAGAGTACGAAGCCCAACCTCGCT<br>GCCATCACCGATTACCTCAACACAGAAATGGACGAGACCGGCATCAAACGTGTGTGTTCC<br>ACGGCGAACGTGTTCAGCAACCCCCGTTATATGAACGGCGCCAGCACGAACCCCGATTTC<br>GATGTAGTGGCACGCGCTATCGTACAAATCAAGAACGCTATCGACGCCGGTATCAAACTC<br>GGAGCAGAGAACTATGTCTTCTGGGCGGTCCGAGGGCTATATGTCGCTCCTCAACACC<br>GACCAGCGCCGAGAGAAAGACATATGGCTACTATGCTCCGTATGGCGGTGACTACGCG<br>CGTTCAAAGGATTCAAGGGCACCTTCCTCATCGAACCCAAACCGTGTGAGCCGTCCAAA<br>CATCAGTACGATGTGGACACAGAGACCGTCATCGGTTTCCTTAAAGCGCATGGACTCGAC<br>AAGGATTTCAAAGTCAATATCGAGGTCAACCCACGCCACCCCTCGAGGCCACACGTTCGAA<br>CACGAACTGGCTTGGCTGTAGATGCCGGCATGCTCGGTTTCGATTGACGCCAACCCGCGGT<br>GACGCCAGAACGGATGGGACACCGACCAATTCCCATTGATAAATTTCGAACTCACTCAG<br>GCTTTCATGCAGATCGTCCGCAACGGCGGTTTCGGAACAGCGGTACGAACTTCGACGCC<br>AAGACACGCCGTAACTCCACCGACTTGGAGGACATCTTCATCGCCCATATCAGCGGCATG<br>GACGCTTGGCTCGTGGTTGCTCAATGCTGTCGAAATCCTCGAGAAAGAGCCCGATCCCG<br>GCTATGCTCAAAAGAGCGTTATGCTTCCCTTGACGGCGGCATCGGAAAGGACTTTGAGGAG<br>GGCAAACTGACTTTCGAGCAGGTCTATGAGTACGGCAAGAAGGTCGGAGAACCCCAACAG<br>ACCAGCGGCAACAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATGA |
| 5586MI207_002 | <i>Neocallimastigales</i> | Amino Acid       | 68         | MKEYFPEIGKIQFEGPESKNPMAFHYDYDAERVVAGKTMKEWMRFAMAWHHTLCAEGGDQF<br>GGGTKHFFPNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGESVEEYEA<br>AITDYLKQKMDETGIKLLWSTANVFSNPRYMNGASTNPDFDVVARAIVQIKNAIDAGIKL                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| TABLE 2       |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------------|---------------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|               |                           |                  |            | GAENYVFWGGREGYMSLLNTDQRREKEHEMATMLRMARDYARSKGFKGTFLIEPKPCEPSK<br>HQYDVDTETVIGFLKAHGLDKFKVNIENVNHATLAGHTEHEHELACAVDAGMLGSIDANRG<br>DAQNGWDTDQFPIDNFELTQAFMQIVRNGGFGTGGTNFDAKTRRNSTDLEDIFIAHISGM<br>DACARALLNAVEIILEKSPIPAMLKERYASFDDGGIGKDFEEGKLTFEQVVEYGKKVGEPKQ<br>TSGKQELYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI209_003 | <i>Neocallimastigales</i> | DNA              | 69         | ATGAAAGAGTATTTCCCTGAGATCGGTAAGATCCAATTTGAAGGCCCGGAGTCCAAGAAC<br>CCGATGGCGTTTCACTACTATGACGCAGAGCGGTAGTACCGGTTAAACAATGAAAGAA<br>TGGATGCCGTTTCGCCATGGCATGGTGGCACACCTCTGTGCAGAAGCGCGACCACTTC<br>GGAGGAGGAACGAAGCATTTCCCGTGGAATGAAGCGCTAACGCTTTGGAGATCGCCAAA<br>CACAAAGCCGATGCGGGATTGAGATCATGCAGAAACTCGGCATCCCCTATTTCTGTTTC<br>CATGACGTGGATCTCATCGCCGAGGCGGATTCCGTGGAGGAGTACGAAGCTAACCCCGCT<br>GCCATCACCGATTACCTCAAACAGAAAAATGGACGAGACCGGCATCAAACGTGCTGTGTC<br>ACGGCGAACGTTTCAGCAACCCCGTTACATGAACGGTGCAGACGAACCCGGATTTC<br>GATGTAGTGGCACGCGCTATCGTACAAATCAAAGAACGCTATCGACGCCGTATCAAACTC<br>GGAGCAGAGAACTATGTCTTCTGGGCGGTCCGAGGGCTATATGTCGTCCTCAACACC<br>GACCAGCGTCGCGAGAAAGACATATGGCTACTATGCTCCGTATGGCGGTGACTACGGG<br>CGTGCCAAAGGATTCAAGGGCACCTTCCTCATCGAACCCAAACCATGTGAGCCGTCCTCAA<br>CATCAGTACGATGTGGACACAGAGACTGTATCGGTTTCCTCAAAGCGCATGGACTCGAC<br>AAGGATTTCAAAGTCAACATCGAGGTCAACCAACGCCACCTCGCAGGTCACACGTTTCGAA<br>CACGAACTGGCTTGGCTGTAGATGCCGGCATGCTCGGTTTCGATTGACGCCAACCCGCGT<br>GACGCCCAGAACGGATGGGACACTGACCAGTTCCTATTGATAAATTCGAACTCACACAG<br>GCTTTCATGCAGATCGTCCGCAACGGCGGTTTCGGAACAGCGGTACGAACTTCGACGCC<br>AAGACACGCCGTAACTCCACCGACTTGGAGGACATCTTCATCGCCCATATCAGCGGCATG<br>GACGCTTGTGCCGTGCGTTGCTCAACGCCCATCGAAATCCTCGAAGAGAGCCCGATCCCC<br>GCTATGCTCAAAGAGCGTTACGCTTCCTTGACGGCGGCATCGGAAAAGGACTTTGAGGAT<br>GGTAAACTGACTTTCGAGCAGGTCTATGAGTACGGCAAGAAGTCCGAGAACCCCAACAG<br>ACCAGCGGCAACAGGAGTCTACGAAACCATCGTCGCCCTCTATGCCAAAGTAA |
| 5586MI209_003 | <i>Neocallimastigales</i> | Amino Acid       | 70         | MKEYFPEIGKIQEGPESKNPMAFHYDAERVVAGKTMKEWMRFAMAWHHTLCAEGGDQF<br>GGGTKHFFPWNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGDSVEEYEANPA<br>AITDYLKQKMDETGIKLLWSTANVFSNPRYMNGASTNPDFDVVARAIVQIKNAIDAGIKL<br>GAENYVFWGGREGYMSLLNTDQRREKEHEMATMLRMARDYARAKGFKGTFLIEPKPCEPSK<br>HQYDVDTETVIGFLKAHGLDKFKVNIENVNHATLAGHTEHEHELACAVDAGMLGSIDANRG<br>DAQNGWDTDQFPIDNFELTQAFMQIVRNGGFGTGGTNFDAKTRRNSTDLEDIFIAHISGM<br>DACVRALLNAIEIILEKSPIPAMLKERYASFDDGGIGKDFEDGKLTFEQVVEYGKKVGEPKQ<br>TSGKQELYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI214_002 | <i>Neocallimastigales</i> | DNA              | 71         | ATGAAAGAGTATTTCCCTGAGATCGGAAAGATCCAATTCGAAGGCCCGGAGTCCAAGAAT<br>CCTATGGCATTTCACTACTATGACGCAGAGCGGTGTAGTACCGGTTAAACAATGAAAGAG<br>TGGATGCCGTTTCGCTTTGGCATGGTGGCACACGCTCTGCGCAGAAAGCGGCACCACTTC<br>GGAGCGGCACGAAGCATTTCCCTTGGAATGAAGGTGCAACCGCTTTGGAGATCGCCAAG<br>CACAAAGCCGATGCAGGCTTCGAGATCATGCAGAAACTCGGCATCCCCTATTTCTGTTTC<br>CATGACGTGGATCTGATCGCCGAGGCGGTTCCGGTAGAAGAGTATGAAGCTAATTTAACG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| TABLE 2       |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------|--------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|               |                    |                  |            | GCTATACCCGATTACCTCAAACAGAAAAATGGACGAGACCGGCATCAAAC TGCTGTGGTCC<br>ACTGCGAACGTTGTCGGTAACGCACGTTATATGAACGGCGGAGCACGAACCCCGATTTC<br>GATGATGGCACGGCTATCGTGCAGATCAAGAACGCTATCGACGCCGGCATCAAAC TG<br>GGCGAGAGAACTACGTCCTCTGGGGCGGTGCGAGGGATATATGTCGTCCTGAACACC<br>GACCAGAAAGCGTGAGAAAGAGCATATGGCTACCATGCTCCGTATGGCGGTGACTACGCG<br>CGTCCAAAGGATTCAAAGGTACGTTCCCTCATCGAGCCCAAACCGGTGAGCCGTCCTCAA<br>CATCAGTACGACGTGGACACTGAGACCGTCATCGGTTTCCTCAAAGCCCATGGTCTCGGC<br>AAGGATTTCAAAGTGAACATCGAGGTGAATCACGCCACCCCTCGCAGGGCACACGTTCGAA<br>CACGAACTGGCTTGCGCCGTAGATGCCGGCATGCTCGGTTCGATCGACGCCAACCCGGGT<br>GACGCACAAAACGGATGGGACACCGACCCAGTTCCTATTTGATAAATTTCGAACTCACCCAG<br>GCATTCATGCAGATCGTCCGCAACGGCGGTTTCGGAACAGGCGGTACGAACTTCGACGCC<br>AAGACACGCCGTAATTCACCCGACTTGAGGACATCTTCATCGCCCATATCAGCGGCATG<br>GACGCTTGTCGCCGTGCTCAATGCTGTCGAAATCCTTGAAAAGAGCCCGATCCCCG<br>GCGATGCTCAAAGAGCGTTACGCCCTCCTTTGACAGCGGTATGGGTAAGGACTTTGAGGAG<br>GGCAAGCTGACCTTCGAGCAGGTCTATGAGTACGGCAAAACAGGTCGGCGAACCCAAACAG<br>ACCAGCGGCAAGCAGGAGCTCTACGAAACCATCGTCGCCCTCTATGCCAAATAG                                                                                  |
| 5586MI214_002 | Neocallimastigales | Amino Acid       | 72         | MKEYFPEIGKIQFEGPESKNPMAFYHYDAERVVAGKTMKEWMRFALAWHHTLCAEGGDQF<br>GGGTHFFPNEGANALEIAKHKADAGFEIMQKLGIPYFCFHDVDLIAEGGSVEEYEANLT<br>AITDYLKQKMDETGIKLLWSTANVFGNARYMNGASTNPDFDVVARAIVQIKNAIDAGIKL<br>GAENYVFWGREGYMSLLNTDQKREKEHMATMLRMARDYARSKGFKGTFLIEPKPCEPSK<br>HQYDVDTETVIGFLKAHGLGKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDANRG<br>DAQNGWDTDQFPIDNFELTQAFMQIVRNGGFGTGGTNFDAKTRRNSTDLEDIFIAHISGM<br>DACARALLNAVEIILEKSPIPAMLKERYASFDSGMGKDFEEGKLTFFEQVVEYVKQVGEPKQ<br>TSGKQELYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5751MI3_001   | Neocallimastigales | DNA              | 73         | ATGAAAAGAGTATTTCCACAAATCGGCAAGATCCCATTTGAGGGACCAGAGTCAAAGAAC<br>CCAAATGGCATTCCACTACTATGACGCAGAGCGCGTAGTTGCCGGTAAGACAAATGAAGGAA<br>TGGATGCCGTTTCGCTATGCCCTGGTGGCACACTCTCTGTGCTGAGGGTAGCGATCAGTTC<br>GGCCTGGTACAAAAGAAGTTCCTTGGAACGAGGGCGGAGACAGCCCTTGAGCGCGCTAAG<br>CACAAGGCAGATGCTGGCTTCGAGGTTATGCAGAAAGCTCGGCATCCCATATTTCTGCTTC<br>CACGATGTAGACCTTATCGACGAGGGTGCTAACGTGGCTGAGTATGAGGCAAACCTCGCT<br>GCTATCACTGACTACCTGAAGGAGAAAGATGGAGGAGACTGGCGTAAAGCTCCTCTGGTCT<br>ACAGCCAAACGTGTTCGGTAAACGCTCGCTATATGAACGGTGCTTCTACAAATCCTGACTTC<br>GACGTTGTGGCTCGTGCCATCGTACAGATTAAGAACGCTATCGACGCTGGTATCAAGCTT<br>GGTGCTGAGAACTACGTGTTCTGGGGCGCGGCACATGAGCCTTCTGAAACACT<br>GACCAGAAAGCGGAGAAGGAGCACATGGCAACTATGCTCGGCATGGCTCGCGACTATGCC<br>CGCGCTAAGGGATTACCCGGTACCTTCCTCATTGAGCCAAAGCCAAATGGAGCCAAACAAG<br>CATCAGTATGATGTTGACACAGAGACCGTTATCGGTTTCCTCAAGGCTCACGGTCTGGAC<br>AAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCTACTCTCGCCGGTCAACACCTTCGAG<br>CACGAGCTCGCTTGCCTGTTGACGCTGGTATGCTCGGTTCTATCGACGCTAACCCGGGT<br>GACGCTCAGAACCGGATGGGATACCGACCAAGTTCCTCAATCGACAACTTCGAGCTGACACAG<br>GCTTGGATGCAGATTGTTCCGCAATGGCGGTCTTTGGCACAGGTGGTACCAACTTCGACGCA |

| TABLE 2     |                           |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-------------|---------------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism         | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|             |                           |                  |            | AAGACCCGTCGTAACTCTACCGACCTCGAGGACATCTTCATCGCTCACATCTCCGGTATG<br>GACGCTTGTCACGCGCTCTCCTCAACGCAGTAGAGATACTCGAGAACTCTCCAATCCCA<br>ACAAATGCTGAAGGACCGCTATGCAAGCTTCGACTCAGGTATGGGTAAGGACTTCGAGGAC<br>GGCAAGCTCACACTTGAGCAGGTTTATGAGTATGGTAAGAAGGTCGACGAGCCAAAAGCAG<br>ACCTCTGGTAAGCAGGAACTCTATGAGACCATCGTTGCTCTCTATGCAAAATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5751MI3_001 | <i>Neocallimastigales</i> | Amino Acid       | 74         | MKEYFPQIGKIPFEGPESKNPMAFHYIDAERVVAGKTMKEWMRFAMAWHHTLCAEGSDQF<br>GPGTKKFPWNEGETALERAKHKADAGFEVMQKLGIPIYFCFHDVDLLIDEGANVAEYEANLA<br>AITDYLKEKMEETGVKLLWSTANVFGNARYMNGASTNPDEFDVARAIVQIKNAIDAGIKL<br>GAENYVFWGGREGYMSLLNTDQKREKEHMATMLGMARDYARAKGFTGTFLIEPKPMEPTK<br>HQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANRG<br>DAQNGWDTDQFPIDNFELTQAWMQIVRNGGLGTGGTNFDAKTRNSTDLEDIFIAHISGM<br>DACARALLNAVEILLENSPIPTMLKDRYASFDSGMGKDFEDGKLTLEQVVEYGKKVDEPKQ<br>TSGKQELYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5753MI3_002 | <i>Prevotella</i>         | DNA              | 75         | ATGGCTAAAGAATACTTCCCCTCCATCGGCAAAATCCCCTTTGAAGGAGGCGACAGCAAA<br>AATCCCCTCGCTTCCATTATTATGACGCGGACGCGTGGTTATGGCAAGCCCCATGAAG<br>GAATGGCTTAAATTGCCCATGGCTGGTGGCACACGCTGGGCCAGGCCCTCCGGAGACCCC<br>TTCGGGGGCCAGACCCGAGCTACGAATGGGACAAGGCGGAATGCCCCTACTGCCCGGCC<br>AAAGCCAAAGGCCGACGCCGGTTTGAATCATGC AAAAGCTGGGTATCGAATACTTCTGC<br>TTCCACGATGTGGACCTTATCGAGGATTGCGATGACATTGCCGAATACGAAGCCCGCATG<br>AAGGACATCACGGACTACCTGCTGGAAGAAGATGAAGGAGACCGGCATCAAGAACCTCTGG<br>GGCACCGCCAATGTCTTCGGCCACAAGCGCTACATGAACGGCGCGGCACCAATCCGCAG<br>TTCGATGTGGTGGCCCGTCCGCCGTCAGATCAAGAACGCCCTGGACGCCACCATCAAG<br>CTGGCGGCTCCAACTATGTGTTCTGGGGCGGCCGGAAGGCTATTACACCTCCTCAAC<br>ACCCAGATGCAGCGGGA AAAAGACCACTGGCCAAGTTGCTGACGGCGCGCCCGGACTAT<br>GCCCCGCCAAGGGCTTCAAGGGCACCTTCCATTAGCCCCAAACCCATGGAACCCACC<br>AAGCACCAGTACGACGTGGATACGGAGACGGTCA TCGGCTTCCTCCGTGCCAACGGCCTG<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAAGCCACCTGGCCGGCCACACCTTC<br>GAGCATGAGCTCACCGTGGCCCGCGAGAACGGTTTCTGGGCTCCATCGGTGCCAACCGC<br>GGCGACGCCCAGAACGGCTGGGACACGGACCAAGTTCCCTGTGGACCCCGTACGATCTTACC<br>CAGGCCATGATGCAAGGTGCTGCTGAACGGCGGCTTCGGCAACGGCGGCACCAACTTCGAC<br>GCCAAACTCCGCCGCTCCTCCACCGACCTGAGGACATCTTCATCGCCCATATTTCCGCC<br>ATGGATGCCATGGCCACGCTTTGCTTAACGAGCTGCCGTGCTGGAAGAGAGCCCCCTG<br>TGCCAGATGGTCAAGGAGCGTTATGCCAGCTTCGACGGCGGCTCGGCAACAGTTCGAG<br>GAAGGCAAGGCTACCCCTGGAAGACCTGTACGAATACGCCAAGGTC CAGGGTGAACCCGTT<br>GTCGCCCTCCGGCAAGCAGGAGCTTTACGAGACTCTCCTGAACCTGTATGCCGTCAAGTAA |
| 5753MI3_002 | <i>Prevotella</i>         | Amino Acid       | 76         | MAKEYFPSIGKIPFEGGDSKNPLAFHYIDAGRVVMGKPMKEWLKFAMAWHHTLQASGDP<br>FGGQTRS YEWDKGECPYCRAKAKADAGFEIMQKLGI EYFCFHDVDLIEDCDDIAEYEARM<br>KDITDYLLEKMKETGIKNLWGTANVFGHKRYMNGAGTNPQFDVVARA AVQIKNALD ATK<br>LGGSNVVFWGGREGYYTLLNTQMQRKDH LAKLLTAARDYARAKGFKGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI GANR<br>GDAQNGWDTDQFPVPYDLTQAMMQVLLNGGFGNGGTNFDAKLRRSSTDPE DIFIAHISA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|             |                   |                  |            | <u>MDAMAHALLNAAA</u> <u>VL</u> <u>EE</u> <u>SP</u> <u>LC</u> <u>Q</u> <u>M</u> <u>V</u> <u>K</u> <u>E</u> <u>R</u> <u>Y</u> <u>A</u> <u>S</u> <u>F</u> <u>D</u> <u>G</u> <u>G</u> <u>L</u> <u>G</u> <u>K</u> <u>Q</u> <u>F</u> <u>E</u> <u>E</u> <u>G</u> <u>K</u> <u>A</u> <u>T</u> <u>L</u> <u>E</u> <u>D</u> <u>L</u> <u>E</u> <u>Y</u> <u>A</u> <u>K</u> <u>V</u> <u>Q</u> <u>G</u> <u>E</u> <u>P</u> <u>V</u><br><u>V</u> <u>A</u> <u>S</u> <u>G</u> <u>K</u> <u>Q</u> <u>E</u> <u>L</u> <u>Y</u> <u>E</u> <u>T</u> <u>L</u> <u>I</u> <u>N</u> <u>L</u> <u>Y</u> <u>A</u> <u>V</u> <u>K</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 1754MI1_001 | <i>Prevotella</i> | DNA              | 77         | ATGGCAAAAGAGTATTTCCGTTTACCGGTAAGATTCCTTTCGAAGGAAAGACAGTAAG<br>AAATGTAATGGCTTTCCTACTACTACGAGCCTGAGAAAGTCTGATGGGAAAAGAAGATGAAG<br>GACTGGCTGAAGTTCGCTATGGCTTGGTGGCATACACTGGGTGGCGCTTCTGCTGACCCAG<br>TTTGGTGGTCAGACTCGTTCATACGAGTGGACAAGGCTGGTGACGCTGTTCAGCGCGCT<br>AAGGATAAGATGGACGCTGGCTTCGAGATCATGGACAAAGCTGGGCATCGAGTACTTCTGC<br>TTCCACGATGTTGACCTCGTTGAAGAGGGTGACCATCGAGGAGTATGAGGCTCGCATG<br>AAGGCCATCACCGACTACGCTCAGGAGAAGATGAAGCAGTTCCCCAACATCAAGCTGCTC<br>TGGGGTACCGCAAAACGTATTCCGTAACAAGCGCTATGCTAACGGTGCTTCTACCAACCCCC<br>GACTTCGACGTAGTGGCTCGCGCCATCGTTCAGATCAAGAACGCTATTGATGCTACCATC<br>AAGCTGGTGGTACCAACTATGTGTTCTGGGTGGTCGTGAGGGCTATATGAGTCTGCTG<br>AACACCGACCAAGAGCGTGAGAAGGAGCACATGGCTACTATGCTGACCATGGCTCGCGAC<br>TATGCTCGCGCCAAAGGATTCAAGGGTACATTCTCTATTAGCCCGAAGCCCATGGAGCCCC<br>AGCAAGCACCAAGTATGATGTGGATACAGAGACCGTTATCGGCTTCCTGAAGGCACACAAC<br>CTGGACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCTACACTCGTGGTCATACC<br>TTCGAGCACGAGCTGGCTTGGCTGTTGACGCTGGTATGCTTGGTTCTATCGACGCTAAC<br>CGTGGTATGCTCAGAACGGTTGGGATACCGACCAAGTTCGCCATCGACAACTACGAGCTG<br>ACACAGGCTATGCTCGAGATCATCCGCAATGGTGGTCTGGGCAATGGTGGTACCAAATTTC<br>GATGCTAAGATCCGTCGTAACAGCACCGACCTCGAGGATCTCTTCATCGCTCACATCAGT<br>GGTATGGATGCTATGGCACCGGCTCTGATGAACGCTGCTGACATCCTTGAGAACTCTGAG<br>CTGCCCCGAATGAAGAAGGCTCGCTACGCAAGCTTCGACCAGGGTGTGGTAAGGACTTC<br>GAAGATGGCAAGCTGACCCCTTGAGCAGGTTTACGAGTATGTTAAGAAGGTGGGTGAGCCC<br>AAGCAGACTTCTGTGAAGCAGGAGAAGTACGAGACCATCGTTGCTCTCTATGCAAAAATAA<br><br>MAKEYFPFTGKIPFEGKDSKNVMAFHYYEPEKVVMGKKMKDWLKFAMAWHHTLGGASADQ<br>FGGQTRSYEWDKAGDAVQRAKDKMDAGFEIMDKLGIYFCFHDVDLVEEGDTIEEYEARM<br>KAITDYAQEKMKQFPNIKLLWGTANVFGNKRANGASTNPDFVVARAIVQIKNAIDATI<br>KLGGTNYVFWGGREGYMSLLNTDQKREKEHMAI <del>MT</del> MARDYARAKGFKGTFLIEPKPMEP<br>SKHQYD <del>VD</del> TETVIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGWDTDQFPIDNYELTQAMLEIIRNGGLNGGTNFDAKIRRNSTDLEDLFLAHIS<br>GMDAMARALMNAADILENSELPAMKKARYASF <del>D</del> QGVGKDFEDGKLTLEQVVEYGKKVGE<br>P<br>KQTS <del>G</del> KQEKYETIVALYAK |
| 1754MI1_001 | <i>Prevotella</i> | Amino Acid       | 78         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 1754MI3_007 | <i>Prevotella</i> | DNA              | 79         | ATGGCAAAAAGAGTATTTCCGTTTACCGGTAAGATTCCTTTCGAAGGAAAGAGAGCAAG<br>AACGTAATGGCTTTCCTATTACTATGAGCCTGAAAAGGTGGTCATGGGCAAGAAAATGAAG<br>GATTGGCTGAAATTCGCCATGGCTTGGTGGCACACCCCTCGGTGGAGCCAGCGCCGACCAG<br>TTCGGTGGACAGACCCGCAGCTATGATGGGACAAGCCGAGGATGCCGTACAGCGTGCT<br>AAGGACAAAGATGGACGCCGGCTTCGAGATCATGGACAAACTGGGCATCGAGTATTCTGC<br>TTCCACGATGTCGACCTCGTCGACGAGGGTGCTACCGTTGAGGAGTATGAGGCTCGCATG<br>AAAGCCATCACCGACTATGCCCAGGTCAAGATGAAGGAATATCCCAACATCAAAC <del>T</del> GCTC<br>TGGGGCACCCGCAACGTTTCCGGCAACAAGCGTTATGCCAACCGCGCTTCCACCAACCCC<br>GACTTCGACGTGGTGGCACGGCTATCGTTTCAGATCAAGAATGCCATCGACGCTACCATC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------------|-------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|             |                   |                  |            | AAGCTCGGCGGTCAAGAACTACGTGTTCTGGGGCGGACGCGAGGGCTACATGAGCCTGCTC<br>AATACCGGATCAGAAACGTCAGAAGGAACACATGGCCACCATGCTCACCATGGCGCGGAC<br>TATGCTCGCAGCAAGGGATTCAAGGGCACCTTCCTCATCGAACCACCAACCCATGGAGCCT<br>TCCAAGCACCCAGTATGATGTCGACACCGAGACGGTCATGGGCTTCCTCCGGCCCCAAC<br>CTCGACAAAGGACTTCAAGGTGAACATCGAGGTC AACCCACGCCACGCTCGCCGGCCACACC<br>TTCGAGCACGAACTGGCTTGCGCCGTCGACGCCGGCATGCTCGGCAGCATCGACGCCAAC<br>CGCGGCGACGCACAGAACGGCTGGGATACCGACCAGTTCCTCCATCGACAACTACGAAC TG<br>ACACAGGCCATGCTGGAGATCATCCGCAATGGCGGCTTCGGCAATGGTGGTACCAACTTC<br>GAGCCAAAGATCCGTGCTAACAGCACCGACCTCGAAGATCTCTTCATCGCTCACATCAGC<br>GGTATGGATGCCATGGCTCGCGCGCTGCTCAACGCCGCCCATCTCTCGAGGAGAGCGAA<br>CTGCCCCGCATGAAGAAGGCCCGCTACGCTTCCTTCGACGAAGGTATCGGCAAGGACTTC<br>GAAGACGGCAAACTCACCTTCGAGCAGGTTTACGAGTACGGCAAGAGGTAGGCGAGCCC<br>AAGCAGACCTCCGGCAAGCAAGAGAACTACGAGACCATCGTGGCTCTCTACAGCAATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 1754MI3_007 | <i>Prevotella</i> | Amino Acid       | 80         | <u>MAKEYFPFTGKIPEFGKESKNVMAFHYYEPEKVVMGKKMKDWLKFAMAWHHTLGGASADQ</u><br><u>FGGQTRSYEWDKAEDAVQRAKDKMDAGFEIMDKLGI EYFCFHDVLDLVEGATVEEYEARM</u><br><u>KAITDYAQVKMKEYPNIKLLWGTANVFGNKRYANGASTNPDFDVVARAIVQIKNAIDATI</u><br><u>KLGGQNYVFWGGREGYMSLLNTDQKREKEHMTMLTMARDYARSKGFKGTFLEPKPMEP</u><br><u>SKHQYDVDTETVIGFLRAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN</u><br><u>RGDAQNGWDTDQFPIDNYELTQAMLEIIIRNGGLNGGTNFDAKIRRNSTDLEDLFLAHIS</u><br><u>GMDAMARALLNAAAAIL EESL PAMKKARYASFDEGIGKDFEDGKLT LLEQVVEYGKKVGE P</u><br><u>KQTS GQKEKYETIVALYSK</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 1754MI5_009 | <i>Prevotella</i> | DNA              | 81         | ATGAAAGAGTATTTCCCGCAAATTTGAAAGATTCCTTCGAGGGACCAGAGAGCAAGAGT<br>CCATTGGCGTTCATATTATTATGAGCCGGATCGCATGGTGCTCGGAAAGAGGATGGAGGAT<br>TGGCTGAAATTCGCCATGGCATGGTGGCACACCCCTTGGCAGGCCAGCGGCCACCA GTTC<br>GGCGACACAGACACGTGAGTACGAGTGGGATAAGGCTGGAGATCCGATACAAAGGGCAAAG<br>GATAAGATGGACCGCCGGATTCCGAGATCATGGAGAAATTGGGTATCAAGTACTTCTGCTTC<br>CATGATGTGGATCTCGTCGAGGAAGCTCCCACCATCGCCGAATATGAGGAGCGTATGAGG<br>ATCATCACCGACTATGCGCTCGAGAAGATGAAAGCCACTGGCATCAAAC TCTTTGGGGT<br>ACAGCCAAATGTTTTCGGACATAAGAGATATATGAATGGGCCCGCCACCAACCCGGAGTTC<br>GGTGTGTCGCCAGGGCTGCTGCCAGATCAAGAACCGGATCGACGCCACCATCAAGCTG<br>GGAGGAACAAACTATGTGTTCTGGGTGGCCGCGAGGGCTACATGAGCCTGCTCAACACC<br>CAGATGCAGAGGGAAGGACCATCTCGCCAAATATGCTCAAGGCTGCTCGTGACTATGCT<br>CGGCCAAGGGATTC AAGGGCACATTCCTCATCGAGCCGAAGCCGATGGAACCTACTAAG<br>CATCAGTACGATGTCGACACTGAGACCGTGATCGGCTTCCTCCGCGAAACGGTCTTGAC<br>AAGGATTTCAAGGTCAACATCGAGGTCAATCACGCCACTCTTGGGGTCACACTTTCGAG<br>CATGAGCTCGCCGTGGCTGTCGACAAATGGTCTCCTTGGCTCAATCGATGCGAACACAGGGGA<br>GATTATCAGAACGGTTGGGACACCGACCAGTTCCCTGTTGATCTCTTTGATTGACCCAG<br>GCCATGCTCCAGATCATCCGTAACGGAGGCCCTCGGTAATGGTGGATCCAAC TTCGACGCC<br>AAGCTTCGCCGTAAC TCCACTGATCCTGAGGATATATTCATTGCCCATATTTGCGGTATG<br>GACGCTATGGCCAGGGCTCTCCTTGGCCGCCCGGATCGTGGAGGAGTCTCCTATCCCCG<br>GCTATGGTCAAAGAGCGGTTACGCATCCTTCGACGAAGGTGAGGGCAAGAGATTCGAGGAT |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 1754MI5_009 | <i>Prevotella</i> | Amino Acid       | 82         | GGTAAGATGAGTCTGGAGGAAC TTGTGATTACGGGAAGACTCACGGAGAGCCCCGCCAG<br>AAGAGTGGCAAAACAGGAGCTCTACGAAACCC TTGTCAACATGTACATCAAAATAA<br>MKEYFPQIGKIPFEGPESKSP <sup>1</sup> AFHY <sup>2</sup> YEPDRMVLGKRMEDWLKFAMAWHHTLGQASGDQF<br>GGQ <sup>3</sup> REYEW <sup>4</sup> DKAGDPIQ <sup>5</sup> RAKDKMDAGFEIMEKLGIKYFCFHDVDLV <sup>6</sup> EEAPTIAEY <sup>7</sup> EERM <sup>8</sup><br>IITDYALEKMKATGIKLLWGTANVFGHKRYMNGAATNPEFGVVARAAVQIKNAIDATIKL<br>GGTNYVFWGREGYMSLINTQM <sup>9</sup> QREKDHLANMLKAARDYARAKGFKGTFLIEPKPMEPTK<br>HQYD <sup>10</sup> VD <sup>11</sup> TETVIGFLRANGLDK <sup>12</sup> FKVNI <sup>13</sup> EVNHATLAGH <sup>14</sup> TFEHELAV <sup>15</sup> DNGLLSIDANRG<br>DYQNGW <sup>16</sup> DTDQFPVDLFDLTQ <sup>17</sup> AMLQIIRNGGLNGGSNFD <sup>18</sup> AKLRNSTDPEDIFIAHICGM<br>DAMARALL <sup>19</sup> AAAAIVEESPI <sup>20</sup> PAMVKERYAS <sup>21</sup> FDEGEGKRFEDGKMSLEELVDYAKTHGEPAQ<br>KSGKQ <sup>22</sup> ELYETLVNMYIK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI1_003 | <i>Prevotella</i> | DNA              | 83         | ATGGCAAAAGAGTATTTTCCGTTTACCCGGTAAGATTCC <sup>1</sup> TTTCGAGGGAAAGCACAGTAAG<br>AATGTAATGGCGTTCCACTACTACGAGCCCCGAGCGGTGGTAATGGGCAAGAAGATGAAG<br>GAGTGGCTGAAGTTTGCCATGGCCTGGTGGCACACGCTGGTGGAGCCAGTGCCGACCAG<br>TTTGGCGGACAGACCCGCAGCTACGAGTGGGACAAGGCTGAAGACGCCGTGCAGCGGTGCC<br>AAGGACAAAGATGGATGCCGGCTTCGAGATCATGGACAAAGCTGGGCATCGAGTATTTCTGC<br>TTCCATGATGTCGATCTCGTTGACGAGGGTGCCACTGTCGAGGAGTATGAGGCTCGCATG<br>CAGGCCATCACCGACTATGCGCAGGAGAAGATGAAGCAGTATCCTGCCATCAAGCTGCTG<br>TGGGTA <sup>1</sup> CGGCCAA <sup>2</sup> TGTCTTTGGCAACAAGCGTTATGCCAACGGTGCCTCTACCAAATCCC<br>GACTTCGATGTGGTGGCCCCGCCCATCGTGCAGATTAAAGAATTGCCATTGATGCCACCATC<br>AAGCTGGCGGCAGCAACTATGTGTTCTGGGCGGTCCGAGGGCTACATGTCGTGCTC<br>AACACCGACCAAGAGCGTGAGAAGGAACACATGGCCCCGATGCTGACCATGGCCCCGCAC<br>TATGCCCGCTCGAAGGGCTTCAAGGGCAACTTCCTGATTAGCCCCAAGCCCATGGAGCCG<br>TCGAAGCATCAGTACGACGTGGACACCGAGACGGTTATCGGATTCTCCGCGCACATGGC<br>CTTGACAAAGGACTTCAAGGTGAACATCGAGGTGAACCATGCCACGCTGGCCGGTCATACC<br>TTCGAGCACGAACTGGCTTGGCCCGTAGATGCCGGCATGCTGGGCAGCATTGATGCCAAC<br>CGCGCGACGCACAGAACGGATGGGACACCGACCAAGTTC <sup>1</sup> CCCATCGACAACTATGAGTTG<br>ACACAGGCCATGATGGAGATTATCCGCAATGGCGGTCTGGGTCTTGGCGGTACCAATTTC<br>GATGCCAAGATTCCCGGTAACTCCACCGACCTGGAAGACCTCTTCATCGCCACATCAGT<br>GGCATGGACGCCATGGCTCGTGGCTCCTTAATGCTGCCGACATCTGGAGAACAGCGAA<br>CTGCCCGCCATGAAGAAAGCGCGCTACGCCCTCGTTCGACAGTGGCATGGGCAAGGACTTC<br>GAGGACGGCAAACTGACCCCTTGAGCAGGTTTACGAATACGGCAAAAAGTCGGCGAACCT<br>AAGCAGACCTCCGGCAAGCAGGAGAAGTACGAGACCATCGTGGCTCTCTATGCCAAGTAA |
| 5586MI1_003 | <i>Prevotella</i> | Amino Acid       | 84         | MAKEYFPFTGKIPFEGKDSKNVMAFHY <sup>1</sup> YEPERVVMGKKMKEWLKFAMAWHHTLGASADQ<br>FGGQ <sup>2</sup> TRSYEWDKAEDAVQRAKDKMDAGFEIMDKLGIEYFCFHDVDLV <sup>3</sup> DEGATVEEYEARM<br>QAITDYAQEKMKQYPAIKLLWGTANVFGNKR <sup>4</sup> YANGASTNP <sup>5</sup> DFVVARAIVQIKNAIDATI<br>KLGGSNYVFWGREGYMSLLNTDQ <sup>6</sup> KREKEH <sup>7</sup> MARMLTMARDYARSKGFKGNFLIEPKPMEP<br>SKHQYD <sup>8</sup> VD <sup>9</sup> TETVIGFLRAHGLDK <sup>10</sup> FKVNI <sup>11</sup> EVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGW <sup>12</sup> DTDQFPIDNYELTQAMMEIIRNGGLGLGGTGN <sup>13</sup> FDAKIRRNSTDLEDLFLAHIS<br>GMDAMARALL <sup>14</sup> NAADILENSELPAMKKARYAS <sup>15</sup> FDSGMGKDFEDGKLTLEQVYEYGKKVGE <sup>16</sup> P<br>KQTS <sup>17</sup> GKQEKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 5586MI2_006 | <i>Prevotella</i> | DNA              | 85         | ATGGCAAAAGAGTATTTCCGTTTACAGGTAAATTCCTTTCGAAGGAAAGACAGTAAG<br>AACGTAATGGCTTTCCTACTACTACGAGCCCCGAAAAGGTCGTGATGGGAAAGAAAATGAAA<br>GACTGGCTGAAGTTCCCCATGGCCTGGTGGCACACACTGGGTGGCGCCGACCGACAG<br>TTTGGCGGCCAGACACGAGCTATGAGTGGGACAAAGGTGCCGATGCCGTGCAGCGGCA<br>AAGGACAAAGATGGACGCCCGGCTTCGAAATCATGGACAAAGCTGGGCATCGAGTATTTCTGC<br>TTCACGACGTGGACCTCGTTGAGGAGGGAGCCACCATCGAGGAGTATGAGGCCCGCATG<br>AAGGCTATCACCGACTATGCCCAGGAGAAGATGAAACAGTATCCCAGCATCAAGCTGCTC<br>TGGGCAACGCCCAATGTGTTTGGCAACAAGCGTACGCCAAGCGCGCCAGCACCAACCCC<br>GACTTCGACGTCGTGGCCCGTGCCATCGTGCAGATCAAGAAGCCCATCGATGCCACCATC<br>AAGCTGGCGGCACCAACTACGTGTTCTGGGCGGACGCGAGGGCTACATGAGCCTTGCTC<br>AACACCGACCAAGAGCGCGAGAAGGAGCACATGGCCACCATGCTCACCATGGCCCCGGAC<br>TAGCCCCGGCAAGGATTCAAGGGCACCTTCCTCATCGAGCCCCAAGCCCATGGAGCCG<br>TCGAAGCACCAAGTACGACGTGGACACCGAGACCGTCATCGGTTTCCTGAAGGCCACGGT<br>CTGGACAAAGGACTTCAAGGTGAACATCGAGGTGAACCAAGCCAGCTGGCCGGCCACACC<br>TTCGAGCATGAGCTGGCCTGCGCCGTCGACCGCGGTATGCTGGGCAGCATCGATGCCAAC<br>CGCGCGACGCCCAAGACGGCTGGACACCCGACCATGTCCTCCCATCGACAACTTCGAGCTC<br>ACCCAGGCCATGATGGAATTATCCGCAACGGCGGCCCTCGGCAACGGCGGCACCAACTTC<br>GACGCTAAGATCCGCCGCAACTCCACCGACCTCGAGGACCTCTTCATCGCCACATCAGC<br>GGCATGGACGCCATGGCCCGGCACTGATGAACGCTGCCGACATTCGGAACACAGCGAG<br>CTGCCCGCCATGAAGAAGGCACGCTACGCCAGCTTCGACGCCGGCATCGGCAAGGACTTT<br>GAGGATGGCAAGCTCTCGCTGGAGCAGGTCACGAGTATGGCAAGAGGTGGAAGAGCCC<br>AAGCAGACCAAGCGCAAGCAGGAGAAAGTACGAGACCATCGTCGCCCTCTATGCCAAGTAA |
| 5586MI2_006 | <i>Prevotella</i> | Amino Acid       | 86         | MAKEYFPFTGKIPEFGKDSKNVMAFHYYEPEKVMGKKMKDWLKFAMAWHHTLGGASADQ<br>FGQTRSVEWDKAADAVQRAKDKMDAGFEIMDKLGEYFCFHDVDLVEEGATIEEYEARM<br>KAITDYAQEKMKQPSIKLLWGTANVFGNKRKYANGASTNPDFDVVARAIVQIKNAIDATI<br>KLGNTNYVFWGGREGYMSLNTDQKREKEHMA <del>TML</del> TMARDYARAKGFKGTFLIEPKPMEP<br>SKHQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSDAN<br>RGDAQNGWDTDQFPIDNFELTQAMMEIIRNGGLNGGTTNFDAKIRRNSTDLEDLFLAHIS<br>GMDAMARALMNAADIMENSELPAMKKARYASFDAGIGKDFEDGKLSLEQVVEYGKKVEEP<br>KQTS <del>GKQ</del> EKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 5586MI8_003 | <i>Prevotella</i> | DNA              | 87         | ATGGCAAAAGAGTATTTCCGCTTTACAGGCAAGATTCCCTTTCGAGGGAAAAGACAGTAAG<br>AACGTGATGGCTTTCCTACTACTACGAGCCGGAGCGTGTGGTGATGGGCAAGAAGATGAAG<br>GAGTGGCTGAAGTTCGCCATGGCCTGGTGGCACACACTGGGTGGCGCATCGGCCGACCCAG<br>TTCGGAGGCCAGACACGCACTACGAGTGGGACAAAGGCCCGCCGACGCCGTGCAGCGCGCC<br>AAGGACAAAGATGGACGCCCGGCTTCGAGATTATGGACAAAGCTGGGCATCGAGTACTTCTGC<br>TTCACGATGTAGACCTCGTTGAGGAGGGTGAGACCATAGCCGAGTACGAGCGCCGCATG<br>AAGGAAATCACCGACTACGCACAGGAGAAGATGAAGCAGTTCCCCAACATCAAGCTGCTC<br>TGGGGCACAGCCAAACGTGTTCCGCAACAAGCGCTACGCCAACGGCGCATCGACCAACCCC<br>GACTTCGACGTTGTGGCACGCGCCCATCGTGCAGATCAAGAACGCCCATCGACGCCACCATC<br>AAGCTCGCGCGCTCCAAC <del>TAT</del> GTGTTCTGGGGCGGACGCGAGGGCTATATGAGCCTGCTC<br>AACACCGACCAAGCGCGGAGAGGAGCACATGGCCACCATGCTCACCATGGCCCCGGAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

| TABLE 2      |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|              |                   |                  |            | TATGACGGCCAAAGGATTCAAGGGCACATTCTCTCATCGAGCCGAAGCCCATGGAGCCC<br>TCGAAGCACCATGACGCTAGACACAGAGACCGTTCATCGGCTTCCTCCGTGCACACGGG<br>CTGGACAAAGGACTTCAAGGTGAACATCGAGGTAAACCAAGCCACACTGGCCGGCCACACC<br>TTCGAGCACGAGCTGGCTTGCGCCGTCGACGCTGGCATGCTGGCAGCATCGACGCCAAC<br>CGTGGCGACGCACAGAACGGATGGACACCCGACCAAGTCCCCATCGACAACTTCGAGCTC<br>ACACAGGCCATGATGGAATCATCCGCAATGCGGACTGGGCAATGGCGGCACCAACTTC<br>GAGCCAAAGATCCGTCGTAACAGCACCGACCTCGAAGACCTCTTCATCGCCACATCAGC<br>GGCATGGACGCCATGGCACCGGCACTGCTCAACGCTGCCGACATCCTGGAGCACAGCGAG<br>CTGCCAAAGATGAAGAAGGAGCGCTACGCCAGTTCGACGCAGGCATCGGCAAGGACTTC<br>GAAGACGGCAAGCTCACACTCGAGCAGGTCTACGAGTACGGCAAAGGTCGAAGAGCCCC<br>CGTCAGACCAAGCGCAAGCAGGAGAAAGTACGAGACCATCGTCGCCCTCTATGCCAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5586MI18_003 | <i>Prevotella</i> | Amino Acid       | 88         | MAKEYFAFTGKIPFEGKDSKNVMAFHYYEPERVVMGKKMKEWLKFAMAWHHTLGGASADQ<br>FGGQTRS YEWDKAADAVQRAKDKMDAGFEIMDKLGI EYFCFHDVDLVEEGETIAEYERRM<br>KEITDYAQEKMKQFPNIKLLWGTANVFGNKRYANGASTNPDEFDVVARIVQIKNAIDATI<br>KLGGSNYVFWGGREGYMSLLNTDQKREKEHMA TMLTMARDYARAKGFKGTFLIEPKPMEP<br>SKHQYD VDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGWDTDQFPIDNFELTQAMMEIIRNGGLNGGTNFDAKIRNSTDLEDLFLIAHIS<br>GMDAMARALLNAADILEHSELPKMKERYASFDAGIGKDFEDGKLTLEQVYEYGKKKVEEP<br>RQTS GKQEKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5586MI14_003 | <i>Prevotella</i> | DNA              | 89         | ATGGCAAAAGAGTATTTCCGTTTACTGGTAAAGATTCCTTTCGAGGGGAAAAGGATAGTAAG<br>AATGTAATGGCTTCCACTATTACGAGCCCCGAGAAAGTCGTGATGGGAAAGAAGATGAAG<br>GACTGGCTGAAGTTCGCAATGGCTTGGTGGCATACTACACTGGGTGGTGCACTGCAGACCCAG<br>TTCGGTGGAGAGACCCGCAGCTACGAGTGGAGCAAGGCTGCTGATCCCGTTCAGCGCGCC<br>AAGGACAAAGATGGACGCCGGCTTTGAGATTATGGATAAAGCTGGGCATCGAGTACTTCTGT<br>TTCACGATATAGACCTCGTTCAGGAGGCAGATACCATGCAGAAATATGAGGAGCGCATG<br>AAGGCAATTACCGACTATGCTCTGGAGAAGATGAAGCAGTTCCTCCCAACATCAAGTTGCTC<br>TGGGTACCGCTAACGTATTAGCAACAAGCGCTATATGAACGGTGCTTCTACCAATCCC<br>GACTTCGACGTGGTGGCCCGTGCCATCGTTCAGATCAAGAACGCTATTGATGCAACCATC<br>AAACTCGGTGGTACCAACTATGTATTCTGGGTGGTCTGTGAGGGTTACATGAGCCTATTG<br>AATACCGACCAAGAGCGTGAAAAGGAGCACATGGCAATGATGCTCGGTATGGCTCGCGAC<br>TATGCCCCGAGCAAGGATTCAAGGGTACGTTCTCTCATCGAGCCGAAGCCGATGGAGCCC<br>TCTAAGCATCAGTATGATGTCGATACGGAGACTGTGATTGGTTTCCTGAAGGCACACGGT<br>CTGGACAAGGACTTCAAGGTGAACATCGAGGTGAACCAAGCTACACTGGCTGGTCATACC<br>TTCGAGCATGAGCTGGCTTGCGCTGTTGACCGCAGGTATGCTGGGCTCTATCGACGCTAAC<br>CGCGGTGATGCCCCAGAACGGCTGGGATACCGACCAAGTCCCCATCGACAACTACGAGCTG<br>ACACAGGCTATGATGGAATCATCCGCAACGGTGGTCTGGGCAATGGTGGTACCAACTTC<br>GACGCTAAGATCCGCCGTAACTCTACCGACCTCGAGGATCTGTTTCATCGCTCATATCAGT<br>GGTATGGATGCTATGGCCCGTGCTTGTGTAATGCTGCCGACATCTTGGAGAACTCTGAA<br>CTGCCCGCTATGAAGAAGGCCCGCTACGCCAGCTTCGACAACTGTTAAGAAAGTTGAAGAGCCG<br>GAGGATGGCAAGCTGACCTTCGAGCAGGTTTACGAATATGTTAAGAAAGTTGAAGAGCCG<br>AAGCAGACCTCTGGCAAGCAGGAGAAAATACGAGACCATCGTTGCTCTGTATGCTAAATAA |

| TABLE 2      |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.    | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5586MI14_003 | <i>Prevotella</i> | Amino Acid       | 90         | MAKEYFPFTGKIPFEGKDSKNVMAFHYYEPEKVVMGKKMKDWLKFAMAWWHHTLGGASADQ<br>FGGETRSYEWSKAADPVQRAKDKMDAGFEIMDKLGI EYFCFHDIDL VQ EADTIAEYEEERM<br>KAITDYALEKMQQFPNIKLLWGTANVFSNKRYMNGASTNPDFVVARAIVQIKNAIDATI<br>KLGGTNYVFWGGREGYMSLLNTDQKREKEHMMMLGMARDYARSKGFKGTFLEPKPMEP<br>SKHQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGWDTDQFPIDNYELTQAMMEIIRNGGLNGGTFNFDAKIRRNSTDLEDLFLIAHIS<br>GMDAMARALLNAADILENSELPAMKKARYASFDNGIGKDFEDGKLTTFEQVVEYGKKVVEEP<br>KQTSKGQEKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5586MI26_003 | <i>Prevotella</i> | DNA              | 91         | ATGGCAAAAGAGTATTTTCCGTTTACCGGTAAATAATTCCTTTCGAGGGAAAGGACAGTAAG<br>AATGTAATGGCTTTC CACTACTACGAGCCTGAGCGCGTAGTGATGGGAAAAGAAGATGAAG<br>GATTGGTTGCGATTTGCAATGGCTTGGTGGCACACACTGGGTGGCGCTTCTGCCGACCAG<br>TTTGGTGGTCAGACCCGCAGTTACGAATGGGACAAAGGTGCTGATGCTGTTCAGCGTGCT<br>AAGGACAAAGATGGATGCCGGCTTCGAGATTATGGATAAGCTGGGAATCGAGTTCTTCTGC<br>TGGCACGATATCGACCTCGTTGAAGAGGGTGAGACCATTTGAAGAGTATGAGCGCGCGCATG<br>AAGGCTATCACCGACTATGCTCTTGAGAAGATGCAGCAGTATCCCAACATCAAGAAACCTC<br>TGGGGAACAGCCAAATGTGTTTGGCAACAAGCGTTATGCCAACGGTGCCAGCACAAACCCA<br>GACTTTGACGTCGTTGCTCGTGCTATCGTACAGATTAAAGAAIGCTATCGACGCTACTATC<br>AAGTTGGGTGGTCAGAAATTATGTGTTCTGGGGTGGCGTGAGGGCTACATGAGCCTGCTC<br>AATACTGACCAGAAAGCGTGAGAAGGAGCACATGGCTACAAATGCTGACCATGGCACGCGAC<br>TATGCCCGCAGCAAGGGATTCAAGGGTAACTTCCTCATTGAGCCCAAGCCCATGGAGCCG<br>TCAAAGCACCAAGTATGATGTTGACACCGAGACCGTATGCGGTTTCCTGCGTGCCCAAC<br>CTTGACAAAGGATTTCAAGGTAAATATCGAGGTTAACCATGCTACTCTGGTGGTCATACT<br>TTCGAGCACGAACTGGCATGGCTGTTGACGCTGGTATGCTTGGTTCATCGATGCTAAC<br>CGTGGTGATGCCCCAGAATGGCTGGGATACCGACCAAGTTC C C C C A T C A A C A A C T A T G A A C T C<br>ACTCAGGCTATGCTTGAGATCATCCGTAATGGTGGTCTGGGTCTTGGCGGCACAAACTTC<br>GATGCCAAGATTCTGTCGTAAC T C A A C A G A T C T T G A G G A T C T C T T C A T C G C T C A C A T C A G T<br>GGTATGGATGCCATGGCCCCGTGCTCTGCTGAATGCTGCTGCTATTTCTGGAGGAGAGCGAG<br>CTGCCTAAGATGAAGAAGGAGCGTTATGCTTCTTTTCGATCCGGTATCGGTAAGGACTTC<br>GAGGATGGCAAGCTTACCCTTTGAGCAGGCTTACGAGTATGGTAAAGGTTGAGGAGCCCC<br>AAGCAGACTTCAGGCAAGCAGGAGAAAGTACGAGACCATCGTTGCTCTGTATGCAAAATAA |
| 5586MI26_003 | <i>Prevotella</i> | Amino Acid       | 92         | MAKEYFPFTGKIPFEGKDSKNVMAFHYYEPERVVMGKKMKDWLRFAMAWWHHTLGGASADQ<br>FGGQTRS YEWDKAADAVQRAKDKMDAGFEIMDKLGI E F F C W H D I D L V E E G E T I E E Y E R R M<br>KAITDYALEKMQQYPNIKNLWGTANVFGNKRYANGASTNPDFDVVARAIVQIKNAIDATI<br>KLGQNYYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARSKGFKGNFLIEPKPMEP<br>SKHQYDVDTETVCGFLRAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGWDTDQFPINNYELTQAMLEIIRNGGLGLGTFNFDAKIRRNSTDLEDLFLIAHIS<br>GMDAMARALLNAAAIL E E S E L P K M K K E R Y A S F D A G I G K D F E D G K L T L E Q A Y E Y G K K V E E P<br>KQTSKGQEKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI86_001 | <i>Prevotella</i> | DNA              | 93         | ATGAAAACAGTATTTTCCCAAGATTGAAAAGATACCCCTTCGAGGGGTGTAGAGAGCAAGAAT<br>GTGATGGCTTTC CACTATTATGAGCCAGAAAAGAGTAGTCA T G G G C A A G C C T A T G A A A G A A<br>TGGCTGGCGCTTCGCTATGGCGTGGTGGCACACGCTGGGGCAGCGGAGCGGGACCCCTTC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|               |                   |                  |            | GGCGGACAGACCCGCAGCTACGAGTGGGACCCGTGCGGCCGACGGCTACAGCGCGCCAAG<br>GACAAGATGGATCGGGCTTCGAGCTGATGGAGAAAGCTTGGCATTGAGTACTTCTGCTTC<br>CAGACGTGGACCTCGTAGAAGAGGGCGCCACGGTGGAGGAATACGAGCGGGATGGCT<br>GCCATCACCGACTACGCGGTAGAGAAAGATGCGCGAGCATCCGAGATACACTGCTGTGG<br>GGCACGGCCAATGTCTTCGGCCACAAAGCGCTACATGAACGGAGCCGCCACCAACCCCGAC<br>TTCGACGTGGTGGCGGTGCGGTGGTGAGATAAAGAACAGCATCGACGCCACGATCAAG<br>CTGGCGCGGAGAACTATGTGTCTGGGGCGGACGCGAGGGATATATGAGCCTGCTCAAC<br>ACCGACAGCGCGGAGAAAGGAGCACCTGGCCATGATGCTTCCGAAGGCCCGGACTAT<br>GGCGCGCCACGGCTTCAAGGGCACCTTCTGTATAGAGCCCAAGCCGATGGAGCCCATG<br>AAGCACCAGTACGACGTGGACACCGAGACGGTGATAGGTTTCTGCTGCCACGGACTG<br>GACAAGGACTTCAAGGTGAACATCGAGTGAACCCACGCCACGTTGGCGGCCACACGTTTC<br>GAGCACGAGCTGCCCTGTCCCTCGATGCCGGCATGCTGGGCAGCATCGACGCCAACCGT<br>GGGACCGGCAGAACGGATGGGATACGGACCAAGTCCCCATAGACTGCTACGAGCTCACG<br>CAGCGTGGATGGAGATCATTCGTGGCGCGGCTTCAACACCGCGGCACCAACTTCGAC<br>GCTAAGCTGCGCCGCAACTCGACCGACCCCGAGGATATCTTCATAGCTCACATCAGCGGC<br>ATGGATGCTATGGCCCGCCCTGCTCTGCGCCCGGACATCTTGGAGCACAGCGAGCTG<br>CCGGAGATGAAGCGGAAGCGCTATGCCCTCGTTCGACAGCGGCATGGGCAAGGAGTTCGAA<br>GAGGCAATCTCAGCTTCGAGCAAACTCTATGCCCTACGGCAAGCAGGGCGGCAACCGGCC<br>ACGACCAGCGGCAAGCAGGAGAAAATACGAAGCCATTGTTTCACTTTATACCCGATGA |
| 5586MI186_001 | <i>Prevotella</i> | Amino Acid       | 94         | MKQYFPQIGKIPFEGVESKNVMAFHYYPEPERVVMGKPMKEWLRFAMAWWHTLGQASGDPF<br>GGQTRSYEWDRAADALQRAKDKMDAGFELMEKLGIEYFCFHDVDLVEEGATVEEYERRMA<br>AITDYAVEKMHREHPEIHCLWGNTANVFGHKRYMNGAATNPDFDVVARAVVQIKNSIDATIK<br>LGENYVFWGGREGYMSLLNTDQRREKEHLAMMLAKARDYGRAHFKGTFLIEPKPMEPM<br>KHQYDVDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSI DANR<br>GDAQNGWDTDQFPIDCYELTQAWMEIIRGGGFTTGGTNFDAKLRRNSTDPEDIFIAHISG<br>MDAMARALLCAADILEHSELPEMKRKRYASFDSGMGKEFEENLSFEQIYAYGKQAGEPA<br>TTS GKQEKYEAI VSLYTR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI108_002 | <i>Prevotella</i> | DNA              | 95         | ATGGCAAAAGAGTATTTTCCGTTTATCGGTAAGGTTCCTTTCGAAGGAACAGAGAGCAAG<br>AACGTGATGGCATTCCTACTACTATGAGCCCCGAAAAGGTGGTCATGGGTAAGAAAATGAAG<br>GACTGGCTGAAGTTCGCTATGGCTTGGTGGCACACACTGGGTGGTCCAGCGCCGACCAG<br>TTTGGTGGTCAGACTCGCAGCTACGAGTGGGACAAGGTGCTGATGCCGTTCAGCGCGCC<br>AAGGACAAAGATGGA TGCTGGCTTCGAGATCATGGATAAGCTCGGCATTGAGTACTTCTGC<br>TTCCATGACGTAGACCTCGTTGAGGAGGGTGAAACCGTCGCTGAGTATGAGGCTCGCATG<br>AAGGTCATCACCGACTATGCCCTGGAGAAGATGCAGCAGTCCCCCAACATCAAAC TGCTC<br>TGGGTTACTGCTAACGTGTTCCGCCACAAGCGCTATGCCAACGGTGCCAGCACCAATCCC<br>GACTTCGACGTCGTGGCCCCGTGCTATCGTTCAGATCAAGAA TGCCATCGATGCTACCAT<br>AAGCTCGGCGGTACGAACTATGTGTTCTGGGGTGGTCGTGAGGGCTACATGAGCCTTCTC<br>AACACCGACCAGAAAGCGGAGAGGAGCACATGGCAACGATGCTGACCATGGCTCGCGAC<br>TATGCCCGCGCCAAGGGATTCAAGGGCACGTTCCCTCATCGAGCCGAAGCCCATGGAGCCC<br>TCGAAGCATCAGTACGACGTCGACACCGGAGACCGTCATCGGCTTCCTCCGTGCCACGGT<br>CTGGATAAGGACTTCAAGGTGAACATCGAGGTTGAACCCACGCCACGCTGGCCGCTCATACC                                                                                                                                                                                                                                                                                                   |

| TABLE 2        |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|----------------|-------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.      | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|                |                   |                  |            | TTGAGACGAAC TGGCTTGGCCCGTTGATGCCGGCATGCTCGGCTCTATCGATGCCAAC<br>CGCGCGACGCTCAGAACGGCTGGACACCGACCACTTCCCCATCGACAACTACGAGCTC<br>ACTCAGGCCATGATGGAAATCATCCGTAATGGCGGCTGGGCAACGGCGGCACGAACTTC<br>GATGCCAAGATCCGTGTAACAGCACCGACCTCGAGGACCTCTTCATCGCTCACATCAGC<br>GGCATGGATGCCATGGCACGGCTCTGATGAACGCTGCTGCCATCCTCGAAGAGAGCGAG<br>CTGCCCCCATGAAGAAGGCCCGCTATGCTTCGTCGACGAGGGTATCGGCAAGGACTTC<br>GAGGACGGCAAGTTGTCACTTGAGCAGGCTACGAAATATGTTAAGAAAGTTGAGGAGCCC<br>AAGCAGACCTCGGCAAGCAGGAGAACTACGAGACCATCGTGGCCCTCTATGCCAAGTAA<br>MAKEYFPFIGKVPFEGTESKNVMAFHYYEPEKVVMGKKMKDWLKFAMAWWHTLGGASADQ<br>FGGQTRSYEWDKAADAVQRAKDKMDAGFEIMDKLGI EYFCFHDVDLVEEGETVAEY EARM<br>KVITDYALEKMQQFPNIKLLWGTANVFGHKRYANGASTNPDFDVVARAI VQIKNAIDATI<br>KLGGTNYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARAKGFKGTFLIEPKPMEP<br>SKHQYD VDTETVIGFLRAHGLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDAN<br>RGDAQNGWDTDQFPIDNYELTQAMMEIIRNGGLNGGTNFDAKIRRNSTDLEDLFLAHIS<br>GMDAMARALMNAAAIILEESEL PAMKKARYASFDEGIGKDFEDGKLSLEQVY EYGKKVEEP<br>KQTS GKQEKYETIVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5586MI1182_002 | <i>Prevotella</i> | Amino Acid       | 96         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 5586MI1182_004 | <i>Prevotella</i> | DNA              | 97         | ATGGCAAAAGAGTATTTTCCGTTTGTGGTAAGATTCTCTTTCGAGGGAAAGGATAGTAAG<br>AATGTAATGGCTTTCCTACTATTACGAAC CAGAGAAGGTCTGATGGGAAA GAAGATGAAG<br>GACTGGCTGAAGTTCGCCATGGCATGGTGGCACACACTGGGACAGGCCAGTGCCGACCCG<br>TTTGAGAGTCAGACCCGCAGCTACGAGTGGGACAAGGCTGACGATGCTGTGCAGCGCGCA<br>AAGGACAAAGATGGATCCCGGATTTGAGATCATGGACAAAGCTGGGCATCGAGTACTTCTGC<br>TTCCACGATGTAGACCTCGTTGAGGAGGGAGCAACTGTTGAGGAGTACGAGGCTCGCATG<br>AAGCCATCACCGACTATGCATTGGAAGAAGATGAAAGATATCCCAACATCAAGAACTC<br>TGGGTACAGCCCAATGTATTCAGCAACAAGCGCTATATGAACGGTGCCAGCACCAACCCC<br>GACTTCGACGTTGTTGCACGTGCCATCGTACAGATAAAGAACGCCATTGACGCTACCATC<br>AAGCTCGGCGGT CAGAACTACGTGTTCTGGGCGGACGTGAGGGATACATGAGCCTGCTC<br>AACACCGACCAAGAGCGGAGAGGACACATGGCAACCATGCTGACCATGGCTCGCGAC<br>TAGCTCGCAAGAACGGTTTCAAGGGCACATTCTCTCATCGAGCCTAAGCCCATGGAACCC<br>TCAAAGCACCAAGTACGACGTAGACACAGAGACCGTATGCGGTTTCTCTCCGCGCCCATGGT<br>CTTGACAAAGGATTTCAAGGTGAACATTGAGGTGAACCAACGCTACCTCGCCGGCCACACC<br>TTTGAGCATGAAC TGGCTTGGCCCGTCGACAAACGGCATGCTCGGCAGCATCGATGCCAAC<br>CGGGCGACGTT CAGAACGGCTGGACACCGACCACTTCCCCATCGACAACTACGAGCTG<br>ACTCAGGCCATGCTCGAAATCATCCGCAACGGTGGTCTGGGCAACGGCGGTACCAACTTC<br>GACGCCAAGATCCGTCTGTAAC TACCGACCTCGAGGATCTGTTCATCGCCCA CATCAGC<br>GGTATGGACGCCCATGGCACGTGCAC TGTCAATGCAGCAGCCATAC TGGAGGAGAGCGAG<br>CTGCTGCCATGAAGAAGGAGCGTTACGCCAGCTTCGACAGCGGCATCGGCAAGGACTTC<br>GAGGACGGCAAGCTCACACTTGAGCAGGCCCTATGAGTATGGTAAGAAGTTGAGGAGCCA<br>AAGCAGACCTCTGGCAAGCAGGAGAAGTATGAGACTATAGTAGCCCTCTACGCTAAGTAG<br>MAKEYFPFVGKIPFEGKDSKNVMAFHYYEPEKVVMGKKMKDWLKFAMAWWHTL GQASADP<br>FGGQTRSYEWDKADDAVQRAKDKMDAGFEIMDKLGI EYFCFHDVDLVEE GATVEEY EARM<br>KAITDYALEKMKEYPNIKLLWGTANVFSNKR YMNASTNPDFDVVARAI VQIKNAIDATI |
| 5586MI1182_004 | <i>Prevotella</i> | Amino Acid       | 98         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|               |                   |                  |            | <u>KLGGQNYVFWGGREGYMSLLNTDQKREKEHMA</u> <u>TMLTMARDYARKNGFKGTFLIEPKPMEP</u><br><u>SKHQYDVDTE</u> <u>TVCGFLRAHGLDKDFKVNIEVN</u> <u>HATLAGHTFEHELACAVDNGMLGSIDAN</u><br><u>RGDVQNGWDTDQFPIDNYELTQAMLEII</u> <u>IRNGGLNGGTFNDAKIRRNSTDLEDLFIAHIS</u><br><u>GMDAMARALLNAAAIL</u> <u>EESELPAMKKERYAS</u> <u>FD</u> <u>SGIGKDFEDGKLT</u> <u>LEQAYEYGKKVVEEP</u><br><u>KQTS</u> <u>GKQEKYETIVALYAK</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI193_004 | <i>Prevotella</i> | DNA              | 99         | <u>ATGACTAAAGAGTATTTCCCTACCAT</u> <u>TGGCAAGATTCCCTTTGAGGGACCTGAAAGCAAG</u><br><u>AACCCGCTTGCA</u> <u>TTC</u> <u>CATTACTATGAGCCCCGACCGCTGGT</u> <u>CATGGGCAAGAAGATGAAA</u><br><u>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACAC</u> <u>CCTGGGCCAGGCCTCCGGCGACCAG</u><br><u>TTCGCGGCCAGACCCGCCACTATGCCTGGGATGAT</u> <u>CCGGATTGCCCGTATGCACGTGCC</u><br><u>AAAGCCAAAGGCCGACGCCGGTTTCGAAAT</u> <u>CATGCAGAAAC</u> <u>TGGGCATTGAATTCTTCTGC</u><br><u>TTCCACGACATCGACCTGGTCGAGGATGCCGATG</u> <u>AAATCGCCGAGTACGAGGCCCGGATG</u><br><u>AAGGACATCACCGACTATCTGCTCGTCAAGATG</u> <u>AAAGAGACCGGCATCAAGAACCTTTGG</u><br><u>GGAACGGCCAAACGTATTTGGCCACAAAGCGT</u> <u>ACATGAACGGCGCCGCCACCAACCCCGAT</u><br><u>TTCGACGTGCTGGCCCGTGCCGCCGTCCAGAT</u> <u>CAAGAAACGCCATCGACGCCACCATCAAG</u><br><u>TTGGGCGGTCAGAACTATGTGTTCTGGGGCGG</u> <u>CGCTGAAGGTAC</u> <u>CAGACCCCTGCTCAAT</u><br><u>ACCCAGATGCAGCGCGAGAAAGGAACACAT</u> <u>GGCCGTATGTTGGCACTGGCCCGGACTAT</u><br><u>GGCCGTGCACACGGTTTCAAGGGCACGTTCT</u> <u>CATCGAGCCCAACCGATGGAGCCGACC</u><br><u>AGCACCAAGTACGATCAGGATACGGAAACCGT</u> <u>CATCGGCTTCTTGCGCCCGCATGGCCTC</u><br><u>GACAAGGACTTCAAGGTCAACATCGAGGTGA</u> <u>ACCATGCTACCCTGGCGGCCACACCTTC</u><br><u>GAGCACGAGCTGGCTTGGCCGTCGACCA</u> <u>CGCATGCTGGCAGCATCGACGCCAACCCGG</u><br><u>GGTGATGCCCAGAACGGCTGGGACACCGAC</u> <u>CCAGTCCCGATCCGATAACTATGAGCTGACG</u><br><u>CTGGCCATGCTCCAGATCATCCGCAACGG</u> <u>CGGCTGGCACCCGGCGGCTCGAACTTCGAT</u><br><u>GCGAAGCTGCGTCGAACTCCACCGATCCG</u> <u>GAAGATATCTTCATCGCGCACATCAGCGCC</u><br><u>ATGGATGCCATGGCCCGCCCTGGTCAATG</u> <u>CTGTGCGCAATCTCGAGGAATCGCCCATC</u><br><u>CCGGCCATGGTCAGGGAACGTTACG</u> <u>CCCTCGTTCGACAGCGGAAAGGCGAGGGAATATGAG</u><br><u>GAAGGCAGGCTGTCTCTCGAAGACATCGT</u> <u>GGCCTATGCCAAAGCCACGGCGAACCCGAAA</u><br><u>CAGATTTCCGGCAAGCAGGAACTCTACG</u> <u>AAACCATCGTGGCTCTCTATTGCCAAGTAG</u> |
| 5586MI193_004 | <i>Prevotella</i> | Amino Acid       | 100        | <u>MTKEYFP</u> <u>TTIGKIP</u> <u>FE</u> <u>GPESKNPLAFHY</u> <u>EPDR</u> <u>LRV</u> <u>MGKMKD</u> <u>WLR</u> <u>FAMAWWHT</u> <u>LGQASGDQ</u><br><u>FGGQ</u> <u>TRHYAWDDPD</u> <u>CP</u> <u>YARAKAKAD</u> <u>AGFEIMQ</u> <u>KLGI</u> <u>EFFCFHDI</u> <u>DLVEDADEIAEY</u> <u>EARM</u><br><u>KDITD</u> <u>YLLVKMKETGI</u> <u>KNLWG</u> <u>TANVFG</u> <u>HKRYMNGA</u> <u>ATNP</u> <u>DFDVL</u> <u>ARA</u> <u>AVQIK</u> <u>NAID</u> <u>ATIK</u><br><u>LGQNYVFWGGREGYQ</u> <u>TLLNTQM</u> <u>OREKEH</u> <u>MGRMLAL</u> <u>ARDYGR</u> <u>AHGF</u> <u>KGTF</u> <u>FLIEPKP</u> <u>MEPT</u><br><u>KHQYDQD</u> <u>TETVIG</u> <u>FLRRHGLDKF</u> <u>KNIEVN</u> <u>HATLAGHTFEHELACAVDHGMLGSIDANR</u><br><u>GDAQNGWDTDQFPIDNYELT</u> <u>LAMLQIIRNGGL</u> <u>APGGSN</u> <u>FD</u> <u>AKLRN</u> <u>STD</u> <u>PD</u> <u>EDIFIAHISA</u><br><u>MDAMARALVN</u> <u>AVAIL</u> <u>EE</u> <u>SP</u> <u>IPAMV</u> <u>RRERYAS</u> <u>FD</u> <u>SGK</u> <u>GREYEEGR</u> <u>LSLEDIVAYAKAHGE</u> <u>PK</u><br><u>QISGKQEL</u> <u>YETIVALYCK</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5586MI195_003 | <i>Prevotella</i> | DNA              | 101        | <u>ATGGCAAAAGAGTATTTCCCGCAGATCGG</u> <u>AAAGATCGGCTTTGAGGGTCTCTGCAAGCAAG</u><br><u>AACCCGCTGGC</u> <u>ATTCCATTAT</u> <u>TATGACCGCGAGCGCGTGGT</u> <u>GATGGGTAAACCCATGAAA</u><br><u>GACTGGTTTAA</u> <u>ATTCGCCCTCGCGTGGTGGCACAGCCCTCGGCCAGGCCTCCGGCGACCCG</u><br><u>TTCGGGCGGCAGACCCGCTCCTACG</u> <u>AGTGGGACAAGGGCGAATGCCCCTACTGCGCGGCC</u><br><u>CGGCCAAAGGGGACGCCGGCTTCG</u> <u>AGATCATGCAAAAGCTCGGCATCGGCTATTTCTGC</u><br><u>TTCCACGACGT</u> <u>CGACCTCATCGAAGACACGGAC</u> <u>GACATCGCCCGAATATGAGGCCCCGCTC</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------|-------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|               |                   |                  |            | AAGGACATCACGGACTACCTGCTCGAAAGGATGCAGGAAACCGGCATCAAGAACCTCTGG<br>GGCAGGCCCAATGTCTTCGGTCACAAAGCGCTACATGAACGGCGCGCACCAATCCGCAG<br>TTCGACATCGTCGCCCGCGCTGCCGTCCAGATCAAGAACGCCCTCGACGCCACCATCAAG<br>CTCGGTGGCTCGAACTACGTCTTCTGGGGCGGCCGGAAGTTATTACAGCTGCTCAAC<br>ACCCAGATGCAGCGCGAGAAAGACCACTCGCCAAGCTCCTCACCGCCGCCGCGACTAT<br>GCCCCGCCAAAGGCTTCCAGGGCACCTTCTGATCGAGCCCAAGCCGATGGAGCCGACC<br>AAGCACCAAGTACGATGTCGACACGGAGACTGTAATCGGATTCCTCCGCGCCAACGGACTG<br>GACAAGGACTTCAAGGTCAACATCGAGGTCAACCACGCCACCTCGCCGGCCATACCTTC<br>GAGCATGAGCTGACCGTCGCCCGCGAGAACGGATTCTCGGCAGCATCGACGCCAAACCGC<br>GGTGACGCCCAGAACGGCTGGGACACCGACCAAGTTCCCGTGGACGCCTACGACCTCACC<br>CAGGCCATGATGCAAGTGCTCCTGAACGGCGGTTTCGGCAACGGCGCACCAATTTTCGAC<br>GCCAAGCTCCGTCGCAGCTCCACCGATCCCGAGGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCCCACGCCCTGCTGAACGCCCGGCCATTCTCGAGGAGAGCCCCGTG<br>CCCGCATGGTCAAGGAGCGTTACGCTCCTTCGACAGCGGTCTGGCAAGCAGTTCGAG<br>GAGGAAAAGGCCACGCTGGAGGACCTCTACGACTACGCCAAGGCCCATGGCGAGCCCCGTC<br>GCCGCTCCGGCAAGCAGGAACTGTGTAAACTTACCTGAATCTGTATGCAAAAGTAA                                                                                    |
| 5586MI195_003 | <i>Prevotella</i> | Amino Acid       | 102        | MAKEYFPQIGKIGFEGPASKNPLAFHYIDAERVVMGKPMKDWKFALAWHSLGQASGDP<br>FGQTRSIEWDKGECPCYCRARAKADAGFEIMQKLIGYFCFHDVDLIEDTDDIAEYEARL<br>KDITDYLLERMQETGIKNLWGTANVFGHKRYMNGAGTNPQDIVARAAVQIKNALDATIK<br>LGGSNVVFWGGREGYYTLLNTQMQRKDHAKLLTAARDYARAKGFQGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANR<br>GDAQNGWDTDQFPVDAYDLTQAMQVLLNGGFNGGTNFDAKLRSSSTDPEDIFIAHISA<br>MDAMAHALLNAAAILEESPLPAMVKERYASFDSGLGKQFEEGKATLEDLYDYAKAHGEPV<br>AASGKQELCETYLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5586MI196_003 | <i>Prevotella</i> | DNA              | 103        | ATGACAAAAGAGTATTTCCCTACCATCGGCAAGATCCCCTTTGAGGGACCCGAGAGCAAA<br>AACCCCTCGCTTTTCATTACTATGAGCCCCGACCGCTTGGTCATGGGCAAGAAGATGAAA<br>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACACCTTGGGCCAGGCCTCCGGCGACCAG<br>TTTGGCGGCAGACCCGCCACTATGCTGGGATGATCCGGATTGCCCGTATGCACGTGCC<br>AAAGCCAAAGCCGACGCCGGTTTCGAAATCATGCAGAAAACCTGGGCATTGAATTCTTCTGC<br>TTCACGACATCGACCTGATCGAGGATACCGATGACATCGTCGAGTATGAGGCCCGGATG<br>AAGGACATCACCGACTATCTGCTGGTCAAGATGAAAGAGACCGGCATCAAGAATCTCTGG<br>GGAACGGCCAAACGTATTCGGGCACAAAGCGCTATATGAACGGCGCTGCCAACCAACCCGAT<br>TTCGACGTGCTGGCCCGTGCGCCGTCCAGATCAAGAACGCCCATCGACGCCACCATCAAG<br>CTGGCGCGCCAGAAATTATGTGTTCTGGGCGGGCGTGAAGGCTACCAGAGCCTGCTCAAT<br>ACCCAGATGCAGCGCGAAAAGGAACACATGGCCGTATGTTGGCACTAGCCCGGACTAT<br>GGCCGTGCACACCGGTTTCAAGGGCACGTTCTTCATCGAGCCCCAAACCGATGGAGCCGACC<br>AAGCACCAAGTACGATCAGGATACGGAGACCGTTCATCGGTTTTCTGGCGCCCATGGCCTC<br>GACAAGGACTTCAAGGTCAACATCGAGGTGAACCATGCTACCTTGGCGGCCACACCTTC<br>GAGCACGAGCTGGCCTGCGCCGTCGACCACCGCATGCTGGGCAGTATTGACGCCAAACCGC<br>GGTGACGCCCCAGAACGGCTGGGACACCGACCAAGTTCCCGATCGATAACTATGAGCTGACG<br>CTGGCCATGCTCCAGATCATCCGCAACGGCGGCCTGGCACCCCGCGGCTCGAACTTCGAT |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|               |                   |                  |            | GCGAAGCTGCGCAACTCCACCGATCCGGAAGATATCTTCATCGGCACATCAGCGCC<br>ATGGATGCCATGGCCCGCCCTGGTCAACGCTGTCGCCATTCTTGAGGAATCGCCCAT<br>CCGACATGGTCAAGGAGCGCTACGTTCTGTCACAGCGGAAAAGGCAGGAGTACGAA<br>GAGGGAAACTTTCCTTCGAGGACCTCGTGGCTATGCCAAAGCCACGGCGAACCGAAA<br>CAGATTCCGGCAAGCAGGAACTCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 5586MI196_003 | <i>Prevotella</i> | Amino Acid       | 104        | MTKEYFPTIGKIPFEGPESKNPLAFHYYPEPDRILVMGKKMKDWLRFAMAWHHTLGQASGDQ<br>FGGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDI DLIEDTDDIVEYEARM<br>KDITDYLLVKMKETGIKNLWGNTANVFGHKRYMNGAATNPDFDVLARAAVQIKNAIDATIK<br>LGGQNYVFWGGREGYQSLNLTQMQREKEHMGRLALARDYGRAHGFKGTFLIEPKPMEPT<br>KHQYDQDTETVIGFLRRHGLDKDFKVNIEVNHATLAGHTFEHELACAVDHGMLGSI DANR<br>GDAQNGWDTDQFFIDNYELTLAMLQIIRNGGLAPGGSNFDAKLRNSTDPEDIFIAHISA<br>MDAMARALVNAVAILEESPIDPMVKERYASFDSGKGREYEEGKLSFEDLVAYAKAHGEPK<br>QISGKQELYETIVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI197_003 | <i>Prevotella</i> | DNA              | 105        | ATGACAAAAGAGTATTTCCTACCATCGGCAAGATCCCCTTTGAGGGACCCGAGAGCAAA<br>AACCCCTCGCTTTTCATTACTATGAGCCCCGACCGCTTGGTCATGGGCAAGAAGATGAAA<br>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACACCCCTGGGCCAGGCCTCCGGCGACCCAG<br>TTTGGCGGCCAGACCCGCACTATGCCTGGGATGATCCGGATTGCCCGTATGCACGTGCC<br>AAAGCCAAAGGCCGACGCCGTTTCGAAATCATGCAGAAACTGGGCATTGAAATTCTTCTGC<br>TTCACGACATCGACCTGATCGAGGATACCGATGACATCGTCGAGTATGAGGCCCGCGATG<br>AAGGACATCACCGACTATCTGCTGGTCAAGATGAAAGAGACCGGCATCAAGAATCTCTGG<br>GGAACGGCCAAACGTATTTCGGGCACAAAGCGCTATATGAACGGCGCTGCCAACCAACCCCGAT<br>TTCGACGTGCTGGCCCGTCCCGCCGCCAGATCAAGAACCCCATCGACGCCACCATCAAG<br>CTGGCGGCCAGAAATTATGTGTTCTGGGCGGCGTGAAAGGTACAGAGCCTGCTCTCAAT<br>ACCCAGATGCAGCGCGAAAAGAACACATGGGCCGTATGTTGGCACTAGCCCGCGACTAT<br>GGCCGTGCACACCGGTTTCAAGGGCACGCTCCTCATCGAGCCCCAAACCGATGGAGCCGACC<br>AAGCACCACTACGATCAGGATACGGAGACCGTCTCATCGGTTTTCTGGCCCGCCATGGCCTC<br>GACAAGGACTTCAAGGTCAACATCGAGGTGAACCATGCTACCTGGCGGGCCACACCTTC<br>GAGCACGAGCTGGCCTGGCCGTGACACCGCATGCTGGGCAGTATTGACGCCAAACCGC<br>GGTGACGCCCAGGACGGCTGGGACACCGACCCAGTTCCTCCGATCGATAACTATGAGCTGACG<br>CTGGCCATGCTCCAGATCATCCGCAACGGCGGCTGGCACCCCGGCTCGAACTTCGAT<br>GCGAAGCTGCGTCGCAACTCCACCGATCCGGAAGATATCTTCATCGCGCACATCAGCGCC<br>ATGGATGCCATGGCCCGCCCTGGTCAACGCTGTCGCCATTCTTGAGGAATCGCCCAT<br>CCGACATGGTCAAGGAGCGCTACGCTTCGTTTCGACAGCGGAAAAGGCAGGAGTACGAA<br>GAGGGAAACTTTCCTTCGAGGACCTCGTGGCTATGCCAAAGCCACCGCGCAACCGAAA<br>CAGATTCCGGCAAGCAGGAACTCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG |
| 5586MI197_003 | <i>Prevotella</i> | Amino Acid       | 106        | MTKEYFPTIGKIPFEGPESKNPLAFHYYPEPDRILVMGKKMKDWLRFAMAWHHTLGQASGDQ<br>FGGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIEFFCFHDI DLIEDTDDIVEYEARM<br>KDITDYLLVKMKETGIKNLWGNTANVFGHKRYMNGAATNPDFDVLARAAAQIKNAIDATIK<br>LGGQNYVFWGGREGYQSLNLTQMQREKEHMGRLALARDYGRAHGFKGTLLIEPKPMEPT<br>KHQYDQDTETVIGFLRRHGLDKDFKVNIEVNHATLAGHTFEHELACAVDHGMLGSI DANR<br>GDAQDGWDTDQFFIDNYELTLAMLQIIRNGGLAPGGSNFDAKLRNSTDPEDIFIAHISA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|               |                   |                  |            | <u>MDAMARALVNAVAII</u> EE <span style="font-family: monospace;">SPI</span> <span style="font-family: monospace;">P</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">M</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">F</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">L</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">F</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">L</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">H</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">P</span> <span style="font-family: monospace;">K</span><br><u>QISGKQ</u> ELYETIVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI199_003 | <i>Prevotella</i> | DNA              | 107        | ATGACAAAAGAGTATTTCCCTACCATCGGCAAGATCCCCTTTGAGGGACCCGAGAGCAAAA<br>AACCCCTCGCTTTTCATTACTATGAGCCCGACCGCCTGGTCATGGGCAAGAAGATGAAA<br>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACACACCTGGGCCAGGCCTCCGGCGACCCAG<br>TTTGGCGGCCAGACCCGCCACTATGCCTGGGATGATCCGGATTGCCCGTATGCACGTGCC<br>AAAGCCAAAGGCCGACGCCGGTTTCGAAATCATGCAGAAACTGGGCATTGAATTCTTCTGC<br>TTCCACGACATCGACCTGATCGAGGATACCGATGACATCGTCGATGAGGCCCGGATG<br>AAGGACATCACCGACTATCTGCTGGTCAAGATGAAAGAGACCGGCATCAAGAATCTCTGG<br>GGAACGGCCAACGTATTTCGGGCACAAAGCGCTATATGAACGGCGCTGCCACCAACCCCGAT<br>TTCGACGTGCTGGCCCGTCCGCCGTCAGATCAAGAACGCCCATCGACGCCACCATCAAG<br>CTGGCGGCCAGAAATTATGTGTTCTGGGGCGGCGTGAAGGTACCAGAGCCTGCTCAAT<br>ACCCAGATGCAGCGCGAAAAGGAACACATGGGCCGTATGTTGGCACTAGCCCGCGACTAT<br>GGCCGTGCACACGGTTTCAAAGGCACGTTCTTCATCGAGCCCAAAACCGATGGAGCCGACC<br>AAGCACCAGTACGATCAGGATACGGAGACCGTCATCGGTTTTCTGGCCCGCCATGGCCCTC<br>GACAAGGACTTCAAGGTCAACATCGAGGTGAACCATGCTACCTGGCGGGCCACACCTTC<br>GAGCAGAGCTGGCCTGGCCCGTCGACCACGGCATGCTGGGCAGTATTGACGCCAAACCGC<br>GGTGACGCCCAGAACGGCTGGGACACCGACCACTTCCCGATCGATAACTATGAGCTGACG<br>CTGGCCATGCTCCAGATCATCCGCAACGGCGCCTGGACCCCGGCTCGAACTTCGAACTTCGAT<br>GCGAAGCTGCGTCGCAACTCCACCGATCCGGAAGATGTCTTCATCGCGCACATCAGCGCC<br>ATGGATGCCATGGCCCGCCCTGGTCAACCGCTGTCCCATTTCTTGAGGAATCGCCCAT<br>CCGACATGGTCAAGGAGCGCTACGTTCTGTCGACAGCGGAAAAGGCAGGGAGTACGAA<br>GAGGGAAACTTTCCTTCGAGGACCTCGTGGCCTATGCCAAAGCCACGGCGAACCCGAAA<br>CAGATTTCCGGCAAGCAGGAACTCTACGAAACCATCGTGGCTCTCTATTGCAAGTAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5586MI199_003 | <i>Prevotella</i> | Amino Acid       | 108        | MTKEYFPPTIGKIPFEGPESKNPLAFHYYPEPDRLVMGKKMKDWLRFAMAWHHTLGQASGDQ<br>FGGQTRHYAWDDPDCPYARAKAKADAGFEIMQKLGIFFCFHDIDLIEDTDDIVEYEARM<br>KDITDYLLVKMKETGIKNLWG <span style="font-family: monospace;">TAN</span> VF <span style="font-family: monospace;">G</span> H <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">M</span> <span style="font-family: monospace;">N</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">T</span> <span style="font-family: monospace;">N</span> <span style="font-family: monospace;">P</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">F</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">L</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">Q</span> <span style="font-family: monospace;">I</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">N</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">I</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">T</span> <span style="font-family: monospace;">I</span> <span style="font-family: monospace;">K</span><br><u>LGQN</u> YVFWGGREGYQSL <span style="font-family: monospace;">LNT</span> QMQREKEHMG <span style="font-family: monospace;">RML</span> ALARDYGRAHGF <span style="font-family: monospace;">KGT</span> FLIEPKPMEPT<br><u>KHQYDQ</u> DTETVIGFLRRHGLDKDFKVNIEVNHATLAGHTFEHELACAVDHGMLGSI <span style="font-family: monospace;">DANR</span><br><u>GDAQ</u> NGWDTDQFPIDNYELTLAM <span style="font-family: monospace;">LQI</span> IRNGGLAPGGSNFD <span style="font-family: monospace;">AKLR</span> NSTDPEDVFI <span style="font-family: monospace;">AHISA</span><br><u>MDAMARALVNAVAII</u> EE <span style="font-family: monospace;">SPI</span> <span style="font-family: monospace;">P</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">M</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">F</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">R</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">L</span> <span style="font-family: monospace;">S</span> <span style="font-family: monospace;">F</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">D</span> <span style="font-family: monospace;">L</span> <span style="font-family: monospace;">V</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">Y</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">K</span> <span style="font-family: monospace;">A</span> <span style="font-family: monospace;">H</span> <span style="font-family: monospace;">G</span> <span style="font-family: monospace;">E</span> <span style="font-family: monospace;">P</span> <span style="font-family: monospace;">K</span><br><u>QISGKQ</u> ELYETIVALYCK |
| 5586MI200_003 | <i>Prevotella</i> | DNA              | 109        | ATGGCAAAAAGAGTATTTCCCACAAATCGGAAAAGATCCCCTTCGAGGGCGTTGAGAGCAAG<br>AATCCCTTGCTTTCCATTATTATGACGCCGAGCGCGTGGTCATGGGCAAGCCCATGAAG<br>GACTGGTTCAAGTTCGCGATGGCCTGGTGGCACACACCTGGGCCAGGCTTCGGCGGACCCG<br>TTCGGCGGCCAGACCCGCTCCTACGAGTGGGACAAGGGCGAGTCCCCCTACTGCCCGGCC<br>CGGCCAAGGCTGACGCCGGCTTCGAGATCATGCAGAAAGCTCGGAATCGGCTACTATTGC<br>TTCACGACATCGACCTGGTGGAGGACACCGAGGACATCGCCGAATACGAGGCCCGCATG<br>AAGGACATCACCGACTACCTCGTCGAGAAGCAGAAGGAGACCGGCATCAAGAACCTCTGG<br>GGCACCGCGAACGTGTTCCGCAACAAGCGCTACATGAACGGCGCCCGCACGAACCCGCGAG<br>TTCGACATCGTCCGCCCGCGGCCCTGCAGATCAAGAACCGCATCGATGCCACCATCAAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|               |                   |                  |            | CTCGGGCACCGGTACGTGTTCTGGGGCGCGGGAAGGCTACTACACCCCTGCTGAAC<br>ACCCAGATGCAGCGCGAGAAAGGACCACCTCGCCAAGATGCTCACCGCGCGCCGACTAC<br>GCGCGCCCAACGGCTTCAAGGGCACCTTCTCATCGAGCCCAAGCCGATGGAGCCACCC<br>AAGACCAATACGACGTGGACACGGAGACCGTGATCGGCTTCTCGCGCCAATGGCCTG<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCTCGCCGGCCACACCTTC<br>GAGCACGAGCTCACCGTGGCCGTTGACAACGGCTTCTCGGCAGCATCGACGCCAACCGC<br>GGGACGCCCCAGAACGGCTGGGATACCGACCAAGTTCCTGGGTGGATCCGTACGATCTCACC<br>CAGGCGATGATCCAGATCATCCGCAACGGCGGCTTCAAGGACGGCGGCACCAACTTCGAC<br>GCCAGGCTCGCGCGCTTTCACCGACCCGGAGGACATCTTCATGCCCACATCAGCGCG<br>ATGGACGCCATGGCCACGCCCTGCTGAACGCCCGCCGCTCATCGAGGAGAGCCCGCTC<br>TGGGAGATGGTCGCCAAGCGTTACGCTTCTTCGACAGCGGCTTCGGCAAAAAGTTCGAG<br>GAAGGCAAGGCCACCTCGAGGAACTCTACGAGTATGCCAAGGGAACGGTGAGGTC AAG<br>GCCGAATCCGGCAAGCAGGAGCTCTACGAGACCTTCTGAACCTCTACGCCGAAATAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI200_003 | <i>Prevotella</i> | Amino Acid       | 110        | <u>MAKEYFPTIGKIPFEGVESKNPLAFHYIDAERVVMGKPMKDWFKFAMAWHHTLQASADP</u><br><u>FGGQTRSYEWDKGECPYCRARAKADAGFEIMQKLIGIGYCFHDIIDLVEDTEDIAEYFARM</u><br><u>KDITDYLVEKQKETGIKNLWG</u> TANVFGNKRYMNGAATNPQFDIVARAALQIKNAIDATIK<br>LGGTGYVFWGGREGYYTLLNTQMØREKDHAKMLTAARDYARANGFKGFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKNIEVNHATLAGHTFEHELTVAVDNGFLGSI DANR<br>GDAQNGWDTDQFPVDPYDLTQAMIQIIIRNGGFKDGGTNF DARLRRSSTDPEDIFIAHISA<br><u>MDAMAHALLNAAA</u> AVIEESPLCEMVAKRYASFDSGLGKKFEEGKATLEELYEYAKANGEVK<br>AESGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 5586MI203_003 | <i>Prevotella</i> | DNA              | 111        | ATGGCACAAAGCGTATTTTCTACCATCGGGAAAAATCCCCTTCGAGGGACCCGAAAGCAAG<br>AATCCCTTGGCATTCATTATTATGAGCCCCGACCGCCTGGTCCTGGGCAAGAAGATGAAG<br>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACACACGCTGGGCCAGGCTTCGGCGGACCAG<br>TTCGGCGGCCAGACCCGCCACTACGCTGGGACGAGCCCGCCACGCCCTTGGAAACGGGCC<br>AAGGCCAAGGGGATGCCGCTTCGAGATCATGCAGAAACTGGGCATCGAATTCCTCTGC<br>TTCCACGATGTGGACCTCATCGAAGAGGGCGCCACGATCGAGGAATACGAGACGGGATG<br>CAGCAGATCACGGATTATCTGCTGGTCAAGATGAAAGAGACCGGCATCCGCAACCTCTGG<br>GGTACGGCCAACGTGTTCCGACACGAGCGCTACATGAACGGCGCGGCCACGAACCCCGAT<br>TTCGATGTCGTGGCCCGCGCGCGTGCAGATCAAGACGGCCATCGACGCCACCATCAAG<br>TTGGCGCGCGAGAACTATGTGTTCTGGGGCGCGCGGAAGGCTATATGAGCCTGCTCAAT<br>ACGAGATGCACCGGAGAAAGCTGCACTGGGCAAGATGCTCGCCGGCGCGCGGACTAC<br>GGACGCGCCACGGCTTCAAGGGGACCTTCTCATCGAACCCAAAGCCGATGGAACCCACCC<br>AAGCATCAGTATGACCAGGATACGGAGACGGTTCATCGGTTTCTGCGCCGCTACGGCCTG<br>GACGAAGACTTCAAGGTGAACATCGAGGTCAACCACGCTACGCTGGCGGCCATACCTTC<br>GAACACGAACTGGCCACGGCGTGCATGCCGGCCTGCTGGGCAGCATCGACGCCAACCGC<br>GGGACGCCCAGAACGGCTGGGATACCGACCAAGTTCCTGGATCGACAACTACGAACTGACC<br>CTGGCGATGCTGCAGGTCA TCCGCAACGGCGGTCTGGCCCGGGGCTCGAATTTTCGAT<br>GCCAAGCTCCGCGCGGAAC TCCACCGATCCGGAAGACATCTTCATTGCCCACATCAGCGCG<br>ATGGATGCGATGGCGCGGCCCTGCTCAATCGGGCGGCCCTCTCGGAGACGTCCTCCCGATT<br>CCGGCGATGGTCAAGGCGCGTTACGCTTCGTTTCGACAGCGCGCGCGCAAGGATTTTCGAA |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 5586MI203_003 | <i>Prevotella</i> | Amino Acid       | 112        | GAGGGAAGGATGACGCTGGAAGACCTCGTGGCCTATGCCAGGACCCACGGCGAGCCGAAGCGGACCTCGGGCAAGCAGGAACTCTATGAGACCCCTCGTGGCGCTTATTGCCAAATAGMAQAYFPTIGKIPFEGPESKNPLAFHYEPDRLVLGKKMKDWLRFAMAWWHTLQASGDQFGGQTRHYAWDEPATPLERAKAKADAGFEIMQKLGIEFFCFHVDVLIIEGATIEEYEQRMQQITDYLLVKMKETGIRNLWGTANVFGHERYMNGAATNPDEFDVVARAAVQIKTAIDATIKLGGENYVFWGGREGYMSLLNTQMHREKLHIGKMLAAARDYGRAHGFKGTFLIEPKPMEPTKHQYDQDTETVIGFLRRYGLDEDFKVNIEVNHATLAGHTFEHELATAVDAGLLGSIDANRGDAQNGWDTDQFPIDNYELTLAMLQVIRNGGLAPGGSNFDAKLRRNSTDPEDIFIAHISAMDAMARALLNAAALCETSPIPAMVKARYASFDSGAGKDFEEGRMTLEDLVAYARTHGEPRTSQKQELYETLVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5586MI205_004 | <i>Prevotella</i> | DNA              | 113        | ATGACCAACGAGTATTTTCCCGGAATCGGTGTGATTCCGTTTGAAGGACAGGAAAGCAAGAAATCCCTGGCTTCCATTATTATGACGCCAACCGCGTAGTAGTGGGCAAAACCCATGAAGGAATGGTTCAAATTTGCCATGGCCTGGTGGCATACGCTGGGCAGGCATCGGCCGATCCC TTCCGGCGGACAGACCCCGCTCCTACGCATGGGACAAAGGCGAGTGCCCTTACTGCCGTGCCCGCCAGAAAGGCCGACGCCGGCTTTGAACATGATGCAGAAAGCTGGGAATCGGCTATTTCTGCTCCACGATGTGAATATCATCGAGGACTCGGAGGACATTGCCGAGTATGAGGCCCGTATG AAGGACATCACGGAATCTGCTGGTGAAGATGAAGAAACGGGCATCAAGAAATCTGTGGGGCACGGCCAAACGTCCTCGGCCACAAAGCGCTATATGAACGGCGCGCCACCAACCCGCAA TTTCGACGTGGTAGCCCGCTGCGGTCCAGATCAAGAAACGCCCTGGACGCCACCATCAAGCTGGGGCGCAGCAATTATGTGTTCTGGGGCGCGCGGAGGCTACTACACCTTTTGAACACGCAGATGCAGCGGGAGAAAGCACCACTGGCCAGATGCTCAAGGGCGCCCGGACTATGCCCCGGCAAGGGATTCAAGGGCACGTTCTCATTTAGCCCCAAGCCCATGGAGCCCCACCAGCACCAAGTACGACGTACGATACGAGACCAAGTCCCTGGCAGCATCGACGCCAAACCGC GAGCACGAGCTCACCGTGGCCCCGGAAAACGGCTTCTGGCAGCATCGACGCCCAACCGC GGAGACGCCCAGAACGGCTGGGATACAGACCAAGTCCCCGTGGACGCTTTGACCTCACC CAGGCCATGATGCAAGTCTGCTCAACGGCGGATTGGCAACGGCGGCACCAACTTCGACGCCAAACTGCGCCGTTCTCCACGGATCCCGAGGACATCTTCATCGCCCATCAGCGCCATGGACGCCATGGCCACGCCCTCCTGAACGCCCGCCCATCCTGGAAGAGAGCCCCCATG CCGGGCATGGTGAAGGAGCGCTACGCTTCTTCGACAAATGGCCTTGGCAAGAAGTTCGAGGAAGGAAAGGCCACGCTGGAAGAGCTGTACGACTATGCCAAGAAGAACGGCGAGCCGTG GCCGCTTCCGGAAAGCAGGAACTGTACGAAACGCTGCTGAACCTGTACGCCCAAGTAA |
| 5586MI205_004 | <i>Prevotella</i> | Amino Acid       | 114        | MTNEYFPGIGVIPFEGQESKNPLAFHYIDANRVVMGKPMKEWFKFAMAWWHTLQASADPFGGQTRSYAWDKGECPCYCRARQKADAGFELMQKLGIGYFCFHDVNI IEDCEDIAEYEFARMKDI TDYLLVKMKETGIKNLWGTANVFGHKRYMNGAATNPQDFDVVARAAVQIKNALDATIKLGGSNYVFWGGREGYYTLLNTQMQRKDHQLAQMKAARDYARGKFGKGTFLIEPKPMEPTKHQYDQDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANRGDAQNGWDTDQFPVDAFDLTQAMMQVLLNGGFGNGGTNFDAKLRRSSTDPE DIFIAHISAMDAMAHALLNAAALILESPMPGMVKERYASFNDNGLGKKFEEGKATLEELYDYAKKNGEPV AASGKQELYETILNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---------------|-------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5586MI206_004 | <i>Prevotella</i> | DNA              | 115        | ATGGCAAAAGAGTATTCCCGACTATCGGCAAGATTCCCTTCGAGGGCGTCGAATCCAAG<br>AACCCGATGGCATTCACACTATTATGACGCGAAACGCGTCGTGATGGGCAAGCCCATGAAG<br>GACTGGCTCAAGTTCGCGATGGCCTGGTGGCACACCTGGGACAGGCTTCGGCGACCCG<br>TTCGGCGGCAGACCCGTTCTACGAGTGGGACAAGGGCAGTGCCCTACTGCCGCGCC<br>AAGGCCAAGGCCGACGCCGCTTCGAGATCATGCAGAAACTGGGCATCGAGTACTACTGC<br>TTCCATGACATCGACCTGGTGGAGGACACCGAGGACATCGCCGAGTACGAGGCCCGCATG<br>AAGGACATCACCGACTACCTCGTCGAGAAGCAGAAGGAGACCGGTATCAAGAACCTCTGG<br>GGACGGCCAAACGTGTCGGCAACAAAGCGTACATGAACGGCGCCGCCAACCCCGCAG<br>TTCGACGTGTCGCCCGCGCGCGTCCAGATCAAGAAAGCCATCGACGCCACCATCAAA<br>CTCGCGGCACCTCTTACGTGTTCTGGGCGCGCTGAAGGCTACTACACCTCCTCCTGAAC<br>ACCCAGATGCAGCGCGAGAAAGGACCACTCGCCAAGATGCTCACCGCGCCCGCGACTAC<br>GCCCCGCCACGGCTTCAAGGGCACCTTCTCATCGAGCCCAAGCCCATGGAGCCCAAC<br>AAGCACAGTACGACGTGGACACGGAGACCGTGTATCGGCTTCTCCGCGCCAAAGGCTG<br>GACAAGGACTTCAAGGTCAATATCGAAGTGAACACGCCACCTCGCGGCCACACCTTC<br>GAGCATGAGCTCACCGTGGCGTCGATAACGGCTTCTCGGCTCCATCGACGCCAAACCGT<br>GGGACGCCCCAGAACGGCTGGGATACCGACCACTTCCCGGTGGATCCGTACGACCTCAC<br>CAGGCCATGATGCAGATCATCCGCAACGGCGGCTTCAAGGACGGCGGCACCAACTTCGAC<br>GCCAAACTCCGCCGCTCTCACCGACCCGGAGGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCACGGCTCCTGAACGCCCGCGCTCATCGAGGAGAGCCCGCTC<br>TGCAAGATGGTCGAGGAGCGCTACGCTTCTTCGACAGCGGTCTCGGCAAGCAGTTCGAG<br>GAAGGCAAGGCCACCTTGAGGACCTCTACGAGTATGCCAAGAAACGGCGAGCCCGCTC<br>GTCGCTTCCGGCAAGCAGGAGCTCTACGAGACCTTCTGAACCTCTACGGGAAGTAG |
| 5586MI206_004 | <i>Prevotella</i> | Amino Acid       | 116        | MAKEYFPITIGKIPEGVESKNPMAFYHYDAKRVVMGPKMKDWLKFAMAWHHTLGQASGDP<br>FGGQTRSVEWDKGECPYCRAKAKADAGFEIMQKLGIEYYCFHDI DLVEDTEDIAEYEARM<br>KDITDYLVEKQKETGIKNLWGTANVFGNKRVMNGAATNPQFDVVARAAVQIKNAIDATIK<br>LGTSYVFWGGREGYYTLLNTQMQRKDHIAKMLTAARDYARAHGFKGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKFKVNIENVNHATLAGHTFEHELTVAVDNGFLGSI DANR<br>GDAQNGWDTDQFPVPYDLTQAMMQIIRNGGFKDGGTNFDAKLRRSSTDPEIDIFIAHISA<br>MDAMAHALLNAAAIVIEESPLCKMVEERYASFDSGLGKQFEEGKATLEDLIEYAKKNGEPV<br>VASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5586MI208_003 | <i>Prevotella</i> | DNA              | 117        | ATGTCAACTGAGTATTTCCTACAAATCGGCAAGATTCCCTTCGAGGGACCCGAGAGCAAG<br>AACCCCATGGCCTTCCACTACTATGAACCCGAAAAGTTGGTGATGGGCAAGAAGATGAAG<br>GACTGGCTGCGTTTCGCAATGGCCTGGTGGCACACCTTGGAGCCGCATCCGGCGACCAG<br>TTCGGCGGACAGACCCGCAGTTACGCCTGGGACAAGGGCGCACTGCCCTTACAGCCCGGCC<br>CGGCCAAAGGTCGACGCCGCTTCGAGATCATGCAGAAAGCTCGGCATAGAGTCTTCTGC<br>TTCCATGACATCGACCTGGTCGAGGATACCGACGACATCGCCGAGTATGAAGCCCGGATG<br>AAAGACATCACGGAATCTGCTGGAAAAGATGGAGGCTACCGGCTCAAGAACCTCTGG<br>GGACGGCCAATGCTTCGGTCACAAAGCGTTATATGAACGGTGCAGCCACAAACCCCGAT<br>TTCGCAGTGGTCGCAAGGGCGCGCTGCAGATCAAGAAACGCCCATCGACGCCACCATCAAG<br>CTGGGTGGTGAGAACTATGTGTTCTGGGTGACGCGAGGGTTATATGAGCCTGCTCAAC<br>ACCCAGATGCAGAGGGAGAAAGAACACCTTGCCCAAGATGCTCACCGCCCGACGTGACTAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|               |                   |                  |            | GCACGGCCAAAGGTTTCAAGGGCACGTTCTCATCGAACCCAAAGCCGATGGAACCCACC<br>AAGCACAGTATGACCAGGATACCGAGACCGTTATCGGATTCTTCGCAGCCACGGCCTG<br>GACAAGGACTTCAAGGTCAACATCGAGGTGAACCAACGCCACCTGGCGGCCATACCTTC<br>GAGCAGAACTGGCCACCGCGTCGACAACGGCATGCTGGCAGCATCGACGCCAACCGC<br>GGAGACGCCCAGAACGGCTGGGACACCGACCAAGTCCCGATCGACAACTTCGAGCTCACG<br>CTTGCCATGATGAGATAATCCGCAACGGCGCCTGGACCGGGCGGTTCGAACTTCGAC<br>GCAAAGCTGGCCGCAATTCCACCGATCCCGAGGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCCGCCCTCGTCAACGCCGCCCATCTCTGGCGAGTCGCCCGTT<br>CCGGCTATGTTCAAGGACCGCTATGCTTCGTTTCGACTGGGCAAGGCAAGGACTTCGAA<br>GACGGCAAACTGACTCTCGAAGACATCGTCGCCTACGCCAGGAGAATGGCGAGCCGAAA<br>CAGATTCCGGCAAGCAGGAACTCTACGAAACTATCGTCGCTCTTTACTGCAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5586MI208_003 | <i>Prevotella</i> | Amino Acid       | 118        | MSTEYFPTIGKIPFEGPESKNPMAFHYEPEKLVMGKKMKDWLRFAMAWHHTLGAASGDQ<br>FGGQTRSVAWDKDCPYSRARAKVDAGFEIMQKLGIEFFCFHDI DLVEDTDDIAEYEARM<br>KDITDYLLLEKMEATGIKNLWGTANVFGHKRYMNGAATNPDAVVARAAVQIKNAIDATIK<br>LGGENYVFWGGREGYMSLLNTQMQRKEKHLAKMLTAARDYARAKFGKGTFLIEPKPMEPT<br>KHQYDQDTETVIGFLRSHGLDKDFKVNIEVNHATLAGHTFEHELATAVDNMGMLGSI DANR<br>GDAQNGWDTDQFPIDNFELTLAMQIIRNGGLAPGGSNFDAKLRNSTDPEDIFIAHISA<br>MDAMARALVNAAAII LGESVPAMVKDRYASFDCGKGKDFEDGKLTLEDIVAYARENGEPK<br>QISGKQELYETIVALLYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 5586MI210_002 | <i>Prevotella</i> | DNA              | 119        | ATGTCATATTTTCCTACTATCGGTAACATCCCCTTTGAGGGTGTAGAGAGCAAGAAATCCC<br>CTTGCCCTTCCATTATTTATGACGCTTCCCGCGTAGTTATGGGCAAGCCCATGAAGGAGTGG<br>CTCAAGTTTGCCATGGCCTGGTGGCACACGCTGGGTCAAGCATCGGCCGACCCCTTTCCGGC<br>GGACAAAACCGCAGCTATGCTGGGACAAAGCGAGTGCCCTACTGCCGTGCCCGTGCC<br>AAGCCCGACGCCGGCTTCGAGCTCATGCAGAAACTGGGCATCGAGTATTTCTGTCTCCAC<br>GACATTGACCTCATCGAGGACTCGGACGACATTCAGAGTACGAGGCCCGCTCTGAAGGAC<br>ATTACGGACTACCTCCTGGAGAAGATGAAGAACCGGTATCAAGAACCTGTGGGGTACG<br>GCCAATGTGTTCGGTAACAAGCGTTACATGAACGGTGCTGTCTACCAACCTCAGTTTGAC<br>GTTGTGGCCCGCGTGCCGTCCAGATCAAGAACGCCATTGACGCTACCATCAAGCTGGGC<br>GGTTCCAACTATGTGTTCTGGGTGGCCGTGAGGGTTACTACACGCTTCTGAACACCCAG<br>ATGCAGCGTGAGAAAGAATCACCTGGCTGCCATGCTCAAGGCTGCCCGGACTATGCCCGC<br>GCCAACGGTTTCAAGGGCACCTTCTCATTGAGCCCAAGCCCATGGAGCCCAACCAAGCAC<br>CAGTACGACGTAGACACGGAGACCGTGATTGGATTCTCCGGCCCAACGGTCTGGAGAAG<br>GACTTCAAGGTGAACATTGAGGTGAACCAACGCTACTCTTGCCGGTCACACCTTCGAGCAC<br>GAGTCAACCGTGGCCCGTGAGAACGGCTTCTGGGTTCCATTGACGCCAACCCGCGGAGAT<br>GCCCAGAACGGCTGGGACACCGACCAAGTTCCTGGTAGATGCCCTTGACCTCACCCAGGCC<br>ATGATGCAGATTCTCCTCAACGGAGGCTCCGGCAATGGCGGTACCAACTTGACGCCCAAG<br>CTGGCCCGTTCTCCACCGACCCCGAGGACATCTTCATCGCGCACATCAGCGCCATGGAT<br>GCCATGGCTCACGCCCTGCTCAATGCAGCTGCCGTGCTGGAGGAGAGCCCGCTTTGCAAG<br>ATGGTCAAGGAGCGTTACGCTTCTTCGACACGCGGTCTTGCCAAAGCAGTTCGAGGAAGGA<br>AAGGCTACGCTGGAAGATCTGTATGCCATATGCCGTCAAGAACGGTGAGCCCGTGGTGGCT<br>TCCGGCAAGCAGGAACTGTACGAAACCTTCTTGAACCTCTATGCAAAATGGTAA |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 5586MI210_002 | <i>Prevotella</i> | Amino Acid       | 120        | MSYFPTIGNIPFEGVESKNPLAFHYDASRVVMGKPMKEWLKFAMAWWHHTLGQASADPFG<br>GQTRSYAWDKGECPCYCRARAKADAGFELMQKLGIEYFCSHDIDLIEDCDDIAEYEARLKD<br>ITDYILLEKMKKTGIKNLWGTANVFGNKRVMNGAATNPQFDVVARAAVQIKNAIDATIKLG<br>GSNYVFWGGREGYYTLLNTQMQRKNHLAAMLKAARDYARANGFKGTFLIEPKPMEPTKH<br>QYVDTEETVIGFLRANGLEKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANRGD<br>AQNGWDTDQFPVDAFDLITQAMMQILLNGSGNGGTNFDAKLRRSSTDPEIDIFIAHISAMD<br>AMAHALLNAAAALVEESPLCKMVKERYASFDSGLGKQFEEGKATLEDLYAYAVKNGEPPVVA<br>SGKQELYETFLNLYAKW                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5586MI212_002 | <i>Prevotella</i> | DNA              | 121        | ATGTCAAACGTGAGTATTTCCCTACAAATCGGCAAGATTCCCTTCGAGGGACCCCGAGAGCAAG<br>AACCCCATGGCCTTCCACTACTATGAACCCGAAAGTTGGTGATGGGCAAGAAGATGAAG<br>GACTGGCTGCGTTTCGCAATGGCTGGTGGCACACCCCTTGGAGCCGCATCCGGCGACCAG<br>TTCGGCGGACAGACCCGCAGTTACGCCCTGGGACAAGGGCGACTGCCCTTACAGCCCGGCC<br>CGGCCAAGGTCGACGCCGCTTCGAGATCATGCAGAAAGCTCGGCATAGAGTTCTTCTGC<br>TTCCATGACATCGACCTGGTCGAGGATACCGACGACATCGCCGAGTATGAAGCCCGGATG<br>AAAGACATCACGGACTATCTGCTGGAAGAAGATGGAGGTTACCGGCATCAAGAACCCTCTGG<br>GGCAGGCCCAATGTCTTCGGTCACAAAGCGTTATATGAACGATGCAGCCACAAACCCCGAT<br>TTCGCAGTGGTCGCAAGGGCGCGTCGAGATCAAGAACGCCATCGACGCCACCATCAAG<br>CTGGTGTTGAGAACTATGTGTTCTGGGTGGACGCGAGGGTTATATGAGCCTGCTCAAC<br>ACCCAGATGCAGAGGGAGAAAGGAACACCTTGCCAAGATGCTCACCGCCGACGTGACTAT<br>GCACGCCCAAAGGTTTCAAGGGCACGTTCTCATCGAAACCCGAGCCGATGGAACCCACC<br>AAGCACCAAGTATGACCAGGATACCGAGACCGTTATCGGATTCTCCGAGCCACGGCCTG<br>GACAAGGACTTCAAGGTCAACATCGAGGTGAACCAAGCCACCCCTGGCGGGCCATACCTTC<br>GAGACGAACTGGCCACCGCCGTCGACAAAGGATGCTCGGAGCATCGACGCCAACCGC<br>GGAGACGCCAGAACGGCTGGGACACCGACCAAGTTCCCGATCGACAACTTCGAGCTCACG<br>CTTGCCATGATGCAGATAATCCGCAACGGCGGCTGGACCGGGGTTTCAACTTCGAC<br>GCAAAGCTGCGCCGCAATTCCACCGATCCCGAGGACATCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCCGCCCTCGTCAACGCCCGCCCATCCTCGGCGAGTCGCCCCGT<br>CCGGCTATGTTCAAGGACCGCTATGTTGTTTCGACTGGGCAAGGGCAAGGACTTCGAA<br>GACGGCAAACCTGACTCTCGAAGACATCGTCCGCTACGCCAGGAGAAATGGCGAGCCGAAA<br>CAGATTCCCGCAAGCAGGAACCTCTACGAAACTATCGTCGCTCTTTACTGCAAGTAA |
| 5586MI212_002 | <i>Prevotella</i> | Amino Acid       | 122        | MSTEYFPTIGKIPFEGPESKNPMAFHYYEPEKLVMGKKMKDWLRFAMAWWHHTLGAASGDQ<br>FGGQTRS YAWDKGDCPYSRARAKVDAGFEIMQKLGIEFFCFHDI DLVEDTDDIAEYEARM<br>KDITDYLLLEKMEVTGIKNLWGTANVFGHKRYMNDAAATNPDFAVVARAAVQIKNAIDATIK<br>LGGENYVFWGGREGYMSLLNTQMQRKEKEHLAKMLTAARDYARAKGFKGTFLIEPEPMEPT<br>KHQYDQDTE TVIGFLRSHGLDKDFKVNIEVNHATLAGHTFEHELATAVDNMGMLGSI DANR<br>GDAQNGWDTDQFPIDNFELTLAMQIIRNGGLAPGGSNFDKLRNSTDPEDIIAHISA<br>MDAMARALVNAAAAILGESVPAMVKDRYASFDCGKGKDFEDGKLTLEDIVAYARENGEPPK<br>QISGKQELYETIVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 5586MI213_003 | <i>Prevotella</i> | DNA              | 123        | ATGACCAACGAGTATTTCCCGGAATCGGTGTGATTCCGTTTGAAGGACAGGAAAGCAAG<br>AATCCCCCTGGCTTCCATTATTATGACGCCAAACCGCGTAGTATGGGCAAAACCCATGAAG<br>GAATGGTTCAAATTTGCCATGGCCTGGTGGCATACGCTGGGGCAGGCATCGGCCGATCCC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

| TABLE 2       |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|---------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.     | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|               |                   |                  |            | TTGGGGACAGACCCGCTCCTACGCATGGGACAAAGGGCGAGTGCCCTTACTGCCGTGCCCGCAGAAAGGCCGACGCCGGCTTTGAACATGATGCAGAAAGCTGGGAATCGGCTATTTCTGCTCCACGATGTGGATATCATCGAGGACTCGAGGACATTCGGAGTATGCCGCGGTATG AAGACATCAGGACTATCTGCTGGTGAAGATGAAGGAAACGGGCATCAAGAAATCTGTGGGCACGGCCAAACGTCTTCGGCCACAAAGCGCTATATGAACGGCGCCGACCAACCCGCAA TTCACGTGGTAGCCCGCTGCGGTCCAGATCAAGAAACGCCCTGGACGCCACCATCAAGCTGGGGCGCAGCAATTATGTGTTCTGGGGCGCCGGGAAGGTACTACACCTTTTGAACACGAGATGCAGCGGGAGAAAGCACCTGGCCAGATGCTCAAGGGCGCCGCGACTATGCCGGCAAGGATTCAAAGGCACGTTCTCATTTAGGCCCAAGCCCATGGAGCCACCAAGCACAGTACGTAAGTACGGAGACCGTGATTGTTTCTGCGCGCCAAACGGGCTG GACAAGGACTTCAAGGTGAATATCGAAGTGAACCAACGCCACCTTGCCGGCCATACCTTCGAGCACGAGCTCACCGTGGCCCGGAAAACGGCTTCTGGGCAGCATCGACGCCAAACCGCGAGACGCCAGAACGGCTGGGATACAGACCAGTCCCTGGCGCGCCATACCTTCGGAGCGCCATGGCCACGCCCTCCTGAACGCCCGCCCATCCTGGAAGAGAGCCCCATGCCGGCATGGTGAAGGAGCGCTACGCTTCTTCGACAAATGGCTTGGCAAGAAGTTCGAGGAAGAAAGGCCACGCTGGAAGAGCTGTACGACTATGCCAAGAAACGGCGAGCCTGTGCGCTTCCGGAAAGCAGGAACTGTACGAAAACGCTGCTGAACCTGTACGCCAAGTAA |
| 5586MI213_003 | <i>Prevotella</i> | Amino Acid       | 124        | MTNEYFPGIGVIPFEGQESKNPLAFHYDYDANRVVMGKPKMKWFKFAMAWWHHTLGQASADP_FGGQTRSYAWDKGECYPYCRARQKADAGFELMQKLIGYFCFHDVDIIEDCEDIAEYEARMKDITDYLLVKMKETGIKNLWGTA NVFGHKRYMNGAATNPQFDVVARA AVQIKNALDATIKLGGSNVYFWGGREGYYTLLNTQM QREKDH LAQMLKAARDYARGKFKGTFLIEPKPMEPTKHQYD VDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANRGDAQNGWDTDQFPVD AFDLTQAMMQVLLNGGFGNGGTNFDAKLRRSSTDPEDIFIAHISAMDAMAHALLNAAAIILEESPMPGMVKERYASFDNGLGKFFEKGKATLEELYDYAKKNGEPV AASGKQELLYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5586MI215_003 | <i>Prevotella</i> | DNA              | 125        | ATGGCAAAAGAGTATTTCCCGCAGATCGGAAAAGATCGGCTTTGAGGGTCTTGAGAGCAAGAACCCGATGGCATTC CATTTATATGACGCCCGAGCGTGTCTGTCGTCGCGAAAAGAGATGAAGACTGGCTGAAGTTCGCGATGGCTGGTGGCATACGCTCGGACAGGCTTCGGCGACCCCA TTCCGGCGCCAGACTCGCAGCTATGAGTGGGACAAGGGCGAGTGCCCCTACTGCCGTGCCCGGCCAAAGGCCGACCGGGCTTCGAGCTCATGCAGAAAGCTCGGCATCGAGTACTTCTGCTCCACGACATCGACCTCATCGAGGACTCGGACGACATCGACGAGTACGAGGCCCGGATGAAGACATCACCGACTACCTGCTGGAGAAGATGAAGGAGACCGGAAATCAAGAATCTCTGGGAAACGGTCTTCGGTCACAAAGCGCTACATGAACGGCGCGCTACCAATCCGCAGTTTGAAATCGTCGCCCCGGCTGCCGTCCAGATCAAGAACGGCTCGACGCCACCATCAAGCTCGGGCGCTCCA ACTAGTCTTCTGGGGCGCGCGAGGGCTATTACACGCTGCTGAATACCCAGATGCAGCGCGAGAAAGGACCATCTCGCCAGGCTCCTTACCGCGCCCGGACTATGGCGCGCCAAAGGGGTTCAAGGGGACCTTCCCCCATCGAGCCGGAAGCCGATGGAGCCGACC AAGCACCAAGTATGACGTGACACGAGACCGT CATCGGTTTCTCCGCCAGAATGGCCTC GACAAGGACTTCAAGGTCAATATCGAGGTGAACCAACGCCACCTCGCCGGCCATACCTTC                                                                                                                                              |

| TABLE 2        |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------------|-------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.      | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|                |                   |                  |            | GAGCAGAGCTGACCGCGGCCCGGAGAACGGCTTCCTCGGCAGCATCGACGCCAACCGC<br>GGCAGCCCCAGAACGGCTGGGACACCGACCAAGTTCCTGGGTGGACGCCTTCGATCTCAG<br>CGGCCATGATGCAGATCCTGCTCAATGGCGGTTTCGGCAACGGCGGCACCAACTTCGAC<br>GCCAAGCTGCGCCGAGCTCCACCGATCCGAGGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCACGCCCTGCTGAATCGGCCGCCATCCTCGAGGAAAGCCCGCTG<br>CCGCCCTGGTCAAGCAGCGCTATGCTCCTTCGACAGCGGTCTCGGCAAGCAGTTCGAG<br>GAGGTAAGGCCACGCTCGAGGACCTGTACGCATACGGAAGGAGCACGGCGAGCCCCGTC<br>CGGCCCTCCGGCAAGCAGGAGCTCTGCGAGACCTATCTCAACCTCTACGCCGAAATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5586MI1215_003 | <i>Prevotella</i> | Amino Acid       | 126        | <u>MAKEYFPQIGKIGFEGLESKNPMAFHYIDAERVVLGKKMKDWLKFAMAWWHTLQASGDP</u><br><u>FGGQTRSYEWDKGECPYCRARAKADAGFELMQKLGIEYFCFHDIDLIEDCDDIDEYEARM</u><br><u>KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAATNPQFEIVARAAVQIKNALDATIK</u><br><u>LGSNYVFWGGREGYYTLLNTQMOREKDHLLRLTAARDYARAKGFKGTFPIEPKPMPT</u><br><u>KHQYDVDTETVIGFLRQNLDKDFKNIEVNHATLAGHTFEHELTAARENGFLGSIDANR</u><br><u>GDAQNGWDTDQFPVDAFDLTRAMQILLNGGFNGGTFNFDKLRRSSTDPEDIFIAHISA</u><br><u>MDAMAHALLNAAAIILEESPLPALVKQRYASFDSGLGKQFEEGKATLEDLIYAYAKEHGEPV</u><br><u>AASGKQELCETYLNLYAK</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 5607MI1_003    | <i>Prevotella</i> | DNA              | 127        | ATGAGTAAAGAGTATTTTCTCTGGGATTGGCAAAATCCCCTATGAGGGAGCCGAGAGCAAG<br>AATGTGATGGCATTCACATAATTATGATCCGAACGCGTGGTCATGGGCAAGAAAAATGAAA<br>GACTGGTTCAAAGTTCGCTATTGCTTGGTGCCATACCCCTGGGCAGGCCAGTGTGACCCAG<br>TTTGGCGGACAGACCCGTTTCTATGAATGGACAAAGCCGAGGACCCCTTGCAGCGTGCC<br>AAGGACAAAGATGGATGCCGTTTGAATCATGCAGAAAGCTGGGCATCGAGTATTTCTGT<br>TTCCATGATGTGGACCTCATCGAGGAGGCCGATACCATCGAGGAATATGAAGCCCCGATG<br>CAGCGATTACCGACTACGCGCTGGAGAAGATGAAGGCAACGGGTATCAAAGTTGCTGTGG<br>GGCACTGCCAACGTTGTCGGCCACAAAGCGTTACATGAACGGCGCGCCACCAATCCCAGC<br>TTCAATGTCGTGGCACGTGCAGCCGTGCAGATCAAGAACGCCCTCGATGCTACCATCAAG<br>TTGGCGGAAACGAGCTACGTCCTCTGTTGGGCGGTGCTGAAGGCTATCAGAGCCTGCTCAAC<br>ACCCAGATGCAGCGCGAGAAACCACCTGGCCAAAGATGCTCAGGCAGCCCCGTGACTAT<br>GCCGTGCTAAGGGCTTCAAAGGCACCTTCCTGATTGAGCCCAAAGCCGATGGAACCCACC<br>AAGCACCAAGTATGACCAGGACACCGAGACCGTTATCGGCTTCTTGGTGCCCAATGGCCTT<br>GACAAGGACTTTAAGGTCAACATTGAGGTCAACCATGCCACGCTGGCTGGCCACACCTTT<br>GCACATGAGTTGGCAGTGGCTGTGGATAACGGTATGCTGGGCAGCATCGATGCTAACCGT<br>GGTGACCAACCAGAACGGCTGGGATACAGACCAAGTTCCTCATCAACAGTTATGAACTCACC<br>AATGCTATGCTGCAGATCATGCACGGCGCGGTTTCAAGGACGGCGGTACCAACTTTGAC<br>GCCAAGCTGCGCCCGCAACAGTACCGACCCCGAGGACATCTTTACCGCTCACATCAGTGGT<br>ATGGACGCTCTGGCCCCGTCCCTGTTGAGTGTCTGCCGATATCCTTGAGAAAGCGAGTTG<br>CCTGAAATGCTCAAGGAACGCTATGCCAGCTTTGACGGGGTGAAGGCAAGCGCTTTGAG<br>GATGCCCAGATGACTCTTGAGGAACCTGGTTGCCATATGCCAAGTCCCATGGCGAGCCTTGCT<br>ACCATCAGTGGCAAGCAGGAAAAATATGAAGCCATCGTGGCTTTGCACGTCAAGTAA |
| 5607MI1_003    | <i>Prevotella</i> | Amino Acid       | 128        | MSKEYFPGIGKIPYEGAESKNVMAFHYIDPERVVMGKKMKDWFKFAIAWWHTLQASADQ<br>FGGQTRFYEWDKAEDPLQRAKDKMDAGFEIMQKLGIEYFCFHDVLDLIEEADTIEEYARM<br>QAITDYALEKMKATGIKLLWGTANVFGHKRYMNGAATNPDENNVARAAVQIKNALDATIK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------|-------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|             |                   |                  |            | <u>L</u> GTSYVFWGGREGYQSL <sup>1</sup> LNTQMQREKNHLAKMLTAARDYARAKGFKGTFLIEPKPMEPT<br><u>K</u> HQYDQDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFAHELAVAVDNGMLGSIDANR<br><u>G</u> DHQNGWDTDQFPINSYELTNAMLQIMHGGGFKDGGTNFDAKLRRNSTDPEDIFIAHISG<br><u>M</u> DALARALLSAADILEKSELPEMLKERYASFDAGEGKRFEDGQMTLEELVAYAKSHGEPA<br><u>T</u> ISGKQEKYEAI <sup>2</sup> VALHVK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5607MI2_003 | <i>Prevotella</i> | DNA              | 129        | ATGAGTAAAGAGTATTATCCTGAGATTGGCAAAATCCCCTTTGAGGGTCCCCGAGAGCAAG<br>AATGTGATGGCGTTCCATTACTATGAACCCGAACGCGTCGTCATGGGTAAGAAGATGAAA<br>GACTGGCTCAAGTTTGCCATGTGCTGGTGGAACAGCCTGGGTCAAGCCAGTGCCGACCAG<br>TTCCGGGACAGACACGTTTCTACGAGTGGGACAAGGCCGATACCCCCCTGCAGCGTGCC<br>AAGGACAAAATGGATGCCGGATTTGAAATCATGCAGAAAGTTGGGCATCGAGTACTTCTGC<br>TTCCACGATGTGGACCTCATCGAGGAGGCCGATACCATCGAGGAATACGAGGCCCGCATG<br>AAGCCATTACCGACTATGCGCTGGAGAAGATGCAGGCCACCGGCATCAAGTTGCTGTGG<br>GGCACTGCCAATGTGTTCCGCCACAAAGCGTACATGAACGGCGCCGCCACCAATCCC<br>GATTTCAATGTCGTGGACGTGCCGCCGTCCAAATCAAGAAIGCCATCGATGCCACCATCAAG<br>CTGGCGGCACGAGTTACGTCCTTCTGGGTGTCGTGAGGGCTATCAGAGTCTGCTCAAC<br>ACGCAGATGCAGCGCGAGAAAGGACCATCTGGCCCGCATGCTGGCGGAGCCCGCGACTAT<br>GGCGTGCCCATGGCTTCAAGGGCACTTTCCTGATCGAGCCCCAAACCCATGGAGCCCCACC<br>AAGCACCAAGTATGATGTGGACACCGAGACCGTGCTCGGCTTCTGCGTGCCACGGCCTG<br>GACAAGGACTTCAAGGTTAAACATCGAGGTCAATCATGCTACGCTGGCGGGACACACTTTC<br>AGCCACGAACTGGCTGTGGCCGTGGACACCGATCGTGGCAGCATCGACGCCAAACCGC<br>GGCGATTATCAGAAATGGCTGGGACACCGACCATGTTCCCCTCGACAGCTTCGAGCTCACC<br>CAGGCCATGCTGCAGATCATGCGCGCGCGCGGCTTCAAGGACGGAGGTACCAACTTCGAT<br>GCCAAGCTGCGTCGCAACAGTACCGACCTGAGGACATCTTCATCGCCACATCAGCGGT<br>ATGGATGCCATGGCACGCGGCTGTTGAGCGCTGCCGCTATCCTCGAGGATGGCGAGTTG<br>CCCGCGATGCTCAAGGCACGTTATGCCAGCTTTGACCAGGCGGAGGGTAAGCGCTTTGAG<br>GACGGCGAGATGACGCTCGAGCAGCTGGTGGATTATGCAAAGGATTATGCCAAATCGCAC<br>GGCGAGCCTGATGTCATCAGCGGCAAGCAGGAGAAAGTTTGAAACCATCGTGGCCCCTTTAC<br>GCCAAGTAA |
| 5607MI2_003 | <i>Prevotella</i> | Amino Acid       | 130        | MSKEYYPEIGKTPFEGPESKNVMAFHYYEPERVVMGKKMKDWLKFAMCWWHSLGQASADQ<br>FGGQTRFFYEWDKADTPLQRAKDKMDAGFEIMQKLGIEYFCFHDVDLIEEADTIEEYEARM<br>KAITDYALEKMQATGIKLLWGTANVFGHKRYMNGAATNPDENVVARA <sup>3</sup> AVQIKNAIDATIK<br>LGGTSYVFWGGREGYQSL <sup>1</sup> LNTQMQREKDH <sup>4</sup> LARMLAAARDYGRAHGF <sup>5</sup> GTFLIEPKPMEPT<br>KHQYD <sup>6</sup> VD <sup>7</sup> TETV <sup>8</sup> LGFLRAHGLDKDFKVNIEVNHATLAGHTFSHELLAVAVDNGMLGSIDANR<br>GDYQNGWDTDQFPIDSFELTQAMLQIMRGGGFKDGGTNFDAKLRRNSTDPEDIFIAHISG<br>MDAMARGLLSAAAI <sup>9</sup> LEDGELPAMIKARYASFDQGE <sup>10</sup> GKRFEDGEMTLEQLVDYAKDYAKSH<br>GEPDVISGKQEK <sup>11</sup> FETIVAIYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5607MI3_003 | <i>Prevotella</i> | DNA              | 131        | ATGACCAAACGAGTATTTCCCGGAAATCGGTGTGATTCCGTTTGAAGGACAGGAAAGCAAG<br>AATCCCCCTGGCTTCCATTATTATGACGCCAAACCGCGTAGTATGGGCAAAACCCCATGAAG<br>GAATGGTTCAAATTTGCCATGGCCTGGTGGCATACGCTGGGCGAGGCATCGGCCGATCCC<br>TTCGGCGGACAGACCCCGCTCCTACGCGATGGGACAAGGCCGAGTGCCTTACTGCCGTGCC<br>CGCCAGAAAGGCCGACGCCGGCTTTGAACTGATGCAGAAAGCTGGGAATCGGCTATTTCTGC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------|-------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|             |                   |                  |            | TTCCACGATGTGGATATCATCGAGGACTCGAGGACATTGCCGAGTATGAGGCCCGGTATG<br>AAGGACATCACGGACTATCTGCTGGTGAAGATGAAGGAACGGGCATCAAGAATCTGTGG<br>GGCAGGCCAACGTCTTCGGCCACAAAGCGTATATGAACGGCGCCGCCAACCCGCAA<br>TTCGACGTGGTAGCCCGCTGCGGTCCAGATCAAGAACCCCTGGACGCCACCATCAAG<br>CTGGCGGCAGCAATTATGTGTTCTGGGCGCGCCGGAAGGCTACTACACCTTTTGAAC<br>ACGCAGATGCAGCGGGAGAGACCACTGGCCCAGATGCTCAAGCGGCCCGCGACTAT<br>GCCCCGGCAAGGATTCAAGGGCACGTTCTCATTTAGCCCCAAGCCCATGGAGCCCACC<br>AAGCACAGTACGACGTAGATACGGAGACCGTGATTGGTTTCTTGGCGCCCAACGGGCGG<br>GACAAGGACTTCAAGGTGAATATCGAAGTGAACCAACGCCACCTGGCCGCCATACCTTC<br>GAGCACGAGCTCACCGTGGCCCGCGAAACGGCTTCTGGGCAGCATCGACGCCAAACCGC<br>GGAGACGCCCAGAACGGCTGGGATACAGACCAAGTTCCCCGTGGACGCCTTTGACCTCACC<br>CAGGCCATGATGCAAGTCTGCTCAACGGCGGATTTCGGCAACGGCGGCACCAACTTCGAC<br>GCCAAACTGGCCCGTTCTCCACGGATCCCGAGGACATCTTCATCGCCCATCAGCGCC<br>ATGACGCCATGGCCACGCCCTCTTGAACGCCCGCCCATCTTGAAGAGAGCCCCATG<br>CCGGCATGGTGAAGAGCGCTACGCTTCTCTCGACAAATGGCCTTGGCAAGAAGTTCGAG<br>GAAGGAAAGGCCACGCTGGAAGAGCTGTACGACTATGCCAAGAAACGGCGAGCCTGTG<br>GCCGCTTCCGGAAGCAGGAACTGTACGAAACGCTGTGAACCTGTACGCCAAGTAA                           |
| 5607MI3_003 | <i>Prevotella</i> | Amino Acid       | 132        | MTNEYFPGIGVIPFEGQESKNPLAFHYDANRVVMGPKMKEWFKFAMAWWHTLQASADP<br>FGGQTRS <del>Y</del> AWDKGECPIYCRARQKADAGFELMQKLIGYFCFHDVDIIEDCEDIAEYEARM<br>KDITDYLLVKMKETGIKNLWG <del>T</del> ANVFGHKRYMNGAATNPQFDVVVARAAVQIKNALD <del>ATIK</del><br>LGGSNVYFWGGREGYYTLLNTQM <del>Q</del> REKDHIAQMLKAARDYARGKFGKTF <del>LIEPK</del> MEPT<br>KHQYDVDTETVIGFLRANGPKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI <del>DANR</del><br>GDAQNGWDTDQFPVD <del>AFDL</del> TQAMMQVLLNGGFNGGTNFD <del>AKLR</del> SSTD <del>PEDIFIA</del> HISA<br>MDAMAHALLNAAAII <del>LEES</del> PM <del>PGM</del> VKERYASFDNGLGK <del>KE</del> GKATLEELYDYAKKNGEPV<br>AASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5607MI4_005 | <i>Prevotella</i> | DNA              | 133        | ATGACTAAAGAGTATTTCCCTTCCGTCGGCAAGATTGCCTTTGAAGGACCCCGAAAGCAAG<br>AACCTATGGCCTTCCATTATTATGACGCCAATCGCGTGGTAATGGGAAAGCCGATGAAA<br>GAATGGCTTAAATTTGCCATGGCCTGGTGGCACACCTTGGCCAGGCCTCTGCAGACCCC<br>TTCGGCGGT <del>CAGAC</del> CCCGCTCCTACGAGTGGGACAAAGGCGAGTCCCCCTACTGCCCGCGCC<br>AAGGCCAAGGCCGATGCCGGCTTTGAAC <del>TGATGCAGAA</del> ACTGGGCATCGAGTATTTCTGC<br>TTCCACGATATAGACCTGGTGAAGACTCGGATGATATCGCCGAATACGAGGCCCGCATG<br>AAGGACATCAGGACTATCTCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG<br>GGAACCGCCAAACGTGTTCCGCCACAAAGCGTATATGAACGGCGCGCCACCAACCCCTCAG<br>TTCGACATCGTGGCCCCGTGCCGCTGTCCAGATCAAGAAACGCCCTGGATGCCACCATCAAG<br>CTGGCGGGCTCCAAC <del>TATGTGTTCTGGGCGCGCTGAGGGCTACTATA</del> CCCTCCTGAAAC<br>ACCCAGATGCAGAGAGAAAGGACCACTGGCCAAAGATGCTCACCGCCGCCCGGACTAT<br>GCCCCGTGCCAAGGGCTTCAAGGGCACCTTCTCATCGAACCCAAAGCCGATGGAGCCCCACC<br>AAGCACCA <del>GTACGAC</del> GTAGATACGGAGACCGTGATCGGCTTCTTCCGCGCCAAACGGCCTG<br>GACAAGGACTTCAAGGTGAATATTGAGGTGAACCAACGCCACCTTGGCCGGCCACACCTTC<br>GAGCACGAGCTCACCGTGGCCCCGGAGAACGGCTTCTTGGGCAGCATCGACGCCAACCGC<br>GGAGACGCCCCAGAACGGCTGGGATACGGACCAAGTTCCCCGGTGGATGCCTTCGACCTCACC |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|             |                   |                  |            | CAGGCTATGATGAGATCCTTCTGAACGGAGGCTTCGGCAACGGCGGTACCAACTTCGACGCCAAACTGCGCCGCTCTCCACGGACCCCGAGGACATCTTCATCGCCACATCAGCGCTATGGATGCCATGGCCACGCCCTGCTGAATGCAGCCGCATCCTGGAGGAAAGCCCGCTTCCGAAGATGCTGAAAGAGCGTTATGCCAGCTTTGACGGCGGTCTGGGCAAGAAGTTCGAA GAAGGCAAGGCCTCTCTGGAAGAACTCTACGAGTATGCCAAGAGCAACGGAGAGCCCGTG GCCGCTTCCGGCAAGCAGGAGCTCTGCGAAACGTACCTGAACCTCTACGCTAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5607MI4_005 | <i>Prevotella</i> | Amino Acid       | 134        | MTKEYFPSVGKIAFEGPESKNPMAFYHDANRVVMGKPMKEWLKFAMAWHHTLGQASADP FGGQTRSIEWDKGECPYCRAKAKADAGFELMQKLGIEYFCFHDIDLVEDCDDIAEYEARM KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAATNPQDIVARAAVQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMQRKDHAKMLTAARDYARAKGFKGTFLIEPKPMEPT KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSIDANR GDAQNGWDTDQFPVDAFDLTQAMQILLNGGFNGGTNFDAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAAILLEESPLPKMLKERYASFDGGLGKKFEEGKASLEELYEYAKSNGEPV AASGKQELCETYLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5607MI5_002 | <i>Prevotella</i> | DNA              | 135        | ATGGCTAAAGAATACTTCCCCTCCATCGGCAAAATCCCCTTTTGAAGGAGCCGACAGCAAA AATCCCCTCGCTTCCATTATTATGACGCCGACGCGTGGTTATGGGCAAGCCCATGAAG GAATGGCTTAAATTGCCCATGGCCTGGTGGCACACGCTGGGCCAGGCCTCCGGAGACCCC TTGGCGGCCAGACCCGCAGCTACGAATGGGACAAGGGCGAATGCCCTACTGCGCGGCC AAGGCCAAGGCCGACGCCGGTTTGAATCATGCAAAAAGCTGGGCATCGAATACTTCTGC TTCCACGATGTGGACCTTATCGAGGATTGCGATGACATGCCGAATACGAAGCCCGCATG AAGGACATCACGGACTACCTGCTGAAAAGATGAAGGAGACCGGCATCAAGAACCCTCTGG GCACCGCCAATGTCTTCGGCCACAAGCGTACATGAACGGCGCGGCACCAATCCGCAG TTCGATGTGGTGGCCCGTGC CGCGTCCAGATCAAGAAACCCCTGGACGCCACATCAAG CTGGCGGCTCCAACTATGTGTTCTGGGCGCGCGGAAGGCTATTACACCTCCTCTCAAC ACACAGATGCAGCGGGAAAAAGACCACCTGGCCAAAGTTGCTGACGGCGCGCCGACTAT GCGCGGCCAAAGGCTTCAAGGGCACCTTCTCATTGAGCCCAAAACCATGGAACCCACC AGCACCAAGTACGACGTGGATACGGAGACGCTCATCGGCTTCTCCGTGCCAACGGCCTG GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCTGGCCGGCCACACCTTC GAGCATGAGCTCACCGTGGCCCGCGAGAACGGTTTCTGGGCTCCATCGATGCCAAACCGC GCGACGCCCCAGAACGGCTGGGACACGGACCAAGTCCCCTGTGGACCCGTACGATCTTACC CAGGCCATGATGCAGGTGCTGCTGAACGGCGGCTTCGGCAACGGCGGCACCAACTTCGAC GCCAAACTCCGCCGCTCTCCACCGACCTGAGGACATCTTCATGCCCCATATTTCCGCC ATGGATGCCATGGCCACGCTTTGCTTAACGCAGCTGCCGTGCTGGAAGAGAGCCCCCTG TGCCAGATGGTCAAGGAGCGTTATGCCAGCTTCGACGATGGCCTCGGCAAAACAGTTTCGAG GAAGGCAAGGCTACCCCTGGAAGACCTGTACGAATACGCCAAGGCCAGGGTGAACCCGTT GTCGCCTCCGGCAAGCAGGAGCTTTACGAGACTCTCCTGAACCTGTATGCCGTCAAGTAA |
| 5607MI5_002 | <i>Prevotella</i> | Amino Acid       | 136        | MAKEYFPSIGKIPFEGADSKNPLAFHYDAGRVMGKPMKEWLKFAMAWHHTLGQASGDP FGGQTRSIEWDKGECPYCRAKAKADAGFELMQKLGIEYFCFHDVLDLIEDCDDIAEYEARM KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAGTNPQDVVARAAVQIKNALDATIK LGGSNYVFWGGREGYYTLLNTQMQRKDHIAKLTLTAARDYARAKGFKGTFLIEPKPMEPT KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSIDANR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|             |                   |                  |            | GDAQNGWDTDQFPVPYDLTQAMMQVLLNGGFGNGGTNFDAKLRRSSTDPE <sup>1</sup> FI <sup>2</sup> AHISA<br>MDAMAHALLNAAA <sup>3</sup> VL <sup>4</sup> EE <sup>5</sup> SPL <sup>6</sup> CQ <sup>7</sup> MV <sup>8</sup> KERYAS <sup>9</sup> FD <sup>10</sup> DGL <sup>11</sup> GK <sup>12</sup> Q <sup>13</sup> FE <sup>14</sup> E <sup>15</sup> GKAT <sup>16</sup> LED <sup>17</sup> LYE <sup>18</sup> YAKA <sup>19</sup> QGE <sup>20</sup> PV<br>VASGKQELYETLLNLYAVK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5607MI6_002 | <i>Prevotella</i> | DNA              | 137        | ATGACCAAAGAATA <sup>1</sup> TTCCCTACCGTCGGGAAGATCCCCTTCGAGGGCCCCGAAAGCAAG<br>AACCTATGGCGTTCCATTACTATGACCCCAACCGTCTGGTGATGGGCAAGAAGATGAAA<br>GACTGGCTGCGTTTCGCCATGGCCTGGTGGCACACCCCTCGGCCAGGGCTCGGGCGACCCAG<br>TTCGGCGGCCAGACCCGCAGTTATGCGTGGGACGAGGAGAAATGCCCGTACGAGCGCGCC<br>CGTGCCAAAGGCTGACGCCGGCTTCGAGATCATGCAGAAACTCGGTATCGAGTTCTTCTGC<br>TTCACGACATCGACCTGATCGAGGATACCGACGACATCGCCGAGTATGAGGCCCGCCTG<br>AAAGACATCACGGACTATCTGCTCGAGAAGATGAAAGCCACTGGCATCAAAAATCTCTGG<br>GGAACGGCCAAACGTGTTCCGCCACAAGCGTTGCATGAACGGCGCGCCACCAACCCCGGAC<br>TTCGCCGTGCTGGCCCCGGCTGCCGTCCAGATCAAGAACGCCATCGACGCCACCATCAAG<br>CTGGCGGCGAGAACTATGTGTTCTGGGTGGCCGGAAGGCTACACGAGCCTGCTCAAC<br>ACCCAGATGCAGCGTGAGAAAGAGCACCTGGCGCGCCTGCTGTCCCTGGCCCGGACTAT<br>GGCCGCGCCACGGCTTCAAGGTACCTTCTGATCGAGCCCAAGCCGATGGGACCCGACG<br>AAACACCAAGTACGACCAAGGATACGGA <sup>1</sup> AACTGT <sup>2</sup> CATCGGTTTCTGCGCCGCCACGGTCTA<br>GACAAGGACTTCAAGGTCAATATCGAGGTGAACCATGCCACGCTGGCGGGCCACACCTTC<br>GAACACGAACTGGCCTGGCCGTGGATCACGGTATGCTGGGCAGCATCGACGCCAACCGC<br>GGTGACGCACAGAA <sup>1</sup> CGGCTGGGATACCGACCA <sup>2</sup> AGTTCCCGATCGACAACTTCGAGCTGACC<br>CTTTCATGCTCCAGATCATCCGCAACGGTGGCCTGGACCCCGCGGCTCGAATTTTCGAT<br>GCCAAGCTGCGCCGCAACTCCACCGATCCCGAAGACATTTTCATCGGCACATCAGCGCC<br>ATGACGCCCATGGCCCCGGCATTGGTCAATCGGGCCGCCATCCTGGAGGAGCGCTATT<br>CCGAAGATGGTCAAGGAGCGTTACGCTTCGTTCCACAGCGGCA <sup>1</sup> AAAGGCAAGTAACGAG<br>GAAGGCAAGCTGACGCTCGAAGACATCGTGGCCTATGCCAAGGCGA <sup>1</sup> ACGGAGAACCCGAAG<br>CAGATTTCCGGCAACAGGA <sup>1</sup> ACTCTACGAGACGCTTGTGCGCACTCTATAGCAAAATAA |
| 5607MI6_002 | <i>Prevotella</i> | Amino Acid       | 138        | MTKEYFPTVGKIPFEGPESKNPMAFHYYDPNRLVMGKKMKDWLRFAMAWHHTLGQASGDQ<br>FGGQTRS <sup>1</sup> YAWDEGEC <sup>2</sup> PYERARAKADAGFEIMQKLGI <sup>3</sup> EFFCFHDI <sup>4</sup> DLIEDTDDIAEYE <sup>5</sup> EARL<br>KDITDYLLLEKMKATGIK <sup>1</sup> NLWGTANVFGHKRCMNGAATNPDFAVLARA <sup>2</sup> AVQIKNAIDATIK<br>LGGENYVFWGGREGYTSL <sup>1</sup> NTQM <sup>2</sup> REKEHLGRLLSLARDYGRAHGFKGTFLIEPKPMGPT<br>KHQYDQDTETVIGFLRRHGLDKDFKVNIEVNHATLAGHTFEHELACAVDHGMLGSI <sup>1</sup> DANR<br>GDAQNGWDTDQFPIDNFELTSLMLQIIRNGGLAPGGSNFDAKLRNSTDPEDIFIAHISA<br>MDAMARALVNAAAILEESAIPKMVKERYASFDSGKGKEYE <sup>1</sup> EGKLTLEDIVAYAKANGEPK<br>QISGKQELYETLVALYSK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 5607MI7_002 | <i>Prevotella</i> | DNA              | 139        | ATGACCAAAGGTA <sup>1</sup> TTCCCTACCATCGGCAGGATTCCCTTCGAGGGAACTGAAAGCAAG<br>AATCCCCCTCGCATTC <sup>1</sup> CATTACTATGAGCCCCGACCGGCTCGTACTGGGCAAGAAAATGAAA<br>GACTGGCTGCGTTTCGCGATGGCCTGGTGGCACACCCCTGGGCCAGGGCTCCGGCGACCCAG<br>TTCGGCGGCCAGACCCGCAGCTATGCTGGGACAAGGCCAGTGCCTCTATGAGCGCGCC<br>AAGGCCAAAGCCGACGCCGGCTTCGAGATCATGCAGAAACTCGGCATCGAGTTCTTCTGT<br>TCCACGACATTGACCTCGTTGAGGATACCGACGACATCGCCGAGTATGAGGCCCGGATG<br>AAGGACATTACCGACTATCTCCTGTTCAAGATGAAGGAGACCGGAATCAAGAACCTCTGG<br>GGTACGGCCCAATGTCTTCGGCCACAAAGCGCTATATGAACGGCGCGCCACCAATCCCAGAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|             |                   |                  |            | TTCGACGTTGGTGGCCCGCCGCTCCAGATCAAGAACGCCCTCGATGCCACCATCAAGCTGGGCGGTGAAAACATATGTGTTCTGGGGCGCCGCGAAGGCTATATGAGCCTGCTCAACACGAGATGCAGCGTGAGAAAGAGACACCTGGCGCGGATGCTGTCGCCCGCCGGACTACGCCGCGCCACGGCTTCAAGGTACCTTCTCATCGAGCCCAAACCGATGGAACCGACC AAGCACAGTACGACCAAGGATACGGAAACCGTGATCGGCTTCCTTCGCCGCCACAGGCTTGACAAGGATTTCAAGGTGAACATCGAAGTGAACCAACGCCACGCTGGCCGCCACACCTTCGAGCACGAACTGGCCACCGCGTCGACTCGGCGCTGCTGGGCAGCATCGAGCCCAATCGCGGACGCTCAGAACGGCTGGGATACCGACCACTTCCCGATCGACAACTTCGAACTCAGGCTGGCCATGCTGCAGATTATCCGCAACGGCGGTCTGGCACCCGGGCTCGAACTTCGACGCCAAACTGCGCCGTAACCTCCACCGATCCGGAAGATATCTTCATCGCCACATCAGTGGCATGGACGCGATGGCCCGTGGCTGGTCAACGCCCGCGCAATCTGGGAAGAGTCTCCCATCCGAGATGAAGAAAGAACGCTACGGCTGTTCCGACAGCGGCAAGGAATTCGAAAGGGCAAAGCTCTGCCTCGAAGACCTCGTGGCCTATGCCAAGGGCAACGGAGAACCGAAA CAGATCTCCGGCAGGCAGGAACATATATGAGACCATCGTCGCCCTTTATTGCAAAATAG                                                                                                                                                                                                                                                                                                                                         |
| 5607MI7_002 | <i>Prevotella</i> | Amino Acid       | 140        | MTKGYFPTIGRIPFEGTESKNPLAFHYEYEDRLVLGKKMKDWLRFAMAWWHHTLGQASGDQ FGGQTRSYAWDKAECPIYERAKAKADAGFEIMQKLGIEFFCFHDIIDLVEDTDDIAEYEARM KDITDYLLVKMETGIKNLWGTVANVFGHKRYMNGAATNPFDVVRAAVQIKNALDATIK LGGENYVFWGREGYMSLLNTQMQRKEHGLRMLVAARDYARAHGFKGTFLIEPKPMEPT KHQYDQDTETVIGFLRRHGLDKDFKVNIEVNHATLAGHTFEHELATAVDCCGLLGSIDANR GDAQNGWDTDQFPIDNFELTLAMLQIIRNGGLAPGGSNFEDAKLRNSTDPEDIFIAHISA MDAMARALVNAAAIIWEESPIPMKKERYASFDSGKGKEFEFGKLCLEDLVAYAKANGEPK QISGRQELYETIVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5608MI1_004 | <i>Prevotella</i> | DNA              | 141        | ATGACCAACGAGTATTTTCCCGGAATCGGTGTGATTCCGTTTGAAGGACAGGAAAGCAAG AATCCCATGGCTTCCATTATTATGACGCCAAACCGCGTAGTGATGGCAAAACCCATGAAG GAATGGTTCAAAATTGCCATGGCTGGTGGCATACGCTGGGCAGGCATCGGCCGATCCC TTCCGGCGGACAGACCCGCTCCTACGCATGGGACAAGGCGAGTGCCCTTACTGCCGTGCC CGCCAGAAAGGCCGACGCCGGCTTTGAACTGATGCAGAAAGCTGGGTATCGGCTATTTCTGC TTCCACGATGTGGATATCATCGAGGACTCGGAAGACATTGCCGAGTATGAGGCCCGTATG AAGGACATCACGGACTATCTGCTGGTGAAGATGAAGGAAACGGGCATCAAGAACCCTGTGG GGACGGCCAAACGCTCTCGGCCACAAGCGCTATATGAACGGCGCTGCCACCAACCCGCAG TTCGACGTGGTGGCCCGCGCTCGGTCCAGATCAAGAACGCCCTGGACGCCACCATCAAG CTGGCGGCAGCAATTACGTGTTCTGGGGCGCGCGAAGGCTATTATACCTTTTGGAAC ACGCAGATCGGGCGGAGAAAGGACCACCTGGCCCAGATGCTCAAGGCAGCCCGTGACTAT GCCCGCGCAAGGGATTCAAGGGCACGTTCCTCATTGAGCCCAAGCCCATGGAGCCCACC AAGCACCAGTACGACGTAGATACGGAGACCGTGATTGGCTTCCTGCGCGCAACGGACTG GACAAGGACTTCAAGGTGAATATCGAAGTGAACCAACGCCACCTGGCCGGCCACACCTTC GAGCACGAACTCACCGTGGCCCGGAAACGGCTTCTTGGGCAGCATCGACGCCAACCGG GGAGACGCCCAGAACGGTTGGGATACAGACCAGTTCCCCATAGATGCCTTTGACCTCACC CAGGCCATGATGCAGGTCTCTGCTCAACGGCGGATTGGCAACGGCGGACCAACTTCGAC GCCAAACTGCGCCGTTCTCCACGGATCCCGAGGACATCTTCATCGCCCATCGGCGCC ATGGACGCCATGGCCACGCCCTCCTGAACGCCCGCCGCATCCTGGAAGAGAGCCCCATG |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|             |                   |                  |            | CCGGCATGGTGAAGGAGCGCTACGCTTCCTTCGACAAATGGCCTTGGCAAGAAGTTCGAG<br>GAAGGAAAGGCCACGCTGGAAGAGCTGTACGACTATGCCAAGAAACGGCGAGCCTGTG<br>GCCGCTTCCGGCAAGCAGGAACTGTACGAAACGCTGCTGAACCTGTACGCCCAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5608MI1_004 | <i>Prevotella</i> | Amino Acid       | 142        | MTNEYFPGIGVIPFEGQESKNPMAFHHYYDANRVVMGKPMKEWKFAMAWWHTLQASADP<br>FGGQTRSYAWDKGECPYCRARQKADAGFELMQKLGIYFCFHDVDIIEDCEDIAEYEFARM<br>KDITDYLLVKMKETGIKNLWGTANVFGHKRYMNGAATNPQFDVVARAAVQIKNALDATIK<br>LGGSNYVFWGGREGYYTTLWNTQMRREKDHLAQMLKAARDYARGKFGKTFLEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANR<br>GDAQNGWDTDQFPIDAFDLTQAMMQVLLNGGFGNGGTNFDAKLRRSSTDPEDI FIAHIGA<br>MDAMAHALLNAAAAILLEESPMPGMVKERYASFNDNGLGKKFEEGKATLEELYDYAKKNGEPV<br>AASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5608MI2_002 | <i>Prevotella</i> | DNA              | 143        | ATGAAAGAAATACTTCCCTACCATCGGAAAAAATCCCTTTCGAGGGCCCTCAGAGCAAGAAT<br>CCGCTCGCATTCCACTTACTATGACGCCAACCGCGTTGTGCGCGGCAACCCATGAAGGAC<br>TGGCTCAAAGTTCGCCATGGCTTGGTGGCACACCCCTGGCGCAGCATCGGCAGACCCCTTC<br>GGCGCCAGACCCGCAGCTACGAGTGGGACAAAGCCGAGTGCCCTTACTGCCGTGCCCGT<br>GAAAAGCCGACGCCGGCTTCGAGATCATGCAGAAACTTGGAATCGAGTACTTCTGCTTC<br>CATGACATCGACCTTGTGGAAGACTGGAGGACATTGCCGAGTACGAGGCCCGCATGAAG<br>GACATCAGGACTACCTCCTGGAGAAAGATGAAGGCCACCGGCATCAAGAAACCTGTGGGGC<br>ACCGCCAAACGTCTTTGGCAACAAGCGCTACATGAACGGCGCAGCCACCAACCCCTCAGTTC<br>GACATCGTTGCCCGTGAGCTGTCCAGATCAAGAACGCCATCGACGCAACAATCAAGCTG<br>GGCGGTACCGGTTACGTATTCTGGGGCGCGCGAGGGCTACTACACCCCTCCTGAACACC<br>CAGATGCAGCGCGAGAAGGACCACCTTGCCAAAGATGTCACCGCAGCCCGCGACTACGCC<br>CGGCCAAAGGATTCAAGGGCACATTCTCTCATCGAGCCCAAGCCCATGGAGCCCAACAAG<br>CACAGTACGATGTTGACACGGAAACCGTCAATCGGCTTCCTCCGCGCAACGGCCTGGAC<br>AAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCCCTGGCGGCCACACCTTCGAG<br>CACGAGCTCACCGTGGCCGTGGACAAACGGCTTCCTGGGCAGCATCGACGCAAAACCGCGGC<br>GACGCCAGAACGGCTGGGACACTGACCAGTTCCCTGTGGATCCTTACGACCTCACCCAG<br>GCAATGATGCAGATTATCCGCAACGGCGGCTTCAAGGACGGCGGCACCAACTTCGACGCC<br>AAACTCCGCCGAGCTCCACGGACCCCGAGGACATCTTCATCGCCCAACATCAGCGCAATG<br>GATGCAATGGCACACGCCCTCATCAACGCTGCTGCAGTGTCTGAGGAAAGCCCTCTGTGC<br>GAGATGGTTGCAAAGCGCTACGCCAGCTTTGACAGCGGTCTTGGCAAGAAAGTTCGAGGAA<br>GGCAAAGCCACTCTCGAGGAGATCTACGAGTATGCCAAGAGGCCCGCCACCCGTCGCC<br>GCCTCCGGCAAGCAGGAGCTCTACGAGACACTGCTCAATCTGTACGCTAAATAA |
| 5608MI2_002 | <i>Prevotella</i> | Amino Acid       | 144        | MKEYFPTIGKIPFEGPQSKNPLAFHYDANRVVAGKPMKDWLKFAMAWWHTLGAASADPF<br>GGQTRSYEWDKAECPYCRAREKADAGFEIMQKLGIEYFCFHDIDLVEDCEDIAEYEFARMK<br>DITDYLLLEKMKATGIKNLWGTANVFGNKRYMNGAATNPQFDIVARAAVQIKNALDATIKL<br>GGTGYVFWGGREGYYTTLNTQMQRREKDHAKMLTAARDYARAKGFKGTFLIEPKPMEPTK<br>HQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVAVDNGFLGSI DANRG<br>DAQNGWDTDQFPVDPYDLTQAMMQIIRNGGFKDGGTNFDAKLRRSSTDPEDI FIAHISAM<br>DAMAHALLNAAAIVLEESPLCEMVAKRYASFDSGLGKKFEEGKATLEEIYEVAKKAPAPVA<br>ASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 5608MI3_004 | <i>Prevotella</i> | DNA              | 145        | ATGACCAAAGAGTATTTCCTACAAATCGGAAAAGATTCCCTTCGAAGGCCCGGAGAGCAAG<br>AATCCGCTGGCATTCATTACTATGAACCCGACAGAAATCATCCTCGGACGGAAGATGAAG<br>GACTGGCTGGCTTCGCCGTGGCTGGTGGCACACCTCGGCCAGGCGTCCGGCGACCAG<br>TTCGAGGCCAGACCCGCAACTATGCTGGGACGAGCCCGAATGCCCGGTAGAGCGCGG<br>AAAGCCAAAGGCCGACGCCGGCTTCGAGCTGATGCAGAAAGCTGGGCATCGAGTATTCTGC<br>TTCACGACGTAGACCTCATAGAGGAGGCCCGCAACCATCGAAGAAATATGAGGAGCGCATG<br>GGCATCAATAACCGACTACCTGCTCGGGAAGATGAAGGAGACAGGTATCAAGAACCTCTGG<br>GGACCGCCAAACGTGTCGGCCACAAAGCGTTACATGAACGGAGCCGCCAACCAACCCCGAC<br>TTCGACGTGGTGGCCCGTGGCCCGTGCAGATCAAGAAAGCCCATCGACGCCACCATCAAG<br>CTGGCGGCGAGAAATTACGTATTCTGGGCGGACGCGAGGGCTATGCAAGCCTGCTCAAC<br>ACTCAGATGCAGCGCGAGAAAGACCACTGGGACGCATGCTGGCTGCAGCCCGGACTAT<br>GGCCGCCCCACGGATTCAAGGGCACTTTCCTCATCGAGCCCAACCCATGGAGCCTACC<br>AAGCACAGTACGACCAAGGATACCGAGACCGTTATCGCCTTCCTGGCAGGAACGGCCTC<br>GACAAGGATTTCAAGGTAAACATCGAGGTGAACACGCCACCTGGCGGGCCACACCTTC<br>GAGCACGAACTGGCGTGGCAGTGGACAAACGGCCTGCTTGGCAGCATCGACGCCAAACCGC<br>GGCAGCGCGCAGAACGGATGGGACACCGACCAAGTTCCTCCATCGACCAACTTCGAGCTCACC<br>CAGGCCATGCTGCAGATAATCCGCAACGGCGGACTGGGAACCGGCGGATCGAACTTCGAC<br>GCCAAGCTGGCCCGCAATTCACCGACCTGAGGATATCTTCATCGCCACATCAGTGCG<br>ATGACGCCATGGACGGCGCTGGCAACGCCCGCAATCATCGAAGAGAGCCCCATC<br>CCCGCAATGCTGAAGGAGCGCTACGCATCGTTCGACAGCGGCAAGGGCAAGGAGTTCGAG<br>GACGGCAAACTGAGCCTCGAAGAACTGGTAGCCTACGCCAAGCGGAACGGCGAGCCGAAG<br>CAGATTTCCGGCAAGCAGGAACTCTACGAAACCATAGTGGCCCTCTATTGCAAGTAA |
| 5608MI3_004 | <i>Prevotella</i> | Amino Acid       | 146        | MTKEYFPTIGKIPEGPESKNPLAFHYEPDRRIILGRMKDWLRFVAVWHTLQASGDQ<br>FGQTRNVAWDEPECPVERAKAKADAGFELMQKLGIEYFCFHDVDLIEEAAATIEEYEERM<br>GIITDYLLGKMKETGIKNLWGTANVFGHKRYMNGAATNPDFDVVARAAVQIKNAIDATIK<br>LGGENYVFWGGREGYASLLNTQMQRKDHIGRMLAAARDYGRAHFGKGTFLIEPKPMEPT<br>KHQYDQDTETVIAFLRRNLDDKFKVNIENVNHATLAGHTFEHELAVAVDNLGLLSIDANR<br>GDAQNGWDTDQFPIDNFELTQAMLQIIRNGGLGTGGSNFDKLRNSTDPEDIFIAHISA<br>MDAMARALANAAAIIIESPIPAMLKERYASFDSGKGKEFEFDGKLSLEELVAYAKANGEPK<br>QISGKQELYETIVALYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5609MI1_005 | <i>Prevotella</i> | DNA              | 147        | ATGGCACAAAGAATACTTCCCTACCATTTGGGAAAAATCCCCTTCGAGGGCACTGAGAGCAAG<br>AATCCCCTTGCTTTCATTACTATGAGCCGGAGCGCATTTGTCTGGGCAAAACCCATGAAA<br>GAATGGCTCAAGTTTGCCATGGCCTGGTGGCACACGCTGGGCAGGCATCGGCCGATCCC<br>TTCGGCGGCCAAACCCGCAGCTATGCCTGGGATAAGGGCGAATGCCCTACTGCCGTGCC<br>CGGCCAAAGGCGGACGCCGGCTTCGAGATTATGCAAAAAGCTGGGCATCGAGTACTTCTGC<br>TTCCACGATATCGACCTGGTAGAAGACTGTGACGATATTGCGGAATACGAAGCCCGCATG<br>AAGGACATCAGGACTACCTCCTGGAGAAGATGAAGGAAACCGGTATCAAGAACCTCTGG<br>GGCACCGCCAATGTGTTTGGTCACAAAGCGTACATGAACGGCGCGCCACCAACCCCGCAG<br>TTTGACGTAGTGGCCCGTGCCGCTGTTCAGATTAAAGAACGCCATTGACGCCACCATCAAG<br>TTGGGCGGTGCCAATTACGTGTTCTGGGGCGCGCGAGGGCTATTACAGCCTCCTGTAAC<br>ACCCAGATGCAGCGGGGAGAAAGGACCACCTGGCCCAAGCTGCTCACGGCAGCCCCGACTAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------|-------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|             |                   |                  |            | GCCCGGCCAACGGCTTCAAGGGAACCTTCCTGATTGAGCCCAAGCCCATGGAGCCCACC<br>AAGCACAGTACGACGTGGATACGGAGACGGTCATTGGCTTCCTCCGCGCAACGGCCTG<br>GACAAGGACTTCAAGGTGAATATCGAGGTGAACCAAGCCACGTTGGCCGCCACACCTTT<br>GAGCACGAGCTCACCGTGGCCCGCGAGAACGGCTTCCTGGGCAGCATCGACGCCAACCGC<br>GGCGATGCCCAGAACGGCTGGGATACGGACCAAGTCCCGGTAGACGCTTATGAGCTCACC<br>CAGGCCATGATGCAAGTGCTCCTGAACGGAGGCTTCGGCAACGGCGGCACCAACTTCGAC<br>GCCAAGCTGGCCCGCTCCTCCACGGACCCGGAGGACATCTTCATCGCCCATATCAGTGCG<br>ATGGATGCCATGGCCACGCCCTGCTCAACGCCCGCCGCTGCTGGAGGAAAGCCCCCTG<br>TGCCAGATGGTGAAAGGAGCGCTACGCCAGCTTTGACAGCGGTCGCGGCAAGCAGTTCGAG<br>GAAGGAAAGGCCACCCCTGGAGGACCTGTACAACTACGCCAAAGCCACCGGTGAACCCCGTG<br>GTTGCCTCCGGCAAGCAGGAACCTTTACGAGACCCCTCCTGAACCTCTATGCAAAAGTAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 5609MI1_005 | <i>Prevotella</i> | Amino Acid       | 148        | MAQEYFPTIGKIPFEGTESKNPLAFHYYPERIVCGKPMKEWLKFAMAWHHTLGQASADP<br>FGGQTRSVAWDKGECPYCRARAKADAGFEIMQKLGIEYFCFHDIDLVEDCDDIAEYEARM<br>KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAATNPQFDVVARAAVQIKNAIDATIK<br>LGGANVVFWGGREGYYSLNLTQMQRKDLAKLLTAARDYARANGFKGTFLLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSI DANR<br>GDAQNGWDTDQFPVDAYELTQAMQVLLNGGFGNGGTNFDAKLRSSSTDPEIDIFIAHISA<br>MDAMAHALLNAAA VLEESPLCQMVKERYASFDSGPGKQFEEGKATLEDLYNYAKATGEPV<br>VASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5610MI1_003 | <i>Prevotella</i> | DNA              | 149        | ATGGCACAAAGAATACTTCCTACCATTGGGAAAATCCCTTCGAGGGCACTGAGAGCAAG<br>AATCCCTTGCTTCCATTACTATGAGCCGGAGCGCATTGTCTGCGCAAAACCCATGAAA<br>GAATGGCTCAAGTTTGCCATGGCCTGGTGGCACACGCTGGGGCAGGCATCGGCCGATCCC<br>TTGGCGGCCAAAACCGCAGCTATGCCTGGGATAAGGGCGAATGCCCTACTGCCGTGCC<br>CGTGCCAAAGCGGACGCCGCTTTTGAGATTATGCAAAAAGCTGGGCATCGAGTACTTCTGC<br>TTCACGATATCGACCTGTTAGAAGACTGTGACGATATTCGGAATACGAAGCCCGCATG<br>AAGGACATCAGGACTACCTCCTGGAGAAGATGAAGAAACCGGCATCAAGAACCTCTGG<br>GGCACCGCCAATGTGTTGTCACAAAGCGCTACATGAACGGCGCGGCACCAATCCGCAG<br>TTTGACGTGGTGGCCCGTGTGCCGTGCAAAATCAAGAAACCCATTGACGCCACCATCAAG<br>TTGGGCGGTGCCAATTACGTGTTCTGGGCGCGCGGAGGGCTATTACAGCCTCCTGAAC<br>ACCCAGATGCAGCGGGAGAAAGGACCACCTGGCCAAGCTGCTCACGGCAGCCCGGACTAT<br>GCCC GCGCCAACGGCTTCAAGGGAACTTCCTGATTGAGCCCAAGCCCATGGAGCCCACC<br>AAGCACCAGTACGACGTGGATACGGAGACGGTCATTGGCTTCCTCCGCGCCAACGGCCTG<br>GACAAGGACTTCAAGGTGAATATCGAGGTGAACCAACGCCACGCTGGCCGGCCACACCTTT<br>GAGCACGAACTCACCGTGGCCCGCGAGAACGGCTTCCTGGGCAGCATCGACGCCAAACCGC<br>GGCGATGCCCAGAACGGCTGGGATACGGACCAAGTCCCGGTAGACGCTTATGAGCTCACC<br>CAGGCCATGATGCAAGTGCTCCTGAACGGAGGCTTCGGCAACGGCGGCACCAACTTCGAC<br>GCCAAGCTGGCCCGCTCCTCCACGGACCTGGAGGACATCTTCATCGCCCATATCAGTGCG<br>ATGGATGCCATGGCCACGCCCTGCTCAACGCCCGCCGCTGCTGGAGGAAAGCCCCCTG<br>TGCCAGATGGTGAAGGAGCGCTACGCCAGCTTTGACAGCGGTCGCGGCAAGCAGTTTCGAG<br>GAAGGAAAGGCCACCCCTGGAGGACCTGTACAACTACGCCAAAGCCACCGGTGAACCCCGTG<br>GTTGCCTCCGGCAAGCAGGAACCTTTACGAGACCCCTCCTGAACCTCTATGCAAAAGTAG |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5610MI1_003 | <i>Prevotella</i> | Amino Acid       | 150        | MAQEYFPTIGKIPFEGTESKNPLAFHYYPEPERIVCGKPKMEWLKFAMAWWHHTLGQASADP<br>FGGQTRSYAWDKGECPCYCRARAKADAGFEIMQKLGIEYFCFHDIDLVEDCDDIAEYEARM<br>KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAGTNPQFDVVARAAVQIKNALDATIK<br>LGGANVYFWGGREGYYSLNLTQMQRKDLAKLLTAARDYARANGFKGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSIDANR<br>GDAQNGWDTDQFFPVDAYELTQAMMQVLLNGGFGNGGTNFDAKLRSSSTDLEDIFIAHISA<br>MDAMAHALLNAAAACLEESPLCQMVKERYASFDSGPGKQFEEGKATLEDLYNYAKANGEPV<br>VASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5610MI2_004 | <i>Prevotella</i> | DNA              | 151        | ATGGCAAAAGAATAATTTCCCTACCATCGGCAAAGATTCTTTTGAAGGAACCGACAGCAAG<br>AGTCCCTCGCCTTCCATTACTATGACGCCCAAGCGCGTTGTGATGGGCAAAACCCCATGAAG<br>GAATGGCTCAAGTTCGCCATGGCTGGTGGCACACCCCTGGGCCAGGCATCGGCCGACCCC<br>TTCGGCGGTCAGACCCGCCACTATGCCTGGGATGAAGCGGAATGCCCCTACTGCCCGGCC<br>AAAGCCAAAGGCCGACGCCGGCTTCGAGATCATGCAGAAACTGGGCATCGAGTACTTCTGC<br>TTCACGATGTGGACCTGGTGAAGACTCGACGACATCGCCGAGTACGAAGCCCGCATG<br>AAGGACATCACGGACTACCTGCTGGAGAAGATGAAGGAAACCCGGCATCAAGAACCTCTGG<br>GGCAGGCCCAATGTGTTCCGCCACAAAGCGTTACATGAACGGCGCGCGGACCAACCCCGCAG<br>TTTGACATTGTGGCCCGCGCTGCCGTCCAGATCAAAAACGCCCTGGACGCCACCATCAAG<br>CTGGCGGTTCCAACTACGTGTTCTGGGGCAGCCGCGAAGGCTACTACACCTCCTGAAAC<br>ACCCAGATGCAGCGGGAGAAAGACCACCTGGCCAAGCTCCTGACCGCGCCCGCGACTAC<br>GCCCCGCCAAAGGCTTCAAGGGAACCTTCTCATCGAGCCCCAAACCCATGGAGCCCCACC<br>AAGCACCAAGTACGACGTGGACACCGAGACCGTAATCGGCTTCTGCGTGCCCAACGGCCTG<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCCCTGGTGGCCACACCTTC<br>GAGACGAACTCACCGTCGCCCGTGAAAACGGCTTCTTCGGATCGATCGACGCCAAACCGC<br>GGCAGCCCCAGAACGGCTGGGACACCGACCAAGTTCCCCGTAGACGCCTATGACCTCACC<br>CAGGCCATGATGCAAGTGCTGCTGAACGGCGGTTTCGGCAATGGCGGTACCAACTTCGAC<br>GCCAAGCTCCGCCGCTCCTCCACGGATCCGGAAGACATCTTCATCGCCACATCAGCGCC<br>ATGGACGCCATGGCCACGCCCTGTGAACGCCCGCGCGCTGTGGAAGAAAGCCCGCTT<br>CCGCCATGGCGAAAGAGCGCTACGCTCCTTTGACAGCGGACTTGGCAAGAAGTTCGAA<br>GAGGAAAAGGCCACCTCGAAGAGCTGTACGACTATGCCAAGGCTAACGACGCCCCCTGTC<br>GCCGCTCCGGCAAGCAGGAACCTTACGAAAACCTTCTTGAACCTCTATGCAAAATAG |
| 5610MI2_004 | <i>Prevotella</i> | Amino Acid       | 152        | MAQEYFPTIGKIPFEGTDSKSPAFHYIDAQRVVMGKPKMEWLKFAMAWWHHTLGQASADP<br>FGGQTRHYAWDEGECPCYCRAKAKADAGFEIMQKLGIEYFCFHDVDLVEDCDDIAEYEARM<br>KDITDYLLLEKMKETGIKNLWGTANVFGHKRYMNGAGTNPQDIVARAAVQIKNALDATIK<br>LGGSNYVFWGSREGYYTLLNTQMQRKDLAKLLTAARDYARAKGFKGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSIDANR<br>GDAQNGWDTDQFFPVDAYDLTQAMMQVLLNGGFGNGGTNFDAKLRSSSTDLEDIFIAHISA<br>MDAMAHALLNAAAACLEESPLPAMAKERYASFDSGLGKKFEEGKATLEELYDYAKANDAPV<br>AASGKQELYETFLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 5751MI1_003 | <i>Prevotella</i> | DNA              | 153        | ATGGCAAAACAGATAATTTCCGCAAAATCGGAAAAGATTAAATTCGAAGGAACAGAGCAAG<br>AATCCGCTTGCGTTCCATTATTATGACGCCAAACAGGGTAGTCCTCGGAAAGGCAATGGAG<br>GAGTGGCTCAAGTTCGCAATGGCTTGGTGGCATACTCTCGGACAGGCTTCCGGAGACCAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|             |                   |                  |            | TTCGGGGCCAGACCCGCAGCTACGAGTGGGATCTTGAGCCACCCCCGAGCAGCGCGCA<br>AAGGACAAAGCTCGACGCCGGCTTCGAAATAATGGAGAAACTTGGAATCAAGTATTTCTGT<br>TTCCACGATGTTGACCTTATCGAAGACAGCGACGATATTCGACATATGAGGCTCGTCTC<br>AAGGACCTTACAGACTACGCTGCAGAGCAGATGAAGCTCCACGACATCAAGCTCCTCTGG<br>GGTACAGCGAATGTATTCGGCAACAAGCGCTACATGAACGGTGGGTACAAACCCCTGAT<br>TTCGATGTAGTTGCCCGCGCAGCCGTTCAGATTAAGAACGCTATCGACCGACCATCAAG<br>CTCGGTGGTACCAGCTATGTATTCGGGGCGGTCTGAGGGATATCAGAGCCTGCTCAAC<br>ACTCAGATGCAGCGTGAGAAAGGACCCTCGCAACCATGCTTACAATCGCTCGCGACTAT<br>GCTCGCAGCAAGGGCTTTACCGGAACTTCCTTATCGAGCCTAAGCCGATGGAGCCTACA<br>AAACACCAGTACGACGTAGATACAGAGACTGTTGTGGCTTCCTCAAGGCACACGGCCTG<br>GACAAGGACTTCAAGGTAATAATCGAGGTTAACCACGCAACTCTCGCAGGCCACACCTTC<br>GAGCACGAACTCACCGTTGCTGTGGATAACGGAATGCTCGGTTCTATCGACGCTAACCGC<br>GGTGATGCACAGAACGGCTGGGATACAGACCAGTTCCTGTAGCGCTGAGGAGCTTACC<br>CTCGCTATGATGCAGATTATCCGTAATGGTGGCCTTGGCAACGGAGGATCCAACCTCGAC<br>GCAAAGCTTCGCCCGCAACTCTACCGATCCTGAAGACATCTTCATCGCACACATCTGCGGT<br>ATGGATGCAATGGCACACGCTCTCCTCAATGCAGCTGCAATTATCGAGGAGTCTCCTATC<br>CCTACAAATGGTTAAGGAGCGTTACGCTTCCTTCGACAGCGGTATGGTAAGGACTTCGAG<br>GATGAAAAGCTTACCCCTCGAGGATCTCTACAGCTACGGCGTGAAGAACGGAGAGCCAAAG<br>CAGACCAGCGAAAGCAGGAGCTCTATGAGACTCTCATGAATATCTATTGCAAGTAA |
| 5751MI1_003 | <i>Prevotella</i> | Amino Acid       | 154        | MAKQYFPQIGKIKFEGTESKNPLAFHYIDANRVVLGKAMEEWLKFAMAWWHHTLGQASGDQ<br>FGGQTRSYEWDIAATPEQRAKDKLDAGFEIMEKLGIKYFCFHDVDLIEDSDDIATYEARL<br>KDLTDYAAEQMKLHDIKLLWGTANVFGNKR YMNGAATNPDEFVVARAAVQIKNAIDATIK<br>LGTSYVFWGGREGYQSLNLTQMÖREKDH LATMLTIARDYARSKGFTGTFLIEPKPMEPT<br>KHQYDVDTETVVGF LKAHGLDKDFKVNIEVNHHATLAGHTFEHELTVAVDNMGLGSI DANR<br>GDAQNGWDTDQFPVSAEELTLAMMQIIIRNGGLNGGNSF DAKLRNSTD PEDIFIAHICG<br>MDAMAHALLNAAAII EESPIPTMVKERYASFDSGMGKDFEDGKLTLEDLYSYGVKNGEPK<br>QTSAKQELLYETLMNIYCK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5751MI2_003 | <i>Prevotella</i> | DNA              | 155        | ATGGCAAAAGAAATTTTTCACAAAGTAGGCAAGATTC CATTTGAGGGTCCCTGAAAAGTACT<br>AACGTACTCGCATTC CACTACTATGATCCAGAACGCGAAGTTC TTGGTAAGAAAAATGAAA<br>GATTGGCTGAAGTATGCTATGGCTTGGTGGCACACACTCGGT CAGGCAAGTGCGGACCAA<br>TTCGGTGGTCAAACTCGTTCGTATGAATGGGATGAAGCCGACGATGTTCTTCAACGCGCA<br>AAGGATAAAATGGA TGCTGGTTTGAATTGATGACCAAAC TTGGCATTGAATACTACTGC<br>TTCCATGATGTCGACCTTATTGAAGAAGGTGC AACAAATTGAAGAAATATGAAGCTCGTATG<br>CAAGCTATCACCCGACTACGCATTAGAAAAACAAAAAGAAACCCGGCATTAAGCTCCTTTGG<br>GGTACTGCTAATGTGTTTGGTCATAAGCGTTATATGAATGGTGGGCAACAAACCCCTGAC<br>TTTGATGTAGTGGCTCGCGCTGCTGTACAAAATCAAGAACGCTATCGATGCAACTATCAAG<br>CTTGGTGGTCAAAACTATGTATTC TGGGGTGGCCGCGAAGGTTATATGAGTTTGCTCAAC<br>ACTCAAAATGCAACGCGAAAAAGACCACCTTGGCAAAAGATGCTTACCGCAGCTCGCGACTAT<br>GCTCGTGCTAAGGGCTTCAAGGGTACATTCCTCGTTGAACCTAAGCCTATGGAACCAACT<br>AAGCATCAATATGATACCGATACAGAAACTGTGATTGGTTTCCTCCGTGCAAAATGGTCTT<br>GAAAAAGACTTCAAGGTGAACATTGAAGTGAACCATGCTACTCTCGCTCAGCACACTTTC                                                                                                                                                                                                                                                                                                 |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-------------|-------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|             |                   |                  |            | GAACACGAACTCGTGTGGCTGTCGACAATGGCATGTCGGTTCTATCGACGCTAAACCGT<br>GGCGATGCTCAAAATGGCTGGGATACCGACCAAT'TCCCAATCGACAAC'TACGAACTCACC<br>CTCGCTATGCTCCAAATCATTCGCAATGGTGGTCTTGGCAATGGCGGTAGCAACCTCGAC<br>GCTAAGATTGTCGTAATAGCACCGACCTTGAAGACCTCTTTATCGCTCACATCAGTGGT<br>ATGGATGCTATGGCTCGTGCACTTCTCAATGCTGCTGCAATCGTTGAAAAGAGCGAAATT<br>CCTGCTATGTTGAAGCAGCGTTATGCAAGCTCTGATGAGGTATGGTAAGGACTTCGAA<br>GAAGGAAAAC'TCACTCTCGAACAAC'TCGTAGACTATGCTAAGGCTAACGGCGAACCTGCT<br>ACAGTAAGCGGCAAGCAAGAAAAGTATGAAA'CTCTCGTTGCTCTCTACGCTAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5751MI2_003 | <i>Prevotella</i> | Amino Acid       | 156        | MAKEFFPQVGKIPFEGPESTNVLAFHYYDPEREVLGKKMKDWLKYAMAWWHHTLGQASGDQ<br>FGGQTRSYEWDEADDVLQRAKDKMDAGFELMTKLGIEYYCFHDVDLIEEGATIEEYEARM<br>QAITDYALEKQKETGIKLLWGTANVFGHKRYMNGAATNPDFDVVARAAVQIKNAIDATIK<br>LGGQNYVFWGGREGYMSLLNTQMOREKDHAKMLTAARDIYARAKGFKGTFLVEPKPMEPT<br>KHQYD'TD'TETVIGFLRANGLEKDFKVNIEVNHATLAQHTFEHELAVAVDNGMLGSI DANR<br>GDAQNGWDTDQFPIDNYELTLAMLQIIRNGGLNGGNSLDAKIRRNSTDLEDLFIAHISG<br>MDAMARALLNAAAIVEKSEIPAMLIKQRYASSDAGMGKDFEEGKLTLEQLVDYAKANGEPA<br>TVSGKQEKYETLVALYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 5752MI1_003 | <i>Prevotella</i> | DNA              | 157        | ATGACTAAAGAGTATTTCCCGGGAATCGGAAAAGATTCCGTTTGAAGGAACCAAGAGCAAG<br>AACCCCTTGGCCTTCCATTATTATAACGCCTCCAGGTAGCGATGGGCAAGCCCATGAAG<br>GACTGGCTCAAGTATGCCATGGCCTGGTGGCACACCCCTGGGCCAGGCCTCTGCAGACCCCC<br>TTTGGCGGCCAGACCCGCTCCTACGAATGGACAAGGGCGAGTGCCTTATTGCCGCGCC<br>AAGCAGAAAGGCCGATGCCGGCTTTGAGCTCATGCAGAAAGCTGGGCATCGAGTACTACTGC<br>TTCCACGACGTGGACATCATCGAGGACTGCGAGGACAT'TGCCGAGTACGAGGCCCGCATG<br>AAGACATCACGGA'CTACCTGCTGGAGAGCAGAAAGACCGGCATCAA'GAACCTCTGG<br>GGCACCGCCAACG'TGTTGGCCACAAAGCGTACATGAACGGCGCGCCACCAACCCCTCAG<br>TTTGACATTGTGGCCCGTCCGCGCTCCAGATCAAGAAACGCCCTGGATGCCACCATCAAG<br>CTGGTGGTACCAACTACGTGTTCTGGGTGGCCGCGAAGGCTACTACACGCTGCTCAAC<br>ACCCAGATGCAGCGGGAGAAACCACCTGGCCAAAGATGCTCACCGCCGCCCGGACTAC<br>GCCGCGCCAAGGGCTTCAAAGGCACCTTCTCATTTAGCCCAAACCCATGGAGCCACCC<br>AAGCACCAGTACGACGTGGACACCGAGACCGTGATTGGTTTCATCCGCGCCAACGGCCTG<br>GACAAGGACTTCAAGGTAACAT'TGAGGTAAACCCACGCCACCCCTGGCCGGCCACACCTTT<br>GAGCACGAGCTCACCGTGGCCCGCGAGAACGGCTTCTTGGGTCCATCGACGCCAACCCGC<br>GGAGATGCCCAGAACGGCTGGGATACGGACCAAGTTCCCCATCGACGCCCTGGATCTCACC<br>CAGGCTATGATGCAAGT'CATCCTCAACGGTGGCTTCGGCAATGGCGGCACCAACTTTGAC<br>GCCAAGCTCCGCGCGCTCCTCCACCGATCCCGAGGACATCTTCATCGCCCAATCAGCGCC<br>ATGGATGCCATGGCACACGCCCTCCTGAACGCAGCGCCCATCCTGGAAGAGAGCCCCCTG<br>CCCCCATGGTCAAGGAGCGTTACGTTCTCTCGACAGCGGTCTGGGCAAGAAGTTCGAA<br>GAAGGCAAGGCCTCCCTGGAAGAACTTTACGAATATGCCAAGAAATGGAGAGCCCCGTG<br>GCCGCTTCCGGCAACAGGAGCTCTGCGAAA'CTTACTTGAACCTCTATGCAAAAGTAG |
| 5752MI1_003 | <i>Prevotella</i> | Amino Acid       | 158        | MTKEYFPGIGKIPFEGTKSKNPLAFHYYNASQVAMGKPMKDWLKYAMAWWHHTLGQASADP<br>FGGQTRSYEWDKGECPYCRAKQKADAGFELMQKLGIEYYCFHDVDI IEDCEDIAEYEARM<br>KDITDYILLEKQKETGIKNLWGTANVFGHKRYMNGAATNPQDIVARA'AVQIKNALD'ATIK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|             |                   |                  |            | <u>LGGTNYVFWGGREGYYTLLNTQM</u> <u>QREKNHLAKMLTAARDYARAKGFKGTFLIEPKPMEPT</u><br><u>KHQYDVD</u> <u>TETVIGFIRANGLDKDFKVNIEVNHATLAGHTFEHELTVARENGFLGSIDANR</u><br><u>GDAQNGWDTDQFPIDALDLTQAMMQVILNGGFGNGGTNFD</u> <u>AKLRRSSTD</u> <u>PEDIFIAHISA</u><br><u>MDAMAHALLNAAA</u> <u>IL</u> <u>EE</u> <u>SPLPAMVKERYAS</u> <u>FD</u> <u>SG</u> <u>LG</u> <u>LK</u> <u>KK</u> <u>FE</u> <u>EG</u> <u>KAS</u> <u>LE</u> <u>EL</u> <u>YE</u> <u>YAK</u> <u>KN</u> <u>GE</u> <u>PPV</u><br><u>A</u> <u>A</u> <u>S</u> <u>G</u> <u>K</u> <u>Q</u> <u>E</u> <u>L</u> <u>C</u> <u>E</u> <u>T</u> <u>Y</u> <u>I</u> <u>N</u> <u>L</u> <u>Y</u> <u>A</u> <u>K</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 5752MI2_003 | <i>Prevotella</i> | DNA              | 159        | <u>AT</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>A</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>T</u> <u>T</u> <u>G</u> <u>A</u> <u>G</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u><br><u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u><br><u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u><br><u>T</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>C</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u><br><u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>T</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>T</u> <u>C</u> <u>G</u><br><u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>G</u> <u>A</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u><br><u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>T</u> <u>C</u> <u>T</u> <u>G</u><br><u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>T</u> <u>T</u> <u>T</u> <u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u><br><u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>A</u><br><u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u><br><u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>C</u><br><u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u><br><u>G</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u><br><u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u><br><u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u><br><u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>C&lt;/</u> |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------|-------------------|------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|             |                   |                  |            | AAGGACATCACCGACTACCTCGTCGAGAAGCAGAAGAAACCGGCATCAAGAACCTCTGG<br>GGCAGGCCCAACGTGTTCCGCAACAAGCGCTACATGAACGGCGCGCCACGAACCCGCAG<br>TTCGACGTGTCGCCCCGCCCGCTCCAGATCAAGAACGCCATCGACGCCACCATCAAG<br>CTGGCGGTACCGGTTACGTGTTCTGGGGCGGCGTGAAGGCTACTACACCTCCTGAAC<br>ACCCAGATGCAGCGCGAGAAAGGACCACTCGCCAAGATGCTCACCGCGCCCGCGACTAC<br>GCCCCGCCCACGGCTTCCAGGGCACCTTCTCATCGAGCCCAAGCCCATGGAGCCCACC<br>AAGCACAGTACGACGTGGACACGGAGACCGTGATCGGCTTCTGCGGCCAACGGTCTG<br>GACAAGGACTTCAAGGTCAATATCGAGGTGAACCAAGCCACCTCGCCGGCCACACCTTC<br>GAGCACGAGTCACCGTGGCTGTCGATAACGGCTTCTCGGCTCCATCGACGCCAAACCGC<br>GGCACGCCCAGAACGGCTGGGACACCGACCAAGTTCCTCCGTGGACCCGTACGACCTCACC<br>CAGGCCATGATGCAGATCATCCGCAACGGCGGTTTCAAGGACGGCGCACCAACTTCGAC<br>GCCAAGCTCGCCCGCTCTTCCACCGACCCGGAGGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCCATGGCCCCACGCCCTGCTGAACGCCCGCCCGCTCATCGAGGAGAGCCCGCTC<br>TGCAAGATGGTCGAGGAGCGTACGCTTCTTCGACAGCGGCGCTCGGCAAGCAGTTCGAG<br>GAAGGCAAGGCCACCCCTCGAGGACCTCTACGAGTATGCCAAGAAGAATGGCGAGCCCGTC<br>GTCGCCCTCCGGCAAGCAGGAGCTCTACGAGACGCTGCTGAACCTTACGGCGAAGTAG                                                                                |
| 5752MI3_002 | <i>Prevotella</i> | Amino Acid       | 162        | MAKEYFPTIGKIPFEGVESKNPMAFHYYDANRVVMGKPMKDWLKFAMAWHHTLGOASGDP<br>FGGQTRS YEWDKGECPYCRAKAKADAGFEIMQKLGIEYYCFHDI DLVEDTEDIAEY EARM<br>KDITDYLVEKQKETGIKNLWGTANVFGNKR YMNGAATNPQFDVVARA AVQIKNAIDATIK<br>LGGTGYVFWGGREGYYTLLNTQMQR EKDHLAKMLTAARDYARAHGFQGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVAVDNGFLGSI DANR<br>GDAQNGWDTDQFPVDPYDLTQAMMQI IIRNGGFKDGGTNFD AKLRRSSTD PEDIFIAHISA<br>MDAMAHALLNAAA VIEESPLCKMVEERYASFDSGLGKQFEEGKATLEDLYEYAKKNGEPV<br>VASGKQELYETLLNLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 5752MI5_003 | <i>Prevotella</i> | DNA              | 163        | ATGGCAAAAGAGTATTTCCCGACAATCGGTAAGATCCCTTCGAGGGACCCCGAGTCCAAG<br>AACCCGATGGCATTCAC TACTATGACCGGGAGCGCGTGGTGATGGGCAAGAAGATGAAG<br>GACTGGTTCAAGTTCGCGATGGCCTGGTGGCACACCTTGGGCCAGGCTTCGCGCCAGCCCG<br>TTCGGCGGCAGACCCGCTCTACGAGTGGGACAAGGGCGAAGGCCCTGTCTCCCGCGCC<br>CGCGCCAAAGGCTGACGCCCGGTTTCGAGATCATGCAGAAACTGGGCATCGGCTACTACTGC<br>TTCACGACATCGACCTGGTGGAGGACACCGAGGACATCGCCGAGTATGAAGCCCGCATG<br>AAGGACATCACCGACTACCTCGTGGAGAAGCAGAAGGAGACCGGCATCAAGAACCTCTGG<br>GGCAGGCCCAACGTATTCGGCAACAAGCCCTACATGAACGGCGCCGCCACGAACCCGCAG<br>TTCGACATCGCCGCCCGCGCCCTGCAGACCAAGAACGCCCATCGATGCCACCATCAAG<br>CTGGCGGCACCCGGTTACGTGTTCTGGGCGCGCCGTGAAGGCTACTACACCTCCTGAAC<br>ACCCAGATGCAGCGCGAGAAGGACCACTTGCCAAGATGCTCACCGCGGCTCGCGACTAT<br>GCCCCGCCCAACGGCTTCAAGGGCACCTTCTTCATCGAGCCGAAACCGATGGAGCCCACC<br>AAGCACAGTACGACGTGGACACGGAGACCGTGATCGGCTTCTTCGCGGCCAACGGCCTG<br>GACAAGGACTTCAAGGTGAACATCGAAGTGAACCAAGCCACCTCGCCGGCCACACCTTC<br>GAGCACGGGCTCACCGTGGCCGTTGACAACGGCTTCTTCGGCAGCATCGACGCCAAACCGC<br>GGAGACGCCCAGAACGGCTGGGATACCGACCAAGTTCCTCCGGTGATCCGTACGACCTCACC<br>CAGCGGATGATCCAGATCATCCGCAATGGCGGCTTCAAGGACGGCGGTACCAACTTCGAC |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------|-------------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|             |                   |                  |            | GCCAAGCTCCGCCGCTCTTCCACCGACCCGGAGGACATCTTCATCGCCACATCAGCGCG ATGACGCCATGGCCCCACGCCCTGCTGAACGCCCGCCGCGTCTCGAGGAGAGCCCCGCTC TCGAGATGGTTGCAAAGCGTTACGTTCTTTCGACAGCGGTCTCGGCAAGAAGTTCGAG GAAGGCAACGCCACCTCGAGGAACTCTACGAGTACGCCAAGGCGAAGGCGAGGTCGTT GCCGAATCCGGCAAGCAGGAACTCTACGAGACCCCTGCTGAACCTCTACGCGGAAGTAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 5752MI5_003 | <i>Prevotella</i> | Amino Acid       | 164        | MAKEYFPTIGKIPFEGPESKNPMAFHYYDAERVVMGKKMKDWFKFAMAWHHTLGQASADP FGGQTRSYEWDKGEGPCSRARAKADAGFEIMQKLIGIYCFHDI DLVEDTEDIAEYEARM KDITDYLVEKQKETGIKNLWGTANVFGNKPYMNGAATNPQFDIAARAALQTKNAIDATIK LGGTGYVFWGGREGYYTLLNTQMQRKDLAKMLTAARDYARAHGFKGTFFIEPKPMEPT KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHGLTVAVDNGFLGSI DANR GDAQNGWDTDQFFPYDLTQAMIQIIRNGGFKDGGTNFDAKLRRSSTDPEDIFIAHISA MDAMAHALLNAAA VLEESPLCEMVAKRYASFDSGLGKKFEEGNATLEELYEYAKAKGEVV AESGQELLYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 5752MI6_004 | <i>Prevotella</i> | DNA              | 165        | ATGGCAAAAAGAGTATTTCCCGCAATCGGAAAAGATCCCCTTCGAGGGCGCTGAGAGCAAG AATCCCCCTTGCTTCCACATATTATGACGCCGAGCGTGTGGTCATGGGCAAGCCCCATGAAG GACTGGTTCAAGTTCGCGATGGCCTGGTGGCACACCCCTGGGCCAGGCTTCCGCCGACCCG TTCCGGCGGCAGACCCGCTCCTACGAGTGGGACAAGGGCGAGTCCCCCTACTGCCCGCGCC CGCCAGAAAGGCTGACGCCGGTTTCGAGATCATGCAGAAAGCTCGGCATCGGCTACTACTGC TTCCACGACATCGACCTGGTCGAGGACACCCGAGGACATCGCCGAGTACGAGGCCCGCATG AAGGACATCACCGACTACCTCGTCGAGAAGCAGAAGGAGACCGGCATCAAGAACCTCTG GGCACGGCCAAACGTGTTCCGCAACAAGCGCTACATGAACGGCGCCGCCACGAACCCGCAG TTCGACATCGTCGCCACCGGCCCTGCAGATCAAGAACCGGATCGGCGCCACCATCAAG CTGGCGGCACCGGTTACGTGTTCTGGGCGGCCGTGAAGGTTACTACACCTCCTGTAAC ACCCAGATGCAGCGCGAGAAAGGACCACTCGCCAAGATGCTCACCGCCGCCCGGACTAC GCCCGGCCAAACGGCTTCAAGGGCACCTTCTCATCGAGCCGAAAGCCGATGGAGCCACCC AAGCACCATATGACGTGGACACGGAGACCGTGATCGGCTTCTCCGCGCCAACGGCCTG GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAAGCCACCTCGCCGGCCACACCTTC GAGCACGAGCTCACCGTGGCGGTGACAAACGGCTTCTTCGGCAGCATCGACGCCAAACCGC GGTGACGCCCAGAACGGCTGGGATACCGACCAAGTTCCC GGTTGGATCCC GTACGATCTCACC CAGGCGATGATCCAGATCATCCGCAACGGCGGCTTCAAGGATGGCGGACCAACTTCGAC GCCAAGCTCCGCCGCTTCCACCGACCCGGAGGACATCTTCATCGCCCATCAGCGCG ATGGACGCCATGGCCACGCCCTGCTGAACGCCCGCCGCTCATCGAGGAGAGCCCCGCTC TCGGAGATGGTCGCCAAGCGCTACGCTTCTTCGACAGCGGTCTCGGCAAGAAGTTCGAG GAAGGCAACGCCACCTCGAGGAACTCTACGAGTACGCCAAGGCGAACGGTGAGGTCAAG GCCGAATCCGGCAAGCAGGAGCTCTACGAGACCTTCTGAACCTCTACGCGGAAATAG |
| 5752MI6_004 | <i>Prevotella</i> | Amino Acid       | 166        | MAKEYFPTIGKIPFEGAESKNPLAFHYYDAERVVMGKPMKDWFKFAMAWHHTLGQASADP FGGQTRSYEWDKGECPYCRARQKADAGFEIMQKLIGIYCFHDI DLVEDTEDIAEYEARM KDITDYLVEKQKETGIKNLWGTANVFGNKRMYMNGAATNPQFDIVAHAALQIKNAIGATIK LGGTGYVFWGGREGYYTLLNTQMQRKDLAKMLTAARDYARANGFKGTFLIEPKPMEPT KHQYDVDTETVIGFLRANGLDKDFKVNIEVNHATLAGHTFEHELTVAVDNGFLGSI DANR GDAQNGWDTDQFFPYDLTQAMIQIIRNGGFKDGGTNFDAKLRRSSTDPEDIFIAHISA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|-------------------|------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|             |                   |                  |            | <u>MDAMAHALLNAAA</u> VIIEESPLCEMVAKRYASFDSGLGKKFEEGNATLEELYEYAKANGEVK<br>AESGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 5753MI1_002 | <i>Prevotella</i> | DNA              | 167        | ATGGCAAAAGAGTATTTCCCCACTATCGGGAAAGATTCCTTTCGAAGGAGTCGAGAGCAAG<br>AACCCCTTGCAATCCATTATATGACGCAAAACCGCATGGTCATGGGCAAGCCCATGAAG<br>GACTGGTTCAAAGTTCGCCATGGCATGGTGGCACACCCCTGGGACAGGCCTCCGCAGACCCG<br>TTCGGCGGCCAGACCCGCTCCTACGAATGGGACAAGGGCGAATGCCCCTACTGCCCGGCC<br>AGGGCAAAAGGCCGATGCCGGCTTCGAGATCATGCAGAAACTGGGTATCGAGTATTTCTGC<br>TTCCATGACATCGACCTGGTAGAGGACTCGGACGACATCGCCGAGTACGAGGCCCGCATG<br>AAGGACATCACGGACTATCTCCTGGAGAAGATGAAGGAACCCGGCATCAAGAACCCTCTGG<br>GGCACCGCCAACGTTTCGGCAACAAGCGTTACATGAACGGCGCGGCACCAATCCGCAG<br>TTCGACGTAGTGGCCCCGCGCTGCCGTCCAGATCAAGAACGCCCATCGACGCCACCATCAAG<br>CTCGGCGGTTCCAACATATGTGTTCTGGGGCGCGCTGAAGGATACTACACCTGCTGAAC<br>ACCCAGATGCAGCGCGAGAAAGGACCACTCGGCAAACTGCTCACCGCGCCCGCGACTAT<br>GCCCCAAAGAACGGCTTCAAAGGCACCTTCCTCATCGAGCCCAAAGCCGATGGAGCCCCACC<br>AAGCACCAGTACGACGTAGACACGGAGACCGTGATCGGCTTCCTCCGCGCCAACGGCCTG<br>GAGAAAGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCTTGGCCGGCCATACCTTC<br>GAGCATGAACTCACCGTGGCCGTGGACAACGGCTTCCTGGGATCCATCGACGCCCAACCGC<br>GGGACGCCCAGAACGGCTGGGATACGGACCAAGTTCCCGGTAGACCCGTACGACCTCACCC<br>CAGGCCATGATGCAGATCATCCGCAACGGCGGCTCGGCAACGGCGGTACCAACTTCGAC<br>GCCAAACTGCGCCGTTCTCCACCGATCCTGAGGACATCTTCATCGCCACATCAGCGCC<br>ATGGACGCCATGGCCCCACGCCCTGCTCAACCGACGCCCGCTGCTGGAAGAAAGTCCGCTC<br>TGTGAGATGGTCAAGGAGCGCTACGTTCTTCGACAGCGGTCTCGGCAAGAAAGTTCGAA<br>GAGGCAAGGCTACCTTGGAAGAAATCTACGATATGCCAAGAAAGAGCGGCGAACCCTGTG<br>GTCGCTTCCGGCAAGCAGGAGCTCTACGAAACCTCTGCTGAACCTCTACGCCCAAGTAG |
| 5753MI1_002 | <i>Prevotella</i> | Amino Acid       | 168        | MAKEYFPITIGKIPFEGVESKNPLAFHYIDANRMVMGKPMKDWFKFAMAWWHHTLGQASADP<br>FGGQTRSYEWDKGECPYCRARAKADAGFEIMQKLGIEYFCFHDIDLVEDCDDIAEYEARM<br>KDITDYLLLEKMKETGIKNLWGTVANVFGNKRVMNGAGTNPQFDVVARAAVQIKNAIDATIK<br>LGGSNYVFWGGREGYYTLLNTQMQREKDLGLKLLTAARDYARKNGFKGTFLIEPKPMEPT<br>KHQYDVDTETVIGFLRANGLEKDFKVNIEVNHATLAGHTFEHELTVAVDNGLGSI DANR<br>GDAQNGWDTDQFPVPDYPDLTQAMMQII RNGGGLNGGTFNFDAKLRRSSTDPE DIFIAHISA<br><u>MDAMAHALLNAAA</u> VL EESPLCEMVKERYASFDSGLGKKFEEGKATLEEIYEYAKKSGEPV<br>VASGKQELYETLLNLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5753MI2_002 | <i>Prevotella</i> | DNA              | 169        | ATGGCTAAAGAATACTTCCCCCTCCATCGGCAAAATCCCCTTTTGAAGGAGGCGACAGCAAA<br>AATCCCCCTCGCTTCCATTATTATGACGCCGACGCGTGGTTATGGCAAGCCCCCATGAAG<br>GAATGGCTTAAATTCGCCATGGCCTGGTGGCACACCGTGGGCCAGGCCTCCGGAGACCCC<br>TTCGGCGGCCAGACCCGAGCTACGAATGGGACAAGGGCGAATGCCCCTACTGCCCGGCC<br>AAAGCCAAGGCCGACGCCGTTTGAATCATGCAAAAAGCTGGGTATCGAATACTTCTGC<br>TTCACGATGTGGACCTTATCGAGGATTTCGATGACATTCGCCGAATACGAAGCCCGCATG<br>AAGGACATCACGGACTACCTGCTGGA AAAAGATGAAGGAGACCGGCATCAAGAACCCTCTGG<br>GGCACCGCCAATGTCTTCGGCCACAAGCGCTACATGAACGGCGCCCGCATCAAGCAACCCGAG<br>TTCGACGTGGTCCGCCGCGCGCTCCAGATCAAGAACCGCATCGACGCCACCATCAAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|-------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|             |                   |                  |            | CTCGGGGTACAGTTATGTATTCTGGGGCGCGGAGGGCTACTACACCCCTCCTGAAC<br>ACCCAGATGCAGCGTGAGAAAGACCACCTGGCCAAAGATGCTCACCGCAGCCCCGCGACTAC<br>GCCCCGCCAAGGGCTTCAAGGGCACCTTCTCATCGAGCCCAAGCCGATGGAGCCCCACC<br>AAGCACAGTACGACGTTGACACGGAGACCGTGATCGGCTCCCTGCGGCCAACGGCCTG<br>GACAAGGACTTCAAGGTGAACATCGAGGTGAACCAACGCCACCTTGGCCGCCACACCTTC<br>GAGCACGAACTCACCGTGGCTGTTGACAACGGCTTCTGGGCTCCATCGACGCCAACCGC<br>GGGACGCCCCAGAACGGCTGGGATACGGACCAAGTCCCGGTAGACCCGTACGACCTCACCC<br>CAGGCCATGATGCAGATTATCCGCAACGGCGGCTTCAAGGACGGCGGCACCAACTTCGAT<br>GCCAAACTGCGCCGCTTTCACCGATCCGGAAGACATCTTCATGCCCACATCAGCGCT<br>ATGGATGCCATGGCACACGCCCTGCTCAACGCCCGCGCTGCTGGAAGAGAGCCCCGCTG<br>TGCAACATGGTCAAGGAGCGTTACGCCGGCTTCGACAGCGGCTTGGCAAGAAGTTTCGAG<br>GAAGGAAAGGCAACGCTGGAGGAAATCTATGACTATGCCAAGAAAGACGGCGAACCCGTC<br>GTGGCTTCCGGCAAGCAGGAACTCTACGAAACCATCCTGAACCTCTATGCCAAGTAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 5753MI2_002 | <i>Prevotella</i> | Amino Acid       | 170        | <u>MAKEYFP</u> <u>SI</u> <u>GKI</u> <u>PF</u> <u>EG</u> <u>GS</u> <u>KN</u> <u>PL</u> <u>AF</u> <u>HY</u> <u>D</u> <u>AG</u> <u>RV</u> <u>VM</u> <u>GK</u> <u>PM</u> <u>KE</u> <u>WL</u> <u>K</u> <u>F</u> <u>AM</u> <u>AW</u> <u>HT</u> <u>LT</u> <u>Q</u> <u>AS</u> <u>GD</u> <u>P</u><br><u>FG</u> <u>Q</u> <u>T</u> <u>R</u> <u>S</u> <u>Y</u> <u>E</u> <u>W</u> <u>D</u> <u>K</u> <u>G</u> <u>E</u> <u>C</u> <u>P</u> <u>Y</u> <u>C</u> <u>R</u> <u>A</u> <u>K</u> <u>A</u> <u>K</u> <u>A</u> <u>D</u> <u>A</u> <u>G</u> <u>F</u> <u>E</u> <u>I</u> <u>M</u> <u>Q</u> <u>K</u> <u>L</u> <u>G</u> <u>I</u> <u>E</u> <u>F</u> <u>C</u> <u>H</u> <u>D</u> <u>V</u> <u>D</u> <u>L</u> <u>I</u> <u>E</u> <u>D</u> <u>C</u> <u>D</u> <u>D</u> <u>I</u> <u>A</u> <u>E</u> <u>Y</u> <u>E</u> <u>A</u> <u>R</u> <u>M</u><br><u>K</u> <u>D</u> <u>I</u> <u>T</u> <u>D</u> <u>Y</u> <u>L</u> <u>L</u> <u>E</u> <u>K</u> <u>M</u> <u>K</u> <u>E</u> <u>T</u> <u>G</u> <u>I</u> <u>K</u> <u>N</u> <u>L</u> <u>G</u> <u>T</u> <u>A</u> <u>N</u> <u>V</u> <u>F</u> <u>G</u> <u>H</u> <u>K</u> <u>R</u> <u>Y</u> <u>M</u> <u>N</u> <u>G</u> <u>A</u> <u>T</u> <u>N</u> <u>P</u> <u>Q</u> <u>F</u> <u>D</u> <u>V</u> <u>V</u> <u>A</u> <u>R</u> <u>A</u> <u>A</u> <u>V</u> <u>Q</u> <u>I</u> <u>K</u> <u>N</u> <u>A</u> <u>I</u> <u>D</u> <u>A</u> <u>T</u> <u>I</u> <u>K</u><br><u>I</u> <u>G</u> <u>G</u> <u>T</u> <u>S</u> <u>Y</u> <u>V</u> <u>F</u> <u>W</u> <u>G</u> <u>G</u> <u>R</u> <u>E</u> <u>G</u> <u>Y</u> <u>T</u> <u>L</u> <u>L</u> <u>N</u> <u>T</u> <u>Q</u> <u>M</u> <u>Q</u> <u>R</u> <u>E</u> <u>K</u> <u>D</u> <u>H</u> <u>L</u> <u>A</u> <u>K</u> <u>M</u> <u>L</u> <u>T</u> <u>A</u> <u>R</u> <u>D</u> <u>I</u> <u>Y</u> <u>A</u> <u>R</u> <u>A</u> <u>K</u> <u>G</u> <u>F</u> <u>K</u> <u>G</u> <u>T</u> <u>F</u> <u>L</u> <u>I</u> <u>E</u> <u>P</u> <u>K</u> <u>P</u> <u>M</u> <u>E</u> <u>P</u> <u>T</u><br><u>K</u> <u>H</u> <u>Q</u> <u>Y</u> <u>D</u> <u>V</u> <u>D</u> <u>T</u> <u>E</u> <u>T</u> <u>V</u> <u>I</u> <u>G</u> <u>S</u> <u>L</u> <u>R</u> <u>A</u> <u>N</u> <u>G</u> <u>L</u> <u>D</u> <u>K</u> <u>F</u> <u>K</u> <u>V</u> <u>N</u> <u>I</u> <u>E</u> <u>V</u> <u>N</u> <u>H</u> <u>A</u> <u>T</u> <u>L</u> <u>A</u> <u>G</u> <u>H</u> <u>T</u> <u>F</u> <u>E</u> <u>H</u> <u>E</u> <u>L</u> <u>T</u> <u>V</u> <u>A</u> <u>V</u> <u>D</u> <u>N</u> <u>G</u> <u>F</u> <u>L</u> <u>G</u> <u>S</u> <u>I</u> <u>D</u> <u>A</u> <u>N</u> <u>R</u><br><u>G</u> <u>D</u> <u>A</u> <u>Q</u> <u>N</u> <u>G</u> <u>W</u> <u>D</u> <u>T</u> <u>D</u> <u>Q</u> <u>F</u> <u>P</u> <u>V</u> <u>D</u> <u>P</u> <u>Y</u> <u>D</u> <u>L</u> <u>T</u> <u>Q</u> <u>A</u> <u>M</u> <u>Q</u> <u>I</u> <u>I</u> <u>R</u> <u>N</u> <u>G</u> <u>G</u> <u>F</u> <u>K</u> <u>D</u> <u>G</u> <u>G</u> <u>T</u> <u>N</u> <u>F</u> <u>D</u> <u>A</u> <u>K</u> <u>L</u> <u>R</u> <u>R</u> <u>S</u> <u>T</u> <u>D</u> <u>P</u> <u>E</u> <u>D</u> <u>I</u> <u>F</u> <u>I</u> <u>A</u> <u>H</u> <u>I</u> <u>S</u> <u>A</u><br><u>M</u> <u>D</u> <u>A</u> <u>M</u> <u>A</u> <u>H</u> <u>A</u> <u>L</u> <u>L</u> <u>N</u> <u>A</u> <u>A</u> <u>A</u> <u>V</u> <u>L</u> <u>E</u> <u>S</u> <u>P</u> <u>L</u> <u>C</u> <u>N</u> <u>M</u> <u>V</u> <u>K</u> <u>E</u> <u>R</u> <u>Y</u> <u>A</u> <u>G</u> <u>F</u> <u>D</u> <u>S</u> <u>G</u> <u>L</u> <u>G</u> <u>K</u> <u>K</u> <u>F</u> <u>E</u> <u>E</u> <u>G</u> <u>K</u> <u>A</u> <u>T</u> <u>L</u> <u>E</u> <u>E</u> <u>I</u> <u>Y</u> <u>D</u> <u>Y</u> <u>A</u> <u>K</u> <u>K</u> <u>S</u> <u>G</u> <u>E</u> <u>P</u> <u>V</u><br><u>V</u> <u>A</u> <u>S</u> <u>G</u> <u>K</u> <u>Q</u> <u>E</u> <u>L</u> <u>Y</u> <u>E</u> <u>T</u> <u>I</u> <u>N</u> <u>L</u> <u>Y</u> <u>A</u> <u>K</u> |
| 5753MI4_002 | <i>Prevotella</i> | DNA              | 171        | ATGTCAAAAGAGTATTTCCCTACAAATCGGCAGGGTCCCCTTCGAGGGACCTGAGAGCAAG<br>AATCCGCTGGCGTTCACCTATTACGAGCCGGACCGGCTCGTCTTGGCAGGAAAAATGAAG<br>GACTGGCTGCGCTTCGCAATGGCCTGGTGGCATACGCTCGGCAGGGCTTCGGCGACCCAG<br>TTCGGCGGACAGACCTGCACATACGCTGGGATGAAGCGAGTGTCCCGTCTGCCGGGCA<br>AAGGCCAAGGCTGACGCCGGCTTTGAACGTATGCAGAAACTGGGCATCGGGTATTTCTGC<br>TTCCACGACGTGGACCTGGTCGAGGAGGCCGACACCATTGAAGAAATACGAGGAGCGGATG<br>CGGATCATCACCGACTACCTGCTCGAGAAGATGGAAGAGACCGGCATCCGCAATCTCTGG<br>GGAACCGCCAATGTCTTCGGACACAAAGCGCTATATGAACGGCGCCGCCACCAATCCCAGC<br>TTCGACGTCGTGGCCCCGTGCCCGGTCAGATCAAGAAATGCCATCGATGCCACCATCAAA<br>CTGGGTGGTGAGAACTATGTGTCTGGGGTGGCCGCGAGGGCTATACGAGCCTGCTCAAC<br>ACGCAGATGCACCGGGAAAAACACCACTCGGAAATATGCTCAGGGCAGCCCCGGACTAT<br>GGCCGTGCCACGGTTTCAAGGGAACGTTCTGTATCGAGCCCAAGCCGATGGAGCCGACC<br>AAGCATCAGTACGACCCAGGATACGGAGACGGTTCATCGGTTTCTTGGCTGTACGGCCTG<br>GACAAGGATTTCAAGGTGAACATCGAGGTGAACCAACGCCACGCTCGCCGGACACACCTTC<br>GAGCACGAACTGGCCACTGCGGTGATGCCGGCCTGCTGGGCAGCATCGATGCCAACCGC<br>GGGACGCCCAGAACGGCTGGGATACCGACCACTTCCCGATCGACAACTACGAACTCACG<br>CTGGCGATGCTGCAGATCATCCGCAATGGCGGACTCGACCCGGGGATCGAACTTCGAT<br>GCCAAGTTGCGCCGCAATTCACCCGATCCGGAAGACATCTTCATCGCCACATCAGCGCG<br>ATGGACGCGATGGCCCCGTCCCCTGCTCAATCGGGCGGCATCTGGACCGGAATCGCCGATT<br>CAGGATATGGTCAGGGACCGCTATGCTTCCTTCGACAGCGGAAAGGCAGGGAGTTCGAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

| TABLE 2     |                   |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|-------------------|------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.   | Class of organism | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 5753MI4_002 | <i>Prevotella</i> | Amino Acid       | 172        | GAAGGCAGACTCAGTCTGGAAGACCTCGTGGCCTATGGAAGGAGCACGGTGAGCCGCGC<br>CAGATCTCCGGCAGGCAGGAACCTTATGAAACCATCGTAGCGCTTACTGCAGGTAA<br><br>MSKEYFPTIGRVPFEGPESKNPLAFHYEPDRLVLGRKMKDWLRFAMAWWHTLQASGDQ<br>FGGQTCTYAWDEGECPCRAKAKADAGFELMQKLIGYFCFHDVDLVEEADTIEEYEERM<br>RIITDYLLLEKMEETGIRNLWGTANVFGHKRYMNGAATNPDEFDVVARAAVQIKNAIDATIK<br>LGENYVFWGGREGYTSLLNTQMHREKHHIGNMLRAARDYGRAHGFKGTFLIEPKPMEPT<br>KHQYDQDTETVIGFLRCHGLDKFKVNIENVNHATLAGHTEHELATAVDAGLLGSIDANR<br>GDAQNGWDTDQFPIDNYELTLAMLQIIRNGGLAPGGSNFDKLRNSTDPEDIFIAHISA<br>MDAMARALLNAAAIIWTESPIQDMVRDRYASFDSGKGREFFEEGRLSLEDLVAYAKEHGEPR<br>QISGRQELYETIVALYCR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 5752MI4_004 | <i>Prevotella</i> | DNA              | 173        | ATGACTAAAGAGTATTTCCCGGGAATCGGAACGATTCCGTTTGAAGGAACCAAGAGCAAG<br>AACCCCTGGCCTTCCATTATTATAACGCCTCCCAGGTAGTGATGGGCAAGCCCCATGAAG<br>GACTGGCTCAAGTATGCCATGGCCTGGTGGCACACCTGGCCAGGCCTCTGCAGACCCC<br>TTTGGCGGCCAGACCCGCTCCTACGAATGGGACAAAGGCGAGTGCCCGTACTGCCCGCGCC<br>AAGCAGAAAGGCCGATGCCGGCTTTGAGCTCATGCAGAAAGCTGGGCATCGAGTACTACTGC<br>TTCCACGACGTGGACATCATCGAGGACTCGGAGGACATTGCCGAGTACGAGGCCCGCATG<br>AAGGACATCAGGACTACCTGCTGGAGAAGCAAAAGAGACCGGCATCAAGAACCTCTGG<br>GGACCGCCAAACGTGTTTGGCCACAAAGCGTACATGAACGGCGCGCCACCAACCCCTCAG<br>TTTGACATTTGTGGCCCGTGCCGCCGTCCAGATCAAGAAACGCCCTGGATGCCGCCATCAAA<br>CTGGGTGGTACCAACTACGTGTTCTGGGTGGCCGCGAAGGTACTACACGCTGCTCAAC<br>ACCCAGATGCAGCGGGAGAAACCACTGGCCAAGATGCTCACCGCCGCCGCGACTAC<br>GCCCCGCCAAAGGCTTCAAGGGCACCTTCTCATTTAGCCCCAAACCATGGAGCCCCACC<br>AAGACCAGTACGACGTGGACACCGAGACCGTGATTGGTTTCATCCGCGCCAACGGCCTG<br>GACAAGGACTTCAAGGTAACAATTGAGGTAAAACCAACGCCACCTGGCCGGCCACACCTTT<br>GAGCACGAGCTCACCGTGGCCCGCGAGAACGGCTTCTGGGCTCCATCGACGCCCAACCGC<br>GGAGATGCCCAGAACGGCTGGGATACGGACCAAGTTCCTCATCGACGCCCTGGATCTCACC<br>CAGGCTATGATGCAGGTTCCTCAACGGTGGCTTCGGCAATGGCGGCACCAACTTTGAC<br>GCCAAGCTCCGCCGCTCCTCACCGATCCCGAGGACATCTTCATCGCCCATCAGCGCC<br>ATGGATGCCATGGCACACGCCCTCCTGAACGCAGCGCCCATCCTGGAAGAGAGCCCCCTG<br>CCCCCATGGTCAAGGAGCGTTACGCTTCTTCGACAGCGGTCTGGGCAAGAAGTTCGAA<br>GAAGGCAAGGCCTCCCTGGAAGAACTTTACGAATATGCCAAGAAGAAATGGAGAGCCCCGTG<br>GCCGCTTCCGGCAAACAGGAGCTCTGCGAAACTTACTTGAACCTCTATGCAAAGTAG |
| 5752MI4_004 | <i>Prevotella</i> | Amino Acid       | 174        | MTKEYFPGIGTIPEGTKSKNPLAFHYYNASQVVMGKPMKDWLKYAMAWWHTLQASADP<br>FGGQTRSYEWDKGECPYCRAKQKADADAGFELMQKLGI EYCFHDVDI IEDCEDIAEY EARM<br>KDI TDYLLLEKQKETGIKNLWGTANVFGHKRYMNGAATNPQFDIVARA AVQIKNALDAAIK<br>LGGTNYVFWGGREGYTYLLNTQM QREKNHLAKMLTAARDYARAKGFKGTFLIEPKPMEPT<br>KHQYDQDTETVIGFIRANGLDKDFKVNIEVNHATLAGHTEHELTVARENGFLGSI DANR<br>GDAQNGWDTDQFPIDALDLTQAMMQVILNGGFNGGNTNFDKLRNSTDPEDIFIAHISA<br>MDAMAHALLNAAAII EESPLPAMVKERYASFDSGLGKKFEEGKASLEELYEYAKKNGEPV<br>AASGKQELCETYINLYAK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| TABLE 2    |                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|------------|--------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clone No.  | Class of organism  | Type of Sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 727MI4_006 | <i>Rhizobiales</i> | DNA              | 175        | GTGACTGATTTCTTCAAGGGCATCGGCCCCGTCAAGTTTGAGGGGCCGACAGAGTCCAAT<br>CCGCTGGCCTATCGCCACTATAACAAGGACGAAATCGTCCTCGGCAAGCGGATGGAAGAC<br>CATATCCGTCCCGCGTTGCCCTATTGGCACACCTTCGCCTATGAGGGCGCGATCCGTTT<br>GGGGCCGCACCTTCGATCGCCCTGGTTCGACAAGGTATGGACGGGCCCGCCTCAAG<br>GCCGACGTGGCCTTCGAACTGTCGACCTGCTCGACGTTCCTTTCTTCTTCCACGAT<br>GCTGATATCGCTCCCGAAGCGCAACGCTGGCCGAGAGCAACCGCAATGTGCGCGAGATT<br>GGCGAGATCTTCGCTCGCAAGATGGAACACGACCGCACCAAGCTGCTCTGGGGTACGGCA<br>AACCTGTTCTCCAATCGCCGCTACATGGCCGGCGGCCACCAACCCGGACCCGGAAATC<br>TTGCCTATGCCGTGGCAGGTGAAGAACGTGCTGGAACCTGACCCACGAACTGGCGGC<br>GCCAACTATGTGCTGTGGGCGGTGCGAGGGTTATGAAACCTGCTCAACACCAAGATC<br>GGCAGGAAATGGACCAGATGGCCGTTTCTGTCGATGGTCGTCGAGCATGCCGAAAAG<br>ATCGGCTTCAAGGGCCAGATCCTGATCGAGCCCCAAGCCGACGAGCCGAGCAAGCACCCAG<br>TATGACTTCGACGTTGCAACCGTTTACGGCTTCCTCAAGAAGTATGGTCTCGAAACCAAG<br>GTGAAGTGCAATATCGAGGTGGCCATGCCTTCCTCGCCAATCACTCCTTCGAGCATGAA<br>CTGGCTTTGGCCGCATCGCTGGGCATTTCTCGGCTCGGTCGACGCCAATCGCAACGATCTA<br>CAGTCCGGCTGGGATACCGACCACTTCCCCAATAATGTCCCCGAACCGCACTCGCCTTC<br>TATCAGATTCTCAAGGGCGGGACTGGGCAATGGCGCTGGAACCTCGACGCCCGCGTG<br>CGCCGCCAGTCACTTGATCCGGCCGACCTGCTGCACGGCCATATCGGCGGCTCGACGTG<br>CTGGCGCGGGCCTCAAGGCCCGCGGCGCTGATCGAGGACGGCACCTATGACAAAGTCTC<br>GTCGACGCCCGCTATGCCGGCTGGAACCAAGGCTCGACGCCGAAAACCTCAACCCGACGCTAGG<br>CTGAACCTTGCCGACCTGGCTGCCAAGGTGACGCCGAAAACCTCAACCCGACGCTAGG<br>TCCGGCCAGCAGGAATATCTCGAAAACCTGATCAACCGGTTCGTTTAG |
| 727MI4_006 | <i>Rhizobiales</i> | Amino Acid       | 176        | MTDFFKGIAPVKEGPPQSSNPLAYRHYNKDEIVLGRMEDHIRPGVAYWHTFAYEGGDPF<br>GGRTFDRPWFDKMGD GARLKADVAFELEFDLLDVPFFCFHDADIAPEGATLAE SNRNVREI<br>GEIFARKMETSR TKLLWGTANLFSNRRY MAGAATNPDP EIFAYAAGQVKNVLELTHELGG<br>ANYVLWGGREGYETLLNTKIGQEMDQMRFLSMVVEHAEKIGFKGQILIEPKPQEP SKHQ<br>YDEDVATVYGFLKKYGLETKVKCNIEVGHAFLANHSFEHELALAA SLGILGSVDANRNDL<br>QSGWDTDQFPNNVPETALAFYQILKAGGLNGGWNFDARVRRQSLDPADLLHGHIGGLDV<br>LARGLKA AAAALIEDGTYDKVVDARYAGWNQGLGKDI LGGKLN LADLAAKVDAENLN PQPR<br>SGQQEYLENLINRFV                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

### 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 300 $\mu$ l of lysis buffer: approximate concentrations (50 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 8.0), 300 mM NaCl, 10 mM imidazole (Sigma, #I5513), to which was added about 2  $\mu$ l/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  $\mu$ m). 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.

| <b>Table 3</b>               |                            |              |
|------------------------------|----------------------------|--------------|
| <b>XI activity at pH 7.5</b> |                            |              |
| <b>SEQ ID NO:</b>            | <b>Volumetric Activity</b> | <b>FIOPC</b> |
| 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         |

| <b>Table 3</b>               |                            |              |
|------------------------------|----------------------------|--------------|
| <b>XI activity at pH 7.5</b> |                            |              |
| <b>SEQ ID NO:</b>            | <b>Volumetric Activity</b> | <b>FIOPC</b> |
| 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         |

| <b>Table 3</b>               |                            |              |
|------------------------------|----------------------------|--------------|
| <b>XI activity at pH 7.5</b> |                            |              |
| <b>SEQ ID NO:</b>            | <b>Volumetric Activity</b> | <b>FIOPC</b> |
| 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 OD<sub>600</sub> 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.

| <b>Table 4</b>                   |                 |           |           |           |                     |              |
|----------------------------------|-----------------|-----------|-----------|-----------|---------------------|--------------|
| <b>Anaerobic EtOH Production</b> |                 |           |           |           |                     |              |
|                                  | <b>Time (h)</b> |           |           |           |                     |              |
| <b>SEQ ID NO:</b>                | <b>0</b>        | <b>24</b> | <b>48</b> | <b>72</b> | <b>EtOH (g/L/h)</b> | <b>FIOPC</b> |
| 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  |

| <b>Table 4</b>                   |                 |           |           |           |                     |              |
|----------------------------------|-----------------|-----------|-----------|-----------|---------------------|--------------|
| <b>Anaerobic EtOH Production</b> |                 |           |           |           |                     |              |
|                                  | <b>Time (h)</b> |           |           |           |                     |              |
| <b>SEQ ID NO:</b>                | <b>0</b>        | <b>24</b> | <b>48</b> | <b>72</b> | <b>EtOH (g/L/h)</b> | <b>FIOPC</b> |
| 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.

| <b>Table 5</b>                        |                                |                        |                          |                                |
|---------------------------------------|--------------------------------|------------------------|--------------------------|--------------------------------|
| <b>XI activity at pH 6 and pH 7.5</b> |                                |                        |                          |                                |
| <b>SEQ ID NO:</b>                     | <b>Organism Classification</b> | <b>VA, pH 6 (U/ml)</b> | <b>VA, pH 7.5 (U/ml)</b> | <b>Percent activity (pH 6)</b> |
| 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        | caccattaattaaAGCTTTGTAAATATGATGAGAGAATAATATA<br>AATCAAACG        |
| 131.5AR                               | 183        | GGCGCGCCTCTAGAAAGCTTAATCGACAAGAACACTTCT<br>ATTTATATAGGTATGAAA    |
| 131.5BF                               | 184        | GCAGGGATATCGGTACCCACCAGCGGCCGCTGAAGAAG<br>GTTTATTTTCGTTTCGCTGT   |
| 131.5BR                               | 185        | caccattaattaaCCCAGGTGAGACTGGATGCTCCATA                           |
| ABMCSF                                | 186        | GCCTCTAGAAAGCTTACGCGTGAGCTCCCTGCAGGGATA<br>TCGGTACCCACCAGCGGCCGC |
| ABMCSR                                | 187        | CGCTGGTGGGTACCGATATCCCTGCAGGGAGCTCACGCG<br>TAAGCTTTCTAGAGGCGCGCC |

[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: TGCGTGTGCCGCGAGTCCACGTCTACTCGCGAACCGAGTGCAGGCGGGTCTTCGGCCAGGAC GGCCGTGCGTGACCCCGGCCGCCAGACGAAACGGACCGCGCTCGCCAGACGCTACCCAGCC CGTTCATGCCGGCCGCGAGCCGACCTGTCTCGGTCGCTTCGACGCACGCGCGGTCTTCGG GTACTCGCCTAAGAC (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        | caatagcgccgcggtaccTGCGTGTGCCGCGAGTCCAC        |
| R88-BamHI-R                           | 190        | TGTTAGGATCCGTCTTAGGCGAGTACCCGAAAGG            |
| BamHI-ura-F                           | 191        | caataggatccAGGCATATTTATGGTGAAGAATAAGT         |
| ura-Xho-R                             | 192        | TGTTACTCGAGAAATCATTACGACCGAGATTCCCG           |
| XhoI-R88-F                            | 193        | caatactcgagTGCGTGTGCCGCGAGTCCAC               |
| R88-NotI-R                            | 194        | TGTTAGCGGCCGCGGTCTTAGGCGAGTACCCGAAAGG         |

[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<br>ATTATCAATACTGCCATTTCAAAGA       |
| overlap-TDH-R                         | 197        | AACGTCGACCTCGAGGGATCCACTAGTTCGAAACTAAGT<br>TCTTGGTGTTTTAAACT               |
| overlap-PGK-F                         | 198        | GTGGATCCCTCGAGGTCGACGTTTAAACATTGAATTGAA<br>TTGAAATCGATAGATCAAT             |
| PGK-R                                 | 199        | CACCAGCGGCCGCGGTACCGATATCCCTGCAGGGAGCTC<br>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 pYDABF0006 was replaced with a subset of the XI genes of Table 2 by digestion of pYDABF0006 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        | TTTCCCATTTCGATATTTCGAGCTCC |
| 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 OD<sub>600</sub> = 0.2. Samples were incubated overnight at about 30°C and about 220 rpm. Cultures were harvested when OD<sub>600</sub> 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.

| <b>Table 11</b>                                                 |                                |                       |                          |
|-----------------------------------------------------------------|--------------------------------|-----------------------|--------------------------|
| <b>SA of XI Expressed in an Industrial <i>S. cerevisiae</i></b> |                                |                       |                          |
| <b>SEQ ID NO:</b>                                               | <b>Organism Classification</b> | <b>SA, pH6 (U/mg)</b> | <b>SA, pH 7.5 (U/mg)</b> |
| 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][FY LIV]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 T][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        | ATGTCTGAACCAGCTCAAAAGAAACAAAAGGTTGCTAACAACTCTCTAGAACAATTGAAAGCCTCCGGCACTGTCGTTGTTGCCGACACTGGTGATTTTCGGCTCTATTGCCAAGTTTCAACCTCAAGACTCCACAATAACCCATCATTTGATCTTGGCTGCTGCCAAGCAACCAACTTACGCCAAGTTGATCGATGTTGCCGTGGAATACGGTAAGAAGCATGGTAAGACCACCGAAGAACAAGTCGAAAATGCTGTGGACAGATTGTTAGTCGAATTCGGTAAGGAGATCTTAAAGATTGTTCCAGGCAGAGTCTCCACCGAAGTTGATGCTAGATTGTCTTTTGACACTCAAGCTACCATTGAAAAGGCTAGACATATCATTAATTTGTTTGAACAAGAAGGTGTCTCCAAGGAAAGAGTCCTTATTAAAATTGCTTCCACTTGGGAAGGTATTCAAGCTGCCAAAGAATTGGAAGAAAAGGACGGTATCCACTGTAATTTGACTCTATTATTCTCCTTCGTTCAAGCAGTTGCCTGTGCCGAGGCCCAAGTTACTTTGATTTCCCATTTGTTGGTAGAATTCTAGACTGGTACAAATCCAGCACTGGTAAAGATTACAAGGGTGAAGCCGACCCAGGTGTTATTTCCGTCAAGAAAATCTACAACACTACTACAAGAAGTACGGTTACAAGACTATTGTTATGGGTGCTTCTTTCAGAGCACTGACGAAATCAAAAACCTTGGCTGGTGTTGACTATCTAACATTCTCTCCAGCTTTATTGGACAAGTTGATGAACAGTACTGAACCTTTCCCAAGAGTTTTTGGACCCTGTCTCCGCTAAGAAGGAAGCCGGCGACAAGATTTCTTACATCAGCGACGAATCTAAATTCAGATTGACATTGAATGAAGACGCTATGGCCACTGAAAAATTGTCCGAAGGTATCAGAAAATTCTCTGCCGATATTGTTACTCTATTTCGACTTGATTGAAGAAGAAAGTTACCGCTTAA                                                                                                                                                                                            |
| <i>XKS1 (S. cerevisiae)</i> | DNA              | 216        | ATGTTGTGTTTCAGTAATTCAGAGACAGACAAGAGAGGTTTCCAACACAATGTCTTTAGACTCATACTATCTTGGGTTTGATCTTTCGACCACAACACTGAAATGTCTCGCCATTAACCAGGACCTAAAAATTGTCATTCAGAAACAGTGGAATTTGAAAAGGATCTTCCGCATTATCACACAAAGAAGGGTGTCTATATACACGGCGACACTATCGAATGTCCCGTAGCCATGTGGTTAGAGGCTCTAGATCTGGTTCTCTCGAAATATCGCGAGGCTAAATTTCCATTGAACAAAGTTATGGCCGTCTCAGGGTCCTGCCAGCAGCACGGGTCTGTCTACTGGTCCTCCCAAGCCGAATCTCTGTTAGAGCAATTGAATAAGAAACCGGAAAAAGATTTATTGCCTACTGAGCTCTGTAGCATTTGCAAGGCAAACCGCCCCCAATTGGCAAGACCACAGTACTGCAAAGCAATGTCAAGAGTTTGAAGAGTGCATAGGTGGGCCTGAAAAAATGGCTCAATTAACAGGGTCCAGAGCCCATTTTAGATTTACTGGTCCTCAAATTCTGAAAATTGCACAA TTAGAACCAGAAGCTTACGAAAAAACAAAGACCATTTCTTTAGTG TCTAATTTTTTTGACTTCTATCTTAGTGGGCCATCTTGTTGAATTA GAGGAGGCAGATGCCTGTGGTATGAACCTTTATGATATACGTGAAGAAAATTCAGTGATGAGCTACTACATCTAATTGATAGTTCTTCTAAGGATAAACTATCAGACAAAAATTAATGAGAGCACCCATGAAA AATTTGATAGCGGGTACCATCTGTAAATATTTTATTGAGAAGTACGGTTTCAATACAACTGCAAGGTCTCTCCCATGACTGGGGATAATTAGCCACTATATGTTCTTTACCCCTGCGGAAGAATGACGTTCTCGTTTCCCTAGGAACAAGTACTACAGTTCTTCTGGTACCCGATAAGTATCACCCCTCTCCGAACCTATCATCTTTTCAATCATCCAACCTCTGCCAAACCATTATATGGGTATGATTTGTTATTGTAATGGTTCTTTGGCAAGGGAGAGGATAAGAGACGAGTTAAACAAAGAACGGGAAAATAATTATGAGAAGACTAACGATTGGACTCTTTTAATCAAGCTGTGCTAGATGACTCAGAAAGTAGTGAAAATGAA |

| Table 13             |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|----------------------|------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Name        | Type of sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|                      |                  |            | TTAGGTGTATATTTTCCTCTGGGGGAGATCGTTCCTAGCGTAAAA<br>GCCATAAACAAAAGGGTTATCTTCAATCCAAAAACGGGTATGATT<br>GAAAGAGAGGTGGCCAAGTTCAAAGACAAGAGGCACGATGCCAAA<br>AATATTGTAGAATCACAGGCTTTAAGTTGCAGGGTAAGAATATCT<br>CCCCTGCTTTCGGATTCAAACGCAAGCTCACAAACAGAGACTGAAC<br>GAAGATACAATCGTGAAGTTTGATTACGATGAATCTCCGCTGCGG<br>GACTACCTAAATAAAAGGCCAGAAAGGACTTTTTTTGTAGGTGGG<br>GCTTCTAAAAACGATGCTATTGTGAAGAAGTTTGCTCAAGTCATT<br>GGTGCTACAAAGGGTAATTTTAGGCTAGAAACACCAAACATCATGT<br>GCCCTTGGTGGTTGTTATAAGGCCATGTGGTCATTGTTATATGAC<br>TCTAATAAAATTGCAGTTCCTTTTGATAAATTTCTGAATGACAAT<br>TTTCCATGGCATGTAATGGAAAGCATATCCGATGTGGATAATGAA<br>AATTGGGATCGCTATAATTCCAAGATTGTCCCCTTAAGCGAACTG<br>GAAAAGACTCTCATCTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| TKL1 (S. cerevisiae) | DNA              | 217        | ATGACTCAATTCACTGACATTGATAAGCTAGCCGTCTCCACCATA<br>AGAATTTTGGCTGTGGACACCGTATCCAAGGCCAACTCAGGTCAC<br>CCAGGTGCTCCATTGGGTATGGCACCAGCTGCACACGTTCTATGG<br>AGTCAAATGCGCATGAACCCAACCAACCCAGACTGGATCAACAGA<br>GATAGATTTGTCTTGTCTAACGGTCACGCGGTGCTTTGTTGTAT<br>TCTATGCTACATTTGACTGGTTACGATCTGTCTATTGAAGACTTG<br>AAACAGTTCAGACAGTTGGGTTCAGAACACCAGGTCATCCTGAA<br>TTTGAGTTGCCAGGTGTTGAAGTTACTACCGGTCCATTAGGTCAA<br>GGTATCTCCAACGCTGTTGGTATGGCCATGGCTCAAGCTAACCTG<br>GCTGCCACTTACAACAAGCCGGGCTTTACCTTGTCTGACAACCTAC<br>ACCTATGTTTTCTTGGGTGACGGTTGTTTGCAAGAAGGTATTTCT<br>TCAGAAGCTTCCTCCTTGGCTGGTCATTTGAAATTGGGTAACCTG<br>ATTGCCATCTACGATGACAACAAGATCACTATCGATGGTGCTACC<br>AGTATCTCATTCGATGAAGATGTTGCTAAGAGATACGAAGCCTAC<br>GGTTGGGAAGTTTTGTACGTAGAAAATGGTAACGAAGATCTAGCC<br>GGTATTGCCAAGGCTATTGCTCAAGCTAAGTTATCCAAGGACAAA<br>CCAACCTTTGATCAAAATGACCACAACCATTTGGTTACGGTTCCTTG<br>CATGCCGGCTCTCACTCTGTGCACGGTGCCCCATTGAAAGCAGAT<br>GATGTTAAACAACATAAGAGCAAATTCGGTTTCAACCCAGACAAG<br>TCCTTTGTTGTTCCACAAGAAGTTTACGACCACTACCAAAAGACA<br>ATTTTAAAGCCAGGTGTCGAAGCCAACAACAAGTGGAACAAGTTG<br>TTCAGCGAATACCAAAAAGAAATTCCCAGAATTAGGTGCTGAATTG<br>GCTAGAAGATTGAGCGGCCAACTACCCGCA<br>AATTGGGAATCTAAGTTGCCAACTTACACCGCCAAGGACTCTGCC<br>GTGGCCACTAGAAAATTATCAGAACTGTTCTTGAGGATGTTTAC<br>AATCAATTGCCAGAGTTGATTGGTGGTTCTGCCGATTTAACACCT<br>TCTAACTTGACCAGATGGAAGGAAGCCCTTGACTTCCAACCTCCT<br>TCTTCCGGTTCAGGTAACTACTCTGGTAGATACATTAGGTACGGT<br>ATTAGAGAACACGCTATGGGTGCCATAATGAACGGTATTTCAGCT<br>TTCGGTGCCAACTACAAACCATACGGTGGTACTTTCTTGAACCTC<br>GTTTCTTATGCTGCTGGTGCCGTTAGATTGTCCGCTTTGTCTGGC<br>CACCCAGTTATTTGGGTGCTACACATGACTCTATCGGTGTCCGT<br>GAAGATGGTCCAACACATCAACCTATTGAAACTTTAGCACACTTC<br>AGATCCCTACCAAACATTCAAGTTTGAGAGACCAGCTGATGGTAAC<br>GAAGTTTCTGCCGCCTACAAGAACTCTTTAGAATCCAAGCATACT<br>CCAAGTATCATTGCTTTGTCCAGACAAAACCTTGCCACAATTGGAA<br>GGTAGCTCTATTGAAAGCGCTTCTAAGGGTGGTTACGTACTACAA |

| Table 13                    |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----------------------------|------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Name               | Type of sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|                             |                  |            | GATGTTGCTAACCCAGATATTATTTTAGTGGCTACTGGTTCCGAA<br>GTGTCTTTGAGTGTTGAAGCTGCTAAGACTTTGGCCGCAAAGAAC<br>ATCAAGGCTCGTGTTGTTTCTCTACCAGATTTCTTCACTTTTGAC<br>AAACAACCCCTAGAATACAGACTATCAGTCTTACCAGACAACGTT<br>CCAATCATGTCTGTTGAAGTTTGGCTACCACATGTTGGGGCAAA<br>TACGCTCATCAATCCTTCGGTATTGACAGATTTGGTGCCCTCCGGT<br>AAGGCACCAGAAGTCTTCAAGTTCTTCGGTTTCACCCCAAGAGT<br>GTTGCTGAAAGAGCTCAAAGACCATTGCATTCTATAAGGGTGAC<br>AAGCTAATTTCTCCTTTGAAAAAAGCTTTCTAA                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <i>RPE1 (S. cerevisiae)</i> | DNA              | 218        | ATGGTCAAACCAATTATAGCTCCCAGTATCCTTGCTTCTGACTTC<br>GCCAACTTGGGTTGCGAATGTCATAAGGTCATCAACGCCGGCGCA<br>GATTGGTTACATATCGATGTCATGGACGGCCATTTTGTTCCAAAC<br>ATTACTCTGGGCCAACCAATTGTTACCTCCCTACGTCGTTCTGTG<br>CCACGCCCTGGCGATGCTAGCAACACAGAAAAGAAGCCCACTGCG<br>TTCTTCGATTGTCACATGATGGTTGAAAATCCTGAAAATGGGTC<br>GACGATTTTGCTAAATGTGGTGCTGACCAATTTACGTTCCACTAC<br>GAGGCCACACAAGACCCTTTGCATTTAGTTAAGTTGATTAAGTCT<br>AAGGGCATCAAAGCTGCATGCGCCATCAAACCTGGTACTTCTGTT<br>GACGTTTTATTTGAACTAGCTCCTCATTTGGATATGGCTCTTGTT<br>ATGACTGTGGAACCTGGGTTTGGAGGCCAAAAATTCATGGAAGAC<br>ATGATGCCAAAAGTGGAACTTTGAGAGCCAAGTTCCCCCATTTG<br>AATATCCAAGTCGATGGTGGTTTGGGCAAGGAGACCATCCCGAAA<br>GCCGCCAAAGCCGGTGCCAACGTTATTGTGCTGGTACCAGTGTT<br>TTCCTGACGCTGACCCGCACGATGTTATCTCCTTCATGAAAGAA<br>GAAGTCTCGAAGGAATTGCGTTCTAGAGATTTGCTAGATTAG                                                                          |
| <i>RKII (S. cerevisiae)</i> | DNA              | 219        | ATGGCTGCCGGTGTCCCAAAAATTGATGCGTTAGAATCTTTGGGC<br>AATCCTTTGGAGGATGCCAAGAGAGCTGCAGCATAACAGAGCAGTT<br>GATGAAAATTTAAAATTTGATGATCACAAAATTATTGGAATTGGT<br>AGTGGTAGCACAGTGGTTTATGTTGCCGAAAGAATTGGACAATAT<br>TTGCATGACCCTAAATTTTATGAAGTAGCGTCTAAATTCATTTGC<br>ATTCCAACAGGATTCCAATCAAGAACTTGATTTTGGATAACAAG<br>TTGCAATTAGGCTCCATTGAACAGTATCCTCGCATTGATATAGCG<br>TTTGACGGTGCTGATGAAGTGGATGAGAATTTACAATTAATTAAA<br>GGTGGTGGTGCTTGTCTATTTCAAGAAAAATTGGTTAGTACTAGT<br>GCTAAAACCTTCATTGTCTGTTGCTGATTCAAGAAAAAAGTCACCA<br>AAACATTTAGGTAAGAACTGGAGGCAAGGTGTTCCCATTGAAATT<br>GTACCTTCCTCATAACGTGAGGGTCAAGAATGATCTATTAGAACAA<br>TTGCATGCTGAAAAAGTTGACATCAGACAAGGAGGTTCTGCTAAA<br>GCAGGTCTGTGTAAGTGAACAATAAATACTTCATTATCGATGCG<br>GATTTTCGGTGAAATTTCCGATCCAAGAAAATTGCATAGAGAAATC<br>AAACTGTTAGTGGGCGTGGTGGAAACAGGTTTATTCATCGACAAC<br>GCTTCAAAGCCTACTTCGGTAATTCTGACGGTAGTGTTGAAGTT<br>ACCGAAAAGTGA |
| <i>GAL2 (S. cerevisiae)</i> | DNA              | 220        | ATGGCAGTTGAGGAGAACAATATGCCTGTTGTTTCACAGCAACCC<br>CAAGCTGGTGAAGACGTGATCTCTTCACTCAGTAAAGATTCCCAT<br>TTAAGCGCACAATCTCAAAGTATTCTAATGATGAATTGAAAGCC<br>GGTGAGTCAGGGTCTGAAGGCTCCCAAAGTGTTCCCTATAGAGATA<br>CCCAAGAAGCCCATGTCTGAATATGTTACCGTTTCCTTGCTTTGT<br>TTGTGTGTTGCCTTCGGCGGCTTCATGTTTGGCTGGGATACCGGT<br>ACTATTTCTGGGTTTGTGTCCAAACAGACTTTTTGAGAAGGTTT<br>GGTATGAAACATAAGGATGGTACCCACTATTTGTCAAACGTCAGA                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

| Table 13      |                  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|---------------|------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Name | Type of sequence | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|               |                  |            | ACAGGTTTAATCGTCGCCATTTTCAATATTGGCTGTGCCTTTGGT<br>GGTATTATACTTTCCAAAGGTGGAGATATGTATGGCCGTAAAAAG<br>GGTCTTTCGATTGTCGTCTCGGTTTATATAGTTGGTATTATCATT<br>CAAATTGCCTCTATCAACAAGTGGTACCAATATTTCAATTGGTAGA<br>ATCATATCTGGTTTGGGTGTCGGCGGCATCGCCGTCTTATGTCCT<br>ATGTTGATCTCTGAAATTGCTCCAAAGCACTTGAGAGGCACACTA<br>GTTTCTTGTTATCAGCTGATGATTACTGCAGGTATCTTTTGGGC<br>TACTGTACTAATTACGGTACAAAGAGCTATTCGAACTCAGTTCAA<br>TGGAGAGTTCCATTAGGGCTATGTTTCGCTTGGTCATTATTTATG<br>ATTGGCGCTTTGACGTTAGTTCCTGAATCCCCACGTTATTTATGT<br>GAGGTGAATAAGGTAGAAGACGCCAAGCGTTCATTGCTAAGTCT<br>AACAAGGTGTCACCAGAGGATCCTGCCGTCCAGGCAGAGTTAGAT<br>CTGATCATGGCCGGTATAGAAGCTGAAAACTGGCTGGCAATGCG<br>TCCTGGGGGGAATTATTTTCCACCAAGACCAAAGTATTTCAACGT<br>TTGTTGATGGGTGTGTTTGTTCAAATGTTT<br>CAACAATTAACCGGTAACAATTATTTTTTCTACTACGGTACCGTT<br>ATTTTCAAGTCAGTTGGCCTGGATGATTCCTTTGAAACATCCATT<br>GTCATTGGTGTAGTCAACTTTGCCTCCACTTTCTTTAGTTTGTGG<br>ACTGTCGAAAACCTTGGGACATCGTAAATGTTTACTTTTGGGCGCT<br>GCCACTATGATGGCTTGTATGGTCATCTACGCCTCTGTTGGTGT<br>ACTAGATTATATCCTCACGGTAAAAGCCAGCCATCTTCTAAAGGT<br>GCCGGTAACTGTATGATTGTCTTTACCTGTTTTTATATTTTCTGT<br>TATGCCACAACCTGGGCGCCAGTTGCCTGGGTCATCACAGCAGAA<br>TCATTCCCCTGAGAGTCAAGTCGAAATGTATGGCGTTGGCCTCT<br>GCTTCCAATTGGGTATGGGGTTCTTGATTGCATTTTTCACCCCA<br>TTCATCACATCTGCCATTAACCTTCTACTACGGTTATGTCTTCATG<br>GGCTGTTTGGTTGCCATGTTTTTTTATGTCTTTTCTTTGTTCCA<br>GAACTAAAGGCCTATCGTTAGAAGAAATTCAAGAATTATGGGAA<br>GAAGGTGTTTTACCTTGGAAATCTGAAGGCTGGATTCCTTCATCC<br>AGAAGAGGTAATAATTACGATTTAGAGGATTTACAACATGACGAC<br>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        | CACCATTAATTA <u>ACCCGGGGC</u> ACCTGTCACCTTTGGAA               |
| rDNA (R1)-over-R                                  | 222        | CGCGTAAGCTTTCTAGAGGCGCGCCAAGCTTTTACACTCTTG<br>ACCAGCGCA       |
| AB vector-MCS-R                                   | 223        | CCGCTGGTGGGTACCGATATCCCTGCAGGGAGCTCACGCGTA<br>AGCTTTCTAGAGGCG |
| rDNA(R3)-over-R                                   | 224        | CTGCAGGGATATCGGTACCCACCAGCGGCCGCAGGCCTTGG<br>GTGCTTGCTGGCGAA  |
| rDNA(R3)-over-R                                   | 225        | ACCTCTGCATGCGAATTCTTAAGACAAATAAAATTTATAGAG<br>ACTTGT          |
| rDNA(R2)-over-R                                   | 226        | GTCTTAAGAATTCGCATGCAGAGGTAGTTTCAAGGT                          |
| <i>Pac I</i> -rDNA(R2)-R                          | 227        | CACCATTAATTAATACGTATTCTCGCCGAGAAAACTT                         |

[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<br>GGTAAGGACTCTAAGAACCCAATGGCTTTCCACTACTACGACCCA<br>GAAAAGGTTATTATGGGTAAGCCAATGAAGGACTGGTTGAGATTC<br>GCTATGGCTTGGTGGCACACCTTGTGTGCTGAAGGTGGTGACCAA<br>TTCGGTGGTGGTACTAAGAAGTTCCCATGGAACAACGGTGCTGAC<br>GCTGTTGAAATTGCTAAGCAAAAGGCTGACGCTGGTTTCGAAATT<br>ATGCAAAAGTTGGGTATTCCATACTTCTGTTTCCACGACGTTGAC<br>TTGGTTTCTGAAGGTGCTTCTGTTGAAGAATACGAAGCTAACTTG<br>AAGGCTATTACCGACTACTTGGCTGTTAAGATGAAGGAAACCGGA |

| Table 15                                              |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-------------------------------------------------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description                                  | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|                                                       |            | ATTAAGTTGTTGTGGTCTACCGCTAACGTTTTTCGGTAACGGTAGA<br>TACATGAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTTGCT<br>AGAGCTATTGTTCAAATTAAGAACGCTATTGACGCTGGTATTAAG<br>TTGGGTGCTGAAAACACGTTTTTCTGGGGTGGTAGAGAAGGTTAC<br>ATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACACATG<br>GCTACCATGTTGACCATGGCTAGAGACTACGCTAGAGCTAAGGGT<br>TTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCATCT<br>AAGCACCAATACGACGTTGACACCGAAACCGTTATTGGTTTCTTG<br>AAGGCTCACAACTTGGACAAGGACTTCAAGGTTAACATTGAAGTT<br>AACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTGGCT<br>GTTGCTGTTGACAACAACATGTTGGGTCTATTGACGCTAACAGA<br>GGTGACTACCAAAACGGTTGGGACACCGACCAATTCCCAATTGAC<br>CAATACGAATTGGTTCAAGCTTGGATGGAAATTATTAGAGGTGGT<br>GGTTTGGGTACAGGTGGTACTAACTTCGACGCTAAGACCAGAAGA<br>AACTCTACCGACTTGGAAGACATTTTCATTGCTCACATTGCTGGT<br>ATGGACGCTATGGCTAGAGCTTTGGAATCTGCTGCTAAGTTGTTG<br>GAAGAATCTCCATACAAGGCTATGAAGGCTGCTAGATACGCTTCT<br>TTCGACAACGGTATTGGTAAGGACTTCGAAGACGGTAAGTTGACC<br>TTGGAACAAGCTTACGAATACGGTAAGAAGGTTGGTGAACCAAAG<br>CAAACCTCTGGTAAGCAAGAATTGTACGAAGCTATTGTTGCTATG<br>TACGCTTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Codon optimized DNA<br>encoding XI of SEQ<br>ID NO:58 | 239        | ATGGCTAAGGAATACTTCCCAGAAATTGGTAAGATTAAGTTCGAA<br>GGTAAGGACTCTAAGAACCCAATGGCTTTCCACTACTACGACGCT<br>GAAAAGGTTATTATGGGTAAGCCAATGAAGGAATGGTTGAGATTC<br>GCTATGGCTTGGTGGCACACCTTGTGTGCTGAAGGTGGTGACCAA<br>TTCGGTGGTGGTACTAAGAAGTTCCCATGGAACGAAGGTACTGAC<br>GCTGTTACCATTGCTAAGCAAAAGGCTGACGCTGGTTTCGAAATT<br>ATGCAAAAGTTGGGTTTCCCATACTTCTGTTTCCACGACATTGAC<br>TTGGTTTCTGAAGGTAACCTCTATTGAAGAATACGAAGCTAACTTG<br>CAAGCTATTACCGACTACTTGAAGGTTAAGATGGAAGAAACCGGA<br>ATTAAGTTGTTGTGGTCTACCGCTAACGTTTTTCGGTAACGGTAGA<br>TACATGAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTTGCT<br>AGAGCTATTGTTCAAATTAAGAACGCTATTGACGCTGGTATTAAG<br>TTGGGTGCTGAAAACACGTTTTTCTGGGGTGGTAGAGAAGGTTAC<br>ATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACACATG<br>GCTACCATGTTGACCATGGCTAGAGACTACGCTAGATCTAAGGGT<br>TTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCATCT<br>AAGCACCAATACGACGTTGACACCGAAACCGTTATTGGTTTCTTG<br>AAGGCTCACAACTTGGACAAGGACTTCAAGGTTAACATTGAAGTT<br>AACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTGGCT<br>GTTGCTGTTGACAACGGTATGTTGGGTCTATTGACGCTAACAGA<br>GGTGACTACCAAAACGGTTGGGACACCGACCAATTCCCAATTGAC<br>CAATACGAATTGGTTCAAGCTTGGATGGAAATTATTAGAGGTGGT<br>GGTTTGGGTACTGGTGGTACAACTTCGACGCTAAGACCAGAAGA<br>AACTCTACCGACTTGGAAGACATTTTCATTGCTCACATTTCTGGT<br>ATGGACGCTATGGCTAGAGCTTTGGAATCTGCTGCTAAGTTGTTG<br>GAAGAATCTCCATACTGTGCTATGAAGAAGGCTAGATACGCTTCT<br>TTCGACTCTGGTATTGGTAAGGACTTCGAAGACGGTAAGTTGACC<br>TTGGAACAAGCTTACGAATACGGTAAGAAGGTTGGTGAACCAAAG<br>CAAACCTCTGGTAAGCAAGAATTGTACGAAGCTATTGTTGCTATG<br>TACGCTTAA |

| Table 15                                        |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------------------------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description                            | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Codon optimized DNA encoding XI of SEQ ID NO:78 | 244        | ATGGCTAAGGAATATTTCCCATTCACCGGTAAGATTCCATTCGAA<br>GGTAAGGACTCTAAGAACGTTATGGCTTTCCACTATTATGAACCA<br>GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC<br>GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA<br>TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGGTGAC<br>GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT<br>ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC<br>TTGGTTGAAGAAGGTGACACCATTGAAGAATATGAAGCTAGAATG<br>AAGGCTATTACCGACTATGCTCAAGAAAAGATGAAGCAATTCCCA<br>AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTTCGGTAACAAG<br>AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT<br>GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT<br>AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT<br>TATATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACAC<br>ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG<br>GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA<br>TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC<br>TTGAAGGCTCACAACCTTGGACAAGGACTTCAAGGTTAACATTGAA<br>GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG<br>GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCATTGACGCTAAC<br>AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT<br>GACAACTATGAATTGACCCAAGCTATGTTGGAAATTATTAGAAAC<br>GGTGGTTTGGGTAACGGTGGAACCAACTTCGACGCTAAGATTAGA<br>AGAAACTCTACCGACTTGGAAGACTTGTTTCATTGCTCACATTTCT<br>GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGACATT<br>TTGGAAAACCTCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT<br>TCTTTCGACCAAGGTGTTGGTAAGGACTTCGAAGACGGTAAGTTG<br>ACCTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGGTGAACCA<br>AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT<br>TTGTATGCTAAGTAA |
| Codon optimized DNA encoding XI of SEQ ID NO:96 | 245        | ATGGCTAAGGAATATTTCCCATTCATTGGTAAGGTTCCATTCGAA<br>GGTACTGAATCTAAGAACGTTATGGCTTTCCACTATTATGAACCA<br>GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC<br>GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA<br>TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGCTGAC<br>GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT<br>ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC<br>TTGGTTGAAGAAGGTGAAACCGTTGCTGAATATGAAGCTAGAATG<br>AAGGTTATTACCGACTATGCTTTGGAAAAGATGCAACAATTCCCA<br>AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTTCGGTCACAAG<br>AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT<br>GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT<br>AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT<br>TATATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACAC<br>ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG<br>GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA<br>TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC<br>TTGAGGGCTCACGTTTGGACAAGGACTTCAAGGTTAACATTGAA<br>GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG<br>GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCATTGACGCTAAC<br>AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT                                                                                                                                                                                                                                                                                                                                                                                                                                |

| Table 15             |            |                                                                                                                                                                                                                                                                                                                                                                                                                           |
|----------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                  |
|                      |            | GACAACTATGAATTGACCCAAGCTATGATGGAAATTATTAGAAAC<br>GGTGGTTTGGGTAACGGTGGAAACCAACTTCGACGCTAAGATTAGA<br>AGAAACTCTACCGACTTGGAAGACTTGTTTCATTGCTCACATTTCT<br>GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGCTATT<br>TTGGAAGAATCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT<br>TCTTTCGACGAAGGTATTGGTAAGGACTTCGAAGACGGTAAGTTG<br>TCTTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGAAGAACCA<br>AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT<br>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        | CCCCATCGACAACCTACGAGCTCACT |
| N16PCR_R                                 | 229        | CAACTTGCCGTCCTCGAAGTCCTTG  |
| N05PCR_F                                 | 230        | CGAGCCTGAGAAGGTCGTGATGGGA  |
| N05PCR_R                                 | 231        | TACGTCGAAGTCGGGGTTGGTAGAA  |
| N08PCR_F                                 | 240        | TACTTGGCTGTTAAGATGAAG      |
| 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; TTAATTAAGTTAATTACCTTTTTTGCAGGGCATATTTATGGTGAAGAATAAGTTTTGACCATC AAAGAAGGTTAATGTGGCTGTGGTTTCAGGGTCCATAAAGCTTTTCAATTCATCATTTTTTTT TTATTCTTTTTTTTGATTCCGGTTTCCTTGAAATTTTTTTTGATTTCGGTAATCTCCGAACAGAAG GAAGAACGAAGGAAGGAGCACAGACTTAGATTGGTATATATACGCATATGTAGTGTTGAAG AAACATGAAATTGCCCAGTATTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAAACG AAGATAAATCATGTCGAAAGCTACATATAAGGAACGTGCTGCTACTCATCCTAGTCCTGTTG CTGCCAAGCTATTTAATATCATGCACGAAAAGCAAACAACTTGTGTGCTTCATTGGATGTT CGTACCACCAAGGAATTACTGGAGTTAGTTGAAGCATTAGGTCCCAAATTTGTTTACTAAA AACACATGTGGATATCTTGACTGATTTTTCCATGGAGGGGCACAGTTAAGCCGCTAAAGGCAT TATCCGCCAAGTACAATTTTTTACTCTTCGAAGACAGAAAATTTGCTGACATTGGTAATACA GTCAAATTGCAGTACTCTGCGGGTGTATACAGAATAGCAGAATGGGCAGACATTACGAATG CACACGGTGTGGTGGGCCCAGGTATTGTTAGCGGTTTGAAGCAGGCGGCAGAAGAAGTAAC AAAGGAACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCATGCAAGGGCTCCCTAGCTACTG GAGAATATACTAAGGGTACTGTTGACATTGCGAAGAGCGACAAAGATTTTGTTATCGGCTTT ATTGCTCAAAGAGACATGGGTGGAAGAGATGAAGGTTACGATTGGTTGATTATGACACCCG GTGTGGGTTTAGATGACAAGGGAGACGCATTGGGTCAACAGTATAGAACCGTGGATGATGT GGTCTCTACAGGATCTGACATTATTATTGTTGGAAGAGGACTATTTGCAAAGGGAAGGGATG CTAAGGTAGAGGGTGAACGTTACAGAAAAGCAGGCTGGGAAGCATATTTGAGAAGATGCGG CCAGCAAAACTAAAAAACTGTATTATAAGTAAATGCATGTATACTAAACTCACAAATTAGA GCTTCAATTTAATTATATCAGTTATTACCCGGGAATCTCGGTCGTAATGATTTTTTATAATGAC 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</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2                               |
| pBPB008             | yBPA136       | <i>MATalpha, ura3, adh2</i> :: TAL1-XKS1, <i>pho13</i> :: TKL1-XKS1, <i>gre3</i> :: RPE1-RKI1 and YLR388.5:: GAL2                           |
| BD31328             | 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:96)      |
| BD31336             | 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:96)  |
| BD31526             | 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:78)      |
| BD31527             | 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:78)  |
| BD34364             | 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:238)     |
| BD34365             | 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:238) |

| 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<br>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<br>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<br>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<br>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.

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

| <b>Table 20</b>                  |           |
|----------------------------------|-----------|
| <b>Vitamin mixture (1000x)</b>   |           |
| <b>Components</b>                | <b>mM</b> |
| ZnSO <sub>4</sub>                | 100       |
| H <sub>3</sub> BO <sub>3</sub>   | 24        |
| KI                               | 1.8       |
| MnSO <sub>4</sub>                | 20        |
| CuSO <sub>4</sub>                | 10        |
| Na <sub>2</sub> MoO <sub>4</sub> | 1.5       |
| CoCl <sub>2</sub>                | 1.5       |
| FeCl <sub>3</sub>                | 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)).

| <b>Table 21</b>    |              |
|--------------------|--------------|
| <b>Clean sugar</b> |              |
| <b>Component</b>   | <b>Conc.</b> |
| 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        | ATGGCTAAGGAATATTTCCCATTCACCGGTAAGATTCCATTCGAA<br>GGTAAGGACTCTAAGAACGTTATGGCTTTCCACTATTATGAACCA<br>GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC<br>GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA<br>TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGGTGAC<br>GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT<br>ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC<br>TTGGTTGAAGAAGGTGACACCATTGAAGAATATGAAGCTAGAATG<br>AAGGCTATTACCGACTATGCTCAAGAAAAGATGAAGCAATTCCCA<br>AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTTCGGTAACAAG<br>AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT<br>GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT<br>AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT<br>TATATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACAC<br>ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG<br>GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA<br>TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC<br>TTGAAGGCTCACAACTTGGACAAGGACTTCAAGGTTAACATTGAA<br>GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG<br>GCTTGTGCTGTTGACGCTGGTATGTTGGGTTCATTGACGCTAAC<br>AGAGGTGACGCTCAAAACGTTGGGACACCGACCAATTCCCAATT<br>GACAACTATGAATTGACCCAAGCTATGTTGGAAATTATTAGAAAC<br>GGTGGTTTGGGTAACGGTGAACCAACTTCGACGCTAAGATTAGA<br>AGAACTCTACCGACTTGAAGACTTGTTTCATTGCTCACATTTCT<br>GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGACATT<br>TTGGAAACTCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT<br>TCTTTCGACCAAGGTGTTGGTAAGGACTTCGAAGACGGTAAGTTG |

| Table 22                                                    |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------------------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description                                        | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|                                                             |            | ACCTTGAACAAGTTTATGAATATGGTAAGAAGGTTGGTGAACCA<br>AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT<br>TTGTATGCTAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Boles codon optimized<br>DNA encoding XI of<br>SEQ ID NO:96 | 245        | ATGGCTAAGGAATATTTCCCATTCATTGGTAAGGTTCCATTGAA<br>GGTACTGAATCTAAGAACGTTATGGCTTTCCACTATTATGAACCA<br>GAAAAGGTTGTTATGGGTAAGAAGATGAAGGACTGGTTGAAGTTC<br>GCTATGGCTTGGTGGCACACCTTGGGTGGTGCTTCTGCTGACCAA<br>TTCGGTGGTCAAACCAGATCTTATGAATGGGACAAGGCTGCTGAC<br>GCTGTTCAAAGAGCTAAGGACAAGATGGACGCTGGTTTCGAAATT<br>ATGGACAAGTTGGGTATTGAATATTTCTGTTTCCACGACGTTGAC<br>TTGGTTGAAGAAGGTGAAACCGTTGCTGAATATGAAGCTAGAATG<br>AAGGTTATTACCGACTATGCTTTGGAAAAGATGCAACAATTCCCA<br>AACATTAAGTTGTTGTGGGGTACTGCTAACGTTTTCTGGTCCACAAG<br>AGATATGCTAACGGTGCTTCTACCAACCCAGACTTCGACGTTGTT<br>GCTAGAGCTATTGTTCAAATTAAGAACGCTATTGATGCTACCATT<br>AAGTTGGGTGGTACTAACTATGTTTTCTGGGGTGGTAGAGAAGGT<br>TATATGTCTTTGTTGAACACCGACCAAAAGAGAGAAAAGGAACAC<br>ATGGCTACCATGTTGACCATGGCTAGAGACTATGCTAGAGCTAAG<br>GGTTTCAAGGGTACTTTCTTGATTGAACCAAAGCCAATGGAACCA<br>TCTAAGCACCAATATGACGTTGACACCGAAACCGTTATTGGTTTC<br>TTGAGGGCTCACGTTTTGGACAAGGACTTCAAGGTTAACATTGAA<br>GTTAACCACGCTACCTTGGCTGGTCACACCTTCGAACACGAATTG<br>GCTTGTGCTGTTGACGCTGGTATGTTGGGTCTATTGACGCTAAC<br>AGAGGTGACGCTCAAAACGGTTGGGACACCGACCAATTCCCAATT<br>GACAACTATGAATTGACCCAAGCTATGATGGAAATTATTAGAAAC<br>GGTGGTTTGGGTAACGGTGGAACCAACTTCGACGCTAAGATTAGA<br>AGAAACTCTACCGACTTGGAAGACTTGTTTCATTGCTCACATTTCT<br>GGTATGGACGCTATGGCTAGAGCTTTGATGAACGCTGCTGCTATT<br>TTGGAAGAATCTGAATTGCCAGCTATGAAGAAGGCTAGATATGCT<br>TCTTTCGACGAAGGTATTGGTAAGGACTTCGAAGACGGTAAGTTG<br>TCTTTGGAACAAGTTTATGAATATGGTAAGAAGGTTGAAGAACCA<br>AAGCAAACCTCTGGTAAGCAAGAAAAGTATGAAACCATTGTTGCT<br>TTGTATGCTAAGTAA |
| Boles codon optimized<br>DNA encoding XI of<br>SEQ ID NO:38 | 246        | ATGAAGGAAATTTTCCCAAACATTCCAGAAATTAAGTTCGAAGGT<br>AAGGACTCTAAGAACCCATTTCGCTTTCCACTATTATAACCCAGAC<br>CAAATTATTTTGGGTAAGCCAATGAAGGAACACTTGCCATTCGCT<br>ATGGCTTGGTGGCACAACCTGGGTGCTACCGGTGTTGACATGTTT<br>GGTGCTGGTCCAGCTGACAAGTCTTTCGGTGCTAAGGTTGGTACT<br>ATGGAACACGCTAAGGCTAAGGTTGACGCTGGTTTCGAATTCATG<br>AAGAAGTTGGGTATTAGATATTTCTGTTTCCACGACGTTGACTTG<br>GTTCCAGAATGTGCTGACATTAAGGACACCAACAAGGAATTGGAC<br>GAAATTTCTGACTATATTTTGGAAAAGATGAAGGGTACTGACATT<br>AAGTGTTTGTGGGGTACTGCTAACATGTTCTCTAACCCAAGATTC<br>TGTAACGGTGCTGGTTCTACCAACTCTGCTGACGTTTTCTGCTTTC<br>GCTGCTGCTCAAGTTAAGAAGGCTTTGGACATTACCGTTAAGTTG<br>GGTGGTAGAGGTTATGTTTTCTGGGGTGGTAGAGAAGGTTATGAA<br>ACCTTGTTGAACACCGACGTTAAGTTCGAACAAGAAAACATTGCT<br>AGATTGATGAAGATGGCTGTTGAATATGGTAGATCTATTGGTTTC<br>AAGGGTGACTTCTATATTGAACCAAAGCCAAAGGAACCAATGAAG<br>CACCAATATGACTTCGACGCTGCTACCGCTATTGGTTTCTTGAGG<br>GCTCACGGTTTGGACAAGGACTTCAAGTTGAACATTGAAGCTAAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| Table 22                                                |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------------------------------------------------|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description                                    | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|                                                         |            | CACGCTACCTTGGCTGGTCACACCTTCCAACACGACTTGAGAATT<br>TCTGCTATTAACGGTATGTTGGGTTCTATTGACGCTAACCAAGGT<br>GACATGTTGTTGGGTTGGGACACCGACGAATTCCCATTGACGTT<br>TATTCTGCTACCCAATGTATGTATGAAGTTTTGAAGAACGGTGGT<br>TTGACCGGTGGTTTTCAACTTCGACTCTAAGACCAGAAGACCATCT<br>TATACCATGGAAGACATGTTCTTGGCTTATATTTTGGGTATGGAC<br>ACCTTCGCTTTGGGTTTGATTAAAGGCTGCTCAAATTATTGAAGAC<br>GGTAGAATTGACCAATTCATTGAAAAGAAGTATTCTTCTTTCAGA<br>GAAACCGAAATTGGTCAAAAGATTTTGAACAACAAGACCTCTTTG<br>AAGGAATTGTCTGACTATGCTTGTAAGATGGGTGCTCCAGAATTG<br>CCAGGTTCTGGTAGACAAGAAATGTTGGAAGCTATTGTTAACGAC<br>GTTTTGTTTCGGTAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:78 | 247        | ATGGCTAAGGAATACTTTCCATTACCGGAAAGATACCATTTGAA<br>GGTAAAGATTCTAAAAACGTAATGGCTTTTCATTATTACGAACCA<br>GAAAAAGTTGTTATGGGCAAAAAGATGAAAGATTGGTTGAAATTT<br>GCGATGGCTTGGTGGCATACTCGGGGGAGCTTCCGCTGATCAA<br>TTTGGCGGACAAACCAGATCATACGAATGGGATAAAGCAGGCGAT<br>GCCGTGCAGAGAGCAAAGGATAAAATGGATGCTGGTTTCGAAATT<br>ATGGATAAGCTAGGTATCGAATACTTCTGCTTCCATGACGTCGAT<br>TTGGTTGAAGAGGGCGATACTATCGAGGAATACGAGGCGAGAATG<br>AAGGCTATAACAGACTACGCCCAGGAGAAAATGAAACAATTTCT<br>AACATCAAATTACTCTGGGGTACTGCCAATGTGTTTGGTAACAAA<br>AGATACGCAAACGGGGCTTCAACTAATCCTGACTTCGATGTTGTT<br>GCAAGAGCCATTGTTCAAATCAAAAACGCGATAGACGCTACTATT<br>AACTAGGTGGCACGAATTACGTCTTTTGGGGTGGAAGGGAAGGT<br>TACATGTCTCTGCTTAATACAGATCAGAAGAGAGAGAAGGAACAC<br>ATGGCAACAATGCTCACTATGGCCCGTGACTACGCAAGAGCAAAA<br>GGTTTTAAGGGCACTTTCCTTATCGAACCAAAGCCTATGGAACCA<br>TCAAAACACCAATATGATGTTGACACAGAACTGTGATCGGCTTT<br>TTGAAAGCTCATAACTTGGACAAGGATTTCAAAGTAAACATTGAA<br>GTTAATCATGCTACACTAGCAGGACACACATTTGAACACGAACTG<br>GCCTGTGCGGTAGATGCAGGGATGCTGGGTTCATCGACGCTAAT<br>AGAGGGGATGCTCAAAATGGTTGGGATACCGATCAATTTCCAATC<br>GACAATTACGAATTAACACAAGCTATGTTGGAGATTATTAGAAAT<br>GGAGGTTTGGGTAATGGGGGTACAACTTCGATGCTAAGATTTCGT<br>CGAAATTCACAGACTTAGAAGATTTGTTTCATTGCGCATATATCT<br>GGTATGGATGCTATGGCCAGAGCATTGAATGCCGCTGACATC<br>TTAGAAAACAGTGAACCTCCAGCAATGAAAAAGGCCAGATATGCC<br>TCTTTCGATCAAGGTGTAGGAAAAGATTTTGAGGACGGCAAGTTG<br>ACTTTAGAACAAGTCTATGAATACGGTAAAAAGGTGCGGCGAACCT<br>AAGCAAACCAGCGGAAAGCAAGAGAAATACGAGACTATCGTGGCT<br>CTTTATGCAAAATAA |
| DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:96 | 248        | ATGGCCAAGGAGTACTTCCCTTTTATCGGCAAGGTCCCATTTGAA<br>GGGACAGAATCCAAAAACGTCATGGCTTTTCACTACTATGAACCT<br>GAGAAGGTAGTTATGGGTAAAAAGATGAAAGATTGGTTGAAGTTT<br>GCAATGGCATGGTGGCATACTTGGGTGGGGCCTCTGCTGATCAA<br>TTTGGAGGACAACTAGATCATACGAATGGGATAAAGCAGCTGAT<br>GCCGTTCAAAGAGCCAAAGATAAAATGGATGCCGGGTTCGAAATC<br>ATGGACAAATTGGGTATCGAATATTTCTGCTTCCATGATGTAGAC<br>CTTGTTGAGGAGGGTGAAACCGTCGCTGAATATGAGGCGAGAATG<br>AAGGTTATTACGGATTACGCACTAGAAAAGATGCAGCAGTTTCCA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

| Table 22                                                |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------------------------------------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sequence Description                                    | SEQ ID NO: | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|                                                         |            | AACATAAAACTATTGTGGGGTACTGCTAATGTTTTTCGGACATAAA<br>CGTTACGCTAACGGAGCTTCCACTAATCCAGACTTTGATGTTGTC<br>GCGAGAGCTATCGTTCAAATCAAAAATGCAATCGATGCTACAATT<br>AAGTTAGGAGGGACAAATTACGTGTTCTGGGGTGGTAGAGAAGGT<br>TACATGAGCCTGCTTAATACAGATCAAAAGAGAGAAAAGGAGCAC<br>ATGGCAACAATGCTAACAATGGCTAGAGATTATGCCCCGAGCTAAG<br>GGCTTCAAAGGCACTTTTTCTGATAGAACCCTAAACCAATGGAACCA<br>TCTAAACACCAATACGATGTAGACACCGAAACTGTAATAGGCTTC<br>CTTCGTGCACATGGTTTGGATAAAGATTTTAAGGTGAACATTGAA<br>GTGAATCATGCTACTTTAGCCGGTCACACTTTTGAACATGAATTA<br>GCATGTGCTGTTGATGCGGGAATGTTGGGTTCTATCGATGCCAAC<br>AGAGGCGACGCCCAAAATGGTTGGGACACAGACCAGTTTCCTATT<br>GACAATTACGAACTCACCCAAGCTATGATGGAAATTATCAGGAAT<br>GGGGGACTGGGAAATGGTGGTACGAACTTTGATGCGAAGATAAGG<br>AGAAACTCTACTGACTTAGAAGATTTGTTTATAGCACATATTTCA<br>GGTATGGACGCTATGGCAAGAGCTTTAATGAATGCCGCAGCAATC<br>TTGGAGGAAAGTGAACCTCCCAGCTATGAAAAAGGCAAGATACGCA<br>AGTTTTGATGAGGGTATTGGCAAAGACTTCGAAGATGGTAAACTA<br>TCTTTAGAACAAGTGTACGAGTATGGCAAAAAGGTAGAGGAACCA<br>AAACAAACATCAGGCAAACAAGAGAAATATGAAACAATTGTCGCT<br>CTTTACGCGAAGTAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| DNA 2.0 codon optimized DNA encoding XI of SEQ ID NO:38 | 249        | ATGAAGGAAATCTTCCCTAACATCCCAGAGATCAAATTCGAAGGC<br>AAAGACTCTAAAAATCCATTTGCCTTCCACTATTACAACCCAGAC<br>CAGATCATTTTAGGTAAACCAATGAAGGAGCACTTGCCATTTGCT<br>ATGGCTTGGTGGCATAATCTAGGCGCCACTGGTGTGATATGTTT<br>GGTGCAGGCCCTGCGGACAAATCTTTCGGAGCTAAAGTAGGAACT<br>ATGGAACATGCAAAAGCGAAAGTTGATGCTGGGTTTGAGTTCATG<br>AAGAAATTAGGAATCAGATATTTCTGCTTTCATGATGTTGACTTG<br>GTTCCCTGAGTGTGCTGACATTAAGGATACAAACAAGGAAC TTGAT<br>GAAATCTCTGACTACATTTTGGAAAAGATGAAAGGTACTGACATA<br>AAGTGTTTGTGGGGCACGGCTAATATGTTTTCCAATCCAAGATTT<br>TGTAACGGCGCTGGCTCAACTAATT CAGCAGATGTCTTTGCATTC<br>GCTGCTGCACAAGTCAAGAAAGCACTTGACATTACAGTCAAAC TG<br>GGTGGGAGAGGATACGTTTTTCTGGGGTGGTAGAGAAGGCTACGAA<br>ACATTGTTGAATACAGACGTTAAGTTTGAACAAGAGAATATTGCA<br>AGGTTAATGAAAATGGCAGTGG AATATGGGCGTTCTATAGGTTTT<br>AAAGGTGATTTCTACATTGAGCCAAAACCAAAGGAACCTATGAAA<br>CATCAATACGATTTTCGATGCCGCAACAGCAATAGGTTTCCTTAGA<br>GCCCACGGGTGGATAAAGACTTTAAGCTCAATATCGAAGCCAAC<br>CACGCAACACTTG CAGGCCATACATTTCAACATGATCTTAGAATA<br>TCTGCTATTAACGGAATGCTCGGCTCAATTGATGCCAATCAGGGT<br>GATATGCTACTAGGTTGGGATACTGATGAGTTTCCATTTGATGTA<br>TACTCCGCTACACAATGCATGTATGAGGTGTTGAAAAATGGTGGT<br>CTGACCGGTGGCTTCAACTTCGATAGTAAGACCAGACGTCCTTCA<br>TACACTATGGAAGATATGTTTCTGGCGTATATCTTAGGTATGGAC<br>ACATTTGCTTTAGGTCTAATCAAAGCCGCTCAAATCATTGAAGAT<br>GGCAGAATTGACCAGTTTATAGAAAAGAAATACTCCAGTTTTCGA<br>GAAACCGAAATCGGACAAAAGATTCTCAATAACAAAAC TTCA TTG<br>AAGGAATTATCTGATTACGCCTGTAAGATGGGTGCGCCAGAATTA<br>CCTGGAAGCGGTAGACAAGAGATGCTTGAAGCTATCGTGAATGAT<br>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 OD<sub>600</sub> = 0.2. Samples were incubated overnight at about 30°C and about 220 rpm. Cultures were harvested when OD<sub>600</sub> 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<br>SEQ ID NO. | Nucleic Acid<br>SEQ ID NO: | Codon optimization method | SA, pH 7.5<br>(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<br>SEQ ID NO.              | Nucleic<br>Acid<br>SEQ ID<br>NO: | Codon optimization method                 | SA, pH<br>7.5<br>(U/mg) | % activity as<br>compared to<br><i>Orpinomyces</i><br>XI |
| NA                                 | NA                               | Host negative control (no recombinant XI) | 0.005                   | 7%                                                       |
| Genbank Accession No.<br>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 (iv) 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  $\beta$ -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.

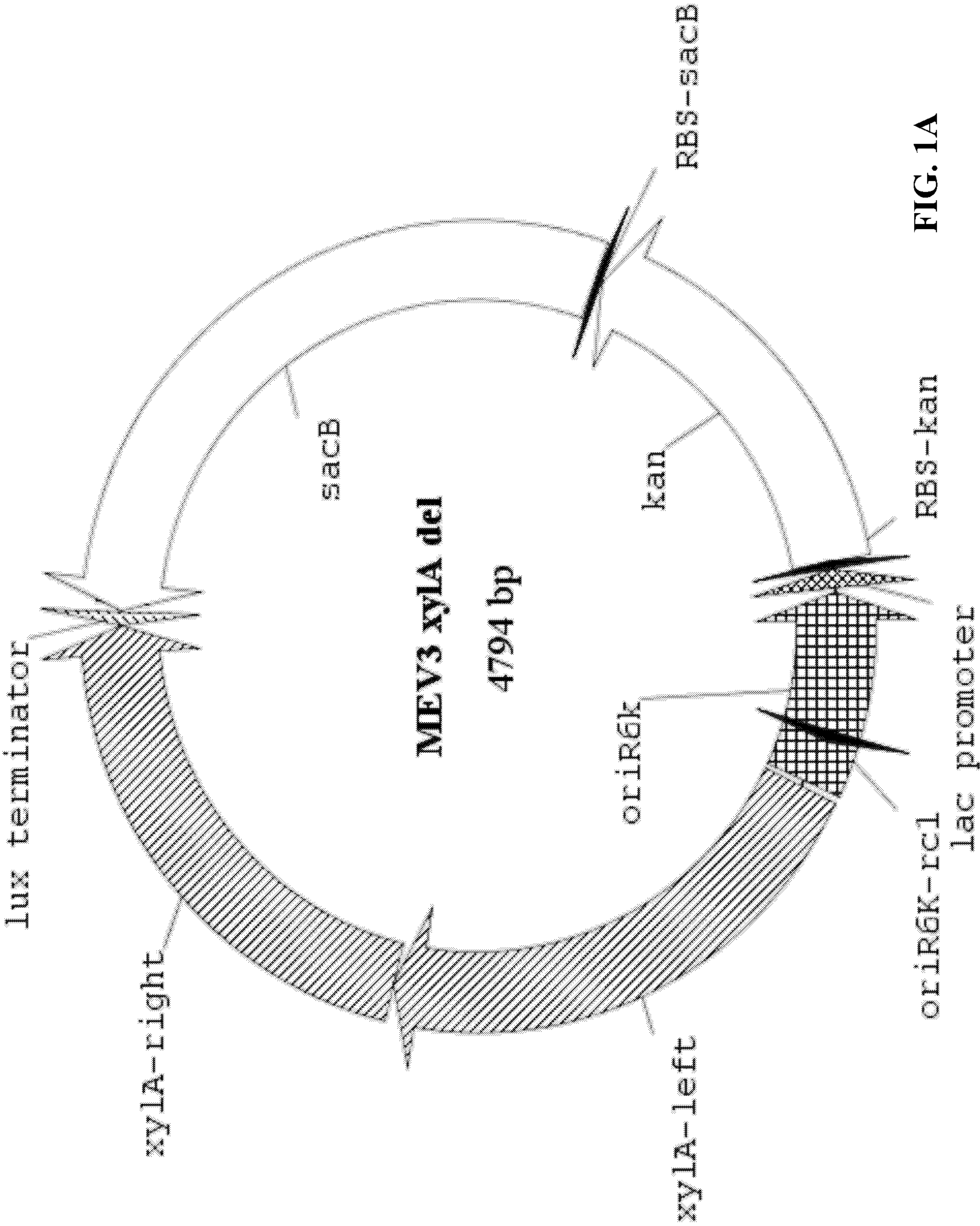
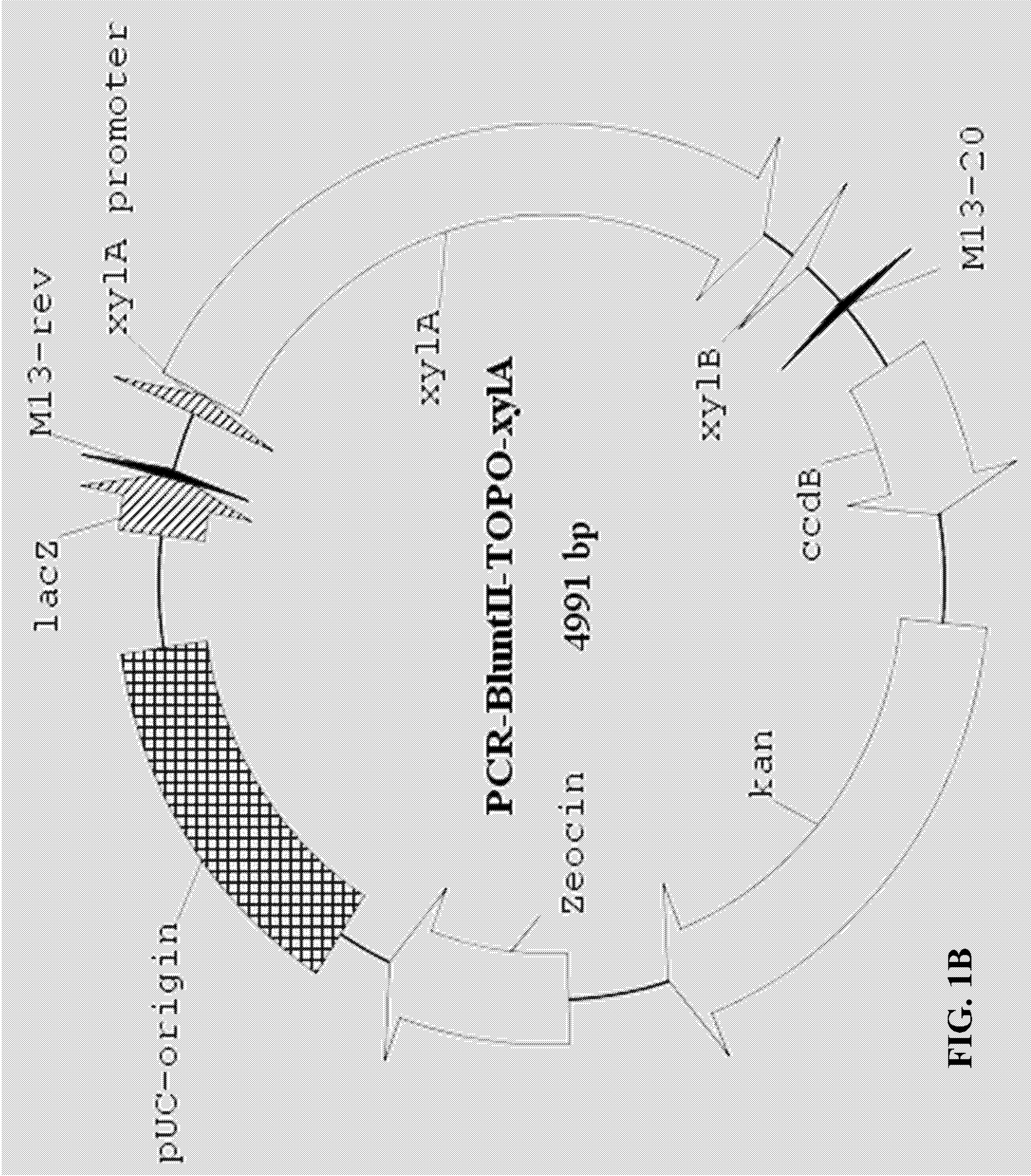


FIG. 1A



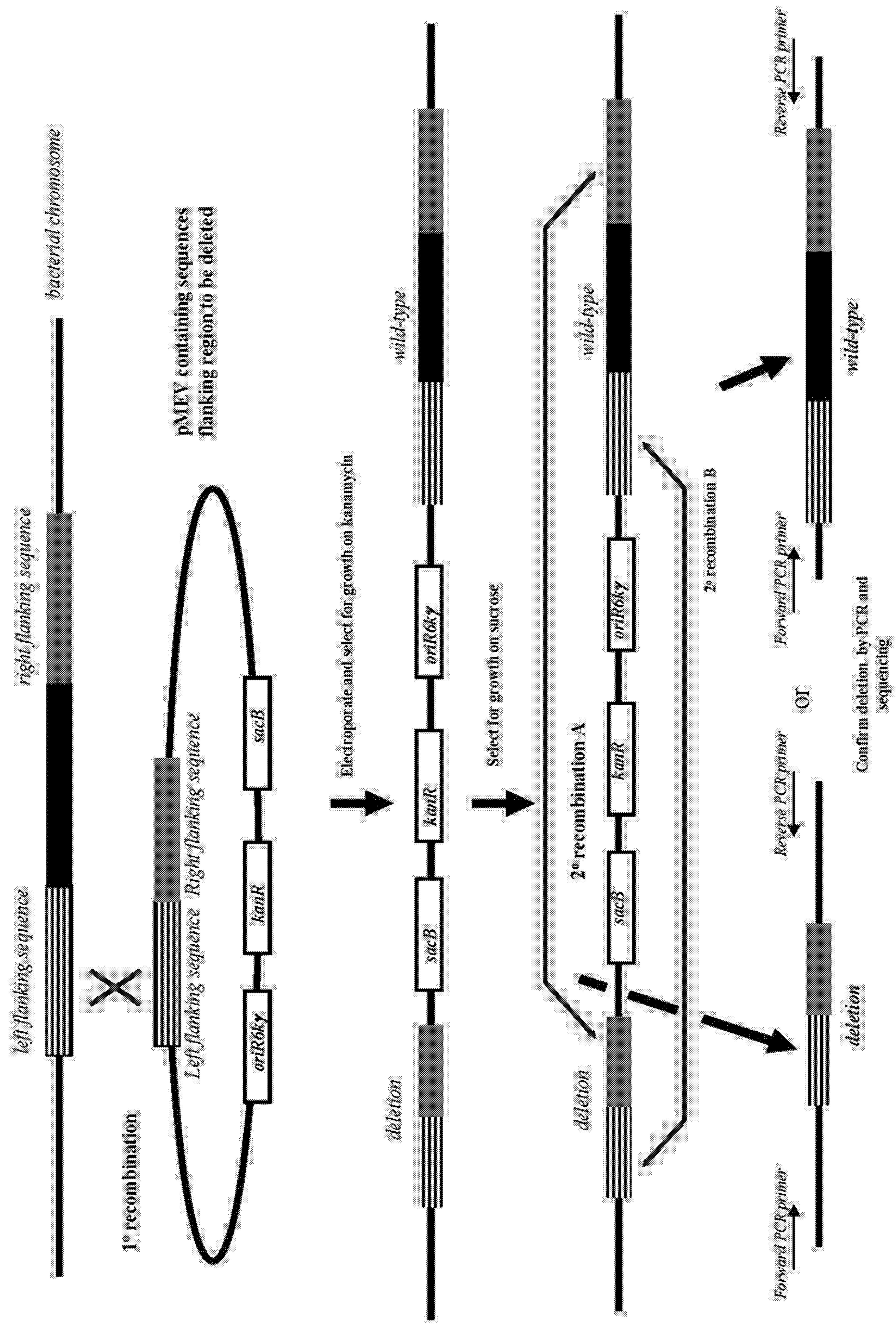


FIG. 2

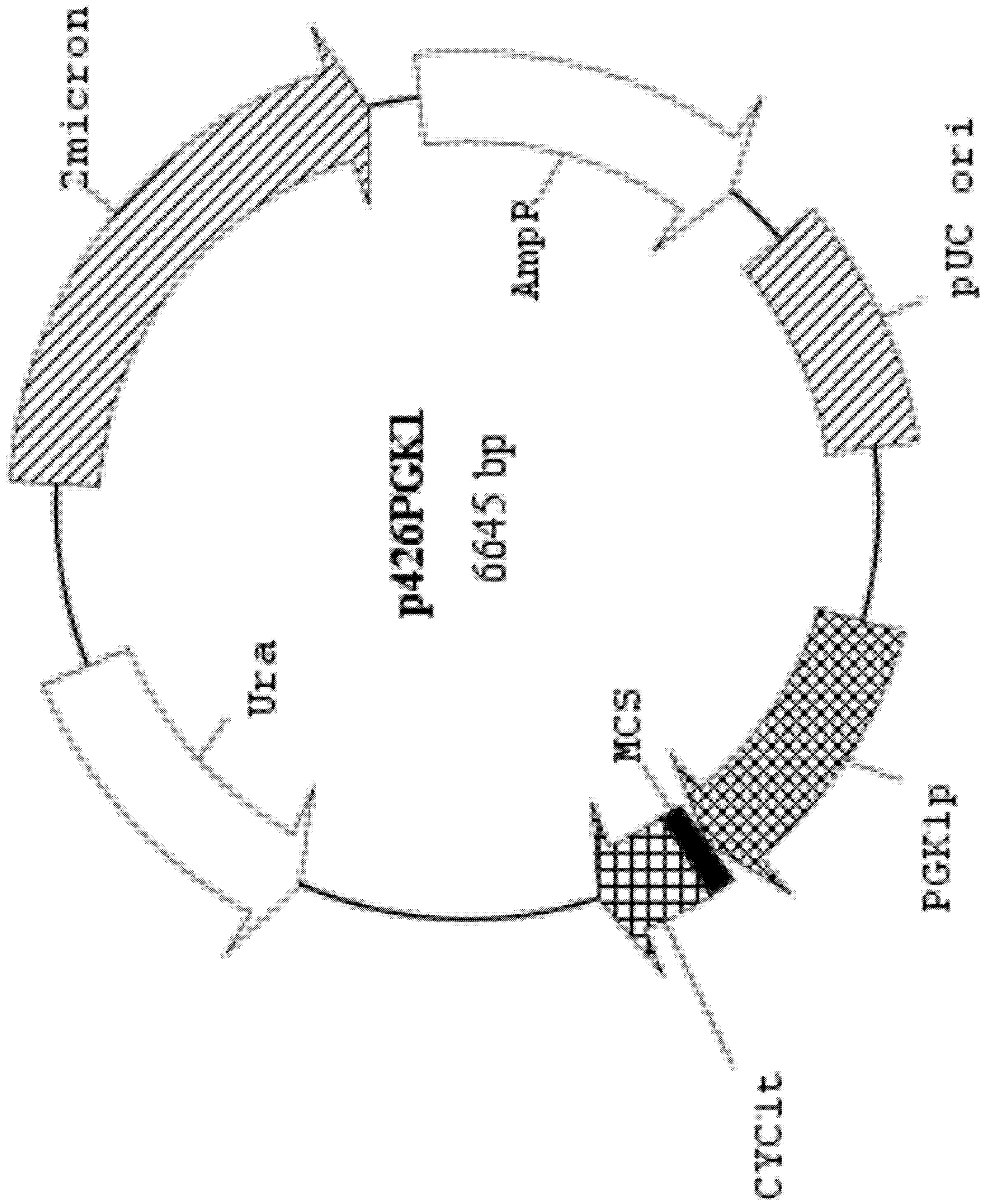


FIG. 3

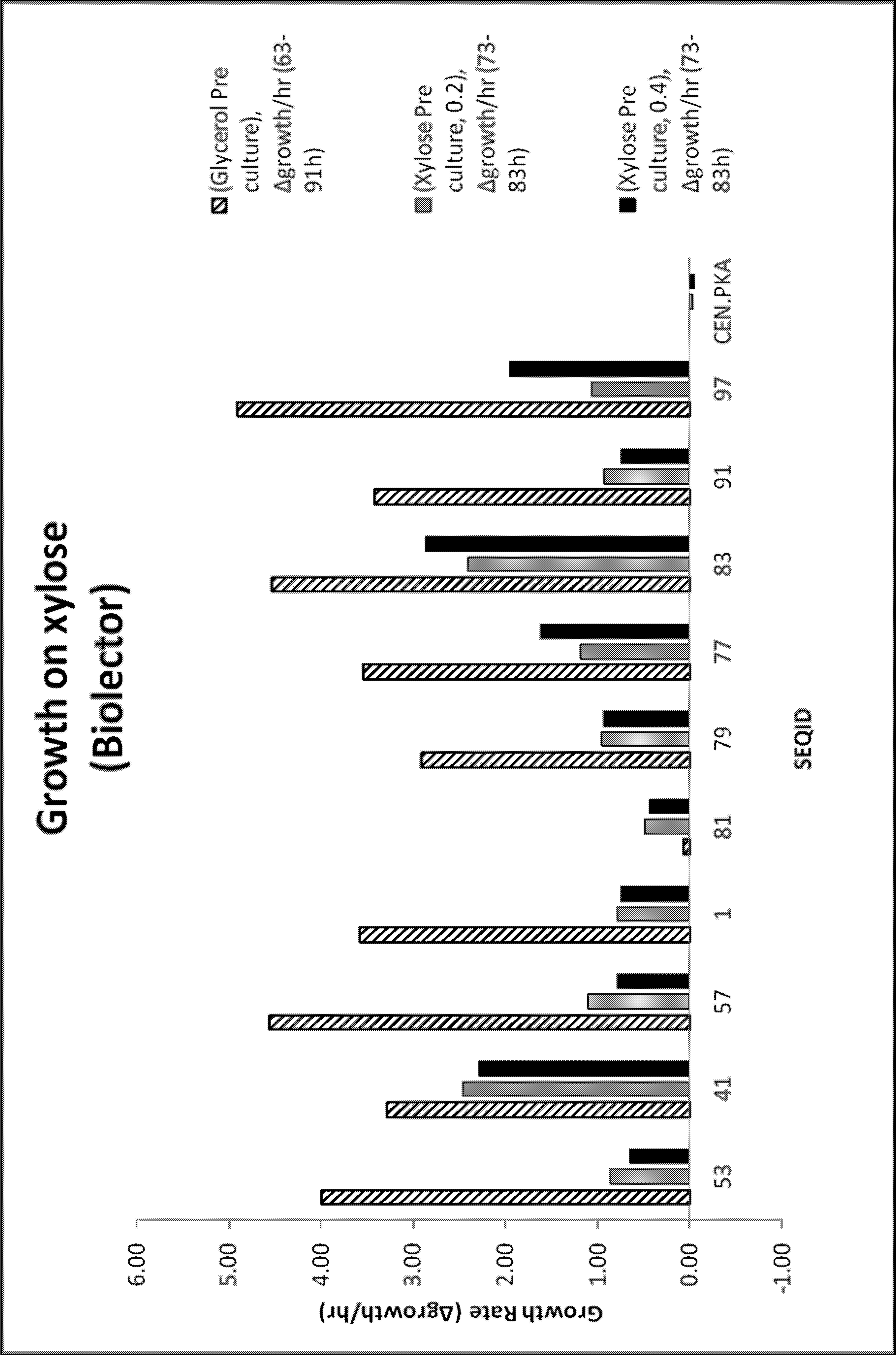


FIG. 4

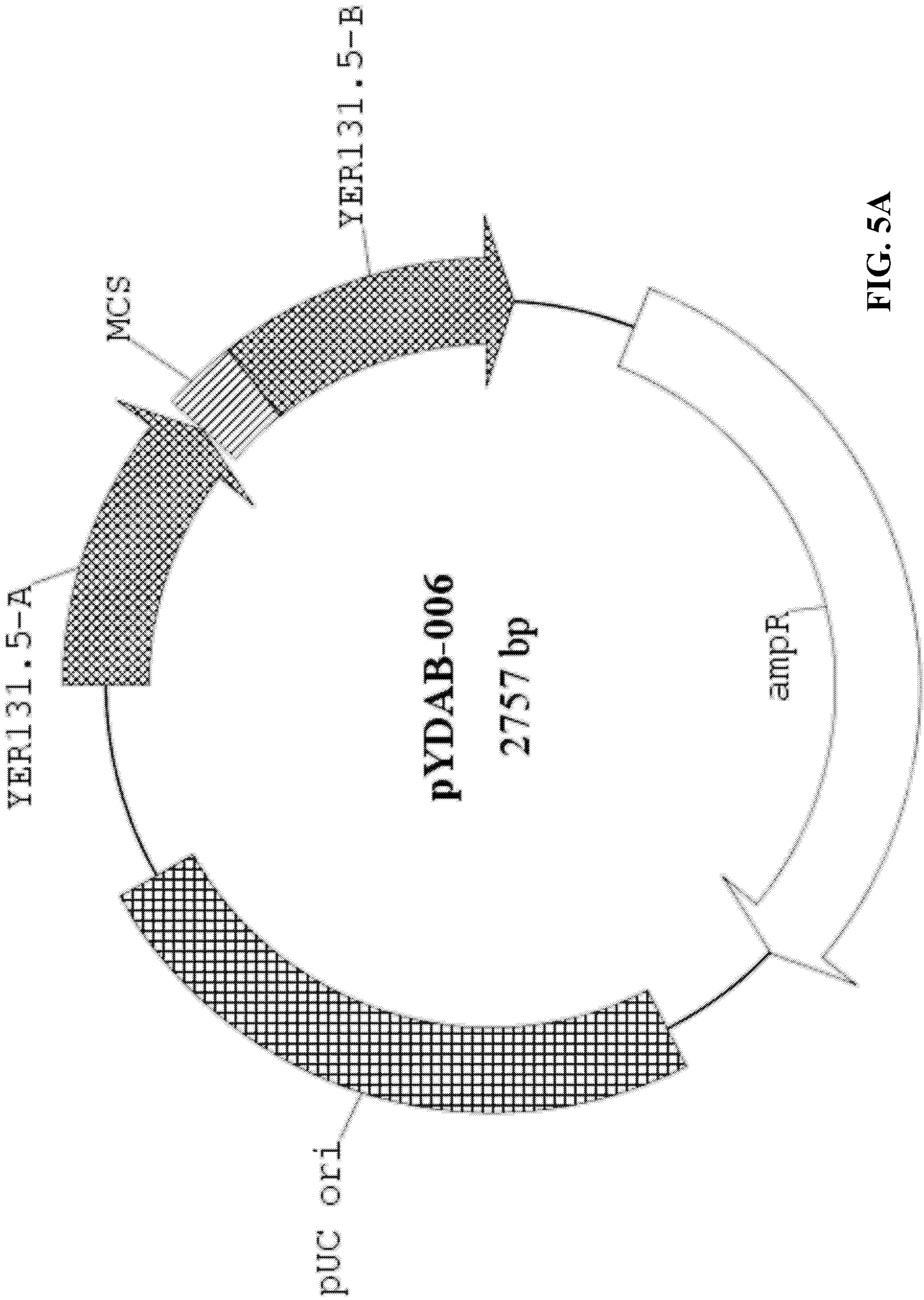
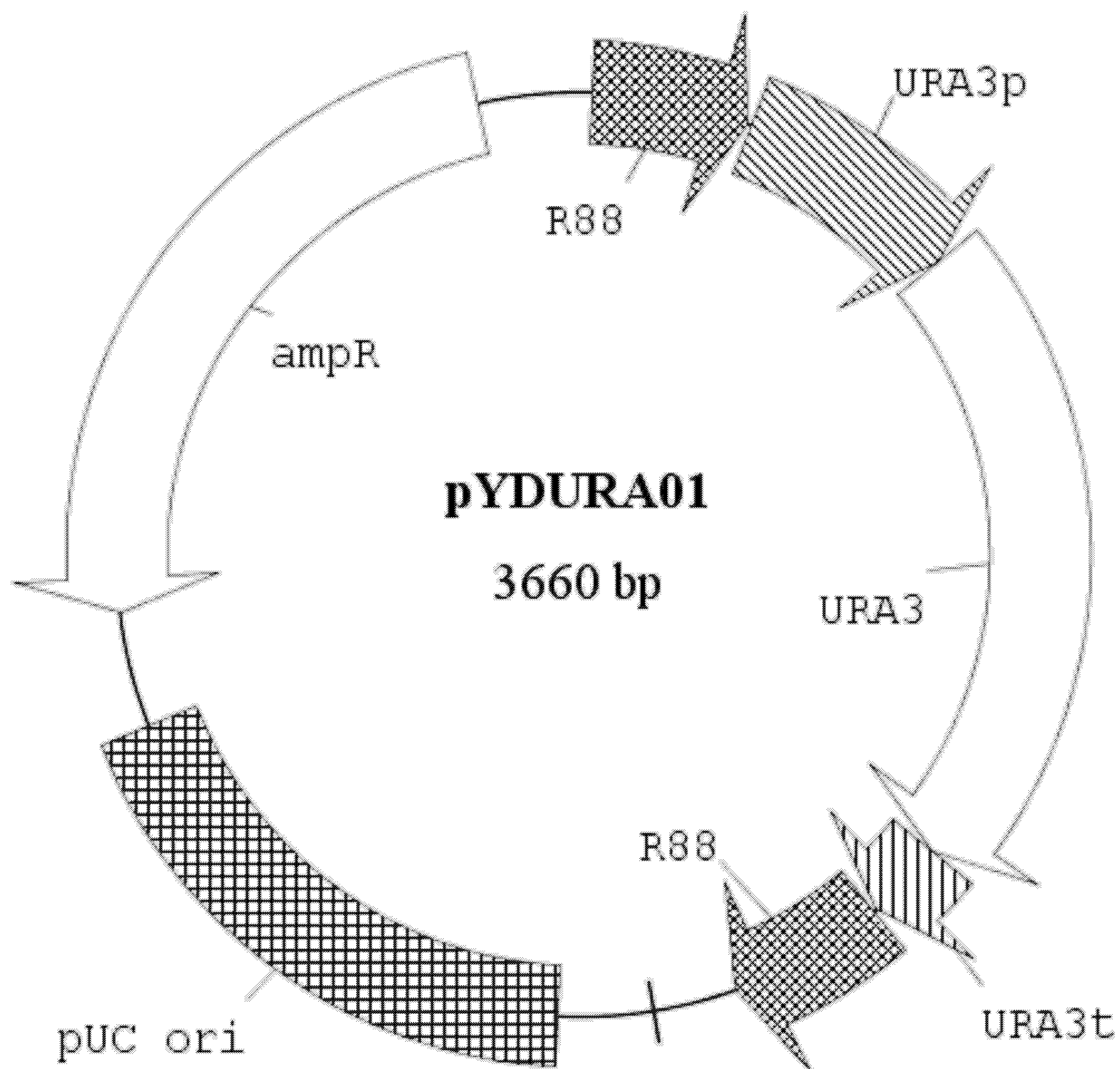


FIG. 5A

**FIG. 5B**

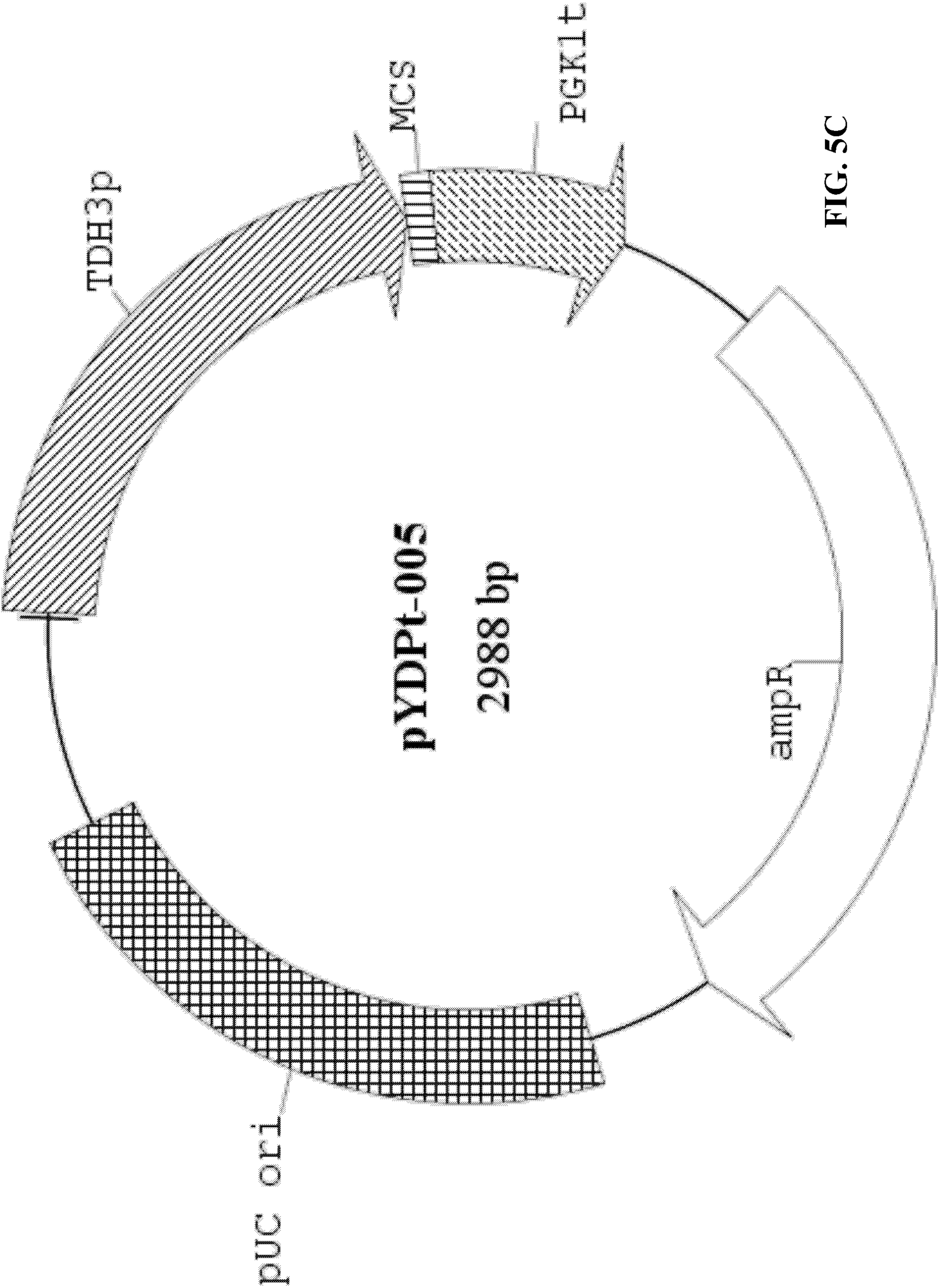


FIG. 5C

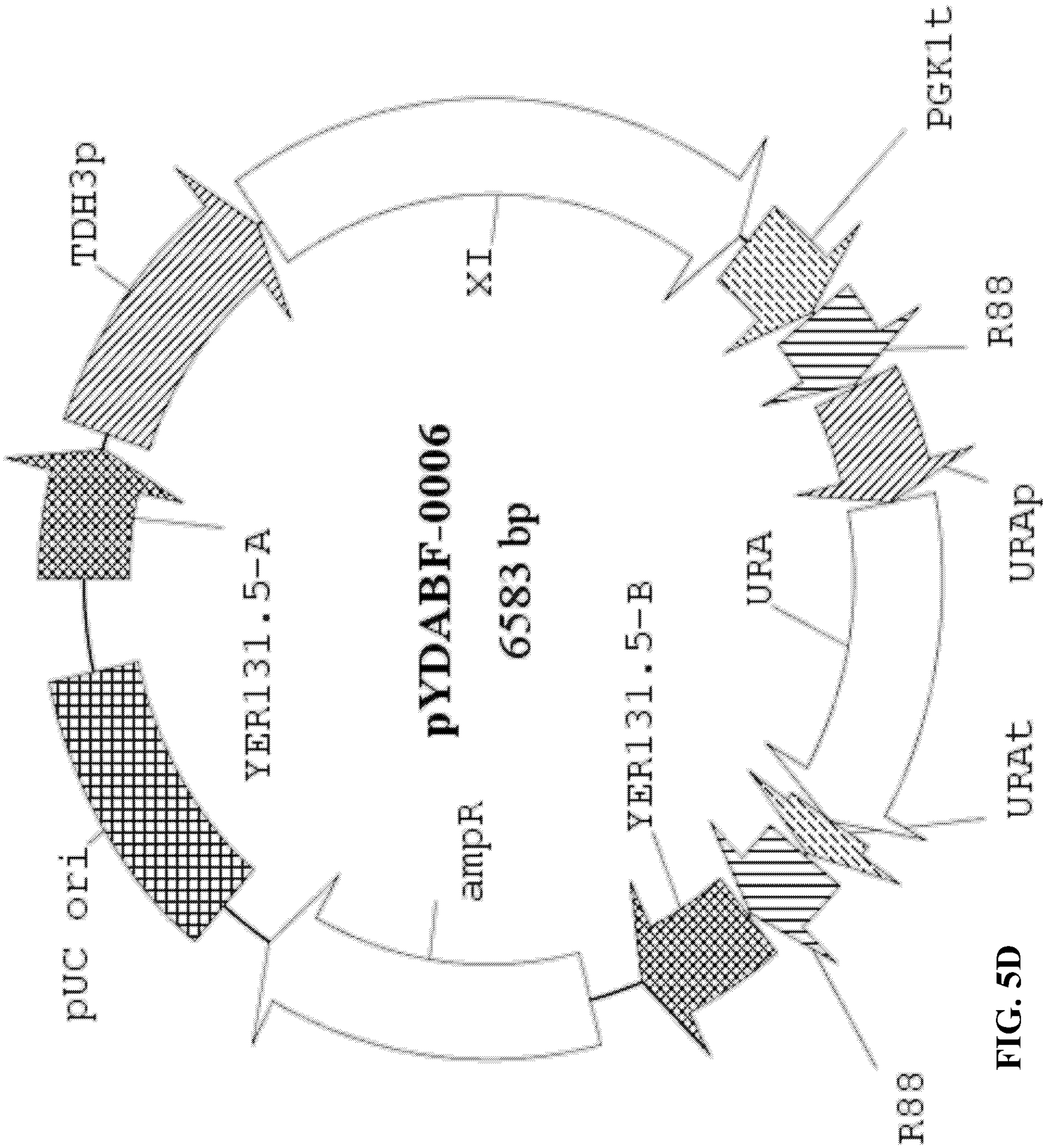
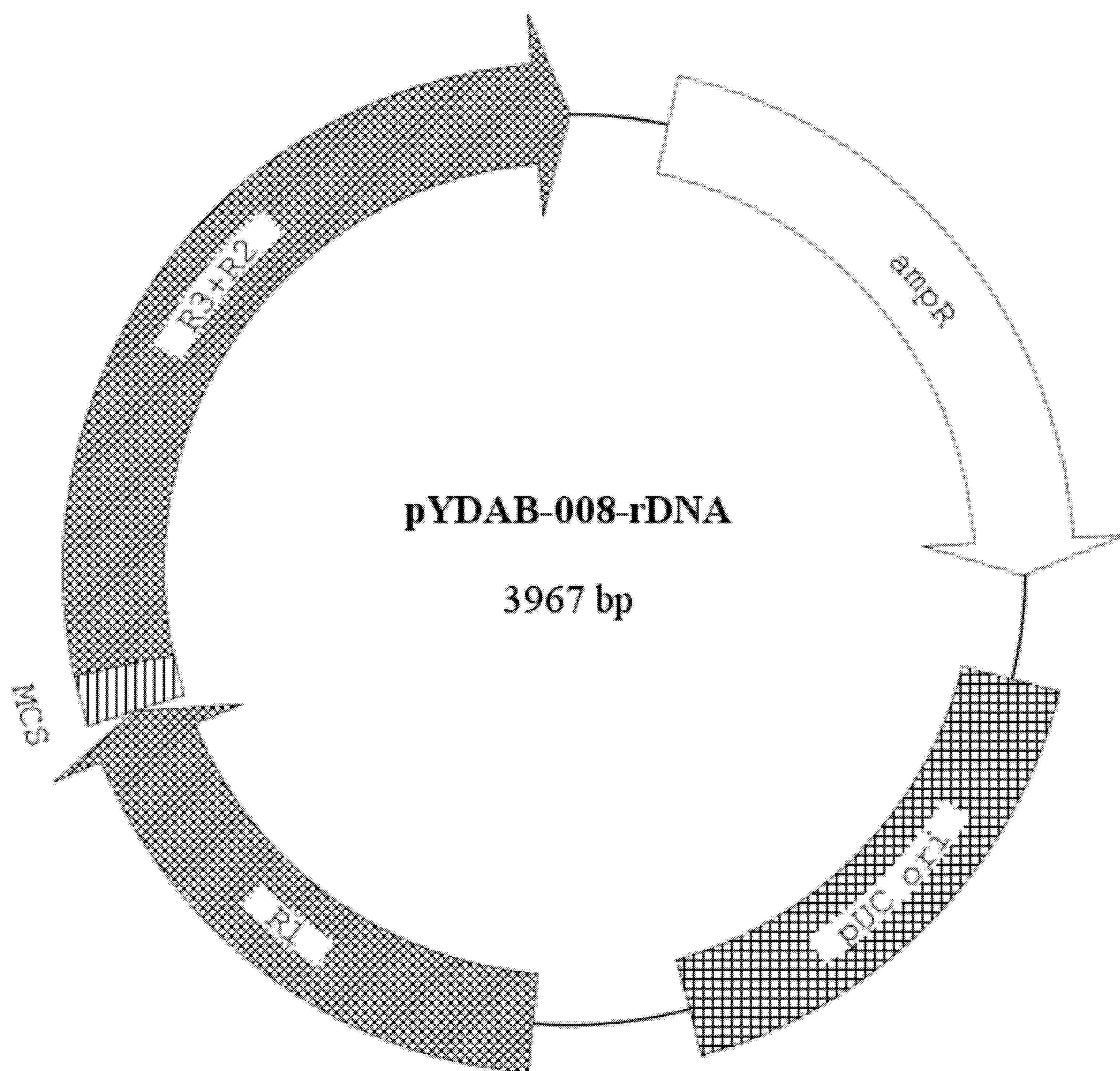


FIG. 5D

**FIG. 6**

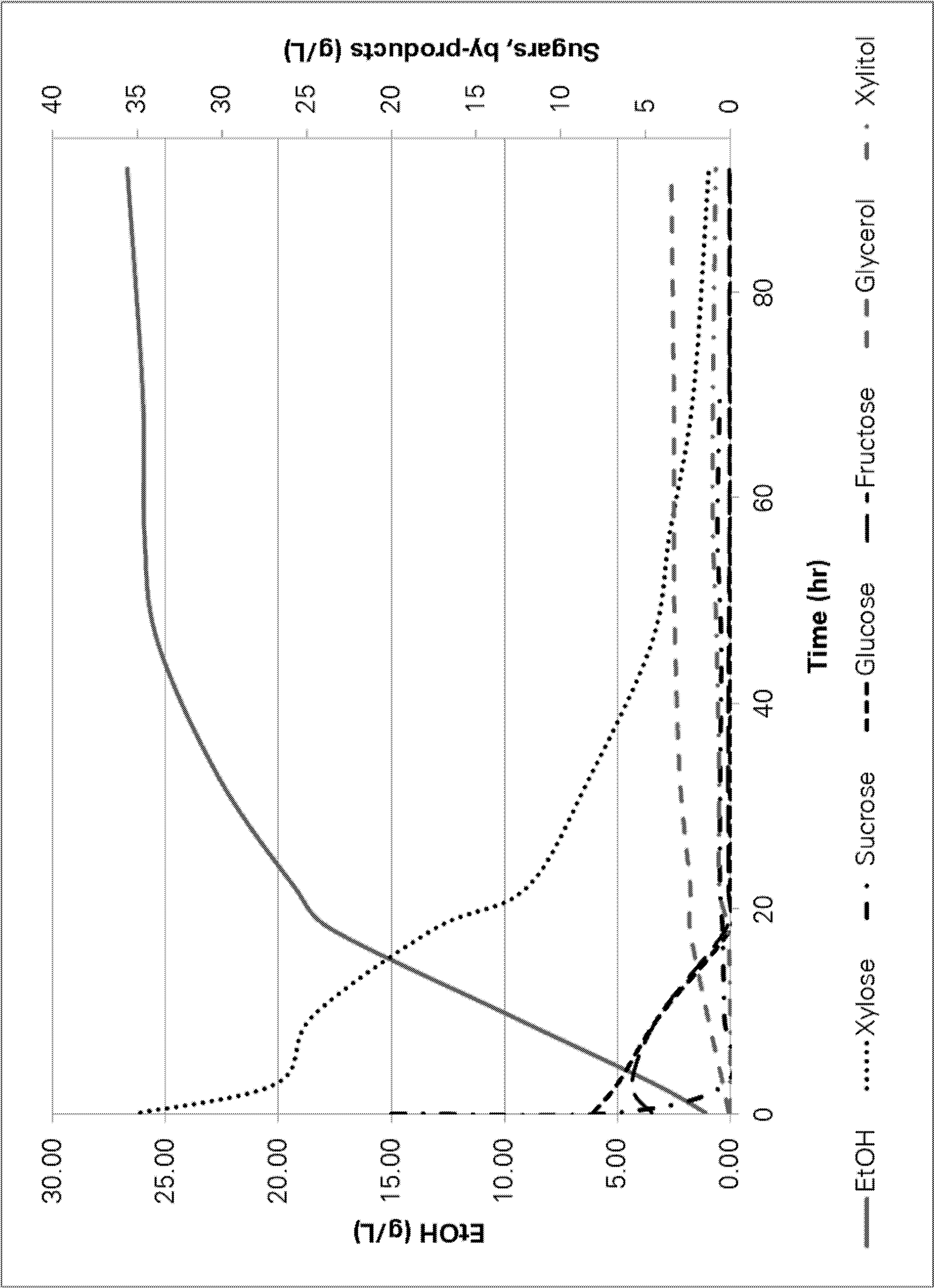


FIG. 7A

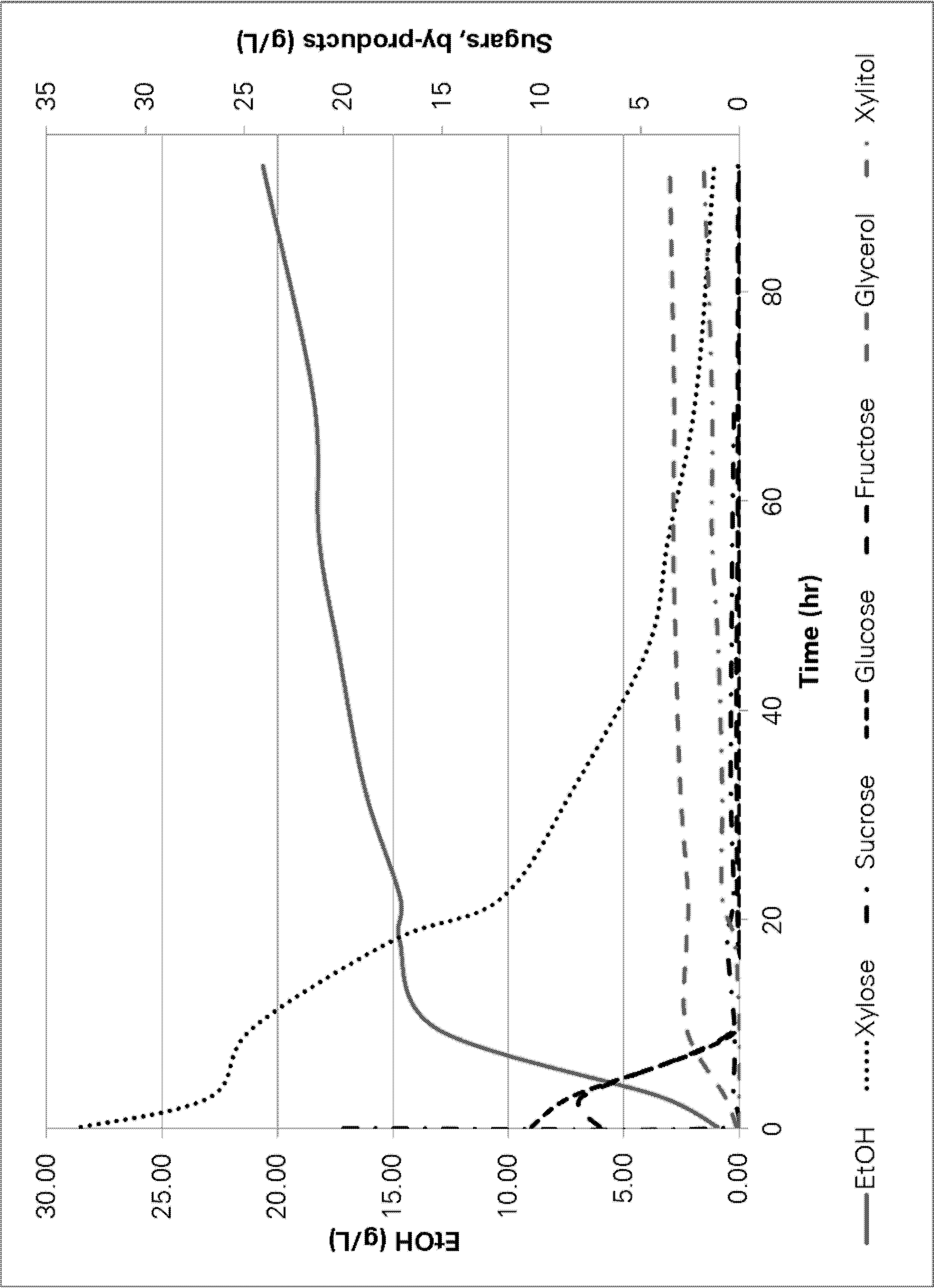


FIG. 7B

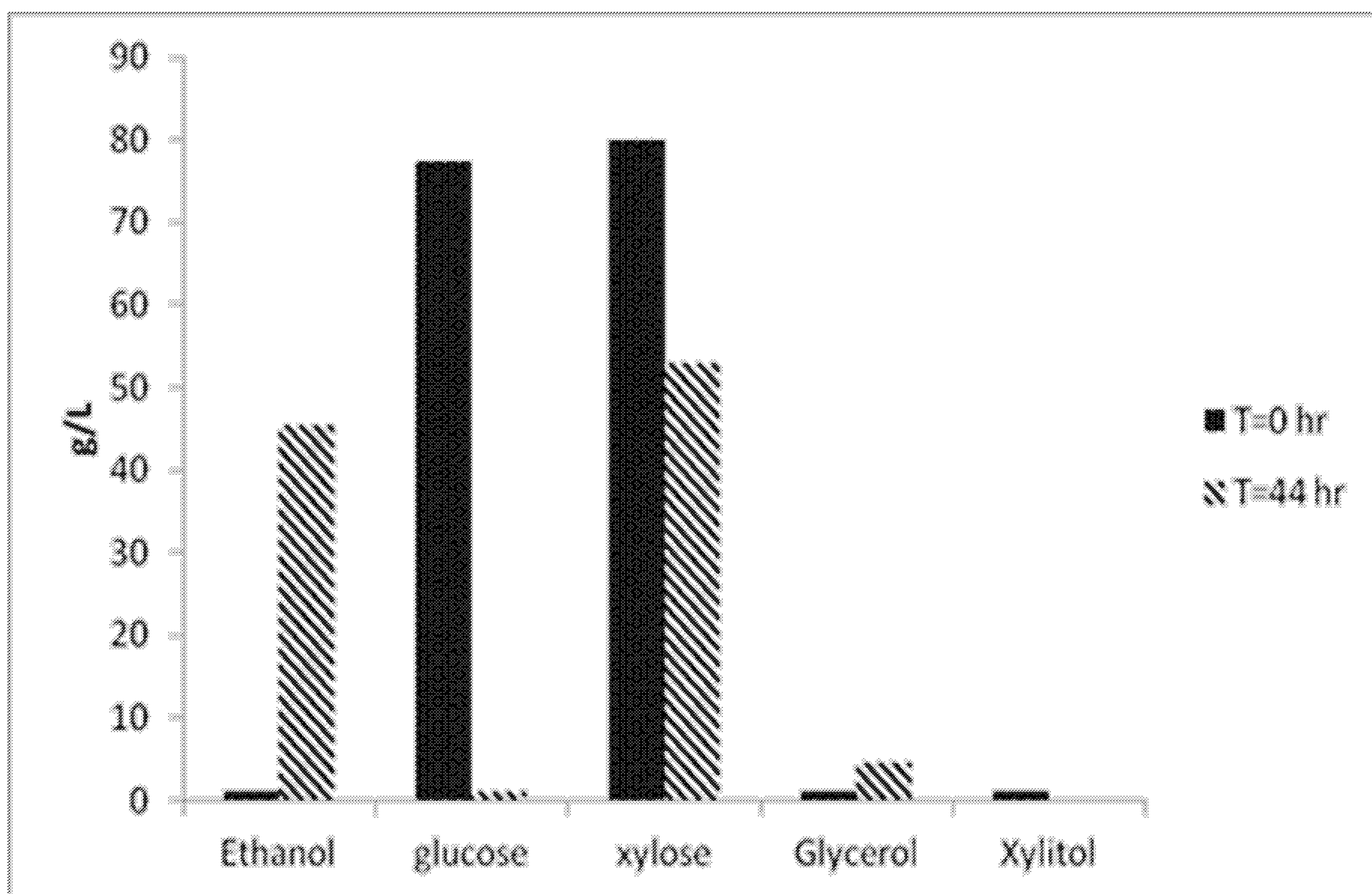


FIG. 7C

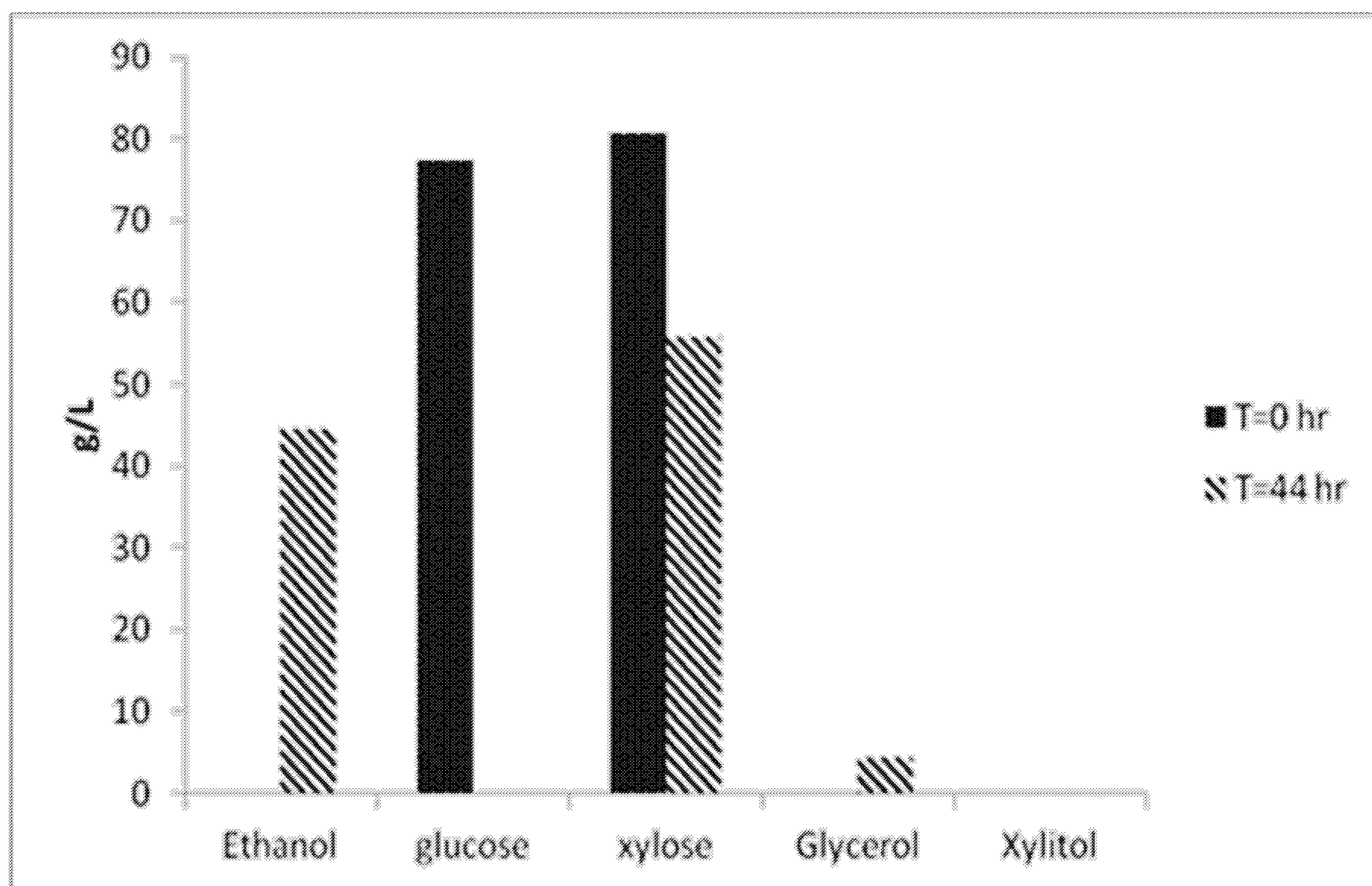


FIG. 7D

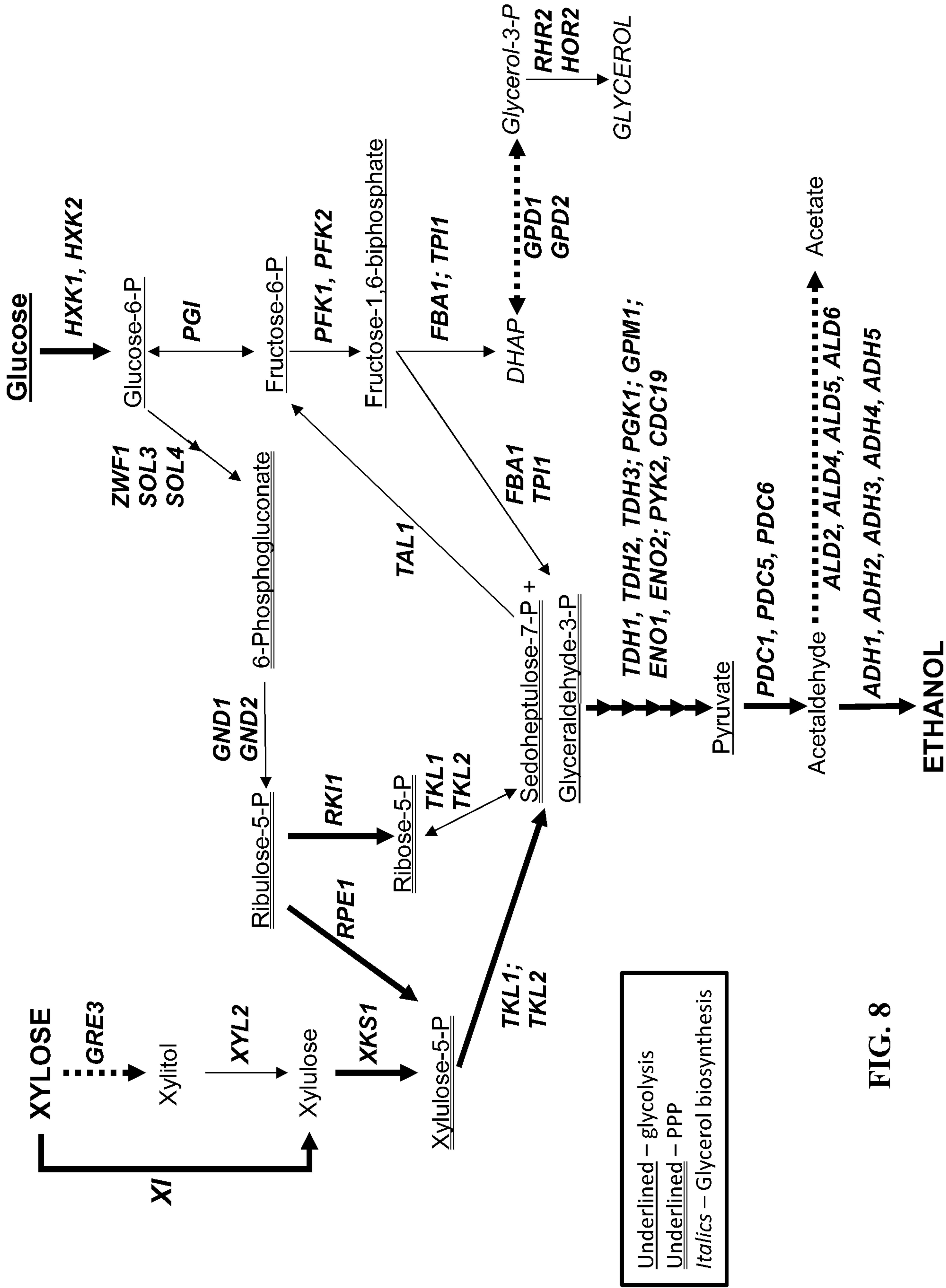


FIG. 8