



(51) International Patent Classification:

CI2N 9/22 (2006.01) CI2N 15/55 (2006.01)
CI2N 15/10 (2006.01) CI2P 19/30 (2006.01)

(21) International Application Number:

PCT/CA2022/051668

(22) International Filing Date:

11 November 2022 (11.11.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/263,896 11 November 2021 (11.11.2021) US

(71) Applicant: **INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE** [CA/CA]; 490 rue de la Couronne, Québec, Québec G1K 9A9 (CA).

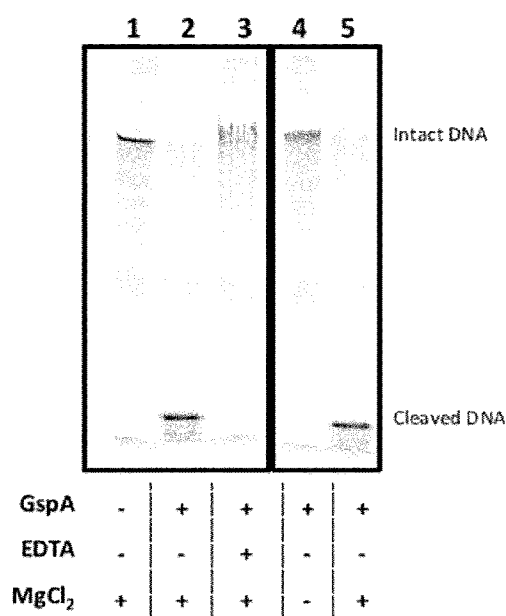
(72) Inventors: **VEYRIER, Frédéric**; 7075 rue de Lanaudière, Montréal, Québec H2E 1Y1 (CA). **CHENAL, Martin**; 9935 rue Lajeunesse, App. 213, Montréal, Québec H3L 2C9 (CA).

(74) Agent: **LAVERY, DE BILLY, LLP**; 1, Place Ville Marie, Bureau 4000, Montréal, Québec H3B 4M4 (CA).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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).

(54) Title: **ENDONUCLEASES THAT SELECTIVELY CLEAVE SINGLE-STRANDED NUCLEIC ACIDS AND USES THEREOF**

**FIG. 2A**

(57) Abstract: The present application relates to endonucleases specific for single-stranded nucleic acid molecules. These endonuclease s having a length of 60 to 150 amino acids or less and comprises a single GIY-YIG domain. The present application also relates to compositions and cells comprising such endonucleases as well as to methods of cleaving single-stranded nucleic acid molecules comprising a nucleotide sequence recognized by the endonucleases.

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

TITLE OF INVENTION

ENDONUCLEASES THAT SELECTIVELY CLEAVE SINGLE-STRANDED NUCLEIC ACIDS AND USES THEREOF

CROSS REFERENCE TO RELATED APPLICATIONS

5 The present application claims the benefit of U.S. provisional patent application serial No. 63/263,896, filed on November 11, 2021, which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

The present disclosure generally relates to the fields of enzymology and molecular biology, and more particularly to endonucleases.

10 BACKGROUND ART

Endonucleases are enzymes that cleave the phosphodiester bond within a polynucleotide chain. Two types of endonucleases are restriction nucleases and homing endonucleases.

15 Restriction endonucleases are enzymes that recognize a specific nucleotide sequence in a double-stranded nucleic acid called a restriction site. Upon binding to the restriction site, the restriction endonucleases cleave within or near the restriction site. These enzymes are routinely used for DNA modification in laboratories, such as for genetic engineering and molecular cloning. For example, they are used to assist insertion of genes into plasmid vectors, to distinguish gene alleles by specifically recognizing single base changes in DNA known as single-nucleotide polymorphisms (SNPs), to digest genomic DNA for gene analysis, and to insert nucleic acid molecules within the genome of an organism. Most of the known restriction enzymes recognize a
20 restriction site, typically comprising from 4 to 8 nucleotides that are often palindromic, within a double-stranded DNA molecule, and produce a double-stranded cut in the DNA.

25 Homing endonucleases are double-stranded DNases that have large, asymmetric recognition sites (12-40 base pairs). However, unlike restriction endonucleases, homing endonucleases tolerate some sequence degeneracy within their recognition sequence, which means that single base changes do not abolish cleavage but reduce its efficiency to variable extents. As a result, their observed sequence specificity is typically in the range of 10-12 base pairs.

30 In contrast to restriction and homing endonucleases that cleaves double-stranded nucleic acids at specific sites, endonucleases that cleave single-stranded nucleic acid molecules are usually non-specific, *i.e.*, they do not recognize a specific sequence (*e.g.*, restriction site) within the single-stranded nucleic acid molecules but rather cleave at various sequences to degrade the nucleic acid molecules in several fragments.

Thus, there is a need for the identification of novel endonucleases that are able to recognize and cleave at specific sequences within single-stranded nucleic acid molecules such as single-stranded DNA molecules.

The present description refers to a number of documents, the content of which is herein incorporated by reference in their entirety.

SUMMARY

The present disclosure provides the following items 1 to 73:

1. An isolated endonuclease specific for single-stranded desoxyribonucleic acid molecules, the isolated endonuclease having a length of 60 to 150 amino acids or less and comprising a single GIY-YIG domain.
2. The isolated endonuclease of item 1, having a length of 70 to 130 amino acids.
3. The isolated endonuclease of item 1, having a length of 80 to 120 amino acids.
4. The isolated endonuclease of any one of items 1 to 3, wherein the GIY-YIG domain is of the following sequence (I):

X1-X2-X3-B1-X4-X5-X6-B2-X7-B3-X8 (I)

wherein

X1 is any amino acid;

X2 is any amino acid;

X3 is Y or H;

B1 is a sequence of 8 to 12 amino acids;

X4 is Y or H;

X5 is any amino acid.

X6 is G or D;

B2 is a sequence of 6 to 15 amino acids;

X7 is R or T;

B3 is a sequence of 30 to 40 amino acids; and

X8 is E.

5. The isolated endonuclease of item 4, wherein X1 is Y, W, V, A, F, I, C, H, R, T or S.
6. The isolated endonuclease of item 5, wherein X1 is Y.
7. The isolated endonuclease of any one of items 4 to 6, wherein X2 is V, I, L, T, A or F.
8. The isolated endonuclease of item 7, wherein X2 is V.
9. The isolated endonuclease of any one of items 4 to 8, wherein X3 is Y.
10. The isolated endonuclease of any one of items 4 to 9, wherein B1 is a sequence of 8 to 10 amino acids.
11. The isolated endonuclease of any one of items 4 to 10, wherein X4 is Y.
12. The isolated endonuclease of any one of items 4 to 11, wherein X5 is I, L, V, T, A, C, or K.

13. The isolated endonuclease of item 12, wherein X5 is I, T or V.
14. The isolated endonuclease of item 13, wherein X5 is I.
15. The isolated endonuclease of any one of items 4 to 14, wherein X6 is G.
16. The isolated endonuclease of any one of items 4 to 15, wherein B2 is a sequence of 6 to 10 amino acids.
17. The isolated endonuclease of item 16, wherein B2 is a sequence of 6 to 8 amino acids.
18. The isolated endonuclease of any one of items 4 to 17, wherein X7 is R.
19. The isolated endonuclease of any one of items 4 to 18, wherein B3 is a sequence of 35 to 40 amino acids.
20. The isolated endonuclease of any one of items 4 to 19, wherein X8 is E.
21. The isolated endonuclease of any one of items 4 to 20, wherein the GIY-YIG domain is of the following sequence (II):

X1-X2-X3-B1-X4-X5-X6-B2-X7-B4-X9-B5-X8 (II)

wherein

X1, X2, X3, B1, X4, X5, X6, B2, X7 and X8 are as defined in items 4 to 20;

B4 is a sequence of 1 to 5 amino acids;

X9 is H, Q or Y; and

B5 is a sequence of 30 to 38 amino acids.

22. The isolated endonuclease of item 21, wherein B4 is a sequence of 2 to 4 amino acids.

23. The isolated endonuclease of item 21 or 22, wherein X9 is H.

24. The isolated endonuclease of any one of items 21 to 23, wherein B5 is a sequence of 30 to 35 amino acids.

25. The isolated endonuclease of item 24, wherein B5 is a sequence of 31 to 33 amino acids.

26. The isolated endonuclease of any one of items 4 to 25, wherein the GIY-YIG domain is of the following sequence (III):

X1-X2-X3-B1-X4-X5-X6-B2-X7-B4-X9-B5-X8-B6-X10 (III)

wherein

X1, X2, X3, B1, X4, X5, X6, B2, X7, B4, X9, B5 and X8 are as defined in items 4 to 20;

B6 is a sequence of 15 to 20 amino acids; and

X10 is N or K.

27. The isolated endonuclease of item 26, wherein B6 is a sequence of 16 to 19 amino acids.

28. The isolated endonuclease of item 26, wherein X10 is N.

29. The isolated endonuclease of any one of items 1 to 28, comprising an amino acid sequence having at least 50% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.

30. The isolated endonuclease of any one of items 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 60% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.
31. The isolated endonuclease of any one of items 1 to 28, wherein the isolated endonuclease
5 comprises an amino acid sequence having at least 70% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.
32. The isolated endonuclease of any one of items 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 80% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.
- 10 33. The isolated endonuclease of any one of items 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 90% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.
34. The isolated endonuclease of any one of items 1 to 28, wherein the isolated endonuclease comprises the amino acid sequence of any one of the sequences set forth in SEQ ID NOs:2-2891.
- 15 35. A composition comprising (i) the isolated endonuclease of any one of items 1 to 34, and (ii) an aqueous saline solution or buffer.
36. The composition of item 35, wherein the aqueous saline solution or buffer comprises a metal.
37. The composition of item 36, wherein the metal is in the form of a metal salt.
- 20 38. The composition of item 36 or 37, wherein the metal is magnesium, manganese or nickel.
39. The composition of item 38, wherein the metal is magnesium.
40. The composition of item 39, wherein the composition comprises magnesium chloride (MgCl_2).
41. The composition of any one of items 35 to 40, wherein the single-stranded nucleic acid
25 molecule is a single-stranded DNA molecule.
42. The composition of any one of items 35 to 41, wherein the single-stranded nucleic acid molecule comprises a nucleotide sequence having at least 50% sequence identity with the sequence: GTCATTCCCNNNNNNNNGGGAATC or GTCATTCCCGCGAAAGCGGGAATC.
43. The composition of any one of items 35 to 42, wherein the single-stranded nucleic acid
30 molecule comprises the following nucleotide sequence: GTCANNCCNGNNNNANNCNGGNNNC.
44. The composition of item 43, wherein the single-stranded nucleic acid molecule comprises the following nucleotide sequence: GTCAYBCCMGYRHAVRCKGGVRNC.
45. The composition of item 43 or 44, wherein the single-stranded nucleic acid molecule comprises any one of the nucleotide sequences depicted in **FIG. 8** and **FIG. 9C**.
- 35 46. The composition of any one of items 35 to 44, further comprising the single-stranded nucleic acid molecule defined in any one of items 41 to 45.

47. A method for cleaving a single-stranded nucleic acid molecule, the method comprising contacting the single-stranded nucleic acid molecule with the isolated endonuclease of any one of items 1 to 34 or the composition of any one of items 35 to 45 under conditions suitable for cleavage of the single-stranded nucleic acid molecule by the isolated endonuclease, wherein the single-stranded nucleic acid molecule comprises a recognition sequence for the isolated endonuclease.

48. The method of item 47, wherein said conditions comprises a temperature of about 20 to about 55°C.

49. The method of item 48, wherein said conditions comprises a temperature of about 35 to about 40°C.

50. The method of item 49, wherein said conditions comprises a temperature of about 37°C.

51. The method of any one of items 47 to 50, wherein said conditions comprises the presence of a metal.

52. The method of item 51, wherein said metal is magnesium, manganese or nickel.

53. The method of item 52, wherein the metal is magnesium.

54. The method of item 53, wherein the magnesium is in the form of magnesium chloride (MgCl₂).

55. The method of any one of item 51 to 54, wherein said metal is at a concentration of at least 5 mM.

56. The method of any one of items 51 to 54, wherein said metal is at a concentration of at least 10 mM.

57. The method of any one of items 47 to 56, wherein said contacting is for a period of at least 2 minutes.

58. The method of any one of items 47 to 56, wherein said contacting is for a period of at least 15 minutes.

59. The method of any one of items 47 to 58, wherein said conditions comprises a pH of about 6 to 8.

60. The method of any one of items 47 to 58, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.00001.

61. The method of item 60, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.01.

62. The method of item 61, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.5.

63. A method for rendering a single-stranded nucleic acid susceptible to cleavage by the endonuclease defined in any one of items 1 to 34, the method comprising incorporating a nucleotide sequence comprising a recognition sequence for the isolated endonuclease into the single-stranded nucleic acid.

64. The method of item 63, wherein the nucleotide sequence comprises one of the sequence defined in any one of items 42 to 45

65. The method of item 64, wherein the method comprises adding a nucleic acid fragment comprising the nucleotide sequence defined in any one of items 42 to 45 at the 5'-end, 3'-end or within the single-stranded nucleic acid.

66. The method of item 64, wherein the method comprises introducing one or more mutations within the sequence of the single-stranded nucleic acid to obtain the nucleotide sequence defined in any one of items 42 to 45.

67. A cell comprising the endonuclease defined in any one of items 1 to 34, wherein the endonuclease is heterologous to the cell.

68. A method for expressing the endonuclease defined in any one of items 1 to 34 in a cell, the method comprising introducing a nucleic acid encoding the endonuclease into the cell.

69. The cell of item 67 or the method of item 68, wherein the cell is a prokaryotic or eukaryotic cell.

70. The cell or method of any one of items 67 to 69, wherein the cell comprises a single-stranded nucleic acid that is cleaved by the endonuclease.

71. The cell or method of any one of items 68 to 70, wherein the nucleic acid encoding the endonuclease is present in a vector.

72. A kit comprising the endonuclease defined in any one of items 1 to 34 or the composition of any one of items 35 to 45, and instructions for cleaving single-stranded nucleic acid molecules using the endonuclease.

73. The kit of item 72, wherein said instructions comprise the method of any one of items 47 to 62.

Other objects, advantages and features of the present disclosure will become more apparent upon reading of the following non-restrictive description of specific embodiments thereof, given by way of example only with reference to the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

In the appended drawings:

FIG. 1A depicts purification of SsnA from *N. meningitidis*, *N. elongata*, *R. felis* and *L. pneumophila* along with three mutant proteins from *N. meningitidis*. Proteins were fused to a N-terminal 6xHis-tag or a N-terminal GST tag by cloning into pET15-MHL and pGEX vectors respectively, then purified to near purity by affinity chromatography.

FIG. 1B depicts the amino acid sequence of SsnA from *N. meningitidis* (SEQ ID NO:2) along with its sequence features. Individually mutated residues are highlighted.

FIG. 2A shows that SsnA is a metal-dependent endonuclease. Results of an endonuclease assay of 1 μ M SsnA on a 100 nt specific ssDNA. Rows #1-3 were performed using commercial NEBuffer2.1 (containing 10 mM $MgCl_2$), supplemented with 10 mM EDTA where indicated (#3). Rows #4-5 were performed using a homemade identical buffer lacking $MgCl_2$, which was supplemented with 10 mM $MgCl_2$ where indicated (#5).

FIG. 2B shows that SsnA only interacts with single-stranded DNA containing a specific sequence (NTS). Gel-shift (top, EMSA) and endonuclease (bottom) assays of 0.5 μ M SsnA with 0.8 μ M of DNA. Magnesium was omitted from the EMSA reaction mix to allow DNA binding by SsnA without nuclease activity. ssDNArev is the reverse complement of ssDNA. dsDNA is the double-stranded annealed product from ssDNA with ssDNArev. The arrow represents the recognition pattern (NTS repeat) of SsnA. Unless otherwise indicated (last well), assayed proteins were his-tagged.

FIG. 2C shows enzymatic activity of SsnA on truncated ssDNA and RNA. Sequences of 75nt, 37nt and 28nt were assayed, all of them containing a complete NTS sequence. The 37nt RNA sequence is the transcribed equivalent to the 37nt DNA. The 75nt ssDNA(T $\rightarrow$ U) sequence corresponds to the 75nt ssDNA with uracils instead of thymines (but with deoxyribose sugars).

FIG. 3 depicts the binding and cleavage of different ssDNA by SsnA. 100 nt oligonucleotides containing the full-length NTS repeat with different flanking sequences were assayed. Each sequence was taken from the *N. meningitidis* genome.

FIG. 3B depicts the assayed ssDNA sequences. The NTS repeat region is underlined.

FIG. 4A depicts the cleavage site determination of SsnA on a 75nt ssDNA containing its target sequence. 5'-labelled oligonucleotides of 18 to 25nt were run on a gel next to the results of an endonuclease assay with SsnA.

FIG. 4B depicts the sequence requirements for binding of cutting of SsnA on ssDNA. Gel-Shift (binding activity) and endonuclease (cutting activity) assays were performed on 75 nt ssDNA containing the target sequence (NTS), with single-nucleotide mutations throughout its length. The relative activities were measured for each individually mutated sequence and illustrated using a heat-map. Cleavage activity was normalized to the binding activity since binding is a prerequisite for cutting. Arrows denote the palindrome within the repeated sequence, which forms the stem of a stem-loop structure. Scissors depict the cutting site identified in **FIG. 5**.

FIG. 5 depicts the absence of binding activity of SsnA on branched DNA. Gel-shift assays (EMSA) of different branched DNA structures in presence of SsnA. H : Holliday junction, D : D-loop, F : fork, Y : pseudo-Y.

FIGS. 6A-D show the nuclease activity of SsnA. **FIG. 6A:** Metal requirements for the nuclease activity of SsnA. 10mM of each metal was used as the sole metal in the reaction mix. GST-tagged SsnA was assayed with nickel to ensure that the activity seen with His-SsnA was not due to nickel interacting with the purification tag.

FIG. 6B: Magnesium and manganese requirements for the nuclease activity of SsnA.

FIG. 6C: Temperature effect on SsnA's nuclease activity. For all of the above experiments, 0.8 μ M of 100 nt ssDNA containing the recognition sequence was used with 1 μ M of His-SsnA WT unless otherwise indicated. Boxed images show the specific cleavage products that were obtained and quantified.

FIG. 6D: depicts the cleavage kinetics of SsnA with the indicated concentrations of ssDNA.

FIG. 6E: Dose-response nuclease assay of SsnA depicting its sensitivity.

FIGs. 7A-C show the maximum likelihood mid-rooted phylogeny of SsnA and the Ssn protein family (GIY-YIG small proteins). Proteins NMV_0044 (SsnA) and NMV_0402 (SsnB, circle) from *N. meningitidis* 8013 2C4.3 was blasted against all bacteria with a threshold of 50% identity. Results were screened and curated to only keep proteins of 80-120 amino acids, corresponding to single GIY-YIG-domain proteins. **FIG. 7B** shows the portion of the tree corresponding to SsnA homologs from the *Neisseriaceae* family. **FIG. 7C** shows the portion of the tree corresponding to SsnB homologs from the *Neisseriaceae* family.

FIG. 8 shows the alignment of the putative recognition sequences of SsnA homologs, all neighbouring the SsnA gene in the corresponding species. The recognition sequence of SsnANm was blasted against the genomic regions (10kbp) encompassing homologous genes from other species, then aligned with AlignX from the Vector NTI software.

FIG. 9A shows an alignment of the amino acid sequences of SsnA nucleases from *N. meningitidis* 8013 2C4.3 (SsnA(NMV0044), SEQ ID NO:2), *Neisseria elongata* subsp. *glycolytica* (SsnA(EFE49965.1), SEQ ID NO:3), *Legionella pneumophila* subsp. *pneumophila* str. Paris SsnA(WP_011213498), SEQ ID NO:4) and *Rickettsia felis* str. URRWXC12 (SsnA(WP_011271370), SEQ ID NO:5), performed using Clustal Omega 1.2.4. Conserved amino acids are indicated by asterisks (*), the residues with strong similarity (PAM250 MATRIX score between amino acids of greater than 0.5) indicated by colons (:), and the residues with weak similarity (PAM250 MATRIX score between amino acids of 0.5 or less) indicated by dots (.).

FIG. 9B depicts the percent identity matrix of SsnA homologs aligned above.

FIG. 9C: Sequences located near the genes encoding SsnA homologs and having similarities with the recognition sequence from *N. meningitidis* were identified and aligned using Clustal Omega. Conserved nucleotides are highlighted, with darker tones highlighting the most conserved positions.

FIG. 10A-C show the nuclease activity of SsnA homologs from other species. Sequences located near the genes encoding these SsnA and having similarities with the recognition sequence from *N. meningitidis* were identified and synthesized with a 5' fluorescent tag, then tested with their respective SsnA using nuclease assays as previously described. **FIG. 10A** depicts the nuclease activity of SsnA from *N. elongata* (EFE49965.1), **FIG. 10B** depicts the

nuclease activity of SsnA from *R. felis* (WP_011271370) and FIG. 10C depicts the nuclease activity of SsnA from *L. pneumophila* (WP_011213498).

FIGs. 11A-11JJJJ depicts the amino acid sequences of various putative Ssn having at least 50% sequence identity with Protein NMV_0044 (SsnA) from *N. meningitidis* 8013 2C4.3 (SEQ ID NO:2).

FIGs. 12A-B show quantitative transformation assays performed on wild-type (WT), SsnA knock-out (KO) and SsnA complemented (Compl) strains of *N. meningitidis*. FIG. 13A: transformation of a plasmid that does not harbour the nuclease's recognition sequence. FIG. 13B: transformation of a plasmid harbouring SsnA's recognition sequence.

FIGs. 13A-B show an alignment of the amino acid sequences of 79 endonucleases comprising a GIY-YIG domain belonging to the GIY-YIG_unchar_3 Conserved Protein Domain Family (CD10448) (Lu S *et al.* (2020). "CDD/SPARCLE: the conserved domain database in 2020.", *Nucleic Acids Res.* 48(D1):D265-D268)

DETAILED DISCLOSURE

The use of the terms "a" and "an" and "the" and similar referents in the context of describing the technology (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context.

The terms "comprising", "having", "including", and "containing" are to be construed as open-ended terms (i.e., meaning "including, but not limited to") unless otherwise noted.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context.

The use of any and all examples, or exemplary language ("e.g.", "such as") provided herein, is intended merely to better illustrate embodiments of the claimed technology and does not pose a limitation on the scope unless otherwise claimed.

No language in the specification should be construed as indicating any non-claimed element as essential to the practice of embodiments of the claimed technology.

Herein, the term "about" has its ordinary meaning. The term "about" is used to indicate that a value includes an inherent variation of error for the device or the method being employed to determine the value, or encompass values close to the recited values, for example within 10% of the recited values (or range of values).

Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All subsets of values within the ranges are also incorporated into the specification as if they were individually recited herein.

Where features or aspects of the disclosure are described in terms of Markush groups or list of alternatives, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member, or subgroup of members, of the Markush group or list of alternatives.

Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in stem cell biology, cell culture, molecular genetics, immunology, immunohistochemistry, protein chemistry, and biochemistry).

Unless otherwise indicated, the recombinant protein, cell culture, and immunological techniques utilized in the present disclosure are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley and Sons (1984), J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T. A. Brown (editor), *Essential Molecular Biology: A Practical Approach*, Volumes 1 and 2, IRL Press (1991), D. M. Glover and B. D. Hames (editors), *DNA Cloning: A Practical Approach*, Volumes 1-4, IRL Press (1995 and 1996), and F. M. Ausubel *et al.* (editors), *Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present), Ed Harlow and David Lane (editors) *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory, (1988), and J. E. Coligan et al. (editors) *Current Protocols in Immunology*, John Wiley & Sons (including all updates until present).

The present inventors have identified a family of endonucleases that preferentially bind to and cleave single-stranded DNA. These endonucleases recognize and cleave specific nucleotide sequences in single-stranded nucleic acids (single-stranded DNA). These endonucleases may be useful, for example, in various genetic engineering and molecular biology applications. These endonucleases are short proteins (typically less than 150 amino acids, and preferably less than 140 or 130 amino acids) and comprises a conserved GIY-YIG domain. The GIY-YIG domain comprises two short semi-conserved motifs "GIY" and "YIG" in the N-terminal part, followed by an Arg residue in the center and a Glu residue in the C-terminal part. The GIY-YIG domain has an α/β -sandwich architecture with a central three-stranded antiparallel β -sheet flanked by three-helices. The three-stranded antiparallel β -sheet contains the GIY-YIG sequence elements.

The present disclosure provides an isolated endonuclease specific for single-stranded nucleic acid molecules (e.g., binds to and cleaves a single-stranded nucleic acid such as single-stranded DNA), the isolated endonuclease having a length of 150 amino acids or less and comprising a GIY-YIG domain.

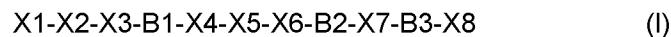
The present disclosure also provides a cell comprising an endonuclease specific for single-stranded nucleic acid molecules as described herein. In an embodiment, the endonuclease is heterogenous to the cell. "Heterogenous" as used herein means that the endonuclease is the

product of a gene that is not naturally present in the cell. For example, if the endonuclease is the endonuclease of SEQ ID NO:2, the cell is not a *Neisseria elongata* subsp. *glycolytica* cell. The cell may be a cell from another bacterial species or subspecies or an eucaryotic cell (mammalian cell, human cell, yeast cell, etc.), for example.

In an embodiment, the endonuclease of the present disclosure has a length of 80 to 130 amino acids. In another embodiment, the endonuclease of the present disclosure has a length of 85 to 120 amino acids. In an embodiment, the endonuclease of the present disclosure does not comprise any additional domain, it only consists of a single GIY-YIG domain.

FIGs 13A-B depicts a sequence alignment of 79 representative GIY-YIG domains from short endonucleases according to the present disclosure. These 79 endonucleases belong to the GIY-YIG_unchar_3 Conserved Protein Domain Family (CD10448) (Lu S *et al.* (2020). "CDD/SPARCLE: the conserved domain database in 2020.", *Nucleic Acids Res.* 48(D1):D265-D268). **FIG. 2A** also depicts the sequences corresponding to the GIY motif, the YIG motif and the conserved Glu residue (putative metal binding site) in SsnA from *N. meningitidis* (SEQ ID NO:2).

In an embodiment, the endonuclease of the present disclosure comprises a GIY-YIG domain of the formula I:



wherein

X1 is any amino acid, preferably G, Y, W, V, A, F, I, C, H, R, T, S, more preferably Y

X2 is any amino acid, preferably V, I, L, T, A, F, more preferably V

X3 is Y or H, preferably Y

B1 is a sequence of 8 to 12 amino acids, preferably 9 to 11 amino acids, for example 10 amino acids

X4 is Y or H, preferably Y

X5 is any amino acid, preferably I, L, V, T, A, C, or K, more preferably I, T or V, even more preferably I.

X6 is G or D, preferably G;

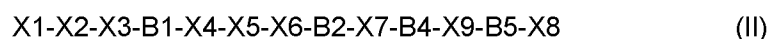
B2 is a sequence of 6 to 15 amino acids, preferably of 6 to 10 amino acids, more preferably of 6 to 8 amino acids, for example 7 amino acids;

X7 is R or T, preferably R;

B3 is a sequence of 30 to 40 amino acids, preferably of 35 to 40 amino acids, for example 35, 36 or 37 amino acids; and

X8 is E, I or A, preferably E.

In another embodiment, the endonuclease of the present disclosure comprises a GIY-YIG domain of the formula II:



wherein

X1, X2, X3, B1, X4, X5, X6, B2, X7 and X8 are as defined above;

B4 is a sequence of 1 to 5 amino acids, preferably of 2 to 4 amino acids, more preferably of 3 amino acids;

5 X9 is H, Q or Y, preferably H; and

B5 is a sequence of 30 to 38 amino acids, preferably of 30 to 35, more preferably of 31 to 33 amino acids, for example 32 amino acids.

10 In another embodiment, the endonuclease of the present disclosure comprises a GIY-YIG domain of the formula III:

X1-X2-X3-B1-X4-X5-X6-B2-X7-B4-X9-B5-X8-B6-X10 (III)

wherein

X1, X2, X3, B1, X4, X5, X6, B2, X7, B4, X9, B5 and X8 are as defined above;

15 B6 is a sequence of 15 to 20 amino acids, preferably of 16 to 19 amino acids, more preferably of 18 amino acids; and

X10 is N or K, preferably N.

20 In an embodiment, the isolated endonuclease of the present disclosure comprises or consists of an amino acid sequence having at least 50% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891).

The term "endonuclease" as used herein refers to an enzyme having the ability to cleave a single-stranded nucleic acid molecule, such as single-stranded DNA, at or near a specific nucleotide sequence (recognition or restriction site).

25 The term "isolated" as used herein refers to a molecule (endonuclease) that is in a milieu or environment that is different from the natural milieu or environment where it is found in nature (*i.e.*, that has been subjected to human manipulation), for example a endonuclease that has isolated from the natural bacteria that normally expressed it. As such, "isolated" does not necessarily reflect the extent to which the endonuclease has been purified, but indicates that the molecule has been separated in some way from the natural environment where it is normally found. An isolated endonuclease may also be produced recombinantly by cloning a nucleic acid encoding the endonuclease in a host cell capable of expressing the endonuclease, and collecting the endonuclease produced.

35 "Identity" refers to sequence identity between two polypeptides. Percent (%) sequence identity with respect to a reference polypeptide sequence is the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the reference polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to

achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are known for instance, using publicly available computer software such as Clustal Omega, BLAST, BLAST-2, ALIGN or
5 Megalign (DNASTAR) software. Appropriate parameters for aligning sequences are able to be determined, including algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

Similarity refers to sequence similarity between two polypeptides. Percent (%) sequence similarity with respect to a reference polypeptide sequence is the percentage of amino acid
10 residues in a candidate sequence that are similar (identical or conserved) with the amino acid residues in the reference polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence similarity. and considering conservative substitutions as part of the sequence similarity. The similarity between amino acids can be defined either by their chemical properties (e.g., hydrophobic, hydrophilic, charged, polar,
15 etc.) or based on a PAM matrix.

Variations in the endonucleases described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the endonuclease that results in a change in the amino acid
20 sequence as compared with the native sequence of the endonuclease. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the endonuclease. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the endonuclease with that of homologous known protein molecules and
25 minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically
30 making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence (e.g., ability to cleave single-stranded nucleic acid molecules).

In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 55% similarity or identity with the any one of the sequences set forth in
35 SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 60% similarity or identity with the any one of

the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 65% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 70% identity similarity or with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 75% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 80% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 85% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 90% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 95% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 96% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 97% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 98% similarity or identity with the any one of the sequences set forth in SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of an amino acid sequence having at least 99% similarity or identity with the any one of the sequences set forth SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811)

and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises or consists of or consists one of the amino acid sequences set forth in any one of SEQ ID NOs:2-5, **FIGs. 11A-JJJJ** (SEQ ID NO:6-2811) and **FIGs. 13A-B** (SEQ ID NO: 2812-2891), preferably SEQ ID NOs:2-5.

FIG. 9A depicts an alignment of the amino acid sequences of SEQ ID NOs:2-5, with the residues conserved between the sequences indicated by asterisks (*), the residues with strong similarity (PAM250 MATRIX score between amino acids of greater than 0.5) indicated by colons (:), and the residues with weak similarity (PAM250 MATRIX score between amino acids of 0.5 or less) indicated by dots (.). In an embodiment, the isolated endonuclease comprises the conserved residues in SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises the conserved residues and the residues with strong similarity in SEQ ID NOs:2-5. In an embodiment, the isolated endonuclease comprises the conserved residues and the residues with strong and weak similarities in SEQ ID NOs:2-5.

FIGs. 13A-B depict an alignment of the amino acid sequences of the GIY-YIG domains from 79 endonucleases belonging to the GIY-YIG_unchar_3 Conserved Protein Domain Family (CD10448), with the residues conserved between the sequences indicated by a # sign above the sequences. In an embodiment, the isolated endonuclease comprises the conserved residues in the sequences depicted in **FIGs. 13A-B**.

The isolated endonuclease may further comprise additional amino acids at its amino- (N) and/or carboxy (C)-terminal end. For example, the isolated endonuclease may be fused to a peptide or polypeptide, for example a peptide or polypeptide that may be used as an affinity tag to facilitate the detection and/or purification of the endonuclease. Examples of affinity tags include polyhistidine (His) tags, polyarginine tags, glutathione-S-transferase (GST) tags, FLAG tags, streptavidin-binding peptide or streptavidin-binding protein (SBP) tags, streptavidin-binding tag (Strep-tag), calmodulin-binding peptide (CBP) tags, chitin-binding tags, Maltose-binding protein (MBP) tags, and natural histidine affinity tags (HAT). The peptide or polypeptide may be fused directly to the N- and/or C-terminal end of the endonuclease, or indirectly via a linker. Such a linker may be a peptide/polypeptide linker comprising one or more amino acids or another type of chemical linker (e.g., a carbohydrate linker, a lipid linker, a fatty acid linker, a polyether linker, PEG, etc.) having suitable flexibility and stability to allow the endonuclease to adopt a proper conformation. The linker may comprise at least 2, 3 or 4 amino acids. The linker may comprise about 100, 90, 80, 70, 60 or 50 amino acids or less, and preferably 20, 15 or 10 amino acids or less.

The isolated endonuclease of the disclosure may be produced by expression in a host cell comprising a nucleic acid encoding the isolated endonuclease (recombinant expression) or by chemical synthesis (e.g., solid-phase peptide synthesis). Peptides and polypeptides can be readily synthesized by manual and automated solid phase procedures well known in the art.

Suitable syntheses can be performed for example by utilizing "t-Boc" or "Fmoc" procedures. Techniques and procedures for solid phase synthesis are described in for example Solid Phase Peptide Synthesis: A Practical Approach, by E. Atherton and R. C. Sheppard, published by IRL, Oxford University Press, 1989. Alternatively, the polypeptides may be prepared by way of segment condensation, as described, for example, in Liu et al., Tetrahedron Lett. 37: 933-936, 1996; Baca et al., J. Am. Chem. Soc. 117: 1881-1887, 1995; Tarn et al., Int. J. Peptide Protein Res. 45: 209-216, 1995; Schnolzer and Kent, Science 256: 221-225, 1992; Liu and Tarn, J. Am. Chem. Soc. 116: 4149-4153, 1994; Liu and Tarn, Proc. Natl. Acad. Sci. USA 91: 6584-6588, 1994; and Yamashiro and Li, Int. J. Peptide Protein Res. 31 : 322-334, 1988). Other methods useful for synthesizing polypeptides are described in Nakagawa et al., J. Am. Chem. Soc. 107: 7087-7092, 1985.

The isolated endonuclease may also be prepared using recombinant DNA technology using standard methods. Accordingly, in another aspect, the disclosure further provides a nucleic acid (e.g., mRNA, cDNA) encoding the above-mentioned endonuclease. The disclosure also provides a vector comprising the above-mentioned nucleic acid. In yet another aspect, the present disclosure provides a cell (e.g., a host cell) comprising the above-mentioned nucleic acid and/or vector. The disclosure further provides a recombinant expression system, vectors and host cells, such as those described above, for the expression/production of an endonuclease of the disclosure, using for example culture media, production, isolation and purification methods well known in the art.

The endonuclease of the disclosure can be purified by many techniques of peptide/polypeptide purification well known in the art, such as reverse phase chromatography, high performance liquid chromatography (HPLC), ion exchange chromatography, size exclusion chromatography, affinity chromatography, gel electrophoresis, and the like. The actual conditions used to purify a particular peptide or polypeptide will depend, in part, on synthesis strategy and on factors such as net charge, hydrophobicity, hydrophilicity, and the like, and will be apparent to those of ordinary skill in the art. For affinity chromatography purification, any ligand or antibody that specifically binds the endonuclease (or to an affinity tag fused to the endonuclease) may for example be used.

The present disclosure also provides a composition comprising (i) an isolated endonuclease specific for single-stranded nucleic acid molecules as described herein, and (ii) an aqueous saline solution or buffer.

The composition according to the present disclosure an aqueous saline solution or buffer. Such aqueous saline solutions or buffers include ingredients that stabilize the endonuclease and provide suitable conditions for the enzymatic activity of the endonuclease (e.g., conditions that permit the cleavage of single-stranded nucleic acids). The aqueous saline solution or buffer present in the composition according to the present disclosure may include suitable salts,

buffering agents, minerals, co-factors, stabilizing agents, anti-oxidants, redox reagent, preservatives, etc.

In an embodiment, the composition comprises a buffering agent. The buffering agent is useful to keep the composition at a desired pH. Buffering agents are well known in the art and include potassium, acetate, citrate, acetate, phosphate, carbonate, succinate, histidine, borate, maleate, tris(hydroxymethyl) aminomethane (Tris), BIS-Tris, piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), 2-(N-morpholino)ethanesulfonic acid (MES), (3-(N-morpholino)propanesulfonic acid) (MOPS), N-(2-Acetamido)-2-aminoethanesulfonic acid (ACES), (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (HEPES), magnesium and hydrochloride buffers. In an embodiment, the composition comprises a Tris buffer. In a further embodiment the Tris buffer is a Tris-HCl or a Tris-acetate buffer. In an embodiment, the buffering agent is at a concentration of about 0.1 mM to 1M. In further embodiments, the buffering agent is at a concentration of about 1 mM to about 500 mM, about 1 mM to about 200 mM, about 1 mM to about 100 mM, about 5 mM to about 100 mM, about 5 mM to about 75 mM, about 5 mM to about 50 mM, or about 5 mM to about 25 or 20 mM. In an embodiment, the buffering agent is at a concentration of about 10 mM. In an embodiment, the buffering agent has a pH of about 5 to about 10. In further embodiments, the buffering agent has a pH of about 6 to about 9, of about 6.5 to about 9, of about 7 to about 9, of about 7.5 to about 8.5, or of about 7.6 to about 8.2. In an embodiment, the buffering agent has a pH of about 7.9 or 8.0. In an embodiment, the pH is the pH at a temperature of about 20 to about 40°C. In an embodiment, the pH is the pH at a temperature of about 20 or 25°C. In an embodiment, the pH is the pH at a temperature of about 37°C.

In an embodiment, the composition comprises a salt, such as a metal salt. Common salt-forming cations include ammonium (NH_4^+), manganese, nickel, calcium, iron, magnesium, potassium, sodium and copper. In an embodiment, the metal salt is a magnesium salt, manganese salt or zinc salt. Common salt-forming anions include acetate, carbonate, chloride, citrate, fluoride, nitrate, nitrite, oxide, phosphate and sulfate. Examples of salts include magnesium chloride (MgCl_2), magnesium acetate, potassium acetate (KCH_3CO_2), potassium chloride (KCl), sodium acetate (CH_3COONa), sodium chloride (NaCl), calcium chloride, zinc chloride, manganese sulfate, manganese chloride, nickel chloride, nickel acetate, and sodium sulfate (Na_2SO_4). In an embodiment, the salt comprises a magnesium, manganese, nickel and/or sodium cation. In an embodiment, the salt comprises a chloride anion. In an embodiment, the composition comprises KCl. In an embodiment, the composition comprises KCl and NaCl. In an embodiment, the concentration of salt in the composition is about 1 mM to about 500 mM. In further embodiments, the concentration of salt in the composition is about 10 mM to about 300 mM, about 20 mM to about 200 mM, about 20 mM to about 150 mM, about 30 mM to about 150 mM. In an embodiment, the composition comprises a salt comprising a magnesium cation (e.g., MgCl_2) at a concentration of about 1 mM to about 100 mM, for example about 1 mM to about 50

mM, about 5 mM to about 20 mM, or about 5 to about 15 mM. In an embodiment, the composition comprises a salt comprising a magnesium cation (e.g., MgCl_2) at a concentration of about 10 mM. In an embodiment, the composition comprises a salt comprising a sodium cation (e.g., NaCl) at a concentration of about 1 mM to about 200 mM, about 10 mM to about 150 mM, about 10 mM to about 100 mM, about 25 mM to about 75 mM. In an embodiment, the composition comprises a salt comprising a sodium cation (e.g., NaCl) at a concentration of about 50 mM.

In an embodiment, the composition comprises a stabilizing agent, such as a protein. In an embodiment, the protein is albumin, such as bovine serum albumin (BSA). In an embodiment, the stabilizing agent is at a concentration of about 1 $\mu\text{g/ml}$ to about 1 mg/ml . In an embodiment, the stabilizing agent is at a concentration of about 10 $\mu\text{g/ml}$ to about 500 $\mu\text{g/ml}$. In an embodiment, the stabilizing agent is at a concentration of about 50 $\mu\text{g/ml}$ to about 200 $\mu\text{g/ml}$. In an embodiment, the stabilizing agent is at a concentration of about 50 $\mu\text{g/ml}$ to about 150 $\mu\text{g/ml}$. In an embodiment, the stabilizing agent is at a concentration of about 80 to about 120 $\mu\text{g/ml}$, for example about 100 $\mu\text{g/ml}$.

In embodiments, the composition further comprises additional ingredients, such as detergents (e.g., non-ionic detergents like Triton® X-100 or Tween® 20) and/or redox reagents (DTT, beta-mercaptoethanol). For storage, the composition may further comprise a metal-chelating agent, such as ethylenediaminetetraacetic acid (EDTA).

Aqueous solutions/buffers for endonucleases are available commercially from several providers such as New England Biolabs Inc., Promega and Thermo Scientific. Examples from New England Biolabs Inc. include NEBuffer 1 (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl_2 , 1 mM DTT, pH 7.0@25°C); NEBuffer 1.1 (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl_2 , 100 $\mu\text{g/ml}$ BSA, pH 7.0@25°C); NEBuffer 2.1 (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl_2 , 100 $\mu\text{g/ml}$ BSA, pH 7.9@25°C); NEBuffer 3.1 (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl_2 , 100 $\mu\text{g/ml}$ BSA, pH 7.9@25°C); NEBuffer 4 (50 mM Potassium acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 1 mM DTT, pH 7.9@25°C); and CutSmart Buffer (50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 100 $\mu\text{g/ml}$ BSA, pH 7.9@25°C). Examples from Thermo Scientific include Buffer B (10 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl_2 , 0.1 mg/ml BSA), Buffer G (10 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl_2 , 50 mM NaCl, 0.1 mg/ml BSA), Buffer O (50 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl_2 , 100 mM NaCl, 0.1 mg/ml BSA), Buffer R (10 mM Tris-HCl (pH 8.5 at 37°C), 10 mM MgCl_2 , 100 mM KCl, 0.1 mg/ml BSA), and Buffer Tango™ (33 mM Tris-acetate (pH 7.9 at 37°C), 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/ml BSA). Examples from Promega include Buffer A (6 mM Tris-HCl (pH 7.5 at 37°C), 6 mM MgCl_2 , 6 mM NaCl, 1 mM DTT), Buffer B (6 mM Tris-HCl (pH 7.5 at 37°C), 6 mM MgCl_2 , 50 mM NaCl, 1 mM DTT), Buffer C (10 mM Tris-HCl (pH 7.9 at 37°C), 10 mM MgCl_2 , 50 mM NaCl, 1 mM DTT), Buffer D (6 mM Tris-HCl (pH 7.9 at 37°C), 6 mM MgCl_2 , 150 mM NaCl, 1 mM DTT), Buffer E (6 mM Tris-HCl (pH 7.5 at 37°C), 6 mM MgCl_2 , 100 mM NaCl, 1 mM DTT), Buffer F (10 mM

Tris-HCl (pH 8.5 at 37°C), 10 mM MgCl₂, 100 mM NaCl, 1 mM DTT), Buffer G (50 mM Tris-HCl (pH 8.2 at 37°C), 5 mM MgCl₂), Buffer H (90 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl₂, 50 mM NaCl), Buffer J (10 mM Tris-HCl (pH 7.5 at 37°C), 7 mM MgCl₂, 50 mM KCl, 1 mM DTT), Buffer K (10 mM Tris-HCl (pH 7.4 at 37°C), 10 mM MgCl₂, 150 mM KCl), Buffer L (10 mM Tris-HCl (pH 9.0 at 37°C), 3 mM MgCl₂, 100 mM NaCl), and MULTI-CORE™ Buffer (25 mM Tris-Acetate, pH 7.5 (at 37°C), 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT).

In an embodiment, the composition comprises an aqueous solution/buffer comprising:

from about 1 to about 100 mM of a buffering agent, such as a Tris-based buffering agent (*e.g.*, Tris-HCl), having a pH of about 7 to about 9;

from about 1 to about 100 mM of a metal salt, such as a salt comprising a magnesium and/or sodium cation (NaCl and/or KCl); and

from about 10 µg/ml to about 500 µg/ml of a stabilizing agent, such as a protein (*e.g.*, albumin).

In a further embodiment, the composition comprises an aqueous solution/buffer comprising:

from about 5 to about 20 mM of a buffering agent, such as a Tris-based buffering agent (*e.g.*, Tris-HCl), having a pH of about 7.5 to about 8.5;

from about 10 to about 100 mM of a metal salt, such as a metal salt comprising a magnesium and/or sodium cation (NaCl and/or KCl); and

from about 50 µg/ml to about 200 µg/ml of a stabilizing agent, such as a protein (*e.g.*, albumin).

In a further embodiment, the composition comprises an aqueous solution/buffer comprising:

about 10 mM of a buffering agent, such as a Tris-based buffering agent (*e.g.*, Tris-HCl), having a pH of about 7.9 or 8.0;

about 50 mM of NaCl; and

about 100 µg/ml of albumin (*e.g.*, BSA).

In a further embodiment, the composition comprises an aqueous solution/buffer comprising:

about 10 mM of a buffering agent, such as a Tris-based buffering agent (*e.g.*, Tris-HCl), having a pH of about 7.9 or 8.0;

about 50 mM of NaCl;

about 10 mM of MgCl₂; and

about 100 µg/ml of BSA.

The present disclosure also provides a mixture comprising the above-described isolated endonuclease or composition and a single-stranded nucleic acid molecule (*e.g.*, single-stranded DNA).

The present disclosure also provides a method for cleaving a single-stranded nucleic acid molecule comprising contacting the single-stranded nucleic acid molecule with the isolated

endonuclease or composition defined herein under conditions suitable for cleavage of the single-stranded nucleic acid molecule by the isolated endonuclease.

The results presented in the Examples below show that the endonucleases according to the present disclosure recognize specific nucleotide sequences within the single-stranded nucleic acid molecules. The recognition motif contains a ≈ 30 nt repeated sequence (NTS) that is predicted to form a stem-loop secondary structure. The cleavage of the single-stranded nucleic acid molecules occurs just outside the NTS.

In an embodiment, the single-stranded nucleic acid molecule cleaved by the endonuclease comprises a nucleotide sequence having at least 50% sequence identity with the sequence: GTCATTCCCnnnnnnnnGGGAATC (SEQ ID NO:2917) or GUCAUUCCCnnnnnnnnGGGAUC (SEQ ID NO: 2918).

In an embodiment, the single-stranded nucleic acid molecule cleaved by the endonuclease comprises a nucleotide sequence having at least 50% sequence identity with the sequence: GTCATTCCCGCGAAAGCGGGAATC (SEQ ID NO: 2919) or GUCAUUCCCGCGAAAGCGGGAUC (SEQ ID NO: 2920).

In an embodiment, the single-stranded nucleic acid molecule comprises the following nucleotide sequence: GTCANNCCNGNNNANNCNGGNNNC (SEQ ID NO: 2921) or GUCANNCCNGNNNANNCNGGNNNC (SEQ ID NO: 2922).

The identification of nucleotides in the present disclosure is according to the standard 1-letter nomenclature from the International Union of Pure and Applied Chemistry (IUPAC):

IUPAC 1-letter code	Base
A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T/U
S	G or C
W	A or T/U
K	G or T/U
M	A or C
B	C or G or T/U
D	A or G or T/U
H	A or C or T/U
V	A or C or G
N	any base

In an embodiment, the single-stranded nucleic acid molecule comprises the following nucleotide sequence: GTCAYBCCMGYRHAVRCKGGVRNC (SEQ ID NO: 2923) or GUCAYBCCMGYRHAVRCKGGVRNC (SEQ ID NO: 2924).

In an embodiment, the single-stranded nucleic acid molecule comprises a sequence having at least 50, 60, 70, 80, 90, 95 or 100% identity with one of the following nucleotide sequences:

	• GTCATCCCCGCGCAGGCGGGGACCC (SEQ ID NO: 2925)	or
	GUCAUCCCCGCGGCAGGCGGGACCC (SEQ ID NO: 2926);	
	• GTCATTCCCGCGCAGGCGGGAATCC (SEQ ID NO: 2927)	or
	GUCAUCCCCGCGCAGGCGGGAAUCC (SEQ ID NO: 2928);	
5	• GTCATTCCCGCGCAAAGCGGGAATCC (SEQ ID NO: 2929)	or
	GUCAUCCCCGCGCAAAGCGGGAUCC (SEQ ID NO: 2930);	
	• GTCATTCCCGCGAAGGCGGGAATCC (SEQ ID NO: 2931)	or
	GUCAUCCCCGCGAAGGCGGGAAUCC (SEQ ID NO: 2932);	
	• GTCATTCCCGCGCAGGCGGGAATCC (SEQ ID NO: 2933)	or
10	GUCAUCCCCGCGCAGGCGGGAAUCC (SEQ ID NO: 2934);	
	• GTCATCCCCGCGCAGGCGGGGACCC (SEQ ID NO: 2935)	or
	GUCAUCCCCGCGCAGGCGGGGACCC (SEQ ID NO: 2936);	
	• GTCATGCCCGCGTAGGCGGGCAACC (SEQ ID NO: 2937)	or
	GUCAUGCCCGCGUAGGCGGGCAACC (SEQ ID NO: 2938);	
15	• GTCATTCCCGCGCAAAGCGGGAAGCC (SEQ ID NO: 2939)	or
	GUCAUCCCCGCGCAAAGCGGGAAGCC (SEQ ID NO: 2940);	
	• GTCATTCCCGCGCAGGCGGGAATCC (SEQ ID NO: 2941)	or
	GUCAUCCCCGCGCAGGCGGGAAUCC (SEQ ID NO: 2942);	
	• GTCATTCCCGTGACACGCGGGAATCC (SEQ ID NO: 2943)	or
20	GUCAUCCCCGUGCACACGCGGGAUCC (SEQ ID NO: 2944);	
	• GTCATGCCCGCGCAGGCGGGCATCC (SEQ ID NO: 2945)	or
	GUCAUGCCCGCGCAGGCGGGCAUCC (SEQ ID NO: 2946);	
	• GTCATCCCCGCGAAGGCGGGGATCC (SEQ ID NO: 2947)	or
	GUCAUCCCCGCGAAGGCGGGGAUCC (SEQ ID NO: 2948);	
25	• GTCATTCCCGCGCAAAGCGGGAATCC (SEQ ID NO: 2949)	or
	GUCAUCCCCGCGCAAAGCGGGAUCC (SEQ ID NO: 2950);	
	• GTCATTCCCGCGAAGGCGGGAATCC (SEQ ID NO: 2951)	or
	GUCAUCCCCGCGAAGGCGGGAAUCC (SEQ ID NO: 2952);	
	• GTCATTCCCGCGCAGGCGGGAATCT (SEQ ID NO: 2953)	or
30	GUCAUCCCCGCGCAGGCGGGAAUCU (SEQ ID NO: 2954);	
	• GTCATTCCCGCGTAGGCGGGAATCC (SEQ ID NO: 2955)	or
	GUCAUCCCCGCGUAGGCGGGAAUCC (SEQ ID NO: 2956);	
	• GTCACTCCCGCGAAGGCGGGAGTCC (SEQ ID NO: 2957)	or
	GUCACUCCCCGCGAAGGCGGGAGUCC (SEQ ID NO: 2958);	
35	• GTCACCCCAGCGAAAGCTGGGGTCC (SEQ ID NO: 2959)	or
	GUCACCCCAGCGAAAGCUGGGGUCC (SEQ ID NO: 2960);	

- GTCATTCCCGCACAGGCGGGAATCC (SEQ ID NO: 2961) or
GUCAUCCCCGCACAGGCGGGAUCC (SEQ ID NO: 2962); and
- GTCATTCCCGCGCAGGCGGGAATCT (SEQ ID NO: 2963) or
GUCAUCCCCGCAGGCGGGAUCU (SEQ ID NO: 2964).

5

As shown in the examples below, the genes encoding the endonucleases according to the present disclosure are typically surrounded by one or several repeats of their own recognition sequences in the genome of the bacteria. Thus, the skilled person would be able to easily identify the recognition sequence of any endonuclease according to the present disclosure by identifying
10 repeating nucleotide sequences/motifs located near (i.e., just before/upstream and/or after/downstream) the gene encoding the endonuclease. Such recognition sequence is expected to have some level of sequence identity with the recognition sequences disclosed herein.

The method according to the present disclosure comprises incubating the single-stranded nucleic acid molecule with the composition defined herein for a period of time and under
15 conditions suitable for cleavage of the single-stranded nucleic acid molecule by the endonuclease.

In an embodiment, the period of time is at least 5 minutes. In further embodiments, the period of time is at least 10 or 15 minutes. In yet further embodiments, the period of time is at least 20, 30 or 45 minutes. In embodiments, the period of time is from about 15 minutes to about
20 2 hours, from about 20 minutes to about 90 minutes, from about 30 minutes to about 60 minutes, or from about 45 to about 60 minutes.

In an embodiment, the conditions for incubation comprise a temperature of about 10, 15 or 20°C to about 60, 55 or 50°C. In further embodiments, the conditions for incubation comprise a temperature of about 25 or 30°C to about 40 or 45°C, such as a temperature of about 35 to about
25 40°C, e.g., about 36, 37 or 38°C.

In an embodiment, the conditions for incubation comprise the presence of a suitable amount of a metal, such as a divalent metal. Examples of divalent metals include magnesium, manganese, cadmium, calcium, cobalt, nickel, zinc, iron and copper. In a further embodiment, the divalent metal is magnesium, manganese or nickel, preferably magnesium. If the divalent metal
30 is not present in the initial composition comprising the endonuclease, a suitable amount of the divalent metal is added to the reaction mixture prior to and/or during the incubation period. The divalent metal may be in the form of a salt, such as the salts listed above. In an embodiment, the divalent metal is magnesium and is in the form of magnesium chloride (MgCl₂). In an embodiment, the concentration of metal salt (e.g., MgCl₂) present during the incubation period is about 1 mM
35 to about 100 mM, for example about 1 mM to about 50 mM, about 5 mM to about 20 mM, about 5 to about 15 mM, or about 10 mM.

In an embodiment, the conditions for incubation comprise a pH of about 5 to about 10. In further embodiments, the conditions for incubation comprise a pH of about 6 to about 9, of about 6.5 to about 9, of about 7 to about 9, of about 7.5 to about 8.5, or of about 7.6 to about 8.2. In an embodiment, the conditions for incubation comprise a pH of about 7.9 or 8.0.

5 The amount of the endonuclease relative to that of the single-stranded nucleic acid molecule may be adjusted to obtain a suitable cleavage efficiency. In an embodiment, the [concentration of endonuclease] / [concentration of single-stranded nucleic acid molecule] ratio is at least 0.00005, 0.0001, 0.001, 0.01, 0.05, 0.1 or 0.5. In an embodiment, the [concentration of endonuclease] / [concentration of single-stranded nucleic acid molecule] ratio is from 0.01 to 100,
10 from 0.05 to 50 or from 0.1 to 10.

In another aspect, the present disclosure provides a method for rendering a single-stranded nucleic acid susceptible to cleavage by the endonuclease described herein, the method comprising incorporating the nucleotide sequence defined above (recognition motif) into the single-stranded nucleic acid. The incorporation of the nucleotide sequence (recognition motif)
15 may be achieved by adding the nucleotide sequence defined above (or a portion thereof) at the 5'-end, 3'-end or within the single-stranded nucleic acid, and/or by introducing one or more mutations (e.g., substitutions) within the sequence of the single-stranded nucleic acid to obtain the desired nucleotide sequence (recognition motif). Methods to modify nucleic acids are well known in the art and include, for example, cassette mutagenesis, PCR site-directed mutagenesis
20 and genome-editing technologies using nucleases such as zinc finger nucleases (ZFNs) (Gommans *et al.*, *J. Mol Biol*, **354**(3): 507-519 (2005)), transcription activator-like effector nucleases (TALENs) (Zhang *et al.*, *Nature Biotechnol*, **29**: 149-153 (2011)), the CRISPR/Cas system (Cheng *et al.*, *Cell Res.*, **23**: 1163-711 (2013)), and engineered meganucleases (Riviere *et al.*, *Gene Ther.*, **21**(5): 529-32 (2014)).

25 The present disclosure also provides a method for expressing the endonuclease defined herein in a cell, the method comprising introducing a nucleic acid encoding the endonuclease into the cell. The cell may be a procaryotic or eucaryotic cell. The nucleic acid may be an mRNA or cDNA molecule, naked or incorporated into a vector or plasmid, and it may be incorporated into the cell using any suitable methods for introducing nucleic acids into a cell (e.g., transfection,
30 transformation, etc.).

In an embodiment, the cell comprises a single-stranded nucleic acid that is cleaved by the endonuclease.

In another aspect, the present disclosure provides a kit or commercial package comprising the endonuclease or composition described herein. In an embodiment, the kit or package further
35 comprises instructions setting forth a method for cleaving a single-stranded nucleic acid with the endonuclease, such as the method described herein. The kit or package may further comprise

various components such as solutions or buffers (e.g., a reaction buffer) such as those described herein, containers, vials, etc.

EXAMPLES

The present disclosure is illustrated in further details by the following non-limiting examples.

5

Example 1: Materials and methods

Gene expression and purification (FIG. 1A). The NMV0044 gene was amplified from *N. meningitidis* 8013 2C4.3 by Phusion PCR using primers containing the *NdeI* and *XhoI* restriction sites, then cloned into pET15-MHL expression vector, generating a recombinant GIY-YIG small protein A (SsnA) with a 6xHis-tag in N-terminal. The E64A, Y6A and Y17A mutants were also generated using site-directed mutagenesis by PCR in conserved residues expected to be critical for enzymatic activity (FIG. 1B). Expression was induced at exponential phase with 1 mM IPTG for 4 hours at 37°C, followed by overnight incubation at 23°C. Cells were pelleted, lysed by lysozyme treatment and sonication, and the soluble fraction was purified by affinity chromatography with Ni-NTA resin (Thermo). >95% pure SsnA was obtained with 300 mM imidazole elution buffer (FIG. 1A). The protein was stored at -80°C.

SsnA(NMV0044) gene Nm2C4.3 (SEQ ID NO:1)

atgcagcctgcggttatatttagcaagccaacgtaatggcacgttatcatctgattgggtgcaacgtattaccaacat
 agggaaacatttgattgagggatttacacgggtacaacgttactatgctggttggtatgaactgcacctacgatggagagcgcaatt
 actcgggaaaaacagttgaagaaatggaacagggccttgaaattgcaactgattgaagaaaataatgtttcttggcaggatttatggtt
 tgatattatttag

GeneID : 12395323 Altname : NMV_RS00225 Accession_number: WP_002216166.1 Prot_locus : CAX49033. EMBL_accession: FM999788.1. The SsnA gene is known as NMB0047 or NMB_RS00250 (CDS WP_002216166.1) in the reference strain *N. meningitidis* MC58, and NMA0292 or NMA_RS01540 in *N. meningitidis* Z2491 (CDS WP_002246543.1).

SsnA(NMV0044) protein Nm2C4.3 (SEQ ID NO:2)

MQPAVYILASQRNGTLYIGVTSDLVQRIYQHREHLIEGFTSRYNVTMLVWYELHPTMESAITREK
 QLKKWNRAWKLQLIEENNVSWQDLWFDII

For SsnA homologs from different species, NEIELOOT_01219 (EFE49965.1) was taken from *Neisseria elongata* subsp. *glycolytica*, WP_011213498 was taken from *Legionella pneumophila* subsp. *pneumophila* str Paris and WP_011271370 was taken from *Rickettsia felis* URRWXC12. Expression and purification of 6xHis-recombinant proteins were done by affinity

35

chromatography using a nickel resin. Expression and purification of GST-recombinant proteins were done by affinity chromatography using a glutathione resin (FIG. 1A).

SsnA(EFE49965.1) *N. elongata* (SEQ ID NO:3)

MQPAVYILASQRNGTLYIGVTSNLTQRVYQHREHLVQGFTNQHHVTLLVWYELHSTMEHAITR

5 EKQLKKWNRQWKLRLIEEKNPSWQDLWFEIHK

SsnA(WP_011213498) *L. pneumophila* Paris (SEQ ID NO:4)

MEEKQYVVYILASKAYGTLYTGITSNLVQRIYQHKKGLEAGFTKRYNVHRLVYYEIHSDVYEAIT

REKRIKKWNRQWKINLIEQKNPQWLDLSIGLC

SsnA(WP_011271370) *R. felis* URRWXCal2 (SEQ ID NO:5)

10 MYWVYILCSDRNGTLYIGITNNILRRITYEHKQKIIKGFTAKYNIKLVTETFDIKEALAREKALKK
WNRAWKIKLIEKINLRWEDLGKCISGFPPTRE

Electrophoretic mobility shift assay (EMSA). Gel shift assays, or EMSA, were performed by diluting the proteins in Diluent A (NEB) to the indicated working concentrations. 5'carboxyfluorescein-tagged oligos corresponding to genomic regions of *N. meningitidis* MC58 were synthesized from Sigma. When needed, complementary oligonucleotides were annealed by mixing equimolar amounts in annealing buffer (10 mM Tris-HCl pH8, 50 mM NaCl, 1 mM EDTA), incubating them 5 minutes at 95°C and letting them slowly cool down. For the gel-shift assays, proteins were mixed with the fluorescent oligonucleotides in a reaction buffer containing 50 mM NaCl, 10 mM Tris-HCl, 100 µg/ml BSA, pH7.9. The mixes were incubated at 37°C for 30 minutes before adding native loading dye. Samples were resolved on native 10% TBE-acrylamide gels and imaged with a Typhoon FLA9500 scanner. For branched DNA binding assays, a similar approach was used but the gel was stained with GelStain (Biotium) and imaged on a GelDoc (BioRad) since the oligonucleotides were not fluorescently labelled.

Nuclease assays. Nuclease assays were performed similarly to gel-shift assays, with the addition of 10 mM MgCl₂ in the reaction buffer. Reactions were stopped by adding formamide loading dye and incubating 3 minutes at 95°C. Samples were resolved on denaturing 17.5% TBE-Urea (8M) acrylamide gels, and imaged with a Typhoon FLA9500 scanner. The sequences of the DNA constructs used in the studies described herein are depicted in the table below.

Name	Sequence
ssDNA	aattcaaaaatattccgtcattcccgcgaaagcgggaatctagaacgtaaaatctaaagaaccggtttatccga taagttccgcaccgacagacctgg (SEQ ID NO: 2965)
ssDNArev	ccaggtctgtcgggtcggaaactatcgataaaacggttcttagattttacgttctagattcccgtttcgcggga atgacggaatatatttgaatt (SEQ ID NO: 2966)
dsDNA	aattcaaaaatattccgtcattcccgcgaaagcgggaatctagaacgtaaaatctaaagaaccggtttatccga taagttccgcaccgacagacctgg (SEQ ID NO: 2965)

ssDNA75	ttttatcaaattcaaaaatattccgtcattcccgcgaaagcggaatctagaacgtaaaatctaagaaaccgt (SEQ ID NO: 2967)
ssDNA37	aaaaatattccgtcattcccgcgaaagcggaatcta (SEQ ID NO: 2968)
RNA37	aaaaauauuccgucauucccgcgaaagcggaucua (SEQ ID NO: 2969)
ssDNA28	ccgtcattcccgcgaaagcggaatcta (SEQ ID NO: 2970)
ssDNA75(T→U)	uuuuuaucaaaaucaaaaauauuccgucauucccgcgaaagcggaucuaagaacguaaaaucua aagaaaccgu (SEQ ID NO: 2971)
NTS#1	gaggtttcgggcaacttctaaccgtcattcccgcgagcggaatccagaccattggacagcagcaatattc aaagattatctgaaagtccgggattc (SEQ ID NO: 2972)
NTS#2	gtcattcccacgaacctacattccgtcattcccacgaaagtgggaatctagaatctctaaagcttcagtaaccttt gaatattgctgtgttctaaggt (SEQ ID NO: 2973)
NTS#3	atcccgtcattcccgcactttccgtcattcccgcgaaagcggaatctagaatctcggactttcagataatcttga atattgctgtgttctaaggtc (SEQ ID NO: 2974)
NTS#4	taaatgacccaacttaaaatcccgtcattcccgcgagcggaatctaggtctgtcggtcggaaacttatcg ggtaaacgggttcttgagattttg (SEQ ID NO: 2975)
NTS#5	gtcattcccacgaacctacattccgtcattcccacgaaagtgggaatccagtgcgttgagttcagtcattccaat aaattgccttagtattgaatgtc (SEQ ID NO: 2976)
Holliday_junc	1- Aggaacctcgagggatccgtcctagcaagccgctgctaccggaagcttctcgaggttct (SEQ ID NO: 2977) 2- Aggaacctcgagaagcttccggtagcagcgagagcggtggtgaattcctcgaggttct (SEQ ID NO: 2978) 3- Aggaacctcgaggaattcaaccaccgctcttctcaactgcagtctagactcgaggttct (SEQ ID NO: 2979) 4- Aggaacctcgagtctagactgcagttgagagcttgctaggacggatccctcgaggttct (SEQ ID NO: 2980)
D_loop	1- ataggaggaacctcgagggatccgtcctagcaagccgctgctaccggaagcttctcgaggttctaa gag (SEQ ID NO: 2981) 2- ctcttaggaacctcgagaagttttttttttttttttttttttccctcgaggttctctat (SEQ ID NO: 2982) 3- ttttttttttttctccggtagcagcggttgctaggacgga (SEQ ID NO: 2983)
Fork	1- aggaacctcgagggatccgtcctagcaagccgctgctaccggaagcttctcgaggttct (SEQ ID NO: 2984) 2- tccttgagctccctaggcaggatcggtcgagagttgacgtcagatctgagctccaagga (SEQ ID NO: 2985) 3- aggaacctcgagaagcttccggtagcagcg (SEQ ID NO: 2986) 4- tctcaactgcagtctagactcgaggttct (SEQ ID NO: 2987)

Pseudo-Y	1- aggaacctcgaggatccgtcctagcaagccgctgctaccggaagcttctcgaggttct (SEQ ID NO: 2988) 2- tccttgagctccctaggcaggatcggtcgagagtgacgtcagatctgagctccaagga (SEQ ID NO: 2989)
ssDNA75_mut	ttttatcaaattcaaaa <u>atattccgtcattcccgcgaaagcggaatctagaacgtaaaatcta</u> aagaaaccgt* (SEQ ID NO: 2990)
ssDNA and dsDNA (<i>N.elongata</i>)	ttcaccctggtacccctgaatttcgctcattcccgcgcaggcggaatccagactcttaggttcaggaatctt (SEQ ID NO: 2991)
ssDNA and dsDNA (<i>L.pneumophila</i>)	gagcattggattatgttgatagatggttgcgcctacgcgggcatgactgatttttgcctcatgtacaaaac (SEQ ID NO: 2992)
ssDNA and dsDNA (<i>R.felis</i>)	tacattttattattatttttctggatcccgccctcgcggggatgacattgagcaggtttaccagtctacg (SEQ ID NO: 2993)

*Underlined nucleotides were individually mutated to a mix of all nucleotides but the native one.

Branched DNA constructs were replicated from Fukui, K. et Al. (2018). *FEBS Lett*, 592: 4066-4077.

5

Example 2: SsnA expression and purification

The NMV0044 gene from *Neisseria meningitidis* 8013 2C4.3 was cloned in pET15-MHL, allowing its expression with a 6xHis-tag in N-terminal. Purification was done by affinity chromatography with a nickel resin. The resulting protein was diluted in reaction buffer and used directly for *in vitro* assays to determine its enzymatic activity.

10

Example 3: SsnA is a specific single-stranded nuclease

SsnA possesses a single functional domain, belonging to the GIY-YIG nuclease superfamily. Its nuclease activity was therefore tested on different nucleic acids (**FIGs. 2B-C**). Using dsDNA does not reveal significant binding nor cleavage activity, even when it contains the recognition pattern (**FIG. 2B**, dsDNA). On the other hand, 100 % of a 100 nt ssDNA containing the recognition pattern is cleaved at a unique specific position, meaning it is an endonuclease (**FIG. 2B**, ssDNA). However, SsnA has no activity at all on the exact reverse complement of the ssDNA that is efficiently cleaved (**FIG. 2B**, ssDNA¹), indicating its sequence specificity. These results show that SsnA is a specific single-stranded endonuclease with no detectable activity on dsDNA. The single GIY-YIG domain of SsnA mediates both its cutting and binding activities, without the need for any co-interacting protein or complex. GST-tagged SsnA exhibits the same

20

ssDNA binding and nuclease activity that the His-tagged protein (**FIG. 2B**), suggesting that the addition of tags or protein domains do not alter the enzymatic activity of SsnA, making it modifiable.

Removing magnesium from the reaction mix or adding the metal chelator EDTA completely abolishes the nuclease activity of SsnA (**FIG. 2A**), but allows DNA binding which can be visualized by electrophoretic mobility shift assays (EMSA) (**FIG. 2B**). Therefore, SsnA is a metal-dependent nuclease. SsnA cannot bind dsDNA, or ssDNA without the recognition sequence, which explains why it cannot cleave these substrates. It binds 100 % of ssDNA harboring the recognition sequence.

The glutamic acid residue in position 64 of SsnA is well conserved within the GIY-YIG superfamily, where it often corresponds to a metal cofactor (e.g., magnesium) binding site (**FIG. 1B**). The E64A mutant was therefore expressed and purified along with the WT protein (**FIG. 1A**), and its cutting and binding activity was assessed against ssDNA harboring the recognition sequence (**FIG. 3**). SsnA E64A can efficiently bind ssDNA, but has completely lost its nuclease activity (**FIG. 3**). This conserved amino acid is therefore confirmed as the magnesium binding site allowing cleavage of single-stranded nucleic acids. The tyrosine residues in position 6 and 17 of SsnA are also well conserved within the GIY-YIG superfamily as part of their catalytic core (**FIG. 1B**). The Y6 and Y17 mutants were also expressed, purified (**FIG. 1A**), and assayed for their ssDNA binding and nuclease activities (**FIG. 2B**). Similarly to the E64 mutation, mutation of amino acids Y6 or Y17 completely prevents the nuclease activity of SsnA. In addition, both mutants show reduced binding to ssDNA, indicating that they are involved in both binding and cutting of ssDNA.

The 100nt ssDNA that is specifically bound and cleaved by SsnA contains a ≈30nt repeated sequence (named NTS) frequently found in the *Neisseria* genomes and predicted to form a stem-loop (hairpin) secondary structure. Reducing the ssDNA length to 37nt did not significantly affect the binding activity nor the nuclease activity of SsnA (**FIG. 2C**). Further reducing the ssDNA length to 28nt, keeping only the repeated sequence, completely prevented cutting by SsnA while allowing efficient binding (**FIG. 2C**). Therefore, the NTS repeat sequence is necessary and sufficient for binding by SsnA, but insufficient for cleavage by SsnA which requires extra nucleotides in 5' and 3'.

To assess whether SsnA could bind to and cleave all types of nucleic acids, its activity on both RNA and ssDNA with uracils instead of thymines was also tested (**FIG. 2C**). While SsnA is unable to interact with RNA, some binding and cleavage activities are detected on ssDNA(T→U). Therefore, SsnA is a DNase that specifically interacts with deoxyribonucleic acids.

To confirm that the enzymatic activity of SsnA is specific to the NTS, the assays were repeated on ssDNA from different regions of the *Neisseria* genome which harbor the full-length NTS with different flanking sequences (**FIG. 3**). With these substrates, the binding activity of SsnA was fully preserved, while its cutting activity was reduced to varying extents. These results confirm

that SsnA binds ssDNA specifically to the repeated sequence, and cuts ssDNA just outside of the NTS with a sequence specificity.

The specific cutting site of SsnA on ssDNA was determined precisely by running the cleaved product from a 75 nt ssDNA next to fluorescent oligonucleotides of increasing length (FIG. 4A), which confirmed that SsnA cuts several nucleotides upstream of the NTS repeated sequence (FIG. 4B). Since the binding and cleavage specificities of SsnA are not identical, individual nucleotides in and around the repeated sequence from the same 75 nt ssDNA were mutated and the binding and cleavage specificities of SsnA on the mutated ssDNA were tested (FIG. 4B). Binding activity was completely lost when the ssDNA mutation occurred immediately downstream of the palindromic region. Moreover, binding activity was significantly reduced in mutated ssDNA with mutations located within the palindromic region of the repeated sequence, suggesting that SsnA binds the stem part of the stem-loop formed by this repeated sequence. In contrast, the cleavage activity of SsnA was highly dependent on the sequence immediately upstream of the palindromic region, but still within the conserved part of the repeated sequence. Therefore, SsnA binds to the ssDNA hairpin formed by the NTS repeated sequence, and needs to interact with the sequence immediately upstream of the hairpin to be able to cleave ssDNA a few nucleotides upstream.

Several proteins from the GIY-YIG superfamily interact with branched DNA to accomplish various functions. Since branched DNA have dsDNA and ssDNA regions just like the hairpin formed by the NTS repeat, it was tested whether SsnA could bind different DNA constructs. No binding activity was detected against Holliday junctions, D-loops, forks and pseudo-Y DNA in absence of the recognition sequence (FIG. 5). Therefore, these results show that SsnA specifically interacts with the stem-loop structure formed by the *Neisseria* NTS repeated sequence, and does not bind other secondary structures.

Several parameters of SsnA's nuclease activity were determined using the 100 nt ssDNA containing its recognition sequence (FIG. 6). SsnA requires a metal cofactor such as magnesium, manganese or nickel to cleave ssDNA, with an optimal activity at 10 mM MgCl₂. (FIG. 6A-B). Manganese however only allows partial nuclease activity. The enzyme is active in temperatures ranging 14 to 54°C, with an optimal temperature of around 37°C (FIG. 6C). Most of the ssDNA substrate is cleaved within the first 15 minutes of reaction (FIG. 6D). ssDNA cleavage is dose-dependent and requires subnanomolar amounts of protein (FIG. 6E).

Example 4: SsnA belongs to a novel family of single-stranded specific endonucleases

Tens of thousands of hypothetical proteins throughout the bacterial domain contain a single GIY-YIG domain covering most of their sequence. Apart from the YhbQ nuclease which only shares about 22% identity with SsnA, none of these single-domain GIY-YIG proteins have been characterized to our knowledge. Since SsnA exhibits a unique enzymatic activity, a search for

potential homologs was performed. Using a cut-off of 50% identity and a strict size threshold, thousands of predicted homologs were found in bacteria from all orders (**FIG. 7A**). This family is referred to herein as the Specific single-stranded nuclease (**Ssn**) family. The inferred phylogeny of SsnA shows two distinct branches or clusters of Ssn proteins. These subfamilies are referred to herein as SsnA and SsnB. Some strains or species of bacteria may encode for multiple Ssn homologs. The sequences of several representative Ssn proteins having at least 50% identity with SsnA are depicted in **FIGs. 11A-JJJJ**.

In *N. meningitidis* and closely related species such as *N. elongata*, the gene encoding SsnA is surrounded by several repeats of its own recognition sequence. Although the enzyme cannot cut genomic DNA in a double-stranded form, SsnA might also act as a mobile genetic element, similarly to transposases. It is shown here that the genes encoding SsnA homologs in unrelated species are located near highly similar sequences, which could be an indication of their nuclease specificity (**FIG. 8**).

To truly define the Ssn proteins as a novel family of single-stranded endonucleases, SsnA homologs from three unrelated bacterial species, *Neisseria elongata*, *Rickettsia felis* and *Legionella pneumophila* (sequence alignment in **FIG. 9A**), were expressed and purified. Apart from the *N. elongata* protein which shares 82% identity with the *N. meningitidis* protein, these SsnA homologs share relatively little (47-60%) identity with each other (**FIG. 9B**) and are spread across the Ssn protein phylogeny. They are therefore representative of the diversity of this novel Ssn protein family. They all contain the conserved features of the GIY-YIG domain, namely the N-terminal tyrosine (Y) residues, the central arginine (R) residue as well as the C-terminal glutamic acid (E) residue.

For each homolog, a DNA sequence similar to the NTS repeat found around the *N. meningitidis* SsnA gene was identified in close vicinity of their respective genes (**FIG. 9C**). Despite major differences, each sequence exhibits a 15-20nt imperfect palindrome likely forming a hairpin in ssDNA form. It was then tested whether each SsnA homolog could cut their respective sequence as ssDNA (**FIG. 10**). SsnA from *N. elongata* efficiently cuts its specific ssDNA at a single location, similarly to the *N. meningitidis* nuclease (**FIG. 10A**). SsnA from *L. pneumophila* also cuts its specific ssDNA, but at two locations (**FIG. 10B**), similarly to the SsnA from *R. felis* (**FIG. 10C**). All of the homologs are unable to cleave dsDNA, even in presence of their respective recognition sequences, confirming that they are specific single-stranded nucleases. Moreover, the tested homologs exhibit different specificities and potentially different ssDNA cutting mechanisms, as shown by the variable number of cutting sites obtained.

Therefore, it may be concluded that Ssn proteins can indeed be grouped as a novel family of enzymes, more specifically specific single-stranded endonucleases (Ssn) with a potentially wide array of unique specificities.

Example 5: SsnA is able to cleave ssDNA *in cellulo*

Mutants of *N. meningitidis* 8013 2C4.3 were generated; a knock-out strain (KO) in which the SsnA gene was deleted, and a complemented strain (Compl) in which the gene is overexpressed. The uptake and homologous recombination of DNA by natural transformation was then compared between these strains. Briefly, 200 ng of linearized plasmid DNA is incubated for 2 hours with bacterial suspensions of OD₆₀₀ = 1. The transformations mixes are then serially diluted, and the appropriate dilutions are plated on non-selective media (GCB agar) and selective media containing either 5 µg/ml chloramphenicol (Cm) or 75 µg/ml spectinomycin (Sp). Colonies are counted after an overnight incubation at 37C with 5% CO₂, and the rate of transformation is determined, corresponding to the number of resistant CFUs divided by the number of total CFUs.

Two plasmids were assayed, each containing an antibiotic resistance gene (chloramphenicol or spectinomycin) flanked by sequences homologous to the *N. meningitidis* genome, which allow integration to the host genome by double recombination. One plasmid did not contain the recognition sequence of SsnA, while the other one contains several repeats of it in the homologous regions.

Neisseria species are naturally competent, meaning they can readily update DNA from their environment by transformation and integrate it into their own genome if there is sufficient homology. Of note, only one strand of DNA (ssDNA) is imported in the cytoplasm of the bacteria during natural transformation. When transforming *N. meningitidis* mutant strains with a plasmid that does not contain SsnA's recognition sequence, a slight decrease of transformation efficiency was observed in the SsnA KO strain (**FIG. 12A**). To the contrary, if the plasmid contains SsnA's recognition sequence, the knock-out strain is transformed much more efficiently than the strains expressing SsnA (**FIG. 12B**). These results indicate that SsnA is able to cleave ssDNA *in cellulo* with a high specificity, such as transforming DNA containing its recognition sequence.

Although the present invention has been described hereinabove by way of specific embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims. In the claims, the word "comprising" is used as an open-ended term, substantially equivalent to the phrase "including, but not limited to".

WHAT IS CLAIMED IS:

1. An isolated endonuclease specific for single-stranded desoxyribonucleic acid molecules, the isolated endonuclease having a length of 60 to 150 amino acids or less and comprising a single GIY-YIG domain.

5 2. The isolated endonuclease of claim 1, having a length of 70 to 130 amino acids.

3. The isolated endonuclease of claim 1, having a length of 80 to 120 amino acids.

4. The isolated endonuclease of any one of claims 1 to 3, wherein the GIY-YIG domain is of the following sequence (I):

X1-X2-X3-B1-X4-X5-X6-B2-X7-B3-X8 (I)

10 wherein

X1 is any amino acid;

X2 is any amino acid;

X3 is Y or H;

B1 is a sequence of 8 to 12 amino acids;

15 X4 is Y or H;

X5 is any amino acid.

X6 is G or D;

B2 is a sequence of 6 to 15 amino acids;

X7 is R or T;

20 B3 is a sequence of 30 to 40 amino acids; and

X8 is E.

5. The isolated endonuclease of claim 4, wherein X1 is Y, W, V, A, F, I, C, H, R, T or S.

6. The isolated endonuclease of claim 5, wherein X1 is Y.

7. The isolated endonuclease of any one of claims 4 to 6, wherein X2 is V, I, L, T, A or F.

25 8. The isolated endonuclease of claim 7, wherein X2 is V.

9. The isolated endonuclease of any one of claims 4 to 8, wherein X3 is Y.

10. The isolated endonuclease of any one of claims 4 to 9, wherein B1 is a sequence of 8 to 10 amino acids.

11. The isolated endonuclease of any one of claims 4 to 10, wherein X4 is Y.

12. The isolated endonuclease of any one of claims 4 to 11, wherein X5 is I, L, V, T, A, C, or K.
13. The isolated endonuclease of claim 12, wherein X5 is I, T or V.
14. The isolated endonuclease of claim 13, wherein X5 is I.
15. The isolated endonuclease of any one of claims 4 to 14, wherein X6 is G.
- 5 16. The isolated endonuclease of any one of claims 4 to 15, wherein B2 is a sequence of 6 to 10 amino acids.
17. The isolated endonuclease of claim 16, wherein B2 is a sequence of 6 to 8 amino acids.
18. The isolated endonuclease of any one of claims 4 to 17, wherein X7 is R.
19. The isolated endonuclease of any one of claims 4 to 18, wherein B3 is a sequence of 35 to
10 40 amino acids.
20. The isolated endonuclease of any one of claims 4 to 19, wherein X8 is E.
21. The isolated endonuclease of any one of claims 4 to 20, wherein the GIY-YIG domain is of the following sequence (II):
- X1-X2-X3-B1-X4-X5-X6-B2-X7-B4-X9-B5-X8 (II)
- 15 wherein
- X1, X2, X3, B1, X4, X5, X6, B2, X7 and X8 are as defined in claims 4 to 20;
- B4 is a sequence of 1 to 5 amino acids;
- X9 is H, Q or Y; and
- B5 is a sequence of 30 to 38 amino acids.
- 20 22. The isolated endonuclease of claim 21, wherein B4 is a sequence of 2 to 4 amino acids.
23. The isolated endonuclease of claim 21 or 22, wherein X9 is H.
24. The isolated endonuclease of any one of claims 21 to 23, wherein B5 is a sequence of 30 to 35 amino acids.
25. The isolated endonuclease of claim 24, wherein B5 is a sequence of 31 to 33 amino acids.
- 25 26. The isolated endonuclease of any one of claims 4 to 25, wherein the GIY-YIG domain is of the following sequence (III):
- X1-X2-X3-B1-X4-X5-X6-B2-X7-B4-X9-B5-X8-B6-X10 (III)

wherein

X1, X2, X3, B1, X4, X5, X6, B2, X7, B4, X9, B5 and X8 are as defined in claims 4 to 20;

B6 is a sequence of 15 to 20 amino acids; and

X10 is N or K.

5 27. The isolated endonuclease of claim 26, wherein B6 is a sequence of 16 to 19 amino acids.

28. The isolated endonuclease of claim 26, wherein X10 is N.

29. The isolated endonuclease of any one of claims 1 to 28, comprising an amino acid sequence having at least 50% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.

10 30. The isolated endonuclease of any one of claims 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 60% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.

31. The isolated endonuclease of any one of claims 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 70% similarity or identity with any one of the
15 sequences set forth in SEQ ID NOs:2-2891.

32. The isolated endonuclease of any one of claims 1 to 28, wherein the isolated endonuclease comprises an amino acid sequence having at least 80% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.

33. The isolated endonuclease of any one of claims 1 to 28, wherein the isolated endonuclease
20 comprises an amino acid sequence having at least 90% similarity or identity with any one of the sequences set forth in SEQ ID NOs:2-2891.

34. The isolated endonuclease of any one of claims 1 to 28, wherein the isolated endonuclease comprises the amino acid sequence of any one of the sequences set forth in SEQ ID NOs:2-2891.

35. A composition comprising (i) the isolated endonuclease of any one of claims 1 to 34, and
25 (ii) an aqueous saline solution or buffer.

36. The composition of claim 35, wherein the aqueous saline solution or buffer comprises a metal.

37. The composition of claim 36, wherein the metal is in the form of a metal salt.

38. The composition of claim 36 or 37, wherein the metal is magnesium, manganese or nickel.

39. The composition of claim 38, wherein the metal is magnesium.

40. The composition of claim 39, wherein the composition comprises magnesium chloride (MgCl_2).

41. The composition of any one of claims 35 to 40, wherein the single-stranded nucleic acid molecule is a single-stranded DNA molecule.

42. The composition of any one of claims 35 to 41, wherein the single-stranded nucleic acid molecule comprises a nucleotide sequence having at least 50% sequence identity with the sequence: GTCATTCCCNNNNNNNGGGAATC or GTCATTCCCGCGAAAGCGGGAATC.

43. The composition of any one of claims 35 to 42, wherein the single-stranded nucleic acid molecule comprises the following nucleotide sequence: GTCANNCCNGNNNANNNGGNNNC.

44. The composition of claim 43, wherein the single-stranded nucleic acid molecule comprises the following nucleotide sequence: GTCAYBCCMGYRHAVRCKGGVRNC.

45. The composition of claim 43 or 44, wherein the single-stranded nucleic acid molecule comprises any one of the nucleotide sequences depicted in **FIG. 8** and **FIG. 9C**.

46. The composition of any one of claims 35 to 44, further comprising the single-stranded nucleic acid molecule defined in any one of claims 41 to 45.

47. A method for cleaving a single-stranded nucleic acid molecule, the method comprising contacting the single-stranded nucleic acid molecule with the isolated endonuclease of any one of claims 1 to 34 or the composition of any one of claims 35 to 45 under conditions suitable for cleavage of the single-stranded nucleic acid molecule by the isolated endonuclease, wherein the single-stranded nucleic acid molecule comprises a recognition sequence for the isolated endonuclease.

48. The method of claim 47, wherein said conditions comprises a temperature of about 20 to about 55°C.

49. The method of claim 48, wherein said conditions comprises a temperature of about 35 to about 40°C.

50. The method of claim 49, wherein said conditions comprises a temperature of about 37°C.

51. The method of any one of claims 47 to 50, wherein said conditions comprises the presence of a metal.

52. The method of claim 51, wherein said metal is magnesium, manganese or nickel.
53. The method of claim 52, wherein the metal is magnesium.
54. The method of claim 53, wherein the magnesium is in the form of magnesium chloride (MgCl_2).
55. The method of any one of claims 51 to 54, wherein said metal is at a concentration of at least 5 mM.
56. The method of any one of claims 51 to 54, wherein said metal is at a concentration of at least 10 mM.
57. The method of any one of claims 47 to 56, wherein said contacting is for a period of at least 2 minutes.
58. The method of any one of claims 47 to 56, wherein said contacting is for a period of at least 15 minutes.
59. The method of any one of claims 47 to 58, wherein said conditions comprises a pH of about 6 to 8.
60. The method of any one of claims 47 to 58, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.00001.
61. The method of claim 60, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.01.
62. The method of claim 61, wherein the [concentration of endonuclease] / [single-stranded nucleic acid molecule] ratio is at least 0.5.
63. A method for rendering a single-stranded nucleic acid susceptible to cleavage by the endonuclease defined in any one of claims 1 to 34, the method comprising incorporating a nucleotide sequence comprising a recognition sequence for the isolated endonuclease into the single-stranded nucleic acid.
64. The method of claim 63, wherein the nucleotide sequence comprises one of the sequence defined in any one of claims 42 to 45
65. The method of claim 64, wherein the method comprises adding a nucleic acid fragment comprising the nucleotide sequence defined in any one of claims 42 to 45 at the 5'-end, 3'-end or within the single-stranded nucleic acid.

66. The method of claim 64, wherein the method comprises introducing one or more mutations within the sequence of the single-stranded nucleic acid to obtain the nucleotide sequence defined in any one of claims 42 to 45.

5 67. A cell comprising the endonuclease defined in any one of claims 1 to 34, wherein the endonuclease is heterologous to the cell.

68. A method for expressing the endonuclease defined in any one of claims 1 to 34 in a cell, the method comprising introducing a nucleic acid encoding the endonuclease into the cell.

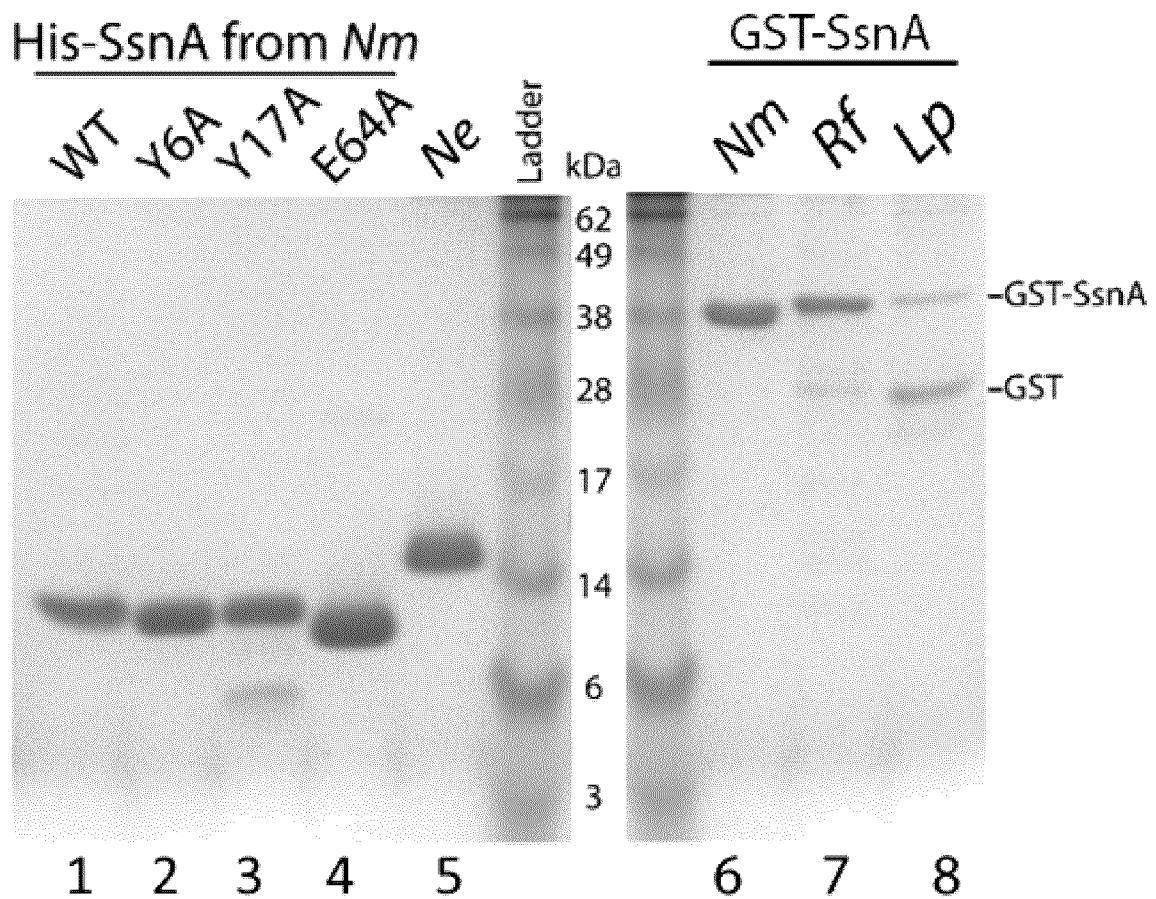
69. The cell of claim 67 or the method of claim 68, wherein the cell is a prokaryotic or eukaryotic cell.

10 70. The cell or method of any one of claims 67 to 69, wherein the cell comprises a single-stranded nucleic acid that is cleaved by the endonuclease.

71. The cell or method of any one of claims 68 to 70, wherein the nucleic acid encoding the endonuclease is present in a vector.

15 72. A kit comprising the endonuclease defined in any one of claims 1 to 34 or the composition of any one of claims 35 to 45, and instructions for cleaving single-stranded nucleic acid molecules using the endonuclease.

73. The kit of claim 72, wherein said instructions comprise the method of any one of claims 47 to 62.



- 1- *N. meningitidis* His-SsnA WT
- 2- *N. meningitidis* His-SsnA Y6A
- 3- *N. meningitidis* His-SsnA Y17A
- 4- *N. meningitidis* His-SsnA E64A
- 5- *N. elongata* His-SsnA WT
- 6- *N. meningitidis* GST-SsnA WT
- 7- *R. felis* GST-SsnA WT
8. *L. pneumophila* GST-SsnA WT

FIG. 1A

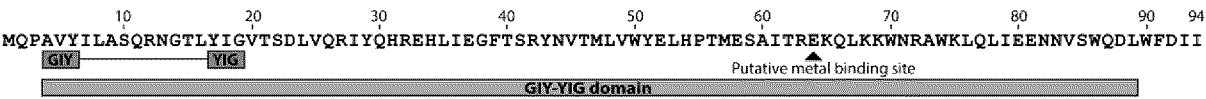


FIG. 1B

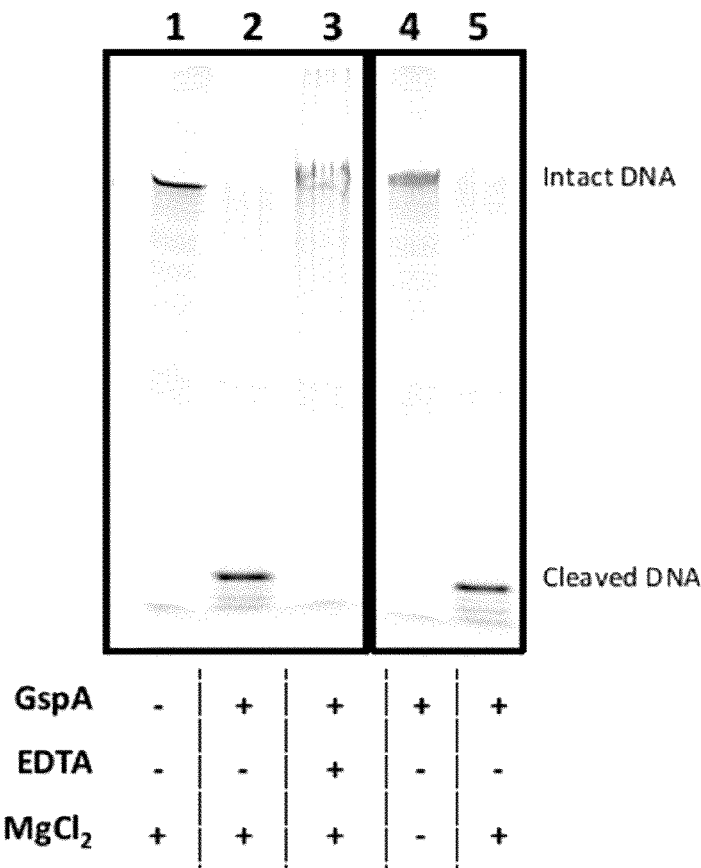


FIG. 2A

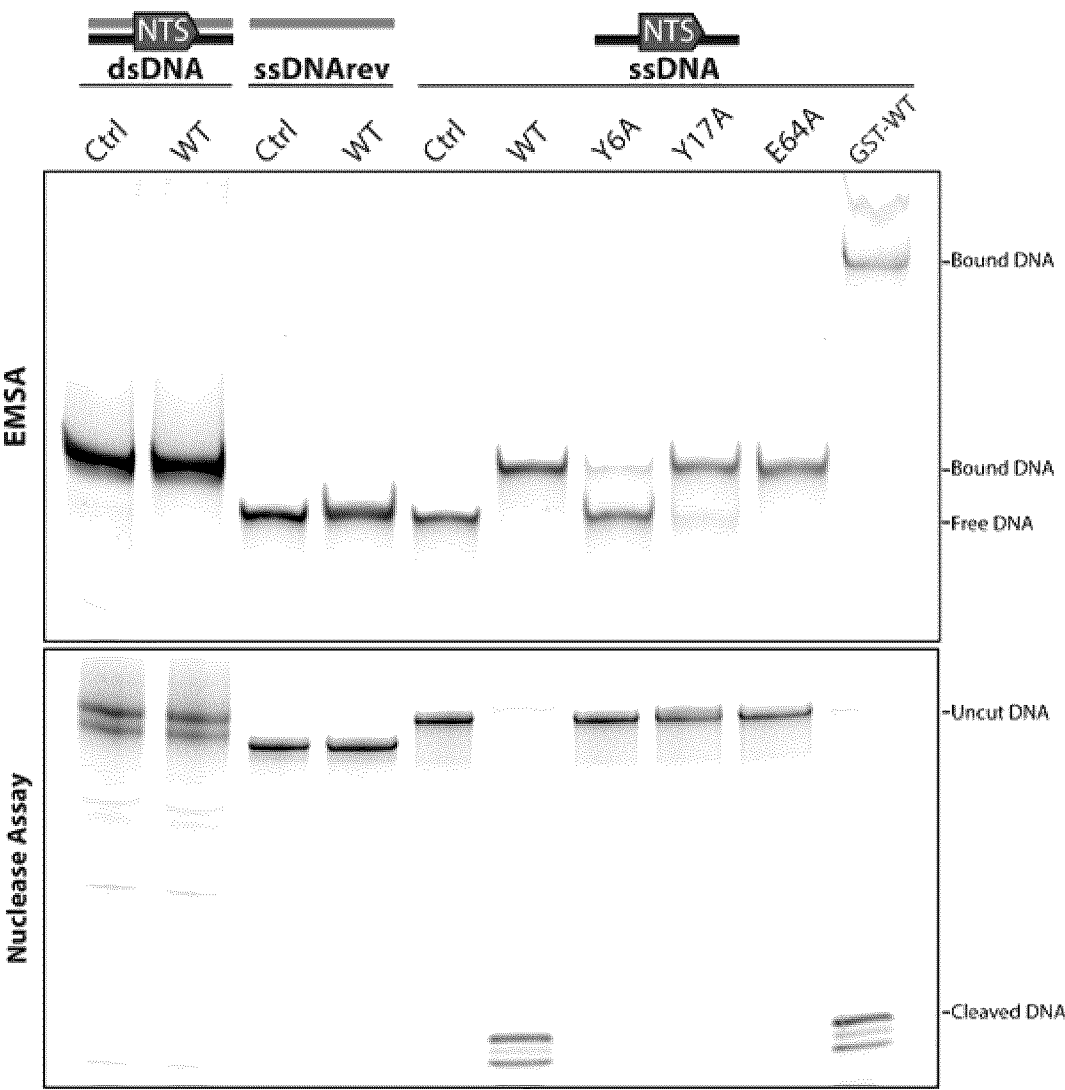


FIG. 2B

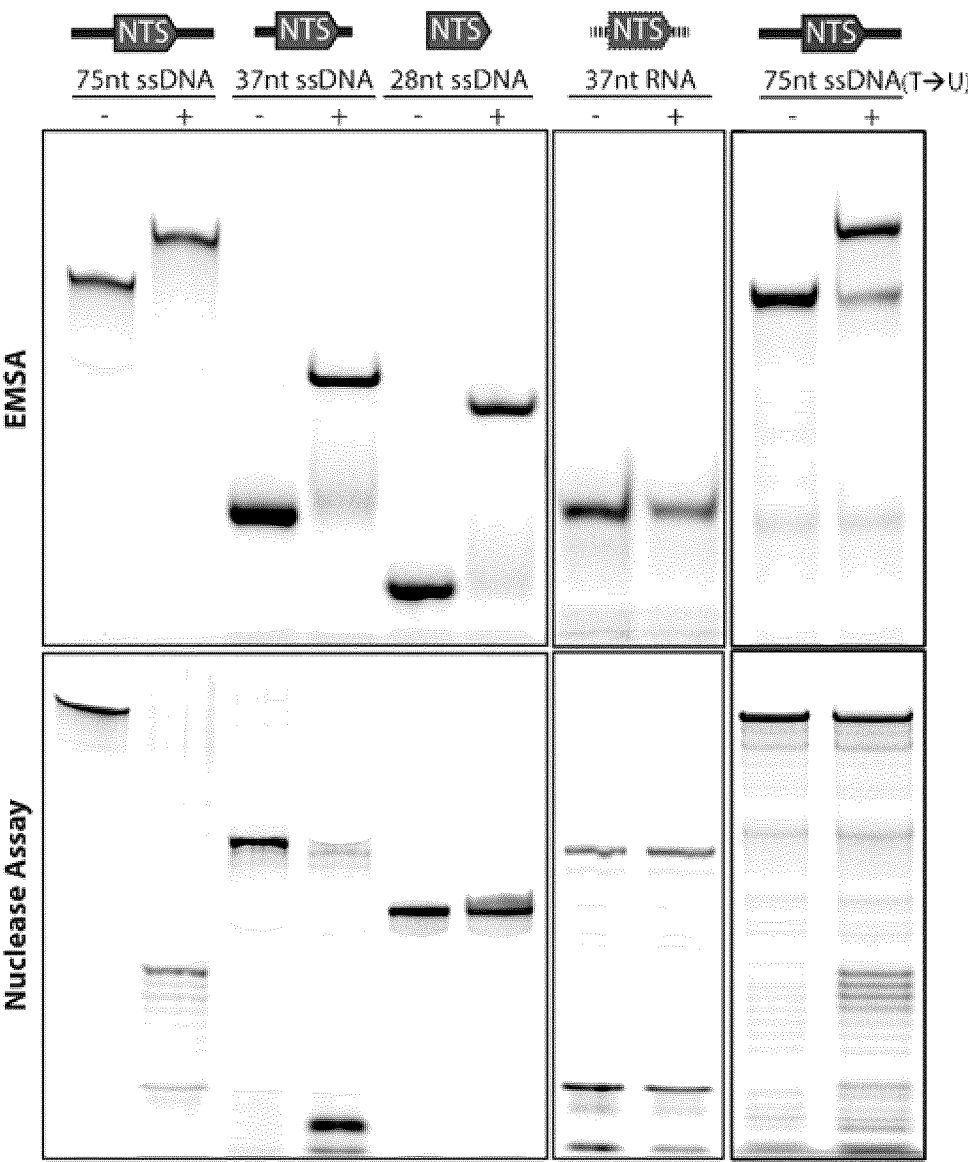


FIG. 2C

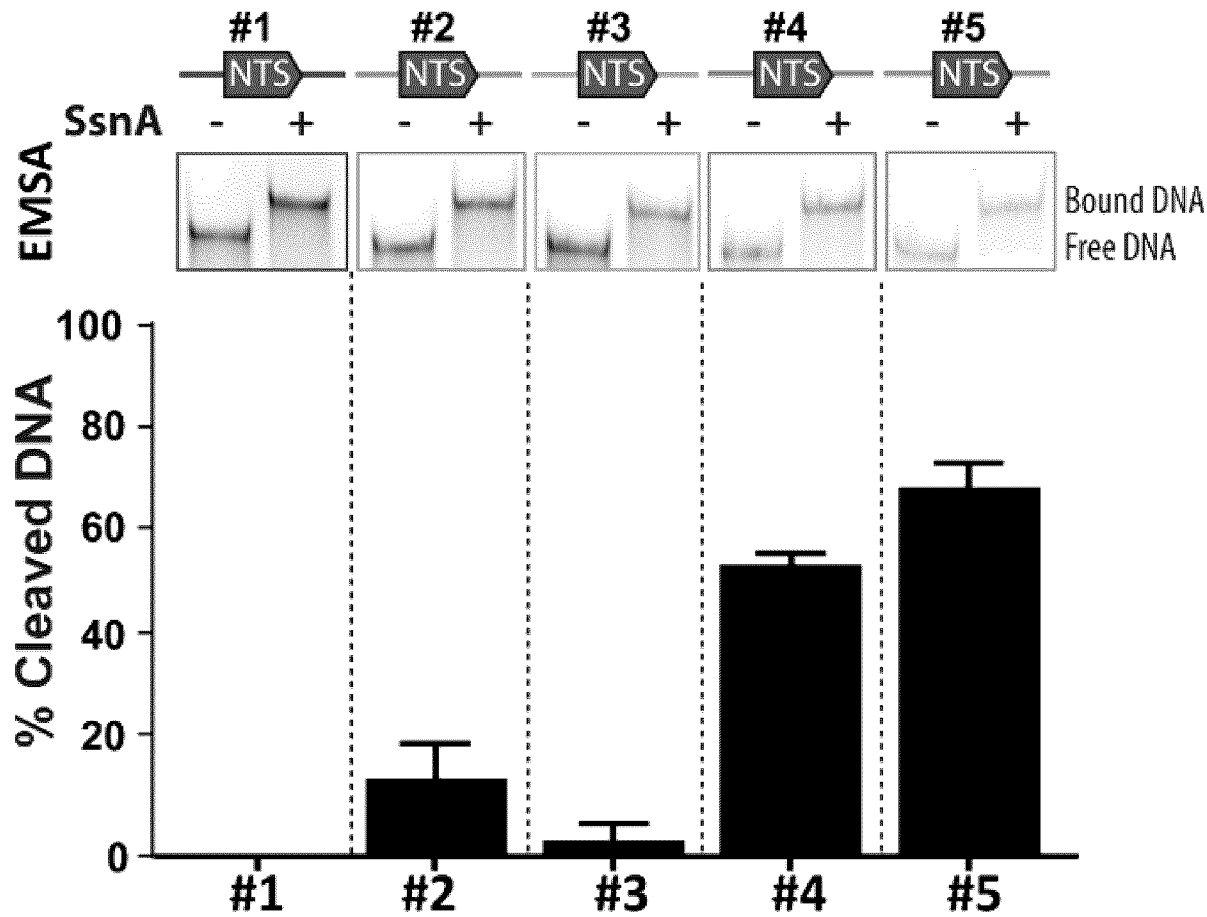


FIG. 3A

NTS#1 GAGGTTTCGGGCAACTTCTAAACCGTCATTCCCGCGCAGGCGGGAATCCAGACCATTGGACAGCAGCAATATTCAAAGATTATCTGAAAGTCCGGGATTC
NTS#2 GTCATTCCCACGAACCTACATTCCGTCATTCCCACGAAAGTGGGAATCTAGAATCTCTAAAGCTTCAGCTAACCTTTGAATATTGCTGTTGTTCTAAGGT
NTS#3 ATCCCGTCATTCCCGCCACTTCCGTCATTCCCACGAAAGCGGGAATCTAGAATCTCGGACTTTCAGATAATCTTTGAATATTGCTGTTGTTCTAAGGT
NTS#4 TAAATGACCCCAACTTAAATCCCGTCATTCCCACGCGCAGGCGGGAATCTAGGTCGTGTCGGTGCGGAACTTATCGGGTAAACGGTTTCTTGAGATTTTG
NTS#5 GTCATTCCCACGAACCTACATTCCGTCATTCCCACGAAAGTGGGAATCCAGTGCCTTGAGTTTCAGTCATTCCAATAAATTGCCTTAGTATTGAATGTC

FIG. 3B

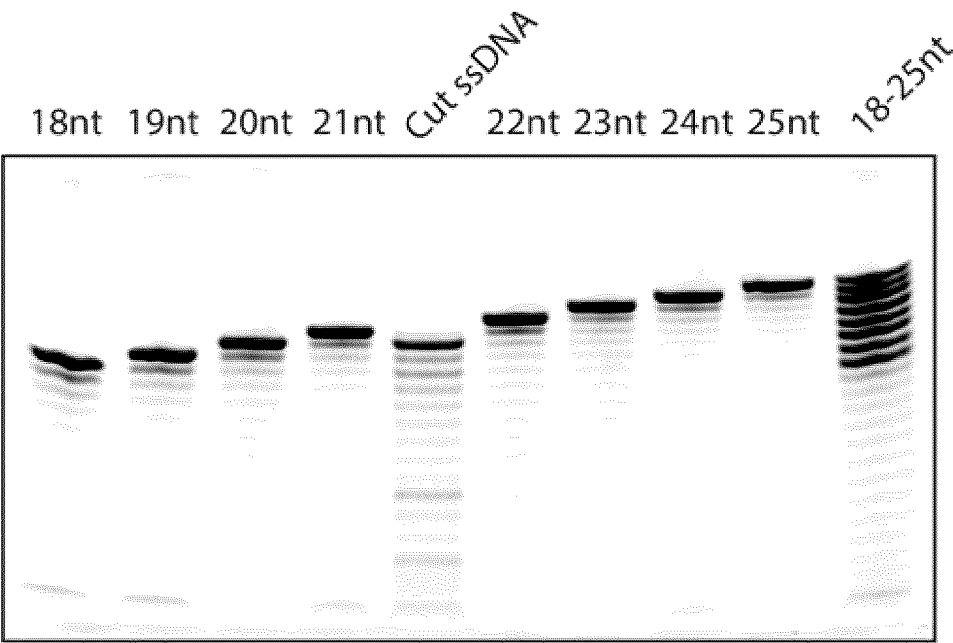


FIG. 4A

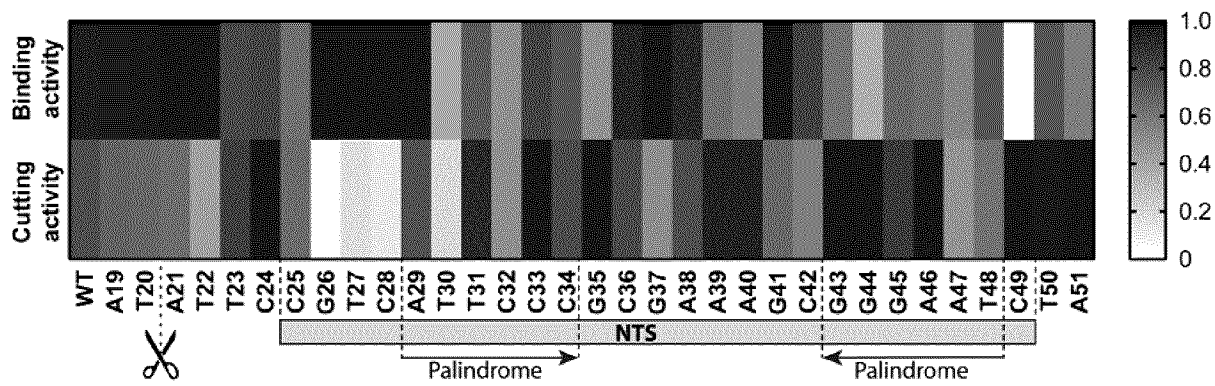


FIG. 4B

7/106

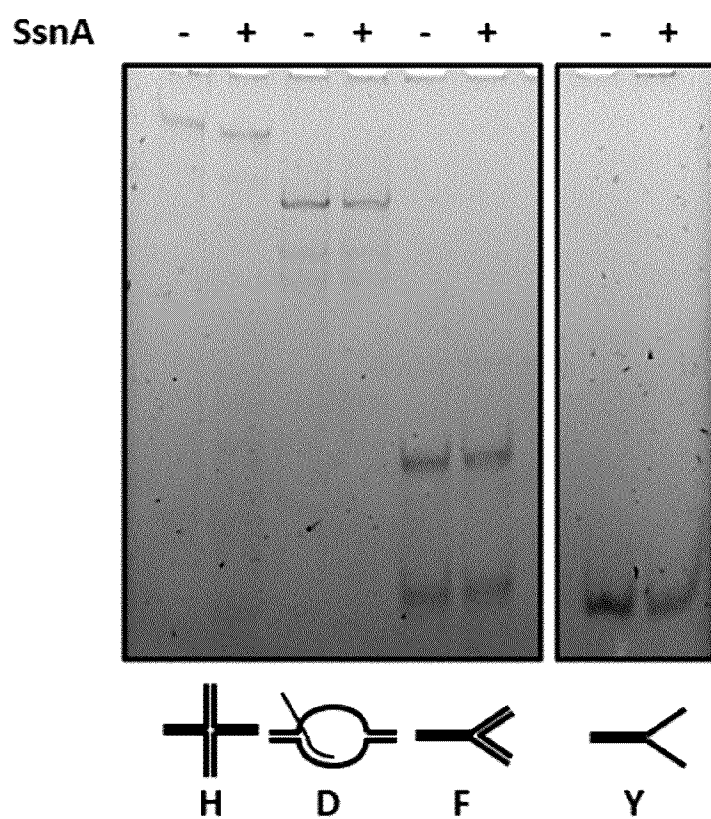


FIG. 5

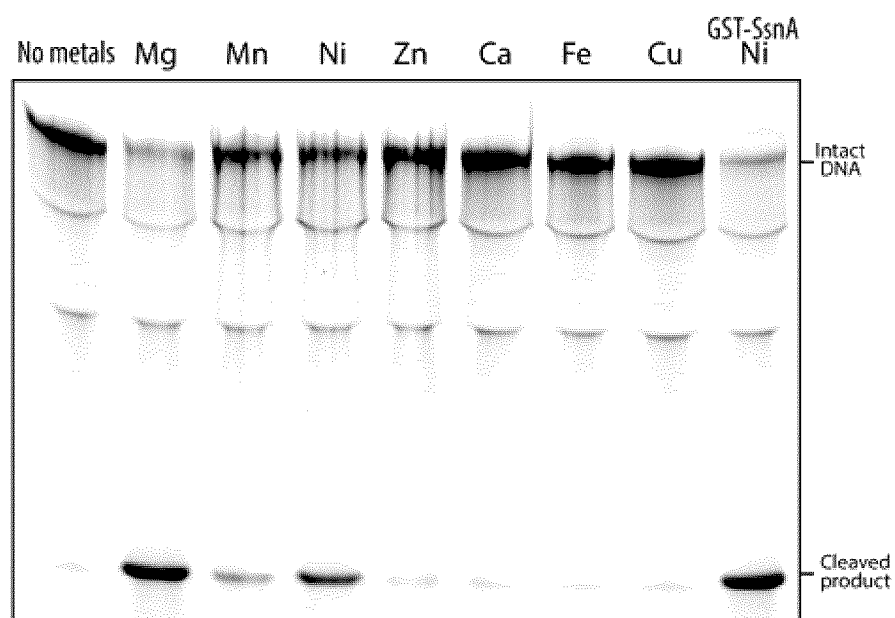


FIG. 6A

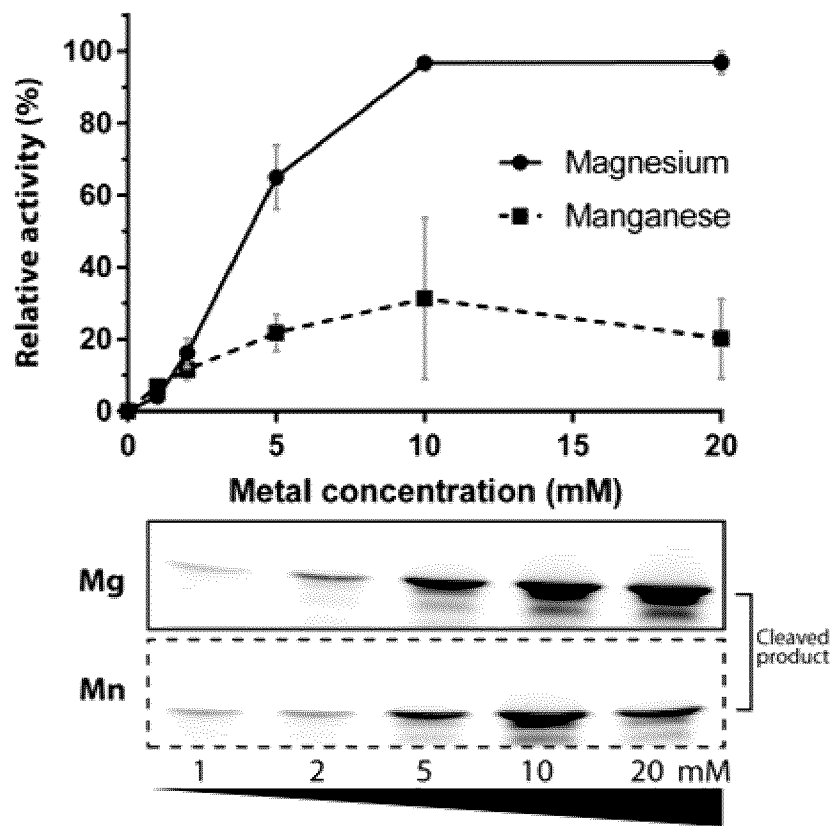


FIG. 6B

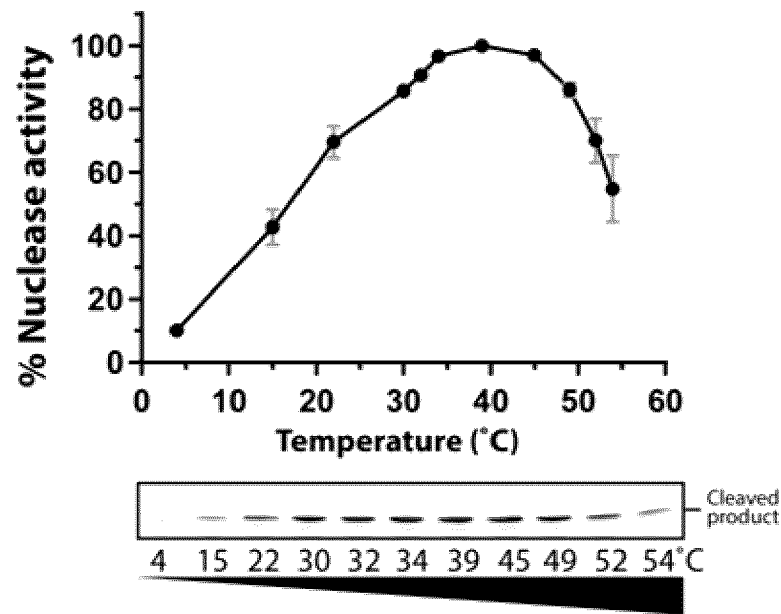


FIG. 6C

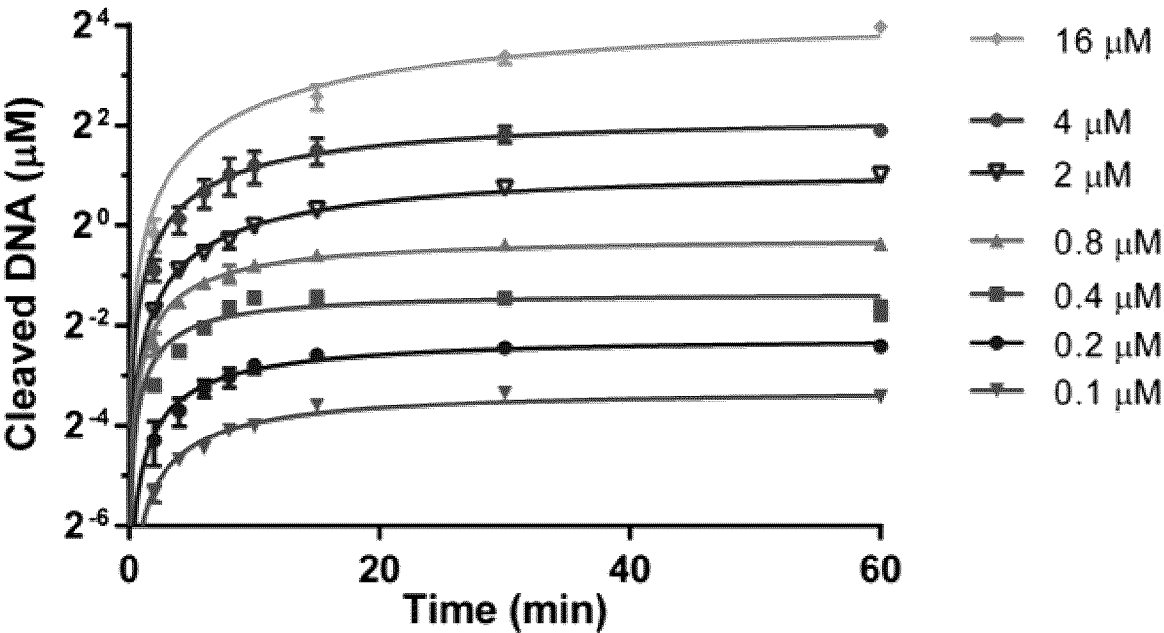


FIG. 6D

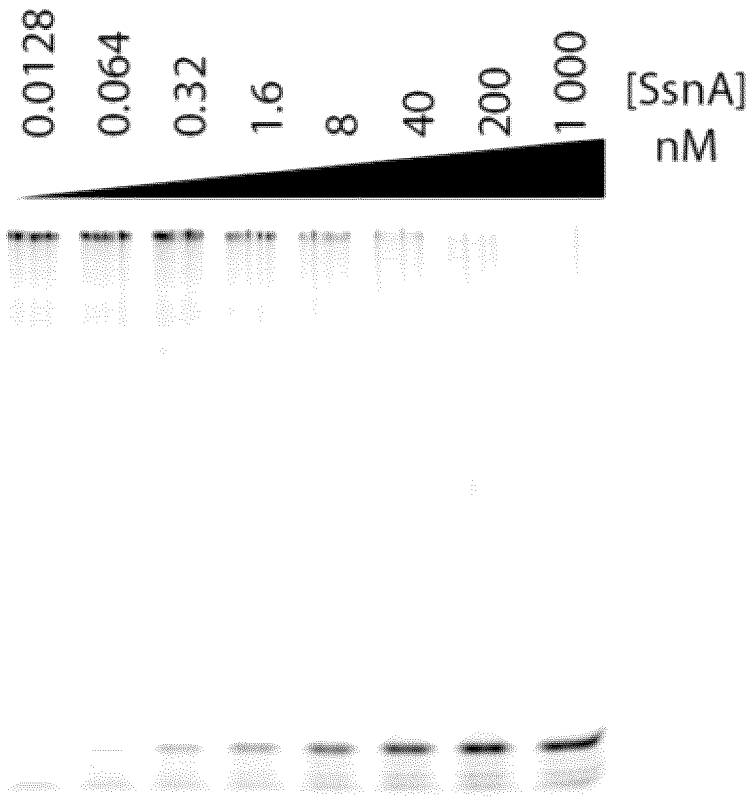


FIG. 6E

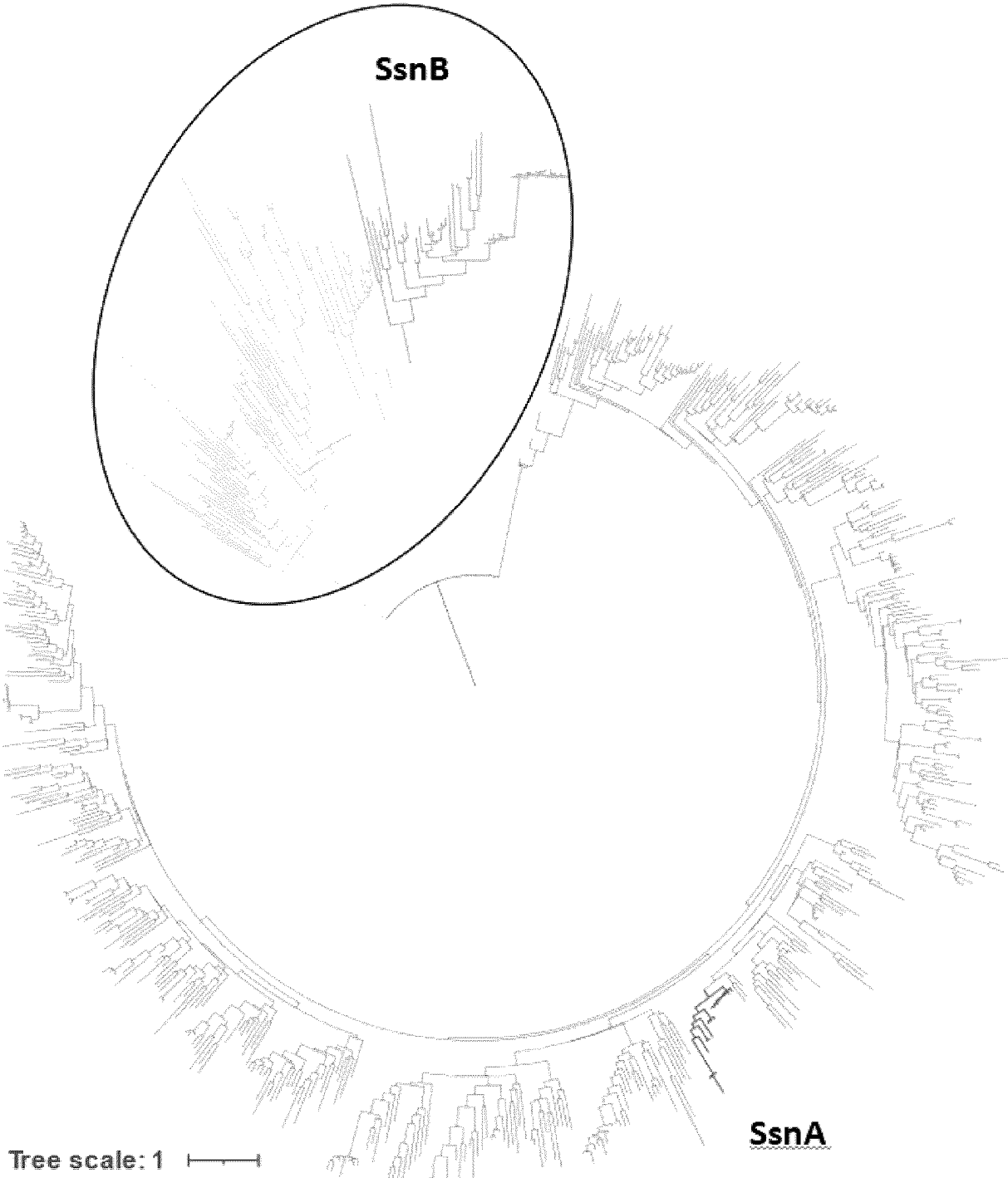
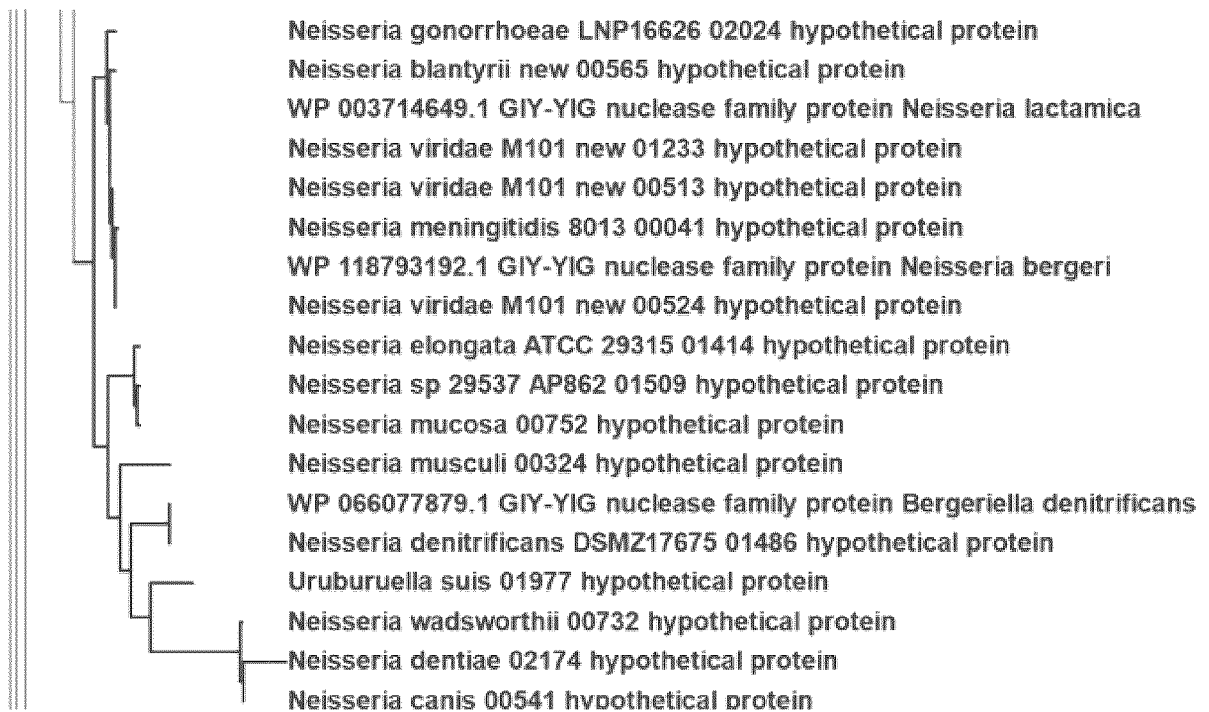


FIG. 7A

**FIG. 7B**

12/106

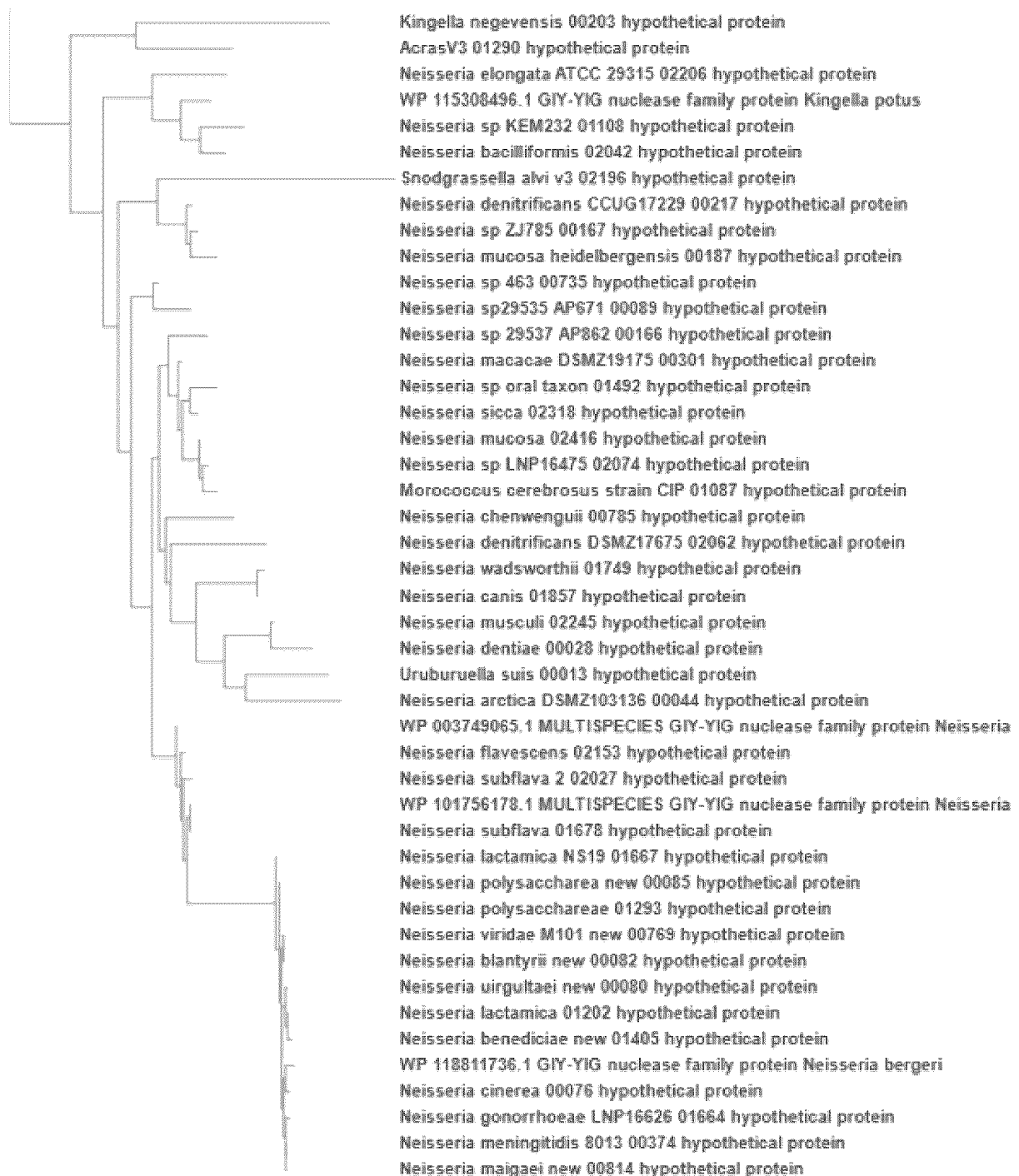


FIG. 7C

13/106

	(1)	1	10	20	35	SEQ ID NO:
<i>Ahrensia sp.</i>	(1)	GACTGTCAT	CCCCGCGCAGGCGGGGACCCACTTTT			2892
<i>Nitrosomonas sp.</i>	(1)	TCGCGTCAT	TCCCGCGAAAGCGGGAATCCCGTTTT			2893
<i>Bermanella marisrubri</i>	(1)	CCTCGTCAT	TCCCGCGAAGGCGGGAATCCATTTAG			2894
<i>Shewanella putrefaciens</i>	(1)	TTATGTCAT	TCCCGCGCAGGCGGGAATCCACAGTA			2895
<i>Hyphomicrobium denitrificans</i>	(1)	GACTGTCAT	CCCCGCGCAGGCGGGGACCCACTTTT			2896
<i>Legionella pneumophila</i>	(1)	ATCAGTCAT	GCCCGCGTAGGCGGGCAACCATCTAT			2897
<i>Sphingobium japonicum</i>	(1)	CCTCGTCAT	TCCCGCGAAAGCGGGAAGCCATCTCC			2898
<i>Francisella novicida</i>	(1)	ACTTGTCAT	TCCCGCGCAGGCGGGAATCCTTTCGA			2899
<i>Vibrio vulnificus</i>	(1)	AGCCGTCAT	TCCCGTGCACACGGGAATCCATTCAT			2900
<i>Bradyrhizobium sp.</i>	(1)	GACCGTCAT	GCCCGCGCAGGCGGGCATCCAGTACG			2901
<i>Rickettsia felis</i>	(1)	CAATGTCAT	CCCCGCGAAGGCGGGGATCCAGAAAA			2902
<i>Sphingomonas wittichii</i>	(1)	ATTGTCAT	TCCCGCGAAAGCGGGAATCCACGGTT			2903
<i>Chlorobaculum parvum</i>	(1)	TTCTGTCAT	TCCCGCGAAGGCGGGAATCCATCTGA			2904
<i>Methylothermobacter sp.</i>	(1)	TGTCGTCAT	TCCCGCGCAGGCGGGAATCTAGGCAA			2905
<i>kangiella coraensis</i>	(1)	ATTGTCAT	TCCCGCGAAGGCGGGAGTCTATTCTT			2906
<i>Oceanithermus profundum</i>	(1)	CTTCGTCAT	CCCCAGCGAAAGCTGGGGTCCAGACCG			2907
<i>Xanthomonas massiliensis</i>	(1)	GACCGTCTT	TCCCGCGCAGGCGGGGACCCAGCGCC			2908
<i>Aquaspirillum sp. LM1</i>	(1)	CTCGTCAT	TCCCGCGCAGGCGGGAATCCAGGGGC			2909
<i>Lysobacter concretionis Ko07</i>	(1)	CATCGTCAT	TCCCGCGAAGGCGGGAATCCATGGAC			2910
<i>Desulfobulbaceae bacterium A2</i>	(1)	TTTCGTCAT	GCCCGCGCAGGCGGGCATCCAGTAAG			2911
<i>Neisseria elongata</i>	(1)	CCGCGTCAT	TCCCGCGCAGGCGGGAATCCGGAAAT			2912
<i>Chitinilyticum aquatile DSM 21506</i>	(1)	CCTCGTCAT	TCCCGCGAACGCGGGAATCCAGTGAC			2913
<i>N. meningitidis</i>	(1)	GCCCGTCAT	TCCCGCGCAGGCGGGAATCCGGCTTG			2914
NMB0248	(1)	TGCCGTCAT	TCCCGCACAGGCGGGAATCCATTCTT			2915
Consensus	(1)	CGTCAT	TCCCGCGCAGGCGGGAATCCA			2916

FIG. 8

Nmeningitidis	---MQPAVYILASQRNGTLYIGVTSDLVQRIYQHREHLIEGFTSRYNVTMLVWYELHPTM	57
Nelongata	---MQPAVYILASQRNGTLYIGVTSNLTQRVYQHREHLVQGF TNQHHVTLLVWYELHSTM	57
LpneumophilaParis	MEEKQYVVYILASKAYGTYTGITSNLVQRIYQHKKGLAEGFTKRYNVHRLVYVEIHTDV	60
RfelisURRWXcal2	----MYWVYILCSDRNGTLYIGITNNILRRTYEHKQKIIKGFTAKYNIKLVYTEEFTDI	56
	****.*. **** *:.*.: : * *:.*.: : :*** :.: : **: * . :	
Nmeningitidis	ESAITREKQLKKWNRAWKLQLEENNVSWQDLWFDII-----	94
Nelongata	EHAITREKQLKKWNRQWKLRLIEKNPSWQDLWFEIHK-----	95
LpneumophilaParis	YEAITREKRIKKWNRWKINLIEQKNPQWLDLSIGLC-----	97
RfelisURRWXcal2	KEALAREKALKKNRAWKIKLIEKINLRWEDLGKICISGFPPTRE	100
	*:;*** :***** *:;.*; : * * * * :	

FIG. 9A

1: Nmeningitidis	100.00	81.91	59.57	50.54
2: Nelongata	81.91	100.00	57.45	46.81
3: LpneumophilaParis	59.57	57.45	100.00	52.69
4: RfelisURRWXcal2	50.54	46.81	52.69	100.00

FIG. 9B

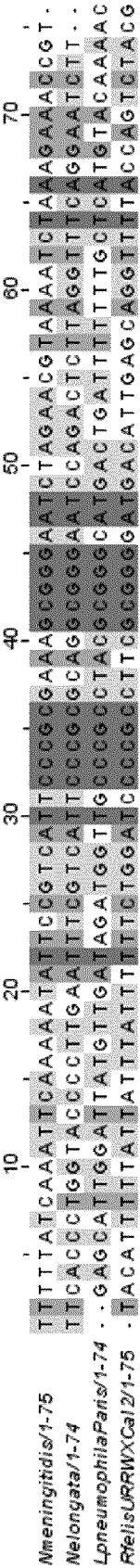


FIG. 9C

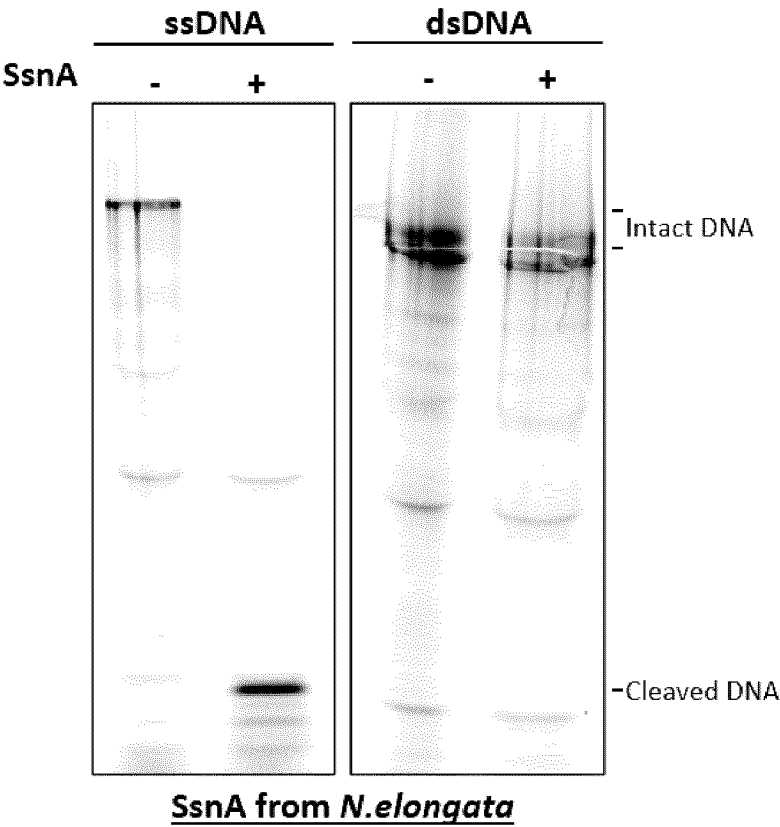


FIG. 10A

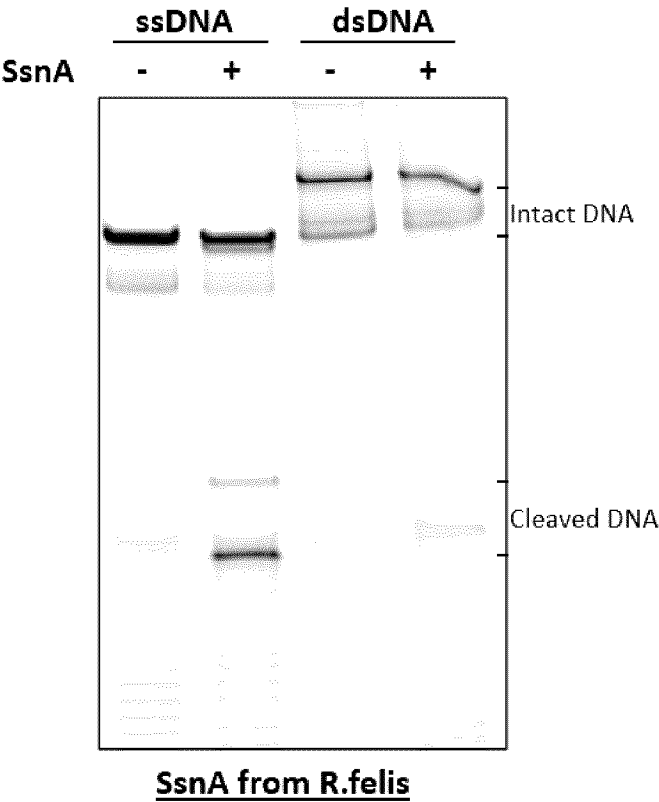


FIG. 10B

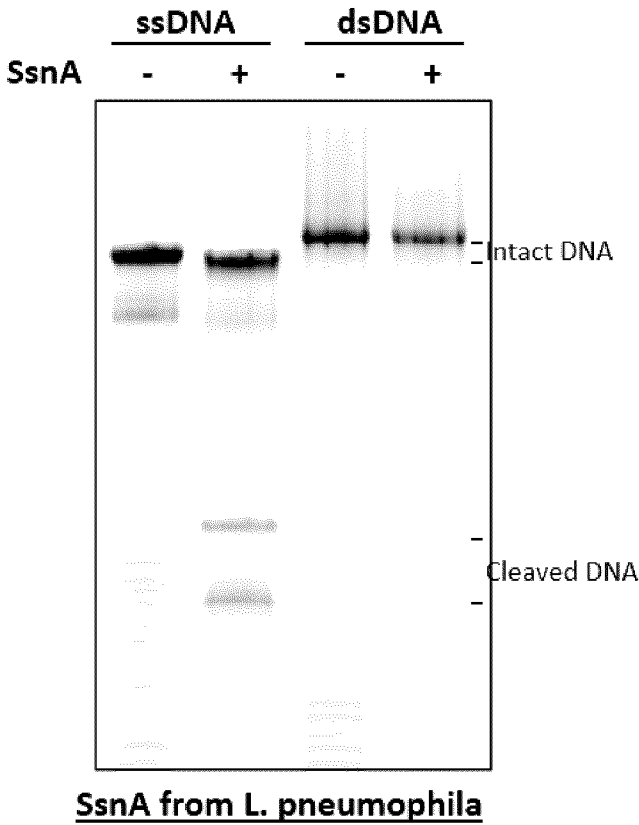


FIG. 10C

<WD_RS01615--Wolbachia_endosymbiont_of_Drosophila_melanogaster--74.71 (SEQ ID NO:6)
MKLSYIYILTANKANSTLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVYFEEFEDISLAIARERRLKEWPRQWKINLINQKNPSWIDLYYLALR
<IL_RS13510--Idiomarina_loihiensis_L2TR--55.84 (SEQ ID NO:7)
MSWAVYLIKCADNTFTYGTGTTDIERRVNEHNTDDKKGAKYTRARRPVVALVWSEEHPNRTKACQREYEIKRWPRAKKVELTGGI
<SCA_RS00630--Staphylococcus_carnosus_subsp._carnosus_TM300--54.32 (SEQ ID NO:8)
MENKYFIYIVRCRDGSLYTGTYNDVIARVEKHNSGKGAKYTKTRRPVKLVYQEAFTKSEAMKREYQIKTYTRMQKLQLIERGVIS
<MAFF_RS18255--Mesorhizobium_japonicum_MAFF_303099--67.42 (SEQ ID NO:9)
MTGYVYMTASQKRGTIYIGVTNDLGRVPEHKAGQGSNFTSRYGVRQRLVWYEEHFDITDAIQREKSLKRWPRLWKIELIEKANPEWFELFRGTGW
<MAFF_RS09315--Mesorhizobium_japonicum_MAFF_303099--65.56 (SEQ ID NO:10)
MTGYVYMTASQKRGTIYIGVTNDLGRVPEHKSGEGSRFTSQYGVQRLVWYEEYFDITDAIQREKSLKRRRLQWKIDLIEKANPDWFELFRGTGW
<AMB_RS20825--Magnetospirillum_magneticum_AMB-1--67.42 (SEQ ID NO:11)
MTPLQRHFYVYMLSSVSGVLYGVGTNDLARRLQEHRSIADGFTKRYHCHRLVWFEMTGDVNAAIAREKQIKGWRRSKKEHLIRETNPWADLAQDF
<RF_RS05395--Rickettsia_felis_URRWXCal2--79.57 (SEQ ID NO:12)
MKQPIVYIVTNKNGILYTGVTNSLIKRIYEHKNYTIKGFSKYNCKILVFEIHQTMNSAINREKQIKAGSRDKKIKLIEQTINWDDLYSEIIL
<RF_RS05645--Rickettsia_felis_URRWXCal2--81.18 (SEQ ID NO:13)
MYWVYILCSNRNGTLYIGITNNILRRYEHKQKIIKGFTAKYNI IKLVYTEEFTDIKEALAREKALKKWNRAWKIKLIEKINLRWEDLGKCSIGFPPTRE
<RF_RS02295--Rickettsia_felis_URRWXCal2--78.82 (SEQ ID NO:14)
MYWVYILCSNKHGTYIGITNNIIRRAYEHKQKIIKGFTSKYNI IKLVYTEEFTDIKEALTREKNLKKWKNWKIELIEKANPEWKDLIE
<RF_RS05075--Rickettsia_felis_URRWXCal2--75.29 (SEQ ID NO:15)
MYWVYILCSNKHGTYIGITNNIIRRAYEHKQKIIKGFTSKYNI IKLVYTEEFTDIKEALTREKNLKKWKNWKIELIEKANPEWKDLIE
<CPS_RS13870--Colwellia_psychrerythraea_34H--82.80 (SEQ ID NO:16)
MKQPAVYILANKNGTLYIGVTSDLVKRIWQHKNHVVSGFTTEEYEVHQLVYFEQLDDMENAITREKVIKKWNRAWKIRLIEQNNPKWLDLYNDIV
<RBE_RS06750--Rickettsia_bellii_RML369-C--75.27 (SEQ ID NO:17)
MKQYVYILASKRNGTLYTGITSNIIKRIYEHKNKIIKGFTSKYNTNKLIVYFEQYHYVIAAISREKLLKEYKRQWKINLIESQNPWLDLYEKIIL
<RBE_RS02025--Rickettsia_bellii_RML369-C--78.82 (SEQ ID NO:18)
MYWVYILCSNRNGTLYIGVTNNIIRRSYEHKQKIIKGFTAKYNI IKLVYTEEFADIEALREKALKWKNWKIELIEKANPEWKDLIE
<NWI_RS05430--Nitrobacter_winogradskyi_Nb-255--64.44 (SEQ ID NO:19)
MAGLVPAIHVLQNSCMKGGYVYFLTNRPNGLYVGVGTNDLVRRVFEHRSRGFVAGFTKRHCLKRLVYFETFDDIRNAIQREHNIKHWPRTWKVRTIVAVNPEWD
DLYPTIV
<SARO_RS04060--Novosphingobium_aromaticivorans_DSM_12444--75.00 (SEQ ID NO:20)
MRDERQPCVYILASQPRGTLYIGVTSNLIGRLWQHRESMIPGFTSQYAVHRLVYFEQFDEMIPAITREKQLKRWHRQWKINLIESANPDWHDLPGLGLGPPA
PRTPRDGP

FIG. 11A

<SARO_RS04780--Novosphingobium_aromaticivorans_DSM_12444--65.12 (SEQ ID NO:21)
MEKGGVYIMANRYRGGMYVGVTSDLVRRVWQHREGVGGSSHVDGFKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAALRALR
GVVER
<SARO_RS14240--Novosphingobium_aromaticivorans_DSM_12444--65.12 (SEQ ID NO:22)
MERGGVYIMANRYRGGMYVGVTSDLIRRVWQHREGVGGSSHVDGFKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAAMALR
GVVER
<SARO_RS04465--Novosphingobium_aromaticivorans_DSM_12444--62.79 (SEQ ID NO:23)
MDREQRGGVYIMADHYRGSYVGVTSHLARRVVFQHRNGTGSYCARHSLVRLVWAERGENIEACIVHEKRLKRWREWKFDLIEGANPDWADLYDMLV
<RPC_RS15955--Rhodopseudomonas_palustris_BisB18--74.19 (SEQ ID NO:24)
MAYVYVYLLASSKHGTYLGLVGTNDLVRRCEYHRTKAAPGFTSRYGVDKLVWFYIYDNIIVTAITREKQLKKWRDWDKIRLIEEVNPEWSDLYPGISH
<RPC_RS15930--Rhodopseudomonas_palustris_BisB18--74.19 (SEQ ID NO:25)
MAYVYVYLLASSKHGTYLGLVGTNDLVRRCFEHRSKAATGFTSRYGIDKLVWFYIYDNIIVTAITREKQLKKWRDWDKIRLIEEVNPEWTDLYPGISN
<RPC_RS06360--Rhodopseudomonas_palustris_BisB18--69.77 (SEQ ID NO:26)
MAGLVPAIHVGQICWMRAGYVYFLTNRPNGILYVGVTSDLARRVYEHRTGAVEGFTTRYGLKRLVYFEVYDDIRLAIQREHNKIHWSRTWKVRLILANNPEWN
DLFETL
<NHAM_RS00505--Nitrobacter_hamburgensis_X14--74.16 (SEQ ID NO:27)
MKQPCVYIVASKRNGTLYTGVGTANLQRRIFEHRESLVAGFSKYGCKILVWYEAHETMTDAITREKQIKAGSRVKKLALIEGLNPEWNDLFETLF
<NHAM_RS07200--Nitrobacter_hamburgensis_X14--75.56 (SEQ ID NO:28)
MSYWIYILTSAPGTYLGVGTNNLVRRRIYEHREDLADGFTHKYNVKKRLVYFEAHDVTMSAIQREKNIKHWSREWKIDLIVRDNPTWSDLYEDIVR
<SALA_RS03675--Sphingopyxis_alaskensis_RB2256--73.03 (SEQ ID NO:29)
MHADFQPCAYIMASARHGTYLGTVTSHLIQRICQHRDGSFDFGFTKRYGVKTLVWFELAGTMHAAITREKQIKNWQRKWKIELIEAANPDWRDLAEELGLPPLG
A
<SALA_RS11335--Sphingopyxis_alaskensis_RB2256--72.73 (SEQ ID NO:30)
MTRPGYLYLMASARNGTLYLGVTSDDLARVWQHRNEVVESFTKKYGCHTLVWYEAFFDIQARQRELQMKKWKRAWKIELIERDNPQWLDLDFRDLSL
<SALA_RS03855--Sphingopyxis_alaskensis_RB2256--67.78 (SEQ ID NO:31)
MGRGGVYIMTNKRGGVLYIGVTADIAARIHQHKTGRGSAFCRKCGLTMLVLAEEYPTIEEAIAREKAMKAWKRQWKIELIEASNPNWDDLAPNIL
<SALA_RS08380--Sphingopyxis_alaskensis_RB2256--54.93 (SEQ ID NO:32)
MSFWAYMLHCRGGAFYTGHTDNLFPYRVAQHESGAIQGFADKLPVTLVWSQEFPTIEALEAERRIKGWSRAKKMALIRGDWDSISNLAKGNGPSTGSGRTE
RD
<SALA_RS04950--Sphingopyxis_alaskensis_RB2256--54.67 (SEQ ID NO:33)
MAFWTYILLCADGLYTGHTDNLDFRIAQHQSLIEGFTSSRLPVRLMWSQDFATRYEALDAEARIKKWSRAKEALIRGDWEAVSHFAKPPKERPTSPFASS
EVEMPEVETAPSPIGVTSGLDTSNRLGFSGGGNP
<AIE_RS03290--Rickettsia_canadensis_str._McKiel--82.86 (SEQ ID NO:34)
MGLVENMLVYIYEHETMDSAITREKQIKSGSRAKKLNLIE
<SWOL_RS12580--Syntrophomonas_wolfei_subsp._wolfei_str._Goettingen_G311--57.89 (SEQ ID NO:35)
MSGEEGYIYILSCSDGTYLGTSTNNLLARLEKHNGEGARYRGRRPVTLVYSEFAASKSDALKRERQIKMTRAAKLALIK
<MMAR10_RS04585--Maricaulis_maris_MCS10--64.13 (SEQ ID NO:36)
MTRRKAKPAGATLAKHPCVYILASQRNGTLYIGVTGNIAMRVMLHREGKGSFAFMRYGVKHLVWYEFHQDFGFAIEREKRLKVWRKWKWLTIEQMNPGWEDL
YESIQMWAPMD
<MMAR10_RS04555--Maricaulis_maris_MCS10--73.03 (SEQ ID NO:37)
MFHVYLLTSHKNGTLYCGVTRELLARVYAHKVRKAGGFTARYGVDQLVWFESHELVNNAITREKRIKRWRAWKVNIETNPDPWRDLWEMGGIDPREVVD
GMRWESDKV
<MMAR10_RS10255--Maricaulis_maris_MCS10--65.17 (SEQ ID NO:38)
MARHPCVYIILASQRNGTLYIGVTGNIAMRIMLHKEGKGSFAFTRYKVDRVLVWYEFHQDFGFAIEREKRLKNWHRQWKINLIERTKPGWEDLFETIQMWAPIG
<MMAR10_RS10670--Maricaulis_maris_MCS10--65.56 (SEQ ID NO:39)
MLAPEAVYLLFSRPHGALYVGRTKDLIWRWQHRTGAIPGHTKRHGKQLGWYEWHDGFESAAREYLIKRWRAWKIALIEAMNPGWKDLWFDIRGDFDEQV
PIVTRLED
<MMAR10_RS12185--Maricaulis_maris_MCS10--75.58 (SEQ ID NO:40)
MYVYVYLLASRRHGTLYCGVTNALLRRVYEHKSAQVSGFTKKHHVHRLVWWEAHEDIGEAIAREKRIKWRWDWKVAILIEQDNPWDWDLVPLGGTDPSEIAGE
GIRRSYD
<MMAR10_RS04280--Maricaulis_maris_MCS10--71.11 (SEQ ID NO:41)
MEDREGTIAVYMMASRYDGLYTGVTANLLWRVRQHKDGSGASGFTRKYKCTSLVWWEHPGVIVEAIRREKRIKKWPRRWKVNLIEDNPNHWRDLWDELWAVFP
NPGLKL
<NEUT_RS09255--Nitrosomonas_eutropha_C91--78.26 (SEQ ID NO:42)
MKDPCVYIMANKRNGTLYTGVTSDLVQRAWQHRESVVEGFTRYQCKLLVWYEQHGSMSAILREKQIKAGSRARKLALIEAQNPDWRDLVYEEIV
<MLG_RS11615--Alkalilimnicola_ehrlichii_MLHE-1--75.27 (SEQ ID NO:43)
MNQPCVYIMTNGRNGTLYVGVTSNLVQRVWQHKSQGVDFGFTTRYGLHRLVWYEVHETMTSAIGREKQLKSGNRRKVGLESMPNPGWRDLYHDII
<CPHA266_RS12005--Chlorobium_phaeobacteroides_DSM_266--67.44 (SEQ ID NO:44)
MRDKAYCVYILANQANNVLYVGMTNDLQRRVLEHKSRLVDGFSKKYNVDKLVYFEETSDVLRATAREKEIKKWRREKKNLLIATMNPPEWDLACECS
<SPUTCN32_RS13085--Shewanella_putrefaciens_CN-32--75.53 (SEQ ID NO:45)
MQPYVYIMASQFNGTLYIGVTSHEQRIWQHKNVIGDFTHKYKVDQLVYFEVHASMYDAICREKQLKKWRREWKINLINSQNPKWEDLWYTVIVGSPHR
<FJOH_RS15120--Flavobacterium_johnsoniae_UW101--66.32 (SEQ ID NO:46)
MKPGFVYIITNKYQTVVYVGVTSNLPKRILEHKNKKYKPSFSKYNLNLVYVEQFQWIEDAISREKQIKAGSREAKNNLIRSINPKWEDLFDEIENIMTM
<SWIT_RS12310--Sphingomonas_wittichii_RW1--80.68 (SEQ ID NO:47)
MEKQPSVYIILASGRHGTLYIGVTSDLPRGVHQRHRESLIKGFSTRYGVSRVLVHYEGFDDMTAITREKQLKKWNRRAWKIELVEAQNPFYWDLVMTFGFPFPGP
<SWIT_RS13200--Sphingomonas_wittichii_RW1--78.41 (SEQ ID NO:48)
MDRQPCVYIILASDRNGTIYIGVTSDLPRKRLYQHRSGTLDGFTARYGVAKLVRYELFEDMSTAIAREKQLKRWHRSWKLNLIQDNADWHDLSIGLGFEP
<SWIT_RS03315--Sphingomonas_wittichii_RW1--74.73 (SEQ ID NO:49)
MLNLIQHPLPQRSIMERNPCVYILASGRNGTIYIGVTSDLMLRLHQHRTKAIPGFTSRYNVFRVLVYFEMFDDMPAAIHREKQLKAWRRDWKIALVEEHNPFWE
DRALDLGFPRFA

FIG. 11B

<SWIT_RS23650--Sphingomonas_wittichii_RW1--68.89 (SEQ ID NO:50)
 MKGGHVYIMSNVSGTLYLVGTADLAARVYQHRIGEGSAFCREHGLTKLVYAEHPHDRIENAIAREKALKNWKRAWKINLIEKANPGWRDLYEEIIA
 <SWIT_RS23595--Sphingomonas_wittichii_RW1--62.07 (SEQ ID NO:51)
 MGGWYIYLASKPFGVLYVGTADLAARLT VHREGRGSD FVKRFNCNRLVLVEPHATIEEAIAREKALKNWRRDWKRLRIAEANPGWDDLFDRFNQ
 <AMET_RS09265--Alkaliphilus_metalliredigens_QYMF--67.47 (SEQ ID NO:52)
 MLTNWNNKVLTYGTNNLERRIYEHKNKLVKGFTAKYNVGVKLVYFDYTSVISAIAREKEIKGWARKKKNELIESINPQWNDLSKELFR
 <AMET_RS11220--Alkaliphilus_metalliredigens_QYMF--66.10 (SEQ ID NO:53)
 MYKSCVVRGFTKYNVSKLVYFDSTDDVKAATREKQIKGCTRGKKIDLIESINPEWDDLSEIFNN
 <CBUD_RS01460--Coxiella_burnetii_Dugway_5J108-111--80.00 (SEQ ID NO:54)
 MEESFYILASQRNGTLVVGSTDLIKRIWEHKNNVPGFTSKYNVHQLVYVEVHQTIMKAAQERRRFKNWCRKWKNLIEKNNPTWRDLYDEICS
 <PHZ_RS08620--Phenyllobacterium_zucineum_HLK1--71.08 (SEQ ID NO:55)
 MASRRNGTLTYGTVSDLAARAYEHREALTPGFTSRYGVTLVWYEPFEDIGLAIRREKRIKKYPRQWKINLIEARNPHWVDLYNLLNR
 <PHZ_RS01255--Phenyllobacterium_zucineum_HLK1--66.67 (SEQ ID NO:56)
 MAGYVYILASAPYGTLYVGTNDIAFRVWQHREGAGSEFCRRYGVHRLVHYELFDDISNAIHREKRLKKWPRRWKINLIEAGNPTWRDLYETLNL
 <PLAV_RS01465--Parvibaculum_lavamentivorans_DS-1--76.74 (SEQ ID NO:57)
 MKDYVYILANGRNGTLTYGTVNDIQRVSEHREIDKGFTSRYGVKHLVWYEVHSEALTAIQREKNLKHWSRAWKIALIEEANPAWHDLYPGLF
 <PLAV_RS14615--Parvibaculum_lavamentivorans_DS-1--73.12 (SEQ ID NO:58)
 MKRPAVYIVANKRNGTLTYGTVSHLERRIYEHREGVLSGFSCRKYGVHRLVHYELFDDISNAIHREKRLKKWPRRWKINLIEAGNPTWRDLYETLNL
 <PLAV_RS03925--Parvibaculum_lavamentivorans_DS-1--69.66 (SEQ ID NO:59)
 MSYFVYILASRKHGTLTYVGTNDIARRIDEHRTGVASGFTTKYRIHQLVYAEHTDDVNEAIGREKAIKEWKREWKVQLIEKANPEWDDLTAASLNR
 <PLAV_RS03595--Parvibaculum_lavamentivorans_DS-1--66.29 (SEQ ID NO:60)
 MSYFVYILNQKNGTLTYVGTNDIARRAWEHREGTSGQFTSKRYGVHRLVHYELFDDISNAIHREKRLKKWPRRWKINLIEAGNPTWRDLYETLNL
 <PLAV_RS06910--Parvibaculum_lavamentivorans_DS-1--68.54 (SEQ ID NO:61)
 MSYFVYILASRKHGTLTYVGTNDIARRVFEHREGTSGFTTKYRVHRLVHYELFDDISNAIHREKRLKKWPRRWKINLIEAGNPTWRDLYETLNL
 <XAUT_RS10915--Xanthobacter_autotrophicus_Py2--76.40 (SEQ ID NO:62)
 MERRYFVYILASKPRGTLYVGVNTDLTRRVFEHRLGTRPGFTSRYHIHQLVWFKIHETAMAAIQREKRLKAWLRQWKIDLVEKHNPWEWDLFDEIALG
 <XAUT_RS15830--Xanthobacter_autotrophicus_Py2--75.56 (SEQ ID NO:63)
 MRGGWYIILNRPNGTLTYVGTNDLVRRVWEHRQGSADGFTTKRYGLTKLVYFEDHTEIALAIQREKAIKHWSRAWKARLIDTNPQWRDLYEDIIT
 <VIBHAR_RS28220--Vibrio_campbellii_ATCC_BAA-1116--72.41 (SEQ ID NO:64)
 MDKTPCVYIILTSKNNSVLTYIGVTSQLPQRIWQHKNVVEGFTKKYHVHKLTVYFELFEDMATAINREKQLKQWKREWKLDLVRESNPDFDLDSNEF
 <VIBHAR_RS05215--Vibrio_campbellii_ATCC_BAA-1116--65.57 (SEQ ID NO:65)
 MKQPAVYIVTCKTKSALYIGVTGNLKKRIWQHKNAVEKGFSRKYKTNKLVYFELFDDFYTAISVSAP
 <RCAS_RS11495--Roseiflexus_castenholzii_DSM_13941--69.41 (SEQ ID NO:66)
 MRNRYFVYIILTNWNNKVMYIGMTNNLERRIYEHKNKMDVGFTKKYNVKNLVYEHHTDVYAAIAREKEIKKWRREKKNKLVMSMNPWKDLSSEFGVGEDFS
 LRSK
 <DOLE_RS08500--Desulfococcus_oleovorans_Hxd3--84.62 (SEQ ID NO:67)
 MERLPVYIILANERNGLTYIGVTANLIRKRVWEHKNDCEFGFSKRYNVHSLVWYEHHTMESAISREKELKKWKRKWKLDLIEQNPNPYWEDLYEELKGE
 <DOLE_RS15425--Desulfococcus_oleovorans_Hxd3--65.06 (SEQ ID NO:68)
 MTNNTNSVLTYGTVSHLKKRVYKHKEKITGGFTKKYNAHKLTVYEVFQDMKAAIAREKQIKNGSRQKVDLIDNMNAEWVDLYELL
 <MPOP_RS25165--Methylobacterium_populi_BJ001--72.22 (SEQ ID NO:69)
 MICAVYLLAHRDGLTYLVGTDKLSRRIWEHKTAAAPGFSARYAVDRVLVWYEVYDRIDEAIAREKALKKKWRRAWKVALIEAMNPWSDLYERLNA
 <OTER_RS11370--Opitutus_terrae_PB90-1--58.82 (SEQ ID NO:70)
 MVVHFAEVNAPAPNPGEYVYIILESSDGALVVGQTNVPERLKRKHGIRIGSKFTRDHIRPTLVVVERRATLQEAVSREAQLKRWRAKKEALIRGDAQSLHRL
 SRSRD
 <CPAR_RS09915--Chlorobaculum_parvum_NCIB_8327--81.72 (SEQ ID NO:71)
 MREPAVYIILASKPNGTLTYIGVTSDLVKRVWEHKNNLLEGFTTKYHVHHLVYFQHNMDPSAFTREKQLKKNRAWKIALIEKENPDWRDLWEDICH
 <CTHA_RS10740--Chloroherpeton_thalassium_ATCC_35110--72.94 (SEQ ID NO:72)
 MKDYVYIILANKNGTLTYGTMNNIQRMYEHKKHLIDGFSKKYGLNHLVLEICQDVSAIAREKQIKGWKREKKVALIESRNPWKDLCDEFGFCFESRT
 <CAGG_RS12975--Chloroflexus_aggregans_DSM_9485--69.41 (SEQ ID NO:73)
 MRNRYFYIYVLTNNWNNKVMYIGTNNLERRIYEHKNKMDVGFTKKYNVKNLVYEHHTDVHAAIAREKEIKKWRREKKNKLVMSMNPWKDLSLEWG
 <ACP_RS03830--Acidobacterium_capsulatum_ATCC_51196--63.53 (SEQ ID NO:74)
 MPQSWYIILASRTGLTYGTVSDLTTRRVLEHRSGVRSASFASKYGCRLVYVEESDDIAVAIAREKAIKGWTRAKKIALIEAKNPWNDLAEHLGSRIIMPNQS
 MAEDKKEAQRIRLPLPETRKQE
 <KOLE_RS00725--Kosmotoga_olearia_TBF_19.5.1--67.37 (SEQ ID NO:75)
 MNPGFYIIMSNPGRRTFYIGVTNNLLRRVKEHKSKEIEGASKYNCVDIVYVESFEQISSAIEREKQLKTWKRQWKIALIEKVNPEMRDLYEDILCGS
 <HBAL_RS02100--Hirschia_baltica_ATCC_49814--71.43 (SEQ ID NO:76)
 MTAVYIVANRKNALTYGTVSHLIQRIHQHKQGVADGYTKRYGCKNLVWYELHEDIREAIRREKSLKSYRRQKNINMIVKNNPHWSDLYFSLL
 <HBAL_RS15440--Hirschia_baltica_ATCC_49814--73.40 (SEQ ID NO:77)
 MSYAIYIVASKDQGTIYIGVTNNIARRIWEHKQGVVDSYTKRYAVNRLVYVEWFDNIEAAIKREKQLKNWQRRWKISLIESDNPNYWRDLYFDLCPE
 <HBAL_RS06805--Hirschia_baltica_ATCC_49814--73.40 (SEQ ID NO:78)
 MTYALYIMASKPYGTLYIGVTNSLQRIWEHKQGGIGCTFTAKHNIKTFFWFEFMDNIETAIQAEKRMKKWKRWDKINLIERNNPHWQDLFDSIGHS
 <HBAL_RS05615--Hirschia_baltica_ATCC_49814--72.04 (SEQ ID NO:79)
 MRYAVYIVSSQDHGTLTYGTVNNLARRIWGHKQGVAKFTTKYSVSLRVWYENHDCIHTAIKREKNIAWQRQWKINMIERDNPYWLDMYASLGQL
 <PHEP_RS04485--Pedobacter_heparinus_DSM_2366--67.03 (SEQ ID NO:80)
 MDRGGCIYIMSNLARTFYIGVTSYDLVARIQEHKTKYPNSFTAKYKLHFCIYYEMQPSIEAIAAREKEVKKWSRSKKIKLITLNPKNWDLWEEIQW
 <KKOR_RS08620--Kangiella_koreensis_DSM_16069--71.91 (SEQ ID NO:81)
 MKEYYVIMASTKNGVLYIGVTGNLLKRVYQHREGIVEGSEYKTKRLVYFESSDTISCIEREKQLKNWKRQWKELIEKTNPQWEDLWHKIAS
 <KKOR_RS08955--Kangiella_koreensis_DSM_16069--70.65 (SEQ ID NO:82)
 MKQYVYVIMASRKGVIYIGVTGNLIRKRVHQRHREGLVDGFSKQYKTKRLVYFESSDILKCIEREKQLKGWKRQWKIGLIETVNPWEDLWIDIAS

FIG. 11C

<KKOR_RS11085--Kangiella_koreensis_DSM_16069--75.28 (SEQ ID NO:83)
MENNYVYIMTTKRNGALYVGVTSDLIKRVWQHKNNVESFTQKYRVHQLVYEEQCDSEIETALNREKRLKEWKQWQVDLIESKNPYWNDLYPKIT
<KKOR_RS03535--Kangiella_koreensis_DSM_16069--62.22 (SEQ ID NO:84)
MTKKGYVYILSNPRKTVLTGVTGNLKGQVFAHKNKIAEGFTRSYDCNDLVYIEIKDSLEKAQVREKEIQSECRLGKEALI KRFNPEWKDLSDELLAIKKR
<APRE_RS06475--Anaerococcus_prevotii_DSM_20548--65.26 (SEQ ID NO:85)
MSYFVYLLTDKTNSTRYYIGYTNNLIRRTYEHNRHLKGSFSDKYNIGKLVYEEVFEDYENAREREIQMKKYSRTKKEALVNSLNPWKDLYYDIK
<VEA_RS23010--Vibrio_antiquarius--69.77 (SEQ ID NO:86)
MTDRTHCVYILTSNNNAVLYIGVTSQLPQRIWQHKNGVVEGFTKKYNVHKLTVYFELFEDMATAINREKQLKQSKREWKLALVRETNPFLDLSTEYQ
<CAKA_RS16615--Coraliomargarita_akajimensis_DSM_45221--52.94 (SEQ ID NO:87)
MPFYTYILRCENDRLYVSGCQKLQERFQAHLERGAFTRRHKPLSIAFSKCYPSRAEAMARETQIKKWSRAKKQALIEQNSSELKRLSRSHDQ
<GFO_RS02490--Gramella_forsetii_KTO803--58.06 (SEQ ID NO:88)
MKEYFVYITTNHKKTVLVGVGTNDLGIRLNQHYQDSKMKKESFAGKYNCYHLVYFEEKYNDPNDAILREKVIKKWRREKKNKLIQEFNSNWQFLENDIL
<SJA_RS16315--Sphingobium_japonicum_UT26S--76.34 (SEQ ID NO:89)
MKREISPTVYILASRRNGTLTGVTSDLVKRLYEHRNGLIEGFTSDYGVKRLVWFEMHDAIDGAILREKRIKKWNRQWKIELIGRENPDWRDLALDFGFEPLK
SRRIVKASGD
<SJA_RS00510--Sphingobium_japonicum_UT26S--73.86 (SEQ ID NO:90)
MKEPCVYILASQPYGTIYIGVTSNLLGRLHQHRTNGVPGFTGRYAVHNLVRFEMCDTMEQAIRREKQLKRWHRQWKINLIETENPRWVDLAVGLGLPSLDQHQ
RHDGC
<SJA_RS13915--Sphingobium_japonicum_UT26S--71.91 (SEQ ID NO:91)
MSKPGVYIMASRRNGTLVGVTSDLVKRAWQHRNGLGSEFSAEYRCHFLVWYEAyddLQEARLRLKRMKKWKRWDWKLRLIETANPGWRDLFEDVAQ
<SJA_RS11530--Sphingobium_japonicum_UT26S--74.70 (SEQ ID NO:92)
MANRRNGTLVGVTSDLAARTYQHRNGLIDGFTKDHDCKLLVWFAYEDLQEARHRELRMKKWKRAWKIELIERDNPQCRDLFETLF
<SJA_RS15690--Sphingobium_japonicum_UT26S--61.63 (SEQ ID NO:93)
MERKQKGGWYIMSDRYRGTIYVGVTSDLPARIHQHRSNGSGDFCARYGLNRLVWAERSDDITRCIEHEKRVKRWREWKFDLIERANPDWQDLFHLLA
<SVI_RS09835--Shewanella_violacea_DSS12--80.65 (SEQ ID NO:94)
MQPCVYILASKNGTLYGVTSDLIKRIWEHREHLVTGFTDRYHVTQLVYELTDCMYSALISLEKKLKKWRRQWKINLIENQNPEWHDLDWEIKS
<LLO_RS02970--Legionella_longbeachae_NSW150--76.67 (SEQ ID NO:95)
MEKYSVYIVTNQKNGTLYGVTSDLVKRIWEHKNKFVESFTKKYNCRLVYEEQHIDVLEAIKREKRLKKYLSRWKIELIESVNPLWRDLYEDIC
<LLO_RS06035--Legionella_longbeachae_NSW150--70.59 (SEQ ID NO:96)
MQSFYVYLMASRRNGTLDTGSTSDLKKRVREHKKNRVPGFTAQHNHLLVYEEVHEAYAEAAARHEKRFKNWCRQWKLNWIEELNPEWRDLDEETCHCLRVSGC
HGQVAARRWWG
<LLO_RS03580--Legionella_longbeachae_NSW150--66.67 (SEQ ID NO:97)
MPYLMASRRKATLDTGSTSDLMKRVWEHKNNRVPRFRARYNVFLVYEEVHEACVAAARRDKRFKNWYRQWALNRIEELNPEWRDLDEEICHKLAPG
<LLO_RS08945--Legionella_longbeachae_NSW150--67.19 (SEQ ID NO:98)
MFLWQAGNGLDTGSTDLMKRVWKHTKNRVPGFTAQYNVHLLVYVVDAYVEAARCDKRFKN
<RHE_RS11510--Rhizobium_etli_CFN_42--67.78 (SEQ ID NO:99)
MTGFVYVMSDKPRGVIYTGVTSDLHGRIWEHRNEIHKGFTQKYRAKNLVWFENHPNIVLAIQREKSLKRYMREWIKLIEELNPTWIDLYERIDEIENVYRPH
PNTREWNDYN
<TRAD_RS10075--Truepera_radiovictrix_DSM_17093--60.00 (SEQ ID NO:100)
MPWHVYVLCRSGALYTGVTDPFLRLTEHNAGRGAAFTKRHGAELVYLEAAESRGAALRREHAIKRLSRARKLELVASASNQLREVHGAR
<M301_RS10825--Methylobacterium_thermautotrophicum_ATCC_35061--80.00 (SEQ ID NO:101)
MQPCVYIMASRRNGTLYGVTSDNLITRVWQHRNNAIDGFTQKNKVHQLIWEYEPHENMESAI SREKALKWNRIWKRLIEQYNPEWQDLYEQLT
<M301_RS02070--Methylobacterium_thermautotrophicum_ATCC_35061--82.80 (SEQ ID NO:102)
MKQPCVYILSNQVRGTLYGVTSDNLITRVWQHKQGLIDGFSKQHEHLHLLVYFEQFDDMTNATITREKQLKKNRWAKWIELIESVNVDWRDLYDLIGG
<M301_RS12960--Methylobacterium_thermautotrophicum_ATCC_35061--77.27 (SEQ ID NO:103)
MIKQPALYIASKYNGTLYGVTSDNLITRVWQHKQGLIDGFSKQHEHLHLLVYFEQFDDMTNATITREKQLKKNRWAKWIELIESVNVDWRDLYDLIGG
<HDEN_RS04670--Hyphomicrobium_denitrificans_ATCC_51888--70.97 (SEQ ID NO:104)
MRVVKPASYILASKRNGTLYGVTGDLAARVSIHQDLEDGFTKRYGVHMLVHYEQFDTMPDAIAREKQLKKFRRADKLALIESGNPQWHDLTSEVQAWRS
<HDEN_RS11290--Hyphomicrobium_denitrificans_ATCC_51888--76.14 (SEQ ID NO:105)
MKRPCVYILASKPYGTLVGVTSDLPHRMAQHDQGLIDGFTKRYRVKTLVYVYEHMETMDAAIVREKRLKRWQRTWKYRLIEQMNPEWSNLVDAATGEIEFGSI
DAPNLVDPSPSNLDGSPPTRG
<HDEN_RS01385--Hyphomicrobium_denitrificans_ATCC_51888--76.67 (SEQ ID NO:106)
MERRFYIYIMANKRNGTIYIGVTNDLARRVYEHREGLVKGFTSRYGLKSIVYEEAFDSVSLAIQRETSKRWPRRWKLALIEKANPQWKDLVDEIV
<DEHLY_RS00795--Dehalogenimonas_lykanthroporepellens_BL-DC-9--77.01 (SEQ ID NO:107)
MASKPNGTLYGVTNNLIRRVYEHKNDLADGFTKKYGIHTLVHFEQCDNILAAITREKQMKKWKRVWKIELIEQANPEWRDLYPEILG
<HMPREF0766_RS22180--Sphingobacterium_spiritivorum_ATCC_33861--56.16 (SEQ ID NO:108)
MYFLYILYSVQADRYIYIGVTHDPEGLRRHNSHYKKGFTGKFMDCWICVYQECYETKNEALIRERQLKSWKSRLKIEELINKSD
<HMPREF0766_RS22175--Sphingobacterium_spiritivorum_ATCC_33861--56.16 (SEQ ID NO:109)
MYFLYILYSVQADRYIYIGVTHDLEGCLRRHNSHYKKGFTGKFMDCWICVYQECYETKNEALIRERQLKSWKSRLKIEELINKSDELVQIPIKSEGSLVRIPVA
STKASARMLFLLSLNYVLIPIQVSDNTSDIFT
<BRESU_RS04135--Brevundimonas_subvibrioides_ATCC_15264--73.03 (SEQ ID NO:110)
MAKIFYVYLLASGPGWMMYVGVTNDLVRRVAEHKEGLIPGYTSEHRIDQLVWYEAHQYIDKAI VREKRIKRWLRWKFWALVEKDNPRWLDLYGEVRLRRPGS
AEPTGPHS
<BRESU_RS01380--Brevundimonas_subvibrioides_ATCC_15264--67.47 (SEQ ID NO:111)
MASGRNGWYIGSTSNLHARVWQHKGTGTHPDGYSAEHGCVHLVWYEQHQWVVDARI RRRRLKKWMDWKLKLI EDANPQWLDLADWYFVNDPNWTPPSEHD
<HMPREF9225_RS00400--Peptoniphilus_duerdenii_ATCC_BAA-1640--71.08 (SEQ ID NO:112)
MSNSSNTIYIYGFTNDLRRVYEHNRNLDVDFGSKKFHTHKLTVYFEVYSTSYDAIAREKQLKKWTREKKNLSIEKDNPEWIDLWDRINNFKKL
<HMPREF9225_RS00365--Peptoniphilus_duerdenii_ATCC_BAA-1640--68.67 (SEQ ID NO:113)
MSNSSNTIYIYGFTNDLRRVYEHNMNLDVDFGSKKFHTHKLTVYFEVYSTAYDAIAREKQLKKWTREKKNLSIEKDNPEWIDLWDRINNFEKL

FIG. 11D

<FSU_RS15000--Fibrobacter_succinogenes_subsp._succinogenes_S85--61.36 (SEQ ID NO:114)
MNSYTYMTNQNNTTLYVGVNDISRRALEHIEGCGATFTSRYKINKLIYFERYTEITDAIKREKQLKGWRREKKEILIMQMNPLWKDLLQDDS
<THEBA_RS05915--Mesotoga_prima_MesG1.Ag.4.2--69.66 (SEQ ID NO:115)
MNSGFVYIMTNPSRNVFYIGVTNNLFRRVFEHKNKMI EGFTHKYNCVVVYVEYFETIKEAIRREKQLKSWKREWKRLRIKTRNPDRLDLWEDIRESYR
<SPHCH_RS11770--Sphingobium_chlorophenolicum_L-1--76.40 (SEQ ID NO:116)
MKREISPTVYILASRRNGTLYTGVTSDLAKRQYEHNRNGLIEGFTSDYGVRRVLVWFEPHDVMESAIQREKRIKKWNRQWKIEELIEAGNPWDWDLARGFGFEPLP
SLRCD
<OB2597_RS08390--Pseudooceanicola_batsensis_HTCC2597--64.04 (SEQ ID NO:117)
MTACVYILASRPHGAIYIGCTADLRRRVAQHRAGAVKAHTRIYGIHTLVWFETHGDLAEARLRERRLRWRRRWKDALIGEHNPHWQDITAFIPD
<MED217_RS13605--Leeuwenhoekiella_blandensis_MED217--72.94 (SEQ ID NO:118)
MEELFYIYILTNNKHTVLYIGVTSDLKRRLLQHKQKSNKGFTRKYNVEKLLYFESSPYVENAIKREKQLKKWNRWKENLINEMNPDWDLWSWMLDE
<MED217_RS00725--Leeuwenhoekiella_blandensis_MED217--62.50 (SEQ ID NO:119)
MKQSYVYILTNTYRTTFYIGVTSNLQQRIVDHKEKKGSFTSKYNVCDLIYFEEFSDIIQAI SREKQLKNWKKDWKLNLIKEKNPKLQTLNIFS
<MED217_RS07740--Leeuwenhoekiella_blandensis_MED217--64.36 (SEQ ID NO:120)
MWVNYLVYIILTNEYRTTYIGITNNIQRRLDQHFFDSINAKKSFAGKYNCIYLVYVEVDNPDTAIAREKEKWKREKKNDLISSFNCPWKTLNNEVF
<MED297_RS10570--Reinekea_blandensis_MED297--57.14 (SEQ ID NO:121)
MPTPTPGWVYFVQCADGTFYTGITTDPHRRLRQHNGELVGGARYTASRRPVTLVFYEPHPSRSDAARAEYRLRKLRSKQKLARIETFEVPSG
<NB231_RS13810--Nitrococcus_mobilis_Nb-231--74.16 (SEQ ID NO:122)
MSKTACVYIILASKNGTLYTGVTDDLQRRVYEHHKHLVEGFTRKYNVTRLVWFQQGDDITEAISVEKKIKNRPRKWKVALIEKTPNGWADLAADWVK
<SI859A1_RS02440--Aurantimonas_manganoxydans_SI85-9A1--71.11 (SEQ ID NO:123)
MAGTVYIMASQKNGTLYTGVTSDLSRRVYEHSGLTKGFVKAHVTRLVWYEDHVDIGSAIQREKSLKRWRRQWKIDLIEAMNPEWNELYRGMGW
<SI859A1_RS04305--Aurantimonas_manganoxydans_SI85-9A1--58.75 (SEQ ID NO:124)
MPLEATVYIILRCADGSYYTGLTKQPV DARLWEHNEGLYDGYTRKRRPVVLVFTETYNRMLDAIARERQIKGWTRAKKEAL IARRYHRLPALAATRTGPDLRQ
PPATLD
<P700755_RS17030--Psychroflexus_torquis_ATCC_700755--67.42 (SEQ ID NO:125)
MPKRAYHNYWVYILTNNKPNGLIYIGVTGGIDYMERHTTGTGSKFTSKYNLKIMVYLEEFQYIEDAIAREKQLKSWHRQWKNLIEKDNPNWKLWVKLG
<P700755_RS03685--Psychroflexus_torquis_ATCC_700755--64.44 (SEQ ID NO:126)
MNHVYIMTNKNNRVIYIGVTNNLDERVKEHKSRTYPNSFTSRYNCDKL VYFEEFEKSDEACKRERRFKKWNRAWKIDLIEEMNPNWFDLSDNWNLD FSI SRL
RK
<P700755_RS03680--Psychroflexus_torquis_ATCC_700755--65.06 (SEQ ID NO:127)
MTNKNRVIYIGVTNNLDERVKEHKSRTYPNSFTSRYNCDKL VYFEEFEKSDEACKRERRFKKWNRAWKIDLIEEMNPNWFDLSDNWNLD FSKSRLRK
<P700755_RS01210--Psychroflexus_torquis_ATCC_700755--56.32 (SEQ ID NO:128)
MKT LGTHNYYYIILTNNKRVLYTGVTNDLQARLYYHINPNALSKAFTTKYKCYFLIYYKHFDINQAIAREKQIKGYRRSKKETLINNINPEWFLNDKI
<RED65_RS07360--Bermanella_marisrubri--76.09 (SEQ ID NO:129)
MNKTPCVYIMTNYRKGTIYVGVTSDLKRRVWQHKNNAVDSFTARYGLKTLVYVEVTENMESAIQREKRLKHWNRSWKLNLDISFNPSWRDLYIDL F
<HP2506_RS08285--Fulvimarina_pelagi_HTCC2506--66.67 (SEQ ID NO:130)
MTGFVYIMASRRGGTLYTGVTSDLSHRVFEHKNALTPGFTSRYGCKRLVWYEEHFDVRDAIAREKTIKRWPRQWKIDLIEATNPDWTELYRGVGV
<FP2506_RS08800--Fulvimarina_pelagi_HTCC2506--65.56 (SEQ ID NO:131)
MTGYVYITNNKGGTLYTGVTSDLANRVYKHKEGLTLGFTSRYGCTRLVWYEEHFDIHDACREKAIKNWRRQWKIDLIESINPEWYEL YRGMGW
<HPDFL43_RS16705--Hoeflea_phototrophica_DFL-43--70.59 (SEQ ID NO:132)
MRKGYVYILASRRNGTLYTGVTSDLPGRLYEHQNNLTPGFTSKYGVKTLVWFEHDLVTEAIAREKTIKNWPRQWKIDLIETMNPWDNDIAHFLHCL
<HPDFL43_RS20480--Hoeflea_phototrophica_DFL-43--68.82 (SEQ ID NO:133)
MTGYVYIVTNHKGHTLYIGVTSDLARRIHEHREAGTPGFASKYGCRLVWYEEHFNIAAGAIQREKSLKRWYRKWKIEELIESVNPWDWRDLYDELW
<KAOT1_RS01675--Kordia_algicida_OT-1--67.05 (SEQ ID NO:134)
MKKPFVYFLTNKNNTVIYIGVTSNLMKRIYQHKTKSYKGFTSKYNCDKLVYFEEFQ SINEAIIREKQLKSGNRKRKEELINKENPEWNDLSEGLWLFYFD
<KAOT1_RS19570--Kordia_algicida_OT-1--72.29 (SEQ ID NO:135)
MTNKNPGLVYIGVTKDLKRRIRQHKNNKVHPTTFSARYNLNKL VYFEKYNSKDEALLREKRMKKWNDRKVELIEKENPTWEDLVGKIK
<KAOT1_RS21380--Kordia_algicida_OT-1--65.88 (SEQ ID NO:136)
MDIHYVYILTNNKSHSVLYVGRTNQLKRRNLQHKNNAKDSFTGKYNVYKLVYFEKTKYVNNNAIIREKQIKKWKQWKINLINGLNPDWKLSEYV
<KAOT1_RS21400--Kordia_algicida_OT-1--61.18 (SEQ ID NO:137)
MRKSYIYIITNKYRTTFYIGITANLKRTRIAHANSKGSVFTTKYNLTDLVYFEEFLDINQAIAREKQLKNWKRDKWIKELIKLNPTLETLEI
<MC7420_RS23830--Coleofasciculus_chthonoplastes_PCC_7420--68.60 (SEQ ID NO:138)
MTNRSKTLYTGVTNDLYRRVYEHKQKLI PGFTQKYNINRLVYQETSDIREAIAREKQIKGWLRAKKIALIESINPQWRDL SADFET
<MC7420_RS16060--Coleofasciculus_chthonoplastes_PCC_7420--52.87 (SEQ ID NO:139)
MPYMYILECSDSSPYTGSTKDLPRRLWQHQNGLGANHTKRLPVKL VYAEYYDRVEDAFYREKQVQGWNRKKKIALMNSDWNRLHVLAECQNESHYRRGLVER
SRNQPLQGSRNQLKPSNDED
<MITSMUL_RS00080--Mitsuokella_multacida_DSM_20544--56.58 (SEQ ID NO:140)
MAEAYTYILECKDGSlyTGWTNDLTRRVKAHNEGRGAKYTRSHRPVHLVYFESFATKEEALAREAAIKKLTRA AKLALIDTAEEKKL
<MDMS009_RS05390--Methylophaga_thiooxydans_DMS010--53.85 (SEQ ID NO:141)
MSKAEBFVYIIEADNGKLYTGITTDIDRRFHQHQTHRGGARFFHTSQA KKL VYQETQPD RSSASKREAAIKKLSRQAKLTIAQY
<HMPREF0072_RS06245--Anaerococcus_lactolyticus_ATCC_51172--68.42 (SEQ ID NO:142)
MAAYFYMLANATNEVLYKGSTTNLIRRIYEHKEHLLKDSFTDKYNVTRLVYVECYDDIQSARAREHQVGKWSRKKDKIISKMPKWKIDLYWGLLG
<HMPREF0072_RS00565--Anaerococcus_lactolyticus_ATCC_51172--64.84 (SEQ ID NO:143)
MYVYVYILLNKSNDRTYVGHTDDLKKRIWEHKNHVPYSH TAKYNI TKLVYVECGEDFYESRAREKQIKKYSRLKKYALIEKENPYFRELYYDLF
<SADFL11_RS09545--Labrenzia_alexandrii_DFL-11--73.03 (SEQ ID NO:144)
MRRFWYVYILAGQKNGTLYTGVTNNLSRRLYHQQGTGSKFTQKYSVKRLVWYEEHATAERAISREKAIKHSRAWKIELIEATNPEWFDLGKDLNR
<HMPREF0077_RS07915--Anaerococcus_tetradis_ATCC_35098--68.18 (SEQ ID NO:145)
MYVYVYILLSNYNNKIYTGVTNNLERLYEHKHLADGFTSKYNNVKNLYV FETGDSIDGAIAREKQIKSYRREKKDAIVNSVNPWDKLSLDWNK
<REIS_RS12820--Rickettsia_endosymbiont_of_Ixodes_scapularis--76.14 (SEQ ID NO:146)
MLASGRNGTLYIGVTSNIIKRIYEHKHKI IKGFSSKYNINKLVYFEQYDNIIIAINREKLLKEYRQWKLNLI ESQNPWEHDLYEKIIIL

FIG. 11E

<REIS_RS12160--Rickettsia_endosymbiont_of_Ixodes_scapularis--73.75 (SEQ ID NO:147)
MLCLYTLINVTVLVYIGVTSIIKRIWQHKKQIRKGFSTTKYDIDKLVEYQFQNDVNLAIGREKRLKEYKRKWKLDLINSSNPSWHDLYEDFLK
<REIS_RS12400--Rickettsia_endosymbiont_of_Ixodes_scapularis--82.98 (SEQ ID NO:148)
MGYILCSNKHGTLVIGITNNIIRRAYEHKQKIIKGFSTSKNNIIKLVYTEEFTKI
<HMPREF0063_RS00810--Aeromicrobium_marinum_DSM_15272--50.00 (SEQ ID NO:149)
MPFTYILLCSDKTFYTGSTRDLIVRLRQHREGGAEYTKRRRPVKLVYEEHEHVAHAFWREKKIQGWTHGKKRMLVRDGPVGVKVEDDSDLFGTG
<ALPR1_RS07475--Algoriphagus_machipongonensis--67.03 (SEQ ID NO:150)
MQKGGAVYILTNTNTVLTGVTNDLFRVSEHRDGLNPHAFSTSKYNLNRVYERFHSIBEAIRNEKQIKGKSRKKKEELINSINPKWKDLWDEIKEW
<M23134_RS16440--Microscilla_marina_ATCC_23134--58.33 (SEQ ID NO:151)
MNSNYFVYITNPKRTTLYIGVTNDLVRLEEHASQAGNPKFTTGRYPCYHLIYWERFDTQDAIDREKELKGWRREKKEALIDAENPQWIFLNDEI
<M23134_RS25755--Microscilla_marina_ATCC_23134--57.14 (SEQ ID NO:152)
MNGNYFVYITNPKRTTLYIGVTNDLVRLEEHASQAGNPKFTTGRYPCYHLIYWERFDTQDAIDREKELKGWRREKKEALINAENPQWIFLNDEI
<M23134_RS10325--Microscilla_marina_ATCC_23134--57.14 (SEQ ID NO:153)
MNGNYFVYITNPKRTTLYIGVTNDLVRLEEHASQAGNPKFTTGRYPCYHLIYWERFDRPQDAIDREKELKGWRREKKEALIDAENPQWIFLNDEI
<CFE428DRAFT_RS25220--Chthoniobacter_flavus_Ellin428--61.36 (SEQ ID NO:154)
MNREYVYVILTDKGRKTLVIGVTNDLPTRLWQHRNPEKASFSQRYHCIVLVYEEQFPDPLSAIAREKQLKGWRREKKNALIETMNPWEELSADWP
<ANHYDRO_RS01145--Anaerococcus_hydrogenalis_DSM_7454--69.32 (SEQ ID NO:155)
MSDYVYVILANKSNSTIYGYTHNLEKRLYEHRNHVVEGFTDKYNCTKLVEYELDNKYEALSREKQLKRRSREKKNLINSINPKRRDLSFDWYS
<VIA_RS11385--Vibrio_orientalis_CIP_102891 = ATCC_33934--71.91 (SEQ ID NO:156)
MEVKQPAVYILANESGSLYIGVTSNLKARIWQHKNSSVDGFSKRYNIDRLVYELYLDITDAIKREKQMKWKREWKDNLIAETNRSWADLYDTL
<MMEL_RS0114960--Citromicrobium_bathymarinum_JL354--77.42 (SEQ ID NO:157)
MTRDFQPTVYLLASRRNGTLVGVTSDLRKRIAEHRDGLIEGFTKDYGVTRLVWFQEHATMDAAILREKRIKKWKREWKINLIREQNADWRDLAVDLGFAMLP
PIPSHRP
<MMEL_RS0105650--Citromicrobium_bathymarinum_JL354--73.86 (SEQ ID NO:158)
MPTEKQPCVYILASKPFGTLYIGVTSDLIARLWQHREDTGEFTKRYAVHRLVHFELFGTMESAIKREKQLKNWHRDWKLNLTRENPDWRDLAVGLGFEPIL
PDEPQAGP
<ACIX8_RS07760--Granulicella_mallensis_MP5ACTX8--67.42 (SEQ ID NO:159)
MPRREYHFWYILSSRSRLYIGMTNSLIRRTSTHREVPGTHTARYSIHRLVYFEYFRYVRSIAIAREKELKHWTREQKIELIERVNPWTWEDLFTLQEPVGF
EG
<ACIX8_RS00170--Granulicella_mallensis_MP5ACTX8--68.24 (SEQ ID NO:160)
MQNREYQFYVYILSSRSRDLYTGMTNNLRLRISQHREMRPGTYTAQYNIHRLVYFERFQYVRSIAIAREKELKDWTREKKIVLIEKVNPWTWEDLAASW
<ACIX8_RS21785--Granulicella_mallensis_MP5ACTX8--62.35 (SEQ ID NO:161)
MSGGYTYILGSVTGTLVIGVTSNLYLRVMEHKDGRFEGFAATHDCNRLLYFEGHEDIRISIAREKQLKGWRREKKNLISTVNPEFKDIAETWGWKMITVHEE
MNP
<ACIX9_RS03840--Granulicella_tundricola_MP5ACTX9--66.28 (SEQ ID NO:162)
MEKGGHYIILASKGRLYTGVTSNLQIRVAQHKS KLHPTSFRTARYNIDRLVYQGFETIEBAI LRESQIKNMHRTKIQLVVSHPNTWQDLSAAWGTPAPSPD
ETKLSAPVTF
<OCEPR_RS07310--Oceanithermus_profundus_DSM_14977--82.22 (SEQ ID NO:163)
MQPAVYIVANRKNAGLYVGVTSDLKRAWHRQGLVEGFSGSKRYGLKRLVYFELHPTMASAIQREKQLKAWKRAWKIDLIQKENPNWNDFHLLD
<OCEPR_RS06985--Oceanithermus_profundus_DSM_14977--77.42 (SEQ ID NO:164)
MEPAVYILASRKNGLVYIGVTAVLKRRVWEHKEGIGSAFTRRYGVHKLVEYFEAHTMESAI SREKQLKNWKRAWKIALIESANPDWDDLWGL
<OCEPR_RS09050--Oceanithermus_profundus_DSM_14977--75.56 (SEQ ID NO:165)
MAYIVYILSNRKRGTLYVGLTRDLRRRVWEHKEGLSSFTRRYGVHRLVYEVHEDANAALQREQQKLRWRSWKVELIEQSNPEWEDLFSSLA
<CELAL_RS04150--Cellulophaga_algicola_DSM_14237--70.33 (SEQ ID NO:166)
MKPGFYVILANKNNTLYVGVTSNLLQRIERHKTIFYPKSFSARYNNVNLVYEEAFQDIGAIIAREKQLKAGSRTKKIALIEKENPDWEDLTEQATDCFG
<BACSA_RS14335--Bacteroides_salanitronis_DSM_18170--65.88 (SEQ ID NO:167)
MHQYVYILSNKTRSVLYIGVTNNLYKRIFIEHKTGIIPGFTSKYKCHYLVYFEEYHDINEAIVREKELKGWTRYKKEKLI RNNNPSLKNLAQEMNWE
<PEDSA_RS11730--Pseudopedobacter_saltans_DSM_12145--63.44 (SEQ ID NO:168)
MIKSVIKQYSVYIMTNRNNTVLYTGVTNDLERRIWEHKFENGSKFTSLYKCHKLVYEDYASVKEAIIAREKQIKAGPRSKKLELINLENPNWSDLSKDW
<PEDSA_RS06295--Pseudopedobacter_saltans_DSM_12145--65.06 (SEQ ID NO:169)
MYKGGFYVYILTNNKTVLYIGVTSDLKGRIVEHMNSNNQSFTRYNVLYLVYEEFLDSISQAI DREKQLKRSREKKEDLINKMNPWTGFLNDEVAEDVYSL
I
<METTU_RS13635--Methylobacter_tundripaludum_SV96--76.09 (SEQ ID NO:170)
MEKQPCVYILASRPNGLTYIGVTSNLIGRVYQHRQNLIAAGFTQRYNVHNLVWYEPHEQLESAILREKQLKNWSRVAKKRLIELTNPQWHDLSPLDGLPSLAAG
SRHSLPG
<Y88_RS14045--Novosphingobium_nitrogenifigens_DSM_19370--63.33 (SEQ ID NO:171)
MKHEPGGRVYIMADRYRGTMVYGTADLPRIIFQHRQGTGSDFCARYRLHRLVWAERGEAIDVCIAHEKRIKQWRREWKFDLVERVNPWDIDLYDSL
<SL003B_RS11590--Polymorphum_gilvum_SL003B-26A1--77.17 (SEQ ID NO:172)
MPAWVYIMASKPFQTLVGVTSDLPRRIYEHREGLVEGFTEKYSVKSLVWYEEYPGVRDAIEKEKQIKRWRRQWKFLIEKANPEWRDLYQDLNR
<CA2559_RS05265--Croceibacter_atlanticus_HTCC2559--69.23 (SEQ ID NO:173)
MLKGQHQYVYVYILSNKMYGVLYIGMTNNLERMMIVHKNKLVEGFSGKQYNLTCKLMCFEQFYVNDIAIKREKQLKNWKRDWKIELIEHNKDWHDLAIDWDF
<CA2559_RS05295--Croceibacter_atlanticus_HTCC2559--64.37 (SEQ ID NO:174)
MSKWTYVIMTNKPNGLVLYIGVTDIHSRVVEHKVKFYPKSFTAKYNCCLKLIYFEEFSEVABASKRERQFKKWKREWKINLIEBMPNPSWTDLNLNWNLDLNKHR
TNQ
<CA2559_RS05285--Croceibacter_atlanticus_HTCC2559--76.44 (SEQ ID NO:175)
MTKWYVYIMASKFDGLIYIGVTNDLDSRVLEHKKHKNPKSFSARYNCCLKLYFEELEDESASTRERQMKKWKREWKINLIEBMPNPSWTDLNLNWNLDIGKSR
LFRK
<CA2559_RS05305--Croceibacter_atlanticus_HTCC2559--60.98 (SEQ ID NO:176)
MTHKRNLHLYVGVTDIKRRTFEHKS GTYGSHTGRYNDIKLVYEEHQWIEQAIKRVKTIKKWKREWKQNIITELNPEWRDLQAQWDLSEFEPNQETNN
<CA2559_RS08900--Croceibacter_atlanticus_HTCC2559--60.56 (SEQ ID NO:177)
MNDEVKIYYCYILSNKNRGLVYIGYTENLKVRLKHKHKTGAVFTKMYNVHDLVYEEKFDIKKQAKQREKQLKNWHDWK

FIG. 11F

<AL1_RS12165--Alistipes_shahii_WAL_8301--62.35 (SEQ ID NO:178)
 MESSYVILTNKHLTVLVGVTTSLSRMKEHAAGQSAFTRRYNVHKLIIYAEYRDVRDAIAREKQIKSWSRKRKIGLIDSQNSLWKDLSDGFV
 <METAL_RS02500--Methylomicrobium_album_BG8--82.80 (SEQ ID NO:179)
 MEKQPCVYIILASGRNGTLYVGVTSDDLGRVYQHRQKMIQGFTARYGVDQLVWYETHEQMESAIMREKQIKKWSRVFKRRLIEDLNPAWRDLWPDLVSSSVAAR
 <METAL_RS15675--Methylomicrobium_album_BG8--78.26 (SEQ ID NO:180)
 MNKQPCVYMLASQNGTLYIGVTSDDLIRRVYQHRNEQAEFGFTKRYGHDLVWYEVHEKMEISAILREKQLKKWPRIAKIRLIEQSNPIWHDLWLELASPSLASG
 FRQSMPE
 <THICY_RS04625--Thiomicrospira_cyclica_AIM1--76.60 (SEQ ID NO:181)
 MQPAVYILSNENNVLYIGVTSNLVQRVYQHRNHLVDGFTKQYNVTKLVYFEVHEDMSEAIKREKALKNWHRDWNKRLVSEMNPDPNDLWDSIV
 <THEIN_RS08900--Thermodesulfatator_indicus_DSM_15286--66.29 (SEQ ID NO:182)
 MQKQYVYIILTNKNNTVLTGTISDLKRRIFQHKNGIGSKFTKKYKVTCLVYIEIFEDPENAILREKKIKSGSRKKKIELININFPPEWRDLDEL
 <GMC_RS0117255--Aquimarina_agarivorans--61.90 (SEQ ID NO:183)
 MYIKEQDQLFFVYIILTNMHTVLYIGFTGNIKQRLKQHRSKSNKGFNSRNCNIHKLIVYFEFTFIYPNEGIKREKQLKKWNREWKNNLINANNPDWEDWNWRIDSF
 <OCA5_RS07355--Oligotrophia_carboxidovorans_OM5--68.54 (SEQ ID NO:184)
 MDSTFFVYVLANCPRGVLYVGVTSDDLRRLESEHRVKLVAGFTATYGVTLNVHVEYASIIERAREHSLKRWRRRAWKFDLIEKDNPWTWRDLSGDI
 <OCA5_RS18885--Oligotrophia_carboxidovorans_OM5--70.00 (SEQ ID NO:185)
 MAGLVPAIHAFHNYRMTNGGYVYIILTNQPNGVLYVGVTDNLVRRVFEHRSFGIDGFTKRHLKRLVYFEQFDDIQTAIQREHNIKHWPRAWKIRKIMVHNPEW
 HDLYETLI
 <F7308_RS09180--Francisella_salina--71.11 (SEQ ID NO:186)
 MAFVYIILASKNGTLYIGVTSNLIKRVYEHKQGYSDGFTKKYDVKKLVYEQFNNTDAIYRVKRLKTWQRKWKLDLINKFNPRWLDLYASILA
 <RH054_RS02605--Rickettsia_heilongjiangensis_054--77.65 (SEQ ID NO:187)
 MYWVYIILCSNKHGTLYIGITNNIIRRAYEHKQTIIGFTAKYNIHKLIVYEEFTDIREALTREKNLKKWQSRGKIELIEKTNPEWKDLIK
 <MURRU_RS10565--Muricauda_ruestringensis_DSM_13258--69.15 (SEQ ID NO:188)
 MSFYVYIISNKKMGTYIIGYTNDLKKRMYRHKHGVSGKFASRYNLKTLVYEEKFNFFMPAIRREKQLKKWNRSWKINLINDFNPDWKDLTYFFD
 <BZARG_RS05915--Bizionia_argentinensis_JUB59--67.37 (SEQ ID NO:189)
 MQPGFYIILTNKNNTTLYVGVTAYLKERIIQHKEKFYKKSFTARYNIDKLVYEAHQMGIDAI GREKQLKAGSRAEKIALIESINPDWDLFDEV
 <BZARG_RS02165--Bizionia_argentinensis_JUB59--62.77 (SEQ ID NO:190)
 MKHYVYIITNKLKGVLYIGETGRKIRIYQHKTKAHPNTFSKYNLDKLVYEEFENETEAKLKEKQMKKWYRAWKVELIERLNPYWKDLYEELK
 <BZARG_RS02105--Bizionia_argentinensis_JUB59--65.06 (SEQ ID NO:191)
 MRNKLNGVIYIGVTDSEITERIKEHKLKVYPKSFTAKYNCCLKLVYEEFESGDAEKRRERQFKKWKRDWIKI KLIEDMNPNSDLIVNWNLYIKLRK
 <MICA_RS06660--Micavibrio_aeruginosavorus_ARK-13--75.56 (SEQ ID NO:192)
 MEKKFVYMLASKEYGTLYIGVTSDDLKRIWEHKSVAEGFTEKYDVQNLVWVEEHQNAESAITREKQMKWEKREKLNRIKDMNPEWRDLYEDICK
 <MICA_RS05095--Micavibrio_aeruginosavorus_ARK-13--77.78 (SEQ ID NO:193)
 MEKRFVYMLASQKGTLYIGVTSDDLKRIWEHKNKVVEGFTKDYVHNLVWFEDHATADAAIRKEKRLKKYERQWKINLIBEKNPEWVCLYDEICK
 <MICA_RS04050--Micavibrio_aeruginosavorus_ARK-13--74.44 (SEQ ID NO:194)
 MVRGVSVMFRFVYIILASKKDGFLYTLGTSDDLKRIWEHKEGVAEGHTKYSYIKRLVYIEIHDTFEQAEIREKRIKRWREWKDRLIEENPEWNDLYSEICK
 <MICA_RS03740--Micavibrio_aeruginosavorus_ARK-13--78.82 (SEQ ID NO:195)
 MKQYFYIILASDKVGTLYVGVTSDDLKRRVVEHKGKTLGFTKQYDVSILVYFEVYDDPENAILREKRVKKWNRSWKIDLIQKSNPDWTDLYDTL
 <GNIT_RS04890--Glaciecola_nitratreducens_FR1064--79.57 (SEQ ID NO:196)
 MKQPCVYMI TNRNNTTLYIGVTSNLVQRVYQHKNKLIKGFSAKYNVDKLVYFELFEDMENAITREKRLKLWKRDWKNRIINEVNPSWQDLYPDLIR
 <KK7_RS0107045--Sphingomonas_elodea_ATCC_31461--74.70 (SEQ ID NO:197)
 MANKRGTIYVGVTSDDLPAAYQHRTGMIEGFTKDNGCTLLVWYEAHEDLDGARLRERQMKKWKRAWKIELIERDNPQWKDLYGTFLG
 <THITH_RS05780--Thioalkalivibrio_paradoxus_Arh_1--80.65 (SEQ ID NO:198)
 MKQPCVYMI TNRNNTTLYIGVTSNLVQRVWQHKSGVVEGFTKQYSLHKLVSFETHTMESAISREKQLKAGSRQRKNALIDSLNPEWDELAPY
 <HMPREF9370_RS01370--Neisseria_wadsworthii_9715--84.95 (SEQ ID NO:199)
 MKQPAVYIMANRKNGLVLYIGVTSNLVQRAYQHKQHIVKGFTDRYNLEMLVWVLEHKTMTESAIQREHTMKHWLRKWKLELIEKNPNPDWRDLWQEIIG
 <KKY_RS11100--Pelagibacterium_halotolerans_B2--70.21 (SEQ ID NO:200)
 MGGYVYIIANRRRGATYIGVTSDIARLYQHREKQTTGGHASRYNISRLVYFEVHATMPLAIQREKNLKSQRQAWKFELIENVNPHWDDLWEEIVKSPF
 <KKY_RS07555--Pelagibacterium_halotolerans_B2--64.52 (SEQ ID NO:201)
 MGGFVYIMAHRRRGATYIGVTNNLVRRVYEHRSKGGGHARSYSIRLVWFEGHADIRDAIHMEKRLKKWQQAQWKFELIENANPHWDDLWEEVAGSF
 <R2APBS1_RS00235--Rhodanobacter_denitrificans--78.49 (SEQ ID NO:202)
 MDRRQPCVYMLASKNGVLYIGVTSDLAKRIWQHRNDVVEGYTHRSVHTLVWYELHLMTESAILREKTLKAWKRDWKMRLINEGNAEWRDLPTIL
 <R2APBS1_RS17530--Rhodanobacter_denitrificans--79.35 (SEQ ID NO:203)
 MRERLPCVYLLASQRNGTLYVGVTSDDLPAIWOHRNNVAGFTQRHHVHGLVWFEVRDTMTESAIAREKAIKAWKRTWKIELIEKSNPYWRDLYAEICG
 <DSC_RS01625--Pseudoxanthomonas_spadix_BD-a59--82.80 (SEQ ID NO:204)
 MAKQPCVYLLASNRNGTRYVGVTSNLVGRTWQHRAHAVQGTDRHDIRLVYQMHDTMTESAI VREKRIKKWRRAWKQLIDVFNPSWRDLWPDIVGEVPPKAN
 VHGFPHSRE
 <KUC_RS17050--Halomonas_boliviensis_LC1--65.06 (SEQ ID NO:205)
 MLTNRTHRVLVYIGVTNNLERRLYEHRQGLMPGFTKKYRCHKLVWLEESNSIEDAIRREKQLKAGSRQRKNALIDSLNPEWDELAPY
 <OWEHO_RS03685--Owenweeksia_hongkongensis_DSM_17368--61.63 (SEQ ID NO:206)
 MYFTYIILTNKNNTVLTGTSDKLRLRFQHKNMVGTSTARYNVGKLIVFQSFETFPPEAFELERRIKGWKRVRWKIELIEKENPDWDLAEDWDFNEYLK
 <BFO_RS02520--Tannerella_forsythia_92A2--56.70 (SEQ ID NO:207)
 MRPSGTHNYVYIILTNKNKVLTYGTVNDLKNRLYFHCNPKPFSAFTARYKCVYLVYERYSDIYHAI SREKQIKGLSREKKEKMI TEFNKQWSFLNDEI
 <G4Z01_RS08920--Pseudoalteromonas_citrea_DSM_8771--74.19 (SEQ ID NO:208)
 MQPTTYIILASKPQGTLYIGVTRDLKRIYEHKHHFVEGFTKRYKVTCLVYFETHCDMYSAITREKQLKNWRREWKINLIEKVNPNDWDDL YLSI
 <ON00_RS14795--Sphingomonas_echinoides_ATCC_14820--67.07 (SEQ ID NO:209)
 MASQRNGTYIILGSTINLAKRAWHRNGLVDGFTKRYGCKLLVWYEARADWEDARQRELQMKWEKRRWKVDEIEGLNPDWEDLYDRIAQP
 <ON00_RS15725--Sphingomonas_echinoides_ATCC_14820--62.34 (SEQ ID NO:210)
 MTNKPFGVLYIGVTADLARIQAHREGRGSSFAKRWDCRKLVLIESHDRIEDAILREKRLKDWNRVWKCRLIAEANPNWDDLWETING

FIG. 11G

<HMPREF9453_RS10590--*Dialister succinatiphilus*_YIT_11850--68.24 (SEQ ID NO:211)
MGYVYILSNBYRNLYIGVTNNLQRRYQEHLSGAVDGFTKKYHVHCLLYEYQYNHIEEAIEREKQLKRWRSRKKEWLIKQINPEMVDLAEWQ
<SOLCA_RS15360--*Solitalea canadensis*_DSM_3403--61.05 (SEQ ID NO:212)
MRDHQYFVYILTNNKNAVLYVGVTTNNLERRLVEHYFGKHSFTKKYNVHFLIYYERFYIEDAINREKELKGWTRAEKESLIENENSNDWNFWNELVMEAWPPSR
EMLIEYKE
<NIASO_RS00745--*Niabella soli*_DSM_19437--66.67 (SEQ ID NO:213)
MQKGGFVYILTNNKLRVLYIGVTSNLPRLWEHEHHEIPGSFTDKYNVTLIIYEWFDTINSAIEREKQLKGWIRKKKEQLIARKNPQWKVFNEEISKEIYSL
LY
<NIASO_RS01430--*Niabella soli*_DSM_19437--70.27 (SEQ ID NO:214)
MRTVLYTGVTTNDIVRRVQEHRSKENKGFTSRYNCCYLYMYEYSDVRNAIAREKQLKRWHRWKNLNRIGQNPENKDLAEWLS
<NIASO_RS10780--*Niabella soli*_DSM_19437--60.23 (SEQ ID NO:215)
MKNNTCFVYILQSKKDRSFYVQGQCNLDLRRMSKHFDGFSKYTASKRPLSLVYFEVFLSRSEAIKRERQIKKTKSRKYILQLIENWRKDWQ
<GILLI_RS12875--*Gillisia limnaea*_DSM_15749--58.62 (SEQ ID NO:216)
MLSNKTRSMYLVGVTDLEKRISQHKNGEGSIYTKKYHLTHLMNFKQFTAIQVVIAREKQLKWNLIKSNPELRDLFEDLR
<GPOI_RS24255--*Gordonia polyisoprenivorans*_VH2--55.06 (SEQ ID NO:217)
MSGYVYILECSDGSYYVGSTVDLERRVNEHNDGKAAYTSRRRPVRVVFSEQPPTVEEAYRLEKRIQGWSRGKRRALIDGDFDLRLDARRPGARGKGIGGAE
PAP
<GFUN_RS06415--*Glaciecola punicea*_ACAM_611--76.34 (SEQ ID NO:218)
MKQPCVYITNRRNTTLYIGVTSNLVQRIYQHKIKLFGKFTAQYNVDKLVYFEQFEDMESAITREKRLKVWKRWDKNRLINEMNPCWKELYFDLMC
<PUV_RS08155--*Parachlamydia acanthamoebae*_UV-7--58.54 (SEQ ID NO:219)
MKIDEWIVYIIQAESGKLYTGITNDLQRRFQDHKQLKKGARFHFHSSPDKIVFSEKHPNRSEALKRESAIKKMNRQEKLKLTSMNMDL
<RHEL_RS05855--*Rickettsia helvetica*_C9P9--82.42 (SEQ ID NO:220)
MRRLGSKNRKIQGVRSRKNYETTCIYLIANKRNGTIYVGVTSNLVKRIYEHKNNNTNGFSKKYDCKILVFYIEIHEMTDSAIAREKQIKSGSRAKKLNLIEM
NVNWKDLYDEII
<RHEL_RS06570--*Rickettsia helvetica*_C9P9--77.42 (SEQ ID NO:221)
MKQPIVYIVTNKKNGLYTGVTNSLRVYQHKNDIIGKFSKYNCKILVFYIEIHEMTNSVINREKQIKAGSRDKKIKLIEQTNINWDDLYNEIIL
<RHEL_RS11805--*Rickettsia helvetica*_C9P9--77.11 (SEQ ID NO:222)
MDILCSNRNGTLCIGVTNNIIRRAYEHKQKTVKGFTAKYNIILKVYTEEFTDIKEALTREKALKKWKRSWKIELIEKTNPEWKDLIE
<RHEL_RS08700--*Rickettsia helvetica*_C9P9--81.54 (SEQ ID NO:223)
MKAYYVYILSNKYDGLTYIGVTSDIIRKRVWQHKQKIIKGFTTKYNLNKLIYFEQFNNINLAISREKFRNS
<BGIM_RS0115645--*Zavarzinella formosa*_DSM_19928--54.67 (SEQ ID NO:224)
MSEKPAATTWVYILRCRGGTLYTGITDPARRLKQHNAGTAARYTRSHRPVKMVYQEDQPGRSEALKREAAIKRLTRAACKLIAG
<ON09_RS00760--*Pedobacter agri*_PB92--60.44 (SEQ ID NO:225)
MERGGCVYIMTNSFNILYVGVTSDLQIRVSQHKNKVFEKSFISKYKCFKLAYFMFFPRIEDAIVEEKRIKGGSRQKIDILIELNPSWKDLYSDLD
<ON07_RS16365--*Gillisia marina*--60.53 (SEQ ID NO:226)
MNSYVYILKSDNTFYTGVTNSNLHKLRFHQSSKYPDSYTSQRLPVALVYAEFTDINSIAIATEKQVKWNWSRAKKVALIAGEFSELPNLAKNFSK
<ON07_RS06745--*Gillisia marina*--58.75 (SEQ ID NO:227)
MKFYTVYILECSDNTFYVGITSNLDRLEQHKNGNNTSSYTNKRRLILKWFQNFNPABAIKFEKQLKGWSRAKKIALIEEDWGRLILYSKNHTDFGKPDNTN
SI
<JOOMADRAFT_RS14810--*Joostella marina*_DSM_19592--59.14 (SEQ ID NO:228)
MLYKYFVYITNTQYRTTLYIGMTNNIQRRLSQHYFDSQNAKKSFAGKYKCIYLVYIEGYETPSEAIQREKELKKWRREKKNILIQNFNPKWETLNNDII
<Q7C_RS03890--*Methylophaga frappieri*--60.24 (SEQ ID NO:229)
MKKTDWTLIYVSTEDGALYTGITDRLARRLKQHRQELAGGAKFFRRSPALALLYQEKHADRRSASQREYAIKQLSKQQLKLINDKTAS
<Q7A_RS05040--*Methylophaga nitratireducens*--56.58 (SEQ ID NO:230)
MTKTEWTLYIIQAANGHLYTGITDRLERRLKQHQNHTGGARFFHTSDAEKVYVYQEQHPDRSSASRREAAIKLSRQQLNLIIGNPSDHY
<IC32_RS34730--*Bradyrhizobium liaoningense*_CCBAU_05525--78.65 (SEQ ID NO:231)
MAYYVYILASKKHGTLTYIGVTSDLVRRVYEHRTKAVPGFTTKYDVSKLVLFBIYDDAVTAIAREKELKKWRRDWKTRLIEBQNPWNDDLYPGIAN
<IC32_RS17050--*Bradyrhizobium liaoningense*_CCBAU_05525--80.23 (SEQ ID NO:232)
MSTRCYVYILASKIGGTLTYIGVTNDLIRVAEHKSKVIESFTEKYDVARLVYFEQFDDPENAIKREKRLKKWNRWVRLIEQHNPWNDDLYPGMAGPP
<IC32_RS32390--*Bradyrhizobium liaoningense*_CCBAU_05525--77.65 (SEQ ID NO:233)
MPYFVYILASRRNGTLTYIGVTNDLVRRIGEHKAKVIPGFARQYDVTLVYFEAFESVLEARSREHSLKRWRRRAWKLIIEELNPDWRDLMDLQSLAN
<IC32_RS11880--*Bradyrhizobium liaoningense*_CCBAU_05525--69.05 (SEQ ID NO:234)
MVSEDRSEHVLRYILASRRHGTLYIGITNSLQKRLEEHRAQGSEFVKRYGVHRLVYTESYERAEAEIAREKQVQKRWKRDWKIELIERDNPEWRDLSDLLV
<MICLODRAFT_RS28160--*Microvirga lotononidis*--75.58 (SEQ ID NO:235)
MHERYAVYILASQKNGTLTYIGVTGNLVQRVEQHKSLAIPGFTRRYRVTRLVYVEMFADVIDAIAREKQLKGWNRRAWKIKLIERSNPDWIELDPTAR
<MICLODRAFT_RS32655--*Microvirga lotononidis*--75.56 (SEQ ID NO:236)
MPYSVHLLASRRHGTLYLGVTTNDLARRVDRHKRVVPGFSARYDVCRVLVYEVHDIRIGEAIAREKSLKKWRRDWKITLIEEMNPDWSDLYPTLNR
<MICLODRAFT_RS25870--*Microvirga lotononidis*--66.29 (SEQ ID NO:237)
MKVYVYILASRPGGALYVGVTTNDLARRIWEHKTGHGSDHARRYRIDRLVYVESYPSIADAIQRETSIKRWPRVWKTNLILEMNPWTRDLYEELNA
<UU9_RS10505--*Rhodanobacter fulvus*_Jip2--84.04 (SEQ ID NO:238)
MTRELHPAVYIMASARNGTLYIGVTSNLVQVRVWQHREGVVEGFTTHRYRIKTLVWYEQHDTMESAITREKALKKWNDRDWKRLIEASNPDWRDLWPEIVGDEAP
HSSQPPARHSREGNPERDQPSIPQALDSRLRGNDGEAP
<UU9_RS08445--*Rhodanobacter fulvus*_Jip2--79.57 (SEQ ID NO:239)
MRERQPCVYILASRRNGTLTYAGVTSNLPARIWQHRSNVGEGFTRRHVHDLVWFELHETMESAIAREKAIKAWKRIWKIELIEKSNPYWRDLYAEICD
<UUA_RS10200--*Rhodanobacter thiooxydans*_LCS2--87.23 (SEQ ID NO:240)
MKRELQPAVYILASQRNGTLTYIGVTSNLVQVRVWQHREGVVEGFTTEKHNKMLVWYEQHATMESAITREKSLKKWNRVWKLRLIEAANPEWRDLWAEIVDAPSA
DVVRAEALDSRLRGNDERGTP
<SIDU_RS16385--*Sphingobium indicum*_B90A--78.65 (SEQ ID NO:241)
MKREISPTVYILASRRNGTLTYTGVTSDLVKRLYEHRNGLIEGFTSDYGVKRLVWFEMHDAIDGAILREKRIKKWNRQWKIELIEAGNPEWDDLALAFGEPLR
ERAGDGFPLSRE

FIG. 11H

<SIDU_RS02960--Sphingobium_indicum_B90A--75.90 (SEQ ID NO:242)
MANRRNGTLYLGVTSDLAARAYQHRNGLIDGFTKDHDCLLVWEAYEDLQEARHRELRMKKWKRAWKIELIERDNPQWRDLFETLF
<SIDU_RS00270--Sphingobium_indicum_B90A--71.91 (SEQ ID NO:243)
MSKPGVYIMASRRNGTLYLGVTSDLVKRAWQHRNGLGSEFSAEYRCHFLVWYEAYYDDLQEARLRELRMKKWKRDWKLRIETANPGWRDLFEEVAQ
<SIDU_RS13755--Sphingobium_indicum_B90A--61.63 (SEQ ID NO:244)
MERKEKGGWVYIMSDRYRGTIYVGVTSDLPAIHQHRSGNGSDFCARYGLNRLVWAERSDDITRCIEHEKRVKRWREWKFDLIERANPDWQDLFHLA
<AEQSU_RS10980--Aequorivita_sublithincola_DSM_14238--63.22 (SEQ ID NO:245)
MKVPHYITNTNTYKNVIYVGVTSDIVKRVKQHKYKGFSSRYNCDKLVFYREFPTMLEAIAFEKKLKAGNRKNKEKLINEMNPEWKDLSEGWYE
<TERRO_RS18865--Terriglobus_roseus_DSM_18391--64.71 (SEQ ID NO:246)
MPKVHGYWVYILSSRSKTLTYGVTSNLSVHRVLQHRDGGKSVFTTRYRVARLVHFESFQYIQDAIAREKDWKDWRRERKVALIEKGNPSWSDLAEGWTRGVETP
QYDWKRARE
<DESTI_RS14725--Desulfomonile_tiedjei_DSM_6799--65.17 (SEQ ID NO:247)
MQRQYCIYILTNRNTVLTGITRDLKKRVHQHREKLSKGFTKLYNIEKLVFYEFTDDVRAAIAREKQIKAGPRQKKIDLIEDMNPDWLDLYEEL
<MROS_RS02740--Melioribacter_roseus_P3M-2--72.04 (SEQ ID NO:248)
MNQYVYIMANKHNTLTLYGVTNNLIRRVYEHKNKLTNGFTKRYNCKLVWYIEHDTVPNAIKREKQIKAGSRKKKLDLINKLNPEWKDLYFDLI
<MROS_RS13825--Melioribacter_roseus_P3M-2--75.00 (SEQ ID NO:249)
MTNKSSTLYTGITNNLIRRVYEHKNKLMFGFTSKYNI TKLIYYEEFNDVKKAIEREKQIKGWLKKKKIELIEKENPEWKDLSKEWY
<O5O_RS0108155--Methylocystis_parvus_OBBP--77.27 (SEQ ID NO:250)
MRNPVYIVASKRNGTLYTGVTNNLPRRIWEHRDELDRDGTFRYGLNRLVWYEPHGDRTAIQREKTMKHWPRAWKVRLLIQMNPWDSLYETLAN
<O5O_RS0117310--Methylocystis_parvus_OBBP--75.82 (SEQ ID NO:251)
MRNPVYIVASKRNGTLYTGVTNNLPRRAWEHREGLVDGFTKQYGCRLLVWCEYYETMPEAIAREKQIKAGSRKKKALIEAMNPWTRDLYEELH
<O5O_RS0107155--Methylocystis_parvus_OBBP--72.09 (SEQ ID NO:252)
MRGGWVYILASRRNGTLYLGVTSNLDLARRAFEHRESRISGFTQRYGLNRLVWYEPHGDRTAIQREKTMKHWPRAWKVRLLIQMNPWDSLYETLAN
<RJP_RS02595--Rickettsia_japonica_YH--76.47 (SEQ ID NO:253)
MYWVYILCSNKHGTLTYIGITNNIIRRAYEHKQTIKGF TAKYNI IKLVYKEEFTDIQEALTRETNLKKWQSSVKIELIEKTNPWKDLIK
<AZOLI_RS22735--Azospirillum_lipoferum_4B--77.91 (SEQ ID NO:254)
MEERFFVYILASRRNGTLYLGVITSNLRGRMEHRDGIIPGFTDRYGVKRLVHVEPYATAEEAINREKRLKKWNDRWKIRLIEEHNPDWADLYERLNA
<MCC_RS03070--Rickettsia_rhipicephali_str_3-7-female6-CWPP--78.82 (SEQ ID NO:255)
MYWVYILCSNKHGTLTYIGITNNIIRRAYEHKQTIKGF TAKYNI IKLVYTEELTNIQEALTRETNLKKWQSSVKIELIEKTNPWKDLIE
<CSE_RS07445--Caldisericum_exile_AZM16c01--74.44 (SEQ ID NO:256)
MDEKFFVYIILSNKHNNVLYVGTNDLKRRIYEHKNKLVGFTKKNVHKLWYIEFEDPVNAISREKQLKDFRREKVKLINDFNPWEDLYEKIL
<BJ6T_RS12980--Bradyrhizobium_japonicum_USDA_6--73.12 (SEQ ID NO:257)
MKQPCVYMFANRRSGTIYVGVTSNPPRAFEHHEGLVKGFSKYGCKLLVWYELHATMTDAIAREKQIKGGSRAKKLALIESANPDWTDLYETLI
<BJ6T_RS13695--Bradyrhizobium_japonicum_USDA_6--77.27 (SEQ ID NO:258)
MLASRRDGAIVGTINDLVRVYEHRIKAVPSFTTKYNI TRLVWFEAYDDPISAI SREKELKKWKRAWKTQLIEKKNPNWDDLYESICN
<A33Q_RS07875--Indibacter_alkaliphilus_LW1--65.52 (SEQ ID NO:259)
MQKGGAVYIILTNKHHTTLYIGVTNDLQRRLYEHQNGLKPNFSFVKYNLKNLVYEGFHSIDEAIMREKQLKGWTRQKKNILVNTMNPWEWDLSDVRSW
<B879_RS10790--Cecembia_lonarensis_LW9--64.84 (SEQ ID NO:260)
MRRHEYFVYIILTNQKNTVLYVGMTNDLSRRLSEHTDKYNPEGFTAKYKCKFLHWSYQYVRDAIAREKEIKKWRREKKEKLISENNPNWDFLNKEVVEN
<S7S_RS09595--Alcanivorax_pacificus_W11-5--53.93 (SEQ ID NO:261)
MLEDEMFGSRRYVYIILTNRFHSTLYTGMTDPDLRRRVIIHKLRSTTCFYNPRSNKLVYFEAQYAYLAVRRERMIKAMHRPDLALISRHNPWRDLAMDLD
IIGTS
<P24_RS11995--Oceanibaculum_indicum_P24--74.42 (SEQ ID NO:262)
MAYYVYILASRRNGTLYVGTGNLPRRIYEHREELVPGFTKQYGVKRLVYAEAHEDVRIALQREKNLKHWSRAWKLALIEAGNPSWRDLF
<P24_RS13240--Oceanibaculum_indicum_P24--73.26 (SEQ ID NO:263)
MPYTYIILASKPGGTLYIGITNDLIRRIEHREAAVPGFTKRYKIHLVYFEEHATSEQAIRREKRLKTWLRWVQLIETDNPWWDLYPLLL
<P24_RS12675--Oceanibaculum_indicum_P24--70.93 (SEQ ID NO:264)
MAFFVYILASGEQGTLYIGVTNDIARRVHEHKGAGADGFTKRHGVSRVLYYEMFDTAADAIQREKAMKRWKRDWKNLIDRTNPWDRDLYETINA
<TH3_RS04880--Thalassospira_xiamenensis_M-5_DSM_17429--78.02 (SEQ ID NO:265)
MKQPCVYILASNRNGTLYIGVTADLRARIWQHRTGTDKGFTQRYNVGQLVHFELYSMENAILREKRLKKWRRAWKLDLIEKHNPDWDLFDSLNA
<NA2_RS15380--Nitratireductor_pacificus_pht-3B--69.15 (SEQ ID NO:266)
MGGYVYILASRKNGTLYTGVTSDLEQRMWEHREGITGGFVKRYGVHRLVWYRDYADIRDAIEDEKRIKRWRRKWKLELIDGINPEWDDLYDIINH
<B176_RS0117755--Pedobacter_arcticus_A12--64.84 (SEQ ID NO:267)
MERGGCIYILTNNKNTLYVGVTSDLKSKRLNEHKNHIYPRFSARYNLTICYIYETFTSIVEAIAREKEIKKWRREKKEKLINQLNPKWDDLSIGIQEW
<ON41_RS14210--Nocardia_transvalensis_NBRC_15921--75.00 (SEQ ID NO:268)
MGRRPWQPPCVYILASRRDGLVYVGATRNVLGRVWQHKNDVASFTRRYGVHTLVWYEAHETLECALLREKRLKKWNRAWKIRLIEAKNPWDDLYPTIL
<BN396_RS06190--Occidentia_massiliensis--76.47 (SEQ ID NO:269)
MYWVYILCSNKHGTLTYIGITNNIISAYEHKQKIIKGF TAKYNI IKLVYTEEFTDIQKALTEKNLKKWQSSVKIELIEKTNPWKDLIK
<HGAL_RS09795--Alistipes_senegalensis_JC50--63.33 (SEQ ID NO:270)
MESGYVYIILTNKHLTVLYVGVTTCLSRMKEHAAGLSAFTRRYNVHKLIVVEEYSVIRDAIAREKQIKSWSRKRKMELVDSRNPLWKDLNSGIV
<BN60_RS12615--Reynanella_massiliensis_521--60.00 (SEQ ID NO:271)
MREHRYWVYIILVSHSLSMAYVGVTDNLGRRVFEHRNGLGSEFAKKYQVRRLVYTEEFQIEEATAAREKQLKGWKIRKKNELVRAANPEWEDLMPEETRISPLS
S
<BN30_RS07845--Fenollaria_massiliensis--57.69 (SEQ ID NO:272)
MMNKIYIYMLRCKDGSYITGYTDNIYKRIKKHRSKGAKYTRGRGPFELVHYESFLTKEEAMQREAAIKKLNKEAKLKLIEENAKKE
<HMPREF9695_RS13465--Afipia_broomeae_ATCC_49717--79.57 (SEQ ID NO:273)
MAYYVYILASKNGTLYLGVTDNIVRRGYEHRTKAVDGTTRYGVDKLVWFEVHDDALTAITREKELKKWRRAWKIRLIEENNCEWVDLYPEIAS
<HMPREF9695_RS16140--Afipia_broomeae_ATCC_49717--71.76 (SEQ ID NO:274)
MTEDGVTYHIYILASQPNGTLYIGITNNLALRLSQHKGGRGSEFAQRYAVNRLVYVEGYDTPSDAIRREKQLKKWNRAWKIELIEKNPDWQDLSDLAIFGIS
<HMPREF9695_RS05595--Afipia_broomeae_ATCC_49717--68.54 (SEQ ID NO:275)
MDQTFVYIILANRARGVLYVGVTDNLSRRLAEHRAKLVPAFTTRYGVTCVLHIETYSILEAREHSHLKRWRRTWKFELIEKDNPEWRDLSDMSS

FIG. 111

<GARC_RS10445--Paraglaciecola_arctica_BSs20135--78.49 (SEQ ID NO:276)
MKQPCIIYILTNDNNSTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQLEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<GARC_RS17625--Paraglaciecola_arctica_BSs20135--78.49 (SEQ ID NO:277)
MKRPCLYILTNDNNSTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQLEDMDSAILEKRLKKWNREWKNNRLINELNPSWKDLYFDLVG
<GARC_RS09705--Paraglaciecola_arctica_BSs20135--76.34 (SEQ ID NO:278)
MKQPCIIYILTNDNNSSLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWKDLYFDLVG
<HMPREF9718_RS09240--Sphingobium_anoikuyae_ATCC_51230--79.78 (SEQ ID NO:279)
MTREISPTVYMLASRRNGTLYGVGTSDLIKRLYEHRNGLIEGFTKDYGVKRLVWFEPHDSMDSAILEKRIKKWNQWKIELIEAGNPWDLDLALGFGFAPLP
RD
<HMPREF9718_RS01145--Sphingobium_anoikuyae_ATCC_51230--72.04 (SEQ ID NO:280)
MRPAVYIMASGRNGTLYTGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQLEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<HMPREF9718_RS10945--Sphingobium_anoikuyae_ATCC_51230--70.65 (SEQ ID NO:281)
MKQPCVYILASQPYGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWKDLYFDLVG
<HMPREF9718_RS05515--Sphingobium_anoikuyae_ATCC_51230--60.00 (SEQ ID NO:282)
MTGGWVYIMTNRAYGVLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWKDLYFDLVG
<ASHG_RS05985--Anaerococcus_senegalensis_JC48--68.18 (SEQ ID NO:283)
MSDYYVYILANNSNSTIYIGYTYNLEKRLYEHRNHVVEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<TVNIR_RS031575--Thioalkalivibrio_nitratireducens_DSM_14787--80.65 (SEQ ID NO:284)
MKQPCVYIMTNRRNGTLYGVGTSDLIKRLYEHRNGLIEGFTKDYGVKRLVWFEPHDSMDSAILEKRIKKWNQWKIELIEAGNPWDLDLALGFGFAPLP
<ECHVI_RS13365--Echinicola_vietnamensis_DSM_17526--68.13 (SEQ ID NO:285)
MKKGGCVYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<ECHVI_RS05155--Echinicola_vietnamensis_DSM_17526--58.67 (SEQ ID NO:287)
MYILECSNGAFYTGSTVDLDRLEKHQAGHGAFNFKRYLPVKLVYFEEYARIDQAFYREKQIQRWGRSKKIALINGDLDRLEHLSACQNNNSFKYYNSKKE
<C943_RS06545--Mariniradius_saccharolyticus_AK6--67.39 (SEQ ID NO:288)
MEKGGAIYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<C943_RS12305--Mariniradius_saccharolyticus_AK6--63.74 (SEQ ID NO:289)
MRKHEYFYVYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<FAES_RS04190--Fibrella_aestuarina_BUZ_2--59.77 (SEQ ID NO:290)
MKNYAFYVYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<C900_RS06425--Fulvivirga_imtechensis_AK7--67.42 (SEQ ID NO:291)
MSKGGVYIIVSNKHRTVLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<DDD_RS13120--Nonlabens_dokdonensis_DSW-6--64.37 (SEQ ID NO:292)
MAFKSLHLYWYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<DDD_RS10455--Nonlabens_dokdonensis_DSW-6--60.23 (SEQ ID NO:293)
MKTMYTYIMAHINKKLFYVGTNNIKQRVFDHKNATFETHVGYKNIKALVWFEPHDSMDSAILEKRIKKWNQWKIELIEAGNPWDLDLALGFGFAPLP
ERFKKD
<DDD_RS10485--Nonlabens_dokdonensis_DSW-6--67.44 (SEQ ID NO:294)
MASYEHKYYVYLLSNKKDGLVYVGMTDDIKHRINQHKKEKYKAFSKRYQTCKLVYFEEFLWVQDAQAREKKLKKWKRWQWKIDLEKDNPEWDLALAEALWDDE
NY
<DDD_RS06510--Nonlabens_dokdonensis_DSW-6--65.22 (SEQ ID NO:295)
MTPGYVYIILTNNKNTTLYGVGTSDLIKRLYEHRNGLIEGFTDKYNCTKLVEYELDDKYEALSREKQLKRRSREKNNRLINSINPKWRDLSLDWYR
<DDD_RS00990--Nonlabens_dokdonensis_DSW-6--62.35 (SEQ ID NO:296)
MPRDYAFHYVYIMTNFKNGTLYIGFTGNLQRIILHKGAGGKFTSKYELDRVLYKEFFQYADKGIAREKQLKTKWKREWKNNRLINELNPSWEDLYFDLVG
EYKED
<ON35_RS02405--Wolbachia_endosymbiont_str_valisugana_of_Drosophila_suzukii--74.71 (SEQ ID NO:297)
MKLSYIYIILTNNKNTTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<C693_RS0119170--Dyella_ginsengisoli_LA-4--79.57 (SEQ ID NO:298)
MRERRPCVYILASGRNGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<HMPREF9696_RS08305--Afipia_clevelandensis_ATCC_49720--75.27 (SEQ ID NO:299)
MAYYVYMLASKKNGTLYLGVGTNDIVRRGYEHRTKAVDGTTRYGVDKLVWFETDYDDALTAITREKEFKKWRVWVKIRLIEESNREWIDLPGIAS
<HMPREF9696_RS01700--Afipia_clevelandensis_ATCC_49720--73.33 (SEQ ID NO:300)
MSYVYILASAPGGTLYIGMTNDLIRRVYEHREGFADGFTKRYGVKTLVYFEAHTASGAIQREKNIKHSREWKIDLIIRDNPWDRDLYEDIVR
<HMPREF9696_RS10500--Afipia_clevelandensis_ATCC_49720--75.86 (SEQ ID NO:301)
MTNQNGTLYTGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<HMPREF9696_RS12170--Afipia_clevelandensis_ATCC_49720--68.89 (SEQ ID NO:302)
MPHYVYILASQNGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<C725_RS13040--Pacifimonas_flava--64.04 (SEQ ID NO:303)
MPGYTYIMASRKQGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<C725_RS08550--Pacifimonas_flava--67.44 (SEQ ID NO:304)
MTYGGHYVYIMTNRRHGTLYTGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
<S58_RS23140--Bradyrhizobium_oligotrophicum_S58--76.60 (SEQ ID NO:305)
MAHYVYLLASKKHGTLYLGVGTNDLIRRVYEHRTKSVPSFTSRDYIGKLVWFEBIYDDAIAAISREKELKKWRRGWKIRLIEEQNPWDDLYLGICG
<S58_RS33580--Bradyrhizobium_oligotrophicum_S58--74.16 (SEQ ID NO:306)
MSARSYVYILASHIGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG
SGILDRFVEPGDDSD
<S58_RS34635--Bradyrhizobium_oligotrophicum_S58--68.24 (SEQ ID NO:307)
MYVYVYILASGRHGTLYIGVTSNLVQRIYQHKNKQKGFSSKYNLTKLVYFEQFEDMDSAILEKRLKKWNREWKNNRLINELNPSWEDLYFDLVG

FIG. 11J

<F752_RS0100795--Nafulsella_turpanensis_ZLM-10--65.17 (SEQ ID NO:308)
MTGKGYYIYILTNKGRITVLYIGVTANLYARIQQHKDGSVFTQRYRCTDLLYYEFYPSIEEAIQREKQLKNWKREWKEQLIQKVNPLLTDLSGEIEDMQ
<F752_RS23840--Nafulsella_turpanensis_ZLM-10--68.12 (SEQ ID NO:309)
MKHHNYFVYIILTNKNKTVLYTGLTNDLERRVFEHKEKSRKSAFTARYHIGKLVYERFSSHIEEAIEREKQIKAGSRI
<F752_RS0100765--Nafulsella_turpanensis_ZLM-10--54.02 (SEQ ID NO:310)
MSRQHNYVYIITNPKKTVLYTGMTNNLTARLQQHYANRGKKETFAGSYFCYCLLYYERHQYVQHAIEREKEIKGYSRQQEKDLISEFNPEWNFNLVNE
<F752_RS24465--Nafulsella_turpanensis_ZLM-10--68.97 (SEQ ID NO:311)
MSNKNRTTLYIGVTSDELVRVLQHKAGLSKFTKRYNLFDDL YFERIMGMENAIAREEPAPLN
<F752_RS0100720--Nafulsella_turpanensis_ZLM-10--52.27 (SEQ ID NO:312)
MSHHNYIYFLMNSAKTGYTGTNGLRNCLDQHFSERGERRKPLQEKYFCFHLVYWERHPNVLDAIRKEKEIKGWRREKKVKLIEDFNPEWKFLNGEV
<D778_RS13980--Xanthomarina_gelatinilytica--70.45 (SEQ ID NO:313)
MYKTIHQYYLVLTNKTNGTLYIGVTNDLERRMFEHKNKL VAGFTKKYGLNKL IYFETYQYVND AIKREKNMKKWKRWQWKIDLIEENPCWNDLSKDWVYLID
<D778_RS05950--Xanthomarina_gelatinilytica--58.24 (SEQ ID NO:314)
MKTGSHNYVYIILTNKSKTVLYTGTNNLKERLYAHTNPLPFSKAF TAKYKCFFLIYYEHFFEVDEAIEREKQIKGWLRIKKEELIATKNPEWRFLNDDV
<ADICEAN_RS13380--Cesiribacter_andamanensis_AMV16--68.89 (SEQ ID NO:315)
MQRGCCVYIISNKNRSVFYIGVTANLYTRTSEHKQIGS AFTKKYNCTDLVYETFGSIEEAIKREKQLKEWKRAWKERLIKESNPEMRDLYDDLK
<ADICEAN_RS10850--Cesiribacter_andamanensis_AMV16--67.39 (SEQ ID NO:316)
MANRTYVYLLSNKSGSVLYVGTNDLQRRVLEHKSGYNPGSFTSRYKVHKL VVWEELPSFEEAIAREKQLKGGSRARKVQLIQAHNPQWLDLSQGWYD
<MPL1_RS012800--Methylophaga_lonarensis_MPL--81.72 (SEQ ID NO:317)
MKQPAIYIMTNKRNGTLYVGVTSHLIKRVWEHKNNVVEGFTQKYRLHLQVLVYFELHDSMIGAISREKQIKKWRREWKNINLIEKLNADWRDLYSLL
<MPL1_RS08895--Methylophaga_lonarensis_MPL--72.83 (SEQ ID NO:318)
MQKLPVYIILTNRRNGTLYTGVTSNLP HRIWQHKN SVIDGFSKRYATHQLV FYETHESMESAILREKQIKSGSRKRKLELIEAFNPRWEDLYNSIL
<MPL1_RS12000--Methylophaga_lonarensis_MPL--73.32 (SEQ ID NO:319)
MKNYVYIILSNASNQVLYIGVTSHLIRRIYEHKNHLIDGFTKKYNVKKLVYFEQTS DVSDAIAREKQLKRWREKKNHLISQSNPDWHDLYPDLVESGG
<MPL1_RS06665--Methylophaga_lonarensis_MPL--73.33 (SEQ ID NO:320)
MTVKNYVYIILSNASNQVLYIGVTSHLIRRIYEHKHLIDGFTKKYNVKKLVYFEQTS DVSDAIAREKQLKRWREKKNHLISQSNPDWHDLYPDLVESGG
<MPL1_RS05595--Methylophaga_lonarensis_MPL--54.29 (SEQ ID NO:321)
MKTTDWTLYII EAANGHLYTGITD LTRRLQQHQSGKGARFFASSQAKQVYQESYFDRSSASRREAAIKKLSRQAKLDLIASSISQSR
<MESS2_RS13445--Mesorhizobium_metallicurans_STM_2683--64.04 (SEQ ID NO:322)
MTGYAYMTASQKRGTIYIGVTNDLGRRMPEHKS GHGSRFTNRYGAQCLVWYEEHFDIRDAIQREKSLKRWRPQWKIELIERTNPEWFNLFRTGW
<MESS2_RS06060--Mesorhizobium_metallicurans_STM_2683--65.17 (SEQ ID NO:323)
MTGYAYMTASQKRGTIYIGVTNDLGRRMPEHKS GSGSRFTSRYGVQRLVWYEEFFDITDAIQREKSLKRWRPQWKIELIEKANPEWFELFRGTGW
<N646_RS12100--Vibrio_alginolyticus_NBRC_15630--ATCC_17749--71.26 (SEQ ID NO:324)
MDKTPCVYIILTSNNNAVLYIGVTSQLPQRLWQHKN GVVVEGFTKKYNVHKL VYFELFDDMANAINREKQLKQWKREWKLALVRETNPNFLDLSTEYQ
<PGOND44_RS10920--Psychroflexus_gondwanensis_ACAM_44--67.82 (SEQ ID NO:325)
MPKRNHHNYSVYIILTNKPKGSLYIGVTGGLDDRMEHRLKGC GSKFTSTYKLRVLVYFEFQYIEDAITREKQLKNWHRQWKINLIEQDNPEWEDLYKPLD
<H142_RS0105915--Robiginitomaculum_antarcticum_DSM_21748--66.67 (SEQ ID NO:326)
MQPCVYMMASKRRGTIYIGVTSDLASRDWEHWSRANPKSFTARTNARYLVWYEEHEIMTDAIQREKSLKRWRDWDKIELIEKTNPQWHPINPDTGEFMYDGDW
S
<H142_RS15060--Robiginitomaculum_antarcticum_DSM_21748--68.54 (SEQ ID NO:327)
MFFTYIITNRKQALYTGHTEDLVFRMQQHIEGVFEGFSKEHDCKKLMWFELHETRDGAFERERRIKEWKRRWKIEMLEQLNTNWDLYINISEADIYNPARY
FDQNTYNPKSGDGPRFSTG
<H142_RS0111245--Robiginitomaculum_antarcticum_DSM_21748--73.56 (SEQ ID NO:328)
MHFADSLAYVYMMASMKNGTLYLGMTNDLERRVWEHKNNENPKSYTARYH I HRLVYFETFDSPDTAIRREKYLKGKSRQFKIELLEMKNPHWIELYEFKMRG
GNI
<H142_RS15015--Robiginitomaculum_antarcticum_DSM_21748--52.33 (SEQ ID NO:329)
MLEQPTVYVYDES DLHAAMFYTYIILTNRPKGTLYFGHADDLPATIAAHKAGKVDDL NPKHMACKMLVWFEAHPDLGSAFNRTQDMLTWNRSQVIKRIKLLNPKW
ADLSDTLTPDILADPARKFPSPDNALPSV
<H142_RS16595--Robiginitomaculum_antarcticum_DSM_21748--57.97 (SEQ ID NO:330)
MYIGVTSDLASRDWAHYSRNPKSFTARANARYLVWYEEHEIMADAIQFEKSLKRRYRDRKIKLIKTTNRVGTR
<SX1_RS00320--Sphingobium_xenophagum_NBRC_107872--76.09 (SEQ ID NO:331)
MDRQPAAYIMASDRNGTLYIGVTSNLMARIYQHRTKAMDGFTARYDVNRLVWFEMCDTMEVAIVREKQLKNWKRKWKLDLIEAGNPRWRDLAEDLGFDPLSPS
S
<SX1_RS05670--Sphingobium_xenophagum_NBRC_107872--76.40 (SEQ ID NO:332)
MKREISPSVYILASRRNGTLYTGTVDLVKRLYEHRNGLIDGFTKDYGVKRLVWFEPHDLRESAILREKRIKKWNRQWKIGLIEAENPDWDDLALGFGEPLP
SLRSD
<SX1_RS13610--Sphingobium_xenophagum_NBRC_107872--71.74 (SEQ ID NO:333)
MKQPCVYILASQPNGTLYIGVTSNLI GRMLQHRTDALAGFTSRYGVHRLVRFEACDTMEQAILREKQLKRWHRQWKINLIESENPDWVDLAVNLGLPPVHPQL
THDGC
<SX1_RS01860--Sphingobium_xenophagum_NBRC_107872--75.90 (SEQ ID NO:334)
MASRRNGTLYLGVTSDLAGRAYQHQNALVEGFTKRHGCTMLVWFESYDDLQEARLRELQMKKWKRAWKIELIERENPQWRDLFETLF
<SX1_RS17025--Sphingobium_xenophagum_NBRC_107872--74.16 (SEQ ID NO:335)
MAKPGYVYLMASSQNGTLYLGVTSDLVQRAWQHRNGFGGGEFS AKYRCHLLVWY EAYDDIQEARQRELRMKKWKERKWL RMI BEVNPTWRDLFEDIAT
<SX1_RS11670--Sphingobium_xenophagum_NBRC_107872--59.04 (SEQ ID NO:336)
MPFTVTPAKAGVSGADALQ LDMARGGYTCVMTNKP RGVLYIGVTADIATRAQHHRNGTGS AFCKKYGLTRVLVPEPHDDIRIAIAREKALKAWKREW KIRLVE
QANPDWRDMSDVI F
<SX1_RS00490--Sphingobium_xenophagum_NBRC_107872--54.29 (SEQ ID NO:337)
MLHCAGAFVYGHGTDNLEQRIAGHEQSGSIAGFTSDRLPVKLVWCECTQRIEAL ENERRLK GWSRAKMMALIRGDWDEISVLARSRKDAPSTGSG
<B051_RS0107600--Pontibacter_roseus_DSM_17521--59.77 (SEQ ID NO:338)
MPNHN YFVYIITNRTKKVLYIGMTNTLQRRLEE HGANAGIPDTFAGKYHCHYLVYERFTYVQQAI DREKELKKWSRKKKEALI QSMNPTWNFLNEEVKEHE

FIG. 11K

<B067_RS0102355--Dasania_marina_DSM_21967--78.49 (SEQ ID NO:339)
MAKQPAVYIITNKNRGTLYIGVTGHTFQRIWQHRRNNELDGFTKKYQVHKLVYFELTDCITTAIAREKQLKSWRRSWKIELIEQDNPKWNDLWFEIID
<B094_RS0102150--Legionella_shakespearei_DSM_23087--71.43 (SEQ ID NO:340)
MNSFYTYIMASKRNGTLVVGSTSDLIKRVWEHKNKVI PESFTAQYNNVHMLVYELHEHYMEAAREQRFKNWRRKWKLNLIHVNPKWNDLYEEICR
<B094_RS0113505--Legionella_shakespearei_DSM_23087--52.81 (SEQ ID NO:341)
MFWVYILGCNDKSYTGHTDNLENRLYQHHRNRMISGCTATRLPVTLLYTASFPTRMEALAAENQIKGWSRKKKEALIKENWHEELSFHARRKKRES
<B153_RS0134370--Spirosoma_panaciterrae_DSM_21099--57.14 (SEQ ID NO:342)
MNGYAFYVYIVTNSSEKTVLYTGMTNDLVRRLEEHYESRGQDPKFAGKYCHRLIYWEHHQYVYNAISREKEIKGGREKKVALINSINPEWRFLNDDIQS
<B154_RS0111705--Segetibacter_koreensis_DSM_18137--59.77 (SEQ ID NO:343)
MKYAGCIYIILNSHHTVLYTGVTSDLKKRIFQHQNKEFSNSFTSKYNCNCKLVYVENFIQIBEAIAREKQIKAGSRKAKETLINNFNPEWNDLGELIKNW
<B156_RS0107960--Spirosoma_luteum_DSM_19990--59.77 (SEQ ID NO:344)
MNGYAFYVYIVTNPRTALYTGMTNDLVRRIQEHYESRGQVGKFGAGKYCHRLIYWEHHQYVRNAITREKEIKGWRREKKVELINSINPEWRFLNDEIQE
<B158_RS0101330--Kangiella_aquimarina_DSM_16071--80.65 (SEQ ID NO:345)
MKTQGCVMYMLASKRNGTLTGVTSNLLQVRVWQHKNNAKGFTEKYDVQQLVYVEVTEDIVSAITREKQIKWNNRVWKIRLIBEMNPEDLYELIDEKVRK
<B158_RS0103990--Kangiella_aquimarina_DSM_16071--77.42 (SEQ ID NO:346)
MDKQFYVYIILANQRNGTLVVGVTSDLVKRVWQHKSNNVKGFTQKYGVHQLVYVEQHGNSYSAIQKEKRLKEWKRKWKLDLIEGTNPDWKDLYDEII
<B158_RS0110705--Kangiella_aquimarina_DSM_16071--57.78 (SEQ ID NO:347)
MRKGYVYIILNPRKTVLYTGTTNNLGKQLFEQKNRISERLTRSYGCNELVYIEITESLEKAREEEEEIKSECRIGKETLIQRFNPEWKDLSEGELLAVKKR
<B160_RS0101475--Niabella_aurantiaca_DSM_17617--65.48 (SEQ ID NO:348)
MQKGGFVYIILNKRTRTVLYIGVTSNLRARLWEHEHETPGSFDTKYNVTFLLIYEFWDTINSIAIEREKQLKGWVRKKKEQLIAQKNPDWKFVNEEISKEIYSL
LY
<C504_RS0104500--Ahrensia_kielensis_DSM_5890--71.76 (SEQ ID NO:349)
MKRKGVYIILASRNGTLTGVTSDLARRLLEHEHIGSGQFTAKYNNVKRLVWYEEYDLVTTAI TREKTIKKWPRQWKINAIEEINPDWNNLKSLL
<C507_RS0105140--Amorphus_coralli_DSM_19760--76.74 (SEQ ID NO:350)
MAFYVYLLANRRNGTLYLGMTDDLARRVHEHREAINRGFSQRYGVKQLVWYEVHDSREAAFTREERQMKAWKRGWKLRAIEEMNPDWADLYETLQ
<C507_RS0118770--Amorphus_coralli_DSM_19760--72.09 (SEQ ID NO:351)
MAFYVYILASEPNGTLTGHTDDLARRVHEHRTLRLPGFTRTYGVAGLVWYEVHDTREAAFVREERQIKKWRRRWKRLRIEGLNPTWSDLYDSL
<C507_RS0108135--Amorphus_coralli_DSM_19760--72.09 (SEQ ID NO:352)
MAFYVYMLASRRNRTRYVGMTDDLARRIYEHKTGALPGFTRTHRVKTLVWYEVHDSREAAFIRERQIKKWREWKIALVPQQNGEWRDLYETLPL
<C508_RS0103935--Anaeromusa_acidaminophila_DSM_3853--70.00 (SEQ ID NO:353)
MERKGYVYIILNCTRSVLTYIGVTSNLRQWYIQRHRLMDGFTDQYNNVILVYECYBSIREEIIREKQLKAGSRLLKKEALINEKNPSWQDLYTEVVG
<B155_RS0112595--Balneola_vulgaris_DSM_17893--69.41 (SEQ ID NO:354)
MANIYVYIILSNYSKMLYIGVTNDLERRVYEHKNKLI EGYSKRYNLFQLVYVESTHDIGRAIEREKQLKGWTRAKKVNLIIKSNPYWNDLAAEWY
<B155_RS0100110--Balneola_vulgaris_DSM_17893--67.78 (SEQ ID NO:355)
MSKGYVYFLTNFQRSTLYIGITSNLTARRIWEHKQKGSAFTTRYKQTLIYAEEDFSIVDAIDREKQLKNWHKEWKNLVKQHNPRLEDWYETLRA
<WNO_RS03555--Wolbachia_endosymbiont_of_Drosophila_simulans_wNo--75.56 (SEQ ID NO:356)
MKGYIYIILASERNGTLTYIGITSNLTAKRIWEHKSASGFTSKYNNVCKLVYFEFQDINLAISREKLLKSWQRKWKINLIEKTNQEWKDLYDEIVSG
<F550_RS0112240--Henriciella_marina_DSM_19595--77.65 (SEQ ID NO:357)
MAFYVYIMTNKPNGLVYTGMTDDINRAWEHREHVIGGFTAKYNCMLVWVEVHETRESAFERERRIKKWERAWKIRLIEEVNPDWRDLADALM
<F550_RS0101350--Henriciella_marina_DSM_19595--71.59 (SEQ ID NO:358)
MRSSWLYIVTNKRNGTLTGVTSDLTARIWQHRNGHYPGFSRRYGCTRLVWFETHDDISEAIAREKRIKRWRDVKLALIEKDNLDWVDFDELGTGWVPSRF
AASGMTKIGSDR
<F550_RS0107880--Henriciella_marina_DSM_19595--67.44 (SEQ ID NO:359)
MARFDLIATYIMANKNGTYTGSTSDLWTRIGQHKSGKGLFTAKYGCRLVWVEFHDSMAAAHRRERRLKTWPRQWKINLIEEVNPDWKDLSMTLPFT
<F550_RS0111470--Henriciella_marina_DSM_19595--67.02 (SEQ ID NO:360)
MKPFYVYILASQKNGTLTYTGSTDDISKRVWEHKQGHGAFTTRYKAVHRLVWLEPYDSRDSAKAHEYQIKAWKRSWKIELIETTNPDWQDLYLKLNH
<F455_RS0101380--Rudaea_cellulosilytica_DSM_22992--78.49 (SEQ ID NO:361)
MEKHPCVYIILASGRNGTLTYIGVTSDLIKRIYEHNRDVVEGFTNKYIVHDLVWYEQHAEMLSAIARERQKWKSRKAKIVLVERHNPQWKDLWSDLIEG
<F458_RS19360--Salinimonas_chungwhensis_DSM_16280--73.53 (SEQ ID NO:362)
MLYLGVTNLIQRIYQPKKKLVKGFTSQYGLTKRVYFERWKDMENAI VRENRLKHWNLQRKNKLI SEVNSD
<F551_RS0114470--Hirschia_maritima_DSM_19733--70.11 (SEQ ID NO:363)
MVYAVYIVTSQDFGLTYTGVTNNLVRRRIWEHKQCGSDFTRKHGVDRLVWFENHDNINAAIKREKNIAWQQRQWKINLIEKDNPHWIDLYPEQGF
<F598_RS0104920--Psychroflexus_tropicus_DSM_15496--63.33 (SEQ ID NO:364)
MSYVYVMTNKSNSLIYIGVTNDLDERVKEHKLKVHPNSFTSRYNCKNLVYFEFENDSDASLRERRFKKWNRAWKVDLIEEMNPWSDLTDNWNLDSLNTRL
SK
<F598_RS0106700--Psychroflexus_tropicus_DSM_15496--58.62 (SEQ ID NO:365)
MKTGLGTHNYVYIILNKNKTLYTGVTNDLKSRLHFHKKPTAPSKAFTTKYKCYFLIYYEHFPDINQAI DREKQIKGYRRLKKNLINEFNPEWKFLNDEI
<OREMA_RS0109955--Orenia_marismortui_DSM_5156--70.93 (SEQ ID NO:366)
MRRSYAYMTNKNNTVLYIGVTNNIKRRVYEHKEKLVDFGFTKKYNC SKLVWAEFETMEBAIQAEKRMKKWKREYKENLIKELNPDWNDLYDYF
<BRAEL_RS0137595--Bradyrhizobium_elkanii_USDA_76--67.78 (SEQ ID NO:367)
MAGLVPALHVLQYPHMAGGYVYFLTNRPNGVLYVGVTSDLVRRIFEHRS GFVDGFTKRYGLKRLVYFEHFDDIRTAIQREHTIKHWPRAWKIRKI IAAPNWD
DLYDTIVQ
<BRAEL_RS0111260--Bradyrhizobium_elkanii_USDA_76--62.35 (SEQ ID NO:368)
MSTAGSYVYIILASRHHGTIYIGVTNDLPRLELHRSRGSKFVYQYQVFLRVHVEEFATPQEAIAREKQLKFWKRDWKIKLIEAENPDWNDLSGLI
<A37W_RS0103630--Wenxinia_marina_DSM_24838--64.04 (SEQ ID NO:369)
MPHYVYILASRPGGAIYIGRTNRLRQRIEAHRAGLSLHTSKYGIRTLVWFEAHDDFETSLRRERAIKRWHRAWKDQLITGANPDWKDVIHALP
<A390_RS0105005--Lamprocystis_purpurea_DSM_4197--78.26 (SEQ ID NO:370)
MERKPAVYIILASKRNGTLVVGVTNLPKR VWEHKLHDLVEGFTTRYHVHRLVWYAQLETMPDAIRLEKQMKGWQRAWKLRVIEEMNPAPWRDLYDDFAAGD
<G342_RS0116945--Kaistia_granuli_DSM_23481--77.42 (SEQ ID NO:371)
MQPCAYILASGYNGTLTYIGVTSALPQVRVWQHREEIFDGFTKRYGVKTLVYVEPHDSIGEAIRREKQLKHWNRAWKIRLIEKDNPDWRDLYPEIAI

FIG. 11L

<H145_RS0108940--Hymenobacter_aerophilus_DSM_13606--63.53 (SEQ ID NO:372)
MKTYYYVLLTNKNKNTVLYIGVTNDLQRRVGEHKAGRHPGFTTKYNLHLYLVYFESYDINAAIAAREKQKAGSRQKKEALINAENPEWDLDATEF
<H145_RS0116975--Hymenobacter_aerophilus_DSM_13606--54.02 (SEQ ID NO:373)
MNAGYVYIITANPTKTALYTGLTNDLHTRLRQHFNENRGTKSTFAGRYSCYQLLYFEHFQDVNQAIAREKEIKGWTREKKNMLITQLNPHWQPIDPTAWQDGAQ
<H149_RS0111370--Solimonas_variicoloris_DSM_15731--76.34 (SEQ ID NO:374)
MRQPCVYIMASRRNATLYVGVTSDDLVRVWQHRSGVVGFTQRYGCKLLVWYERHETMESAVQREKQIKGGSRAKRLALIEALNPQWRDLYEAIL
<H151_RS0107530--Woodsholea_maritima_DSM_17123--65.93 (SEQ ID NO:375)
MEERVIAVYMLASKPFGTLTYIGVTSLLQRIWQHKNNVYPGFAQRYDCKTLVWYETHCEITHAIAREKSLKSYRRQNKINLIETMNPWCDSLEDFMR
<H151_RS0113345--Woodsholea_maritima_DSM_17123--65.56 (SEQ ID NO:376)
MQAYVYILASKPHGTLTYIGSCADPALRIEQHRQGHGSQFVKYVSVYTLVYLEGCEETHDAITLERRMKRWRSSWKIDRIEKDNPWNDDLYPHLNQYLP
<F812_RS21990--Hyphomicrobium_zavarzinii_ATCC_27496--78.65 (SEQ ID NO:377)
MKQPVVYILASKPLGTLTYIGVTSDDLKRMAEHTQGLIQGFTKRYDVKRLVYVYETHDQMLAAIAAREKQIKEWRRRAWKIRLIEQMNPEWKDLFDKQSGETAFGDT
DTHAFAEPTSPGSGSRPASG
<G372_RS0112120--Thioalkalivibrio_thiocyanoxidans_Arh2--75.28 (SEQ ID NO:378)
MERHPAVYIITNRPHTGLTYIGVTSNLPARITQHRQDLVPGFTQRYRLHFLVWFEQHAQMDIAIREKQIKRWRHPWKIHLIERDNPWTQDLYEGLF
<A30Q_RS0106040--Methyloferula_stellata_AR4--72.53 (SEQ ID NO:379)
MSGYVYIMSNKQGTLYIGVTSLSRRVYEHREKLVGFTSRYGLHRLVYVYAFDDIRAAIQREKNLKHWPRAWKLALIHENPEWDDLYLRLE
<A30Q_RS0112130--Methyloferula_stellata_AR4--71.91 (SEQ ID NO:380)
MRQPCVYIMASQRNGLTYIGVTSANLSRRAFEHREGLVKGFTSAYDCKRFVWYVYDLMIDAIAREKQIKGGSRAKRLALIESFNPEWDDLYDSLA
<ADIS_RS01790--Lunatimonas_lonarensis--69.57 (SEQ ID NO:381)
MEKGGCIYFMTNTHNTLYIGVGVASNLIKRVYEHKTVGHKTSFTARYNLTKLVYVEIFHSHIEBAIDREKQLKAGSRKKIELVNDNFPEWKDLYEQICRDW
<ADIS_RS11925--Lunatimonas_lonarensis--70.65 (SEQ ID NO:382)
MEKGGCIYFMTNTHNTLYIGVGVASNLIKRVYEHKTVGHKTSFTARYNLTKLVYVEIFHSHIEBAIDREKQLKAGSRKKIELVKDFNPEWKDLYEQICRDW
<L338_RS03565--Euryhalocaulis_caribicus--77.53 (SEQ ID NO:383)
MHADAIEKAAWYIMANHRNGTLYSGSTVDLVRRTWEHREGFIPGFTTRYGCKMLVWFEVHPDLETALVREKRLKRWRRRAWKLKLIEMNPHWHDLWEDLSL
<L338_RS03475--Euryhalocaulis_caribicus--78.65 (SEQ ID NO:384)
MRAFYVYILASARNGLTYIGVTNDLIRRVWEHREGLGSKFCRKYGVSMLVFETFTVDAAINREKRLKRYLRKWKLELIERNNPQWRDLYGEVASGWLGPQI
<L338_RS10465--Euryhalocaulis_caribicus--74.44 (SEQ ID NO:385)
MHEKLFWVYLLANRPRGAIYLGVTSNLVKRVWEHRLLEIPGHTRKYKITQLVWFEQHASAEAAIGREKQMKKWKREWKVALIEGMNPTWKDLYEDALRNQPF
<L858_RS0106945--Sphingobium_bisphenolivorans--77.53 (SEQ ID NO:386)
MKREISPTVYILASRKNGTLYTGVSDDLKRLYEHRNGLIEGFTSDYGVKRLVWFEHFDAMDSAILREKRIKKWNRQWKIELIEAGNPERDDLALSFGEPLP
STRRPIERSGDFPRSR
<L858_RS0108910--Sphingobium_bisphenolivorans--64.71 (SEQ ID NO:387)
MERERQGGWYIMADRYRGTMYVGVTDLAARVHQHRSGAGSDFCARYGLARLVWAERGDMMSCIEQEKRLKRWRRDWFELIERGNPDWQDMFELLV
<L858_RS0107505--Sphingobium_bisphenolivorans--63.22 (SEQ ID NO:388)
MTNKPFGVLYIGITSSTLTVRVDHRRGVGSDFCRKYNLHRLVLAEEHGRIDEAIAAREKALKAWKRNWKIQLIEENNPQWRDLFTLLI
<M472_RS39050--Sphingobacterium_paucimobilis_HER1398--64.77 (SEQ ID NO:389)
MERGGVYIMLTNVTNRTVLYIGVTSDDLTVLVEHREKRYANSFTSKYNSCLCVYYESFPTIEBAIDREKQLKKNRSKKENLINKLNQWDLWQVVKW
<M472_RS35235--Sphingobacterium_paucimobilis_HER1398--64.77 (SEQ ID NO:390)
MERGGCVYIITNNTTFTYIGVTSIDQSRVYEHINHTHEHAFSKRYKLSKLVYKFPSTIEAINEEKRIKGSRLKKITLIEITLNPRWEDLYQKEVKHW
<SALLO_RS0109830--Salisaeta_longa_DSM_21114--60.67 (SEQ ID NO:391)
MARRTYVYIITNAHHTVLYTGMTNDLRRRVAEHRQEGSAFVRRYNATKLVYAEAYSEVRDAIRREKQLKAGTRAQKVACIQRONPSWRDLADL
<YWA_RS0106810--Flavobacterium_antarcticum_DSM_19726--66.28 (SEQ ID NO:392)
MKLWYIYIMTNKPNGLVLYIGVTNDLLEERVKEHKLKVYRNSFTAKYNCCKLVYFEWKKGDAAIREQQFKKWKRDWKIKLIVDMNPSWMDISLNWNFDKDLR
E
<G405_RS0112795--Oceanicaulis_alexandrii_DSM_11625--75.58 (SEQ ID NO:393)
MYFVYILMASQKNGTLYCGVTNDLRRVVEHREGLAPGFTESHGVKRLVWFESYDSVEHAIRREKRIKRYPRQWKINLIEAENPDWADLWNRLSL
<G405_RS0103245--Oceanicaulis_alexandrii_DSM_11625--65.56 (SEQ ID NO:394)
MTRAWYMLSNRPGVLYITGMCNADLRRRWEHRQSGGSSFADKYNCCKRLVWFAFSDVTAABEHERMKRWKRAWKVELIEQANPDWRDLYSEISG
<G410_RS0103340--Novosphingopyxis_baekryungensis_DSM_16222--61.18 (SEQ ID NO:395)
MRGGWYIMANKPFGTTYIGVTANLPARVHQHSGTSGSIHCRDHGLHHLVYAEYPTIAEAIAREKAMKKWKRRWKTDLISKANPDWSDLLPY
<G410_RS0107505--Novosphingopyxis_baekryungensis_DSM_16222--52.86 (SEQ ID NO:396)
MAFWTYMLHCHGGQYTGTDGLENRMAQHAGSIPGFTADRLPVKLVAQDFPTRYEALSMERRIKGWSRAKKKALIRGDWKEISRLAKSRGKDEASTGSAR
TDEGPVPVKR
<G513_RS23090--Nevskia_ramosa_DSM_11499--77.17 (SEQ ID NO:397)
MDKEPCIYILASRNGTLYIGVTSNLQORLWQHRDGAFTGFTSKHRTQRLVWFERCDDMLAAITREKQLKKWNRMWKRLRIEENPEWRDLAVDMGLFADDDG
TGSKTGFPARE
<G523_RS0103400--Ponticaulis_koreensis_DSM_19734--68.82 (SEQ ID NO:398)
MAFYVYLLASKKNGTLYCGSTDDIAFRVWVSHKDGRGAFTKYNVTRLVWYELHTTREAATREYQIKAWKRDWKIRLIEETNPHWNLDLYQLNM
<G523_RS0106300--Ponticaulis_koreensis_DSM_19734--73.26 (SEQ ID NO:399)
MSAYFYVYILASQRNGLTYIGVTNNLFRRVWEHRRNNEGVSTFRKHVHRLVWFEHGDINRAIYREKRIKRWKRDWKLRIETENPNWTDLYDERFVV
<G523_RS0102295--Ponticaulis_koreensis_DSM_19734--65.12 (SEQ ID NO:400)
MAKYDFIACYILANRKNGLTYIGVTSDDLQRIEQHKLKGSVFAAKYGCDDLWVYARFPEMAPALQRETIKGWPRQWKINLIEEHRNFWDDLSGQLSEYGW
<G536_RS0106410--Hymenobacter_norwichensis_DSM_15439--58.24 (SEQ ID NO:401)
MSYCVYLLTNPARTVLYTGVTNDLERRLYEHSQGLNAGKFTGRYQCNLLVYFELCPDAQQAIVRGKEIKGWSRSKKEALVTTLPNDWKFIDICTWSG
<G536_RS0112330--Hymenobacter_norwichensis_DSM_15439--57.47 (SEQ ID NO:402)
MYIYIITNPTKTVLYTGVTNDLRRRLYEHSSGRGNAGKFTGRYQTDLLVYFELCPDAMQAIAREKQLKGWTRSKKDALITAFNPNWKTIDLETWAG
<K355_RS0110460--Thalassospira_lucentensis_MCCC_1A00383_=_DSM_14000--78.49 (SEQ ID NO:403)
MQQPFVYILASRKNGTLYIGVTSDLRQRIWQHRTGAVPGFSHTYNVTQLVYEMHDTMADAILREKRLKKWRRRAWKIRLIEELNSDWHDLFDHINA
<G626_RS0100140--Treponema_bryantii_NK4A124--77.53 (SEQ ID NO:404)
MQEYFYVYILANNSHTLYVGMTNDLKRRLYEHQNLKVPGFTEKYNVNKLIVYVYKTTDVKAAIQREKNLKKWNRRAWKLELIEKNNPTFKDLSLECE

FIG. 11M

<H536_RS0104185--Aurantimonas_coralicida_DSM_14790--70.00 (SEQ ID NO:405)
MAGTVYIMASQKNGTLYIGVTSLSRRVYEHKSGLTRGFVKAHGVTRLVWYEDHVDIGSAIQREKSLKRWRQWKIDLIEAMNPEWNELRYGMGW
<H536_RS22340--Aurantimonas_coralicida_DSM_14790--60.00 (SEQ ID NO:406)
MSLSKDQPTTEGRTPEPPDRDQMPLEATVYILRCVDGSYYTGLTKQPVPEARLWEHNEGLYDGYTRKRCPPVVLVFTETTYDRMLDAIAREQQIKGWTRAKKEALIA
RRYHCLPTLAATRTGPDLLRRQPPATLD
<M527_RS0129685--Sphingobium_chinhatense_IP26--72.04 (SEQ ID NO:407)
MTRDFDPTVYILASHKNGTLYIGVTSNLVQRIYQHRSGLIEGFTKDYGVNRLVWFESGSGMEEAIRREKQIKKWNQWKIGLIGRENPDWRDLALDFGFEPLK
SRRIVKASGDWNRVNGFPFPPRE
<M527_RS51110--Sphingobium_chinhatense_IP26--73.86 (SEQ ID NO:408)
MKEPCVYILASQPYGTIYIGVTSNLLGRLHQHRTNGVPGFTGRYAVHNLVRFEMCDTMEQAIRREKQLKRWHRQWKINLIETENPRWVDLAVGLGLPSLDQHQ
RHDGC
<M527_RS57140--Sphingobium_chinhatense_IP26--71.91 (SEQ ID NO:409)
MSKPGYVYIMASRRNGTLYLGVTSDLVKRAWQHRNGLGSEFSAEYRCHFLVWYAYDDLQEARLRELRMKKWKRDWKLRLIETANPGWRDLFEDVAQ
<H527_RS56115--Sphingobium_chinhatense_IP26--75.90 (SEQ ID NO:410)
MANRRNGTLYLGVTSDLAARAYQHRNGLIDGFTKDHDCLLVFEAYEDLQEARHRELMMKKWKRAWKIELIERDNPQWRDLFETLF
<M527_RS32965--Sphingobium_chinhatense_IP26--63.64 (SEQ ID NO:411)
MERKEKGWYVYIMSDRYRGTIYVGVTSNLTSLRIHQHRSIGSGSDFCARYGLNRLVWAERADDITSCIEHEKRLKRWRREWKFDLIERANPDWQDLFHLA
<H562_RS0110835--Chryseobacterium_palustre_DSM_21579--63.75 (SEQ ID NO:412)
MIEFTEGIYDFYVYILTNRNRTVLYTGVNTDIRRLKEHREKKNPKSFTARYNVEFLIYYEHFGWIQLAIAAREKEIKDLRRELKLDLIKSLNPDCLKFLNWHFE
IPTEG
<H562_RS0109105--Chryseobacterium_palustre_DSM_21579--59.77 (SEQ ID NO:413)
MKTGLTHNYFIYILTNNKMTVLYTGVNTINLAERVWCWNCPEAISKHFTTKYKCFYLIYFEQFQDIETAI SREKQIKGWTRKKKEDLINLVNPTWRFLNDEI
<H562_RS0106570--Chryseobacterium_palustre_DSM_21579--55.88 (SEQ ID NO:414)
MDFVYVYILSKYKQYVGVTSNLTSLRHFHSHNSLSTKGFTLRYCLWEVIYVEFYNSKSEAMAREKFFKSGAGRDWINLNIAFSDE
<H562_RS0113095--Chryseobacterium_palustre_DSM_21579--56.96 (SEQ ID NO:415)
MKGYTYILECADGSYYTGSTTDLERRVQQHQGEGANHTSKNLPVKLVYEEYQRIDEAFFREKQIQGWSRRKKEALINGWDEELKKAECMNE SHYKNK
<H573_RS0116350--Spongobacter_marinus_DSM_19753--64.77 (SEQ ID NO:416)
MLEQLPAVYILANRQSGSLHSGVAANLRQHLHDYRQALSEIRKEVFELVYFELFEDLEQAYMREKQIQEYSHRRKTALIDSINPRWDDLYQTGEEA
<G507_RS0111745--Flavobacterium_gelidilacus_DSM_15343--76.47 (SEQ ID NO:417)
MYELYQKDVVYIILANKNSVLYVGVTTNNIQRRILEHKDRVNGFTTYRYNVDKLVYEVHGYINDAILREKRIKKWNKWKQELIEHNKEWIDLSLEW
<G507_RS0100530--Flavobacterium_gelidilacus_DSM_15343--68.89 (SEQ ID NO:418)
MKTGFVYIITNNKNTVLYIGVTSNLPQRIMQHKYKGSFSARYNLDKLVYFEQFPIMGDAIAREKQLKAGNRAAKEKLINSMPDWRDLYEDIKNYIP
<G507_RS0110685--Flavobacterium_gelidilacus_DSM_15343--58.06 (SEQ ID NO:419)
MIKGYVYIILNKNHSLVYVGTADKLYRIECHKEGTGAVFTTKYSATFLIYFEEYDNYRDAFLREKQLKNWHKDWKWNLAKILNPELKDLYLEL
<G507_RS0100950--Flavobacterium_gelidilacus_DSM_15343--57.32 (SEQ ID NO:420)
MQQSFYVYILKCSDDSYTYGVTSNLSQRFFQHQSGFYPCYTYKRRPVLELVYCEFTDINLAIEKEKQIKKWSRAKKEALINEDYDDLPLNLAKKKFL
<G509_RS0114890--Flavobacterium_subsaxonicum_WB_4.1-42_=_DSM_21790--70.53 (SEQ ID NO:421)
MKPGFYIIVANYTNTTLYVGVTSNLDHLRIKQHKNKQHPDSFSARYNMNKLIVYFESFQMIEDAITREKQIKGGSRAKKIQLIETINPEWNDLYEIDNYL
<G512_RS0109030--Nesiotobacter_exalbescent_DSM_16456--74.44 (SEQ ID NO:422)
MAYYVYMTASRRNGTLYTGVNTDLARRLYEHREKVAMSFTARYGVSQLVYEEYERIGDAIAREKKLKRWRKGWVELIEGFNPAWRDLYQELA
<G521_RS0108975--Pedobacter_oryzae_DSM_19973--71.11 (SEQ ID NO:423)
MYFVYMI TNNKNTVLYIGVTTNLERVYEHQKINKGFTSKYCNKLVWYEEYSTVKEAI SREKQLKNWQRNWNKNQELIEKKNPEWNDLSEKWF
<G522_RS0124155--Pleomorphomonas_oryzae_DSM_16300--67.37 (SEQ ID NO:424)
MNPQVYIVTNKPNRTVLYIGVTSNLDLGRISDRHNHVYRGFTDSYCNPTLRVWYEDNGDITEAIEREKQLKHWKREWKIQLIEMNPEWEDLFETIMS
<G530_RS0104125--Gramella_echinicola_DSM_19838--64.13 (SEQ ID NO:425)
MKKSYTYILSNKNRTVYIIGVTDGLNKRI TEHKNGKGSDFTKKYNLTDLLYFEEFTDINQAI SREKQLKNWRKEWKDLIKDFNPELKDLPFEI
<G530_RS0115920--Gramella_echinicola_DSM_19838--55.91 (SEQ ID NO:426)
MQDYFVYIVTNEKRSVLYIGITNNLNI RLDQHYRDSVHKSLSFAGKYNCYHLIYFEKFNNAEAI IREKTIKKWRREKKLNIGDFNPEWQFLERDIL
<G530_RS0111995--Gramella_echinicola_DSM_19838--61.36 (SEQ ID NO:427)
MLSNDQSQSYTYIILTNRNKTVLYIGVTTNNLQRRILIEHHQCARLAKRTFTAKYHCIFLIHYEKFVSWQDAIAREKQLKGWLRSKVGLINENNPEWKFLNDEFS
L
<G530_RS16535--Gramella_echinicola_DSM_19838--61.76 (SEQ ID NO:428)
MKKGYTYILSNKNKTVKYIIGVTDGLNKRI TEHKNGKGSDFTKKYNLIDLLYFEEFTDINQAI SREKQLKNWRKE
<G537_RS0130380--Inquilinus limosus_DSM_16000--73.40 (SEQ ID NO:429)
MQPAVYITANRPYGTLYTGVTSNLLRRIHQHREGTVPGFSSRYGCKILVWYETLPSMKEAIAREKQIKGGSREKKLALIQLVNPWEWRDLYPGLV
<G539_RS0121750--Kaistia_adipata_DSM_17808--80.65 (SEQ ID NO:430)
MQPCVYIMASARNGTLYIGVTSNLPQRVWQHREGFLGGFTSRYRIRTLVYVEGHD TMESAIRREKQIKMLWKRWSWKIELVERSNNPWL DLYGEIAS
<G539_RS0102075--Kaistia_adipata_DSM_17808--78.89 (SEQ ID NO:431)
MPRSLRPCVYIVASRRHGTLYIGVTSNIAARAQQHRDGLVEGFSKRYGKTIVVFFETHDTMEAAILREKQLKEWRRDWKIRLIEESNPDWSDLYSLTME
<G543_RS0104175--Legionella_moravica_DSM_19234--77.91 (SEQ ID NO:432)
MEEKHYCVYILASKAYGTLYTGITSNLIQRVYQHKGLSEGFTKRYHVHRLVYVEIHSVDYQAITREKRIKKWNQWKINLIEQNNPQWLDLIGLGF
<G543_RS0109760--Legionella_moravica_DSM_19234--68.82 (SEQ ID NO:433)
MQNQSYMYFVANAYNTVLYVGVTTNNLIRRIYEHKNKLIKGTQKYNVDRVLYVEIFDTIVLAIEREKQIKGWSRNKKNL VNSFNNSNWN DLYQSIL
<G543_RS19080--Legionella_moravica_DSM_19234--74.47 (SEQ ID NO:434)
MIHHGSKAYGTLYTGLTSNRIQRVYQHKKEELNEGFTKRYTVHRLVYVEIYSDV
<G551_RS0111000--Mesonia_mobilis_DSM_19841--63.22 (SEQ ID NO:435)
MPQRTYHNYTVYIILTNKPKGVLYIGVTTGGIDDRMERHIEGKASQFTSKYKLRVLYVEEFQYIEDAIAREKKLKNWHRQWKINLIEQENPEWKDLYKPLATDP
ETSSG
<G551_RS0102295--Mesonia_mobilis_DSM_19841--64.37 (SEQ ID NO:436)
MKTSYVYILTNYKRTTFYIGITSNLTAKIAEHQLGKGSFTAKYDLKFLIYEEFLDIQHAIKREKQLKSWKDWKIDLVKKKNPELKTIEFI

FIG. 11N

<G551_RS0103645--*Mesonía mobilis*_DSM_19841--55.95 (SEQ ID NO:437)
MKT LGTHNYVYIILTNRNKTVLYTGVTTNNLKDRLYFHRNPNSKNSFTKKYCNCFLLVYLEHFQNIQEAIKREKQIKGYSRTKKTETLINNLNPEWHFFNDQFE
<G563_RS0122690--*Runella zeae*_DSM_19591--67.06 (SEQ ID NO:438)
MKDHSYVYIILTNYSKTVLYVGVTTNNLVRRLEDEHKQGLSGFTAKYCNCFLLVYWEHHKYVYNAIEREKEIKGWRRSKKETLINEFNPWKFLNDEIQDQ
<G564_RS0100895--*Salinimicrobium xinjiangense*_DSM_19287--69.32 (SEQ ID NO:439)
MKLSYVYIILSNKKRSLTYIGVTSNLYNRVVEHKS GKGSVFTAKYHLMMLLYFEEFTSIDQAIAREKQQLKNWHKQWKWNLIKELNPELEDLFDKLRP
<G564_RS0112145--*Salinimicrobium xinjiangense*_DSM_19287--64.71 (SEQ ID NO:440)
MDFSYYTMMNSHNTVYIYGLTSLNLIKRVYQHKTKAYKGFTSKYNC DKLVFYQEFTNINEAIALEKKIKAGSRSKKEKMI FEQNPEWKDLS DGWLFII
<G578_RS25060--*Thalassobaculum salexigens*_DSM_19539--71.11 (SEQ ID NO:441)
MRDPIRPCVYILASRRNGTLYTGVGTGLVRRVALHRAATAGFCATYGVHRLVYFELHGRMEDAILRETRVKKRRRAWKLALIEGVNPGWEDLYQSLRTGDLNL
APFAAPPPIRGDRRDALGPGSPDQVRSRPG
<G578_RS0116490--*Thalassobaculum salexigens*_DSM_19539--63.33 (SEQ ID NO:442)
MQPTVYIILTNHRRGTLTYIGVTANPVSRIQLQHRREGSEFVRRYRLHRLVYVEQFARITDAILREKQQLKAWRRRAWKIELIEGANPEWTELWPTPGVEACDRGGP
<G578_RS0110095--*Thalassobaculum salexigens*_DSM_19539--76.67 (SEQ ID NO:443)
MRVWVYILASRRNGTLYVGLTTDLARRIREHREKRLAGFTSRHHVTRILVYVEEQERLDHALIRERRLLKRWRRDWKLDLIEQQNPDPWRGLSYDILGPPDSG
<G581_RS0106510--*Thermodesulfovibrio thiophilus*_DSM_17215--74.19 (SEQ ID NO:444)
MKEQYVYIMTNKYNTVLYVGVTTNDLVRRVYEHKKNLKVGF TAKYNI VKL VY YETHEDIKEAIKREKQIKGWLRAKKEELIKTMNPSWKDLYEKIV
<H148_RS0116580--*Salinimicrobium terrae*_DSM_17865--65.48 (SEQ ID NO:445)
MFKSYTYMMTNVHHTVYIGVTSNLYNRVYQHKTKAFKGFTSKYNCQKL VFFQEFSNVSEAI DFEKKIKAGNRAKKEKLI SEQNPEWKDLS DGWLFVV
<H344_RS0112990--*Arenibacter latericius*_DSM_15913--73.63 (SEQ ID NO:446)
MKPGYVYILANKNNTLYVGVTSNLYVRIQQHKTKYKKSFSARYNVDKLVYVEVQRIGEAIGREKQQLKAGNRARKVALIEQENPEWRDLTDCFG
<H510_RS0116910--*Pedobacter glucosidilyticus*_DSM_23534--66.28 (SEQ ID NO:447)
MQKGGVYIMTNKYNTVLYIGVTSNLYNRVYQHKTKAFKGFTSKYNCQKL VFFQEFSNVSEAI DFEKKIKAGNRAKKEKLI SEQNPEWKDLS DGWLFVV
<H510_RS0112375--*Pedobacter glucosidilyticus*_DSM_23534--63.95 (SEQ ID NO:448)
MQKGGVYIMTNKYNTVLYIGVTSNLYNRVYQHKTKAFKGFTSKYNCQKL VFFQEFSNVSEAI DFEKKIKAGNRAKKEKLI SEQNPEWKDLS DGWLFVV
<H510_RS0116970--*Pedobacter glucosidilyticus*_DSM_23534--65.22 (SEQ ID NO:449)
MKTYSIYLLTNKNNTVYIGVTSNLYNRVYQHKTKAFKGFTSKYNCQKL VFFQEFSNVSEAI DFEKKIKAGNRAKKEKLI SEQNPEWKDLS DGWLFVV
<H593_RS0105550--*Psychroserpens burtonensis*_DSM_12212--67.02 (SEQ ID NO:450)
MKHYVYIITNKSGLVYI GETGGLKKRIYQHKKNVHPTTFSAIYNLDMLIYYEYNSEVEAKLEKQMKKWNREWKLELIEKANPDWKYLYEEI
<H593_RS0108015--*Psychroserpens burtonensis*_DSM_12212--62.65 (SEQ ID NO:451)
MTNKKEGVLYIGCTNHIIERVKEHKFIKYPRRAFTAKYKRNLLVYFEEFEDGEEASMEIRIQFKWKRDWKIKLIEDMNP SWMDLSMNWNLGLNKFRD
<H593_RS0101015--*Psychroserpens burtonensis*_DSM_12212--62.07 (SEQ ID NO:452)
MKTIGTHNYIYIILTNKNKTVLYTGVTTNNLVERLYPHKNPLPFSKAFTTRYKCYLVYVEHYFGIDDAIKREKQIKGISRFKKEKMI SEFNSQWDFLNDNI
<H598_RS0112350--*Ferrimonas kyonanensis*_DSM_18153--77.53 (SEQ ID NO:453)
MKQPCVYIITNDRNGTLYIGVTSNLMKRIWEHKNQFIDGFSHSHQLTQLVYFETHQSMYQAITREKQQLKH WKREWKIQLIEQENPYWQDLWETLM
<H599_RS0105105--*Flavobacterium daejeonense*_DSM_17708--68.42 (SEQ ID NO:454)
MKPGFYIITNKYQTVVYVGVTSNLPQRILEHKNKKYPTSF SARYD VNVLYVEEQFQWIEDAIAREKQIKAGSRAAKNKLQISINPTWKDLDFDEIKDIMTE
<H599_RS0103600--*Flavobacterium daejeonense*_DSM_17708--65.06 (SEQ ID NO:455)
MSNKHKGVIYIGVTDNLIERVKEHKLKVYSNSFTAKYNC DKLYFEEWENGFDATKREKQQLKKGKRDWKVKLIEDMNP T WLDISLNNWFNFNKIRE
<H599_RS0113245--*Flavobacterium daejeonense*_DSM_17708--58.14 (SEQ ID NO:456)
MKFYFYIILCKKDESLYVGITSDLERRIMEHNAGKYPEAYTHSRRPVELVYQDFTPNQAITFEKRIKKWSRAKKQCLIDENFDKQLSLAECRNATHHKYKP
DEE
<H601_RS0111815--*Flavobacterium filum*_DSM_17961--64.77 (SEQ ID NO:457)
MENPQTGYHTYVYIITNKYRSTYYVGVTTNNLRERLQHKENSTDTTTFASKYNLEFLVFEKYTWIEBAIAREKELKKWRREKKLDLIRNSNPKFEFLNHL
FS
<H602_RS0116505--*Flavobacterium tegetincola*_DSM_22377--65.26 (SEQ ID NO:458)
MKPGFYIITNSNNTLYIGVTSNLPQRILEHKEKLYPNFSF SARYN LNKLYVDFDQFMIGDAIAREKQQLKGGSR LAKFTLIEKTNPLWNDFDEIKDIMI
<H602_RS0111145--*Flavobacterium tegetincola*_DSM_22377--63.75 (SEQ ID NO:459)
MIEHTEGITYFFVYIILTNKNVLYTGVTTNDLKRRLKEHKS HLNPSSTSRYNVYLVYFEKFTWIQQAIEREKDKIKNWPRSKDNLILEFNP KHSF
<H603_RS0111715--*Flavobacterium frigidarium*_DSM_17623--60.47 (SEQ ID NO:460)
MIEYTEGIYNFYVYIITNKNKKVLYTGVTTNNLRRLGEHKNKNLPNSFTARYNVEFLLYVEHFGWIQLAIAAREKEIKLMNRDLKEELINTINPEWNFLNELF
<K358_RS0117260--*Vibrio litoralis*_DSM_17657--70.79 (SEQ ID NO:461)
MMNQPCVYIILSNNKQSLYIGVTSQKLQVRVQHKNKQVNGFSYRYNTINLVYIEIHNVMESAI SREKQQLKH WKRQWKS NLI TEHNKEWVDLYNQ
<K384_RS07755--*Thermonema rossianum*_DSM_10300--59.09 (SEQ ID NO:462)
MKPLGCHNYFVYIITNKS KKVLYVGTNDLRRRLYEHKLDAYTTKSHTGKYNAYFLVYWERFEYVEHAIAREKEIKKWRRAKKEALINSFNPEWRFLNDEI
<METSZ_RS26780--*Methylocaldum szegediense*_O-12--79.27 (SEQ ID NO:463)
MPNKQSESAQTSPISEADLNVI PAGTPESSVQDGRRRNSVDVNGTLYVGVTSKLVRRRAFQHRNGLVEGFTNKYRVHMLVWFELHERMETAILREKLIK KWPR
AKIGLIEASNSEWRDLWTNVIGIS
<SPHPH_RS0113935--*Sphingomonas phyllosphaerae*_FA2--78.41 (SEQ ID NO:464)
MERQPCVYIILASRPHGTLYIGVTSNLLARLIQHRHGT FEGFTSRYDVTRLVWYETADTMAAAI LREKQQLKKWNRDWKRLRIEETNPQWLDLAPSIGLDGTHDG
<SPHPH_RS0111505--*Sphingomonas phyllosphaerae*_FA2--75.58 (SEQ ID NO:465)
MKSPGVYVYILMASQRNGTLYLGVTSNLYVARAFQHRGTGAIDGFSKRYGCTLLVWFVEVHDDIQDARQREARLKKRKRDKWLELIERNTNPQWKDLYDTLF
<SPHPH_RS0117435--*Sphingomonas phyllosphaerae*_FA2--72.09 (SEQ ID NO:466)
MASRRNGTLYAGVTSNLAQRVWQHREGAVAGFTQRHGCKTLVWFESADTMDAAI TREKQIKAGSRAKIALIEAMNPGWHDLYEQIAHG
<YWC_RS0115565--*Massilia alkalitolerans*_DSM_17462--76.40 (SEQ ID NO:467)
MGKHSFVYIILASGRYGTLYIGVTSDLIRRVQHNERFVEGFTKTHVHQLVYVEIHTDIIAAI TREKQIKKWRNRWKNLIQSTNP WHDLYPTIAA
<A37S_RS0110425--*Olivibacter sitiensis*_DSM_17696--64.84 (SEQ ID NO:468)
MKRGGVYIITNKS HRVLYTGVTSDDL SRIKKHRDKEFPNSF SARYN VYKLVYRFFSTIQEAIABEKKIKGGNRKQKLRLIEGLNPDWNLWDEIVEW
<C515_RS20545--*Chitinilyticum litopenaei*_DSM_21440--81.91 (SEQ ID NO:469)
MAPAVYILASKRNGTLYIGVTSNLLKRIWEHKNDVTEGFTRRYGIHLLVYVEYQHGSM DAAI TREKQQLKKWNRDWKQLIESVNPQWRDLWDDICGIP TGCPPA
RA

FIG. 110

<F569_RS0112355--Hellea_balneolensis_DSM_19091--69.66 (SEQ ID NO:470)
MIRPCVYILASGIHGTLYIGVTS DLPGRLEQHENKHSFGFTARYNVRLVCFEEYEDMATAIRREKTMKKWLRAWKIELIEKANPHWHRLDAESGEFIYDE
<F569_RS18245--Hellea_balneolensis_DSM_19091--68.24 (SEQ ID NO:471)
MFYTYIVASKRNGTLYTGQTNDICRRAEHEHLGI FPGFQGEHHCIYLVWYEEHPTRQAAFI RERRIKKWERAWKLSMIEAVNPLWLDLGLCPMWPPLDKRVYP
EIWDNALATAVPR
<F569_RS0109015--Hellea_balneolensis_DSM_19091--70.93 (SEQ ID NO:472)
MNDDAYAYVYILASHKNGTLYIGVTRDLGTRLEQHKSKADAKSFTARYNVTRLVWFEEFDLIVEAIAYEKRLKKYKRQWKINLIEEHNPDWRELDGSGEFL
<F569_RS0113650--Hellea_balneolensis_DSM_19091--67.42 (SEQ ID NO:473)
MSYGYVYLLCNRKNGTLYLGVTADLYGRLQQHRSFHNPKSFTAQHNATRLVWFERNDLVAEIAQEKTMDKWPRQWKKNVIERDNPHWDELVLVDYHD
<F569_RS19285--Hellea_balneolensis_DSM_19091--64.71 (SEQ ID NO:474)
MFYTYIITNRYKGKLYTGHTEDLVQRMEQHRYKIFPGYAAKHHCQYLLLYEEHESRDAAFKRERRIKKWKRDWKLNLIESSNPEWLDIMACPTWPLPEGDLFD
DLRAKAYHEARIPIGKPPHICRTKLDMSLPLPAQG
<F569_RS17675--Hellea_balneolensis_DSM_19091--58.82 (SEQ ID NO:475)
MYELTFWTYIITNAYKGTLYTGHTDDIGTRMEQHITGAFESFSKRYKLKHLIYWQEFETRDEAFKKERQIKWKRWKLNLIEDMNPLWVDIHSSPMWPLPGM
GMHPELYKRCLEHCLDPPDLRRDERRYFD
<G391_RS0104915--Brevundimonas_aveniformis_DSM_17977--73.26 (SEQ ID NO:476)
MPTHRTFIATYIMASQRNGTIYTGMTANLYARVWKHKQHVWQGTDDYACERLVWFEEQYAWVVDIAIAREKRIKKWLRWKLALIEKDNPDWDLAADWYPVND
PDWQPPPEAD
<G394_RS0102520--Desulfomicrobium_escambiense_DSM_10707--79.78 (SEQ ID NO:477)
MDASKSYFVYILASKNGTLYTGVTSSLKARIWQHKSNNVEGFTQKYQVHKLVYFEMHSSPNDAILREKQIKKWNDRWKIRLIEENNPWKDLWEEISSM
<G407_RS0117230--Salinarimonas_rosea_DSM_21201--76.47 (SEQ ID NO:478)
MTGVALVTQVAYSGESRHEIAPVYVYMLASRMHGTYLGVTRDLARRLHEHKLGVYKGFTERYGVHRLVWYEGHERIDEAISREKRLKRWRDWKIRLVEDM
NPLWSDLTPLPSDW
<G441_RS0116185--Aestuariibacter_salexigens_DSM_15300--70.97 (SEQ ID NO:479)
MKQPAVYLLKNHHTSVLYVGVTSNLIQRIYQHKTEQAEGFSQRYHCHELVYVELWNDMYSAICREKCIKKWRRRAWKERLINSINPQWRDLYDELL
<G457_RS0102410--Algoriphagus_marincola_DSM_16067--58.24 (SEQ ID NO:480)
MEAFMYILKCGNGLYVVGSTKNLEKRLQHLLEGANFTRKHLPLVELVYETFERIDEAFKREKQVQNWSSRAKKKALINRDFEALKNLSKGKT
<G466_RS0120960--Arenibacter_certesii_DSM_19833--71.43 (SEQ ID NO:481)
MKPGFYVILANKNNTLYVGVTSNLLRRIQQHKEKHKKSFSAARYNVDKLVYEEAFQDIEGAI GREKQLKAGNRARKVALIEKENPGWDDLSEQATDCFG
<G466_RS0108270--Arenibacter_certesii_DSM_19833--61.36 (SEQ ID NO:482)
MFFNHFYIITNEYRTLYVGVTVNNIQRRLSQHYFDSQNSKKSFAKYNCIYLVYVEGFENPKTAISREKEIKKWRREKENSLSISFNPKWETLNNQIF
<G473_RS0114185--Niveispirillum_irakense_DSM_11586--78.65 (SEQ ID NO:483)
MTGYVYILASQKGLTLYIGVTSDIGRIWEHRDGLVDGFTKQYDVKMLVYVELFDRIDDAIFRDKQLKKWNRQWKINLIEASNPDWDLSTLL
<G489_RS0106460--Cucumbacter_marinus_DSM_18995--74.19 (SEQ ID NO:484)
MSAWVYIMANKARGTLYVGVTSDLVRRVWQHREGEMEGFTRQHGKLVLYVEEHPTMPLAIQREKNLKRWRQWKLDLIEGMNSKWEDLWPRIST
<G489_RS0115375--Cucumbacter_marinus_DSM_18995--73.03 (SEQ ID NO:485)
MSGWVTIMANKPRGTLYVGVTSDLVRRVWQHREGVEGFTRQHGKLVLYVEELPTMPLAIQREKNLKRWRQWKLDLIEGMNPKWEDLWPRISG
<G489_RS20730--Cucumbacter_marinus_DSM_18995--77.63 (SEQ ID NO:486)
MCAWVYIMANKPRGTLYVGVTDGLVRRVWQHREGEMEGFTRQHGKLVLYVEEHPTMPLAIQRVKNLKRWRQWKLDLIEA
<G489_RS0115975--Cucumbacter_marinus_DSM_18995--59.76 (SEQ ID NO:487)
MQGFERASVYVLRCADGTFYTGITRKPVESRLWEHNQGLGDGYSARRPVELVFVEAFDRIDEAIAERERQIKGRSRRKKLALIAQDYERLPALSSRVAAKRKL
<G494_RS0121070--Desulfobulbus_mediterraneus_DSM_13871--77.78 (SEQ ID NO:488)
MKQPCFYVILASKRNGTLYTGVTSDLLKRVWEHKNQADGFTQKYGIHLLVYEEQCEDMVSAIAREKQIKKWHRSWKIQRIEELNPQWHDLYETLL
<H526_RS0112950--Aquimarina_latercula_DSM_2041--78.02 (SEQ ID NO:489)
MLKPEHQYSVYILSNKNGTIYIGITNNIERRIFEHKNKLVEGFSKKYNLILYFEVFRNVNEAILREKRLKNWNRQWKINLIEKENPNWTDLAKDWYP
<H527_RS0100215--Aquimarina_muelleri_DSM_19832--78.02 (SEQ ID NO:490)
MLKPKHQYAVYILTNNKNGTLYIGMTNNLERRIFEHKNKSVEGFTKKYGVDKLVYFEVYQVNDIAITREKRLKSWNRQWKIDLIEEENSNNWKDLAYDWYD
<H527_RS0108615--Aquimarina_muelleri_DSM_19832--63.86 (SEQ ID NO:491)
MTNKPdGVIYIGCTDDIEARVLEHKNKIYKNSFTAKYNCdKLVYFEYKSGLEASKREKQMKKWKRWKINLILDMNPSWMDISLNWSENVNSVFKIKRSLPT
QE
<H559_RS19150--Chitinilyticum_aquatile_DSM_21506--82.98 (SEQ ID NO:492)
MSFAVYILASKRNGTLYIGVTSNLLKRIWEHKNdVTEGFTRRYNIHLLAYYEQHASMDAAITREKQLKKWNRDWKQLIESVNPQWRDLWDDICIGIPTGCPPA
RA
<H560_RS0108980--Chryseobacterium_daeguense_DSM_19388--65.56 (SEQ ID NO:493)
MKAGFYIMTNKNNTLYTGVTSDLPKRIQQHKESHYQQSFTTRYNVNKLIVWESFQEI GDAIFREKQIKAGSRQNKLNLIISMNPEWDLADDIKDIMNPF
<L288_RS31415--Sphingobium_quisquiliarum_P25--60.87 (SEQ ID NO:494)
MKEPCVYILASQPCGALYIGVTSNLPGRLAQHRAIGPSTERYGVHRLVRVEACDAMERAILREKQLKNWRRQGKINLVERENPAWVDLAVGLGLPLDRSV
GRHGC
<RLDS_RS36625--Sphingobium_lactosutens_DS20--77.27 (SEQ ID NO:495)
MRDRHPCVYILASKRNGTLYIGVTSDLIRRMHEHRTGSVAGFTREYGIKRLMWFEQATMEEAILREKRLKKWNRRAWKITLIEDLNPHWDDLAVGLGFDRLPD
<RLDS_RS41835--Sphingobium_lactosutens_DS20--74.70 (SEQ ID NO:496)
MASGRNGTLYLGVTSDLAGRTYQHRNGLVDGFTKEHGCTMLVWFEEAYDDLQEARRELQMKKWKRAWKIELIERDNPRWRDLFETLF
<RLDS_RS36775--Sphingobium_lactosutens_DS20--65.88 (SEQ ID NO:497)
MREGAKGGWYIMADRYRTIYVGVTDIVARVDHKKRIGSGDFCARYGLGLRVWMEFGEDIRACIEQEKRLLKRWRRRAWKFDLIERDNPLWRDISETAV
<RLDS_RS53755--Sphingobium_lactosutens_DS20--63.10 (SEQ ID NO:498)
MARGGYTIMTNKAHGVLYIGVTANLAARVLQHRNGTGSTFCRYGLTCLVLVEPHDDMTLAIARKKALKVWKREWKIRLIEGSNPDWDLSDLIA
<RLDS_RS53375--Sphingobium_lactosutens_DS20--54.29 (SEQ ID NO:499)
MPFWTYMLHCNGGALYVGHTDDLERRIASHEHGAVPGFTANRLSVKLIWSEYFATRIEALEMERRVKWSRAKKLALVRGDWDRISSLARGKGKGRASTSSAR
TEAGK
<M529_RS34155--Sphingobium_ummariense_RL-3--72.09 (SEQ ID NO:500)
MACDGRFAFVYIMACRRNGTIYIGSTVDLPKRAWEHRNSVVPGFTRKYGCRLLVWYEVDRGWEEAARLRERQMKWKRAWKLREIEADPEWADLYDRIARP

FIG. 11P

<M529_RS0125050--Sphingobium_ummariense_RL-3--62.22 (SEQ ID NO:501)
MRGGWYIMNRRAYGVLYIGVTTSLPHRTMQHRDETGSEFCRRYNLRLVYAEHETIIEAIAAREKAMKAWKQWKIELIETINPQWQDLFDLHLE
<KQS_RS07955--Flavobacterium_indicum_GPTSA100-9=_DSM_17447--67.44 (SEQ ID NO:502)
MAKRSYHNYWVYIILTNKPKGTLYIGVTTGGIDDRMERHLKGEGSKFTAKYKLMVLVYEEFYIHDIAIAREKQLKNWHRQWKINLIEADNPNWENLWKGDVEV
<MBEBAB_RS06445--Brevundimonas_abyssalis_TAR-001--73.03 (SEQ ID NO:503)
MAFYTTYTLASRRNGTLYTGSTDDLIARVSQHKAKTFQGFATAQYGVVDQLVWFEVHDSRDGAFRRERQIKKWKRRWKVQMI EENNPDPWRDLYDEVRLGDLKDASD
WVPACAGMSGVKNDG
<MBEBAB_RS01450--Brevundimonas_abyssalis_TAR-001--71.76 (SEQ ID NO:504)
MAFFTYIVASGRNGTIYVGSTDDLVTNRVNHKNKTFGGFTAKYGVDRVLVWYEVFDSREAAFKRERRIKKWNRVWKLDLIEKANPAWDDLHDLRLGLDLDVDAKD
WAPPEED
<MBEBAB_RS08630--Brevundimonas_abyssalis_TAR-001--65.12 (SEQ ID NO:505)
MGRERPLIAAYIVTNKPFGTLYTGVTSNLYQRVHQHRTGAFAGFTKEHGLKRLVWFEPFELMTSAIRREKALKRYLRAWKINLIERENPDWDDLSEWDRPPV
WKHGPTDE
<M666_RS02070--Cellulophaga_baltica_18--70.93 (SEQ ID NO:506)
MKTGFVYILANKNNTTLVVGITSNLLQRIERHKTKFYPKSFSARYNTNKLVYEAQFDIGAAIAREKQLKAGNRAKKIALIEKENPDWEDLIKQATDCFGKSA
SQ
<BRAGE_RS0102480--Bradyrhizobium_genosp._SA-4_str._CB756--78.72 (SEQ ID NO:507)
MAYVYVILASRRDGAIVGITNDLVRVVEHRIKAVRGFTAKYNI TQLVWFEVYDDPISAI SREKELKKWKRAWKIQLIEKDNPDWNDLYESICK
<BRAGE_RS0107855--Bradyrhizobium_genosp._SA-4_str._CB756--75.56 (SEQ ID NO:508)
MAYVYLASKKHGTLTYIGVTSDLVRVVEHKTAVPGFTTKYGVDKLVLFEIYDDAVTAIAREKELKKWRRDWKTRLIEEQNPNDWDLPGITN
<BRAGE_RS0135980--Bradyrhizobium_genosp._SA-4_str._CB756--77.53 (SEQ ID NO:509)
MSARHYVYVILASRIGGTLTYIGVTNDIIRRVTOHKSCLIESFTEKYDVARLVYFEQFDDPENAI KREKRLKKWNRAWKIRLIEQHNPNDWDPYPIAGPP
<BRAGE_RS0139850--Bradyrhizobium_genosp._SA-4_str._CB756--76.40 (SEQ ID NO:510)
MTYVWYIPASKPGGTLTYIGVTNNLVRVVEHCEGLADGFTKRYGVKTLVYFEAHD TIAAALQREKNMKHWSREW KIDLIVKSNPAWQDLYDEIAR
<ADHAQ_RS0118995--Adhaeribacter_aquaticus_DSM_16391--57.47 (SEQ ID NO:511)
MFNSNYFVYITTNPNRTTFYIGVTNDLPRRLTEHFEQSGKPFQTFAGKYFCYNLLYFEPHTTPEHAIAREKQLKGWSRSKKEALIKSQNPFLKFLNHLVMEG
<M595_RS11975--Lyngbya_aestuarii_BL_J--72.09 (SEQ ID NO:512)
MINYYVYRMNTNKSCTLYTGVTNDRIRRVVEHKNKLIPGFTQKYNI TLLVYEEPTDISAAIAREKQIKGWLKQKKIDLIESVNPNDWDL SLKWYD
<RGRA_RS0102485--Rickettsia_gravesii_BWI-1--75.56 (SEQ ID NO:513)
MVYIVTNKTNGLTYIGVTFNLIKRIYQHKNDTIKGFSSKYNCKILVYFEIHHQTMNSAINREKQIKAGLRDKKIKLIEQTNINWDDLYNEIIL
<N508_RS05735--Mucispirillum_schaedleri_ASF457--69.66 (SEQ ID NO:514)
MEDKFKSII LSHNQATASSRNQYIYIITNYHNTVLYVGVTSNLKR IYEHKNKVVKGFTNTYCNKLVYFEIFSDIQEAIKREKYIKGKRSFKLDLIQSLNP
NWL DLYDMITQ
<N177_RS06435--Lutibaculum_baratangense_AMV1--70.59 (SEQ ID NO:515)
MEQKFWAYIILTNKRNGALYVGMTSDLVRRMDQHRSGVVPGFTRRYGVKRLVHAEQFSDVYEA LYRERQLKAWRRAWKMR LIEANPQWRDLAEELFI
<MGMO_RS17305--Methyloglobulus_morosus_KoM1--79.57 (SEQ ID NO:516)
MTKQPCVYLLASQONGTLYTGVTSDLLKRVYQHRGTGMVEGFTKRYGVHMLVWYIEHEKMESPIQREKQLKNWPRIAKIRLIEQQNPNDWDLWHALVV
<P855_RS0100440--Thalassobacillus_devorans_MSP14--56.96 (SEQ ID NO:517)
MADNHFVYIILCKDDTLYTGVTNDLEQRLAKHESGKGAKYTRGRGPFELVFYERFPPTKIAAMQREFQIKQWSRSKKEALIALKNEG V
<FSS13T_RS07850--Flavobacterium_saliperosum_S13--66.32 (SEQ ID NO:518)
MKPGFIYIITNKNNTTLTYIGVTSNLPKRI TEHIEKRFNDSFSARYNKLVLVYEAQFMIGDAIAREKQLKAGSRAKKMALINGMNPDPWKDLYEDIKDVMMWF
<RMONA_RS08830--Rickettsia_monacensis--76.14 (SEQ ID NO:519)
MLASGRNGTLYIGVTSNIIKRIYEHKHKI IKGFTSSKYNINKLVYFEQYDNIIA AINREKLLKEYKRQWKLN LIESQNP EWHDLYEKII L
<RMONA_RS08175--Rickettsia_monacensis--74.03 (SEQ ID NO:520)
MGVTSNIIKRIYEHKHKI IKGFTSSKYNINKLVYFEQYDNIIA AINREKLLKEYKRQWKLN LIESQNP EWHDLYEKII L
<RMONA_RS07995--Rickettsia_monacensis--73.75 (SEQ ID NO:521)
MIYFLINVTVLYIGVTSDI IKRIVQHKQKIRKGFTTKYDIDKVVYFEQFNDVNLA I GREKRLKEYKRQWKLDLINSNPSWHDLYEDFLK
<RMONA_RS00360--Rickettsia_monacensis--84.78 (SEQ ID NO:522)
MILCSDRNGTLYIGITNNILRRAYEHKQKIKGFTTKYNVIKLVYTEEFADIKLL
<RMONA_RS08435--Rickettsia_monacensis--86.96 (SEQ ID NO:523)
MLCSNKHGTLTYIGITNNIIRRAYEHKQKIKGFTSKYNTKLVYTEEFTKI
<W911_RS05140--Hyphomicrobium_nitrativorans_NL23--74.73 (SEQ ID NO:524)
MKRQPCVYILAGRRNGTLYVEVTSDLHKRMAEHTQGLVPGFTKKHGVKHLVYEMHDL L TEAIRREKQLKEWQRIWKIRLIETMNL EWQDLFDATSGEIAFGPA
DLQRETGEQ
<HMPREF2087_RS07520--Helicobacter_canis_NCTC_12740--76.34 (SEQ ID NO:525)
MKQAYVYILFSQPNGLTYGVTSDLTRRVYEHKQKIVKGFTQKYGVDKLGYEYVESMEMAITREKQLKAGNRKKKLELITQSNPQWKDLYDDL F
<CUL_RS17420--[Clostridium]_ultunense_DSM_10521--67.11 (SEQ ID NO:526)
MYVYVMMANLNKNKVLTYEVTNNLEHSVYEHKNKLIKGTAKYNNVNLVYFEYTS DVNLAIDREKQIKGWTRQKKNEVIQGIYLE
<LOA_RS05105--Legionella_oakridgensis_ATCC_33761=_DSM_21215--79.55 (SEQ ID NO:527)
MEEKQYVYVILASKKYGTLYTGITSNLVQRIWQHKEGLAEGFTKKYNVHHLVYIEIHS DVEYEAITREKRIKKWNRQWKINLIEQNNPQWLDLGVGLF
<LOA_RS07015--Legionella_oakridgensis_ATCC_33761=_DSM_21215--79.55 (SEQ ID NO:528)
MKHGRKQYVYVILASKKYGTLYTGITSNLVQRIWQHKEGLAEGFTKKYNVHHLVYIEIHS DVEYEAITREKRIKKWSRQWKINLIEQNNPQWLDLIGLGF
<LOA_RS05625--Legionella_oakridgensis_ATCC_33761=_DSM_21215--76.40 (SEQ ID NO:529)
MRERQPAVYIMANKRNGTIYTGVTSNLIKRVYEHKYADLDGFTKKYGCYKLVYELIEDMASA I SREKQLKGGSRKKKLALIEKINPLWEDLYETLT
<LOA_RS01360--Legionella_oakridgensis_ATCC_33761=_DSM_21215--66.29 (SEQ ID NO:530)
MRTESC VYLLTNKHNNVLYTGVTNDLIRRVYEHKNKLVAGFTQKYNVDRVLVYFEVCSGIVMAIEREKQIKGWSRKKKHD LINALNPEWNDLYPSLL
<LOA_RS04300--Legionella_oakridgensis_ATCC_33761=_DSM_21215--66.29 (SEQ ID NO:531)
MHTESC VYLLTNKHNNVLYTGVTNDLIRRVYEHKNKLVAGFTQKYNVDRVLVYFEVCSGIVMAIEREKQIKGWSRKKKHD LINALNPEWNDLYPSLL
<LOA_RS11430--Legionella_oakridgensis_ATCC_33761=_DSM_21215--64.52 (SEQ ID NO:532)
MHTESC VYLLTNKHNNVLYTGVTNDLIRRVYEHKNKLVAGFTQKYNVDRVLVYFEVCSGIVMAIEREKQIKGWSRKKKQD LINALNPEWNDLYPSLL

FIG. 11Q

<LOA_RS02885--Legionella_oakridgensis_ATCC_33761 = DSM_21215--64.52 (SEQ ID NO:533)
 MRTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS01265--Legionella_oakridgensis_ATCC_33761 = DSM_21215--65.17 (SEQ ID NO:534)
 MHTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS11600--Legionella_oakridgensis_ATCC_33761 = DSM_21215--67.44 (SEQ ID NO:535)
 MHTESVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS03070--Legionella_oakridgensis_ATCC_33761 = DSM_21215--67.44 (SEQ ID NO:536)
 MRTESVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS03750--Legionella_oakridgensis_ATCC_33761 = DSM_21215--65.17 (SEQ ID NO:537)
 MHTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS06100--Legionella_oakridgensis_ATCC_33761 = DSM_21215--65.17 (SEQ ID NO:538)
 MHTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMTIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS12830--Legionella_oakridgensis_ATCC_33761 = DSM_21215--65.56 (SEQ ID NO:539)
 MHTESVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS05835--Legionella_oakridgensis_ATCC_33761 = DSM_21215--64.04 (SEQ ID NO:540)
 MHTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMTIEREKQIKGWSRKKKHDLINALNPEWNDLYPSLL
 <LOA_RS00030--Legionella_oakridgensis_ATCC_33761 = DSM_21215--69.23 (SEQ ID NO:541)
 MRTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKHDLINALNPE
 <LOA_RS06340--Legionella_oakridgensis_ATCC_33761 = DSM_21215--71.43 (SEQ ID NO:542)
 MHTESCVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWS
 <LOA_RS08680--Legionella_oakridgensis_ATCC_33761 = DSM_21215--70.00 (SEQ ID NO:543)
 MHTESVYLLTNKHNNVLYTGVTHDLIRRVYEHKKNLVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWS
 <MESLO2_RS011420--Mesorhizobium_loti_R88b--67.42 (SEQ ID NO:544)
 MTGYVYMTSSKKRGTIYIGVTNDLGRMPPEHKSQGSQTSRYGVQRLVWVYEEYFDITDAIQREKSLKRWLRLQWKIELIEKANPDWFELFRGTGW
 <DESACE_RS00185--Desulfurella_acetivorans_A63--64.04 (SEQ ID NO:545)
 MKRGFVYIMTNKYNVLYTGVTSNLAKRVYQHKNKTIIEGFTKQYNIVKLVIYAEAFDDIAEAIKKEKQIKGKTRQKKIDLINAFNPDWVLDYDDIAT
 <M890_RS0111525--Olleya_marilimosa_CAMO30--63.53 (SEQ ID NO:546)
 MKTSYTYIITNKYRTFFYGVGTADLNKRI TEHQNGTGSVFTKKYNLKLIIYIEIFTDINQAILREKQIKNWKKEWKNLNIKEQNPTLKTLM
 <DZ60_RS0118215--Aquimarina_pacifica--65.91 (SEQ ID NO:547)
 MKKSHVYFMTNKNNTVIYIGVTSNLLKRVYQHKMKVYKGFASKYCNCDLVIYFEFDDINQAITREKQLKAGNRKRKEDLINANNLEWNDLSNGWLFNFT
 <Z053_RS0106170--Aquimarina_megaterium_XH134--79.79 (SEQ ID NO:548)
 MLKPKHQYAVYIITNKNGTLYIGMTNNLERRVFEHKNKLIDGFTKKYELDMLVYFEIYKYVNDATREKRLKSWNRQWKINLIEEDNPNWKLACDWYD
 <Z053_RS0107815--Aquimarina_megaterium_XH134--73.40 (SEQ ID NO:549)
 MLKPRHQYAVYIITNKNGTLYIGMTNNLERRVFEHRSKYVDGFTKKYDLKLIIYFGIYQYVNDATREKRLKAWKQWKINLIEEDNKDWNLDASDWYD
 <Y931_RS0126355--Aphanizomenon_flos-aquae_NIES-81--69.74 (SEQ ID NO:550)
 MTNNHNKVLYTGVTHDLIRRVYEHKEKLEIEGFTKKYNVTKLVYIEIFIDATSAIAREKQIKAGSRQKKIDLINSIN
 <METLA_RS0106410--Methylosarcina_lacus_LW14--78.26 (SEQ ID NO:551)
 MNIKQPCVYLLASSRNGTLYVGVTSDLIGRIYQHRNDLAEFGTKRFVHQVLVWFELHETMESAI RREKQIKKWNRAKLNLI ESANPAWDDLWFSLSL
 <N825_RS018460--Skermanella_stibiirensistens_SB22--71.91 (SEQ ID NO:552)
 MSGGWYIMTNRPDGTYIGVTADLARRAQHREPVGEGFTKRYGLRMLIWAWEHDSILLAIQREKTINHWRSSWKVALIVKNNPTWANLYGSLNI
 <A058_RS0115525--Flavobacterium_succinicans_LMG_10402--57.30 (SEQ ID NO:553)
 MEKGFTYIITNKNYTVLYVGATKNLKNRIDCHRNGTGAVFTKKYNATILIIYFEAFDDLADAFKREKQIKNWKKEWKNLEKLSNPELKDLYLNL
 <A058_RS0109955--Flavobacterium_succinicans_LMG_10402--60.53 (SEQ ID NO:554)
 MKLSYVYILKCADGSYYTGVTSNLTQRVFRHNTAYYPDCYTANRRPLELVFYCEFTDITVAIEKEKQIKKWSRAKKEALINGDPEALVNLAKKKFSI
 <JCM21793_RS19135--Nitratireductor_aquibiodomus_NL21 = JCM_21793--69.66 (SEQ ID NO:555)
 MGGYVYILASRKNGTLYTGVTSLEQRI CEHREGITGGFTKKYGVHMLVWYRDYSIDRAIEDEKRIKRWRRQWKLELIERINPDWDDLSSALNH
 <Q320_RS0115285--Deefgea_rivuli_DSM_18356--81.72 (SEQ ID NO:556)
 MKQPCVYLLCNKNGTLYVGVTSNLVQRIWQHKEGLAEFGSKYSKLTIIWYEQHATMESAIAREKAIKEWKRAWKLTIEKMNPDRDLYPDIV
 <Q327_RS0104035--Campylobacter_upsaliensis_DSM_5365--71.79 (SEQ ID NO:557)
 MKPKRCPVYILCNKNGTLYVGVTSDLQSRIVYQHKNKTFSGFTVKYNVSRILAYFEFFETMEEAILREKQLKGGSRACKIALILSVLA
 <T345_RS0104120--Legionella_fairfieldensis_ATCC_49588--82.35 (SEQ ID NO:558)
 MEEKHYVYVILASKYGTLYTGVTSNLIQRIYQHKELVEGFTKKYNHHLVYIEIHSVDYEAITREKRIKKWNRQWKINLIEQNPNQWLDLSIGLF
 <T345_RS0103285--Legionella_fairfieldensis_ATCC_49588--73.33 (SEQ ID NO:559)
 MKSFYVYIMASQRNGTLYIGTSDLIKRVWEHKNKVI PGFTARYNVHRLVYIEIHETIYIEAAREKCFKNWRRQWKINLIEKLNPEWCDLYEEICR
 <T345_RS0109650--Legionella_fairfieldensis_ATCC_49588--58.14 (SEQ ID NO:560)
 MEWVYILKCRDQSYTGHNDLLEHRIVQHQNKLIPGCTYSTKLPKILVYSESFPSREELATERQLKKWSRRKKEALMQGDWKNLSDYAKRSMKY
 <T367_RS0117135--Crocinitomix_catalasitica_ATCC_23190--60.00 (SEQ ID NO:561)
 MEFVYVYILYSQKADKIYIGFTSDLISRIYSHNNLQKGHTQKYRFPWSVYVYEFYNTKIEAMHREKNLKGQGRAWIHDHLNKGAGFLSA
 <FH5T_RS16090--Draconibacterium_orientale--66.29 (SEQ ID NO:562)
 MKGFFVYIMSNERTAFYIGVTSDLKTRIAADHKNGIGSAFTSKYKLTDLLYIEDFPDIYQAI DREKQLKNWHRQWKINLIKSKNPQMNDLYWEIR
 <FH5T_RS13305--Draconibacterium_orientale--65.88 (SEQ ID NO:563)
 MKDHNHYVYIVTNKNTVLYIGVTNDLQRRVFEHENGLIPGFSKKYCNLYLVFYEHFQNI BYAIBREKEIKKWRREKKNLVKNFNP SWEFLNQKIYASL
 <U859_RS0103770--Pedobacter_jeongneungensis--65.91 (SEQ ID NO:564)
 MERGGTYIITNFTHTTLYTGVTSDIMFRVKEHKDKKEYSKSFSAKYNTDKLVYFCSYDSIEAIVEEKRIKGGSRKKIKLIQSMNPEWKDLWEIVKDL
 <AL00_RS10660--Sphingobium_lucknowense_F2--72.04 (SEQ ID NO:565)
 MTRDFDPTVYILASHKNGTLYIGVTSNLVQRIYQHRSGLIEGFTKDYGVNRLVWFVSESGSMEEAIRREKQIKKWNRRQWKIGLIGRENPDWRDLALDFGFEPLK
 SRRIVKASGDWNRVNGFPPPRE
 <AL00_RS07155--Sphingobium_lucknowense_F2--73.86 (SEQ ID NO:566)
 MKEPCVYILASQPYGTIYIGVTSNLLGRLHQHRTNGVPGFTGRYAVHNLVRFEMCDTMEQAI RREKQLKRWHRQWKINLIETENPRWVDLAVGLGLPSLDQHQ
 RHDGC

FIG. 11R

<AL00_RS06495--Sphingobium_lucknowense_F2--71.91 (SEQ ID NO:567)
MSKPGVYVIMASRRNGTLYLGVTSDLVKRAWQHRNGLGSEFSAEYRCHFLVWYEAADDLQEARLRELRMKKWKRDWKLRLIETANPGWRDLFEDVAQ
<AL00_RS00010--Sphingobium_lucknowense_F2--75.90 (SEQ ID NO:568)
MANRRNGTLYLGVTSDLAARAYQHRNGLIDGFTKDHDCKLLVWFEAYEDLQEARHRELRMKKWKRAWKIELIERDNPQWRDLFETLF
<AL00_RS12785--Sphingobium_lucknowense_F2--63.64 (SEQ ID NO:569)
MERKEKGGWYIMSDRYRGTIYVGVTSNLTSRIHQHRSIGSGSDFCARYGLNRLVWAERADDITSCIEHEKRLKRWREWKFDLIERANPDWQDLFHLLA
<DA69_RS12885--Brevundimonas_naejangsanensis--71.26 (SEQ ID NO:570)
MNERRPFIATYILASRPLGVLYVGMTSNLYERIRQHREDGIEGFAQRYGCKTLVWYEAHGFTVEAIRREKSLKRWNRWTKLELIEKTNPEWQDLYPSLFLGLAP
DPRVKPEDEGGREGGVEAFLKRLDRL
<DA69_RS13565--Brevundimonas_naejangsanensis--68.13 (SEQ ID NO:571)
MATHRTFIATYIMASGRNGTLYVGSTSNLESRVWKHRTAFPGFTDAYDCTRLVWLEQHALMLDAIRRERRRIKKWLRDWKLSLIEKDNPDWRDLAADWFPDPN
WTPPPGVD
<HY30_RS03520--Hyphomonas_chukchiensis--69.41 (SEQ ID NO:572)
MAFFTYIVTNKPRGSLYCGQTDDINRAWEHRNRVLKGYSHQYNCTALVWFETPFESRENALARERQIKNWKRWDWKIELIEANPDWRDLADTLI
<HY36_RS04415--Hyphomonas_atlantica--71.11 (SEQ ID NO:573)
MANEDRIIAVYMMANRQKQALYTGVTSHLVQRVWQHREGIFGGFTQKYGCKTLVWYELHDEISTAIAREKSLKSYRRKNKINLITERNPDKDLWFETIG
<HY36_RS06200--Hyphomonas_atlantica--71.11 (SEQ ID NO:574)
MAFYTYIVCNRRNGTLYTGHTDDLRRCEWHQNRIRPGFASKYHCARLVWFEEHDTRESAFTRERRRIKDWRRQWKNLIEKANPDWRDLYLELLDWTPTPPFH
NTAHPGEGRGFAPHRPD
<HY36_RS05620--Hyphomonas_atlantica--67.44 (SEQ ID NO:575)
MANRKGALYAGHTDNIARVLAHREGGAHTSKYNIKRLVWMEAHDTREAAKTREYQIKKWKRDWKIRLIEETNPDKDLWFSLNQ
<HY29_RS06305--Hyphomonas_beringensis--70.59 (SEQ ID NO:576)
MAFYVYIMTNKPRGTLTYCGQTDDINRAWEHRNRILKGYTYKNCMLVWYAGFDTREAAIDRERQIKNWTRAWKIDLVEKTNPEWRDLATALI
<HJA_RS11330--Hyphomonas_jannaschiana_VP2--72.09 (SEQ ID NO:577)
MESYWVYILASKKDGAIYVGITSDLLGRVWQHRQGTGGHTKKYHIRRLVWFEDEFNPDAIAFEKKLKRWRRAWKDELIAKENPEWVDLYDQLTKP
<HOC_RS07940--Hyphomonas_oceanitis_SCH89--75.56 (SEQ ID NO:578)
MAFYVYMLASQRNGTLYIGQTDLIRRVWQHNRKVLRGFTSKYDCHMLVWFADFPDRDSARSRRERQMKAWKRKWKLDLIEANNPRWADLYEDLVSYVPVSPAN
AGAQSCKSELRS
<HOC_RS05255--Hyphomonas_oceanitis_SCH89--72.22 (SEQ ID NO:579)
MANEDRIIAVYMMANRKNNGTLYIGVTSHLVQRTWQHREGTAAGFTKTHGCKRLVWYEVHEWIEATIAREKSLKSYRRQKNINLIEKMNPDWKDLWFETIG
<HOC_RS08800--Hyphomonas_oceanitis_SCH89--69.89 (SEQ ID NO:580)
MAFYVYIMASQKIGTLYTGHTDDIAYRTWNHREGGAFTRQHSVKRLVWYEVHESRESAKTREYQIKKWKRAWKIRLIEEANPDWRDLYFTLNQ
<HOC_RS04175--Hyphomonas_oceanitis_SCH89--70.59 (SEQ ID NO:581)
MAFFTYIVTNKPRGSLYCGQTDDINRAWEHRNRVLKGYSHQYNCTALVWFETPFESRENALARERQIKNWKRWDWKIELIEEANPDWRDLADTLI
<HOC_RS12240--Hyphomonas_oceanitis_SCH89--60.47 (SEQ ID NO:582)
MAFYVYVMNPTTGALYTGHTDDISRAHEHRHGLPGWSRDHGCKKLWVYADFESRDAKQREWQMKKWKRDWKLQIAEDNPDGVDLYERLNV
<BP92_RS16515--Flavobacterium_sasangense_DSM_21067--69.77 (SEQ ID NO:583)
MNRIGVYIMTNKNTTLYIGVTNNLCRRRIYEHKNHLIKNAFTAKYNLEYCYIFEFOYFDLAIQREKELKKWNRQKKEALINKNPEWKEIATEKGIVQNTT
PYTQQVKELMELLNQTKQIPDDRDK
<BP27_RS0104050--Marinospirillum_insulare_DSM_21763--77.91 (SEQ ID NO:584)
MKSSYVYIMTNKNGTLYIGVTSNLIKRVYEHKNGLCDGFTKKYALKTLVYVEISEDITSASIREKQLKAGNRRRKIELIESMNINWLDLYPGLLG
<WGMM_RS02305--Wolbachia_endosymbiont_of_Glossina_morsitans_morsitans--73.56 (SEQ ID NO:585)
MKLSYIYILTNKANSTLYIGVTANLIKRIWEHKNKIGSGFTSRYNVNKLIVYEEFEDISLAIAREKRLKEWPQWKINLINQKNPSWIDLYLALR
<FEM21_RS14495--Flavobacterium_seoulense--62.50 (SEQ ID NO:586)
MFEYKEGYHSYVYIITNRYRSTFYIGMTNNLRIRLQQHENIVENNKTFAKYGLEFLVYEEKFTWVQLAIAREKELKGWNRKTKITLIRSFNPEFEFLNDS
FKSRDVIPL
<REISMN_RS08895--Rickettsia_buchneri--76.14 (SEQ ID NO:587)
MLASGRNGTLYIGVTSNIIKRIYEHKHKIKGFSSKYNNKLVYFEQYDNIIAANREKLLKEYKRQWKNLIESQNPEWHDLYEKIIL
<REISMN_RS09080--Rickettsia_buchneri--73.75 (SEQ ID NO:588)
MLCLYTLINVTLYIGVTSDIKRIWQHKQKIRKGFTTKYDIDKLVYFEQFNDVNLAIGREKRLKEYKRKWKLDLINSSNPSWHDLYEDFLK
<BU07_RS0107065--Legionella_wadsworthii_DSM_21896 = ATCC_33877--71.59 (SEQ ID NO:589)
MEEKQYVYLLASEKGTLYVGVTNRNLEQRFYQHCNDLVEGFTKKYKVNHLVYEEAHSVDVNEAIRREKQIKKWNQWKNINLIEEKNPHWINLAIGLF
<BU07_RS0111690--Legionella_wadsworthii_DSM_21896 = ATCC_33877--74.44 (SEQ ID NO:590)
MKSFYVYIMASRPNGTLYTGSSNLIKRVWEHKNVPIPGFTAKYKVYTLVYFEMHETEEAARREKRFKNWCRQWKNINLIEERNPEWRDLYEEICH
<T342_RS0116915--Acidiphilium_angustum_ATCC_35903--74.19 (SEQ ID NO:591)
MRVERQFVYMMANRRNGTIYVGVTSDLVRRVAQHRGEGSGFTARYGCRMLVWFEMWGSMEGAIAREKRLKTGSRRGKLALIEAGNPEWEDLGFGLV
<T363_RS0105045--Agathobaculum_desmolans_ATCC_43058--73.33 (SEQ ID NO:592)
MQYVYVYVLTNNKHHVLYVGVTNNLIRRMQEHKSHIITGTARYQVDRLLYEQTTDQVYAAIAREKQIKGWTRKKKLELIRQMNPTWRDLYPDIIQ
<EO99_RS0115285--Rhizobium_marinum--73.12 (SEQ ID NO:593)
MTGYVYIVTNHKGHTLYIGVTSDIARIYEHREGLTPGFTSRYGLTQLVWYEEHWDIRDALQREKTMKQWYRKWKIALIEERNPDWRDLYLELW
<AW17_RS21430--Pedobacter_borealis_DSM_19626--70.45 (SEQ ID NO:594)
MERGGCVYILTNNHTHTVLYIGVTSDFRIVEHREKKYIDSFTAKYNSCKLVYEQFDSIEEATAKEKQLKNWKRIWKINLINQNNPNWEDLFKTELE
<AW17_RS05305--Pedobacter_borealis_DSM_19626--64.77 (SEQ ID NO:595)
MERGGCVYIMTNFNTVYIYIGVTSDLHSRVIDHKCKRYPTSFATKYNCKLVYFMFYPGIEEATAEKKLKKWNRDWKIKLISDHNPQWDLDFNEDL
<AW17_RS07590--Pedobacter_borealis_DSM_19626--68.18 (SEQ ID NO:596)
MERGGTVYIITNFHTTLYVGVTSDLMFRIKEHRDKEYPKSFAKYNMDKLVYFCSYDSIEEATAEKKRIKGSRRKKIKLIESMNPWKDLWEIVKEW
<AW17_RS10805--Pedobacter_borealis_DSM_19626--59.52 (SEQ ID NO:597)
MANSFNTVFTGTGVTSDLQGRVWQHKNNEFPKSFSTSKYKCYKLVYVYTFPNIEEAIDEEKRIKGGSRQKIDLVASANPNWVDLYDELDS
<AW17_RS22980--Pedobacter_borealis_DSM_19626--57.61 (SEQ ID NO:598)
MDHNYFVYIITNNKNTVVYTGVTNDLEARLRQHIENKKNFAFTYKYNCHYVYFERFEHIEYAIEREKEIKGWTRAKKNALIGRKNKNWDFLNDMIVEE

FIG. 11S

<AT45_RS09135---Chryseobacterium_hispalense_DSM_25574--61.63 (SEQ ID NO:599)
MFELNEGIYKFYVYILTNTKTRTVLYTGTGNLHQRLFFQHVHKVNDPSTSKYNVCFIYYESFGWIQQAIEREKEIKLLKRVKRLRLIRSTNPNLEFLNHLFE
<EC86_RS0110795--Polaromonas_glacialis--77.78 (SEQ ID NO:600)
MPYFVYMLASKPKGTLYIGVTNHLVRRIRHEHKSACDGGTARYDVHHLVWFEBITESIEAAIQREKQLKNWKRWDWKIELIEKENPLWDDLYSKIL
<U741_RS0103240--Polycyclovorans_algicola_TG408--77.42 (SEQ ID NO:601)
MNYVYVLLASERNGTLYVGVTTNNLIRRVWEHKNDVVDGFTKRYSVHQLVWFEEADNPAAILREKQIKKWNRSWKLLIEEKNPQWIDLNSIAG
<BS69_RS13050--Sphingomonas_astaxanthinifaciens_DSM_22298--71.79 (SEQ ID NO:602)
MSAHAEVVERQPCVYILATAPRGTLTYIGVTSDIAGRLWQHRNGVTGGFVSRYKVMRLVRFEMCGDMIAAIAAREKQLKRWHRQWKVNLI
<BS69_RS0111705--Sphingomonas_astaxanthinifaciens_DSM_22298--58.67 (SEQ ID NO:603)
MHFWVYMLRCSGSGSYVVGQTDDLEKRLAEHQDGLHEGYTRRRRPVRLVWSESTGRREEALAFERRIKGWTRAKKEALIIIGDWETIKALSEPPHERASTSLSTN
EPCKMIFNPRSS
<BS19_RS10055--Vibrio_pacinii_DSM_19139--73.03 (SEQ ID NO:604)
MDKQPCVYILSSANRNVLYIGVTSQQLQRIWQHRTSISDSFTSKYNVNCIFYEIHQTMEAAIYREKQLKSWRRNWKDELIKSINPYWRDLTYDI
<BP26_RS0122855--Nevskia_solis_DSM_19509--77.53 (SEQ ID NO:605)
MEKHFFVYILASTRNGILYVGVTSNLPQRIWQHKNDEVGFTKRYVHTLVWFEEQHESESATREKQVKKWNRMWKIREIELLNPGWRDLTYDDITA
<BP26_RS0110400--Nevskia_solis_DSM_19509--72.22 (SEQ ID NO:606)
MGKKHRYFYVYILTNSRVLYIGVTNDLIQRLYQHRNKLVEGFTKRYNITELVYFEETSNNVDALAREKQLKGWLRRAKKIDLI EASNPEWRDLGSELGCPDPSL
GSG
<EH31_RS15385--Erythrobacter_longus--70.79 (SEQ ID NO:607)
MKRVIEPCTYLMASGFRGTIYTGVTSDLMARIAQHRDGTGDTGFTKRYTAHLVHFEMFGTMDQAITREKQIKNWRQWKINLIEANPRWRDLACELGFAPLS
GRSAGDGS
<EH31_RS05680--Erythrobacter_longus--67.39 (SEQ ID NO:608)
MAKHPCVYILASGRGYIYIGVTSIAARYHQHQKGTFFGFAKRRNCKRLVHIEMFATMDEAITREKQLKNWHRDWKINLVEGDNPEWRDLGLEMLR
<IG63_RS0103140--Marmoricola_aequoreus--57.14 (SEQ ID NO:609)
MAWTYILLECDGSLVVGSTTDLMARLHQHQQLGAIRTRRLPVRLVWSAEFETIPEAFAFEKQVRKWGRAKRLALIRGDFDALPDLASRSHHGNNRRRGEYEG
CAE
<GV67_RS12800--Pseudorhizobium_pelagicum--72.73 (SEQ ID NO:610)
MRQGYVYILASKRNGTLTYTGVTSDLPRLYEHQNDLTPGFTSRYGVKTLVWFIEHDLVVDIAIRREKTIKKWPRAWKLNLIESVNPWDDISPWLL
<BN863_RS13850--Formosa_agariphila_KMM_3901--65.88 (SEQ ID NO:611)
MHVYVYILTNNKHTVLYIGRTKQLKVRLLIQHKNNSIKFTTGRYNVYKLVYFETTKYINNAIKRERQLKKNREWKINLINGLNPDWDLSEYI
<HY57_RS00850--Dyella_japonica_A8--79.57 (SEQ ID NO:612)
MRIRQPCVYILASGRNGTLTYTGVTSNLPARIWQHRNDVVEGFSHRQYVHMLVWYEVHELMEISAI SREKAIKKWNRAWKLGIEKTNPCWCDLYPDLP
<HY57_RS17255--Dyella_japonica_A8--79.57 (SEQ ID NO:613)
MDRRQPCVYILASGHNGTLTYVGVTSDLVKRVWQHRNEVIESFTSQHHVHNLVWYELHETMESAILREKRLKDWKRAWKRLIEGENPKWADLYPTLL
<RG540_RS04970--Neorhizobium_galegae_bv._orientalis_str._HAMBI_540--73.03 (SEQ ID NO:614)
MKGYVYILASRRNGTLTYIGVTDLPGRLYEHQNELTPGFTSRYGVKTLVWFEEYDILLVDIAITREKTMKKWPRKWLNLIEERNPDWEDIASHLSL
<IL45_RS06275--Nonlabens_ulvanivorans--69.05 (SEQ ID NO:615)
MSSKNRILYVGATDDIKRRTFEHKTCKYKTSFSSKYNVDWLVYEVVKGVTWEEAHKREKQLKKWNRAWKMLRIQEGNQDWLDLSQGWNLDIYKD
<IE88_RS0105880--Parvularcula_oceani--68.13 (SEQ ID NO:616)
MARFDYIAYVYMLTNQKRGTLTYVGVTSDLPLQRLAQHRSKGSTVGTGKYLGLTRLIWFERHDMRTVARQRERRLKEWQRLWKIEIEAGNPDWEDLGGDLL
<IE88_RS0104135--Parvularcula_oceani--71.91 (SEQ ID NO:617)
MTYVYVYILANKRGTLYTGVTNDLSRRIAHEHREGGATFPAQHGCTKLWVFAFESVDAAI THEKRVKKWRRRAWKEEBIEKTNPDWRDLFHDMMGGWTG
<IA57_RS10175--Mangrovimonas_yunxianensis--72.73 (SEQ ID NO:618)
MKNGYQFYVYILSNKNGTLTYGMTNDLERRMFEHKNKLRAGTGKYEELDKLVYFEEQFYVNDIAIKREKQLKNWNRQWKIDLIEKDNGKWNDAKDWIY
<IA57_RS04955--Mangrovimonas_yunxianensis--65.88 (SEQ ID NO:619)
MRKSYVYLLTNKNNTVIYIGVTSNLLKRMVYEHKAKIRKGFTYRNCDKLVYEEFSDINQAIAREKQLKSGNRKRKEALINKENPEWNLSDGWLFYFD
<IA57_RS09510--Mangrovimonas_yunxianensis--63.95 (SEQ ID NO:620)
MVKRAYHNYWVYILTNPNGTRYIGVTGGIDDRMERHVEGKSSFLAKYKLKHLVYEEFQYVHEAIAIKREKQLKNWHRQWKINLIEKVNPEWQNLWKP
<IA57_RS09445--Mangrovimonas_yunxianensis--64.47 (SEQ ID NO:621)
MKRTYHNYWVYIITNKPRGVLYIGVTGGIDDRMERHIKGEAAFAKYNLTRLVYEEFQYVEDAIKREKQLKNWHRGWKINLIE
<CC94_RS0109630--Methylomicrobium_agile--82.80 (SEQ ID NO:622)
MEKQPCVYILASGRNGTLTYVGVTSDLIGRVYQHRQKMIQGFTARYGVDQLVWYETHEQMESAIMREKQIKKWSRVFKRLIELDNPAWRDLWPDLVSSVAAR
<CC94_RS0103045--Methylomicrobium_agile--78.26 (SEQ ID NO:623)
MNKQPCVYMLASQCNGTLTYIGVTSDLIRRVYQHRNEQAEFGTKRYGIHDLVWYEVHEKMESAILREKQLKKWPRIAKIRLIEQSNFIWHDWLWELASPSLASG
FRQSMPE
<EP25_RS0117295--Methylomarinus_vadii--80.43 (SEQ ID NO:624)
MNKQPCVYILASRNGTLTYIGVTSNLIQRVYQHRTGMIIEGFTKRYGVHNLVWYESHEMMESAIQREKQLKKWARTAKIRLIEQQNPEWQDLWQDLT
<JE46_RS10640--Flavobacterium_chungangense_LMG_26729--67.09 (SEQ ID NO:625)
MLIYEGSHTYIYIITNKSKTVPYTGVTNNLKI RL IQHKENITDGNKSFASKYKVEFLLYYEKFTWIQEAISREKEIKGWSRNKKIELIKTLNPDLDLNNLF
D
<JE46_RS02660--Flavobacterium_chungangense_LMG_26729--64.56 (SEQ ID NO:626)
MQLQQGYHTYIYIILTNSKTVPYTGVTNNLKI RL NQHKENNVNEKNTFTAKYNVIFLIYYEKFTWIQEAIAAREKEIKDWRREKKIELIKTLNPDLDLNDLF
A
<IW15_RS04465--Chryseobacterium_solis--57.95 (SEQ ID NO:627)
MKIAAAHNYVYILTNNKRTVLYTGVTNDLKSRLYWHNLPEIESKIHFTTKYKCFYLVYEHSDIEQSI IREKQIKGYSRMKKEKLINNFNPSWDFLNETIE
KISHYISLRK
<M2A_RS12705--Tepidicaulis_marinus--73.33 (SEQ ID NO:628)
MTDYVYVYILANRRRGTLTYVGVTDLLRRVVEHKSEAADGFTKCYGVKQLVWFEEATPDIEEAIRAERIKRWRDYKIAMIERNPDWLDLDFDELVR
<M2A_RS05840--Tepidicaulis_marinus--71.59 (SEQ ID NO:629)
MSGEARLPCVYMLTNRPGGTLYAGVTRDVIRRVWEHKTDVADGFSKKYQLHFTLVYAEFHAIMPDAILREKQIKGWRRRAWKVALIEKDNQWRDIYEELAG

FIG. 11T

<SP6_RS17925--*Sphingomonas paucimobilis*_NBRC_13935--75.58 (SEQ ID NO:630)
MTKAGVYVIMASRLNGTIYIGVTSDDLAKRAYQHRNALIDGFTTKYGCFLVWYEAHDELDGARLRERQLKKWNRGWKRLRIEKNPDWDLYESLF
<JP75_RS14745--*Devosia riboflavina*--77.91 (SEQ ID NO:631)
MGKTYFVYIILASRYGALYIGVTNDLVGRVYTHREELVGRQTSRYHIHSLVYFEVFEDIDEAILREKRLKKWRRAWKIELIEQGNPEWTDLYPELAHW
<JP75_RS03630--*Devosia riboflavina*--76.67 (SEQ ID NO:632)
MYFVYIMANRKHGALYTGVTNDLLRRTSEHREGLPGFTKTNGCKTLVWFEPHDDIEFAITREKRIKKWHRDWKERLIEEQNPWRDLWWEIIGPQSAENTP
Y
<JP75_RS21685--*Devosia riboflavina*--68.09 (SEQ ID NO:633)
MGGYVYILADFRGRGTIYIGVTNDLVRRVYEHREGAAAGYTKRNKIKRLVYFEEHABIEAAIQRETSLSKRWYRQWKEALIEESNPDRDLWDEIVK
<DM09_RS07795--*Ghiorsea bivora*--80.43 (SEQ ID NO:634)
MIKQPCVYIILASKKHGRLYVGVTSNLVHRVWQHKNNDYIEGVTKKYNVHILVYAAQYETMVSATITREKQMKKWNRRNWKIRLIEENNPNWDDLWNSIH
<MB29_RS06810--*Rickettsia tamurae*--83.87 (SEQ ID NO:635)
MKQPTVYIILANKRNGTIYLGVTSNLIKRIYEHKNNNTNGFSKYSCKMLVYFIEHETMDSAITREKQIKSGSRAKKLNLI EQMNVNWKDLYDEII
<MB29_RS06165--*Rickettsia tamurae*--78.49 (SEQ ID NO:636)
MKQPIVYIVTNKKNGILYTGVTSNLIKRIYQHKNDTIKGFSSKYNCKILVYFIEHQTMSAINREKQIKAGSRDKKIKLIEQTNINWDDLYNEIIL
<MB29_RS01960--*Rickettsia tamurae*--76.34 (SEQ ID NO:637)
MKQNYVYIILASRRNGTIYIGVTSNIKRTYEHKHKIIGFSSKYNINKLVYFEQYDDIIAAISREKLLKEYKRQWKNLIESQNPWEHDLYEKIIL
<FPF3_RS01280--*Flavobacterium psychrophilum*_FPF3--65.91 (SEQ ID NO:638)
MNNTMIEYNESHYIYVYIITNKYRSTYYIGVTNNLSIRLQKHENIEKGIKTFASKYNIEFLVYIEKFTWIQLAIAREKELKGRREKKDLIKSFNNNFEF
LNDRFVKNNAIPVSE
<BN1847_RS08520--*Erythrobacter vulgaris*--73.40 (SEQ ID NO:639)
MRYFQPTVYIVASQRNGTIYVGVTSNLPQRAFQHREGAADGFSKRHGCKLLVWYEHETMESAILREKQIKGGSRAKKLALIEALNPQWLDYFEEVCT
<BN1847_RS03880--*Erythrobacter vulgaris*--63.33 (SEQ ID NO:640)
MGEKNRGGWYIMADRYRGTMVGVTAIPLRIQQHRKGTGSDFCAYGLKKLVWAEFAENIEACIAEQEKRIKRWQREWKFELIERANPDWDDLFDTLV
<BN1847_RS07280--*Erythrobacter vulgaris*--60.00 (SEQ ID NO:641)
MPFHAYLLRCNDGSYYTGHTDNLDLRIAQHQTGAIGGYTAQRLPVRLIWSSETFETREEALSARQIKGWSRAKKQALIDSNWELVSELAKNRQR
<BN1094_RS08180--*Legionella massiliensis*--65.59 (SEQ ID NO:642)
MHTESVYLLTNKHNNVYTGVTNDLIRRIYEHKKNKLVGFTQKYNVDRLVYIEICSCIVVAIEREKQIKKWSRKKKQALINVLNPQWDDLWLSLC
<LG45_RS05980--*Flavobacterium aquatile*_LMG_4008_=ATCC_11947--68.60 (SEQ ID NO:643)
MPNRVYHNYWYIVTNKPNGTLYIGVTGGIDDRMERHSSNGSIFTSKYKLFKLVYFEFQYIDDAIGREKQLKNWHRQWKINLIEESNPWNEDLWKLKKE
<LG45_RS01295--*Flavobacterium aquatile*_LMG_4008_=ATCC_11947--66.32 (SEQ ID NO:644)
MKPGFIYIVTNKYNTLYIGVTSNLAKRILEHIEKRYENSFSARYNLNKLIVYEQYQMI GDAIAREKQSKAGSRATKLNLIKFNPLWNDLYADIEDIMTI
<LG45_RS10230--*Flavobacterium aquatile*_LMG_4008_=ATCC_11947--67.05 (SEQ ID NO:645)
MIEYNESYTFYVYIITNKHRSYYIGMTNNLGVRLQKHENIEKGIKTFASRYNLEFLVYIEKFSWQLAIDREKELKSWNRDRKLELIRSFNENFEFLNDR
FVKTE
<IX84_RS02455--*Phaeodactylibacter xiamenensis*--69.77 (SEQ ID NO:646)
MSRAGFVYIMTNKNTTLYIGVTNDLCRRYIEHKNHLVENAFTDKYNLEYCIYEEYYPFDLAIKREKELKKWNRQKKEALINKNPEWKVLATEKGFMHKQV
PFAQQVADLLRELEENQIIPPCSGNDTPS
<LF63_RS01415--*Oleigrimonas soli*--77.42 (SEQ ID NO:647)
MRSREPCVYMLASRRNGTIYIGVTTDLPLDRVWQHRNDLVAGFTRHYRVHRLVWFESHATMESAIAREKTLKNWKRAWKLQLEETNPYWRDLVPSRL
<LS64_RS04715--*Helicobacter saguini*--77.89 (SEQ ID NO:648)
MKKAYIYIITNKNKGTLYTGVTSDLIKRIYEHKKNKIVKGFSSKYDLKNLVYEMFENIESAIIREKQIKAGSRKKKLELINNFPNTWKDLYEDLL
<GY21_RS01165--*Cryobacterium roopkundense*--58.97 (SEQ ID NO:649)
MPYMYIVQCADGSFYVGTWNIEERLWQHTAGDGAAYTRRRRPVTLVYLEEGPRIEDAYAREKQVQGWWSRAKRLALIEGRPADLAPLSVRRHRPEKSPP
<MC45_RS17520--*Sphingomonas taxi*--82.95 (SEQ ID NO:650)
MEWQPCVYIILASRRHGTLYVGVTSNLIGRLIQHREGLIPGFTSRYGVHRLVWYDVADTMEAAIHREKQLKWRNRDWKRLRIEENPGWDDLALGLGLPPVKES
DT
<MC45_RS03525--*Sphingomonas taxi*--76.74 (SEQ ID NO:651)
MSKPGVYVMMASQNGTIYLGVTSDLPARVYQHRNGLVEGFTTKYGCFLVWYEVHDELESARYRELQLKKWNRGWKRLRIEKNPDWDDLWDLFESLF
<MC45_RS03505--*Sphingomonas taxi*--74.70 (SEQ ID NO:652)
MASGRNGTIYLGVTSDLIRRAYEHRNRLVDGFTKEHGCIILLVWYDASDDIQAARQRELQMKKWKRAWKVDLIEQGNPQWKDLYETLF
<MC45_RS08030--*Sphingomonas taxi*--62.22 (SEQ ID NO:653)
MADGRGGWVYMLSDRYRGAMYVGVTQAQLEARIHQHRTGNGSDFCSRYGLHRLVWAEHGETIEACIAHEERLKRWRREWKFALIERGNPDWRDLFDSVV
<N792_RS01745--*Lysobacter concretions*_Ko07_=DSM_16239--84.95 (SEQ ID NO:654)
MERQPCVYLLASARNGTIYVGVTSDLVTRIWHQHREHVAGGFSARHDTVRLVWYEQHDAMESAIQREKRIKKWNRWKRLRIDSSNPSWRDLWPEIIGAVPTAN
VHGSPPSRG
<Q763_RS15205--*Flavobacterium beibuense*_F44-8--63.22 (SEQ ID NO:655)
MFQTVGEHNYFVYIILTNKSKSVLYIGFTTNLKDRLHFHQNEAHSKSFTARYNCFYLIWEHYTDVQLAIDRETQLKKWSRIKKDDLINKINPDWRFLNEEI
<SP5_RS03305--*Sphingomonas parapaucimobilis*_NBRC_15100--73.03 (SEQ ID NO:656)
MDGPAYVYIMANRRNGTIYLGSTINLPKRAWHEHRNGVVEGFTKRYGCHLLVWYEVHESLEFARQRELQMKKEWRRGWKLQIEGLNPDWDDLWDLRIALP
<RG22_RS09335--*Flaviramulus ichthyenteri*_Th78--64.71 (SEQ ID NO:657)
MKSSYAYILTNKYRTTFYIGVTANLSNRLEKHYDGTASKFTEKYNIKDLIYFEIFSNIEQAIAREKQLKNWHKE
<NW74_RS00250--*Parvimonas micra*--67.06 (SEQ ID NO:658)
MFYVYIISNNFRNLYIGVTRNTERRLYEHKKNLIDGFSKKNLNLVLFYETTEYALDAIAREKQLKKWNRREKKFDLIKSLNPELRDLSEDW
<LK12_RS06435--*Novosphingobium malaysiense*--60.47 (SEQ ID NO:659)
MGGWYIILTNKPRGVLYIGVTSNIAARIAQHRSGTGSFACRRYGLNRLVYEVHABEILSAIAREKALKAWKRDWKIRLIERNPDWTDLFERLL
<PK98_RS02965--*Croceibacterium mercuriale*--60.92 (SEQ ID NO:660)
MAPGAHGTLYVGVTAAIARVMQHREGRGSVFCKEHGCPRLVYVERHEGIEBAIRREKQVKAWRQWKLRLIRAGNPEWRDLFEDVLA
<SKP52_RS11820--*Sphingopyxis fribergensis*--80.22 (SEQ ID NO:661)
MHENFGPTVYLLASRRNGTIYLGVTSNLTQRIHQHREGLIPGFTREYGVKLLVWFEGHATMTAIQREKRIKKWNRRAWKLELIESANPDWRDLAEDFGFAPIG

FIG. 11U

37/106

<SKP52_RS16450--Sphingopyxis_fribergensis--73.03 (SEQ ID NO:662)
MHADFQPCVYIMASARHGTLYTGVTSHLIQRICQHRGGSFDGFTKRYGVKTLVWFELAGTMHAATREKQIKNWQRKWKIELIEAANPEWRDLAEELGLTPLG
G
<SKP52_RS03260--Sphingopyxis_fribergensis--69.41 (SEQ ID NO:663)
MASGRNGTTYLGVTSDLLTRAWQHRNEVIDSFSKKYGCMLRVWYEAFFDDIQARQRELQMKWKRAWKVELIERDNPQWLDLFDRLSL
<NJ75_RS15995--Novosphingobium_subterraneum--60.00 (SEQ ID NO:664)
MDRERLGGWVYIMADRYRGAMYVGVTSIAARIHQHRNGCGSDFCARYGLKRLVWAERGDRIEDCIAHERRLKRWREKFDLIERGNPDWLDLYDVIVG
<OC25_RS23550--Pedobacter_kyungheensis--69.57 (SEQ ID NO:665)
MERGGCIYILTNTFQKTTLYIGVTSDLRSRIYEHQNKIYPKSFSAKYNLTHCIYEEVFSSIEEADREKQLKKWNRKKTETLINDFNPWNWDLTDIINEWD
<OC25_RS20115--Pedobacter_kyungheensis--66.67 (SEQ ID NO:666)
MERGGWVYIMTNVHNTVYIYIGVTSDLRSRVFEHKNKEYPTSFTAKYNCTKLVFYEFYFDIEEATAEAEKRLKKWNRQWKIELINSKNPDWLDLFNEDL
<OC25_RS22345--Pedobacter_kyungheensis--64.84 (SEQ ID NO:667)
MERGGCVYILTNTLNTVLYIGVTSELYFRIIEHREKKYPNSFTAKYNCKLIYYEQFFIEEATAKEKQLKNWKRQWKINLINEHNPWDDLFNLLE
<RM51_RS07650--Chryseobacterium_taiwanense--64.21 (SEQ ID NO:668)
MKPGFVYIMTNKNNTVLYTGVTSNLPKRQVQEHQEKAYELSFSTSKYNLNKLVWESFQEIIGNAIFREKQIKAGSRQKINLINSNPWDLTEEEKIDMNP
<RM51_RS08260--Chryseobacterium_taiwanense--61.70 (SEQ ID NO:669)
MIEFSEGIYKFVYIILTNRKKTVLYTGVTSNHLQRLYHHSIKRNSNSTARYNVEFLIYYEKFGWVQQAIEREKEIKNFSRNKLELIRSVNPNLDFWNYLFK
DIM
<NELON_RS07125--Neisseria_elongata_subsp._glycolytica_ATCC_29315--93.62 (SEQ ID NO:670)
MMQPAVYILASQRNGTLYIGVTSNLTQRVYQHREHLVQGFTNQHHVTLVWYELHSTMEHAITREKQLKKWNRQWKRLRIEKNPSWQDLWFEIIK
<BN1125_RS00410--Aeromonas_fluviialis--81.72 (SEQ ID NO:671)
MQPFVYIMASQYHGTYIGVTSQLLQVRVWQHKNIGTGGFTHQYNVMTLVYFEAHKTMYYAAITREKQLKKWRREWKINLIETQNPWTQDLWYDITDGSPPSRG
<BN1174_RS05750--Rickettsia_hoogstraalii--77.78 (SEQ ID NO:672)
MVYIVTNKKNGILYTGVTSNLIKRIYQKNDIINGFSKKYNCKILVFYIEHQTMNSAINREKQIKAGSRDKKIKLIEQTNINWDDLYNEIIL
<TD18_RS02945--Psychroserpens_mesophilus--67.44 (SEQ ID NO:673)
MTFWYVYIMAYKPNVGIYIGVTDNLDERVREHKYKIYPKSFTARYNCENLVYFEEDNGEDASTRERRFKKWKREWKVNLIEEMNPSTWDLNINWLDNNREF
RVNK
<SB78_RS04230--Rickettsia_asembonensis--83.33 (SEQ ID NO:674)
MKQPAVYIISNKRNGTIYGVTSNLIKRIYEHKNNTTNGFSKKYDCKMLVFYIEHETMDSAITREKQIKSGSRTKKLNLEIEMNVN
<SB78_RS01325--Rickettsia_asembonensis--77.92 (SEQ ID NO:675)
MKQYVYILASGRNGTLYIGVTSNIIKRTYEHKKHIIKGFSSKYNINKLVYFEQYDDIIAIAINREKLLKEYKRQWKLNHL
<GL4_RS09030--Methyloceanibacter_caenitepidi--75.86 (SEQ ID NO:676)
MLASRIGGTLYVGVTDNLIIRRIEHRTKAVEGFTKRYGVARLVYFEQFDDITVAIEREKRLKKWRRAWKVLIKEHNPWDDLYASITSA
<NMS_RS08240--Nonlabens_marinus_S1-08--60.47 (SEQ ID NO:677)
MPRRSFHNTYVYFLSNKKDGVLYIGVTGGIEGRLLRRHLRGDSTFTKKYQAHKLIVYFEDFQYVNDIAIREKQLKNWHRQWKINLIEAENPEWDLSPERETS
<SD69_RS05615--Gilvimarinus_agarilyticus--78.57 (SEQ ID NO:678)
MRKRNGSLYVGVTSHLVQRVWQHKQHQVARFTQKYKVTYLVYFEMHDDMLAAIAIREKQIKHWQRRWKVELIETHNPYWRDLYGDIL
<AW25_RS08025--Francisella_tularensis_subsp._novicida_U112--74.16 (SEQ ID NO:679)
MKKGYVDILTNNNSVLYVGVTSDIVKRIYEHKNKFVDGFSRRYNLEKLVYFELYQNIEDAIIREKQLKKWNRWKNRIINGMNPKWDDLYNEISS
<BZ13_RS08320--Francisella_philomiragia_subsp._philomiragia_ATCC_25015--72.22 (SEQ ID NO:680)
MAFVYILASKCNGTLYIGVTSNLIKRVYEHKQGYSDGFTKKYDVKKLVYEQFNNTDAIYREKRLKTWQRKWKLDLINKFNPRWLDLYASILA
<jaqu_RS09505--Jannaschia_aquimarina--73.33 (SEQ ID NO:681)
MWTYILASQRNGTLYVGVTSDLVRRVWEHREHVAPNSFTARCNTVTKLVWFERHEAAPYAIKREKSIIKKWKRAWKISLIERENPEWCDLWSGISRLS
<LZ09_RS22190--Desulfonatronum_thioautotrophicum--74.44 (SEQ ID NO:682)
MCIYSPVNGTVLFIIGVTSNLLKRVWEHKNVIEGFTKKYCVHNLVYEEHHDMAALMREKQLKKWNRWIKIRLIEEFNPSTWDKYELLKDGAFPSARE
<TS85_RS23435--Sphingomonas_hengshuiensis--71.26 (SEQ ID NO:683)
MRERTPCVYILASRPGTLYIGVTSDLFGRLYQHRNGVTAGFVARYNVQRLVYFETCEAMDGAIAIREKQLKRWHRWKNLIERANPEWDLASGLGIAEPLS
PGD
<TS85_RS15460--Sphingomonas_hengshuiensis--71.26 (SEQ ID NO:684)
MDKLPCVYMFASGRHGTLYVGVTDLMARVHQHRSGSLGGFTKRHGKRLVWYQTGGDMAGAILREKQLKAWRRDWKVALIEQDNPCWEDLAIGLGFPLG
<TS85_RS22645--Sphingomonas_hengshuiensis--70.11 (SEQ ID NO:685)
MNGGLLRDGLAASLRVIERTPCVYILASGYNGTLYVGVTSNLVGRVMQHRDGAFAHGTTRRYGIRRLVYFEVAETMDAAIHREKQLKRYRDRWKNLIERQNP
EWNLDLVGFLPPLG
<TS85_RS05945--Sphingomonas_hengshuiensis--69.77 (SEQ ID NO:686)
MTNKPRGVLYIGVTAALNQVRVQHQHREGKGSFCARYNLTRLVLAEPHDSIDDAIVREKRLKEWRRAWKIALIEETNPEWRDLFGEVWE
<TS85_RS21840--Sphingomonas_hengshuiensis--61.80 (SEQ ID NO:687)
MENGEQGGWYIMANRYRTMYAGVTAHLAERMHQHRSDEGSEFCQYGLNRLVWAEHAPSITDCIAQEKRLKRWREWKFALIERANPEWCDLYDQIR
<UQ52_RS01915--Rickettsia_raoultii--76.67 (SEQ ID NO:688)
MYVIVTNKKNGILYTGVTSNLIKRIYQHKNDTIKGYSKYNCKILVFYIEHQTMNSAINRAKQIKAGLRDKKIKLIEQTNINWDDLYNEIIL
<AW14_RS00875--Siansivirga_zeaxanthinifaciens_CC-SAMT-1--65.88 (SEQ ID NO:689)
MDVHYVYILTNNKNTVLYVGRTKQLKTRISQHKNNSLKTFTGKYNVDKLVYFETTKYVNSIKRERQIKKWNREWKINLINGLNPDKDLTEYI
<AW14_RS03385--Siansivirga_zeaxanthinifaciens_CC-SAMT-1--68.60 (SEQ ID NO:690)
MNHVYVYIITNKSHGVLYIGVTNHIDERVKEHKLKVPKSFSTAKYNCCKLIYFETFNQADNASLREKQFKKWKRDWKIKLIEDMNPWSWDLNINWNLNFKLR
<PK35_RS01135--Tamlana_nanhaiensis--74.73 (SEQ ID NO:691)
MYNRQYSFYVYILSNKNGTLYIGMTNDLERMFEHKNKLVEGFTKKYSLNKLIYLEHFQYVNDIAIKREKQLKNWKRQWKI GLIEKDADWDLSSDWFN
<PK35_RS14820--Tamlana_nanhaiensis--67.06 (SEQ ID NO:692)
MDVHYVYILTNNKNTVLYVGRTKQLKRLKQHKNNSSKTFTGKYNVTKLVYFETTKYVNSIKRERQIKKWNRAWKISLINSNPDKDLSEYL
<PK35_RS05290--Tamlana_nanhaiensis--68.60 (SEQ ID NO:693)
MKLWYVYIMANKPNGVLYIGVTDNLDERVKEHKLKVPKSFSTAKYNCCKLIYFEFSEGGIAERRERQFKKWKRDWKIELIEEMNPWSWDLNINWNLNFKIR
K

FIG. 11V

<PW52_RS14700--Tamlana_sedimentorum--67.06 (SEQ ID NO:694)
MDVHYVYILTNNKHTVLYVGRTKQLKRLKQHKTNPKFTGTGKYNVTKLVYFETTKYVNNNSIKRERQIKKWNRSWKISLINSNPDKDLSEYI
<LH29_RS11855--Draconibacterium_sediminis--59.34 (SEQ ID NO:695)
MKAWTYITTNNSHNRVLYTGVTSDLKGRIDSHKIKKYPNAFTAKYNADKLWVYQEFNSIIDARAREKQLKAGNRAKKIKLIETINPKWEDLFEKL
<UB38_RS07805--Photobacterium_iliopiscarium--69.89 (SEQ ID NO:696)
MSREPAVYIILTSELKTVLYIGVTSQVLARIWLHKNDVVVGFSQKYQVHHLVHYEVFDTMHAAIAREKQLKKWRRSWKEQLITENNPOWLDLYDSL
<LFA_RS08575--Legionella_fallonii_LLAP-10--77.78 (SEQ ID NO:697)
MKESYVYILANKHNGTLYIGLTSDLIKRVWEHKNKFAPGFTQKYNVTILVYIEIFTDITLAAAREKRLKEWKQWKVDLIQKNNPKWKDLYLEIVQ
<LFA_RS04440--Legionella_fallonii_LLAP-10--62.22 (SEQ ID NO:698)
MSKEYIYIYLLTNTHTVLYTGVTTNNLVRRYIEHKNKAQLGFTQKYNVNLVYIEICSDIVVAITREKQIKAWTRKKKEDLIATLNPEWADLYRSIC
<LMI_RS05950--Tatlockia_micdadei--69.77 (SEQ ID NO:699)
MREYIYIYLLTNKYKNVLYTGITNNLIRRVYEHKHLKQGGFTQKYNVDRLVYIEIFTDIKEAITREKQIKGWARCKKNNTLIETINPDWKDLYESLLC
<LHA_RS05625--Legionella_hackeliae--76.74 (SEQ ID NO:700)
MKSIFYIYIASKRNGTLYTGSTDLIKRIWEHKNQVPGFTAKYNHMLVYIEMHETVYEAACREKRFKNWPRQWKLNLIKMNPAWDDLYEMICS
<UG56_RS27160--Nocardioides_luteus--57.83 (SEQ ID NO:701)
MAYTYMLLCADRTLYVGSTDLERRLSEHRVGLGSEYTRHRLVQVLWYEQYDSVEHAFEREKQIQNWSRAKRLALVRGRPGDLPDLANTSQTPHP
<VZ94_RS16955--Methylococcus_oryzae--84.88 (SEQ ID NO:702)
MYILASKKNGTLYVGVTSNLVQRIYQHRENLVDGFTKRYGVHDLVWYEPHETMESAILKEKRLKTWERAAKIRLIESTNPTWQDLWPTLTDTSL
<VC82_RS06970--Muricauda_lutaonensis--69.88 (SEQ ID NO:703)
MSNRPNGLVLYGVTDNIDERVKEHKLKVNPKAFTSRYNCDKLVYFEIYENGAAVREKRLKKWKREWKIELIEDMNPSSDLSMNWKLNNYRLRK
<VC82_RS07380--Muricauda_lutaonensis--65.85 (SEQ ID NO:704)
MFNEGWHYVYIITNQHKTLYTGMTNNLAKRLYQHRENIRESKTFVARYKCSHLLYEEFTWVQEAIAAREKEIKGWVRRKKLDLIKGSNPDCLKFIEKLF
YQDV
<RAMDARK_RS03300--Rickettsia_amblyommatidis_str._Darkwater--77.78 (SEQ ID NO:705)
MVYIVTNKNGILYTGVTSNLIKRIYQHKNDTIKGFSSKYNCILVFYIEIHTQMHSAINREKQIKAGLRDKKIKLIEQTNNWDDLYNEIIL
<REIP_RS05625--Rickettsia_endosymbiont_of_Ixodes_pacificus--78.49 (SEQ ID NO:706)
MKQPIVYIVTNEKNGILYTGVTSNLIKRIYQHKNDTIKGFSSKYNCILVFYIEIHTQMHSAINREKQIKAGSRDKKIKLIEQTNNWDDLYNEIIL
<REIP_RS04255--Rickettsia_endosymbiont_of_Ixodes_pacificus--74.68 (SEQ ID NO:707)
MYIGVTSNLIKRIYEHKHKIKIGFSSKYNNKLVYFEQYDNI IAAINREKLLKEYKRQWKLNLIKESQNPWEHDLYEKIL
<REIP_RS05340--Rickettsia_endosymbiont_of_Ixodes_pacificus--73.42 (SEQ ID NO:708)
MYIGVTSNLIKRIYEHKHKIKIGFSSKYNNKLVYFEQYDNI IAAINREKLLKEYKRQWKLNLIKESQNPWEHGLYEKIL
<REIP_RS07605--Rickettsia_endosymbiont_of_Ixodes_pacificus--73.75 (SEQ ID NO:709)
MFIYFLINVTLYIGVTSDIKKRIWQHKKQKIRKGFSTKYDIDKLVYFEQFNDVNLAIGREKRLKEYKRKWKLDLINSNPSWHDLYEDFLK
<REIP_RS09195--Rickettsia_endosymbiont_of_Ixodes_pacificus--80.43 (SEQ ID NO:710)
MLCSNKHGTYITNNIIRRAYEHKQKIKIGFTSKYNI ICLVYTEEFTKI
<TW84_RS23360--Vibrio_neptunius--68.54 (SEQ ID NO:711)
MYVKQPAVYILANQSGNCLYIGVTSDLKARVWQHKNQSGVGGFSKRYNIDCLVYELFLDITDAICREKQLKKWKRKWKENLIAKTNRWCVDLYDSL
<VE25_RS13055--Devosia_geojensis--82.22 (SEQ ID NO:712)
MHRFTFCVYIMASRRNGTLYIGVTNDIVRRAWEHREGAVEGFTRRKHGIGHLVWFVHVHDNIDAAITREKLIKQWQRAWKGRLEETNPDPWRDLWDLTDQR
<VE25_RS21215--Devosia_geojensis--78.65 (SEQ ID NO:713)
MGTFDNYFVYILASRKYGAIYIGVTRDLVRRVYTHREKIVPGFTGKYNIRQLVHFVEQYAAPQDAIAREKQLKKWHRAWKIELIEQSNPEWRDLWEDIAHP
<VW29_RS01880--Devosia_limi_DSM_17137--73.03 (SEQ ID NO:714)
MSNGREFVYILASGKNGTLYIGVTGNLIGRAYTHRNDLLDSFTHRSVHDLVYFEQFEDPISAITREKRIKKWNRDQWVRLIEESNPQWDLRFHDIAGWLD
<VW29_RS17865--Devosia_limi_DSM_17137--70.33 (SEQ ID NO:715)
MGRIFTVYMMASMKNGTLYTGVTGNPAGRIWQHKTTHAFQSFRTARHKVERLVWYEVHADPLAAIQREKRIKEWQREWKRIRLITENPDWRDLSEAEMQV
<VW29_RS02980--Devosia_limi_DSM_17137--67.42 (SEQ ID NO:716)
MAAWVYIMANRPFGTIYVGATSNLIRRVYEHNRDLVPSFTREHQLHLLVYFEQHQDIHRAMQREHNIKHWSRTWKDQLIEAGNPEWDDLYLSIAPG
<VW35_RS09870--Devosia_solis--68.54 (SEQ ID NO:717)
MKGYVYILASKKNGTLYTGVTSNPSRRLFEHQHALTPGFTSRYGVKTLVWLEEHDLVVDARREKAIKKYPRQWKINLIEAVNPWDADISSWLFDT
<VW23_RS12300--Devosia_insulae_DS-56--79.07 (SEQ ID NO:718)
MDSEKDFYVYMLASKMGGTIYIGVTSDLIGRTYIHRNLVPGFTSKYGVHRLVWFEEQGDVEYAILREKQLKKWNRRAWKIRLIEEKNPAWVDLYPTLFP
<VW23_RS18025--Devosia_insulae_DS-56--74.73 (SEQ ID NO:719)
MDKQFFVYMMASGKRGTYITGVTSDLIGRAWQHRTHAIGSFRTARYGVTLVWFEEQGDAAEAITREKRIKEWRRAWKVELIEATNPRWDDLWDEIVG
<VW23_RS20940--Devosia_insulae_DS-56--77.91 (SEQ ID NO:720)
MSTSGRFFVYILASRKYGTYITGVTDLPVRIYIHRDLIPGFSRYHVHKLTVYFEQHQDDPYNAITREKQIKKWLRAWKIELIEKDNPHWDDLYDGLFG
<WH91_RS21070--Devosia_psychrophila--70.33 (SEQ ID NO:721)
MGKTYMVYIMANRKHGAVYTGVTGNPAGRVWQHKNHVFKGSFTDQWGCCKLVWYEVHDDPESAIIREKRIKDWKRAWKDRIIEEMNPDRDLWDLVDFG
<WH87_RS02675--Devosia_epidermidihirudinis--69.57 (SEQ ID NO:722)
MKQFCVYILTNNKHNVTYTGVTSNLVGRYTHRNDLLDGFSRRYKVDLVWYEHYETAESAITREKQIKKWRRAWKDELVAAMNPDRDLWDLVDAKHVY
<WX88_RS11560--Hymenobacter_terrenus--61.90 (SEQ ID NO:723)
MYVYMLTNPTRAVLYIGVTNNLTRLAEHADNLGNPNKFTGRYQINLLVYFETQTSQAQAIAREKELKGGSRRAKKEVLIKFNPTWAAIDLETWAGEPKES
<WX88_RS12585--Hymenobacter_terrenus--52.94 (SEQ ID NO:724)
MSYVYVYITTPAKTVLYIGITNNLSNRMTQHYTQRGSTQTFAGRYFCYNLVHVEEYQDINAAIAREKELKGWTRAKKDLVLIADNPERLFIQT
<PKOR_RS24105--Pontibacter_korlensis--51.65 (SEQ ID NO:725)
MSSCNVYVYITTPRKTYITGVTDLPRLQEHFVNSGNASTFAGKYCYNPFVYHAIEREKEIKKWNHSSKESLITRFNCPWKFNLNEVQVE
<OYT1_RS07215--Ferriphaselus_amicola--77.42 (SEQ ID NO:726)
MEKQAAVYIILASRNGTLYVGVTSDLVRRVWEHKNLDLVEGFSKRYGVHNLVYFELGENMFAAIREKQLKAGSRKKKLQIEAKNPDRDLFDEIV
<OYT1_RS11695--Ferriphaselus_amicola--54.79 (SEQ ID NO:727)
MPSWICYLLECADGTLYCGITNDLEKRLAAHNAGIKAKYTRSTPVALVYMESHADKSATLKREMAIKKLSRSQKLSLIATSTVSAD

FIG. 11W

<SCH01S_RS07760--Sphingomonas_changbaiensis_NBRC_104936--69.23 (SEQ ID NO:728)
MAVERTPAVYIVTNSFHGTLTYGTSTNLPGRVYQHRNGTFRGFTAHEGCKRLVWFIEIHATMEEAIVREKRIKRWREWKYALIEADNPRTWRDLAEDFGFEPILL
RTPGVEKQVGPSSSPG
<AA088_RS17065--Kordia_zhangzhouensis--72.41 (SEQ ID NO:729)
MQIYSVYIITNKPNGVLYIGVTKDLKRRIRQHKNKVQPTTFSARYNLSRLVYFESFETKDEALLREKRMKKWNDRDWKVELIEKENPDCENLVDKIK
<AA089_RS07885--Kordia_jejudonensis--76.92 (SEQ ID NO:730)
MYKTIHQYVYVILSNKNGTLYIGISNNLERRI FEHKYKLVEGFSKKYNLTCLVYFEI FQFVDEAIAREKKLKKWKRWKINLIEEKNPDWNLALDWSIDMD
SATSAE
<AA089_RS06665--Kordia_jejudonensis--72.41 (SEQ ID NO:731)
METYSVYIITNKPNGVLYIGITKDLKRRIRQHKNKVHPTTFSARYNLSRLVYFENHKNKEEALLREKRMKKWNDRDWKVELIEKENPNWEDLINKVN
<AA089_RS19330--Kordia_jejudonensis--67.06 (SEQ ID NO:732)
MDTHYVYILTNSKSHVLYIGRTNQLKRRLYQHKNNGIRTFGTGKYNVHKLIVYFETTKYVNNAIQRRERQLKKWKREWKINLINGLNPDWDLSEYV
<AA089_RS19310--Kordia_jejudonensis--62.35 (SEQ ID NO:733)
MKKSYIYILTNTYRTTFYIGITADLKNRIISHADGKGSEFTKKYNVTDLVYEEFELDINQAIAREKQLKNWKRWDWKIELIKKLNPSLETINL
<AA089_RS19320--Kordia_jejudonensis--62.07 (SEQ ID NO:734)
MKKSYIYILTNTYRTTFYIGVTTDLKRRINDHAAGVSGSKFTKKYNVIDLIYEEFELDVNQAI SREKQLKSWKRWDWKIDLIKLNPTLETILY
<YP76_RS04090--Sphingobium_chungbukense--63.22 (SEQ ID NO:735)
MTNKPCGVLYGITSNLARVDQHRRGVSGDFCRKYNLHRLVLAEEHGRVEDAIAQEKAALTAWKRNWKIRLIEESNPEWRDLFAFIV
<YH63_RS17740--Afipia_massiliensis--77.53 (SEQ ID NO:736)
MSFFVYMLASQRNGTLYIGMTDDLPRRVWQHRTGAVPGFTSRYEVMKLVWYQAHSRESAFMRERQLKKWNRGWKIALIEAGNPEWRDLFEDLSP
<YH63_RS12285--Afipia_massiliensis--78.49 (SEQ ID NO:737)
MKQPCVYIVASQRNGTIYTGVTSTNLPRRTFEHREGLIKGTFTKKYNCKILVWYIEHETMESAILREKRIKAGSREKKLGLIQASNPWDDLYGTILY
<YH63_RS18940--Afipia_massiliensis--75.28 (SEQ ID NO:738)
MKQPCAYIVASERNGLTYTGVTADLPRRAFEHREGAVKGFSTKYGCKILVWYIEHETMLEIAIAREKQIKAGSRAKKLLLIESTNPNWIDLYESLM
<YH63_RS11525--Afipia_massiliensis--70.59 (SEQ ID NO:739)
MSVTTYVYILASQLNGTLYIGITNNLALRLSQHKGGRGSEFAQRYAVNRLVYVEGYDTPSDAIRREKQLKKWNRAWKIELIEKMNPEWRDLSDLAISGIS
<YH63_RS09840--Afipia_massiliensis--68.89 (SEQ ID NO:740)
MAGLVPAIHVLQNCIMTGGFVYILTNRPNGTLYVGVTDNLVRRRIHEHRSGFVDGFTKRHELKLLVYFEQFDDIRDAIQREHNIKHWPRAWKVRKILVENPDWV
DLYPKIIE
<YH63_RS10705--Afipia_massiliensis--69.88 (SEQ ID NO:741)
MLASARNGTLYIGVTNNVAARLELHRLGQGGSEFVKRYRVYRLVYVETYSRAEEAIQREKQMKKNRAWKIRLIEQDNPEWRDLTMDLA
<YH63_RS09225--Afipia_massiliensis--66.67 (SEQ ID NO:742)
MYVYVYILASGKHGTLYIGVTNSIQQRDLHRSRGHSGSEFVKKHEIHRVLHLEEHQSPAAIRREKQLKKWNDRDWKIRLIESENPDWSDRSDLIG
<BN1233_RS00585--Microvirga_massiliensis--73.12 (SEQ ID NO:743)
MKQPAVYIMANRRNGTLYTGVTSSLSQRVWQHREGFIPGFTSRYACKLLVWYEEHEDMTIAIAREKQIKASSRRKKLALIEAQNPDRDLYHDLI
<AAW31_RS00495--Nitrosomonas_communis--84.09 (SEQ ID NO:744)
MLASRRNGTLYVGVTSNLVKRVWEHKNIVKGFTQKYGVHTLVWFELHETMESAIRREKAIKEWKRWDWLKELIEQQNPAPWDLYSEIL
<WYH_RS13790--Altererythrobacter_atlanticus--81.32 (SEQ ID NO:745)
MQRDFGPTVYIMASRRNGTLYIGVTSNLLQRAYQHRTGAI SGFTKEHGVRMLVWFQHSMTDHAI TREKRLKKWNRAWKRLIEDSNSAWRDLAEDFGFEPIC
<WYH_RS16355--Altererythrobacter_atlanticus--64.71 (SEQ ID NO:746)
MKCGYVYILTNRPEGVLYIGVTEDLAKRVAQHRS GAVSSFTKRYNCHHLVWFERFENIHDARLFERRMKKNWRDWKIAIEERNQAWQDLSHALL
<AAW00_RS02450--Erythrobacter_luteus--78.72 (SEQ ID NO:747)
MARDIQPCVYIMANRNGTIYVGVTS DIAQRGWQHRQGLVDGFTQYCKCTLVWYELHPTMESAILREKQLKAGSRKKKIALIVERNPDWRDLYAEVIA
<AAW00_RS06785--Erythrobacter_luteus--71.59 (SEQ ID NO:748)
MPSERQPCVYILASKAYGTLYIGVTS DLVARVWQHRNDAPDFTSRDYDKTLVRYELFATIHEAIAREKQLKRWHRQWKINLIESENPDWHDLPAGIGLPIT
PRLGQDGS
<AAW00_RS09850--Erythrobacter_luteus--66.28 (SEQ ID NO:749)
MGGWVYIMSNRPGGVLYIGVTS DLPRVHQRHTGKGS AFCRRYGLTRLVYAEQHDDI HAAIVREKAMKAWKREWKIELIETINPAWDDLYERLAGL
<AAW00_RS05950--Erythrobacter_luteus--55.22 (SEQ ID NO:750)
MKDFYAYIVRCADGSYYTGHTEDIDARIAALNSGAIDGYTLRRRPVVLVWSEFPTRDEAFAGERRIKGSRAKKEALIAGDWDEI SRLARNRQD
<AA309_RS15735--Microvirga_vignae--74.16 (SEQ ID NO:751)
MTYHVYMLASRRHGTLTYIGVTNDLARRMSEHRSGDGPFTAQYDVTRLVWYESYARIAEAIAREKALKKKWRDWKIRLSEEMNPEWEDLYLALNQ
<TVD_RS07005--Thioalkalivibrio_versutus--73.12 (SEQ ID NO:752)
MEKHPAVYIILTNRREGTLYTGVTSTNLVRIAQHRNDLVPGFTTRYKLHRLVWFQEHESMDEAIRREKQIKRWRRPWKIDLVETENPAWRDLYNDLF
<TVD_RS09195--Thioalkalivibrio_versutus--70.97 (SEQ ID NO:753)
MKQPCVYIMSNRRD GALYVGVTSNLVQRVWQHRSGEASGFTTRYGLHRLVWFESCEEMTSAIEREKQLKAGPRRRKVALIEQTNPEWRDLFDDIVG
<QQ39_RS18370--Pragia_fontium--69.32 (SEQ ID NO:754)
MKQPCVYILASRPKALYVGVTSNLPNRIECHQRGEVLGFTQKYNIKRLVWYELHSSMYQAITREKNIKHWRRIWKIELIENMPDWDDLKSTLSA
<Epro_RS03070--Endomicrobium_proavitum--75.00 (SEQ ID NO:755)
MTKQYVYI IKNRLRGTLTYIGVTNDIARRMYEHKNKLEIGFSKKYSLDKLVYCEICDDINSALRREKILKEWKRVWKIELIEQSNNSNNDLIENDV IIK
<AM2010_RS07090--Altererythrobacter_marensis--74.71 (SEQ ID NO:756)
MLNLFQHPSPGTLRTMPFERNPCVYILASGRYGTLYIGVTS DLIQRLYQHREGITGGFTKRYRVHRLVRYEMFGDMEAILREMQLKRWHRQWKINLIERDNP
HWADLAVGLGFEPALCGGRNGS
<AM2010_RS00015--Altererythrobacter_marensis--67.42 (SEQ ID NO:757)
MTDEPYCWYILASQKRGTLTYIGLTSRLAERIEHRSGRVPGFTQYGVRLVHLEFFSDIHEAIAREKALKKKWRRAWKIALIEAGNPDRDLWHDLNR
<AM2010_RS07090--Altererythrobacter_marensis--65.96 (SEQ ID NO:758)
MAGHIYILTNPKHGVLVYGVTS DLPVRIMQHRDGTGSKFCRRYGLKQLVYVEDHAEIESAIAREKAIKAWKRAWKIELIERTNPSWSDLFVDV I
<AM2010_RS02700--Altererythrobacter_marensis--63.33 (SEQ ID NO:759)
MKPGCVYILTNPMPGVLYIGVTTDLPRRVEQHRS GAVSSFSKRYNCHRLVWFAEFENIHDARSFERQMKKNRAWKVARIAMNPEWRDLAEL
<SQ03_RS07165--Methylobacterium_platani_JCM_14648--69.66 (SEQ ID NO:760s)
MAKGGFVYMLASGRNGTLYLGSTVDLCRRVYEHQQAHPGFTTRYGVHVLVWYEHYDRI TDARHREYTLKKWRRAWKIALIEAMNPDRDLADLNG

FIG. 11X

<QR78_RS16900--Methylobacterium_indicum--75.56 (SEQ ID NO:761)
MASFVYMMASRPHGTLTYIGVTNNLARRVFEHQTKVLGGFTARYDVARLVWYETYESINEAIAREKDLKRWRDDWKLKLVDEFNPGWSDLYSTLAGF
<QR78_RS22195--Methylobacterium_indicum--73.03 (SEQ ID NO:762)
MAKGGFVYMLASGRNGTLYLGSTTDLCRRIHEHQQAHPGFTTRYDVTHLVWYEHYELITDARHREYALKKWRSSWKLSEAMNPDPWRDLSADLNN
<ACM46_RS06095--Chryseobacterium_angstadtii--64.21 (SEQ ID NO:763)
MKPGFVYIMTNKNHTLTLYGTVSNLPRVQHQKETYYEQSFTSRYKLYILVWWSFQEI GDAIFREKQIKAGSRQKKLDLINSINPEWRDLTGDIENIMDPF
<V474_RS13555--Novosphingobium_barchaimii_LL02--76.14 (SEQ ID NO:764)
MQNERQPCVYILASGRYGTLYIGVTSNLIIRRIWQHRSVDLPFGTSTRYDVHRLVHFEMFGEMIPAITREKQLKRWHRQWKINLINAENPEWRDLAVALGFSPLA
SGKTRDGP
<V473_RS14315--Sphingobium_czechense_LL01--60.00 (SEQ ID NO:765)
MKFTVTPAKAGVSGGALRWDMARGGYTYIMTNKPRGVLYIGVTADIAARAEQHRNGTGSFAFKTYGLTRLVLVEPHDDITLAIAREKALKAWKREWKIRLVE
ESNPGWRDMSDVIF
<V473_RS06260--Sphingobium_czechense_LL01--57.35 (SEQ ID NO:766)
MTFWAYMLHCADRSFYVGHTDNLQRIQHQDQGGIAGYTQTRRPVKLVWSQAFGTRMEALEAERQIKGWSRAKKLALIREDWTLISSLARSNKEKKSPTSGSG
RTE
<MALG_RS23320--Marinovum_algicola_DG_898--64.71 (SEQ ID NO:767)
MGHFVYILASRPGGAIYIGRSNLRARIAAHRAGLSRHTESYNIRTLVWFEEHADFETSLRERRLKWRRAWKDALIMEHNPQWRDITHLVPE
<AB843_RS18250--Nitratireductor_solii--72.22 (SEQ ID NO:768)
MGGYVYILASRKNGTLYGTVSDLEQRVWEHREGITGGFTEKYRVKMLVWYRDYHDIRDAIDDEKRIKWRWKWLEMI EAINPQWRDLYSTLNQ
<AB843_RS02000--Nitratireductor_solii--56.67 (SEQ ID NO:769)
MEVTVYILRCADGTYAGLTKQIEGRVWEHNEGIYDGYTKSRPVELMFTETDYDRILDALIRERQIKGWSRAKKEALIALDYALPDLARNRQGEAGAKRPR
GSTGSP
<ACH65_RS02795--Frateruia_defendens--79.35 (SEQ ID NO:770)
MHNRPQCVYILASRRNGTLYIGVTSDLPRVWQHRNDCVDGFTKRYRVHALVWFELHEAMASAI GREKALKAWRRRAWKIELIEKSNPYWRDLYADIRG
<AF337_RS37500--Rhizobium_ecuadorensis--69.15 (SEQ ID NO:771)
MKGYVYILASARNGTLYGTVTRDLAGRSYEHQNELTPGFTSKYGVKTLVWFEEHDLTTAAITREKTIKKWPRQWKLNLIERENPEWEDISFHLGL
<AF337_RS30505--Rhizobium_ecuadorensis--67.78 (SEQ ID NO:772)
MTGFVYIMSDKPRGVIYGTVTSDLHGRIWEHRNEIQKGFTQRYKAKNLVWFYHPNIVLAIQREKSLKRYLREWKIKLIEELNPTWMDLYERIDEIENVYRPH
PNTREWSDDYN
<RN70_RS11650--Staphylococcus_schleiferi--54.55 (SEQ ID NO:773)
MDKHYYIVQCKDASLYTGYTNNAARIAKHNEGKGAKYTKSRFPVQLRYETFTKSAALKRECAIKKMTROQKIQLIEE
<AKN83_RS15855--Bradyrhizobium_embraense--74.47 (SEQ ID NO:774)
MSYVYVYILASRRDGAIVGITNDLVRRYIEHRVKAVPSFTSKCNI TQLVWFEIYDDPISAITREKELKKWKRAWKIELIEATNPQWNDLYESICI
<AKN83_RS30580--Bradyrhizobium_embraense--76.19 (SEQ ID NO:775)
MAYYVDLLASRKHGTRYLGVTNDIVRRVHEHRTKAAPGFTSRYGVDKLVWFEIYDDAATAITREKELKKWRDDWKVRLIEEQNPVWVDRFPGISN
<AKL29_RS21780--Bradyrhizobium_pachyrhizii--78.48 (SEQ ID NO:776)
MGARSYVYVILASHIGGTLTYIGVTNDLIRRVGQHKLKLVEGFTEKYDIVRLVYFEQFDDPENAI RREKRLKKWRRAWKISLIEKDNPD
<AKL29_RS14585--Bradyrhizobium_pachyrhizii--62.35 (SEQ ID NO:777)
MSTAGSYVYVILASRHGTLTYIGVTNDIRTRLELHRSGRGSKFVYKYQVFRVLVHVEEFATPQEAIAREKQLKFWKRDWKIKLIEAENPDWNDLSGLI
<AKN84_RS21730--Bradyrhizobium_tropiciagri--75.27 (SEQ ID NO:778)
MKQPCVYIVANKRHGTIYGTVTANPPKRSFEHREGLVKGFSKAYGCKLLVWYELHETMYDAITREKQIKSGSRAKKIALIEALNPWNDLYDTLI
<AKN84_RS06580--Bradyrhizobium_tropiciagri--74.16 (SEQ ID NO:779)
MGTRSFYVYVILASHIGGTLTYIGVTNDLIRRVGEHKLKLIESFTEKYDVVKLVYFEQFDDPENAIKREKRLKKWPRAWKIALIERNDPDWIDLYPEIAGGGG
<AKN84_RS41710--Bradyrhizobium_tropiciagri--63.53 (SEQ ID NO:780)
MSTVGSYVYVILASRHGTLTYIGVTNDIRLRLHRSGRGSKFVYKYRVFRVLVHVEEFATPQEAIAREKQLKFWKRDWKIKLIERNDPDWNDLSGLI
<ATC1_RS01710--Flexilinea_flocculi--65.06 (SEQ ID NO:781)
MTNSHNTVLYIGVTNNIQRVIEHKS GKGSYFTSKYHLTKLVYFETSSDVYAAISREKQIKAGSRKKIDLINSINPEWKDLYDE
<TBC1_RS02970--Lentimicrobium_saccharophilum--65.12 (SEQ ID NO:782)
MKKGGYVYIMTNTNCSTLYIGVTNDLCRRIGEHKNHLI KNSFTDKYNLEHCYFEEFASIIDAIKREKEIKKWNQKKENLINSQNPDKVLVTEHGFKRENV
SFGRQMDDIVQELQAKGEIPPSSE
<J121_RS02575--Erythrobacter_citreus_LAMA_915--73.91 (SEQ ID NO:783)
MPFEREPVYVILASQPPGTLYIGVTSDLVKRLVQHQHREGVTGGFVARYKVHRLVRFEMFGDMEAAILREKKLKRWHRQWKINLVERDNPWHVDFDILGFDPLA
SSRAKDGSRNFRMTGA
<J121_RS04370--Erythrobacter_citreus_LAMA_915--69.15 (SEQ ID NO:784)
MSAWVYIMSNREDGVLYIGVTADLSRRIVQHREGKGSFACRRYGLTRLVYAEHDSIEDAIAREKAMKAWKRAWKIELIEGVHPAWDDLFDVVL
<J121_RS04580--Erythrobacter_citreus_LAMA_915--67.44 (SEQ ID NO:785)
MDSGTLGGWYIMTDYRGTYLYIGVTSDIATRVSQHREDIGSDFCRYGLHRLVWAEERFDIDEAITFEKRIKWRQWKFEKIEKTNPEWEDLFETLV
<J121_RS01450--Erythrobacter_citreus_LAMA_915--63.33 (SEQ ID NO:786)
MLGGWYVYIMTNSPHGTLTYIGVTADIAARAHAREGQGEFCCKHGLTRLVYAEYEDIDHDAIAREKAMKRWKRQWKRLIRRDNDPDWNDLFETIL
<J121_RS07270--Erythrobacter_citreus_LAMA_915--60.00 (SEQ ID NO:787)
MADRYGTIYVGTADIAARAVQHREGKGSFTHRYGLTRLVYLERDDITDAIAREKTDIAAREKAVKKWNRAWKIRAI EEQNPWHDLFDLTLM
<ALC15_RS08820--Bradyrhizobium_viridifuturi--60.71 (SEQ ID NO:788)
MKAASVYVYVILASRHGTLTYIGVTNDIHTRELHRSGRGSKFVYKYQVFRVLVHVEEFATPQEAIAREKQLKFWKRDWKIKLIEAENPDWSDRSGLI
<BN1969_RS03020--Clostridium_phoceensis--52.63 (SEQ ID NO:789)
MSWWVYLLRCGDGTLYGTITDLDLRRILAAHNAGRGAKYTRSRPVTLVWREBPDKSAALRREYAVKQLTRAKKLALIEGGPLV
<H663_RS40775--Limnolobus_planktonicus_II-D5--78.49 (SEQ ID NO:790)
MKQPYVYVILASARNGTLYIGVTSDLVARIWQHKNHVEFGFSQKYGVNQLVWFEAEHETMESAI VREKQLKSGSRSRKLSLIETNP SWRDLYGDIL
<NC99_RS00145--Sunxiuqinia_dokdonensis--71.76 (SEQ ID NO:791)
MTTFYVYVILANKPNGVLYVGMTNELERRIAEHKKGQIRGFTHKYNIDKLIVYFEETDYDEAFQRRERQMKKWKRDWKIEIENKNPEWKDLANNWY

FIG. 11Y

<NC99_RS09565--Sunxiuqinia_dokdonensis--64.44 (SEQ ID NO:792)
MHKGFVYILSNRNRRTFYIGVTSIDIKRRVLEHKTGKGSFTLKYNLTLDMYYEKIEGMDCEINREKQLKNWHRDWINLIREGNPEMKDLSAEWFIIQDGLV
IADRARAKAYFEKLNRRNSKSDPETSSG
<NC99_RS21655--Sunxiuqinia_dokdonensis--61.54 (SEQ ID NO:793)
MYVYVYIESLRDGSFYIGYTSNLEQRLHYHNTGKSRYSRKSPWEIVWFERAATRSDAIRREKFLKRQRNRKFYRRLIDDFEKLKS
<AM279_RS14655--Flavobacterium_akiainvivens--60.92 (SEQ ID NO:794)
MYKTFGTHNYVYMLTNKAKTVLYIGITNDLKERLYYHQNPPEPHSKAFTAKYKCKYLVIYEHYDDVNVDAIDRETQLKKWRRDKKEALIKGFNPEWKFLNDSI
<AM279_RS11220--Flavobacterium_akiainvivens--60.92 (SEQ ID NO:795)
MQHTFGSHNYVYIITNKIKTVLYIGFTTDLKERLYYHANPEPHSKAFTAKYKCYLIYFEHYQDVQIAIDRETQLKKWNRAKKENLINGFNPQWRFLNNEV
<SU32_RS07670--Ahrensia_marina--73.33 (SEQ ID NO:796)
MSGWVYIMASKRNGTIYIGVTSDLQGRAHEHRTAQSKGFTARYNCKDLVWYERHDNIVLAIEREKKLKKFKRAWKLELIEGFNPDWEDLYENCYVRDNPSEI
RARV
<SU32_RS06435--Ahrensia_marina--68.89 (SEQ ID NO:797)
MSGVYMTASKKGGTIYIGVTADLSQRIYDHKQMDTKSFTAKYGVTRLVWYEEHDDIRDAIQREKSLKRWRQWKIDLIESINPEWHELYKGMGW
<SU32_RS10695--Ahrensia_marina--70.59 (SEQ ID NO:798)
MKRKGYYIILASRKNGLTYIGVTSDLARRLLEHEHIGSGFTAKYNNVRLVWYEEYELVTTAITRETIKKWPRQWKINAIBEINPDLDNLKSWLL
<ADS78_RS05860--Idiomarina_abyssalis--55.84 (SEQ ID NO:799)
MAWSVYLKVCADNTFTYTGVTDDVERRVNEHNVDDKKGAKYTRARRPVELVWSEVHPNRSKACQREYEIKRWSRNRKAELTG
<ABB55_RS12025--Prosthecomicrobium_hirschii--76.67 (SEQ ID NO:800)
MQGRQHEYWYIILASRRNGTLYIGVTNDIVRRVWEHREGLVPGFTSRYGVKTLVWCEYRYVNDIAIAREKAIKSWPRQWKINLIEKDNPNWDDLFDLIL
<ABB55_RS05470--Prosthecomicrobium_hirschii--69.66 (SEQ ID NO:801)
MSAWVYILASRKQGTLYIGVATTDLACRMWEHQERINKGFTERYGVTRLVWSEFGTISEALEFEYRIKRWRRRAWKIQMI EKENPDWHDLSANLL
<ABB55_RS01960--Prosthecomicrobium_hirschii--61.80 (SEQ ID NO:802)
MNAWVYILASRPDGLFVGTAEDELVGRVYHHRARVPAGFSLGHAMQRLVWYAAFDRVEAAMAFARDLRRRHRHAKVRIEAMNPDWDDLYPEVT
<ADN01_RS03330--Levilinea_saccharolytica--64.10 (SEQ ID NO:803)
MTNYNNAVLYTGVNDLGRVLEHKSGRGGVFTCRYHIKKLVYEEAGCDIQTAISRDQKIKAGSRQKKVDLVNSMNPWKDLFEEGVGD
<AC812_RS16235--Bellilinea_caldifistulae--67.06 (SEQ ID NO:804)
MSKLPYFVYIILTNSYNTVLYTGVTDLIRINEHKNIGSKFTSRYRVHKLVIYVESTESVEGAI GREKQIKGWSRAKKVSLIESVNPWKDLSDEL
<I595_RS13795--Croceitalea_dokdonensis_DOKDO_023--62.20 (SEQ ID NO:805)
MSKEGWHNYVYIILTNKNRTVIYTGMTNNLHRRLEHQNAISLNSKSFVAKYKCKHLLYERYSWVQEAI SREKEIKGWLVRVKKVDLIKSTNPNLEFLEGSFL
Q
<JCM15579_RS09335--Marinifilum_fragile_JCM_15579--61.96 (SEQ ID NO:806)
MNTCYVYILSNKNKTVLVYGVTDLSRRLSEHINQVDNINAFTRRYNCHHLLIYFEQYSESIEDAIREKKIKGWRREKKEKLIESFNPDKFLEGDFFL
<JCM12296_RS11720--Desulfosarcina_cetonica_JCM_12296--84.04 (SEQ ID NO:807)
MQPAVYILSSKRNGLTYIGVTSNLVKRVWEHKNNIIAGFTKRYNVHQVLVWYBIHTTMEAAIQREKQLKNWKRKWKLDLIEKVNPAWQDLYGQIV
<LH19_RS00860--Sphingopyxis_macrogoltabida--71.91 (SEQ ID NO:808)
MHADFQPCAYIMASARHGTLTYGTSHLIQRIQCQHRDGSFGGFTDRYGVKTLVWFELAAATMHAATREKQIKNWQRKWKIELIEAANPRWRDLAEELGLPSLF
G
<LH19_RS12345--Sphingopyxis_macrogoltabida--71.43 (SEQ ID NO:809)
MERYPCVYILARSSHGTLTYGTSNLVQRIHQHREGMFKGYTKDYAIKRLVWFERHEMMSAILREKRIKRWLRPWKYDLINAVNPTWRDLAEDLGFDFPLPLN
PRKAGPGSSPG
<LH19_RS03110--Sphingopyxis_macrogoltabida--65.12 (SEQ ID NO:810)
MDQTYHYVYIILASRRNGTLYGTGDLAARCSQHRNGQGSQFTERYGVHRLVHAEAFDPVSEAIAREKAIAKKWRRRAWKLELIERDNPQWLDLYDRNV
<LH19_RS17775--Sphingopyxis_macrogoltabida--53.52 (SEQ ID NO:811)
MSFWAYMLHCRGGAFYTGHTDNLFPYIAQHEAGAI PGFTSDRLPIVLVWSQEFATRFEALEAERRIKGWSRAKKMALIRGDWDKISSLAKGKDGPPSTSSGQTE
<BN1313_RS12210--Phenylobacterium_immobille (ATCC_35973)--75.58 (SEQ ID NO:812)
MERSFYVYIILASQRNGTLYGTGDLSSRVWTHQQREGDGFARHGVTRLVWYEEFPTADEBAITAEKRIKRWRRRAWKIRRIEANPLWLDLYETFNS
<BN1313_RS13510--Phenylobacterium_immobille (ATCC_35973)--73.26 (SEQ ID NO:813)
MERSFYVYVILASQPNGTLYTRVTGDLSSRVWTHKQRESGGFTAQYGVTLVWYKAFPTADGAVTAEKRIKRCRRRAWKLQLIEKSNPQGRDLYETFNQQPRQPG
RSEAESRDPGFVHSP
<BVIR_RS07445--Blastochloris_viridis--81.11 (SEQ ID NO:814)
MRYAVYMLATRKDGPLYVGVTDNLVRRVFEHRSDAVRGFTWQFNVHRLVWFEAHDSEAAITREKQIKRWRRRAWKIEMIETENPEWDDLYERIAGGSSADDPG
SR
<BVIR_RS00445--Blastochloris_viridis--72.22 (SEQ ID NO:815)
MAGWVYILANRPNGTLYVGTADLLRRVWQHRTHTVGGFTVRYGVMLVYFEECATMPDAIQREKNIKHWPRAWKVRLIHSLNPDWDDLGYSL
<BVIR_RS15100--Blastochloris_viridis--64.49 (SEQ ID NO:816)
MEFFVYMLASKPNGTLYIGVTRDVIRRVHQHRTKAVPGFTAKYGVTRLVYVETFDGPESAIGAFSAKVGTFPAEENATNQSQSREKQLKGWNRAWKIRLIEQAN
PSWRDLFEELAG
<LA5095_RS15615--Labrenzia_alba--72.73 (SEQ ID NO:817)
MHVDQYVYVYILASGLGGTLYVGVANALLKRLYEHREGSGAFTAKYGVHRLVYFEIHEEIEQAIKREKQIKRWKRAWKISLIEERNPNWDDLGLSL
<RC62_RS11740--Flavobacterium_aquidurensis--65.82 (SEQ ID NO:818)
MLIYEGFHTYIYIITNKVKTLYTGTNSLSIRLKQHKNNILLGNKTFASRYNVEFLLYEKFETWIEQAIAREKEIKGWTRAKKIALIKVINPDNLFLDSL
F
<RC62_RS06230--Flavobacterium_aquidurensis--64.29 (SEQ ID NO:819)
MQLLEGYHTYIYIITNRSKTVFTYGTNNLEKTRLNQHKENIAGNKSFAKYKIEFLLYEKFETWIEQAIAREKEIKGWTRNKKIELIKTINLDDLFLNYLF
V
<BN2746_RS04990--Lagierella_massiliensis--76.67 (SEQ ID NO:820)
MKNYIYIMANKTNKTIYIGVTSNLIKRVWEHKNKLVDGFTKHYNVNKLVIYEEETDVSNAIAREKQLKGWNRAKKNLIDSNPNWDLIDYDDL
<AQ6619_RS13850--Caulobacter_henrici--75.56 (SEQ ID NO:821)
MPMAFHVYIMASRRNGTLYIGHTDDLARRVFEHQSRERGFTAKYGCTMLVWCELHETRESAQIRERQMKEWRRVWKLRLIEESNPDRDLAQDFLG

FIG. 11Z

<AQ619_RS03280--Caulobacter_henricii--67.02 (SEQ ID NO:822)
MMFHVYIMASRRNGTLYGTDDLGRRVSEHQSRERGFTGRYGTMLVWHEAHETREGTWRRERQIKEWRRVWKLQLEADNPEWRDLALEFLG
<AQ619_RS16595--Caulobacter_henricii--64.37 (SEQ ID NO:823)
MFSQNRSIGTMMVPEAWIAVYMMTDYRGTLTYGTGQFFTRIVQHREGQLRGFTKKYSLKQLVWYEHDLMTETAIQRETSIKRWPRQWKINLIERDNPRWDD
LYSAVFEWTFVPRRV
<AQ619_RS15275--Caulobacter_henricii--71.05 (SEQ ID NO:824)
MSTYAANGRRHLHPSQSPRLTLYIGVTNDLSRRRIWEHREGVGSQFCRKYGVHRLVYEEHEDISFAIHREKRLKKYHRQWKINLIEGFPNTWTDLYETWNC
<ASC46_RS12205--Chelatococcus_sambhunathii--72.94 (SEQ ID NO:825)
MDNYTVYILASGRNGTLYVGLTNLARRIFEHKQKTVPGFTRRYGVDRLVWYEHYRSIAEARQRETSKRWRRRAWKLSLVEAMNPAWDDLNDRLNA
<ASC46_RS09540--Chelatococcus_sambhunathii--72.09 (SEQ ID NO:826)
MRGGWVYIMTNRPNGLTYIGVTSNLRRAWEHHEMIDGFTKRYGLKQLVYEEAYEDIRSAIQRESNMKHWPRRWKIRLILEMNPQWRDLYETLA
<ASC44_RS15470--Lacinutrix_mariniflava--68.48 (SEQ ID NO:827)
MKPGYIYILTNKNNTLYVGVTAYLKERILQHKEKHDLKSFTARYNLDKLVYEAHQMGDAIAREKQLKGGSAQKTGLIESINPEWLDLFEEL
<ASC41_RS13135--Lacinutrix_algicola--76.92 (SEQ ID NO:828)
MFKTQHQQYYIILSNKYNGTLYIGVTNNLERRMFEHKKKIEGFSKRYDLTNLVPHFETFYINDAIKREKNMKKWKRAWKIKLIQEDNPNWSDLAGWDVYE
<ASC41_RS01670--Lacinutrix_algicola--63.64 (SEQ ID NO:829)
MKKSHIYFLTNKNNTVIYIGVTSHLIKRIYQHKTKVYKGFASRYNCDKLVYFEEFDDINLAIIREKQLKAGNRKRKEGLINLENPNWDLGDGWLFSF
<ASC48_RS07250--Pannonibacter_indicus--71.76 (SEQ ID NO:830)
MRCFWVYILAGRKNGTLYIGVTNDLTRLRLEHQSGTGSRTARYSVQRLVWYEPHGTAEBAIAREKAIKSWPRAWKVELIATNPEWYDLGRFLNA
<ASC48_RS15915--Pannonibacter_indicus--69.57 (SEQ ID NO:831)
MPAWVYIMASKPYETLYTGVTAVLARVYEHREGQIEGFTERYGIKWLWVYQEVATMHEAIAEMEKKIKRWRRQWKFDLIEKGNPQWRDLYPDNL
<ASC48_RS00965--Pannonibacter_indicus--65.17 (SEQ ID NO:832)
MPFHVYITVSGMNGTLYIGVTNDLARRIHEHRTICGSFCKHSAVLLVYEFTEFDPETAI SADKRIKLWQRSWKIQLIERFNPGWKDLYLTNLN
<APR42_RS00485--Salegentibacter_mishustinae--62.92 (SEQ ID NO:833)
MKNSYIYILSNSTRIMIYIGVTSDLKKRI TEHRSDMGSLFARKYHLKFLIYFEEFSKIHQAIAREKQLKNWHKEWKNLVKSMNPDLKDYEDL
<APR42_RS06575--Salegentibacter_mishustinae--58.14 (SEQ ID NO:834)
MRVSFYIILECADNTYTGVTSNLEKRMIEHQGGAPHERYTAKRRLPLKLVFAEFTDINFAIIREKQIKKWSRAKKKALINGRFEDLPNLAKKNFS
<ABB29_RS06310--Pseudoxanthomonas_dokdonensis--80.65 (SEQ ID NO:835)
MKQPCVYLLASGFHGTLYLGVTTDLIRRIDQHKRNLVAGFTRRYHVHTLVWYEQHPPTLDSAI PREKALKQWKRAWKIELIEQNNPRWDLPLPGLL
<AST99_RS04780--Formosa_algae--67.06 (SEQ ID NO:836)
MDVYVYIILTNKNHTVLYIGRTKQLKVRILQHKNNSIKTFTGRYNVHKLVYFETTKYINNAIKRERQLKKNRKKWKNINLIKGLNPDWDLSEYI
<AST99_RS17880--Formosa_algae--59.77 (SEQ ID NO:837)
MKTGKNHNYVYIVTNKNKTVLYTGVTGNLKERLYIHNNPNPCSKAFTAKYNCVFLVYEEHFSCIETAIKREKQIKGYSRMKKDKLISAFNPNWHFLNDMI
<AOQ71_RS29780--Bradyrhizobium_manausense--75.53 (SEQ ID NO:838)
MAYYVYILASRRDGAIVGVTDNLVRRVYEHRIKAVPGFTKKYNI TQLVWFETYDDPISAI SREKELKKWKRWKTQLIEKDNPNWNDLYESICK
<CQ12_RS32690--Bradyrhizobium_jicamiae--67.06 (SEQ ID NO:839)
MYVYVYILASRRHGTLYIGVTNSLARRMEQHRSGDGSSFVKYGVYRLVYVEAFERAEEAIAREKQLKRWKRDWKIELIERENLEWRDLSDLIV
<CQ12_RS41070--Bradyrhizobium_jicamiae--67.06 (SEQ ID NO:840)
MYVYVYILASRRHGTLYIGVTNSLSKRMEQHRNNEGSSFVKYGVYRLVYVEAFERPDIAIAREKQLKRWKRDWKIELIERENLEWRDLSDLIV
<CP49_RS37300--Bradyrhizobium_valentinum--77.66 (SEQ ID NO:841)
MSYVYVYILASRRDGAIVGVSKDIVRRIYEHRTKAVPGFTSKYNI TRLVWFETIYDDPISAI SREKELKKWKRWSKVQLIEKDNPDWNDLYESICS
<CP49_RS22800--Bradyrhizobium_valentinum--74.44 (SEQ ID NO:842)
MVPRFAALYLLASDKYGTLYLGVTDNIVRRVYEHKSKAVPGFTKRYAVDKPVWFETIYDDPVTATREKELKKWRDWDKIRLIEERNPQWIDLYPQISV
<CP49_RS23320--Bradyrhizobium_valentinum--77.53 (SEQ ID NO:843)
MGARSYVYILASRIGGTLYIGVTNDLVRRAGEHRLEPMKGFTKKYGVHKLIVYFEQFDDIENAI SREKRLKKWNRWVKVRLIEELNPNWDDLYPSIAGA
<CP49_RS00480--Bradyrhizobium_valentinum--67.06 (SEQ ID NO:844)
MYVYVYILASRRHGTLYIGVTNSLAKRMEQHRNNEGSSFVKYGVYRLVYVEAFERPDIAIAREKQLKRWKRDWKIELIERENLEWRDLSDLIV
<CQ13_RS05150--Bradyrhizobium_retamae--75.28 (SEQ ID NO:845)
MGTRSYYVYILASRIGGTLYIGVTNDLVRRVGEHKLKIAEGFTKKYEVSRIVYFEAFDQIEFAIQREKRLKKWPRAWKISLIEKQNPDWIDLYPEIAGGGG
<AR323_RS00515--Altererythrobacter_troitsensis--76.34 (SEQ ID NO:846)
MHPDFQPTVYLMASARNGTLYIGVTSNLLQ RVAQHREGTFGGFTTKYAVKCLVWFQHSMEHAIVREKQLKKNWNRWKLDIKKANPEWRDLAEDFGFEPLR
<XM53_RS05270--Roseovarius_atlanticus--64.71 (SEQ ID NO:847)
MAFYTYIMACRSNTAIYVGTGNLDRDRAAQHRLGAGSMQTSKYNIAKLIVYFETHETLDAAIARERKLKRWRREWKDLIEHANPSWSDLSMEASFL
<ASU31_RS14470--Pedobacter_ginsenosidimutans--66.28 (SEQ ID NO:848)
MELGGWIYIMTNYEKTLYIGVTSDLRSRVYDHQNKIYPDSFTAKYNLIYCYVFEDFQSI EAINREKQLKKWNRSKKETLINNFKNKNWDLTEVINEWDE
<LC55x_RS12560--Lysobacter_capsici--85.87 (SEQ ID NO:849)
MPAKAEVSKTGQTRNFQVVI PAKAGIQSEADVSKSPAVYILASARHGTLYIGVTSDLIQRVWQHREHLAKGFTERYEVDRVLVWHEHETMESAILREKRLKKW
NREWKIRLIEETNPQWRDLWLDIVAVR
<LG3211_RS12065--Lysobacter_gummosus--83.70 (SEQ ID NO:850)
MARSPAVYILASAKRGTLTYIGVTSDLVRRIQHREHAADGFATRYNVDRLVWYEMHDTMESAILREKRLKKWNREWKIRQIEDTNPTRWDLWPDIASTF
<L21SP5_RS06145--Salinivirga_cyanobacteriivorans--53.75 (SEQ ID NO:851)
MKGYMYILKCANGQYYTGSTKFLDKRIEQHCQCGKANFTRKYLPVELVYEEYDRIDEAFYREKQVQGWGRKKKEALIEGRYSDIPGLAKKHFG
<SD51_RS01620--Alicyclobacillus_tengchongensis--57.33 (SEQ ID NO:852)
MNQDVCRPTWHDHVRPGHFVYIVECSDGTLYTGATDVLRRIAMHNEGRGAKYTRGRGPVRLLYVESYPDKGAALRREKSIKRLTRAQKRLRVATYQAPRGTA
<VO84_RS18500--Terracidiphilus_gabretensis--74.12 (SEQ ID NO:853)
MPANFYFTYIVASKRSTLYIGVTNNLQRIQFHKKQKTFPGFTAKYNCNTRLVWFERIYVRGDAIYREKELKGWLRRAKLLALIAENNPTEWDLSEFPWPHLIANN
SSPEQNRPLSHKSPSS
<VO84_RS16565--Terracidiphilus_gabretensis--68.00 (SEQ ID NO:854)
MPPSYFTYIMASRSHTLTYIGVTSDLQKRVFQHKWREHGDFTGRYNCNRLIWFETHQDVVAIAIDREKQLKGWRREKKIALIEKTNPAWIDLSDWYDLEPADHS
RSLDRMNT

FIG. 11AA

43/106

<VO84_RS15340--Terracidiphilus_gabretensis--69.57 (SEQ ID NO:855)
MAERIYFTYIMASRSHTLYIGMTGNLRLRVFQHKWHEKDGFRTARYNCDRLVWFESFSDVSLAIQREKELKGWRREKKVALIEQANVTWQDLSRGWYDVEPADS
KRALDRMDA
<VO84_RS20395--Terracidiphilus_gabretensis--68.48 (SEQ ID NO:856)
MLERTYFTYIMASRSHTLYIGMTNNLRHRVFQHKWREQDGFRTARYNCDRLVWFQSFSDVSLAIQREKELKGWRREKKIALIEQTNPTWQDLSREWYEIEPVDL
KRALG
<ATY38_RS06830--Nitrosomonas_ureae--78.26 (SEQ ID NO:857)
MKQSPVYILASRGYNGTLTIGVTGHLIQRIVQHKQDLVEGFTKKYSAHFLVYEMHEEMQAAIQREKQLKKWNRQWKINLIEKMNPAAWKDLWEELGKR
<BC008_RS43550--Mastigocoleus_testarum_BC008--69.66 (SEQ ID NO:858)
MSEQYIYILMTNKYNTVLYTGVTNDLIKRIYEHKEKLIAGFTQKYNVNKLIVYIEIFTDINSAIAREKQIKAGSRQKKIDLVSNNPEWRDLYDGLL
<SHM7688_RS19040--Shimia_marina--65.17 (SEQ ID NO:859)
MSHFVYILASRPGGALYVGRRTDLRLRLQAHRAGLSTHTHKYNIKTIVWFEEHESFEPSSLRERAIKWRRRDWIKLITDANPNWQDITAFVPA
<BN3124_RS04060--Legionella_saoudiensis--71.59 (SEQ ID NO:860)
MKEWNPVYIMANKRNGTLYTGVTNNLIQRVYQHKHGLQPGFTKQYECCKLVYIYADMLLAIAREKQIKAGSRKKKLALIESVNPRWIDLPESLM
<BN3124_RS15085--Legionella_saoudiensis--74.44 (SEQ ID NO:861)
MHTESYVYFLNTYNNVLYVGVTTNNLVRVYEHKNRLMKGFTHKYNVNRLVYIEVCPDIIIAIKREKQINGWSREKKNLIIITMNPWKDLYESLF
<DEALK_RS05270--Dehalogenimonas_alkenigenens--78.89 (SEQ ID NO:862)
MHEYYVYIMASKRHGTLTYIGMTNDLVRVVCQHRNDLIEGFTKRYGVHLLVHYEQTTDVAIAIWEKRLKKWNRWAKIELIETQNPWHDLYPELVS
<AT746_RS00545--Legionella_alkaliphilum--80.43 (SEQ ID NO:863)
MKQPCVYILGSRNGTLYIGVTSNLVQRVWQHKTEQSEGFTSQYHVHKLIIYEVFEDMYHAITREKQLKKWNRWAKIRLIESINPYWRDLWEVEQTSN
<Lbir_RS00005--Legionella_birminghamensis--76.34 (SEQ ID NO:864)
MRERQPAIYIMANKRNGTLYTGVTSDLLKRVYEHRYAELNGFTKKYNCKILVYFELTDSMTAISREKQLKSGSRKKKLLIERFNPSWSDLYETLI
<Lwor_RS02415--Legionella_worsleiensis--74.44 (SEQ ID NO:865)
MNVFYVYIMASKRNGTLYVGSMSDVIKRVWEHKNKLI SGFTARYTPVPLVYEMHSTYLEAARREKRFKNWCRQWKINLIEDLNPEWRDLYEEICS
<Lwor_RS12565--Legionella_worsleiensis--68.60 (SEQ ID NO:866)
MRTESYVYLLANKHNNVLYTGVTSDLIRRVYEHKNKLIVAGFTQKYNVDRVLYFEVCSGIVMAIEREKQIKGWSRKKKQNLINSLNPEWDDLYSSL
<Lwor_RS04215--Legionella_worsleiensis--89.47 (SEQ ID NO:867)
MNVFYVYIMASKRNGTLYVGSMSDLIKRVWEHKNKLI PGFTARYNVHMLVYEMHSTYLEA
<Lwor_RS06975--Legionella_worsleiensis--64.63 (SEQ ID NO:868)
MTVLTNGTLYVGATTDLNKRTWEHKNKFI PGFMAKYNLHMLVYIEIDELYEAARREKRFKNWPRQWKINLIEKINSERCDLYEEICQ
<Lcin_RS09440--Legionella_cincinnatiensis--77.78 (SEQ ID NO:869)
MEKYSYVYIVTNQKNGTLYIGITTDLVKRVWEHKNKFVESFTKKYNCTRLVYIEQHVDVLEAIKREKRLKKYLRAWKIELIESINPSWRDLYEDIC
<Lcin_RS17715--Legionella_cincinnatiensis--79.52 (SEQ ID NO:870)
MLASKYGTLYTGVTSNFVQRIWQHKGLAEGFTKKYHVHHLVYIEHTDVYEAITREKRIKKWNRQWKINLIEQNNPHWLDLDIKLF
<Lcin_RS09420--Legionella_cincinnatiensis--59.21 (SEQ ID NO:871)
MEFWYILKCRDQSYIIGTNDLNDLRFQHQNKLI PGCYTSRKLPVQLVYSEDFPTREEALAAERQIKGWSRRKKEAYIRGDWTALSEYSKRTSNKKA
<Lery_RS13305--Legionella_erythra--73.33 (SEQ ID NO:872)
MGERHYFVYILSSKSYGTLTYIGVTSNLTQRINQHKNGWVGDFTKKYCVHRLVYIEIHADIYEAITREKRIKKWNRQWKINLIEKNPQWLDLSIGLV
<Lgor_RS10885--Fluoribacter_gormanii--79.55 (SEQ ID NO:873)
MEEKQYVYVNLASKYGTLYVGVTSNLAQRIYQHQKGLQGFTHKYNVHLVYIEIHSVDYEAITREKQIKKWNQWKINLIEQNNPQWLNLNKELF
<Lgor_RS08470--Fluoribacter_gormanii--74.44 (SEQ ID NO:874)
MKAFYVYIMASQRNGTLYVGVTSNLIKRVWEHKNKVI SGFTAQYNVHMLVYIEHTSTYIEAARREKRFKNWCRQWKINLIEQINPEWFDLYEEICS
<Lgor_RS09975--Fluoribacter_gormanii--69.23 (SEQ ID NO:875)
MHKESIYVYLLANKHNNVLYTGVTNDLIRRVYEHKNKLIVAGFTKKYNVDRVLYIEACRCIIIVAIEREKQIKGWSRKKKHDNLINALN
<Lgor_RS14195--Fluoribacter_gormanii--70.67 (SEQ ID NO:876)
MHKESIYVYLLANKHNNVLYTGVTNDLIRRVYEHKNKLIVAGFTKKYNVDRVLYIEACRCIIIVAIEREKQIKGWSRKKKHDNLINALNPE
<Lgra_RS03850--Legionella_gratiana--77.78 (SEQ ID NO:877)
MEKYSYVYIVTNQKDGTLTYIGVTNLVKRIWEHKNKFVDGFTKKYNCRLVYIEQHFDVLEAIKREKRLKKYLRAWKIELIESVNPTWRDLYQDIC
<Ljam_RS04585--Legionella_jamestowniensis--78.82 (SEQ ID NO:878)
MEEKRYFVYILASKYVYALYTGVTSNLVQRVYQHKGLAEGFTKRYNVHRLVYIEIHPDVYEAITREKRIKKWNRQWKINLIEQNNPQRLDLSIGLF
<Ljam_RS02140--Legionella_jamestowniensis--77.78 (SEQ ID NO:879)
MKESYVYILASKYNGTLYIGLTSNLIKRVWEHKNKFVPGFTQKYNVNTNLVYIEFMNDISLSAAREKRLKEWKQWKIDLIQKNPEWKDLYHEIVQ
<Ljam_RS04750--Legionella_jamestowniensis--66.29 (SEQ ID NO:880)
MHTESYVYLLTNKHHNVLYTGVTNNLIRRIYEHKNKLVRGFTQKYNVDRVLYIEACENIVVAIGREKQIKGWSRKKKDDLISTFNPQWEDLYRSLC
<Llon_RS11280--Legionella_londiniensis--77.33 (SEQ ID NO:881)
MEEMQYFVYILASKAYGTLYTGITSNLVQRIYQHKGLAEGFTKRYSVHRLVYIEIHSVDYEAITREKRIKKWNRQWKINLIEQNNPQST
<Llon_RS11315--Legionella_londiniensis--80.00 (SEQ ID NO:882)
AYGTLTYGITSNLVQRIYQHKGLAEGFTKRYSVHRLVYIEIHSVDYEAITREKRIKKWNRQWKINLIEQNNPQST
<Lmac_RS11990--Legionella_maceachernii--72.22 (SEQ ID NO:883)
MRSAYVYILASRPNGLTYVGVTSNLIKRTWEHKNKVI PGFTAQYNVHMLVYIEAHELYVEAARREKRFKNWPRQWKINLIEKLNPEWRDLYEEICQ
<Lmac_RS03510--Legionella_maceachernii--67.44 (SEQ ID NO:884)
MPSLWIRGKVLHARKSHEGILHLSINKHKNVLYTGITNNLIRRVYEHKHLQQGFTHKYNVDRVLYIEIYSGIEEAITREKQIKGWARHKEALINTINPLW
VDLYESLW
<Lqua_RS02680--Legionella_quateirensis--74.44 (SEQ ID NO:885)
MEQYSYVYIVTNKKNGTLYVGITSNLIKRVWEHKNKATEGFTSKHGCTRLVYIEQHTDILEAIKREKRLKKYLRAWKINLIESINPKWDDLYLDICR
<Lqua_RS16695--Legionella_quateirensis--70.97 (SEQ ID NO:886)
MPTKQPAIYIMANKRNGTLYTGVTNLIKRVYEHKYADEPGFTQQNGCKYLVYIEIEMDTSAIAREKQLKGGSRKKKLALIEQINPYWEDLYEQLI
<Lqua_RS09085--Legionella_quateirensis--67.74 (SEQ ID NO:887)
MQNQSYMYFVTNAYNTVLYVGVTTNDLIRRIYEHKNKLIKGFTQKYNVDRVLYIEIFDTIVLAIEREKQIKGWSRNKKNELVNAFNSDWDNDLYQSIL
<Lqui_RS14695--Legionella_quinlivanii--73.33 (SEQ ID NO:888)
MKSFYVYIMTSKRNSTLYTGSTSDLIRRVWEHKNQVI PGFTAKYNVHLLVYIEHMDTYEAARRERRFKKWPRQWKINLIEKMNPWDADLYEEICS

FIG. 11BB

<Lqui_RS04125--Legionella_quinlivanii--75.27 (SEQ ID NO:889)
MREYQPAIYIMANKNGITYTGVTSDDLKRVYEHKFAEVEGFTKRYSCSKSLVYYELAEDITSAISLEKQLKGGSRQKKIKLIEKNPLWEHLYPKLV
<Lste_RS02390--Legionella_steelei--54.65 (SEQ ID NO:890)
MEWVYILKCKDNSYIIGHTDNLDNRLFQHQNKLI PGCYTSTKLPIQLVYTQDFPTREEALAAERQLKGWSRRKKEAFINGDWTALSEYSKRLIK
<Lsan_RS17380--Legionella_santicrucis--68.42 (SEQ ID NO:891)
GSTSDI IKRAWEHKNKVI PGFTAKYNVHMLVYYEVHKLIEAARREKRFKNWPRQWKINLIEGLNPEWRDLYEIEICQ
<Lsan_RS21315--Legionella_santicrucis--86.79 (SEQ ID NO:892)
MRPSYVYILASQRNGTLYVGSTSEI IKRAWEHKNKVI PGFTAKYNVHMLVYYE
<Lsan_RS17385--Legionella_santicrucis--71.70 (SEQ ID NO:893)
MGSTSDI IKRAWEHKNKVI PGFTAKYNVHMLVYYEVHKLIEAARREKRFKNWPMI
<NS334_RS05955--Sphingomonas_endophytica--80.68 (SEQ ID NO:894)
MERQPCVYIILASRPHGTLYIGVTSNLLARLI QHRDGTGFGFTDRYDI TWLVWYDVADTMEAAI HREKQVKKWNRDWKLRLIEETNPQWLDLAARI GLDSAHDG
<NS334_RS09205--Sphingomonas_endophytica--75.90 (SEQ ID NO:895)
MANRRNGTLYVGVTSNLAQRVWQHREGVADGFTQRYGCKMLVWFEVGDMEVAIVREKQIKGGSRRARKVSLIEAMDPGWNDLFG
<NS334_RS08670--Sphingomonas_endophytica--73.03 (SEQ ID NO:896)
MIGAAVYVILMASRRNGTLYMGSTDLPRRVWEHRNAVVEGFTKTHGCHLLVWYEAHESWEAARQRELRMKEWRRSWKLQEI EBLNPDWDDLYDRIALP
<NS334_RS08795--Sphingomonas_endophytica--65.17 (SEQ ID NO:897)
MPLMAWVYIMTNAPSGTLYIGVTAHLAARIQQHREGSGDFCKEHGLTRLVYVESHDITMLEAITREKALKKWKRAWKLNLI GSVNPDWRDLWADINM
<NS355_RS01025--Solirubrum_puertoriconensis--73.03 (SEQ ID NO:898)
MERPGYVYIMASDRNGTLYLGVTSNLAQRVLQHRDGVVAGFTKRKHCHRLVWFEAHSIESARQRELRMKEWKRAWKLEIEGLNPDWDDLYDRIALP
<AVJ23_RS14500--Ponticoccus_marisrubri--73.91 (SEQ ID NO:899)
MPARPVYIMTNPRGTLYVGVTGNLVRVWEHRSHAMPGFTDRYNLEHLVWFEAHDPRGTAIQREKSLKRWRRQWKIELVEAHNPAWRDLWAQIAGP
<ASU33_RS01305--Solirubrum_puertoriconensis--60.00 (SEQ ID NO:900)
MYVYIILTNPTQTVLYIGVTNDLARRLYEHSQGRGDANKFTGRYQADLLVYFEQCPDAKQAIAREKQLKGWTRQKKDALIAAFNPWTQAI DLETWDG
<ASU33_RS19455--Solirubrum_puertoriconensis--75.47 (SEQ ID NO:901)
MLYVGVTNDLGRMAEHRSAGHKGFTKRYNVTKLIYFEEHTDINAAIAREKQLKAGRCGQPEPKLGRAAALGRFRLVLGER
<APZ00_RS01305--Pannonibacter_phragmitetus--72.94 (SEQ ID NO:902)
MRCFWYIILAGRKNGTVYTGVTNDLIRRLHEHQSGTGSRTARYSVQRLVWYEPHGTAEADAIAREKAIKSWPRAWKVELIEATNPEWYDLGRFLNA
<APZ00_RS19090--Pannonibacter_phragmitetus--70.00 (SEQ ID NO:903)
MAGYVYIMSDKPRGVIYTGVTSDLEKRIWEHREGTRKGFTSQYKARNLVWFEIHPNIVLAIQREKSLKRYLREWIKIRLIEGLNPTWLDLYEKMGEIENVYNPH
ENTQMWNNDYN
<BN3265_RS08600--Anaerococcus_rubeinfantis--70.45 (SEQ ID NO:904)
MADYVYIILANKSNSTIYIGYTYNLEKRLYEHRNHVVEGFTDKYNCTKL VYEECLDDKYEALSREEQLKRWRSREKKNRLINSINPKWRDL SLDWYR
<AR686_RS09230--Chryseobacterium_greenlandense--64.47 (SEQ ID NO:905)
MMEFTEGVYTFYVYIILTNKNRTVLYTGVTNNLHKRLQHQHKTRSDPKSFTAKYNVEFLVYEEKFGWI HEAIEREKEIKNLSRIKKLELIRIQNPNDVLDLDFQ
LEF
<AUC43_RS20070--Hymenobacter_sedentarius--69.41 (SEQ ID NO:906)
MLRTYVYIILTNANHTVLYTGVTNDLARRVAEHKAGVHEGFTKQYNNVRLVHFETYPDINAAIAREKQLKAGSRARKLALINATNP TWKELAG
<AUC43_RS08470--Hymenobacter_sedentarius--60.00 (SEQ ID NO:907)
MVVYIILTNTRGTVLYIGVTNDLPRMYEHSTGRGDASKFTGRYQADLLVYFELLDPATQAIAREKVLKGWSRAKKDALVLA FNPTWQAI DLESWA AEATKR
<AUC43_RS20105--Hymenobacter_sedentarius--58.62 (SEQ ID NO:908)
MGFYVYIILTNPARTVLYTGVTNNLERRLHEHSMGLGEAGKFTGRYQCNLLVYFEVSPDARQAIAREKQIKNWARGKKEFLISTINPAWEPIDLQTWTG
<TAGGR_RS01620--Thermodesulfovibrio_aggregans--70.00 (SEQ ID NO:909)
MSKKGYVYIMTNPANTVLYTGVTSDLKKRVYEHKNKLADGFTKKYNNVVKLVYFEVYENI TDAIMREKQIKGGSRRKKIELINKFNPQWKDLYDDIL
<AWU67_RS05585--Microterricola_viridarii--60.29 (SEQ ID NO:910)
MYIVECSDGSYVGSADVEQVRVWQHNEGGAAYTRRRQPVTLVYFEEYDRIDEAYGREKQVQVRSRAKRRALIDGVGAELL PGLARKRWGIEHTSPDAHPLV
E
<Lupro_RS02260--Lutibacter_profundi--67.44 (SEQ ID NO:911)
MKENFIYIILSNKNRKVLYIGVTSNLLKRIEEHKNGIGSIFTKYNNVHDLVYEEFNNDINLAIKREKQLKNWNKAWENLIKETNPNDLIDYDQL
<Lupro_RS03755--Lutibacter_profundi--60.23 (SEQ ID NO:912)
MIEFTEGIYSFYVYIILTNTRYTTFYIGVTNNLNRRLQEHKNTTESTKKTFAAKYNLKYLVYQKFGWVQQAIAAREKELKGWRREKKNLIRAFNP TFKSLNYL
FE
<Maq22A_RS21640--Methylobacterium_aquaticum--75.56 (SEQ ID NO:913)
MAYFVYIMMASRRHGTLYIGVTNNLARRVYEHRTKAQDGTARYDVARLVWYEAYESINEAIAREKDLKRWRRDWKLKLVEDFNPGWSDLYSTLAGF
<KORM41_RS10815--Kordiimonas_lipolytica--76.40 (SEQ ID NO:914)
MDKQGYTYIILASKKHGTLYVGVTSDLVKRVWEHKEDHADGFTKKYGVHALVHFEAFDDINS AIEREKQLKRWHRQWKINLIEEHNPDWKDLYPSLTA
<KORM41_RS13130--Kordiimonas_lipolytica--72.09 (SEQ ID NO:915)
MRSWGNAALTYVYIILTNRPHGTLYIGVTNNNARRITEHRDGKGSRF AHKYNLRLVYVEPYDTVEEAIAREKAMKEWQRAWKVR LIEEHNPEWDDLYGR LNW
<AT959_RS17565--Dechloromonas_denitrificans--80.65 (SEQ ID NO:916)
MAKHPAVYIILASQRNGTLYIGVTSDLVKRIWEHKNLAEGFTKEYGCHTLVYFEQHAEMSAAI LREKQMKKNRAWKLALIEKENPTWRDLWPEICA
<ATB72_RS22590--Vibrio_coralliirubri--72.04 (SEQ ID NO:917)
MKQPAIYIILANRSNSVLYIGVTGNLQRTWQHKAQDLEGFTKKYNNVVKLVYFEVDFRAAIEREKQLKKNWRSWKEMLI SELNPSRSDLYDELL
<AL538_RS11505--Vibrio_harveyi--71.26 (SEQ ID NO:918)
MDKVPVYIILTSKNSVLYIGVTSQLAQRIWQHKNQVVEGFTKKYHVHKL VYFELFEDMATAINREKQLKQWKREWKDLVRKTNP DFLDLSNEF
<SGRAN_RS12400--Sphingopyxis_granuli--77.42 (SEQ ID NO:919)
MREPPQPAVYIMASRRNGTLYTGVTSNLAQRVWQHREGLGGFSGRYDCGMLVWFEI HANMEFAIAREKQLKAGSRAKKLALIEAENPLWRDLWPDILDGSE
<SGRAN_RS20495--Sphingopyxis_granuli--77.27 (SEQ ID NO:920)
MRERNPCVYIMASGRHGTIYIGATSDLMARVYQHREGLIDGSTKRYGVKRVLYFENMHETMDAAIVREKRLKAWKRDWKIALIETDNPFWEDLAIHLGFPPPT
<SGRAN_RS11485--Sphingopyxis_granuli--77.53 (SEQ ID NO:921)
MYVYVIMASQKNGTIYIGVTNDIIRRTYEHREGQLTGFSQKYRVKHLVHYEEFQDIRDAIQREKRLKKNRAWKIQLIETQNPDRDLWFDLNR

FIG. 11CC

<SGRAN_RS04685--Sphingopyxis_granuli--76.47 (SEQ ID NO:922)
MVPFPGVYIMASRNGTIYIGVTSDDLPRVWQHREGMIEGFTKRYGCKLLVWFDCEFIHDAQLEVRMCKKNRAWKINRIEHNPEWNDLSETLN
<SGRAN_RS06165--Sphingopyxis_granuli--70.97 (SEQ ID NO:923)
MRDVFQPCIIYILANHKGGTLYIGVSSNLAQRLWQHHSVGTGKFAAKYRTYRLVHFELFGTMELAIAREKQLKNWHRQWKINLIDAGNPEWRDLAFEVGAESIG
QTMDETSSA
<SGRAN_RS09145--Sphingopyxis_granuli--72.83 (SEQ ID NO:924)
MHPDAQFVYIMANRKDGLVLTGVTSDLIGRTWQHQRKASGFSARYRCDKLVWFEIYDDMLSAIAREKQIKAGSRARKIALIEAKNPGWRDLFPFELAGG
<SGRAN_RS15820--Sphingopyxis_granuli--56.76 (SEQ ID NO:925)
MAFWSYILLCADGRYTGHTDELERQIAQHQSSEIEGFTSSRLPVRLMWSQDFPTRTEALETELRIKKWSRAKKEALIRGDWAAISHFAKPPKERSTAPSSVR
VERSRDTNGNESEGLS
<ALA01S_RS32095--Azohydromonas_lata_NBRC_102462--84.78 (SEQ ID NO:926)
MNPTGQFVYILASQRHGTLTYIGVTSNLIQRSWVHREEIEGFSQRYGTHLLVYFELHATMQAAITREKQIKKWHRAWKVRLEAENPNWDDLWPAITGESSG
G
<CAter10_RS18145--Collimonas_arenae--75.53 (SEQ ID NO:927)
MASYVVMVSNKLRGTLTYIGVTSDLIKRTWQKHEAFVDGCEKHLKSLVWYIEHEDVTEAIKREKQLKKWNRNWKIHLIEQSNPNWQDLYPGLLN
<AZF00_RS01030--Zhongshania_aliphaticivorans--77.42 (SEQ ID NO:928)
MKQPCVYILASKPHGTLTYIGVTSNIVARICQHKADLVTFGSQRYSVHRLVYIELHQEMYAAITREKQLKRWKREWKIRLIEEKNPQWQDLYPELF
<NL2_RS09570--Novosphingobium_lentum_NBRC_107847--71.59 (SEQ ID NO:929)
MSRADRQPCPYILASRPYGTLYIGVTSDLVRLWQHRNGALPGFSDRYAIYRLVRYELFGDMDLAIAREKQLKNWHRPWKITLIQTDPNDWHDLPGLGLPL
APRGPRNGL
<NL2_RS01775--Novosphingobium_lentum_NBRC_107847--64.84 (SEQ ID NO:930)
MRQERVPCSYILASGTGYTLTYIGVTSNLPARMWQHRTGALPGFSSRNRTYCLVHFEFHATMEHAIVREKQLKNWHRPWKINLINEGNPHWADLAMGLGLPLA
PGKPRNGS
<NL2_RS02180--Novosphingobium_lentum_NBRC_107847--68.60 (SEQ ID NO:931)
MAKGGVYIMANRYRGGGLVVGVTSDLIRRVHQRHREGSGSAHVADFCKTMLVHAERHDTIEAAIAREKLVKKWRDWDKWFALIEAVNPDPWRDLWEWFAADSGQG
<NL2_RS01915--Novosphingobium_lentum_NBRC_107847--64.04 (SEQ ID NO:932)
MGGWVYIMTNGAHTLYIGMTAHLPARIGQHRAGRGSEFCAEYGLTRLVYVERHDRIEDAIAREKAMKAWKRQWKLRLIRHNPDPWRDLWEDINGL
<NL2_RS09490--Novosphingobium_lentum_NBRC_107847--62.22 (SEQ ID NO:933)
MKPGYVYLLINKPNGVLYIGVTDLSKRIAQHSREVAGFSKRYNCARLVCFEGFDDIHDARQFEVRMCKKNWNGWKLRLRIAGQNPDPWRDLTDTLPV
<NL2_RS15895--Novosphingobium_lentum_NBRC_107847--57.14 (SEQ ID NO:934)
MAFFAYLLRCSGSGYAGHTDDELRFAQHQCAGLGGYTAKRRPVMLVWTDSTFTRDDAFAVERKLKWSRAKKDALIADNWELISQLARGRTGRNDAGTDGG
RPSTSSG
<NN5_RS01715--Novosphingobium_naphthalenivorans_NBRC_102051--75.53 (SEQ ID NO:935)
MRDRDFQFVYIMASERNGLTYIGVTSWLPKRAWEHRESMIGGFTKRYDCKLLVWVFEHATMEYAIAREKQIKGSGSRKKLALIEAANPQWRDLFEEICLP
<NN5_RS13575--Novosphingobium_naphthalenivorans_NBRC_102051--64.10 (SEQ ID NO:936)
MTSDRYRGGIYTGTAHLQRRWQHREGTGSQFARKYGFTRLVYAEYDEIEDAIAREKAIKKWRDWDKIRLIETANPDWEDMFERLGGVV
<SSA01S_RS03615--Sphingomonas_sanguinis_NBRC_13937--73.03 (SEQ ID NO:937)
MDGPAHYIIMANRRNGTLVIGSTINLKPRAWEHRNGVVEGFTKRYGCHLLVWYEVHESLEFARQRELQMKEWRRGWKLQIEGLNPWDWDLVDRIALP
<SSO01S_RS04170--Sphingomonas_solis_NBRC_100801--78.02 (SEQ ID NO:938)
MHRTPAVYIMASGRNGTLTYIGVTSDDLGRAHQHREGMIDGFSKRYGCKLLVWYEVHEVESAIVREKRIKEWKRAWKLDLIERENPRWRDLAEDLGFDPPLPP
TRHPGERDDPG
<SSO01S_RS14205--Sphingomonas_solis_NBRC_100801--72.41 (SEQ ID NO:939)
MPDRNPNPCVYILASAPRGTYIGVTSDDLPGRLYQHRHGTEGFATRCRVHRLVYFEMAETMDAAIAREKQLKRWHRDWNKLNLIERNPEWNDLAPVLGIAERL
<SSO01S_RS14765--Sphingomonas_solis_NBRC_100801--69.66 (SEQ ID NO:940)
MEDQGFVYIMASRRNGTLTYIGSAHALPKRAWEHRNGVVEGIIKKYGCNLLVWYEAHESLDAARLRERQMKEWRRRAWKIREIEGLNPEWNDLYNAVAP
<SSO01S_RS00155--Sphingomonas_solis_NBRC_100801--65.88 (SEQ ID NO:941)
MAHPGQWYIMADRYRTIYVGVTSDLFRRVHQRHSGEGSDFCRYGLHRLVWADHAPSITDAIAREKRVKRWRRRAWKFSLIETGNPDWRDLSDTLF
<SSO01S_RS16220--Sphingomonas_solis_NBRC_100801--58.43 (SEQ ID NO:942)
MGGVYIMTNGKPCGVLYTGVAANLAQRVDQHRDAGSAFCRKYGLKTLVLAEFPHDRIEDAIAREKMLKAWKRWDWKIELIERTNPGWRDLRDDIR
<VME01S_RS20520--Vibrio_mediterranei_NBRC_15635--75.28 (SEQ ID NO:943)
MKQASVYILSSRSNKVLYVGVTSQKQRIWQHKNVVDGFSKRYHVNKLVPFEVHQSMATAIAREKQIKRWKRDWKENLIKSMNPWNMDLYDSL
<VME01S_RS20935--Vibrio_mediterranei_NBRC_15635--68.48 (SEQ ID NO:944)
MTKQPCVYITASKSGVYIYIGVTSNLLQQRWQHREGVGTGGFVKKYNVYHLVHYELFDCMEDAIAREKQIKHWNCSWKNDLVTSHNPYWRDLTPDLW
<SAD01S_RS10860--Sphingomonas_adhaesiva_NBRC_15099--76.14 (SEQ ID NO:945)
MRERQPCVYILASDRNGTLVGVTSNLI GRVMQHRTGTFGGFTARYGVHRLVWYDVADTMDAAIAAEKRIKRWRRWKVELIERDNPHWNDIAPGIGLPPQ
<SAD01S_RS13605--Sphingomonas_adhaesiva_NBRC_15099--61.96 (SEQ ID NO:946)
MPWVYVMTNAPSGTLTYIGVTAHLAARVMQHREGSGGFCREHGLTRLVHVEEHATMLDAIAREKALKKWKRGWKLNLIGRTNPWDWDLWDVINM
<SAD01S_RS05345--Sphingomonas_adhaesiva_NBRC_15099--69.66 (SEQ ID NO:947)
MITGAAVYIMASRRNGTLVIGSTVNLKPRAWEHRTGAVDGFSGKRYGCHRLVWYEVHESWEAARQRELQMKEWRRNWKLREIEGLNPWDWDLVDRIARP
<SAD01S_RS16805--Sphingomonas_adhaesiva_NBRC_15099--64.37 (SEQ ID NO:948)
MANHRRGQTYLGVTSNLPQRAWQHRNGVIEGHSRDKGCVRVWYEFDDLQACACEYRMKKWKRAWKLRLEERNPDWRDLFDDVCA
<AWN68_RS05680--Rosei virga_echinicomitans--64.77 (SEQ ID NO:949)
MNSGCVYIMSNKNRTLYIGVTSNLTARVLEHRNGTGSKFTSKYCNCFDLIYFYYPSIVVAIAREKQLKNWHRDWDKINLIKSENTELLDLSNSVGVGD
<AWN68_RS15230--Rosei virga_echinicomitans--66.67 (SEQ ID NO:950)
MPKGGFVYIVTNLTHSTLYTGVTSDLSRVIIQHREKHLPSFTAKYKCYKLVYQAFGHI ESAIEEEKRIKGGSRVKKIELVESMNPWDWEDLYDDIKDW
<AWN68_RS09045--Rosei virga_spongicola--65.17 (SEQ ID NO:951)
MPKGGYVYIMSNKTRTVLYIGVTSDES RVYQHKNGEKSKFTKLYRCYFLVYEVFDHIESAIAREKVLKKWNRWKEELITAFNPNRLDLWEEIDGFN
<AWN68_RS08645--Rosei virga_spongicola--69.05 (SEQ ID NO:952)
MTNKYHTTLVGVTSDLVSRVQHREQHFPKSFTSKYCNCFKLVYRVFDHIESAI TEERLKGGSQQKLDLINEMNPNWDLNLFTEVQYW
<AWN68_RS09055--Rosei virga_spongicola--65.53 (SEQ ID NO:953)
MGKNGYVYIVSNTHRTMFYIGVTSDES RAYQHENGMLKGFSSKYNCDLIYYEVFNHIEFAIAREKVLKKWKREWKERILEMNPPEMRLNDEVFRFN

FIG. 11DD

<AWW68_RS04385--Roseivirga_spongicola--61.96 (SEQ ID NO:954)
MTHAKGGSVYILTNKNTTLYTGVTSDDLASRLREHITGLYPKSFTARYNLKLVYMEHFSSEIEAIKREKAVKKGKTRINKEKLINSVNPNDLSEEVLSW
<AWW68_RS04690--Roseivirga_spongicola--55.95 (SEQ ID NO:955)
MAGLHVYILNTSKKVLVVGVTNDLEKRLIQHYQNRGEKRTFAGRYCYLYIFYEYHHNNIGGIEREKEIKGWSRQKVDLINSYNPKWFLNDLQITKWPPDP
GTF
<MB14_RS13030--Roseivirga_ehrenbergii--64.13 (SEQ ID NO:956)
MAYGGSVYILTNKNTTLYTGVTSDDLARLRQHRNKAFPNSTARYNLTLVLYMEHFSSEIEAIQREKAIKAKTRANKEKLIDSVPNDLSEEVLSW
<MB14_RS13095--Roseivirga_ehrenbergii--64.13 (SEQ ID NO:957)
MAYGGSVYILTNKNTTLYTGVTSDDLARLRQHRDKAFPKSFTAKYNLTLVLYMEHFSSEIEAIQMEKAIKAKTRINKEKLINSVNPNDLSEEVLSW
<A1704_RS14875--Chryseobacterium_cucumeris--65.12 (SEQ ID NO:958)
MTAFSGRYTYIYILTNKNNRVLYTGVGTGNLHRRLYQHKTKLNPGSFTARYNLEFLVYVEKYDWHHAIQREKEIKNWSRNKKLELIRSTNPNLEFLNYLFE
NWI
<AO498_RS13790--Algoriphagus_sanaruensis--63.04 (SEQ ID NO:959)
MKKGGSYIIMTNYSNTVLYTGVTDNLIRRVTDHKDGKLNLSFTYRYNLKLVYFESFHTIEAIAKEKQIKGGSRKKNENLINSNPWDLWEEILKW
<A3774_RS01385--Neptuniibacter_marinus--58.97 (SEQ ID NO:960)
MVKDTGLSRWYVYIVKCVDTLYTGVTTDIERRLRQHNGDLVGGAKYTRVRRPVQMVWREBQPDQSSACQREAAIKKLSRAEKLRLVD
<A3773_RS03185--Neptuniibacter_pectenicola--57.53 (SEQ ID NO:961)
MSSWYVYIVECVDGTLTYGTVDLQRLRQHNGDLVGGAKYTQVRPVRIVWNESQQNRSAACQREAAIKKLSRSDKLALVKQYN
<SAB01S_RS07885--Sphingobium_abikonense_NBRC_16140--59.30 (SEQ ID NO:962)
MARGGCAYVMTNKPRGVLYIGVTADLAARIEQHCNGAGSIFCKTYGLTRLVLAEPHDEIAFAIAREKTLKAWKRAWKIRLIEETNPDPDRDIDSLML
<SAB01S_RS17525--Sphingobium_abikonense_NBRC_16140--58.90 (SEQ ID NO:963)
MARGGYTYIMTNKARGVLYIGVTCNLAAARMLQHRNGAGSTFCKKYGLTCLVLVEPHDDITLAIAREKALKAWKREWIRL
<SAB01S_RS17205--Sphingobium_abikonense_NBRC_16140--56.72 (SEQ ID NO:964)
MARGGYTYIMTNKARGVLDIGVSANRAARVLEHRNGAGSTFCKKYGLTCLVLVEPHDDMTLAIAREKALKAWKRE
<SAS01S_RS19815--Sphingomonas_asaccharolytica_NBRC_15499--74.16 (SEQ ID NO:965)
MMGQAFVYIMASRRNGTIYIGSTDDLRRVWEHRNDVPGFTRDYGCHELLVWYEVHDGLEQARLRERRMKKWKRAWKIREIEGLNADWSDLYERIPSL
<SAS01S_RS03645--Sphingomonas_asaccharolytica_NBRC_15499--67.42 (SEQ ID NO:966)
MTGGFVYILTNKRKHGALYTGVCADIAARMQHRDGTGSKFCKKYGINRLVYVESHDPIYDAITREKAIKKWNRKAWKIALTERNPDWRDLFFDINC
<SAS01S_RS05980--Sphingomonas_asaccharolytica_NBRC_15499--66.67 (SEQ ID NO:967)
MPVQIERTPCVYLLAKAYHSTLYTGVTSNLVGRIIHHREGAFDGTFRKYGYRLVWYLTDDTMEAVIKREKQIKRWREYKYNLIEELNPTWNDLAVGLGLPP
LAP
<SAS01S_RS05195--Sphingomonas_asaccharolytica_NBRC_15499--65.12 (SEQ ID NO:968)
MRGGYTYIVANKPFGVLYIGVTSSLAIEAHRQKRGSSFAAKWDCTRLVLIEAHSIDDAIAREKRLKSWNRVWVKCRILIEQNPNWDDLWDQINGGAESVAL
DNPAAPASAGATIAAQVHPE
<SCA01S_RS16450--Novosphingobium_capsulatum_NBRC_12533--69.57 (SEQ ID NO:969)
MRRERQPCVYILASGVRGTLYIGVTSALMQRLWQHREGITGGFTARYGVYRLVHFIEFADMEQAIIVREKQLKNWRRDVKLNLVHAANPDWRDLAIDLGFPLPG
QKGPHPGP
<SCA01S_RS10325--Novosphingobium_capsulatum_NBRC_12533--66.28 (SEQ ID NO:970)
MEKGGVYVIMANRYRGGIYVGVTSDDLIRRVQHRDGRGSIHVRDFDKTKLVYAEERHEIEPAIAREKLVKKWREWKFALIEANNPDWDLWDQWYPTAARQG
EIAK
<SMA01S_RS21155--Sphingomonas_mali_NBRC_15500--70.79 (SEQ ID NO:971)
MPGGYVYILANRKNKALYTGVCADIAARMQHRDGTGSKFCKKYDISRLVYVERHDAIIDAIAAREKAVKKWHRAWKIALIEQANPDWRDLFFDINR
<SMA01S_RS07460--Sphingomonas_mali_NBRC_15500--65.52 (SEQ ID NO:972)
MAKGGWYIIMASKPRGMLYVGVTAYLAERVDQHRDVGSAVCRRYGTGLVLAELHDEIEYAIAREKALKAWKRAWKIELIEHVNPGWEDLFSRDLVIA
<SPI01S_RS20340--Sphingomonas_pituitosa_NBRC_102491--73.49 (SEQ ID NO:973)
MANKRNGTIYVGSTDLPGRAYQHRSGMIEGFTTRDKGCTLLVWYEAHEDLDGARLRERQMKKWKRAWKIALIERDNPQWKDLYGTILFG
<SPI01S_RS05715--Sphingomonas_pituitosa_NBRC_102491--73.49 (SEQ ID NO:974)
MLASQRHGTLYTGVTSDLMGRIHQHRSQVTPGFTKRNNVVKRLVWYEVGGDMAAIIKREKQLKEWKRDWKIVLIERENPFWEDLAIISLGFPLDAPDP
<SPR01S_RS03425--Sphingomonas_pruni_NBRC_15498--71.28 (SEQ ID NO:975)
MEMPSYVYIMASGRNGTIYMGSTDDLVRAYEHRNGVIDGFTKRYGCKILVWYEEHSELAARLRERQMKKEQRAWKREIEBGMPEWEDLFDRIALP
<SPR01S_RS19610--Sphingomonas_pruni_NBRC_15498--70.79 (SEQ ID NO:976)
MAGGFVYILANRKHGALYTGVCADIAARMQHRDGTGSKFCKKYGISRLVYVESHDIAIEAITREKAIKKWHRAWKIALIERNNPDWRDLFFDINR
<SPR01S_RS07290--Sphingomonas_pruni_NBRC_15498--67.82 (SEQ ID NO:977)
MEIERTPCVYLLAKAYHSTLFTGVTSNLVGRIIQHREGTGGFTKRYGVYRLVWYMTGDTMEAVITREKQIKRWPREYKYNLIEELNPTWNDLAVGLGLPPLA
P
<A4F88_RS09255--Flavobacterium_chilense--65.82 (SEQ ID NO:978)
MEKQGGYHTYIYIITNKSFTVYTGVTNNLRIRLSQHKENIGGESKSFSTSKYKVEFLLYYEKFTWIEAIGREKEIKGWNRNKKIELIKTINPDLDPLNY
<AA2016_RS23115--Aminobacter_aminovorans--70.00 (SEQ ID NO:979)
MSGYVYILASRKNGTLYTGVTSNLEQVRVWEHREEITGGFTKKGYGVKRLVWYRDYQDIGDAIVDEKRIKRWRRKWKLEIEAMNPGWADLYETLNG
<AA2016_RS22085--Aminobacter_aminovorans--67.42 (SEQ ID NO:980)
MTGYVYMTASQKGGTIYIGVTNDLARRLPEHKASEGSRFTARYGVHRLVWYEEHFDIRDIAIAREKSLKRWRPQWKVELIEKTNPWEHELFRGLGW
<AA2016_RS20145--Aminobacter_aminovorans--66.29 (SEQ ID NO:981)
MTGYVYMTASQKGGTIYIGVTNDLARRLPEHKAGKSGRFTARYGVQRLVWYEEHFDIRDIAIAREKSLKRWRPQWKVELIETNPWEHELFRGLGW
<A0W33_RS02915--Pontibacter_akesuensis--59.77 (SEQ ID NO:982)
MASHNYFTYITNPRKTIYLVGVTNDIIRLDEHFKNCNPGSSFAGKYCYNLVYERHSQINHAVEREKEIKKWSRSKKEALISYMNPFQFLNLKVQDE
<NP4_RS30145--Nocardia_pseudobrasiliensis_NBRC_108224--68.48 (SEQ ID NO:983)
MGRRQWIPTYIMADRPQGMLYGATRNLVGRIWQHKNVVPSTYRRHGIHTLVWYEAHATLENALLRERRLKGYSRPRKIALIEASNPWDWDLYRSL
<NM1_RS20430--Nocardia_mexicana_NBRC_108244--69.57 (SEQ ID NO:984)
MGRRYRRPPCVYIMASHPHGVLYIGATSNLVGRAWQHRNDVVESYTKRFRVHMLVWYQSAGTLEAALIQERRLKTRWRERKIRLIEETNPWDWDLYPALL
<AUP43_RS18560--Oceanibaculum_pacificum--65.88 (SEQ ID NO:985)
MAYVYVYILASDVRCVYIGVTHDIARRAYEHRQSAVGKFTKRYAVKRLVHYELFGTALEAIQREKSLKRRRRDVKVNLIERDNPWDLIPGTRPGMTK

FIG. 11EE

<A4G25_RS07075--Staphylococcus_condimentii--53.09 (SEQ ID NO:986)
MENKYFIYIVRCRDGSLYITGYTNDVTARVEKHNAKGAKYTKRRPVKLVYQETTFETKSEAMKREYQIKYTRVQKQLQIERGVIS
<A4A58_RS24575--Tardiphaga_robiniae--80.00 (SEQ ID NO:987)
MTYYVYLLASRKYGTLTYIGVTNDLIRRVYEHRSKVVDFGTRRYDVVDKLVWFEVYDDVLSAISAREKELKKWRRDWWKIRLIEEQNPWNMDLYPVLSG
<A4A58_RS16245--Tardiphaga_robiniae--60.29 (SEQ ID NO:988)
MRQSRIFYVYILASGHHGTLTYIGVTNDIRLELHRCGRGSEFVKRYKVMRLVHVEEFSSIVEAIAREKQLKNWKRE
<AVT10_RS14030--Sphingomonas_hankookensis--75.28 (SEQ ID NO:989)
MRIFEPCVYIVASRRNGTIYTGVTSNLIQRLHQHRSFGSPGFTADHDCHRIVWFERHETMDAAIQREKRIKKWNRDQKQNLVERDNPWSDLSGGLGFFPVPQ
<AVT10_RS17990--Sphingomonas_hankookensis--72.41 (SEQ ID NO:990)
MVERLPCVYILASARNGTLYTEVTSDLVRRIAQHREGTFGGFTAENIKRLVYWEIAATMEAAIYREKTVKRWRRDQKQNLIERDNPWNDLAIIGLGEPLPS
PSRPVDFPGTSPG
<AVT10_RS12825--Sphingomonas_hankookensis--79.41 (SEQ ID NO:991)
MKAGFYIVASRRNGTIYTGVTSNLIQRIHQHRRGGTVEGFTTEHDCCKLLVWYECDDDLQEARLRERRIKKWNRE
<AVT10_RS00535--Sphingomonas_hankookensis--62.22 (SEQ ID NO:992)
MRGGWTYIMANRRFGMTYVGVTDLAARVWQHRRDSGSAYCRKYGIRTLVLAEKHDDIESAIRREKALKAWRAEWKHLIADANPEWRDLFDDIV
<AVT10_RS17700--Sphingomonas_hankookensis--64.71 (SEQ ID NO:993)
MDGERQGGWYIMADRYRGATYVGVTDLLIRIHQHRTGHGSHCAKYNLARLVWAERGGDILACIAQEKMKRWRHRDQKQNLIERGNDPWRDLSTTIFG
<PsAD2_RS00540--Pseudovibrio_axinellae--76.47 (SEQ ID NO:994)
MFYYMYILASRNKHGTLTYIGVTNDLARRVYEHKSKVGCRTAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<PsAD2_RS06425--Pseudovibrio_axinellae--64.44 (SEQ ID NO:995)
MSYFMYILASKPRGTPTVGTNDLARRVYEHKSKVGCRTAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<A6052_RS19810--Pedobacter_panaciterrae--66.67 (SEQ ID NO:996)
MDRGGCIYIMNTIARTVLYIGVTSDLQARIAEHKLAVDQKSFYAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<FBFR_RS07365--Flavobacterium_fryxellicola--65.00 (SEQ ID NO:997)
MIEYTEGIYNYFYVYIMTNKNRTVLYTGVTNNLKRRLNEHKNHINASSFTARYNVEFLIYYEHFGWIQLAIAAREKEIKDMRRELKLNLIQEFNPENMKFLNSHFE
<LPB303_RS16500--Polaribacter_atrinae--67.06 (SEQ ID NO:998)
MKKYHIYFLTNKNNTVLYIGVTNNLLKRIYQHKTKYKGFYAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<LPB303_RS05260--Polaribacter_atrinae--58.24 (SEQ ID NO:999)
MPKQYFIYITNEKKTVLYIGVTNNIQRKISQHYFDSQNAKKSFAKYNKYCYLLYETFFDINLAILREKQLKGLREKKNKIISDFNPNWDFLNNDF
<A7A78_RS14390--Aequorivita_soesokkakensis--69.47 (SEQ ID NO:1000)
MKPGYVYIVTNKNNTVLYIGVTNNLQKRIYQHKTKYKGFYAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<A7X65_RS32920--Bradyrhizobium_stylosanthii--63.10 (SEQ ID NO:1001)
MSTGSYYVYILASRHHGTIYIGVTNDIRARLELHRSRGRGSEFVKRYGVFRLVHVEAFATPQEAIAAREKQLKFWKRDWIKLIEENPDWNRDGLI
<AYJ54_RS20715--Bradyrhizobium_centrolobii--76.67 (SEQ ID NO:1002)
MTYWVYILASKPGGTLYIGVTNNLVRVYEHREGLAEGFTKRYGVKTLVYFEAHEITAAALQPEKNIKHWSREWKIDLVKSNPEWRDLYDEIVR
<AYJ54_RS34345--Bradyrhizobium_centrolobii--77.65 (SEQ ID NO:1003)
MGERVYVYILASKIGGTLYIGVTNDLIRRAEHKSEIIESFTEKYDVARLVYFEQFNDPENAIKREKRLKKWNRRAWKIRLIGQHNPNDWDLPGIAGPQ
<AYJ54_RS25455--Bradyrhizobium_centrolobii--73.12 (SEQ ID NO:1004)
MAYYVYILASKKYGTLYLGVTDNIVRRIHEHKS KAVAGFSKRYAVDRLVWFEIYDDPVTATIEREKELKKWRREWKIRLIEERNPQWIDLIAQIAS
<AYJ54_RS34345--Bradyrhizobium_centrolobii--77.65 (SEQ ID NO:1005)
MPFFVYILASKRNGLTYIGVTNDLARRMTEHKAKLIPGFTQRYGVTLVYFESFDSVLEARAREHSLKRWRRAWKIRLIEEQNPWIDSYPIAG
<AYJ54_RS13130--Bradyrhizobium_centrolobii--68.24 (SEQ ID NO:1006)
MYHVVYILASRRHGTLYIGITNLLQKRLLEEHRNGKSGSEFVKRYGVYRLVHSEAYERVEDAIAREKQLKRWKRDWIKLIEERENPEWRDLSDLLV
<AXW67_RS20865--Bradyrhizobium_neotropicale--72.73 (SEQ ID NO:1007)
MAYCVYILASKKYGTLYLGVTDNIVRRTHEHKS KAVAGFSKRYAVDRLVWFEVYDDPVTATIEREKELKKWRREWKIRLIEEQNPWIDSYPIAG
<AXW67_RS38450--Bradyrhizobium_neotropicale--76.39 (SEQ ID NO:1008)
MDMPFFVYILASKRNGLTYIGVTNDLPRRMAEHKAKLIPGFTQRYGVTLVYLETFSVLEARAREHSLKRWRRAWKIRLIEEQNPWIDSYPIAG
<AAV28_RS29930--Bradyrhizobium_diazoeficiens_USDA_110--77.66 (SEQ ID NO:1009)
MAYYVYILASRRDGAIVYGITNDLVRRTYEHRIKAVPGFTTKYNIIRLVWFEIYDDPISAIAREKELKKWRRAWKIRLIEEQNPWIDSYPIAG
<AAV28_RS25010--Bradyrhizobium_diazoeficiens_USDA_110--76.40 (SEQ ID NO:1010)
MAYYVYILASRKHGALYIGVTNDLVRVYEHRSKAVPGFTTKYGVSKLVFEIYDEPASAIAREKELKKWRRDQKQNLIERGNDPWRDLSTTIFG
<AAV28_RS37585--Bradyrhizobium_diazoeficiens_USDA_110--79.78 (SEQ ID NO:1011)
MGAHHYVYILASKIGGTLYIGVTNDLIRRAEHKSKLIESFTEKYDVARLVYFEQFDDPENAIKREKRLKKWNRRAWKIRLIEEQNPWIDSYPIAG
<A6I66_RS12630--Tamlana_agarivorans--67.06 (SEQ ID NO:1012)
MDVHYVYILTNKNHTVLYVGRSKQLKVRNLQHKNNLSKSFYGRYVTKLVYFETTKYVNNISIKRERQIKKWNREWKINLINGLNDPDKLSEYI
<SY85_RS17195--Flavisolibacter_tropicus--64.04 (SEQ ID NO:1013)
MPKLTYSYVYMLSNVHRSVYFVGVTSNLIQRLHQHRSFGSPGFTADHDCHRIVWFERHETMDAAIQREKRIKKWNRDQKQNLVERDNPWSDLSGGLGFFPVPQ
<SY85_RS14475--Flavisolibacter_tropicus--57.65 (SEQ ID NO:1014)
MHQYVYIVTNPERTVFTYTGVTNDLEQRIIEHYLNKGRKRTFAGKYCYNNLIYFAYEQYINNAIAREKEIKGWNRRKKLTLEAVNPSLTFNNAQLFDGWPPK
EITTR
<AKL11_RS09640--Endozoicomonas_ascidiicola--80.43 (SEQ ID NO:1015)
MSKQPCVYIILVSKKNGTLYIGVTNNLQKRIYQHKTKYKGFYAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<AK963_RS09155--Endozoicomonas_atrinae--79.35 (SEQ ID NO:1016)
MPKQPCVYIILVSKKNGTLYIGVTNNLQKRIYQHKTKYKGFYAKYGVSKLVSEFQFRDNDADIRREKQVKKRWRRDQKQNLIERGNDPWRDLSTTIFG
<A8025_RS07600--Rufibacter_ruber--62.96 (SEQ ID NO:1017)
MKQHNYFVYITSNPTKTYLGVGTNDLRRVYEHREGLADGYTKEHGKQLVHFDVHETVPLAIQREKNLKHVWRRAWKIALIEEKNPDWRDLWVDITK
<TV41_RS05360--Wolbachia_endosymbiont_of_Dactylopius_coccus--77.78 (SEQ ID NO:1018)
MKSYYVYILASERNGLTYIGITSNLIKRIWEHKS KAI SGFTSKYNSKLVYFEFQDINLAI SKEKLLKSWQRKWKINLIEKTNQEWKDLIDEIVSGFQCQAL
E
<A3840_RS13285--Devosia_elaeis--69.89 (SEQ ID NO:1019)
MAGTYIILANHPRGKIYIGVTNDLVRVWEHREGLADGYTKEHGKQLVHFDVHETVPLAIQREKNLKHVWRRAWKIALIEEKNPDWRDLWVDITK

FIG. 11FF

<A3840_RS09940--Devosia_elaeis--69.66 (SEQ ID NO:1020)
MKGYVYILASKKNGTLYTGVTSNPSRRLLEHTRELTPGFTSRYGKTLVWLEVDLVVNARREKAIKKYPRQWKINLIEAVNPEWKDLSAWLLDT
<A6A05_RS12425--Magnetospirillum_moscoviense--72.09 (SEQ ID NO:1021)
MQGGWVYIMTNRPNSTLYVGVTSDLARRAWEHREGVADGFTKRYGLKQLVYAERYEDIRTAIQREKNLKHWPRAWVKLILTDNPNWDDLYERLA
<A6A05_RS05940--Magnetospirillum_moscoviense--70.93 (SEQ ID NO:1022)
MQGGWVYIMTNRPNSTLYVGVTSDLARRAWEHREGVADGFTKRYGLKRLVYVERIEDIRTAIQREKNLKHWPRAWVKLILADNPNWDDLYERLA
<A6A03_RS06560--Chloroflexus_islandicus--63.64 (SEQ ID NO:1023)
MQHHYFVYIMTNNHHTVLTGMTNDLRRVYQHKTRGSGQFTKRYQATKLVYEMTIDVRAAI TREKQIKAGSRQKKIALIEAMNPEWRDLSVDWA
<A6A04_RS17815--Magnetospirillum_marisnigri--67.74 (SEQ ID NO:1024)
MTRKWKGRPVTKHPAVTILASQPNGILYIGVASDLAKRSGDHKTGAIDGFTKRYGIHTLVWFEIDETMPLTIEREQMKKWRAWKI SLIEQTNPCWDDLYAA
ILR
<A6A04_RS03840--Magnetospirillum_marisnigri--73.26 (SEQ ID NO:1025)
MQGGWVYIMTNRPNSTLYVGVTSDLARRAWEHREGVADGFTKRYGLKQLVYAERYEDIRTAIQREKNLKHWPRAWVKLILTDNPNWDDLYERLA
<AU381_RS10665--Ensifer_glycinis--70.97 (SEQ ID NO:1026)
MAGYVYIVTNHKGRTLYIGVTSNLERRI FEHREGLTPGFASKYGCNRLVWYEEHLQIGTAIQREKSLKRWRQWKIELIEKMNPDRDLYFELW
<ATB98_RS27575--Sinorhizobium_saheli--74.12 (SEQ ID NO:1027)
MHAAANEGYVYILASKRNGTLYTGVTRDLSHRLFEHQNDLTPGFTSKYGVKTLVWFEEDLLTAAI TREKNIKWKPRKWKLNLI EQMNPWEDIAHYLHGL
<ATB98_RS00950--Sinorhizobium_saheli--72.04 (SEQ ID NO:1028)
MAGYVYIVTNHKGRTLYIGVTSNLERRI FEHREGLTPGFASKYGCNRLVWYEEHLQIGTAIQREKSLKRWRQWKIELIEKMNPDRDLYFELW
<A5893_RS03795--Pedobacter_psychrophilus--57.95 (SEQ ID NO:1029)
MKSYYVYLLTNKNNTVIYTGFTNDLKRRIWEHKFDKGSKFTTKYCNCKLIYYEENFTPYDGIAREKQLKGGSRAKKEALINLENSQWSDLSIDWY
<I597_RS07135--Dokdonia_donghaensis_DSW-1--65.59 (SEQ ID NO:1030)
MBFNQFYTYII TNKNNTVLYLGFITDNLKRRINQYKSKSGKGFSGKDYTEKLVWFEVFTQYVENAIKREKQIKKWNREWKENLINDMNPKWEDLYWML
<I597_RS02385--Dokdonia_donghaensis_DSW-1--62.50 (SEQ ID NO:1031)
MFTNFYIYITNTQYRTTVYIGVTNNIQRRIQGHYFDSQNEKKSFAGKYNCYELVYVEGFQYAKDAI TREKQLKRWSRQKKNQLIETVNPKWESLNNEVF
<I597_RS02370--Dokdonia_donghaensis_DSW-1--59.30 (SEQ ID NO:1032)
MPYTYIILTNNKNTVLYVGATRNLKIRMVQHKTKYYPKSFSAQYNCYNLMYYELFETHNEVLAREKQLKSGNRKRKISLIESKNPDWDLASSYASLGSEDYL
IKEIW
<A8C56_RS08070--Niabella_ginsenosidivorans--64.84 (SEQ ID NO:1033)
MNEHHYTVYIVSNYLRITVLTGVTDIVRRMSEHKSGEQKGFTSRYRCYYLVYFEEFSDIRNATAREKQLKRWHREWKLNLIKTQNPDMKDLSEWFW
<A8C56_RS03905--Niabella_ginsenosidivorans--65.48 (SEQ ID NO:1034)
MQKGGFIYIILTNNKNTVLYIGVTSNLSRLWEHQHETPGSFTDQYNATLLIYYEFWNTINSATIEREKQLKGWIRKKKEQLIARKNPQWKVFNEEISKEIYSL
LY
<A8B35_RS08235--Labilibacter_aurantiacus--65.93 (SEQ ID NO:1035)
MERGGAVYIILTNKTHQVLYVGVTSKLKQRIYEHKTKFHAKSFTARYGVNLLVYKGFHSIEEAIAYEKQLKSGNRKRKLDLI TEQNPWSDLYNGID
<A9D15_RS16025--Croceicoccus_pelagius--82.80 (SEQ ID NO:1036)
MAPTVYILASRRNGTLYIGVTSNLIARLHQHRSQDIDGFTKRYDVHLVWVFEQHEAMENAI VREKRMKKWNRAWKRLIEEHNPDWRDLAVDLGFEPMARAGF
PPSRERRRLYGLILNRLRQDQPRACQAARGWVSRARNALRLR
<A7X80_RS07370--Rhizobium_bangladeshense--69.66 (SEQ ID NO:1037)
MSGYVYIITNQKNGTLYIGVTSLEHRIYEHREGLTPGFAWKYGCTRLVWYEEHWGIGSAIQREKSLKRWRQWKIDLVEAMNPEWDDLYHTLW
<A9D12_RS10910--Porphyrobacter_neustonensis--79.12 (SEQ ID NO:1038)
MTKVPAVYIMASRRNGTLYVGVTSDLPGRAWQHRDGI VQGFTTKRYGCKLLVWFERHETMESAIMREKQIKAGNRKAKALIEADNSTWRDLFEIEIAQ
<A9D12_RS06815--Porphyrobacter_neustonensis--75.86 (SEQ ID NO:1039)
MRQERNPCVYILASQPHGTLYIGVTSDLIGRLWQHRNGVVEGFTSRYKVHRLVRYELFGDMEHAILREKQLKRWHRQWKINLIESENPEWHDLAAGLGFPPVG
PRRIPDGS
<AMJ99_RS11640--Rhizobium_esperanzae--67.78 (SEQ ID NO:1040)
MSGFVYMMSDKPRGVIYTGVTSDLQGRIVEHRNEIHKGFTSEKRAKNLVWFEYHPNIVLAIQREKSLKRYLREWKIKLIEELNPTWMDLYERIDEIENVYRPH
PNTREWSYDN
<BA052_RS03150--Wolbachia_endosymbiont_of_Nomada_flava--77.78 (SEQ ID NO:1041)
MKSYYVYILASERNGTLYIGITSNLIKRIWEHKS KAI SGFTSKYNVSKLVYFEEFQDINLAISKEKLLKSWQRKWKINLIEKTNQEWKDLIDEIVSGFQCQAL
E
<BA051_RS03185--Wolbachia_endosymbiont_of_Nomada_leucophthalma--77.78 (SEQ ID NO:1042)
MKSYYVYILASERNGTLYIGITSNLIKRIWEHKS KAI SGFTSKYNVSKLVYFEEFQDINLAISKEKLLKSWQRKWKINLIEKTNQEWKDLIDEIVSGFQCQAL
E
<BA050_RS03115--Wolbachia_endosymbiont_of_Nomada_panzeri--77.78 (SEQ ID NO:1043)
MKSYYVYILASERNGTLYIGITSNLIKRIWEHKS KAI SGFTSKYNVSKLVYFEEFQDINLAISKEKLLKSWQRKWKINLIEKTNQEWKDLIDEIVSGFQCQAL
E
<BA053_RS06005--Wolbachia_endosymbiont_of_Nomada_ferruginata--77.78 (SEQ ID NO:1044)
MKSYYVYILASERNGTLYIGITSNLIKRIWEHKS KAI SGFTSKYNVSKLVYFEEFQDINLAISKEKLLKSWQRKWKINLIEKTNQEWKDLIDEIVSGFQCQAL
E
<A6E01_RS17660--Vibrio_breoganii--69.32 (SEQ ID NO:1045)
MKAPCVYILSSINRKVLYIGVTSNLRGRVWQHKSHYFSPSFTSRYHVTRLVYFEVHDEMEHAIEKEKKLKFWRRSWKDALITEMNPEWNDLYETL
<I603_RS13190--Porphyrobacter_dokdonensis_DSW-74--67.78 (SEQ ID NO:1046)
MDREQQGGWYIMADRYRWITVYGSTADLAFRI TQHRDQGSKFCAEYGLTRLVWMELLPTMQEAMDQEKRIKRWRREWKFALVEKGNPDWRDLFDDVLK
<A9996_RS11145--Gelidibacter_algens--68.13 (SEQ ID NO:1047)
MSWSIYIITNIPNGVLYIGQSKRLKERIYQHKIKSHPNTF SARYNLKLKLVYFENHDIQDNARMREKQLKKNRWDWKIELIENTNPEWLDLYKNL
<A9996_RS03140--Gelidibacter_algens--67.39 (SEQ ID NO:1048)
MKPGFIYIILTNKNNTTLYVGVTSITVIRVYQHKDKHDPKSFTARYNLDKLVYEAQFMIGDAIGREKQLKGGSRAKKEALINSINPEWKDLFDELKKEFGDD
DYDDDDDDDE

FIG. 11GG

<A9D35_RS03700--Formosa_haliotis--67.06 (SEQ ID NO:1049)
MKTSYVFMNKNNTVIVGVTSNLKRRVFQHKSKTYKGFTSKYCNCDLVYFEIFQDIKLAIAAREKQIKAGNRKRKISLITSTNPNWTDLSNGWIFDV
<A9D35_RS14965--Formosa_haliotis--68.24 (SEQ ID NO:1050)
MKTSEFYILTNSYRTTFVVGVTNNLLKRIHEHQNNVGSVFTKYNLKDLYVEEFETDINQAIAREKQQLKNWKRWSKLNLIKETNPTLKALVI
<FLP_RS19030--Flavobacterium_piscis--64.77 (SEQ ID NO:1051)
MFNPQEGFHSYVYILTNKYRSTFYIGVTNNLKLRLQQHKDNLNGNGKTFASKYSIEFLVYFEKFWVIQEAIAREKELKKWRREKKIELIQSFNSGMEFLNYH
FE
<FLP_RS20310--Flavobacterium_piscis--63.29 (SEQ ID NO:1052)
MHLEVGFHTYYIYIITNKAHSVFTYGTNDLRSRLQHKNNIVSKNKTFAKYNVEFLLYEKFETWIEAIAREKEIKGWRREKKLALIKIINPELKFLEY
<A6B35_RS04315--Mesorhizobium_amorphae_CCNWGS0123--66.29 (SEQ ID NO:1053)
MTGYVYMTASQKRGTIYIGVTNDLGRIPEHKAGEGASFTSRYGVQRLVWYEEFFDITDAIQREKSLKRWRPQWKIELIEKTNPEWFEFGRGTGW
<ADU59_RS23080--Pararhizobium_polonicum--73.12 (SEQ ID NO:1054)
MAGYVYIVTNHKGRTLYIGVTADLERRIYEHREGVTPGFTSKYVGCTQLVWYEEHGSIGTATQREKSLKRWRPQWKIELIEAMNPDPWRDLYHELM
<ADU59_RS02795--Pararhizobium_polonicum--74.16 (SEQ ID NO:1055)
MKGYVYILASKRSGTLYIGVTNLLPGRLEFHQNELTKGFTSRYGVKTLVWFEYEDLVTTAITREKTIKKWPRQWKLNIEAMNPDIWAHYHLGL
<A6F65_RS01295--Altererythrobacter_namhicola--63.33 (SEQ ID NO:1056)
MDREKQGGWCYIMADRYRGTLYVGVTSDLAARVTQHREGRGSEFCKYGLTRLVWVERGNTIDECIAHEKRIKKWRVRWKFDLIEKANPDWDDLFDLLT
<LMTR13_RS00990--Bradyrhizobium_icense--74.16 (SEQ ID NO:1057)
MTYVYVILASKRPGTLYIGVTNNLIRRVYEHREGADGFTKRYGVKTLVYFEAHTIAVALQREKNIKHWSREWKIDLIVAGNPDPWRDLYEDIAR
<LMTR21_RS01080--Bradyrhizobium_paxllaeri--74.44 (SEQ ID NO:1058)
MTYVYVILASKRPGTLYVGVTTNNLIRRVYQHREGADGFTKRYGIKTLVYFEAHDIAIATALQREKNIKHWSREWKIDLIVAGNPDPWKDLYDEIVR
<LMTR21_RS23425--Bradyrhizobium_paxllaeri--67.06 (SEQ ID NO:1059)
MYVYVILASRRHGTLYIGVTNSLSKRMEQHRNREGSSFVKTYGVRLVYFEAFERDEAIAREKQLKRWRDKWIELIERENLEWRDLSDLIV
<AXE80_RS02165--Wenyngzhuangia_fucanilytica--60.53 (SEQ ID NO:1060)
MIYVYVILHCSNDNTYTGTTNNLNKRIEHHQSGKNFNSYTSRRLPVQLMFYCTFTNVEVAISKEKQIKKWSKAKKEALINNQFDKLPNFARKKFPNKK
<A7A08_RS06395--Methylobacterium_haloferans--71.59 (SEQ ID NO:1061)
MKSYYVYILASRRHGTLYIGVTADLPRRVWEHRQDAVDGFTRRHAVHRLVYMEEFSEPLAIAAREKSLKRWRPDWKINLIEQGNPDWADLFDGLNR
<AWL63_RS09215--Sphingomonas_panacis--69.77 (SEQ ID NO:1062)
MASGRNGTIYIGVTNSLPVPRAYQHRNGLVEGFTKKYGCCLLVWYEAHTIETDARQRELMMKKWRKWLKSEIERLNPNWTDLYETIA
<AWL63_RS01070--Sphingomonas_panacis--63.95 (SEQ ID NO:1063)
MQRGWYVILMASKPHGMLYIGVTAHLAARVDQHRDVGSAFCRGYGIKTLVYAEFDTIQAISREKALKAWKRAWKQDLIDSANPEWRDLYDTLA
<A6302_RS21045--Methylobrevibacter_pamukkalensis--71.11 (SEQ ID NO:1064)
MAGWVYIMANRMRGTLYIGVTSDLPRAYEHREGLTGPFASRHGCRRLVWTEEFPTITEAISREKSLKKWTRAWKALVERDNPHTDLDFTLNM
<A9E74_RS06110--Methylobacterium_muralis--56.58 (SEQ ID NO:1065)
MTKTEWTLIIEAANGHLYTGITDLELRLKQHQHNGGARFFHTSQAQRVYVYEHQHPDRSSASRREAAIKKLSRQQLNLIKGPSDHY
<BFL28_RS12630--Sphingomonas_turrisvirgatae--80.23 (SEQ ID NO:1066)
MEKAGYVYIMASRRNGTLYIGVTSDDLARAHQHRNKLDIGFTKEHDCITLLVWFEGHGDIEAARTRELQMKRWKRAWKIELIERENPQWRDLYATLF
<BFL28_RS13330--Sphingomonas_turrisvirgatae--71.26 (SEQ ID NO:1067)
MRERLPCVYIMASGYNGTLYVGVTSNLIIGRMVQHRAGTFGGFTARYGVTRLVWFEAGESMEAIAAREKQVKKWRREWKRNLIERENPMWNLAVGLGLEPLVV
VSDGAVDPGTRPG
<AUC68_RS11850--Methyloceanibacter_methanicus--68.54 (SEQ ID NO:1068)
MRGGWYVILTNKPNGLYIGVTRDLARRAFEHRDGVAPGFTTRYGLKRLVWCEHHETITAAIQREKNIKHWSRAWKVRLLIADNPWDDLYVSLNQ
<AUC71_RS11505--Methyloceanibacter_marginalis--74.44 (SEQ ID NO:1069)
MRQFYVYILAKHRHGALYIGVTNDLIRRVYEHREGLAGPFTKRYGVKTLVWFEYEQHATAATAIQREKTMKRWPRAWKIRTERTNPTWADLYEEICR
<AUC71_RS09130--Methyloceanibacter_marginalis--71.11 (SEQ ID NO:1070)
MRQFYVYILAKHRHGALYIGVTNDLIRRVYEHREGLAGPFTKRWQGIKQLVYVYEQHATAATAIQREKTMKRWPRAWKIQTEEBLNPTWTDLYEEICR
<AUC71_RS11250--Methyloceanibacter_marginalis--66.28 (SEQ ID NO:1071)
MRGGWYVILTNKPNGLYIGVTRDIARRAWEHREGMVPSFTTRYGVKRLVWCERHETIAGAIQREKRMKHWPRAWKVRLLIAMPNPEWDDLYETLNQ
<AUC70_RS12180--Methyloceanibacter_stevinii--76.40 (SEQ ID NO:1072)
MKDRRYVYVILASRIGGTLVGVTTNDLIRRVYEHRTKTAEGFTKRYGVARLVYFEQFDDITAAIEREKRLKKWNRRAWKIELIEERNANWDDLYPSIATT
<BHK69_RS02780--Bosea_vaviloviae--70.65 (SEQ ID NO:1073)
MREPSVYIRNGTLYVGVTSQLGRAWQHRTGIVLGFTTRYGCKLLVWYERYERMEDEAIAREKQIKAGPRARKLTIEAINPTWRDLYDELA
<BES08_RS15355--Novosphingobium_resinovororum--77.42 (SEQ ID NO:1074)
MRETFDPCTYLLASHRNGTLYVGVTSNLLRLHQHREGLEGFTSDYGVHRLVWFERHATIESAILREKRIKTWNRRAWKALIEQANPDWRDLAVDFGVPLP
PPRIVRG
<BES08_RS13820--Novosphingobium_resinovororum--61.63 (SEQ ID NO:1075)
MPGWYVILTNKPHGVLYIGVTADLPARMMQHRGGAGSAFCRKYGLKRLVLEPHGRIEDAIRREKAMKAWKREWKIRLIEEANPEWADLFRLL
<BES08_RS00690--Novosphingobium_resinovororum--65.48 (SEQ ID NO:1076)
MEKGGFSYIMANRPFVLYVGVTTNDIHRVMHQHYEGRGSRFCQRYGIDRLVLVPEPHGSIQAIAREKQQLKAWREWKLELIEQVNPWDDLSGLL
<Verru16B_RS08915--Lacunisphaera_limnophila--60.00 (SEQ ID NO:1077)
MPELFPVIPAPAIADLAWYVILQTAGALYIGQTRDIGERLKRHRGLGSKFTTDSHGPRLVFCEGTMSLDEAVVREHQLKRWSRPKKEALIRGDLPRRLREL
SRSRERPGNGARA
<BFP71_RS09340--Fabibacter_misakiensis--67.42 (SEQ ID NO:1078)
MPKGGYTYIMSNKQRTTLYIGVTSDBLRTVYQHRNREGSKFTKRYCYHLYYFEVFEAIESAIEREKVLKRWNRSWKEELIKGFNPQLRDLWDQIQGFN
<BFP71_RS16245--Fabibacter_misakiensis--63.04 (SEQ ID NO:1079)
MYKGGTYVYIITNKNKTTLYVGVTSDIRSLREHLEKKYPRSFSTRYNLFLLIYMEHFTIEAIFREKYIKGKRRNWKENLIKSENPNWKDLGPEVLSW
<OLTU_RS01440--Oligoflexus_tunisiensis--66.67 (SEQ ID NO:1080)
MRMYVYVILSDRFRRLLYTGVTNNLRLHEHRTKHNPNSTARYDITLVYFETHGHAGQANIREKEIKRMNRWERERLIREFEEQERTCPHKLKNPYAPPG
RQRRRPPSCPPWPGASLPHD
<S4054249_RS21465--Pseudoalteromonas_luteoviolacea--76.34 (SEQ ID NO:1081)
MQPSVYVILASKPYGTIYIGVTKDLKTRVYQHKHNLTKSFTSQYNVHLLVYFECYDSITTAIEREKQLKNWKRAWKIQLESFNPWKQDLYHKI

<BHS39_RS06995--*Salagentibacter salarius*--71.59 (SEQ ID NO:1082)
MNNDYQYHVYIITNRKNGTLYIGITRNLKQRLYQHRNRNLKSFSSRYRLSKLVFYEKFEYPSLAIKREKQQLKKNWNLWKINLINDFNPDWEDLTGRFK
<BHS39_RS04865--*Salagentibacter salarius*--62.92 (SEQ ID NO:1083)
MQNSYIYILSNSTRMTIYIGVTSDLKKRI TEHKSGMGSFLFARKYHLKFLIYFEEFSEIHQAIAREKQQLKNWHKEWKWNLVKSMPDLKDLIEDL
<BHS39_RS04875--*Salagentibacter salarius*--61.54 (SEQ ID NO:1084)
MQKSYIYLLSNSTRMTIYIGVTSDLKKRI TEHKTEGTGSSFTKRYHLEFLIYFEEFSEIHQAIAREKQQLKNWHKEWKWDLAKSINPDLKDLIEEL
<BI364_RS01155--*Acidihalobacter prosperus*--83.91 (SEQ ID NO:1085)
MCTSLRAKKGTLTYGVTSNLPGRVWQHKEKVAVAGFTKKYDVSLLVYFEMHPVAESAI LREKQIKKWNRAWKISLIEENNPAPWRDLYDEIV
<BHQ40_RS11400--*Sphingobium phenoxymethylphosphatibacter*--79.31 (SEQ ID NO:1086)
MMKERAPTIVYILASRKNGTLYTGVSDDLKRLYEHNRGLIEGFTTDYGVNRLVWFEPHDTMETAIQREKRIKKWNRQWKIDLIEANPDWDDLALTFGFDPLP
SLRTDDADPADRPDGFPPSRE
<BHQ40_RS22295--*Sphingobium phenoxymethylphosphatibacter*--77.27 (SEQ ID NO:1087)
MPRQPCITIIASLRKNGTLYIGVTSDDLQRIHQHRI GAVPGFTQDYGVKRLMWFEEHGTMDIAI IREKRLKKNRAWKIALIEALNPHWDDLALIGLGFERLPD
<BHQ40_RS21220--*Sphingobium phenoxymethylphosphatibacter*--75.86 (SEQ ID NO:1088)
MPDLIRHPLALHILVMPNQPCVYILASGRHGTLYGVGTGLHIERIHQHRTTMDGFTARHVSVHRLVWFYQADFPPTAIALEKKLKWRDRDWKVSLEKDNPFW
EDRAIILGLPPLGSS
<BHQ40_RS20495--*Sphingobium phenoxymethylphosphatibacter*--70.93 (SEQ ID NO:1089)
MRGGWVYILASRYRGTLTYIGVTANIAARAYQHKTGIGADFTKRYGVTLVYAEQAHATIEEAIAREKALKKWKWDKFRLEEGNPDWIDLYDRING
<BHQ40_RS18265--*Sphingobium phenoxymethylphosphatibacter*--64.04 (SEQ ID NO:1090)
MARGGWYITMNTKPRGVLYVGTADIAARVLQHRSGSGDFCRKYNLTRLVLAEPHEDIEWAIAREKALKAWRREWIRLIEEVPNPEWEDLIERII
<BHQ40_RS24060--*Sphingobium phenoxymethylphosphatibacter*--55.71 (SEQ ID NO:1091)
MPFWTYMLHCGGSLYVGHDTNLEQRIGQHEQGEIPGFTRDHLPVKLVWSEFSQSRIEALERMERRIKGWSRTKKLALIRSDWEEISRLSKGKGKPTSSG
<BFC17_RS18235--*Alteromonas lipolytica*--78.13 (SEQ ID NO:1092)
MIDKYPVYIILTNQYKTVFYTGVTNNLQRIYQHKMQMADGFSKKYGLKVLVYFEMHEEIEAAI QREKRLKWRWSWKIKLIEQQNPNRDLYPLILG
<DUPY_RS12185--*Duganella phyllosphaerae*--75.58 (SEQ ID NO:1093)
MEKLSVYMLASGPYGTLYIGVTSDLIRRVWQHREALAHGFTEQYHVKRLVWYEQHSDIRAAITREKHLKWRREWIKDLVHQSNPRWHDLYPTLTA
<LPB138_RS18235--*Urechidicola croceus*--68.13 (SEQ ID NO:1094)
MKPGFIYILTNKNNTTLYGVGTANLEQVRVQHKSKMNRKSFTARYNLDKLVYFESFQMGDAIAREKQQLKAGSRAKKIALIEGLNPNWDDL
<LPB138_RS09685--*Urechidicola croceus*--64.71 (SEQ ID NO:1095)
MKNSFVYILTNKYRTTFYIGVTNLSKRLYEHYEGKASVFTKKYQIINLVYEQFIDVEQAIAREKQQLKNWHKEWKLNLIKEMNPKELTLDYQ
<LPB138_RS09570--*Urechidicola croceus*--61.29 (SEQ ID NO:1096)
MLHNYFVYIIVTNQYRTTYIGVTNNIQRLSEHYFDSQNSKSFAGKYNCFYLIYYEGFENIESAITREKQQLKWRREKKNNLINRFPKWESENNQII
<BJP26_RS04090--*Sphingomonas melonis*_TY--79.78 (SEQ ID NO:1097)
MRDVFQPCVYILASRRHGTLYIGVTSNLLARLVQHREGLIPGFTKAYGVHVLIIWYEVHATMEAAIVREKRLKEWRRRAWKIELIEAQNEDWGDLGVGLGLAPLP
PVPTKESMT
<BJP26_RS07890--*Sphingomonas melonis*_TY--75.58 (SEQ ID NO:1098)
MTKIGYVYIMASRLNGTIYLGVTSDLAKRAYHHRNALIDGFTKIYGCRLLVWYEAHDELD SARLRERQQLKKNRDNWKLRLIEDKNPDWLDLYESLF
<BJP26_RS03575--*Sphingomonas melonis*_TY--72.04 (SEQ ID NO:1099)
MRPSIYMMASGRNGTIYVGVTSDLPRRAWEHREGTINGFTKRYGCKLLVWFETADTMDVAIAREKQIKAGSRAKKVALIEAMNPLWNDLYEQIAHG
<BJP26_RS04260--*Sphingomonas melonis*_TY--60.00 (SEQ ID NO:1100)
MANDRGWVYIMSDRYRGTLTYGVTAHLAARIHQHRTGDGSDFCRRYGLHRLVWAERSTAIADCIAHEKRLKWRHRDWKIALIERGNPDWHDLSQDLA
<LPB03_RS07295--*Polaribacter vadi*--75.00 (SEQ ID NO:1101)
MKNVHQYVYIILASKIRGTLYIGMTNDLQRRVYEHKMGITKGFTSKYGVNKLIIYFETFKDVNEAIEREKRLKKNRAWKIELFEKNPNPSWQDLAKDWYDDILT
D
<LPB03_RS02990--*Polaribacter vadi*--67.03 (SEQ ID NO:1102)
MKPGYVYIILTNKNNTTLYGVTSNLAQRLKQHRSKANEKSFTAKYNLDKLVYKAFQMGDAIAKEKELKAGKRAKKIALIEGVNPNWENLIEP
<LPB03_RS06840--*Polaribacter vadi*--67.06 (SEQ ID NO:1103)
MHKKVFIHTYVYIITNKYRSTFYIGVTNNLKVRLKHKNNILQKKKTFAAKYNIEFLIYYEKYNWIIQLAIAREKELKWRREKKLVLI REMNPTFVFLNSDF
E
<LPB03_RS01040--*Polaribacter vadi*--57.83 (SEQ ID NO:1104)
MNYCYIILSNKRNRTLLYVGSTRDLEKRIKVHKSNGANFTKRYNVFELIYFETFDNKMAREKSQLKNWHKEWKWNLVKKSNPDLKILEI
<EM308_RS13560--*Flavobacterium gilvum*--65.91 (SEQ ID NO:1105)
MIEFTEGYHSYVYIITNKYRSTYYIGMTNNLGI RLQQHKENIEKGIKTASKYNIEFLVYIEKYTWVQQAIAREKEIKGWRRDKKELIRSFNDNFELNHH
FVNSNVI PPPSE
<BEN49_RS17675--*Hymenobacter coccineus*--68.24 (SEQ ID NO:1106)
MRTYVYVYLLTLNHTVLYGVGTNDLRRVAEHKAGSQAGFTKQYNVHKL VHFETYP SIADAIAREKQQLKAGSRAKKLALINASNPWEEELAA
<BEN49_RS23910--*Hymenobacter coccineus*--60.24 (SEQ ID NO:1107)
MYVLNTQTQTVLYIGVTNDLIRLHEHSSNRGDASKLTGRYQADLLVYFETAPDATQAI TREKQQLKGMRRKKDALINALNPS
<BEN47_RS08645--*Hymenobacter lapidarius*--58.62 (SEQ ID NO:1108)
MYVYIILNTGRVLYIGVTNDLNRRLYEHRI GNGDASKFTGRYQADLLVYFELAPDATQAIAREKQQLKGWTRAKKEALINTLNPDWNTIDLETW
<BEN48_RS08475--*Hymenobacter glacialis*--54.32 (SEQ ID NO:1109)
MAYFVYITTPAKTVLYIGVTSGLNRLAQHYDNRGDAKTFAGRYFCYHLLYAEQFADITS AIAREKELKGWTRAKKDALIESVNPEREILT SY
<LT090_RS10745--*Thalassotalea crassostreae*--66.67 (SEQ ID NO:1110)
MNKAAYIILTNKPNGTLYIGATSNLENRI DQHRFKVTGSFSKRYNLDKLVYFELFDTFYDALTREKQQLKNWRRQWKINLIEKVNPHWDDLCLNTLI
<BHE75_RS02010--*Sphingomonas haloaromaticamans*--73.86 (SEQ ID NO:1111)
MDRQPFAYIILASGKGTLYIGATSDPAKRLHEHREGLIKGFTSRYGITRLVYFESFDDMHSAIAREKQQLKWRREWKNLIEASNPGWDLAPRLRLGSAIPQ
R
<BHE75_RS07390--*Sphingomonas haloaromaticamans*--73.86 (SEQ ID NO:1112)
MIERQPCVYIILASSFNGALYGVTFDILLRIARHRDGMFDGFTERYGVKYL VYWEAFGTMEAAI LREKQIKWRWRDWKNLIERNNPSWHD LAVPLGLEPLAP
SPA

FIG. 1111

<BHE75_RS16710--Sphingomonas_haloaromaticamans--72.73 (SEQ ID NO:1113)
MSRQPCVYIILASDRYGTLYIGVTSNLPVRLHQHRTGTIPGFTSRYGVHRLVRLMFEDMQSATAREKQLKRWHRWIKINLIESDNPWNWDLAIGLGFEPAGRH
MPHDGP
<BHE75_RS02915--Sphingomonas_haloaromaticamans--68.97 (SEQ ID NO:1114)
MERNPCVYIILASGRHGTLYIGVTSEPMKRLHAHRSGSVPGFTARHGVRLVHFETFGDMVSAIAREKQLKAWRRDWKVALIEAANPFWEDRALEWGFPLSPF
<BHE75_RS15115--Sphingomonas_haloaromaticamans--63.64 (SEQ ID NO:1115)
MTRGGWTCILSNKPKGILYIGVTANLAARMQHRGTGSAFCRKYGDLRLVFAEPHDSMVEAIAREKALKAWKRDWKIRLIEQGNPEWEDLFDRI
<BIW53_RS20015--Pseudoalteromonas_hyunsanensis--76.67 (SEQ ID NO:1116)
MQPAVYILCSKPKGVLYIGVTSNLKKRVYEHQRHYVDGFTKKYNVTQLVYFECHEQMSALCREKQLKNWRRQWKNLIIQQKNPEWMDLYNFI
<BET10_RS12660--Pseudoalteromonas_amlolytica--75.56 (SEQ ID NO:1117)
MQPTVYILCSKPKGVLYIGVTSNLKKRIYEHQRHYVDGFTKKYNVTQLVYFECHEQMSALCREKQLKNWRRQWKNLIIQQENPEWIDLYNFI
<BIW12_RS11520--Flavobacterium_commune--68.42 (SEQ ID NO:1118)
MKPGFYIITNKYQTVVYGVTSNLPERILEHKNKKYPTSFSAFYDVNIIIVYEQFQWIEDAIAREKQIKAGSRAAKNKLIQSINPTWKDLFDEIKDIMTE
<BIW12_RS03110--Flavobacterium_commune--63.64 (SEQ ID NO:1119)
MYELNEGHTYIYVYIITNSYRSTFYIGMTNNLKQRLQOHHQNIQVGNKTFATKYNLEFLVYEQFTWVQQAIVREKELKNWRREKKNLIRSFNSDFEFLNYR
FNKDNEIPSASE
<LO55_RS24520--Massilia_timonae--75.56 (SEQ ID NO:1120)
MDKSSVYIILASRPYGTLYIGVTSDLIKRVWQHKAEVPSFTSKHGKRLVWYETHDTITQAIVREKQLKRWNRWIKINLIQSTNPTWHDMYAAICV
<BG455_RS14305--Roseinatronobacter_thiooxidans--54.12 (SEQ ID NO:1121)
MVYLVYIAACRSNTAIHIGVTSDLRAEMRLRLGQGHPCARYKICKLVWSEACADLGAARRRADQLKRWRRWTKDALIAEHNQGWDRDLAATLKQDRE
<BLX24_RS24560--Arsenicibacter_rosenii--58.33 (SEQ ID NO:1122)
MPDFAFYIITNPEKHLYTGVNTDLVRRMQEHYDARGKDDTFAGKYFCYNLVWFEHFKYIRNAIDREKEIKGWRRAKKIALIESFNPWFLLNGEIQG
<BLA27_RS15860--Ochrobactrum_cytisi--60.24 (SEQ ID NO:1123)
MEVTVYILRCADGSYITGLTKQNI EARLWEHNEGIYESYTNKRFPVVLVFTETIDYDRIVDAIARERQIKGWSRAKKEALIALNYEGLPALS KR G
<BKY27_RS02410--Niella_nitratireducens--70.59 (SEQ ID NO:1124)
MRDHSYFTYIITNPNGLTYTGVNTLERRVVEHRQGRANSFTRRYNLHRLVWFEPHADIREAILREKRIKKWNRWVKVEMI EAMNPDWRDLATELRL
<AX760_RS01400--Pararhizobium_antarcticum--71.91 (SEQ ID NO:1125)
MKGYVYILASKRNGTLYTGVTSDLPGRLFEHQNELTPGFTSRYGVKTLVWFEEYDLVTTTITREKTIKKWPRQWKNLIEAMNPDWDDIAHYLHGL
<BMF35_RS01560--Erythrobacter_gangjinensis--73.40 (SEQ ID NO:1126)
MFQPAAYIMASQKGGTLVVGVTSDPVKRIHQHREGLVDGFTKRYRCKRLVWIEFHENMDSAIIREKQLKAGSRKTKIALIEADNADWRDLWDIAIK
<BMF35_RS12315--Erythrobacter_gangjinensis--73.86 (SEQ ID NO:1127)
MPRERQPCVYILASKPHGTLYIGVTSDLMTRLWQHRSEVNGFTSRYGVGTIVRYEQFATMDAAIAREKQLKRRHRPWKNILIQSENPDWHDLDVGLGLPPLA
TR
<BMF35_RS08520--Erythrobacter_gangjinensis--67.78 (SEQ ID NO:1128)
MRGGWYIMANRYRGTIHI GLTPDIARRVAQHRDGKGSFAFIQRYGCKAIVHAERYDSIDDITREKALKKWNRDWKIRLIEEQNPWDWEDRYADLVIP
<BMF35_RS11360--Erythrobacter_gangjinensis--65.12 (SEQ ID NO:1129)
MAKGGWYIMANRYRGGMYVGVTSDLQARILQHRNGEGSQHVADSGKTQLVYAEHREDIETAI VREKRLKRWREWKFALIEESNPDWRDLWEWHPTPEE
<BMF35_RS03085--Erythrobacter_gangjinensis--64.44 (SEQ ID NO:1130)
MDRETHGGWLYIMADRYRGTMYVGMTADLAARIVQHRDGSQSKFCABEYGLTKLVWAEKIAPIHLCIEQEKRMKKWREWKIALVEKANPDWKELFEQLI
<CabyS_RS15705--Caldithrix_abyssi_DSM_13497--76.67 (SEQ ID NO:1131)
MALKNYVYIMTNKPHGVLYTGTNNLIRRTYQHKHKLIDGFTKKYNVDKLVYIEIHHDPMSAIIREKRIKKWKRAWKIELIEGMNPEWKDLYDDLGGKGDE
GCIGEGFPLARE
<CabyS_RS02395--Caldithrix_abyssi_DSM_13497--68.54 (SEQ ID NO:1132)
MKDFKYFVYIMTNWNNKVMYVGTNNLMRRVYEHKNKLIDGFTKKYNINKLVYIEYTFQDVRAAIEREKEIKKWREKKKNLVNTKNPEWRDL SKDFE
<CabyS_RS02390--Caldithrix_abyssi_DSM_13497--68.54 (SEQ ID NO:1133)
MKDFKYFVYIMTNWNNKVMYVGTNNLMRRVYEHKNKLIDGFTKKYNINKLVYIEYTFQDVRAAIEREKEIKKWREKKKNLVNTKNPEWRDLSEDFE
<LPB144_RS06990--Gramella_salexigens--55.68 (SEQ ID NO:1134)
MKEYLYITTNKKTVLYTGTNNLAKRIDQHYYESKNDKRSFAGKYNCYHLIYYECFNDPNEAILREKTIKKWREKKDKLIQKFNPDWKFL ENEL
<LPB144_RS06170--Gramella_salexigens--59.76 (SEQ ID NO:1135)
MFYQGWHTYTYITLNKKTVLYTGTNNLKRRLDEHWSGRNAIAFTKRYRCHYLIHYEKYTWIQEAI VRENELKGWLSKKEQLINEHNPDRNFLNDHFS
<LPB136_RS01415--Tenacibaculum_todarodis--71.58 (SEQ ID NO:1136)
MKPGYVYILTNKNNTTLVVGVTSNLKQRILOHKEKHDKKSFSARYNLTKLVYIEAHQMIGDAIGREKQLKAGNRKAKKIALIESINPAWDLDFDDL
<LPB136_RS00570--Tenacibaculum_todarodis--63.86 (SEQ ID NO:1137)
MANKPKGVYIIGVTDNIEERVKEHKLKLYPNSFTARYNCNKLVYFEEILKAVQANERERQMKKKWRNWKTKLIEDMNPSPWMDLSMNWNFNNTNKRFEHR
<BSQ44_RS11630--Mesorhizobium_oceanicum--66.29 (SEQ ID NO:1138)
MTGYVYMTASQKGTIYIGVTSDLGRRIPEHKSSEGSFRFTGKYRVTRLVWYEEHFDIRDAIQREKSLKRWRRDWKIALIESMNPHEELFRGTGW
<BSL82_RS12545--Tardibacter_chloracetimidivorans--75.00 (SEQ ID NO:1139)
MERQPAVYIILASGKHGTLYIGVTSNLMQRLYQHRGTGLVGFTKRYEVHHLVHFEMFDDMSLAIAREKQLKRWHRDWKINLIERDNPEDWRDLAIELGFRSLR
<BMW22_RS02690--Rhizobium_leguminosarum--74.47 (SEQ ID NO:1140)
MKGYVYILASKRNGTLYVGVTDLAERLYEHQNELTPGFTSRYGVKILVWFEEYDLTITREKTIKKWPRQWKNLIERENPEWEDISFHLGL
<BMW22_RS08175--Rhizobium_leguminosarum--68.89 (SEQ ID NO:1141)
MSGFVYMMSDKPRGVYIIGVTSDLHGRIWEHRNEIQKGTQKYRAKNLVWFEEHPNIVLAIQREKSLKRYMREWKIKLVEELNPTWMDLYERIDEIENIYRPH
PNTREWS DYN
<F7310_RS02285--Francisella_uhiginis--75.58 (SEQ ID NO:1142)
MKNGYVYILTNKNGVLYIGVTSDIKRVYEHKNKFPVDSFSKKYSLNKLIVYFVYENIEEAIIREKQLKKWNRWAKDRIINTMNSDWVDLYDKLLS
<F7310_RS02265--Francisella_uhiginis--74.44 (SEQ ID NO:1143)
MKNGYVYMLTNKNSVLYRVTSDIVRRIYEHKSKEFVDGFTKRYNLNKLIVYFIEYENIEDAIIREKQLKNWHRNWKIDLKSVNPTFEDLYERIVT
<F7310_RS00845--Francisella_uhiginis--71.91 (SEQ ID NO:1144)
MYTYIMTNKFYGTLYIGVTSNLRVFEHKNGLVDGFTKKYSIKYLVYFESFEDVRAAVHREKRLKAWLRKWKIDLINSVNSQWHDLYEEIIR

FIG. 11JJ

<F7310_RS06965--Francisella_uliginis--55.29 (SEQ ID NO:1145)
MSAYMYILKNDGSSYTGSTNNLERYLQHMTGEGAKYTSKRLLPVKLVYFEECNRIDEAFGREKQVQKWARAKKNALINNFSEELVRLAECMNETHSHKNNRP
FDSAQGGSSYTCGVQQNEKI
<RG47T_RS23560--Mucilaginibacter_polytrichastri--69.23 (SEQ ID NO:1146)
MKVHSYVYIILCNKANTVLYVGVTSIEERVWQHKKQVFKGFSKAYDCNKLVEEDYQWVQDAIAREKQLKGGSRQSKINLIVNENPAWNDLSNSWYD
<RG47T_RS20230--Mucilaginibacter_polytrichastri--59.78 (SEQ ID NO:1147)
MERGGCVYIMTNKRQSVLYTGITSDIVARTDRDHKNRVHPSFSDRYNCDRLVYHCFYPHIEEAIATAEKAIKGGNRKSKEALINSNMPQWKDLYDELLVQ
<RG47T_RS16100--Mucilaginibacter_polytrichastri--59.14 (SEQ ID NO:1148)
MLFKQGHVYIMSNQYRTTFYIGVTSDLRSRVWQHLLNNNEGSAYVKYKLFDLVYEHFDRI TDAIDREKQLKNWHHEWKNLIKSVNPEMKDLKVEIDYQE
<RG47T_RS24155--Mucilaginibacter_polytrichastri--60.92 (SEQ ID NO:1149)
MPNYNFFTYILTSYNKRVLYTGVTDLDRRLFHYFGLDQPNSTSKYKCYLVWFERHQYIDHAIQREKEIKGWIRNKKNKLI ELENPSWRFLNTEIMDWPP
VR
<RG47T_RS19610--Mucilaginibacter_polytrichastri--58.51 (SEQ ID NO:1150)
MPNYNFYTYMLTNSSKTVLYTRVTNNLDRRLYEHYFNEGSKNFTSKYKCYLVWFERHQYIDHAIEREKEIKGWSREKKSNLIEDENPFWNFLNKDIMDWPP
LKMS
<BRX40_RS02685--Sphingomonas_koreensis--62.22 (SEQ ID NO:1151)
MRGGWTYIMASKPRGMLYIGVTAYLAERVDQHRDRGSFAFCRYGIKTLVLAERHDSIEDVIRREKALRAWKRDWKIKLIEQANPAWDDLFQIA
<BUQ68_RS10245--Mongoliimonas_terrestris--77.27 (SEQ ID NO:1152)
MPLTTGAPLKEPAVYILASGRRGTLYTGVTSNLPNRVWEHKKHVKRLGFTAKYRIDRLVWFEMHETMETAIRREKQIKDWRRLWKVELVEASNPDWRDLSDTLA
<BUQ68_RS17315--Mongoliimonas_terrestris--68.89 (SEQ ID NO:1153)
MEKGGWYIYITNRPDGVLTYGTADLPRRIGEHRHGLVPGFARKYGCRTLWVYEHHADIRDAIYREKLIKTNRAWKVRMIREANWGWDLDWGRVLGEE
<BLX89_RS02735--Rhizobium_arenae--70.00 (SEQ ID NO:1154)
MSGFVYMMSDKPRGVYIYIGVTSDLTARIWEHRNHIHEGFTAKYRAENLVWFETHPNIVLAIQREKSLKRYRREWKIELIEGLNPTWLDL FERIGEIDNIYRPH
PSTKMWNNDYN
<ASM33_RS01965--Wolbachia_endosymbiont_of_Folsomia_candida--72.22 (SEQ ID NO:1155)
MKNYYYITASEPNTGLYIGMSTNLKQIWEHKSIISSFTSKYNVCKLVLYEEFQDINLAVAKEKLLKGWQRKWKIDLVEQENQDWKDLYDEVIK
<ASM33_RS00605--Wolbachia_endosymbiont_of_Folsomia_candida--72.62 (SEQ ID NO:1156)
MIIFIYILQVDTALYIGITSNLKIRVWEHKNKVVSFGFTSKYNVQKLVYFEEFQDIDSALNREKLLKSWKREWKIDLVARKNPNWYDLYEYI IK
<ASM33_RS07360--Wolbachia_endosymbiont_of_Folsomia_candida--72.29 (SEQ ID NO:1157)
MIIFIYILQVDTALYIGITSNLKIRVWEHKNKVVSFGFTSKYNVQKLVYFEEFQDIDSALNREKLLKSWKREWKIDLVARKNPNWYDLYEYITQ
<BVC88_RS20425--Algoriphagus_marinus--65.52 (SEQ ID NO:1158)
MFVYIILTNKTKTYLIGVTNDLQRLYEHRTNLGSSNSFTSKYKCHFLVYFEEYGTPEKAI TREKELKKWRREKKNNLIEAENPNWDFLQIE
<BVC88_RS20295--Algoriphagus_marinus--64.77 (SEQ ID NO:1159)
MKKGGSTYIMTNKNTVLYIGVTSDLIRRVSEHKVKTNPSTFTAKYNLDKLVYFENFQSI BEAINREKQLKGGSRKKEILINSINPEWIDLFDELDRHFEGF
<BVC88_RS01185--Algoriphagus_marinus--60.98 (SEQ ID NO:1160)
MEAFMYILKCGNGDYITGSKTNLEKRIAHEHQMGLGANFTKKHLFPVKLVYFEVDFRIDLAFSREKQVQWNSKAKKEALIDKDYDKLKALSKGKS
<BVC87_RS16170--Tenacibaculum_agarivorans--63.64 (SEQ ID NO:1161)
MRKSFYIILNTYRTTFYIGVTNLSKRLTEHYEGTGSFTTKYNIITDLIYFEEFNDINRAIAREKQLKNWHKEWKNLNIREVNPFTKTLDY
<BJF92_RS06400--Rhizobium_rhizosphaerae--67.74 (SEQ ID NO:1162)
MCGFVYMMSDKPRGVYITGVTSDLQRIWEHRNEVREGFTSKYQAKNLIWFERHPNIVLAIQCEKSLKRYRREWKIRLIEELNPTWLDLYEEIDRIDNPYMPH
RNTTRWNNDYN
<BJF95_RS12795--Rhizobium_oryzihadicis--73.40 (SEQ ID NO:1163)
MAGFVYIMTNHKGHTLYIGVTSDLARRVYEHREAITGGFTTRYDLRQLVWYEEHYEIGSAIQREKSLKRWYRQWKIDLIEAFNPDWDRDLYSELM
<BIY22_RS03955--Vibrio_panuliri--70.65 (SEQ ID NO:1164)
MKQPVVYILSNQTRRSIYVGVTTNLKARVWQHQSICGEGFTKRYRINRLVYFELFSDIREAIAREKQLKWKWREWKNNLIEKVNPTWSDLYPEI
<BVL55_RS00695--Salaquimonas_pukyongi--73.33 (SEQ ID NO:1165)
MTGFVYIMASKRNGTLYTGVTSDLARRVYEHREGLIEGFTKRYDCKMLVWYEEYFNVAIDAIQREKSIKRYRQWKLDLIESINPEWNDLYETLGW
<BVL55_RS08560--Salaquimonas_pukyongi--75.56 (SEQ ID NO:1166)
MGSFVYIMASRNGTLYTGVTTLKRRAREHRDKVLGGFTARYNCTLLVWYREYENVVDAIADEKRIKKWRRAWKIQMI EAMNPDWDRDLYPNLA
<Dform_RS03570--Dehalogenimonas_formicexedens--70.51 (SEQ ID NO:1167)
MTNAAKTVTYTGVTSDLIKRVWQHKNLVEGFTKRYAIHSLIYFEETNNAYSAI SREKEIKGWRRSKKVLINEFNPNWSDLYDGI CGREA
<RA876_RS08225--Rhodoferax_antarcticus--73.63 (SEQ ID NO:1168)
MLKLAYVYIILASKNGTLYTGMTNNLERRIYEHKHHLLDGFTHKYDVTRLVWFTQGEDIRAAITLEKKIKNRSRQWKIDLIEKTNPRWNDLATSWLDSGSCDC
AQDDALCTMQDDALCIMQNAESKKEE
<BKY59_RS14955--Sphingomonas_montana--77.17 (SEQ ID NO:1169)
MERQPAVYIILASGRNGTLYVGVTSDLMRVYEHRAKVRDGFARHASTLVWFEMHDGMDGAIDREKLLKNWRRSWKLALIEAANPLWHD LAVDLGFEPQADP
TVPNERQRGGPGS
<BWR22_RS08790--Lacinutrix_venerupis--77.27 (SEQ ID NO:1170)
MKGSYHQYMYIILSNKNGTLYIGVTNDLERRIFEHKKKLI EGFTKKYGLNKLIIYFEQFYINDAIEREKQLKNWNRSWKIELIEKENKDWTDLSSNWNF
<BW723_RS10540--Polaribacter_reichenbachii--71.28 (SEQ ID NO:1171)
MPKTQHQYVYIITNKKDGVLYIGVSNLDERRMFEHKNKLVKGFSEYRNLDKLIYFETYQYVADAIKREKNLKKWKQWKINLIEBNAEWKDL SKDWFD
<Ga0080574_RS13930--Pelagibaca_abyssi--64.47 (SEQ ID NO:1172)
MGRSGNLRARVVAHRAGLSAHTSKYKIRRLVWFQEHEDFETSLRRERQLKRWRAWKDRLITEANPNWHDISYDIP
<SPHI_RS05975--Sphingomonas_jeddahensis--82.61 (SEQ ID NO:1173)
MPKQPCVYIILASGKHGHTLYIGVTSNLLTSLHQHREGLIEGFTTRYGVHRLVWFDMADEAAITREKQLKKWNDRDWKLRMIEESNFEWRDLAIDLGFALVEKP
ARMD SRLRGNDGIMRGNDGTGKHDRD
<SPHI_RS11240--Sphingomonas_jeddahensis--73.86 (SEQ ID NO:1174)
MRDRQPCVYIILASDFNGTLYIGVTSNLI GRMLQHRTDGTSGFTSRYGVLRLVWYDVADTMEGAISAEKRIKKWPRQYKLNLI ERENPFWEDLAI SLGLSPL
<SPHI_RS13120--Sphingomonas_jeddahensis--76.74 (SEQ ID NO:1175)
MEKAGFYIIMASQRNGTIYIGVTSNLAKRVWEHRNGI VPGFTTKYRCHHLVWYEVHDSIEAAQQRELQMKEWKRNWKLREIEGFNPDWNDLYERIA

FIG. 11KK

FIG. 11LL

<BUW93_RS07365--*Salagentibacter_sediminis*--58.95 (SEQ ID NO:1208)
MQKSWYIILNKNNTVFYTGDTSNIQKRMVQHNNKIYEGFAAKYNCEKLVFVQEFSEINQAIIEYEKKIKSGNRSRKEKLINGQNPWDIDLKSKMILG
<B6S01_RS02345--*Sphingobium_herbicidovorans*--61.73 (SEQ ID NO:1209)
MTRDQGGWVYDRYRGTMYVGVTSNLAARTHQHRSGERSDFCARYGLTRLVWAEHSGDIANCIAHEKRLKRWRRQWKFDLIERGNPDWRDMFDTLA
<B6S01_RS09995--*Sphingobium_herbicidovorans*--61.05 (SEQ ID NO:1210)
MRRSWTYIMTNKPRGVLYIGVTSNLAVRVDQHRRGVGSVFCISKYNLRLCLVLAEEHQRIIDAIAREKALKALKAWKRDWKIELIEAGNPEWRDLFELIM
<B6S01_RS10550--*Sphingobium_herbicidovorans*--58.57 (SEQ ID NO:1211)
MAFYAYLLRCSGDSFYAGHTDDLESRIAQHERGVIPGYTAARRPLALVWSESGFTRDEALAAERRIKGWSRAKKLALIDGDWERMRLARNRQGS
<B8783_RS14870--*Henriciella_litoralis*--67.02 (SEQ ID NO:1212)
MKPFYVYILASQKNGTLYTGHTDDISTRIYAHKEGRGSAFTRKYGVKRLVWVEHHQTRDGAKTREYQIKAWKRAWKIRLIEETNPWDDDLKLNH
<B8746_RS00675--*Henriciella_pelagia*--68.60 (SEQ ID NO:1213)
MAQYDFIATYLLANRKNKGTLYVGSTSDLTRMGQHRSGAGSHFTGKYGVARLVWFERFPEMEPALRAERRMKKEWKRWKIDLIEKGNPDWDDLTMSTLTPY
<B8784_RS03655--*Henriciella_aquimarina*--72.94 (SEQ ID NO:1214)
MAFYVYITTNKPYGVLYTGMTDDINRAWKHPHEHLIKGFTDRYNCEMLVWYEPHESRRSAFDRERRIKWKRSWKVKLIEETNPWDRLGDVLL
<B8784_RS06185--*Henriciella_aquimarina*--66.28 (SEQ ID NO:1215)
MARFDLIAVYMLANRKNKGTIYTGVSDDLWARIQGHKSGKGSHTARYGCDRLVWYIEHTDISPARHRERRIKTWPRQWKINLIETNRRHWDLSMTLPFV
<SAZ01S_RS18635--*Sphingomonas_azotifigens*_NBRC_15497--77.27 (SEQ ID NO:1216)
MDKQPCVYIILASRRHGTLYTGVTSNLLGRIVQHREKVTHTGFTAHEIARLVWFEVHEDMSAAILREKRIKKWNRDWKLALIERSNPQWVDLAIGLGFQPSRTV
NGKR
<SAZ01S_RS21675--*Sphingomonas_azotifigens*_NBRC_15497--74.70 (SEQ ID NO:1217)
MANKRNGTIYVGSTSDLPGRAYQHRTGMIEGFTTDKGCRLLVWYEAHEDLDGARLRERQMKKWKRAWKIELIERDNPQWKDLYGTFLG
<SAZ01S_RS02220--*Sphingomonas_azotifigens*_NBRC_15497--71.08 (SEQ ID NO:1218)
MLASGHRGTLYTGVTSDLMGRVHQHRSCTVPGFTRKNKVKRLVWYEVGGDMAAAIRREKQLKEWKRDWVKVLIERENPFWEDLAI GLGFPPLEVPDL
<IIF7_RS13810--*Zunongwangia_atlantica_22II14-10F7*--61.54 (SEQ ID NO:1219)
MNEYYTYITCNSSKFTFYIGVTNNLNKRLSQHYEDSIKEKKSFAKYNKYNLIYFEVFNQVDAIAREKQLKKWSRAKENLIRLKNPNLEFLSL
<B7364_RS03555--*Aquidulcibacter_paucihalophilus*--72.22 (SEQ ID NO:1220)
MSASEPEDRYVYVLTNRPRGTFTYVGVTSDLVKRVWQHRTHQKPSFTDRYNLERLVWVEQYREVGVAIQREKRIKRWRRDWKIQLVETNNPEWRDLWNDISF
<BST97_RS11600--*Nonlabens_spongiae*--67.44 (SEQ ID NO:1221)
MPKRSFHNYSYIMTNKPNGVYIYIGVTGGLEDVRARHKLIGSRFTSKYCNCKLVYFEDFYINDAINREKQLKNWRRSWKIDLIEKDNPEWKDLSNGWALDL
NIENYK
<BST97_RS05660--*Nonlabens_spongiae*--58.67 (SEQ ID NO:1222)
MPIGYMYILECGNGTYTGSTIDLEKRIEEHKIGQGANTHKKHGVDKLVYVEEYPRIDLAFNREKQVQKWRREKKEALILGKHQNLPELAKAYRDL
<CAK95_RS13335--*Pseudorhodoplanes_sinuspersici*--70.97 (SEQ ID NO:1223)
MKQPAVYIMASKRNGTLYTGVTSNIQQRAWQHREGIGTGFAARYGCKLLVWYELHDMHIAAIEHEKQIKGGSRAKKIALITANPEWRDLYSSLF
<CAY62_RS09820--*Photobacterium_damselae_subsp._damselae*--71.26 (SEQ ID NO:1224)
MFAYVILTSTNSNLYVGVTSRLKERVWQHKQHVTSFGTDKYNVEKLVYVEFDDISSAIKREKQLKWRRSWKNELITGFNRYVVDLYDYV
<BUU34_RS14665--*Novosphingobium_panipatense*--73.86 (SEQ ID NO:1225)
MRDERQPCVYIILASERYGTLYIGVTSNLMHRLWQHRNDALPGFTARYAVHRLVHFELFGDMERAI VREKQLKRWHRQWKINLINAENPEWRDLAMALGLPSMV
PGKPRDGS
<BUU34_RS15975--*Novosphingobium_panipatense*--58.97 (SEQ ID NO:1226)
MAFHTYILRCRDGTYYVGHTDDLERRIAEHQSAAFQGYTSRRLPAVLLWAQDFQTRDEAFAAERKLKGSRAKKEAMMAGDWILVSELARSASPQGASTSSA
<BUU34_RS14525--*Novosphingobium_panipatense*--57.69 (SEQ ID NO:1227)
MAFHTYILRCRDGTYYVGHTDDLERRIAEHGSGMIQGYTSRRLPVTFWLAQDFQTRDDAFAVERKLKGSRAKKEAMMAGDWVLVSQ LACSATPRGASTGSA
<BZF59_RS05530--*Porphyrobacter_donghaensis*--67.42 (SEQ ID NO:1228)
MERGGYIYIMTNKRYGVLYTGVTSDLPARIMQHREGKGGAFSTRYKCTRLVYIEHPTIEDAIAREKAIKAWQRMWKIEAIEKQNP SWDDLFLTINA
<BZF59_RS08985--*Porphyrobacter_donghaensis*--66.67 (SEQ ID NO:1229)
MDREQHGGWVYIMADRYRGTIYVGSTADLAFRVSQHREGGSRFCAEYGLTRFVWAE LPTMHEAMDQEKRVKRWREWKFALIEKANPEWRDLFDDILK
<BZF59_RS15270--*Porphyrobacter_donghaensis*--64.44 (SEQ ID NO:1230)
MGAWVYIMSGAHGTLYIGVTSDLAARVCAHREGRGSAFCRKYDLVRLVYAEHHDITIDAIAREKAMKKWKRWKRLRLRQONPDWDLDFETINA
<BZF59_RS06040--*Porphyrobacter_donghaensis*--63.95 (SEQ ID NO:1231)
MEQGGWVYIMANRFRGMYVGVTADALARVYQHREGTGSRH IKDFAKTRLVYVERHDTIEGAIAREKLVKKWRWDWKFALIEADNPEWNDLWEQWFAPHTEQK
<BZF59_RS12530--*Porphyrobacter_donghaensis*--61.63 (SEQ ID NO:1232)
MTSKRGVLYIGVTADLPARIMQHKKHKGSTFCRRYGLDRLVYAE PHAEIADAIAREKAMKAWKRAWKIELIEKTNPEWGD LFEQIH
<BZF59_RS13190--*Porphyrobacter_donghaensis*--65.12 (SEQ ID NO:1233)
MEKGGWVYIMANRYRGGMYVGVTADALARVYRHREGTGSRHVADFAKTRLVYVERHDTIEGAIAREKLVKKWLREWKFALIEASNPDWDDLWEQWFAPAEAGQK
<BZF59_RS02920--*Porphyrobacter_donghaensis*--58.14 (SEQ ID NO:1234)
MTSKRGVLYIGVTADLPARIMQHKKHKGSTFCRRYGLDRLVYAE PHAEIADAIAREKAMKAWKRAWKIELIEKTNPEWGD LFEQIH
<BZF57_RS15200--*Porphyrobacter_colymbi*--77.42 (SEQ ID NO:1235)
MDQNRQFQPAVYIMANHKNGTLYVGVTSDLPRIWEHREGHVEGFTRYGCKRLVWFELHATMEHAIAREKQIKGGSRAKKIALIEALNPEWRDLFFELNP
<BZF57_RS06910--*Porphyrobacter_colymbi*--73.26 (SEQ ID NO:1236)
MEKVYFVYIILASDRNGTLYVGITSDIAARVWQHRTGAVPGFTSRYGVRHLVHMEAFSDVDEAILREKRIKKWRRRAWKLELIERDNPQWLDLYERINA
<BZF57_RS07000--*Porphyrobacter_colymbi*--67.05 (SEQ ID NO:1237)
MDRETFGGWYIMADRYRGTMYVGVTSDIAARVSQHRSGAGSDFCARYGLTRLVWAE LCGSIDAAI AHEKRVKRWREWKFALIEKANPDWDDLFDHLLG
<BZF57_RS15220--*Porphyrobacter_colymbi*--67.44 (SEQ ID NO:1238)
MGGWVYIMTNGPHGTLYIGVTADLAARVFAHREGRGSAFCRKHNLTRLVYAEHHDITIDAIAREKAMKKWKRWKRLRLRAANPAWIDL FETING
<BZF57_RS11740--*Porphyrobacter_colymbi*--62.79 (SEQ ID NO:1239)
MKGGWVYIMANRYRGGMYVGVTADAIDRVYQHRTGTGSRHVADFAKTRLVYVEHHDITIEGAIAREKLIKKWREWKFALIEAANPDWDDLWEQWFAPDTEQK
<CBP12_RS00455--*Oceanisphaera_avium*--75.00 (SEQ ID NO:1240)
MTKSPTIYLLTNKPGVLYIGVTSQ LIRQIWHQKHNQVSGFTARYNLHKL VYFEVFDMMYLAI TREKQLKNWNRRAWKVDLIEQTNPQWRDLWEDICG

FIG. 11MM

<CBP31_RS03510--*Oceanisphaera profunda*--80.43 (SEQ ID NO:1241)
MLKSPAVYIISNKPQGVLYIGVTSDDLQVRVWQHNRNHDIAAGFTARYNVNKLIVYFEMLSDMYSASIREKQLKNWRRRAWKVELIEQSNPQWRDLWNVCG
<LOKVESMR4R_RS16585--*Yoonia vestfoldensis*--62.92 (SEQ ID NO:1242)
MRIYHLFTTCSQNCMTHEFVYIMASRPNGAIYIGRTRNLHSRVQHHRDGRSVHTAKYIKTLVWFETHDDFDSSLRERAIKRWRRRAWKNQLIHTANPHWQDI
TAHIPV
<CAP51_RS11235--*Acinetobacter populi*--83.87 (SEQ ID NO:1243)
MQIKQPAIYIMASQRNGTLYIGVTSVLKRIWEHKNSIIQGFQKYHVHLLVYELHSSMEEAILREGRLKKWKRWNKLHLEENNPQWLDLYPEIIK
<B0E34_RS09940--*Chryseobacterium mucoviscidosis*--66.32 (SEQ ID NO:1244)
MKPGFYIIMANKNNTVLYAGVTSNLPKRIYQHKKEIFELSFTSKYNIDKLVYVESFQEIGDAIFREKQIKAGSRQKIDLINSINPEWRDLTEDIKILNPF
<SPMU_RS10450--*Sphingomonas mucosissima*--73.86 (SEQ ID NO:1245)
MRERQPCVYILASGRYGLTVGVTSNLASRLMQHRDGTAGFTSRYGVRLRVYDVADTMESAITVEKRIKKWPRDYKFNLIERENPFWEDLATSLGLPPL
<SPMU_RS01135--*Sphingomonas mucosissima*--72.73 (SEQ ID NO:1246)
MLDLFQHPWGRVSEGEVAKQPCVYLLAGQRLGTLYIGVTSNLVARLWQHRSGTTAGFTSRYGVFLLVRYEFFADMPSSAIAREKQLKRWHRQWKINLIESDNP
WVDLAPALGLEPVARVRDGC
<SPDO_RS09860--*Sphingomonas dokdonensis*--81.52 (SEQ ID NO:1247)
MARQPCVYILASGKHGTLYIGVTSNLLARLHQHRAGLIEGFTSRYGVHRLVWFEMADTMEEAIGREKQLKKWDRDWKLRVLEENNPWVDLAINLGFQVENT
ARMDRLRGNDGIVTDDR
<SPDO_RS09055--*Sphingomonas dokdonensis*--73.86 (SEQ ID NO:1248)
MRERQPCVYILASGRNGTLYIGVTSNLIQRLQHRDGTAGFTSRYGVQRLVYDVADTMEEAIAATEKRLKKWPRQYKLVNLIERENPFWEDLAI SLSPPL
<SPDO_RS07130--*Sphingomonas dokdonensis*--73.86 (SEQ ID NO:1249)
MKRQPCVYLLASQRRGTLYIGVTSNLVARLWQHRSGTTAGFTSRYGVFLLVRYEFVADMPSSAIAREKQLKRWHRQWKINLIESDNPWVDLAPSLGLEPLPSRA
RRDGC
<B7088_RS17175--*Pseudomonas aestusnigri*--79.55 (SEQ ID NO:1250)
MTKHPCVYILASQRNGTLYIGVTSNLLKRIWEHRNGVSGFTQRYRIHHLVYFEQAADMLSAISREKQLKRWKRQWKINLIEQSNPAWDDLYTRLAAPQERC
<CCO29_RS02660--*Polaribacter haliotis*--71.88 (SEQ ID NO:1251)
MKTIHQYVYLLASKIRGTLYIGVTDNLQRRVYEHKMGIKGFTQKYGVNRLVYFETYQNI EEAITRENNMKWKRDWKIKLIEEENIQWLDLAKDWYDDIL
<CCO29_RS05865--*Polaribacter haliotis*--69.47 (SEQ ID NO:1252)
MKPGFYIILTNKNNTLYIGVTADLEQRIKQHKSKINKKSFTARYNLSKLVYEEAFQMI GDAIAREKQLKSGNRAKKIALIESMNPWEELFYPL
<CCO29_RS01875--*Polaribacter haliotis*--57.65 (SEQ ID NO:1253)
MITNKNKTVLVGLGTNNLKDRLYYHSHPEANSKHFSHKYCNCKYLIYFEHYQDVEIAINREKQIKKWNKKKEFLISTKNPDWNFLNDEL
<CDA63_RS14090--*Hymenobacter amundsenii*--61.18 (SEQ ID NO:1254)
MIFSMYVYIILTNQQTLYIGVTDNLTRRLYEHSSGRGDAGFTGRYQADLLVYFEIAPDATQAIAREKQLKGWNRKKDALISAFNPWKADLETWEA
<CDQ91_RS18175--*Sphingopyxis witflariensis*--75.27 (SEQ ID NO:1255)
MRDVFQPCYIILASQRNGTLVGVTSNLIQRIWQHNRNGIAEGFSKHYRAYRLVHFELLGTMELAITREKQLKNWHRQWKINLIEADNPEWRDLAFDYGAGPIG
QTMDAETSSA
<CDQ91_RS14960--*Sphingopyxis witflariensis*--69.66 (SEQ ID NO:1256)
MHADFQPCAYIMASARYGTYLTGVTSHLIQRIQCQRDGSFDFGFTKRYGVKTLVWFELAGTMYGAI TREKQIKNWQRKWKIALIEANPDWRDLAEELGLPPLG
G
<CDQ91_RS14950--*Sphingopyxis witflariensis*--68.54 (SEQ ID NO:1257)
MHADFQPCAYIMTSARQGTLYTGVTSHLIQRIVQHRDGSFDFGFTKRYGVKTLVWFELAGTMHGAI TREKQIKNWQRKWKIDLIEANPDWRDLAEELGLPPLG
E
<CDQ91_RS12490--*Sphingopyxis witflariensis*--75.00 (SEQ ID NO:1258)
MREERQPAVYMMSNRKDGVLVGVTSDLIARAWQHRQKQHGFSARYRCDKLVWFELHGMHSAITREKQIKAGSRAKKVALIEVENPEWRDLYPELTGG
<CDQ91_RS11940--*Sphingopyxis witflariensis*--67.44 (SEQ ID NO:1259)
MEQTYHVYIILASGRNGTIYIGVTGNLSQRIWQHNRNGEGSKFAAKYDVHRLVHAEAFADVNEAIAAREKAMKKWRRDWKLELIERDNPQWLDLYEHLNG
<CEP68_RS22110--*Brevundimonas vesicularis*--69.89 (SEQ ID NO:1260)
MPTHRFTIATYIMASGRNGTLYTGMTANLEARVWKHRSVFEFGFSASHGCTRLVWFERHAWVVEAIAAREKRIKKWLRDWKCLKLIESENPEWRDLAEDWFPVND
PNWRPPVEED
<CEP68_RS30415--*Brevundimonas vesicularis*--64.84 (SEQ ID NO:1261)
MSERRPIATYIQASRPHGVITYGMTSNLYERGHQHRGTGAFEGFASRYGCGMLVWYEQFELVTNAIRQEKALKCWNRVWKLELIERTNPEWVDLYPSLCGWAP
DPRVKPEDDGEGSVA AFLKSLESGA
<BSZ04_RS22315--*Vibrio rotiferianus*--74.71 (SEQ ID NO:1262)
MDKVPVYIILTSNNNAVLYIGVTSQLPQRIWQHKNQIVGFTKKYNVHKLIVYFELFEDIPTAISREKQLKQWKREWKLELVRESNPNFLDISSSVQ
<VVCECT4999_RS15350--*Vibrio vulnificus*--71.59 (SEQ ID NO:1263)
MSLPCVYIILASQPYGTLYIGVTDLMKRIWQHKNGETDGFSSKYRVVHLVYFEQFEAMTDAILREKRLKKWNRQWKIQLIESANPHWLDLYESLRTTA
<B0A81_RS13280--*Flavobacterium plurextorum*--63.41 (SEQ ID NO:1264)
MESLQGYHTYYIYIITNKFKTLFYTGVTNNLRVRLCQHKENVTGENKFTFSKYKVEFLLYYEKFTWIQEAIIREKEIKGWRDCKKIELIKTINPDLDLNFHIF
LKIIDLLI
<VCASEI_RS01460--*Vibrio casei*--75.00 (SEQ ID NO:1265)
MNKQPCVYIILSSNLHRSYIGVTSQQLQVRVWQHKNQVDGFSRRYNTVLLVYEQYSNMELAI VREKQLKKWKREWKNELIQKVPNPNFTDLYFDL
<B0A64_RS08865--*Flavobacterium araucanum*--63.16 (SEQ ID NO:1266)
MKPGFYIITNKYQTVVYTGVTSNLPQRILQHKNNKYPKSFARSYDVNILVYEQFQWIGDGTREKQIKAGYREAKNDLIRSINPTWKDLFEEIENIMIM
<B0A64_RS18855--*Flavobacterium araucanum*--68.35 (SEQ ID NO:1267)
MQLQEGYHTYYIYIITNKSKTYFTYTGVTNNLRIRLSQHKDNIMKENKTFASYNVYLLYYEKFTWIQEAISREKEVKDWRREKKIELIKTINPNLDFLNDMF
A
<AREALGSMS7_RS00205--*Arenibacter algicola*--74.47 (SEQ ID NO:1268)
MHKTIHQYVYIILTNKNKGTLYIGMTNDLQRRYIEHKNKLIDGFTKKYNLKLIVYFENHQFVEDAIRREKNLKKWKRWDKINLISKDNPQWQDLASNWYDQ
<CHH26_RS05825--*Erythrobacter flavus*--67.78 (SEQ ID NO:1269)
MTGERGGYVYIMADRYRGTYTGVTADIAARTVQHSEGRGGVFTSRYALTRLVYVERHDDIEAAIAREKAVKKWNRRAWKIRAEIEQNPEWNDLFDTLNM
<CHH26_RS11525--*Erythrobacter flavus*--62.22 (SEQ ID NO:1270)
MDRETHGGWVYIMANRYRGALYGVGTADLASRVSQHCDGSGSDY CARRGLTRLVWAEFSNPIDAAIAHEKRLKRWRCWKALIEKANPDWNLFDTLN

FIG. 11NN

<B7H23_RS08820--Notoacmeibacter_marinus--71.91 (SEQ ID NO:1271)
MTGYVYMLASQKNGTLYIGVTNSLARRIPEHREATKGFATARYGCHRLVWYEEHFDMDAIQREKTMKRWPRAWRVALIEEHNPDWNEFLVGVGWGE
<clem_RS02690--Legionella_clemsonensis--81.11 (SEQ ID NO:1272)
MKDSYVYILASQKNGTLYIGLTSDLIKRIWEHKNKFVPGFTNKYNVTILVYVEIFNDITLAAGREKRLKEWKQWKIDLIQKHNPWKDLYHEIVQ
<clem_RS08830--Legionella_clemsonensis--64.04 (SEQ ID NO:1273)
MYTESYVYLLTNKHLNVLYTRVTNNLIRRIYEHRNKLVRGFTQKYNVDRLVYEEACENIIVAIAREKQIKGWSRKKKDALISVANPQWEDLYRSLY
<B6S08_RS00975--Oceanimonas_doudoroffii--84.95 (SEQ ID NO:1274)
MPKQPAVYMLTNKRNGTLYIGVTSNLRVVRVQHKQEAEGFSARYGLHWLVYFELHDDMYAAITREKQLKKWNRAWKRLRIETQNPQWGRDLWEEII
<B6S08_RS00910--Oceanimonas_doudoroffii--83.70 (SEQ ID NO:1275)
MPKAPAVYIIVTNKRNGTLYIGVTSNLRVQRIWQHROHEADGFTAKYDLHRLIYFESHSSMYDAITREKQLKRWKRDWKIRLIEEFNPHWQDLWNDII
<CHT99_RS19085--Sphingobacterium_cellulitidis--57.47 (SEQ ID NO:1276)
MGHIYFVYIITNKNKVLVYIGVTNSLKRRLHEHQFDDGYFNSFTKKYVVFYLYIELFSDVNDAINRETQLKGWTRKKKNALINKKNPNWDFLNNEI
<CHU93_RS12355--Sandarakinorhabdus_cyanobacteriorum--59.30 (SEQ ID NO:1277)
MARGGWYIIMANRYRGGMVGVGTADVLRRVQQHSAGAGSIHVDFAKLRLVYAEHRDEIERAIAREKLKVRREWKPFALIGADNPGWDDLWDTWFTPSEQQOE
<CHU95_RS05815--Niveispirillum_lacus--76.40 (SEQ ID NO:1278)
MPGYVYILASARNGTLYIGVTSMDGRRIWEHREGLVEGFAKRYGIKTLVYVEAFERIDDAIHREKRLKKWNRAWKLELIESSNPDWNDLSSTLL
<CHU95_RS21685--Niveispirillum_lacus--74.42 (SEQ ID NO:1279)
MERVGYYIIMANRRNGTLYIGWTDLTRRVYEHNRGLLPGFTRKYGCKFLVYELFGDPGGAVRREKQLKEWQRAWKIRLIEQDNPLWVDLYPTLA
<BSZ36_RS03410--Rubricoccus_marinus--80.40 (SEQ ID NO:1280)
MVQKSAVYILASQRNGTLYIGVTSDLVKVRQHKEGAHEGFSTRYGVHRLVYFEQHPISIVDAIAREKRLKKWRRAWKLDLIEGFNPGWEDLYERLVFGE
<BSZ36_RS10005--Rubricoccus_marinus--66.67 (SEQ ID NO:1281)
MKPMWYILANRYDVTLYVGVTDIERRLFETRSGRSGFTSRYRVTKLVYAEHRWARDIAAREKEIKGWRRAKKLALVQRENPDRLDLSPEASGGTA
<BSZ36_RS03815--Rubricoccus_marinus--60.00 (SEQ ID NO:1282)
MPSRRSVWYILASRTRVLYTGVTTNNIERRLNEHREGAPGSFTARYRARRLVHLEEYPNARDIAAREKQIKAWRREKKRALIEAENPEWRDLAAPEAER
<BSZ36_RS16025--Rubricoccus_marinus--63.53 (SEQ ID NO:1283)
MPTRDHQYVYILASQSRRLYTGVTDNLARVEEHRAGVGSFATTRYKIHRLVHFEEHNRVDAIQREIEIKAWRREKKRLIESHSAGWDLAPSSPAAR
<BEUL_RS03935--Bifidobacterium_eulemuris--69.41 (SEQ ID NO:1284)
MYVYVYILANTSNSVIYIGMTSDIHRMFEHKNKLIDGFTKQYNVHKLVYAEYPTAIEAIAAREKQLKKWRDKKNALVNQNPWLWDDLSLVM
<BAQU_RS05460--Bifidobacterium_aquikefiri--65.88 (SEQ ID NO:1285)
MYVYVYMLTNQHTVLYIGVTNNIMRRLQEHRRGEGRAFTAQYKLTKVYLETSESIMAAIEREKQLKHWRSRTKKDALINAKNPQWKDSLQW
<CIT40_RS29515--Bradyrhizobium_amphicarpaceae--79.78 (SEQ ID NO:1286)
MGARHYVYVYILASKIGTLYIGVTNDLIRVAQHKLKLIESFTEKYNVRLVYFEQFDDPENAIKREKRLKKWNRAWKVRLIEQHNPWDDLYHELGPFP
<CIT40_RS20155--Bradyrhizobium_amphicarpaceae--76.34 (SEQ ID NO:1287)
MAYVYVYLLASKKYGTLYLGVTDIVRRIYEHKSKVAVGFSKRYSVDRLVWYIYDDPITAITREKELKKWRREWKIRLIEEKNPEWIDLYPQIAV
<CA834_RS09075--Winogradskyella_aurantia--66.28 (SEQ ID NO:1288)
MKRAYHNYVYILTNKPNGTLYIGVTGGLDDRMERHLDGEGSKFASKYKLKLIVYEEFQFMDDAIKREKQLKNWHRQWKINLIERENPDWKNLWKPLGNN
<CJO11_RS00400--Altererythrobacter_mangrovi--67.78 (SEQ ID NO:1289)
MDRENRGWYIMTDYRGTYIGVTSDLASRIHQHRTGTGSDFCARYRLTRLVWAEFGESITECITHEKRIKRWHRWKIEMIERANPSWDDLDFSLV
<CJO11_RS12815--Altererythrobacter_mangrovi--63.33 (SEQ ID NO:1290)
MGWVYIMTNRGGLVLYIGVTADLPARVMQHRTGKGSFACRRYGLDRILVCAESHTNIVEAISREKAMKAWKRAWKIELIEKANPGWGDPLFGHLL
<CIT37_RS26200--Bradyrhizobium_ottawaense--76.40 (SEQ ID NO:1291)
MAYVYVYILASKKHGTLYIGVTNDLVRRVYEHKTKAVPGFTTKYGVDKLVLFYIYDDPTTAIAREKELKKWRWDKTRLIEEQNPWDDLYPGIAN
<CIT37_RS26630--Bradyrhizobium_ottawaense--74.16 (SEQ ID NO:1292)
MPTAFFVYILASQRNGTLYIGVTNDLARRMAAHKAKRVPGFTKQYGVTLVYFESYESILEARAREHSLKRWRAWKLIKIEDLNPDWDLTDLNLSAN
<CLC93_RS01110--Paraferriromonas_sedimenticola--76.34 (SEQ ID NO:1293)
MKQPCIYILSSQPNGTIYIGVTSDLVKRIWQHKNKFVSGFTEHYEVNRLVYELCDDMEQAILREKQLKKWNRAWKIRLIEKQNPQWADLYSQIL
<BSZ37_RS18540--Rubrivirga_marina--64.71 (SEQ ID NO:1294)
MGAARTYVWYMLTNRSGTLYVGVTTNNLERRLAHVASGDERFTGRYALDRILIYVESYPEARDGIAREKQLKGWRRSKKVALVNAANPSWRDLRGRAELP
<CIT31_RS04440--Mesorhizobium_wenxiniae--66.29 (SEQ ID NO:1295)
MTGYAYMTASQKRGTYIGVTNDLGRMPHKSSEGSFRFTSRYGVQRLVWYEEHFDIRDAIQREKSLKRWPRQWKIELIEKANPEWFELFRGTGW
<CIT31_RS19370--Mesorhizobium_wenxiniae--65.56 (SEQ ID NO:1296)
MTGYAYMTASQKRGTYIGVTNDLGRMPHRSQGSFRFTSRYGVQRLVWYEEHFAIQREKSLKRWPRQWKIELIEKTNPEWFELFRGRGGRSVLLRRSTAA
R
<CJF12_RS07405--Chryseobacterium_piperi--55.17 (SEQ ID NO:1297)
MKTLGTHNYVYIITNRIKTVLYIGVTNDLKARLYWHKNPEAGNIHFTTKYKCFYLLYYEHFQDVEQAILREKQIKGYSRMKKENLINRFPNPEWNFLNETIE
<BA173_RS05505--Rickettsia_sp._MEAM1_(Bemisia_tabaci)--77.65 (SEQ ID NO:1298)
MYWVYILCSDRNGTLYIGVTNNILRRAYEHKQQVIKGFTSKYNI IKLVYTEEFADIEALAREKALKKWNRSWKIKLIESINPKWNDLGECISGFPTTRE
<BA173_RS05075--Rickettsia_sp._MEAM1_(Bemisia_tabaci)--75.29 (SEQ ID NO:1299)
MYWVYILCSDRNGTLYIGVTNNILRRAYEHKQKVIKGFISKYNI IKLVYTEEFADIEALAREKALKKWNRSWKIKLIESINPKWNDLGECISGFPTTRE
<BA173_RS00285--Rickettsia_sp._MEAM1_(Bemisia_tabaci)--85.25 (SEQ ID NO:1300)
MITNKRNGTLYVGVTSNLTIRIYEHKNELLKGFSSKYDCKMLVYELHDTMISAITREKQIYY
<CJD35_RS01395--Sphingobium_hydrophobicum--77.27 (SEQ ID NO:1301)
MRIRQPCVYILASQRNGTLYIGVTSDLIARLHQHRTSAVPGFTRDYGVNRLMWFEQHGTMDEAILREKRLKKWNRAWKISMI EELNPHWDDLAI GLGFERLPD
HATPRVTRPNRT
<CJD35_RS10915--Sphingobium_hydrophobicum--74.71 (SEQ ID NO:1302)
MPDLIRHPFGASCAMSNQPCVYILASGRHGTLYIGVTGHLVERVHQHRTGIVAGFTSRHGVRRVLVWFEHQDPDPTAIALEKKLKKWRDWDKVALIEKDNPFW
EDRAIMLGFPPLERG
<CJD35_RS04910--Sphingobium_hydrophobicum--70.65 (SEQ ID NO:1303)
MKQPCVYILASQPYGTLYIGVTSNLI GRMLQHRTDALAGFTSRYGVHRLVFEACDTMEQAILREKQLKRWHRQWKINLIESENPDWDLAVNLGLPPVHPRL
THDGC

FIG. 1100

<CJD35_RS01590--Sphingobium_hydrophobicum--74.70 (SEQ ID NO:1304)
MAGRRNGTLYLGVTSDLAGRAYQHRNALVEGFTKRHGCTMLVWFEAYDELQEARLRELQMKKWKRAWKIELIERENPQWRDLFETLF
<CJD35_RS09345--Sphingobium_hydrophobicum--74.16 (SEQ ID NO:1305)
MTKPGYVYLMASQNGTLYLGVTSDLVQRAWQHRNGFGGFEFSAKYRCHLLVWYEAyddIQEARQRELRMKKWERKWKLRVIEEMNPTWRDLFEDIAT
<CJD35_RS07465--Sphingobium_hydrophobicum--61.11 (SEQ ID NO:1306)
MRGGWVYVMTNRAYGVLYVGVTSLLPHRAMQHREGIGSDFCRRYRLDRLVYAEERHETIMDAIAREKAMKAWKREWKVELIESINPKWVDFDLHLHD
<CJD35_RS05770--Sphingobium_hydrophobicum--60.71 (SEQ ID NO:1307)
MPFTVTPAKAGVSGADALQDLMARGGYTYVMTNKPGRGVLYIGVTADIATRAQHHRKGTGSAFCKKYGLTRLVLVEPHDDIRLAIAREKALKAWKREWKIRLVE
EANPDWRDMSDVIF
<CJD35_RS07245--Sphingobium_hydrophobicum--59.52 (SEQ ID NO:1308)
MNFTVTPAKAGVSGADALQDLMARGGYTYVMTNKPGRGVLYIGVTADIATRAQHHRNGTGSAFCKKYGLTRLVLVEPHDDIRLAIAREKALKAWKREWKIRLVE
EANPDWRDMSDVIF
<SJ05684_RS14290--Ensifer_sojae_CCBau_05684--72.04 (SEQ ID NO:1309)
MAGYVYIVTNQKRGTYLIGVTSNLERRIFEHREGLTPGFASKYGCNRLVWYBEHLQIGTAIQREKSLKRWYRQWKIELIETMNPDRDLYFELW
<CK510_RS12060--Rhizobium_elsteri_CCALA_953--68.13 (SEQ ID NO:1310)
MNNYFVYIMTNHSTLYTGITNDLNRRVYEHKQKLVPGFTQKYNI TKLVYFEETSDVNDIAAREKQIKGWLREKKIVLIESMNSEWKDLSAGWYD
<AWC36_RS11920--Brenneria_goodwinii--76.34 (SEQ ID NO:1311)
MSKQPCYIILSNRPYGSGLYVGVSSHLKRIWQHKNHI TGGFSQKYDLNMLVYEFFEDINNAILREKQLKKWKRAWKIRLIESINPDWEDLYSYII
<CLD06_RS23665--Rhizobium_endosymbiont_of_Drosophila_subpulchrella--74.71 (SEQ ID NO:1312)
MKLSYIYILTNKANSTLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVYFEFEDISLAIAERERLKEWPRQWKINLINQKNPSWVDLYLALR
<CNQ84_RS07800--Pseudomonas_abyssi--78.65 (SEQ ID NO:1313)
MTKQPCVYIILASQRNGTLYVGVTSNLLKRIWEHRNGVSGFTQRYRIHHLVYFEQAADMLSAIAAREKQLKRWKRKWKINLIEQNNPAWDDLYTTLTAPKQALL
<CMV14_RS23665--Rhizobium_dicambivorans--75.00 (SEQ ID NO:1314)
MEKQPVYIILASGKHGRLYVGVTSNLLPGRHLQHRESLIKGFSTRYGVTRLVHFELFGDMIMAIAREKQLKEWRDWNKLNIEASNPDWIDLAEGLGFDLSGPE
<CMV14_RS14285--Rhizobium_dicambivorans--78.72 (SEQ ID NO:1315)
MLNLIQHPLPQRPPIERQPCVHILASGRHGTIYIGVTSNLMGRHLQHRTATPFGFTSRDYVFRDLPLWEDRALDLGFPRQS
<GJW30_RS21725--Variibacter_gotjawalensis--76.67 (SEQ ID NO:1316)
MIYVYIILASARRGTYLIGVTDNLIIRVYEHRESAVPGFTKRHDVKRLVYFEVHDDISEAIEREKQIKRWREWKIKLVEEQNEDWHDWPSLVASGEVSK
<GJW30_RS04385--Variibacter_gotjawalensis--73.33 (SEQ ID NO:1317)
MGGWVYFMSNGRDGILYVGVMSDLIRRVYEHREGLRKGFSGQYGLKRLVYFERHDDIRDAIQRETRIKKWPRAWKVRILHASNPNWDDLYFTIAA
<SVA_RS14160--Sulfurifustis_variabilis--68.24 (SEQ ID NO:1318)
MQERCYFVYILTNPSNRVLYTGVTSDLERRLAEHRDGLTPGFASRYRCRKLWYETIPDIAAAIAAREKQIKGGSRAKKIALIETMNPWTQDLAPSAKQ
<MgSA37_RS17555--Mucilaginibacter_gotjawali--64.13 (SEQ ID NO:1319)
MGFYQYGGCVYIMTNKMHYALYVGVSSDLTGRVWEHKNNVYPGYTARYNCHKLIVYMFYSRIEBAI GVEKTLKAGNRKNKIKLINAMNPKWDLDDLMRE
<MgSA37_RS12650--Mucilaginibacter_gotjawali--59.78 (SEQ ID NO:1320)
MFPKGGCVYIMTNKMHYALYVGVSSDLTGRVWEHKNNVYPGYTARYNCHKLIVYMFYSRIEBAI GVEKTLKAGNRKNKIKLINAMNPKWDLDDLMRE
<CFK21_RS02915--Thiohalobacter_thiocyanaticus--81.72 (SEQ ID NO:1321)
MKRQPCVYIILASQRNGTLYTGVTSDLVKRVWEHRTHAVEGFTRKYGVTMLVWYEWHGNDMSAICREKAIKKWRRAWKLAIEKMNQWRDLYSDDL
<CFK21_RS14540--Thiohalobacter_thiocyanaticus--76.34 (SEQ ID NO:1322)
MEKKPAVYIILCNRPNGTYLIGVTSNLPKRIWEHRSKTVKGFTAKYNNVRLIWFELFETMDLAIEREKQLKGRPRKAKVALIEDGNPVWRDLFEEICG
<CFK21_RS09090--Thiohalobacter_thiocyanaticus--75.27 (SEQ ID NO:1323)
MEKQPAVYIILCNRPNGTYLIGVTSNLPKRIWEHRSKTVKGFTAKYNNVRLIWFELFETMDLAIEREKQLKGRPRKAKVALIEDGNPVWRDLFEEICG
<CFK21_RS10350--Thiohalobacter_thiocyanaticus--74.44 (SEQ ID NO:1324)
MYILCNRPNGTYLIGVTSNLPKRIWEHRSKTVKGFTAKYNNVRLIWFELFETMDLAIEREKQLKGRPRKAKVALIEDGNPVWRDLFEEICG
<CA729_RS11895--Calothrix_parasitica_NIES-267--69.32 (SEQ ID NO:1325)
MNKQYIYIMTNQYNKVLTYGTNDLQRRVSEHKSCLIEGFTKKYNI TKLVYIEIYEDSYTAISREKQIKAGSRQKKIDLNVQNMHQWKLDSNEL
<CA741_RS24145--Fremyella_diplosiphon_NIES-3275--69.66 (SEQ ID NO:1326)
MNQYIYIMTNQYNKVLTYGTNDLQRRVSEHKSCLIEGFTKKYNI TKLVYIEIYEDSYTAISREKQIKAGSRQKKIDLNVQNMHQWKLDSNEL
<MKLM6_RS00300--Methylobionas_koyamae--83.70 (SEQ ID NO:1327)
MKQPAVYIILASQRNGTLYIGVTSNLPKRIWEHRSKTVKGFTAKYNNVRLIWFELFETMDLAIEREKQLKGRPRKAKVALIEDGNPVWRDLFEEICG
<CN311_RS28185--Mesorhizobium_sanjuanii--66.67 (SEQ ID NO:1328)
MTGYVYMTASQKRGTYIYIGVTNDLGRMPPEHKSQSGTSRFTSQYGVQRLVWYEEYFDITDAIQREKSLKRWPRQLKIELIDKTNPEWELFRGTGW
<CO657_RS14610--Rhizobium_acidisoli--71.28 (SEQ ID NO:1329)
MKGYVYIILASGRNGTYLIGVTRDLAVRLYEHQNELTPGFTSKYGVKTLVWFEEHDLTTAI TREKTIKKWPRQWKNLIEQENPEWEDISFHLSSL
<CRI93_RS01190--Longimonas_halophila--64.71 (SEQ ID NO:1330)
MARTYHVYIMSNRHTVLYIGVTGNLAHRVHAHQDGTGSAFTRRYNVDLIYAEPYADVDEALRREKQLKRWTRAKKLGLIRRQNPQLTIQGPYG
<BB934_RS05850--Microvirga_ossetica--78.49 (SEQ ID NO:1331)
MAYHVYMLASGRHGTLYVGVTDNLAIRIHEHKEKLIKGFSTQYGVTRLVWYETIERIDEAIAAREKALKKWHNRWKIRLIEEMNPEWEDLYLTNQ
<BB934_RS17560--Microvirga_ossetica--81.18 (SEQ ID NO:1332)
MAEAFVYIILATRKDGPLYVGITSNLPKRVFEHKTTHAIPGFTARYNVDRVLVYFEVFEFDPVTASREKQLKKWRRAWKVELIERDNPEWRDLAEFIP
<BB934_RS27000--Microvirga_ossetica--67.47 (SEQ ID NO:1333)
MTNRRNGTYLIGVTSNLPKRIWEHRSGVTPGFTGQYGRKRLVWYEHFPEIRDAIQRETMTKHWPRAWKVRILHAMNPDWDDLYELLNG
<BFP76_RS03140--Amylibacter_kogurei--56.32 (SEQ ID NO:1334)
MKDCYTYIIVTNRRHRLCAVGTRDLVNRPLKLGKDNPGASAWDEKMSRLVWYEHKHTLDAALEREERIANLRDWFYLVVERENPAWRDLAQRGVRETHTF
NAAIGGGARAH
<CJ305_RS15875--Leeuwenhoekiella_nanhaiensis--68.24 (SEQ ID NO:1335)
MEELFYTYIILTNKNTVLYIGVTADLKKRLQHTQKSNKGFTGKYNVEKLVYFESTPYVNNAIHREKQLKKWNREWKENLISDMNPDWDLWSMLE
<CSW64_RS00760--Caulobacter_mirabilis--74.44 (SEQ ID NO:1336)
MSFHVYIVASRRNGTYLIGSTNLAIRVWEHREKLRPFTSRYGVLVWCRVCDTREAAKDLERRMKEWRRSWKRLIEESNPDWDLDETINQ
<CSW64_RS19075--Caulobacter_mirabilis--72.29 (SEQ ID NO:1337)
MASGPYGTLYVGHTESLAKRAWQHREKQLSGFTAKYEVVRLVWYEHVDTREAAKAREQKLLKWNRAWKIRLIEALNPQWDLDTLNA

FIG. 11PP

<CSW64_RS16780--Caulobacter_mirabilis--68.60 (SEQ ID NO:1338)
MRRRHVYVLSRRNGTLYVGVTGNLAYRVWLHREGRGATFTTTRYGVKLLVWYVAFDEPLDGI AF EKRLKRWRRAGKLALIEKDNP TWRDLYETLNG
<CQW49_RS02660--Methylosinus_trichosporium_OB3b--69.57 (SEQ ID NO:1339)
MKQPAVYIIASSRNGTLYTGITSNLSQRAYQHREGLTPGFTTRYGCKLLVCYEEYERMVEAIAREKQ LKGGSR RQKLALIESMNPISWSDLYEDLAR
<CR152_RS14615--Massilia_violaceinigra--75.28 (SEQ ID NO:1340)
MEKYSVYMLASDPYGTLYTGVTSDLVKRVWEHRNAVVEGFTTKRYDVHKLWVYEMHTDIMEA IKREKQIKKWRNWKVNLVQAGNP TWRDLYDFTAD
<CR103_RS18540--Massilia_psychrophila--77.53 (SEQ ID NO:1341)
MEVCRGSDVDRASVYMLASECYGTLYVGVTSDLVKRIWQHREGGLADGFTTKKYKVKRLVWFETHGDIIEAISREKQLKKWARAWKVALIQEANPDWRDLYDDF
TA
<CJ014_RS26645--Pleomorphomonas_carboxyditropha--68.42 (SEQ ID NO:1342)
MNPGLWYIVTNKPGQVLYTGVTSDLEGRIWEHRNHVYRGFTDSYNCTRLVWYEDYGDITDAIQREKQLKHWKRAWKIRLIEEMNPEWKDLFETIMS
<CJ014_RS22770--Pleomorphomonas_carboxyditropha--66.32 (SEQ ID NO:1343)
MNPGLWYIVTNKPGQVLYTGVTSDLEGRIWEHRNAVHVGFTHSYNCTRLVWYEDYGDITDAIYREKRLKTWKRAWKIGLIEEMNPGWNDLFETIMS
<CVM73_RS38135--Bradyrhizobium_forestalis--79.78 (SEQ ID NO:1344)
MGASHYVYVYIIASKIGGTYIGVTNDLIRRVAEHKS KLI ESFTEKYDVARLVYFEQFDDPENAIKREKRLKKWNRRAWKIRLIEQHNPNDWDLYEIAGPP
<CVM73_RS07635--Bradyrhizobium_forestalis--76.47 (SEQ ID NO:1345)
MHTAFFVYIIASKRNGTLYIGVTNDLARRMAAHKAKLVPGFTKQYGVTLVYFETYESILEARAREHSLKRWRAWKLLIEDLNSDWRDLTDELSALAG
<CLV73_RS02840--Chryseobacterium_geocarposphaerae--59.77 (SEQ ID NO:1346)
MYKGTHTNYYIILTNKNTVLYIGVTNSLKERLYWHQNPDPVTRHFTHKYSCCHLIYYEDFKDIDEAIKREKQLKGRREKKEILINEFNPEWKFLNNEI
<CLV45_RS02730--Hymenobacter_chitinivorans_DSM_11115--61.18 (SEQ ID NO:1347)
MYVYLLTNPAHTVLYTGVTNDLQRRLYEHGQNLGTAGNFTGRYQC NLLVYFEVCADAVQAI GREKEIKGWSRAKKEALITTTNPEWKVIDLQTWSG
<CLV45_RS12455--Hymenobacter_chitinivorans_DSM_11115--55.17 (SEQ ID NO:1348)
MPYVYVLTNPAKTVLYIGITNNLPERMAQHYANRSTDTQTFAGRYFCYNLYVFEETD SKAAIAREKELKGWTRAKKEALITATNAEAWFLQA
<CLV57_RS15785--Mucilaginibacter_auburnensis--60.87 (SEQ ID NO:1349)
MRQYNFYTYILTTASKKVLYVGVTNNLERRLYEHYYGVGNKDSFTNKYKCYLVWYERHQYVNHAIEREKEIKGWLRAKKEKLINEENPQWRFMNSDIMEWPP
KKICHSESEESYTPC
<CDH50_RS08485--Confluentibacter_lentus--61.18 (SEQ ID NO:1350)
MKKCYCYILTNKNRSVLYIGYTDNLQKRISQHQKGTGAVFTKKYNVVDLVYEEFEEMKVAKSREKQLKNW HKDWKWNLVKETNP KLETILILN
<CDH49_RS12145--Confluentibacter_citreus--61.18 (SEQ ID NO:1351)
MKKCYCYILTNKNRSVLYIGYTDNLQKRISQHQKGTGAVFTKKYNVVDLVYEEFEEMKVAKSREKQLKNW HKDWKWNLVKETNP KLETILILN
<CSH92_RS12210--Alteromonas_flava--70.79 (SEQ ID NO:1352)
MMKQPCVYIITNTYNTTVYIGVTGNLIQRIYQHKTKQINGFSKNYNLQKLVYFEQFTSMENAIQREKRLKNWRRRAWKDELIASINPGWEDLYEILLG
<CH352_RS05115--Leptospira_hartskeerlii--56.41 (SEQ ID NO:1353)
MSFVYIILCSDLSFYTGCTCDLDRRLYEHNAGIGSLYTKVRRPVCLIFYEFVSVAEEAFRRERQIKNWSRAKKISLIQGD FKL LSEL SRPSTSPSLTLRTSSG
RAHSD
<CWM47_RS18480--Spirosoma_pollinicola--58.33 (SEQ ID NO:1354)
MREYAFYVYIVTNPEKTVLYTGMTNDLIRRLQEHYESRGQPEKFAGKYCHQLIYWEQH QYVLNAIAREKEIKGWRREKKVELINSINPEWCFLNDEIQG
<CWD77_RS03900--Rhodohalobacter_barkolensis--72.09 (SEQ ID NO:1355)
MKCYTYIITNRKSGTIYTGMTNDLAVRIYQHKQKLVPGFSKKYNLNLKLVWFEHRT PQDAIDREKQIKGWKRYKKIALIEEMNPQWMDLYKGLF
<CW740_RS10280--Kangiella_profundi--77.42 (SEQ ID NO:1356)
MDKQFYVYVMASKNGTLYIGITSNLV KRVWEHKNNAVEGFTKKYEVHQLVYEEQHESYS AIQREKRLKEWKRWKWL DLIEGMNPEWKDLYDEII
<CW740_RS08020--Kangiella_profundi--73.03 (SEQ ID NO:1357)
MKEYYYIMASKRDGVLYVGVTGNLQQRVHQHREGLVEGFSKQYKTKRLVYFESSDIKACIEREKQLKNWKRQWKELIEKINPQWEDLWSSIAS
<CW740_RS03210--Kangiella_profundi--56.67 (SEQ ID NO:1358)
MKKGYVYILTNPRKTVFYTGATNNLGKQLFEQKNRISERLRSYGCNELVYIEITESLEKAREREEIKSECRIGKETLIERFNP EWKDLSGDLLAVKKR
<CSW08_RS01525--Confluentibacter_flavum--63.95 (SEQ ID NO:1359)
MKRAYHNWYVYILTNKPRGTLYIGVTGGIDDRMERHVAGEGSKFTKKYRLKQLVYFEFQYIDDAIKREKQLKNW HREW KINLI EADNPEWENLWVALDPKTI
EGDKAKVILSCLTR
<CSW08_RS10560--Confluentibacter_flavum--61.18 (SEQ ID NO:1360)
MKKYCYILTNKNRIVLYIGYTDNLKKRINQHQKGAGALFTKKYSVVDLVYEEFEEMKIAKSREKQLKNW HKDWKWNLVKETNP KLETILILI
<BZG02_RS06295--Labilibaculum_filiiforme--66.29 (SEQ ID NO:1361)
MKGGYIYILSNKNRTTLYIGVTSDLD RRIAEHVEGEGSKFSSKYNLTDLVYFEFE DDIGDAISREKKLKNWKREWKMNLIKSLNPELKD LKMDFR
<BZG02_RS06300--Labilibaculum_filiiforme--65.17 (SEQ ID NO:1362)
MKGGYIYILSNKNRTTLYIGVTSDLD RRIAEHVEGEGSKFSSKYNLTDLVYFEFE DDIGDAILREKQLKNWRREWKMNLIKSLNPELKD LKMDFR
<CWS31_RS13265--Colwellia_echini--79.57 (SEQ ID NO:1363)
MKKPCVYILSSKNGTLYIGVTSNLQKRIWQHKNKVVS GFTEDYRINTLVYFEQFDDMENAI SREKVLKKWNRRAWKVRLIEKNPNWDL DYDGIL
<BZG01_RS03340--Labilibaculum_manganireducens--60.92 (SEQ ID NO:1364)
MKTYVYVYILTNKNKTVLYVGVTNNLQRR LSEHVSHVDHINAF TKRYNCHHLIYFEEHDDVNLAIEREKKLKGWRREKKNILISEFNPEWKFL EGEFFV
<BZG01_RS03360--Labilibaculum_manganireducens--59.77 (SEQ ID NO:1365)
MKTYVYVYILTNKNKTVLYVGVTNNLQRR LSEHVSHADHINAF TKRYNCHHLIYFEEHDDVNLAIEREKKLKGWRREKKNILISEFNPEWKFL EGEFFV
<CWS72_RS06670--Telmatospirillum_siberiense--72.09 (SEQ ID NO:1366)
MPGGWYVYFMNRPNGTLYVGVTSDLARRVWEHREGQADGFTKRYGLRRLVYVERHEDISTAIQREKLNKHSRAW E VALILKANPNWDDLYEQLM
<BD749_RS01090--Pontibacter_ramchanderi--58.89 (SEQ ID NO:1367)
MKRHNVFYVYITTNPGKTVLYVGVTNDLVYRL EQHKENRGKPEAFVGRYFCYKLLYERYTYVQHAI DREKELKSW SREKKMDLIK NENPYLRF LMDP
<SGCZBJ_RS06765--Caulobacter_zeae--67.82 (SEQ ID NO:1368)
MDLLFSVYIILASAPYGTLYCGHTDY LARRIWQHREEHFRGFTRRHGVHRLVWYENGDSREGGLIRERQIKKWNRAWKIRLIEAMNPHWDL DYETLNC
<C0039_RS11450--Halioglobus_lutimaris--73.12 (SEQ ID NO:1369)
MAKQAAVYIILASRSTGALYIGVTGNLMQRIWHHRNIRAE GFTSRYAVHKL VYFEMHRNV RDALAREKRLKRW HQTWRLRIIEEQNPQWMDLWDQVIQ
<CV013_RS02720--Altererythrobacter_flavus--72.73 (SEQ ID NO:1370)
MTANHRNGTLYTGVTSNLVKRIWQHREGTIEGFTQCYGCKRLVWFEHHATMEHAIARKKRIKAGPRKRKIALIETANPNWRDLWFDICR

FIG. 11QQ

<CV013_RS09675--*Altererythrobacter flavus*--63.53 (SEQ ID NO:1371)
MDREKGVWVYMMANRYRTIYIGVTSDDLPRRIHQHIAAGNSNFCPRYGLTRLVWAEGERDIAATCIAHEKRIKHWRREWKFALVEKGNPGWEDLSGWL
<CV013_RS14725--*Altererythrobacter flavus*--60.87 (SEQ ID NO:1372)
MKSGYVYILNRPMTGLYIGVTSDDLARMEQHKSAVSSFTKRYNCHRLVWFAQFENIHDARAFEHRMCKWNR
<C1H87_RS05285--*Flavivirga eckloniae*--73.86 (SEQ ID NO:1373)
MYKTIHQYYLYILNKTCTGLYIGVTDNLERRMFHEHKNKLIEGFTKKYGLYKLLYFETYQYVEDAIKREKNMKKWKRWQKIDLIEKENPNWKDLSKDWTYLIE
<C1S73_RS09320--*Vibrio jasicida*_090810c--71.26 (SEQ ID NO:1374)
MDKTPCVYIILTSKNSVLYIGVTSQPLQRIWQHKNQVVEGFTKKYNVHKLIVYFELFEGMASAINREKQLKQWKREWKLDLVRKTNPDFLDLSNEF
<B7R93_RS05545--*Salinibacter altiplanensis*--67.06 (SEQ ID NO:1375)
MGERYYAYIMASRSRVLYVGMNTNLLARSVREHKQKDADGFTARYNVTRLVWYRAFPRPQDAIAAEKKIKGWLKKKAAALIEKNPSWNLDLGPVMEEAR
<B7W82_RS05860--*Salinivenuus iranica*--72.09 (SEQ ID NO:1376)
MSQSRGGFVYILASKPNGTLYVGVTDIVRRVYQHRTGERSEFVEQYGVTRLVHLERFDQIEEAIKREKQLKPKWKRAWKLDLIRENNPAWEDLYEACRTYGR
ERGATGGLSAP
<B7W82_RS01085--*Salinivenuus iranica*--72.22 (SEQ ID NO:1377)
MATRGVYIILTNKPNGTLYVGVTSDDPIERIRRHKEGRGSQFAAQYNLTRLVHIEEHDTMPAALRREKRLKRWREWKLSLIREHNPTWEDRYEEVCRLLAGGG
RLP
<B7W82_RS14335--*Salinivenuus iranica*--62.35 (SEQ ID NO:1378)
MPPGPPYYVYILNSRSTVLYVGMNTNGLRRRLREHRESTGRSFEQYNTDILVYFEEYDSFQAAIEREKQLKRWREKKEALIRSANPDLEALSPPVE
<C1T31_RS09615--*Hanstruopera neustonica*--70.45 (SEQ ID NO:1379)
MHRNTHQYYLYILSNKYKGTLYIGVTDNLERRMFHEHKNKLVEGFTKTYGLDKLIYFETFYQYVNDAIKREKNMKKWKRWQKINII EEDNPWDDL SKDWPC LID
<C1T31_RS12045--*Hanstruopera neustonica*--68.24 (SEQ ID NO:1380)
MKMSYVYFLANKNNTVIYVGVTSDDLKRIYQHKTKAYKGFAKYNCCKLVYETFTDINQAIAREKQIKAGNRTRKMRLEADNPQWNDLSDGWLIFYFD
<C1637_RS19000--*Chryseobacterium lactis*--65.46 (SEQ ID NO:1381)
MKAGFIYIMTNKNHTTLYTGVTNLPKRVQQHKESYFSQSFTSRYNLHKLIVYWEAFQEIGDAIGREKQIKAGSRQKKLNINAMNPNWLDLADDIENIMDV
<C1637_RS18380--*Chryseobacterium lactis*--60.92 (SEQ ID NO:1382)
MKTGLGTHNYVYMLTNKVKTVLYTGVTNDLKNRLFHWQNEPAVGKSFSTSKYKCFYLIYFEHFDSEIEIAIAREKQIKGWTREKKENLITEFNPDWKFLNEEVE
<C1634_RS03850--*Chryseobacterium viscerum*--67.78 (SEQ ID NO:1383)
MKAGFYIMTNKNHTTLYTGVTNLPKRVQQHKESYFSQSFTSRYNLYILVYWEAFQEIGDAIFREKQIKAGSRQKKLDLINSINSDWKDLTGDEVIMDV
<CX42_RS02930--*Tabrizicola aquatica*--62.35 (SEQ ID NO:1384)
MAHFVYIMASRPGGAIYIGRTGNLAARVEAHRAGASVHTAKYKIRTLVWFECHEFNTSLQRRERAMKRWPRWKNALIAERNPNWQDVTSHIPV
<C1949_RS03750--*Pseudomonas oceani*--71.91 (SEQ ID NO:1385)
MTKQPCVYIILASKRNGTLYVGVTSNLLKRIWEHRNGMVSGFTQRYQIHHLVYLEQAADTLASAIAREKQLKHWKRWQKINLIEQTNPGDDLYTLTAPKQALL
<CLV41_RS17905--*Labrenzia marina*--75.28 (SEQ ID NO:1386)
MAFFVYILASRKRGTYLTGVTNDLARRVYEHREKTASGFTRRYQATQVLYEYFDPDPSAIAREKRLKRWSEWQKQAI EAFNPQWRDLYEDLNR
<CLV41_RS22450--*Labrenzia marina*--68.24 (SEQ ID NO:1387)
MKRYWVYMLASRKNNTLYTGVTNDLMMRIHEHQTGRGSKFATRYGALRLVYDEHPSAGLAIAREKAIKSWPRKWKIDLIEMLPDWYDLKRFLNS
<CWD90_RS16795--*Arenibacter catalasegens*--72.53 (SEQ ID NO:1388)
MHKTNHQYYVYILNKNKGTLYIGMTNMLQRIYEHKNKFTDGTNKYGLDRLVYFETHQFVDDAIKREKNLKKWKREWKINLIIKDNPDWQDLATDWQPO
<CWD90_RS08395--*Arenibacter catalasegens*--68.18 (SEQ ID NO:1389)
MKKSYVYLLTNKNNTVIYVGVTSNLLKRIYEHKIGIFHGFSTSRYNCKNLIYFEFSDIRKAIAREKQIKAGNRRRKEQLVNRDNPWEDLAI GWFIFYD
<CWD90_RS13050--*Arenibacter catalasegens*--55.29 (SEQ ID NO:1390)
MTKGYFYILSNSTNRLLYAGATRYVIKRLRLHELGMGVSFTKKYHVYKLIYFEEFEEIGDAFKREKQIKNWHKDWKNLIIKEMNPPELLDLRHRFE
<CWD91_RS19080--*Arenibacter hampeyensis*--70.59 (SEQ ID NO:1391)
MNGSYVYMLASRKNNTVIYVGVTSNLLKRIYEHKTVGFQGFASRYNCNLIYFEFSDIQKAIAREKQIKAGNRRRKEQLVNSNPWDLAEGWIFYD
<CWD91_RS14350--*Arenibacter hampeyensis*--56.99 (SEQ ID NO:1392)
MFKYKHFYIITNEHRTTLYVGMNTDIQRRLSQHYFDSQNSKKSFAKGYSCIYLVYEGFENPKDAIAREKEIKKWRREKKNLSI SAFNPKWETLSNHHI
<C3B73_RS01760--*Gracilimonas amylolytica*--64.21 (SEQ ID NO:1393)
MKKRFYVYMLSNSTSKMIYTGMTGLERRIYEHKNKLVEGFTKKYNLHLYVYDWDTHDVGYAIQREKQIKGWSRKKKIELIESLNPKWEDLAKDWS
<C3B73_RS02845--*Gracilimonas amylolytica*--54.95 (SEQ ID NO:1394)
MKAYTYIVTNPKKTLYTGVTNDLRRRIVEHYMDRGTSKSFAGKYCYLLIWEAFPTMNEAIQAEKLIKGNRKWKEDLIRKNRPYWNLSNENVLGWPPYK
TLPPESS
<C3B73_RS07495--*Gracilimonas amylolytica*--59.34 (SEQ ID NO:1395)
MPYTYIITNSSKTLYVGMNTSLPSRILEHGQONAGDSKTFTGRYRCYFVLWYEEYKYVQDSINREKQIKRWSRKKKFELIKRTNPKLRLNEDLGLYPFGDY
ADERNIRLSKRGRQ
<ATY29_RS20750--*Rhizobium hidalgonense*--69.32 (SEQ ID NO:1396)
MLVSGRNGTLYTGVTDLAAGRYEHQNELTPGFTSKYGVKMLVWFEEDLLITAI TREKTIKKWPRQWKHNLI ERENPEWEDISFHLLGL
<CPJ18_RS14430--*Agrobacterium rosae*--71.91 (SEQ ID NO:1397)
MKGYVYILASQRNGTLYTGVTDLASRLYEHQNELTPGFTSRYGKTLVWFEEDLLTAAITRETTIKKWPQWKHNLI ESNPDWQDISHYLLSL
<C2I35_RS19435--*Roseovarius confluentis*--63.53 (SEQ ID NO:1398)
MAFYTYIMACGSNTAIYVGTGNLDRAAQHRLEARSVHTSKYNI TKLVYFETHETLDAAVAREKRLKRLRREWKNLIEHANPSWSDLSMEASFL
<C3L50_RS08960--*Flavobacterium alvei*--65.67 (SEQ ID NO:1399)
MIEFQEGYHTYVYIILTNKAKTVLYTGVTNNLKNLQQHKDSLNPSFTAKYNVHFLLYFEKFTWIIQLAIAREKEIKGWKKDKKII LIKTINPELNFNLNLF
<CWC60_RS10810--*Minwua thermotolerans*--67.42 (SEQ ID NO:1400)
MARDLVYIILTDHNRGTFYVGMNTDLVRRMEQHRAGTAGSFSEYGLRHLVWYERHDDPNMAIRREKRIKRWPRWKLKTI EASNPDWRDLAYS LGGAAIHDC
PRIPDPRTRG
<CRJ03_RS13935--*Roseovarius nitratireducens*--68.60 (SEQ ID NO:1401)
MPKDMFVYIVTNKPRGTLYIGVTSDLAKRAYQHRTHAIRGFTDCYNLERLVWFERHAEPRAHQREKSLKRWYRDWKIALVETANPQWRDLWQNMRA
<LY01_RS07470--*Nonlabens xylanidelens*--70.33 (SEQ ID NO:1402)
MKPGYVYIMTNKNNTLYTGVTSDLERRVKEHKSNDPKSFTARYDLNKLIVYEAHQMIGDAIAREKQLKAGNRAKKIALIESINPEWQDLIIGDE
<LY01_RS08750--*Nonlabens xylanidelens*--63.53 (SEQ ID NO:1403)
MAIRSLHLYWIYILTNKKGGVLYIGVTGGIDDLRRRLREGSAFTKKYNAHKLIIYFEFQYINDAIAREKQLKNWHREWKINLIEENPEWQDLAAGWVLE

FIG. 11RR

<LY01_RS09500--Nonlabens_xylanidelens--69.05 (SEQ ID NO:1404)
MSSKNRILVVGATDDIKRRTFEHKTKKYKTSFSSKYNVDWLVIYEVVKGWEEAHKREKQLKKWNRAWKMRLIQEYNQDWDLDLSQGWNLDIYKD
<LY01_RS09465--Nonlabens_xylanidelens--65.22 (SEQ ID NO:1405)
MASYEHKYYVYLLSNKKGVLVYLGMTDDVKHRTGQHKHEERYKKAFAFSKRYQTCKLVYFEEFQWVQEAQAREKKLKKWKRQWKIDLIEKDNPEWDLDAEAWWDDE
NY
<GY24_RS05635--Microterricola_pindariensis--57.89 (SEQ ID NO:1406)
MYILECSDGSFYVGSTFDVEARLWQHNEGGAAAYTRRRQPVKLLYFEEYERIDEAYAREKQVQNWNSRAKRRALIDGVMELLPELAKRRRPPPLVE
<CCR94_RS14925--Rhodoblastus_sphagnicola--68.54 (SEQ ID NO:1407)
MAGLAPAIHAVEHASMKGGWVYIVTNKPNIGPYTGVATDLARRIHQHRTGVHEGFTQYGLKRFVHYEFHDDIVAAVQCEKNIKHWSRAWKARMIVEFNESWR
DLYEDLA
<CCS01_RS06830--Rhodopila_globiformis--69.66 (SEQ ID NO:1408)
MGGWVYMMTNKPFGLTVGVNTDIVRRAWEHRQGTGSCFSARYNLKRLVVFVEYHDEILAAIHRETRLKHWPRLWKLNLIEQQNPLWEDLFEQIAG
<CW354_RS09610--Marinicaulis_flavus--76.40 (SEQ ID NO:1409)
MAFYVYILASKKYGLTYIGMTDDLEKRIWERKEKVLKGFTAKYDVNQVLVWYEMHDTRESAFIRERSMKKWSRNWKNLIERENPDWEDLSETWSS
<CW354_RS15225--Marinicaulis_flavus--63.95 (SEQ ID NO:1410)
MPRFDFIAAYMLANRKNGLTYAGSTADLVARMDAHTGRGSWFAAKYGCNRLVWYERFDEMEPAIAYEKRLKRWPRKWKIDLIESVNPNDLPLF
<CW354_RS01885--Marinicaulis_flavus--64.44 (SEQ ID NO:1411)
MSGVYVYFIEGKRDGLTYGTADIHERVFQHKEGRGSSFCRRYGLKLVYEFNFRIEDAIHEKRVKKWRRRAWKNRLVDAMPDWRDLYESL
<BST83_RS05850--Polaribacter_filamentus--62.34 (SEQ ID NO:1412)
MLSNKNKTVLYIVAADKDLVERIKVHKSGBKANFTKKYSVFELVYFELHSDNKSARKREKQLKNWHKWKWNLVKKFNPDLITLIEIKNRP
<BST85_RS09555--Aureitalea_marina--68.60 (SEQ ID NO:1413)
MDKGIYVYILSNKRNGLTYVGVNTDIFRRVYEHRSKGNKSSFSFRYGLIRLVYEFHLDPAVGIRREKLIKGWKRKWKVELIEQVNPKWADISEDWYS
<AL537_RS05850--Vibrio_diabolicus--71.26 (SEQ ID NO:1414)
MDRTPCVYILTSNNNAVLYIGVTSQLPQRIWQHKNQVVEGFTKKYNVHKLTVYFELFEDMATAVNREKQLKQWKRKWKLDLVRETNPFLDLSAEYQ
<BTO14_RS03620--Polaribacter_butkevichii--70.83 (SEQ ID NO:1415)
MKTIHQYVYLLASKIRGTYIGVNTDLQRRVYEHKMGIKKGFTQKYGVNRLVYFETFENIEEAI TRENNMKWKWRDWKIKLIEEENVQWDLASDWYDDIL
<BTO14_RS06630--Polaribacter_butkevichii--68.24 (SEQ ID NO:1416)
MKKYHIYFLTNKNNTVLYIGVTSNLLKRIYQHKTKYKGF TAKYCNCAKLVYFEEYDDINSIAIAREKQLKAGNRKRKNEIVNLNPNWDLSDGWLQF
<BTO16_RS01365--Polaribacter_glomeratus--67.02 (SEQ ID NO:1417)
MKPGFYVILTNKNNTLYIGVTSNLEQRIKQHKSKIDKKSFTARYNLDKLVYEAQFMIGDAIAREKQLKAGSRSKKIALIERINTGWEDLSLD
<BTO18_RS07010--Polaribacter_porphyrae--72.34 (SEQ ID NO:1418)
MFAKLFKTHQYVYIITNKKDEVLYVGVNTDLERRMFEHKNKLVVEGFSRRYNDLKLIFYEAYQFINDAIKREKNIKWKREWKINLIVEDNPDWKDLSDF
D
<BTO18_RS07025--Polaribacter_porphyrae--69.15 (SEQ ID NO:1419)
MFKTQHQQYVYIITNKKDGLVLYIGVSNLKRRI FEHKNKLIKGFSEYRNLDKLIYFEVQYVSDAIKREKNLKNWKRQWKINLIVEDNADCNDLSEDF
<BTO18_RS01605--Polaribacter_porphyrae--65.88 (SEQ ID NO:1420)
MKKYHYVFLTNKNNTVLYIGVTSNLLKRIYQHKTKYKGF TAKYCNKLIYFEEFDDINSAIVREKQLKAGNRKRKNDLVNLENPNWDLSDNWLQF
<BST92_RS08730--Nonlabens_arenilitoris--63.95 (SEQ ID NO:1421)
MTGVMSFLVNLKFMKRAHHNYTVYIITNKKGLVLYTGMTGGLDRLRLHRLGEGSKFTKKYNASKLVYFEEFQYVNDAI TREKQIKNWHRQWKINLIEKDN
NWDDFSENWVL
<BST92_RS09380--Nonlabens_arenilitoris--59.72 (SEQ ID NO:1422)
MSHVENKLFYVGVNTNIIKQRFVDHKNATFETHVGFNIKQLVWFEKHDNPTAIAKREKLIKWKREYKINLI
<CLV33_RS14715--Jejuia_pallidilutea--75.27 (SEQ ID NO:1423)
MKVNYHQYVYLLSNKKNGLTYIGLTDLERRMFEHKNKLVSGFTSKYGVDKLMYFEHYQVNDAIKREKQLKNWNRQWKIDLIEQDNNKWEDLSSDWIY
<CLV33_RS09640--Jejuia_pallidilutea--68.24 (SEQ ID NO:1424)
MFNYETYYIYIITKKKDGILYIGMTGNIKQRI TQYKKKHTRGFANKYGLDILYIYEFKFSATEIAIKREKQIKKWNQWKINLINDLPNDWDLSSHFI
<CLV33_RS07705--Jejuia_pallidilutea--68.60 (SEQ ID NO:1425)
MKYVYVYIMTNKPNGLYIGITDNIERI KEHKLKVPKSF TAKYCNCKLVYFEEFENGEEAVVREKRFKKWRDWKIELIEDMNSSWDLDSINWNLNYNKLR
D
<CLV33_RS10995--Jejuia_pallidilutea--61.45 (SEQ ID NO:1426)
MNYYCYILSNKNRTVLYIGFTSNIKERIKAHQSGRGAFFTKYVNFLLIYIYDDSKLARKRERQLKNWNKEWKWNLIKTSNPDLIKLEI
<B1R32_RS12240--Abditibacterium_utsteinense--63.53 (SEQ ID NO:1427)
MVQHRFYFTIVANANRTIYIGVTSDLQRIQQHSAKTFAGFTAQYGLSGLVYFEEFGDVRDAIEREKQLKGFRRAKVALIEAQNPHEWEDVAADWFS
<S2091_RS04280--Solimicrobium_silvestre--82.42 (SEQ ID NO:1428)
MKAPAIYLLANQRNGTYIGVTSDLIQGAWQHREGIAQGFTHQHQVKTIVWHEQHETMESAITREKALKKNRKKWLRLIEKFNPWKDLWLEITGQESSH
<C5748_RS18935--Phyllobacterium_phragmitis--66.67 (SEQ ID NO:1429)
MSRYVYMMASKKGGTYIGVNTDLSHRVWEHKNGITPGFTSEYGCARLVWYEEHDDIRREBTLKRWNRRWKIDLIESINPDWAELYRGIGW
<C5L14_RS21320--Labrys_okinawensis--72.94 (SEQ ID NO:1430)
MRNGRHEAAGEANQRNSILYTGVTSNLSERVYQHRKGLIRGFTSRYGCTLLVYEFYEEEMPALIREKQIKAGSRKKLELIESLNPTWRDLYLDF
<C6Y40_RS20570--Alteromonas_alba--76.14 (SEQ ID NO:1431)
MIEKYPVYILTSYKSVFYTGVTNNLLQRIYQHKMHLADGFSKKYGLKLLVYFEMHEDIEAAIQREKRLKRWRDWKIKLIEQNPHWQDLYPLIQG
<CVO77_RS15595--Sphingopyxis_lindanitolerans--74.16 (SEQ ID NO:1432)
MHADFQPCAYIMASARHGTLYTGVTSHLIQRIQHRDGSGLDGFTKRYGVKTLVWVELAGTMHAAITREKQIKNWQRKWKIELIATNPWRDLAEELGLASLS
PPSSSRT
<CVO77_RS06985--Sphingopyxis_lindanitolerans--74.42 (SEQ ID NO:1433)
MANRKNGAIYTGTSKLVQRVYQHREMLVEGFTKTHGCVLLVWFEAHEDLQEARRELOIKRWKRSWKVRMVEETNPDRDLWVEIQG
<CVO77_RS06970--Sphingopyxis_lindanitolerans--76.14 (SEQ ID NO:1434)
MNKPGYLYLMSARRNGTYIGVTSDDLKRVWQHRNEVIDSFSSKYGCTLLVWYEAFFDDIQARQRELQMKKWKRAWKVELIERDNPQWDLDFRLSL
<CVO77_RS15270--Sphingopyxis_lindanitolerans--60.23 (SEQ ID NO:1435)
MLTNRGNQLFYIGVTSDI AARMQHRAGAGSAHCQRYNIRKLVHVEFCSTMEHAIAREKALKWHRDWRRLIDESNPEWDDLDFRILG

FIG. 11SS

61/106

<BST86_RS08165--Nonlabens_agnitus--68.89 (SEQ ID NO:1436)
MPHYVYIITNKNNGTLYIGETGNIKRRI SQHKRKVHPAKFSARYNLDKLVYFEILD SL DARRIREKQ LKKWN RDWKIRIIEEKNPDLIDL SLNWDLSFKMMEK
EI
<BST86_RS02480--Nonlabens_agnitus--62.50 (SEQ ID NO:1437)
MAKRWFHNYTVYILTNKKNGLVYIGITGGLEDRLRRHRLGEGSYFTKKYKAHRLVYFEFQVLVHDAIAREKQ LKNWHRQWKINLIESENPDWNDLAENWKL
<BST86_RS08540--Nonlabens_agnitus--58.23 (SEQ ID NO:1438)
MKGWMIYLLCSNGSYTGSTNDLNRRI IKHQAGLGANHTKKYLPVELIYFEFDRVQDAFYREKQVQGSRAKKEALINGFSD ELKRLAKCLNETSHEKWRKE
<CLV81_RS19735--Muricauda_pacifica--58.82 (SEQ ID NO:1439)
MKSSFYILNTYRTTFYIGVTSDLKKRI TEHIEGFGSAFIKRYNLKDLIFEEFSNINQTIEREKQ LKNWHRWKINLIREQNPSLSTLII
<CLV81_RS06475--Muricauda_pacifica--57.30 (SEQ ID NO:1440)
MNQEYCYILTNKNKTVLVGATNDLERIAEHKKGIHPNAFTKRYNCTFLVYFETFGDVKDAFRREKQ LKAGNRKRKEALINSINPEWHDLSEWQF
<CLV58_RS27515--Spirosoma_oryzae--58.70 (SEQ ID NO:1441)
MNQYAFYVYIVTNPEKTVLYTGMTNDLIRLQEHYESRGQAEKFAGKYCHRLVYWEYHQVVRNAIDREKEIKGWRREKKVALIQ SINPEWRFLNDDILG
<CLW00_RS04560--Mongolibacter_ruber--62.79 (SEQ ID NO:1442)
MKRHEFYVYILTNFAKTVLYGVTNNLSKRLAEHNEKLNPESTSKYNCCHLVHVESYQYVREAIAREKEIKKWRSSKEALIAEYNPGWKFLNEEVLGN
<C7H52_RS09085--Mesoflavibacter_aestuarii--63.29 (SEQ ID NO:1443)
MQLSYIYILTNKYRTTFYVGVTSNLLNRI TEHQNGTGSVFTKKYNLKDLYFEFETDINQAIAREKQ IKNWRKDWKFNLIKETNPKELETIIL
<C7H61_RS07120--Mesoflavibacter_sabulilitoris--67.39 (SEQ ID NO:1444)
MKPGFTYIMTNKTYITLTVGVTNNIVKRVQQHKDKVMKSFRTARYNLNKL VYVEAFQSIGDAIAREKQ LKAGSRAKKIALIVATNPDPYTDLFDRAKEVYGD
<C7H61_RS13620--Mesoflavibacter_sabulilitoris--67.47 (SEQ ID NO:1445)
MTNKPGVLYIGVTTNIGRIKHKLVYPKSFTARYNCDKLVYFEEDNGLIAEKREKQ FKKWKRNWKIELIEEMNPNTDLSLNNLNFNKIRE
<C7H61_RS12940--Mesoflavibacter_sabulilitoris--59.34 (SEQ ID NO:1446)
MKTGLHNHYVYILTNKNTVLYTGVTNNLQERLHYHNNLPFSKSTSKYKCFYLVVYEHFNVEIAITREKQ IKGWLRVKKENLINSFNPNWEFLNNDI
<C7H61_RS03135--Mesoflavibacter_sabulilitoris--59.77 (SEQ ID NO:1447)
MLKPLGTHNYVYILTNKNRTVLYTGVTSNLVERLHYHNNPLPFSKSTAKYKCFFLIYFEHFEIEAAIKREKEIKGKSRVKKENLISFNPNNWFLNQIKI
<C6569_RS03795--Phreatobacter_cathodophilus--73.03 (SEQ ID NO:1448)
MPGAGFFYVYILASGRNGTLYVGTNDIGRRSFEHRSGAGAVTTRRYRVHRLVWVEPYDGI AAAIAREKQ IKGWNRAWKRLIERDNPQWRDLYGELAGP
<CU100_RS02905--Phyllobacterium_eophyticum--67.78 (SEQ ID NO:1449)
MSGYVYILTNKNGTVYVGTADLAGRIYQHREGHTPGFAWKYGCTRLVWVEERDLVVD AIQREKSLKRWYRQWKIELIESINPNWNDLYATLA
<CU103_RS03290--Phyllobacterium_sophorae--70.00 (SEQ ID NO:1450)
MSGYVYILANHHKGTLYIGVTSDLAGRMVQHREGYTPGFRAWYGCARLVWVEEHLVVD AIQREKSLKRWHRQWKIELIESINPNWEDLYATLS
<C7C56_RS03810--Massilia_glaciei--77.42 (SEQ ID NO:1451)
MDKQSYVYIMASARNGTLYIGVTSDLVKRVWQHREGQADSFTKMYGVKMLVWFEAHTDIAEVI TREKQ LKKWNRSWKVDLIQKDNVRWRDFYEDILSR
<C7I84_RS17950--Mesorhizobium_ephedrae--68.24 (SEQ ID NO:1452)
MGRRGYVHILASQRNGTLYTGVTSDLAGRLLQHQTGSGSKFAKRYGVMRLVWVEEHAWVVD AIKREKAIKKWPRRWKIKLIEENPNHWHDI AWHLL
<C7I84_RS15130--Mesorhizobium_ephedrae--66.29 (SEQ ID NO:1453)
MTGYVYMTASQKRGTIYIGVTDNLGRMP EHKSGQGSRFSAARYGVQRLVWVEERFVDRDAIQREKSLKRWP RQWKIELIEKANPEWEHLFRGLGW
<C7I56_RS08635--Sphingosinicella_vermicomposti--74.71 (SEQ ID NO:1454)
MERLPYVYILASGRNGTLYIGVTSDLIRRIHEHRAGERSGFVSRY SAYRLVYHETHQTMPDAIAREKQ LKNWHRRAWKLN LIEKNPHWEDLAIGLGFAPLPPA
PHRPYA
<C7I56_RS02085--Sphingosinicella_vermicomposti--60.61 (SEQ ID NO:1455)
MTFWTYMLHCNAGRLYIGHTDNLERRIAQHEQGQIPGFTRNHLPVTLVWSETLPSREEALAAERKIKGGSRAKKLALIRGDFEEI SRLAKSKDRPSTSSGRTE
FYVK
<C8348_RS10080--Pannonibacter_carbonis--72.22 (SEQ ID NO:1456)
MSGYVYMMSDKPRGVIYTGVTSDLPKRIWEHREGTRKGFTSQYKAKNLVWFEVHPSIVLAIQREKSLKRYLRDWKIRLIEELNPTWLDLYERMEEIENVYKPH
ENTKMWDDYN
<C8348_RS09455--Pannonibacter_carbonis--74.16 (SEQ ID NO:1457)
MYHVYILASRMHGTLYVGTNDLPRRVDEHRQGVGSRFTREHGVSRVLYFEGFEDVEAAIAAEKRLKRWHRAWKVQLVERTNPLWKDLSLELNM
<B0I18_RS13530--Taibaiella_chishuiensis--52.81 (SEQ ID NO:1458)
MPPPYFVYITTNKNKTVLYTGVTNNLVRMSQHVYQANSRNSATFTGRYACCVLYVFEHRYIPEAIKREKQ IKGWTRARKKALIE TINPGICFLDAG
<C7S18_RS17360--Ahniella_affigens--74.19 (SEQ ID NO:1459)
MPKRRREPAYVILRNYRNGTLYVGVTSNLIARLWIHREGLLPGFTQRYGLHRLAYFELRETM TDAIQREKELKRF SRVKKLN LIESINPQWRDLYPELDDVTS
QDSPGSRNPLRGFQDDKV
<C7293_RS19305--filamentous_cyanobacterium_CCT1--81.52 (SEQ ID NO:1460)
MDKQPTVYILASKPYGTLYTGVTSNLTKRVWEHRNNVVEGFTSRYRVHQLVYFEQHTSMEEAILREKRIKKWNRAWKIKLIEEKNPGWKDLWLDL
<C8B47_RS11600--filamentous_cyanobacterium_CCP4--81.52 (SEQ ID NO:1461)
MDKQPTVYILASKPYGTLYTGVTSNLTKRVWEHRNNVVEGFTSRYRVHQLVYFEQHTSMEEAILREKRIKKWNRAWKIKLIEEKNPGWKDLWLDL
<C7J89_RS00215--Staphylococcus_kloosii--58.44 (SEQ ID NO:1462)
MDSHYVYIVRCKDGLSYTGAKDVEQRVKKHNLGQAKYTKNRRPVTLVYHETHDTKSSALKREYEIKYTSRNKKLQLIKEAQ
<C6368_RS14805--Brumimicrobium_mesophilum--74.12 (SEQ ID NO:1463)
MKTPKTFFVYMLSNYNNKLLYIGFTSDLAQRVHKHKTGFYDGTKKYNIKRLLYFETYPNAAEAITREKQ IKSYNRSKKNLIDDVNPQWREIKAPINNF
<C6368_RS10055--Brumimicrobium_mesophilum--70.59 (SEQ ID NO:1464)
MKTIKSYVYIITNRNNTLYVGTNNLARRI FEHKTINKGFTSKYNINKLVYETFSLVIDAIKREKQ LKRLSRVNKEEIIHVNSDWKELKPPIHIWSNR
<CV103_RS10060--Sphingomonas_fennica--73.86 (SEQ ID NO:1465)
MDRQPFAYILASGKPGTLYIGATSDPAKRLHEHREGLIKGFTSRYGITRLVYFESFDDMHSIAIAREKQ LKKWRREWKNLIEASNPGWVDLAPGLRLGSAIPQ
R
<CV103_RS20430--Sphingomonas_fennica--73.86 (SEQ ID NO:1466)
MSRQPCVYILASDRYGTLYIGVTSNLPVRLHQHRTGTIPGFTSRYGVHRLVRLEMFEDMQSAIAREKQ LKRWHRREWKNLIE SDNPNWIDL AIGLGFEPAGRH
MPHDGP

FIG. 11TT

<CV103_RS10825--Sphingomonas_fennica--72.41 (SEQ ID NO:1467)
MIERQPCGYILASSFNGALYVGVTSDLLRRIARHRDGMFDGFTKRYGVKYFVYWEAFGTMEAAI LREKQI KRWRDWRNLI ERDNP SWHDQAVPLGLEPVAP
SPA
<CV103_RS03965--Sphingomonas_fennica--62.50 (SEQ ID NO:1468)
MERGGWYIMTNKPRGVLYIGVTAHLAARVAQHREGLGS AFCRRYNLQRLVLVEPHSSIEEAIAREKALKAWKRAWKVALIEMVNPCWHD LADDIGLAS
<CV103_RS12780--Sphingomonas_fennica--63.64 (SEQ ID NO:1469)
MTRGGWTCILSNKPKGILYIGVTANLAARMQHRGTGTS AFCKRYGLDRLVFAEPHDSMVEAIAREKALKAWKRDWKIRLIEQGNPEWEDLDFRIL
<DBZ16_RS06915--Neisseria_sicca--74.70 (SEQ ID NO:1470)
MANGKNGTLYIGVTMNLPERVWQHKNHVNVDGFTARYDVHDLVWYEFFENMPEAVAKEKTMKKWRREWIKLIEEQNP EWLDLSGVLFV
<C8P69_RS05175--Phreatobacter_oligotrophus--75.29 (SEQ ID NO:1471)
MRSRGPFFYVYILANRRHGTLYVGVTTNDISRRTFEHRTGVGAVFTRRYSVHRLVWMEPYDRVAEAIAREKQIKGWNRAWKIRLIESSNPEWRDLHEELAGRGQS
G
<C8P69_RS16000--Phreatobacter_oligotrophus--71.11 (SEQ ID NO:1472)
MERARGGWYIITNRRDGVLYTGVTDIARRGHEHRTGAYEGFTKHYGLKLLVWMBERYDDIRDAIQRERSIKRWPRAWKVR LIHESNP DWRDLYDDL
<C8J24_RS03525--Sphingomonas_aerolata--79.55 (SEQ ID NO:1473)
MARQPAVYIMASGRNGTLYIGVTS DLMARIVQHRDGMFDGFSKRYGVNRLVWFAMADTMDAAI TREKQLKVWKR SWKTTLIETDNPTWRDLAEDWGF DPLPTP
SS
<C8J24_RS05845--Sphingomonas_aerolata--77.42 (SEQ ID NO:1474)
MRETFQPCVYILASRRHGTLYVGVTSNLLARLVQHREGLIPGFTRRYGISRLAYYEMFDTMEAAIVREKRLKEWRRRAWKIELIEARNENWDDLGI GLGLPLRP
DLPTEPVLGV
<C8J24_RS09790--Sphingomonas_aerolata--66.29 (SEQ ID NO:1475)
MEEQAFVYIMANFPNGAIYLGSTLNLPRVWEHRNGLSGFTKRYGCKLLVWYEAACGEWEAARQRELQMK EWKRAWKVREIEGLNPAWDDLYDRIALP
<C8J48_RS17050--Desmospora_activa_DSM_45169--54.67 (SEQ ID NO:1476)
MREKNGYTVYILECGDNTLYTGITNHLQQRINQHHQKGAKYTRGRGPFKLVWTEHPDRASASRREREIKALSRKDKRQLIAAGQQERNRNGETEEL YR
<C7449_RS11550--Mycoplana_dimorpha--69.66 (SEQ ID NO:1477)
MKGYVYILASQRNGTLYGTVTTDLPRRLYEHQNDLTPGFTSRYGKTLVWFEEDLVTDIAIAREKAIKNWPRKWKLDLIEAANPDWADIAHHLGL
<C8N47_RS18080--Mangrovibacterium_marinum--66.28 (SEQ ID NO:1478)
MERGGYIYILTNNKNTVLYTGVTSALT KRLYEHKTFGYKDSFSTRYNVSKLVYEECFSTIEEAI IREKQIKAGSRKKKLDLINRFPNPEWKDLSGTVTE
<CLG96_RS11135--Sphingomonas_oleivorans--62.92 (SEQ ID NO:1479)
MQGGWYIYIMTDKPHGTLYIGVTADLAARIYQHRTGTGSDFCRKYGLKRLVYVEPYPTIGEAIMREKAMKAWKRIWK TALISKTNP DWEDLYPRING
<CLG96_RS11070--Sphingomonas_oleivorans--65.17 (SEQ ID NO:1480)
MAGGFVYIMTNGPSGTLVYGVGTANLAARIHQHRTGQGSDFCRKHRLNRLVYIERHETMLEAIAAREKALKKWQRTWKNLNGKTNPEWRDLFEEING
<C8J26_RS03600--Sphingomonas_aurantiaca--80.68 (SEQ ID NO:1481)
MARQPAVYIMASGRNGTLYIGLTS DLMARIVQHREGMFDGFSKRYGVNRLVWFAMADTMDAAI TREKQLKVWKR SWKITLIETDNPTWRDLAEDWGF DPLPPP
SS
<C8J26_RS00375--Sphingomonas_aurantiaca--80.68 (SEQ ID NO:1482)
MDRQPCVYILASGKHGTLYIGVTSNLLGR LQQHREGLIKGFTSRYDVRLVHYEMADTMDAAIAREKQLKKNWRD WKVTLIERDNP DWDDLAI GLGLDFVSGR
SSR
<C8J26_RS10430--Sphingomonas_aurantiaca--72.09 (SEQ ID NO:1483)
MTKAGVYVIMASRRNGTYLGVTSDLAARSYQHRNGLIIGFTKTHYCKLLVWYEVHDDLENARLRELQMKKNRQWKLSEIERMNP DWNDFESLF
<C8J26_RS04130--Sphingomonas_aurantiaca--60.47 (SEQ ID NO:1484)
MAQGGWYVYIMTNPYGMLYIGVTAHLASRVQQHRIDTGSAYCRRYGLKTLVFAEQHETIQDAIAREKRLKAWRRSWKDDLIATINPTWADLYDQLH
<C8P68_RS08500--Mucilaginibacter_yixingensis--60.64 (SEQ ID NO:1485)
MITNIHQYNVYIVTNKAHTVYVGVTS DLFTRILQHKNKAYPGFTAKYECHKL VYFEDFQWIDQAIAREKQLKGGSRQNKIDLIVNSNP DWKDL SKDWFD
<C8P68_RS05170--Mucilaginibacter_yixingensis--56.82 (SEQ ID NO:1486)
MNQGYVYIMTNAYRTTFYIGVTANLRSRVWQHTNGBGSVFKRYRLFDLVYHEYFEKI TDAINREKQLKNWHHEWKINLIKAANPEMKDLKSELELI
<C8P68_RS08925--Mucilaginibacter_yixingensis--58.70 (SEQ ID NO:1487)
MSSYNFYTYILTNKSKRVLYTGTINNLERLYEHYTGADQPN SFTHKYQCYYLIWFERYQHVQQAIDREKEIKGWL RKKKVGLIGETNP DWRFLNEDIVEWPP
SK
<C8R30_RS03595--Nitrosomonas_nitrosa--77.42 (SEQ ID NO:1488)
MEKQPAVYILASQRNGTLYTGVTSN LVKRVWEHKNDQVEGFTKRYRVHTLVYFELCESMETAIMREKQIKAGSRLKKLKLIEAKNP GWQDLYDSIV
<CJD38_RS10555--Stenotrophobium_rhamsosiphilum--69.33 (SEQ ID NO:1489)
MSSFTRTLYVGVTDNLP RRVLEHKS KLVAGFSAKYNI TDLVYAEYWDVRE AIDREKQIKSWRREKKIALIESVNTGWTDLAL
<DBO86_RS18715--Pseudomonas_indoloxydans--83.87 (SEQ ID NO:1490)
MKQPAVYIMASQRNGTLYTGV TADLLRRVWQHREGVVEGFTLRYGVKLLVWFEQHEMSGAIAREKAIKKWNRAWKRLRIEERNPQWRDLWSEIVGESSGAAG
SPPSRG
<C8J25_RS10360--Sphingomonas_faeni--78.41 (SEQ ID NO:1491)
MTTWGNDVARQPAVYIMASWRNGTLYIGVTS DLMARIVQHREGAFAGFTKRYGVNQLVWFAMADTMEAAI TREKQLKVWKR SWKKELIETDNPTWRDLAEDWG
FDLPPTPIPS
<C8J25_RS00600--Sphingomonas_faeni--74.16 (SEQ ID NO:1492)
MRETFQPCVYILHSRRYGTLYIGVTSNLLARLVQHREGLIPGFTRRYGIKRLAYYEMFDTMEAAIVREKRLKDWRRRAWKIDLIEARNENWDDLGI GLGLPRLT
EPTLGV
<C8J25_RS02170--Sphingomonas_faeni--73.49 (SEQ ID NO:1493)
MASSRNGTLYLGVTSDLLARAYQHRNRLIDGFSKAHGCI LLVWFETSDDIQ SARQRELQMKWKRAWKIDLIEQTNPHWKDLYETLF
<C8J25_RS02150--Sphingomonas_faeni--65.17 (SEQ ID NO:1494)
MGEQAFVYIMANFPNGAIYLGSTVNLPRVWEHRNGFGGFTKRYGCKILVWYEAACGEWEAARQRELQMK EWKRAWKVREIEGFNPAWDDLYDRIALP
<C8N35_RS11140--Breoghanian_corrubedonensis--73.33 (SEQ ID NO:1495)
MERPHDYFIYVLASGLGGTLYVGVTDNLVRRIS EHRQGIASNFTRKYRVHHLVYEHHRDINEAIRREKQVKRWNRWKIQLIEERNPNWDDL FKS LI
<C8N40_RS03330--Pontibacter_mucosus--59.77 (SEQ ID NO:1496)
MKSHNYFVYITTNPGKTVLYVGVTDNLVVRLEQHKENRGKPFETFAGRYCYCKLLYERHTYVQYAIEREKELKLLSREQKLELIRKENPHLQFLKVED

FIG. 11UU

<C8N40_RS21535--Pontibacter_mucosus--58.02 (SEQ ID NO:1497)
MKSHNYFVYIITNPRKTVLYTGMTNDLQVRVEQHKANRGKPETFAGRYCYKLLYFERSTYVQHAIEREKETKLSREDKLQIRKENPMLMFLDV
<C8N44_RS10720--Roseivivax_pacificus--67.42 (SEQ ID NO:1498)
MKKDPPLFVYIMTNRPRGTLYIGVTGDIARRAWEHRTHATPGFTDRYNLERLVHFEVFEEPESEAIRREKALKRWRRDWKIALIEKGNADWRDLWEEITP
<C8P64_RS02890--Gramella_gaetbulicola--66.28 (SEQ ID NO:1499)
MKTSYIYILSNKNRSLVLYIGIASSELIKRMTEHKNKIGSDFCRKYNLEDLLYFEEFTDINKAIAAREKQLKNWHKEWKWNLIKESNPDLRDLFPPEL
<C8N46_RS02455--Kordia_periserrulae--79.12 (SEQ ID NO:1500)
MYKTIHQYVYIILSNKNGTLYIGITNNLERRVFEHKNKLVDGFTTKYNLTCLMYEYIFQHVDEAIKREKRLKKWKRWQKINLIEEKNPNWKDIASDWATNMD
SATSAE
<DDE77_RS04570--Devosia_submarina--76.67 (SEQ ID NO:1501)
MGQTYFVYMMANRKHGALYTGVTSDLVRRTYEHREGLADGFTKTHGCKHLVWYEVHEDIEAAIILHEKRLKKWQRHWKEELIETMNPNNWRDMWWEIEGGPTT
<DDE77_RS01525--Devosia_submarina--67.74 (SEQ ID NO:1502)
MAGYTYILADMRGRRTYIGVTNNLHRIYEHREGLADGYTKNNIKRLVHFETHEVPLAIQREKTLKRWYRQWKEALIEEHNPGWQDLWDEINK
<DDE77_RS13385--Devosia_submarina--67.42 (SEQ ID NO:1503)
MKGYVYILASKKNGTLYIGVTSHPSRRLFEHQHELTGFTSRYGVKTLVWVFEQHDLVVDVDAIRREKAIKKYPRQWKINLIEAVNPEDWDLSS
<C7G61_RS35030--Bradyrhizobium_algeriense--68.24 (SEQ ID NO:1504)
MYVYVILASRQHGTLYIGVTNSLSKRLEQHRNGEGSSFVKTYGVYRLVYVETFERPDEAITREKQLKRWKRWDWKIELIERENLEWRDLSDLIV
<B9Z37_RS14050--Limnhabitans_parvus_II-B4--76.60 (SEQ ID NO:1505)
MAYFVYILANQKNGTLYIGVTNHLIRRVHEHKQHAVDGFRTARYNVTQLVWVFEQTESIQEALAYEKRLKHWKREWKIALIEKENPDWFDLYESIL
<DDF62_RS17725--Caulobacter_radicis--78.65 (SEQ ID NO:1506)
MAFYVYIVASRRNGTLYIGTDDDLARRIWEHQEKVRPGFTKRYDVGMVLVWYEVHDTRESALIRERQMKKWRRLWKLELIERTNPGWRDLAQDLNC
<DDF62_RS18260--Caulobacter_radicis--77.91 (SEQ ID NO:1507)
MERVSVYIILASRNGTLYIGVTNNDLSRRTWEHREKLMGFTKRYGVMTLVWYEVHDEIGAAIQREKLIKRWRRVWKLELIERMNPLWDLLEYTLNC
<DDF62_RS13785--Caulobacter_radicis--70.11 (SEQ ID NO:1508)
MAVYIMANERLKTLYIGATGYLNRRWEHREGLIPGFTKKYGLTRLVWYEPHDDMATAFQRRERSLKRWRPQWKINLIERTNPFWDLYPLGLG
<HYN59_RS09880--Flavobacterium_album--61.05 (SEQ ID NO:1509)
MKPGFVYIMTNSNNTVLHVEVTSDLHERVADHKSRYQDSFARYNLKLVYFEAFQMGDAIGREKQMKGSGSRKKVLDLIRSMNPEWNDLTFSINDFL
<HYN59_RS05125--Flavobacterium_album--63.22 (SEQ ID NO:1511)
MYKTLGTHNYVYMLTNKAKTVLYTGVTNELKERIYYHRNPEPYSKAFATKYKCFYLVYFEHYQDIEIAIDREKQIKGWTRKKNKLIDDFNPGWKFLNDEI
<HYN59_RS17450--Flavobacterium_album--60.92 (SEQ ID NO:1511)
MYKTLGTHNYVYIILTNKIKTLYIGFTNNMKERLYYHQHPEAFCKAFTTKYKCYNLVYWEHYQDVDAIDRETQLKKWSRKKKDALITAFNPEWKFLNDEI
<DD559_RS04400--Sphingomonas_pokkali--77.27 (SEQ ID NO:1512)
MDKQPCVYIILASRRHGTLYIGVTSNLLGRLYQHREKVTQSFTAEDHDIARLVWFEAHEDMSAAIILREKRIKKWNRDWKLALIERNSNPQWVDLAIGLGFEPQPT
GKWIIPASAGMTE
<C8E01_RS06245--Pontibacter_virosus--64.20 (SEQ ID NO:1513)
MKSHNYFVYIITNPSKTVLYVGVGTNDLLYRLQEHKQNRGKPETFSGRYFCYQLVYERYTYVQHAIDREKELKSWSRKKMELIKSENPLYRFLKADP
<C8E01_RS05500--Pontibacter_virosus--60.49 (SEQ ID NO:1514)
MKSHNYFVYIITNPNKTVLYVGVGTNDLVYRLQEHKQNRGKPETFAGRYFCYKLLYERYTYVQHAIDREKVLKSWSRDCKMELIKSENPLYRFLKADP
<P3BBH22_RS05350--Lawsonibacter_asaccharolyticus--53.95 (SEQ ID NO:1515)
MSWWVYLLRCGDGTLTYGTITDDLDRLAAHNAGRGAKYTRSRRPVMVWREEQPDKSAALRREYAVKQLTRAKKLALIEGGPLV
<C5689_RS12480--Methylosinus_sporium--79.78 (SEQ ID NO:1516)
MDRLFVYIILASQRNGTLYIGITNNLSRRVFEHKKRLTPGFTSRYGVMDLVFYESYTDVTEAIAAREKSLKKWNRAWKLQILIERINPTWRDLYQDLNR
<DDT93_RS01990--Ruegeria_kandeliae--67.06 (SEQ ID NO:1517)
MPHFVYILASRPGGALYIGRTSNLRRRFEQHRSGAVPGHTSLYIGIRTLVWFEIHAEFEPSSLRERRLLKRWRRRAWKDDLILQCNPQWQDI THLIPD
<DDV73_RS15625--Nocardioides_spelunca--58.82 (SEQ ID NO:1518)
MAWAYLLECDGSGFYVGSTTDIERRVWEHNEGLGASYTRRRRPVRLAWSAEYERVSADFAVEKQVQGWWSRAKRLALIEGRLQDLPLASRAWRSP
<HYN56_RS09575--Flavobacterium_crocinum--67.05 (SEQ ID NO:1519)
MLSPQNGYHTYVYIITNRYRSTFYIGVTNNLRERLNRQHKENIKKNEKTFAAKYGIEBFLVYEEKFTWIQEAITREKELKGWIRDKKLELIRDFNPTFEFLNYY
FE
<HYN56_RS24470--Flavobacterium_crocinum--64.63 (SEQ ID NO:1520)
MYLQEGFHTYFIYIILTNKAKSVFYVGVGTNNLKNRLKHKENILGGNKTFASKYNVEFLLYEEKFTWIQEAIKEKEIKGWRREKKIALIKTMNPD
<HYN56_RS01420--Flavobacterium_crocinum--65.82 (SEQ ID NO:1521)
MKSIEGYHTYIYIILANRHKTLYIGVTNNLRRLRLKQHKENISSNNNTFAAKYNAQYLLYEEKFTWIQEAIAREKEIKGWIRDKKIELIKTINPDLDLYLF
E
<DF286_RS10540--Sphingosinicella_humi--77.11 (SEQ ID NO:1522)
MLASRRHGTLYGVGTSDLPKRIHEHRIGAVPGFTRRHRIKHLVYFEAHEGIEAAIIREKRIKEWRRDWKIELIERDNPEWNDLAVVLGFEELLPETPIPHRHP
GESRDPRTDWW
<DF286_RS14130--Sphingosinicella_humi--56.96 (SEQ ID NO:1523)
MQFWVYILRCRDGSYYTGHTDDLEKRLWEHREGLASDWTKRRRPVELVWCDWTGERYDALAFERRVKWNSRAKKQALIAGDWERVGYFARPPHERPSTTLGTN
GLDVSIRIGPSTSEAPTSPFVSEVEARPSASVKNTP
<DBR16_RS01700--Hyphobacterium_indicum--70.65 (SEQ ID NO:1524)
MKRPCVYIILASRRNGTLYIGVTSNLARVEEHRQGRASKFTRKYGVTRLVHFELFEEIGEAILREKRIKKWKRAWKLDLIEQYNPDWQVSGEIPYD
<DHV10_RS07030--Occallatibacter_savannae--70.45 (SEQ ID NO:1525)
MREHVYFTYIVASKRSTLYIGVTSDLRHRVFEHKTAYEGFTSRYNCNRLVWFERENNVSAAIQREKELKGWLSKKVELIQAANPTWEDLSADWFPPLTREV
LETGAYGKADPSLR
<DIR46_RS05215--Massilia_oculi--71.43 (SEQ ID NO:1526)
MDKPSCVYIILASRPGTLYIGVTSDLIKRVWQHKAEVVPSTSKHGKRLFWYETRTDI THAILREKQLKRWNRDWINLIQSTNPTWHDLYAAICG
<DI487_RS06475--Flavobacterium_sediminis--66.28 (SEQ ID NO:1527)
MSFRNYHNYVYIILTNKPKGTLYIGVTGGIDDRMERHISGQSKFTAKYKLLVYCEEFFQFINDAIAAREKQLKNWHRNWKINLIEEHNPNWENLWNFSESEK
<LX78_RS05325--Formosa_spongicola--65.88 (SEQ ID NO:1528)
MKTSYVYFMTNKNNTLYGVTDHLIKRVFEHKNETYKGFTTKYCNKLVYFEEFTSINEAIVREKQIKKGNRKRKEDLINRENPNWNLSDGWLKFKPDT

FIG. 11VV

<LX78_RS13150--Formosa_spongicola--69.88 (SEQ ID NO:1529)
MANKPNGVTYIGMTDDI EERI KEHKLKIRPKSFTAKYNC DKLIYFEEYEVDEKAIIRERQLKKWKRDWKIKLIEEMNPSWTDLSTNWSLDFNKIRNER
<LX78_RS00920--Formosa_spongicola--63.22 (SEQ ID NO:1530)
MQLSYTYILTNKYRTTFYIGVIANLSKRLYEHYEGVASVYTKYNIKDLIYYEEFTDINQAIAREKQLKNWHKEWKLNLIRKFNP TLKTLTSYQ
<LX92_RS13275--Maribacter_polysiphoniae--71.00 (SEQ ID NO:1531)
MQPLQHQQYYVYILTNKKHGTLYIGMTNNLKRRIYEHKNKLVEGFTQKYNLDKVVFMESFLYVQDAI LREKRMKKWKRWQKINLIEEDNPEWMDLAIDWYD
<C7455_RS15235--Roseicyclus_mahoneyensis--76.67 (SEQ ID NO:1532)
MEVYMMTNKPDGTLYIGVTNDLVRRVHEHRTHAVKGFTDRYNLERLVWFEPHDSPTFAIRREKALKSWNRAWKVDLIETANPGWRDLWDEIIGSSPTMTEGVA
R
<LX99_RS11550--Mucilaginibacter_oryzae--55.43 (SEQ ID NO:1533)
MWNHNYVYITNSNRTVLYIGVTNDIAVRVSQHHENKGTQKSFTGKYCYCHLVYEEHFTHIEHAIAREKEIKKWREKKEALIASTNPTWTFLNDSVIN
<DEM25_RS14280--Oceaniradius_stylonematis--67.07 (SEQ ID NO:1534)
MASKKGGTLVGVTSDLAQRVWEHENGISGKFTAKHGCKRLVWFEEHDLVTRAITREKTIKKWPRPWKINAIEAINPEWNDLRPWL
<DDQ68_RS03775--Hymenobacter_nivis--61.18 (SEQ ID NO:1535)
MYVYIILTNQTQTVLYIGVTNDLTQRLYEHSEN RGNAAKFTGRYQADLLVYFEIAPDATQAIAREKQLKGWTRRRKDALITAFNP TWQAI DL ETWEG
<DKG77_RS01850--Flagellimonas_aquimarina--60.22 (SEQ ID NO:1536)
MFHAYFVYITNENKTTLYIGITNNIQKRLSQHYFDSQNAKKS FAGKYNCYLVYVEGFEDEIETAIQREKELKKWREKKNKLI TSFNPDWEFLNHQVI
<DKB74_RS15370--Sphingosinithalassobacter_portus--73.86 (SEQ ID NO:1537)
MERQPCVYIILASGRVGTIYIGVTSDLIARLHQHRAGTVPWFTSKYSVDRLVRFE LFGTMEQAIAREKRLKRWRREWKLNLIESDNP DWLDLAVGLGF EAPRLD
VRKSGC
<DKB74_RS15910--Sphingosinithalassobacter_portus--76.83 (SEQ ID NO:1538)
MASGRNGTLYIGVTGDLVKRAWQHREGALDGFTKRYGCKYLVWFEVYDRIDS AIRRERQMKEWRRRAWKLREIEEMNP DWHLFP TLM
<DKB74_RS11290--Sphingosinithalassobacter_portus--67.42 (SEQ ID NO:1539)
MEGRVEKRGYVYIMANGPSGTLYTGV TADLAARIHQHRS GTSGDFCRDHGLTRLVYVETHERIDAAIAREKAIKKWKRAWKLNLIGQANPDWRDLWEDMQR
<DKB74_RS11065--Sphingosinithalassobacter_portus--53.33 (SEQ ID NO:1540)
MFWVYILRCS DASYTGHTDD LERRVAQHQS GEIPGYTHDRRPVSLMWQECFQTREEAIRAELRIKPWSRAKKEAL IAGDWTALKRAAMP PGERESRVAAAPR
LRSGRTEEDR
<DKG80_RS14095--Reichenbachiella_versicolor--63.74 (SEQ ID NO:1541)
MQKGGAVYIILTNKNKTVLYTGVTSDLISRMYEHRNKQHPSNFTARYNVDRLVYVEGFHSIEBAI IREKQIKSGSRKKKEAL IYSTNP EWNDLYDEI KEW
<DHW03_RS04185--Pedobacter_yonginense--65.56 (SEQ ID NO:1542)
MERGGWVYILMTNAHHTTLYTGVTSNLVRRIKQHKDKVYRNSFTYKYNI EKLVYFNFFPSIQEAISEEKRIKGSRLKKLKLINDQNP EWIDLWTEIEKW
<AOX55_RS12210--Sinorhizobium_fredii_CCB AU_25509--67.03 (SEQ ID NO:1543)
MKGYVYILASKRNGTLYMGVTRDLSHRLF EHQN DLTPGFTSKYGVKTLVWFEEFDLLTTAITREKTIKKWPRKWKLNLIEQMNPEWEDVAHYLHGL
<DFR52_RS22995--Hoeflea_marina--69.41 (SEQ ID NO:1544)
MRKGYVYILASRRNGTLYTGVTS DLPGRLYEHQNDLTPGFTSKYGVKTLVWFEEHDLVTD AIAREKAIKNWPRRWKIDLIEAMNP DWDDIAHYLHGL
<CR164_RS02215--Prosthecochloris_marina--75.58 (SEQ ID NO:1545)
MSSEFFYYVYILANRAYGT VYVGMTNDILRRVHEHKSGLVEGFTKKYGVHSLVYVEFTKDPYNAISREKTLKKWKREWKVALIEKDNPEWIDL YDSLGL
<DL796_RS07570--Kangiella_spongicola--70.97 (SEQ ID NO:1546)
MLQGYVYILTNKAKTVLYTGVTS DLIQRVHQHKQRSSGFSKYNCNLVYVEVIDGIEBAI IREKQIKAGSRANKEKLINSINPSWKDLYYSLT
<C7450_RS13775--Sinorhizobium_fredii_CCB AU_25509--80.23 (SEQ ID NO:1547)
MGARYAVYILASRRNGTLYIGVTS DILLRRVDQHRAKAVPGFTQRYDVGRVLYFERYGEISEAIAREKQLKGWNRAWKIRLIEQSNPEWRDL DATM
<C7450_RS12930--Chelatococcus_asaccharovorans--76.74 (SEQ ID NO:1548)
MDDRYAVYILASRRNGTLYIGVTNDLLRRVQHRAAAVPGFTQKYGVSHLVYFERFRDISEAIAREKQLKGWNRAWKIRLIERSNPEWHDLDATA
<C7450_RS03920--Chelatococcus_asaccharovorans--67.39 (SEQ ID NO:1549)
MKQPCVYILASRPGGAIYVGVTSHLARRIWEHRQGLADSHTKRYGITRLIYVEQYDRMDAIQRERTMKHWPRVWKT RL IASMPNPEWRDLYDDLA
<C7451_RS11350--Blastomonas_natatoria--76.14 (SEQ ID NO:1550)
MDREPCVYILASDAYGTLYIGVTSNLVGR LHQHRTEEIKSFTSR YHVHRLVRFE LFD TMEGAIEREKQIKRWHRQWKINLIESDNP NWADLAVGLGF DPLPPR
GRKNGP
<C7451_RS15940--Blastomonas_natatoria--65.88 (SEQ ID NO:1551)
MKGGWVYMMADRYRGTYTGV TADIGRVEQHRSDGGSRFVKRYGLDRVLYTEWHDIEIFAIAQEKA IKKWCRQWKVDLIEQANPNWEDLSHLINA
<C7451_RS14795--Blastomonas_natatoria--53.01 (SEQ ID NO:1552)
MAFWVYILKCADGKCYTGQIDDLERRLGEHQYGGFCFTSRRRPVELVWSEYYQTRGDAIEVEKQVKSWSRAKKIGLINGDWQM VSHFARPPHERPSTSLGTN
GGGENSIEAGGAEP
<C7451_RS06255--Blastomonas_natatoria--54.29 (SEQ ID NO:1553)
MAFYAYLLRCS DGSYYAGHTDDLE YRMSQHTSGSCGGYTAKR RPVELVWSENFPSREDAFGVERQIKGWSRAKKQALIAGDWELI SRLSRGLRNSADDGAGPS
TSSG
<DMY87_RS19095--Rhizobium_wuzhouense--70.21 (SEQ ID NO:1554)
MAGYVYIVTNHKGRTLYIGVTS DLERRIYEHRECQTPGFASKYGC RQLVWYEEHDSIRTAIQREKSLKRWYRQWKIALIEGFNP DWRDLYYELM
<C7477_RS02705--Phyllobacterium_leguminum--70.00 (SEQ ID NO:1555)
MAGYVYIVTNHKGRTLYIGVTS DLSGRIYEHREGMTPGFTSRYGCTRLVWYEEHERIIDAIQREKSLKRWL RQWKIELIESINPGWDDLYPRLW
<BJ122_RS15710--Rhodopseudomonas_faecalis--79.78 (SEQ ID NO:1556)
MAYYVYLLASRKDGALYVGVTNDLVRR IYQHRTKTIPGFTSRYNITRLVWFEVYADPISAITREKELKRWRRAWKIALIERDNPGWNDLYDSIAQ
<BJ122_RS15765--Rhodopseudomonas_faecalis--70.79 (SEQ ID NO:1557)
MAGLVPAIHVLRSSMAGGYVYVLTNRPNGLYVGVTSDLVRRVYQHREGFVDGFTKRYGLKRIVYFEHFEDIRTAIQREHNIKHWPRAWKVRMILATNP SWD
DLFPTIAQ
<DK847_RS04755--Aestuariaiivirga_litoralis--72.53 (SEQ ID NO:1558)
MEKYPVYIMASRYRGTLVGVTSNLYSRVCDHKNGTTPGFTSDYAVHTLVWYEHMTVDAAI LREKRIKEWKRNWKIELIEKMNDRDWRDLHDDIDVLATLVE
DDPL
<DK847_RS15605--Aestuariaiivirga_litoralis--69.77 (SEQ ID NO:1559)
MASGFCYILAHRRNGVLYVGVTANLPRIYEH RMGLLGFTSR YGVTRLVLYEEYPRFMDAVQREKNIKHWPRAWKINLIEAGNPEWIDL YDQLM

FIG. 11WW

<DK847_RS07445--Aestuariivirga_litoralis--64.86 (SEQ ID NO:1560)
MYIGVTSDDLTRVCAHKNGTTPGFTSTYGTNRLVWYEHRLLMVAIRREKQLKAWKRDWKIEIEKMNPDWRDLHDEIDVLATLVSE
<DN068_RS19415--Taibaiella_soli--67.71 (SEQ ID NO:1561)
MQKAGFVYVMTTLANSVLYIGVTSNLSQRVWQHRTKYYPNCFTAKYNCVKLVYYYRFDTSIGDAIAGEKRLKDRHRAFKIDLIKRNPNWQDLWEEINE
<DN068_RS11510--Taibaiella_soli--65.62 (SEQ ID NO:1562)
MQKAGFVYIMTTITNSVLVYGVSTLSQRVWQHRTKYYPNCFTAKYNCVKLVYYYRFDTIAEAIAGEKRLKRSRAFKIVLIKRNPDWKDLWEEINE
<CPJ17_RS17790--Rhizobium_tumorigenes--71.91 (SEQ ID NO:1563)
MKGYVYILASKRNGTLYIGVTRDMAGRLYDHQNELTPGFTSRHGVKTLVWFEEHDLTAATREKTMKKWPRQWKNLNIETNPWDQDISHYLHGL
<CPJ17_RS11760--Rhizobium_tumorigenes--69.89 (SEQ ID NO:1564)
MTGYTYIVTNHKGHTLHIGVTSDLARRIHERREALTPGFTATYGCRLQAWYEEHFNIQDAIQREKSLKRWYRQWKIELEKTNPDWRDLYYELW
<CPY51_RS16955--Rhizobium_tubonense--65.59 (SEQ ID NO:1565)
MSGVVYMSDKPRGVIYTGVTSDLQGRVWEHRNETQQGFTQKYKVKNLVWFETHPNIVLAIQREKSLKRYLREWKIELEGFNPTWIDLYESIAALENVYRPH
PNTRQWGDYN
<LX95_RS03580--Mesonia_algae--68.35 (SEQ ID NO:1566)
MKTSYVYIMNEYRTTLYIGVTSNLSKRITTEHQLGKASLFTAKYNLKYLIYFEEFLDIQEAIKREKQLKSWKKNWKIDLVKKNLPELKTIELT
<B5V02_RS17660--Mesorhizobium_kowhaii--67.42 (SEQ ID NO:1567)
MTGYVYLTASQKRGTIYIGVTNDLGRRIPEHKSQGGSRFTSRYGVQRLIWIYEEYFDIVDAIQREKSLKRWPRQWKNIELEKTNPEWFELFRGTGW
<LX81_RS09275--Hwanghaeicola_aestuarii--68.24 (SEQ ID NO:1568)
MAHYVYLLASRPGGALYLGSTNDLRRLEQHRSGAIPGHTQRYGIRTLVWFAAFETYSEALTQERRMKRWRRSWKIDLIAESNPWDQDVSAQIPHA
<LV85_RS18650--Algoriphagus_chordae--67.39 (SEQ ID NO:1569)
MKKGGSIYLLTNKNNTVYIGVTEDLIRRVQEHKSKFDPNSFTAKYNLDKLVYFEGFHTIDEAIAREKQIKAGSRACKEKLINSVNPWKDLYNELLSE
<LV85_RS21715--Algoriphagus_chordae--56.92 (SEQ ID NO:1570)
MDEFVYILYSKSAKTYLGMTSDLTRFHFHNYKSTKGTFLYRPLVYIVHVEFFTTKKEALIREKELTKGGRDWIKREILINYY
<DJ021_RS04195--Phenylobacterium_hankyongense--68.97 (SEQ ID NO:1571)
MADRMGTLYIGVGTQGLVRRVWEHREGIVPGFTSKYGLKRLVWFEPFEGMLNAIQREKSLKRWPREWKNLIERENPHWDDLYDYGIV
<DJ021_RS12040--Phenylobacterium_hankyongense--70.11 (SEQ ID NO:1572)
MASRRNGTYAGSTDDLRTVGHQKTHDKTSFCGRYGVHMLVWFETHESREAAFRERRRIKEWRRRAWKLRMIEAQNPDPWRDLSNNLINLLAL
<DJ019_RS10820--Phenylobacterium_kunshanense--65.56 (SEQ ID NO:1573)
MRDRTYVWYILASDKLGTLYVGVTSNLERRIAHRAGEAEGFTRRHGGRRLVWSRGFGEVTAIRFEKQLKRWRDWDKIRLIEADNPHWADLYLEMMALPPLH
PDLAS
<DMB71_RS04440--Flavobacterium_tistrianum--61.36 (SEQ ID NO:1574)
MHLQEGFHTYFIYILTNAKSVFYVGVVTNNLKNRLIKHKDNIASRNNTFASKYNVEFLLYYEKFTWIEAIAREKEIKGWRREKKIALIKTMNPNLEFLDY
<DMA14_RS02300--Flavobacterium_sharifuzzamanii--63.64 (SEQ ID NO:1575)
MFNPQKGYHTYIYIITNRHRSTFYIGVTQNLKVLKQHKENIENNKTFAKSYGIEBFLVYYEKFIWTQEAIAREKELKGWIRAKKIDLIKEFNPTFEFLNHH
FE
<DMA14_RS05880--Flavobacterium_sharifuzzamanii--64.56 (SEQ ID NO:1576)
MHQQEGYHTYIYIILTNNKYKTVFYTGVTNNLKIRLSQHKENHTGNKFTAKYNATYLLYYEKFTWIEAIAREKEIKDWREKKIELIKTINPDLDLFLNYLF
E
<DMA14_RS22095--Flavobacterium_sharifuzzamanii--65.85 (SEQ ID NO:1577)
MHLQEGFHTYFIYILTNAKSVFYVGVVTNNLKKRLIKHKENI LLNNKTFVSKYNVEFLLYYEKFTWIEAIAREKEIKGWRREKKIDLIKINPELEFLEY
<DM444_RS05355--Flavobacterium_ginsenosidimutans--67.05 (SEQ ID NO:1578)
MFNSQQGYHTYIYIITNRHRSTFYIGVTQNLKERLQHKENIENNIKTFAKSYGIEBFLVYYEKFTWIEAIAREKELKGWIRAKKIELIRDFNPTFEFLNHH
FE
<DM444_RS13990--Flavobacterium_ginsenosidimutans--65.82 (SEQ ID NO:1579)
MHQQEGYHTYIYIILTNNKYKTVFYTGVTNNLIRLSQHKENNTGNKSFTSKYNATYLLYYEKFTWIEAIREKEIKDWREKKIELMKTINPDLDLFLNYLF
E
<DJ018_RS15040--Phenylobacterium_deserti--72.73 (SEQ ID NO:1580)
MLASRKHGTYITGVSGRFLHRIVEHREGLREGFTSRYRVKRLVWFELHHAIEPAIRREKSLKKYPAWKINLIERDNPWHDLDFGLLQEDGPLSNVQR
<CH338_RS20575--Rhodoplanes_elegans--74.12 (SEQ ID NO:1581)
MEQRRFVYILASKRNGTLYVGVTDNLMRRVAEHKAKRVPGFTRQYGVQDLVHAEPFDSILEARARERTLKRWHRAWKALIEQSNPGWRDLSEWEL
<CH338_RS25775--Rhodoplanes_elegans--69.77 (SEQ ID NO:1582)
MHGGWVYIFTNRPNGTLYVGVTSDLVRRVAEHRDGVVPGFTRKYRIGKRLVYFEAYDAIRDAIQREHNMKHWPRTWKVRILLATNHEWKDLFETLF
<CH339_RS23385--Rhodobium_orientis--72.04 (SEQ ID NO:1583)
MKRPAVYIMASKRNGTLYTGVTSDLNARIHQHRTGTSGGFASQYGCRLVYAEILDDMETAIAREKQIKAGSRVRKRLRIEENPDWRDLAEDLL
<CH339_RS20105--Rhodobium_orientis--63.44 (SEQ ID NO:1584)
MSAWVYVLASRPGTLYVGVTTDLRRRVQEHREEQRRGFTKAYRVKRLVHYEHYDRNDAIQREKNIKHWPRQWKMDLVEAANPKWRDLWNDIASP
<CH341_RS28415--Rhodoplanes_roseus--75.29 (SEQ ID NO:1585)
MQRRFVYIILASRRNGTLYVGVTDNLMRRVAEHKAKQVPGFTRQYGVDTLVHAEPFDSILEARARERTLKRWHRAWKALIEQNNPSWRDLSEWEL
<CH341_RS29410--Rhodoplanes_roseus--75.95 (SEQ ID NO:1586)
MFAGGRWLRRDRLRTLYVGVTRDLVRRTWQHKSHDVPGFTRYDVRDLVWFECYDDPVTAIEREKDIKKWRRRAWKIALIEKDNPDWRDLYPEITG
<DK188_RS03475--Algoriphagus_litoralis--59.34 (SEQ ID NO:1587)
MEAFMYILRCNGLYYVGSTTNLEKRLQHQHQLGEGANFTKKHLPVSLVYFESFDRIDEAFEREKQVQNSRAKKEALISRNFKALKELSKGKT
<LV92_RS11250--Arenibacter_echinorum--58.33 (SEQ ID NO:1588)
MTKGYFYILSNSTRNLLYAGATRDVIRKRLRLHELGMASVFTTKYHIKLYIYFEEFKEIEDAFKREKQIKNWHKDWKNLIKEMNPPELLDLRHRFEQLVMHY
<LV92_RS11290--Arenibacter_echinorum--55.91 (SEQ ID NO:1589)
MFYNHFVYIITNHEHTTLYVGVTNHIIQRRLSQHYFDSQNSKSFAGKYNIYLVYVEGFENPKDAIAREKEIKKWRREKKNSLVSSFNPKWESLNNQIV
<LV83_RS18020--Algoriphagus_yeomjeoni--65.88 (SEQ ID NO:1590)
MVIKIVFVYIILTNKTKSTLYIGVTNDLQRRLYEHKNGLGASNSFTTKYKCHLIYFEEYATPIEAI SREKELKKWRREKKNNLIEKENPNWDFINIECE
<DDY94_RS13210--Winogradskyella_tangerina--72.73 (SEQ ID NO:1591)
MYKTHQYYLYILSNKYNGTLYIGITNNLERRMFEHKNQLIDGFSKYKLNRLVYFETFYITDAIRREKNMKKWKRWKINLIEEENKNWSDLSKHW

FIG. 11XX

<DN752_RS20065--Echinicola_strongylocentroti--68.13 (SEQ ID NO:1592)
MKKGGCIYILTNNKNTTLYVGVTSNLIKRIHEHKFTSTKSFYRYNINKLVYVYEFTHSIEEAIIREKQLKGGSRKKKELINSFNPDKDLWDDIQKW
<DN752_RS11790--Echinicola_strongylocentroti--58.23 (SEQ ID NO:1593)
MKGYMYILACANGAFYTGSTVDLGKRLAEHQAGFGANFTRKYLPELVVYFEEYARVDQAFYREKQIQGWGRSKKIALINGDLKHLNLSACQNKTHFKNYSK
KE
<HY3_RS02135--Hyphomonas_pacifica--72.22 (SEQ ID NO:1594)
MANEDRIIAVYMMANRKNKNGALYTGVTSNLVGRIWQHREGIFRGFTQKHGCKTLVWYEVHEWVEGAIAREKSLKSYRRQNKINLIEAKNPHWKDLWFEITG
<TZ35_RS14335--Vibrio_maritimus--67.05 (SEQ ID NO:1595)
MNAKTKKPCVYITTSKGRGVYIGVTSNLSKRIWKHKHGVTEGIVKKYQVNLVHYEYETMEEAASRELQRLQWKRAWKNDLVTNNPYWKDLSHELW
<TZ48_RS04610--Vibrio_variabilis--64.77 (SEQ ID NO:1596)
MNPKNKTKPCVYITTSKGRGVYIGVTSNLSKRIWKHKHGMTEGIVKKYQVNLVHYEWFDTMEEAASRELQRLHWKRAWKNDLVTNNPYWKDLSHELW
<CA260_RS14215--Dyella_jiangningensis--80.90 (SEQ ID NO:1597)
MDWRRPCVYILASGRNGTLYVGVTSDLVKRVWQHRNEVTGKGFSTQHVVYNLVWYELHETMESAILREKRLKDWKRVWKLRIEENPEWVDLYPTLT
<B0I10_RS02265--Flavobacterium_lacus--55.32 (SEQ ID NO:1598)
MQKGYTYILTNKNTLVLYVGVTDLKRNRI DCHKDGPVTFTKRYNVTKLIVYFEEFDNYHDAFEREKQLKNWHEKWNLA KILNPELKDLYLEL
<CLV55_RS03260--Flavobacterium_aciduliphilum--66.32 (SEQ ID NO:1599)
MKPGFYIYIVTNKNNNTTLYIGVTSNLPKRI LEHKEKRYKNSFSARYNLDKLVYFEQFQMI RDAI RREKQLKGGNRAVKLNINNFNPDWNDLYDEIKDSMI I
<HME9304_RS09235--Flagellimonas_maritima--55.81 (SEQ ID NO:1600)
MTKGYCYILTNKNTLVLYIGVTSNLPKRI SEHKNGKYSNSFTKRYNVTKLIVYFEEFDTIKDVFAREKQLKSGNRKRKEDLITSINPDWKDLSLEWQAD
<CU669_RS11530--Magnetospirillum_kuznetsovii--76.34 (SEQ ID NO:1601)
MQRPCVYILASKPNGTLYIGVTSNLPKRI SEHKADAVPGFTSRYGVHSLVYVFEHDLMA DAILREKRI KKNRAWKLQMSRGNPTWRDLFDIVR
<CU669_RS08925--Magnetospirillum_kuznetsovii--73.33 (SEQ ID NO:1602)
MQGGWYIILTNKNTLVLYIGVTSNLPKRI SEHREGVADGFTKRGHLKRLVHYFEEFDI RLAIQREKQLKHWPRAWKVRIILDANPNWDDLYDVLA
<CU669_RS17080--Magnetospirillum_kuznetsovii--65.88 (SEQ ID NO:1603)
MRDHTYTYVYLLSSSLGVLYCGVTNNLERRIAEHRSGTADGFTKRYLVNRLVWFECSGDVRAAIAREKQIKGWKRCRQVALIETLNPDWRLSKNSQSS
<DOU38_RS05425--Rhodothermaceae_bacterium--77.91 (SEQ ID NO:1604)
MPKRGYVYILASRRNGTLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DIU31_RS19985--Mucilaginibacter_rubeus--62.77 (SEQ ID NO:1605)
MHQYFVYIILTNKANTVLYVGVSDDLHNRI SEHKKIYKGYTAKYACDKWVYFEQFQLIDDAIAREKQLKAGSRQKKIDLIVATNP SWGDLSDVDWYD
<DIU31_RS04445--Mucilaginibacter_rubeus--58.75 (SEQ ID NO:1606)
MPDKQGHVYIMTNTYRTTYIIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DIU38_RS20195--Mucilaginibacter_gossypii--62.77 (SEQ ID NO:1607)
MHQYFVYIILTNKANTVLYVGVSDDLHNRI SEHKKIYKGYTAKYACDKWVYFEQFQLIDDAIAREKQLKAGSRQKKIDLIVATNP SWGDLSDVDWYD
<DIU38_RS04495--Mucilaginibacter_gossypii--58.75 (SEQ ID NO:1608)
MPDKQGHVYIMTNTYRTTYIIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DIU38_RS04495--Mucilaginibacter_gossypii--58.75 (SEQ ID NO:1608)
MPDKQGHVYIMTNTYRTTYIIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DPM33_RS14570--Mesorhizobium_hawassense--65.29 (SEQ ID NO:1609)
MTGYVYMTASQKGGTIYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DPM33_RS14415--Mesorhizobium_hawassense--65.59 (SEQ ID NO:1610)
MTGYVYMTASQKGGTIYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DPM33_RS14415--Mesorhizobium_hawassense--65.59 (SEQ ID NO:1610)
MTGYVYMTASQKGGTIYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<CDH04_RS07835--Francisella_adeliensis--72.94 (SEQ ID NO:1612)
MSQEQNKIGYIYILSNKNNIVYGVGTANIVKRVHEHKQLVGEFTKHKHLKDLVYFESYDDIRLAIAREKVLKCLDMSKLNLVKKENPKFEDLI
<DPN68_RS08415--Flavobacterium_tibetense--66.29 (SEQ ID NO:1613)
MKPGFYIYIILTNKNTLVLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DQ400_RS18825--Halomonas_sulfidaeris--63.86 (SEQ ID NO:1614)
MLTNHTRHVLVYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DR980_RS09245--Flavobacterium_psychrolimnae--55.81 (SEQ ID NO:1615)
MKFYIYVILLCNDDSLYVGVTSDIDRRVI EHNVGKYPEAYTNSRPVTLVYQDFDPNQAI EFEEKLKKWSRVKKQALIDGNFDKQLDISECRNATHHKYKS
DFEDIE
<DMZ65_RS08120--Primorskyibacter_marinus--67.06 (SEQ ID NO:1616)
MAHFVYIMASRPGAMYTGRTTDLARVETHRAGLSGHTAKYNIIRLVWFEHDCFEASIRRERAIKRWRRDWKIALITQSNPNWHELTPIQIPG
<DUT69_RS01660--Pedobacter_namyangjuensis--75.00 (SEQ ID NO:1617)
MERGGCVYIILTNVNTVLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT69_RS08730--Pedobacter_namyangjuensis--56.67 (SEQ ID NO:1618)
MIGHNYFVYMLTNFERTVLYTGVTDNLEIRLEQHRAGENSYSFTKKYKCFYLVYFERFQYVDHAIDREKEIKGWSRAKKNALIETENKFWRFLNEDIKSD
<DUT68_RS06080--Nubsella_zeaxanthinifaciens--67.03 (SEQ ID NO:1619)
MIRGGSVYIILTNVNTVLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT70_RS06725--Pedobacter_chitinilyticus--65.91 (SEQ ID NO:1620)
MERGGCVYIIMTNTVNTVLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT70_RS13645--Pedobacter_chitinilyticus--68.48 (SEQ ID NO:1621)
MVRGGCIYITNTYQRTSTLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT70_RS13170--Pedobacter_chitinilyticus--66.67 (SEQ ID NO:1622)
MERGGVYIILTNKNAVLTYGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT70_RS05050--Pedobacter_chitinilyticus--58.43 (SEQ ID NO:1623)
MKKIVYIVTDNRNRTLHVGSADLMKTLDFYKMPNLFDDSAQQLTRLVYFEEFRTEEQALSFRFKMVSFRFTRAQKERLVRSCNPDWI DLTAGLDFESMYMHQK
INNQSLLPFTV
<DUT71_RS13185--Pedobacter_namyangensis--64.84 (SEQ ID NO:1624)
MVRGGSVYIILTNVNTVLYIGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG
<DUT71_RS16205--Pedobacter_namyangensis--67.78 (SEQ ID NO:1625)
MERGGVYIILTNKNAVLTYGVTSNLPKRI SEHREGVADGFTKRYLVNRLVVEEHDDI QRAIVREKQMKKWRRAWKLELIESLNPWDDLFEQLRG

FIG. 11YY

<DUT71_RS17015--Pedobacter_nanyangensis--67.39 (SEQ ID NO:1626)
MVRGGCIYIMTNYQRSTLYIGVTSNQLNRLYEHRTKIYPASFATKYELIFCIYYEYTTIEEAIAREKQLKKWNRKTKMEKLINQVNRWKDLSPEILEW
<DFR50_RS13020--Roseiarcus_fermentans--71.59 (SEQ ID NO:1627)
MRSPCVYIMASRRNGTLVVGVTNIGRRAYEHRTGAVEGFTKRYGCKLLVWLEPHERMDEAIAREKQIKSGSRKDKLALIERDNPQWHDLFERLNM
<DFR50_RS29140--Roseiarcus_fermentans--73.56 (SEQ ID NO:1628)
MTERRYVYIMASRRNGTLVVGVTNHLGARAFQHRNGVGSQFTAKYSVKTLVWYEHYDDVREAIAREKQLKKWERRWKLELIERFNPDWRDLYETLNN
<DFR50_RS20825--Roseiarcus_fermentans--71.08 (SEQ ID NO:1629)
MANKRNGTLVAGVTSDLRSRRAYEHRQGLIPGFTSRYGCKRLVFEHYERVVDIAIAREKQIKGGSRAKKIALIEAMNPWDNDLYETLA
<DFR50_RS22235--Roseiarcus_fermentans--67.44 (SEQ ID NO:1630)
MRGGWVYIMTNRPNGLTYGTVTSNIGRRAYEHREGLVEGFTKRYGLKRLVYMEFYERITDAIQRESNMKHWPRAWKAQLIVNYPGWVDLFETLNG
<DRW42_RS27440--Pedobacter_miscanthi--65.12 (SEQ ID NO:1631)
MELGGWIYIMTNYEKTTLTYIGVTSDLRSRIYEHQNKTYNSFTAKYNLIYCVYYEAFLSIEBEAIDREKQIKKWNRSKKDALINSFNENWADLTEIINWEDE
<DRW42_RS16785--Pedobacter_miscanthi--65.91 (SEQ ID NO:1632)
MERGGTVYIITNFTHHTLTGTVDIMFRIKEHKDKKEYPKSFSAKYNTVKLVYFCSYDSIEBEAIAEEKRIKGGSRKKKIKLIDSMNPEWKDLWEIVKEW
<DRW42_RS20850--Pedobacter_miscanthi--58.67 (SEQ ID NO:1633)
MDHNYFVYIILTNKNTVVYTGVDLEARLKQHIIDNKENKFAFTYKYNCHYLVEYEGFQYIEHAIEREKEIKGWTKQKRP
<ACPOL_RS27715--Acidisarcina_polymorpha--69.41 (SEQ ID NO:1634)
MPRSYFTYIMVASRSHTLVGVTSNMEQRIWQHRNKEYEGVSARYNCNRLVCFEQYSNAGTAIAREKELKGWRRARKIELIAISNPTWVDLSEGWGQPINSTAA
PSLGSG
<RHODOSMS8_RS17650--Rhodobiaceae_bacterium--73.33 (SEQ ID NO:1635)
MEKVYVYLLASERNGLTYAGMTSDLARRVHEHKSGATDGFSGRYGVKRLVVEQTHDVPVAIAREKRIKKYPRAWKINLIERANPTWRDLYEEILG
<DRB19_RS21555--Microvirga_aerophila--78.82 (SEQ ID NO:1636)
MSEAFYVHILATRKDGPLYAGITNNLHKRVYEHKTHAVPGFTARYNVDRLVYFEI FENPEAAIAREKQLKKWHQAWKVALIERENPEWRDLAEFVPLERIRF
R
<DRJ53_RS05490--Paracnuella_aquatica--64.37 (SEQ ID NO:1637)
MNKQGVYIILTNHHTTLTYGTVSNLYQRMQQHMNIYPHSFTAKYNCNKLIVYCYAYATMMDAIEEEKRIKGSRRMDKVHLISQQNPWKDLWQTEIS
<DRJ53_RS19135--Paracnuella_aquatica--62.79 (SEQ ID NO:1638)
MLRNSFPYAVYIMSNKHLTTFYIGVTGNLQQRWQHKNDANFTGKYNCCELI FYETYTDVQQAIAREKNLKNWHREWKINLARTINPELKDAAEWY
<DRJ53_RS04770--Paracnuella_aquatica--58.62 (SEQ ID NO:1639)
MNRGYSFYVYIILTNQKQKTVLYTGTNNLHERII THWKGGEPTQTGTGRYRVHFLFYEAHTHIDQAILREKEIKGWVRKKKCDLISGFNPEWKFLNEEVFGW
PPKY
<DK368_RS13460--Oleisolibacter_albus--77.91 (SEQ ID NO:1640)
MRSYFYVYILASQRNGTLTYIGVTDNLARRVYEHKCATDGFTRIYGVNLLVYIEFDATFAHAIGREKQLKKWNRWAKIRLIEERNPQWLDLYGTLH
<DFP77_RS04180--Marinomonas_foliarum--79.35 (SEQ ID NO:1641)
MSKQPCVYIILASKRNGTFYVGVTNLIQRIWQHKNKQVQGFTEYRGVDQLVYFELYEDMSEAILREKRIKKWNRQWKVNLIEKDNLHWEDLWRQIQ
<DFQ08_RS04540--Winogradskyella_arenosi--56.47 (SEQ ID NO:1642)
MSKKYYTYIILTERNHIFVVGVTNNIERRMSEHKAATFGTHVRHYNIKKLVYFEAHSIDIRIAIKREKTIKKWKVWKRNLITKMNPEWMYLSLGNLSKYKNK
KD
<DFR58_RS02230--Anaerobacterium_chartisolvans--70.79 (SEQ ID NO:1643)
MSEHCYVYIILTNKNTVLYTGTVNDLVRRIFEHKEKMVEGFTKKYNVNKLIVYGVFSDIKSAIQREKQIKGITRQKKINLINIMNKDWDKLYRDICG
<DFR58_RS07350--Anaerobacterium_chartisolvans--70.11 (SEQ ID NO:1644)
MLTNKNTVLYTGTVNDLVRRIFEHKEKMVEGFTKKYNVNKLIVYIEFSDIKSAIQREKQIKGITRQKKINLINIMNKDWDKLYRDICG
<C7476_RS06370--Phyllobacterium_bourgognense--67.78 (SEQ ID NO:1645)
MSGYVYILGNKNGTLTYIGVTSNLIQRIWQHKNKQVQGFTEYRGVDQLVYFELYEDMSEAILREKRIKKWNRQWKVNLIEKDNLHWEDLWRQIQ
<DFR48_RS17555--Ciceribacter_lividus--71.91 (SEQ ID NO:1646)
MKGYVYILASQRNGTLTYIGVTSNLIQRIWQHKNKQVQGFTEYRGVDQLVYFELYEDMSEAILREKRIKKWNRQWKVNLIEKDNLHWEDLWRQIQ
<DT096_RS03765--Ephemeropterica_cinctiostellae--82.80 (SEQ ID NO:1647)
MKQPAVYIILASQRNGTLTYIGVTSNLIQRIWQHKNKQVQGFTEYRGVDQLVYFELYEDMSEAILREKRIKKWNRQWKVNLIEKDNLHWEDLWRQIQ
<Wcon_RS02565--Wolbachia_endosymbiont_of_Cylisticus_convexus--74.44 (SEQ ID NO:1648)
MKSYYVYILASERNGLTYGTITSNLTKRIWEHKSIAISGFTSKYNVCKLVYFEFQDINLAISREKLLKSWQRKWKIDLEKTNQEWKDYDEIVSG
<CGC45_RS02985--Francisella_opportunistica--57.14 (SEQ ID NO:1649)
MYILKGSNDSYTGSTNNLEIRLYQHNMGEAGAKHTAKRLPVELVYFEEYSRIDEAFGREKQIQKWSRAKKEALINEFFENLQALAECMNETHCKNKD
<DDF70_RS11080--Microvirga_calopogonii--77.91 (SEQ ID NO:1650)
MPYVYVYLLASRRHGTLYLGVTDNLARRVDEHKRVIPGFSALYDISRLVWYEVYDRIDEAIAREKSLKKWRRDWKVTIEEMNPWDSDLYPSLNG
<DXK31_RS21195--Niabella_yanshanensis--63.75 (SEQ ID NO:1651)
MDKGGFIYIILTNKRLTVLYTGTANLRRRIWEHEHHEKPGSFDTQYVNCFLIYEWFDNIVTAIAREKQIKKWRREKKEALISTKNPDREFLNDEVYNDIYRL
LY
<DXK21_RS11550--Brevundimonas_bullata--71.76 (SEQ ID NO:1652)
MYFYTYIILAGRRNGVLYTGSTDDLKRVVEHREKLRGGFTAKYGVSRVWYELHGTREDAFRERQIKERRAWKLDLIERMNPQWDDLPTLL
<DXK21_RS09500--Brevundimonas_bullata--71.76 (SEQ ID NO:1653)
MAFFTYIMASRRNGTLTYGTSTDDIGLRAWQHKEGAGSKFTAEHDVERLVWFETHETREAAFRERQIKKWNQWKLRLIEEMNPQWRDLVETLSA
<DXK21_RS07160--Brevundimonas_bullata--66.67 (SEQ ID NO:1654)
MNERPFIAAYILASRPLGVLYGMTSNLYERIRQHREDRVEGFAQRYGCKTFVWYETHGFTVEAICEKALKRWKRTWKLGLVEKTNPEWQDLYPSLFLGLAP
DPRVKPEDDGEGLDLFLKRLDGA
<DXK21_RS06090--Brevundimonas_bullata--68.13 (SEQ ID NO:1655)
MVNRDLIGVYMMANRMHGAIFYGVSSDLILRTDQHRRGVVEGFTRRYGLKRLVWYEAHESIVDAIQREKSLKRYPRDWKTNLIERMNPWDSDLYPALVTGGA
ESAPNVEAYRDWTPAEQAAFDLRPLPKNWRDDE
<DXK21_RS00200--Brevundimonas_bullata--65.91 (SEQ ID NO:1656)
MANRRNGAIYTGVTSDLHARVWKHRLGTGKGTDDYGCGLVWFEPQLGMTEAIAREKQIKKLRRAWKLELIEAANPVWEDLAADWYD

FIG. 11ZZ

<DXK21_RS04320--Brevundimonas_bullata--65.52 (SEQ ID NO:1657)
MREDRLFTVYIVTNTTRGVLYVGVTSDDLGRIGQHRDRYSGFSCQWGLRRLVWYQVFDDAESAINFEKRLKRWRRDWKFGIVERMNPEDLWEAAFKSEGG
VLGGPG
<DES45_RS00630--Microvirga_subterranea--74.42 (SEQ ID NO:1658)
MQERYAVYILASRKNGTLTYIGVTGNLPQRMEQHKALAVPGFTRRYRVTTLVYVERYADVTEAIAREKQLKGWNRWIRLIERSNPDWTDLDPTAC
<DUL71_RS00060--Mucilaginibacter_endophyticus--62.77 (SEQ ID NO:1659)
MHQYFVYILTANKANTALYGVSSDLNRILEHKGKAYKGYTAKYSCDKLVYFEQFQWIDDIAIAREKQLKAGSRQKKIDLIATNPSWGDLSIDWYD
<DN064_RS00030--Winogradskyella_aurantiaca--67.44 (SEQ ID NO:1660)
MVRAYHNYWVYIITNKPRGTLTYIGVTGGIDDRMERHKDGKSGFASKYKLDLLVYEEFQFINDIAIAREKQLKNWHRDWKIDLIETNPDKDLWNKTKGKGL
RC
<DL238_RS02585--Erythrobacter_lutimaris--76.34 (SEQ ID NO:1661)
MCEAFQPTVYLLASRRNGTLTYVGATSDLIKRIAEHRDGLIEGFTKDYGVRRVLVWFQAHATMDAAILREKRIKKWKREWKINLIRETNADWRDLAVDLGFAMLP
PIPSHRP
<DL238_RS01190--Erythrobacter_lutimaris--75.00 (SEQ ID NO:1662)
MLNLFQHPAGNLRIMPTEKQPCVYLLANKPFGTIYIGVTSDLVARLWQHREGAVEGFSKRYAVHRLVHFEFDTMESAIKREKQLKNWHRDWKLNLI TRENP
DWRDLAVGLGFDPPRSDEKAGP
<DL238_RS09360--Erythrobacter_lutimaris--65.12 (SEQ ID NO:1663)
MEKGGWVYIMANRYRGVYVGTADLIARAHQHRTGEGSDHVAQTGKTLVYAEHREDIEAAIAREKLVKKWRREWKFALIEAENPEWKDLWHVWFGGDSARP
<DL238_RS14720--Erythrobacter_lutimaris--66.28 (SEQ ID NO:1664)
MDRERRGWVYIMADRYGAIYVGVTSDIARVTLHRSGKGSYD CARRGLTRLVWAEHGEFIDSCIAQEKRKRWHRAWKIELVEKANPDWDDLYEMLI
<DL238_RS14750--Erythrobacter_lutimaris--65.56 (SEQ ID NO:1665)
MHGGWVYIMTNAHPGTLTYIGVTANLPARMMAHREGRGSEFCRWSLTRLVLEHYDRIEDAIKREKSVKRWKRQWKLKLIIRDNPDWNDLFETINA
<DL238_RS06570--Parabacteroides_acidifaciens--68.18 (SEQ ID NO:1666)
MNYNSQYVYIMSNYSNTVLYIGFTNDLMRRVMEHHNGAIEGFSKKYHCHKL VYEEYDMMVDACKREKQLKEWNRWAKDELINKMNPTRKDLSENWDKEGFI
CDSY
<DWU89_RS04915--Parabacteroides_acidifaciens--61.63 (SEQ ID NO:1667)
MLTNVHGNVLYTGVTNNLERRMQKHSGLISGFTSKYKCSKLVYFEEFNKIEEAIAREKQLKNWKRWNKLVISMNPEWKDLSLDWDI
<DXT99_RS18400--Pontibacter_diazotrophicus--58.24 (SEQ ID NO:1668)
MQSHNVFVYITTNPRKTTLYVGVNTDLPRLQEHFMNCGNPSTFAGKYCYNLIYFERHTRIEHAIEREKEIKKWSRKKKEVLIKTNRFEWFLNLQIQDE
<DW352_RS20840--Pseudolabrys_taiwanensis--81.61 (SEQ ID NO:1669)
MLASKRNGTLTYIGATDDLVRVWQHRNGVIDGFTKEHAVKTLVWYIEIHETREAALTRERQMKKWNRIWKLELIEKANPVWRDLWSEITR
<DW352_RS25390--Pseudolabrys_taiwanensis--80.00 (SEQ ID NO:1670)
MFYVYLLASRREGTLVGVTRDIVRRTYQHREKLLPGFTSTYDVRRLVWFYIYDDPTNAIEREKQLKKWRRWAKIELIEKENPDWKDLYPEIIR
<DW352_RS23890--Pseudolabrys_taiwanensis--70.79 (SEQ ID NO:1671)
MAAYWVYLLTNNPRSTLYVGVNTDLVRRITYEHREGLVKGFTQSYGLKMLVHFEQYDAPYDAIQREKNIKHWPRWLKQLIEASNPQWRDLYNDITR
<DW352_RS25255--Pseudolabrys_taiwanensis--72.22 (SEQ ID NO:1672)
MAGLVPALHGSNHRPMGGYVYILTNRNGTLTYVGVTSDLVRRVAEHREGAVKGFTKTYSLKRLVFEIYDDIRIAIQREKNMKHWTRTWKVRLLIQSNPQWND
LFEALL
<BX611_RS07455--Lutibacter_oceani--55.81 (SEQ ID NO:1673)
MKKSFYILKCSNDNTYYTGVTNLEKRLFQHSSGFYKNCYTNKRFPVMTVYAEFTDINLAIEKEKQLKKWSRVKKEALINGCYDDLVLNLAKKKFE
<BX611_RS07460--Lutibacter_oceani--53.49 (SEQ ID NO:1674)
MKKSFYILKCSNDNTYYTGVTNLEKRLFQHSSGFYKNCYTYKKRFPVMTVYAEFTDINLAIEKEKQLKNWSRVKKEALINGGYDDLVLNLAKKKFE
<C8P67_RS12485--Flavobacterium_aquicola--62.20 (SEQ ID NO:1675)
MQFQEGYHTYIYIITNKAKTVLYTGVTNHLRIRLEQHKSNIVTKNKTFASTYNVHLLIYIEKFTWIQEAIAIEKEIKGWRREKKMDLIKTINPDLDLNYLFE
E
<DRF62_RS05800--Chryseobacterium_piscium--63.22 (SEQ ID NO:1676)
MKTGLGTHNYVYIITNKTKTVLYTGVTNDLKNRLYWHQNP EADKNFTTKYKCFYLVYFEQFSDIETAIKREKQIKGWTRIKKENLIKEFNPTWKFLNDEV
<DRF62_RS05565--Chryseobacterium_piscium--63.22 (SEQ ID NO:1677)
MIEFSEGIYTFYVYILTNRKRAVLYTGTNNLHKRLFQHSTKQNSFTAKYNVEFLIYIEKFGWIHKAIEREKEIKNFSRKKLELIRCVNPNLDFWNHFLFE
M
<DRF57_RS11455--Chryseobacterium_rhizosphaerae--63.95 (SEQ ID NO:1678)
MIGFSEGIYTFYVYILTNRKRSVLYTGVTGNLHKRLYQHKTKKNPGSFTAKYNVEFLIYIEKYSWIHHAIEREKEIKSWPRIKKLELIREINPNLEFLNYLFE
HWF
<DRF60_RS04110--Chryseobacterium_elymi--64.21 (SEQ ID NO:1679)
MKPGFYIMTNKNHTTLYTGVTNLSKRIQHQHVEYEQSFTSRKYLYLVYWFESFQEI GDAIFREKQIKAGSRQKKLDLINSINPEWRDLTDDIENIMDVF
<DRF59_RS06410--Chryseobacterium_flavum--63.95 (SEQ ID NO:1680)
MNEFSKGSHTYIYIITNKNRKVLTYGTGNLHMRLYQHKTKLNP GSFTARYNLEFLIYIEKYDWIHHAIEREKEIKNWSRIKKLELIRANPNLEFLNYLFE
NWL
<C7460_RS16395--Marinoscillum_furvescens_DSM_4134--64.04 (SEQ ID NO:1681)
MKGGYVYIITNKNTVLYTGVTNLYARINQHRDGVNPGFTKRYNCKKL VYWEFLETITS AIHREKRLKKYTRLMKENLINQMNPWDDLYDSIEEMQ
<DFQ10_RS09465--Winogradskyella_eximia--75.00 (SEQ ID NO:1682)
MHKNVHQYLYILSNKYNGTLTYIGTNDLERRMFEHKNKLIDGFTKKYSLNRLIYFETFYQVNDIAIKREKNMKKWRQWKIDLIEKENPNWRDL SKDWPSL
<DFQ02_RS08705--Seonamhaeicola_aphaedonensis--57.50 (SEQ ID NO:1683)
MKIYYVYILKCSDKTFYTGITSNLTKRIMDHKTGKHTSSYTYKKRPVTLQFYAQFTNVELAISTEKQIKKWSRAKKLALINNDFEKLPNLAKKTFK
<C8N41_RS00235--Winogradskyella_sediminis--65.12 (SEQ ID NO:1684)
MNYWYVYIMTNKPNVYIYIGTNDIEERVKEHKLKVYPKSFTAKYNC DKLVYFEEFQGN EAEI PERQFKKWRDWKIELIEEMNPNTWDISTNWNLYNKLK
K
<C8N41_RS00175--Winogradskyella_sediminis--59.09 (SEQ ID NO:1685)
MSKKYTYIILTERNHIFVGVGTNNMERRMHEHKAATFGTHVGHYNIKKLVYFEEHSDIRVAIRREKTIKKWKAWKKNLITQMNPDWMDLSLDWDSKYKNK
<C7448_RS06865--Tenacibaculum_gallaicum--76.14 (SEQ ID NO:1686)
MHKTNHQYIYIITNKDGVLTYIGVTNNLERRIFEHKNKLKVGFSKYNLDKLIYFEEYQFIKEAIKREKNMKKWKRAWKINLVIKENPNWDDLAKNWN

FIG. 11AAA

<C7448_RS01840--*Tenacibaculum gallaicum*--59.76 (SEQ ID NO:1687)
MLDKYVYIITNKNKTTLYIGVTNNLQKRLSQHYFDSEHAKKSFVGYKNCYLIYYEAFETTNEAIYREKELKKWRREKKESLINSFNPE
<C7448_RS01855--*Tenacibaculum gallaicum*--61.84 (SEQ ID NO:1688)
MKIYFVYILQCSKSYTGVTSNLSQRI FQHONGTFKDSYTSKRRPVLEVFYCEFTNVEMAIEKEKQIKKWSRAKKEALINNEFEKLPNLAKKKFNS
<DX905_RS12075--*Sphingomonas crusticola*--75.00 (SEQ ID NO:1689)
MLNLFRRHPSRGRFRVRERQPSVYILASGPYGTLYIGVTSDLMARLHQHRQGVTTGGFTSNYEVKRLVRFEILDDMPAAIAREKQLKRWHRDWKINLIEQDNPD
WADLAIGLGLPLDSPWPNGP
<DX905_RS08050--*Sphingomonas crusticola*--67.42 (SEQ ID NO:1690)
MKRGYVYIMTNMPNGTLYIGVTSDIARITQHRTGSGSDFCKRYDLRLVYVELYPTIQEAIAREKAMKAWQRAWKLKLI SAANPRWDDLYLTNA
<DX905_RS03805--*Sphingomonas crusticola*--57.53 (SEQ ID NO:1691)
MTFWAYMLRCADERFYVGHDTNLEGRMSQHGHGRCYDFTARRRPVALVWSQEFPTREEALSARERLKGWTRAKKLALIEGDWARI SWLAI PPKERPSTSLRTN
GEGAITALPSPDPSPRKREGNS
<DXU93_RS02115--*Brumimicrobium aurantiacum*--68.24 (SEQ ID NO:1692)
MRINRVYFVYILTNPSKTTLYVGYSDDLKRVHQHKTGYYESFSKRYNI TKLVYYETFNDAKLAFSRRERQIKKYRRSKKEALIEIDNPKWEEIIPISN
<DXU93_RS02110--*Brumimicrobium aurantiacum*--67.06 (SEQ ID NO:1693)
MRINRVYFVYILTNPSKTTLYVGYSDDLKRVHQHKTGYYESFAKRYNI TKLVYYETFNDAKLAFSRRERQIKKYRRSKKEALIEIDNPKWEEIIPISN
<DZC52_RS15335--*Wenzhouxiangella sediminis*--69.66 (SEQ ID NO:1694)
MYVYVLTNKPRTLYIGVTNNLVRVWEHRGRFVSGFTRKYHLSKLVHYEGFDELVEAIQREVRLKRWRRQWKIQLIERENPEWMDLYPGIADGAVDVSSCY
<DXN05_RS08935--*Blastomonas fulva*--63.89 (SEQ ID NO:1695)
MFIVYVLYSERFDKIYIGYTSNLPRLRSHNELGKKGWITRYRPWLLIHTTETFTKQEALSREKELKNKAWRAWIRDALIPR
<DZ858_RS10830--*Marixanthomonas ophiurae*--61.36 (SEQ ID NO:1696)
MLYSYFVYIVTNTYRTTLYIGVTNNIQRISQHYFDSQNGGKSFTGKYNCHYLVYEGFESPKDAIAREKELKKWRREKKNKLI TSFNPKWESLNNQVI
<B5J99_RS17215--*Blastomonas fulva*--68.60 (SEQ ID NO:1697)
MANRRKGTLYIGVTNDLVRRADEHRNGQIDGFTKRHGCTRLGWFEAFEDLHEARLVEAQMKKWKRAWKVRVIEELNKDWDLLWQQLGL
<B5J99_RS01825--*Blastomonas fulva*--63.86 (SEQ ID NO:1698)
MMADRYRGTIYTGVTADIGRRVEQHRSEGGSRFAKRYGLSRLVYTEWHEIISAIEREKAIKKWRRQWKIDLVEEANPEWDLDTPLMNA
<B5J99_RS17215--*Blastomonas fulva*--54.67 (SEQ ID NO:1699)
MAFYAYLLRCNDGSFYAGHTDDLEHRLSQHQAGTLGGYTAKRPPVTLVWSEDFGTRDDAFAVERQIKGWSRAKKEALIAGDWQLISILAKNRMPSGTSGAGPS
TGSG
<MXM03_RS13560--*Maritalea myrionectae*--77.78 (SEQ ID NO:1700)
MGGWVYIMSNKMGNTLYVGVTSDLVRRVAEHREGVVEGFTKDHELKRLVYFEAHEIGAIIQREKSLKRSRAWKVELIQKFNPFWDDLFQICG
<MXM03_RS09770--*Maritalea myrionectae*--74.19 (SEQ ID NO:1701)
MGGWVYIMTNKINGTLYVGVTADLVRRAWEHREGVVEGFESHYSRLKSLVYFEQHDSSIAAVQREKTIKGWPRRWKVELIEQSNPRWDELYDEIAG
<DXA39_RS06215--*Anaerococcus nageyae*--70.93 (SEQ ID NO:1702)
MSNDYVYIVSNKTDSTVYIGMTHDLEKRLYEHKNHLAKGSFTDKYNCYKLVYEWIDDDYSALAREKQIKKWNRNKKNKLINSINPEWKDLSLEWYK
<ASB58_RS20170--*Pseudomonas gallaeciensis*--61.36 (SEQ ID NO:1703)
MTKQPCVYILASQRNGTLYIGVTSNLLKRVWEHRNGVVSFGFTQRYQIHHLVYFEQAADMLSAISREKQLKRWKRQWKINLIEQNNPAWDDLYTTLTAPKQALL
<DWB85_RS08145--*Halioglobus sediminis*--78.49 (SEQ ID NO:1704)
MDFQPAVYMLASKRHGTIYIGVTSNLRVRIWQHRRNHVTAGFTARYCVNRLVWFELCSSMYDAIQYERLKNWKREWKIELIERNPAWDLWGEVSSG
<DZK30_RS20170--*Taibaiella koreensis*--61.36 (SEQ ID NO:1705)
MRSYFVYILTNKTKTLYTGVTNDLSRRIDEHMQDALGNKQSFSGKYNCIHLLYEYRDIREAISREKEIKGWTRAKKVTLIKLANPEMQFLNIG
<DZK30_RS22570--*Taibaiella koreensis*--56.25 (SEQ ID NO:1706)
MKAYVYILKCSNGMYTGSTTNMERRLEQHHQGIGAQFTSKFLPVTLVYLESFTRIVDAFAREKQIQGWSRKKKEALIKGRYDQLPPELAKSYTKKNQON
<DFR49_RS09405--*Hephaestia caeni*--76.09 (SEQ ID NO:1707)
MEKQPAVYIVASARNGTIYTGVTSHLLGRVHQHRTGEVEGFTKRYHVRRLVWFEPHEIMESAIIVREKRIKNWPWRWKDLIEKDNPTWRDLAEDFGFAPLPQA
LDGSTPHRQRSTASVWDSGS
<DFR49_RS20660--*Hephaestia caeni*--68.97 (SEQ ID NO:1708)
MRERLPCVYMMASGFLGTLYVGVTSNLVGRVMQHREGTFEGFTKRYGCELLVWYDVGETMETAIAREKTIKKWPRDWTCLNIERTNPEWNDLAVGLGLEPLS
<DFR49_RS14085--*Hephaestia caeni*--68.97 (SEQ ID NO:1709)
MQARRVTIERTPCVYILARSYHGTLYTGVTNSLIGRI IQHREGSFEGFTKRYAIRRLVHYEVAETMEVAIHREKQIKRWPREYKYNLIERLNPTWNDLAVGLG
LPPLPT
<DFR49_RS18110--*Hephaestia caeni*--66.67 (SEQ ID NO:1710)
MSGGYLYILANRKNALYTGVTADIAARMQHKAGKGSKFCREYIGIDTLVYVEPAEHI EDIAAREKAIKKWRRRAWKIDLIEKANPDWRDLFFDLV
<DFR49_RS07815--*Hephaestia caeni*--66.28 (SEQ ID NO:1711)
MRRGYVYIVTNEPSGTLYIGVTSNIAARVHQHRSFGTSDFCCKYRLRLVLEPYDTIEEAIARERALKAWKRSWKNLIGRSNPDWSDLYPTLHLD
<DFR49_RS21135--*Hephaestia caeni*--62.92 (SEQ ID NO:1712)
MRGEDRGWLYIMADRYRGTMVYGVGTADIRSRVHQHRTGHGSDFCCHRYGLVWADHTPLITDAIAHEKRLKRWRAWKFDLIERANPEWHDLFYDLA
<DFR49_RS15250--*Hephaestia caeni*--56.41 (SEQ ID NO:1713)
MSFWAYILHCRGGVYTGHTDNLVDVRI GQHQDSTMPGFTRKYAPVTLVWSQEFPSRLEALEAERRIKGWSRAKKLALIRGDWTTISALAKSKNGPSTSSRAA
EGELLSKSVRAELVEAPLFSPQGSVK
<BXY53_RS12430--*Dichotomicrobium thermohalophilum*--72.04 (SEQ ID NO:1714)
MKPIHNDRQAAYVILANRRDGTLYTGVTSNLVKRIHEHKTGVSEGFARYGCDRLVHFELFDHIADAIAREKQIKAGSRAKKIALIEAHNPGWADLYPEIIA
<D3H65_RS02495--*Paraflavitalea soli*--58.51 (SEQ ID NO:1715)
MKTHTHYVYILTNNSKTTLYIGVTNNLEQRIIEHYLNRGNSTSFAGKYFCHFLLYESFGMIQAAIAREKELKRWSRHKKTILINI TNPEWKFLNKDLFDQW
PPREPISHRGIY
<D3H65_RS24085--*Paraflavitalea soli*--63.75 (SEQ ID NO:1716)
MAFFVYIYSEKLDKYIYGVTEENVEERLRKHQCNHSGFTGKASDWSIRFTELHPDKTAALKRERQLKGWKSRI RVEQLIQKNSHQATPHL
<D5I41_RS01965--*Vibrio superstes*--73.03 (SEQ ID NO:1717)
MKQPSVYILSSINRNVLYIGVTSNLKGRVWQHKRHLLPGFTLRYHVTRLVYFEVHDDMEHAIEREKRLKCWRRRAWKEALIESKNPKWNDLYDGL
<D1164_RS20550--*Mariniphaga sediminis*--61.54 (SEQ ID NO:1718)
MKAWTYLTNNKANKVLYAGVTSNLKSRMESHRTKKYPNSFTARYNADKLWFQEFDSIIDARDRERQLKAGNRARKIKLIEEMNPEWKDLYETL

FIG. 11BBB

<D3Y59_RS00065--Hymenobacter_oligotrophus--67.06 (SEQ ID NO:1719)
MRHHRFVYIILTNQNHMVLVYGVGTNDLGRRVTEHRSGVHQGFTRKYNVTKLIYFEEYTDINAIAAREKQFKSGSRQKKDLAVASLNPNDLELP
<D3Y59_RS12335--Hymenobacter_oligotrophus--58.82 (SEQ ID NO:1720)
MYVYIILTNPTQTVLYIGVTNDLARRLYEHSQGRGDAGRFTGRYQADLLVYFEQCPDATQAIAREKQKLGWTRQKKDALIAAFNPWTQAI DLQSWSM
<D3Y59_RS06700--Hymenobacter_oligotrophus--56.32 (SEQ ID NO:1721)
MPYVYVITTNPAKTVLYIGVTNDLANRLAQHYANRGEPTGAGKYFCYNLLYFEQHQSSSLNAIAAREKELKGWTRAKKDALIDTLNPAREFLQP
<D1222_RS06020--Henriciella_algicola--77.65 (SEQ ID NO:1722)
MAFYVYIMTNKPHGVLYTGMTDDINRRAWEHREHVIGGTAKYNCEMLVWVEVHETRESAFERERRIKKWERAWKIRLIEEANPDWRDLADALM
<D1222_RS00340--Henriciella_algicola--71.59 (SEQ ID NO:1723)
MAFYVYIVTNQNRNGTLYIGHTDDIERRAHEHRNGVIRGFASKYGCRLVWYDTFGTRDAALTRERQMKAWRREWKISLIEKDNPGWRDLADDWHLLGVRHYRM
PYPPAAMPNPDFRQEGGQSSAHPKLISATPAKAGAHGGWGGTSKVAR
<D1222_RS06735--Henriciella_algicola--68.09 (SEQ ID NO:1724)
MKPFYVYILASQKNGTLYTGSTDDISKRVWEHKQGGAAFTTRKYAVHRLVWLEPHDSRDSAKTREYQIKAWKRAWKIKLIETENPDWDLVLYLKNH
<D1224_RS14610--Henriciella_barbarensis--72.53 (SEQ ID NO:1725)
MASWDKVTAVYMLANRRNGTLYTGVTSDLVRRVGEHREGHIPGFTKRYGCKHLVWVEVHDLVSQAIYREKSIKSYRQNKINLIEAQNPDPDRDLWFELVR
<D1224_RS07810--Henriciella_barbarensis--72.94 (SEQ ID NO:1726)
MAFYVYIMTNKPRGVLYTGMTDDINRRAWEHREHIKGTDRYCNCEMLVWVEVHETRESAFERERRIKWKRDWKIELVEGRNPDWDLADSLM
<D1224_RS00180--Henriciella_barbarensis--71.59 (SEQ ID NO:1727)
MAFYVYIVTNQNRNGTLYIGHTDDIERRAHEHRNGVIRGFASKYGCRLVWYDTFDRDAALTRERQMKAWRRDWKISLIEKDNPGWRNLADDWHPLGARHYRM
PYHPAAMPDPFRQKGGQSNSTQSKTITAAPAKAGAHGGWGDTPPEAR
<D1224_RS07085--Henriciella_barbarensis--68.09 (SEQ ID NO:1728)
MKPFYVYILASQKNGTLFTGSTDDISKRVWEHKQGLGAFTTRKYAVHRLVWLEPHDSRDSAKTREYQIKAWKRAWTKLIEAENPDWDLVLYLKNH
<D1224_RS03730--Henriciella_barbarensis--66.28 (SEQ ID NO:1729)
MARYDLIATYMMANRNGTLYTGSTSDLWTRIGQHKSGQSVFTAKYGCNRLVWCFHQDI SAALHRERRLKAWPRQWKINLIEEVNPDWDLTMTLPFT
<D1614_RS04095--Maribellus_luteus--80.00 (SEQ ID NO:1730)
MANFYVYILANKPNGVLYIGLTNELERRLGEHKDKVLKGTFSKYNVDKLGYYEEFDTYEEAFSRERRLKKNREWKIELIEKENANWEDLSAE
<D1614_RS11815--Maribellus_luteus--61.18 (SEQ ID NO:1731)
MKDHNFYIYIVTNKNTVLYIGVTNDLQRRTFEHENELIPKFSKKYNCRHLVYEHFHQIEYAIAREKEIKKWRREKKENLINQFNPKWNFLNSKIYEAL
<D1614_RS13615--Maribellus_luteus--63.22 (SEQ ID NO:1732)
MTAWTYITTNKHNVLYTGVTSELKERIKSHKDKKYPNAFTARYHADKLWVFPKFDISI DAREREKQIKAGSRAKKIKLIEEMNPEWKDLFETL
<DUW71_RS05440--Salinibius_halmophilus--80.68 (SEQ ID NO:1733)
MKVPSVYILASKNGTLYVGTSDPIKRMWQHKSQYFSEGFTSKYGCCELLVYIEHCEIESAIAAREKQLKKWNRWAKISLIEANNPEWQDLSTSLDH
<D2V08_RS16715--Muricauda_lutimaris--64.77 (SEQ ID NO:1734)
MKESFVYIMTNKYRTTFYIGVTNVLKRI TEHHEGVGSNFTGKYKLNLDLVYFEVYSDMNQAIAREKQLKNWRHDWKLNLNMEQNPTLKTLDYG
<D0T11_RS03540--Hymenobacter_rubripertinctus--59.77 (SEQ ID NO:1735)
MRYHQYVYVYVLTNPRTVLYIGVTNDLVRRLLEHVANEGKPHFTAGRYFCSQLMYWELFGNINQAIAREKEIKGWRREKKDALIATLNPDPDRDLTSEVTP
<D0T11_RS09825--Hymenobacter_rubripertinctus--55.68 (SEQ ID NO:1736)
MSYVYVYITTNPAKTVLYIGVTNDLANRMAQHYANRGSQTFAGKYFCYNLLYFEEHQNSKAAIAAREKELKGWTRAKKEALIAVDNSNWVLET
<D0X99_RS18800--Algoriphagus_lacus--65.93 (SEQ ID NO:1737)
MQKGGAVYIMTNSHNTVLYIGVTSDLFRRVFEHRNGLNPNFSKYNVSKLVYFESFHSIEBAIVREKQIKGGRNRKEDLINSINPEWKDLWEEIQNW
<D2V17_RS01295--Erythrobacter_xanthus--77.11 (SEQ ID NO:1738)
MANRKHGAIYTGVTSDLQRRVWQHREGMVDGFTRYGCKLLVWFELHATMELAILREQQIKGGSRRKKVALIEEDNPDPDFDLWHKMS
<D2V17_RS13605--Erythrobacter_xanthus--65.56 (SEQ ID NO:1739)
MDREQGGWYIMADRYRGAMYVGMTADLASRIDQHRRGVGSDFCAKYGLKRLVWAEQVTPIDLCIAQEKRIKRWKREWKFALIEKANPDWRDLFDLLV
<D2V17_RS02720--Erythrobacter_xanthus--68.60 (SEQ ID NO:1740)
MEKGGWVYIMANRYRGGMYVGVASDLIRRVHRHRTGEGSAHVAERDKTMLVYAEHHEIEQAI VREKRIKRWREWKFALIEENPDWRDLWEVWFSKE
<D2V17_RS02975--Erythrobacter_xanthus--62.79 (SEQ ID NO:1741)
MDKGGWVYIMANRYRGGMYVGVTSNAMARIARHREGTGSRHVADF GKTRLVYIERHEEIEQAIAREKLVKKWRREWKFALIEADNPDWDDLWERWFAPED
<D2V93_RS18225--Muricauda_taeansensis--65.91 (SEQ ID NO:1742)
MEQSFVYIMTNTYRTTFYIGVTNLANRVVEHQEGSGSKFTKKYNLTDLVYFEVFSDDIDQAIAREKQLKNWHRDWKINLIKEQNPTLRTLDY
<D2V93_RS04400--Muricauda_taeansensis--64.63 (SEQ ID NO:1743)
MNKYVYVYILTNKNTVLYTGVTNDLTRLIEHIEENIEQQRKRTFAAKYCKKHLVYEEYAWIQDAISREKEIKGWVRIKKLQLIKTFNPDMDFLDGLFLR
<D2V07_RS07145--Erythrobacter_zhengii--66.28 (SEQ ID NO:1744)
MDREMHHGGWLYIMADRYRGTMVGM TDLP IRI TKHREGKGSEFCAKYGLRLLVWAEQVTPIDLCIAQEKRIKRWKREWKFALIEKANPDWEDLYERLA
<C2846_RS13895--Pseudomonas_jilinsensis--75.00 (SEQ ID NO:1745)
MREPCVYIMSNKPNGLTVGVTSNLPARVWQHRTGAVPGFTSRYGLTKLVWYERHEQMAEAI SREKQLKAGNRAGKIRLIEESNPHWNLYTSL
<D4L85_RS19955--Chryseolinea_solii--63.64 (SEQ ID NO:1746)
MAKGCCVYIILVNNRHTVFYVGATSNLYARITQHREKVNRSFTARYNINKLVYETFD SRKEAIAREKQLKTYSRICKIELIVGFNPENWDLNTYVKYW
<C0730_RS03935--Flavobacterium_anhuiense--62.50 (SEQ ID NO:1747)
MFNPQNGYHTYIYIITNKHRSTFYIGVTQNLKVLKQHKENIENNNKTFAAKYGIEFLVYEEKFTWIEQAIAREKELKGWIRTRKIDLIKKFNPTFEFLNYH
FE
<D5I55_RS05490--Aestuariisphingobium_litorale--71.74 (SEQ ID NO:1748)
MPKQPCVYIILASRRHGTIYGVTS DVLVGR LHAHRTADGFTKEHRVARLVRYEQFPDMLSAIAAREKQLKNWRRRAWKINLIEQDNLDWHD LAVDFGFEPLISA
ARFMPDPETSSG
<D5I55_RS07175--Aestuariisphingobium_litorale--66.30 (SEQ ID NO:1749)
MAKQPCVYIILASNRHGTIYIGVTGDLQLRLHHHRSGASSGFTQEYRVFRLVRYEIFEDMLSAIAAREKQLKNWRRDWKINLIERDNRDWHDLATDFGFAPISRT
LRFMPDPETSSG
<D5I55_RS17320--Aestuariisphingobium_litorale--74.71 (SEQ ID NO:1750)
MASQPNGTIYLGVTSNLAQRAYQHRHGLVEGFSKKHGTLLVWYEQHDDIQDARIREYQMKSWKRAWKLR LIEGRNPQWRDMFGLMR
<D5I55_RS09815--Aestuariisphingobium_litorale--65.59 (SEQ ID NO:1751)
MGGFVYIMTNKPLGVLYGVTS LAERVM AHREGRGSVFCRQWDLRLRLVYEHHP TIEGAIAREKTLKKWKRVWKLRLIAKSNPNWDDLQYSINA

FIG. 11CCC

FIG. 11DDD

<BC961_RS04575--Flavobacterium weaverense--63.95 (SEQ ID NO:1785)
MIEYTEGIYNFYVYILTNKNKTVLYTGVTTNNLRRRLNEHKEKINPHSFSAKYNVEFLIYYEHFGWIQLALTREKEIKDFRRELKLDLIKSFNPQLQFLNSHFD
<BC961_RS04930--Flavobacterium weaverense--63.75 (SEQ ID NO:1786)
MIEFQEGHHTYVYVYILTNKAKTVLYTGVTTNNLKI RLKQHKHDSLNPNSTAKYNVHLLYFEKFTWIIQIAIAKEKEIKGWRKDKKII TLIKTINPDNLNLDLFE
<EAX61_RS04700--Dokdonia sinensis--60.23 (SEQ ID NO:1787)
MKTMYTYIMAHENRFLVYGVVTNNIKQRTFDHKNATYETHGTGRYNIQQLVWFEVHSSPMDAIRREKTIKKWKREWKINLIELMNPYVWDLAADWDFNTYITNK
ERFKDL
<EAX61_RS15595--Dokdonia sinensis--60.00 (SEQ ID NO:1788)
MTYTYTYILTNKYNKVIYVGFTNDLKKRILEHKNKKYKNSFSKGYNVDKLVYFESHKTAQAARKREILIKNNWREWKNNLVNDMPDNDLGMWL
<EAX61_RS08670--Dokdonia sinensis--62.79 (SEQ ID NO:1789)
MPYTYTYILTNKNKTVLYVGATSDMRKRI FQHKTKFYPKSFAQYNCYNLMYEVEFDTWKEALAREKQLKAGNRKRKEKLINSRNPWKDLAGSFASLGEVPSL
HSELFYIEEIIW
<EAX61_RS08375--Dokdonia sinensis--57.14 (SEQ ID NO:1790)
MGVPGYFYMLTNHKKRLLYAGATKDLRKLRLVHKTGYGARFPKRYGVVDLVYVEKFPDVHDAFLREKQIKNWHKEWKWNLIKTKPNLEELSDFID
<C1707_RS02690--Caulobacter flavus--74.19 (SEQ ID NO:1791)
MAFYVYIVASRRNGTLYIGHTDDLARTWEHQEKVYPGFTRKRYDVGILVWYEVHDTRESVLIRERQMKKWRRLWKLELIERTNAGWRDLAQDLNC
<C1707_RS17445--Caulobacter flavus--67.82 (SEQ ID NO:1792)
MAVYIMANERLKTLYIGATGYLNRRAWEHREGLIPGFTKKYGLKRLVWYEPHDDMATAFQQEKS LKRWPRQWKINLIERTNPFWDLLYPLGA
<EFK50_RS16095--Marmoricola ginsengisoli--56.96 (SEQ ID NO:1793)
MPYVYMVECCDGSYYVGSTWDLERRIEHNLKLSRYTRRRTPVRLVWFEEYSRIEDAFLEKQVQGWGRPKRQALIDDRDLDDLPLSKSRDVG
<D7004_RS17830--Pedobacter jejuensis--68.18 (SEQ ID NO:1794)
MEQGGCVYILSNFKHTVLYIGVTSELYFRIKEHKEKFPKSFSAKYNCNTLGFFEQFSTIEEAI IREKQIKNWKRWSWKINLINEFNPNWDDLFLTLE
<D7004_RS11575--Pedobacter jejuensis--62.07 (SEQ ID NO:1795)
MERGGFVYIMTNKNHKVLYTGVTSNLRNRVFEHINNHPKSFSTQYKCYKLVYNNFFWSIEEAI AEEKRIKGGSR LAKLKLINLFNPIWNDLYETLD
<ED236_RS10825--Pseudomethylobacillus aquaticus--77.42 (SEQ ID NO:1796)
MKQPCVYILASKLNGTLVYGVTS DLMRRIWAHKHDLIDGFSKKYAVHDLVWYEQHQEMRTAIQREKHI KSWKRAWKIELIEQLNPAWSDLYNDLL
<EG854_RS11430--Brevundimonas halotolerans--64.84 (SEQ ID NO:1797)
MGTHRFTFIATYIMASGRNGTIYTGSTSNLESRVWKHKTRAFPGFTDENDCDQLVWFERHTWVTD AIKRERRLKWKMRDWKLALIEAHNPEWRDLAADWFPEND
LDWTPPPEED
<EG854_RS07830--Brevundimonas halotolerans--68.60 (SEQ ID NO:1798)
MREHSYVWYIVASGRCGWLYVGMTNDLIRRI GEHRHGQVDGYSRDHGTDRLVWYEPHRYVDQAILREKRI KRWRPWKFALIEAANPDWADLYPALALGPGVA
AGSR
<D9T21_RS18760--Alteromonas oceani--75.00 (SEQ ID NO:1799)
MIEKYPAYVYILTNYSKSVFYTGVTNLLQRIYQHKMHLADGFSKKYGLKLVYFEMHEDIEAIAAREKRLKRWRDWKIKLIEKQNPQWNDLYPLIQG
<EG028_RS07760--Chitinophaga barathri--65.52 (SEQ ID NO:1800)
MERGGTVYIMTNKHHHTVLYTGVTS DVS RVIEHHEKYYPQSFTAKFNVNKLVIYYRTFTSIEEAIMEEKRIKGGSR AKKIALINSVNPWSKELSEEL
<EIY84_RS07750--Amniculibacterium aquaticum--62.12 (SEQ ID NO:1801)
MFLVYILFSGTLGKFYIGMTSDIEKRNQKHLEHNDGFTSKAKDWKIVFTQEFDDKSSALKREAQIKNWKSKKMILNLISS
<EGC82_RS14475--Shewanella livingstonensis--76.14 (SEQ ID NO:1802)
MDKQSYIYIMANRPNGTLYTGVTS DLI RRVNWEHKNGLSDGFTKEHSIKMLVYEVFDDIYNAI TREKQLKNWKREWKIQLIEKFNPRWEDLEI
<EHO51_RS13325--Methylocystis rosea--69.66 (SEQ ID NO:1803)
MRGGWYVYIMTNRPNGTLYLGVTTNDLARRAFEHREGRSGFTQIYGLRRLVWYETHDDIRTAIQREKTMKHWPTWKVRLILEMNPDPWRDLYEDLAN
<EIB18_RS18615--Caulobacter vibrioides--60.44 (SEQ ID NO:1804)
MFSPGRTRMVPMDAWAVYMMADHYRGTL YAGVTAQFITRITD HREGRGS AFTRRYGLTRLVWYEEHMLMTEAI QRETSIKRWPRQWKINLIERNPRWDDLY
AAVFEWTPVPKWTGPFPPPS
<EG849_RS02915--Flavobacterium macacae--55.06 (SEQ ID NO:1805)
MEKGFTYILTNKI FTVLVYGATRD LKTRIDCHRN GTGALFTKKYNAPIL IYFEEFDDYNDAFRREKQLKNWHKDWKNLAKLSNPQLKDLYLEL
<EFJ44_RS05425--Brevundimonas fluminis--69.66 (SEQ ID NO:1806)
MAFFT YIVASRRNGTLYTGSTDGLYVRTLQHREGTFRGFTKRHGVA MLVWFEKHP TREAAFI RERRIKEWKRPWKLELIERANPGWRDLFADLEAGHGRQAKD
WTPPSDDPAAP
<EI424_RS04760--Tenacibaculum litoreum--61.36 (SEQ ID NO:1807)
MLENYIVYIITNKNKTTLYIGVTNDLQKRLSQHYFDSEHAKKS FAGKYNCYYL VYENFESVKEAI QREKELKKWRREKKENLINSFNPEWRFLNNEVF
<EFR00_RS17440--Rhizobium sophoriradicis--69.89 (SEQ ID NO:1808)
MAGYVYIVTNQKNGTLYIGVTS DLERRI HEHREGLTPGF AWKYGCTRLVWYEEHWDIGSAI QREKSLKRWNRQWKDLVEAMNPGWDDLYLTW
<EFR00_RS20755--Rhizobium sophoriradicis--67.02 (SEQ ID NO:1809)
MRGYVYILASARNDTLYTGVT RDLAGRLYEHQNELTPGV TWKFGVKT LVWFEEHDLTTAI TREKTIKKWPRQWKLNLIERENPEWEEISFHLLGL
<EI291_RS21945--Hymenobacter rigui--58.62 (SEQ ID NO:1810)
MYVYIILTNTTKTVLYIGVTNDLARRLYEHSQSRGDTGKFTGRYQTDLLVYFEACPDAMQAI AREKQLKGWTRSKKDALISAFNPSWNPIDLETLS
<EI291_RS11460--Hymenobacter rigui--60.71 (SEQ ID NO:1811)
MYIYLLKNPAHTVLYTGITNDLERRLYEHAQNLGNASRFTGRYQC NLVYFEVCADAVQAI AREKEIKGWSRAKKKALIATTNSDW
<EI293_RS15345--Hymenobacter perfusus--58.82 (SEQ ID NO:1812)
MYVYIILTNTTKTVLYIGVTNDLVCRLYEHSQSRGDTGKFTGRYQTDLLVYFENC PDATQAI AREKQLKGWTRSKKDALITAFNPEWQCIDLETWQ
<EI293_RS15495--Hymenobacter perfusus--58.24 (SEQ ID NO:1813)
MSFYVYILTNARTVLYTGVTNNLERLHEHNKGLGEAGKFTGRYQC NLVYFKTHPNATQAI AREKQIKNWTRNKKEYLINLLNPSWEQIDL
<EDE15_RS07215--Edaphobacter aggregans--70.79 (SEQ ID NO:1814)
MRDYEFFVYILCSR SRKLYVGITNDLIRRVGEHKS GAVEGFTKRYATRLVYFERFQYVSNIAI AREKELKDWRELKVQLIEAENPTWEDLAADM
<EHW67_RS08495--Arenibacter aquaticus--58.06 (SEQ ID NO:1815)
MFYNHFVYITNEHRTTLYIGVTNNIQRRLSQHYFDSQNSKKS FAGKYNCYYL VYEGFESSKDAI AREKEIKKWRREKKKSLVTSFNPKWETLNNQII
<ELY12_RS00120--Altererythrobacter rigui--68.89 (SEQ ID NO:1816)
MSKGGWYIILANRRNGALYTGVTADIAARMVQHRDGTGS AFCEHGSRLVHLEP YDSIEDAI AREKAVKKWRRAWKIELIERGNPDWRDRWFELIS

FIG. 11EE

<ELY12_RS10945--*Altererythrobacter rigui*--62.35 (SEQ ID NO:1817)
MDRETGGWVHILADRFRGALYVGVTSDLAARMVQHRDNGSAYCARRGITRLVWVAERGDSDIEDCIEDEKRLKRWRREWKFDLIEKTNPDWDDLFEHLT
<ELY13_RS12095--*Altererythrobacter xinjiangensis*--82.09 (SEQ ID NO:1818)
MHGFAFPTVYLLASRRNGTLYIGVTSNLLTRLHQHRRERLLPGFTSKYGVKHLVWFEPHGTMEQAILPEKRIK
<ELY13_RS05610--*Altererythrobacter xinjiangensis*--64.10 (SEQ ID NO:1819)
MRGGWYIILTNKPHGVLYIGVTSIDIAARMQHRVGTGSFACRRYGLKRLILVERHATIYEAIAREKALKAWQRAWKIRLIEGSNPD
<BLTE_RS14300--*Blastochloris tepida*--77.91 (SEQ ID NO:1820)
MEFVYMLASKNGLYIGVTRNLIRRVHEHRIKAVPGFTAKYDVTRLVHVETFDNPETAITREKQLKGWNRAWKIRLIERENPEWRDLFDELALG
<BLTE_RS04935--*Blastochloris tepida*--73.33 (SEQ ID NO:1821)
MAGWVYILSNRPNGTLYVGVTGDLIRRVWQHKTTHVVASFTARYDLHRLVYFEEFATLPDAIQREKNIKHWPRACKVRLIHAVNPDWDDLYESLI
<SAMIE_RS15905--*Sphingobium amiense*--75.27 (SEQ ID NO:1822)
MQSDFDPTVYLLASHKNGTLYVGVTSNLAQRLHQHRSGLIEGFTSDYGVHRLVWFSEGGSMEEAIRREKQIKKWNQWKINMIERNPNPNWRDLALDFGFPELG
SRADQTPPASDSAGDGFPPSRE
<SAMIE_RS12430--*Sphingobium amiense*--63.53 (SEQ ID NO:1823)
MQGGWVYIMADRYRGTMVGTADIAARIYQHRSGAGSDFCARYGLRRLVWAERGEDIRACIEQEKRVRKRWRREWKFDLIERVNEPWRDLSETLA
<EL388_RS00705--*Sulfuritortus calidifontis*--87.10 (SEQ ID NO:1824)
MQPCVYMLASRRNGTLYIGVTSVLKRVWQHKNLVEGFTSRHAVHTLIWHELHETMESAIVREKQLKKWNRAWKRLIEESNPEWQDLYETIA
<EF888_RS09855--*Silicimonas algicola*--58.43 (SEQ ID NO:1825)
MHVYVYMLASRPFAGLYVGVTSSLSHRLDQHRSGRGSVHTAKYGITTLVYRVFETRSDARDFERRLKRWRRGWKEALIAEINPRWRDLGTSIHHDH
<EJC49_RS10675--*Mesorhizobium carbonis*--66.29 (SEQ ID NO:1826)
MTGYVYMTASQKGGTIYIGVTNDLARRIPEHKSAGSRFTGRYGVTRLVWYEEHFDIRDAILREKTLKRWRRKWKVALIESMNPGEELLRGTCW
<EJC49_RS14795--*Mesorhizobium carbonis*--67.78 (SEQ ID NO:1827)
MAGYVYMSDKPRGVYIYIGVTSDLHDRVWEHRNEAHRGFTSKYSAKNLVWFELHPNIVLAIQREKSLKRYRREWKVNLIENLPTWLDLYERIDEIENVYQPH
PNTRMSSDYV
<EJV47_RS09340--*Hymenobacter gummosus*--67.06 (SEQ ID NO:1828)
MHCYVYVILTNQNHVTVVYVGVTNDLTRRVAEHREAHQKGFTRKYNLNLKIYFEEYTDAAKAIAREKQMKAGSRQKKLDLVNSTNPDWAEELLPV
<EJV47_RS23515--*Hymenobacter gummosus*--63.22 (SEQ ID NO:1829)
MSFFIYILTNPARTVLYIGVTNDLERRLYEHGQGLGEAGKFTGRYQCNNLIYFEISPDAVQAIAREKQLKNWSRAKKEWLIGTLNPNWEPIDPRTWNG
<EJV47_RS08795--*Hymenobacter gummosus*--64.38 (SEQ ID NO:1830)
MSYFVYILSNPARTVLYVGVTNNLERLCEHSLGLGDAGKFTGRYQCNNLIYFEISPDAVQAIAREKQLKNWSRAKKGVAQYAKSRLEND
<EJV47_RS27485--*Hymenobacter gummosus*--60.00 (SEQ ID NO:1831)
MKYHNYVYVITNPKKTVLYIGVTNDLVRRLDEHLNNAAGRPHTFAGRYWCHLLMYWELFADVHQAIAREKELKGWSRAKKEALIAAQNPEWRELNQEVGGF
<EJV47_RS27310--*Hymenobacter gummosus*--61.90 (SEQ ID NO:1832)
MYIYLLTNPARTVLYTGVTTNNLERQLYEHSQLGFAGRFTGRYQCNNLIYFEICPDATQAIQKEIKGWSRAKKEALIASLNPDWVPIDVRTCAG
<EJV47_RS13245--*Hymenobacter gummosus*--54.95 (SEQ ID NO:1833)
MYIYIVTNPACTVLYIGVTNNLRQRTEQHYANRQGRGTAGRYCYNLIYYESFPDATAIIVREKELKGWTRAKKEELIAMNPTWLSYNSV
<EFB14_RS03915--*Rhizobium fabae*--67.78 (SEQ ID NO:1834)
MSGFVYMSDKPRGVYITGVTSDLHGRIWEHRNAIQKGTQKRYAKNLVWFENHPNIVLAIQREKSLKRYLREWKIKLVEELNPTWIDLYERIDEIENIYRPH
PNTREWSYDYN
<EFQ99_RS05175--*Rhizobium vallis*--69.89 (SEQ ID NO:1835)
MAGYVYIVTNQKNGTLYIGVTSDLERRIYEHREGLTPGFAWKYGCTRLVWYEEHWDIGSAIQREKSLKRWYRQWKIDLVEAMNPDWKDLYLTLLW
<CWI84_RS02390--*Idiomarina tyrosinivorans*--60.26 (SEQ ID NO:1836)
MSDSNPTQWVYIILCCADDTLYTGVTDLERRLHQHNLVSGARYTRARRPVTLVWHCRADSRSAACQEQYRIKQLSRAEKQQLIEGKLS
<E0131_RS00170--*Muricauda beolgyonensis*--70.21 (SEQ ID NO:1837)
MSFYVYIISNKKLGTIYIGYTNDLKKRMYRHKHGVSGSKFASRYNLITLVYEEKFKFMPAIRREKQLKKWNRNWKVNLINDFNPDWKDLYTFD
<E0131_RS07085--*Muricauda beolgyonensis*--61.36 (SEQ ID NO:1838)
MKQSFYIMTSQYRTTFYIGVTSDLNRILEHTEGKGSNFTEKYNLTDLVYFEEFSEINQAIAREKQLKNWHDWKNINLIKEKNPKLETLDY
<E0131_RS01205--*Muricauda beolgyonensis*--61.36 (SEQ ID NO:1839)
MLSPYFVYITNQNKSIVLYIGVTNNIQDRLSQHQFDSQNAKSFAGRYNCIHLVYERFDSIEMAIAREKELKKWREKKDRLSINFPNPNWEFLNNEMV
<E0130_RS08175--*Muricauda aquimarina*--78.41 (SEQ ID NO:1840)
MTKNFHQYVYIILSNKRNGTLYTGMTNNIERRILEHKQGIKGFTRKYGLHMLVYFETHQSVNDAILREKRIKKWKREWKIEMIEKDNPFWDLDSNEWDL
<E0130_RS12945--*Muricauda aquimarina*--65.88 (SEQ ID NO:1841)
MRHSYVYIMTNKYRTTFYIGVTSNLKKRIAHEHYQGLGSKFTAKYQLTDLVYFEIFTDIIQAIAREKQLKNWRHDWKNVNIKEQNPTLETLDYG
<EJJ38_RS05600--*Aquabacter cavernae*--75.56 (SEQ ID NO:1842)
MGEAYVYILTNRPNGTLYVGVTNNIVRRVWEHREGVADGFTKRYGLKLLAYFERHEAISVAIQREKAMKHWPRACKVRLILDLNPTWQDLYPSIV
<EMQ25_RS01195--*Arsenicitalea aurantiaca*--67.42 (SEQ ID NO:1843)
MKGYVYILASGRNGTLYVGVTNPSRRWFHRHALTPGFTSRYGVHRLVWLEEHDLVVSVAIQREKTIKHYPRQWKINLIEAVNPDWSDLRGWLMDDAD
<X907_RS06480--*Glycocalis alkaliphilus*--82.22 (SEQ ID NO:1844)
MRERLFPAPVYILASQRNGTLYTGVTSDLVRRRIYEHREGMVPGFTRKYGVTLVWFEMHATMEHAIVRETRIKGWKRAWKLELIEAANPDWRDLYETLL
<X907_RS01755--*Glycocalis alkaliphilus*--73.33 (SEQ ID NO:1845)
MREREFQPAVYMLASRRNGTLYCGVTSDLVRRVFEHREGRASAFTRRYGVNQLVWFEMHSMMDHAIARETHIKGWKRQWKLELIEAMNPDWRDLYETLF
<X907_RS12055--*Glycocalis alkaliphilus*--65.59 (SEQ ID NO:1846)
MSAYIYILASKRNGTLYVGSCANPAERIAQHMRMGKSEFVWKYKFDLVYLEHCGTTHDAIARERQVKHWNRAWKIALIEQHNPWEWRDLYPELNNLMVF
<X907_RS09790--*Glycocalis alkaliphilus*--68.97 (SEQ ID NO:1847)
MDRAWVYILTNKPYGVLYTGMCNVDLVRRIWEHREGDVRGFTKRYNLKTLVFFEPHQDVTYAAERERRIKRWREWKIALIGKVNPDWRDLYPHLLG
<ENE74_RS06385--*Sphingobium algorifonticola*--81.61 (SEQ ID NO:1848)
MTRDPCVYIILASARHGTYIGVTSNLMQRLHQHREGRIEGFTSQHNKRLVHYEMADTMEAAIIREKQVKTWRDRWKIALIEKDNPFWEDLAVSLGFPPLSQA
<ENE74_RS01720--*Sphingobium algorifonticola*--74.12 (SEQ ID NO:1849)
MANKRNGTLYIGVTSNLIARAYQHRNGLIDGFTKDHDCLLVWFEYHDDLQDARLRELQMKKWKRAWKIELIERTNPRWHDLDLSLA

FIG. 11FFF

<ENE74_RS06640--*Sphingobium_algorifonticola*--70.45 (SEQ ID NO:1850)
MKEPCVNI LASQPYGTLYIGVTSNLI GRMLQHRSGAVLGFTAQYGVYRLVRFPCDTMEQAIMREKQLKRWHRQWKINLIESENPHWIDLAVGLGLPLVPSA
GRDGC
<EOD41_RS17905--*Mucilaginibacter_limnophilus*--61.36 (SEQ ID NO:1851)
MVFKQGYVYIMSNQYRTTFYIGVTSNLDSDRVWQHNGDGSVFVKYLLFDLVYEHFDCIEDAIAREKQLKNWHKQWKVDLIQSVNPEMDLKDITLNS
<EPB12_RS13010--*Vibrio_chagasii*--69.66 (SEQ ID NO:1852)
MKQPTVYIMTNRNTNSVLYGVGTGNLKRRIWLHKTGEVDGFTKKYNVHKL VYFEVDFEDFRAAIEREKQLKRWNRTWKDNLINKRNPSSRDLYDELT
<EC328_RS01425--*Gudongella_oleilytica*--69.66 (SEQ ID NO:1853)
MSFRPSSLCHPDLFPVIPTGVEGSHLLRRVIVYVWYIILTNNFNVLVYIGVTNDLQRRVFEHKNKLVEGFTQKYNVDKLVYESTEDVESAILREKQLKNWRK
EKKLVLIERMNPKWFDLSEDIY
<EQG53_RS03460--*Brevundimonas_diminuta*--68.13 (SEQ ID NO:1854)
MNERRPFIATYILASRPLGVLYVGMTSNLYERTRQHREDRIEGFAQRYGCKTLVWYERHGFVTEAIKREKALKRWNRTWKLELVEQTNPQWQDLYPSLCGLAP
DPRVKPEDDGEGE GSVRALLRRLLDDV
<EQG53_RS01715--*Brevundimonas_diminuta*--67.03 (SEQ ID NO:1855)
MATHRTLIATYIMASGRNGTLYAGSTSNLESRVWKRHRTAFPGFSDAYDCTRLVWFEQHALMLDAIRREKRRIKKWLRDWKLALIEKANPDWRDLAADWFSDPN
WTPPGAH
<EK403_RS06735--*Hansschlegelia_zhihuaiae*--70.21 (SEQ ID NO:1856)
MLSAFYIMASGPNGLVYGVTSDLVRRVFEHRSGAEGFTRRYGVKTLVWFETHDAIVAAIAREKALKHWTRARKVRLIRETNPDWRDLWDDIVR
<EK403_RS07600--*Hansschlegelia_zhihuaiae*--70.79 (SEQ ID NO:1857)
MGGWVYILASGPGGTLYGVTSDLVRRVFEHRTDAVPGFTKRYGVKDLVYFERHETIMAAIQREKNIKHWPREW KIDLIVSMNPYWRDLFDEIAC
<DSM03_RS16050--*Leeuwenhoekiella_aestuarii*--70.59 (SEQ ID NO:1858)
MEELFYTYILTNRNHTVYIGVTSDLRKRLKQHTQKSNEGFTGKYNVEKLIYFESTPYVENAIKREKQLKKWNPWAKENLINEMNPDWKDL SMMVA
<DSM02_RS08680--*Leeuwenhoekiella_polynae*--65.88 (SEQ ID NO:1859)
MQNSYVYIVTNTYRTTFYIGVTSDLQRMLDHLNRSKFTSKYNLHDLIYFEFTNIDQAILREKQLKNWRKSWKNLNIKEKNPTLQTLNIFE
<DSM02_RS10660--*Leeuwenhoekiella_polynae*--54.65 (SEQ ID NO:1860)
MKNYFYVILKCDNDSFYTGVTNNLEQRIEQHSSGKIKDSYTYKRRPVTLVWNE SFNEIGQAILWEKKIKAWSRRKKIALIEGNWDDLILFSKNYTQFGNKIEK
<EAV90_RS00990--*Bradyrhizobium_vignae*--77.53 (SEQ ID NO:1861)
MYILASKPGGTLYIGVTNNLVRRVYEHREGLADGFTKRYGVKALVYFEAHDITIAALQREKNMKHWSREW KIDLIVKSNPTWQDPYDEIAR
<X268_RS31690--*Bradyrhizobium_guangxiense*--77.53 (SEQ ID NO:1862)
MTYWVYILASKPGGTLYIGVTNNLIRRVYEHREGLADGFTKRYGVKTLVYFEAHDITIAALQREKNMKHWSREW KIDLIVKGNPAWQDLYDEIAR
<XH85_RS44020--*Bradyrhizobium_zhanjiangense*--59.52 (SEQ ID NO:1863)
MSTGSVYIILASRHGTLYGVGTNDVRRARFELRRSGRGSKFVHRCKVRLVHVEALATPQEAIAREKQLKFWRDWKIKLIEQENPDWNDRAGLI
<XH91_RS09170--*Bradyrhizobium_guangzhouense*--81.11 (SEQ ID NO:1864)
MAYYIYILASRRDGAIVGVTDNLIIRRIYEHQIKAVPGFTAKYNI TRLVWFEIYDDPI SAI SREKELKKWKRAWKTQ LVEKDNPNWNNDLYESICK
<XH91_RS04815--*Bradyrhizobium_guangzhouense*--73.03 (SEQ ID NO:1865)
MGARSYVYIILASKYGGTLYIGVTNDLIRVAEHKLQVLVESFTKEYGVVKL VYFEQFNDPENAIKREKRLKKWRDWKIALIEKDNPDWNDLYAEIAGPP
<XH91_RS04655--*Bradyrhizobium_guangzhouense*--64.71 (SEQ ID NO:1866)
MNAPAAVYVYILASRRYGTLYIGVCNDLGNRLTLHRSGRGSKFVKYGVTRLVYVETAYSPSEAIAREKALKKEWQRDWKIRLIEQENPDWDL SHLL
<GRAN_RS12315--*Granulicella_sibirica*--67.06 (SEQ ID NO:1867)
MASREYHFWYILSNRSHGTLYIGVTNDLIRSRAAQHREREP GPTHRTARYKVGVRVYVEFYQYVNAIAREKELKHWTQAQKIALIESVNPTWEELPS
<ESZ48_RS06465--*Gelidibacter_gilvus*--70.65 (SEQ ID NO:1868)
MKPGYVYILTNKNNTTLYGVGTAYLKERIIQHKEKHDKKSFTARYNLDKLVYEVYQMGDAIGREKKIKAGSRAKKITLIER TNSDWKDLFDEL
<ESZ48_RS09615--*Gelidibacter_gilvus*--69.47 (SEQ ID NO:1869)
MKPGFIYILTNKNNTTLYGVSSDIVERVQGHMKMLKNKNCFTARYNLDKLVYEVYQMGDAIGREKQLKGGSRADKIVLINRLNPEWRDLYDEVKIEYNGQ
<ESZ48_RS12685--*Gelidibacter_gilvus*--65.88 (SEQ ID NO:1870)
MFQEKNFVYIMINKNNSVIYVGFSGQVYERVQGHKKYSGKFANKYNLKNLVYVEKFEYPNEAIKREKQIKKWNREWKINLINDFNPSWDDL SHFLQP
<ESZ48_RS02025--*Gelidibacter_gilvus*--63.86 (SEQ ID NO:1871)
MTNKMNGVLYIGISDNI EERVKEHKLKVPKSF TAKYNCNDRLVYLEEYNNGLEASKREI SLKKWKREWKIKLIEEINPSWMDISLNWNLYNRLRE
<ESZ00_RS05205--*Acidipila_dinghuensis*--65.88 (SEQ ID NO:1872)
MSRCYFVYLIASRTRVLYCGVTGNLEKRTYEHQNGIFEGFSSKYHCTRLVWFEAFEDVRNAIDRETQIKRWRREKKLALITRMNPSWTDLSESWQR
<ESZ00_RS17315--*Acidipila_dinghuensis*--68.24 (SEQ ID NO:1873)
MASDRHYFVYLMASRTHVLYCGITNNLERRTYEHQHGLLEGFSSEYHCTRLVWFESFGDVRNAIGREKQIKRWRREKKLALIRKMNPCWVDLSEAWRN
<ESZ00_RS05390--*Acidipila_dinghuensis*--62.50 (SEQ ID NO:1874)
MDRHYFVYLMASRTHVLYCGVTGNLERRTHQHQRSLGLESSHYPCTRLVWFEFGDDVRSAGREKQIKRWRREKKLALIRKMNPCWVDLSEAWRPEK
<EO081_RS13550--*Sphingomonas_desiccabilis*--70.00 (SEQ ID NO:1875)
MDRLSCVCLLASGRNGTLYGVTSNLMGRHLQHRSGLFLPGFTNRYGVKRLVWYEAETIPDAIAREKQLKAWRRDWKVALIERENPFWEDRAIDLGFPLPQT
<EO081_RS13595--*Sphingomonas_desiccabilis*--69.57 (SEQ ID NO:1876)
MLNLFQHPCLPTRGATLLAMSKQPCVYILASRPLGTLYIGVTSDLVRLRHEHRS GDHAGFTSVYAVHRLVRYELCEDMPSAIGREKQLKRWHRNWKINLIESE
NPDWHD LAPGILPLAPRRGRGP
<CWB41_RS02800--*Methylovirgula_ligni*--72.09 (SEQ ID NO:1877)
MASQRNGTLYTGVTSDLARRAQHREGLISGFTKQYGCKLLVWYEHYERMDEAIAREKQVKAGSRAKKLAL IETVNPQWNDLYEELSR
<CWB41_RS02415--*Methylovirgula_ligni*--71.91 (SEQ ID NO:1878)
MAGLVPAIHAGHP EAMRGWVYIMTNKPNGTLYTGVTNDLSRRAYEHTGLYQGFTKQYGLIRLVYLER YDDIRAAIQRERMNKHWPRAWKVRLILESNPDWR
DLYDDLA
<DN53_RS05820--*Muricauda_olearia*--73.12 (SEQ ID NO:1879)
MPHNIHQYVYI LTNHKNGTLYIGMTNNLQRRVQEHKSGKLKGFTQRYGLHLVYVEFHYQYVSNAILREKQLKKWKR SWKIRLIEEQNP HVDLAKGWF
<DN53_RS10700--*Muricauda_olearia*--65.67 (SEQ ID NO:1880)
MQQSFYVIMTNKYRTSFYIGVTSDLYKRVIHESKGIGSKHTKRYNLTLVYFETFS DINNAIAREKQLKN
<DN53_RS07640--*Muricauda_olearia*--54.65 (SEQ ID NO:1881)
MKLSYVYILKCS DGTGYTGVTSNLEKRFQEHSSGKHPRSYTFSRPTELVYFEQFTNI AFQIREKQVKKWSRAKKEALIQERQNTLPNLAKKKFQ

FIG. 11GGG

<EWH08_RS09045--*Sphingobium francense*--78.65 (SEQ ID NO:1882)
MKREISPTVYILASRRNGTLTGVTSDLVKRLYEHRNGLIEGFTSDYGVKRLVWFEMHDAIDGAILREKRIKKWNRQWKIELIEAGNPEWDDLALAFGEPLR
ERAGDGFPLSRE
<EWH08_RS00275--*Sphingobium francense*--71.91 (SEQ ID NO:1883)
MSKPGYVYIMASRRNGTLTLGVTSDLVKRAWQHRNGLGSEFSAEYRCHFLVWYEAYDDLQEARLRELRMKKWKRDWKLRLIETANPGWRDLFEDVAQ
<EWH08_RS11540--*Sphingobium francense*--65.52 (SEQ ID NO:1884)
MTNKP HGVLYGITSSLAARVDQHRRGVGSFAFCRKYNLRLVLAEEHDRVDAIAREKALKAWKRDWKLRLIEESNPEWRDLFDLIV
<EWH08_RS09600--*Sphingobium francense*--61.63 (SEQ ID NO:1885)
MERKEKGGWVYIMSDRYRGTIYVGVTSDLPARIHQHRSGNGSDFCARYGLNRLVWAERSDDITRCIEHEKRVKRWREWKFDLIERANPDWQDLFHLA
<EWH12_RS05930--*Sphingobium cupriresistens*--60.00 (SEQ ID NO:1886)
MKFTVTPAKAGVSGGALRWDMARGGYTIMTNKPRGVLYIGVTADIAARAEQHRNGTGSFAFKTYGLTRLVLVEPHDDITLAIAREKALKAWKREWKIRLVE
ESNPGWRDMSDVIF
<ERX46_RS11065--*Brumimicrobium glaciale*--69.41 (SEQ ID NO:1887)
MKTTKLYFTYMLSNYSNKLIVGSSSNLSQVRHRHKIGYFEGHTKKYNINRLIYFETYTSIDEAIKREQQIKSYNRSKKNKLINKVNPQWRELEAPLNFF
<EWM57_RS19440--*Hymenobacter persicinus*--57.95 (SEQ ID NO:1888)
MNPFRYSLYIITNPTKTVLYTGVTNDLVTRLQHQHCQNRGSAAFTAGTYHCYQLMYFEHYQQVDQAIAREKEIKGWTRAKKDALIVQLNPDWQVLDPLWEDAP
Q
<EWM57_RS20445--*Hymenobacter persicinus*--58.54 (SEQ ID NO:1889)
MNPFRYSLYIITNPTKTVLYTGVTNDLVTRLQHQHCQNRGSAAFTAGTYHCYQLMYFEHYQQVDQAIAREKEIKGWTRAKKDALIVQLNPD
<EWM62_RS08120--*Mucilaginibacter terrigena*--60.44 (SEQ ID NO:1890)
MKIHQYVYVYIITNANNNVLYTGVTNDLADRVLQHKDKMNKGSAQYQCGKLVYEEFQWIQDAIAREKQIKAGSRKKKIALIVNENPDNDLSIGWYD
<EWM62_RS18710--*Mucilaginibacter terrigena*--59.78 (SEQ ID NO:1891)
MKQHNYTYTILNTNKTLYTGVTNDIEKRLYEHYFGLNQESSFTSKYKCFYLLWYERYQYVNHAIEREKEIKGWTRAKKIYLIQENPNWTLNNDIMEWPP
INTHTEYGGLTVGD
<ELE36_RS16440--*Pseudolysobacter antarcticus*--73.12 (SEQ ID NO:1892)
MKNPCVYVLASARYGTLIGVTSDLIKRICKHRNDAAEGFTKKYCVHALVWFEQHEMESAISREKALKSWRRDWKLNLIQSQNPWRDLYLDIIL
<DEJ70_RS02655--*Wolbachia pipientis* wAlbB--66.67 (SEQ ID NO:1893)
MFLAFSLTSNPQLGLYIGITSNLTKRIWEHKSATISGFTSKYNVCKLVYEEFQENLAISREKLLKSWQRKWKIDLIEKTNQEWKDLDEIVSG
<EWE75_RS21380--*Sphingomonas populi*--65.12 (SEQ ID NO:1894)
MQRGGWYTLMAKPHGMLYIGVTAHLAARVDQHRRDVGSAYCR <RYGIKTLVYTERFDTIEEAI SREKALKAWKRAWKQDLIDSANPEWRDLYDTLA
<EWE75_RS11420--*Sphingomonas populi*--62.22 (SEQ ID NO:1895)
MGGFTYIMTNKPFGLIYIGVTSDIARIQAHRKRGSAFAAKWDCITRLVLI EHHDRIEDALIREKRLKDWNRWLKCRILAEANPNWDDLWETINA
<EVS81_RS03490--*Leucobacter triazinivorans*--72.94 (SEQ ID NO:1896)
MNRHAYVYLLASGNGTLVGVTSDLARRVYEHNRHLVPGFTDRYDVTSVLVWYVAGENIVDAIALEKKIKNRSRRWKIDLIERENPSWEDLAAGWYE
<CWO91_RS09105--*Bradyrhizobium genosp.* SA-3--65.88 (SEQ ID NO:1897)
MNAPAAYYVYILASRRYIIGICNDLNNRLTLHRSGRGSKFVTKYGVTRLVYVETYSFSEAREKALKEWCRDWKIRLIEQENPDWRDLSHLI
<EV197_RS06765--*Aquimarina brevivita*--64.37 (SEQ ID NO:1898)
MINDETLYFYVYILNKNHVTLYVGVTSDLKNRLKQHHNGSKGSFTKRYKVDKLVYFESTKYINNAIKREKQIKKWNREWKINLINDINSWDEDISYRLG
<BDD14_RS23635--*Edaphobacter modestus*--68.24 (SEQ ID NO:1899)
MRCYAGLDERSHVLHLHCASRSHNFYVGMTSEIELRVEQHKQGRFEGYSKKYCNRLVWFERHPYVQDAIAREKQKLGWSRAKKI GLIERENPAWVDLSEEWG
QPVKMIGEV
<BDD14_RS24055--*Edaphobacter modestus*--66.30 (SEQ ID NO:1900)
MPPDQAYVYILASRRRLYIGVTRLADRIREHKTSIDPEAFTARYNIRSLVYEFYESISRAIARETELKSWLRRAKKIQLIVVNNPTWRDLSLDWGKPIATF
SESQRRPEPMF
<EXU30_RS03630--*Shewanella maritima*--80.00 (SEQ ID NO:1901)
MQGLQVPIGVYVYMMANKKDGSIYIGVTSNLINRVYQHKAKLVKSHTQKYNINLLVYEVHDEIGSAIEREKKLKRWRSSWKQCLIEKQNPDWLDLYQEIL
<EYD46_RS03800--*Hyunsoonleella pacifica*--74.73 (SEQ ID NO:1902)
MHKNIHQYLYILSNKKNGTLYIGVTNDLERRMFEHKNKLVEGFTSKYGLDKLMEFESYQYVNDAIKREKNMKKWKRWKINLIEENPNWDDLSDADWTYLIE
<EYD46_RS15500--*Hyunsoonleella pacifica*--70.59 (SEQ ID NO:1903)
MFNDETYIYIITNKKDGLIYIGMTGNIKQRMTOYKQKRYTRGFANKYGLDILVYEEKFESIIEAIKREKQIKKWNRWKINLINDLNPWNDSLHFFI
<EYD46_RS03435--*Hyunsoonleella pacifica*--67.39 (SEQ ID NO:1904)
MKPGYIYILTNKNNTLYGVGTAYLKERIIQHKEKRYKKSFSARYNLNKLIVYESSFQMIGEAIGREKQLKAGSRAKKIALIESLNPDWEDLFDLSL
<EYD46_RS03705--*Hyunsoonleella pacifica*--63.95 (SEQ ID NO:1905)
MAKRAYHNYVYILNRPNGTLYIGVTGGLDRMERHIAEGNSFSAKYKLRVLYFEFFQFINDVIKREKQLKNWHRQWKINLIEKENPNWDDLWDPIRFN
<EYD46_RS07930--*Hyunsoonleella pacifica*--58.62 (SEQ ID NO:1906)
MKTIGNHNYVYIITNKSCTVLYTGMTNNLKERLYFHNNPTAFSKAFTTKYRCFYLIYHEHFFEI KDAITREKQIKGWSRLKKENLINSFNPSWSFLNDEI
<EYS08_RS20020--*Pedobacter kyonggii*--61.36 (SEQ ID NO:1907)
MERGGCVYIMANSFNTVFYTGVTSDLQGRVWQHKNNEFPKSFTSKYKCYKLVYTFPPNIEEAI DEEKRIKRGSRQKQKIDLVASENPNWVDLYDELDS
<EYV96_RS05830--*Dyella terrae*--74.44 (SEQ ID NO:1908)
MSKRKI TPCVYMLASRPQGTLYIGVTGNLLRRVWEHRTYVADGFTKRYNIHQLVWFEEHPIMMSAIKRETNLKDWRRAWKFQLIQAENPHWEDLYPRLLDDSR
<EYW49_RS15080--*Siculibacillus lacustris*--78.65 (SEQ ID NO:1909)
MYFVYILASRRNGTLYIGVTNDLVRRIGEHCSGAVPGFTRRYGVTRLVYFEIYDRIDEAITREKQLKGWNRWTKIRLIEASNPEWRDVGGDIPWS
<EZL74_RS02055--*Flavobacterium silvisoli*--67.82 (SEQ ID NO:1910)
MLCLHFSQKKNGLYIGVTNSIERIEEHKNKLSNCFATAKYDVKLLVYETFTQYINDAIDREKQLKKWNRQWKINLIEAENKDWKDLSENWIN
<EZL74_RS10070--*Flavobacterium silvisoli*--60.98 (SEQ ID NO:1911)
MKPSYVYILKCSGDGTYTGYFDSNLTRIYRHDGTGYFPDCYTYKRRPVSILVYCEFTNINLAFEKEKQIKKWSKAKKEALIDGHYDALINLAKKNFK
<EOD97_RS06655--*Oricola cellulositytica*--67.02 (SEQ ID NO:1912)
MGGYVYMLASKKGGTIYIGVTSLSQRIYDHKQRINRKSFTSTYGAVALVWYEDYDDITDAIQREKSLKWRWRQWKIELIEAMNPEWDELYKGIGW
<EOE05_RS09735--*Roseitalea porphyridii*--69.51 (SEQ ID NO:1913)
MASKRNGTIYTGVTSDLQGRAQEHRDGLTPGFSRRYGCTTLVWYERHDNIVEAIAAREKELKKWRRAWKLYLIEGFNPDWSDLLLESCYARDNPSDAIRDRGEKA
REGGLDSGFRCAAPE

FIG. 11HHH

<E0E05_RS13635--Roseitalea_porphyriddii--66.67 (SEQ ID NO:1914)
MGGYVYMLASKKGGTIYIGVTGDLSSRIHEHRQREAGFTARYGAIRLVVYENHDDIRDALQREKSLKRWRRQWKIDLIESINPEWNDLYRGIGW
<E0E05_RS01840--Roseitalea_porphyriddii--65.17 (SEQ ID NO:1915)
MTGFVYILANKRNGTLYAGVTSDEGRMHDHRNGLTPGFATKYGCDRLVWLERRPNIVLAIRREKSLKDYKRAWKLDLIAKLNPWDDLTDSDIHVIDNPYKPK
AGSRWSELYE
<EZM97_RS33835--Dyella_soli--79.57 (SEQ ID NO:1916)
MRIRQPCVYILASRRNGSLYTGVTSNLVARVWQHONGVVEGFSHRYRVHRLVWYEVHETMESAIRREKAIKKWMRAWKIELVERSNPYWNDLFDGIHS
<EZM97_RS05040--Dyella_soli--74.44 (SEQ ID NO:1917)
MSKRKITPCVYMLASRPQGTLYIGVTGNLLRRVWEHRTYVADGFTKRYNIHQVLVWFEEHPIMMSAIKRETNLKDWRRAWKFQLIQAENPHWEDLYPRLLDDSR
<EV216_RS15975--Rhodovulum_steppense--69.41 (SEQ ID NO:1918)
MVHWVYIMASRPGGAIYIGRTHDLRQRVEAHRAGLSAHTARYNIRTLVWFEAHAEFEPSSLRRERQLKRWRAWKQDLITESNPWKMDMTHAIP
<EV200_RS06685--Pedobacter_psychrotolerans--64.84 (SEQ ID NO:1919)
MERGGTYIITNFTHTTFTYTGVTSDLLSRMKHEKHEKHSNSFSARYNVNKLIVYFCFYSTIEDATAAEKRIKGGSRLLKKIKLIENLNPEWKDLWDEVKEW
<EV200_RS06685--Pedobacter_psychrotolerans--60.92 (SEQ ID NO:1920)
MNHNYFVYIITNKNKTIVYTGVTNDLYTRLKQHIENINNRYAFTKKYNCHFLIFFERYQYIQAIEREKEIKGWSRAKKNALIALENPDWKFLEEIFSE
<EV196_RS13945--Mariniflexile_fucanivorans--64.71 (SEQ ID NO:1921)
MKKSYVYFLTNKNKTIVYIGVTSNLLKRVYQHKTKSYKGFTQKYNCDILVYFEEFEDINQAI SREKQLKSGNRKRKEALINLENPNWSDLSGWLIFYFD
<EV196_RS15575--Mariniflexile_fucanivorans--62.22 (SEQ ID NO:1922)
MKRAYHNYWVYILTNKARGTLYIGVSGGIDDRMERHVSNGSKFTKYYQLKHLVYFEEFYIDDAIKREKQLKNWHREWKINLIESQNPPEWKNLWTPVVTCLR
P
<EV196_RS04945--Mariniflexile_fucanivorans--56.82 (SEQ ID NO:1923)
MQHNYFVYITANKYRTTLYIGFTNNIQNRLSQHYFDSQNAKKSFAKGYACIYLVYETFEFFPKQGIAREKELKGWRREKKNKLI SDFNPKWEDLNMQIF
<EV196_RS13230--Mariniflexile_fucanivorans--55.29 (SEQ ID NO:1924)
MLSHYCYILSDKNRKFVGYTCDIKRMTQHKSGNGSNFTKKYFVFDLIYFEIFNSSILGRNREKQLKNWHKEWKNLVRMSNPCLKTEI
<EV680_RS06270--Uruburuella_suis--86.17 (SEQ ID NO:1925)
MQPAVYIMASQRNGTLYIGVTSNLIQRVYQHRNHQVAGFTDRYQISLLVWYELHENMESAI GREKQLKKWQRKWKLDLIEQHNSAWRDLWPDIIIG
<EV665_RS24845--Shinella_granuli--72.04 (SEQ ID NO:1926)
MTGYVYIVTNHKNGLTYIGVTADLERRIHEHREKITPGFSSKHDCVRLVWYEEHDSIGTAIQREKSLKRWYQWKVELIEAMNPDRDLYLEFW
<EDD52_RS13985--Primorskyibacter_sedentarius--62.92 (SEQ ID NO:1927)
MAHFVYIMASRPGCALYTGRTOQLGARVEAHTGMCVHTAKYNIIRRLVWFERHDCYETSLRRERAIKWRDRDWKIALITQSNPNWHDITSQISG
<EV148_RS17615--Dokdonella_fugitiva--72.83 (SEQ ID NO:1928)
MAIRQPAVYILANRPRGTLYVGTGWLRQRIHEHREGLIDGFTKRHGKALVWFEWHDNFARRETQLKPKWRAWKLELIETANPQWRDLWADIAW
<EDC25_RS10810--Pseudofulvimonas_gallinarii--84.95 (SEQ ID NO:1929)
MAMEKQPAVYILASDRNGTLYIGVTSDLVKRIWQHRQHAVEGFTARYGVTRLVWYELHGMVSAISREKQVKKWRRSWKLRLEELNPDWTDLWGEICGFETG
APRSD
<EV195_RS09535--Tenacibaculum_skagerrakense--65.91 (SEQ ID NO:1930)
MKKSYVYFITNKNNSVIYIGVTNLLKRVYEHKTKTRKGFTSKYNCEKLVYFEEYNSIKEAIIREKQLKAGNRKRKEQLINTLNPDWDLAVDWVFNIN
<EYY60_RS08250--Flavobacterium_zhairuonensis--66.67 (SEQ ID NO:1931)
MFNPQIGFHSYVYIITNEHRTSFYIGVTNNLKQRLAKHKENIDLNISSPAAYKNIQFLVYEEKFAWIQEAIAREKELKKWRDKKIELIRSFNETFEFLNPH
FE
<EYY60_RS01210--Flavobacterium_zhairuonensis--62.20 (SEQ ID NO:1932)
MKLHEGYHTYIYIITNPKFTVYTGVTNNLRVRLSQHKENITGENKTFASKYKVEFLLYEKFETWIQEAI GREKEIKGWRDRKIELIKTINPDLDLNYIF
S
<EZJ43_RS01920--Pedobacter_changchengzhani--63.74 (SEQ ID NO:1933)
MERGGCVYIILTNAHTVFYIGVTSELYFRIKEHKEKYPNFSFAKYNCKNLVYEEQFSSIQEAILKEKQLKNWRRGWKINLINELNPNWDDLFLTLE
<E2A64_RS06350--Pseudohoeftlea_suaedae--71.11 (SEQ ID NO:1934)
MTGYVYIVTNHKNGLTYIGVTSDLARRAYEHREGLTPGFSSKYGCSRLVWYEEHFDIRDAIQRETSKRWYREWKEILIEQMNPDRDLYQELW
<E2A64_RS07995--Pseudohoeftlea_suaedae--64.71 (SEQ ID NO:1935)
MVRFVYIMASRRGGTIHTGVTDLPRLVDEHQSGDSVFTGKYAVQRLVWHEEHFDIRDAIQLERTIKGWPRQWKVELIERENPDWFDLMRGIGWYPARGCVP
<E2F50_RS13265--Rhizobium_deserti--71.76 (SEQ ID NO:1936)
MRKGYVYILASRRNGTLYTGVTSDLPGRLYEHQNELTPGFTSRYGVKTLVWFEEHDLIVDAIRREKTIKKWPRAWKLNLIETMNPWSDIAHYLHGL
<E2I14_RS14440--Sapientia_aquatica--84.78 (SEQ ID NO:1937)
MKQPAVYLLASKRNGTLYVGVTNLIQRIWQHREGSVGGFSSDYSVKALVWYEMHSTMESAIRREKALKKWNRAWKCLKLIEERNPEWIDLWPEINGVVTYLPK
LFSPPPREN
<E2I14_RS10095--Sapientia_aquatica--83.70 (SEQ ID NO:1938)
MKQPAVYLLASKRNGTLYVGVTNLIQRIWQRREGSVGGFSSDYSVKALVWYEMHSTMESAIRREKALKKWNRAWKCLKLIEERNPEWIDLWSEINGEVYLPK
LFSPPPREN
<E2F49_RS04845--Luteimonas_terrae--76.09 (SEQ ID NO:1939)
MRDRLPCTYILASAPRGTYIYIGVTSDLVQRIWQHRNHLAPGFTSQHRTHHLVWFETHDTMEAAITREKRLKAWQRLWKIQLVESTNPHWDLYPDLV
<E2I05_RS14375--Parashewanella_spongiae--81.72 (SEQ ID NO:1940)
MEKQPCVYILASKRNGTLYIGVTNLIKRIWEHKNNIVEGFTQKYSVHRLVYFEQHDNMEASAITREKQLKAWKRQWKINLIEETNPWNDLWKSIL
<E2000_RS11370--Qipengyuania_sediminis--76.34 (SEQ ID NO:1941)
MKRDFQFAGYMVANKRNGTIYIGVTSNLPQRIAQHRDGIIDGFTKRHGCKMLVWFELHATMKFAILREKQLKGGSRSKKLALIEGANPRWDLFAEICA
<E2000_RS02460--Qipengyuania_sediminis--64.29 (SEQ ID NO:1942)
MEREKRGWVYIMSNRYRGTYLFGVTSDLAARIHQHRTGRGSEFCRENTRLVWAGFIEPIDAAIAHEKRLKRWPRGWKFALVEEANPEWRDRFDELAYS
<EV664_RS10305--Stakelama_pacifica--75.90 (SEQ ID NO:1943)
MANKRNGTLYLGVTSDLARVHQHRLIDGFSKQYGCTLLVWFEAHDDLQEARRELQIKKWKRDWKLELIEAENPQWKDLYATLF
<EV664_RS10710--Stakelama_pacifica--61.36 (SEQ ID NO:1944)
MGGWVYIMTNKREGVLYIGVTSDLAARVLQHRAGKGSFACRRYGLNRLAYVERHDDIASAIARERALKAWKRAWKIDLVESVNPDWLDLFDHLA

FIG. 11III

<BC659_RS03855--Sediminibacterium_goheungense--68.18 (SEQ ID NO:1945)
MTKLGVVYIITNSSKTMVYGVTSDDLTKRIRQHQQKVYSKSYTAKYNIHLIIYYETALVVDALKREKEIKGWSRSKKEKLIKVNPNWEILCI
<DFS08_RS10790--Sphingobium_barthaii--63.44 (SEQ ID NO:1946)
MHPTACIRASHENGTLCIEGTSNRLQRRHQHRSGLIEGFTKDYGVHRLVWFENGSGMEEAIRRGKQIKKWNRRQGTIDLIERNNAEWQDLALDLGFEPLESRI
G
<DFS08_RS12725--Sphingobium_barthaii--67.42 (SEQ ID NO:1947)
MGKPGYVYIMASGRNATLMVGTSDLVRRWQHRRNGFGSEFSAEHRCRFLVWYEAHDDLQEARLRELRMKKWKRDWKLHLIEATNPDPWRDLFDDIAR
<EDD54_RS19470--Otharobacter_diazotrophicus--73.33 (SEQ ID NO:1948)
MKSGETVYIVTNRPNGLIYTGVTSDLVGRIFEHREGLIDGFTKRYGCTRLVWYEDHEEIGDAIWRERRIKTWNRAWKVRILGLNPWEWRDLYAEVAGMVEGR
<DFQ04_RS05810--Algoriphagus_boseongensis--68.13 (SEQ ID NO:1949)
MEKGGAIYIMTNFKNTVLYTGVTNDLIRRVFEHKNGLNPNFSFVKYKLTCLVFFESFHSIEEAIAREKQIKGGSRSKKEKLINSINPEWKDLWEEIQSF
<DFQ04_RS06780--Algoriphagus_boseongensis--57.95 (SEQ ID NO:1950)
MYFTYIILTNFSKTVLYIGVTNNLARRLAEHKEDAFNRKKSFCGRYKVFYLLYFEFFQWIEAIAREKELKNWRREKKERLISSQNPWNDFLDI
<DFQ07_RS06960--Tenacibaculum_caenipelagi--76.14 (SEQ ID NO:1951)
MHKTNHQYYIITNKNDGVLYIGVTNNLERRIFEHKNKLIKGFSSKYNLDKLIYFEEHQFIEEAIKREKNMKKWKRAWKINLVKENPNWDDLAKNWN
<DFQ07_RS05220--Tenacibaculum_caenipelagi--59.55 (SEQ ID NO:1952)
MLDKYVYIITNKNTTLYIGVTSNQLKRLSQHYFDSEHAKKSFAKYNCYLIYYETFTETINEAIYREKELKKWRREKKETLINNFNPEWRFLNNEVF
<ATL17_RS03515--Maritalea_mobilis--80.65 (SEQ ID NO:1953)
MSAFVYIMASQRNGTLYVGVTDLVRRVWEHREGQVDGFTKQHDVLDKLVHFEQHDFNSAI TREKALKKWNRRKWLKLIKGNPNWNLDYDEVSR
<ATL17_RS09545--Maritalea_mobilis--75.53 (SEQ ID NO:1954)
MGGWVYIMSNKMGNTLYVGVTSDLVRRVYEHREGLVDGFTKANELKRLVYFEAHDEIGAAIQREKGLKRWSRAWKVELIQKLNPFWDLDLFDQICG
<ATL17_RS02195--Maritalea_mobilis--72.04 (SEQ ID NO:1955)
MGGWVYIMTNKNGTLYVGVTDLVRRVWEHREGQVDGFTKQHDVLDKLVHFEQHDFNSAI TREKALKKWNRRKWLKLIKGNPNWNLDYDEVSR
<A8950_RS08880--Dongia_mobilis--73.17 (SEQ ID NO:1956)
MTDGPNGTLYVGVTDLARRVSRHAGEIVGFTKRYDLKQLVYFEPHEDINAAIHREKRLKKWPRAWKVRILTRENPEWRDLYHDLFR
<CLV99_RS00270--Sphingobacterium_yanglingense--65.91 (SEQ ID NO:1957)
MERGGVYIMLTNINRTVLYIGVTSDLVSRVLEHREKHYLNSFTSKYNSCLCVYYESFPAIEEAI DREKQLKKWSRAKKEMLINKLNPNQWDLWEVVKW
<CLV99_RS09880--Sphingobacterium_yanglingense--65.91 (SEQ ID NO:1958)
MERGGCVYIITNKNTTLYIGVTSNQLSRVYEHNLQKYKHSFSAKYNLVNLVYKYFPTIDEAIAEKRKVKAGNRQNKELINLFPQWEDLYGKEVKHW
<C8D91_RS07030--Marinicella_litoralis--83.70 (SEQ ID NO:1959)
MKNPAVYILTNQRNGTLYIGVTSDLIQRIYQHKSNAVEGFSKKYQTHMLVYFEQHPMTESALLREKQMKWKRAWKINLIEKLNPFWKDLYQDII
<C8D91_RS07035--Marinicella_litoralis--82.61 (SEQ ID NO:1960)
MKNPAVYVLTNQRNGTLYIGVTSDLIQRIYQHKSNAVEGFSKKYQTHMLVYFEQHLTMSALLREKQMKWKRAWKINLIEKLNPFWKDLYQDII
<C8D80_RS07950--Flavobacterium_cheniae--66.28 (SEQ ID NO:1961)
MIEFTEGVYINFYVYILTNKAKTVLYTGVTNNLKRRLQEHNEKLNPFQSTAKYNLNLYIYYEKYQWAIAREKEIKLLRRDLKLELIKEFNPNMEFLNNQFS
<C8D80_RS03360--Flavobacterium_cheniae--59.21 (SEQ ID NO:1962)
MQISYVYILKCADDSYTGVTSNQLNQRMYQHDSAYFPDCYTASRRPVLFVYAEFTNIGFAIDKEKQIKKWSRAKKEALINGDFDELNLAKKRFDK
<DFR44_RS00730--Hydromonas_duriensis--69.41 (SEQ ID NO:1963)
MAYVYLMSTNHTVLYIGVTNDLRRVAEHKAKIQKGFTEKYNVDKLVYFESTNAMIDAI TREKQLKNWKRAWKNDLIETNPNWDLADSIGVTQEYIDAVK
THYASYGTNPASRR
<EV668_RS13010--Enterovirga_rhinocerotis--77.78 (SEQ ID NO:1964)
MAFYVYILASKRNGTLYIGSTDDLRRVWQHKTKELRGFTSKYDVKLLVWYVDFDTREAAAFVRRQMKEWQRRWKLELIEKNPNPDWEDLYETLNR
<EV668_RS02030--Enterovirga_rhinocerotis--77.01 (SEQ ID NO:1965)
MYHVYITASRRHGTYLGVTSNLSQRIWQHKNKVPFGFSKYGVDRLVWYEAAYERIDEAISREKALKWRRDWKRLIEEMNPDWQDLYLLNLR
<EV668_RS09715--Enterovirga_rhinocerotis--60.00 (SEQ ID NO:1966)
MPFWASMLRCTDGSFHVGTSDLDTRIGQHGGEMDGTARRRPVVLVWAEFQDILDIAAERRIKGWSRAKKEALIGDWERIRLLGSRARRDPPA
<B0I21_RS18725--Sphingobacterium_paludis--61.36 (SEQ ID NO:1967)
MVFGGTVYIITNFNTTLYTGVTSDLIQIRMYEHINHKHRSFSKRYKLYKLVYFKFSFSSIEEAIHEEKRIKAGSRAKKEALINKMNPWEDLYQREVKYW
<CLV90_RS03440--Maribacter_spongiicola--57.50 (SEQ ID NO:1968)
MKGYVYITTNKNTVLYVGATDDIKRMNEHKNKIYPNAFTKKYNCIDLIVYFEFENVEDAFKKELRLKAGNRKKKEELINSINPE
<DFQ03_RS01250--Maribacter_caenipelagi--69.23 (SEQ ID NO:1969)
MKPGYVYILTNKNTTLYVGVTVNLGERIQHKNKHDKKSFTARYNVDKLVYFEAFQEIIGDAIAREKQIKAGNRKAKIALIEKKNPDWEDLT
<BXY82_RS06550--Gelidibacter_sediminis--69.23 (SEQ ID NO:1970)
MSWSVYIITNKPNGVLYIGETKQLKRIYQHINKAHPGFTSARYNLDKLVYVEHHKSEGDAKKRELQLKKWNRAWKIELIEKSNPEWKDLYQNL
<LY04_RS06990--Oceanimonas_baumannii--83.87 (SEQ ID NO:1971)
MHKQPAVYMLTNQRNGTLYIGVTSNQLVQRIWQHRRQRAVEGFSARYGLHWLVYFELHDDMYSAITREKQLKKWKRAWKIRLIEVQNPGRDLWEEII
<EDC55_RS07000--Allofrancisella_inopinata--74.42 (SEQ ID NO:1972)
MKKGYTYILTNKNNAVLYIGVTSDLIKRVYEHKQKFVDGFSKKYNLDKLVYFETYEDIEDAIVREKQLKKWNRIWKERINSENPSTWDLYEGLLT
<A8975_RS01080--Meridianimaribacter_flavus--69.88 (SEQ ID NO:1973)
MSNRANGVLYIGVTDNIEERVKEHKLKVPKSFSAKYNTMLVYFEFENGKDAEKRRERQFKKWKRNWKVELIEEMNPWNIDISANWNLNFKLRE
<EDC23_RS06325--Thiohalophilus_thiocyanatocydans--83.87 (SEQ ID NO:1974)
MEKQPCVYLLASKRYGTLYTGVTNNLKRVEHKTTHSVGFTKKYDVTVLVWYEVHESMESAIKREKSLKNWNRAWKIRLIEETNSQWRDLYKQLL
<C8D99_RS10675--Aminivibrio_pyruvaticus--68.24 (SEQ ID NO:1975)
MERSYVYLLANANRVLYTGVTNNLKRRLFEHKGRLVEGFTKKYRVHKLVEETNDVNTAIAAREKEIKGWIRAKKNALVEKENSQWSDLSKGWEVDSSELP
K
<DFQ06_RS18195--Algibacter_wandonensis--65.88 (SEQ ID NO:1976)
MDVHYVYILTNKNHTVLYVGRTKQLKVRNLQHKNNSSKSFSGKYNVDKLVYFETTKYVNNSIKRRERQIKKWNRAWKINLINSINPDWDLSEYI
<E2R65_RS02880--Mucilaginibacter_phyllosphaerae--62.07 (SEQ ID NO:1977)
MWNYNVYVYIITNHDKTLYIGVTNDLNRRLYEHREHKGTSFSAKYCYCNLYVEYEHFSKIEYAIEREKEIKKWRREKKNALIKATNPDKFLNNEIDK
<E2R65_RS03025--Mucilaginibacter_phyllosphaerae--59.78 (SEQ ID NO:1978)
MWNYNVYVYIITNHDKTLYIGVTNDLNRRLFEHRENKGTGKSFTAKYCYCNLYVHYEHFFNIEHAIEREKEIKKWRREKKDALINSTNPDPWRFLNNEVG

FIG. 11JJJ

<E2488_RS01565--Gramella_jeungdoensis--70.83 (SEQ ID NO:1979)
MHQYYVYILASKIRGTLTYIGITNNLQRRVYEHKTTGGIEGFTQKYGVHKLMYFETFAQSIDDAINREKNLKKWKRDKWIKLIEEENKKWLDLSNDWYNDILID
<E2488_RS01640--Gramella_jeungdoensis--67.39 (SEQ ID NO:1980)
MKESYIYILTNKNKTVTYIGVTSNLERITQHKNNTGCSFTTKYNNVHFLVYIEMFTSIEQAIAREKQLKKWNSWKWELIKKVNPKLLDLYSDL
<E2488_RS06940--Gramella_jeungdoensis--67.05 (SEQ ID NO:1981)
MKKSYVYFMTNKNNTVIYIVVTSNLLKRI FQHKHKIYKGFSTKFNADKLVEEFEDVT LAISREKQLKSGNKRKRKEQLIFEMNPNWIDLSEDWVLNE
<E2488_RS06825--Gramella_jeungdoensis--62.50 (SEQ ID NO:1982)
MIEFNEGYHTYVYVYIITNKYRSTYYIGVTNNLKRRLNEHIQNIENGTKTFASKYQLKYLVEYEFKFSWVQQAIAREKELKGWRREKKLNLRSSNNKFEFLNAH
FIESNT
<E2488_RS04070--Gramella_jeungdoensis--53.49 (SEQ ID NO:1983)
MKKSFYVILKCSNDNTYYTGVTSNLENLRFQHSSGFYKNCYTFKRRPVTMI FYAEFTDINLAIEKEKQLKKWSRVKKEALINSYDDLVNLAKKKFE
<E2P86_RS11305--Sphingobacterium_psychroaquaticum--69.23 (SEQ ID NO:1984)
MERGGGIYIMTDINRGTLTYIGVTSDDLFRVREHRTKAYQKSFTSRYNLTVCVYEFYSSIEEAINREKELKGWRREKKKNLINESNPSWNDLWQDIENW
<E2R66_RS20515--Mucilaginibacter_psychrotolerans--64.29 (SEQ ID NO:1985)
MSNKYNSVLYTGVTSELISRIWDHKNKTYPNSTAKYNCTKL VYCYFPHIEEAI AAEKTIKGSSEYKRQLVKSMNPWDWKNLYDGLLEE
<E2R66_RS22740--Mucilaginibacter_psychrotolerans--61.70 (SEQ ID NO:1986)
MRQYNFYTYILTNYTKTVLYTGVTNDLERRLYEHYLGKGQSDSFTAKYKCFYLVLYERHQYINHAIEREKEIKGWLRAKKIKLIEEENRNWNFLNSEIMDWPP
VFTG
<E2R66_RS17600--Mucilaginibacter_psychrotolerans--59.77 (SEQ ID NO:1987)
MWNYNFYVYITNDRKSVLYVGVNTDLTRRLDEHRHNTDAKSSFTGKYHCYLVYVEHYTHIEHAIEREKQIKKWRREKKDALIASTNPEWRFNKEINDD
<E2R66_RS01100--Mucilaginibacter_psychrotolerans--58.62 (SEQ ID NO:1988)
MWNHNYMYIITNPGKTTLYVGMTNDVGIRAFQHKENRGTTNNSFAGKYCYCLVYEHFTHVEQAIAREKEVKKWRREKKLALIESINPEWRFNDEVGG
<E2R66_RS10415--Mucilaginibacter_psychrotolerans--60.71 (SEQ ID NO:1989)
MTNKYNSVLYTGVSSELISRVWDHKNKTYPKSFTARYNCYKLVYCYFTHIEEAI AIEKVIKGSSEYKKQLVNAMNPEWKDLYDGLLEE
<E3C22_RS06320--Jiella_endophytica--76.74 (SEQ ID NO:1990)
MGGWVYIMASQRNGTLYVGVTSDLHRVLEHREGDVKGFTARYGCRTL VWHERRDDIEIAI QRETSIKRYRRRAWKLRLIEDLNPQWDLDFERFSDDEDNFPYRPS
FDGALDDLPEPDPALDSRGARE
<E3C22_RS03915--Jiella_endophytica--73.33 (SEQ ID NO:1991)
MTGSVMMSDKPRGVIYIGVTSDDLGRVWEHKNEVQEGFTSRYEAKNLVWFELHPNIVLAIRREKSLKRYSAWKINLIEGLNPTWLDLYERIDEIENVMPH
CNSREWSLYN
<E3C22_RS02530--Jiella_endophytica--70.33 (SEQ ID NO:1992)
MAGHVYIMTNKKHGTLYTGVTSDLSQRVYEHKTHAFPNNGFSAKFSCTRLVWYVEHYDIRDAI IREKSIKWRREWKIELIETMNPENWELYRGMGW
<ELS82_RS19340--Vibrio_ouchi--73.03 (SEQ ID NO:1993)
MBELELKQPAVYILANESGCLYIGVTSNLRKARIWQHKNSSVDGFSKRYNIDRLVYELYLDITDAIKREKQMKKWKREWKESLIAETNLSWADLYDTL
<CU048_RS05150--Beijerinckiaaceae_bacterium--73.49 (SEQ ID NO:1994)
MSNKRSTLYIGVTSDLSSRVWEHREGVIDGFTKRYELKRLVYETDYDIRNALQREKNLKHWPVWKLALIHGMNPWDWEDLYERLW
<E4S40_RS15010--Algoriphagus_kandeliae--65.93 (SEQ ID NO:1995)
MEKGGAVYIKTNSNHTVLYTGVTSDLLRRVHEHKTGLNKNSTSKYKITKL VFYEFHFSVEEAINREKQIKGGRKKKESLINIINPEWKDLWDEIREW
<E4T55_RS03010--Legionella_israelensis--80.00 (SEQ ID NO:1996)
MKESYVYILASKRHGTLYIGLTSNLIKRVWEHKNKCVPGFTSRYNVAVLVYVEIFNDITLAAAREKRLKEWKQWKIELIEKMNPDRDLYHEI IK
<E4T54_RS01860--Legionella_geestiana--75.61 (SEQ ID NO:1997)
MENGRTILCLHITSKKYGTLYTGITSNLVQRIWQHKEGLAEGFTKKYNVHHLVYVEIHNDAYQAITREKRIKKWNRQWKINFIDQNNPDWDLDFGIGLF
<EGY25_RS01315--Brevundimonas_intermedia--68.13 (SEQ ID NO:1998)
MYTADSLIGCYMMANRMHGVVYVGVSSQLVTRVQGHRAGVIDGFTKRYGLKRLVWVEFHETIVGAI QREKSLKRWPRDWKANLIERANPHWDDL FADLVRASG
AEPDPDAYRNWTPPEE
<EGY25_RS02745--Brevundimonas_intermedia--70.79 (SEQ ID NO:1999)
MANSFYVYLLASGCPGLYLRGVNTDIVRRVAEHKSGLIPGYTADRIGIDQLVWYETHQYIDQAILREKRIKRWLREWKFALVEKDNPRWLDLYGEVRAKAEPS
AAPTGLG
<E0W73_RS08470--Flavobacterium_foetidum--66.67 (SEQ ID NO:2000)
MLNPQIGFHSYVYVYITNEHRSSFYIGVTNNLKKRLAKHQENIDLNISSFAAKYNI RFLVYEFKFTWIQEAIAREKELKKWRDKKIELIRSFNENFEFLNPH
FE
<EU557_RS24125--Hymenobacter_wooponensis--63.22 (SEQ ID NO:2001)
MSYFVYMLSNPARTVLYIGMTNNLERRLYEHSGKLGDAKGFTGRYQCNNLLVYFETSPDAGHAIAREKQIKNWSRTKKEWLIETLNPTWEPIDLRTWK
<EU557_RS09750--Hymenobacter_wooponensis--55.29 (SEQ ID NO:2002)
MLRQAQHDVLLFLSSYMYVYIITNPIKTVLYIGVTNHLTRRLYEHSSGRGAANKFTGRYQADLLVYFELCSDATQAIAREKQLKGWTRRRKKNVLI TAFNPTWQ
VIDLETWIG
<EU557_RS21170--Hymenobacter_wooponensis--55.17 (SEQ ID NO:2003)
MLYTYITTNPARNVLYIGVTNNLTTRLGQHYANRGNAQT FAGRYFCYHLLYFEEYPSNAAAIAREKELKGWTRMKKEALIVTQNPWFELSTG
<E5J99_RS00195--Hymenobacter_elongatus--60.00 (SEQ ID NO:2004)
MRAHKYFVYIVTNPKTVLYTGITNDLETRLHQHCENRDLPTTFAGRYHCYLLLYFEHYREVEQAIAREKEIKGWTRLKKEALVRQFNPEWQIIDPQLWQPAE
E
<E5J99_RS03975--Hymenobacter_elongatus--52.87 (SEQ ID NO:2005)
MFAHCYVYIVTNPARTSVYTGVTNDVRRRIQEHFDNRGNPGTFAGRYFCYNLVYVEPFEHIDGAIAREKEIKGWMRVKKNALIAVKNPQWLALNYAER
<E5K00_RS11815--Hymenobacter_aquaticus--59.55 (SEQ ID NO:2006)
MAVYVYILTNTQTQTVLYIGVTNDLTRLRYEHSSDRGDAGKFTGRYQTDLLVYFEIAPDATQAIAREKQLKGWTRKKKDALI TEFNPTWQAI DLETWSG
<E5K00_RS01900--Hymenobacter_aquaticus--54.02 (SEQ ID NO:2007)
MFAHCYVYIVTNPARTSVYTGVTNDVARRIQEHFDNRGNRSTFAGRYFCYNRVYVEPFEHIEGAIAREKEIKGWTRAKKDALIEMANPQWLTLNYAEL
<EN848_RS02435--bacterium_M00.F.Ca.ET.205.01.1.1--67.74 (SEQ ID NO:2008)
MTGYVYMTASQKGTIYIGVTNDLGRMPHKAQAPGFTSRYGVRQLVWYEEHFDIADAI QREKSLKRWPRQWKIELIEKANPEWFELFRGIGW
<EN848_RS03510--bacterium_M00.F.Ca.ET.205.01.1.1--65.59 (SEQ ID NO:2009)
MTGYVYMTASQKGTIYIGVTNDLARRMPEHKTKGSSFTSRYGVRQLVWYEEYFDITDAI QRETSKRWPRQWKIELIEKTNPEWFELFRGIGW

FIG. 11KKK

<EM932_RS16535--Flavivirga_rizhaonensis--69.41 (SEQ ID NO:2010)
MDKHYYILTNKNHTTLVYGRSKQLKQRLVQHKNNLSKSFSGKYNVTKLVYFETTKYVNNNSIKREKQIKKWNREWKINLNSLNPDKWDLSEYI
<EM932_RS01860--Flavivirga_rizhaonensis--61.36 (SEQ ID NO:2011)
MKKSFVYFMTNKNNTVIYVGVTSDDLKRVYQHKTKAYKGFTYKYNCCKLIYFEEFNNDINQAIIVREKQIKGGNRKRKEYLIHSINPEWNDLSDGWLFYFD
<EM932_RS18605--Flavivirga_rizhaonensis--65.88 (SEQ ID NO:2012)
MKSSYVYILTNNKYRTTFYIGITANLSKRLEEHYDETASKFTKKYNIKDLIYFEEFSDIEQAIAREKQKLNWHKEWKNLNIKELNPTLKTLDY
<EM932_RS13245--Flavivirga_rizhaonensis--59.77 (SEQ ID NO:2013)
MKTGKTHNYVYVLTNNKMTVLYTGVTTNNLKERLYFHKNEPFPKSAFTSKYKCFYLIYYEHFFEIENAIKREKQIKGWSREKKNLSSINPKWSFLDDTI
<EM932_RS18745--Flavivirga_rizhaonensis--57.95 (SEQ ID NO:2014)
MLHKYFVYITNNKYRTTLYIGFTNNIQNRLSQHYFDSQNAKKSFTGKYGCIYLVVYETFDYPKLGIAREKELKKWRREKKNKLVSDFNPKWENLNNQIF
<EN795_RS06845--bacterium_M00.F.Ca.ET.152.01.1.1--67.74 (SEQ ID NO:2015)
MTGVVYMTASQKRGTIYIGVTNDLGRMPPEHKAGQAPGFTSRYGVQRLVWVYEEHFDIADAIQREKSLKRWRPQWKIELIEKANPEWFELFRGIGW
<EN795_RS07920--bacterium_M00.F.Ca.ET.152.01.1.1--65.59 (SEQ ID NO:2016)
MTGVVYMTASQKGGTIYIGVTNDLARRMPEHKTGKGSFTSRYGVQRLVWVYEEYFDITDAIQRETSKLRWRPQWKIELIEKTNPEWFELFRGIGW
<EN799_RS63715--bacterium_M00.F.Ca.ET.156.01.1.1--68.54 (SEQ ID NO:2017)
MTGVVYMTASQKRGTIYIGVTNDLGRVPEHKSQGSFTSRYGVQRLVWVYEEYFDITDAIQREKSLKRWRPQWKIELIEKANPEWFELFRGTGW
<EN799_RS13200--bacterium_M00.F.Ca.ET.156.01.1.1--68.54 (SEQ ID NO:2018)
MTGVVYMTASQKGGTIYIGVTNDLGRMPPEHKSQGSYFTSRYGVQRLVWVYEEYFDITDAIQREKSLKRWSRQWKIELIEKANPEWFELFRGTGW
<E5A73_RS01100--Sphingomonas_gei--75.58 (SEQ ID NO:2019)
MASGRNGTLYTGVTSQLLGRIHQHNRKTFDGTAKYDVNRVWFVHAEMEPAIIVREKRIKEWRREWKLELIEKKNPTWRDLAEDFGDFPLQDDDP
<E5A73_RS12635--Sphingomonas_gei--67.82 (SEQ ID NO:2020)
MIERAPCVYLLANRYNGALYVGVTSNLVARIHQHREGTFDGHTKKYGIHRLVYFETGDSMEGAIAREKRVKLWRDWRKNLIERQNPEWNDLAVGLGLPLAE
<E5A74_RS05470--Sphingomonas_naasensis--74.19 (SEQ ID NO:2021)
MGGYVYIVTNRKNGALYTGVTADIAQRTWQHREGTGSRTAKHRIHLRVYVERHDDIENAIIREKRIKKWNRRAWKIALIEKSNPEWRDLWFDLNNK
<E5A74_RS15320--Sphingomonas_naasensis--70.79 (SEQ ID NO:2022)
MEDKGYVYIMASGRNGTLYTGSAFDLAKRAWEHNRGVVDGFTKKYGCKLLVWYEPHSSLEAARQREKPMKEWQRAWKLREIEGMNPEWEDLYDRIATP
<E5A74_RS11100--Sphingomonas_naasensis--68.97 (SEQ ID NO:2023)
MRKAASAGRPAKGRIRIGAVSRPLRHPGLVPGSTARQAVRVKIDGIPCVYILASRYNGALYVGVTSNLPGRVLQHRDGTGYEGHTKKYGCRLVYEEVWETMTE
AIPREKRLKWRDRDWRKNLIERRNPANWDLAAGLGLPEWVP
<EN805_RS06855--bacterium_M00.F.Ca.ET.162.01.1.1--67.74 (SEQ ID NO:2024)
MTGVVYMTASQKRGTIYIGVTNDLGRMPPEHKAGQAPGFTSRYGVQRLVWVYEEHFDIADAIQREKSLKRWRPQWKIELIEKANPEWFELFRGIGW
<EN805_RS07930--bacterium_M00.F.Ca.ET.162.01.1.1--65.59 (SEQ ID NO:2025)
MTGVVYMTASQKGGTIYIGVTNDLARRMPEHKTGKGSFTSRYGVQRLVWVYEEYFDITDAIQRETSKLRWRPQWKIELIEKTNPEWFELFRGIGW
<EJE47_RS05680--Wolbachia_endosymbiont_of_Aedes_albopictus--66.67 (SEQ ID NO:2026)
MFLAFSLTSNPLGLYIGITSNLTKRIWEHKSKAISGFTSKYNVCKLVYFEEFQEIINLAISREKLLKSWQRKWKIDIEKTNQEWKDLDEIVSG
<EJB00_RS02260--Wolbachia_endosymbiont_of_Drosophila_mauritiana--75.00 (SEQ ID NO:2027)
MKGYIYILASERNGTLYIGITSNLTKRIWEHKSKAISGFTSKYNVCKLVYFEEFQDINLAISREKLLKSWQRKWKINLIEKTNQEWKDMNKYCLVWSKRR
<GS03_RS05355--Flavobacterium_sangjuense--66.32 (SEQ ID NO:2028)
MKPGFIYIVTNKYNNFTFYIGVTSNLPQRILEHIEKRYENSFSARYNKLKVYEEYQYQMGIDAIAREKQKGGSSREKKIELIRSINPNWIDLFEIIDIITI
<E6V41_RS04770--bacterium_B66--72.04 (SEQ ID NO:2029)
MKQPCVYIITNKHNTTLYIGVTSNLPQRIYQHKAKLVAGFSKRYNLQKLVYFEQFSTMEDAILREKRLKNWRRAWKDDLSAENPNWKDLYENCL
<B1991_RS13540--Rhodanobacter_lindaniclasticus--83.87 (SEQ ID NO:2030)
MRERQPCVYLLASQRNGTLYVGVTSNLPARIWQHRSNVAEGFTAHHHVHVLVWFELHETMESAITREKAIKAWRRRAWKVELIEKSNPYWRDLAEICG
<E7Z59_RS02670--Robertkochia_marina--69.47 (SEQ ID NO:2031)
MFETSCVYIILTNAHTVLYVGVTSNLTTRRIYQHKSKMYPGFTTRYNCSSLVYETFLDMQMAIAREKQIKKRSRTYKINLIEADNPEWNDLSEGWIFDVI
<E5988_RS05530--Sphingomonas_olei--82.61 (SEQ ID NO:2032)
MERQPCVYIFASGRHGTLYIGVTSNLPVRLHQRHREGLIEGFTTRYGIHRLVWFDMADTMEAAITREKQLKKWNRDWRKLRLIEEGNPEWRDLAIDLGFARVEKT
MRMDSRLRGNDVARWNDGTGKHDRD
<FAA86_RS10875--Rhizobium_rosettiformans_W3--68.89 (SEQ ID NO:2033)
MSGFVYMMSDKPRGVIYTGVTSDLQRIWEHRNEERKGFTSRYNAKLLVWFQHPNIVLAIIRREKSLKRYLRDWKIKLIEGLNPTWLDLYDRIYEIENVSPH
PSSPQWSDYN
<FAA97_RS04405--Rhizobium_ipomoeae--69.15 (SEQ ID NO:2034)
MAGFVYIVTNHKGRTLYIGVTSDSLQRIYEHREGQTPGFASKYGCRLVWYEDHDRIGTAIQREKSLKRWRPQWKIELIEGFNPDWRDLYCELM
<FAA97_RS03970--Rhizobium_ipomoeae--69.66 (SEQ ID NO:2035)
MSGFVYMMSDKPRGVIYTGVTSDLQRIWEHRNEVRKGFTSRYHAKLLVWFQHPNIVLAIIRREKSLKRYLRDWKIKLIEGLNPTWLDLDCDRIYEIENIYSPH
PSSPQWSDYN
<EZV76_RS03305--Muricauda_alvinocaridis--59.49 (SEQ ID NO:2036)
MKLSYVYILKCSGDGTYTGVTSLEKRIEBHIVGKRSTSYFSRRPVELVYAKFTQIELAIQKEKQIKKWSKAKKEALINGQYDALPNLAKKKFQ
<E5222_RS15350--Altererythrobacter_aquimixticola--78.49 (SEQ ID NO:2037)
MPRDFQPAVYLLTSRRNGTLYIGVTSNLLQRIAQHREGTFDGFQKQHGCRILVWYESHATMESAIRREKQLKKWNRRAWKMRILIEANRPWRDLAADFGFEALH
<E5222_RS00215--Altererythrobacter_aquimixticola--76.09 (SEQ ID NO:2038)
MRPAPVYIMANKRNGAIYTGVTNLPQVRVQHQREGAAGFTKRYNCKLLVYERHETMEAILREKQVKAGSRKRKLALIEAMNSQWRDLFEDIAHG
<E5222_RS14550--Altererythrobacter_aquimixticola--64.44 (SEQ ID NO:2039)
MGGWVYILTNRDGLVYIGVTSGLAARMMQHRAGTGSFACRKYGLKTLVYARHRDIEDAIAREKAMKAWKRAWKIELIEKTNPGWDDLDFTLV
<E5222_RS01720--Altererythrobacter_aquimixticola--54.29 (SEQ ID NO:2040)
MAFYAYLLRCNDGSYYAGHTDDIDLVRVARHNSGALGGYTAARLPVTLVWSDTFPTREEAFAAERRIKGSRAKKEALIAAGDWDLVSRLLARNRQG
<E8M12_RS05925--Thalassotalea_mangrovi--77.42 (SEQ ID NO:2041)
MKQPAVYIITNTSHSVIYIGVTSNLPQRMQHRQKLDGFSKYNLVCLVYFEAFDDMVNAITREKQLKKWSRSWKDELIGRLNPEWRDLYDDIL
<E8M01_RS02220--Phreatobacter_stygius--73.26 (SEQ ID NO:2042)
MYVYVYMMASDRNGTLYVGVTTNDIARRSQHRTGTGAFTRKYGVVKLWVMEPHDSIVEAIGREKQIKGWNRAWKLIERNPNPHWEDLYERLNG

FIG. 11LLL

<E8M01_RS00460--Phreatobacter_stygius--65.85 (SEQ ID NO:2043)
MTNRPNGTLTYGVASDLRRRAWEHREGAMPGFTRYGLKMLVFFERFEDIRDAIQREKNIKHWPRAWKVRLLILAENPDWDLDDQLGGRVPTGGVTDGWSSAV
AG
<FAZ90_RS14135--Vibrio_cyclitrophicus--75.28 (SEQ ID NO:2044)
MDRQPCVYVLSSPNNVLYIGVTSQQLQRIWQHKNIMHSFTSKYNNVHNLVYELFEDIEAAIAREKQQLKTWKRAWKNELVKSVPNSWQDLYHTL
<CFBP5473_RS04315--Agrobacterium_larrymoorei--77.78 (SEQ ID NO:2045)
MSGYVYIVASQKRGTYLIGVTSNLEQRIYEHREGLTKGFTSRYGCTRLVWYEEHMSIGTAIPREKSLKRWRDWWKIKLIEDFNPEWRDLYEEFG
<EZV72_RS05365--Salinimonas_lutimaris--67.05 (SEQ ID NO:2046)
MKQPAVYILSNARNTVLYVGVTSDISMRIETHQNAKMTGFAKRYNLYKLIVFEYIDSMEQAIIVREKRLKKWPRKWKEELIDNVNPNGRNDISHLIDPVY
<FDP09_RS09155--Echinicola_rosea--68.13 (SEQ ID NO:2047)
MEKGGCIYILTQNNTTLYVGVTSNLIKRIHEHKFTTESTKSFTYRYNINKLVYETFHSHIEDAIREKQIKGGRKKKKELINSFNPDKDLWDDIQKW
<E0502_RS07725--Altererythrobacter_lutipelagi--60.00 (SEQ ID NO:2048)
MVSETKGGWVYIMADRYRGTMYVGVTAYLAEVQQHSGSGSDFCTKYGLDRLVWAEQAGADVSDCIAHEKRIKRWREWKFEELIERGNPDWSDLSGSL
<EBL82_RS10125--Chakrabartia_godavariana--72.04 (SEQ ID NO:2049)
MDAAFQPCVYLLASQRNGTIYIGVTSRFPVQRISQHRTRSVKGFATRHHGALRLVRYEFFGTMEIAIAREKQQLKRWHRDWWKINLIEHDNPEWRDLAPDIGASPIV
RGMDAETSSA
<EBL82_RS11260--Chakrabartia_godavariana--70.65 (SEQ ID NO:2050)
MPKQPCVYILASRRHGTIYVGVTSDLVGSGLHAHRTRTAGFTKEHRVARLVRYEQFPDMLSAIAREKQQLKNWRRRAWKINLIEQDNLWDHDLAVDFGFEPLISA
ARPMDEPETSSG
<EBL82_RS04705--Chakrabartia_godavariana--74.71 (SEQ ID NO:2051)
MASQPNGTIYLVGTNSNLAQRAYQHRHGLVEGFSKKHGCTLLVWYEHKDDIQDARIREYQMKSWKRAWKRLRIEGRNPQWRDLFGELMR
<EBL82_RS03315--Chakrabartia_godavariana--62.77 (SEQ ID NO:2052)
MGGFVYIMTNKPLGLVLYVGVTSLSAEVMAHLEGRGSFACCKWDLRCLVYIEPHPTIEGAIVREKAIAKNWRRWLKRLRISEANPNWDDLFDSSINA
<EBL82_RS01100--Chakrabartia_godavariana--65.12 (SEQ ID NO:2053)
MEKGGWVYIMANRYRGMVGVTDLARRVGQHQHREGIGSTHVVDKRLVYAEHRHEDIAAIAAREKLVKKWRREWKFALIEADNPEWVDLWEQWYPTSEGPK
PDGA
<FDY95_RS10670--Hymenobacter_jeollabukensis--69.41 (SEQ ID NO:2054)
MHYYVYILTQNSTVLYVGVTDNLIRRVAEHREGQHKGFTKRYNLNKLIVFEYEPDAKAAIAREKQQLKPGNRQKKLDAVNAVNPNAELLPE
<FDY95_RS23750--Hymenobacter_jeollabukensis--63.10 (SEQ ID NO:2055)
MGFFVYILPNQARTVLYIGVTDNLERRLHEHSGQLGDAGKFTGRYQCNLLIYFESSPDATQAIAREKQQLKKWSRAKKEALITALNPGWTPIDLHTWAG
<FDY95_RS08210--Hymenobacter_jeollabukensis--62.07 (SEQ ID NO:2056)
MSYFVYILSNPARTVLYIGVTDNLERRLHEHGGQLGDAGKFTGRYQCNLLVYFEISPDVQAIAREKQQLKNWSRTKKEWLVGTNLNPRWEPIPRWTQGG
<FE810_RS08945--Thalassotalea_litorea--75.27 (SEQ ID NO:2057)
MKQPAVYIITNQTSNLIYIGVTSNLPRIYQHREKLIAGFSLKYHLNKLIVFEIYDMPNIAIAREKQQLKWLRAWKNELINKENPDWQDLYPTVVL
<FFT13_RS02005--Wolbachia_endosymbiont_of_Drosophila_santomea--74.71 (SEQ ID NO:2058)
MKLSYIYILTNNKANTSLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVFEFEDI SLAIARERRLEKEWPRQWKINLINQKNPSWIDLYYLALR
<FFT11_RS02240--Wolbachia_endosymbiont_of_Drosophila_yakuba--74.71 (SEQ ID NO:2059)
MKLSYIYILTNNKANTSLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVFEFEDI SLAIARERRLEKEWPRQWKINLINQKNPSWIDLYYLALR
<FFT12_RS03560--Wolbachia_endosymbiont_of_Drosophila_teissieri--74.71 (SEQ ID NO:2060)
MKLSYIYILTNNKANTSLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVFEFEDI SLAIARERRLEKEWPRQWKINLINQKNPSWIDLYYLALR
<FCN14_RS11125--Aliifodinibius_saliphilus--67.78 (SEQ ID NO:2061)
MEQKYHVYIMTNPHHTVLYTGMDGDIKRGWQHKQLTDGFTKRYNCTKMVYIDEYDHPQDAIDREKQIKAGSRQKKIDLIBELNPDWRDLYEELIR
<FCN14_RS14175--Aliifodinibius_saliphilus--64.44 (SEQ ID NO:2062)
MEKKYHVYIMTNPHHTVLYTGITSDIIRGWQHRKREVEGFTKRYNCTKMIYVEEYDHPQEAIDRENQIKGWTRKKKEALIKRLNPEWKDLYGEFI
<FC091_RS04620--Winogradskyella_algicola--55.29 (SEQ ID NO:2063)
MRLYYVYILKCSGDSGYTGMTNDLERRLIEHQTKASYGYTYKRLPVKLWVYLECSKPDEAIGIEKKIKGWSRKKKEALINENWDDLVRFSKNYTEYGPKTSS
TSSD
<CWC18_RS19635--Pseudoalteromonas_aurantia--74.19 (SEQ ID NO:2064)
MQPTTYILASKPQGTLYIGVTRNLKKRIYEHKHHFVEGFTKRYQVTELVYFETYSMDYSAITREKQLKNWRREWKINLIEKVNPNWDDL YLSI
<FDP22_RS17675--Paraoceanicella_profunda--64.04 (SEQ ID NO:2065)
MAFYVYILASRPGGALYTGSTDDLRRRVEQHRSGALAGHTDYGIRTLVWFEAHDTRAELLRERRIKRWRLWKEQLIMQHNPQWRDITADIPW
<FF124_RS11395--Martelella_lutilitoris--72.22 (SEQ ID NO:2066)
MAGYVYIVTNRPHGTLYIGVTSDLERRVFEHREGLIKGFTSRYGCRRLVWYVEHFEIGSAIAREKALKGWNRAWKIALVSDFNPGWDDL YERFV
<E6A44_RS14520--Pedobacter_ureilyticus--64.77 (SEQ ID NO:2067)
MKRGGCVYIMTNVNTVLYIGVTSDLFVRVTQHKDKIYPRSFYAKYCNKLVYCYFYSTIEEAI AVEKQMKAWNREWKINLIVKNPNQWTDLYQEAL
<E6A44_RS07335--Pedobacter_ureilyticus--67.03 (SEQ ID NO:2068)
MQRGGCIYIMTNKNTTLYIGVTSNLSRIHQHKTHIFPNSFTAKYNLEVCYVYDVFSSIEAILREKEIKKWRREKKEELINRINPEWDLWAEISKW
<E6A44_RS08715--Pedobacter_ureilyticus--63.74 (SEQ ID NO:2069)
MQKGGSIYILTNNVNTVLYIGVTSDLRSRLIAHREKKYPKSFTAKYGAVICIYYETFFSIEEAIEREKEIKKWRREKKNLINSMPNSWKDLWDEKVEK
<E6A44_RS04720--Pedobacter_ureilyticus--67.53 (SEQ ID NO:2070)
MQRGGLVYILTNNKNTVLYVGVTSDIARYIWOHKNHIYPNSFAAKYNCEKLVYCYFYPTIEEAI FEEKRIKSGSRLAKITLIKRYQSALERFV
<FGG15_RS19235--Flagellimonas_algicola--66.30 (SEQ ID NO:2071)
MSFYVYIISNKKNGTIYIGYTDNLGKRMKQHKNGTGSFRFAKRYALDKLVYEEKFKFMSAIKREKQLKKWNRQWKINLINDFNPKWEDLTWFFE
<FGG15_RS19270--Flagellimonas_algicola--63.53 (SEQ ID NO:2072)
MKSSFVYIITNQYRTTFYIGVTADLYKRVSEHENGIGSTFTKKYNLKDLYFEEDSDINQAIAREKQIKNWHKEWKINLIEKNPSLSTLKL
<FGG15_RS10000--Flagellimonas_algicola--57.65 (SEQ ID NO:2073)
MNLEKGYVYMLTNASRTIYIGATRNIKKIRKXIDGTGAVFTKKYHLKILYIEEFDDVLELAFKREKQLKNWKKDWKWNLVKSNNKDLDDLNLVFC
<FGG15_RS12510--Flagellimonas_algicola--60.98 (SEQ ID NO:2074)
MNSHFRTYYIYILTNNKNTVLYIGVTSNLTBRLSEHHENIILKRRTFAARYNCRHLLYYEQFSRIIDAIAREKEIKGWLRVKKLKLKSMNPEMEFMEHLFL

FIG. 11MMM

<FE772_RS13650--Lysobacter_enzymogenes--71.74 (SEQ ID NO:2075)
MDANPNPLRLSARQGTLYAGVTSDDPTQRIWQHNRDLTEGFTRRHGVHVLVWYELHPDMLSAIIREKRIKSWKLELIESANPQWRDLYSAIIGASAAPAS
<FHG64_RS07800--Antarcticibacterium_flavum--60.23 (SEQ ID NO:2076)
MRIQGHHTYVYIILTNQHKTVLYTGVTDNLRIRLQQHKEGAEKREQSFTARYNCFLLVHYEYKWTWQEAIAAREKEIKSWTRKKKEELISRNPNKWEFLNHHFF
INL
<OUY_RS04150--Wolbachia_endosymbiont_of_Leptopilina_clavipes--74.44 (SEQ ID NO:2077)
MKSYYIYLASERNGLTYTGITSNLTKRIWEHKSKAISGFTSKYNVCKLVYFEFQEIINLAISREKLLKSWQRKWKINLIEKTNQEWKDYDEIVSG
<FHY50_RS02230--Sphingomonas_japonica--77.01 (SEQ ID NO:2078)
MEKVPTVYIVASQRNGTVYTGVTSDLMARIHQHRTKAFKGFTAKHSCDRLVWFVHETMASAILREKQVKNWRDWDKLALIEQDNPRWYDLAETLGFELIAP
QPSRST
<FHY50_RS00720--Sphingomonas_japonica--70.11 (SEQ ID NO:2079)
MANHRRGQTYLGVTSNLVQRAFQHRNKLFDSKDKGCTLLVWYEEYDDLQEARACEYRMKKWKRAWKLRLIEHNPGWKDYGEIIA
<FIL70_RS16200--Sphingobium_fuliginis_ATCC_27551--78.65 (SEQ ID NO:2080)
MKREISPMVYILASRRNGTLYTGVTSDLVKRLYEHRNGLIEGFTSDYGVKRLVWFEMHDAIDGAILREKRIKKWNRQWKIEIEAGNPEWDDDLALAFGEPLR
ERAGDGFPLSRE
<FIL70_RS15010--Sphingobium_fuliginis_ATCC_27551--70.45 (SEQ ID NO:2081)
MKEPCVYILARQPYGTLYIGVTSNLPGRHLHQHRTNGIPRFSGRYAVHSLRVFEMCETMEQAIIDREKQLKRWHRQWKINLMESENPRWIDLAPGLGLSPLHPNQ
RHDGC
<FIL70_RS01395--Sphingobium_fuliginis_ATCC_27551--75.00 (SEQ ID NO:2082)
MANRRNGTLYLGVTSDLAARAYQHRKGLIDGFTKDHCKLLVWFAYDDLQEARHCELMMKKWKRAWKIELIERDNPQWRICSKRCSDDALGSRFRGSTAME
<FIL70_RS18025--Sphingobium_fuliginis_ATCC_27551--65.85 (SEQ ID NO:2083)
MTNKRGRVLYVGITTNLAVRVDQHRGMSGDFCRKYNLHRLVLAEEHGRVDDAIAREKALKAWKRDWKLQLEIQQSNPEWRDLSDFIA
<EAH84_RS00460--Sphingomonas_oligophenolica--70.65 (SEQ ID NO:2084)
MNGPAYVYIMANHRNGTIYLGSTNLPKRAWHRNGVVEGFTKRYGCHLLVWYETHDSWEAARQLERQMKWEWRNWKVQIEIGENPDWDDLYDRIARP
<EAH76_RS01585--Sphingomonas_glacialis--66.28 (SEQ ID NO:2085)
MRGGWYIYIMSNRYRGTYIGVTSIAGRSFQHRDKTGSDFCRRYDLRLVYADFAENIADAIAREKAMKKWRRRAWKIELIERANPKWLDLFFETLNGAAEEAAG
FPLSRE
<EAH76_RS18275--Sphingomonas_glacialis--61.11 (SEQ ID NO:2086)
MGGYTYIMTNKAFGLVYVGTADARIQAAREGRSAFARKWGTTRVYIEPHERIEDAIVREKRLKDWDRVWKCRLIVEANPNWDDLWETINN
<FIA58_RS15600--Flavobacterium_jejuense--74.12 (SEQ ID NO:2087)
MASKNNKVLVYGFNTNIIIRRTKEHKDRINKGFTYRYNVDKLVYVEFGYVEDAILREKRLKKWNRWKIQLIEENNPIWDDLSLDW
<FIA58_RS11665--Flavobacterium_jejuense--61.36 (SEQ ID NO:2088)
MKELTNGYHTYVYIILTNKHRTTFYIGVTSNLLKIRLQQHIDSVTENTHSFVSRYKVEYLVYEEKHTWTWQEAIAAREKELKGWRREKKLHLIQSFNESFAFLNHH
FLDK
<VIN01S_RS18180--Vibrio_inusitatus_NBRC_102082--76.34 (SEQ ID NO:2089)
MSKHPVCVYIILSSNNRNVLYIGVTSNLLGRVWQHKNKVVEGFTYQYNASILVYEMHDSMESAILREKQLKKWNRRAWKEHLVNKFNQWRDLYGDLF
<VIN01S_RS18210--Vibrio_inusitatus_NBRC_102082--70.45 (SEQ ID NO:2090)
MKAPCVYILSSINKRVLYIGVTSNLLKGRVWQHKSIFYPSFTSRVHVTRLVYVEVHDDMEHAIEKEKRLKFWRRRAWKDALIDEMNPKWNDLYETL
<VCO01S_RS16065--Vibrio_comitans_NBRC_102076--70.45 (SEQ ID NO:2191)
MKTPCVYILSSINKRVLYIGVTSNLLKGRVWQHKSIFYPSFTSRVHVTRLVYVEVHDDMEHAIEKEKRLKFWRRRAWKDALIDEMNPKWNDLYETL
<FKV25_RS06600--Lysobacter_aestuarii--69.89 (SEQ ID NO:2092)
MKHPCVYMLASDHGHTLYVGVTSNLLTRVWHHRNDVTEGFTRKYGVHDLVWFESHESSMAPAIQREKAIKAWRREWKIQLIEERNPRWLDLYPTLLQARKHCI
AVLPNARRAEPRGAKSNAKAGSQRSIA
<FB551_RS10710--Chryseobacterium_aquifrigidense--66.28 (SEQ ID NO:2093)
MTAFSKGRHTYIYIILTNKSRRVLYTGVTGNLHRRLYQHKTCLNPGSFTARYNLEFLIYEEKYNWIHNIAIRREKEIKNWSRIKKLELIRNTNPNLEFLNYLFE
NWL
<FWH45_RS18845--Sphingomonas_trueperi--74.70 (SEQ ID NO:2094)
MANKRNGTIYVGSTDLAARAYQHRTGMIEGFTKDKGCTLLVWYEAHEDLDGARLRERQMKWKRAWKIELIERDNPQWKDYGLTIFG
<FAF07_RS02450--Changchengzhania_lutea--65.91 (SEQ ID NO:2095)
MKKSHMYFMTNKNNTVIYVGVTSNLLKRVYQHKTKTHKGFTFKYNCCKLVYFEVFNNDINQAITREKQIKGSRKQKEDLINSNPNPEWKDLSGWLIFYFD
<FAF07_RS07820--Changchengzhania_lutea--57.95 (SEQ ID NO:2096)
MQHKYFVYITTNKYRTTLYIGFTNNIQRNLSQHYFDSQNAKKSFAKYGCVLYVYETFEYPKLGISREKELKKWRREKKNKLSDFNPKWENLNQIF
<FNA46_RS25280--Rhizobium_straminoryzae--76.34 (SEQ ID NO:2097)
MRGYVYIVTNQRNGTLYIGVTSDIARRIWEHREGLNEGFAAKYGCRMLVWYEEHDRIETAIQREKSLKRWRKWKIQLIESMNPDPWRDLYPELW
<FLL45_RS14310--Aliikangiella_marina--60.67 (SEQ ID NO:2098)
MSYVYVYILTSTHNNVYVGVSNLERRIFEHSNAQYNSFCRRYNIKLVYFEFCFNSLDAIQREKQIKKWRREKKIKLIEAANPTWEDLQTCGGSRLYGRDDG
<BVG17_RS05330--Wolbachia_endosymbiont_of_Laodelphax_striatellus--70.45 (SEQ ID NO:2099)
MILTSHKHKALYIGITSNLTKRIWEHKNKVISGFTSKYNVQKLVYFEIFQDVGSALEKREKLLKSWKREWKIDLVEREKNPWADLYDEIIK
<BWZ41_RS02355--Wolbachia_endosymbiont_of_Nilaparvata_lugens--74.44 (SEQ ID NO:2100)
MKSYYIYLASERNGLTYTGITSNLTKRIWEHKSKAISGFTSKYNVCKLVYFEFQEIINLAISREKLLKSWQRKWKINLIEKTNQEWKDYDEIVSG
<FN961_RS07190--Shewanella_hanedai--81.91 (SEQ ID NO:2101)
MQPCYIILANKRNGTLYIGVTSSELVKRVWEHKEHLVEGFTKKYSVNLVYFEATDTMYSAITREKQLKNWCRQWKINLIEETNPEWNLDFDQIR
<FNB15_RS03665--Ferrovibrio_terrae--67.78 (SEQ ID NO:2102)
MAYAWLYIMTNKPNGLIYVGVTDIQRRAHYEHREGLIAGFTKRYNKLRLVYEAHDDIQLALQREANMKHWPRAWKIRLILAQNPDWTDLYDTLV
<FND99_RS01170--Flavobacterium_daemonense--66.32 (SEQ ID NO:2103)
MKPGFYIITNKYQTVVYTGVTNLPKRILEHKNNKPKSFSARYDVNILLVYEQFQWIEDAITREKQIKAGSREAKNNLIRSINPTWKDLFEEIEDIMI
<FND99_RS12140--Flavobacterium_daemonense--65.88 (SEQ ID NO:2104)
MHYQEGYHTYIYIILTNKAKTVYTGVSNNLRKRLNQHQENILSKNKTFAQYNAQFLLFYEKFTWQEAISREKEIKGWNRNKKLELIKINADFCFLNYLF
E
<FKM97_RS12040--Rhodoligotrophos_appendicifer--70.93 (SEQ ID NO:2105)
MLGGFYIILASGRNGTLYVGVTDLAARVHQHRRERLTPGFTSRHGVARLVHYERYEDIQREKTLKHLRLSWKIALIERDNPQWLDLYEQLNL

FIG. 11NNN

<FOF46_RS29595--Aquimarina_algiphila--78.02 (SEQ ID NO:2106)
MLKPKHQYAVYILTNKNGTLYIGMTNNLERRIFEHKQKCIDGFSKYNLDLIVYFELYQFVEDAIRREKRMKAWKRQWKINLIEEDNKDWNLDASDWYY
<FOF46_RS29590--Aquimarina_algiphila--76.60 (SEQ ID NO:2107)
MLKPKHQYAVYILTNKNGTSLIGMTNNLERRIFEHKQKCIDGFTKKNLDLIVYFELYQFVEDAIRREKRMKAWKRQWKINLIEEDNKDWNLDASDWYD
<FOF46_RS22295--Aquimarina_algiphila--65.17 (SEQ ID NO:2108)
MKKPCVYILTNKNNVTIYTGVTSNIVRRVYQHKTKFYKGAFSARYNCNKLIVYFEELNSMGLAIAREKQLKAGNRKRKEALINMENPEWKDLSDGWLFVFD
<FNZ56_RS11505--Lysobacter_lycopersici--80.65 (SEQ ID NO:2109)
MEKQPAVYIILAHRRYGVLVGVTSDLIARTWQHRTHAVEGFTSKFNVAMLVWYEHDTMYSAITREKQIKKWRREWKEVIEKENPEWIDLWPKIIGADN
<FNZ56_RS02340--Lysobacter_lycopersici--73.91 (SEQ ID NO:2110)
MKLPCTYILASRWGHTLYTGVTSDLVKRVWEHKHVDVAGGFTRKYRVHDLVWFERHATMELAIAREKAIKEWRRRAWKEVLIETANPKWVDLYPDIIDR
<FNZ56_RS06175--Lysobacter_lycopersici--66.67 (SEQ ID NO:2111)
MERTYFVYILSSRSRNLVGVGTGDLARRMWEHRTHAVPGHTARYRINRLVHFETTADPASAIAREKQIKGWLRSRKIALIERSNPTWDDLAEGWFADSASGKA
GPSLRSG
<PPF71_RS00650--Algibacter_amyololyticus--63.95 (SEQ ID NO:2112)
MKYVYVYIMSNKPKGVLYIGFTDSIEERVKEHKLKVYPSFTAKYNCNKLIYFEEFDNGNEAEKRRERQFKKWKRDWKIKLIEDMNPWSWSDLSINWNLNYNKLR
K
<OF66_RS05175--Seleniivibrio_woodruffii--75.56 (SEQ ID NO:2113)
MREGVYVILSNKNNVDVLYGVTSDIVKRIWEHKNKVTDGFTSRYNVTKLVIYECYNDIETAIVREKQIKGGSRRKKKIDLVESVKNKDWDLYPDIK
<FBZ95_RS23665--Bradyrhizobium_sacchari--73.12 (SEQ ID NO:2114)
MKQPCVYMFANRRNGTIYTGVTSNPPRAFEHREGLVKGFSRYGCKLLVWYELHDTMIDAIAREKQIKGGSRAKKLALIESLNPDMWMDLYETLI
<FBZ95_RS12630--Bradyrhizobium_sacchari--80.00 (SEQ ID NO:2115)
MVYVYVYLLASKKYGTLYLGVTDNDIVRRIHEHKTQVAGFSKRYSDRLVWFEIYDDPITAITREKELKKWRRQWKVRLIEEKNPQWIDLSPQIAS
<FBZ95_RS23835--Bradyrhizobium_sacchari--74.29 (SEQ ID NO:2116)
MSAGSYVYVHILASRHHGTIYIGVTNDLRLTRLELHRSGRGSKFVQKYKIFRLVHVEAFATPQEAIAREKRLKFWKRDWKIKLIEKNPDWDRSGLV
<IQ13_RS12090--Lacibacter_cauensis--65.22 (SEQ ID NO:2117)
MAHGGSVYIMTNILRTVVYIGVTSELRLTRILQHREHYFKGSFTDKYNCITCVYYENFSRIEEAIAREKELKKWNRTKKNELITLFSQWIDLWTEIEKW
<LY60_RS17405--Sedimentibacter_saalensis--73.03 (SEQ ID NO:2118)
MYIYIILTNKYNKVLVGVTDNDLIRRIYEHKNSLDGFTAKYKHKLVIYDCTEDVISAITREKQIKGWTREKKIKLIESMNPDKDLWSDLN
<IQ17_RS06375--Bradyrhizobium_daqingense--77.65 (SEQ ID NO:2119)
MPFFVYILASKRNGTLYGVTDNDLTRRMTEHKAKLIPGFTQYDVTLVYFETFEALSAREHTLKRWRRAWKLIKIEDLNPGRDLTDELNTLAT
<JMN93_RS04850--Roseibium_hamelinense--64.71 (SEQ ID NO:2120)
MRKRYVYVILTRKGGTLYTGVTNDLPRRLVEHMSGTSGQFTTKRYSVKRLIYWEHETAPQAIAREKAIKRWRKWKIDLIEGLNPDWFDLGKTLNR
<JN11_RS04965--Mucilaginibacter_frigoritolerans--63.04 (SEQ ID NO:2121)
MDSKNCVYITTNIAHSLVLYTGVTSDIVRRVYHKNKTYPNSTAKYDCNKLIVYHAFYSTIEEALAAEKVIKSGNREYKIKLINELNPEWKELYDDLIKS
<JN11_RS00365--Mucilaginibacter_frigoritolerans--61.36 (SEQ ID NO:2122)
MYFQRGCVYITTNKLHVTLYTGVTSDITGRIWEHKNKIHPNSFTAKYNCDKLVYFFYSIEEAIASEKAIKGGNRKYKQLVESINPEWIDLVDKLMEE
<IQ16_RS25380--Bradyrhizobium_huanghuaihaiense--77.53 (SEQ ID NO:2123)
MGGRHYVYVILASKIGGTYIGVTNDLIRVAEHKSKLIESFTERYDVARLVYFEQFDDPENAIKREKRLKKWNRRAWKVRLEIKENPNWDDLYPGIAGPP
<IQ16_RS10100--Bradyrhizobium_huanghuaihaiense--72.04 (SEQ ID NO:2124)
MTYVYVYILASRKYGTYLIGVTSDLVRRVYEHKTKVVPGFTRKYGVDKLVLFELFDDPTSIAIAREKELKKWRRDWKTRLIDEQNPWDDLYSGISS
<IQ16_RS09465--Bradyrhizobium_huanghuaihaiense--75.82 (SEQ ID NO:2125)
MQTAFFVYVILASKRNGTLYGVTDNDLVRMVAHKAKLVPGFTKQYGVTLVYFETYGSILEARAREHSLKRWRRAWKLIKIEDLNPDWDRDLTDELSFWAG
<IQ16_RS18905--Bradyrhizobium_huanghuaihaiense--65.88 (SEQ ID NO:2126)
MYVYVYILASRHHGTLYIGVTNSLQKRLQHRAGKGSEFVKRYGVNRLVYTEAYERAEAAIAREKQLKRWKIELIERDNPWEWRDLSDLLV
<IQ16_RS43615--Bradyrhizobium_huanghuaihaiense--63.24 (SEQ ID NO:2127)
MSTGSYVYVILASRHHGTIYIGITNNIRMRLEQHRGRGSKFVQKYKIFRLVHIEAFATPQGAIAREKQLKFWKRD
<CBW56_RS07245--Denitratisoma_oestradiolicum--79.35 (SEQ ID NO:2128)
MKQPCVYILASRRNGTLYTGVTSNLIQRIWQHRNNLVGGFTKKYGVHTLVWFELHETMESAIGREKAIKEWRRRAWKMALIEKNPMWCDLYDSIT
<FQY83_RS03795--Luteimonas_marina--82.80 (SEQ ID NO:2129)
MKNQPATYIILASGRNGTLYGVTSDLIGRVWQHRNHLAPGFTDRYRVTRLVWYEVHDAMEAAITREKQLKKWNRRAWKIRLVEESNPYWNLDWPSLIGESQDTG
SPPSRG
<FQY83_RS10635--Luteimonas_marina--81.52 (SEQ ID NO:2130)
MDKQPCVYIILASDRNGMLYIGVTSDLIGRTWQHREHIVDGFTTKRYGVARLVWYEMHATMESAILHEKRIKKWRRAWKAQLIDEKNPSWRDLWPDITGQVPKA
<FQY83_RS03960--Luteimonas_marina--77.42 (SEQ ID NO:2131)
MKQPCVYIILASRWNGTLYTGVTSNLVKRVWEHRNDVVPGFTRKYRVHNLVWFQHDMSMSAIAAREKALKKEWKRAWKIELIDRTNPGWRDLYPDMIGT
<FPZ08_RS18430--Devosia_ginsengisoli--80.00 (SEQ ID NO:2132)
MGRDYFVYVILTNRKHGTLYTGVTNDLVRVVEHREGLVDGFSRRHGCKHLVWFEIHGDI EAAILREKQVKKWLRPWKDALIEQTNPDWRDLWWEIIK
<FPZ08_RS17285--Devosia_ginsengisoli--70.79 (SEQ ID NO:2133)
MAWVYVYILTNRKFGTLYIGATTDLVQVRVQHKGEFIPAFTRKYHLNRLVYFEEHQTYGAMQORERTMKHWKAWEKVALIEKNPEWDDLYWVIAPG
<FQY64_RS10590--Vulcaniibacterium_thermophilum--82.80 (SEQ ID NO:2134)
MNRQPAVYIILASGRNGTLYTGVTSDLIRRVQCHREHAVDGFRTARYGVACLVWYELHATMDSIAIVREKRIKKWRRAWKLALIEASNPHYWNLDWPEIIGQVVRP
IGKFDPC
<ESZ36_RS16215--Colwellia_demingiae--82.80 (SEQ ID NO:2135)
MKQPCVYILASEKNGTLYGVTSNVLKRIWQHKHNVSGFSEKHEVKRLVYFELQDDMVNAITREKVLKKWNRRAWKIRLIEKTNLWNDLYDDIV
<ESZ26_RS07215--Colwellia_hornerae--71.74 (SEQ ID NO:2136)
MKQPHYIYIMTNVKNNTLYIGVTSQLAQRVYQHKSKLTGFTKYNLNLKLVYFESFETMYDAITREKQLKNWRRRAWKNLIEQINPNWQDLSKDL
<FRF71_RS04790--Novosphingobium_ginsenosidimutans--65.56 (SEQ ID NO:2137)
MSGGWYVILANRKGALYTGVTADLPARMVQHRSGIGRKFCAEYIDIKLVHVEWHDRIEADIVREKAIKRWRRAWKIELIEAANPDWRDRWFECG
<FRZ54_RS23320--Mucilaginibacter_ginsenosidivorans--61.25 (SEQ ID NO:2138)
MTNPYHNVLVYGVTSNLIARAWDHKNKTYPGSFTAKYNCNKLIYQYFRIEEAIAAEKTLKGSNRKKESLAILSNPDWDLVYDDIIKE

<FRZ54_RS10525--Mucilaginibacter_ginsenosidivorans--58.62 (SEQ ID NO:2139)
MWNYNFYMYITTNPEKKVLYVGMTNNLRVRLQHTENKGPDSFAGKYCYCNLVYVEHFTVHDHAIAREKQVKKWRREKKVALIKSENPTWRFLNNEI
<FSB76_RS13705--Mucilaginibacter_ginsenosidivorax--64.84 (SEQ ID NO:2140)
MKVHEYIYILSNKANTVFYIGVTNDLQSRQAQHKQKVNTGFSQAQYNCNKLIVYEEFQWVQDAIAREKQLKAGSRQKKIDLIVGFNPSWDDLSDGWYD
<FSB76_RS13820--Mucilaginibacter_ginsenosidivorax--65.93 (SEQ ID NO:2141)
MGRGGYIYIVTNMTHSVLYIGVTSDIVSRWDHKNKTYPDSFTAKYNCNKLIVYEEGYPHIEEIAAEKLLKGGSRANKIRLINEMNPEWKDLYDGLVDDL
<FSB76_RS13285--Mucilaginibacter_ginsenosidivorax--58.33 (SEQ ID NO:2142)
MWNHNYIYITTNPAKTVLYIGVTNDLCVRLQHQRENKGTGKSFQGYCYCNLVYVEHFTHIEHAIEREKEIKKWRREKKDALIATNPNWTFNKEVCDDD
<RAS_RS02220--Rickettsia_asiatica--77.42 (SEQ ID NO:2143)
MKQPIVYIVTNKKNGLYTGVTNLIKRKYQHKNDIIRGFSKKYNCKILVFYIEHETMNSVINREKQIKAGSRDKKIKLIEQTNINWDDLYNEIIL
<RAS_RS04070--Rickettsia_asiatica--82.76 (SEQ ID NO:2144)
MYIFIANKRNGTIYVGVTSNLVKRIYEHKNNNTNGFSKKYDCKILVFYIEHETMDSIAIAREKQIKSGSRAKKNLNIEQMNVNWKDLYDEII
<RAS_RS04285--Rickettsia_asiatica--81.16 (SEQ ID NO:2145)
MYWVYILCSNKNGLTYVGTNNIIRRVYEHKQKIIKGFTLKYNITKLIVYTEBFTDIQEALTRKKNLKKWKRS
<RAS_RS03950--Rickettsia_asiatica--75.00 (SEQ ID NO:2146)
MGVTNNIIRRAYEHKQKTVKGFTAKYNIKLAHTEBFTDIKEALTREKALKKWKRSWKIELIEKTNPEWKDLIE
<SAQ01S_RS02865--Sphingomonas_aquatilis_NBRC_16722--79.78 (SEQ ID NO:2147)
MRDVFQPCVYILASRRHGTYIGVTSNLLARLIQHREGLPGFTKAYGVHVLVWYEVHATMEAAIVREKRLKEWRRRAWKIELIEAQNEDWGDGLVGLGLAPLP
PVPTKESMT
<SAQ01S_RS18870--Sphingomonas_aquatilis_NBRC_16722--76.74 (SEQ ID NO:2148)
MTKVGQVYMMASRLNGTYILGVTSDLAKRAYQHRNALIDGFTKIYGCRLLVWYEAHDELD SARLRERQLKKWNRDWKLRLIEEKNPDWLDLYESLF
<SAQ01S_RS18540--Sphingomonas_aquatilis_NBRC_16722--69.32 (SEQ ID NO:2149)
MRERQPCVYILASGNGTYIGVTSNLIQRLQHRDGTGEGFTARYGVHLVWFETDTMADAIASEKRIKAWRRDWKKNLIEKANPGWLDLAVNLGLPPLG
<ODE01S_RS01970--Oceanithermus_desulfurans_NBRC_100063--81.91 (SEQ ID NO:2150)
MEPAVYILASKKDGVLTYIGVTSSELKRRVWEHREGIGATFTRKYHVHRLVYERHPSMESAIVREKRLKKWERAWKVRLEIAMNPGWRDLYDL
<NSE01_RS14355--Novosphingobium_sediminis--76.34 (SEQ ID NO:2151)
MRQDFQPTVYLLASSRNGTYIGVTSNLIQRIHQHREGTFGGFTARHTIHRVLVWFQAHATMEHAIIREKRFKKWERAWKLNLEKDNPTWRDLAEDFGFPPL
Q
<NSE01_RS06390--Novosphingobium_sediminis--68.18 (SEQ ID NO:2152)
MLNLFQHPSCFMPCCAMPTEKNPCVYILASGHYGTTYIGVTSNLMKRLGQHWQGVTRGFTSRAYVYRLVHFEMFGDMERAI SREKQLKNWRREWKNINLINGVNP
EWRDLAVGLGFPPLAPVRDRDP
<AAE02nite_RS03250--Adhaeribacter_aerolatus--58.02 (SEQ ID NO:2153)
MWNKNYFVYITTNPKKTAYVGMTNNLERLEEHRENKGPETFAGKYCYCNLVYERHTTANHAIEREGEVKLMNREEKEKLIKANSANPKLLFLKIEG
<SAE01_RS00280--Segetibacter_aerophilus--64.84 (SEQ ID NO:2154)
MKKGGAIYILTNLHQVYFTGVTSDVVKRIQEHKEKKYSNSFTATYQLDKLVYYENYLTIEEIAAEKRIKAGNRASKIKLIISKNAEWKDLWDEISQW
<MHA02_RS01670--Methylobacterium_haplocladii--77.91 (SEQ ID NO:2155)
MKTYAVYLLARKRDGTYLGVTSNLSRRVWEHRSAKIPGFSARYGVDRVLVWYELYDRIDEAIAAREKAMKKWRRRAWKVALIEEANPTWDLAEQLNA
<MHA02_RS02820--Methylobacterium_haplocladii--71.60 (SEQ ID NO:2156)
MPYFYMLGSRRHGTYIGVTSNLDLARRVWEHKTTHAHRGFTARYRVERLIWYETERYIDEAITREKALKTWRRDWKVVLIEE
<RS001_RS45675--Reyranella_solii--65.93 (SEQ ID NO:2157)
MERNPVVHILASPNRRALYIGITTDLPRLLEHRQGVVHTARYNIKLIYFEAHTAPDAIAREKQLKNWSREKKILILNRSNPDWRDLSEIH
<RS001_RS25480--Reyranella_solii--60.47 (SEQ ID NO:2158)
MSERYFVYIMMSTNSLNTMYIGVTNDIGRRVFEHRNRKGSEFAAKYRVTRLVYVEEFENPTEAIAAREKQLKGWKIRKKNQLVRAANPTWADLAPSS
<RS001_RS31215--Reyranella_solii--60.00 (SEQ ID NO:2159)
MHAYYVYVLTNVKRNVMYIGVTNDLERRVADHCDGRGAFTRYRLNTLVHVEEFQYIDEAIAAREKEIKGWRRSKKDALVEASNPTWADLADHGVAKGPSLRS
G
<FRY97_RS13455--Phaeodactylibacter_luteus--68.60 (SEQ ID NO:2160)
MSRAGFVYIMTNKNRTTYIGVTNDLCRRYEHKNNHLVENAFTDKYNLEYCIYEEYFDFLAIKREKELKKWNRQKKEALINRKNPEWKVLATEKGFMHKQV
PFAQQVADLLRELEENQEIPPCSGNDTPS
<FBZ93_RS28765--Bradyrhizobium_macuxiense--73.40 (SEQ ID NO:2161)
MSYVYVYILASRRDGAIVYGTNDLVRRVYEHKMKAVPSFTSKYNI TQLVWFIEYDDPI SAISREKELKKWKRQWKTQLIEATNPEWNDLYESICV
<FBZ93_RS00095--Bradyrhizobium_macuxiense--74.19 (SEQ ID NO:2162)
MKQPCVYIVANKRHGTIYTGVTANPPKRSFEHREGSVKGFSAKYGCRMLVWYELHETMYDAITREKQIKAGSRAKKIALIEALNPGWNDLYDTLI
<FBZ93_RS22310--Bradyrhizobium_macuxiense--70.59 (SEQ ID NO:2163)
MDAAVYVYVYILASRRHGTYIGVTNDLRNRLTLHRAGRSQFVKKYGVTRLVYVEVYASPOEAIAREKALKEWRRDWKIRLIEQDNPEWNDLSHLP
<FBZ93_RS13885--Bradyrhizobium_macuxiense--69.41 (SEQ ID NO:2164)
MHYVYVYILASRRHGTYIGVTSNVQKRLEEHRAIGSSFVKTYGVYRLVYVEAYERAEEAIAAREKQLKRWKRDWKIELIERENLEWRDLSDLLV
<FRX97_RS06330--Luteibaculum_oceani--65.56 (SEQ ID NO:2165)
MPTTAKEFYLYITTNKYRGTYIGVTNNLPRRILEHQSGYGGKFSYTYNCKLLVYFBPHDCYKVLQRREQVIKRWRRNWKLEKKNPDWRDLCEEILPKNWE
EIKLRKFIRKEY
<FRX97_RS05115--Luteibaculum_oceani--63.95 (SEQ ID NO:2166)
MANYHVYILTNKRKTVLYIGVCNNLKRRLYEHQQGKLLNAFSKRYGCOFLIYERFEDINQAI GREKELKKWRRDKKEALIEKTNPNWDFLNDEV
<FSB78_RS17530--Sphingomonas_ginsenosidivorax--80.68 (SEQ ID NO:2167)
MERQPCVYILASAKHGTYIGVTSNLLGRLHQHREGLIKGFTSRYGVRVLVHYEMADTMDAAITREKQLKTWNRDWKVTLEIRDNPDWDDLAIGLGLDPVPSR
GAR
<FSB78_RS12560--Sphingomonas_ginsenosidivorax--75.00 (SEQ ID NO:2168)
MRERQPCVYILASGRNGTYIGVTSNLVGRIMQHRDCSFEFGFTKRYGVTRLVWFDMDTMEGAIEAEKRIKKWPRRYKLTMIERDNPDWRDLALDLGLPLD
<FSB78_RS00110--Sphingomonas_ginsenosidivorax--70.45 (SEQ ID NO:2169)
MPRDRQPCVYILASGVRGTYIGVTSDLGARLYQHRTGVTGRFVGQYAVYRLVRYEMSDDMESAI SREKQLKRWHRGWKINLVESQNPPEWNDLAPGLGLAEPL
TPRR

FIG. 11PPP

<FSB78_RS05410--Sphingomonas_ginsenosidivorax--71.59 (SEQ ID NO:2170)
MAREPCVYIILSSGRHGTLYIGVTSLSMARLYQHRTGAFPGFTKKYDVKRLVHFEMFADMPSAIAREKQLKAWRRDWKIELIEVENPFWEDLAIGLGLLEPLPTR
<FSB78_RS01320--Sphingomonas_ginsenosidivorax--72.22 (SEQ ID NO:2171)
MEKPGFVYIMASGRNGTIYIGVTSGLAKRVWEHRNGVVPGFTHKYGCHDLVWFEAHDTEAARQRELRMKWEKRDWVKVEIEWMNPWNDDLFDTLTPD
<FSB78_RS12040--Sphingomonas_ginsenosidivorax--69.77 (SEQ ID NO:2172)
MVHGGWYIIMASKPHGLLYIGVTNHAAARVDQHRREDEGSAFCCRYYGIRMLVFAEAHETIQAAIWREKALKAWNRRAWKIRLIEELNPEWRDLYDTLA
<FSZ31_RS12460--Sphingorhabdus_soli--74.42 (SEQ ID NO:2173)
MANRKQGAITYGSTSNLLARIFQHRERLVEGHSKTYGCTRLVWFEIHDDLQNAHRELQIKKWKRSWKIRLIESGNPDWKLWFDIRA
<ESV24_RS01575--Aequorivita_lipolytica--65.15 (SEQ ID NO:2174)
MERRIFEHKKATHKTHAGYNNIKKLIVYFEEHQWIEQAIKREKTIKKWKREWKINII TENNPEWKDLSGDWDL SKFETNKE
<ESU54_RS11925--Aequorivita_antarctica--62.03 (SEQ ID NO:2175)
MKTYYVYILSCSDGSYYTGVTSNLTQRLLEDHQHRKHLNGYTSSRLPVLAYCTTTEITIAIAAEKQIKKWSRAKKEALINEAYELLPNLAKKKFKK
<ESY88_RS17170--Subsaximicrobium_wynnwilliamsii--64.21 (SEQ ID NO:2176)
MKPGFAYLLTNKNNTTLVVGSTSRVLQRIKQHKKEKFFPKSFSAQYNLNLKLVYFEAFQMGIDAVGREQQLKAGNRKAKVALINGANPEWLDLYEEVKQLFGSDA
EE
<ES711_RS14825--Gelidibacter_salicanalis--69.57 (SEQ ID NO:2177)
MKPSFVYILTNKNHTTLYIGVTSNIVNRIHQHKDKYYPKSFTARYNLDKLVYFEAFQMGIDAGREKQLKAGSRAKKIALIDSINPEWVDLYDVVRAEFGDGE
LE
<ES711_RS00600--Gelidibacter_salicanalis--65.88 (SEQ ID NO:2178)
MDVFFVYILTNKNHCVLYVGRTKQLKKRLYQHAHNTKKSFTGRYNVHKLVYFETTRYVNNAIQREKQIKKWNREWKISLINSINPDWKLSEYISD
<ES711_RS13075--Gelidibacter_salicanalis--61.18 (SEQ ID NO:2179)
MSREYYTYILTHERNHIYLVGVTNNMARMSEHKAATFKTHVGHYNNIKKLIVYETHDRLAIKREKTIKKWKREWKINII TQMNPEWNLDSNGWDL SKY
<FUA26_RS08640--Seonamhaeicola_algicola--67.06 (SEQ ID NO:2180)
MEVHYVYILTNKNHTVLYIGRTKQLKERLQKHNNSKSFTGKYNVTKLVYFETTKYVNNSIKREKQIKKWNREWKISLINSINPDWKLSSYI
<CS053_RS07515--Rhodanobacter_glycinis--80.65 (SEQ ID NO:2181)
MDRRQPCVYMLASKHNGTLYIGVTSDLIKRVWQHRNQVVEGFTSRHRIHDLVWYEQHETMESAILREKQLKEWKRVWKLRLIEGENPEWRDLYPTLL
<CS053_RS17545--Rhodanobacter_glycinis--79.35 (SEQ ID NO:2182)
MTERNPCVYLLASHRNGTLYVGVTSDLPARVWQHRGEVVDGFTQRHHVHGLVWFEQHETMESAIAREKAIKAWKRAWKIELIEKSNPYWRDLYAEICG
<FTW19_RS07090--Terriglobus_albidus--68.24 (SEQ ID NO:2183)
MAREYFVDFVYILASRSRTLYIGMTNGIGRRQRQHAKMQGFSFSTRYNIYRLVYERHQYVLNATAREKELKTWSRQKIELIERENPAWVDLSEAFYRNYKGE
RIT
<EA652_RS02230--Wolbachia_endosymbiont_of_Drosophila_ananassae--74.71 (SEQ ID NO:2184)
MKLSYIYILTNKANSTLYIGVTANLIKRIWEHKNKIVSGFTSRYNVNKLIVYFEEDFEDISLAIARERRLEKWPQWKINLINQKNPSWVDLYLALR
<FVF61_RS05945--Formosa_maritima--69.32 (SEQ ID NO:2185)
MYKAIHQYYLYVLTKNTNGTLYIGVTNDLERRIFEHKNKLAVGFTKKYGLNKLIYFETFQYVNDIAIKREKNMKKWRQWKIDLIEENPCWNLDSKDWIYLI
<FVF61_RS05715--Formosa_maritima--70.53 (SEQ ID NO:2186)
MKPGFVYILTNKNNTTLVVGVTSNIVLRVQHKKEHMLKAFTAKYNLEKLVYFEAFQMGIDAGREKQLKAGSRKVALINSNPEWLDLYEEVKRAFGGFD
YGDEEE
<FVF61_RS08725--Formosa_maritima--55.81 (SEQ ID NO:2187)
MLTHERNHIYVGVNTNNIERCISEHKSASFGIHVGHYNNIKKLIVFEEHSDIRIAI KREKTIKKWKREWKKNQITEMNPDWKLNDLWDLSKYK
<FUA24_RS04670--Seonamhaeicola_marinus--65.88 (SEQ ID NO:2188)
MDIHVYILTNNTNTVLYIGRTKQLKERLQHRNNSKSFTGKYNVTKLVYFETTKYVNNSIKREKQLKKNWREWKISLINSINPDWKLSEYI
<FUA24_RS09335--Seonamhaeicola_marinus--63.95 (SEQ ID NO:2189)
MSKRVYHNYWVYIITNKPGQVLYIGVTGGIDDRMERHVEGTGSVFSAKYKLLKLVYFEFQFIEDAIKREKQLKNWRDWKINLIESLNPWENLWKPCHPET
SSG
<FUA24_RS12150--Seonamhaeicola_marinus--63.95 (SEQ ID NO:2190)
MKHWYVYIMNRPNGVLYIGVTDSDIDRVKEHKLKVYPKAFTRAYNCDKLVYFEFENGSEAEIRERQFKKWKRDWKIDLIEDMNPWNIDLTENWNLFNFKLR
I
<FUA24_RS21960--Seonamhaeicola_marinus--53.33 (SEQ ID NO:2191)
MKEGFFVYILTNKNKTVLYVGATGNLKNRVKEHISGKGSVFTRKYSSYLIVFEKIAKYEDAFKREKQIKKNWNEWKNWNLISKNPDLKNLYLDLLE
<ES674_RS09805--Bizionia_myxarmorum--67.74 (SEQ ID NO:2192)
MKQGFVYILTNKNNTTLYIGVTADILTRVKQHRDKLDSKSFTAKYNLNLKLVYFEVFMIGDAIGREKQLKAGSRATKLALIHGFNPEWLDLYADLEEAFGV
<FMN50_RS12005--Rhodobacterales_bacterium--75.28 (SEQ ID NO:2193)
MMFFVYVILASRQRTLYTGVTNDLARRVHEHREKTASGFTTRYDVTRLVYETFDTPDAAIAREKRLKRWPREWKIQAIENFPNGWDLYEDLNR
<FXB38_RS30240--Bradyrhizobium_cytisi--74.19 (SEQ ID NO:2194)
MKQPCVYMFANRRNGTIYTGVTSNPPRAFEHREGLVKGFSSQYGCCKLLVWFELHATMTDAITREKQIKGGSRAKKLALIESVNPWGWDLYETLI
<FXB38_RS05700--Bradyrhizobium_cytisi--76.40 (SEQ ID NO:2195)
MGGRQYVYVILASRIGGTYIGVTNDLIHRVAEHKSKLIESFTEKYDVARLVYFEQFDDPENAIKREKRLKKNRAWKVRILIEETNPNWDDLYSGIAGPP
<FXB38_RS14695--Bradyrhizobium_cytisi--75.27 (SEQ ID NO:2196)
MAYYVYVLLASKKYGTLYLGVNTNIVRRTHEHKSIVIAGFTKRYSVSRVLVWFEIYDDPVTATITREKELKKWRREWNIRLIEEKNPQWIDLYPQIAV
<FXB38_RS05725--Bradyrhizobium_cytisi--75.28 (SEQ ID NO:2197)
MAGLVPAIHVRQAHMGGWYVILASAAAGLYLVGVTNDIVRIYEHNRGLVPGFTKKYGVKLLVYFERYEDIQTAIQREHNIKHSRTWKVRLILKDNPEWK
DLYDTITQ
<FXB38_RS14980--Bradyrhizobium_cytisi--65.88 (SEQ ID NO:2198)
MNAAPAYVYILASRRYGTLYIGVCNNISNRLALHRSRGSKFVKYGVTRLVYVETIYASPSAIAREKALKEWQRDWKIRLIERDNPDRDL SRL
<LX73_RS07465--Fodinibius_salinus--67.78 (SEQ ID NO:2199)
MATKRGFYIILSNKSRVLYIGVTSDDLQRIHQHRIGSGSSFTKRYNLRPIYEEYKSIKTAIEREKQLKNWERTWKDLIREVNPTLDDIWTDIANNR
<LX73_RS05035--Fodinibius_salinus--66.67 (SEQ ID NO:2200)
MRFQNVITKRGYVYILSNKTRSVLYIGVTNNLVKRIWMHRHSRGSEFTQKYRVKYLIVYYESFKHLERAINREKQLKNWHREWKFNIRKHNPSLKDLPKIVT
SK

FIG. 11QQQ

<LX73_RS04725--Fodinibius_salinus--69.77 (SEQ ID NO:2201)
MGYKSSSGKNGFVYILSNKTKSTLYIGVTAFLDKRIFDHVYGEAGAKTAKYNNVNIILLYEYKYPNINEAIEREKQLKNWHREWKLNLIKSENPTLENLWDS
<C7447_RS02550--Tenacibaculum_adriaticum--65.52 (SEQ ID NO:2202)
MKTSPSYFLTNKNNTVIYIGITSNLIKRIYEHKTGKYKGFAFYKNCNKLVIYETTFQSI TEAIAAREKQLKAGNRKRKEDLINKENPDWKDLSEGWLNFV
<FOE74_RS18755--Rufibacter_glacialis--58.02 (SEQ ID NO:2203)
MKQHNYFVYITSNPTKTVFYVGMTNDLSRRLIEHRENRGKPETFAGRYCYKLLYYERFTYVQHAIEREKELKLMIREEKI TLIKSLNPKMLFLKIAD
<FZC33_RS18535--Labrys_neptuniae--77.17 (SEQ ID NO:2204)
MKQPAIYIMASRRNGTLYTGVTSNLSERIYQHKQGLIKGFTSRYGCTSLVYVEFYEEMVPAILREKQIKAGSRRKKLELIESLNPTWRDLYLDF
<FZC33_RS29700--Labrys_neptuniae--75.56 (SEQ ID NO:2205)
MRHRYWVYLMASKRNGTLYVGMTNDLPRRVHEHEEKSFSFGTAQYDVNLLVWYEEHRYVGNAIQREKNIKHWPRAWKLKIEDLNPEWDDLDLHLI
<FHQ18_RS04005--Deferribacter_autotrophicus--70.89 (SEQ ID NO:2206)
MSLLCNTFLPAFYFLQPVYLLDENVIRHWGFLRAFYGWGIIVPLTTVLYTCVTSNLLKRVYEHKHELVDGFTKRYKIKKLVYVEFVDSIEBAIKREKQIKGGS
RKKKIALIESINPNWSDLYYQLD
<F0R75_RS06350--Francisella_marina--70.00 (SEQ ID NO:2207)
MAFVYILLESKNGTLYIGVTSNLIKRVYEHKQGYSDGFTKKYDVKKLVYVEQFNNTDAIYREKRLKTWQRKWKLDLINKFNPRWSDLYASILA
<FW780_RS07480--Chryseobacterium_sediminis--65.12 (SEQ ID NO:2208)
MIEFSTGRYTYIYILTNKNRKVLYTGVGTGNLHRRHLQHKTCLNPSSTFARYNLEFLIYYEKYDWIHHAIEREKEIKNWSRIKKLELIRKTNPNLEFLNYLFE
NWL
<F0361_RS04545--Maribacter_flavus--55.21 (SEQ ID NO:2209)
MKGICYILTNNKNTVLYVGATNNLKRRIRDHKEGKYPNAFTKRYNCDKLVIYFEFEDVKDAFEREQQLKAGNRKRKEESINSKNPEWKDLSDFWDFREVDSE
<F0361_RS05490--Maribacter_flavus--58.54 (SEQ ID NO:2210)
MFKEGWRYYYYIILTNRNKTVLYTGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<F0361_RS05495--Maribacter_flavus--57.32 (SEQ ID NO:2211)
MYKEGWRYYYYIILTNRNKTVLYTGVTSNLSQNRRLRQHQNIDISLNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<F1193_RS07070--Blastochloris_sulfovirodis--81.11 (SEQ ID NO:2212)
MRYAVYMLATRKDGLYVGVTDNLVRRVFEHRSGAVRGFTWQFNIHRLVWFEPHDSIEAAI IREKQIKRWHRAWKVEMI EAENPGWDDLYERAGAFPEDDPG
SRLTPRPG
<F1193_RS02885--Blastochloris_sulfovirodis--74.74 (SEQ ID NO:2213)
MRPAYVYILASRPGAIYVGVTDNLIRRVFEHRSGAVDGFTKRYNIKQLVYFEQFDSIYDAIQREKNIKHWPRAWKTRLIYQONPTWRDLFPELLS
<F1193_RS10700--Blastochloris_sulfovirodis--75.56 (SEQ ID NO:2214)
MAGWVYILSNRPNGLTYVGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<F1193_RS13035--Blastochloris_sulfovirodis--72.22 (SEQ ID NO:2215)
MTRAYVYILASRPGAIYVGVTDNLVRRVFEHRTEALDGFTKRYNIKLLVHFEQFDRI IDAIQREKNIKHWPRVWKTRLIAQTNPTWRDLYPELVM
<FZFO1_RS18180--Pseudomonas_salina--77.42 (SEQ ID NO:2216)
MRMKQPAVYILMASSRGTYLVGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<EMO90_RS11920--Bifidobacterium_vespertilionis--71.26 (SEQ ID NO:2217)
MLTNRHNTVLYIGVTTNNIVRRMHEHRNHTADGFTARYNVSKLVYMETTPDPIAAIAREKQLKGWTRAKKNALIEAANPEWRDLWPDLAGELSDFSS
<F4U95_RS15655--Sphingobium_limneticum--78.65 (SEQ ID NO:2218)
MKREISPTVYILASRRNGTLYTGVTDNLIRKLYEHRNGLIEGFTKDYGVKRLVWFEPHDSMSAILREKRIKKWNRQWKIELIEAGNPWDLDLALGFGFEPLS
SVRSE
<F4U95_RS04095--Sphingobium_limneticum--73.03 (SEQ ID NO:2219)
MESQAFVYIMANRRDGTIYIGSTVELPRRVWEHRNSVVPFGFTKRYECHLLVWYEVHDNREAAARLRERQMKKEWKRAWKLREIEGFNPEWADLYDRIAQP
<F4U95_RS13365--Sphingobium_limneticum--58.21 (SEQ ID NO:2220)
MAFHAYLLRCGDGSFYAGHTEDETRVAQHQQGLIPGYSIARRPVVLVWTESEFGTRDEALAAERRIKGWRRAKKQALIGGD
<F0184_RS04060--Roseibium_aquae--73.03 (SEQ ID NO:2221)
MPAWVYILASRPYQTYLTGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<F0184_RS02175--Roseibium_aquae--76.71 (SEQ ID NO:2222)
MAFYVYLLASRKNGTFTYTGVTNTLARRIHEHREKQADSFTRRYDVTRLVYETYTEAEAIAREKRLKRWPRAWKREDPASRAG
<F0184_RS02170--Roseibium_aquae--65.88 (SEQ ID NO:2223)
MHKRYWVYIILTRKGGTLYTGVTDNLPRRLNEHISGAGSQFTKRYSVKRLIWEHEHETAPQAIAREKAIKRWPRKWIDLI EDLNPDWFDLVRHLNR
<F0155_RS09205--Roseibium_sediminis--77.53 (SEQ ID NO:2224)
MAYYIYILASRQRTLYVGVTDNLARRVHEHREKQADGFTRQHGVTTRLVYETTFGSPFEEAIAREKRLKRWNRRAWKIQAIEAFNPWDWEDLYEDLNR
<F0155_RS13045--Roseibium_sediminis--71.91 (SEQ ID NO:2225)
MPAWVYIMASKPYETLYTGVTSNLERRIYEHREGLEGFTNIYGLKRLVWYEECAEMI EAIKREKQIKRWREWKFELIERMNPKNRDL
<F0155_RS10845--Roseibium_sediminis--71.11 (SEQ ID NO:2226)
MSGFVYIMMSDKPRGVIYTGVTSDLQKRIWEHKEIRKGFTSKYKAKNLVWFETHPNIVLAIQREKSLKRYLREWKIRLIEELNPTWNLDYERMGEIENVYEPH
ENTRKWDDYN
<FRZ61_RS02305--Hypericibacter_adhaerens--75.56 (SEQ ID NO:2227)
MRTCYVYLLASQPNGLTYIGVTNNLVRRVYEHREGALEGFTKRYGVKNLVHFEFADVYLAIQREKTLKKWPRAWKLDLIEKANPQWRDLYEEVV
<FRZ44_RS02235--Hypericibacter_terrae--74.44 (SEQ ID NO:2228)
MKVYVYCLLASQRNGTLYIGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<FRZ44_RS23265--Hypericibacter_terrae--71.76 (SEQ ID NO:2229)
MSFYVYILASDRNGTLYVGVTDNLIRRVAEHKGAKVPGFTKRYGVNRLVYFETCDNPEGAILREKHIKNWYRRWKIELIEKHNPWDHDLAPGLVA
<FEZ63_RS01545--Microvirga_brassicacearum--75.82 (SEQ ID NO:2230)
MVYCVYLLASRRHGTYLVGVTDNLSSRRVGEHKAKIIPGFSARYDVHRLVWYETRYERIDEAI TREKSLKKWRRDWKIRLIEDINPDWSDLYLTINR
<FEZ63_RS24415--Microvirga_brassicacearum--72.09 (SEQ ID NO:2231)
MQKGGWVYIMSNKRNGLTYLVGVTSNLSQNRRLRQHQNIDILNKKTFSRYKCHHLVYVEEFTWIQEAIAAREKEIKGWLRIKKLNLIKTKPNLDFLEDSFL
<F6X38_RS05010--Aureimonas_leprariae--60.00 (SEQ ID NO:2232)
MLEATVYILRCADGSYYTGLTKQPIESRVWEHNEGIEYEGFTKRRRPVTLVFTETYDRMLDAIAREKQIKGWSRAKKEALIALDYKALPALSRRNRQEQAVSTAP
SPFPHAELVEASGKVPAYSDTQHASRVDPSTGSG

FIG. 11RRR

<DPR14_RS14670--Skermanella_pratensis--72.09 (SEQ ID NO:2233)
MPKEAGAWVYIMTNRPNGLTYGVTTNLAARRIHEHREEAHDGFTMRYGLKRLVWVVRHDEILRAVQREKNIKHWSRAWKVLLILKDNPEWDDLYERLQI
<DPR14_RS26760--Skermanella_pratensis--68.89 (SEQ ID NO:2234)
MELGGVWYMMANRPDGTLYVGVTTADLARRAYEHRKRTSNRFTKRYNVQRLVWLERYYDDIRLAIQREKNIKHWPRAWKVALIVVANPEWRDLYEDLMK
<F8237_RS12495--Bradyrhizobium_betae--78.65 (SEQ ID NO:2235)
MGARCYVYVILASRIGGTYIGVTNDIIRRVAEHKSKEVIESFTEKYDVARLVYFEQFDDPENAIKREKRVKKWNRRAWKVRLIEKDNPNWDDLYPGIAGPP
<EKH55_RS14130--Ensifer_alkalisoli--73.12 (SEQ ID NO:2236)
MAGYVYIVTNQKRGTYIGVTNLERRIFEHREGMTPGFASKYGCRLVWVYBEHLQIGTAIQREKSLKRWYRQWKIELIEKNMNPWDRDLYFELW
<DEF52_RS05440--Wolbachia_endosymbiont_of_Nasonia_oneida--69.23 (SEQ ID NO:2237)
MSCLYEFKRKSIVSGFTSRYNVKNLVYFEFEDISLAIARERRLKEWPRQWKINLINQKNPSWIDLYYLALR
<FNH24_RS23875--Microlunatus_speluncae--55.13 (SEQ ID NO:2238)
MPYTYLLECVDGTFYVGVSTRNLDQRMQDHYLGSADGYTSKRRPVRLVWFAECDRIEAYALERKIKGWRRAKRVALIEGRFTDLPGLSRTSRARGSDGGQVG
R
<GCU34_RS04475--Flavobacterium_haoranii--56.79 (SEQ ID NO:2239)
MPYFVYILHSLTKDKFYTGSTANLSERIIRHNQKSKGFTGSTNEWKLVYKEEFTSQSEAILREKQIKNWKSKSKIQLIQHNNI
<GCU34_RS04460--Flavobacterium_haoranii--56.79 (SEQ ID NO:2240)
MPYFVYILHSLTKDKFYGYTANLSERIIRHNQKSKGFTGSTNDWKLKYKEEFTSQSEAILREKQIKNWKSKSKIQLIQHNNI
<GCU34_RS04470--Flavobacterium_haoranii--58.02 (SEQ ID NO:2241)
MPYFVYILHSLTKDKFYGYTANLSERIIRHNQKSKGFTGSTNDWILVHKEEFISQSEAILREKQIKNWKSKSKIQLVQNNI
<GCU42_RS12560--Sphingomonas_ginsengisoli_An_et_al._2013--73.56 (SEQ ID NO:2242)
MKLPVCYILASQPRGTLYVGVTSDDLVARLYQHRSGEVKGFTTRYGVMLLVREFPFETMLDAIEREKQLKRWHRQWKINLIESENPDWHDLDVGLGDFDMPSP
<GCU42_RS02640--Sphingomonas_ginsengisoli_An_et_al._2013--70.65 (SEQ ID NO:2243)
MQREPCVNIILASGFKGTLYTGVTSDLVARMFKHREGTTRGFTTRYGVTRLVWFERHEMLTAIAREKTIKRWREWKIELIVEANPSWRDLAEDFGFEPLIGK
RQVDPGSSPG
<GCU42_RS09870--Sphingomonas_ginsengisoli_An_et_al._2013--65.88 (SEQ ID NO:2244)
MPGWVYILTNPFGVLYTGVTADLVERVFQHRAGRSRFAHKYQCHTLVYAEHEHERIEDAIQREKRIKEWPRLWKLRLIHEANPEWIDLSGS
<EY643_RS02755--Halioglobus_maricola--64.52 (SEQ ID NO:2245)
MANKAAVYILASRHNGSLYIGVTRNLMKRIWHHRNVAEGFSTPCTRHRLVYFEFLPDMGEALRREARLKTWRPWPTERIIREQNPDMWDLWDQVQI
<GCU44_RS12840--Cyclobacterium_xiamenense--68.13 (SEQ ID NO:2246)
MKKGGAVYIMTNQNTTLYVGVTENLLKRVMEHKSRLTPSSFTSRYKIKKLVYFELFHSIEEATCREKQLKSGSRKKKNDLINAFNPSWKDLYEEVKNW
<EA082_RS15470--Pseudomonas_pelagia--75.82 (SEQ ID NO:2247)
MRLPAVYMMASGVRGTLYVGVTSDDLVRVQHRAGTFGGFTGRHQVKELVWYERHGSMEAAICREKVLKKWRRLWKLELVEELNPEWRDLWAEIAE
<FZ934_RS07775--Rhizobium_grahamii--69.66 (SEQ ID NO:2248)
MKGYVYILASQQRNGTYTGVTNLPRLFEHNDLTPGFTSRYGVKTLVWVEEYDVLVTDIAIAREKAIKNWPRKWKLDELIEAMNPWDQDIAHFLHGL
<F8A92_RS14070--Cumulibacter_manganitolterans--59.21 (SEQ ID NO:2249)
MYILRCADDTLYVGSTTDLESRMWQHQNQGGGAAYTRRRRFEVLVYTESFDSADAAYFREKQVQVQWSRAKRDALIEGRLLDDITSNMSTTWAKNRAERDARGRR
PG
<F8A92_RS15395--Cumulibacter_manganitolterans--58.44 (SEQ ID NO:2250)
MPAMYILRCSDATYVVGSTWDVESRLWQHQNQGEAAAYTRRRRPVELVYSEYIESIAAAFVREKQVQVQWSRAKREALIRGDYAEIKRLGRSSRGDLSGK
<F3168_RS16040--Polymorphobacter_fuscus--69.79 (SEQ ID NO:2251)
MDPGGFVYILASRKQGTLYIGVTSDLVRRISEHRLDLTPGFTSRHGVKRLVWMERYDDIALAIGREKQMKRWNRDWKIRLVEAGNPWDWDLAVSVLGFAPLPY
TLSSGSGPPP
<GUJ39_RS23305--Pedobacter_petrophilus--64.84 (SEQ ID NO:2252)
MERGGCIYILTNFNNTVLYIGVTSDLYSRIIEHREKKYPQSFTSKYCNKLIYYEQFSTITEAIGREKQLKNWRREWIKIKLIQTQNNQWDDLFLSLED
<GN111_RS21210--Desulfofarcina_widdelii--80.68 (SEQ ID NO:2253)
MLASKRNGTYIGVTSDLAKRVWEHKNDVVGFTKRYKVHQLVWYIEIHATMESAIRREKQLKNWKRWNKLDLIEKNPTWKDFYNEIV
<GN111_RS18855--Desulfofarcina_widdelii--75.56 (SEQ ID NO:2254)
MNGKHYYVYILCNKRNGLTYGTISNLIKCIWQHKEAVEGFTKQYGLKCLVHFEQYQDVKEALLREKRIKKWNRQWKINLIEQNNPNWDDLYQKLV
<GN112_RS17470--Desulfofarcina_ovata--81.91 (SEQ ID NO:2255)
MQPAVTILASKRNGTYIGVTNVLKRVWEHKNNIADFTKRYNVHQLVWYIEIHATMEAAIQREKQLKNWKRWKWLDLIEKVNPAWQDLYGQIV
<E1180_RS09255--Roseibium_denhamense--70.79 (SEQ ID NO:2256)
MAYFVYILAGRKHGTYIGVTNDLARRLDQHRRGDAASFTHKNVTRLVFAEAFHHVDEAIAEMEKLKKWRRSWKIQLEKSNPDWDLVSVGL
<E1180_RS00705--Roseibium_denhamense--65.88 (SEQ ID NO:2257)
MQKRFVYIILTRKGGTYIGVTNDLPRRLFEGTSGTGAQFTKRYSVKRLIWEYEEHATAPEAVAREKAIKRWPWKWIDLIEAMNPWDFDLGKTLNM
<E1140_RS16655--Fulvivirga_lutimaris--67.44 (SEQ ID NO:2258)
MEFYVYMLTNKNHSLVLTGFTDNLDRRVFEHKNQLVSGFTSKYCNKLVYFEFIEKDEALHRERQLKRYKRSWKENLINEKNPEWEDLYGNY
<E1140_RS17105--Fulvivirga_lutimaris--61.80 (SEQ ID NO:2259)
MAKGGYVYIMTNKTRTVLYIGVTNLTARTYQHKNQEGSHFTIKYKCYDLYIEFFERIESAILREKQMKKWKREWKENLIVSFNPSKRDLFDEIQDF
<E1140_RS20570--Fulvivirga_lutimaris--63.64 (SEQ ID NO:2260)
MERGGAVYIMTNQHHTVFYTGVTNLTARVQEHMKVYKPSFTSKYCNKLVYWFESFHSIEEAIMREKQIKVYRREKKLALINALNPQGVDFDLSLE
<E1140_RS17340--Fulvivirga_lutimaris--61.25 (SEQ ID NO:2261)
MAILQPYTVYILKCSNDSYVGVTDIDRLKEHQEGLIKGYTSSRPLELVFTEHFQEVNRAIAFEKQIKGWRRAKKEALISGSWHLLPELSKAYKK
<GJ689_RS13970--Rhodoplanes_serenus--77.53 (SEQ ID NO:2262)
MSFYVYLLASRRNGTYAGMTDDLVRVWEHRTGVVAGFTRMVAVKTLVWVEPHTATRASAFRLERQVKKWNRRAWKQLIEASNPWQDLWTELGP
<GJ689_RS03910--Rhodoplanes_serenus--69.77 (SEQ ID NO:2263)
MHSGWYIILNHPNGTYIGVTNDLVRVVAEHREGVVPGFTRRYGIKRLVYFEGYDAIRDAIQREHNMKHWPTWKVRLILATNPEWKDLFETLV
<GLX25_RS08590--Agromyces_luteolus--59.15 (SEQ ID NO:2264)
MPFMYMLRCADGTYVGVSTVDLDRRLAQHSAGEGADYTRRLPVELIYFEEYSRIDDAFTREKQVQVQWSRAKRLALAADDLDALGPASRKRWW
<FOB90_RS08540--Staphylococcus_fleurettii--54.55 (SEQ ID NO:2265)
MDKHYYIIVQCKDASLYTGYTNNAARIAKHNEGKGAKYTKSRPVLRYETTFETKSAALKRECAIKMTRQQKIQLIEE

FIG. 11SSS

87/106

<FLP08_RS06900--Gramella_aestuarii--67.44 (SEQ ID NO:2266)
MKTSYIYILSNKNRSVLYIGVTSSELIKRMTEHKNIGSDFCRKNLEDDLYFEETINKAIAAREKQKLNWHKEWKNLIKESNPDLEDLFQEL
<GN157_RS06505--Flavobacterium_rakeshii--60.92 (SEQ ID NO:2267)
MYKTLGTHNYIYILTNNKNTVLYVGFTNNLKERLYYHENPEAHKSFTSKYKCCYLIYWEHYNDVDIAITREKEIKAWRREKKEKLIQAQFNPEWNLNNQI
<GOY17_RS03670--Lysobacter_soli--83.70 (SEQ ID NO:2268)
MRRAACGLRNYATASCSSFPRRRESSGFASGTNVKQPCVYLLASGRNGTYAGVTSDLVKRVWEHRTHATEGFSHKYDATMLVWFELHETMESALTREKRIK
KWRREWKVELIEGMNPLYWRDLWEIHGCS
<GOY17_RS07710--Lysobacter_soli--77.42 (SEQ ID NO:2269)
MTSSEAPVKQPCVYILASDRNGTLYTGVTSGLIRRIQQHRSRVVEGFTRRHRVHSLVWFECDSMYAAISREKAIKEWRRRAWKLQIESTNPQWRDLYPDIL
<GOY04_RS02940--Sphingomonas_prati--74.42 (SEQ ID NO:2270)
MEKPGFVYIVTNKRAGTLYIGVTSDLVRRYIAHREALIDGYSKRHGCRILVWYQAFEMIEDARYREVQMKEWQRAWKIRLIEVMNPDWSDLYPSLL
<GOY20_RS11280--Sphingomonas_formosensis--71.59 (SEQ ID NO:2271)
MGRQPSVYIILASNRNGTLYIGVTSNLTARLHQHRTSDRRGFTARYDVKHVVFHFEIFEDMTCARAREKQKLNHWRREWKLTIEQENPDWQDLAVALGFEPPIAGD
<GOY20_RS07960--Sphingomonas_formosensis--72.41 (SEQ ID NO:2272)
MFGATLPRHSRQGLIKGFSSRYGSSRLVYFEMFQDMPNAILREKQHKWRRDWKLNLEGSNPSWADLAEGLVNGTYG
<E2H98_RS05310--Permianibacter_aggregans--76.67 (SEQ ID NO:2273)
MAKPGWVYIILASGRNGTLYIGVTDLPRLRVQHKKEKNVPGFTNRYDQVLFVWFESCEEISGAILREKQKLNHWRREWKLTIEQENPDWQDLAVALGFEPPIAGD
<GPL20_RS27755--Bradyrhizobium_cajani--75.27 (SEQ ID NO:2274)
MKQPCVYMFANRRNGTIYIGVTSNLPVRVYEHREGLADGFTRRYGVKTLVWYELHATMIDAITREKQIKGGSRADKLALIESVNPDPWDLTYDTLI
<GPL20_RS06895--Bradyrhizobium_cajani--78.65 (SEQ ID NO:2275)
MSGRCYVYIILASQIGGTLYIGVTDNLIRRVQHKSKLIESFTSEKYKVSRLVYFEQFDDPENAIKREKRLKKWNRTWKIRLIEQHNPNWDDLHYEIAAGP
<GPL20_RS36120--Bradyrhizobium_cajani--76.67 (SEQ ID NO:2276)
MTYVWVYILASKPGGTLYIGVTSNLPVRVYEHREGLADGFTRRYGVKTLVWYELHATMIDAITREKQIKGGSRADKLALIESVNPDPWDLTYDTLI
<POF45_RS14065--Lysobacter_panacisoli--80.90 (SEQ ID NO:2277)
MEKQPAVYMLSSRNGTIYVGVTSNLTGRVQHREHLVDGFTRRHDVAVLWYELWDDMAAIIQREKQKKNRAWKIRLIEQENPEWSDLWPLGLASI
<GH266_RS22020--Stappia_indica--74.16 (SEQ ID NO:2278)
MRDHNLYVYIILASGVGGTLYTGVTDVARRVSEHRDKVARSFTRHGVTHLVYFEHHTDVEAAIRREKQIKRWKRAWKIRHIEEHNPDWRDLFHGLP
<Mpaci_RS01715--Marinicauda_pacifica--76.67 (SEQ ID NO:2279)
MDRYSVYILDRPRGVLYTGVTSDLIRRAAEHRDGVVGGFTRRYNAKILVWYEMHADVEAAILREKRIKRWRRRAWKLELVEAMNPGRWDLWFDVGTGQGHAPG
<Mpaci_RS03650--Marinicauda_pacifica--71.11 (SEQ ID NO:2280)
MYRAFVYILTNKPHGVLYTGMCNDTLIRRIWQHRTKAFRGFTAKYNCTRLVYEAHDHVTLAAERERRMKRWHRRAWKIDLIESMNPDPWRDLYEDLTV
<GR138_RS02905--Shinella_kummerowiae--70.92 (SEQ ID NO:2281)
MAGYVYITTNQKHGTLYIGVTSDLERRIHEHREGRTPGFSSTYGC SRLVWYEEHDSIGTAIQREKSLKRWYRQWKIALIEAMNPDPWRDLYLDLW
<GRI47_RS09230--Erythrobacter_pelagi--77.78 (SEQ ID NO:2282)
MQRHFQPCVYIMASSRNGTIYVGVTSNLTQRAAFQHREGVVDGFTKRYGCKLLVWFELHATMEQAILREKQIKGGSRSKLLALIESRNPQWKDLFEQVCA
<GRI47_RS02765--Erythrobacter_pelagi--62.92 (SEQ ID NO:2283)
MDRARHGGWVYILANRYRGKLYVGVTSDLAARTAAHRVGHGSEYCAKHEIDRLVWAEFTFIDVAIAHEKRLKRWREWKIALVEKANPEWRDLFDDLNQ
<GRI62_RS01250--Erythrobacter_arachoides--72.73 (SEQ ID NO:2284)
MRQERQPCVYILVNGAYGTYIGVTSDLIARVWQHRTDAVPGFTSRYRVKTLVRYELLATMDEAIAAREKQKLNHWRWKNINLIESDNPDPWDLAPGLGLPPLA
PHQMRDGS
<GRI32_RS05980--Altererythrobacter_aestuarii--80.65 (SEQ ID NO:2285)
MQREFQPAVYILASARNGTLYLGVTSNLTQRIAQHRAGVEGFAKRHGCKMLVWFELHATMDHAI TREKQKLNHWRWKNINLIESDNPDPWDLAPGLGLPPLA
<GRI32_RS10365--Altererythrobacter_aestuarii--70.11 (SEQ ID NO:2286)
MQFERNPCAYILASGHYGTLYIGVTSDLMLGRHLQHRS GTTGFTGRYKIHRLVHFDMFDTMEAAIAAREKQKLNHWRWKNINLIESDNPDPWDLAPGLGLPPLA
QRAPGDGS
<GRI89_RS05470--Altererythrobacter_salegens--75.53 (SEQ ID NO:2287)
MLRYARNHERVFFVARSEATKQSRVAKREFQPAVYIVANRRNGAIYTGVTSDLVQHIWQHREGVADGFTKQYGCKMLVWYESHSTMEYAILREKQIKGGPRRR
KLELIEQANLKWRLYAEICA
<GRI89_RS05780--Altererythrobacter_salegens--67.05 (SEQ ID NO:2288)
MRQERQPCVYLLTNGLYGAIYIGVTSDDLGRVQHREGLISGFTSRFKIHRLVHFHFEIDTMESAIAREKQKFNWRRQWKINPINSNPNQWVDLAVGFLPPLA
RSGRTNGS
<GRI35_RS00350--Altererythrobacter_aestiae--78.49 (SEQ ID NO:2289)
MAREINPHVYLLASKRNGTLYLGVTSDLLARIHQHRTGTIEGFSKEHNVMKLVWFELHATMDHAI TREKQKLNHWRWKNINLIESDNPDPWDLAPGLGLPPLA
<GRI35_RS03810--Altererythrobacter_aestiae--72.73 (SEQ ID NO:2290)
MQFDKQPCVYILASQPHGTLYIGVTSDLITRLGQHRDGAASGFVSRYGVKQLMRFMFDDMQMAISREKQKLNHWRWKNINLIQSENPDWDLAPGLGLPPLA
PRDPNGS
<GRI35_RS10400--Altererythrobacter_aestiae--64.44 (SEQ ID NO:2291)
MSKGRWVYIVTNRYRGTIYIGVTSLSQRTMQHRDGTGSKFAGKYGCDRLVFAEHCDSEDAIAAREKQIKKNWREWKIRLIEQENPDWEDRFHDLIDV
<GRI72_RS13770--Altererythrobacter_marinus--73.56 (SEQ ID NO:2292)
MLNLFQHPSPGTLRMPFPERNPCVYILASGRYGTLYIGVTSGLIQRLYQHREGIAGGFTKRYRVHRLVRYEMFGDMEGAILRERQLERWHRQWKINLIERDNP
HWADLAVGLGFEPALRGGNGS
<GRI72_RS13885--Altererythrobacter_marinus--66.29 (SEQ ID NO:2293)
MTDEPYCWVYILASQKRGTYIGLTSRLAERIEEHRNGRVPFGFTQYGVRRPVHLEFFGDIHEAIAAREKALKKWRRAWKIALIEAGNPDPWRDLWHDLNR
<GRI68_RS02760--Altererythrobacter_halimionae--76.92 (SEQ ID NO:2294)
MPFERTPCVYILASKRNGTSYIGVTSDLPGRIHQHRS GTFEGFSQRHVDKTLVWFELHATMETAIQREKRLKKWERDWRRLVEDGPNPEWRDLAEDVGFNNLD
<GRI68_RS12605--Altererythrobacter_halimionae--66.29 (SEQ ID NO:2295)
MRTFETPVYILASHDRGRLYTGVTSDDLARLHRHRELFDDGYTRDREIKRLVWFERHDEIEIAIRREKTIKRWPRQWKFNVEAENPEWRDLAEAFGFPLSQ
K
<GRI38_RS05990--Altererythrobacter_aurantiacus--75.56 (SEQ ID NO:2296)
MEKRGVYIIMASKRNGTLYIGVTSNLAARVWQHKQGDIEGFATKYGCKMLVWFELHATMDHAI TREKQKMKWRRVWKLRIEELNPDPWDDLFEFESIC

FIG. 11TTT

<GRI58_RS09020--*Porphyrobacter_algicida*--62.77 (SEQ ID NO:2297)
MSAWTYILTNKPHGVLYIGVTDALISARMAQHRAQGQSAFCRRYGLDRLVLINHASITGAIAAREKAMKAWKRAWKIELIEANNCWEWDLFDVVLAA
<GRI59_RS05460--*Erythromicrobium_amosum*--72.83 (SEQ ID NO:2298)
MRQDRQPCVYTLASERHGTLYIGVTSDLIARLWQHHRNGTIEGFTRRYRVHRLVRYELFGDMERAILREKQLKRWHRQWKINLIESENPDWHDLEVELGLPQLN
HL
<GRI59_RS00335--*Erythromicrobium_amosum*--68.54 (SEQ ID NO:2299)
MERGGYVYIMTNRRYGVLYTGVTSDLPARITQHREGKGGFTTSRYKCTRLVVDHYPTIENAIAREKAIKAWKRSWKIEAIERQNPTWDDLFLTLNA
<GRI59_RS08370--*Erythromicrobium_amosum*--62.65 (SEQ ID NO:2300)
MANRYRGMYIGVTTADAIRRVHQHREGSGQHVHDFGKRRLLVYVEHETIEGAIAREKLVKKWRREWKFALIEADNPWDWDLWDQWAFAPRAPED
<GRI59_RS13710--*Erythromicrobium_amosum*--63.53 (SEQ ID NO:2301)
MQRETRGGWYIMTNRYRGTYLVGSTDLAQRMVYEHREGRGSEFCDEHGLTRLVWAERASIVDAKHHERRLKRWHRWKFALIEKENPDWDDLFGQLN
<GRI75_RS03065--*Altererythrobacter_soli*--72.41 (SEQ ID NO:2302)
MRERDPFCVYILASGHYGTLYTGVTSNLLQRLVQHREGLIPGFTSQYKVHSLVHFEMFGDMERAITREKQIKNWHRQWKINLINAENPEWRDLAVGFGLPPLC
PPGGRDGS
<GQA99_RS10315--*Shewanella_litoralis*--77.27 (SEQ ID NO:2303)
MDKQSYIYIMANRPNGLTYIGVTSDLIRRNWEHKNGLADGFTKEHSIKMLVYVEFDDIYNATREKQLKNWKREWKIKLIEKFNPRWEDLEI
<GO003_RS25845--*Methylicorpusculum_oleiharenae*--81.52 (SEQ ID NO:2304)
MTKQPCVYLLASQRNGTLYIGVTSDLIDRVYQHRKELAEFGTKRYKIHDLVWYEPHQTMSAIREKQLKKWDRKAKIGLIEQNPHWKDLWPELSIP
<GOQ17_RS00585--*Sphingobacterium_composti_Ten_et_al._2007_non_Yoo_et_al._2007*--68.97 (SEQ ID NO:2305)
MERGGVYFMTNKNAILYVGVTSNLRERVQHKKKFYPNSTAKYNCTKIVFYLFPSIEEAINEEKRIKAGSRQQKINLIEKMNLLWEDLYETLDD
<GOQ16_RS03905--*Sphingobacterium_bovisgrunniensis*--60.00 (SEQ ID NO:2306)
MIHNYFVYILTNANRSVLTYIGVTNDIQTRLVYEHYIDSQTTKRHFTGKYNVYLYIYERYHNINDAIDREKELKGWRRSKEELINVFNPISWIFLNDV
<ABD07_RS02670--*Nitrosomonas_oligotropha*--79.57 (SEQ ID NO:2307)
MTEKQPAVYILASKRNGTLYIGVTSNLIIRRVWEHKNLADGFTKKYQVHQLVWYELHETMSAISREKAMKNWKRAWKISRIETVNPDKDLYQSL
<IMCC3317_RS10580--*Kordia_antarctica*--73.63 (SEQ ID NO:2308)
METYSVYIITNPKPGVLYIGVTKDLKRRIQHKKNVHPTTFSARYNLNLYVFENHEEKESALLREKRIKKWNRDWKVELIEKTNPNWEDLISEIK
<IMCC3317_RS00550--*Kordia_antarctica*--64.77 (SEQ ID NO:2309)
MNHPIYVFLTNKNNTVIYIGVTSNLTNRVYQHKAKIYKGTSTKYNCGLVYFEFQSSINEAIAAREKQLKAGNRKRKEELININENPEWDLSDGWLIFYFD
<IMCC3317_RS20030--*Kordia_antarctica*--63.53 (SEQ ID NO:2310)
MDTHYVYILTNKSHVTLVVGSRKQLKRLIQHKNNSTSTFTGKYNVHKLIVYFETTKYVNNISIKRERQIKKWKREWKINLINGLNPDKDLSEYI
<GTQ45_RS00590--*Pyruvaticibacter_mobilis*--71.26 (SEQ ID NO:2311)
MASGRNGTLYIGVNDIARRAHEHRTGLVEGFTROYGVHRLVHVEAFADVREAIRREKRLKDWREWKALALIEHNPRWRDLWDDMLS
<GTQ45_RS11795--*Pyruvaticibacter_mobilis*--65.96 (SEQ ID NO:2312)
MAFFVYILTNKKGHTLYVGHDTDLALRLTEHRSWAVPGFTKAQGLGLVYFEQHDREAAARKRERRLKHWLRAWKIQLEDVNPWEWDLSDALL
<GTQ45_RS10755--*Pyruvaticibacter_mobilis*--64.71 (SEQ ID NO:2313)
MAFFVYILTNKPGHTLYVGVTDLARRLTEHQSGAPGFSARYNLRLVHFEAFDSREEARQREERRLKHWIRDWKVSLEADNPWDWDLTGRVPHF
<GWI72_RS09565--*Microvirga_tunisiensis*--71.91 (SEQ ID NO:2314)
MSAYVYILASGRNGTLYIGVTTSLPQRIHQHRQGYTPGFASRYRVNLLVYABAEHIDEAIAAREKRLKTWKRWSWKIDLIEASNPEWQDLASAALL
<GWI72_RS01905--*Microvirga_tunisiensis*--73.12 (SEQ ID NO:2315)
MTYHVYILASRMHGTLYVGVTDLARRVDEHRRGLGSRFTREHGVTRLVVFYEGFEDVEAAIAAEKRLKRWHRWAKVQLVERNPLWKDLYLELNM
<SPSYN_RS04830--*Sporotomaculum_syntrophicum*--59.74 (SEQ ID NO:2316)
MYVYVYILICADGTYTGYTTDLVRLHEHNRGQGSKYTRARLPVELVYEVLPDKSSSLKREIAIKKLSRREKIRLVKDCQSN
<GZN30_RS01970--*Vibrio_ponticus*--70.65 (SEQ ID NO:2317)
MKQPAVYILSNHTRAIYGVTTNLKARVWQHKSGVCEGFTKRYRINRLVYELFSDIRDAIAAREKQIKKNWKRAWKNLIEKVNPSWKDLYSDI
<CSC73_RS17435--*Pseudoxanthomonas_sacheonensis*--86.96 (SEQ ID NO:2318)
MEKQPCVYLLASSRNGTLYAGVTSNLIIGRTWQHREHVVEGFSKRYEVTCLVWYELHPTMESAIQREKRIKKWNWAKVRLIDESNPSWRDLWPDITGQVPPKAN
VHGFPSPRE
<CSC74_RS16560--*Pseudoxanthomonas_yeongjuensis*--85.87 (SEQ ID NO:2319)
MGKQPCVYLLASSRNGTLYAGVTSNLIIGRTWQHREHVAEGFSKRYEVIRLVWYELHPTMESAIQREKRIKKWNWAKVRLIDESNPSWRDLWPDITGQVPPKAN
VLGFPSPRE
<CSC74_RS04380--*Pseudoxanthomonas_yeongjuensis*--85.23 (SEQ ID NO:2320)
MLATGKRGTLYIGVTSDLVARTWQHREHIVEGFTRRYNVTMLAWYELHGTMESAILREKQLKKWNREWKRLVQEFNPEWRDLWDDIVG
<CSC74_RS02340--*Pseudoxanthomonas_yeongjuensis*--73.91 (SEQ ID NO:2321)
MQRERIPCVYILASGIRGTLYVGVTSNLPGRVWQHKSDRAEGFARKYRIHTLVWYEVHGTMESAILREKSLKAWKRAWKLELIETSNPEWRDLYIEIL
<CSC78_RS08200--*Pseudoxanthomonas_japonensis*--82.80 (SEQ ID NO:2322)
MEKQPATYMLASGRNGTLYVGVSSDLIGRIWQHRTHAVPGFTDRYRATRLVWYELHPTMEAATREKRIKKWNWAKVRLIDESNPSWRDLWPDITGQVPPKAN
<CSC78_RS15030--*Pseudoxanthomonas_japonensis*--74.73 (SEQ ID NO:2323)
MPRERIPCVYILASQPRGTLYIGVTSNLPARIWQHRNDQADGFTRRYGVHTLVWYEVHQMESAILEKTLKAWKRLWKIQLVEDANPAWRDLYPDIT
<DT603_RS07480--*Pseudoxanthomonas_gei*--89.25 (SEQ ID NO:2324)
MKQPAVYILASDRNGTLYAGVTSNLIARTWQHREHLDVGTFRYDVTMLVWYELHATMDTAITREKQIKKNWRAWKRLRIQEANPRWRDLWPDITGQVPPKAN
<DT603_RS04145--*Pseudoxanthomonas_gei*--84.78 (SEQ ID NO:2325)
MDKQPATYILASERNGTLYIGVTSNLIARVWQHREHVEGFTSKYEVSVLVWYEFHSAMEEAILREKRIKKWNWAKVRLIEEKNPYWNLWNEIAG
<DT603_RS15385--*Pseudoxanthomonas_gei*--70.97 (SEQ ID NO:2326)
MSRKQPCVYLLASGRNGTLYAGVTANLLARISQHREGLTEGFTKKYRVHTLVWFEFHSMTAEAIAREKSIKEWRIRKIELIEACNPHWLDLYPSLV
<CSC70_RS09940--*Pseudoxanthomonas_kalamensis_DSM_18571*--84.78 (SEQ ID NO:2327)
MKQPAVYMLASKRNGTLYIGVTSNLIARVWQHREHVEGFTRRYGVHTLVWYEQHVMEEAIRREKQLKKWNWAKVRLIESNPSWRDLWDDIVG
<CSC70_RS07905--*Pseudoxanthomonas_kalamensis_DSM_18571*--81.72 (SEQ ID NO:2328)
MKKPCVYILASRYHGTLYIGVTTDLRRVWEHKNDCVEGFTSKYQVHDLVWYEMHATLDTAIPREKAIKEWKRDWKIELIEKQNPWRDLYPEMM
<GTH33_RS06215--*Sphingomonas_insulae*--79.55 (SEQ ID NO:2329)
MDREPCVYILASRPHGTLYVGVTSNLPGLVQHREGLIPGFTSRYGVHRLVWYDVADTMEAVIVREKQLKRWNRDWKRLRIEERNPAWDDLAIIGLGLPPAPVK
ESNT

<GTH33_RS03975--*Sphingomonas insulae*--73.12 (SEQ ID NO:2330)
MRPAIYMMANGRNGTIYVGVTSNLTTRRAWEHREGSSDGTFRYCKLLVWFEEAADTMEAAILREKQLKAGSRAKKITLIEAMNPKWHDLYEQIAHG
<GTH33_RS15955--*Sphingomonas insulae*--69.88 (SEQ ID NO:2331)
MLASGVNGTIYTGVTNSNLIGRIVQHRDGTFFGFTSRYGVTTLVWFDMDATMEVAIASEKRIKAWRRDWKKNLIERDNPWGADLVALGLPPLD
<GT349_RS21725--*Rhizobium daejeonense*--75.00 (SEQ ID NO:2332)
MRQGYVYILASRRNGTLYTGVTSDLARRLYEHKHDLTGFTSRYGVTTLVWFEEVHDLTISALTREKTIKKWPRAWKLNLIESVNPDWDDISHYLL
<GT360_RS13325--*Vibrio astriarenae*--69.23 (SEQ ID NO:2333)
MEKPPFVYMMASANKALYIGVTNLTQGRVWQHKGNGVVDGFSKKYQTHRLVFEEQYQDIIAAIAAREKQLKNWHRQWKITLLSNETRNGWIYRSAFKAAASIAH
PP
<GXM17_RS18600--*Flavobacterium fluviatile*--65.26 (SEQ ID NO:2334)
MKPGFYIITNKYQTVVYTGVTNSNLPQRILEHREKKYPKSFARSYNNVILVYEQFQWIEDAIAREKQIKAGSRETKNDLIRSINPSWKDLFEEIENIMIE
<GXM17_RS19850--*Flavobacterium fluviatile*--67.78 (SEQ ID NO:2335)
MLNTQTGFHSYYVYITNEHRTSFYIGVTNNLKLRLQKLAHKENIDLNVSSFAAKYNIQFLVYEEKFTWIEAIAREKELKKWRDCKIELIKSSNEIFEFLNFH
FE
<GXM17_RS08950--*Flavobacterium fluviatile*--64.77 (SEQ ID NO:2336)
MLNPQEGFHTYVYIITNKNRSTFYIGVTNNLKLRLQKHKNLEIENKTFASKYIGIEFLVYEEKFVWIEEAIAREKEIKKWRDCKELSLIQIFNPEFEFLNYH
FE
<GXM17_RS20470--*Flavobacterium fluviatile*--63.16 (SEQ ID NO:2337)
MFNTQEGFHSYYVYITNSYRSTFYIGVTNNLKLRLQKHKNLEIENKTFASKYGLKFLIYEEKFVWIEEAIAREKELKKWRDCKIELIKSSNEIFEFLNFH
FDSI
<GXM17_RS17130--*Flavobacterium fluviatile*--65.06 (SEQ ID NO:2338)
MKLSEGYHTYWIYIITNKSCTVLYTGVTNNLKLRLQKHKNLEIENKTFASKYGLKFLIYEEKFVWIEEAIAREKELKKWRDCKIELIKSSNEIFEFLNFH
N
<GXM16_RS04345--*Flavobacterium limi*--65.26 (SEQ ID NO:2339)
MFNPQEGFHSYYVYIITNKNRSTFYIGVTNNLKLRLQKHKNLEIENKTFASKYGLKFLIYEEKFVWIEEAIAREKELKKWRDCKIELIKSSNEIFEFLNFH
FDSI
<GWK10_RS07850--*Spongiivirga citrea*--59.09 (SEQ ID NO:2340)
MKESYIYVLTKNYRTTFYIGVTGNLKKRLVEHTSGTSGKFTGKYNVVDLIYFEVFTDINQAIAREKQLKNWRKEWKNLIRKVNPTLATLEY
<GWK08_RS08995--*Leptobacterium flavescens*--64.71 (SEQ ID NO:2341)
MRKYHIYIITNKNNTVLYTGVTNSLTKRIYQHKVKAYKSFSSRYNCNLVYFEEHNSMRAAIAAREKQIKAGSRRKKEQLINSINPEWNDLSDGWLIFYFS
<G3008_RS13645--*Cryomorpha ignava*--64.29 (SEQ ID NO:2342)
MPKTGYTYVLSNKNRTVLVGVTSRIRLRMLQHKSGHGVFSSRYNVYDLMFETIPIVNAIKREKQLKNWHEKWKWNFIKENNPDLDLACECLPMRKYS
IRRRRWIRRKGRSD
<G6N06_RS13440--*Gelidibacter japonicus*--70.33 (SEQ ID NO:2343)
MEKNIWTSVYIITNKPNGVLYIGESKRLKKRIYQHKNVHPSTFSARYNNLKLVYFENYDTESDAKLREVQLKKWNRRAWKIELIEKLNPNWLDLYDNL
<G6N06_RS13420--*Gelidibacter japonicus*--71.08 (SEQ ID NO:2344)
MSNKPQGVYIYGVTTNNIDQRVKEHKLKIYQESFTAKYNCNDRVLYFEIYDNDKEALIRERRMKKWKRDWKIQLIESMNPSSWMDLSMNWNLDFNNYRNERK
<FNJ47_RS36995--*Bradyrhizobium uaiense*--76.40 (SEQ ID NO:2345)
MGARTYVYVILASHIGGMLYIGVTNDLIRRVGEHRLKLIKSFTEKYDVVKLVYFEQFDDPENAIKREKRLKKWNRRAWKARLIEENNPWEDLYPGIAGV
<FNJ47_RS32515--*Bradyrhizobium uaiense*--76.34 (SEQ ID NO:2346)
MAYYVYLLANEKYGTLYLGVTTNDIVRRVYEHKSKIVAGFTKQYDVRVWFEIYDDALTAIAREKELKRWRDCKIELIEKNPQWIDLYPEIAAGHIG
<FNJ47_RS16425--*Bradyrhizobium uaiense*--63.10 (SEQ ID NO:2347)
MSTAGSYVYVILASRHGTYIYGVTTNDIRLRLLELHRSGRGSKFVYKYRVFLVHIEEFATPQEAIAAREKQLKFWKRDWKIKLIEADNPWDNDRSDLI
<G3M57_RS20830--*Caulobacter rhizosphaerae*--68.13 (SEQ ID NO:2348)
MEERYPWIAVYILSDHYRGTYLIGVTSDLISRVIKHREGVDFGTQYGLHRLVWYETFGSMTETAIQREKTLKHWRDCKINLIERDNPWNLDLFDPIVNWTP
VPRQV
<G3M57_RS05005--*Caulobacter rhizosphaerae*--65.52 (SEQ ID NO:2349)
MVGREPWIAVHIMADHYRGTYLIGVTSAPFTRMIQHREGAFPGFTRRYGLRLVWFELHDLMTETAIQREKSLKEWRRQWKINLIERDNPWDDLYSATFEWTP
VPRQV
<G3M57_RS05175--*Caulobacter rhizosphaerae*--75.28 (SEQ ID NO:2350)
MGFYTYIVASRRNGTLYGVSTDDLVRRIIEHREYRRDGTAKHGCTILVWYEVFESREETFRRERRRIKEWRRSWKMLIEAENPTWRDLALDLPNA
<G3M57_RS12705--*Caulobacter rhizosphaerae*--65.52 (SEQ ID NO:2351)
MVAADAWIAVYMTDHYRGTYLIGVTSANFQSRRIQHREGVGGFTRRYDLKRLVWYIEIHDRMTDAIQRETSIKRWRRQWKINLIERDNPWDDLYEAVFNWTPV
PRVI
<G3M57_RS03295--*Caulobacter rhizosphaerae*--63.41 (SEQ ID NO:2352)
MASKPYGTLYTGSTNDIWRRAEHRAMENPCFTQKYGVTRLVWVEQHENRIDAFRRERRIKEWRRRAWKIRLIEEMNPGWVDDLHLGESFSF
<G4223_RS16465--*Magnetospirillum aberrantis* SpK--73.26 (SEQ ID NO:2353)
MQGGWVYIMTNRPNGLTYGVTSDLARRAWEHREGVADGFTKRYGLKRLVYAERYEDIRTAIQREKNLKHWPRAWKVKLILADNPNWDDLYERLA
<G6N74_RS29860--*Mesorhizobium zhangyense*--72.22 (SEQ ID NO:2354)
MSGYVYILASRKNGTLYTGVTNSLERRVWEHREGLTDGFTKKYGVTRLVWYRDYVDIGDAIMDEKRIKKWRRKWKLELIEAMNPEWDDLYEVLNG
<G6N73_RS12390--*Mesorhizobium camelthorni*--61.25 (SEQ ID NO:2355)
MEVTVYILRCSGSGYYTGLTKQIEARVWEHNEGIYDGYTRKRPPVVLVFTETDYDRIVDAIERERQIKGWSRAKKEALIALDYEGLPELARNRQGGYVRRWP
PVAG
<G6N80_RS17865--*Rhizobium rhizoryzae*--69.77 (SEQ ID NO:2356)
MVGFAYMMSDKPRGVIYTGVTSDLAGRAWQHRNEITKGFTSKYNARLLVWFERHPNIVLAIQREKSLKRYKREWKIKLIEGFNPTWLDLYEALDQRDNPYMPH
PNTREWHYHN
<G3569_RS05570--*Alifodiniobius halophilus*--67.42 (SEQ ID NO:2357)
MGYKNSGKCGYVYIITNKSXSVLYIGVSSSLDKRIIEHKGQGSKFTIKYNCITILLYEYKYPNINEAIAAREKQLKNWHRSWKLNLIKSNKPTLNNLWKNLS
<G3569_RS12765--*Alifodiniobius halophilus*--66.67 (SEQ ID NO:2358)
MRFQNIYVTRQGYVYILSNKTRRVLYIGVTGNLVKRVWLHRQGLGSEFTQKYRTKYLIHYESFKHIERAINREKQLKNWHRWKNLIREHNPFRDLWPKILL
RRNYF

FIG. 11VVV

<G3569_RS17900--*Aliifodinibius halophilus*--65.88 (SEQ ID NO:2359)
MARYYYIMSNISKMIYTGMTNSLTRRVREHKEKVNNEGFTKRYNIHRLVYEEFQDVREAIAREKEIKGWRREKKVALIEKVNPKWKDLAANE
<G3569_RS12615--*Aliifodinibius halophilus*--72.15 (SEQ ID NO:2360)
MAKKYHVYIMSNNAHTVLYTGMTSDIIRGWQYKQQLIDGFSKRYNCTKMVIIEYNHPQDAIDREKQIKAGSRKKKLELIEELNPN
<GYM46_RS05565--*Brevundimonas mediterranea*--69.77 (SEQ ID NO:2361)
MPTHRFTIATYIMASGRNGTLYVGMTANLTARVWKHRTGAFDGF GKAYGCTQLVWFEQHAWVVEAIRREKRIKKWLRDWKLQLIEAANPDWRDLAEGWYPVHD
PNWTPPVEDEGFPGSASQPEG
<PL9631_RS00130--*Planktothrix paucivesiculata*_PCC_9631--69.32 (SEQ ID NO:2362)
MKNYYVYIMTNKSKTLYTGVTNDLTRRVYEHKHKLIPGFTQKYNI TKLVYEEETSDVHAIAAREKQIKGWLEKKINLIESMNPWNVDLSLNL
<AGR9A_RS15030--*Agrobacterium salinitolerans*_str._Hayward_0363--71.91 (SEQ ID NO:2363)
MKGYVYILASKRNGTLYTGVTKNLPNRVFEHQNELTPGFTSRYGVKMLVWFEEYDLLTSAITREKTIKKWPRQWKLNLIESMNPENWDISSHLHLGL
<AGR5A_RS03185--*Agrobacterium genomosp.*_5_str._CFBP_6626--76.67 (SEQ ID NO:2364)
MTGFVYIVASGRHGTYLIGVTSNLEQRIYEHREGLTKGFTSIYGCTRLVWVEEHMSIGTAIQREKSLKRWRDWWKIKLIEDFNPEWRDLYGEFG
<AGR5A_RS15535--*Agrobacterium genomosp.*_5_str._CFBP_6626--71.91 (SEQ ID NO:2365)
MKGYVYILASKRNGTLYTGVTDLPRLEHQNELTPGFTSQYGVKMLVWFEEYDLLTSAITREKTIKKWPRQWKLNLIERTNREWRDISSHLHDL
<AGR5A_RS15770--*Agrobacterium genomosp.*_5_str._CFBP_6626--72.94 (SEQ ID NO:2366)
MTFAHAFLMREGYVYILASKRNGTLYTGLTSDLSHRLYEHQNELTPGFTTRYGVKTLVWSETYDLVTDIAIAREKTIKNWPRAWKIKLIEGLNPAWDDIAHFL
<AGR13A_RS23060--*Agrobacterium genomosp.*_13_str._CFBP_6927--75.29 (SEQ ID NO:2367)
MRQGYVYILASKRNGTLYTGVTSDLSHRLYEHQNELTPGFTTRYGVKTLVWSETYDLVTDIAIAREKTIKNWPRAWKIKLIEDLNPAWDDIAHFL
<AGR7B_RS22070--*Agrobacterium deltaense*_RV3--73.86 (SEQ ID NO:2368)
MRQGYVYILASKRNGTLYTGVTSDLSHRLYEHQNELTPGFTTRYGVKTLVWSETYDLVTDIAIAREKTIKNWPRAWKIKLIEDLNPAWDDIAHFL
<BN3794_RS04110--*Xanthomonas massiliensis*--82.80 (SEQ ID NO:2369)
MEKQPATYIILASKRNGTLYTGVTSDLSHRLYEHQNELTPGFTTRYGVKTLVWSETYDLVTDIAIAREKTIKNWPRAWKIKLIEDLNPAWDDIAHFL
<C7S05_RS11740--*Neisseria gonorrhoeae*--98.94 (SEQ ID NO:2370)
MQPAVYILASQRNGTLYIGVTSDLVQRIYQHREHLIEGFTSRYNVTMPVWYELHPTMESAITREKQLKKWNRRAWKLQIEENNVSWRDWFDII
<VTH8203_RS17275--*Vibrio thalassae*--71.74 (SEQ ID NO:2371)
MTKQPCVYIITASKSGVYIIGVTSNLLQRIWKHNRNGVTDGFEVKKYVYHLVHYELFDCMEDAIAREKQIKHWNRSWKNELVTSHPYWRDLTPDLW
<VAT7223_RS05100--*Vibrio atlanticus*--72.04 (SEQ ID NO:2372)
MKQPAIYILANRSNSVLYIGVTSNLLQRIWKHNRNGVTDGFEVKKYVYHLVHYELFDCMEDAIAREKQIKHWNRSWKNELVTSHPYWRDLTPDLW
<GA0061102_RS22230--*Rhizobium miluonense*--70.97 (SEQ ID NO:2373)
MGGYTYIITNHKRGTYLIGVTDLERRIYEHREALTPGFADWKYCTRLVWYEDHADIVSAIQREKSLKRWRQWKIEELIEKANPDWRDLYYELW
<GA0061099_RS04040--*Bradyrhizobium yuanmingense*--72.04 (SEQ ID NO:2374)
MKQPCVYMFANRRNGTIYTGVTSNPPREHREGLVKGFSSKYGCKLLVWYELHATMSDAIVREKQIKGSRRAKKLALIESLNPNWDDLYATLI
<GA0061103_RS1590--*Rhizobium multihospitium*--70.97 (SEQ ID NO:2375)
MGGYTYIITNHKRGTYLIGVTDLERRIYEHREALTPGFADWKYCTRLVWYEDHADITSATQREKSLKRWRQWKIEELIEKTNPDWRDLYYELW
<GA0061098_RS07005--*Bradyrhizobium shewense*--76.40 (SEQ ID NO:2376)
MSPRCYVYVILASKIDGTLYIGITNDLIRRAVQHKLKLIESFTKDYGVVRLVYFEQFDDPENAIRREKRLKKGNRWTKVRLIEQHNPWNWDDLYHEIAALHEAV
VPREGASSTPRRV
<GA0061098_RS07030--*Bradyrhizobium shewense*--73.33 (SEQ ID NO:2377)
MHGGWYVILASAPNGILVVGVTSDIVRRIYQHRNGFVPGFTTKYGVKLLVYFERYEDIQTAIQREHNIKHWSRTWKVRLILRDNPEWKDLYDTIVQ
<BLR30_RS01730--*Chryseobacterium taeanense*--62.11 (SEQ ID NO:2378)
MKPGFIYIMTNINNTVLYTGVTSNLPKRINEHKEFFELSFSTSKYNKNKLIYVESFQEIQDAIFREKQIKAGSRQKKIDLINSINPEWRDLAEDIKDIMNPF
<BLP81_RS12660--*Psychroflexus sediminis*--66.67 (SEQ ID NO:2379)
MPKRSHHNYTVYIILTNKRTGLYIGVTTGGIDDRMERHLKGEQSFTSKYKLRVLYFEFQYIEDAIAREKQILKNWHRQWKINLIEQDNPEWKDLYKPLGKDP
DLDPETSLPACRSQAGSG
<BLP81_RS02785--*Psychroflexus sediminis*--59.77 (SEQ ID NO:2380)
MKTGLGTHNYVYIILTNKNKTALYTGVTNDLQRLYFHKNPALSKAFTTKYKCFYLIYFEHFQDVSQAIAREKQIKGYRRSKKEALIDGFNPEWEFLNDKI
<BLP81_RS02755--*Psychroflexus sediminis*--60.92 (SEQ ID NO:2381)
MKTGLGTHNYVYIILTNKNKTALYTGVTNNLQRLYFHKNPALSKAFTTKYKCFYLIYFEHFQDVSQAIAREKQIKGYRRSKKEALIDGFNPEWEFLNDKI
<BLS90_RS13910--*Flavobacterium omnivorum*--62.11 (SEQ ID NO:2382)
MKPGFIYIITNKYQTVVYTGVTSNLPQRILEHKEKRYPASFSARYNANLLVYEQFCIGDAIAREKQIKAGSREAKNELIRTNLPTWNLDFEIKDIMTE
<BLS90_RS09255--*Flavobacterium omnivorum*--58.62 (SEQ ID NO:2383)
MKFYVYVYIILCNDSDSLVVGVTSDIDRRVIEHNAGKYPDAYTHSRRPVTLFYQDFDTPNQAIQFEKELKKWSRVKKQALIDGNFDKQLQDLSECRNATNHKYKS
DFEDIEE
<BLT80_RS17565--*Winogradskyella thalassocola*--70.33 (SEQ ID NO:2384)
MKHYVYIITNKKDGLVYIGETKRLKKRIYQHKNVHPTTFSARYNLDLTVYFKEYNTEAEAKLREKQMKKWNRVWKIEELIEQFNPDWKDLYETI
<BLT80_RS01380--*Winogradskyella thalassocola*--62.50 (SEQ ID NO:2385)
MHKTLGTHNYVYIILTNKNKTALYTGVTNDLNNRVRFHRENAGINSKAFSTSKYKCFYLYVEHYNSVETAIQREKQIKGWKRFKKDIMISEFNCNWKFLNDSI
<BLS98_RS15175--*Chryseobacterium jejuense*--65.12 (SEQ ID NO:2386)
MKAGFIYIMTNKNHTTLYTGIIISNLLKRVQQHKEAYYSQSFTSRYKLYILVYWEAFQEIQDAIAREKQIKAGSRQKKLNINSINPEWDDLTNTIQDIMDFV
<BLT04_RS13405--*Chryseobacterium soldanellicola*--58.14 (SEQ ID NO:2387)
MIELNDGIYKIFYVYIILTNKSRRVLYTGVTNNLHLRLAQHKNNKINPNSFTAKYNTCFLIYYESFGWIQAIEREKEIKLLKRDNKFKLIRNFNPNLEFLNYLFQ
<BLR17_RS03015--*Flavobacterium glycinis*--66.28 (SEQ ID NO:2388)
MKLWFVYIMSNKHKGVYIIGVTDNLIERVKEHKLKVYSNSFTAKYNCCKLIYFEFEWENGFDATKREKQLKKWKRWDKVNLIEDMNPTWLDVSLNWNFNFNKIR
E
<BLR17_RS12770--*Flavobacterium glycinis*--66.32 (SEQ ID NO:2389)
MKPGFVYIITNKYQTVVYTGVTSNLPQRILEHKNKKYPTSFARSYDITSLVYEQFQWIGDAIAREKQIKAGSRASKNKLQISINPTCKDLFEEIRDLMTE
<BLR17_RS11045--*Flavobacterium glycinis*--56.96 (SEQ ID NO:2390)
MKFYVYVYILRCSNDSLYVGVTSDVERRLMEHNAGKYPEAYTHSRRPVELVYFQEFTEPNQAIQFEKELKKWSRVKKQALIDENYDELTRLSQCRNDTHYKNKD
EKEEKDL

FIG. 11WWW

<BLT26_RS02355--Microbulbifer_yueqingensis--79.27 (SEQ ID NO:2391)
MHRQPCVYIILASKNGTLYTGTSSSLAQRIWQHRDGLSEGFTHKYGVRLVWYELHSTMTSAITWEKQIKSWRRRAWKLMIERFNPA
<BLR31_RS00530--Pelagibacterium_luteolum--75.58 (SEQ ID NO:2392)
MGAQRYYYVYILASGIGGTLVGVGTNDLFRRI GEHRDKAVAGFTRAYDVERLVYVYETHESIDGAIEREKRLKRWRRGWKVQLIEEQNPWVDLFAGLRP
<BLS22_RS03735--Natronincola_ferrireducens--69.41 (SEQ ID NO:2393)
MYVYVYILTNRSNKVLYTGTNNLERRLYEHNKLIKGFTEKYNVKNLVYFGSTDDVEAALVREKQIKGWTRCKKINLIESINPEWKDLSQGFNY
<BLQ12_RS21025--Desulfoluna_spongiiphila--81.91 (SEQ ID NO:2394)
MTPCVYILASGKNGTLYIGVTSDLVQRIWQHKEFDVDFGFTSRHKIHALVWYEVHETMESAITREKRLKKWNRRAWKISLIEAMNPQWLDLYGSIL
<BLS44_RS01645--Niabella_drilacis--67.44 (SEQ ID NO:2395)
MKKGGFIYIILTNKLTVLYIGVTSNLRRRIWEHAHHENENSFTDQYNVTGLVYVEWFDVIVTAIEREKQLKRWRSRTKKVALIKVKNPEWRLNNEEVYDEVYSL
LY
<BLS44_RS19835--Niabella_drilacis--55.70 (SEQ ID NO:2396)
MKRYVYVYILKCDSDNSYYTGITSNLEQRLRQHESGYEGSYTSSRRPVTLVYVECFHNVRSAIAFEKQIKGWSRKKKEALMRSDWDTLKLSSRNYTQYGDPSDD
HPSSGSGGRERPV
<BLR37_RS12285--Pedobacter_soli--65.91 (SEQ ID NO:2397)
MERGGTVYIITNFTHTTLYTGTSDIMFRIKEHKDKKEYPKSFSKYNNTDKLVYFCSYDSIEEAI AAEKRIKGGSRKKIKLIESMNPWKDLWEIVKDW
<BLR37_RS11930--Pedobacter_soli--58.70 (SEQ ID NO:2398)
MDRNYFVYIILTNKNTVYVYTGVTNDLEVRILTQHRENDKDNKFAFTYKYNCYLVYFEHFQYIEHAIEREKEIKGWTRAKKNALIE TENKNWDFLNDTITEE
<BLT76_RS14045--Nocardioides_lianchengensis--58.57 (SEQ ID NO:2399)
MAYTYMLLCSGDSGFVVGSTRDLEHRLYEHQIGEGAA YTRHRLPVQLVWHEEHENVGAAFHREKQVQNWRSRVKRLALIRQDYAGLPALAKKDFSRRRDRSDESG
<BLQ71_RS07285--Mameliella_alba--66.29 (SEQ ID NO:2400)
MAHFVYIMAARPGGALYIGRSTNLRDRVAHRAGLSVHTARYGIHTLVWFEHAEFESSLRERRRLKRWRRRAWKDLITEFNPWQDMIHAIP
<BLT59_RS14930--Nucilaginibacter_pineti--64.77 (SEQ ID NO:2401)
MKVHEYVYVYITNKYNVLYVGVTSDLQSRINEHKLKIYKGFSSKYBELNKL VYEEFQWIQDAIAREKQLKAGSRQKKDELI IANNPLWNGWYDD
<BLP55_RS07530--Flavobacterium_caeni--55.17 (SEQ ID NO:2402)
MKTKVWVLAQKATVMPYVYILSNPRRTTLYINFTQDIDITCLAAHRAGTESPFTKKYILVHLIYFEPFEKIE TAQARARQLKNWHRRAWKWNLIMAA NPKLRAL
STTDAEPSSAQPYTL
<BLQ00_RS02915--Ulvibacter_litoralis--65.88 (SEQ ID NO:2403)
MKRSFYVIMNKYRTTFYIGVTANLQKRILEHIEGVGSKFTSKYNLKDLYFEEFSDIEQAIAREKQLKNWHKGWKNLNI RETNPKLETLDIF
<BLQ00_RS10285--Ulvibacter_litoralis--66.27 (SEQ ID NO:2404)
MANKPNGVLYIGCTHNI EERVKEHKLKIY PNSFTARYNCDKI IYFEEINAGDIASNREKQMKKWKREWKVKLIEEMNP SWMDLSLNWNLFNFKLRNE
<BLQ00_RS10270--Ulvibacter_litoralis--59.09 (SEQ ID NO:2405)
MPREHKYTYILTHENNHIFVYGVTTNMMKRRTAEHKASTHGTHVGRYNIKKLVYFEEHQWIEAVRREKTIKKWKREWKINI ITEMNPEWEDLSDDWDL SKFA
IENE
<BLP10_RS20585--Microvirga_guangxiensis--80.00 (SEQ ID NO:2406)
MAEAFVYIILATRDGPLYVGVTSDDLHKRVFEHKTTHAIPGFTARYNVDRVLVYFEVFNPEAAIAREKQLKKWRAWKVELIERENPEWRDLAEFFVP
<BLT30_RS13070--Chryseobacterium_hungaricum--55.43 (SEQ ID NO:2407)
MKTGLGTHNYVYIILTNKNTVLYTGTNDLKGRLFWHINPESHNTNHFYKYKCFLLIYEFQDQVDAIAREKQIKGYSRMKKENLINNINPNWDFWNDKIE
<BLP56_RS00725--Novosphingobium_stygium--75.00 (SEQ ID NO:2408)
MRDERQPCVYILASQPRGTLYIGVTSNLIGRLWQHRESMIPGFTSQYAVHRLVYFEQFDEMI PAITREKQLKRWHRQWKINLIESANPDWHD LAPGLGLGPPA
PRTPRDGP
<BLP56_RS01450--Novosphingobium_stygium--65.12 (SEQ ID NO:2409)
MEKGGWVYIMANRYRGGMYVGVTSDLVRRVWQHREGVGSSHVDDFGKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAAI ALR
GVVER
<BLP56_RS18080--Novosphingobium_stygium--65.12 (SEQ ID NO:2410)
MERGGWVYIMANRYRGGMYVGVTSDLIRRVWQHREGVGSSHVDDFGKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAAMALR
GVVER
<BLP56_RS01135--Novosphingobium_stygium--62.79 (SEQ ID NO:2411)
MDREQRGWYIMADHYRGS MYVGVTSHLARRVFQHRNGTGS DYCARHSLRLVVAERGENIEACIVHEKRLKRWREWKFDLIEGANPDWADLYDMLV
<BLP56_RS20105--Novosphingobium_stygium--63.51 (SEQ ID NO:2412)
YVGVTSDLIRRVWQHREGVGSSHVDDFGKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAAMALRGVVER
<BLP56_RS18620--Novosphingobium_stygium--61.97 (SEQ ID NO:2413)
VTSDLIRRVWQHREGVGSSHVDDFGKTRLVYAERHEEIEPAIAREKLVKKWRREWKFALIEAGNPDWLDLWEQWYPAAMALRGVVER
<BLS34_RS13790--Thalassobaculum_litoreum_DSM_18839--71.11 (SEQ ID NO:2414)
MRDPIQPCVYILASRRNGTLYTGTGDLVRRVALHRAATAGFCATYGVHRLVYFELLGRMEDAILREKRVKRWRRRAWKLALIEGVNPGWEDLYQSLRTGDLNL
APFAAPPPIRGDRRDALGPGSPDQVRSRPG
<BLS34_RS22270--Thalassobaculum_litoreum_DSM_18839--64.44 (SEQ ID NO:2415)
MQPTVYIILTNHRRGTLYIGVTANPVSRILQHRRGEGSQFVRRCLHRLVYVEQFARMTEAILREKQLKAWRRRAWKVELIEQANPAWSELWMPMGVETVDRGDA
<BLS34_RS13155--Thalassobaculum_litoreum_DSM_18839--76.67 (SEQ ID NO:2416)
MRVWVYILANRRNGTLYVGLTTDLARRIREHREKRLAGFTSRNDVTRLVYVEEQERLDHALIRERRLKRWRRDWKDLIEQQNPDWDRDLSYDILGPPDSG
<BLS08_RS10505--Pseudomonas_chengduensis--82.80 (SEQ ID NO:2417)
MKQPAVYIMASQRNGTLYTGTADLLRRVWQHREGVVEGFTLRYGVKLLVYEQHESMSGAIAREKAIKKWNRRAWKLRLIEERNPQGRDLWSEIVGESSGAAG
APPSRG
<BLS64_RS09620--Salinimicrobium_catena--64.77 (SEQ ID NO:2418)
MHFSYVYIILTNKNHTVLYIGVTSNLIKRIYQHKIKAYKGFTARYNCDQLVYVYQEFNPNI SQAIEYEKKIKAGSRKAKEKLIALSNPEWHDLAADWLFFFQ
<BLS64_RS11730--Salinimicrobium_catena--65.91 (SEQ ID NO:2419)
MKKSYTYLLSNKSRITMLYIGVTADLISRI TQHKNGTGSNFRKYGLQVLLYFEEYTAIDQAIAREKQLKNWHKEWKWNLIKSNNP ELKDLFDELYPWPSIYFK
VFYFRS
<BLS64_RS14900--Salinimicrobium_catena--60.22 (SEQ ID NO:2420)
MLYDYFLYIVTNKNKSTLYIGVTNNIQRRIAQHYFDSEHFRRSFAGRYNXYLVYVEGFATPSEAIQREKEIKKWRREKKNKLISDFNPNWEFLNSDVF

FIG. 11XXX

<BLQ18_RS15995--Muricauda_zhangzhouensis--72.94 (SEQ ID NO:2421)
MTKNVHEYYVILNLNIPNGTLYIGMTNDLERRILEHKTGSVEGFTKKYGLKTSIYFEVHQNVNEAILREKQQLKKWKRQWKIELVEKENPKWKDLASGWYS
<BLQ18_RS03060--Muricauda_zhangzhouensis--62.92 (SEQ ID NO:2422)
MDKYCVYILTNRNKSVLYTGTINDLSRRLIEHIECIKNNKKTFAAKYKCKHLVYVEYDWWQEAIISREKEIKGWVRAKVKLIKTTNPNFEFLEDGFINI
<NSMM_RS14800--Nitrosomonas_mobilis--81.72 (SEQ ID NO:2423)
MEKQPAVYIILASRYNGTLYIGVTSNLIQRIWQHKNLDVEGFAEKYGVHSLVYELHEQMLAAIEREKQQLKKWNRQWKINLIEKVNPSWKDLWLIDL
<BLS33_RS15620--Kriegella_aquimaris--68.24 (SEQ ID NO:2424)
MKKSIFYILTNKNNTVIYCGVTNLSLVKRVYQHKTKYKGTATRYNCGKLIVYFEEYNSIVQAIAREKQQLKAGNRKRKELLIQSINPHWKDLSEGVWFAM
<BLS48_RS29805--Mesorhizobium_qingshengii--66.29 (SEQ ID NO:2425)
MTGVYVMTASQKRGTYIGVTNLDLGRMPHEKSGQSRFTSIYGVQRLVWVEYFDITDAIQREKSLKRWRPQWKIELIEKGNPDWFELFRGTGW
<BLQ49_RS22510--Methylobacterium_phyllostachyos--78.26 (SEQ ID NO:2426)
MREPAVYIMASRRNGTLYTGVTSNLP RR VYEHREVLLVGFTSRYACKLLVWYEPYPTMAEAITREKQIKAGSRKKKLALIEVMSPDWRDLYPDLA
<BLT42_RS12075--Maricaulis_salignorans--70.79 (SEQ ID NO:2427)
MLYFVYLLASGRHGTYLGVGTNNLVRQVFEHKTAQVPGFTKKYRVERLVWWEAHEDINVAITREKRIKRWRDWTMTIERENPDWIDLFASIGGTEPAGLAP
LHLRNW
<BLR27_RS17110--Pedobacter_ruber--71.11 (SEQ ID NO:2428)
MKEYFVYMLANKYHSVLYIGFTNNIRRVYEHKRLVYGFSTKYNCNLLVWYEVFADVNLAIAAREKQQLKVWQRSWKNLIEKHNPWDHDLAADWFVKLGK
<BLR27_RS08865--Pedobacter_ruber--70.45 (SEQ ID NO:2429)
MREYFVYMLANKYHNVTLYIGFTNNIRRVYEHKRLVYGFSTKYNCNLLVWYEFKFDVNLAIAAREKQQLKVWQRSWKNLIEKHNPWDYDLAVDWFVKSSK
<BLR27_RS00280--Pedobacter_ruber--71.76 (SEQ ID NO:2430)
MRQYFVYILSTPDNTIVYIGVTNLDLIRRVREHKEKLNKCFSTQFNCNKLWVEVFRDIRVAIRREKQQLKNWHRSWKNLIEKENPEWMELTPNAQ
<BLP93_RS13015--Desulfonatronum_thiosulfatophilum--83.87 (SEQ ID NO:2431)
MSKQPTVYIILASRNGTLYVGVTSDLVKRVWQHKNNI VEGFTTRYAVHQLVWYELHDTMQSAINREKRIKEWKRWKLDLIEELNPEWIDLYHEL V
<BLS61_RS16360--Pseudomonas_guguanensis--81.61 (SEQ ID NO:2432)
MASQRNGTLYVGVTADLLQRIWQHREGSLEGFTTRYGVKLLVWYEQHGSMSSDAIAAREKAIKKWNRAWKRLRIEERNPHWQDLWPVIIGDSSLAAGSPPSRG
<BLS30_RS17360--Filomicrobium_insigne--69.77 (SEQ ID NO:2433)
MRRFWYILASSPRGTYIGMTNGLVRRVEHRDGTGSVFTRYAIHRLVWVEELPNVQQAIQREKTLKEWPRDWKLNLIERNPNWIDLYPSLPGVAPLRKL
<BLT60_RS02795--Bauldia_litoralis--74.42 (SEQ ID NO:2434)
MVFCYLLASRRNGTLYVGMTDDLARRMWEHQGTGAFPGFTRKYSVRTLVWFEVHQTREAAFTREKQIKKWNRRWKLDLIEAANPEWNDLAAEFWPESGSPLSR
G
<BLT60_RS04260--Bauldia_litoralis--71.11 (SEQ ID NO:2435)
MQHFAGRGFGPCTYILASRKYGTLYVGVTSSVARRVWEHRAGRVRGFTQRYGVKTLVHYEFHQTMFDAIKREKQIKWKRLWKIELIESRNPEDWDLFEELPF
<BLT60_RS18510--Bauldia_litoralis--71.91 (SEQ ID NO:2436)
MAYHVYILTNRPFGTLYVGVTNLDLVRRLSEHRARIVDGFTKTYGLDRLVYERFDRIADAIHREKRLKHWNRAWKVELIESINPNWEDLNPSFL
<BLT60_RS17780--Bauldia_litoralis--72.94 (SEQ ID NO:2437)
MAFFVYIVTNRPKGTLYIGVTNNLARRVGEHKSSHHEGFTKRYKLGLLVWFAAHDVNVNAIAHEKRLKRWNRWAKIDLIEASNPGWEDLTPPFQ
<BLV09_RS00585--Bradyrhizobium_canariense--78.65 (SEQ ID NO:2438)
MKQPCVYIMASRRNGTLYAGVTSDLPKRAFEHREGLAKEGSAKYSCKLLVWYELHDSMIEAITREKQIKAGSRADKLALIEGLNPQWKDLYETLA
<BLV09_RS17260--Bradyrhizobium_canariense--76.67 (SEQ ID NO:2439)
MGYYVYILASRKDGAMYIGVTNLDLVRRIYEHRTKAVAGFASKYNI TRLVWFEIYDDPVSVAISREKELKKWKGWVQLIEAQNPGWDDLYESICN
<BLV09_RS28460--Bradyrhizobium_canariense--75.27 (SEQ ID NO:2440)
MSYYVYLLASRKHGTYLYLGVTSDIVRRGYEHRSKAIGGFTTKYGVDKLVWFEIYDDIATAIAAREKELKKWRDWDKIRLIEBQNPWEDVLYPGISN
<BLU33_RS18585--Mucilaginibacter_mallensis--64.89 (SEQ ID NO:2441)
MELQQYHVYIILANKSNTALYTGTSDLKDRILQHKQKVYKGTAKYECNKLIVYERFQWIDDAIAAREKQLKGGSRKRKIDLIADNPSWGDLTEDWYD
<BLU33_RS00235--Mucilaginibacter_mallensis--61.36 (SEQ ID NO:2442)
MLLKQGHYIYMSNQYRTTYIIGVTSDLERSRVWQHVNGECSAFVKYKLYDLVFYEFERITDAIDREKQLKNWHKDWKI KLIQSMNPEMKDLKQDLKT
<BLU33_RS01335--Mucilaginibacter_mallensis--55.70 (SEQ ID NO:2443)
MNQYFVYIILCSDNSYYTGTINDLERRLYEHDNGISPSYTSKRRPVKLVFHERFPDPAQAI AF EKQIKGWNRTKKEALINRDWHLLPELSVRKKQT
<BLV76_RS05320--Nocardioidea_exalbidus--58.67 (SEQ ID NO:2444)
MIAYVYMLRCSGSLYVGSTRLLQERVHQHQMGMGANYTRSRPLVELVWHEHENP SHAFAREKQVQVNSRAKRLALIA GDYEGLPALAKKDFSRPGTADGSG
PGDLTP
<BLT86_RS09115--Pseudomonas_sihuiensis--82.80 (SEQ ID NO:2445)
MKQPAVYIMASQRNGTLYTGTADLLQRVWQHREGAVEGFTLRYGVKLLVWYEQHESMSGAIAREKAIKKWNRAWKRLRIEERNPQWRDLWSEIVGDSLSLAG
SPPSRG
<BLU40_RS06440--Lutibacter_oricola--67.78 (SEQ ID NO:2446)
MKESYIYILSNKNHTVLYIGVTSNLDRIQQHKNGTGSIFTKKYNIVMLMYFETFTNINEAIKREKQQLKKWNKKWKEWELIQSINPQLKDLYQDLIH
<BLW21_RS08665--Hymenobacter_psychrophilus--64.29 (SEQ ID NO:2447)
MYIYIILTNPARTVLYTGVTDNLERRLHEHSQNLGEAGRFTSRYQCNNLVYFEACATSQQAI IREKEIKGWSRAKKE TL IARLNPTWIAIDLETWDGGTES
<BLW21_RS10520--Hymenobacter_psychrophilus--55.17 (SEQ ID NO:2448)
MPYVYVYITNPSRQVLYIGVTRNLATMAQHYQNRGTAQTFAGRYFCYHLLYFEDYPTSTAAIAAREKELKGWTRVKKEALIAVQNPQWFLKV
<BLW21_RS04605--Hymenobacter_psychrophilus--55.95 (SEQ ID NO:2449)
MPYVYVYITNPAKRVLYIGITNDLANRMSQHYANRGSQAQTFAGKYFCYNLYFEEHQDSKAAIAAREKELKGWTRAKKQALATDNPEWIFLKA
<BLV85_RS00855--Nitrosomonas_halophila--79.78 (SEQ ID NO:2450)
MKQPCVYIILASQRNGTLYTGVTSDLIKRIWEHKNDQTKDFTNKYAVHRLVYFEHEDMI SAITREKQIKKWNRAWKRLRIEERNPGWNDLYEKL
<BLV84_RS09155--Psychroflexus_halocasei--70.11 (SEQ ID NO:2451)
MKRSYHNYSVYI LTNKQGTLYIGMTGGIDRIERHRQKGKSF TAKYNLT KL VYQEDFQYVNDIAIAAREKQQLKSWQRDWKIELIEKDNPEWKDLYKPLRGDSE
<BLU78_RS12435--Alistipes_timonensis_JC136--62.22 (SEQ ID NO:2452)
MESSYVYI LTNKHLTVLYVGVTSLSRMKHAAGQSAFTRRYNVHKLIIYEGYRVDVDAIAAREKQIKSWSRKRKMLVDSRNSLWKDLSDGIV
<BLV00_RS20795--Thalassobacillus_cyri--56.96 (SEQ ID NO:2453)
MASDHFVYIILCKDGTLYTGTNDLYQRLAKHESGKGAKYTRGRGPFLVYFERNCTKTAALQREFQIKQWSRSKKEALIALKNEG

FIG. 11YYY

<C4S60_RS20180--*Vibrio_hangzhouensis*--69.23 (SEQ ID NO:2454)
MVKGPCLYIITSRNHSVLIYIGVTSNLQQRIWKHKHGVGTGGFVKYKYNVHLLVHFEQFDTEMAAIAREKQIKHWNRPWKNEIIVTSHNPFYWRDLSLELW
<C4S29_RS00535--*Chryseobacterium_humi*--67.50 (SEQ ID NO:2455)
MNETTEGIYTFYVYILTNQRKTVLYTGVTTNNLHIRLQQHSKNINEKSFTSRYNHFLIHYEKFQWIGQNAI LREKEIKSFSRDKKLELIRKQNPNDLFLNHLFE
LK
<C4S73_RS15825--*Bosea_lathyri*--68.60 (SEQ ID NO:2456)
MTNRPNGLTYIGVTSDLARRAYEHREGLCAGFTRKYGLKQLVWYEWHDIAAAIHRETSMKRWYRAYKTRAIMARNPDWRDLYEELA
<C4S73_RS03700--*Bosea_lathyri*--74.44 (SEQ ID NO:2457)
MMANKPRGTLYIGVTAELVRRVSEHRLQGRNFTARYNLDKLVYETHDHDILAIQRETSILKRSEARWKIELVEKSNPDWLDLDFESVI
<C4S73_RS24995--*Bosea_lathyri*--65.17 (SEQ ID NO:2458)
MTREVRAMVYIMASGWNGLTVGVTSNLSRRVTEHREGRSGFANKYAVTHLVYAERFERADEAIAQEKLKMRWRERERKISLIESCNPDWRDLYDTIHLD
<BLW20_RS13890--*Flavobacterium_marinum*--61.36 (SEQ ID NO:2459)
MKQYIYILTNRFRTTFYIGVTTNNLNIRISQHADGRGSKFTSKYQLVDLVYIYNDINSVIAREKQKLNWRKDWKINLIKTINPKLETITIL
<BMX42_RS12765--*Flavobacterium_terrigena*--58.06 (SEQ ID NO:2460)
MIKGYIYIMSNKNLSVLIYGATKDLKSRIDSHLNGTGAVFTKKYNANILLYFEEDDYQDAFKREKQKLNWHKEWKNLIKISNPQLKDLIYLDL
<BMX57_RS09805--*Cribrihabitans_marinus*--62.35 (SEQ ID NO:2461)
MPHFVYIMASRPAGALYTGRTRDLRRRVEAHRAGFSNHTAKYKIRTLVWFEGLDDEASLARERQIKRWRDVKIALITRNPDKDMTAHVDP
<BMX43_RS15275--*Deinococcus_reticulitermitis*--75.29 (SEQ ID NO:2462)
MTNRRNGTLTYIGVTTNNLRRVWHRQKVPVPGFTQKYNLTRLVYFEETDVRSAIAREKQKLGWLRAKKVALIEGMNPQWRDLYGDTTP
<BMX21_RS07085--*Bosea_lupini*--71.11 (SEQ ID NO:2463)
MAYFVYLMASRSGTLYLGI TNDLARRAHEHKSRI GFTSRYGVDRLVWYQFEDVRDAIDREKVLKWRRAWKIKLIEDMNPWKDFYERLS
<BMX77_RS09465--*bacterium_A37T11*--64.84 (SEQ ID NO:2464)
MERGGCVYIMTNKGRSVLYIGVTSDLRSRIYEHQHHKYPSSFTARYNLECCYIYEFSDIEBAIQREKQKLGWSRSKKNDLISSVNPWKDLWIEIETW
<BMX77_RS19195--*bacterium_A37T11*--55.06 (SEQ ID NO:2465)
MNRFIYILTDCRKRSLHVMAADLLHTVTTFYRENAHLFFDAANPTRLVYFEELTDDTTACARFNQLQRFTRMQKERLIRAVNVNWDVLIQ
<BMW64_RS18650--*Chryseobacterium_taichungense*--62.22 (SEQ ID NO:2466)
MIEFKDGIYTFYVYILTNKSRTVFYTGVTKNLHQRLYQHKIKSNSNSFTAKYNVEFLVYEFKFGWIGQAIEREKEIKNYTRSKKLELIRSVNPNLDSWNLDFQ
QL
<BMW64_RS00220--*Chryseobacterium_taichungense*--62.07 (SEQ ID NO:2467)
MKTGLGTHNYVYIILTNKIKTVLYTGVTTNDLEKRLYWHKNPEADIKSFTFKYKCFVLVYFECFKDIDTAIKREKQIKGWVRNKKEALINEFNPTWKFLNDEIK
<BMX64_RS13145--*Flavobacterium_sinopsychrotolerans*--58.43 (SEQ ID NO:2468)
MEKGFYIYILTNKNATVLYVGATKNLKNRVDCNRNGTGAVFTKKYNATILYFEVFDLSDAFKREKQKLNWHKDWKNIAKLSNPDLKDLIYLN
<BMX71_RS07100--*Rhodopseudomonas_pseudopalustris*--78.65 (SEQ ID NO:2469)
MAYYVYLLASRKDGPLYLGVTTNNLRRVYEHRTKAVPSFTSRYNITKLWVFEIYDDPI SAITREKEIKKWRRAWKV ALIERENPGWEDLYDGIQ
<BMX71_RS18140--*Rhodopseudomonas_pseudopalustris*--73.12 (SEQ ID NO:2470)
MSYGGFYLLASNGKHTLYIGVTTNDLVRRCYEHRTKAAPGSSRYGIDKLWVFEIYDDVVSATITREKQKLRWRDVKVRLIEENPDWVDLTGLISH
<BMX71_RS10095--*Rhodopseudomonas_pseudopalustris*--69.66 (SEQ ID NO:2471)
MRSYDVYIILASRYGTLYVGVTNDLPRRIAEHKGAGADGFTKKYNIKILVHVEEYSSILEARTREHVLKWRWRDVKIALIEKLNPDWRDLSNDL
<BMX86_RS08560--*Flavobacterium_urocaniciphilum*--65.26 (SEQ ID NO:2472)
MKPGFYIYILTNKNATVLYVGATKNLKPRIAEHKEKRYQNSFTARYNLDKLVYFEKFEIISDAIAREKQKLGKGNRTNKEKLNLSNPEWNDLYEIEIKDIMIV
<BMX04_RS22085--*Friedmanniella_flava*--58.97 (SEQ ID NO:2473)
MPFVYLLRCSGDSFVVGSTRDLQRLTEHALGVVKGYSKRRPVLELLWSLETDRVDDAAELERQIKGWRRAKRLALVEGRGLDRLARAGPHSEGR
<BMW63_RS01140--*Flavobacterium_frigoris*--63.75 (SEQ ID NO:2474)
MIEFQEGYHSYVYIILTNKAKTVLYTGVTTNNLIRLQHQHKSINPNSTAKYNVHLLYFEKFTWIGQAIAREKEIKGWKKDKKITLIKTINPELNFLNYLFE
<BMX24_RS19270--*Chryseobacterium_wanjuae*--58.14 (SEQ ID NO:2475)
MFELNEGIYKFYVYIILTNKNRTVLYTGVTTGNLHQRLKEHVKNINPSTAKYNTCFIYIEKFGWIGQAIEREKEIKLLKRNKYFELIRNVNPNLEFLNHLFQ
QAV
<BMX67_RS07955--*Hymenobacter_actinosclerus*--61.18 (SEQ ID NO:2476)
MYVYIILTNQQTLYIGVTTNDLRRLYEHSSGRGDEGKFTGRQTDLVYFEIAPDATQAIAREKQKLGWTRRKKDALVTA FNPTWQAI DLETWEG
<BM107_RS03010--*Pedobacter_suwonensis*--75.86 (SEQ ID NO:2477)
MERGGCIYIMTNYQRTTLYIGVTSDLRSRIYEHREKVYPKSFTAKYNLTICVYVEYSTIEBAIDREKQKLNWRRAKDTLIDKFNKSWDLTAI INEWED
<BM107_RS22390--*Pedobacter_suwonensis*--60.23 (SEQ ID NO:2478)
MERGGCIYIMANSFNTVFYMGVTSDLQGRVWQHKNNEFSNSFTARYKCKLVYFAFFFGIEBAINEEKRIKEGNRQQKIDILVSKNPNWIDL YDGLDA
<BM107_RS13795--*Pedobacter_suwonensis*--62.50 (SEQ ID NO:2479)
MERGGTVYIITNFTHHTFTYTGVTSDITSRIIEHKEKAHPKSFSKYNLDKLVYFCSYDSIDEAIAEENRIKSGSRKRKIKLIESINPGWTDLWEIVKEW
<BM053_RS05625--*Flavobacterium_swingsii*--78.41 (SEQ ID NO:2480)
MPKTTHQYIYILASQKNGTLYIGVTTNDLERRILEHKQKINEGFTSKYDINMLVYFESFYQYINDAILREKRLKKNRWQKINLIEENKDWKDLSEDWYN
<BM291_RS05500--*Algoriphagus_aquimarinus*--64.84 (SEQ ID NO:2481)
MEKGNAIYIMTNARKTVLYTGVTTNNLRRVYEHKSGLNSTFTFKYNLKLIVYFESFHTIDEALIREKQIKGGSRRKKVDLINSINPEWKDLWEEIEKW
<BM127_RS15190--*Nocardioides_alpinus*--56.58 (SEQ ID NO:2482)
MAWTYILECSDGSFVVGSTVHLBQRLYQHGEGLGAAYTKRRRPVKLVWAADFARVDEAYFYEKQIQGWSRAKRIALIEGREEDLP ELARGYWRPSPV
<BM201_RS16270--*Zunongwangia_mangrovi*--54.12 (SEQ ID NO:2483)
MKKSFVYIILTNKYRTVYFTGVTTNDLRRRI SEHNNGTGSKFTEKYTKDLIYFEEDFNDQAVAREKQKLNWKKEWKNLIKSFPNKLQLEY
<BM201_RS02215--*Zunongwangia_mangrovi*--61.54 (SEQ ID NO:2484)
MNEYTYITCNSSKLTIFYIGVTTNNLNRKLSQHYEDSIKEKRSFAGKYNCLYIYFEVFNVDQAIAREKQKLGWSRAKKENLIRLKNPNLEFLSL
<BM031_RS08605--*Massilia_yuzhufengensis*--75.56 (SEQ ID NO:2485)
MDKHSYVYIILASRPYGTLYTGVTSDLIRRIWQHKEGFVAGFTKTYNIKQLVWYEVHSDIAAI TREKQIKEWPRDWKINLIQSTNPNWHDLYAQIIS
<BM296_RS05355--*Maribacter_stanieri*--64.71 (SEQ ID NO:2486)
MKTSYIYFMTNKNHTVIYTGITSNLLRRVYQHKTKYFIGFTSKYNCNLFVYEEFDDVNLAIEREKQIKSGNRKRKEELVNSVNPWNWEDLSDGWI FECK
<BM246_RS17430--*Algibacter_lectus*--65.88 (SEQ ID NO:2487)
MDVHYVYIILTNKNHTVLYVGRTKQLKVRNLQHKNNSSKSTGKYNVDKLVYFETTKYVNNSIKRERQIKKNWRRAWKINLINSNPDWDLSEYI

FIG. 11ZZZ

<BM289_RS11390--Flavobacterium phragmitis--69.41 (SEQ ID NO:2488)
MLNTQIGYHSYYVYIIITNEHRSSFYIGVTNNLRKRLAKHKENIDLNISSFVAKYNIQFLVYVEKFTWIQEAITREKELKKWRRDKKIELIKNFNETF
<BM172_RS07550--Dyella_marensis--78.65 (SEQ ID NO:2489)
MDRRQPCVYMLSSGHNGILYIGVTSDDLKRVWQHKGHVIDGFSKQYDVHDLVWYELHETMESAILREKALKSWRRQWKLRLIEGENPEWRDLYPAL
<BMZ76_RS02285--Flavobacterium xueshanense--57.30 (SEQ ID NO:2490)
MEKGFTYILTNKNSTVLVVGATKNLKNRIDCHRNGTGAVFTKKYNATILIYFEEFEDLSDAFKREKQLKNWKKDWKWNVAKLLNPELKDLYLNL
<BMZ76_RS10370--Flavobacterium xueshanense--56.98 (SEQ ID NO:2491)
MKFYVYILRCIDDSLYVGLTSDIERRVMEHNMKGYPEAYTYSRRPILLVYFQDFTDPNQAIEFEKKIKKWSRMKKQALIDENFDKQLDSECRNATHHKYKS
DFEDIEE
<BM156_RS04175--Sunxiuginia_elliptica--65.56 (SEQ ID NO:2492)
MHKGFVYILSNKNRTTFYIGVTSIDIKRRIEHTTGIGSAFTKKYNLTDLMYEKEIEGMNECINREKQLKNWHRDWKINLIREDNPENKDLASWFIQDGLL
LSDR
<BM040_RS16305--Pontibacter_chinhatensis--58.51 (SEQ ID NO:2493)
MASHNVFYIITNPKRRTTYTGVTNDLPRRLKEHFLNSGNPASFAGKYFCYNLIYERHTHVEHAIVREKEIKKWNRSKKEBELINSFNPFWKFLNLEVEDE
<BM360_RS18780--Sphingomonas_jatrophae--72.73 (SEQ ID NO:2494)
MRQERQPCVYMLASGFNGTLCIGVTSIDLPCRLHQHRAGTVPFGTCRYGVHHLVWFEMADTMEAAIAREKQLKAWRRDWKMNLIERNPGWRDLSVELGLTPRG
GC
<BM043_RS05965--Algoriphagus_hitonicola--58.46 (SEQ ID NO:2495)
MDEFVYILYSPSTRKTYTGMTSDLI TRFHFNQKSTKGYTIIRYPWIVIVVEFFTKKDALAREKELKSGKGRDWIKKYILPFYK
<BM249_RS09835--Methylobacterium_gossipiicola--58.90 (SEQ ID NO:2496)
MFTEGIGRKPQYGTAVPGTQDYGVTLVWYEGYPEIAMARQREHTLKRWRSSWKLALIEAMNPEWRDLAADLNG
<BM207_RS01675--Pedobacter_insulae--57.95 (SEQ ID NO:2497)
MGKGGCVYIMTNVHTVLTGVSDDLSESRVWEHKNHQPFSFSAKYRCHKLIVYFYFIIHEEADIEEKIKGGSRKAKEKLINALNPEWLDLYESLFD
<BM238_RS02725--Chryseobacterium_frigidisoli--66.32 (SEQ ID NO:2498)
MKPGFIYILTNKYNTTLVGVTSNLPKRIKDHEKRYENSFSSRYNLTKLVYWEAFQEIIGDAIAREKQIKGGSRKKNLNLINNFNAEWKDLYEITIEFESLG
AGF
<BM238_RS06375--Chryseobacterium_frigidisoli--65.22 (SEQ ID NO:2499)
MIETSDGIYTFYVYILMNQRKTVLYTGVTNNLHRLQOQSTNVNLSFTSRYNVHFLVHYEKFDWIQNAILREKEIKSFSRQKKMDMIRKYNPNLDFQNHLEF
LK
<BM238_RS02575--Chryseobacterium_frigidisoli--62.82 (SEQ ID NO:2500)
MIEFTEGIYNKYVYILTNKNRTVLYTGVTNNLYRRLDEHFNNQNPQSFSRYNTKYLIYYQKFGWIQLAIEREKEIKNLSREKKLKLIREINPNLDFLNDLFK
NNITN
<BMZ47_RS20535--Methylobacterium_brachiatum--73.86 (SEQ ID NO:2501)
MRAPAVYIMASRRNGTLYTGVTADLSRRVYEHRESLHPGFTTRYGCKLLVWYEGYPTMNEAIAAREKQIKGGSRKKKLALIEAMNPDWCDLYPGLA
<BMZ09_RS05875--Olleya_namhaensis--77.27 (SEQ ID NO:2502)
MKVNNRQYVYIILSSKNGTLYMGVTNNLERRLIEHKQLASVFTSKYDVNRLMYPIADINEAIKREKQLKNWNKQWKIELIEKENPNWKDLSINWN
<BMZ09_RS10675--Olleya_namhaensis--69.47 (SEQ ID NO:2503)
MKPGYIYMLTNKNNTTLVGVGTAYLKKRILOHKEKHDLSFTARYNLDKLVYHEAHQMGDAIAREKQLKGSRAQKIALIESVNPWETDLDFEL
<BM231_RS08595--Pseudovibrio_ascidiaceicola--76.40 (SEQ ID NO:2504)
MPYIYITASQRNGTLYTGVTNNLARRIHEHREKVATSFTRYGVSGVLVYEEFERIGDAIAREKQIKKRWRAWKLIETLNPWRDLYQDLA
<BM231_RS13415--Pseudovibrio_ascidiaceicola--67.82 (SEQ ID NO:2505)
MYILASKPRGTLYTGVTNDLARRVYEHKSKVGCRTAKYGVSKLVYFEQFRDVNDAIRREKQVKRWRRVWKIDAIEGLNPKWDDLYLRFNH
<F3H98_RS03780--Mesorhizobium_albiziae--72.94 (SEQ ID NO:2506)
MSRRGVYLLASQRTLYGVVTSDLAGRLHQHRTGTGSKFAARYGVSMVLVWFDEYELITDAIAREKAIKKWPRQWKIKLIEBNPDWRDITWHLV
<BT151_RS02910--Flaviramulus_basaltis--62.35 (SEQ ID NO:2507)
MKFSYVYILTNKYRTTFYIGVTANLPKRLEEYHEITSKFTKKYNTKDLIYFETFSNIEQAIAKEKQLKNWHEKWLNLINTINPTLKTLDY
<BT151_RS03475--Flaviramulus_basaltis--67.07 (SEQ ID NO:2508)
MNLWYVYIITNKPQGVYIIGVTDNIDERVKEHKLKANPKAFTAKYNCVSVLVYFEFENGGAEKRRERQFKKWKRDWKIKLIQEMNSS
<BM228_RS03220--Methylobacterium_pseudosasicola--75.00 (SEQ ID NO:2509)
MRAPAVYIMASRRNGTLYTGVTADLSRRVYEHREALRPGFTTRYGCKLLVWYEGYPTISEAIAAREKQIKAGSRKKKLALIEAMNPDWDRDLYPGLA
<BM228_RS03750--Methylobacterium_pseudosasicola--70.00 (SEQ ID NO:2510)
MAFRVYLLASGLHGTLYIGVTNNLARRVYAHRSRHSPGFTAPYGVTRLVWYEAIDRIIEAIEREKALKKWSRDWKIRLIEEMDPDWSLYGTLAGWTGS
<BM400_RS11365--Granulicella_pectinivorans--66.29 (SEQ ID NO:2511)
MPEIAYVYIMASTFQVLYIGITTHIEQRVTQHKEAKDPDSFTARYKVNKLIVYEEFAWITNAIAREKQLKRSRIKKIRLIVAMNPTWQDLSADWGKPFPLLT
PTK
<BM316_RS03870--Yoonia_litorea--61.80 (SEQ ID NO:2512)
MQHFVYIMASGPRGAIYIGRSRNLARVEQHRAGLSVHTSRYKIRWLWVFEVQDDFATSLQRERALKRWRRDWKNQLEATNPHWQDITAHIP
<BM116_RS11525--Pseudomonas_pachastrellae--75.00 (SEQ ID NO:2513)
MPKQPCVYIILASKRNGTLYIGVTSNLIKRIWEHKNLVPGFTQRYRIHHLVYVEQAADMLSAIAREKQLKKWNRRWKLNLIEQANPGWADLYTSLANPEAASP
<BMZ23_RS05960--Pseudomonas_yangmingensis--83.87 (SEQ ID NO:2514)
MKQPCVYIISNRNGTLYTGVTNNLPQRVWQHREGIIEGFSRYRILKMLVWFEPHTNMETAIQREKAIKKWNQWKLRLIEQHNPDRDLWADICR
<F3H54_RS21140--Pleomorphomonas_diazotrophica--70.00 (SEQ ID NO:2515)
MNAGWLYIVTNKPRGVLYTGVTSDLEERRIWEHRNHIYRGFTDSYNCTRLVWYEDHGDITDAIYREKQMKHKKREWIKIRLIEEMNPWQDLYETIM
<F3H54_RS16135--Pleomorphomonas_diazotrophica--68.24 (SEQ ID NO:2516)
MKAGWYIVITNKPQGVLYIGVTSDLGRIWEHRNHIYRGFTDSYNCTRLVWYEDCGDITDAIHRENRLKFWKRAWKIRLIEEMNPWDDLFETIMS
<F3H54_RS19565--Pleomorphomonas_diazotrophica--66.17 (SEQ ID NO:2517)
MNPGWYIVTNKPRGVLYGVTSDLGRIWEHRNGVYRGFTHSYNCKRLVWYEDYDEIPDAIHREKRLKTWNRDWKIRLVEEMNPWDDLFETIMW
<BM072_RS04100--Algoriella_xinjiangensis--67.06 (SEQ ID NO:2518)
MNLVYVYILTNKYRTVYIGVTNNLSKRIIEHKENNGSIFTKKYNVHSLIYFEEFSSIHEAIAAREKQLKNWKREWKINLIQFNPTLKDINIFE
<BM125_RS05885--Dokdonella_immobilis--78.49 (SEQ ID NO:2519)
MRERQPCVYVYLASGQRGTLYTGVTSNLPKRIFEHRNDLVDFGSSRYHVHDLVWFELHETMASAILREKAIKAWKRPWKIRMI EESNPDRDLVPSLVGEAPE

FIG. 11AAAA

<BM253_RS09700--Flavobacterium_ummariense--63.53 (SEQ ID NO:2520)
MLCFIITNKFRFTTYPVGVSNLNNRISQHNNDGSIFTSKYNLTDLIYFEVENEINQAIIREKQLKNWKKWKLNLKSTNPNLETLKLF
<BM197_RS03405--Mycetocola_miduiensis--52.27 (SEQ ID NO:2521)
MPYMYMLRCSGDTLYTGSTWDLSHRMDQHDGSLGAKYTSRRRLPLTLVYFEYFDRIDEAYYRERQIHGWVRWKKDRLIAEGPGGRDPWDLV
<BMZ91_RS12495--Roseovarius_lutimaris--68.89 (SEQ ID NO:2522)
MPERMHVYIMTNRPRTLYIGVTGDLINRIWQHRIHALGGFTDRYNLESLVYLEPHQAPAVAIQREKSLKRWRRNWKIALIEKANPEWRDLWEIEICSG
<BM073_RS03995--Algoriphagus_ornithinivorans--58.46 (SEQ ID NO:2523)
MDEFVYILFSPSTGKTYTGMTSDLI TRFNHFNKSTKGTTRYPWMVVFVEFFDSKKDALKREKELKSGKGRDWVKNYVLPNFH
<BM059_RS05765--Pseudomonas_composti--80.65 (SEQ ID NO:2524)
MKRPVVYIMAGQRNGTLYTGVTADLLQRVWQHREGAVEGFTRRYGVKLLVWYERHEDMGSAITREKAIKKWNRTWKLRLIEERNPQWRDLWSEIVGDSSVAAG
SPPSRG
<BM145_RS11980--Sphingomonas_rubra--78.65 (SEQ ID NO:2525)
MRENFQPCVYLLASGRHGTLYVGVTSNLLARVSQHRERSISGFTARYRVDRLVWYAVADTMDAAITREKQLKKWNRDWWKRLIEENNPWDLDLATALGLPPRG
IPAKADVIGDSEKRRD
<BM145_RS13575--Sphingomonas_rubra--71.59 (SEQ ID NO:2526)
MRERQPCVYMLASGFYGTLYVGVTSNLIGRLIQHREEMRGFTSEHAVKRLVWYEVADTMDAVIAHEKRLKRWLRDWKIALIERDNPRWDDLAI GFGLPPLPP
PR
<BM145_RS07935--Sphingomonas_rubra--69.32 (SEQ ID NO:2527)
MMASGRNGTLYLGSTGNLAKRAWHRNGVVS GF TKKYGCKALVWYEVLGSWEAARQRELQMKWEKRSWKVREIEGLNPDWEDLYDRIAQP
<BM145_RS12670--Sphingomonas_rubra--60.00 (SEQ ID NO:2528)
MVNDHRGGWLYIMADRYRGAMYGTASL MARVHQH RAGTGS DCHRYGLTRLVWAERGGD DIGACIAHEKRLKRWREWK FALIERGNPNWDLDTQLA
<BM015_RS06270--Parafilimonas_terrae--64.84 (SEQ ID NO:2529)
MTKGGSVYIMTNKLTLYIGVTS DLQARIRQHKNHFFKGSFTDEYNLEYCVYYENFFSIEBAILREKQIKKWGRLLKKEKLINSIKPNWIDLWIEIEKW
<BM015_RS05740--Parafilimonas_terrae--62.50 (SEQ ID NO:2530)
MKVVNQIMTNKLTLYIDVTSNLQARIRQHKNHFFKGSFTDEYNLEYCVYYENFFSIEBAILREKQIKMWSRLKKEKLINSINPNWDLWLKLKNGKILFK
LSQSQEI SPIVEMTFEVIYIYSHL
<BMZ66_RS19205--Siccationidurans_arizonensis--69.41 (SEQ ID NO:2531)
MRAYVYLLTNANHTVLYIGVTNNLRRVAEHKAGMHQGF TKKYNVHKL VHFETFPD VNAAIAREKQLKAGSRAKKALVHINNPTWQELAE
<BMZ66_RS06035--Siccationidurans_arizonensis--55.95 (SEQ ID NO:2532)
MPYVYVITTNPGKSVLYIGVTNNIADRMAQHYANRGSTQTFAGRYFCYNLVHVEVYSDITAAITREKELKGWTRAKKDTLIANDNPHWLFLQT
<BM335_RS14565--Lutibacter_maritimus--73.40 (SEQ ID NO:2533)
MKTIHQYYVILASKIRGTLYIGITNDLQRRVYEHKKELVKGFTQKYGVHKL MYFETFKSID EAIKREKNLKKWKRDKWIKLIEEENKKWDLSSDWFNDTLK
D
<BM375_RS14790--Sphingobacterium_wenxiniae--64.77 (SEQ ID NO:2534)
MERGGCVYIMTNNRCTVLYTGVTSDLQSRMYEHQHKIHPNGFTNKYKVNRLVYRFFPNIEETISEEKRLKGGSRHQKIRLIEELNPNWDDLYEKEVKLW
<BM391_RS17470--Pseudovibrio_denitrificans--75.56 (SEQ ID NO:2535)
MPYVYVITASQRNGTLYTGVTNNLSRIIEHKEKTATSF TARYDVSQLVYFEEFERIGDAIAREKKLKRWRRAWKLNLIETLNPTWRDLVLDLVQGC
<BM391_RS19165--Pseudovibrio_denitrificans--74.16 (SEQ ID NO:2536)
MYILASRKHGTLTYGTVNNLPERVHQHKTMMNSGFTAKYGVSKLVYYELYEFPHEAIRKEKQVKRWREWKIQAIEEMNPHWEDLSLSLHFEHT
<BM386_RS01305--Algoriphagus_locisalis--64.84 (SEQ ID NO:2537)
MFVYIILTNKTKSTLYIGVTNDLQRRLYEHINGLGNPNSFTSKYKCHYL VYFEEYATPKEA IAREKELKKWRREKKKNKLIENPNWFEFLNSEIE
<BM383_RS09990--Pustulibacterium_marinum--64.77 (SEQ ID NO:2538)
MKTSYIYILTNKNTTLYIGVTSNLKHRIDQHNKKRGAIFTTKYNLT ELIYVEEFNEIEQAI RREKQLKNWRREWKWNLIKQKNPNLETIYIE
<BM428_RS05610--Bradyrhizobium_arachidis--78.72 (SEQ ID NO:2539)
MAYVYVILASRRDGAIVGITNDLVRRVYEHRIKAVRGFTAKYNI TQLVWFEVYDDPISAI SREKELKKWKRAWKIQLIEKDNPDWNDLYESICK
<BM428_RS09225--Bradyrhizobium_arachidis--75.56 (SEQ ID NO:2540)
MAYVYLASKKHGTLTYIGVTS DLVRRVYEHKTKAVPGFTTKYGVDKLVLEFIYDDAVTAIAREKELKKWRDWDKTRLIEEQNPNWDDLYPGITN
<BM428_RS39735--Bradyrhizobium_arachidis--77.53 (SEQ ID NO:2541)
MSARHYVYVILASRIIGTLYIGVTNDIIRV TQHKSKLIESFTEKYDVARLVYFEQFDDPENAIKREKRLKKWNRWKIRLIEQHNPNWDDPYPGIAGPP
<BT112_RS14935--Sinomicrobium_oceani--69.41 (SEQ ID NO:2542)
MTQKSYVYFMTNQNTVLYVGVTSDLIKRVYQHKTKAYKGFTAKYNC DKLVYFEFSDINEA IAREKQIKKGNRKRKEILIQKTNT EWKDLSEGWIFDVS
<BT112_RS07815--Sinomicrobium_oceani--57.47 (SEQ ID NO:2543)
MKTLSHNYVYIILTNRRKTVLYIGVTNNLKT RLYYHKNTDTNSTAFTARYRCFYLIYWEHFD SIIKAIKREKELKGWRREKKENLINNFNPEWFLNDKID
<BQ4451_RS03720--Anaerococcus_mediterraneensis--67.37 (SEQ ID NO:2544)
MAGYFYILSNASNKVIYKGSTTDI IKRTYEHKNHLIKNSFTSKYNVT KLVYFEIYNDIR DARGREVQI GKWSRAKKDKLINKMNPKRKDLYRDLI
<BQ7474_RS06945--Anaerococcus_urinomassiliensis--67.05 (SEQ ID NO:2545)
MKIYYTYILTNKYNSLYIGVTGNIVNRLF EHKHLLRGYTDKYNCTKL VYEEFADSNEAIKREKTLKSKLRSKKIALIESTNPNWEDLSKDWGIPILSPKE
RQIYKENLQEHNRKMDAIII
<BUQ90_RS04015--Parasphingorhabdus_marina_DSM_22363--73.03 (SEQ ID NO:2546)
MTKPGVYVIMANKRNGTIYLGVASDLPKRIYEHNRGLIEGFTQKYGCKMLVWFQEYDDLEEARAAELRMKNWNRQWKLKRIERMNPEWRDLWDEIAR
<BUQ90_RS04030--Parasphingorhabdus_marina_DSM_22363--73.17 (SEQ ID NO:2547)
MANKRNGTLYLGVTSRLEQRAWQHREGLMEGFSKEHGCVKLVWFEAFDDIQEARRRELQMKWKRGWKLELIEKDN PQWKDLFETLF
<BUB91_RS01725--Leeuwenhoekiella_marinoflava_DSM_3653--67.06 (SEQ ID NO:2548)
MQNSYVYIVTNSYRTTFYIGVTS DLKQRVLDHLENKGSKFTSKYNLHDLIYFEQFTSIDQAILREKQLKNWRKSWKLNLIKEKNPTLQTLNIFE
<BUB02_RS14055--Cnuella_takakiae--62.20 (SEQ ID NO:2549)
MGNAGHTTYIGITGNLHQRIWQHKTGHYGGFTHKYNCHHLLYFETYTDVKQAIAREKKNLKNWRDWDKINLVQKNPELRDLAAGWYAKN
<BUB02_RS09685--Cnuella_takakiae--60.92 (SEQ ID NO:2550)
MTKQGWVYIILSYTHHTVLYTGVT SNLQRMEQHRSHLYPYSFTARYNCTKL VYCYADILEAICEEKRIKGGSRKAKELLININPNKVDLWELRLNLI
<BUB67_RS00220--Desulfotomaculum_putei_DSM_12395--56.58 (SEQ ID NO:2551)
MYVMLECSGDCLYGTITNDLKRRLAMHRQ GKASKFTRGRPLV KLVYFEAGYTKGQALSREKVIKGMTRARKLQLTQSFEF

FIG. 11BBBB

<BUA89_RS01095--Chryseobacterium_arachidis--63.22 (SEQ ID NO:2552)
MIEFSEGIYTFYVYILTNRKRTVLYTGVTTNNLHKRLFQHATKENKNSFTAKYNIEFLIYYEKFQWIIQQAIEREKEIKDFPRNKKLELIRSVNPNLDFWNLYLFE
QEF
<BUA53_RS20290--Flavisolibacter_ginsengisoli_DSM_18119--58.06 (SEQ ID NO:2553)
MYTKPYFVYILTNEMLTGVLYTGVTTNDLEQRITEHYKQGRGHNTSFTGRYVNVYHLLYEPYKYINEAIAAREKEIKGLTREKKMKLINEMNPRLEFLNKEICGCWP
PRDLPGRY
<BUA53_RS19250--Flavisolibacter_ginsengisoli_DSM_18119--57.61 (SEQ ID NO:2554)
MYTKPYFVYILTNEMLTGVLYTGVTTNDLEQRITEHYKQGRGQNTSFAGRYNVYHLLYEPYKYINEAIAAREKEIKGLTREKKMKMINEMNPRLEFLNKEICGCWP
PRDLPGRY
<BUB04_RS16055--Desulfacinum_infernum_DSM_9756--80.90 (SEQ ID NO:2555)
MTKQFVYVILASRKNGLTVVGVTSDLKRRVWEHKSGLVEGFTKKYGVKDLVYFEVHESAEAAITKEKRIKKWRRKWKIALIEKENPYWQDLYENL
<BUA76_RS13850--Chryseobacterium_takakiae--62.11 (SEQ ID NO:2556)
MKPGFIYIMTNKNTVLYTGVTSNLPKRVRHEHKEKFFELSFTSKYNNVNLVYVESFQEIQDAIFREKQIKAGSRQKKIDLINSINPKWGDLTEDIKDIMPDPF
<BUA76_RS13515--Chryseobacterium_takakiae--63.22 (SEQ ID NO:2557)
MKTGLGTHNYVYIILTNRKYTVLYTGVTTNNLKRRLFEHRNPEAFSKAFTTKYKCFFLIYFELFADIDTAIAREKQIKGWIRSKKDNLINEFNPKWKFLNDEI
<BUA71_RS14130--Salegentibacter_echinorum--62.07 (SEQ ID NO:2558)
MKTGLGTHNYVYIILTNRKNTVLYTGMTNDLKRRLFEHRNPEAFSKAFTTKYKCFFLIYFELFADIDTAIAREKQIKGWIRSKKDNLINEFNPKWKFLNDEI
<BUA71_RS09195--Salegentibacter_echinorum--60.92 (SEQ ID NO:2559)
MKTGLGTHNYVYIILTNRKYTVLYTGVTTNNLKRRLFEHRNPEAFSKAFTTKYKCFFLIYFELFADIDTAIAREKQIKGWIRSKKDNLINEFNPKWKFLNDEI
<BUB79_RS05760--Tenacibaculum_mesophilum--61.36 (SEQ ID NO:2560)
MLNKYVYVYIILTNRKNTVLYTGVTTNNLKRRLFEHRNPEAFSKAFTTKYKCFFLIYFELFADIDTAIAREKQIKGWIRSKKDNLINEFNPKWKFLNDEI
<BUB79_RS05745--Tenacibaculum_mesophilum--61.84 (SEQ ID NO:2561)
MKIYFVYILQCSDKSYTGVTSNLSQRIHQHNGTGFQDSYTSKRPPVELVYCFEFTNVEMAIEKEKQIKKWSRSKKEALINNEFEKLPNLAKKKFN
<BUC16_RS08850--Flavobacterium_fluvii--64.63 (SEQ ID NO:2562)
MLNPQLGFHSYVYVYIILTNRKYSTFYIGITNNLKI RLQHQHKNNIETNTKT FASKYNI EFLVFEYKFTWVQEAIAAREKELKKWRDRKKLALIRDFNPKLEFLDHH
FE
<BUC16_RS02605--Flavobacterium_fluvii--58.14 (SEQ ID NO:2563)
MIEYTEGIYSFYVYIILTNRKNTVLYTGVTTNDLKRRLQEHKEKTNPNSTFARYNVEFLLYEHEFTWIIQLAIAAREKEIKDLKRDKKEILINTINPDWNFLNDLF
<BUC29_RS14515--Flavobacterium_defluvii--64.56 (SEQ ID NO:2564)
MLKQEGYHTYIYIILTNRKYTVFYTGVTNNLKI RLQHQHKNNIETNTKT FASKYNI EFLVFEYKFTWVQEAIAAREKEIKDWRRREKKIELIKTINPDLDLFLNYLF
E
<BUB81_RS06575--Marisediminitalia aggregata--72.53 (SEQ ID NO:2565)
MSERYSAVYILSNFTRTVLYVGVTSNLPQVRVYQHKMSMASGFCSTRYVNDLVYEMHEEYAAITREKQLKRWRRSWKEKLI TQKNPQWLDLYPLIVG
<BUA99_RS06160--Flavobacterium_segetis--67.11 (SEQ ID NO:2566)
MIEYTEGIYNFYVYIILTNRKNTVLYTGVTTNNLKRRLNEHKNHINAKSFTARYNVEFLIYYEHFGWIIQLAIAAREKEIKDWRRRELKLNLI TAFNPEMI FLNSHFE
<BUA99_RS05820--Flavobacterium_segetis--65.00 (SEQ ID NO:2567)
MIEFQEGYHTYVYVYIILTNRKAKTVLYTGVTTNNLKI RLQHQHKSLSNPDSFTAKYNVHYLYFYEKFAWIIQIAI SKEKEIKGWKKDKKI TLIKTINPELNFLNDLFE
<BUA99_RS05810--Flavobacterium_segetis--64.94 (SEQ ID NO:2568)
MEFQEGFHTYVYVYIILTNRKAKTVLYTGVTTNNLKVRLQHQHKSLSNPDSFTAKYNVHYLYFYEKFAWIIQIAI SKEKEIKGWKKDKKI TLIKTINPELNFLNDLFE
<BUB42_RS06840--Flavobacterium_micromati--65.26 (SEQ ID NO:2569)
MKPGFIYIILTNRKYTVVCTGVTSNLPRIINHKNKYPNSFSARYNVNII LVYEQFQSIGDAVTREKQIKAGSREFKNNLVGSVNPAPWDLFEDIKDITME
<BUB42_RS08850--Flavobacterium_micromati--62.50 (SEQ ID NO:2570)
MIEYTEGIYNFYVYMTNKNRTVLYTSVTNNLKRRLNEHKNHINASSFTARYNVEFLIYYEHFGWIIQLAIAAREKEIKDMRRELKLNLIKEFNPEMKFLNSHFE
<BUC10_RS08710--Flavobacterium_micromati--60.76 (SEQ ID NO:2571)
MKIYFYVYILKCADQTYTGVTSNLTNRKIQHTEGKYQNSYTFRRPIIILVYFSEFTCIKTAIEKEKQIKKWSKAKKEALI QENYHLLPNLAKKKFQ
<BUB25_RS02995--Flavobacterium_granuli--67.06 (SEQ ID NO:2572)
MKPGFIYIILTNRKYTVVYVGVTSNLPQRIIEHKNKKYPKSFSARYDVNLLVYEQFQSIGDAVTREKQIKAGSRAAKNKLIQSINPT
<BUB30_RS25065--Chryseolinea_serpens--65.91 (SEQ ID NO:2573)
MYRGCVYIIVNNRHTVYVGVTSNLTNRKIQHTEGKYQNSYTFRRPIIILVYFSEFTCIKTAIEKEKQIKKWSKAKKEALI QENYHLLPNLAKKKFQ
<BUB03_RS05205--Chryseobacterium_oranimense--65.52 (SEQ ID NO:2574)
MTNRHHTTLYTGVTSNLPQRLQHKERFFEQSFTSRYNLYLLVYVESFQEIQDAISREKQIKAGSRQKKLDLINSNPEWRDLTEDIENIMDVF
<BUA96_RS03295--Leeuwenhoekiella_palythoae--71.76 (SEQ ID NO:2575)
MBELFYAYIILTNRKNTVLYIGVTTNDLKRRLQHQHKSLSNPDSFTAKYNVHYLYFYEKFAWIIQIAI SKEKEIKGWKKDKKI TLIKTINPELNFLNDLFE
<BUA96_RS05630--Leeuwenhoekiella_palythoae--61.36 (SEQ ID NO:2576)
MKQSYVYIILTNTYRTTFYIGVTSNLPQRIIVDHKEKKGSFTSKYNVCGLIYFEEFSDINQAILREKQKLNWKKDWKLNLIKEKNPKLQTLNIFS
<BUR19_RS10455--Chryseobacterium_zeae--61.63 (SEQ ID NO:2577)
MIEFPEGAYTFYVYIILTNRKRTVLYTGVTTNNLHILRHQHKTKTNSNSFTAKYNVEFLVYVEKFGWIIQHAIAAREKEIKDLNRNKKLEFIRIQNPNEFLNHLF
<BUB62_RS11830--Hymenobacter_daecheongensis_DSM_21074--54.02 (SEQ ID NO:2578)
MFAHCYVNVITNPAKTSVYTGVTNDIRRIQEHFGNRGNPNTFAGRYFCYNLVYYPFEHIDGAIAREKEIKGWTRAKKDALIELDNPQWLTNLNAYEL
<B6K08_RS09630--Anaerovibrio_lipolyticus_DSM_3074--55.95 (SEQ ID NO:2579)
MRAFTYILECADGLTYTGTTDEKRLSTHNSGKASKYTRARLPVKLVYLECKDKETAMSRWHIKQLSRAEKLKLIENKVK
<BUC07_RS05105--Tangfeifania_diversioriginum--73.33 (SEQ ID NO:2580)
MANFYVYILANKPNGVLYIGVTNELERRLGEHKNKILKGFTYKYNVDKLVYFEEFDTYEEAFLRERRLKKWNRDWKIQLIEKENPDWIDLAEWF
<BUC07_RS12955--Tangfeifania_diversioriginum--62.07 (SEQ ID NO:2581)
MRDHNFYIYIVTNRKNTVLYLGVTTNDLQRRRIYEHHEGLMPGFSKKYNCHFLIYYEHLQDINEAIAAREKEIKKWRREKKEKLIINRPNPDWWFLNNEIYEL
<BUC03_RS06710--Mesonia_phycicola--67.09 (SEQ ID NO:2582)
MKTSYVYIILTNRKYRTTFYIGITSELTKRIAEHQLGKGSLSFTAKYNLYLYYEEFLDIQQAIREKQKLSWKDKDWKIDLVKKKNPELKTIELI
<BUB46_RS12215--Aequorivita_viscosa--65.06 (SEQ ID NO:2583)
MTNKSNNVLYIGVTTDDIDSRVEEHKKHVPQSFRTARYNCNKLIYFEEFDNGHKASQREKQMKKNWREWKIKLIEEMNPSWFDLSLNWNNIDLIYKTKRISPKN
LKL

FIG. 11CCCC

<BUB46_RS014250--Aequorivita_viscosa--63.64 (SEQ ID NO:2584)
MIEFNEGSHYYVDIITNKYRSTYYIGVTNNLGRRLRQHRENIKGIKTFASRYQIQHVLVYYEEFSWQLAIAREKELKGVRRDRKKLALIRSFNASFEFLEDR
FLVE
<BUB46_RS08435--Aequorivita_viscosa--56.67 (SEQ ID NO:2585)
MPGYTYITTNKNKTVLVGATDDIRRRIMENKTKIYKNAFTAKYNCNLIIVYFEEYEDVNDADFLEKQQLKAGNRKRKEALINEMNPGWNDSINWKNDEFELR
<BUC09_RS03550--Flavobacterium_terrae--65.91 (SEQ ID NO:2586)
MNEFNPQIGFHSYYVYIITNKYRSTLYIGITNNLKLRLQHKENIRLSNKTTFASNYNLEFLVYEEKFVWVEDAIAREKELKKWRDRKKTTELIKNFNPKFEFLN
HHFEQL
<BUC09_RS13015--Flavobacterium_terrae--61.84 (SEQ ID NO:2587)
MENSKFYVYILYSVTVDKFYIGYTSNLEERIIRHNQKSKGFTGGFNDWELKYHEVFSKHEAMTREKQIKNWKSRTKILELINKI
<BUB56_RS13225--Geosporobacter_subterraneus_DSM_17957--68.54 (SEQ ID NO:2588)
MYVYMLANWSNKVIYTGMTNNLERRIYEHKHLVKGFTKYNINKLVYFESTTDVTAAILREKQIKGWTRSKKNELIKMMNPQWKDLSRDFYEDSSLHSE
<BUB99_RS18955--Reichenbachiella_agariperforans--63.95 (SEQ ID NO:2589)
MSRAKGGYIYIMSNVHRTVLVYIGVTADLYTRAYQHKSGHGSATHRYNVCVLLYYEIFTDIKSAIAAREKRLKKYKREWKWELIDGFNPDRVDLFDQVSEM
<BUR94_RS05615--Oceanicola_litorea--69.23 (SEQ ID NO:2590)
MAVYMTNKPRTGLYIGVTSGLVGRVWQHRTHAFPGFTDLYNLDKLVWYETHEEPTVAIQREKSLKRWYRRWKIDLEKANPGWRDLWYIECPGAEEADV
<BUA97_RS05945--Chryseobacterium_molle--54.43 (SEQ ID NO:2591)
MKQYFAYILRCCDNSYTGVTSDLESRLIKHQSGHYPESTHSRRPTQLVYFNEFNEIEQAIAFEKQVKGWSRKKKEAIIKNKW
<B5522_RS03215--Bradyrhizobium_lablabi--76.60 (SEQ ID NO:2592)
MSYVYIILASRKDGATYIGVTNDLVRRRIYEHRTAPGPGFSRYNI TRLVWFEIYDDPVSAISREKELKKWKRAWKVQLIEAQNPWADDLYESICQ
<B5522_RS10375--Bradyrhizobium_lablabi--74.44 (SEQ ID NO:2593)
MSYVYIILASKPGGTYIGVTNNLVRRVYEHREGLADGFTERYGKTLVYFEAHESIAVALQREKNIKHWSREWKIDLIVAGNPDRDLWYEEIVR
<B5522_RS35050--Bradyrhizobium_lablabi--74.12 (SEQ ID NO:2594)
MNGPFVYVILASQRNGTYIGVTNNLARMSSEHAKLVRGFTKKYGVDDLHVHEAFDSILEAREHSLKRWRAWKIKLIEESNPGWSDLTSQLNS
<B5522_RS12815--Bradyrhizobium_lablabi--68.24 (SEQ ID NO:2595)
MYVYVILASRRHGTLYIGVTNSLAKRMBQHRNNEGSSFVKTYGVYRLVYVESYERABEAIAREKQLKRWKRDWKIELIERENLEWRDLSDLIV
<B5522_RS08670--Bradyrhizobium_lablabi--63.53 (SEQ ID NO:2596)
MSKARAYVYVILVSGQHTLYIGVTNNLRARLEHRAGRGSDGFQQYKYVRLVHVETFATALEAIAREKQLKNWQKWKIRLIEENNLDWDDLHLL
<BUE15_RS14235--Hymenobacter_psychrotolerans_DSM_18569--59.77 (SEQ ID NO:2597)
MYVYILTNPQTQVLYIGVTNDLQRRLYEHSSRRGDARKFTGRYQADLLVYFEICPHAVQAIAREKQLKGWTRAKKDALIAAFNPSPWQAIIDLETWAG
<BUE15_RS12425--Hymenobacter_psychrotolerans_DSM_18569--57.65 (SEQ ID NO:2598)
MYVYILTNPQTQVLYIGITNDLTRLRYEHSSRGSDAGKFTGRYQADLLVYFELCQDQAIAREKQLKGWTRSKKNALITAFNPEWQLIDLETWSG
<BUE15_RS09965--Hymenobacter_psychrotolerans_DSM_18569--52.87 (SEQ ID NO:2599)
MPYVYVITTNPAKTVLYIGITNSLASRMAQHYANRGSDQTFAGRYFCYNLVHFEYTDSKAAIAREKELKGWTRAKKEVLIAATGNPAWDFLET
<BUB07_RS10245--Muricauda_antarctica--74.73 (SEQ ID NO:2600)
MIKNVHQYVYVILSNKFGNTLYVGMTNDIQRRLVLEHKQSIDGFTKKYGLNLVYFETHHNVDNAISREKQLKKWKRLWKIELIEKENPKWDLATDWN
<BUB07_RS02900--Muricauda_antarctica--65.85 (SEQ ID NO:2601)
MNKYVYVILTNKNKTVLYTGVTNDLTRLRIEHIENIEQGKRTFAAKYKCKHLVYEEYAWVQDAISREKEIKGWVRMKKLQLIKTFNPNMDFLEDGFLFR
<BUB80_RS21135--Chryseobacterium_polytrichastrii--63.86 (SEQ ID NO:2602)
MKAGFIYIMTNKKNNTLYTGVTNSLNPRIQEHKEKFHELSFTAKYNLEKLVYWSFQEIIGDAIFREKQIKAGSRQKKLNLINSINPEW
<BUA86_RS02860--Chishuiella_changwenlii--62.82 (SEQ ID NO:2603)
MKIYVYVILECSDKSFYVGVTVNDVFKRFIEHQDNDYKSYTFNRKPLKLVFTSFNNIEIAIEKEKQIKKWSRSKKLALIQDYFEHLPNLSSKKF
<BUB88_RS15195--Chryseobacterium_contaminans--63.95 (SEQ ID NO:2604)
MKAGFIYIMTNKKNNTLYTGVTNSLNPRIQEHKEAYYSQSFSTRYKLYILVYWEAFQEIIGDAIAREKQIKAGSRQKKLNLINSINPEWNDLTDITQMFEEIAS
SLRSSQ
<BUB53_RS04765--Roseovarius_pacificus--77.53 (SEQ ID NO:2605)
MPKDMYVYIVTNKRHGTLYIGVTNDLVRRVWEHRTHAISGFTDRYNLTRLVWYEHVPEAAAIRREKALKRWSRDWKIELVEQANPDWDFWQEIISGQPDLLH
R
<BUC11_RS09605--Flavobacterium_xanthum--58.43 (SEQ ID NO:2606)
MEKGFTYILTNNKQTVLVYGATKNLKKRIDTHISGAGAVFTKKYNATILLYYEEFDQYSDAFKREKQLKNWHKDWKLLIKLSNPELKDLYLSL
<BUC11_RS08510--Flavobacterium_xanthum--65.00 (SEQ ID NO:2607)
MIEYTEGIYNFYVYIMTNKNRTVLTYGVTNNLKRRLNEHKNHINASSFTARYNVEFLIYYEHFGWIQAIAREKEIKDMRRELKLNLIQELNPEMKFLNSHFE
<BUC11_RS10260--Flavobacterium_xanthum--62.50 (SEQ ID NO:2608)
MIEFQEGYHTYVYIILTNKAKTVLYTGVTNNLKI RLQHQHQSLSNPSFTAKYNVHYLLYFHKFTWIQAIAREKEIKGWKKDKKILTIKTINPELNFNLHLE
<BUC11_RS15105--Flavobacterium_xanthum--56.98 (SEQ ID NO:2609)
MKFYVYVILRCIDDSLYVGVTSDIERRVMEHNMKGYPEAYTHSRPVLVLFYQDFADPNQAIIEFEKKIKKWSRMKKQALIDENFDKLDLSECRNATHHKYES
DVEDIEE
<BUC11_RS10270--Flavobacterium_xanthum--61.25 (SEQ ID NO:2610)
MIEFQEGYHYKYVYIILTNKAKTVSYTGVTNNLKI RLQHQHQSLSNPSFTAKYNVHYLLYFHKFTWIQAIAREKEIKGWKKDKKILTIKTINPKLNFLNLHLE
<BUE22_RS13160--Flavobacterium_pectinovorum--62.50 (SEQ ID NO:2611)
MQQFQEGFHTYYIITNKTKTVFTYTGVTNNLRI RLNQHKENILLKNQTFASKYIVFEFLYYEKFETWQDAITREKEIKGWRREKKLELIKTINPDFDFNLNLF
E
<BUA94_RS06680--Labrenzia_suaedae--76.47 (SEQ ID NO:2612)
MQRQFWVYILATHKNGTLYTGVTNDLLRLRYEHQSESVPGFTARYQVKRLVWYEEHGTADRAIAREKAIKTWKRWSWKVELIEESNPEWFDLGRFLNG
<BUA94_RS06685--Labrenzia_suaedae--75.29 (SEQ ID NO:2613)
MPRQFWVYILATHKNGTLYTGVTNNLLRLRYEHQSESVLGFTTRYQVKRLVWYEEHATADHAIAREKAIKTWKRWSWKIDLIEESNPEWFDLGRLLSR
<BUA94_RS15300--Labrenzia_suaedae--71.11 (SEQ ID NO:2614)
MSGFVYMSDKPRGVIYTGVTSDLPKRIWEHREVRKGFTTKYKAKSLVWFESHPSIVLAIQREKSLKRYRRDWKIELIEGLNPTWLDLYERMGEIENVYQPH
RNTRMWDDYN

FIG. 11DDDD

<BUA98_RS12755--Flavobacterium_saccharophilum--63.64 (SEQ ID NO:2615)
MLFQEGFHTYYIYILTNNKPKSVFYTGVTNNLRKRSLSQHQENILLNKNTFASKYNVEFLLYYEKFTWIHEAIAREKEIKGWRREKKIELIKTINPDFEFLNYLFE
<BUA98_RS21325--Flavobacterium_saccharophilum--63.29 (SEQ ID NO:2616)
MQLHEGYHTYYIYIANKHKTVFYTGVTNNLKIRLQHQHRENNTTGKNTFASKYHAIYLLYEKFTWVQEAIAREKEIKDWRREKKIELIKTINPNLDFLNDLFS
<BUB89_RS00335--Flavobacterium_flevense--68.42 (SEQ ID NO:2617)
MKPGFYIITNKYQTVVYVGVTSNLPQRRIEHNKKNKHPSTFSARYDVSIILVYEQFQWIEDAIAREKQIKAGSRAAKNKLIQSINPPWKDLFDEIRNIMTE
<BUB89_RS13970--Flavobacterium_flevense--63.64 (SEQ ID NO:2618)
MYELNQGYHTYYIYIITNSYRSTFYIGMTNNLKQRLQQHHQNIIVDSNKTFAAKYNI EFLVYVEKHTWVHQAILREKELKSWRREKKLELIRSFNSDFEFLNDR
FKSRDAIPPASE
<BUB89_RS08400--Flavobacterium_flevense--62.03 (SEQ ID NO:2619)
MQFKEGYHTYYIYIITNKAKTVLYTGVTNNLKIRLQHQHQNIEFNNTFASRYQVHYLLYEKFTWVQEAIVREKEIKGWRREKKLDLIKIINPDNLFLNDLFI
Q
<BUC23_RS08800--Flavobacterium_xinjiangense--63.16 (SEQ ID NO:2620)
MKPGFYIITNKYQTVVYVGVTSNLPQRRIEHNKKNKHPSTFSARYNDIILVYEQFQWIGDAIAREKLIKAGSRETKNELIRSLNPTWQDLFEEIKNIMTE
<BUC23_RS11895--Flavobacterium_xinjiangense--53.49 (SEQ ID NO:2621)
MKFYVYVILLCNDTLYVGVTSNINRRVMEHNAGKYPEAYTHSRSSVSLVYQDFTNPNQAI EFEKKLKKWSRVKKQALIDGGFEMQLDSECRNATHHKYKPF
DEGDIEE
<B5488_RS06880--Salegentibacter_salegens--63.83 (SEQ ID NO:2622)
MQKSYIYLLSNSTRMIYIGVTSNLPKRI TEHKEIGSSFTRKYHLFLIYFEEFSEIHQAIAREKQKLNWHKDWKWNLA KSKNPSLKDLYEDL
<B5488_RS09880--Salegentibacter_salegens--66.67 (SEQ ID NO:2623)
MKASWYIITSKNNTVLYTGVTNLPKRIQQHKEKFFELSFSTSKYNVCKLVYWSQFQEI GDAIFREKQIKAGSRQKKIDLINSINPEWRDLTEDIKDIMNPF
<BUA38_RS15980--Bradyrhizobium_erythrophlei--78.65 (SEQ ID NO:2624)
MRSRNLVYVILASKIGGTVLVGVTDNLVRRVAEHLKAVEGGFTEKYDVGRVLVYFEQFDDPENAIKREKRLKKWNRAWKIRLIEENPNPWDDLRYGIAGGD
<BUR17_RS06445--Chryseobacterium_scophthalmum--63.33 (SEQ ID NO:2625)
MKAGFIYIMTNENNTVLYTGVTNLPKRIQQHKEKFFELSFSTSKYNVCKLVYWSQFQEI GDAIFREKQIKAGSRQKKIDLINSINPEWRDLTEDIKDIMNPF
<BUR17_RS05305--Chryseobacterium_scophthalmum--58.06 (SEQ ID NO:2626)
MKILGTHNYVYVILTNYNKTVLVGVTNDLKMRYHHKNPTTDEKLHFTTKYKCFNLVYVEHFQDVEQAIAREKQIKGYSRMKKENLIK SINPNWDFWNKDIE
<BUR17_RS08100--Chryseobacterium_scophthalmum--63.22 (SEQ ID NO:2627)
MIEFSEGIYTFYVYILTNNKRKTVLYGTINNHLKRLFHQSTKQNSFSTAKYNVEFLIYVEKFGWIHKAIERENEIKNFSRSKKLKLIRCTNPNLDFWNLYLFE
M
<BUA05_RS09500--Porphyrobacter_sanguineus--70.45 (SEQ ID NO:2628)
MARGGWYIITNKPRGVLYIGVTADIVRRVWQHRNSLGSFVRKYGLDRLVVEAYPTIEBAIAREKQKLNWHRQWKIELIENDNPAWKHVEPDGSSPG
<BUA05_RS00050--Porphyrobacter_sanguineus--67.78 (SEQ ID NO:2629)
MEREKQGGWLYIMADRYRTIYVGSTADLAFRMAQHRDQGSGKFCAEYGLTRLVWAEELLPTMQEAMDQEKRVKRWRRDQWKFALVEKGNPDWRDLDFGIIN
<BUA05_RS11300--Porphyrobacter_sanguineus--61.63 (SEQ ID NO:2630)
MTNKRGGVLYIGVTADLPARIMQHKQSGSAFCRRYGLDRLVCAEPHAEATAAIAAREKAMKAWKRAWKIEQIEKSNPEWGD LFEQIN
<BUA05_RS02590--Porphyrobacter_sanguineus--62.79 (SEQ ID NO:2631)
MTNKRGGVLYIGVTADLPARIMQHKQSGSAFCRRYGLDRLVCAEPHAEIADAITREKAMKAWKRAWKIEQIEKSNPEWGLFEQIN
<BUA05_RS02285--Porphyrobacter_sanguineus--63.95 (SEQ ID NO:2632)
MTNKRGGVLYIGVTADLPARIMQHKQSGSTFCRRYGLDRLVYEPHAEIVAAIAAREKAMKAWKRAWKIEQIDKSNPEWGD LFEQIN
<BUA05_RS08265--Porphyrobacter_sanguineus--60.61 (SEQ ID NO:2633)
MTNKRGGVLYIGVTADLPARILQHKQSGSAFCRRYGLDRLVYAEPHAEIADAIAREKAMKAWKHA
<BUA05_RS00010--Porphyrobacter_sanguineus--60.61 (SEQ ID NO:2634)
MTNKRGGVLYIGVTADLPARILQHKQSGSAFCRRYGLDRLVYAEPHAEIVAAIAAREKAMKAWKHA
<BUE95_RS03950--Flavobacterium_cucumis--65.12 (SEQ ID NO:2635)
MIEFTEGIYNFYVYILTNNKAKTVLYTGVTNNLKRLEEHNEKNLNPSTAKYNLYIYVEKGYWQQAIAREKEIKLLRRDLKLELIREFNPNMEFLNKSFS
<BQ8482_RS16200--Mesorhizobium_delmotii--66.29 (SEQ ID NO:2636)
MTGYAYMTASQKRGTIYIGVTNDLGRRLPEHKSGQSGRFTSRYGVRQLVWYEEHFDIRDAIQREKSLKRWPRQWKIELIEKTNPEWFELFRGTGW
<BQ8482_RS37500--Mesorhizobium_delmotii--66.29 (SEQ ID NO:2637)
MTGYAYMTASQKRGTIYIGVTNDLGRRLPEHKSGQSGRFTSRYGVRQLVWYEEHFDIRDAIQREKSLKRWPRQWKIELIEKTNPEWFELFRGTGW
<BW976_RS27120--Mucilaginibacter_lappiensis--71.76 (SEQ ID NO:2638)
MNIHRYFVYIVANKVNTVFYIGVTSNLEQRIWEHKQKIYEGFTAKYDCSKLVYEEYQWIDDAIAREKQLKAGSRKKKIDLIILTENPSWSDLSSEWYN
<BW976_RS27870--Mucilaginibacter_lappiensis--62.07 (SEQ ID NO:2639)
MSQYNFFTYILTNYNKTVLYTGVTNDLNQRLYEHYFGVAQGESFTSKYKCYLIWFERHQFINHAIEREKEIKGWTREKKVKLIEQENPNWDFLNKEIMWPP
IFTD
<BW976_RS20410--Mucilaginibacter_lappiensis--54.76 (SEQ ID NO:2640)
MWNHNYVYIITNPDRTVLYVGVTDNIRVRLIQHTENKGTGRSFTGKYCYLLVYVEHFTHIEHAIEVEKEIKRWRREKKEALIAATTNPDWIFLNEQVSGE
<BW994_RS11780--Acidiphilium_rubrum--74.19 (SEQ ID NO:2641)
MRVERQPAVYMMANRRNTIYVGVTSNLPQRRIEHNKKNKHPSTFSARYNDIILVYEQFQWIGDAIAREKLIKAGSRETKNELIRSLNPTWQDLFEEIKNIMTE
<B0B12_RS00285--Pontibacter_indicus--64.04 (SEQ ID NO:2642)
MISHNYFVYIITNPKKTTLYIGVTNDIVRQLBEHTLNAGNPSTFAGKYHCHLLMYERFSSVEHAIAREKELKKWSREKKEALIESINPNWLPLNAEVQDD
<B0B12_RS19185--Pontibacter_indicus--64.20 (SEQ ID NO:2643)
MKSHNYFVYIITNPSKTVLYVGVTDNLVYRLQHKDNRGKPEFAGRYFCYKLLYERYTYVQHAIDREKELKEWSREKKEALIKRENPLYRFLKADP
<BW972_RS11195--Pontibacter_lucknowensis--60.92 (SEQ ID NO:2644)
MSGHNYFVYIITNPKKTTLYIGVTNELTRMDEHVLNAGVPTSFAGRYFCYKLLYERYTYVQHAIDREKELKEWSREKKEALIKRENPLYRFLKADP
<BW972_RS10080--Pontibacter_lucknowensis--61.73 (SEQ ID NO:2645)
MKSHNYFVYIITNPSKTVLYVGVTDNLVYRLQHKDNRGKPEFAGRYFCYKLLYERYTYVQHAIDREKELKEWSREKKEALIKRENPLYRFLKADP
<BW964_RS13495--Chryseobacterium_shigense--65.26 (SEQ ID NO:2646)
MKPGFYIIMTNKNKTTLYTGVTNLPKRVQEHKEMYQQSFTSRYNLYILVYWSQFQEI GDAIFREKQIKAGSRQKKIDLINSINPDWRDLADDIENIMDIF

FIG. 11EEEE

<BW955_RS11580--Kroppenstedtia_eburnea--57.89 (SEQ ID NO:2647)
MEMKRVERNYVVYILECDGTLTYGTNNLQRLMHQOGTGAKYTRGRGFTTLKWLGETITRSQALRREREIKKMSRTKKLELIQRRPAGEATKEF
<BW932_RS06060--Chryseobacterium_chaponense--61.54 (SEQ ID NO:2648)
MKTLLGTHNYVYIILTNKMTVLYTGVTDNLKARLYWHQNPETAIKSHFTAKYKCFLLYYESFQDIETAI SREKQIKGWTRKKKDDLI TGFNQTRFLNKDIEE
<BW951_RS21160--Chryseobacterium_gambrini--64.21 (SEQ ID NO:2649)
MKPGFVYIMTNKNNTVLYTGVSTLTPKRIYQHKEKIFELSFTSKYNIDKLWYWSFQEI GDAIFREKQIKAGSRQKKIDLINSINPEWRDLTEDIKDILNPF
<BW951_RS09900--Chryseobacterium_gambrini--62.07 (SEQ ID NO:2650)
MKTLLGTHNYVYIILTNKMTVLYTGVTDNLKARLYWHQNPETIDKSFTFRYKCFYLVYFEHFKDIDVAINREKQIKGWIRSKKESLINEFNPTWKFLNDEIK
<BW951_RS04175--Chryseobacterium_gambrini--59.30 (SEQ ID NO:2651)
MFELNEGIYKFYVYILTNKNRTVLYTGVGTGNLHQRLSQHINKLNPDSTAKYNVCFLIYYQSFGWIHQAIEREKEIKLLKRDKKLELIRAINPNLEFLNYLFE
<CZ814_RS03025--Photobacterium_toruni--72.04 (SEQ ID NO:2652)
MNKEPVVYIILASHNKQVLVGVTSQVLVTRIWHQKNNIVRGFTQKYQVHSLVYERFDSMTAIKREKQLKKWRRSWKEQLIKDNNPQWLDLYDSL
<B5D38_RS06970--Consotaelia_salsifontis--71.11 (SEQ ID NO:2653)
MGGHVYIMASQKNGTLYIGVTSDLRRVHEHKTEALNGFSSRYGCKRLVWYBEHADIRDAIAREKALKKWTRARKIALIEALNPDWQELYRGMGW
<B5D40_RS06965--Fibrobacter_intestinalis--73.17 (SEQ ID NO:2654)
MTNFTNTVLYTGVTDNLVRRAYEHKKNLFEGFTKRYNVTKLVYFEIFDDITLAIQREKQLKAGSRQKKIELIEKNNAEWKDLYESIVQ
<B5D62_RS15375--Enhydrobacter_aerosaccus--63.86 (SEQ ID NO:2655)
MGRRPALYIDVTSDLGRRLLEHRRGLSQHTSRYNITKPTYIEAHTAPNAIEREKQLKRWREKKIALIDANPEWQDLSGEIL
<B5X82_RS18860--Sphingomonas_histidinilytica--82.95 (SEQ ID NO:2656)
MEKQPSVYIMASGRHGTLYIGVTSDDLGRVHQHRESLIKGFTSRYGVSRLVYEGFDDMITAITREKQLKKWNRAWKIELVEARNPDWDDLAVTTFGFPFPGP
<B5X82_RS14785--Sphingomonas_histidinilytica--78.41 (SEQ ID NO:2657)
MDRQPCVHILASDRNGTIYIGVTSDLPKRLYQHRSGTLDGFTSRYGVAKLVRYLEFEDMSTAIAREKQLKRWHRSWKLNLI EQDNAGWHDLSIGLGFEP
<B5X51_RS11355--Salegentibacter_holothuriorum--69.35 (SEQ ID NO:2658)
MNNDYQYVYIITNRKNGTLYIGITGNLQRLYQHRNRNLKSFSSRYRLNKLVFYKIEYPSLAIKREKQLKKWNRLWKINLINDFNPWDGLTVRFK
<B5X51_RS11265--Salegentibacter_holothuriorum--62.77 (SEQ ID NO:2659)
MQKSYIYILSNSTRMTIYIGVTADLKKRINEHKS DIGSLFARKYHLKFLIYFEFSEIHQAIAREKQLKNWHKEWKWNLVKSMNPDLDLYEDL
<B5X51_RS08435--Salegentibacter_holothuriorum--61.05 (SEQ ID NO:2660)
MQKSWYIITNKNSTVLYTGVTSNLQKRMFEHKS KVNKSFAAKYNCEKLVLYQEFSDVTQAIEYEKKIKAGNRARKEKLVNAQNPDWIDLKMDILG
<B5X51_RS11250--Salegentibacter_holothuriorum--64.79 (SEQ ID NO:2661)
MKKSYIYILSNSTRMTIYIGVTSDLKKRITEHKSEMGSLFTRRYHLKFLIYFEFSEIHQAIAREKQLKNWHKE
<B5X72_RS18050--Parapedobacter_luteus--53.93 (SEQ ID NO:2662)
MNNYIYILTDNCNRTCLHVGMADQLERAIKTYREAPGLFFDACAKVSRVLVYHEALPTEESALRRFRELSTYTRMQKEKLIKRYNPNVWDLGLVRPLNAPGSSGL
TNARLHLRS
<B5X76_RS15520--Sphingobacterium_nematocida--65.91 (SEQ ID NO:2663)
MERGGAIYILTNVNRVLYIGVTSDLYSRLVEHREKRYNSFTSRYSNLYCVYFESFPTIEAIDREKQLKKWSRAKKENLINKNPQWNDLWEIVKEW
<B5Y19_RS04450--Bosea_thiooxidans--73.33 (SEQ ID NO:2664)
MPYFVYLLARKKDGTLYLGVTSDLARRVHEHKSQTIGFTSRYDVRDLVWYEEFEWVNEAIDREKRVKKWRRRAWKIKLIEDMNPEDADLYEGLA
<B5Y19_RS14425--Bosea_thiooxidans--72.09 (SEQ ID NO:2665)
MANKPRGTLYLGVTVDLVRRVSEHKQGIGRSFTARYNLDKLAWFEVFEHIELAIQRETS LKRWSEWKIELIERKNPGWRDLFEDIYPPSFSAEA
<B5X64_RS04075--Chryseobacterium_balustinum--62.07 (SEQ ID NO:2666)
MIEFSEGIYTFYVYILTNKRKTVLYTGTNNLHKRLFQHS TKQNSNSTAKYNAEFLIYKFGVWHKAIEREKEIKNFSRNKKLELIRCVNPNLDFWNHLE
M
<B7E01_RS10540--Varibaculum_timonense--68.67 (SEQ ID NO:2667)
MLTNWNNKVLYTGTNNLERRIYEHKNKLVKGFTSKYNVNKL VYFEYTS DVNSAIAREKEIKGWVRRKKNELIESINPEWRDLSEEF
<B8779_RS06555--Nitratriuptor_tergarcus_DSM_16512--70.11 (SEQ ID NO:2668)
MTSMNHRAIYIGVTSNLVKRVYEHKNGLCDGFTKRYHCTKL VYFEVFDNIDEAINREKYLGKKRDFKIALIEKQNPWRDLYDDIL
<B9A24_RS13240--Hymenobacter_roseosalivarius_DSM_11622--60.00 (SEQ ID NO:2669)
MYVYIILTNPTQTVLYIGVTNDLIRRLYKHSSGRGDAGKFTGRYQADLLVFELCPDATQAIAREKQLKGWTRHKKDMLITAFNPTWQVIDLETWTR
<B9A24_RS14760--Hymenobacter_roseosalivarius_DSM_11622--54.76 (SEQ ID NO:2670)
MPYVYIITNPAKTVLYIGITNSLPQRMAQHYDNRGTNQT FAGRYFCNLYVFEHMDSKTAIARENELKGWTRAKKEALIASENPEWFLQGS
<B9A12_RS15700--Desulfacinum_hydrothermale_DSM_13146--50.65 (SEQ ID NO:2671)
MWFVYICQRGGKLYTGTITDLQHRMTQHKQALLYEPHPDKFSAARREKQIKGWRRREKKLALCHKKPS
<B9A58_RS11485--Reichenbachella_faecimaris--66.29 (SEQ ID NO:2672)
MAKGGYIYIVSNKNRTVLYTGVTSNLHNRIYEHKNGLGASFTKKYNCTDLYYEFFTDIESAIHREKRIKKYKREWKENLIHEMNPEDRDLFDEISEMK
<B9A58_RS18375--Reichenbachella_faecimaris--65.93 (SEQ ID NO:2673)
MQKGGAVYIILNRNHSTLYVGVTSDLYSRIVEHKEKKYKPSFSSRYNTTKLVYEGFGHSIEEAI VREKQIKGGLRKKKEDLIESINPEWNDLDFEVSSW
<B9A58_RS14890--Reichenbachella_faecimaris--56.32 (SEQ ID NO:2674)
MKQHNYFVYIITNPNKTTLYVGLTNDLPTRLDQH FENRGFKTYAGRYFCNLIYWERHQYIDHAIEREKEIKKWREKKEALINAMNPEWNFLNEKVRD
<B8977_RS04810--Moheibacter_sediminis--67.06 (SEQ ID NO:2675)
MQLSYVYILTNQYRTTFYIGVTSNLKKRIIQHKEKKSIFTSKYNLDILVYEEHLYITLAIQREKQLKNWKREWKNLIRQONPTFAELEI
<B8991_RS06925--Pseudoceanicola_flagellatus--64.71 (SEQ ID NO:2676)
MAHFVYIMASRPGAMYGRATDLRTRVEAHRAGLSVHTAKYNI RRLVWFEEHDCFEASLRRERAIKRWRRDWKIALITQSNPNWHELTPOIPD
<B9A52_RS07995--Aquiflexum_balticum_DSM_16537--62.79 (SEQ ID NO:2677)
MKRHEYFVYIILTNKKTLYIGMTNDLISRLNQHGEKFNPESTAKYQCKLYLVHWSYQYVQDAIAREKEIKKWREKKEKEDIKELNPDWKFLENEVGG
<B9000_RS26125--Tistlia_consotensis_USBA_355--75.00 (SEQ ID NO:2678)
MRQPCVYILASGRNGTLYIGVTSDIVARAWQHRSGVVEGFTDRYAVHRLVHVEFFGTMEEA ILREKKLKRWRRAWKIALIEEGNPQWRDLYEDLQR
<B9000_RS17780--Tistlia_consotensis_USBA_355--79.78 (SEQ ID NO:2679)
MKQPCVYILSNRRHGTLYVGVTSDIVIRRVWEHRTDAVEGFSRRYQLHRLVYFEFHD TMERAIQREKQLKLWRRAWKIQLIEGGNPWRDLYETLMP
<B9000_RS25645--Tistlia_consotensis_USBA_355--73.91 (SEQ ID NO:2680)
MRQPCVYILASRQNGTLYVGVTSDIVARAWQHRTDAVDGFTTRYGVHRLVHVESFGTMEEA ILREKKLKRWRRAWKIALIEQGNDPWRDLCDLQR

FIG. 11FFFF

<B9000_RS13940--Tistlia_consotensis_USBA_355--75.00 (SEQ ID NO:2681)
MRQPCLYI PASKRNTGLYVGTSDIATRVWQHRSGAVEGFTRQYRVHRLVQSAIRSDRIDLIGESPSRIM
<B9N75_RS09585--Sphingomonas_indica--70.00 (SEQ ID NO:2682)
MARTYVWYIMASRRNGTLYTGVTSDIARRVMEHKNKGATAGFTQRCVTVQLVHAEGFENVDEALAREKAVKKWRCWAKLEIEKENPGWRDLYDQVI
<B9N75_RS07485--Sphingomonas_indica--67.69 (SEQ ID NO:2683)
MARIWQHRSGATGGFTARYRVHRLVHYESYDDMEAAI LREKQLKRWHRGWKINLVEAANPQWGD LAIGLGFELPRQ
<B9Y87_RS12160--Arenibacter_troitsensis--72.34 (SEQ ID NO:2684)
MHKTIHQYYIYL TNKNGTLYIGMTNDLQRRRIYEHKNKLNDGFTKKYNLDRLVYFENHQFVDDAIKREKNLKNWKRWDKINLI IKDNPQWEDLASNWDYQ
<B9Y87_RS15155--Arenibacter_troitsensis--65.91 (SEQ ID NO:2685)
MKKSNVYLLTNKYNTVIYVGVTTDLVRRRIYEHKTVGRGFSYRYNCTKL VYVEEYSNMDMAIEREKQI KAGNRRRKEYLIRSTNPEWVDLAKEWLFPPYV
<B9Y84_RS04285--Marivirga_sericea--68.18 (SEQ ID NO:2686)
MVKGGAIYIMANKSHSTLYTGVTS DLYSRVVEHKQKIHPKSFTAKYCNKNLVYFESFHDIEEAIAREKQIKGGSRAKKIALIESMNPENWDLFEKVKEW
<CCE89_RS09620--Altererythrobacter_xiamenensis--68.18 (SEQ ID NO:2687)
MAKEPCVYIILASGRRGYIYIGVTSHPVGRWLQHRNDVFDGYAKRRNCKRLVRIETPTMVEAIAAREKQLKNWHRQWKINLIEBGDNGEWRDLAVEMGLS
<CCK86_RS06615--Sphingopyxis_terrae_subsp._ummariensis--77.01 (SEQ ID NO:2688)
MRERHPCVYIMASGRHGTYIGVTS DLMARVHQHREGLTGGFTRRYGVKRLVFHETHASMDAAITREKQLKAWKREWKIALIEKENPFWEDWAVRLGFRPLP
<CCK86_RS06165--Sphingopyxis_terrae_subsp._ummariensis--75.58 (SEQ ID NO:2689)
MALPGFVYIMASRNGTYIGVTS SELSQRIWQHREGVIEGFTKQYDCKLLVWFEEAFDSIHDARQFEARMKKWNRAWKINRIERNPDWADLFETLV
<CCK86_RS01460--Sphingopyxis_terrae_subsp._ummariensis--69.77 (SEQ ID NO:2690)
MSRTHYHYIILASGRNGTLYTGVTDGLAQRMEQHRDGKSNFTTKYGVTRLVHVEAFADVDTAIAREKAVKKWRRRAWKLELIERGNPQWLDLYDRING
<CCK86_RS12105--Sphingopyxis_terrae_subsp._ummariensis--65.56 (SEQ ID NO:2691)
MAGGHVYIMANRYRGTYIGVTAHLPARIFQHRTSNGSDFCKRYGLNRLVYAEPHNDIRDAIARERAMKKWNRAWKFRLEE QNPDWIDLFDALNA
<CCK86_RS01460--Sphingopyxis_terrae_subsp._ummariensis--58.43 (SEQ ID NO:2692)
MELRVTRNAHYVIMTNKPMGVLYIGVTDNLPARITAHRS GTGSEFCCKWQLKRLYVEFPNRIDEAIAAREKALKKWNRLWLKRLVCEANPEWADLFETISG
<FN756_RS06900--Solitalea_koreensis--69.64 (SEQ ID NO:2693)
MRFHQYFVYMLERRIYEHKSKIFKGTSTKYNCDEL VYFENFDLIDDAISREKQLKAGSRQKKIDLING
<FN795_RS19530--Flavobacterium_resistens--64.44 (SEQ ID NO:2694)
MFNPQIGFHSYVYIITNEHRTSFYNGVTNNLKQRLAKHKNIDNLINSFAGKYNIQFLVYEEKFTWIQEAIAAREKELKKWRDCKIELIRSFNETFEFLNFH
FE
<FN817_RS15345--Chryseobacterium_rhizoplanee--65.12 (SEQ ID NO:2695)
MIEFSGKGRYTYIYILTNKNRVLYTGVGTGNLHRR LHQHKTCLNSGSFTARYNLEFLIYIEKYDWIHHAIEREKEIKNWSRIKKLELIRRTNPNLEFLNYLFE
NWF
<FN815_RS18035--Gracilimonas_mengyeensis--65.17 (SEQ ID NO:2696)
MSKKGYLYIILSNKHRTVYVYTGVTSQLLERVVDHRRGRGCHFTKKYNI RELLYEEFPTMLEAIQAEKRVKRWKQAWKIELIKTINPEMVDLWSEVSTLGQPGL
LDDEDG
<FN815_RS07475--Gracilimonas_mengyeensis--65.56 (SEQ ID NO:2697)
MGNKDLYVYLLTNKNHTVIYCGVSSDLLNRVWVKHKNKVYDGF TKKYNVDKLVYEGFEDPIYAIEREKQIKAGSRQKKIDLIISMNPEWRDLYEDLIEGFK
FDEENDNSDS
<CALSC_RS15920--Calothrix_rhizosoleniae_SC01--70.77 (SEQ ID NO:2698)
QRRRIYEHKSKLIEGFTKKYNI TKLVYETIYEDAYTAISREKQIKAGSRQKKIDLINNLNPGWKDLYEEF
<C1721_RS05415--Parabacteroides_bouchesdurhonensis--62.50 (SEQ ID NO:2699)
MIKNVHQYWIYMLTNEHGNVLYTGVTDNLERRVREHKSGEIPGTAMYNCTKL VYFESFNSIKDAIGREKQLKNWKRKWKNALVVSINPDWNDLAADWDII
<CKW05_RS05640--Legionella_spiritensis--76.14 (SEQ ID NO:2700)
MBEQYVYVYIILASKYGTLYTGVTSNLVQRTWQHKELAGFTKKYNIHHLVYIEIHSEVYEAITREKRIKKWKRWKINLIELNPNQWLDLSIGLF
<CKW05_RS11635--Legionella_spiritensis--65.59 (SEQ ID NO:2701)
MQTESVYVYLLTNKNNVLYTGTINDLIRRIHEHKNKL VAGFTKKYNVDRLIYEVCGDIVGAIEREKQIKGWSRKKKHVLINTLNPENWDLQYSIL
<CKV88_RS02035--Staphylococcus_simiae--58.44 (SEQ ID NO:2702)
MDSHFVYIVKCNDSGLTYGAKDVKARIAKHNDGKGAKYTKTRRPVQLVYHEVYDTS AALKREYEIKTYSRQKKRLIEER
<CKV79_RS01470--Legionella_lansingensis--74.16 (SEQ ID NO:2703)
MKLFYAYIMASKRNGTLYVGSTSDLTRVWEHKNKLTPGFTAKYNVHRLVYIEIHETVYGAARREKRFKNWSRQWKINLIEDLNPEWRDLYEEIYS
<CKV79_RS08660--Legionella_lansingensis--72.15 (SEQ ID NO:2704)
MLTIACSYIWLPMYVGSTSDLIKRVWEHKNKVTS GFTAQHSVHILVYIEVHSTYVEAARREKRFKNWCRQWKINLIEDLNPEWRDLYEEIICS
<CFT68_RS16675--Hymenobacter_gelipurpurascens--58.82 (SEQ ID NO:2705)
MYIYIILTNPTKT VLYIGVTDNLTRLSEHSSGRGDAGKFTGRYQADLLVYFELCPDVTQAITREKQLKGWTRSKKDALITAFNPDWNTIDLETWSG
<CFT68_RS17125--Hymenobacter_gelipurpurascens--55.17 (SEQ ID NO:2706)
MLYTYTYITTNFARNVLYIGVTNNLATRLGQHYANRGNAQT FAGRYFCYHLLYFEEYPSNAAAIAREKELKGWTRAKKEALIASQNPKEWFLTAG
<CHB87_RS19915--Granulicella_rosea--67.44 (SEQ ID NO:2707)
MPREHYSVYILASRSRTLYIGVTNDLSGRVREHREGAGSFTRYQLHRLVYFERFQYVNNAIAREKYLKHMTRA EKIALIETMNPTWEDLESPLPSRIIAA
I
<CHB87_RS02360--Granulicella_rosea--68.24 (SEQ ID NO:2708)
MPPMRREYRFVYIILASRSRNLYVGMNNLAGRVTT HRKAIPGHTARYNIDRLVYEFYKYVNNAIARETELKGWTRAQKIALIEAENPTWEDLAEFPLTHE
ERKADSLRE
<CHB87_RS14290--Granulicella_rosea--67.42 (SEQ ID NO:2709)
MPGNGVYVYILASSFKHLYIGVTSEIENRIWQHKN AIFPDSFTSRYRIDQLVYLETYQGMTAAIAREKQLKRWSRRIKKIRLIVAENPTWRDLS EDWGKPMRLYS
DPSPQPE
<CHB87_RS02375--Granulicella_rosea--68.25 (SEQ ID NO:2710)
MSGQIEIRMWQHKNIGFEGFSKQYHCNRLVYLEGYSEVQSAIAREKQIKGWTRAKKIALIEQQNPKWDDLSEAWGKPMKMHEER
<CHB74_RS18420--Sphingomonas_laterariae--76.14 (SEQ ID NO:2711)
MIERQPCVYIILASAMNGTYIGVTS DLLCRIVQHREGVFEFGFTKRYGIKMLVYWEYAGSMEEAAILREKQIKRWRRDWKRNLIERNNPSWNDLAVSFGLEPLAS
SPA

FIG. 11GGGG

<CHB74_RS18500--Sphingomonas_laterariae--61.80 (SEQ ID NO:2712)
MSEGGWTYIITNKPNGVLYIGVTTDLAARMQHRRGAGSAFCRKYGLDRLVLVEPHDSIVKAITREKALKAWQSAWKIRLIDAANPQWDDLFERIL
<CHB74_RS13265--Sphingomonas_laterariae--61.11 (SEQ ID NO:2713)
MRSGGWTYIITNKPNGVLYIGVTTDLAARMQYRRGAGSAFCRKHGLDRLILAEPHDSIVEAIAREKALKAWQRAWKIRLIDAANPQWDDLFERIL
<CHB92_RS19555--Pontibacter_ummariensis--54.44 (SEQ ID NO:2714)
MHGHNYFYVII TNPGRTVLYTGV TNDLRVRLDQHKVNRGKPEFAGRYCYQLLYEYWTYVQHAIEREKELKLLSREQMTLIKSENPLYRFLKVDD
<CHB69_RS09315--Sphingopyxis_indica--71.91 (SEQ ID NO:2715)
MHRDFQPCAYIMASARHGTLYTGV TSHLIQRV VQHREGTFDGTFRYGVKGLVWFELSGTME LAIAREKQLKNWQRRWKIELEAANPEWRDLAEELGLPPLG
G
<CHB69_RS02560--Sphingopyxis_indica--75.27 (SEQ ID NO:2716)
MRERQPAVYMMANRRNGTIYTGVSLLVRRVWQHREGLGGFSQRYGCKILVWFERDETEAAIAREKQIKAGSRAKKLALIEADNPLWNLDPVPIIDGSE
<CHB69_RS07985--Sphingopyxis_indica--74.12 (SEQ ID NO:2717)
MYRPGVYVIMASQKNGTIYIGATSDLPQRAWQHREGMIEGFTKDYGCKLLVWFEEVFESIYDARQFEARMKKWNRAWKIKRIEERNPDWADLFETLV
<CHB69_RS02510--Sphingopyxis_indica--62.35 (SEQ ID NO:2718)
MRGGVYIMTNKPRGVLYIGVTAIEARTFQHRSGSGSAFCRKYGLDKLVLEVETIYPTIEEAIAREKAMKAWKRQWKIELEASNPWGDDLI PHIL
<CHB83_RS08755--Hymenobacter_mucosus--56.47 (SEQ ID NO:2719)
MYVYIITNPTKTVLYIGTNDLARRLNEHSHNGRGDAGKFTGRYQADLLMYFELCPDAAQAIAREKQLKGWTRQKKDALITAFNPSWETIDLETWTG
<CHB80_RS00635--Dokdonia_pacifica--63.29 (SEQ ID NO:2720)
MKTMYTYIMAHIE NKLFYVGV TNNIRQRTFDHKNATFGTHTGKYNIKQLVWYEVHDSPKIAIRREKTIKKWKREWINLIEEMNPE
<CHB71_RS07125--Ekhidna_lutea--61.18 (SEQ ID NO:2721)
MYVYIILSNYSKTVFYVGMTDDPGRMYEHKNKIYPGFTEKYNCDDL LLYEYEEFEGSDQAAYREHQLKRYRKEWKNRLIDS LNPDWKDLTDRFPFEP
<CHB71_RS07140--Ekhidna_lutea--61.18 (SEQ ID NO:2722)
MYVYIILSNYSKTVFYVGMTDDPGRMYEHKNKIYPGFTEKYNCDDL LLYEYEEFEGSDQAAREHQLKRYRKEWKNRLIDS LNPDWKDLTDRFPFEP
<CHB71_RS03300--Ekhidna_lutea--57.47 (SEQ ID NO:2723)
MKSHNYFVYIITNPARSVIYVGMTNDLPTRLLQHYENRGN SKTWAGKYCYCNLVWYERHQFVANAIDREHEIKKWRREKKNKLIESINPNWKFNLKEIGAWDY
L
<CHB63_RS04465--Lutibacter_flavus--56.98 (SEQ ID NO:2724)
MRLFYVYILKCS DASFYIGVTNNLERLLEHNSNKNKIYTYFSRRPNELVWFEQFTTSDEAIKKEKQLKGWSRKKKEALINENWNNLVNLSKNYTQNKSSSED
<CHB58_RS00710--Desulfurobacterium_atlanticum--70.97 (SEQ ID NO:2725)
MSKQGFVYIVTNKNKNVLYIGVTS DLVRRIQHRNK LIDGFTSKYNC TKLVYEEVFDSEIEAIKREKYLKGKKRQFKVDLINRFPNPEWKDLYESILDSEILRY
RSE
<CRO12_RS02095--Sphingomonas_guangdongensis--77.91 (SEQ ID NO:2726)
MEKRGVYVYMMASGRNGTLYIGVTS DDLKRTWEHREGIADGFTKRYGCR TLVWFEAHDRIDD AIRREKQMKEWRAWKLRQIEASNPWDQDLYPTLM
<CRO12_RS04620--Sphingomonas_guangdongensis--69.32 (SEQ ID NO:2727)
MRQPCVASRRMGTYIGVTS DLVRRILHQHRDGS GSD FVKTYGVHLLVRYELFSDMPAAIAREKQLKRWHRDWKINLIESNPNPDWHD LAPTGLGLEPLRQARKD
GP
<CRO12_RS13795--Sphingomonas_guangdongensis--63.33 (SEQ ID NO:2728)
MSYERGGWYIMADRYRGTMVGV TANLAARILQHRTGTGSDFCARYGLTRMVWABPGDTMEACIEHEKRVKWRREWKFS LIERSNPEWTDISGSIV
<CRO12_RS07735--Sphingomonas_guangdongensis--67.65 (SEQ ID NO:2729)
MWDGQPCVYMMASGCNGTLYVGVTS DLVARVVQH RGGAYAGFTDRYGAKRLVWFETAASMMDAIAVEKRIKR
<CHB64_RS00810--Amphiplicatus_metriothermophilus--66.67 (SEQ ID NO:2730)
MPGYVYMLANKPRGVLYTGAADLARRIHQHREGGS AFCCR YGLDRLVWYAEYQRI TDAI SEEKRIKWRRAWKVRLIEEMNPGWEDLYPLLWGR
<CRO04_RS24085--[Clostridium]_amygdalinum--55.42 (SEQ ID NO:2731)
MNYTYLLKCADNTLYCGWNNHLEKRIAAHNQGKGAKYTKARRPVELVYYESYPTKEAMRREVQIKKLSRKDKLFLIKESGGNGKKD
<CRO16_RS06600--Salinimicrobium_sediminis--69.66 (SEQ ID NO:2732)
MKRSYIYILSNERRAVLYIGV TADLYKRVAEHKS SGRGANFTQKYNLKVLLYFEFESSIDQAIAREKQLKNWHREWKWNLIKLNSELKDLFEDITPH
<CRO16_RS06970--Salinimicrobium_sediminis--64.71 (SEQ ID NO:2733)
MVFSYTYMMTNPHNTVIYTGVT SNLIKRVYQHKTKAYKGFTSKYNC TKLVYFLQFSNI SEADIFEKKI KAGSRAKKEKLI TD LNP EWKDLSDGWVFSF
<CRO54_RS19340--Pedobacter_xixiisoli--67.39 (SEQ ID NO:2734)
MERGGCIYIMTNYQRSTLYIGVTS DLRSRICEHLTKFYPNSFTAKYELTICVYFEIHLTIEEAIAREKQLKKWNRAKKEIILNRSNKEWKDLSPEIMEW
<CRO54_RS11190--Pedobacter_xixiisoli--65.56 (SEQ ID NO:2735)
MERGGYIYILTNVNRSLVLYIGVTS DLVFR LIEHREKVSQSFTAKYNAI ICVYYETFTDIVAAIEREKEIKKWRREKKEQLINTVNP SWKDFWNEVKEW
<CRO54_RS16875--Pedobacter_xixiisoli--66.67 (SEQ ID NO:2736)
MERGGVLYIITNKLNTVLYTGVTSDILSRILQHKNKEYSLGFTAKFNCDKLVYCYFYPTIEDAIEEKRIKGASRLSKIKLISLNP EWKDLFDEINI
<CRO15_RS18800--Flagellimonas_pacifica--74.29 (SEQ ID NO:2737)
MIKNVHQYVYIILTNKKYGTLYVGMTSNLERRI FEHKNKSIDGFTK KYGLDTLVYFETYQYVNDAILREKRLKKWN DLAFDWF
<CRO39_RS02415--Pseudoxanthomonas_wuyuanensis--76.14 (SEQ ID NO:2738)
MRERI PCVYILASKRNGTLYIGVTS NLPARVWQH KSDQVSGFTRRYGVHILVWHEAHD SMESAIAREKALKGWKRAWKIALIENANPYWRDLYPACDSHCLKA
TLPRNFRCSGGFQQPDGSPGRAPSP LHSVQ
<DWA33_RS04620--Rhodopseudomonas_pentothexigena--79.78 (SEQ ID NO:2739)
MPRQFVYIILASRIGTIYGV TNDLVRRI FEHKSKEVEGFTLR YGVDR LVYFEVYDDVENAIRREKRLKKWPREW KIRLIEBHNP DWNDLYPEIAGPP
<DWA33_RS14745--Rhodopseudomonas_pentothexigena--71.11 (SEQ ID NO:2740)
MAYVWYILASAPGGTLYVGVTNDLVRRTFEHREGQVEGFTRK YGVKRLVYFEAHD SIIAAIQREKNIKHWPREW KIDLIVRGNPDWDDLYDEIVR
<CRO48_RS08685--Rhizobium_subbaroonis--71.91 (SEQ ID NO:2741)
MKGYVYILASQRNGTLYTGVT TDLPRRLYEHQNNLT PGFTTRYGVKTLVWFEEHDLVTDAIAREKAIKNWPRKWKL DLEAANPEWEDIAHYLHSL
<CRO46_RS00815--Hoeiflea_halophila--66.29 (SEQ ID NO:2742)
MTGYVYMLASRKG GTIYIGVTNDLARRI PEHTAGTSGS GFTRRYGVRRLVWYEEHFDIGDAIQREKSLKRWRPQWKIDLIEKTNPEWFELFRGTGW
<HDIA_RS07735--Hartmannibacter_diazotrophicus--78.02 (SEQ ID NO:2743)
MKEPCVYILASRRNGTLYIGVTS NLSARVHDHRTGII PGFTSR YHVHRLVWFEAHE TMEAAI LREKQLKRWRRDWKLRVIENTNPEWVDFDGLAR
<HDIA_RS02595--Hartmannibacter_diazotrophicus--70.00 (SEQ ID NO:2744)
MAGWVYIMASRKHGTYLVGVTSDLP RR VHEHRNGLLP GF TERYGCKRLVWSEQFSTIGDAIAREKAVKKWRRAWKVQAI EAANPDWDDLYETLW

FIG. 11HHHH

<HDIA_RS17555--Hartmannibacter_diazotrophicus--72.09 (SEQ ID NO:2745)
MAGWVYIMASRMHGTLYIGVTSDLPRRAWHEHREGIIPGFTSKYGCCKLVWSEEFSEISDAIQREKSLKRWPAWKIALIEKTNPRWVDLYDTLNM
<HDIA_RS04670--Hartmannibacter_diazotrophicus--74.12 (SEQ ID NO:2746)
MSRTYVWYI LANRKDGTLYVGLTSDLQRRVDQHRRAVDGFTKRYGVDRLVHFEAFGEVADAIAREKQLKRWNRWSWKIRLIEAGNPEWKQLDPEW
<HDIA_RS17495--Hartmannibacter_diazotrophicus--68.42 (SEQ ID NO:2747)
MKPGFVYIVTNKPNGLVYTGVTSNLVDVRIYQHRTGVFDGFSKKYGCNRLAWFEAHDDIAEAIHREKRIKTWLRWVKRLILEGNPDWRDLYPELVGLPEGEVD
<CR254_RS02955--Monaibacterium_marinum--72.04 (SEQ ID NO:2748)
MQYVYVLLASRPHGALYLGSTNDLRRRVEQHRAGAIPGHTARYNIGTLVWFETHQTHHEAIVRERRIKKWRDRWKIALIQQHNPDWRDVAFDIPY
<CS899_RS01730--Miniphocibacter_massiliensis--56.96 (SEQ ID NO:2749)
MNYVYILECIDGTLYTGWTNDIERRLISHNIGKAKYTRGRTPVKLIYIEKYPTKVEAMKREFQIKKFTRQKKLHLIEKNSRT
<VTAP4600_RS14315--Vibrio_tapetis_subsp._tapetis--68.48 (SEQ ID NO:2750)
MQKQPCVYFLLSSVNKNALYIGVTSNLKKRVWLHKSNAVNGFSKRYQTHELIYFELYSTMDLAIAREKQVKKWRRQWKNELVDTHNPEWLDLYLSL
<C4Q30_RS00125--Rickettsia_fournieri--78.82 (SEQ ID NO:2751)
MYWVYILCSNKHGTLYIGITNNIIRAYEHKQKI IKGFTAKYNI IKLVYKEBFTDIQEALTREKNLKKWQRSKGIELEKTNPEWKDLIK
<MPLG2_RS15270--Micropoena_glycogenica--67.06 (SEQ ID NO:2752)
MTRDAYVYLMADTHNTVLYIGVTNDLLRRVAEHKLGLQRGFTQRYRCKTLVYFEAGDSITAAIEREKQWKTWRRSWKDDLVASINPTWQDLAPGILTTAVLA
SAGRSRRSPG
<C7D97_RS04135--Anaerococcus_marasmi--69.47 (SEQ ID NO:2753)
MSGYFYILANKSNDVLYKGSTTNI IKRVYEHKNHLDKGSFTDKYNVTKLVYVEFYESI QEARGREVQVSKRSRAKKDKLIDSLNPERKDLWDIVG
<DY411_RS01670--Alysiella_crassa--76.60 (SEQ ID NO:2754)
MTAHPILTRQPAVYILASGRNGTLYIGVTMNLPERIWQHKQGVSLDGFTAKYGVDKLVWYEVSATMPEAIAKEAMKKWRREWKIRLIEEQNPDWLDLFGGLLA
<DY249_RS08680--Anaerococcus_octavius--67.47 (SEQ ID NO:2755)
MITNYNSTLYIGVTNNLYRRLDEHKKKKTNGFTAHYNYKLVYVESTPDYIAAIAAREKQLKRWRSRTKKDKLVNSQNPEWKDLMI EWKN
<DX082_RS06285--Empedobacter_falsenii--75.56 (SEQ ID NO:2756)
MKKGGVYIMANKRNGTLYTGVTADLKRRVYEHKTKYNSEFTAKYNTILVYYSFHSIEDAITTEKRIKAGNRKNKQIETLNPNWIDLWDEINNEW
<DYE40_RS07320--Kingella_potus--70.45 (SEQ ID NO:2757)
MASGRYGVLYIGVTMNLPERVWQHKNGLAPDGFTARYRVYDLVWYELFEHMPAAIAAREKALKKWRREWKIRLVEEQNPWLDLDFSLIR
<DYE47_RS10940--Legionella_beliardensis--81.11 (SEQ ID NO:2758)
MKVSYVYILANQRYGTLYIGLTNNI IKRIWEHKNKLAPGFTAKYNSVQLVYVEIFHDIQTAAAREKRLKEWKQWKIELIETMNPWNHDLVYEEIQ
<DYE47_RS08365--Legionella_beliardensis--77.46 (SEQ ID NO:2759)
MDKQYVYVILASKPYGTLYIGITSNLIQRIYQHQA GLIDGFTKKYNVHQLIYEVHSDAYQAVARERCIKKWNR
<DYE63_RS04385--Neisseria_dentiae--79.57 (SEQ ID NO:2760)
MKQPAVYMSNRKNGVLYLGVTSNLIQRAYQHRQHIAGFTDRYNLESVWYELHETMESAIRREHTMKHWRQWKIDLIENTNPEWRDLWQEIVG
<DYB94_RS05955--Bergeriella_denitrificans--89.36 (SEQ ID NO:2761)
MQPAVYILASRRNGTLYIGVTSDLIKRIYQHKNHQVQSFTERYDVILLVWYELHATMESAITREKQLKKWQRNWKRLIEENPNWDLWAEIIGE
<DYH30_RS10055--Legionella_busanensis--71.59 (SEQ ID NO:2762)
MEEKQYVYVILASTKYGTLYTGVTSNLIQRASQHKADLVGFTKKHRIHRLVYVEIHLDIYEAITREKCIKKWYRQWKINLIEEQNPQWLDLSIGLF
<DQN19_RS04100--Elizabethkingia_meningoseptica--62.50 (SEQ ID NO:2763)
MLNLPKGIYTFYVYILTNKNRTVLYTGVTNNLYKRLYQHTTKSNPKSFTAKYNTMFLVYVEKFGWIQEAIQREKEIKNLTRVKKLELIRSINPLDNLHLFK
SNV
<SHEWBE_RS12090--Shewanella_benthica--82.80 (SEQ ID NO:2764)
MQPCVYILASKKNGTLYVGVTSDLIKRIWEHREHLVTGFTDRYHVTQLVYVELTDCMYS AISREKKLKKWRRQWKINLIENQNPQWRDLWDEIQS
<ROE7235_RS02740--Roseibaca_ekhonensis--64.04 (SEQ ID NO:2765)
MSHYVYIMASRPGGALYTGRTDQLQRRVAHRAGLSTHTAKYNI KTLVWFEPHDDFEQSLRRERAIKWRWRQWKIALITEFNQNWKDMAHAI PD
<RHODPL_RS19690--Rhodoplanes_piscinae--77.53 (SEQ ID NO:2766)
MFYVYLLASRKQGTLYLGITRDLVRRYIYQHKSKSLPGFTSRDYDRLVWFECHDDPITAIAREKDIKKWRRWAKIALIEKDNPDWRDLYPEITS
<E0E36_RS19575--Olivibacter_jilunii--51.69 (SEQ ID NO:2767)
MKRFVYILTDKNRQYLHVGLTDDLFKTMTFYKEHRQLFDDTA AKVSRLVYFEEIDTEELGIARFNVLSAFTRSQKERLIRAKNKDWQDLSLCLGIDNLVANPQ
QPLRFSSMRW
<EL165_RS09050--Chryseobacterium_gleum--65.12 (SEQ ID NO:2768)
MKFKSGSHTYIYIILTNKSRKVLTYTGVTGNLHRRLYQHKTCLNPGSFTARYNLEFLIYIEKYNIWHHAIEREKEIKNWSRIKKLELIRSTNPNLEFLNYLFE
NWL
<EL143_RS02695--Neisseria_canis--84.95 (SEQ ID NO:2769)
MKQPAVYIMSNRKNVLYLGVTSNLIQRAYQHKQHIVKGFTDRYNLEMLVWYELHETMESAIQREHTMKHWRKWKLELIEKNPNPDWRDLWQEII G
<EL201_RS05915--Legionella_pneumophila_subsp._pascullei--73.33 (SEQ ID NO:2770)
MKSFYAYIKASKRNGTLYVGSTSDLIKRVWEHKNKVI PGFSARYNVHILVYVEVHETIYIEAARHEKRFKNWCRQWKINLIEELNPEWRDLYEEICS
<EL201_RS03210--Legionella_pneumophila_subsp._pascullei--74.44 (SEQ ID NO:2771)
MKSFYVYIMASKRNGTLYVGSTVDLIKRVWEHKNNVAPGFTATYDIHMLVYVEVHQSVEAARREKRFKNWPRQWKINLIEQVNSEWRDLYEDICR
<EL220_RS02755--Legionella_sainthelensi--73.91 (SEQ ID NO:2772)
MEKYSYVYIVTNHKNGTLYIGVTTDLVKRIWEHKNKFVESFTKKYNCTRLVYVEYEQHIDVLEAIKREKRLKKYSRTWKIELIESVNPLWRDLYEDIC
<EL220_RS04270--Legionella_sainthelensi--68.89 (SEQ ID NO:2773)
MSLAYVSI LASQHHGTLYVGSTSDIIKRTWEHKNKVSPGFTAQYNVHMLVYVEAYELYVEAARREKRFKNWPRQWKINLIEKLNPPQSRDLYQEICQ
<EL220_RS03595--Legionella_sainthelensi--68.89 (SEQ ID NO:2774)
MSLAYVSI LASQHHGTLYVGSTSDIIKRTWEHKNKVSPGFTAQYNVHMMVYVEAYELYVEAARREKRFKNWPRQWKINLIEKLNPPQSRDLYQEICQ
<EL220_RS02560--Legionella_sainthelensi--68.89 (SEQ ID NO:2775)
MSLAYVSI LASQHHGTLYVGSTSDIIKRTWEHKNKVSPGFTAQYNVHMMVYVEAYELYVEAARREKRFKNWPRQWKINLIEKLNPPQSRDLYEEMCQ
<EL220_RS06140--Legionella_sainthelensi--68.89 (SEQ ID NO:2776)
MSLAYVSI LASQHHGTLYVGSTSDIIKRTWEHKNKVSPGFTAQYNVHMMVYVEAYELYVEAARREKRFKNWPRQWKINLIEKLNPPQSRDLYQEICQ
<EL220_RS10195--Legionella_sainthelensi--68.89 (SEQ ID NO:2777)
MSLAYVSI LASQHHGTLYVGSTSDIIKRTWEYKKNVSPGFTAQYNVHMMVYVEAYELYVEAARREKRFKNWPRQWKINLIEKLNPPQSRDLYEEICQ

FIG. 11III

<EL220_RS09780--Legionella_sainthelensi--67.78 (SEQ ID NO:2778)
MSLAYVSIILASQHHGTLVVGSTSDIIRKTWEHKNKVSPGFTAQYNVHMRVYEAELYVEAARREKRFKNWPRQWKINLIEKLNQSRDLYEEMCQ
<EL220_RS02070--Legionella_sainthelensi--67.78 (SEQ ID NO:2779)
MSLAYVSIILASQHHGTLVVGSTSDIIRKTWEHKNKVSPGFTAQYNVHMMFYEAELYVEAARREKRFKNWPRQWKINLIEKLNQSRDLYEEMCQ
<EL220_RS15060--Legionella_sainthelensi--67.78 (SEQ ID NO:2780)
MSLAYVSIILASQHHGTLVVGSTSDIIRKTWEHKNKVSPGFTAQYNVHMRVYEAELYVEAARREKRFKNWPRQWKINLIEKLNQSRDLYEIEICQ
<EL220_RS05730--Legionella_sainthelensi--66.67 (SEQ ID NO:2781)
MSLAYVSIILASQHHGTLVVGSTSDIIRKTWEHKNKVSPGFTAQYNVHMRVYEAELYVEAARREKRFKNWPRQWKINLIEKLNQSRDLYEIEICQ
<EL220_RS13080--Legionella_sainthelensi--66.67 (SEQ ID NO:2782)
MSLAYVSIILASQHHGTLVVGSTSDIIRKTWEHKNKVSPGFTAQYNVHMRVYEAELYVEAARREKRFKNWPRQWKINLIEKLNQSRDLYQEICQ
<EL218_RS07615--Chryseobacterium_antarcticum--62.50 (SEQ ID NO:2783)
MKPGFIYILTNKTHTVLVYGVTSNLPRIIEHKEKKYQNSFSRRYNDQVLVYVESFQEI GDAIMRERQLKGGSRQKKLDLINRLNPQWTDLYESILDIMNPF5
<EL218_RS02380--Chryseobacterium_antarcticum--63.74 (SEQ ID NO:2784)
MKVGFIYIMVTNKNNTIYVGV TANLPARILEHKTQYKNSFSKYNLDKLVYWEAFQDIGNAIGREKQIKGGSRQKKLDLINSINPEWIDLYESILNIMDL
<EL270_RS01525--Chryseobacterium_jeonii--61.46 (SEQ ID NO:2785)
MKPGFIYIVTNKNNTALYIGVTSKLPTRILEHKKKKFKNSFTAKYNLDKLVYWESTQDIGNAIGREKQIKGGSRQKKLDLINSINPDWNDLYESILSVMDL
<EL297_RS11015--Neisseria meningitidis--98.94 (SEQ ID NO:2786)
MQFAVYILASQRNGTLYIGVTSDLVQRIYQHREHLIEGFTLRYNVTMLVWYELHPTMESAITREKQKKWNRAWKLQIEENNVSWQDLWFDII
<EL297_RS12565--Neisseria meningitidis--74.70 (SEQ ID NO:2787)
MANGKNGTLYIGVTMNLPERVWQHKNHVNIDGFTARYDVHDLVWYQFFENMPEAVAKEKTMKKWRREWIKLIEEQNPFWDLGSLVLFV
<EL294_RS07015--Peptoniphilus_ivorii--73.33 (SEQ ID NO:2788)
MRDYFYILSNKTDSTLYIGVTDLLRRVWEHKEKSVAGFTAKYNVNKL VYVEVAGDVDVAIAREKQLKGWNRKKEALIDAMNPEREDLWESIL
<E1P04_RS01180--Legionella_gresilensis--75.00 (SEQ ID NO:2789)
MEEKQYVYILASKKYGTLYTGITSNLIQVRSQHKADLVEGFTKKHRIHRLVYIEIHLVDYEAITREKCIKKWYRQWKINLIEQHNPQWLDLSIGLF
<E1O41_RS05360--Legionella_impletisoli--73.03 (SEQ ID NO:2790)
MPTKQPAVYIMANKRNGTIYTGVTSDLIKRVYEHKYADEPGFTQKYRCKYLVYIELIEDMVSAIEREKQLKGGSRKKKLALIEKINPYWKDLYETLY
<E1O41_RS11180--Legionella_impletisoli--71.43 (SEQ ID NO:2791)
MTSFYVYILTNKRNGTLVVGSTSDLVKRIWEHKNKFI PNSTAKYNLNLVYIEIHETIYIEAARREARFKWPRQWKINLIEAFNMNWL DLYASICS
<E1O41_RS00005--Legionella_impletisoli--74.44 (SEQ ID NO:2792)
MNDFFYVIMASQRKGTLYVGSTANLIKRVWEHKNKVI PNSTKYNIHKL VYIETDYTAGAAQREKRLKNWHRQWKIDLIEGLNPLWRDLYQEICS
<E1O52_RS01205--Legionella_nagasakiensis--67.78 (SEQ ID NO:2793)
MQTESYVYLLANKHNNVLYTGVTNDLIRRVYEHKNKLVAGFTKKYNVDRLVYIEACGSIVVALEREKQIKGWSRKKKHDLINAFNPEWNDLYQSIL
<E1O47_RS08285--Legionella_yabuuchiae--73.68 (SEQ ID NO:2794)
MHPSYVYILASQRNGTLYIGVTSNIIKRTWEHKNKVI PGFTAKYNVHMLVYIEIHLYIDAARREKRFKNWPRQWKINLIEKLNPEWRDLYEIEICR
<E1O47_RS00795--Legionella_yabuuchiae--70.00 (SEQ ID NO:2795)
MKSFYVYVMAKSRNGTLVVGSTVDLIKRVWEHKNKIAPSTAMYNINILVYIEAHQRYVEAARREKRFKNWPRQWKINIEQVNSEWRDLYEDICR
<E1O47_RS06825--Legionella_yabuuchiae--80.30 (SEQ ID NO:2796)
MLSSKKYGTLYTGITSNLIQRIWQHKEGLAEGFTKKYNVHHWVYIEIHSVDYEAITREKCIKKWNQ
<E1O47_RS02630--Legionella_yabuuchiae--70.97 (SEQ ID NO:2797)
MEGLAEGFTKKYKVHHLVYEVHSDVYEAITREKRIKKWNQWKINLIEQNNPQWLDLIGLL
<E1O47_RS00265--Legionella_yabuuchiae--67.65 (SEQ ID NO:2798)
MTNKHNNHVFYTG VANDLICRVYEHKNKLVAGFTKKNNVDRLVYIEVCGSIITPIEREKQIKRWSRKKK
<E4W33_RS09095--Budvicia_aquatica--73.86 (SEQ ID NO:2799)
MKRPCYVYLLTNKNGTLVYIGVTNLI NRMEQHKSGIISGFTQRYGIKLLVWFEEHSMMTDAIEREKAIKQWQRAWKIELIESINPKWQDLSLTLCF
<MTUNDR4ET4_RS09515--Methylocella_tundrae--72.09 (SEQ ID NO:2800)
MKGGYVYIMSNRRD GYALYIGVTS DLSRRAFEHRQGVVDGFTKRYGLKRLIYERYEDIRDALQREKNLKHVPRVWKLALVHDLNPDWDDLYERVNA
<FGL25_RS01095--Brevundimonas_vancouveri--67.03 (SEQ ID NO:2801)
MATHRTFIATYIMASGRNGTLYAGSTSNLSRVVWKHRTAFPFGSDAYDCTRLVWFEEQHALMLDAIRREIRIKKWL RDWKIALIEKANPDWRDLAADWFS DPT
WTPPPGAD
<FN060_RS11875--Rhizobium_halotolerans--73.12 (SEQ ID NO:2802)
MAGYVYIVTNHKGHTLYIGVTS DLARRIYEHREGLTPGFTSKYGLTQLVWYEEHWDIRDALQREKTMQWYRKWKIALIEERNPDWRDLYLELW
<RTCK_RS17605--Rhizobium_endolithicum--61.25 (SEQ ID NO:2803)
MEVTVYMLRCS DGSYTGTLTKQDVYARLWEHNAGVYDGYTARRRPVELVFMEIYDRIVDAIRARERQIKGWNRAKKGALIAHAYERLPALSRRGFVERQETEQE
R
<FX18_RS06200--Peptoniphilaceae_bacterium--70.93 (SEQ ID NO:2804)
MSNDYVYIVSNKTDSTVYIGMTHDLEKRLYEHKNHLAKGSFTDKYNCYKLVYIEWIDDYDSALAREKQIKKWNRNKKNKLINSINPEWKDLSLEWYK
<FYC17_RS15230--Eggerthellaceae_bacterium--65.88 (SEQ ID NO:2805)
MHEYYVYIMSNPGRTVLVYGV TNSLERRVAEHKSHALPGFTDKYNVIDLVYIEQCGHIEDAIQREKQLKGWSRKKKNALVESLNP DWRDLAAE
<FX040_RS02235--Helicobacteraceae_bacterium--71.59 (SEQ ID NO:2806)
MLFSKPNGTLYVGVTSDLAKRIYEHKNAADGFTKKYGV DKLGYIEAFDDMASAILREKQLKGGNRAKKIELITKSNPQWKDLYEDLICGGF
<FX040_RS07515--Helicobacteraceae_bacterium--75.27 (SEQ ID NO:2807)
MKQAYIYILFNKPNGTLYVGVT SNLIKRIYEHKNKVADGFTKKYGV DKLGYIEFDDIENAILRDKQLKGGSRAKKLELIINSNPNWRDLYEDLV
<AQUSIP_RS07750--Aquicella_lusitana--74.19 (SEQ ID NO:2808)
MKQYFYVYILASQRHGTLYIGVTS DLKKRIWQHKNII PGFTAKYRVHRLVYIEEHSSLVEMVRERRIKNWCRKWKINLIEEFNPTWKDLYEEVCS
<AQUSIP_RS07790--Aquicella_lusitana--71.11 (SEQ ID NO:2809)
MKEYKYVYVYILT SERNGFYVGVGTNNLTRCIYEHKAELVPGFTKKYDV KMLVYIEFEDISEAIRREKIIKKWKRKFVDAIERANPQWKDLYFDLI
<AQUSIP_RS00770--Aquicella_lusitana--71.76 (SEQ ID NO:2810)
MKNYYVYILSNEAKTLYIGITNNLERLYEHKNKLLIGFTKKYNLIKLVYIETTTDVRNAIAREKQLKGWLRIKKIALIEESNPGWDDLSLSWRI
<GFD27_RS01980--Rhizobium_flavum--72.73 (SEQ ID NO:2811)
MRQGYVYILASKRNGTLYTGVTSDLPRLRYEHQNDLTPGFTTRYGVKTLVWFIEIHLVDVAIRREKTIKKWPRAWKLN LIESINPDWDDISPWLL

FIG. 11JJJJ

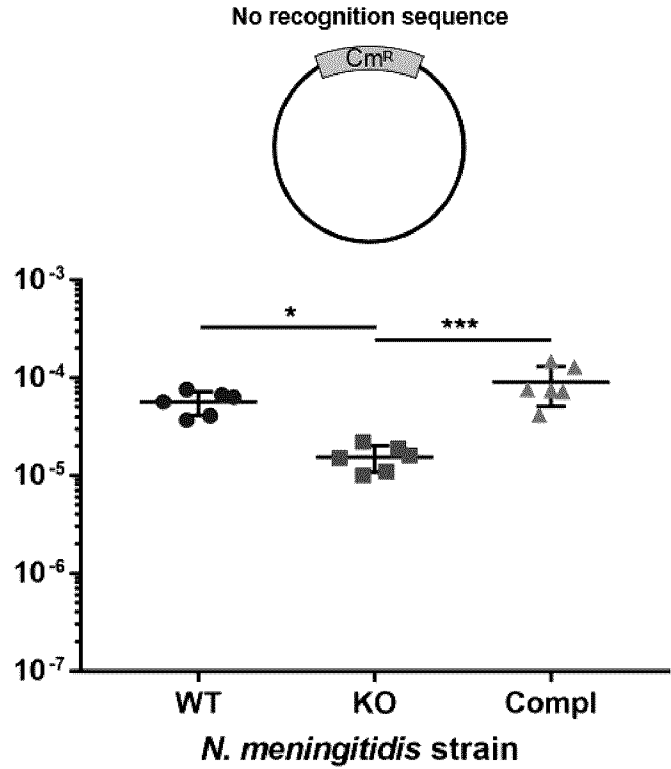


FIG. 12A

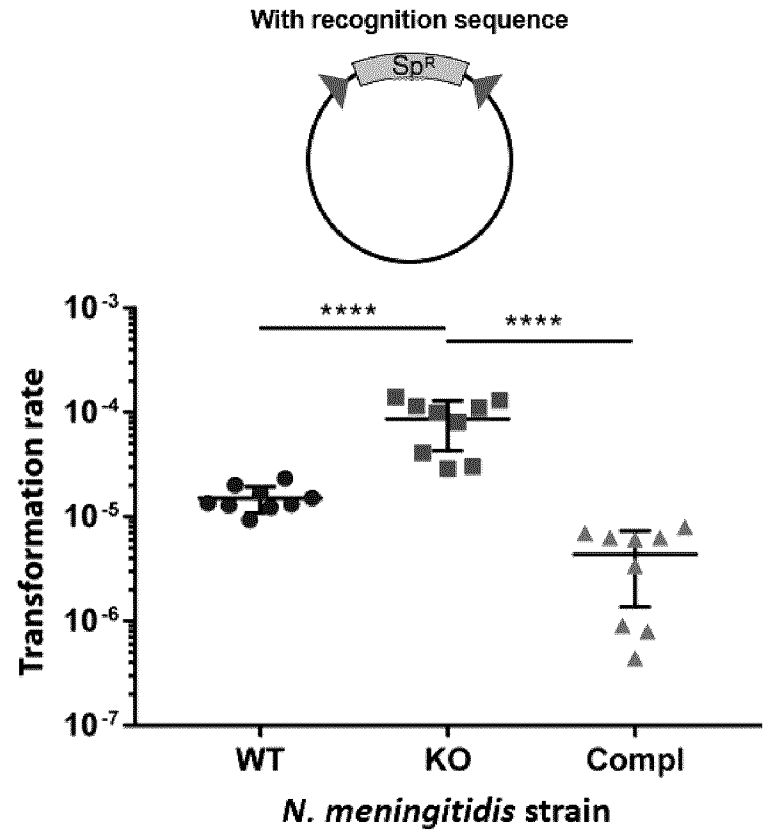


FIG. 12B

Feature 1	#	#	#	#
consensus	1	YVVYILANK-rNGTLVIGVTS-----DLIRRIYEHKEGlg--sgFTSKYNVTRLVYYEEFEDieEAIAREKQKQKWWRR	70	
NP_966153	4	SYVYILTnk-ansTLVIGVTA-----NLIKRIWEHKNKiv---sgFTSRVYMNKLVYFEEFEDisLAIAIRERRLKEWPR	73	
ZP_01916450	30	YWLYILSNskpNGPLVIGDIN-----CLLTRMEHIGQl---rdYAHYHRLDRLVYFEVWTDHDKFVARLRRVRNMPR	100	
ZP_08183153	21	SVVYLLAha-rREAFYIDVasglaisDTRRRRIIVQGEAs-----lPRERVMPLLLWFEACTDLaaAKARATQLRAWPH	94	
ACF12346	4	PAVYILASK-pNGTLVIGVTS-----DLVKRVWEHKNQil---egFTKKYHMHLLVYFEQHNDMPsAFTRKQKQKWWRR	73	
YP_003092183	3	KFYVYVTDn-rRTSLFVGMSS-----DLIKTLDFYKQmpc---lIFDSGQQLTRLVYFEELKTEsQALSREKILISAFTR	72	
YP_004318326	3	RFVYILTdc-nRQYLHVGLTD-----DLFKTMTFYKEHrq---lFFDTAAKVSRLVYFEEIDTEsELGIARFVNLsAFTR	72	
YP_004273595	3	NIYFIVTDa-nRKTSLFVGMSS-----DLIKTMDFYAKIpn---lIFHTDQKLTSLVYFEEFNSIEtANKRFSALSXYTR	72	
YP_617268	27	GVLVILMASa-rNGTLVIGVTS-----DLARVWQHRRNEvv---esFTKKYGCHTLVWYEAFODlQQAIRQRELOQMKNWR	96	
YP_757205	21	PCVYILASq-rNGTLVIGVTS-----NIAMRIMLHKEGkg---saFTRKYKVDRLVWYEFHQDFgFAIEREKRLKNWHR	90	
YP_756058	20	IAYVIMASr-yDGPLYTGVTa-----NLKRVVQHKDGsa---sgFTRKYKCTSLVWMEPHGVIVEAIRREKRIKKWPR	89	
NP_422312	16	IAYVIMADh-yRGTLVAGVTA-----QFIRITIDHREGrg---saFTRRYGLTRLVWYEEHMLMTEAIQRETSIKRWPR	85	
YP_498077	15	GWYVIMANr-yRGGMVGVTS-----DLIRVWQHREGvg---ssHVDDFGKTRLVYAEHHEIEtPAIAIREKLVKQWRR	84	
YP_756108	14	FHVYLLTSh-KNGTLVIGVTR-----ELLARVYAKVKRka---ggFTARYGVQDLVWFESHELvNNAITREKRIKKWRR	83	
EDX79735	8	IATVIMASg-rNGTLVIGVTA-----NLARVWQHRTGaf---dgFGKAYGCTQLVWFEQAWVWVVEAIRREKRIKKWR	77	
ZP_00267500	8	IAAYVIMANr-pQGTLYVGVTS-----NLVQRVQCHKDgtf---pgFTRKHGLKRLVWYEAFAIEtVAIRREKALKRYLR	77	
YP_003058819	8	TAVYIVANr-KNGALVIGVTS-----HLIQRIHIEHQGva---dgYTKRYGCKNLVWYELHEDIEtAIRREKSLKSYRR	77	
YP_004259854	7	YVYVILSNk-rTSVLVIGDIN-----NLKYRFIEHKIGii---pgFTSKYKCHLVYFEEYHDIneAIVREKELKGWTR	76	
ZP_00208415	7	GRVYIMADR-yRGTMVIGVTA-----DLPRRIIFQHROgtg---sdFCARYRLHRLVWAERGEAdVCIaHEKRIKQWRR	76	
YP_003059512	7	YAVYIVSSq-dHGTLVIGVTN-----NLARVWQHKGQvg---akFTTKYVSRLVWYEHDCIhtAIKREKRIKAWQR	76	
YP_757287	7	EAVYLLFSr-pHGALVIGRTK-----DLIKRTWQHRTGai---pgHTKRHGIKQLGWYEHMDGfEsAAHREKILKRWRR	76	
YP_094855	6	PAVYILANK-rNGTLVIGVTS-----DLIKRVWEHKGdgv---pgFTQEHGCKFLVYVELIEDMeAIAIREKQKQKSSR	75	
YP_459696	6	PTVYVLASH-yRGRLYIGVTS-----DLMAKRIYQHDEtm---qgYTSRRDICKRLVWFETHDDIEtAIAIREKTIKKWPR	75	
ACX76174	5	YVYVIMANr-SNTTLVIGVTN-----DISRRALEHIEGcg---atFTSYKINKLIVYFERYETIdDAIKREKQKQWRR	74	
NP_419524	5	FYVYIMANr-rNGTLVIGSTD-----DLARRVFEHQSGvr---gFTAKYGCMTLVWFERHETRESALIREKRLKEWLR	73	
YP_003544434	5	FYVYIMANK-rNGTLVIGVTN-----DLARRVVEHREGlv---kgFTSRVGLKSLVYVEAFDSVsLAIQRETSIKRWPR	74	
YP_003675507	5	PALVILASK-yNGTLVIGVTS-----NLQPRMYCHREGli---dgFSKNYGIKILVYFEMFEDMqTAIAIREKTLKGITR	74	
YP_458324	5	GWYVIMANr-yRGTLVIGVTS-----SLPHRVLQHRTGag---sdFCAQYGCQDLVLAERFMSIdEAIQREKQKQKWWRR	74	
YP_003147384	5	YVYVIMTtk-rNGALVIGVTS-----DLIKRVWQHKNVvv---esFTQKYRVHQLVYVEEQDSIEtALNREKRLKEWRR	74	
YP_003817767	5	YVYVILASg-pCGMYVGVVTN-----DLVRRVAEHKEGii---pgYTSFHRIDQLVWYEAHYIdKAIVREKRIKRWLR	74	
YP_01529536	5	PAVYILANE-rNGTLVIGVTA-----NLIKRVWEHKNDCf---egFSKRYMVHSLVWYELHGTMeAISRKEKELKNWR	74	
YP_001417071	5	YFVYILASK-pRGTLVIGVTN-----DLTRRVFEHRLGtr---pgFTSRVYHIGRLVWFKIHEtAmAAIQREKRLKAWLR	74	
YP_003756409	4	PCVYILASK-pYGTLVIGVTS-----DLPHRMACHDQGi---dgFTKKYRVKTLVYVEHETMDAAIVREKRLKRWQR	73	
ACD34872	4	SWVYILASr-tGTLVIGVTS-----DLTRRVLEHSGvr---sdFASKYGCQDLVLYVEESDDIaIAIAIREKAIKGWTR	72	
ZP_01550813	4	FVYVILASr-KNGTLVIGVTN-----DLTRRIYEHETGgg---skFAARYGAVRLVWYEEHFPvPAQIAIREKAKITWPR	73	
YP_003146897	4	YVYVIMASr-KNGVLYIGVTS-----NLKRVVQHREGiv---egFSEQYKTKRLVYFESSDITsCIEIREKQKQKWWRR	73	
YP_458488	4	GFVYILTnr-KDGVLYIGVTN-----DLARRVVEHRTGtv---ssFTKKYQCHRLVWFEFFENIDHARETEKRMKAWRR	73	
NP_103510	3	GYYVIMASq-KRGTLVIGVTN-----DLGRMPPEHKSgag---srFTSKYGVQQLVWYEEYFDIEtDAIQREKSLKRWLR	72	
NP_273113	3	PAVYILASq-rNGTLVIGVTS-----DLVQRVYQHREHli---egFTSRVYVMTLVWYELHPTMeSAITREKQKQKWWRR	72	
YP_004058011	3	PAVYILASr-KNGVLYIGVTA-----DLKRVVWEHKEGig---saFTRRYGVHKLIVYFEAHPMeAISRKEKQKQKWWRR	72	
ZP_01551056	3	AWVYILASr-pFETLYIGVTT-----DLARRVYEHREGli---egFSNRYGTKSLVWYEQHEEgAAIAIREKQIKRWRR	72	
YP_003059751	3	YALVIMASK-pYGTLVIGVTN-----SLQRRVWEHKEGig---ctFTAKHNIKTfVWFEMFONIEtAIAQREKRMKQKWR	72	
YP_003061436	3	YVYVILASK-dQGTLYIGVTS-----NLARRVYQHKGv---dsYTKRYAVNRLVYVEVDFDIEtAAIKREKQKQKWWRR	72	
YP_001265148	3	GWYVILASK-pFGVLYIGVTA-----DLAARLTVRREGrg---sdFVKRFNCNRLVLEPHATIEEAIAREKALKWRR	72	
YP_459759	3	GFVYVMSDk-pRGVLYIGVTS-----DLKRVVWEHKEGih---kgFTQKYRAKNLVWFEHNPNIvLAIQREKSLKRYMR	72	
ZP_00269375	2	FYVYILASr-pNGTLVIGSTD-----DLKRVVWEHKEKIs---ggFTAKYAVSRLVWYEVHDSReAAAFVREKRIKEWRR	71	
ZP_07825211	2	PCVYILASa-rNGTLVIGVTN-----NVTARLALHRSgag---seFVDYRSVHRLVYIEtYSRATeAIAQREKQKQKWWRR	71	
ZP_03129703	19	LFVYILTSe-rRTTLVIGVTN-----NLGRLVQHONGeg---skFAHYHNLTVLVYETWPOPaSAIAIRETQLKGWTR	88	
YP_004293542	19	YVYVILTnk-nNQVMYIGVTN-----DLKRVVWEHKNKiv---dgFTKKYVMEKLVYFETTSQVlVAIAIREKQIKQWRR	88	
YP_001295178	15	YVYVILTnk-yRSFYVIGITN-----NLSTRLOQHKEIEkgikTFASKYNIEFLVYVEKFTWqLAIAIREKELKGWRR	87	
ZP_01255263	14	YWVYILTnk-pNGILVIGVTS-----GIDYMERHITGtg---skFTSKYNLKMVYEEFYQIEtDAIAIREKQKQKWWRR	83	
EFK33522	11	YVYVILTnk-sRKVLYIGVTS-----NLKRVVWEHKNKiv---gsFTARYNLEFLVYVEKYNIMHAIEREKEIKWRR	81	
NP_774751	9	YVYVILASK-IGGTLVIGVTN-----DLIRVVAEHKSKli---esFTEKYDVARLVYFEQDDPENAIKREKRLKQWRR	78	
YP_004274705	9	YSVYIMTnr-nNTVLYIGVTN-----DLERRVWEHKEFeng---skFTSLYKCHKLVYVEEDYASVKEAIAIREKQIKAGPR	78	
EAP88135	8	YVYVILSNk-mYGVLYIGMTN-----NLKRVVWEHKNKiv---egFSKQYNLTLMCFEQGYvndAIKREKQKQKWWRR	77	
YP_003145891	5	GYYVILSNp-rKTVLYIGVTS-----NLGQVFAHKNKia---egFTRSYDCNDLVYVEIKDSLeKAQVREKEIQSECR	74	
YP_004273654	6	GFVYILTnk-nKTVLYIGVTS-----DLKGRVWEHKNsnr---qsFTFRYVWLYLVYVEFLDSIsQAIDREKQKQKWWRR	76	
CAJ72245	5	YSVYIMTnk-nNTVLYIGVTN-----NLKRVVWEHKEKrv---egFTKKYVMTKLIVYIEFDVrNAIFREKQIKQKSSR	74	
YP_001530877	5	PCVYIMTnn-tNSVLYIGVTS-----HLKRVVYKHKEKIt---ggFTKKYVNAHKLIVYVEVFDQMKAIAIREKQIKQKSSR	74	
ZP_01062387	5	FYVYILTnk-nHTVLYIGVTS-----DLKRVVWEHKEKsn---kgFTRKYNMEKLLVYFESSPVvNAIKREKQKQKWWRR	75	
YP_004317094	5	GCVYILTnk-nRTTLVIGVTS-----NLKRVVWEHKEKfyp---nsFSALYNLNICIYVESFARIEEAIAREKQKQKSSR	74	
YP_003091179	5	GCYVIMSNI-arTTFYIGVTS-----DLVARIQEHKTKyp---nsFTAKYKLHFCIYVEQPSIEEAIAREKEVKQWRR	75	
YP_004216614	5	GWYVILASK--GKRLVIGVTS-----NLQIRVACHKSKIhp---tsFTARYNIDRLVYVQGFETIEEAILRESQIKMWR	74	
ABU58310	6	FYVYILTnk-nNKVMYIGMTN-----DLERRVWEHKNKiv---dgFTKKYVWNLVYVEHDTDvYAAIAIREKEIKQWRR	75	
ACR78865	4	GFVYVMSnp-grTTFYIGVTN-----NLKRVVWEHKEKskI---egFASKYVNCVDIVYVESFEQISsAIAIREKQKQKWWRR	73	
AEB27359	4	GYYVILTnk-nNSVLYIGVTS-----DIVKRIYEHKNKfv---dgFSRRYNLEKLVYFELYQNIeDAIIREKQKQKWWRR	74	
ADV48166	4	GFVYILANK-nNTTLVIGVTS-----NLKRVVWEHKEKfyp---ksFSARYVWNLVYVEAFQDgAAIAIREKQKQKSSR	73	
ZP_00170159	4	YVYVILANK-sNSTLYIGYTH-----NLKRVVWEHKNKiv---egFTDKYVNCVKLVYVEYLDQKYSQALSREKQKQKSSR	73	
ZP_01062150	4	SYVYILTnk-yRTTFYIGVTS-----NLQQRVVDHKEKkg---saFTSKYNVCDLIVYFEEFSDIIEtAIAIREKQKQKWWRR	74	
EAP88139	4	WYVYIMASK-FDGILVIGYTN-----DLKRVVWEHKNKiv---ksFSARYVWNLVYVEELEDsLAsTReRQMKQKSSR	74	
ACV29300	3	YFVYLLTDk-tNSRYVIGYTN-----NLIRRTVYEHKNHlkk---gsFSQKYNIGKLVYVEVFDYvNAIREREIQMKKYSR	73	
YP_004537152	3	PAVYILSNk-nNKVLYIGVTS-----NLVQRVQCHKNHlv---dgFTKQYVMTKLIVYVEHEDMSsEAIKREKALKWRR	72	
ZP_03916739	3	AYVYMLANA-tNEVLKSGST-----NLKRVVWEHKEHlkk---dsFTQKYNVTRLVYVEELEDsLAsTReRQMKQKSSR	73	
YP_004579895	3	HYVYILTnk-KDGVLYIGETK-----NLKRVVWEHKNKiv---tsFSARYNLDKLVYVEEFTKEEAKLREKQKQKWWRR	73	
ZP_01061243	3	YLVYILTnk-yRTTFYIGITN-----NLQRRVVDHKEKkg---saFTSKYNVCDLIVYFEEFSDIIEtAIAIREKELKQWRR	75	
EGC26898	3	YVYVILTnk-tNKVLYIGVTS-----DLKRVVWEHKNHlv---dgFSKYNTTKLVYVESVESIEtAIAIREKQIKSYRR	71	
YP_001677377	2	YTYVIMTnk-FYGTLYIGVTS-----NLKRVVWEHKNKfv---dgFTKKYSIKCLVYVESFEDVrAAIYREKRLKRWLR	71	
YP_001427991	2	AFVYILASK-KNGTLVIGVTA-----NLIKRVVWEHKGf---dgFTKKYDIKILVYVEEQFNQIDAIYREKRLKRWQR	71	
YP_247074	2	WYVYILCSd-rNGTLVIGITN-----NLKRVVWEHKEKii---kgFTAKYNIKLVYVEEFDYvNAIREREIQMKQKSSR	71	
ZP_03915673	2	YVYVILLNK-sNDRTVIGHTD-----DLKRVVWEHKNHlv---ysHTAKYNIKLVYVEEFGEDfYsRAREREKQIKKYSR	72	

FIG. 13A

Feature 1	#	SEQ ID NO:
consensus	71	2812
NP_0966153	74	2813
ZP_01916450	101	2814
ZP_08183153	95	2815
ACF12346	74	2816
YP_003892183	73	2817
YP_004318326	73	2818
YP_004273595	73	2819
YP_517268	97	2820
YP_757205	91	2821
YP_756058	90	2822
NP_422312	86	2823
YP_498077	85	2824
YP_756108	84	2825
EOX79735	78	2826
ZP_08267500	78	2827
YP_003058819	78	2828
YP_004259854	77	2829
ZP_08208415	77	2830
YP_003059512	77	2831
YP_757287	77	2832
YP_094855	76	2833
YP_459696	76	2834
ACX76174	75	2835
NP_419524	74	2836
YP_003754434	75	2837
YP_003675507	75	2838
YP_458324	75	2839
YP_003147384	75	2840
YP_003817767	75	2841
YP_001529536	75	2842
YP_001417871	75	2843
YP_003756409	74	2844
ACQ34072	74	2845
ZP_01508813	74	2846
YP_003146897	74	2847
YP_458488	74	2848
NP_103510	73	2849
NP_273113	73	2850
YP_004058011	73	2851
ZP_01551056	73	2852
YP_003059751	73	2853
YP_003061436	73	2854
YP_001265148	73	2855
YP_469759	73	2856
ZP_08269375	72	2857
ZP_07025211	72	2858
ZP_03129703	89	2859
YP_004293542	89	2860
YP_001295178	88	2861
ZP_01255263	84	2862
EFK33522	82	2863
NP_774751	79	2864
YP_004274705	79	2865
EAP88135	78	2866
YP_003145891	75	2867
YP_004273654	77	2868
CAJ72245	75	2869
YP_001530877	75	2870
ZP_01062307	75	2871
YP_004317094	76	2872
YP_003091179	76	2873
YP_004216614	75	2874
ABU58310	76	2875
ACR70865	74	2876
AE827359	74	2877
ADV48166	75	2878
ZP_08170159	74	2879
ZP_01062150	74	2880
EAP88139	75	2881
ACV29300	74	2882
YP_004537152	73	2883
ZP_03916739	74	2884
YP_004579895	74	2885
ZP_01061243	76	2886
EGC82689	72	2887
YP_001677377	72	2888
YP_001427991	72	2889
YP_247074	72	2890
ZP_03915673	73	2891
	90	Wolbachia endosymbiont of Drosophila melanogaster
	117	Limnobacter sp. MED105
	111	Xanthomonas gardneri ATCC 19865
	90	Chlorobaculum parvum NCIB 8327
	89	Pedobacter heparinus DSM 12366
	89	Sphingobacterium sp. 21
	89	Pedobacter saltans DSM 12145
	113	Sphingopyxis alaskensis RB2256
	107	Maricaulis maris MCS10
	106	Maricaulis maris MCS10
	102	Caulobacter crescentus CB15
	101	Novosphingobium aromaticivorans DSM 12444
	100	Maricaulis maris MCS10
	94	Brevundimonas sp. BAL3
	94	Brevundimonas diminuta ATCC 11568
	93	Hirschia baltica ATCC 49814
	93	Bacteroides salanitronis DSM 18170
	93	Novosphingobium nitrogenifigens DSM 193370
	93	Hirschia baltica ATCC 49814
	93	Maricaulis maris MCS10
	92	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	92	Erythrobacter litoralis HTCC2594
	91	Fibrobacter succinogenes subsp. succinogenes 585
	90	Caulobacter crescentus CB15
	91	Hyphomicrobium denitrificans ATCC 51888
	91	Methylobacter sp. 301
	91	Erythrobacter litoralis HTCC2594
	91	Kangiella koreensis DSM 16069
	91	Brevundimonas subvibrioides ATCC 15264
	91	Desulfococcus oleovorans Hxd3
	91	Xanthobacter autotrophicus Py2
	90	Hyphomicrobium denitrificans ATCC 51888
	89	Acidobacterium capsulatum ATCC 51196
	89	Labrenzia aggregata IAM 12614
	90	Kangiella koreensis DSM 16069
	90	Erythrobacter litoralis HTCC2594
	89	Mesorhizobium loti MAFF303099
	89	Neisseria meningitidis MCS8
	89	Oceanithermus profundus DSM 14977
	89	Labrenzia aggregata IAM 12614
	89	Hirschia baltica ATCC 49814
	89	Hirschia baltica ATCC 49814
	89	Sphingomonas wittichii R61
	89	Rhizobium etli CFN 42
	88	Brevundimonas diminuta ATCC 11568
	88	Afiopia sp. 1NLS2
	105	Chthoniobacter flavus 411n428
	105	Nitrosomonas sp. AL212
	104	Flavobacterium psychrophilum J1P82/86
	100	Psychroflexus torquus ATCC 700755
	98	Chryseobacterium gleum ATCC 35910
	95	Bradyrhizobium japonicum USDA 110
	95	Pedobacter saltans DSM 12145
	94	Croceibacter atlanticus HTCC2559
	91	Kangiella koreensis DSM 16069
	91	Pedobacter saltans DSM 12145
	93	Candidatus Kuenenia stuttgartiensis
	91	Desulfococcus oleovorans Hxd3
	91	Leeuwenhoekiella blandensis MED217
	92	Sphingobacterium sp. 21
	92	Pedobacter heparinus DSM 12366
	91	Acidobacterium sp. MP5ACTX9
	92	Roseiflexus castenholzii DSM 13941
	90	Thermotoga bacterium TBF 10.5.1
	90	Francisella tularensis subsp. novicida GA99-3549
	91	Cellulophaga algicola DSM 14237
	90	Anaerococcus hydrogenalis ACS-025-V-Sch4
	90	Leeuwenhoekiella blandensis MED217
	91	Croceibacter atlanticus HTCC2559
	90	Anaerococcus prevotii DSM 20548
	89	Thioalkalimicrobium cyclicum ALM1
	90	Anaerococcus lactolyticus ATCC 51172
	92	Lacinutrix sp. SM-3-7-4
	92	Leeuwenhoekiella blandensis MED217
	88	Anaerococcus prevotii ACS-065-V-Col13
	88	Francisella philomiragia subsp. philomiragia ATCC 25017
	88	Francisella tularensis subsp. holarctica OSU18
	88	Rickettsia felis URM90Cal2
	89	Anaerococcus lactolyticus ATCC 51172

FIG. 13B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2022/051668

A. CLASSIFICATION OF SUBJECT MATTER

IPC: *C12N 9/22* (2006.01), *C12N 15/10* (2006.01), *C12N 15/55* (2006.01), *C12P 19/30* (2006.01)CPC: *C12N 9/22* (2020.01), *C12N 15/10* (2022.05), *C12P 19/30* (2020.01)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Keywords across the whole IPC

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

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

Databases: Questel Orbit, Canadian Patent Database, PubMed**Keywords:** GIY-YIG motif, endonuclease, ssDNA, monomeric, magnesium, manganese, nickel, Neisseria, SsnA, honing, SEQ ID NOs: 1-2891

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US2012/0070457 (DAUGHERTY ET AL.) 22 March 2012 (22-03-2012) The whole document, specifically SEQ ID NOs: 13759 and 5178	1-36, 41-46, 67-69 and 71 37-40
X Y	WO2005/049642 (BUCHRIESER ET AL.) 2 June 2005 (02-06-2005) The whole document, specifically SEQ ID NO: 5160	1-36, 41-46, 67-69 and 71 37-40
X Y	US2009/298099 (FONTANA ET AL.) 3 December 2009 (03-12-2009) The whole document, specifically SEQ ID NO: 8388	1-36, 41-46, 67-69 and 71 37-40
A	TOMPKINS et al.: 'Molecular underpinnings of ssDNA specificity by Rep HUH-endonucleases and implications for HUH-tag multiplexing and engineering', NUCLEIC ACIDS RESEARCH, January 25, 2021 (25-01-2021), Vol. 49, No. 2, pages 1046-1064, ISSN: 1362-4962 The whole document	1-73

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* "A" "D" "E" "L" "O" "P"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance document cited by the applicant in the international application earlier application or patent but published on or after the international filing date document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed	"T" "X" "Y" "&"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art document member of the same patent family
---	--	--------------------------	--

Date of the actual completion of the international search
05 January 2023 (05-01-2023)Date of mailing of the international search report
16 February 2023 (16-02-2023)Name and mailing address of the ISA/CA
Canadian Intellectual Property Office
Place du Portage I, C114 - 1st Floor, Box PCT
50 Victoria Street
Gatineau, Quebec K1A 0C9
Facsimile No.: 819-953-2476

Authorized officer

Keely Ingrey (819) 639-7697

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2022/051668**Box No. I****Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a. ☒ forming part of the international application as filed.
- b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13^{ter}.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2022/051668

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHANDLER et al.: 'Breaking and joining single-stranded DNA: the HUH endonuclease superfamily', NATURE REVIEWS MICROBIOLOGY, August 2013 (08-2013), Vol. 11, No. 8, pages 525-538, ISSN: 1740-1526 The whole document	1-73
A	GUHA AND EDGELL: 'Applications of alternative nucleases in the age of CRISPR/Cas9', INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, November 29, 2017 (29-11-2017), Vol. 18, No. 12, pages 2565: 1-13, ISSN: 1661-6596 The whole document	1-73
A	DUNIN-HORKAWICZ et al.: 'Phylogenomic analysis of the GIY-YIG nuclease superfamily', BMC GENOMICS, 28 April 2006 (28-04-2006), Vol. 7, No. 98, pages 1-19, ISSN: 1471-2164 The whole document	1-73

INTERNATIONAL SEARCH REPORT

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
US2012070457A1	22 March 2012 (22-03-2012)	None	
WO2005049642A2	02 June 2005 (02-06-2005)	WO2005049642A3 FR2862659A1 FR2862659B1	01 June 2006 (01-06-2006) 27 May 2005 (27-05-2005) 10 February 2006 (10-02-2006)
US2009298099A1	03 December 2009 (03-12-2009)	US8642050B2 AT350394T AU2002302921A1 CA2438080A1 CA2438080C DE60217320D1 DE60217320T2 DK1385876T3 EP1385876A2 EP1385876B1 EP1777233A2 EP1777233A3 ES2278920T3 GB0103424D0 JP2004537977A JP2008263973A PT1385876E US2005260581A1 US7504111B2 WO02079243A2 WO02079243A3	04 February 2014 (04-02-2014) 15 January 2007 (15-01-2007) 15 October 2002 (15-10-2002) 10 October 2002 (10-10-2002) 01 May 2012 (01-05-2012) 15 February 2007 (15-02-2007) 30 August 2007 (30-08-2007) 14 May 2007 (14-05-2007) 04 February 2004 (04-02-2004) 03 January 2007 (03-01-2007) 25 April 2007 (25-04-2007) 30 September 2009 (30-09-2009) 16 August 2007 (16-08-2007) 28 March 2001 (28-03-2001) 24 December 2004 (24-12-2004) 06 November 2008 (06-11-2008) 30 April 2007 (30-04-2007) 24 November 2005 (24-11-2005) 17 March 2009 (17-03-2009) 10 October 2002 (10-10-2002) 13 November 2003 (13-11-2003)