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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 February 2007 (01.02.2007)

PCT

(10) International Publication Number
WO 2007/012131 A1

(51) International Patent Classification:
C12Q 1/04 (2006.01) *C12Q 1/10* (2006.01)
C12Q 1/14 (2006.01) *C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/AU2006/001056

(22) International Filing Date: 27 July 2006 (27.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2005904015 27 July 2005 (27.07.2005) AU

(71) Applicant *(for all designated States except US)*: **DIAGNOSTIC ARRAY SYSTEMS PTY LTD** [AU/AU]; 6 Willunga Way, Bundoora, Victoria 3083 (AU).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **PERERA, Viraj, Navagamuwa** [LK/AU]; 6 Willunga Way, Bundoora, Victoria 3083 (AU). **FRY, Benjamin, Nicholas** [NL/AU]; 74 Coakley Crescent, Lovely Banks, Victoria 3221 (AU).

(74) Agent: **GRIFFITH HACK**; Level 19, 109 St Georges Terrace, Perth, Western Australia 6000 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2007/012131 A1

(54) **Title:** METHOD FOR THE DETECTION AND/OR IDENTIFICATION OF A MICROORGANISM

(57) **Abstract:** The present invention relates generally to a method of identifying micro-organisms in a sample and in particular to the identification of prokaryotic micro-organisms. In some embodiments, the present invention relates to a method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample wherein the test sample is contacted with one or more probes derived from a tuf gene variable region.

METHOD FOR THE DETECTION AND/OR IDENTIFICATION OF A MICROORGANISM

FIELD

The present invention relates generally to a method of identifying micro-organisms in a sample and in particular to the identification of prokaryotic micro-organisms. The method of the present invention is useful, *inter alia*, for the prognosis, diagnosis, prevention, and treatment of diseases associated with pathogens (including human pathogens) and in particular to diseases such as respiratory infections, gastrointestinal tract infections, dermal infection, genital tract infections and nosocomial infections. The present invention also relates to kits used for identifying micro-organisms, particularly prokaryotic micro-organisms.

BACKGROUND

Currently, the standard method used for the diagnosis of bacterial infections is the culture of clinical samples obtained from patients. Culture techniques have been used for many years for clinical diagnosis. Apart from being slow, time-consuming and labour intensive, the reliability of results obtained by culture techniques is very low. Consequently, this has hindered the prescription of effective drugs for treating bacterial infections, resulting in more patient suffering and/or death.

Other shortcomings of present methods of microbial identification include the need for special growth media, the inability to distinguish between closely related species and strains, and the need for large amounts of sample. Importantly, previous methods of microbial detection have found it difficult to simultaneously detect more than one micro-organism in a sample. Consequently,

there is a need to develop methods of identifying micro-organisms in samples which are rapid, sensitive and accurate.

SUMMARY

Accordingly, in a first aspect the invention provides a method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample comprising:

(a) providing a test sample suspected of containing one or more target microorganisms;

(b) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) from said test sample, if present in the test sample, the or each said polynucleic acid comprising at least a portion of a *tuf* gene variable region;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of said one or more said microorganisms in said test sample.

It is preferable that the polynucleic acid in said test sample is amplified, for example, using an amplification technique such as polymerase chain reaction (PCR).

Accordingly, in a second aspect the present invention provides a method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms

in a test sample comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of said one or more said microorganisms in said test sample.

In some embodiments the polynucleic acid is derived from a *tuf* gene variable region as shown in SEQ ID NOs.:1 to 11. Preferably, the polynucleic acid consists essentially of nucleotides 108-220, 393-591, 708-774, 792-816, and/or 885-945 of SEQ ID NOs.:1 to 11.

In other embodiments the polynucleic acid has a sequence which is antisense to a *tuf* gene variable region as shown in SEQ ID NOs.:1 to 11. Preferably, the polynucleic acid is consists essentially of nucleotides antisense to the nucleotides 108-220, 393-591, 708-774, 792-816, and/or 885-945 of SEQ ID NOs.:1 to 11.

In some embodiments, the step of contacting the test sample comprises hybridization with at least a portion of a *tuf* gene variable region with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12 to 84, 208 to 251, 254-315 and 330 to 393.

Accordingly, in a third aspect the present invention provides a method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs: 12 to 84, 208 to 251, 254-315 and 330 to 393;

(d) detecting hybrids formed in step (c); and

(e) identifying said at least one microorganism from said detecting of step (d).

In other embodiments, the at least one probe is selected from the group consisting of SEQ ID NOs: 12 to 84, 208 to 315 and 330 to 393. In other embodiments, the at least one probe is selected from the group consisting of SEQ ID NOs:12 to 392 and 393.

In a fourth aspect, the present invention provides a method for the diagnosis and/prognosis of a disease associated with a target microorganism comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more

probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of the target microorganism in said test sample;

(e) associating the presence of said one or more said target microorganisms with a disease; and

(f) determine diagnosis and/or prognosis for said disease.

In a fifth aspect, the present invention provides a method for the prevention or treatment of a disease associated with a target microorganism comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of the target microorganism in said test sample;

(e) associating the presence of said one or more said target microorganisms with a disease; and

(f) prescribing prevention or treatment of said disease.

In some embodiments, the step of associating the presence of one or more target microorganisms with a disease or the step of prescribing a treatment comprise the step of using a computer program.

In some embodiments, the disease may be a bacterial infection, respiratory infection, gastrointestinal tract infection, dermal infection, bloodstream infection, meningitis or central nervous system infection, urinary tract infection, genital tract infection, wound infection, or nosocomial infection.

In some embodiments, the microorganism detected and/or identified is selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species.

In some embodiments, the step of amplifying at least a portion of a *tuf* gene variable region from the microorganism(s) utilizes at least one primer pair, wherein at least one of the primers is selected from the group consisting of SEQ ID NOs:394-396, or equivalents thereof, provided said equivalents specifically amplify the *tuf* variable region or a part thereof of different species of microorganism at the same time.

In a sixth aspect, the present invention provides a composition comprising at least one probe selected from the group consisting of SEQ ID NOs:12-393 or equivalents thereof, provided that said equivalents specifically bind to a *tuf* gene variable region of a microorganism to be detected.

In a seventh aspect, the present invention provides a composition comprising at least one primer selected from the group consisting of SEQ ID NOs:394-396, or equivalents thereof, provided said equivalents specifically amplify the *tuf* gene variable region or a part thereof of a microorganism to be detected.

In an eighth aspect, the present invention provides an oligonucleotide probe for the detection and/or identification of at least one microorganism found in test sample selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species, wherein said probe is selected from the group consisting of:

for *Haemophilus influenzae:*

TATTATCCGTACAGGTGATGAAGTAGAAA (SEQ ID NO.:12);
AAAGATACAGCGAAAACTACTGTAACG (SEQ ID NO.:13);
CATCGGTGCATTATTACGTGGTAC (SEQ ID NO.:14);
CATACTCCATTCTTCAAAGGTTACCG (SEQ ID NO.:15);
TGACTGGTACAATCGAATTACCAGAA (SEQ ID NO.:16);
CCGTAAATTACTTGACGAAGGTCG (SEQ ID NO.:17);
ACTTCGAATC AGAAGTGTAC (SEQ ID NO.:64);
CAAACGTGAA GAAATCGAAC (SEQ ID NO.:65);
CGATAACATC AAGATGACAG T (SEQ ID NO.:66);

for *Moraxella catarrhalis:*

TGAGCGTGACATCGATAAGTCA (SEQ ID NO.:18);
CAGGCATTATTAAAGTTGGTGATGAAATT (SEQ ID NO.:19);

AAACCAACTACTAAAACCACTTGTACTG (SEQ ID NO.:20);
CGTAAGCTGCTAGACGAAGGT (SEQ ID NO.:21);
TACCCCATTCCTAAATGGCTATCG (SEQ ID NO.:22);
TAACTGGTGCCATCACCCTAC (SEQ ID NO.:23);
AGTTTGATGCAGAAGTATACG (SEQ ID NO.:208)
TATCCTACTACGTGGTACTAA (SEQ ID NO.:209)
ATAACGTTGAGATGAGCG (SEQ ID NO.:210)

for *Enterococcus faecalis:*

AAGGCGACGAGTCTTATGAAGA (SEQ ID NO.:24);
AATTAATGGCTGCAGTTGACGAAT (SEQ ID NO.:25);
GAAGTTCGCGTTGGTGACG (SEQ ID NO.:26);
GACGAAACATCTAAAACAACTGTTACAG (SEQ ID NO.:27);
TGGTGTTGTAGAATTGCCAGAAG (SEQ ID NO.:28);
TAAACCAGCTACAATCACTCCACA (SEQ ID NO.:29);

for *Staphylococcus aureus*:

GTTGACATGGTTGACGATGAAGAATTA (SEQ ID NO.:30);
CTTAGAATTAATGGAAGCTGTAGATACTTACA (SEQ ID NO.:31);
ATCATCGGTTTACATGACACATCTAAAA (SEQ ID NO.:32);
CACCACATACTGAATTTAAAGCAGAAGTA (SEQ ID NO.:33);
CATTCTTCTCAAACTATCGTCCACAA (SEQ ID NO.:34);
CTGGTGTTGTTCACTTACCAGAAG (SEQ ID NO.:35);
TACTGAAATGGTAATGCCTGGTGATA (SEQ ID NO.:36);
CCAATCGCGATTGAAGACGG (SEQ ID NO.:37);
CTGGTTCAGCATTAAAAGCTTTAGAAG (SEQ ID NO.:38);
TGACAACATTGGTGCATTATTACGT (SEQ ID NO.:39);
TGTTACAGGTGTTGAAATGTTCCG (SEQ ID NO.:40);
GTATTATCAAAAGACGAAGGTGGACG (SEQ ID NO.:41);
GCGATGCTCAATACGAAGAAAAAATC (SEQ ID NO.:42);
TGAATTCAAA GCAGAAGTAT AC (SEQ ID NO.:76);
ATTATTACGT GGTGTTGCTC (SEQ ID NO.:77);
GTGATAACGT TGAAATGACA G (SEQ ID NO.:78);

for *Staphylococcus epidermidis:*

TAGACTTAATGCAAGCAGTTGATGATTAC (SEQ ID NO.:43);
CCACACACAAAATTCAAAGCTGAAG (SEQ ID NO.:44);
CTGGTGTTGTAAACTTACCAGAAGG (SEQ ID NO.:45);
GAAATGGTTATGCCTGGCGAC (SEQ ID NO.:46);
CTGGTTCTGCATTAAAAGCATTAGAAG (SEQ ID NO.:47);
TGACAACATCGGTGCTTTATTACG (SEQ ID NO.:48);
CTGTTACTGGTGTAGAAATGTTCCG (SEQ ID NO.:49);
CGTATTATCTAAAGATGAAGGTGGACG (SEQ ID NO.:50);

for *Pseudomonas aeruginosa*:

AAGTTCGAGTGCGAAGT (SEQ ID NO.:51);
AAGGCGTAGAGATGGTAAT (SEQ ID NO.:52);
TTCTTCAAGGGCTACCG (SEQ ID NO.:53);
ATCATCAAGGTCCAGGAAGAAGT (SEQ ID NO.:54);
ACCAAGACTACCTGCACCG (SEQ ID NO.:55);
TGTACGTGCTGTCCAAGGAA (SEQ ID NO.:56);
CCGGTAACTGCGAACTGC (SEQ ID NO.:57);
CACGTTGACTGCCCCGGTCACGC (SEQ ID NO.:85);
ACGCCTTCCGGCAGTTCGCAGTTAC (SEQ ID NO.:86);
ATCGGTCACGTTGACCATGGCA (SEQ ID NO.:87);
GCCGCCTTCGCGGATCGCGAA (SEQ ID NO.:88);

for *Bordetella pertussis:*

GTACATTCTG TCCAAGGAAG (SEQ ID NO.:58);
GACAACGTGG GTATCTTG (SEQ ID NO.:59);
GACAAGGAAA TGGTGCT (SEQ ID NO.:60);

for *Chlamydophila pneumoniae:*

AATTTAAGTC AGCTGTTTAC G (SEQ ID NO.:61);
GAGGTATTGG AAAGAACGAT (SEQ ID NO.:62);
ATAACGTTGA GCTTGATGTT (SEQ ID NO.:63);

for *Legionella pneumophila:*

AGTTTGAAGC AGAAGTGTAT (SEQ ID NO.:67);
TGTTATTACG AGGTACGAAG (SEQ ID NO.:68);
AGATAATGTG CAATTAGTTG TTA (SEQ ID NO.:69);

for *Mycoplasma pneumoniae:*

GAAATTTAAA GCGGAAATCT ATG (SEQ ID NO.:70);
GAACGTGGTC AAGTGTTAG (SEQ ID NO.:71);
GACAATACCT CGATTACAGT (SEQ ID NO.:72);

for *Mycobacterium tuberculosis*
AACATCTCGG TGAAGTTGAT (SEQ ID NO.:73);
TGACAACACC AACATCTC (SEQ ID NO.:74);
ACGAAGGTCT GCGTTT (SEQ ID NO.:75);

for *Streptococcus pneumoniae*

AATTCAAAGG TGAAGTCTAC A (SEQ ID NO.:79);
CAACGTGATG AAATCGAAC (SEQ ID NO.:80);
GACAATCGAC GTTGAGTT (SEQ ID NO.:81);

for *Streptococcus pyogenes*

CAAAGGTGAA GTATATATCC TTTC (SEQ ID NO.:82);
CAACGTGACG AAATCGAA (SEQ ID NO.:83);
CGTGACAATC AACGTTGA (SEQ ID NO.:84);

and combinations thereof.

In a ninth aspect, the present invention provides a method of hybridization comprising contacting a test sample suspected of containing a polynucleic acid with a solid support, which has immobilized thereon one or more probes selected from the group consisting of SEQ ID NOs: 12-392 and 393.

In some embodiments, certain aspects of the invention are provided as kits.

Accordingly, in a tenth aspect, the present invention provides a kit for the detection and/or identification of at least one microorganism or the detection and/or identification of several microorganisms simultaneously in a test sample, comprising

(a) optionally, at least one primer pair allowing amplification of at least a portion of a *tuf* gene variable region;

(b) the composition according to the sixth aspect;

(c) a buffer, or components necessary to produce the buffer enabling a hybridization reaction between the probe(s) contained in said composition and a polynucleic acid present in the sample, or the amplified products thereof;

(d) a solution, or components necessary for producing the solution, enabling washing of the hybrids formed under the appropriate wash conditions;

(e) optionally, a means for detecting the hybrids resulting from the preceding hybridization.

In an eleventh aspect, the present invention provides a method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a sputum sample comprising the steps of

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set

of probes comprising at least one probe selected from the group consisting of SEQ ID NOs: 12 to 84, 208 to 251, 254-315 and 330 to 393;

(d) detecting hybrids formed in step (c);

(e) identifying said at least one microorganism from said detecting of step (d).

In a twelfth aspect, the present invention provides a microarray comprising at least one probe selected from the group consisting of SEQ ID NOs: 12-393 or equivalents thereof, provided that said equivalents specifically bind to the *tuf* gene variable region of a microorganism.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a PCR amplification of the *tuf* gene fragment. Lane 1, DNA standard marker; lane 2, *Streptococcus pneumoniae*; Lane 3, *Moraxella catarrhalis*; Lane 4, *Legionella pneumophila*; Lane 5, *Staphylococcus aureus*; Lane 6, *Haemophilus influenzae*.

Figure 2 shows the detection of *Streptococcus pneumoniae*. Tests 1, 2 and 3-positive test results in duplicate indicating the presence of *Streptococcus pneumoniae*, one of the causative agents of bacterial pneumonia.
Test 1 - SEQ ID NO.:79
Test 2 - SEQ ID NO.:80
Test 3 - SEQ ID NO.:81

Figure 3 shows the detection of *Moraxella catarrhalis*. Tests 1, 2 and 3-positive test results in duplicate indicating the presence of *Moraxella catarrhalis*, one of the causative agents of chronic obstructive pulmonary disease.
Test 1 - SEQ ID NO.:208
Test 2 - SEQ ID NO.:209
Test 3 - SEQ ID NO.:210

Figure 4 shows the detection of *Legionella pneumophila*. Tests 1, 2 and 3-positive test results in duplicate indicating the presence of *Legionella pneumophila*, the main causative agent of Legionnaire's disease.
Test 1 - SEQ ID NO.:67
Test 2 - SEQ ID NO.:68
Test 3 - SEQ ID NO.:69

Figure 5 shows the detection of *Staphylococcus aureus*. Tests 1, 2 and 3-positive test results in duplicate indicating the presence of *Staphylococcus aureus*, one of the causative agents of soft tissue infections and pneumonia in the elderly.
Test 1 - SEQ ID NO.:76
Test 2 - SEQ ID NO.:77
Test 3 - SEQ ID NO.:78

Figure 6 shows the detection of *Haemophilus influenzae*. Tests 1, 2 and 3, positive test results in duplicate indicating the presence of *Haemophilus influenzae*, one of the causative agents of bacterial pneumonia.
Test 1 - SEQ ID NO.:64
Test 2 - SEQ ID NO.:65
Test 3 - SEQ ID NO.:66

DETAILED DESCRIPTION

Before describing preferred embodiments in detail, it is to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only and is not intended to be limiting.

All publications, patents and patent applications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety. However, publications mentioned herein are cited for the purpose of describing

and disclosing the protocols and reagents which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of microbiology and molecular biology which are within the skill of the art. Such techniques are described in the literature. See, for example, Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); "Molecular Cloning: A Laboratory Manual", 2nd Ed., (ed. by Sambrook, Fritsch and Maniatis) (Cold Spring Harbor Laboratory Press: 1989); "Nucleic Acid Hybridization", (Hames & Higgins eds. 1984); "Oligonucleotide Synthesis" (Gait ed., 1984); Remington's Pharmaceutical Sciences, 17th Edition, Mack Publishing Company, Easton, Pennsylvania, USA.; "The Merck Index", 12th Edition (1996), Therapeutic Category and Biological Activity Index; and "Transcription & Translation", (Hames & Higgins eds. 1984).

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a micro-organism" includes a plurality of microorgansims and "a nucleic acid molecule" refers to a plurality of nucleic acid molecules, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

In the description that follows, if there is no instruction, it will be appreciated that cell culture techniques are well-known in this field and any such technique may be adopted.

Although any materials and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred materials and methods are now described.

By "comprising" is meant including, but not limited to, whatever follows the word comprising". Thus, use of the term "comprising" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of". Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

In work leading up to the present invention, the inventors discovered that the *tuf* gene contained variable region which differed between different microorganisms and therefore could be used to identify different microorganisms. Different microorganisms in a test sample can be identified by detecting hybridization between the variable region and a probe specific for the variable region.

Therefore, the present invention relates to polynucleic acid derived from a *tuf* gene variable region. As used herein, a "gene" is a polynucleic acid molecule that contains the information for a specific function, such as the sequence encoding a specific protein. Genes may also comprise non-coding information, such as one or more of a signal sequence, an origin of replication, an enhancer element, a promoter, and/or a transcription termination sequence. The term "polynucleic acid" as used herein designates mRNA, RNA, cRNA, cDNA or DNA.

The *tuf* genes are universally present in all bacteria and generally comprise uniform sequences i.e., sequence in common between all species. Examples of other universal bacterial genes include the 16s ribosomal RNA, 23s ribosomal RNA, and *rpoB*. Examples of *tuf* sequences of different microorganisms are set out in SEQ ID Nos.: 1 to 11.

It will be appreciated by those skilled in the art that the sequences disclosed herein may vary from the sequences provided in any of SEQ ID NOs.:1 to 11, without departing from the invention. In particular, the *tuf* gene sequence might vary by up to 30%, provided that a biological activity of the variant is retained. Specifically, the variant must have a variable region which can be used to identify a target microorganism.

A "*tuf* gene variable region" is those nucleotide sequences which vary in the *tuf* gene between species, subspecies, and substrains of a microorganism. In some embodiments, the *tuf* gene variable region consists essentially of nucleotides 108-220, 393-591, 708-774, 792-816, and/or 885-945 of SEQ ID NOs.:1 to 11.

In some embodiments, the *tuf* gene variable region is from

Berkholderia cepacia, which can be used to specifically identify this species in sputum samples of patients with cystic fibrosis. In some embodiments, one or more of SEQ ID Nos.:269 to 275, which correspond to *tuf* gene variable region of *Berkholderia cepacia*, can be used.

In other embodiments, the *tuf* gene variable region is from the *tuf* gene of *Stenotrophomonas maltophilia*, which can be used to specifically identify this species in sputum samples of patients with pneumonia subsequent to a *Pseudomonas aeruginosa* infection. In some embodiments, one or more of SEQ ID NO.:264 to 268, which correspond to *tuf* gene variable regions of *Stenotrophomonas maltophilia*, can be used.

As used herein, a polynucleic acid comprises at least some part of a *tuf* gene including a *tuf* gene variable region or is derived from a *tuf* gene. The polynucleic acid sequence may have the nucleotide sequence of an entire *tuf* gene *per se* or a part thereof as long as it contains at least one tuf gene variable region or part thereof. The polynucleic acid may also be antisense to the *tuf* gene. In the present invention, the term "antisense" is used to denote a nucleic acid molecule which is DNA or RNA with a complementary sequence to the *tuf* gene variable region of the gene.

According to the invention the *tuf* gene variable region is used to identify a target micro-organism in a test sample. The test sample may be any type of sample suspected of containing a target micro-organism. Consequently, test samples for analysis by the methods of the present invention may be obtained from numerous sources. Such sources include, but are not limited to, samples derived from animal subjects such as humans; flowers; seeds, vegetables; food; and samples derived from the environment. In some embodiments, the test sample is a

"biological sample". Reference to a "biological sample" should be understood as a reference to any sample of biological material derived from an animal such as, but not limited to, blood, saliva, urine, tears, sweat, cerebrospinal fluid, lymph fluid, serum, plasma, mucus, sputum, faeces, and biopsy specimens.

The biological sample which is tested according to the methods of the present invention may be tested directly or may require some form of treatment prior to testing. For example, a biopsy sample may require homogenisation prior to testing. Further, to the extent that the biological sample is not in liquid form, (for example it may be a solid, semi-solid or dehydrated liquid sample) it may require the addition of a reagent, such as a buffer to mobilise the sample.

The test sample may also require processing to release any polynucleic acid prior to testing. For example, the polynucleic acid can be released using a lysis buffer, such as a commercially available lysis solution. A suitable commercially available lysis solution for bacteria is the Lyse-N-Go™ (Pierce Biotechnology Inc, Rockford IL, USA) PCR reagent. Once released the polynucleic acid is available for testing by the methods of the invention.

Alternatively or in addition the test sample may be subjected to a purification step such that the polynucleic acid in the test sample is isolated, concentrated or seperated.

An "isolated" polynucleic acid is one which has been released and/or separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with identification of a target

micro-organism and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In some embodiments, a polynucleic acid will be purified and/or concentrated (1) to a degree sufficient to obtain at least 30 to 45 bases of nucleotide sequence information, and/or (2) to homogeneity by gel electrophoresis.

In isolating the polynucleic acid the test sample may be subjected to a number of different processing steps, such as tissue homogenation, cell isolation and cytoplasmic extraction, nucleic acid extraction and the like. Methods of isolating nucleic acids from cells, tissues, organs or whole organisms are known to those of ordinary skill in the art and are described, for example, in Maniatis *et al.* (Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press, 1989) and in Ausubel *et al*, (Current Protocols in Molecular Biology, John Wiley & Sons, Inc 1998), the content of each are incorporated herein by reference.

In some embodiments of the invention, a cell from a test sample may be used as a source of polynucleic acid for analysis, thereby obviating requirements for purified cultures. The polynucleic acid molecules may be either DNA, RNA or a combination of both. In some embodiments the polynucleic acid is DNA. Preferably, there is at least 100 micrograms of DNA available for analysis; however, considerably smaller amounts of DNA may be used in the methods of the invention. Nucleic acid amplification techniques, such as PCR, may be used to compensate when only low levels of polynucleic acid is available for analysis.

Based on sequence analysis of the target polynucleic acid, a multiplex PCR reaction may be employed with the use of several primers. Multiplex PCR is the term used when more than one pair of primers is used in a polymerase chain

reaction. The goal of multiplex PCR is to amplify several segments of target polynucleic acid simultaneously and thereby to conserve template polynucleic acid, time and expense. Detailed protocols for multiplex PCR can be found in Molecular Cloning, A Laboratory Manual Sambrook and Russell Volume 2, third edition, Coldspring Harbour Laboratory Press. Since clinical samples often contain material that will interfere with PCR, a unique pre-processing method can be developed to obtain optimum amplification. The reaction conditions will be optimised for simultaneous amplification of all polynucleic acid regions of interest.

Other polynucleic acid amplification techniques include self-sustained sequence replication, Q beta replicase, ligation-based thermocycling approaches, real-time PCR, real-time strand displacement amplication (SDA), rolling-circle amplification (RCA) and multiple-displacement amplification (MDA).

Detailed protocols for amplification of polynucleic acid can be found in, among other places, Ausubel *et al* (Short Protocols in Molecular Biology, John Wiley & Sons (1995) volume 2, 5th edition Chapter 15) and in Sambrook and Russell (Molecular Cloning, A Laboratory Manual, volume 2, third edition, Coldspring Harbour Laboratory Press). These publications also provide detailed guidance on how to design nucleic acid amplification primers.

In some embodiments of the present invention, the polynucleic acid is isolated and then amplified using PCR. Any amplification primers may be used to amplify *tuf* genes from target microorganisms as long as the primers specifically amplify the *tuf* gene variable region or a part thereof. Designing primers that are capable of amplifying a large number of *tuf* genes from different species of microorganism is well within the skill in the

art. For example, *tuf* gene sequences for bacterial species are freely available from public databases such as GenBank™. There are a number of databases capable of multiple sequence alignments using the *tuf* gene sequences obtained. Once aligned, persons skilled in the art can determine conserved regions within the various microorganisms by analysing the multiple sequence alignment data. Table 1 shows the multiple sequence alignment of a number of *tuf* genes from various microorganisms.

In some embodiments, the primers used to amplify *tuf* genes from a number of microorganisms are selected from the group consisting of SEQ ID NOs:394-396, or equivalents thereof, provided said equivalents specifically amplify the *tuf* variable region or a part thereof of different species of microorganism at the same time.

The ability to design equivalents of SEQ ID NOs:394-395 is well within the skill in the art. For example, it is well appreciated that the 3′ end of any amplification primer is the most important end of any primer as this needs to properly anneal to a DNA template, for example, for primer extension to take place. Accordingly, the 3′ ends of SEQ ID NOs:394-395 need to be maintained. However, it is further appreciated that the 5′ ends of most primers can be altered, extended and/or shortened without adversely affecting the primers ability to amplify a product. Persons skilled in the art are capable of determining changes that can be made in primers by routine experimentation.

As mentioned above, a polynucleic acid comprising a *tuf* gene or *tuf* gene variable region may be used to identify a target micro-organism in a test sample. Reference to "micro-organism" encompasses any organism that can only be observed with the aid of a microscope and includes

TABLE 1

```
MPNE-HF2                ------------ATGGCAAAACAAAAGTTTGATAGATCAAAAGCTCACGTTAATATTGG 47
MPNE-M129               ------------ATGGCAAGAGAGAAATTTGACCGATCTAAACCCCACGTTAATGTAGG 47
LPNE-Paris              GGGAGACTTTCCGATGGCGAAGGAAAAATTTGAACGTAAGAAGCCGCACGTAAACGTGGG 60
LPNE2-DAS               ------------------------------------------------------------
LPNE-Philadelphia1      ------------ATGGCGAAGGAAAAATTTGAACGTAAGAAGCCGCACGTAAACGTGGG 47
LPNE-Lens               ------------ATGGCGAAGGAAAAATTTGAACGTAAGAAGCCGCACGTAAACGTGGG 47
LLON-DAS                ------------------------------------------------------------
FNUC-DAS                ------------------------------------------------------------
CPNE-CWL029             ------------ATGTCAAAAGAAACTTTTCAACGTAATAAGCCCCATATCAATATTGG 47
EFAE-DAS                ------------------------------------------------------------
EFAE-V583               ------------ATGGCAAAAGAAAAATTTGACCGTTCTAAATCCCATGTTAACATTGG 47
SAUR-Mu50               ------------ATGGCAAAAGAAAAATTCGATCGTTCTAAAGAACATGCCAATATCGG 47
SAUR-N315               ------------ATGGCAAAAGAAAAATTCGATCGTTCTAAAGAACATGCCAATATCGG 47
SAUR-MW2                ------------ATGGCAAAAGAAAAATTCGATCGTTCTAAAGAACATGCCAATATCGG 47
SAUR-MSSA476            ------------ATGGCAAAAGAAAAATTCGATCGTTCTAAAGAACATGCCAATATCGG 47
SAUR-MRSA252            ------------ATGGCAAAAGAAAAATTTGATCGTTCTAAAGAACATGCCAATATCGG 47
SEPI-12228              ------------ATGGCAAAAGAAAAATTTGATCGCTCAAAAGAACATGCCAATATTGG 47
SEPI-RP62A              ------------ATGGCAAAAGAAAAATTTGATCGCTCAAAAGAACATGCCAATATTGG 47
SMIT-DAS                ------------------------------------------------------------
SSAN-DAS                ------------------------------------------------------------
SMIT-NCTC12261          ------------ATGGCAAAAGAAAAATACGATCGTAGTAAACCACACGTTAACATTGG 47
SPNE-R6                 ------------ATGGCAAAAGAAAAATACGATCGTAGTAAACCACACGTTAACATTGG 47
SPYO-M1GAS              ------------ATGGCAAAAGAAAAATACGATCGTAGTAAACCCCACGTTAACATTGG 47
SMUT-DAS                ------------------------------------------------------------
SMUT-UA159              ------------ATGGCAAAAGAAAAATACGATCGTAGTAAACCACACGTTAACATTGG 47
HINF-RdKW20             ------------ATGTCTAAAGAAAAATTTGAACGTACAAAACCGCACGTAAACGTGGG 47
MCAT-DAS                ------------------------------------------------------------
BPER-TohamaI            ------------ATGGCAAAAGGCAAGTTTGAACGTACCAAGCCGCACGTGAACGTGGG 47
BPAR-12822              ------------ATGGCAAAAGGCAAGTTTGAACGTACCAAGCCGCACGTGAACGTGGG 47
BCEP2-DAS               ------------------------------------------------------------
BPSE-1710b              ------------ATGGCAAAAGAGAAGTTTGAGCGGACCAAGCCGCACGTGAACGTTGG 47
SMAL-DAS                ------------------------------------------------------------
PAER-PA01               ------------GTGGCTAAGGAAAAATTCGAACGTAACAAACCGCACGTCAACGTCGG 47
KPNE-DAS                ------------------------------------------------------------
MAVI-K10                ------------GTGGCGAAGGCGAAGTTCGAGCGGACGAAGCCGCACGTCAACATCGG 47
MAVI-paratub            ------------GTGGCGAAGGCGAAGTTCGAGCGGACGAAGCCGCACGTCAACATCGG 47
MAISpig-DAS             ------------------------------------------------------------
MKAN-DAS                ------------------------------------------------------------
MTUB-H37Rv              ------------GTGGCGAAGGCGAAGTTCCAGCGGACCAAGCCCCACGTCAACATCGG 47
MTUB-CDC1551            ------------GTGGCGAAGGCGAAGTTCCAGCGGACCAAGCCCCACGTCAACATCGG 47
MAVI-DAS                ------------------------------------------------------------
CIDP-NCTC13129          ------------GTGGCAAAGGCTAAGTTCGAGCGTACCAAGCCGCACGTCAACATCGG 47
NMEN-MC58               ------------ATGGCTAAGGAAAAATTCGAACGTAGCAAACCGCACGTAAACGTTGG 47
NMEN-Z2491              ------------ATGGCTAAGGAAAAATTCGAACGTAGCAAACCGCACGTAAACGTTGG 47
NMEN-DAS                ------------------------------------------------------------

MPNE-HF2                GACAATTGGACATATTGACCATGGTAAAACAACTTTAACAGCTGCAATCTGTACTTACCT 107
MPNE-M129               TACTATTGGCCACATTGACCACGGTAAAACAACTTTAACAGCAGCTATTTGTACTGTATT 107
LPNE-Paris              CACGATTGGTCACGTAGACCATGGCAAGACGACATTGACAGCAGCTATTACAACGATTAT 120
LPNE2-DAS               ------------------------------------------------------------
LPNE-Philadelphia1      CACGATTGGTCACGTAGACCATGGTAAGACGACATTGACAGCAGCTATTACAACGATTAT 107
LPNE-Lens               CACGATTGGTCACGTAGACCATGGTAAGACGACATTGACAGCAGCTATTACAACGATTAT 107
LLON-DAS                ------------------------------------------------------------
FNUC-DAS                ------------------------------------------------------------
CPNE-CWL029             GACGATCGGGCACGTTGACCATGGTAAAACTACGCTAACAGCGGCAATTACACGCGCGCT 107
EFAE-DAS                ------------------------------------------------------------
EFAE-V583               TACTATCGGACACGTTGACCATGGTAAAACTACATTAACAGCTGCAATTGCTACTGTATT 107
SAUR-Mu50               TACTATCGGTCACGTTGACCATGGTAAAACAACATTAACAGCAGCAATCGCTACTGTATT 107
SAUR-N315               TACTATCGGTCACGTTGACCATGGTAAAACAACATTAACAGCAGCAATCGCTACTGTATT 107
SAUR-MW2                TACTATCGGTCACGTTGACCATGGTAAAACAACATTAACAGCAGCAATCGCTACTGTATT 107
SAUR-MSSA476            TACTATCGGTCACGTTGACCATGGTAAAACAACATTAACAGCAGCAATCGCTACTGTATT 107
SAUR-MRSA252            TACTATCGGTCACGTTGACCATGGTAAAACAACATTAACAGCAGCAATCGCTACTGTATT 107
SEPI-12228              TACTATCGGTCACGTTGACCATGGTAAAACAACTTTAACAGCTGCTATCGCAACTGTATT 107
SEPI-RP62A              TACTATCGGTCACGTTGACCATGGTAAAACAACTTTAACAGCTGCTATCGCAACTGTATT 107
SMIT-DAS                ------------------------------------------------------------
SSAN-DAS                ----------------------------------------------------CTGTATT 7
SMIT-NCTC12261          TACTATCGGACACGTTGACCACGGTAAAACTACTTTGACAGCAGCTATCACAACTGTTTT 107
SPNE-R6                 TACTATCGGACACGTTGACCACGGTAAAACTACCCTAACTGCAGCTATCACAACTGTTTT 107
SPYO-M1GAS              TACAATCGGACACGTTGACCATGGTAAAACTACTTTAACAGCTGCAATCACAACTGTATT 107
```

```
SMUT-DAS                ----------------------------------------------------CTGTACT 7
SMUT-UA159              TACTATCGGACACGTTGACCATGGTAAAACTACCTTAACTGCAGCTATCACAACTGTACT 107
HINF-RdKW20             TACAATCGGCCACGTTGACCACGGTAAAACAACTTTAACAGCAGCAATCACAACCGTATT 107
MCAT-DAS                ------------------------------------------------------------
BPER-TohamaI            TACGATTGGTCACGTTGACCACGGCAAAACGACGTTGACGGCGGCGATCACGACGGTGCT 107
BPAR-12822              TACGATTGGTCACGTTGACCACGGCAAAACGACGTTGACGGCGGCGATCACGACGGTGCT 107
BCEP2-DAS               ------------------------------------------------------------
BPSE-1710b              TACGATTGGTCACGTTGACCACGGCAAGACGACGCTGACGGCAGCGATCGCGACGGTGCT 107
SMAL-DAS                ---------------------------------------------TGACAAGATCGGT 13
PAER-PA01               CACCATCGGTCACGTTGACCATGGCAAGACCACTCTGACCGCTGCACTGACCAAGGTCTG 107
KPNE-DAS                ------------------------------------------------------------
MAVI-K10                GACCATCGGTCACGTTGACCACGGCAAGACCACGCTGACCGCGGCTATCACCAAGGTTCT 107
MAVI-paratub            GACCATCGGTCACGTTGACCACGGCAAGACCACGCTGACCGCGGCTATCACCAAGGTTCT 107
MAISpig-DAS             ----------------------------------------------------------CT 2
MKAN-DAS                 ---------------------------------------------------AAGGTCCT 8
MTUB-H37Rv              GACCATCGGTCACGTTGACCACGGCAAGACCACCCTGACCGCGGCTATCACCAAGGTCCT 107
MTUB-CDC1551            GACCATCGGTCACGTTGACCACGGCAAGACCACCCTGACCGCGGCTATCACCAAGGTCCT 107
MAVI-DAS                ------------------------------------------------------------
NCTC13129               CACCATCGGTCACGTTGACCACGGTAAGACCACCACCACCGCTGCTATCACCAAGGTTTT 107
NMEN-MC58               CACCATCGGTCACGTTGACCATGGTAAAACCACCCTGACTGCCGCTTTGACTACTATTTT 107
NMEN-Z2491              CACCATCGGTCACGTTGACCATGGTAAAACCACCCTGACTGCCGCTTTGACTACTATTTT 107
NMEN-DAS                ------------------------------------------------------------

MPNE-HF2                TGCAAAAAAAGGTGGTGCTAAAG---------CAATGAAATATGATGAAATTGATAAAGC 158
MPNE-M129               AGCAAAAGAAGGTAAATCAGCTG---------CTACTCGTTACGACCAAATCGATAAGGC 158
LPNE-Paris              GGCGAAGAAATATGGTGGTACAG---------CGAAGGCGTACGATCAAATTGATGCTGC 171
LPNE2-DAS               ----------TATGGTGGTACAG---------CGAAGGCGTACGATCAAATTGATGCTGC 41
LPNE-Philadelphia1      GGCGAAGAAATATGGTGGTACAG---------CGAAGGCGTACGATCAAATTGATGCTGC 158
LPNE-Lens               GGCGAAGAAATATGGTGGTACAG---------CGAAGGCGTACGATCAAATTGATGCTGC 158
LLON-DAS                ---------GTTTGGTGGTATAG---------CAAAAGCATACGATCAGATCGATGCTGC 42
FNUC-DAS                ---------------AGCTAAAA---------AAGTAGATTTTGACCAAATTGATGCTGC 36
CPNE-CWL029             ATCAGGGGATGGATTGGCCTCTT---------TCCGTGACTATAGTTCAATTGACAATAC 158
EFAE-DAS                ----------------GGGAAAG---------CACAAAGCTACGATTCTATCGATAACGC 35
EFAE-V583               ATCAAAACACGGTGGCGGGGAAG---------CACAAAGCTACGATTCTATCGATAACGC 158
SAUR-Mu50               AGCAAAAAATGGTGACTCAGTTG---------CACAATCATATGACATGATTGACAACGC 158
SAUR-N315               AGCAAAAAATGGTGACTCAGTTG---------CACAATCATATGACATGATTGACAACGC 158
SAUR-MW2                AGCAAAAAATGGTGACTCAGTTG---------CACAATCATATGACATGATTGACAACGC 158
SAUR-MSSA476            AGCAAAAAATGGTGACTCAGTTG---------CACAATCATATGACATGATTGACAACGC 158
SAUR-MRSA252            AGCAAAAAATGGTGACTCAGTTG---------CACAATCATACGACATGATTGACAACGC 158
SEPI-12228              AGCTAAAAATGGTGACACTGTTG---------CACAATCATACGATATGATTGACAACGC 158
SEPI-RP62A              AGCTAAAAATGGTGACACTGTTG---------CACAATCATACGATATGATTGACAACGC 158
SMIT-DAS                ---------------------------------AAAGACTATGCGTCTATCGATGCTGC 26
SSAN-DAS                GGCACGTCGCTTGCCTTCAGCAGTTAACCAACCTAAAGACTATGCGTCTATCGATGCTGC 67
SMIT-NCTC12261          GGCACGTCGCTTGCCTTCAGCTGTTAACCAACCTAAAGACTATGCGTCTATCGATGCTGC 167
SPNE-R6                 GGCACGTCGCTTGCCTTCATCAGTTAACCAACCTAAAGACTATGCGTCTATCGATGCTGC 167
SPYO-M1GAS              GGCACGTCGCTTGCCTTCATCAGTTAACCAACCAAAAGATTACGCTTCTATCGATGCTGC 167
SMUT-DAS                GGCACGTCGTCTTCCAAGTGCAGTAAACCAACCAAAAGATTATTCATCTATTGATGCTGC 67
SMUT-UA159              GGCACGTCGTCTTCCAAGTGCAGTAAACCAACCAAAAGATTATTCATCTATTGATGCTGC 167
HINF-RdKW20             AGCAAAACACTACGGTGGTGCAG---------CGCGTGCATTTGACCAAATTGATAACGC 158
MCAT-DAS                ----------------------------------------------------CGACTCTGC 9
BPER-TohamaI            GTCGAACAAGTTCGGCGG---------CGAGGCTCGCGGCTACGACCAGATTGACGCGGC 158
BPAR-12822              GTCGAACAAGTTCGGCGG---------CGAGGCTCGCGGCTACGACCAGATTGACGCGGC 158
BCEP2-DAS               ---------GTTCGGCGG---------CGAAGCGAAGAAGTACGACGAAATCGACGCGGC 42
BPSE-1710b              GTCGGCGAAGTTCGGCGG---------CGAAGCGAAGAAGTACGACGAAATCGACGCGGC 158
SMAL-DAS                GCCGAGCGC-TTCGGTGG---------CGAGTTCAAGGACTACTCCTCGATCGACGCCGC 63
PAER-PA01               CTCCGATACCTGGGGTGG---------TTCCGCTCGTGCTTTCGATCAGATCGACAACGC 158
KPNE-DAS                ---------CTACGGTGG---------TTCCGCTCGCGCATTCGACCAGATCGATAACGC 42
MAVI-K10                GCACGACAAGTACCCGGACCTGA---ACGAGTCCCGCGCGTTCGACCAGATCGACAACGC 164
MAVI-paratub            GCACGACAAGTACCCGGACCTGA---ACGAGTCCCGCGCGTTCGACCAGATCGACAACGC 164
MAISpig-DAS             GCACGACAAGTACCCGGACCTGA---ACGAGTCCCGCGCGTTCGACCAGATCGACAACGC 59
MKAN-DAS                GCATGACAAGTTCCCGGACCTGA---ACGAGTCGAAGGCGTTCGACCAGATCGACAACGC 65
MTUB-H37Rv              GCACGACAAATTCCCCGATCTGA---ACGAGACGAAGGCATTCGACCAGATCGACAACGC 164
MTUB-CDC1551            GCACGACAAATTCCCCGATCTGA---ACGAGACGAAGGCATTCGACCAGATCGACAACGC 164
MAVI-DAS                ----------------------------------------------------AACGC 5
NCTC13129               GGCAGACGCTTACCCAGAGCTGA---ACGAAGCTTTCGCTTTCGATGCCATCGATAAGGC 164
NMEN-MC58               GGCTAAAAAATTCGGCGG---------TGCTGCAAAAGCTTACGACCAAATCGACAACGC 158
NMEN-Z2491              GGCTAAAAAATTCGGCGG---------TGCTGCAAAAGCTTACGACCAAATCGACAACGC 158
NMEN-DAS                ------------------------------------------------------------

MPNE-HF2                ACCTGAAGAAAAAGCTAGAGGTATTACTATTAATACTGCTCACGTTGAATATGAAACAGA 218
MPNE-M129               TCCGGAAGAAAAAGCACGGGGAATTACGATTAACTCCGCTCACGTGGAGTACTCCTCTGA 218
LPNE-Paris              GCCAGAAGAAAGAGAGCGTGGTATTACTATTTCTACAGCACACGTTGAGTACGAATCAGC 231
LPNE2-DAS               GCCAGAAGAAAGAGAGCGTGGGATTACTATTTCTACAGCACATGTTGAATACGAATCAGC 101
```

```
LPNE-Philadelphia1  GCCAGAAGAAAGAGAGCGTGGGATTACTATTTCTACAGCACATGTTGAATACGAATCAGC 218
LPNE-Lens           GCCAGAAGAAAGAGAGCGTGGTATTACTATATCTACAGCCCATGTTGAATACGAATCAGC 218
LLON-DAS            ACCAGAAGAAAGAGAGCGAGGCATCACAATCTCAACAGCGCACGTTGAATACGAATCAGC 102
FNUC-DAS            TCCAGAAGAAAAAGAAAGAGGAATTACTATTAATACAGCTCATATTGAATATGAAACAGC 96
CPNE-CWL029         TCCAGAAGAAAAGGCTCGTGGAATTACTATCAACGCTTCTCACGTTGAATACGAAACCCC 218
EFAE-DAS            TCCAGAAGAAAAAGAACGTGGAATCACAATCAACACTTCTCATATCGAATATGAAACTGA 95
EFAE-V583           TCCAGAAGAAAAAGAACGTGGAATCACAATCAACACTTCTCATATCGAATATGAAACTGA 218
SAUR-Mu50           TCCAGAAGAAAAAGAACGTGGTATCACAATCAATACTTCTCACATTGAGTACCAAACTGA 218
SAUR-N315           TCCAGAAGAAAAAGAACGTGGTATCACAATCAATACTTCTCACATTGAGTACCAAACTGA 218
SAUR-MW2            TCCAGAAGAAAAAGAACGTGGTATCACAATCAATACTTCTCACATTGAGTACCAAACTGA 218
SAUR-MSSA476        TCCAGAAGAAAAAGAACGTGGTATCACAATCAATACTTCTCACATTGAGTACCAAACTGA 218
SAUR-MRSA252        TCCAGAAGAAAAAGAACGTGGTATCACAATCAATACTTCACACATTGAGTACCAAACTGA 218
SEPI-12228          TCCAGAAGAAAAAGAACGTGGTATTACAATCAATACTGCACATATCGAATACCAAACTGA 218
SEPI-RP62A          TCCAGAAGAAAAAGAACGTGGTATTACAATCAATACTGCACATATCGAATACCAAACTGA 218
SMIT-DAS            TCCAGAAGAACGCGAACGCGGTATCACTATCAACACTGCGCACGTTGAGTACGAAACTGA 86
SSAN-DAS            TCCAGAAGAACGCGAACGCGGAATCACTATCAACACTGCGCACGTTGAGTATGAAACTGC 127
SMIT-NCTC12261      TCCAGAAGAACGCGAACGTGGTATCACAATCAACACTGCGCACGTTGAGTACGAAACTGA 227
SPNE-R6             TCCAGAAGAACGCGAACGCGGTATCACTATCAACACTGCGCACGTTGAGTACGAAACTGA 227
SPYO-M1GAS          TCCAGAAGAACGCGAACGCGGAATCACTATCAACACTGCACACGTTGAGTACGAAACTGC 227
SMUT-DAS            TCCTGAAGAACGCGAACGCGGTATCACTATCAATACTGCGCACGTTGAGTATGAAACTGA 127
SMUT-UA159          TCCTGAAGAACGCGAACGCGGTATCACTATCAATACTGCGCACGTTGAGTATGAAACTGA 227
HINF-RdKW20         GCCAGAAGAAAAAGCGCGTGGTATTACCATCAACACTTCACACGTTGAATACGATACACC 218
MCAT-DAS            ACCTGAAGAAAAAGCACGTGGTATTACTATTAACACATCTCACATCGAGTATGACACAGA 69
BPER-TohamaI        GCCGGAAGAGAAGGCGCGTGGGATCACGATCAACACCTCGCACGTTGAGTACGAGACGGA 218
BPAR-12822          GCCGGAAGAGAAGGCGCGTGGGATCACGATCAACACCTCGCACGTTGAGTACGAGACGGA 218
BCEP2-DAS           GCCGGAAGAAAAGGCACGCGGCATCACGATCAACACCGCGCACATCGAGTACGAAACGGC 102
BPSE-1710b          GCCGGAAGAAAAGGCGCGCGGCATCACGATCAACACCGCGCACATCGAGTACGAAACGGC 218
SMAL-DAS            GCCGGAAGAAAAGGCTCGTGGTATCACGATCTCGACCGCGCACGTCGAATACGAATCCCC 123
PAER-PA01           GCCGGAAGAAAAGGCCCGCGGTATCACCATCAACACCTCGCACGTTGAATACGATTCCGC 218
KPNE-DAS            GCCGGAAGAAAAAGCTCGTGGTATCACCATCAACACCTCTCACGTTGAATATGACACCCC 102
MAVI-K10            GCCCGAGGAGCGTCAGCGCGGTATCACCATCAACATCTCCCACGTGGAGTACCAGACCGA 224
MAVI-paratub        GCCCGAGGAGCGTCAGCGCGGTATCACCATCAACATCTCCCACGTGGAGTACCAGACCGA 224
MAISpig-DAS         TCCCGAGGAGCGTCAGCGCGGTATCACGATCAACATCTCCCACGTGGAGTACCAGACCGA 119
MKAN-DAS            TCCTGAGGAGCGCCAGCGCGGTATCACGATCAACATCGCGCACGTGGAGTACCAGACCGA 125
MTUB-H37Rv          CCCCGAGGAGCGTCAGCGCGGTATCACCATCAACATCGCGCACGTGGAGTACCAGACCGA 224
MTUB-CDC1551        CCCCGAGGAGCGTCAGCGCGGTATCACCATCAACATCGCGCACGTGGAGTACCAGACCGA 224
MAVI-DAS            GCCTGAAGAGCGTCAGCGCGGTATCACCATCAACATCTCGCATGTTGAGTACCAGACCGA 65
CIDP-NCTC13129      ACCGGAAGAGAAAGAGCGTGGTATTACCATCAACACTCTCCACGTGGAGTACCAGACCGA 224
NMEN-MC58           ACCCGAAGAAAAAGCACGCGGTATTACCATTAACACCTCGCACGTGGAATACGAAACCGA 218
NMEN-Z2491          ACCCGAAGAAAAAGCACGCGGTATTACCATTAACACCTCGCACGTGGAATACGAAACCGA 218
NMEN-DAS            -CCCGAAGAAAAAGCACGCGGTATTACCATTAACACCTCGCACGTGGAATACGAAACCGA 59

MPNE-HF2            AAACAGACACTATGCTCACGTAGACTGTCCAGGTCACGCTGACTACGTTAAGAACATGAT 278
MPNE-M129           CAAGCGTCACTATGCTCACGTTGACTGTCCAGGACACGCTGACTACATTAAGAACATGAT 278
LPNE-Paris          GAGCAGACATTATGCGCACGTAGACTGCCCGGGGCATGCTGACTATGTTAAGAACATGAT 291
LPNE2-DAS           GAGCAGACATTATGCGCACGTAGACTGCCCGGGGCATGCTGACTATGTTAAGAACATGAT 161
LPNE-Philadelphia1  GAGCAGACATTATGCACACGTAGACTGCCCGGGACATGCTGACTATGTTAAGAACATGAT 278
LPNE-Lens           GAGCAGACATTATGCACACGTAGACTGCCCGGGACATGCTGACTATGTTAAGAACATGAT 278
LLON-DAS            GAACAGACACTATGCGCATGTTGATTGCCCAGGACATGCTGACTATGTGAAGAACATGAT 162
FNUC-DAS            AAATAGACACTATGCTCACGTTGACTGTCCAGGGCATGCTGACTATGTTAAAAATATGAT 156
CPNE-CWL029         AAATCGTCACTACGCTCACGTAGACTGCCCTGGTCACGCTGACTATGTTAAAAATATGAT 278
EFAE-DAS            AACTCGTCACTATGCACACGTTGACTGCCCAGGACACGCGGACTACGTTAAAAACATGAT 155
EFAE-V583           AACTCGTCACTATGCACACGTTGACTGCCCAGGACATGCGGACTACGTTAAAAACATGAT 278
SAUR-Mu50           CAAACGTCACTACGCTCACGTTGACTGCCCAGGACACGCTGACTACGTTAAAAACATGAT 278
SAUR-N315           CAAACGTCACTACGCTCACGTTGACTGCCCAGGACACGCTGACTACGTTAAAAACATGAT 278
SAUR-MW2            CAAACGTCACTACGCTCACGTTGACTGCCCAGGACACGCTGACTACGTTAAAAACATGAT 278
SAUR-MSSA476        CAAACGTCACTACGCTCACGTTGACTGCCCAGGACACGCTGACTACGTTAAAAACATGAT 278
SAUR-MRSA252        CAAACGTCACTACGCTCACGTTGACTGCCCAGGACACGCTGACTATGTTAAAAACATGAT 278
SEPI-12228          CAAACGTCACTATGCTCACGTTGACTGCCCAGGACACGCTGACTATGTTAAAAACATGAT 278
SEPI-RP62A          CAAACGTCACTATGCTCACGTTGACTGCCCAGGACACGCTGACTATGTTAAAAACATGAT 278
SMIT-DAS            AAAACGTCACTACGCTCACATCGACGCTCCAGGACACGCGGACTACGTTAAAAACATGAT 146
SSAN-DAS            TAAGCGTCACTACGCTCACATCGACGCTCCAGGACACGCGGACTACGTTAAAAACATGAT 187
SMIT-NCTC12261      AAAACGTCACTACGCTCACATCGACGCTCCAGGACACGCGGACTATGTTAAAAACATGAT 287
SPNE-R6             AAAACGTCACTACGCTCACATCGACGCTCCAGGACACGCGGACTACGTTAAAAACATGAT 287
SPYO-M1GAS          AACTCGTCACTATGCGCACATCGACGCTCCAGGACACGCGGACTACGTTAAAAACATGAT 287
SMUT-DAS            AAAACGTCACTATGCTCACATTGACGCTCCAGGACACGCGGACTATGTTAAAAACATGAT 187
SMUT-UA159          AAAACGTCACTATGCTCACATTGATGCTCCAGGACACGCGGACTATGTTAAAAACATGAT 287
HINF-RdKW20         GACTCGCCACTACGCACACGTAGACTGTCCGGGACACGCCGACTATGTTAAAAATATGAT 278
MCAT-DAS            AGCTCGTCACTACGCACATGTAGACTGCCCAGGCCACGCTGACTATGTTAAAAACATGAT 129
BPER-TohamaI        GACGCGTCACTACGCGCACGTTGATTGCCCGGGTCACGCTGACTACGTGAAGAACATGAT 278
BPAR-12822          GACGCGTCACTACGCGCACGTTGATTGCCCGGGTCACGCTGACTACGTGAAGAACATGAT 278
BCEP2-DAS           GAACCGCCACTACGCGCACGTCGACTGCCCGGGCCACGCCGACTACGTGAAGAACATGAT 162
BPSE-1710b          GAACCGCCACTACGCACACGTGGACTGCCCGGGCCACGCCGACTACGTGAAGAACATGAT 278
```

```
SMAL-DAS          GATCCGTCACTACGCCCACGTTGACTGCCCGGGCCACGCTGACTACGTCAAGAACATGAT 183
PAER-PA01         TGTTCGTCACTACGCCCACGTTGACTGCCCCGGTCACGCCGACTACGTGAAGAACATGAT 278
KPNE-DAS          GACTCGCCACTACGCGCACGTAGACTGCCCGGGCCACGCCGACTATGTTAAAAACATGAT 162
MAVI-K10          CAAGCGGCACTACGCTCACGTCGACGCCCCGGGTCACGCCGACTACATCAAGAACATGAT 284
MAVI-paratub      CAAGCGGCACTACGCTCACGTCGACGCCCCGGGTCACGCCGACTACATCAAGAACATGAT 284
MAISpig-DAS       CAAGCGGCACTATGCCCACGTCGACGCCCCGGGTCACGCCGACTACATCAAGAACATGAT 179
MKAN-DAS          GAAGCGGCACTATGCACACGTCGACGCGCCGGGCCACGCCGACTACATCAAGAACATGAT 185
MTUB-H37Rv        CAAGCGGCACTACGCACACGTCGACGCCCCTGGCCACGCCGACTACATCAAGAACATGAT 284
MTUB-CDC1551      CAAGCGGCACTACGCACACGTCGACGCCCCTGGCCACGCCGACTACATCAAGAACATGAT 284
MAVI-DAS          GAAGCGCCACTACGCACACGTGGACGCCCCCGGCCACGCGGACTACATCAAGAACATGAT 125
CIDP-NCTC13129    GAAGCGCCACTACGCACACGTTGACGCTCCAGGTCACGCTGACTACATCAAGAACATGAT 284
NMEN-MC58         AACCCGCCACTACGCACACGTAGACTGCCCGGGGCACGCCGACTACGTTAAAAACATGAT 278
NMEN-Z2491        AACCCGCCACTACGCACACGTAGACTGTCCGGGGCACGCCGACTACGTTAAAAACATGAT 278
NMEN-DAS          AACCCGCCACTACGCACACGTAGACTGCCCGGGGCACGCCGACTACGTTAAAAACATGAT 119

MPNE-HF2          TACTGGTGCTGCTCAAATGGATGGTGCAATCTTAGTAGTTGCTGCATCTGATGGACCAAT 338
MPNE-M129         TACTGGTGCTGCACAAATGGATGGTGCCATTCTAGTAGTTTCAGCAACTGACAGTGTTAT 338
LPNE-Paris        TACTGGTGCTGCGCAAATGGACGGAGCGATACTGGTTGTATCAGCAGCTGATGGTCCTAT 351
LPNE2-DAS         TACTGGTGCTGCGCAAATGGACGGAGCGATACTGGTTGTATCAGCAGCTGATGGTCCTAT 221
LPNE-Philadelphia1 TACTGGTGCTGCGCAAATGGACGGAGCGATACTGGTTGTATCAGCAGCTGATGGTCCTAT 338
LPNE-Lens         TACTGGTGCTGCGCAAATGGACGGAGCGATACTGGTTGTATCAGCAGCTGATGGTCCTAT 338
LLON-DAS          TACAGGTGCTGCACAAATGGACGGAGCGATACTTGTAGTATCTGCAGCAGATGGTCCTAT 222
FNUC-DAS          AACTGGTGCTGCACAAATGGATGGAGCTATACTTGTTGTATCAGCTGCTGATGGTCCTAT 216
CPNE-CWL029       TACAGGCGCCGCTCAAATGGACGGAGCTATCCTAGTCGTTTCAGCTACAGACGGAGCTAT 338
EFAE-DAS          CACTGGTGCTGCTCAAATGGACGGAGCTATCTTAGTAGTTTCTGCTGCTGATGGTCCTAT 215
EFAE-V583         CACTGGTGCTGCTCAAATGGACGGAGCTATCTTAGTAGTTTCTGCTGCTGATGGTCCTAT 338
SAUR-Mu50         CACTGGTGCTGCTCAAATGGACGGCGGTATCTTAGTAGTATCTGCTGCTGACGGTCCAAT 338
SAUR-N315         CACTGGTGCTGCTCAAATGGACGGCGGTATCTTAGTAGTATCTGCTGCTGACGGTCCAAT 338
SAUR-MW2          CACTGGTGCTGCTCAAATGGACGGCGGTATCTTAGTAGTATCTGCTGCTGACGGTCCAAT 338
SAUR-MSSA476      CACTGGTGCTGCTCAAATGGACGGCGGTATCTTAGTAGTATCTGCTGCTGACGGTCCAAT 338
SAUR-MRSA252      CACTGGTGCTGCTCAAATGGACGGTGGTATCTTAGTAGTATCTGCTGCTGACGGTCCAAT 338
SEPI-12228        CACTGGTGCAGCTCAAATGGACGGCGGTATCTTAGTTGTATCTGCTGCTGACGGTCCAAT 338
SEPI-RP62A        CACTGGTGCAGCTCAAATGGACGGCGGTATCTTAGTTGTATCTGCTGCTGACGGTCCAAT 338
SMIT-DAS          CACTGGTGCCGCTCAAATGGACGGAGCTATCCTTGTAGTAGCTTCAACTGACGGACCAAT 206
SSAN-DAS          CACTGGTGCCGCTCAAATGGACGGAGCTATCCTTGTAGTAGCTTCAACTGACGGACCAAT 247
SMIT-NCTC12261    CACTGGTGCCGCTCAAATGGACGGAGCTATCCTTGTAGTAGCTTCAACTGACGGACCAAT 347
SPNE-R6           CACTGGTGCTGCTCAAATGGACGGAGCTATCCTTGTAGTAGCTTCAACTGACGGACCAAT 347
SPYO-M1GAS        CACTGGTGCCGCTCAAATGGACGGAGCTATCCTTGTAGTTGCTTCAACTGATGGACCAAT 347
SMUT-DAS          TACTGGTGCTGCCCAAATGGATGGTGCTATCCTTGTAGTAGCTTCAACTGATGGTCCAAT 247
SMUT-UA159        CACTGGTGCTGCCCAAATGGATGGTGCTATCCTTGTAGTAGCTTCAACTGATGGACCAAT 347
HINF-RdKW20       TACTGGTGCGGCACAAATGGATGGTGCTATTTTAGTAGTAGCAGCAACAGATGGTCCTAT 338
MCAT-DAS          CACAGGTGCTGCACAGATGGACGGCGCAATCTTGGTTGTTTCTGCAACTGACGGCCCAAT 189
BPER-TohamaI      CACGGGTGCTGCGCAGATGGACGGCGCGATCCTGGTGGTGTCGGCCGCAGACGGCCCGAT 338
BPAR-12822        CACGGGTGCTGCGCAGATGGACGGCGCGATCCTGGTGGTGTCGGCCGCAGACGGCCCGAT 338
BCEP2-DAS         CACGGGTGCAGCGCAGATGGACGGCGCAATCCTGGTGTGCTCGGCCGCAGACGGCCCGAT 222
BPSE-1710b        CACGGGCGCGGCGCAGATGGACGGCGCGATCCTGGTGTGCTCGGCCGCTGACGGCCCGAT 338
SMAL-DAS          CACCGGTGCCGCCCAGATGGACGGCGCGATCCTGGTGTGCTCGGCCGCTGACGGCCCGAT 243
PAER-PA01         CACCGGTGCTGCCCAGATGGACGGCGCGATCCTGGTTTGCTCGGCTGCCGACGGCCCCAT 338
KPNE-DAS          CACCGGTGCTGCGCAGATGGACGGCGCGATCCTGGTTGTTGCTGCGACTGACGGCCCGAT 222
MAVI-K10          CACCGGCGCCGCCCAGATGGACGGCGCGATCCTGGTGGTCGCCGCCACCGACGGCCCGAT 344
MAVI-paratub      CACCGGCGCCGCCCAGATGGACGGCGCGATCCTGGTGGTCGCCGCCACCGACGGCCCGAT 344
MAISpig-DAS       CACCGGTGCGGCCCAGATGGACGGCGCGATCCTGGTGGTTGCCGCCACCGACGGCCCGAT 239
MKAN-DAS          CACCGGCGCCGCCCAGATGGACGGCGCGATCCTGGTGGTCGCCGCCACGGACGGCCCGAT 245
MTUB-H37Rv        CACCGGCGCCGCGCAGATGGACGGTGCGATCCTGGTGGTCGCCGCCACCGACGGCCCGAT 344
MTUB-CDC1551      CACCGGCGCCGCGCAGATGGACGGTGCGATCCTGGTGGTCGCCGCCACCGACGGCCCGAT 344
MAVI-DAS          CACCGGTGCCGCCCAGATGGACGGCGCGATCCTGGTGGTCGCCGCGACCGACGGCCCGAT 185
NCTC13129         CACCGGTGCTGCTCAGATGGACGGCGCAATCCTCGTTGTTGCTGCCACCGACGGCCCAAT 344
NMEN-MC58         TACCGGCGCCGCACAAATGGACGGTGCAATCCTGGTATGTTCCGCAGCCGACGGCCCTAT 338
NMEN-Z2491        TACCGGCGCAGCACAAATGGACGGTGCAATCCTGGTATGTTCCGCAGCCGACGGCCCTAT 338
NMEN-DAS          TACCGGTGCCGCACAAATGGACGGTGCAATCCTGGTATGTTCCGCAGCCGACGGCCCTAT 179

MPNE-HF2          GCCACAAACAAGAGAACACATCTTATTAGCTAGACAAGTTGGTGTTCCTAAAATGGTAGT 398
MPNE-M129         GCCCCAAACCCGTGAACACATTTTGTTGGCCCGCCAAGTGGGTGTGCCACGCATGGTAGT 398
LPNE-Paris        GCCACAAACGAGGGAACACATTCTATTGTCTCGCCAGGTAGGTGTTCCATATATTGTTGT 411
LPNE2-DAS         GCCACAAACGAGGGAACACATTCTATTGTCTCGCCAGGTAGGTGTTCCATATATTGTTGT 281
LPNE-Philadelphia1 GCCACAAACGAGGGAACACATTCTATTGTCTCGCCAGGTAGGTGTTCCATATATTGTTGT 398
LPNE-Lens         GCCACAAACGAGGGAACACATTCTATTGTCTCGCCAGGTAGGTGTTCCATATATTGTTGT 398
LLON-DAS          GCCTCAAACTAGAGAACATATTCTATTGTCACGACAAGTAGGGGTGCCCTACATAGTAGT 282
FNUC-DAS          GCCTCAAACAAGAGAACATATCTTACTTTCTAGACAAGTTGGAGTACCATATATTATTGT 276
CPNE-CWL029       GCCACAAACTAAAGAACATATCTTGCTAGCTCGCCAGGTTGGAGTTCCTTATATCGTTGT 398
EFAE-DAS          GCCTCAAACACGTGAACATATCTTATTATCACGTAACGTTGGTGTACCATACATCGTTGT 275
EFAE-V583         GCCTCAAACACGTGAACATATCTTATTATCACGTAACGTTGGTGTACCATACATCGTTGT 398
SAUR-Mu50         GCCACAAACTCGTGAACACATTCTTTTATCACGTAACGTTGGTGTACCAGCATTAGTAGT 398
```

```
SAUR-N315          GCCACAAACTCGTGAACACATTCTTTTATCACGTAACGTTGGTGTACCAGCATTAGTAGT 398
SAUR-MW2           GCCACAAACTCGTGAACACATTCTTTTATCACGTAACGTTGGTGTACCAGCATTAGTAGT 398
SAUR-MSSA476       GCCACAAACTCGTGAACACATTCTTTTATCACGTAACGTTGGTGTACCAGCATTAGTAGT 398
SAUR-MRSA252       GCCACAAACTCGTGAACACATTCTTTTATCACGTAACGTTGGTGTACCAGCATTAGTAGT 398
SEPI-12228         GCCACAAACTCGTGAACACATCTTATTATCACGTAACGTTGGTGTACCAGCATTAGTTGT 398
SEPI-RP62A         GCCACAAACTCGTGAACACATCTTATTATCACGTAACGTTGGTGTACCAGCATTAGTTGT 398
SMIT-DAS           GCCACAAACTCGTGAGCACATCCTTCTTTCACGTCAGGTTGGTGTTAAACACTTGATCGT 266
SSAN-DAS           GCCACAAACTCGTGAGCACATCTTGCTTTCACGTCAGGTTGGTGTTAAACACTTGATCGT 307
SMIT-NCTC12261     GCCACAAACTCGTGAGCACATCCTTCTTTCACGTCAGGTTGGTGTTAAACACCTTATCGT 407
SPNE-R6            GCCACAAACTCGTGAGCACATCCTTCTTTCACGTCAGGTTGGTGTTAAACACCTTATCGT 407
SPYO-M1GAS         GCCACAAACTCGTGAGCACATCCTTCTTTCACGTCAGGTTGGTGTTAAACACCTTATCGT 407
SMUT-DAS           GCCACAAACTCGTGAACACATTCTTCTTTCACGTCAAGTTGGTGTTAAATACCTTATTGT 307
SMUT-UA159         GCCACAAACTCGTGAACACATTCTTCTTTCACGTCAAGTTGGTGTTAAATACCTCATTGT 407
HINF-RdKW20        GCCACAAACTCGTGAACACATCTTATTAGGTCGCCAAGTAGGTGTTCCATACATCATCGT 398
MCAT-DAS           GCCACAAACTCGTGAGCACATCCTACTGTCTCGTCAGGTTGGTGTACCATATATCATGGT 249
BPER-TohamaI       GCCGCAGACGCGCGAGCACATTTTGCTGTCGCGCCAGGTTGGCGTGCCGTACATCATCGT 398
BPAR-12822         GCCGCAGACGCGCGAGCACATTTTGCTGTCGCGCCAGGTTGGCGTGCCGTACATCATCGT 398
BCEP2-DAS          GCCGCAAACGCGTGAGCACATCCTGCTGGCGCGTCAGGTTGGCGTTCCGTACATCATCGT 282
BPSE-1710b         GCCGCAAACGCGTGAGCACATCCTGCTGGCGCGTCAGGTCGGTGTGCCGTACATCATCGT 398
SMAL-DAS           GCCGCAGACCCGTGAGCACATCCTGCTGTCGCGCCAGGTCGGCGTGCCGTACATCGTCGT 303
PAER-PA01          GCCGCAGACCCGCGAGCACATCCTGCTGTCCCGCCAGGTAGGCGTTCCCTACATCGTCGT 398
KPNE-DAS           GCCGCAGACTCGTGAGCACATCCTGCTGGGTCGTCAGGTAGGCGTTCCGTACATCATCGT 282
MAVI-K10           GCCGCAGACCCGCGAGCACGTGCTGCTGGCCCGTCAGGTCGGTGTGCCCTACATCCTGGT 404
MAVI-paratub       GCCGCAGACCCGCGAGCACGTGCTGCTGGCCCGTCAGGTCGGTGTGCCCTACATCCTGGT 404
MAISpig-DAS        GCCGCAGACCCGTGAGCACGTGCTGCTCGCGCGTCAGGTCGGGGTGCCCTACATCCTGGT 299
MKAN-DAS           GCCGCAGACCCGTGAGCACGTGCTGCTCGCACGTCAGGTGGGTGTGCCCTACATCCTGGT 305
MTUB-H37Rv         GCCCCAGACCCGCGAGCACGTTCTGCTGGCGCGTCAAGTGGGTGTGCCCTACATCCTGGT 404
MTUB-CDC1551       GCCCCAGACCCGCGAGCACGTTCTGCTGGCGCGTCAAGTGGGTGTGCCCTACATCCTGGT 404
MAVI-DAS           GCCGCAGACCCGCGAGCACGTGCTGCTCGGCCGTCAGGTGGGTGTGCCCTACATCCTCGT 245
NCTC13129          GCCTCAGACCCGTGAGCACGTTCTGCTCGCTCGCCAGGTCGGCGTTCCTTACATCCTCGT 404
NMEN-MC58          GCCGCAAACCCGCGAACACATCCTGCTGGCCCGCCAAGTAGGCGTACCTTACATCATCGT 398
NMEN-Z2491         GCCGCAAACCCGCGAACACATCCTGCTGGCCCGTCAAGTAGGCGTACCTTACATCATCGT 398
NMEN-DAS           GCCGCAAACCCGCGAACACATCCTGCTGGCCCGTCAAGTAGGCGTACCTTACATCATCGT 239

MPNE-HF2           TTTCTTAAACAAATGTGATATGGTA------TCTGATGCTGAAATGCAAGACCTAGTAGA 452
MPNE-M129          GTTCCTAAACAAGTGTGACATTGCA------ACTGATGAAGAAGTGCAAGAGTTAGTAGC 452
LPNE-Paris         GTTCATGAACAAAGCGGATATGGTT------GATGACCCTGAGTTATTAGAGTTAGTGGA 465
LPNE2-DAS          GTTCATGAACAAAGCGGATATGGTT------GATGACCCTGAGTTATTAGAGTTAGTGGA 335
LPNE-Philadelphia1 GTTCATGAACAAAGCGGATATGGTT------GATGACCCTGAGTTATTAGAGTTAGTGGA 452
LPNE-Lens          GTTCATGAACAAAGCGGATATGGTT------GATGACCCTGAGTTATTAGAGTTAGTGGA 452
LLON-DAS           GTTTATGAATAAAGCGGACATGGTT------GATGATCCAGAGCTCTTAGAGCTTGTAGA 336
FNUC-DAS           TTATTTAAATAAATCAGATATGGTT------GATGATGAAGAATTACTAGAATTGGTAGA 330
CPNE-CWL029        TTTCTTGAATAAAGTAGATATGATCTCTCAAGAAGATGCTGAACTTATTGACCTTGTTGA 458
EFAE-DAS           ATTCTTAAACAAAATGGATATGGTT------GATGACGAAGAATTATTAGAATTAGTAGA 329
EFAE-V583          ATTCTTAAACAAAATGGATATGGTT------GATGACGAAGAATTATTAGAATTAGTAGA 452
SAUR-Mu50          ATTCTTAAACAAAGTTGACATGGTT------GACGATGAAGAATTATTAGAATTAGTAGA 452
SAUR-N315          ATTCTTAAACAAAGTTGACATGGTT------GACGATGAAGAATTATTAGAATTAGTAGA 452
SAUR-MW2           ATTCTTAAACAAAGTTGACATGGTT------GACGATGAAGAATTATTAGAATTAGTAGA 452
SAUR-MSSA476       ATTCTTAAACAAAGTTGACATGGTT------GACGATGAAGAATTATTAGAATTAGTAGA 452
SAUR-MRSA252       ATTCTTAAACAAAGTTGACATGGTT------GACGATGAAGAATTATTAGAATTAGTAGA 452
SEPI-12228         ATTCTTAAACAAAGTTGACATGGTA------GACGACGAAGAATTATTAGAATTAGTTGA 452
SEPI-RP62A         ATTCTTAAACAAAGTTGACATGGTA------GACGACGAAGAATTATTAGAATTAGTTGA 452
SMIT-DAS           CTTCATGAACAAAGTTGACTTGGTT------GACGATGAAGAATTGCTTGAATTGGTTGA 320
SSAN-DAS           CTTCATGAACAAAGTTGACTTGGTT------GACGATGAAGAATTGCTTGAATTGGTTGA 361
SMIT-NCTC12261     CTTCATGAACAAAGTTGACTTGGTT------GACGACGAAGAATTGCTTGAATTGGTTGA 461
SPNE-R6            CTTCATGAACAAAGTTGACTTGGTT------GACGACGAAGAATTGCTTGAATTGGTTGA 461
SPYO-M1GAS         GTTCATGAACAAAGTTGACCTTGTT------GATGACGAAGAGTTGCTTGAATTAGTTGA 461
SMUT-DAS           CTTCATGAATAAAGTTGATTTGGTT------GATGATGAAGAACTGCTTGAATTGGTTGA 361
SMUT-UA159         CTTCATGAATAAAGTTGATTTGGTT------GACGATGAAGAATTGCTTGAATTGGTTGA 461
HINF-RdKW20        ATTCTTAAACAAATGCGACATGGTA------GATGACGAAGAGTTATTAGAATTAGTCGA 452
MCAT-DAS           ATTCATGAACAAGTGCGACATGGTA------GATGATGAAGAGCTACTAGAATTGGTTGA 303
BPER-TohamaI       GTTCCTGAACAAGGCGGACATGGTT------GATGACGCGGAGCTGCTCGAGCTGGTGGA 452
BPAR-12822         GTTCCTGAACAAGGCGGACATGGTT------GATGACGCGGAGCTGCTCGAGCTGGTGGA 452
BCEP2-DAS          GTTCCTGAACAAGTGCGACATGGTG------GACGACGCTGAACTGCTCGAGCTGGTCGA 336
BPSE-1710b         GTTCCTGAACAAGTGCGATATGGTG------GACGACGCGGAGCTGCTCGAGCTGGTCGA 452
SMAL-DAS           GTTCCTGAACAAGGCCGACATGGTC------GACGACGCCGAGCTGCTCGAGCTGGTCGA 357
PAER-PA01          GTTCCTGAACAAGGCCGACATGGTC------GACGACGCCGAGCTGCTGGAACTGGTCGA 452
KPNE-DAS           GTTCCTGAACAAATGCGACATGGTT------GATGACGAAGAGCTGCTGGAACTGGTTGA 336
MAVI-K10           CGCGCTGAACAAGGCCGACATGGTC------GACGACGAGGAGCTGCTCGAGCTCGTCGA 458
MAVI-paratub       CGCGCTGAACAAGGCCGACATGGTC------GACGACGAGGAGCTGCTCGAGCTCGTCGA 458
MAISpig-DAS        CGCCCTCAACAAGGCCGACATGGTC------GACGACGAGGAGCTCCTCGAGCTCGTCGA 353
MKAN-DAS           CGCGCTGAACAAGGCCGACGCCGTC------GACGACGAGGAGTTGCTCGAACTCGTCGA 359
MTUB-H37Rv         AGCGCTGAACAAGGCCGACGCAGTG------GACGACGAGGAGCTGCTCGAACTCGTCGA 458
```

```
MTUB-CDC1551          AGCGCTGAACAAGGCCGACGCAGTG------GACGACGAGGAGCTGCTCGAACTCGTCGA 458
MAVI-DAS              GGCGCTGAACAAGTCGGACATGGTC------GACGACGAGGAGCTCCTCGAGCTCGTCGA 299
NCTC13129             TGCTCTGAACAAGTGCGACATGGTT------GATGATGAGGAAATCATCGAGCTCGTCGA 458
NMEN-MC58             GTTCATGAACAAATGCGACATGGTC------GACGATGCCGAGCTGTTGGAACTGGTTGA 452
NMEN-Z2491            GTTCATGAACAAATGCGACATGGTC------GACGATGCCGAGCTGTTGGAACTGGTTGA 452
NMEN-DAS              GTTCATGAACAAATGCGACATGGTC------GACGATGCCGAGCTGTTGGAACTGGTTGA 293

MPNE-HF2               AATGGAAGTTAGAGAATTACTTTCTTCTTATGGTTTTGATGGAGATAACACTCCAGTTAT 512
MPNE-M129             AGAAGAGGTACGTGACTTATTAACTTCTTACGGCTTTGATGGCAAGAACACCCTATTAT 512
LPNE-Paris            AATGGAAGTGCGAGATTTATTAAGCAGTTACGATTTCCCAGGGGATGACATACCTATTGT 525
LPNE2-DAS             AATGGAAGTGCGAGATTTATTAAGCAGTTACGATTTCCCAGGGGATGACATACCCATTGT 395
LPNE-Philadelphia1    AATGGAAGTGCGAGATTTATTAAGCAGTTACGATTTCCCAGGGGATGACATACCTATTGT 512
LPNE-Lens             AATGGAAGTGCGAGATTTATTAAGCAGTTACGATTTCCCAGGGGATGATATACCTATTAT 512
LLON-DAS              AATGGAAGTCCGTGATTTGTTGAGCAGTTATGATTTTCCTGGTGATGATATTCCAATTGT 396
FNUC-DAS              AATGGAAGTTAGAGAATTATTAACTGAATATGGATTCCCAGGAGATGACATCCCTGTAAT 390
CPNE-CWL029           GATGGAACTTAGTGAGCTTCTTGAAGAAAAAGGCT--ACAAAGGATG----CCCTATTAT 512
EFAE-DAS              AATGGAAGTTCGTGACTTATTATCAGAATACGATTTCCCAGGCGATGATGTTCCAGTTAT 389
EFAE-V583             AATGGAAGTTCGTGACTTATTATCAGAATACGATTTCCCAGGCGATGATGTTCCAGTTAT 512
SAUR-Mu50             AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SAUR-N315             AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SAUR-MW2              AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SAUR-MSSA476          AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SAUR-MRSA252          AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SEPI-12228            AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SEPI-RP62A            AATGGAAGTTCGTGACTTATTAAGCGAATATGACTTCCCAGGTGACGATGTACCTGTAAT 512
SMIT-DAS              AATGGAAATCCGTGACCTTCTTTCAGAATACGATTTCCCAGGTGATGACCTTCCAGTTAT 380
SSAN-DAS              AATGGAAATCCGTGACCTCTTGTCAGAATACGACTTCCCAGGTGACGATCTTCCAGTTAT 421
SMIT-NCTC12261        AATGGAAATCCGTGACCTATTGTCAGAATACGACTTCCCAGGTGACGATCTTCCAGTTAT 521
SPNE-R6               AATGGAAATCCGTGACCTATTGTCAGAATACGACTTCCCAGGTGACGATCTTCCAGTTAT 521
SPYO-M1GAS            GATGGAAATTCGTGACCTTCTTTCAGAATACGATTTCCCAGGTGATGACCTTCCAGTTAT 521
SMUT-DAS              AATGGAAATCCGTGATCTTCTTTCAGAATATGATTTCCCAGGTGATGATATTCCAGTTAT 421
SMUT-UA159            AATGGAAATCCGTGATCTTCTTTCAGAATATGATTTCCCAGGTGATGATATTCCAGTTAT 521
HINF-RdKW20           AATGGAAGTTCGTGAACTTCTATCTCAATATGACTTCCCAGGTGACGATACACCAATCGT 512
MCAT-DAS              AATGGAAGTTCGTGAACTTCTATCTGACTATGACTTCCCAGGTGATGACACCCCAATCAT 363
BPER-TohamaI          GATGGAAGTCCGCGAACTGCTGAGCAAGTACGATTTCCCGGGCGATGACACGCCGATCGT 512
BPAR-12822            GATGGAAGTCCGCGAACTGCTGAGCAAGTACGATTTCCCGGGCGATGACACGCCGATCGT 512
BCEP2-DAS             GATGGAAGTTCGCGAACTGCTGTCGAAGTACGACTTCCCGGGCGACGACACGCCGATCAT 396
BPSE-1710b            AATGGAAGTGCGCGAACTGCTGTCGAAGTACGACTTCCCGGGCGACGACACGCCGATCAT 512
SMAL-DAS              GATGGAAGTGCGCGAACTGCTGAGCAAGTACGAGTTCCCGGGCGACGACACCCCGATCAT 417
PAER-PA01             GATGGAAGTTCGCGATCTGCTGAACACCTACGACTTCCCGGGCGACGACACTCCGATCAT 512
KPNE-DAS              GATGGAAGTTCGTGAACTGCTGTCTCAGTACGATTTCCCGGGCGACGACACCCCGATCGT 396
MAVI-K10              GATGGAGGTCCGCGAGCTGCTGGCCGCCCAGGAGTTC---GACGAGGACGCCCCGGTGGT 515
MAVI-paratub          GATGGAGGTCCGCGAGCTGCTGGCCGCCCAGGAGTTC---GACGAGGACGCCCCGGTGGT 515
MAISpig-DAS           GATGGAGGTCCGCGAACTGCTGGCCGCCCAGGAGTTC---GACGAGGACGCCCCGGTGGT 410
MKAN-DAS              GATGGAGGTCCGCGAGCTGCTGGCCGCCCAGGAGTTC---GACGAGGACGCCCCGGTGGT 416
MTUB-H37Rv            GATGGAGGTCCGCGAGCTGCTGGCTGCCCAGGAATTC---GACGAGGACGCCCCGGTTGT 515
MTUB-CDC1551          GATGGAGGTCCGCGAGCTGCTGGCTGCCCAGGAATTC---GACGAGGACGCCCCGGTTGT 515
MAVI-DAS              GATGGAGGTCCGCGAACTGCTGGCCGCCCAGGAGTTC---GACGAGGACGCCCCGGTCAT 356
NCTC13129             GATGGAGATCCGTGAGCTGCTCGCTGAGCAGGATTAC---GACGAAGAGGCTCCAATCAT 515
NMEN-MC58             AATGGAAATCCGCGACCTGCTGTCCAGCTACGACTTCCCCGGCGATGACTGCCCGATTGT 512
NMEN-Z2491            AATGGAAATCCGCGACCTGCTGTCCAGCTACGACTTCCCCGGCGACGACTGCCCGATCGT 512
NMEN-DAS              AATGGAAATCCGCGACCTGCTGTCCAGCTACGACTTCCCCGGCGACGACTGCCCGATTGT 353

MPNE-HF2              TAGAGGTTCAGCTTTAAAAGCATTAGAAGGT---GAT----GCTACTTGAGAAGCTAAA- 564
MPNE-M129             TTATGGTTCTGCACTTAAAGCGCTTGAAGGT---GAT----CCTAAGTGGGAAGCTAAG- 564
LPNE-Paris            TGTTGGTTCAGCTTTGAAAGCATTGGAAGGT---GAAGACAGTGATATAGGCGTTAAGGC 582
LPNE2-DAS             TGTTGGTTCAGCTTTGAAAGCATTGGAAGGT---GAAGACAGTGATATAGGCGTTAAGGC 452
LPNE-Philadelphia1    TGTTGGTTCAGCTTTGAAAGCATTGGAAGGT---GAAGACAGTGATATAGGCGTTAAGGC 569
LPNE-Lens             TGTTGGTTCAGCTTTGAAAGCATTGGAAGGT---GAAGACAGTGATATAGGTGTTAAGGC 569
LLON-DAS              TGTTGGATCTGCATTGAAAGCATTAGAAGGT---GATACCAGTGATATTGGGGTGCCAGC 453
FNUC-DAS              TAGAGGTTCATCATTAGGAGCTTTAAATGGT---GA----AGAAAAATGGATAGAAAAG- 442
CPNE-CWL029           CCGTGGTTCTGCTTTGAAAGCTCTTGAAGGT---GATGCAAATTATATCGA---AAAAG- 565
EFAE-DAS              CGCAGGTTCTGCTTTGAAAGCTTTAGAAGGC---GAC----GAGTCTTATGAAGAAAAA- 441
EFAE-V583             CGCAGGTTCTGCTTTGAAAGCTTTAGAAGGC---GAC----GAGTCTTATGAAGAAAAA- 564
SAUR-Mu50             CGCTGGTTCAGCATTAAAAGCTTTAGAAGGC---GAT----GCTCAATACGAAGAAAAA- 564
SAUR-N315             CGCTGGTTCAGCATTAAAAGCTTTAGAAGGC---GAT----GCTCAATACGAAGAAAAA- 564
SAUR-MW2              CGCTGGTTCAGCATTAAAAGCTTTAGAAGGC---GAT----GCTCAATACGAAGAAAAA- 564
SAUR-MSSA476          CGCTGGTTCAGCATTAAAAGCTTTAGAAGGC---GAT----GCTCAATACGAAGAAAAA- 564
SAUR-MRSA252          CGCTGGTTCAGCATTAAAAGCTTTAGAAGGC---GAT----GCTCAATACGAAGAAAAA- 564
SEPI-12228            CGCTGGTTCTGCATTAAAAGCATTAGAAGGC---GAT----GCTGAATACGAACAAAAA- 564
SEPI-RP62A            CGCTGGTTCTGCATTAAAAGCATTAGAAGGC---GAT----GCTGAATACGAACAAAAA- 564
SMIT-DAS              CCAAGGTTCAGCTCTTAAAGCTCTTGAAGGT---GAT----GCTAAATACGAAGACATC- 432
SSAN-DAS              CCAAGGTTCAGCTCTTAAAGCTCTTGAAGGT---GAC----TCTAAATATGAAGACATC- 473
```

```
SMIT-NCTC12261      CCAAGGTTCAGCTCTTAAAGCCCTTGAAGGT---GAC----ACTAAATACGAAGACATC- 573
SPNE-R6             CCAAGGTTCAGCACTTAAAGCTCTTGAAGGT---GAC----TCTAAATACGAAGACATC- 573
SPYO-M1GAS          CCAAGGTTCAGCTCTTAAAGCTCTTGAAGGC---GAC----ACTAAATTTGAAGACATC- 573
SMUT-DAS            TCAAGGTTCAGCTCTTAAAGCTCTTGAAGGT---GAT----ACTGCTCAAGAAGATATC- 473
SMUT-UA159          TCAAGGTTCAGCTCTTAAAGCTCTTGAAGGC---GAT----ACTGCTCAAGAAGATATC- 573
HINF-RdKW20         ACGTGGTTCAGCATTACAAGCGTTAAACGGC---GTA----GCAGAATGGGAAGAAAAA- 564
MCAT-DAS            CAAAGGTTCAGCACTAGAAGCATTGAATGGT---TCTGACGGTAAATATGGTGAGCCTGC 420
BPER-TohamaI        GAAGGGTTCGGCCAAGCTGGCGCTGGAAGGC---GACAAGGGCGAACTGGGCGAGCAGGC 569
BPAR-12822          GAAGGGTTCGGCCAAGCTGGCGCTGGAAGGC---GACAAGGGCGAACTGGGCGAGCAGGC 569
BCEP2-DAS           CAAGGGTTCGGCAAAGCTGGCGCTGGAAGGC---GACAAGGGCGAGCTGGGCGAAACGGC 453
BPSE-1710b          CAAGGGTTCGGCGAAGCTGGCGCTGGAAGGC---GACAAGGGCGAGCTGGGCGAAGTGGC 569
SMAL-DAS            CGCCGGTTCGGCCCGCCTGGCGCTGGAAGGC---GACCAGAGCGACATCGGCGTGCCGGC 474
PAER-PA01           CATCGGTTCCGCGCTGATGGCGCTGGAAGGCAAGGATGACAACGGCATCGGCGTAAGCGC 572
KPNE-DAS            TCGTGGTTCTGCTCTGAAAGCGCTGGAAGGC---GAC----GCAGAGTGGG--AAGCGAA 447
MAVI-K10            GCGGGTGTCGGCGCTCAAGGCGCTCGAGGGC---GAC----GCCAAGTGGGTGGAGT--C 566
MAVI-paratub        GCGGGTGTCGGCGCTCAAGGCGCTCGAGGGC---GAC----GCCAAGTGGGTGGAGT--C 566
MAISpig-DAS         GCGCGTCTCGGCGCTGAAGGCGCTCGAGGGC---GAC----GCCAAGTGGGTGGAGT--C 461
MKAN-DAS            GCGGGTCTCGGCACTGAAGGCCCTCGAGGGC---GAC----CCCAAGTGGGTGGAGT--C 467
MTUB-H37Rv          GCGGGTCTCGGCGCTCAAGGCGCTCGAGGGT---GAC----GCGAAGTGGGTTGCCT--C 566
MTUB-CDC1551        GCGGGTCTCGGCGCTCAAGGCGCTCGAGGGT---GAC----GCGAAGTGGGTTGCCT--C 566
MAVI-DAS            CCGCGTCTCCGCGCTGAAGGCGCTGGAGGGT---GAC----CCGAAGTGGGTCAAGT--C 407
CIDP-NCTC13129      CCACATCTCCGCACTGAAGGCTCTTGAGGGC---GAC----GAGAAGTGGACCCAGT--C 566
NMEN-MC58           ACAAGGTTCCGCACTGAAAGCCTTGGAAGGC---GAT----GCCGCTTACGAAGAAA--A 563
NMEN-Z2491          ACAAGGTTCCGCACTGAAAGCCTTGGAAGGC---GAT----GCCGCTTACGAAGAAA--A 563
NMEN-DAS            ACAAGGTTCTGCACTGAAAGCCTTGGAAGGC---GAT----GCCGCTTATGAAGAGA--A 404

MPNE-HF2            -ATTGATGAATTAATGGCTTCAGTAGATAGCTACATCCCAACTCCAACAAGAGATACAGA 623
MPNE-M129           -ATCCATGATTTAATGAATGCAGTTGATGAATGGATTCCAACTCCTGAACGTGAAGTGGA 623
LPNE-Paris          TATAGAGAAATTGGTTGAAACAATGGATTCATACATTCCTGAGCCAGTTAGAAACATAGA 642
LPNE2-DAS           TATTGAGAAATTGGTTGAAACAATGGATTCATACATTCCTGAGCCAGTTAGAAATATAGA 512
LPNE-Philadelphia1  TATTGAGAAATTGGTTGAAACAATGGATTCATACATTCCTGAGCCAGTTAGAAACATAGA 629
LPNE-Lens           TATTGAGAAATTAGTTGAAACAATGGATTCATACATTCCTGAGCCAGTTAGAAACATAGA 629
LLON-DAS            AATAGAGAAATTGGTAGAGACTATGGATTCTTATATTCCAGAGCCTGTAAGAAACATTGA 513
FNUC-DAS            -ATAATGGAACTTATGGATGCAGTAGATAGCTATATCCCTACTCCTGAAAGAGCAGTAGA 501
CPNE-CWL029         --TTCGAGAACTTATGCAAGCTGTGGATGACAACATCCCTACACCAGAAAGAGAAATTGA 623
EFAE-DAS            -ATCTTAGAATTAATGGCTGCAGTTGACGAATATATCCCAACTCCAGAACGTGATACTGA 500
EFAE-V583           -ATCTTAGAATTAATGGCTGCAGTTGACGAATATATCCCAACTCCAGAACGTGATACTGA 623
SAUR-Mu50           -ATCTTAGAATTAATGGAAGCTGTAGATACTTACATTCCAACTCCAGAACGTGATTCTGA 623
SAUR-N315           -ATCTTAGAATTAATGGAAGCTGTAGATACTTACATTCCAACTCCAGAACGTGATTCTGA 623
SAUR-MW2            -ATCTTAGAATTAATGGAAGCTGTAGATACTTACATTCCAACTCCAGAACGTGATTCTGA 623
SAUR-MSSA476        -ATCTTAGAATTAATGGAAGCTGTAGATACTTACATTCCAACTCCAGAACGTGATTCTGA 623
SAUR-MRSA252        -ATCTTAGAATTAATGGAAGCTGTAGATACTTACATTCCAACTCCAGAACGTGATTCTGA 623
SEPI-12228          -ATCTTAGACTTAATGCAAGCAGTTGATGATTACATTCCAACTCCAGAACGTGATTCTGA 623
SEPI-RP62A          -ATCTTAGACTTAATGCAAGCAGTTGATGATTACATTCCAACTCCAGAACGTGATTCTGA 623
SMIT-DAS            -ATCATGGACTTGATGAACACTGTTGATGAGTACATCCCAGAACCAGAACGCGATACTGA 491
SSAN-DAS            -ATCATGGAATTGATGGACACTGTTGATGAGTACATCCCAGAACCAGAACGCGATACTGA 532
SMIT-NCTC12261      -GTTATGGAATTGATGAACACAGTTGATGAGTACATCCCAGAACCAGAACGTGACACTGA 632
SPNE-R6             -GTTATGGAATTGATGAACACAGTTGATGAGTATATCCCAGAACCAGAACGTGACACTGA 632
SPYO-M1GAS          -ATCATGGAATTGATGGATACTGTTGATTCATACATTCCAGAACCAGAACGCGACACTGA 632
SMUT-DAS            -ATCATGGAATTAATGCATACTGTTGATGACTACATTCCAGATCCAGAACGTGATACTGA 532
SMUT-UA159          -ATCATGGAATTAATGCATACTGTTGATGACTACATTCCAGATCCAGAACGTGATACTGA 632
HINF-RdKW20         -ATCCTTGAGTTAGCAAACCACTTAGATACTTACATCCCAGAACCAGAACGTGCGATTGA 623
MCAT-DAS            TGTTCTAGAACTACTAGACACGCTAGATTCATACATCCCAGAGCCTGAGCGTGACATCGA 480
BPER-TohamaI        GATTCTGTCGCTGGCGCAAGCGCTGGACACGTACATTCCGACGCCGGAGCGCGCGGTCGA 629
BPAR-12822          GATTCTGTCGCTGGCGCAAGCGCTGGACACGTACATTCCGACGCCGGAGCGCGCGGTCGA 629
BCEP2-DAS           GATCATGAACCTGGCCGACGCGCTGGACACGTACATCCCGACGCCGGAGCGCGCGGTGGA 513
BPSE-1710b          GATCATGAACCTGGCCGACGCGCTGGACACGTACATCCCGACGCCGGAGCGTGCGGTCGA 629
SMAL-DAS            CATCCTGAAGCTGGTCGACGCGCTGGACAGCTGGATTCCGGAGCCGGAGCGTGCGATCGA 534
PAER-PA01           CGTGCAGAAGCTGGTAGAGACCCTGGACTCCTACATTCCGGAGCCGGTTCGTGCCATCGA 632
KPNE-DAS            AATCATCGAACTGGCTGGCCACCTGGATACCTATATCCCGGAACCAGAGCGTGCGATTGA 507
MAVI-K10            CGTCGAGCAGCTGATGGAGGCCGTCGACGAGTCGATCCCGGACCCGGTCCGCGAGACGGA 626
MAVI-paratub        CGTCGAGCAGCTGATGGAGGCCGTCGACGAGTCGATCCCGGACCCGGTCCGCGAGACGGA 626
MAISpig-DAS         GGTCGAGCAGCTGATGGAGGCCGTCGACGAGTCGATCCCGGACCCCGTCCGTGAGACCGA 521
MKAN-DAS            GGTGGAGCAGCTGATGGATGCGGTCGACGAGTCGATCCCCGACCCGGTCCGTGAGACCGA 527
MTUB-H37Rv          TGTCGAGGAACTGATGAACGCGGTCGACGAGTCGATTCCGGACCCGGTCCGCGAGACCGA 626
MTUB-CDC1551        TGTCGAGGAACTGATGAACGCGGTCGACGAGTCGATTCCGGACCCGGTCCGCGAGACCGA 626
MAVI-DAS            GGTTGAGGACCTCATGGAGGCCGTCGACGAGTCGATCCCGGATCCGGTTCGCGAGACCGA 467
CIDP-NCTC13129      CATCATCGACCTCATGCAGGCTTGCGATGATTCCATCCCAGACCCAGAGCGTGAGACCGA 626
NMEN-MC58           AATCTTCGAACTGGCTGCCGCATTGGACAGCTACATCCCGACTCCCGAGCGAGCCGTGGA 623
NMEN-Z2491          AATCTTCGAATTGGCTGCTGCATTGGACAGCTACATCCCGACTCCCGAGCGTGCCGTGGA 623
NMEN-DAS            AATCTTCGAACTGGCTGCCGCATTGGACAGCTACATCCCGACTCCCGAGCGTGCCGTGGA 464
```

```
MPNE-HF2                CAAACCTTTCTTATTAGCGGTAGAAGACGTTATGACTATTACTGGTAGAGGTACTGTAGT 683
MPNE-M129               CAAACCCTTCTTGTTGGCAATCGAAGACACCATGACGATTACTGGCCGTGGTACCGTGGT 683
LPNE-Paris              CAAGCCATTTTTGTTACCGATTGAAGACGTATTTTCAATTTCTGGACGCGGAACAGTGGT 702
LPNE2-DAS               CAAGCCATTTTTGTTACCGATTGAAGACGTATTTTCAATTTCTGGACGCGGAACAGTGGT 572
LPNE-Philadelphia1      CAAGCCATTTTTGTTGCCGATTGAAGACGTATTTTCAATTTCTGGACGCGGAACAGTGGT 689
LPNE-Lens               CAAGCCATTCTTGTTGCCGATTGAAGACGTGTTTTCAATTTCTGGACGCGGAACAGTGGT 689
LLON-DAS                CAAATCGTTCTTATTACCGATTGAAGATGTATTTTCAATATCTGGCCGAGGAACAGTAGT 573
FNUC-DAS                TCAACCATTCTTGATGCCAATAGAAGACGTTTTCACTATCACAGGAAGAGGAACAGTTGT 561
CPNE-CWL029             TAAGCCTTTCTTAATGCCTATCGAAGACGTATTCTCAATCTCTGGTCGTGGTACTGTGGT 683
EFAE-DAS                CAAACCATTCATGATGCCAGTCGAAGACGTATTCTCAATCACTGGACGTGGTACTGTTGC 560
EFAE-V583               CAAACCATTCATGATGCCAGTCGAAGACGTATTCTCAATCACTGGACGTGGTACTGTTGC 683
SAUR-Mu50               CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SAUR-N315               CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SAUR-MW2                CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SAUR-MSSA476            CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SAUR-MRSA252            CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SEPI-12228              CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SEPI-RP62A              CAAACCATTCATGATGCCAGTTGAGGACGTATTCTCAATCACTGGTCGTGGTACTGTTGC 683
SMIT-DAS                CAAACCATTGCTTCTTCCAGTCGAAGACGTATTCTCAATCACTGGACGTGGTACTGTAGC 551
SSAN-DAS                CAAGCCATTGCTTCTTCCAGTCGAAGACGTATTCTCAATCACTGGTCGTGGTACAGTTGC 592
SMIT-NCTC12261          CAAACCATTGCTTCTTCCAGTCGAAGACGTATTCTCAATCACTGGTCGTGGTACAGTTGC 692
SPNE-R6                 CAAACCATTGCTTCTTCCAGTCGAGGACGTATTCTCAATCACTGGACGTGGTACAGTTGC 692
SPYO-M1GAS              CAAACCATTGCTTCTTCCAGTCGAAGACGTATTCTCAATTACAGGTCGTGGTACAGTTGC 692
SMUT-DAS                TAAGCCACTCCTTCTTCCAGTCGAAGATGTTTTCTCAATCACTGGTCGTGGTACTGTTGC 592
SMUT-UA159              CAAGCCGCTCCTTCTTCCAGTCGAAGATGTTTTCTCAATCACTGGTCGTGGTACTGTTGC 692
HINF-RdKW20             CCAACCGTTCCTTCTTCCAATCGAAGATGTGTTCTCAATCTCAGGTCGTGGTACTGTAGT 683
MCAT-DAS                TAAGTCATTCCTAATGCCAATCGAAGATGTCTTCTCAATCTCAGGTCGTGGTACTGTTGT 540
BPER-TohamaI            CGGTGCGTTCCTGATGCCGGTGGAAGACGTGTTCTCGATCTCGGGCCGTGGCACGGTGGT 689
BPAR-12822              CGGTGCGTTCCTGATGCCGGTGGAAGACGTGTTCTCGATCTCGGGCCGTGGCACGGTGGT 689
BCEP2-DAS               CGGTACGTTCCTGATGCCGGTGGAAGACGTGTTCTCGATCTCGGGCCGCGGTACGGTGGT 573
BPSE-1710b              CGGCGCGTTCCTGATGCCGGTGGAAGACGTGTTCTCGATCTCGGGCCGTGGTACGGTGGT 689
SMAL-DAS                CAAGCCGTTCCTGATGCCGGTGGAAGACGTGTTCTCGATCTCGGGCCGCGGCACCGTGGT 594
PAER-PA01               CCAGCCGTTCCTGATGCCGATCGAAGACGTGTTCTCGATCTCCGGCCGCGGTACCGTGGT 692
KPNE-DAS                CAAGCCGTTCCTGCTGCCGATCGAAGACGTATTCTCCATCTCCGGTCGTGGTACCGTTGT 567
MAVI-K10                CAAGCCGTTCCTGATGCCGGTGGAGGACGTCTTCACCATCACCGGTCGTGGCACGGTGGT 686
MAVI-paratub            CAAGCCGTTCCTGATGCCGGTGGAGGACGTCTTCACCATCACCGGTCGTGGCACGGTGGT 686
MAISpig-DAS             GAAGCCGTTCCTGATGCCGGTGGAGGACGTCTTCACGATCACCGGTCGTGGCACCGTGGT 581
MKAN-DAS                CAAGCCGTTCCTGATGCCCGTCGAGGACGTCTTCACGATCACCGGCCGCGGCACCGTGGT 587
MTUB-H37Rv              CAAGCCGTTCCTGATGCCGGTCGAGGACGTCTTCACCATTACCGGCCGCGGAACCGTGGT 686
MTUB-CDC1551            CAAGCCGTTCCTGATGCCGGTCGAGGACGTCTTCACCATTACCGGCCGCGGAACCGTGGT 686
MAVI-DAS                CAAGCCGTTCCTGATGCCCGTCGAGGACGTCTTCACCATCACCGGTCGTGGCACCGTGGT 527
CIDP-NCTC13129          CAAGCCATTCCTCATGCCTATCGAGGACATCTTCACCATCACCGGCCGCGGTACCGTTGT 686
NMEN-MC58               CAAACCGTTCCTGCTGCCTATCGAAGACGTGTTCTCCATTTCCGGCCGCGGTACAGTAGT 683
NMEN-Z2491              CAAACCTTTCTTGTTGCCTATCGAAGACGTATTCTCTATTTCCGGTCGTGGTACAGTAGT 683
NMEN-DAS                CAAACCTTTCCTGCTGCCTATCGAAGACGTGTTCTCCATTTCCGGCCGAGGTACAGTAGT 524

MPNE-HF2                AACTGGTAGAGTTGAAAGAGGTACTTTAAAATTAAACGATGAAGTTGAAATCGTTGGTAT 743
MPNE-M129               TACCGGTCGGGTTGAACGTGGTGAATTGAAAGTAGGTCAAGAAATTGAAATCGTTGGTTT 743
LPNE-Paris              AACTGGTCGTGTTGAGAGTGGAATTGTTAAAGTTGGTGAGGAAGTTGAAATTGTTGGAAT 762
LPNE2-DAS               AACTGGTCGTGTTGAGAGTGGAATTGTTAAAGTTGGTGAGGAAGTTGAAATTGTTGGAAT 632
LPNE-Philadelphia1      AACTGGTCGTGTAGAGAGTGGAATTGTTAAAGTTGGTGAGGAAGTTGAAATTGTTGGAAT 749
LPNE-Lens               AACTGGTCGTGTTGAGAGTGGAATTGTTAAAGTTGGTGAGGAAGTTGAAATTGTTGGAAT 749
LLON-DAS                AACAGGACGTATCGAAAGCGGAATTATCAAAGTTGGCGAAGAGATCGAAATTGTAGGAAT 633
FNUC-DAS                TACAGGAAGAGTTGAAAGAGGAATCATCAAAGTTGGAGAAGAAATTGAAATAGTTGGAAT 621
CPNE-CWL029             TACAGGAAGAATCGAGCGTGGAATCGTTAAAGTTTCTGATAAAGTTCAGCTCGTGGGATT 743
EFAE-DAS                TACAGGACGTGTTGAACGTGGTGAAGTTCGCGTTGGTGACGAAGTTGAAATCGTTGGTAT 620
EFAE-V583               TACAGGCCGTGTTGAACGTGGTGAAGTTCGCGTTGGTGACGAAGTTGAAATCGTTGGTAT 743
SAUR-Mu50               TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTT 743
SAUR-N315               TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTT 743
SAUR-MW2                TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTT 743
SAUR-MSSA476            TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTT 743
SAUR-MRSA252            TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTTT 743
SEPI-12228              TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTAT 743
SEPI-RP62A              TACAGGCCGTGTTGAACGTGGTCAAATCAAAGTTGGTGAAGAAGTTGAAATCATCGGTAT 743
SMIT-DAS                ATCAGGACGTATCGACCGTGGTATCGTTAAAGTCAACGACGAAATCGAAATCGTTGGTAT 611
SSAN-DAS                TTCAGGACGTATCGACCGTGGTATCGTTAAAGTCAACGACGAAATCGAAATCGTTGGTAT 652
SMIT-NCTC12261          TTCAGGACGTATCGACCGTGGTATCGTTAAAGTCAACGACGAAATCGAAATCGTTGGTAT 752
SPNE-R6                 TTCAGGACGTATCGACCGTGGTATCGTTAAAGTCAACGACGAAATCGAAATCGTTGGTAT 752
SPYO-M1GAS              TTCAGGACGTATCGACCGTGGTACTGTTCGTGTCAACGACGAAATCGAAATCGTTGGTAT 752
SMUT-DAS                TTCAGGACGTATTGATCGTGGTACTGTTAAAGTTAACGATGAAGTTGAAATCGTTGGTAT 652
SMUT-UA159              TTCAGGACGTATTGATCGTGGTACTGTTAAAGTTAACGATGAAGTTGAAATCGTTGGTAT 752
HINF-RdKW20             AACAGGTCGTGTAGAACGAGGTATTATCCGTACAGGTGATGAAGTAGAAATCGTCGGTAT 743
MCAT-DAS                AACTGGCCGTGTTGAATCAGGCATTATTAAAGTTGGTGATGAAATTGAAATCATCGGTAT 600
```

```
BPER-TohamaI          GACTGGCCGTATCGAGCGCGGCGTGGTGAAGGTTGGCGAGGAAATCGAAATCGTGGGCAT 749
BPAR-12822            GACTGGCCGTATCGAGCGCGGCGTGGTGAAGGTTGGCGAGGAAATCGAAATCGTGGGTAT 749
BCEP2-DAS             GACGGGTCGTGTGGAGCGCGGCGTGGTGAAGGTCGGTGAGGAAATCGAAATCGTCGGTAT 633
BPSE-1710b            GACGGGTCGTGTCGAGCGCGGCGTGATCAAGGTTGGCGAGGAAATCGAAATCGTCGGTAT 749
SMAL-DAS              GACCGGTCGTATCGAGCGCGGCGTGATCAAGGTTGGCGACGAAATCGAAATCGTCGGCAT 654
PAER-PA01             AACCGGTCGTGTAGAGCGCGGCATCATCAAGGTCCAGGAAGAAGTGGAAATCGTCGGCAT 752
KPNE-DAS              TACCGGTCGTGTAGAGCGCGGTATCATCAAAGTAGGTGAAGAAGTTGAAATCGTTGGTAT 627
MAVI-K10              CACCGGCCGTGTCGAGCGCGGCGTGATCAACGTGAACGAGGAAGTCGAGATCGTCGGCAT 746
MAVI-paratub          CACCGGCCGTGTCGAGCGCGGCGTGATCAACGTGAACGAGGAAGTCGAGATCGTCGGCAT 746
MAISpig-DAS           CACCGGTCGTGTCGAGCGCGGTGTGATCAACGTGAACGAGGAAGTCGAGATCGTCGGTAT 641
MKAN-DAS              CACCGGTCGTGTGGAGCGCGGCGTGGTCAACGTCAACGAGGAAGTCGAGATCGTCGGCAT 647
MTUB-H37Rv            CACCGGACGTGTGGAGCGCGGCGTGATCAACGTGAACGAGGAAGTTGAGATCGTCGGCAT 746
MTUB-CDC1551          CACCGGACGTGTGGAGCGCGGCGTGATCAACGTGAACGAGGAAGTTGAGATCGTCGGCAT 746
MAVI-DAS              GACCGGTCGCGTCGAGCGTGGCGTGATCAACGTCAACGAAGAGGTCGAGATCGTCGGCAT 587
NCTC13129             TACCGGCCGTGTTGAGCGTGGCTCCCTGAAGGTCAACGAGGACGTCGAGATCATCGGTAT 746
NMEN-MC58             AACCGGCCGTGTAGAGCGCGGTATCATCCACGTTGGTGACGAGATTGAAATCGTCGGTCT 743
NMEN-Z2491            AACCGGTCGTGTAGAGCGCGGTATCATCCACGTCGGTGACGAGATCGAAATCGTCGGTCT 743
NMEN-DAS              AACCGGCCGTGTAGAGCGCGGTATCATCCACGTTGGTGACGAGATTGAAATCGTCGGTCT 584

MPNE-HF2              ---CCATGATACTAGAAAAGCAGTTGTTACTGGTATGGAAATGTTAAGAAAAACATTAGA 800
MPNE-M129             ---ACGTCCAATCCGTAAAGCAGTTGTTACCGGAATCGAAATGTTCAAAAAGGAACTTGA 800
LPNE-Paris            ---AAGAGACACTCAAAAGACGACTTGTACGGGTGTTGAGATGTTCCGTAAATTACTTGA 819
LPNE2-DAS             ---AAGAGACACCCAAAAGACGACTTGTACGGGTGTTGAGATGTTCCGTAAATTACTTGA 689
LPNE-Philadelphia1    ---AAGAGACACCCAAAAGACGACTTGTACGGGTGTTGAGATGTTCCGTAAATTACTTGA 806
LPNE-Lens             ---AAGAGACACCCAAAAGACGACTTGTACTGGTGTTGAGATGTTCCGTAAATTACTTGA 806
LLON-DAS              ---TCGTGATACTGCAAAGACTACCTGTACTGGTGTTGAGATGTTCCGTAAGTTATTAGA 690
FNUC-DAS              ---TAAACCTACAACTAAAACAACTTGTACAGGGGTTGAAATGTTTAGAAAACTTCTTGA 678
CPNE-CWL029           ---AGGAGAGACTAAAGAAACAATCGTTACTGGAGTCGAAATGTTCAGGAAAGAACTTCC 800
EFAE-DAS              TAAAGACGAAACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 680
EFAE-V583             TAAAGACGAAACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 803
SAUR-Mu50             ---ACATGACACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 800
SAUR-N315             ---ACATGACACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 800
SAUR-MW2              ---ACATGACACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 800
SAUR-MSSA476          ---ACATGACACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 800
SAUR-MRSA252          ---ACATGACACATCTAAAACAACTGTTACAGGTGTTGAAATGTTCCGTAAATTATTAGA 800
SEPI-12228            ---GCACGAAACTTCTAAAACAACTGTTACTGGTGTAGAAATGTTCCGTAAATTATTAGA 800
SEPI-RP62A            ---GCACGAAACTTCTAAAACAACTGTTACTGGTGTAGAAATGTTCCGTAAATTATTAGA 800
SMIT-DAS              TAAAGAAGAAATCCAAAAAGCAGTTGTTACTGGTGTTGAAATGTTCCGTAAACAACTTGA 671
SSAN-DAS              CAAAGAAGAAATCCAAAAAGCAGTTGTTACTGGTGTTGAAATGTTCCGTAAACAGCTTGA 712
SMIT-NCTC12261        CAAAGAAGAAACTCAAAAAGCAGTTGTTACTGGTGTTGAAATGTTCCGTAAACAACTTGA 812
SPNE-R6               CAAAGAAGAAACTCAAAAAGCAGTTGTTACTGGTGTTGAAATGTTCCGTAAACAACTTGA 812
SPYO-M1GAS            CAAAGAAGAAACTAAAAAAGCTGTTGTTACTGGTGTTGAAATGTTCCGTAAACAACTTGA 812
SMUT-DAS              CCGTGATGACATTCAAAAAGCTGTTGTTACTGGTGTTGAAATGTTCCGTAAACAATTGGA 712
SMUT-UA159            CCGTGATGACATTCAAAAAGCTGTTGTTACTGGTGTTGAAATGTTCCGTAAACAATTGGA 812
HINF-RdKW20           ---CAAAGATACAGCGAAAACTACTGTAACGGGTGTTGAAATGTTCCGTAAATTACTTGA 800
MCAT-DAS              ---CAAACCAACTACTAAAACCACTTGTACTGGTGTTGAAATGTTCCGTAAGCTGCTAGA 657
BPER-TohamaI          CAAGCCGACGG---TGAAGACGACCTGCACGGGCGTGGAGATGTTCCGCAAGCTGCTGGA 806
BPAR-12822            CAAGCCGACGG---TGAAGACGACCTGCACGGGCGTGGAGATGTTCCGCAAGCTGCTGGA 806
BCEP2-DAS             CAAGCCGACGG---TGAAGACGACCTGCACGGGCGTTGAAATGTTCCGCAAGCTGCTGGA 690
BPSE-1710b            CAAGGCGACGG---CGAAGACGACCTGCACGGGCGTGGAAATGTTCCGCAAGCTGCTGGA 806
SMAL-DAS              CCGTCCGGTGC---AGAAGACCACCGTGACCGGCGTTGAAATGTTCCGCAAGCTGCTGGA 711
PAER-PA01             CAAGGCGACCA---CCAAGACTACCTGCACCGGCGTTGAAATGTTCCGCAAGCTGCTCGA 809
KPNE-DAS              CAAAGAAACCG---CGAAAACCACCTGTACTGGCGTTGAAATGTTCCGCAAACTGCTGGA 684
MAVI-K10              CCGCCCGTCCAGCACCAAGACCACGGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 806
MAVI-paratub          CCGCCCGTCCAGCACCAAGACCACGGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 806
MAISpig-DAS           CAAGCCGACCAGCACCAAGACCACCGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 701
MKAN-DAS              CCGTCCGACCACCACCAAGACCACGGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 707
MTUB-H37Rv            TCGCCCATCGACCACCAAGACCACCGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 806
MTUB-CDC1551          TCGCCCATCGACCACCAAGACCACCGTCACCGGTGTGGAGATGTTCCGCAAGCTGCTCGA 806
MAVI-DAS              CCGCCCGACCACGACCAAGACCACCGTCACCGGTGTGGAAATGTTCCGCAAGCTGCTCGA 647
CIDP-NCTC13129        CCGCGAGAAGGCTACCACCACCACCGTTACCGGTATCGAGATGTTCCGTAAGCTTCTCGA 806
NMEN-MC58             GAAAGAAACCC---AAAAAACCACTTGTACCGGTGTTGAAATGTTCCGCAAACTGCTGGA 800
NMEN-Z2491            GAAAGAAACTC---AAAAAACCACTTGTACCGGTGTTGAAATGTTCCGCAAACTGCTGGA 800
NMEN-DAS              GAAAGAAACCC---AAAAAACCACCTGTACCGGTGTTGAAATGTTCCGCAAACTGCTGGA 641

MPNE-HF2              CGAAGTAAAAGCTGGGGATAACGCTGGTATCTTATTAAGAGGTATTGATAGAAAAGATGT 860
MPNE-M129             TTCAGCAATGGCTGGGGACAACGCTGGGGTATTACTCCGTGGTGTGGACCGTAAAGAAGT 860
LPNE-Paris            TGAAGGTCGAGCTGGTGATAACGTTGGAGTGTTATTACGAGGTACGAAGCGAGATGAAGT 879
LPNE2-DAS             TGAAGGTCGAGCTGGTGATAACGTTGGTGTGTTATTACGAGGTACGAAGCGAGATGAAGT 749
LPNE-Philadelphia1    TGAAGGTCGAGCTGGTGATAACGTTGGTGTGTTATTACGAGGTACGAAGCGAGATGAAGT 866
LPNE-Lens             TGAAGGTCGAGCTGGTGATAACGTTGGTGTGTTATTACGCGGTACGAAGCGAGATGAAGT 866
LLON-DAS              CGAAGGACGTGCAGGTGATAACGTTGGTATATTACTCCGTGGAACAAAACGAGACGAAGT 750
FNUC-DAS              TCAAGGTCAAGCAGGAGATAATATCGGAGTATTATTAAGAGGAACTAAGAAAGAAGAAGT 738
```

```
CPNE-CWL029     TGAAGGTCGTGCAGGAGAAAACGTTGGTTTACTCCTCAGAGGTATTGGAAAGAACGATGT 860
EFAE-DAS        CTACGCTGAAGCAGGCGACAACATCGGTGCTTTATTACGTGGTGTAGCACGTGAAGATAT 740
EFAE-V583       CTACGCTGAAGCAGGCGACAACATCGGTGCTTTATTACGTGGTGTAGCACGTGAAGATAT 863
SAUR-Mu50       CTACGCTGAAGCTGGTGACAACATTGGTGCATTATTACGTGGTGTTGCTCGTGAAGACGT 860
SAUR-N315       CTACGCTGAAGCTGGTGACAACATTGGTGCATTATTACGTGGTGTTGCTCGTGAAGACGT 860
SAUR-MW2        CTACGCTGAAGCTGGTGACAACATTGGTGCATTATTACGTGGTGTTGCTCGTGAAGACGT 860
SAUR-MSSA476    CTACGCTGAAGCTGGTGACAACATTGGTGCATTATTACGTGGTGTTGCTCGTGAAGACGT 860
SAUR-MRSA252    CTACGCTGAAGCTGGTGACAACATTGGTGCATTATTACGTGGTGTTGCTCGTGAAGACGT 860
SEPI-12228      CTACGCTGAAGCTGGTGACAACATCGGTGCTTTATTACGTGGTGTTGCTCGTGAAGACGT 860
SEPI-RP62A      CTACGCTGAAGCTGGTGACAACATCGGTGCTTTATTACGTGGTGTTGCACGTGAAGACGT 860
SMIT-DAS        CGAATGTCTTGCAGGGGACAACGTTGGTGTGCTTCTTCGTGGTATCCAACGTGATGAAAT 731
SSAN-DAS        CGAAGGTCTTGCAGGGGACAACGTAGGTGTGCTTCTCCGTGGTATCCAACGTGATGAAAT 772
SMIT-NCTC12261  CGAAGGTCTTGCCGGAGATAATGTAGGTGTCCTTCTTCGTGGTGTTCAACGTGATGAAAT 872
SPNE-R6         CGAAGGTCTTGCTGGAGATAACGTAGGTGTCCTTCTTCGTGGTGTTCAACGTGATGAAAT 872
SPYO-M1GAS      CGAAGGTCTTGCAGGAGACAACGTAGGTATCCTTCTTCGTGGTGTTCAACGTGACGAAAT 872
SMUT-DAS        TGAAGGTATTGCAGGGGATAATGTTGGTGTTCTCCTTCGTGGTATCCAACGTGATGAAAT 772
SMUT-UA159      TGAAGGTATTGCAGGGGATAATGTTGGTGTTCTCCTTCGTGGTATCCAACGTGATGAAAT 872
HINF-RdKW20     CGAAGGTCGTGCAGGTGAAAACATCGGTGCATTATTACGTGGTACCAAACGTGAAGAAAT 860
MCAT-DAS        CGAAGGTCGTGCAGGTGAGAACTGTGGTATCCTACTACGTGGTACTAAGCGTGAAGAAGT 717
BPER-TohamaI    CCAGGGCCAGGCGGGCGACAACGTGGGTATCTTGCTGCGCGGCACCAAGCGTGAAGACGT 866
BPAR-12822      CCAGGGCCAGGCGGGCGACAACGTGGGTATCTTGCTGCGCGGCACCAAGCGTGAAGACGT 866
BCEP2-DAS       CCAGGGTCAGGCAGGCGACAACGTCGGTATCCTGCTGCGCGGCACGAAGCGTGAAGACGT 750
BPSE-1710b      TCAGGGTCAGGCGGGCGACAACGTCGGTATCCTGCTGCGCGGCACGAAGCGTGAAGACGT 866
SMAL-DAS        CCAGGGTCAGGCAGGCGACAACGCTGGCCTGCTGCTGCGCGGCACCAAGCGTGACGACGT 771
PAER-PA01       CGAAGGTCGTGCTGGTGAGAACGTTGGTATCCTGCTGCGTGGCACCAAGCGTGAAGACGT 869
KPNE-DAS        CGAAGGCCGTGCTGGTGAGAACGTAGGTGTTCTGCTGCGTGGTATCAAACGTGAAGAAAT 744
MAVI-K10        CCAGGGCCAGGCCGGTGACAACGTCGGTCTGCTGCTGCGTGGTATCAAGCGTGAGGACGT 866
MAVI-paratub    CCAGGGCCAGGCCGGTGACAACGTCGGTCTGCTGCTGCGTGGTATCAAGCGTGAGGACGT 866
MAISpig-DAS     CCAGGGACAGGCCGGCGACAACGTCGGACTGTTGCTGCGTGGCATCAAGCGCGAGGACGT 761
MKAN-DAS        CCAGGGTCAGGCCGGTGACAACGTCGGGCTGTTGCTGCGTGGTGTCAAGCGTGAGGACGT 767
MTUB-H37Rv      CCAGGGCCAGGCGGGCGACAACGTTGGTTTGCTGCTGCGGGGCGTCAAGCGCGAGGACGT 866
MTUB-CDC1551    CCAGGGCCAGGCGGGCGACAACGTTGGTTTGCTGCTGCGGGGCGTCAAGCGCGAGGACGT 866
MAVI-DAS        CCAGGGCCAGGCCGGCGACAACGTCGGTCTGCTGGTTCGTGGCATCAAGCGCGAGGACGT 707
NCTC13129       CTACACCGAGGCTGGCGACAACTGTGGTCTGCTTCTCCGTGGCGTTAAGCGCGAAGACGT 866
NMEN-MC58       CGAAGGTCAGGCGGGCGACAACGTAGGCGTATTGCTGCGCGGTACCAAACGTGAAGACGT 860
NMEN-Z2491      CGAAGGTCAAGCAGGCGACAACGTAGGCGTATTGCTGCGCGGTACCAAACGTGAAGACGT 860
NMEN-DAS        CGAAGGTCAGGCGGGCGACAACGTACGCGTATTGCTGCGCGGTACCAAACGTGAAGACGT 701

MPNE-HF2        TGAACGTGGACAAGTATTAGCTAAACC-TGGTTCAATTAAACCTCACAAACAATTTGAAG 919
MPNE-M129       GGAACGTGGTCAAGTGTTAGCTAAACC-AGGTTCGATTAAACCGCACAAGAAATTTAAAG 919
LPNE-Paris      GGAGCGTGGACAGGTATTGGCAAAGCC-AGGTACCATCAAGCCACACACCAAGTTTGAAG 938
LPNE2-DAS       GGAGCGTGGACAGGTATTGGCTAAGCC-AGGTACCATCAAGCCACACACCAAGTTTGAAG 808
LPNE-Philadelphia1 GGAGCGTGGACAGGTATTGGCGAAGCC-AGGAACCATCAAGCCACACACCAAGTTTGAAG 925
LPNE-Lens       GGAGCGCGGACAGGTATTGGCTAAGCC-AGGAACCATCAAGCCACACACCAAGTTTGAAG 925
LLON-DAS        TGAGCGTGGTCAGGTATTAGCCAAACC-AGGTACAATTAAGCCTCACACTAAATTTGAAG 809
FNUC-DAS        TGAAAGAGGACAAGTTCTTGCTAAACC-AGGAAGTATCCACCCTCATACAAACTTTAAAG 797
CPNE-CWL029     TGAAAGAGGTATGGTGGTTTGTCAGCC-TAACAGCGTGAAGCCTCATACGAAATTTAAGT 919
EFAE-DAS        CGAACGTGGACAAGTATTAGCTAAACC-AGCTACAATCACTCCACACACAAAATTCAAAG 799
EFAE-V583       CGAACGTGGACAAGTATTAGCTAAACC-AGCTACAATCACTCCACACACAAAATTCAAAG 922
SAUR-Mu50       ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCAATTACACCACATACTGAATTTAAAG 919
SAUR-N315       ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCAATTACACCACATACTGAATTTAAAG 919
SAUR-MW2        ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCAATTACACCACATACTGAATTCAAAG 919
SAUR-MSSA476    ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCAATTACACCACATACTGAATTCAAAG 919
SAUR-MRSA252    ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCAATTACACCACATACTGAATTCAAAG 919
SEPI-12228      ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCTATTACACCACACACAAAATTCAAAG 919
SEPI-RP62A      ACAACGTGGTCAAGTATTAGCTGCTCC-TGGTTCTATTACACCACACACAAAATTCAAAG 919
SMIT-DAS        CGAACGTGGACAAGTTATCGCTAAACC-AGGTTCAATCAACCCACACACTAAATTCAAAG 790
SSAN-DAS        CGAACGTGGACAAGTTATCGCTAAACC-AGGTTCAATCAACCCACACACTAAATTCAAGG 831
SMIT-NCTC12261  CGAACGTGGACAAGTTATTGCTAAACC-AGGTTCAATCAACCCACACACTAAATTCAAAG 931
SPNE-R6         CGAACGTGGACAAGTTATCGCTAAACC-AGGTTCAATCAACCCACACACTAAATTCAAAG 931
SPYO-M1GAS      CGAACGTGGTCAAGTTATTGCTAAACC-AAGTTCAATCAACCCACACACTAAATTCAAAG 931
SMUT-DAS        CGAACGTGGTCAAGTTCTTGCTAAACC-AGGTTCAATTCACCCACATACTAAATTCAAAG 831
SMUT-UA159      CGAACGTGGTCAAGTTCTTGCTAAACC-AGGTTCAATTCACCCACATACTAAATTCAAAG 931
HINF-RdKW20     CGAACGTGGTCAAGTATTAGCGAAACC-AGGTTCAATCACACCACACACTGACTTCGAAT 919
MCAT-DAS        TCAACGTGGTCAAGTATTGGCTAAGCC-AGGTTCAATCACCCCACACACCAAGTTTGATG 776
BPER-TohamaI    CGAGCGTGGCCAGGTGCTGGCCAAGCC-GGGTTCGATCAACCCGCACACGGACTTCACGG 925
BPAR-12822      CGAGCGTGGCCAGGTGCTGGCCAAGCC-GGGTTCGATCAACCCGCACACGGACTTCACGG 925
BCEP2-DAS       GGAGCGCGGTCAGGTTCTGGCGAAGCC-GGGTTCGATCACGCCGCACACGCACTTCACGG 809
BPSE-1710b      GGAGCGCGGCCAGGTTCTGGCGAAGCC-GGGTTCGATCACGCCGCACACGCACTTCACGG 925
SMAL-DAS        CGAGCGTGGCCAGGTGCTGGCCAAGCC-GGGCACGATCAAGCCGCACACCAAGTTCGAAG 830
PAER-PA01       AGAGCGTGGCCAGGTACTGGCCAAGCC-GGGCACCATCAAGCCGCACACCAAGTTCGAGT 928
KPNE-DAS        CGAACGTGGTCAGGTACTGGCTAAGCC-GGGCACCATCAACCCGCACACCAAGTTCGAAT 803
MAVI-K10        CGAGCGCGGTCAGGTCGTCACCAAGCC-CGGCACCACCACGCCGCACACCGAGTTCGAGG 925
```

```
MAVI-paratub        CGAGCGCGGTCAGGTCGTCACCAAGCC-CGGCACCACCACGCCGCACACCGAGTTCGAGG 925
MAISpig-DAS         CGAGCGCGGTCAGGTCGTCGTGAAGCC-CGGCACCACCACGCCGCACACCGAGTTCGAGG 820
MKAN-DAS            CGAGCGCGGCCAGGTCGTCATCAAGCC-CGGCACCACAACCCCGCACACCGAGTTCGAGG 826
MTUB-H37Rv          CGAGCGTGGCCAGGTTGTCACCAAGCC-CGGCACCACCACGCCGCACACCGAGTTCGAAG 925
MTUB-CDC1551        CGAGCGTGGCCAGGTTGTCACCAAGCC-CGGCACCACCACGCCGCACACCGAGTTCGAAG 925
MAVI-DAS            CGAGCGTGGCCAGGTTGTGGTCAAGCC-CGGCACCACCACCCCGCACACCGAGTTCGAGG 766
NCTC13129           TGAGCGTGGCCAGGTTGTTGTTAAGCC-AGGCGCTTACACCCCTCACACCGAGTTCGAGG 925
NMEN-MC58           GGAACGCGGTCAGGTATTGGCTAAACC-GGGTACTATCACTCCTCACACCAAATTCAAAG 919
NMEN-Z2491          AGAGCGTGGTCAAGTATTGGCTAAACC-GGGTACAATCACTCCTCACACCAAGTTCAAAG 919
NMEN-DAS            GGAACGCGGTCAGGTATTGGCCAAACCCGGGTACTATCACTCCTCACACCAAGTTCAAAG 761

MPNE-HF2            CAGAAATCTATGCTCTTAAAAAAGAAGAAGGTGGAAGACATACTCCAGTATTAAATGGAT 979
MPNE-M129           CGGAAATCTATGCTTTAAAGAAGGAAGAAGGTGGTCGTCACACCGGTTTCTTAAACGGTT 979
LPNE-Paris          CAGAAGTGTATGTGTTATCCAAGGAAGAAGGCGGACGTCACACACCATTCTTTAATGGAT 998
LPNE2-DAS           CAGAAGTGTATGTATTATCCAAGGAAGAAGGCGGACGTCACACACCATTCTTTAATGGAT 868
LPNE-Philadelphia1  CAGAAGTGTATGTATTATCCAAGGAAGAAGGCGGACGTCACACTCCATTCTTTAATGGAT 985
LPNE-Lens           CAGAAGTGTATGTATTATCGAAGGAAGAAGGCGGACGTCACACTCCATTCTTTAATGGAT 985
LLON-DAS            CGGAAGTATACGTGTTATCTAAAGAAGAAGGTGGACGACATACCCCGTTTTTCAATGGAT 869
FNUC-DAS            GTGAAGTCTATGTATTAACTAAAGATGAAGGAGGAAGACACACTCCATTTTTCACAGGAT 857
CPNE-CWL029         CAGCTGTTTACGTTCTTCAGAAAGAAGAAGGCGGACGTCATAAGCCTTTCTTCAGCGGAT 979
EFAE-DAS            CTGAAGTATACGTATTATCAAAAGAAGAAGGCGGACGTCACACTCCATTCTTCACTAACT 859
EFAE-V583           CTGAAGTATACGTATTATCAAAAGAAGAAGGCGGACGTCACACTCCATTCTTCACTAACT 982
SAUR-Mu50           CAGAAGTATACGTATTATCAAAAGACGAAGGTGGACGTCACACTCCATTCTTCTCAAACT 979
SAUR-N315           CAGAAGTATACGTATTATCAAAAGACGAAGGTGGACGTCACACTCCATTCTTCTCAAACT 979
SAUR-MW2            CAGAAGTATACGTATTATCAAAAGACGAAGGTGGACGTCACACTCCATTCTTCTCAAACT 979
SAUR-MSSA476        CAGAAGTATACGTATTATCAAAAGACGAAGGTGGACGTCACACTCCATTCTTCTCAAACT 979
SAUR-MRSA252        CAGAAGTATACGTATTATCAAAAGACGAAGGTGGACGTCACACTCCATTCTTCTCAAACT 979
SEPI-12228          CTGAAGTATACGTATTATCTAAAGATGAAGGTGGACGTCACACTCCATTCTTCACTAACT 979
SEPI-RP62A          CTGAAGTATACGTATTATCTAAAGATGAAGGTGGACGTCACACTCCATTCTTCACTAACT 979
SMIT-DAS            GTGAAGTTTACATCCTTACTAAAGAAGAAGGTGGACGTCATACTCCATTCTTCAACAACT 850
SSAN-DAS            GTGAAGTTTATATCCTTACTAAAGAAGAAGGCGGACGTCACACTCCATTCTTCAACAACT 891
SMIT-NCTC12261      GTGAAGTTTACATCCTTACTAAAGAAGAAGGTGGACGTCACACTCCATTCTTCAACAACT 991
SPNE-R6             GTGAAGTCTACATCCTTACTAAAGAAGAAGGTGGACGTCACACTCCATTCTTCAACAACT 991
SPYO-M1GAS          GTGAAGTATATATCCTTTCTAAAGACGAAGGTGGACGTCACACTCCATTCTTCAACAACT 991
SMUT-DAS            GTGAAGTTTATATCCTTACTAAAGAAGAAGGTGGACGTCATACACCATTCTTCAATAACT 891
SMUT-UA159          GTGAAGTTTATATCCTTACTAAAGAGGAAGGTGGACGTCATACACCATTCTTCAATAACT 991
HINF-RdKW20         CAGAAGTGTACGTATTATCAAAAGATGAAGGTGGTCGTCATACTCCATTCTTCAAAGGTT 979
MCAT-DAS            CAGAAGTATACGTGCTATCAAAAGAAGAAGGTGGTCGTCATACCCCATTCCTAAATGGCT 836
BPER-TohamaI        CCGAGGTGTACATTCTGTCCAAGGAAGAGGGTGGCCGTCACACGCCGTTCTTCAACGGCT 985
BPAR-12822          CCGAGGTGTACATTCTGTCCAAGGAAGAGGGCGGCCGTCACACGCCGTTCTTCAACGGCT 985
BCEP2-DAS           CCGAAGTGTACGTGCTGAGCAAGGACGAAGGCGGCCGTCACACGCCGTTCTTCAACAACT 869
BPSE-1710b          CTGAAGTGTACGTGCTGAGCAAGGACGAAGGCGGCCGCCACACGCCGTTCTTCAACAACT 985
SMAL-DAS            GCGAAGTGTACGTCCTGTCGAAGGACGAAGGCGGCCGCCACACCCCGTTCTTCAACGGCT 890
PAER-PA01           GCGAAGTGTACGTGCTGTCCAAGGAAGAAGGTGGTCGTCACACCCCGTTCTTCAAGGGCT 988
KPNE-DAS            CTGAAGTGTACATCCTGTCCAAAGACGAAGGCGGCCGTCATACTCCGTTCTTCAAAGGCT 863
MAVI-K10            GCCAGGTCTACATCCTGTCCAAGGACGAGGGCGGCCGGCACACGCCGTTCTTCAACAACT 985
MAVI-paratub        GCCAGGTCTACATCCTGTCCAAGGACGAGGGCGGCCGGCACACGCCGTTCTTCAACAACT 985
MAISpig-DAS         GCAGCGTCTACATCCTGTCCAAGGACGAGGGCGGCCGGCACACGCCGTTCTTCAACAACT 880
MKAN-DAS            GCCAGGTTTACATCCTGTCCAAGGACGAGGGTGGCCGGCACACGCCGTTCTTCAACAACT 886
MTUB-H37Rv          GCCAGGTCTACATCCTGTCCAAGGACGAGGGCGGCCGGCACACGCCGTTCTTCAACAACT 985
MTUB-CDC1551        GCCAGGTCTACATCCTGTCCAAGGACGAGGGCGGCCGGCACACGCCGTTCTTCAACAACT 985
MAVI-DAS            GCAGCGTCTACATCCTGTCCAAGGACGAGGGCGGCCGCCACACGCCGTTCTTCAACAACT 826
NCTC13129           GCTCTGTCTACGTTCTGTCCAAGGACGAGGGTGGCCGCCACACCCCATTCTTCGACAACT 985
NMEN-MC58           CAGAAGTATACGTACTGAGCAAAGAAGAGGGTGGTCGTCACACTCCGTTCTTCGCCAACT 979
NMEN-Z2491          CAGAAGTATACGTACTGAGCAAAGAAGAGGGCGGCCGCCATACCCCGTTCTTCGCCAACT 979
NMEN-DAS            CAGACGTGTACGTACTGAGCAAAGAAGAGGGCGGCCGCCATACTCCGTTCTTCGCCAACT 821

MPNE-HF2            ATAGACCACAATTCTACTTCAGAACTACTGATGTTACTGGACAAATCACACTTGATAAAG 1039
MPNE-M129           ACCGTCCCCAATTCTACTTCCGTACTACAGACGTTACTGGTTCGATTTCCCTACCAGAAA 1039
LPNE-Paris          ACCGTCCACAATTCTATTTCAGAACCACTGACGTGACAGGTACTTGTGACTTGCCATCAG 1058
LPNE2-DAS           ACCGTCCACAATTCTATTTCAGAACCACTGACGTAACA---------------------- 906
LPNE-Philadelphia1  ACCGTCCACAATTCTATTTCAGAACCACTGACGTGACAGGTACTTGTGACTTGCCATCAG 1045
LPNE-Lens           ACCGTCCACAATTCTATTTCAGAACCACTGACGTGACAGGTACTTGTGACTTGCCATCAG 1045
LLON-DAS            ATAGACCACAGTTTTATTTTAGAACAACAGACGTAACAGGAACTTGTGATTTACCATCTG 929
FNUC-DAS            ACAGACCTCAATTCTATTTTAGAACTACTGATATCACTGGTGCAGTAACTTTACCAGATG 917
CPNE-CWL029         ACAGACCTCAGTTCTTCTTCCGTACTACAGACGTGACAGGAGTCGTAACTCTTCCTGAAG 1039
EFAE-DAS            ACCGTCCTCAATTCTACTTCCGTACAACAGACGTTACTGGTGTTGTAGAATTGCCAGAAG 919
EFAE-V583           ACCGTCCTCAATTCTACTTCCGTACAACAGACGTTACTGGTGTTGTAGAATTGCCAGAAG 1042
SAUR-Mu50           ATCGTCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTTCACTTACCAGAAG 1039
SAUR-N315           ATCGTCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTTCACTTACCAGAAG 1039
SAUR-MW2            ATCGTCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTTCACTTACCAGAAG 1039
SAUR-MSSA476        ATCGTCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTTCACTTACCAGAAG 1039
SAUR-MRSA252        ATCGTCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTTCACTTACCAGAAG 1039
```

```
SEPI-12228          ATCGCCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTAAACTTACCAGAAG 1039
SEPI-RP62A          ATCGCCCACAATTCTATTTCCGTACTACTGACGTAACTGGTGTTGTAAACTTACCAGAAG 1039
SMIT-DAS            ACCGTCCACAGTTCTACTTCCGTACAACTGACGTAACTGGATCTATCGAACTTCCAGCTG 910
SSAN-DAS            ACCGTCCACAGTTCTACTTCCGTACAACTGACGTTACAGGTTCAATCGAACTTCCAGCAG 951
SMIT-NCTC12261      ACCGTCCACAATTCTACTTCCGTACTACTGACGTTACAGGTTCAATCGAACTTCCAGCAG 1051
SPNE-R6             ACCGTCCACAATTCTACTTCCGTACTACTGACGTTACAGGTTCAATCGAACTTCCAGCAG 1051
SPYO-M1GAS          ACCGTCCACAATTCTACTTCCGTACAACTGACGTAACAGGTTCAATCGAACTTCCAGCAG 1051
SMUT-DAS            ATCGTCCACAATTCTACTTCCGTACAACTGACGTAACTGGTTCAATTGAGTTGCCAGCAG 951
SMUT-UA159          ATCGTCCACAATTCTACTTCCGTACAACTGACGTAACTGGTTCAATTGAGTTGCCAGCAG 1051
HINF-RdKW20         ACCGTCCACAATTCTATTTCCGTACAACAGACGTGACTGGTACAATCGAATTACCAGAAG 1039
MCAT-DAS            ATCGTCCACAGTTCTACTTCCGTACCACAGACGTAACTGGTGCCATCACCCTACAAGAAG 896
BPER-TohamaI        ATCGTCCGCAGTTCTACTTCCGCACGACGGACGTGACCGGCACGATCGACCTGCCGGCGG 1045
BPAR-12822          ATCGTCCGCAGTTCTACTTCCGCACGACGGACGTGACCGGCACGATCGACCTGCCGGCGG 1045
BCEP2-DAS           ACCGTCCGCAGTT---------------------------------------------- 882
BPSE-1710b          ACCGTCCGCAGTTCTACTTCCGTACGACGGACGTGACGGGCTCGATCGAGCTGCCGAAGG 1045
SMAL-DAS            ACCGTCCGCAGTTCTACTTCCGCACCACCGACATCACCGGCGCAGCTGCGCTGCCGGAAG 950
PAER-PA01           ACCGTCCGCAGTTCTACTTCCGTACCACCGATGTGACCGGTAACTGCGAACTGCCGGAAG 1048
KPNE-DAS            ACCGTCCGCAGTTCTACTTCCGTACTACTGACGTGACTGGCACCATCGAACTGCCGGAAG 923
MAVI-K10            ACCGCCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACGCTGCCGGAGG 1045
MAVI-paratub        ACCGCCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACGCTGCCGGAGG 1045
MAISpig-DAS         ACCGTCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACGCTGCCGGAGG 940
MKAN-DAS            ACCGTCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACGCTGCCGGAGG 946
MTUB-H37Rv          ACCGTCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACACTGCCGGAGG 1045
MTUB-CDC1551        ACCGTCCGCAGTTCTACTTCCGCACCACCGACGTGACCGGTGTGGTGACACTGCCGGAGG 1045
MAVI-DAS            ACCGCCCGCAGTTCTACTTCCGTACCACGGACGTGACCGGCGTGGTGACCCTCCCCGAG- 885
NCTC13129           ACCGCCCACAGTTCTACTTCCGCACCACCGACGTTACCGGTGTTGTGAAGCTTCCTGAGG 1045
NMEN-MC58           ACCGTCCGCAATTCTACTTCCGTACCACCGACGTAACCGGCGCGGTTACTTTGGAAGAAG 1039
NMEN-Z2491          ACCGTCCCCAATTCTACTTCCGTACCACCGACGTAACCGGCGCGGTTACTTTGGAAGAAG 1039
NMEN-DAS            ACCGTCCGCA-TTCTACTTCCGTACCACCGACGTAACCGGCGCGGTTACTTTGGAAGAAG 880

MPNE-HF2            GTGTTGAAATGATTAACCCAGGAGATAACACTAAGATTACTGTTGAACTTATTTCTCCAA 1099
MPNE-M129           ACACCGAAATGGTGCTACCAGGTGACAATACCTCGATTACAGTTGAACTAATTGCACCAA 1099
LPNE-Paris          GAGTTGAAATGGTAATGCCTGGAGATAATGTGCAATTAGTTGTTAGCTTGCATGCTCCGA 1118
LPNE2-DAS           ------------------------------------------------------------
LPNE-Philadelphia1  GAGTTGAAATGGTAATGCCTGGAGATAATGTGCAATTAGTTGTTAGCTTGCATGCTCCGA 1105
LPNE-Lens           GAGTTGAAATGGTAATGCCTGGAGATAATGTGCAATTAGTTGTTAGCTTGCATGCTCCGA 1105
LLON-DAS            GAGTTGAAATGGTAATGCCTGGAGATAACGTACAGTTG-------------------- 967
FNUC-DAS            GAGTAGAAATGGTTATGCCAG------------------------------------- 938
CPNE-CWL029         GAACTGAAATGGTAATGCCTGGAGATAACGTTGAGCTTGATGTTGAGCTCATTGGAACAG 1099
EFAE-DAS            GTACTGAAATGGTAATGCCTGGTGATAACGT--------------------------- 950
EFAE-V583           GTACTGAAATGGTAATGCCTGGTGATAACGTTGCTATGGACGTTGAATTAATTCACCCAA 1102
SAUR-Mu50           GTACTGAAATGGTAATGCCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCCAA 1099
SAUR-N315           GTACTGAAATGGTAATGCCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCCAA 1099
SAUR-MW2            GTACTGAAATGGTAATGCCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCCAA 1099
SAUR-MSSA476        GTACTGAAATGGTAATGCCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCCAA 1099
SAUR-MRSA252        GTACTGAAATGGTAATGCCTGGTGATAACGTTGAAATGACAGTAGAATTAATCGCTCCAA 1099
SEPI-12228          GTACAGAAATGGTTATGCCTGGCGACAACGTTGAAATGACAGTTGAATTAATCGCTCCAA 1099
SEPI-RP62A          GTACAGAAATGGTTATGCCTGGCGACAACGTTGAAATGACAGTTGAATTAATCGCTCCAA 1099
SMIT-DAS            GAACTGAAATGGTAATGCCTGGTGATAACGTGACTATCGACGTT--------------- 954
SSAN-DAS            GTACTGAAATGGTAATGCCTGGTGATAACGTAACAATCGACGTTGAGT------------ 999
SMIT-NCTC12261      GTACTGAAATGGTAATGCCTGGTGATAACGTGACAATCGACGTTGAGTTGATCCACCCAA 1111
SPNE-R6             GTACTGAAATGGTAATGCCTGGTGATAACGTGACAATCGACGTTGAGTTGATTCACCCAA 1111
SPYO-M1GAS          GTACAGAAATGGTTATGCCTGGTGATAACGTGACAATCAACGTTGAGTTGATCCACCCAA 1111
SMUT-DAS            GTACTGAAATGGTTATGCCTGGTGATA--------------------------------- 978
SMUT-UA159          GTACTGAAATGGTTATGCCTGGTGATAACGTTACTATTGACGTTGAATTGATCCATCCAA 1111
HINF-RdKW20         GCGTGGAAATGGTAATGCCAGGCGATAACATCAAGATGACAGTAAGCTTAATCCACCCAA 1099
MCAT-DAS            GCACTGAAATGGTTATGCCAGGTGATAACGTTGAGATGAGCGTTGAGCTAATC-ACCCAA 955
BPER-TohamaI        ACAAGGAAATGGTGCTGCCGGGCGACAACGTGTCGATGACCGTCAAGCTGCTGGCCCCGA 1105
BPAR-12822          ACAAGGAAATGGTGCTGCCGGGCGACAACGTGTCGATGACCGTCAAGCTGCTGGCCCCGA 1105
BCEP2-DAS           ------------------------------------------------------------
BPSE-1710b          ACAAGGAAATGGTGATGCCGGGCGACAACGTGTCGATCACGGTGAAGCTGATCGCGCCGA 1105
SMAL-DAS            GCGTCGAAATGGTGATGCCGGGTGACAACGTCAAGATGGTCGTCAC-------------- 996
PAER-PA01           GCGTAGAGATGGTAATGCCGGGCGACAACATCAAGATGGTTGTCACCCTGATCGCTCCGA 1108
KPNE-DAS            GCGTAGAGATGGTAATGCCGGGCGACAACATCAAAATGG--------------------- 962
MAVI-K10            GCACCGAGATGGTGATGCCCGGTGACAACACCAACATCTCGGTGAAGCTGATCCAGCCCG 1105
MAVI-paratub        GCACCGAGATGGTGATGCCCGGTGACAACACCAACATCTCGGTGAAGCTGATCCAGCCCG 1105
MAISpig-DAS         GCACCGAAATGGTGATGCCCGGTGACAACACCAACATCTCGGTCAAGCTGATCCAGC--- 997
MKAN-DAS            GCACCGAGATGGTGATGCCCGGTGACA--------------------------------- 973
MTUB-H37Rv          GCACCGAGATGGTGATGCCCGGTGACAACACCAACATCTCGGTGAAGTTGATCCAGCCCG 1105
MTUB-CDC1551        GCACCGAGATGGTGATGCCCGGTGACAACACCAACATCTCGGTGAAGTTGATCCAGCCCG 1105
MAVI-DAS            ------------------------------------------------------------
NCTC13129           GCACCGAGATGGTCATGCCTGGCGACAACGTCGACATGTCCGTCACCCTGATCCAGCCTG 1105
NMEN-MC58           GTGTAGAAATGGTAATGCCGGGTGAAAACGTAACCATCACCGTAGAACTGATTGCGCCTA 1099
```

```
NMEN-Z2491          GTGTGGAAATGGTAATGCCGGGCGAGAACGTAACCATCACCGTAGAACTGATTGCGCCTA 1099
NMEN-DAS            GAGTGGAAATGGTAATGCCTGGTGAGAACGTAACCATCACCGTAGA-CTGATTGCGTCTA 939

MPNE-HF2            TTGCTGTTGAAGAAGGAAGTAAATTCTCAATCCG-TGAAGGTGGAAGAACAGTAGGTGCT 1158
MPNE-M129           TTGCTTGTGAAAAAGGTAGTAAGTTCTCCATCCG-TGAAGGTGGTCGAACGGTTGGTGCT 1158
LPNE-Paris          TTGCGATGGATGAAGGTTTAAGATTCGCAATTAG-AGAGGGTGGCCGAACTGTTGGCGCC 1177
LPNE2-DAS           ------------------------------------------------------------
LPNE-Philadelphia1  TTGCGATGGATGAAGGTTTAAGATTCGCAATTAG-AGAGGGTGGCCGAACTGTTGGCGCC 1164
LPNE-Lens           TTGCGATGGATGAGGGTTTAAGATTCGCAATTAG-AGAGGGTGGCCGAACTGTTGGCGCC 1164
LLON-DAS            ------------------------------------------------------------
FNUC-DAS            ------------------------------------------------------------
CPNE-CWL029         TTGCTCTTGAAGAAGGAATGAGATTTGCAATTCG-TGAAGGTGGTCGTACTATCGGCGCT 1158
EFAE-DAS            ------------------------------------------------------------
EFAE-V583           TCGCTATCGAAGACGGAACTCGTTTCTCTATTCG-TGAAGGCGGACGTACTGTAGGTTCA 1161
SAUR-Mu50           TCGCGATTGAAGACGGTACTCGTTTCTCAATCCG-CGAAGGTGGACGTACTGTAGGATCA 1158
SAUR-N315           TCGCGATTGAAGACGGTACTCGTTTCTCAATCCG-CGAAGGTGGACGTACTGTAGGATCA 1158
SAUR-MW2            TCGCGATTGAAGACGGTACTCGTTTCTCAATCCG-TGAAGGTGGACGTACTGTAGGATCA 1158
SAUR-MSSA476        TCGCGATTGAAGACGGTACTCGTTTCTCAATCCG-TGAAGGTGGACGTACTGTAGGATCA 1158
SAUR-MRSA252        TCGCGATTGAAGACGGTACTCGTTTCTCAATCCG-TGAAGGTGGACGTACTGTAGGATCA 1158
SEPI-12228          TCGCTATCGAAGACGGAACTCGTTTCTCAATTCG-TGAAGGTGGACGTACTGTTGGATCA 1158
SEPI-RP62A          TCGCTATCGAAGACGGAACTCGTTTCTCAATTCG-TGAAGGTGGACGTACTGTTGGATCA 1158
SMIT-DAS            ------------------------------------------------------------
SSAN-DAS            ------------------------------------------------------------
SMIT-NCTC12261      TCGCCGTAGAACAAGGTACTACATTCTCTATCCG-TGAGGGTGGACGTACTGTTGGTTCA 1170
SPNE-R6             TCGCCGTAGAACAAGGTACTACATTCTCTATCCG-TGAGGGTGGACGTACTGTTGGTTCA 1170
SPYO-M1GAS          TCGCCGTAGAACAAGGTACTACTTTCTCAATCCG-TGAAGGTGGACGTACTGTTGGTTCA 1170
SMUT-DAS            ------------------------------------------------------------
SMUT-UA159          TCGCTGTTGAACAAGGTACTACTTTCTCTATTCG-TGAAGGTGGACGTACTGTTGGTTCT 1170
HINF-RdKW20         TTGCGATGGATCAAGGTTTACGTTTCGCAATCCG-TGAAGGTGGCCGTACAGTAGGTGCA 1158
MCAT-DAS            TCGCATGGACA---------------------------------------------- 966
BPER-TohamaI        TCGCCATGGAAGAAGGTCTGCGTTTCGCCATCCG-TGAAGGCGGTCGTACCGTCGGCGCC 1164
BPAR-12822          TCGCCATGGAAGAAGGTCTGCGTTTCGCCATCCG-TGAAGGCGGTCGTACCGTCGGCGCC 1164
BCEP2-DAS           ------------------------------------------------------------
BPSE-1710b          TCGCGATGGAAGAAGGTCTGCGCTTCGCGATCCG-CGAAGGCGGTCGCACCGTCGGCGCC 1164
SMAL-DAS            ------------------------------------------------------------
PAER-PA01           TCGCCATGGAAGATGGCCTGCGCTTCGCGATCCG-CGAAGGCGGCCGTACCGTTGGCGCC 1167
KPNE-DAS            ------------------------------------------------------------
MAVI-K10            TCGCCATGGACGAGGGTCTGCGGTTCGCCATCCG-CGAGGGTGGTCGCACCGTCGGCGCC 1164
MAVI-paratub        TCGCCATGGACGAGGGTCTGCGGTTCGCCATCCG-CGAGGGTGGTCGCACCGTCGGCGCC 1164
MAISpig-DAS         ------------------------------------------------------------
MKAN-DAS            ------------------------------------------------------------
MTUB-H37Rv          TCGCCATGGACGAAGGTCTGCGTTTCGCGATCCG-CGAGGGTGGCCGCACCGTGGGCGCC 1164
MTUB-CDC1551        TCGCCATGGACGAAGGTCTGCGTTTCGCGATCCG-CGAGGGTGGCCGCACCGTGGGCGCC 1164
MAVI-DAS            ------------------------------------------------------------
NCTC13129           TCGCTATGGATGAGGGCCTGCGCTTCGCTATCCG-CGAGGGCTCCCGCACCGTCGGCGCT 1164
NMEN-MC58           TCGCTATGGAAGAAGGCCTGCGCTTTGCGATTCG-CGAAGGCGGCCGTACCGTGGGTGCC 1158
NMEN-Z2491          TCGCTATGGAAGAAGGTTTGCGCTTTGCGATTCG-CGAAGGCGGCCGTACCGTGGGTGCC 1158
NMEN-DAS            TCGCTATGGAAGAG---CTGCGCTT-GCGTCGTGTCGCGCGTGTCCAT------------ 983

MPNE-HF2            GGTACAGTAACTAAAGTTATTAAGTAA 1185
MPNE-M129           GGTTCAGTCACGGAAGTGCTTGAATAG 1185
LPNE-Paris          GGTGTAGTCGCTAAAATAATCGAGTAA 1204
LPNE2-DAS           --------------------------
LPNE-Philadelphia1  GGTGTAGTCGCTAAAATAATCGAGTAA 1191
LPNE-Lens           GGTGTAGTCGCTAAAATAATCGAGTAA 1191
LLON-DAS            --------------------------
FNUC-DAS            --------------------------
CPNE-CWL029         GGAACGATTTCAAAGATCAATGCTTAA 1185
EFAE-DAS            --------------------------
EFAE-V583           GGCGTTGTTACTGAAATCGTTAAATAA 1188
SAUR-Mu50           GGCGTTGTTACTGAAATCATTAAATAA 1185
SAUR-N315           GGCGTTGTTACTGAAATCATTAAATAA 1185
SAUR-MW2            GGCGTTGTTACTGAAATCATTAAATAA 1185
SAUR-MSSA476        GGCGTTGTTACTGAAATCATTAAATAA 1185
SAUR-MRSA252        GGCGTTGTTACTGAAATCATTAAATAA 1185
SEPI-12228          GGCGTTGTAACTGAAATCTTTGAATAA 1185
SEPI-RP62A          GGCGTTGTAACTGAAATCTTTGAATAA 1185
SMIT-DAS            --------------------------
SSAN-DAS            --------------------------
SMIT-NCTC12261      GGTATGGTTACAGAAATCGAAGCTTAA 1197
SPNE-R6             GGTATGGTTACAGAAATCGAAGCTTAA 1197
SPYO-M1GAS          GGTATCGTTTCAGAAATCGAAGCTTAA 1197
SMUT-DAS            --------------------------
```

SMUT-UA159	GGTATCGTTTCAGAAATCGAAGCTTAA	1197
HINF-RdKW20	GGCGTTGTTGCGAAAATCATCAAATAA	1185
MCAT-DAS	----------	
BPER-TohamaI	GGCGTCGTCGCCAAGATCATCAAGTAA	1191
BPAR-12822	GGCGTCGTCGCTAAGATCATCAAGTAA	1191
BCEP2-DAS	----------	
BPSE-1710b	GGCGTCGTCGCCAAGATCATCGAGTAA	1191
SMAL-DAS	----------	
PAER-PA01	GGCGTGGTTGCCAAGATCATCGAGTAA	1194
KPNE-DAS	----------	
MAVI-K10	GGCCGGGTCGTCAAGATCATCAAGTAG	1191
MAVI-paratub	GGCCGGGTCGTCAAGATCATCAAGTAG	1191
MAISpig-DAS	----------	
MKAN-DAS	----------	
MTUB-H37Rv	GGCCGGGTCACCAAGATCATCAAGTAG	1191
MTUB-CDC1551	GGCCGGGTCACCAAGATCATCAAGTAG	1191
MAVI-DAS	----------	
NCTC13129	GGTCGCGTTACCAAGATCATCAAGTAA	1191
NMEN-MC58	GGCGTGGTTTCTTCTGTTATCGCTTAA	1185
NMEN-Z2491	GGCGTGGTTTCTTCTGTTATCGCTTAA	1185
NMEN-DAS	----------	

Abbreviations used in Table 1

MPNE	-	*Mycoplasma pneumoniae*
LPNE	-	*Legionella pneumophila*
LLON	-	*Legionella longbeachiae*
FNUC	-	*Fusobacterium nucleatum*
CPNE	-	*Chlamydophila pneumoniae*
EFAE	-	*Enterococcus faecalis*
SAUR	-	*Staphylococcus aureus*
SEPI	-	*Staphylococcus epidermidis*
SMIT	-	*Streptococcus mitis*
SSAN	-	*Streptococcus sanguis*
SPNE	-	*Streptococcus pneumoniae*
SPYO	-	*Streptococcus pyogenes*
SMUT	-	*Streptococcus mutans*
HINF	-	*Haemophilus influenzae*
MCAT	-	*Moraxella catarrhalis*
BPER	-	*Bordetella pertussis*
BPAR	-	*Bordetella parapertussis*
BCEP	-	*Burkholderia cepacia*
BPSE	-	*Burkholderia pseudomallei*
SMAL	-	*Stenotrophomonas maltophilia*
PAER	-	*Pseudomonas aeruginosa*
KPNE	-	*Klebsiella pneumoniae*
MAVI	-	*Mycobacterium avium*
MAIS	-	*Mycobacterium intracellulare*
MKAN	-	*Mycobacterium kansasii*
MTUB	-	*Mycobacterium tuberculosis*
CDIP	-	*Corynebacterium diphtheriae*
NMEN	-	*Neisseria meningitidis*

bacteria, intestinal flora, monerans, monerons, microbes, pathogens, pilus, protists, protistans, protoctists, viruses, rickettsieae, protozoans, algae spirillum and fungi.

In some embodiments, the target micro-organism is a prokaryotic micro-organism. Examples of prokaryotic micro-organisms include *Bordotella pertussis, Haemophilus influenzae, Legionella pneumophila, Mycoplasma pneumoniae, Staphylococcus aureus, Myobacterium tuberculosis, Escherichia coli, Streptococcus pyogenes, Burkholderia cepacia, Corynebacterium, Chlamydophila pneumoniae, Enterococcus faecalis, Fusobacterium nucleatum, Klebsiella pneumoniae, Legionella longbeachiae, Moraxella catarrhalis, Neisseria meningitidis, Pseudomonas aeruginosa, Staphylococcus epidermidis, Stenotrophomonas maltophilia, Streptococcus pneumoniae, Streptococcus mutans, Streptococcus sanguis, and Serratia marcescens.*

As shown in Table 2, the bacteria referred to above display a range of characteristics.

TABLE 2

Gram positive bacterial species	Gram negative bacterial species
Aerobic bacilli	**Aerobic bacilli**
Corynebacterium diphtheriae	*Bordetella pertussis*
	Chlamydophila pneumoniae
	Haemophilus influenzae
	Serratia marcescens
	Moraxella catarrhalis
	Legionella pneumophila
	Legionella longbeachiae
	Mycoplasma pneumoniae
	Klebsiella pneumoniae
	Burkholderia cepacia
	Stenotrophomonas maltophilia
	Pseudomonas aeruginosa
Aerobic cocci	**Aerobic cocci**
Staphylocuccus aureus	*Neisseria meningitidis*
Staphylococcus epidermidis	
Streptococcus pneumoniae	
Streptococcus pyogenes	
Streptpcoccus mutans	
Streptococcus sanguis	
Enterococcus faecalis	
	Anaerobic bacilli
	Fusobacterium nucleatum
	Mycobacterium tuberculosis

The methods of the invention for identifying micro-organisms may be used in a variety of fields including, but not limited to, human medicine, epidemiology, drug development, veterinary medicine, agriculture, the environment, food science, and industrial microbiology. For example, micro-organism identification may be used to determine the identity of a micro-organism found in a patient suffering from an infectious disease. Micro-organism identification may also be used to monitor food safety by testing for pathogens. Similarly, plants may be checked to determine if they harbour phytopathogenic bacteria. Further, the methods of the present invention may be used to identify drug resistant bacteria including antibiotic resistant strains of bacteria.

One method of the present invention for identifying a target microorgansim in a test sample comprises contacting the test sample or a product of the amplification process described *supra*, with a probe derived from the *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place.

The term "probe" as used herein is interchangeably used with the terms "oligonucleotide probe" or "hybridization probe" and refers to a short span of contiguous nucleotides, preferably between 10 and 50 nucleotides, that are complementary to at least a part of a *tuf* gene variable region from target microorganism, such that it is capable of specifically binding to a polynucleic acid as defined herein from the target microorganism. The design, use and labelling of probes is described in Keller *et al.*, *DNA Probes*, pp.149-213 (Stockton Press, 1989).

The terms "specific binding" or "specifically binding" refers to that interaction between a probe of the present invention and the corresponding polynucleic acid from a target microorganism.

As used herein the term "contacting" means to bring the test sample or amplification product into physical proximity with a probe derived from the *tuf* gene variable region such that, if other conditions are suitable, hybridization can take place. The contacting may be direct or indirect. For example, the test sample may directly contact the probe such that hybridisation can take place. Alternatively, the sample may hybridise with another molecule (eg another nucleic acid molecule) which in turn hybridises with the probe derived from a *tuf* gene variable region.

"Hybridisation" is used herein to denote the pairing of complementary nucleotide sequences to produce a DNA-DNA hybrid or a DNA-RNA hybrid. Complementary base sequences are those sequences that are related by the base pairing rules. In DNA, A pairs with T and C pairs with G. In RNA U pairs with A and C pairs with G. In this regard, the terms "match" and "mismatch" as used herein refer to the hybridisation potential of paired nucleotides in complementary nucleic acid strands. Matched nucleotides hybridise efficiently, such as the classical A-T and G-C pair mentioned above. Mismatches are combinations of nucleotides that do not hybridise efficiently.

Defining appropriate hybridisation conditions is within the skill of the art. See eg., Sambrook *et al.*, *supra*. However, ordinarily, "stringent conditions" for hybridisation or annealing of probes are those that

(1) employ low ionic strength and high temperature for washing, for example, 0.015M NaCl/0.0015M sodium citrate/0.1% sodium dodecyl sulfate (SDS) at 50°C, or

(2) employ during hybridisation a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1%

polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750mM NaCl, 75mM sodium citrate at 42°C.

An example of medium stringency conditions for hybridisation is the use of 50% formamide, 5 X SSC (0.75M NaCl, 0.075M sodium citrate), 50mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 X Denhardt's solution, sonicated salmon sperm DNA (50 µg/mL), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 X SSC and 0.1% SDS.

Suitable hybridisation and washing conditions depend on specific applications. The higher the degree of similarity between two micro-organisms that need to be differentiated, the higher the stringency of hybridisation and washing conditions should be. For example, when an application requires differentiating micro-organisms on the genus level, hybridisation and washing conditions of relatively low stringency may be all that is necessary. However, when differentiating between micro-organisms on the species level, wherein the micro-organism share a higher degree of similarity, hybridisation and washing conditions of higher stringency are required. Further, when differentiating between micro-organisms on the strain level wherein the micro-organisms share an even higher degree of similarity, hybridisation and washing conditions of even higher stringency are required.

In some embodiments, the step of contacting the test sample comprises hybridization with at least a portion of a *tuf* gene variable region with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12 to 57, 208 to 209 and 210.

In other embodiments, the at least one probe is selected from the group consisting of SEQ ID NOs:12 to 79, 208 to 209 and 210. In other embodiments, the at least one probe

is selected from the group consisting of SEQ ID NOs:12 to 392 and 393.

In some embodiments, the target microorganism is selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species, wherein the probe is selected from the group consisting of:

for *Haemophilus influenzae:*

TATTATCCGTACAGGTGATGAAGTAGAAA (SEQ ID NO.:12);
AAAGATACAGCGAAAACTACTGTAACG (SEQ ID NO.:13);
CATCGGTGCATTATTACGTGGTAC (SEQ ID NO.:14);
CATACTCCATTCTTCAAAGGTTACCG (SEQ ID NO.:15);
TGACTGGTACAATCGAATTACCAGAA (SEQ ID NO.:16);
CCGTAAATTACTTGACGAAGGTCG (SEQ ID NO.:17);
ACTTCGAATC AGAAGTGTAC (SEQ ID NO.:64);
CAAACGTGAA GAAATCGAAC (SEQ ID NO.:65);
CGATAACATC AAGATGACAG T (SEQ ID NO.:66);

for *Moraxella catarrhalis:*

TGAGCGTGACATCGATAAGTCA (SEQ ID NO.:18);
CAGGCATTATTAAAGTTGGTGATGAAATT (SEQ ID NO.:19);
AAACCAACTACTAAAACCACTTGTACTG (SEQ ID NO.:20);
CGTAAGCTGCTAGACGAAGGT (SEQ ID NO.:21);
TACCCCATTCCTAAATGGCTATCG (SEQ ID NO.:22);
TAACTGGTGCCATCACCCTAC (SEQ ID NO.:23);
AGTTTGATGCAGAAGTATACG (SEQ ID NO.:208)
TATCCTACTACGTGGTACTAA (SEQ ID NO.:209)
ATAACGTTGAGATGAGCG (SEQ ID NO.:210)

for *Enterococcus faecalis:*

AAGGCGACGAGTCTTATGAAGA (SEQ ID NO.:24);
AATTAATGGCTGCAGTTGACGAAT (SEQ ID NO.:25);
GAAGTTCGCGTTGGTGACG (SEQ ID NO.:26);
GACGAAACATCTAAAACAACTGTTACAG (SEQ ID NO.:27);
TGGTGTTGTAGAATTGCCAGAAG (SEQ ID NO.:28);
TAAACCAGCTACAATCACTCCACA (SEQ ID NO.:29);

for *Staphylococcus aureus:*

GTTGACATGGTTGACGATGAAGAATTA (SEQ ID NO.:30);
CTTAGAATTAATGGAAGCTGTAGATACTTACA (SEQ ID NO.:31);
ATCATCGGTTTACATGACACATCTAAAA (SEQ ID NO.:32);
CACCACATACTGAATTTAAAGCAGAAGTA (SEQ ID NO.:33);
CATTCTTCTCAAACTATCGTCCACAA (SEQ ID NO.:34);
CTGGTGTTGTTCACTTACCAGAAG (SEQ ID NO.:35);
TACTGAAATGGTAATGCCTGGTGATA (SEQ ID NO.:36);
CCAATCGCGATTGAAGACGG (SEQ ID NO.:37);
CTGGTTCAGCATTAAAAGCTTTAGAAG (SEQ ID NO.:38);
TGACAACATTGGTGCATTATTACGT (SEQ ID NO.:39);
TGTTACAGGTGTTGAAATGTTCCG (SEQ ID NO.:40);
GTATTATCAAAAGACGAAGGTGGACG (SEQ ID NO.:41);
GCGATGCTCAATACGAAGAAAAAATC (SEQ ID NO.:42);
TGAATTCAAA GCAGAAGTAT AC (SEQ ID NO.:76);
ATTATTACGT GGTGTTGCTC (SEQ ID NO.:77);
GTGATAACGT TGAAATGACA G (SEQ ID NO.:78);

for *Staphylococcus epidermidis:*

TAGACTTAATGCAAGCAGTTGATGATTAC (SEQ ID NO.:43);
CCACACACAAAATTCAAAGCTGAAG (SEQ ID NO.:44);
CTGGTGTTGTAAACTTACCAGAAGG (SEQ ID NO.:45);
GAAATGGTTATGCCTGGCGAC (SEQ ID NO.:46);
CTGGTTCTGCATTAAAAGCATTAGAAG (SEQ ID NO.:47);
TGACAACATCGGTGCTTTATTACG (SEQ ID NO.:48);

CTGTTACTGGTGTAGAAATGTTCCG (SEQ ID NO.:49);
CGTATTATCTAAAGATGAAGGTGGACG (SEQ ID NO.:50);

for *Pseudomonas aeruginosa*:

AAGTTCGAGTGCGAAGT (SEQ ID NO.:51);
AAGGCGTAGAGATGGTAAT (SEQ ID NO.:52);
TTCTTCAAGGGCTACCG (SEQ ID NO.:53);
ATCATCAAGGTCCAGGAAGAAGT (SEQ ID NO.:54);
ACCAAGACTACCTGCACCG (SEQ ID NO.:55);
TGTACGTGCTGTCCAAGGAA (SEQ ID NO.:56);
CCGGTAACTGCGAACTGC (SEQ ID NO.:57);
CACGTTGACTGCCCCGGTCACGC (SEQ ID NO.:85);
ACGCCTTCCGGCAGTTCGCAGTTAC (SEQ ID NO.:86);
ATCGGTCACGTTGACCATGGCA (SEQ ID NO.:87);
GCCGCCTTCGCGGATCGCGAA (SEQ ID NO.:88);

for *Bordetella pertussis*:

GTACATTCTG TCCAAGGAAG (SEQ ID NO.:58);
GACAACGTGG GTATCTTG (SEQ ID NO.:59);
GACAAGGAAA TGGTGCT (SEQ ID NO.:60);

for *Chlamydophila pneumoniae*:

AATTTAAGTC AGCTGTTTAC G (SEQ ID NO.:61);
GAGGTATTGG AAAGAACGAT (SEQ ID NO.:62);
ATAACGTTGA GCTTGATGTT (SEQ ID NO.:63);

for *Legionella pneumophila*:

AGTTTGAAGC AGAAGTGTAT (SEQ ID NO.:67);
TGTTATTACG AGGTACGAAG (SEQ ID NO.:68);
AGATAATGTG CAATTAGTTG TTA (SEQ ID NO.:69);

for *Mycoplasma pneumoniae*:

GAAATTTAAA GCGGAAATCT ATG (SEQ ID NO.:70);
GAACGTGGTC AAGTGTTAG (SEQ ID NO.:71);
GACAATACCT CGATTACAGT (SEQ ID NO.:72);

for *Mycobacterium tuberculosis*
AACATCTCGG TGAAGTTGAT (SEQ ID NO.:73);
TGACAACACC AACATCTC (SEQ ID NO.:74);
ACGAAGGTCT GCGTTT (SEQ ID NO.:75);

for *Streptococcus pneumoniae*

AATTCAAAGG TGAAGTCTAC A (SEQ ID NO.:79);
CAACGTGATG AAATCGAAC (SEQ ID NO.:80);
GACAATCGAC GTTGAGTT (SEQ ID NO.:81);

for *Streptococcus pyogenes*

CAAAGGTGAA GTATATATCC TTTC (SEQ ID NO.:82);
CAACGTGACG AAATCGAA (SEQ ID NO.:83);
CGTGACAATC AACGTTGA (SEQ ID NO.:84);

and combinations thereof.

Once hybridization is believed to have taken place then any hybrids formed are detected. The person skilled in the art would be readily able to detect hybridisation between the probes of the invention and the polynucleic acid of the invention. For example, where neither the polynucleic acid or probe is labelled, hybridisation can be detected using a DNA intercalating dye such as ethidium bromide, SYBR®Green (Applied Biosystems, Scoresby VIC, Australia), or SYTOX Orange (Invitrogen, Mount Waverley VIC, Australia). These dyes bind efficiently to double stranded DNA. Therefore when hybrids between the target polynucleic acid and the probe nucleotide sequence are formed, staining with a dye as described above will facilitate detection of the resulting hybrids.

In some preferred embodiments of the present invention, a polynucleic acid derived from a test sample or a probe is labelled to facilitate hybridisation detection. Methods and materials that can be used to label DNA or RNA molecules are known in the art. In the examples below, a method of labelling the DNA with Cy3 or Cy5 is described. However, other known labelling materials and methods may also be used. For example, other Cy dyes such as Cy3.5 and Cy5.5, Alexa fluorescent dyes and radioactive isotopes such as ^{32}P and ^{35}S and the like can also be used to label the whole genomic DNA to facilitate the detection of hybridisation. Furthermore, a two-colour fluorescent labelling strategy may be used in the present invention. The strategy is described in Ramsay, 1998, *Nat. Biotech.*, 16:40-44) and Shalon *et al.*, 1996, *Genome Res.*, 6:639-645, both of which are incorporated by reference in their entirety. Such multiple-colour hybridisation detection strategy minimises variations resulting from inconsistent experimental conditions and allows direct and quantitative comparison of target abundance among different samples (see Ramsay, 1998, *supra* and Shalon *et al.*, 1996, *supra*). The label is one that preferably does not provide a variable signal, but instead provides a constant and reproducible signal over a given period of time.

In some embodiments, the preferred method of contacting probes and test samples and detecting microorganisms is by using a microarray. A "microarray" is typically a 2-dimensional array upon which two or more probes of the invention have been deposited. In some embodiments there are at least 3 positions on the microarray containing a probe each having a different sequence of the invention. This has the advantage of minimising false positives.

Another method of detecting the presence of polynucleic acid from target microorganisms is specific nucleic acid

microarrays and microchip technology. A microarray is a tool for detecting gene or nucleic acid sequence presence and typically consists of a small membrane or glass slide onto which samples of many probes have been arranged in a regular pattern.

For example, oligonucleotide probes for microorganism polynucleic acid disclosed herein can be deposited at predetermined locations on a glass slide. Test samples or polynucleic acid isolated therefrom can be added to the probes under conditions which allow hybridization or binding between the probes and polynucleic acid if present in the test sample. Alternatively, polynucleic acid from a test sample may be applied to the slide before adding probes for polynucleic acid disclosed herein. Binding between the probes and polynucleic acid can be detected by any means known in the art and specific binding between the polynucleic acid and a probe indicates the polynucleic acid is present in the test sample.

The methodology of hybridisation of polynucleic acids and microarray technology is well known in the art. For example, the specific preparation of polynucleic acids for hybridisation and probes, slides, and hybridisation conditions are provided in PCT/US01/10482, herein incorporated by reference.

Microarray technology allows for the measurement of thousands of genes or polynucleic acids simultaneously thereby presenting a powerful tool for identifying the presence of microorganisms. Two microarray technologies are currently in wide use. The first are cDNA arrays and the second are oligonucleotide arrays. Although differences exist in the construction of these chips, essentially all downstream data analysis and output are the same. The product of these analyses are typically measurements of the intensity of the signal received from

a labelled probe used to detect a polynucleic acid from the test sample that hybridizes to a polynucleic acid at a known location on the microarray. Typically, the signal intensity is proportional to the polynucleic acid quantity. A large number of such techniques are available and useful. Preferred methods can be found in US Pat. Nos. 6,271,002 to Linsley *et al.*; US Pat. No. 6,218,122 to Friend *et al.*; US Pat. No. 6,218,114 to Peck *et al.*; and US Pat. No. 6,004,755 to Wang et al., the disclosure of each of which is incorporated herein by reference.

In a preferred method, biotinylated probes are prepared and hybridized to a microarray such as Affymetrix HG-U133A oligonucleotide microarrays (Affymetrix, Santa Clara, CA) by methods such as those described in Hoffmann *et al.*, 2005, *Mol Biotechnol.*, 29:31-8. Array images can then be reduced to intensity values for each probe (CEL files) using, for example, Affymetrix MAS 5.0 software. Expression measures are extracted using robust multi-array analysis (RMA) for example as described in Irizarry *et al.*, 2003, *Nucleic Acids Res.*, 31:e15 and Dallas *et al.*, 2005, *BMC Genomics.*, 6:59.

Once a microorganism(s) has been detected using the invention described herein, an appropriate diagnosis, prognosis or treatment might be effected. Accordingly, the invention further provides a method for the prognosis, diagnosis, prevention, or treatment in a subject of a disease, comprising the detection of hybridisation between a polynucleic acid in a test sample and a probe derived from a *tuf* gene variable region. Detecting hybridisation between a nucleic acid molecule in a sample and variable region of a gene enables identification of a micro-organism in the sample. This information can be used to determine whether the disease is caused by a pathogenic micro-organism, and hence the prognosis, diagnosis, prevention, or treatment for the disease.

As used herein "prognosis" means a prediction of the course and outcome of a disease or the likelihood of a subject developing a disease.

As used herein "diagnosis" means the process of identifying a disease, such as by its symptoms, laboratory tests (including genotypic tests), and physical findings.

As used herein "prevention" means any prevention of a disease in a subject and includes preventing the disease from occurring in a subject that may be predisposed to the disease, but has not yet been diagnosed as having it. The effect may be prophylactic in terms of completely or partially preventing the disease or sign or symptom thereof.

As used herein "treatment" or "treating" means any treatment of a disease in a subject by administering a medicament to the subject following the detection of a particular microorganism. "Treatment" and "treating" includes: (a) inhibiting the disease, i.e., arresting its development; or (b) relieving or ameliorating the symptoms of the disease, i.e., cause regression of the symptoms of the disease. The effect may be therapeutic in terms of a partial or complete cure of the disease.

Methods of the invention can be used in the prognosis, diagnosis, prevention, or treatment of any subject, such as an animal or plant, as all animals and plants are infected by micro-organisms and are subject to diseases and disorders caused by micro-organisms.

As used herein an "animal" means any animal, such as a human or a mammal of economical importance and/or social importance to humans, for instance, carnivores other than humans (such as cats and dogs), swine (pigs, hogs, and

wild boars), ruminants (such as cattle, oxen, sheep, giraffes, deer, goats, bison, and camels), and horses. Also provided is the treatment of birds, including the treatment of those kinds of birds that are endangered, kept in zoos, as well as fowl, and more particularly domesticated fowl, eg., poultry, such as turkeys, chickens, ducks, geese, guinea fowl, and the like, as they are also of economical importance to humans. Thus, provided is the treatment of livestock, including, but not limited to, domesticated swine (pigs and hogs), ruminants, horses, poultry, and the like. The term does not denote a particular age. Thus, both adult and newborn subjects are intended to be covered.

As used herein a "plant" includes any member of the plant kingdom including, but not limited to, eukaryotic algae, mosses, club mosses, ferns, angiosperms, gymnosperms, and lichens (which contain algae) including any parts (eg pollen, seeds, cells, tubers, stems) thereof, and any cellular components (eg plasmids, ribosomes, etc.) thereof.

"Disease" as used herein is a general term used to refer to any departure from health in which a subject suffers.

In some embodiments the disease is one associated with pathogens (including animal pathogens) and in particular infectious diseases, respiratory infections, including, pneumonia (including community acquired pneumonia), whooping cough, meningitis, typhus, tuberculosis, diptheria, typhoid fever and bacterial diarrhoea; gastrointestinal tract infections, including gastroenteritis; dermal infections; genital tract infections; urinary tract infections; nosocomial infections; typhoid fever; botulism; strep throat; rheumatic fever; syphilis, cystitis; kidney infection; folliculitis; furunculosis (boils); impetigo (school

sore); methicillin resistant *Staphylococcus aureus;* staphylococcal scalded skin syndrome; toxic shock syndrome; cellulitis, erysipelas; necrotising fascilitis; scarlet fever; wound infections ie. dermatitis and scabies.

A medical or veterinary practitioner, after being provided with the results of the above method may prescribe appropriate treatment according to his or her knowledge and skill. However, in an alternative embodiment, a computer software program may be provided to advise on the appropriate treatment of the patient, once the identification of the micro-organism has taken place.

Embodiments of the present invention may also be conveniently provided in kits. The kits of the invention may comprise reagents such as probes or primers, reference polynucleic acids, buffers, enzymes and the like. The kits of the invention may also comprise instructions for carrying out the methods of the present invention. Such instructions may take the form of package inserts, or labels, or referrals to other sources such as books, journals, websites, and compact discs.

The invention will now be further described by way of reference only to the following non-limiting examples. It should be understood, however, that the examples following are illustrative only, and should not be taken in any way as a restriction on the generality of the invention described above.

EXAMPLE 1 IDENTIFICATION OF MICROORGANISMS FROM SPUTUM SAMPLE

Sputum samples were obtained from pathology laboratories and diluted with water. Two microliters of these samples were used to amplify a specific conserved sequence found

in eubacterial species by polymerase chain reaction (PCR). The PCR product was then used as a template for direct incorporation of the fluorescent analogue Cy5 by randomly primed polymerisation reactions Dorell *et al*, 2001 (Genome Res. 11, 1706-15) or by post-PCR labelling.

Oligonucleotide probes were designed containing unique signature sequences chosen from conserved sequences in eubacterial species and printed on micro-array slides.

The labelled PCR product was then hybridised to the micro-array slide. The hybridisation took place under a glass cover slip in a humidified chamber submerged in a 53°C water bath for 30 mins. The slides were washed, dried and scanned for fluorescence by using a laser-activated, confocal scanner.

Background was subtracted for each spot on the scanned array and analysed by the software to give the species present in the clinical sample and to advise on the appropriate treatment of the patient.

Materials and methods

Slides: CodeLink™ activated slides (GE Healthcare).

Oligonucleotide probes: Oligonucleotides ranging from 17 to 24 nucleotides with a 5' amino modification and 10 T attachment at 5' end were designed. Oligonucleotides were designed containing unique signature sequences in the eubacterial species' *tuf* gene.

Spotting solution: 1M sodium phosphate buffer. Oligonucleotides dissolved in Milli-Q water at a concentration of 100 μM. Final concentration of oligonucleotides in spotting solution was 25 μM and that of sodium phosphate buffer was 250 mM.

Printing: Printing of Epoxy slides was performed using *Micro*Grid II Compact Microarrayer (*Bio*Robotics).

Target DNA: A fragment of the *tuf* gene was amplified using the primer pair MAtuf-F4 (forward primer, 5' KYACIGGHGTBGARATGTTC 3', SEQ ID NO.:396) and MAtuf-R5 (reverse primer, 5' GTDCGDCCRCCYTCWCGRAT 3', SEQ ID No.395). A second fragment of the *tuf* gene was amplified using the primers MAtuf-F1 (SEQ ID No. 394) and MAtuf-R5 (SEQ ID No. 395). The primers were designed using conserved regions found in the *tuf* gene. The PCR conditions were 1 x 94°C, 1 min initial denaturation, 35 x 94°C, 30 secs/50°C, 30 secs/72°C, 30 secs, 1 x 72°C, 7 mins final extension. PCR products were purified using the Wizard PCR Preps DNA Purification kit (Promega) prior to labelling.

Labelling: Labelling was performed using the labelling mixture 2mM dATP, dCTP, dGTP, 1.5mM dTTP and 0.5mM biotin-dUTP. Labelling was performed overnight at 37°C for the product obtained using the primers MAtuf-F4 and MAtuf-R5. Direct labelling during the PCR reaction using Cy5-dUTP (GE Healthcare) was performed for the *tuf* fragment amplified using the MAtuf-F1 and MAtuf-R5 primers.

Hybridisation: Hybridisation was performed in a hybridisation chamber at 53°C for 30 minutes. This was followed by 4 wash steps, 5 minutes each as follows: wash 1, 2 x SSC/0.1% SDS; wash 2, 0.5 x SSC/0.1% SDS; wash 3, 2 x SSC; wash 4, 4 x SSC/0.2% Tween 20. After hybridisation, the arrays were scanned using a GenePix 4000B scanner (Affymetrix).

Results

PCR amplification of the *tuf* gene fragment. The size of the amplified PCR product using the primers MAtuf-F4 and MAtuf-R5 was approximately 381 base pairs (Figure 1). The primers MAtuf-F1 and MAtuf-R5 primers produced a PCR amplification product of approximately 1.1 kb.

Scanner outputs. The microarray successfully detected and differentiated organisms tested in this study; namely (a) *Streptococcus pneumoniae*; (b) *Moraxella catarrhalis*; (c) *Legionella pneumophila*; (d) *Staphylococcus aureus*; (e) *Haemophilus influenzae* (Figures 2, 3, 4, 5 and 6)

Discussion

DNA microarrays combine high-precision technology with advanced molecular biology to acquire high-throughput screening of DNA fragments. In this study, the potential of the DNA microarray technique to identify and differentiate PCR derived amplicons from five pathogenic bacteria was investigated; namely, *Streptococcus pneumoniae*, *Moraxella catarrhalis*, *Legionella pneumophila*, *Staphylococcus aureus* and *Haemophilus influenzae*.

A fragment from the *tuf* gene was chosen for amplification. This gene is universally present in all bacterial species. Parts of this gene are highly conserved (conserved regions) whereas the rest exhibit high variation (variable regions). The conserved and variable regions were determined via multiple sequence alignment of the gene using the *tuf* DNA sequences obtained from several different bacterial species. The conserved regions were used to design primers in order to obtain amplicons containing variable regions. These amplicons were used on a DNA microarray to detect and differentiate aforementioned bacterial species.

This indicates that microarray technology can be used as a

method for identifying and discriminating bacterial species. The method has been proven to be highly sensitive and specific. This investigation demonstrates that bacterial *tuf* gene sequences amplified via PCR technique could be successfully coupled with DNA microarrays to serve the above purpose in a clinical diagnostic setup. If current demands can be met, nucleic acid microarray technology will give competitive advantages over other molecular based identification methods. The rapid development in this area will improve the technology and reduce the cost at the same time, making the microarray methodology attractive and competitive.

Sequences

Nucleotide sequences of *tuf* genes from various microorganisms are provided in SEQ ID NOs.:1 to 11. SEQ ID NOs.:12 to 393 are signature (probe) sequences from various microorganisms, in which:

ABAU - *Acinetobacter baumanii*
BCEP - *Burkholderia cepacia*
BPER - *Bordetella pertussis*
CABO - *Chlamydophila abortus*
CDIP - *Corynebacterium diphtheriae*
CJEJ - *Campylobacter jejuni*
CPER - *Clostridium perfringens*
CPNE - *Chlamydophila pneumoniae*
CTET - *Clostridium tetani*
EFAE - *Enterococcus faecalis*
FNUC - *Fusobacterium nucleatum*
HPYL - *Helicobacter pylori*
HHEP - *Helicobacter hepaticus*
HINF - *Haemophilus influenzae*
KPNE - *Klebsiella pneumoniae*
LLON - *Legionella longbeachiae*
LMON - *Listeria monocytogenes*

LPNE - *Legionella pneumophila*
MCAT - *Moraxella catarrhalis*
MTUB - *Mycobacterium tuberculosis*
MPNE - *Mycoplasma pneumoniae*
NGON - *Neisseria gonorrhoeae*
NMEN - *Neisseria meningitidis*
PAER - *Pseudomonas aeruginosa*
SAUR - *Staphylococcus aureus*
SEPI - *Staphylococcus epidermidis*
SSAP - *Staphylococcus saprophyticus*
SMAL - *Stenotrophomonas maltophilia*
SPNE - *Streptococcus pneumoniae*
SPYO - *Streptococcus pyogenes*
SMUT - *Streptococcus mutans*
SSAN - Streptococcus sanguis
SMAR - *Serratia marcescens*
VVUL - *Vibrio vulnificus*
VPAR - *Vibrio parahaemolyticus*
VCHO - *Vibrio cholerae*
YPES - *Yersinia pestis*
YPSE - *Yersinia pseudotuberculosis*

Sequences 394 and 395 are oligonucleotide primers designed against the conserved regions of *tuf* genes of various micro-organisms.

Sequences may also contain alternative bases. These are shown as follows:

R - A or G
Y - C or T
M - A or C
K - G or T
S - G or C
W - A or T
H - A or C or T
B - G or C or T

V - A or G or C

D - A or G or T

I - A base that binds to A, G, C or T

CLAIMS

1. A method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample comprising:

(a) providing a test sample suspected of containing one or more target microorganisms;

(b) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) from said test sample, if present in the test sample, the or each said polynucleic acid comprising at least a portion of a *tuf* gene variable region;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of said one or more said microorganisms in said test sample.

2. A method according to claim 1, wherein the polynucleic acid in said test sample is amplified.

3. A method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of said one or more said microorganisms in said test sample.

4. A method according to claim 3, wherein the polynucleic acid is derived from a *tuf* gene variable region as shown in SEQ ID NOs.:1 to 11.

5. A method according to claim 3, wherein the polynucleic acid consists essentially of nucleotides 108-220, 393-591, 708-774, 792-816, and/or 885-945 of SEQ ID NOs.:1 to 11.

6. A method according to claim 3, wherein polynucleic acid has a sequence which is antisense to a *tuf* gene variable region as shown in SEQ ID NOs.:1 to 11.

7. A method according to any one of claims 1 to 6, wherein the step of contacting the test sample comprises hybridization with at least a portion of a *tuf* gene variable region with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12 to 84, 208 to 251, 254-315 and 330 to 393.

8. A method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a test sample comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one

microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12 to 84, 208 to 251, 254-315 and 330 to 393;

(d) detecting hybrids formed in step (c); and

(e) identifying said at least one microorganism from said detecting of step (d).

9. A method according to claim 8, wherein the at least one probe is selected from the group consisting of SEQ ID NOs: 12 to 84, 208 to 315 and 330 to 393.

10. A method according to claim 8, wherein the at least one probe is selected from the group consisting of SEQ ID NOs:12 to 392 and 393.

11. A method according to any one of claims 1 to 10, wherein the microorganism detected and/or identified is selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species.

12. A method for the diagnosis and/prognosis of a disease associated with a target microorganism comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of the target microorganism in said test sample;

(e) associating the presence of said one or more said target microorganisms with a disease; and

(f) determine diagnosis and/or prognosis for said disease.

13. A method according to claim 12, wherein the microorganism detected and/or identified is selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species.

14. A method for the prevention or treatment of a disease associated with a target microorganism comprising:

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) contacting said test sample or said polynucleic acid from said test sample with one or more probes derived from said *tuf* gene variable region for a time and under conditions sufficient for hybridization to take place; and

(d) detecting any hybridization which has taken place in step (c);

wherein the presence or absence of hybridization is indicative of the presence or absence of the target microorganism in said test sample;

(e) associating the presence of said one or more said target microorganisms with a disease; and

(f) prescribing prevention or treatment of said disease.

15. A method according to any one of claims 3 to 14, wherein the step of amplifying at least a portion of a *tuf* gene variable region from the microorganism(s) utilizes at least one primer pair, wherein at least one of the primers is selected from the group consisting of SEQ ID NOs:394-396, or equivalents thereof, provided said equivalents specifically amplify the *tuf* variable region or a part thereof of different species of microorganism at the same time.

16. A method according to claim 14, wherein the step of associating the presence of one or more target microorganisms with a disease or the step of prescribing a treatment comprise the step of using a computer program.

17. A method according to any one of claims 12 to 16, wherein the disease is either a bacterial infection, a respiratory infection, a gastrointestinal tract infection, a dermal infection, a bloodstream infection, meningitis or a central nervous system infection, a urinary tract infection, a genital tract infection, a wound infection, or a nosocomial infection.

18. A method according to any one of claims 1 to 17, wherein the microorganism detected and/or identified is selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species, *Streptococcus* species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderi* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species.

19. A composition comprising at least one probe selected from the group consisting of SEQ ID NOs:12-293 or equivalents thereof, provided that said equivalents specifically bind to a *tuf* gene variable region of a microorganism to be detected.

20. A composition comprising at least one probe selected from the group consisting essentially of SEQ ID NOs:12-393 or equivalents thereof, provided that said equivalents specifically bind to a *tuf* gene variable region of a microorganism to be detected.

21. A composition comprising at least one primer selected from the group consisting essentially of SEQ ID NOs:394-396, or equivalents thereof, provided said equivalents specifically amplify the *tuf* gene variable region or a part thereof of a microorganism to be detected.

22. An oligonucleotide probe for the detection and/or identification of at least one microorganism found in test sample selected from the group consisting of *Bordetella* species, *Chlamydophila* species, *Haemophilus* species, *Legionella* species, *Mycoplasma* species, *Mycobacterium* species, *Staphylococcus* species,

Streptococcus species, *Pseudomonas* species, *Moraxella* species and *Fusobacterium* species, *Stenotrophomonas* species, *Burkholderia* species, *Enterococcus* species, *Corynebacterium* species and *Neisseria* species, wherein said probe is selected from the group consisting of:

for *Haemophilus influenzae:*

TATTATCCGTACAGGTGATGAAGTAGAAA (SEQ ID NO.:12);
AAAGATACAGCGAAAACTACTGTAACG (SEQ ID NO.:13);
CATCGGTGCATTATTACGTGGTAC (SEQ ID NO.:14);
CATACTCCATTCTTCAAAGGTTACCG (SEQ ID NO.:15);
TGACTGGTACAATCGAATTACCAGAA (SEQ ID NO.:16);
CCGTAAATTACTTGACGAAGGTCG (SEQ ID NO.:17);
ACTTCGAATC AGAAGTGTAC (SEQ ID NO.:64);
CAAACGTGAA GAAATCGAAC (SEQ ID NO.:65);
CGATAACATC AAGATGACAG T (SEQ ID NO.:66);

for *Moraxella catarrhalis:*

TGAGCGTGACATCGATAAGTCA (SEQ ID NO.:18);
CAGGCATTATTAAAGTTGGTGATGAAATT (SEQ ID NO.:19);
AAACCAACTACTAAAACCACTTGTACTG (SEQ ID NO.:20);
CGTAAGCTGCTAGACGAAGGT (SEQ ID NO.:21);
TACCCCATTCCTAAATGGCTATCG (SEQ ID NO.:22);
TAACTGGTGCCATCACCCTAC (SEQ ID NO.:23);
AGTTTGATGCAGAAGTATACG (SEQ ID NO.:208)
TATCCTACTACGTGGTACTAA (SEQ ID NO.:209)
ATAACGTTGAGATGAGCG (SEQ ID NO.:210)

for *Enterococcus faecalis:*

AAGGCGACGAGTCTTATGAAGA (SEQ ID NO.:24);
AATTAATGGCTGCAGTTGACGAAT (SEQ ID NO.:25);
GAAGTTCGCGTTGGTGACG (SEQ ID NO.:26);
GACGAAACATCTAAAACAACTGTTACAG (SEQ ID NO.:27);
TGGTGTTGTAGAATTGCCAGAAG (SEQ ID NO.:28);

TAAACCAGCTACAATCACTCCACA (SEQ ID NO.:29);

for *Staphylococcus aureus*:

GTTGACATGGTTGACGATGAAGAATTA (SEQ ID NO.:30);
CTTAGAATTAATGGAAGCTGTAGATACTTACA (SEQ ID NO.:31);
ATCATCGGTTTACATGACACATCTAAAA (SEQ ID NO.:32);
CACCACATACTGAATTTAAAGCAGAAGTA (SEQ ID NO.:33);
CATTCTTCTCAAACTATCGTCCACAA (SEQ ID NO.:34);
CTGGTGTTGTTCACTTACCAGAAG (SEQ ID NO.:35);
TACTGAAATGGTAATGCCTGGTGATA (SEQ ID NO.:36);
CCAATCGCGATTGAAGACGG (SEQ ID NO.:37);
CTGGTTCAGCATTAAAAGCTTTAGAAG (SEQ ID NO.:38);
TGACAACATTGGTGCATTATTACGT (SEQ ID NO.:39);
TGTTACAGGTGTTGAAATGTTCCG (SEQ ID NO.:40);
GTATTATCAAAAGACGAAGGTGGACG (SEQ ID NO.:41);
GCGATGCTCAATACGAAGAAAAAATC (SEQ ID NO.:42);
TGAATTCAAA GCAGAAGTAT AC (SEQ ID NO.:76);
ATTATTACGT GGTGTTGCTC (SEQ ID NO.:77);
GTGATAACGT TGAAATGACA G (SEQ ID NO.:78);

for *Staphylococcus epidermidis*:

TAGACTTAATGCAAGCAGTTGATGATTAC (SEQ ID NO.:43);
CCACACACAAAATTCAAAGCTGAAG (SEQ ID NO.:44);
CTGGTGTTGTAAACTTACCAGAAGG (SEQ ID NO.:45);
GAAATGGTTATGCCTGGCGAC (SEQ ID NO.:46);
CTGGTTCTGCATTAAAAGCATTAGAAG (SEQ ID NO.:47);
TGACAACATCGGTGCTTTATTACG (SEQ ID NO.:48);
CTGTTACTGGTGTAGAAATGTTCCG (SEQ ID NO.:49);
CGTATTATCTAAAGATGAAGGTGGACG (SEQ ID NO.:50);

for *Pseudomonas aeruginosa*:

AAGTTCGAGTGCGAAGT (SEQ ID NO.:51);
AAGGCGTAGAGATGGTAAT (SEQ ID NO.:52);
TTCTTCAAGGGCTACCG (SEQ ID NO.:53);

ATCATCAAGGTCCAGGAAGAAGT (SEQ ID NO.:54);
ACCAAGACTACCTGCACCG (SEQ ID NO.:55);
TGTACGTGCTGTCCAAGGAA (SEQ ID NO.:56);
CCGGTAACTGCGAACTGC (SEQ ID NO.:57);
CACGTTGACTGCCCCGGTCACGC (SEQ ID NO.:85);
ACGCCTTCCGGCAGTTCGCAGTTAC (SEQ ID NO.:86);
ATCGGTCACGTTGACCATGGCA (SEQ ID NO.:87);
GCCGCCTTCGCGGATCGCGAA (SEQ ID NO.:88);

for *Bordetella pertussis:*

GTACATTCTG TCCAAGGAAG (SEQ ID NO.:58);
GACAACGTGG GTATCTTG (SEQ ID NO.:59);
GACAAGGAAA TGGTGCT (SEQ ID NO.:60);

for *Chlamydophila pneumoniae:*

AATTTAAGTC AGCTGTTTAC G (SEQ ID NO.:61);
GAGGTATTGG AAAGAACGAT (SEQ ID NO.:62);
ATAACGTTGA GCTTGATGTT (SEQ ID NO.:63);

for *Legionella pneumophila:*

AGTTTGAAGC AGAAGTGTAT (SEQ ID NO.:67);
TGTTATTACG AGGTACGAAG (SEQ ID NO.:68);
AGATAATGTG CAATTAGTTG TTA (SEQ ID NO.:69);

for *Mycoplasma pneumoniae:*

GAAATTTAAA GCGGAAATCT ATG (SEQ ID NO.:70);
GAACGTGGTC AAGTGTTAG (SEQ ID NO.:71);
GACAATACCT CGATTACAGT (SEQ ID NO.:72);

for *Mycobacterium tuberculosis*
AACATCTCGG TGAAGTTGAT (SEQ ID NO.:73);
TGACAACACC AACATCTC (SEQ ID NO.:74);
ACGAAGGTCT GCGTTT (SEQ ID NO.:75);

for *Streptococcus pneumoniae*

AATTCAAAGG TGAAGTCTAC A (SEQ ID NO.:79);
CAACGTGATG AAATCGAAC (SEQ ID NO.:80);
GACAATCGAC GTTGAGTT (SEQ ID NO.:81);

for *Streptococcus pyogenes*

CAAAGGTGAA GTATATATCC TTTC (SEQ ID NO.:82);
CAACGTGACG AAATCGAA (SEQ ID NO.:83);
CGTGACAATC AACGTTGA (SEQ ID NO.:84);

and combinations thereof.

23. A method of hybridization comprising contacting a test sample suspected of containing a polynucleic acid with a solid support, which has immobilized thereon one or more probes selected from the group consisting of SEQ ID NOs:12-392 and 393.

24. A kit for the detection and/or identification of at least one microorganism or the detection and/or identification of several microorganisms simultaneously in a test sample, comprising

(a) optionally, at least one primer pair allowing amplification of at least a portion of a *tuf* gene variable region;

(b) the composition according to claim 21;

(c) a buffer, or components necessary to produce the buffer enabling a hybridization reaction between the probe(s) contained in said composition and a polynucleic acid present in the sample, or the amplified products thereof;

(d) a solution, or components necessary for producing the solution, enabling washing of the hybrids formed under the appropriate wash conditions;

(e) optionally, a means for detecting the hybrids resulting from the preceding hybridization.

25. A method for the detection and/or identification of at least one microorganism or for the simultaneous detection and/or identification of several microorganisms in a sputum sample comprising the steps of

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12 to 84, 208 to 251, 254-315 and 330 to 393;

(d) detecting hybrids formed in step (c);

(e) identifying said at least one microorganism from said detecting of step (d).

26. A microarray comprising at least one probe selected from the group consisting of SEQ ID NOs:12-393 or equivalents thereof, provided that said equivalents specifically bind to the *tuf* gene variable region of a microorganism.

27. A method for the detection and/or identification of at least one *Haemophilus influenzae* microorganism or for the simultaneous detection and/or identification of several *Haemophilus influenzae* microorganisms in a test sample comprising the steps of

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion

of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:12-17 and 64-66;

(d) detecting hybrids formed in step (c); and

(e) identifying said at least one microorganism from said detecting of step (d).

28. A method for the detection and/or identification of at least one *Moraxella catarrhalis* microorganism or for the simultaneous detection and/or identification of several *Moraxella catarrhalis* microorganisms in a test sample comprising the steps of

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:18-23 and 208 to 210;

(d) detecting hybrids formed in step (c); and

(e) identifying said at least one microorganism from said detecting of step (d).

29. A method for the detection and/or identification of at least one *Pseudomonas* microorganism or for the simultaneous detection and/or identification of several *Pseudomonas* microorganisms in a test sample comprising the steps of

(a) optionally, releasing, isolating and/or concentrating polynucleic acids from said at least one microorganism(s) to be detected;

(b) optionally, amplifying at least a portion of a *tuf* gene variable region from said microorganism(s) with at least one primer pair;

(c) hybridizing at least a portion of a *tuf* gene variable region from said microorganism(s) with a set of probes comprising at least one probe selected from the group consisting of SEQ ID NOs:51-57 and 85-88;

(d) detecting hybrids formed in step (c); and

(e) identifying said at least one microorganism from said detecting of step (d).

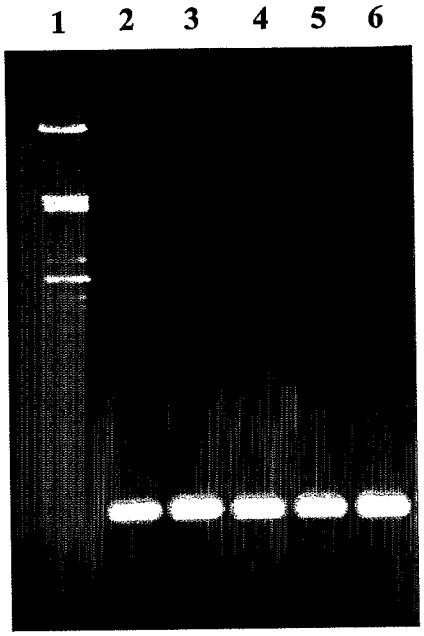


FIGURE 1

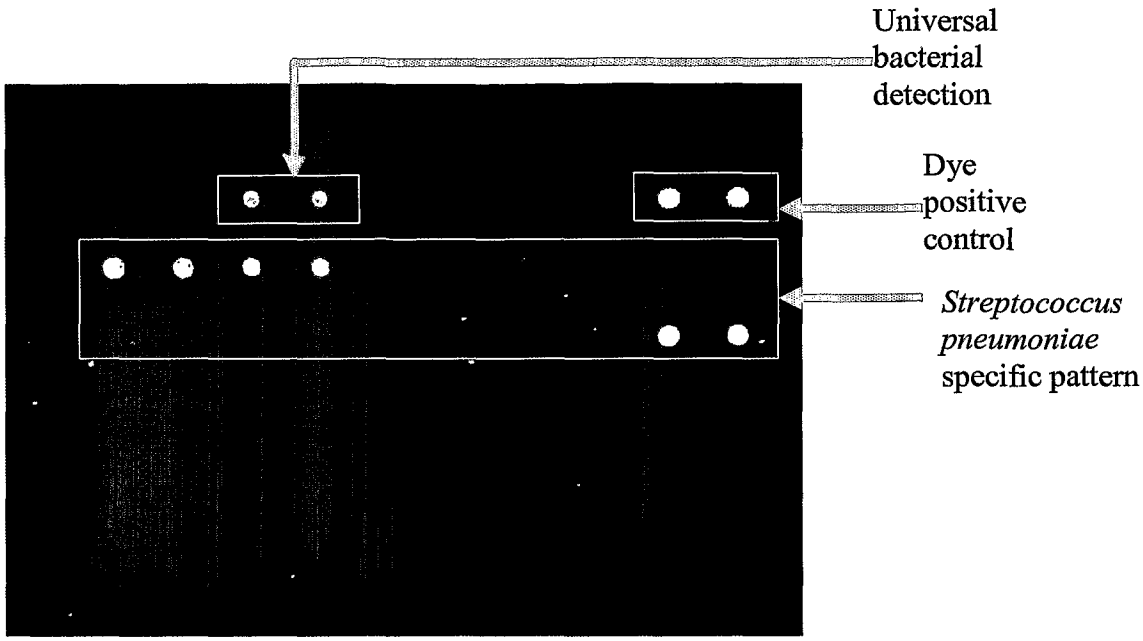

Universal bacterial detection
Dye positive control
Streptococcus pneumoniae specific pattern

FIGURE 2

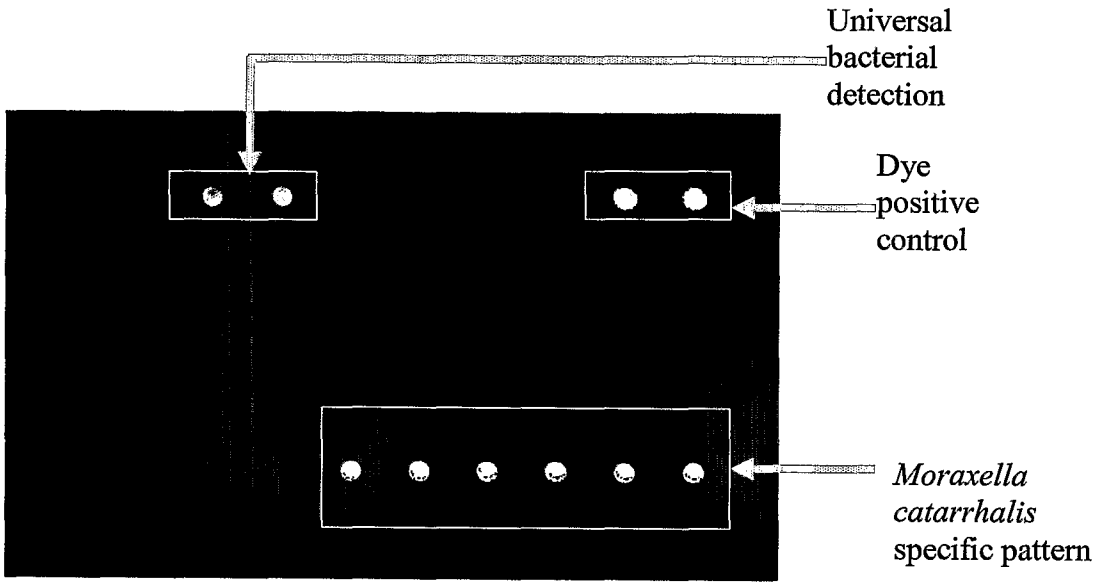

Universal bacterial detection
Dye positive control
Moraxella catarrhalis specific pattern

FIGURE 3

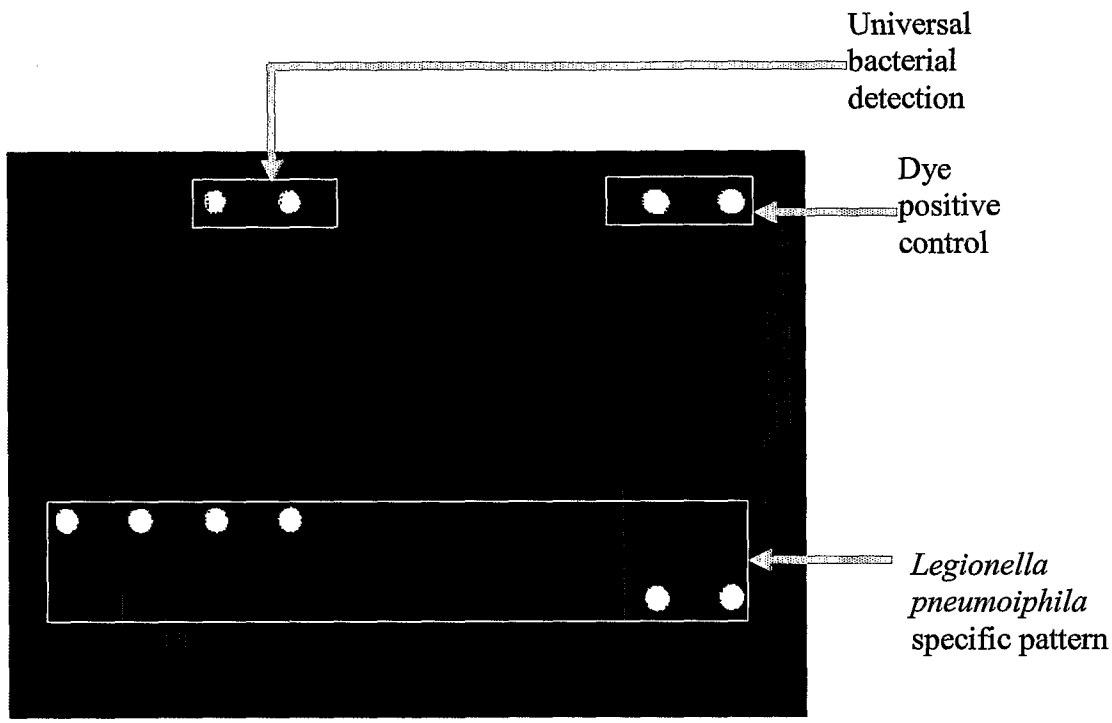

Universal bacterial detection
Dye positive control
Legionella pneumoiphila specific pattern

FIGURE 4

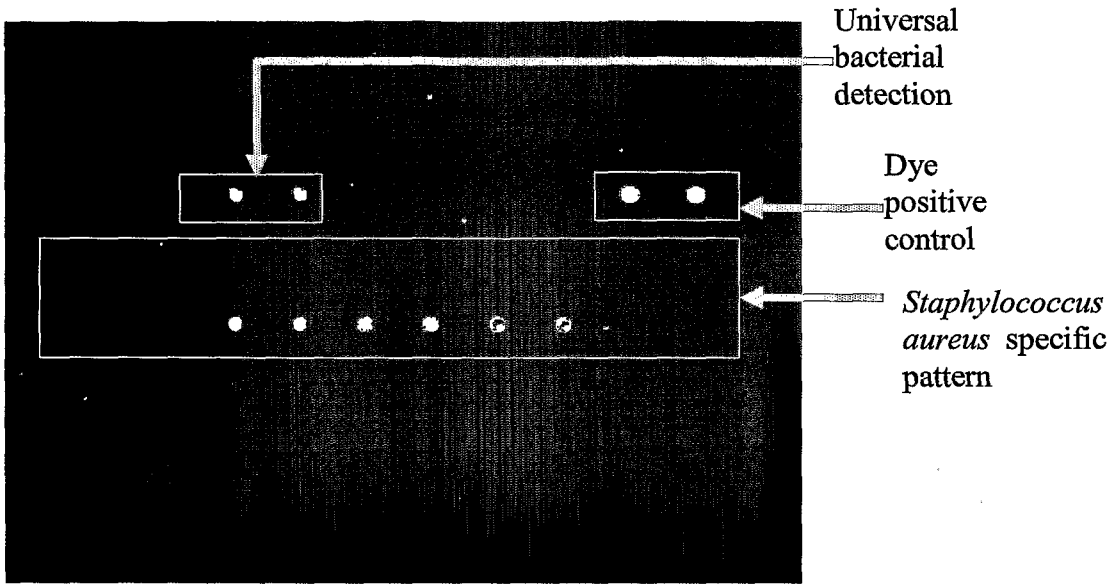

Universal bacterial detection
Dye positive control
Staphylococcus aureus specific pattern

FIGURE 5

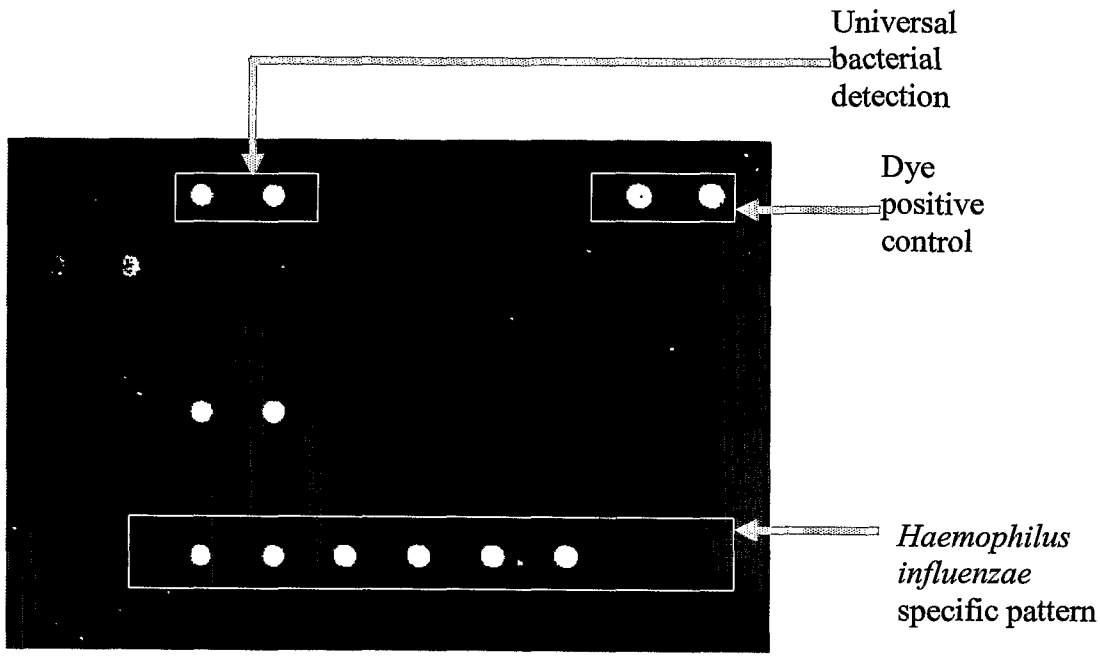

Universal bacterial detection
Dye positive control
Haemophilus influenzae specific pattern

FIGURE 6

SEQUENCE LISTING

<110> Diagnostic Array Systems Pty Ltd

<120> METHOD FOR THE DETECTION AND/OR IDENTIFICATION A MICROORGANISM

<130> FP24311

<140> PCT/AU2006/

<141> 2006-07-27

<150> AU2005904015

<151> 2005-07-27

<160> 395

<170> PatentIn version 3.3

<210> 1

<211> 1191

<212> DNA

<213> Mycobacterium tuberculosis

<400> 1

gtggcgaagg cgaagttcca gcggaccaag ccccacgtca acatcgggac catcggtcac 60

gttgaccacg gcaagaccac cctgaccgcg gctatcacca aggtcctgca cgacaaattc 120

cccgatctga acgagacgaa ggcattcgac cagatcgaca acgcccccga ggagcgtcag 180

cgcggtatca ccatcaacat cgcgcacgtg gagtaccaga ccgacaagcg gcactacgca 240

cacgtcgacg cccctggcca cgccgactac atcaagaaca tgatcaccgg cgccgcgcag 300

atggacggtg cgatcctggt ggtcgccgcc accgacggcc cgatgcccca gacccgcgag 360

cacgttctgc tggcgcgtca agtgggtgtg ccctacatcc tggtagcgct gaacaaggcc 420

gacgcagtgg acgacgagga gctgctcgaa ctcgtcgaga tggaggtccg cgagctgctg 480

gctgcccagg aattcgacga ggacgccccg gttgtgcggg tctcggcgct caaggcgctc 540

gagggtgacg cgaagtgggt tgcctctgtc gaggaactga tgaacgcggt cgacgagtcg 600

attccggacc cggtccgcga gaccgacaag ccgttcctga tgccggtcga ggacgtcttc 660

accattaccg gccgcggaac cgtggtcacc ggacgtgtgg agcgcggcgt gatcaacgtg 720

aacgaggaag ttgagatcgt cggcattcgc ccatcgacca ccaagaccac cgtcaccggt 780

gtggagatgt tccgcaagct gctcgaccag ggccaggcgg gcgacaacgt tggtttgctg 840

ctgcggggcg tcaagcgcga ggacgtcgag cgtggccagg ttgtcaccaa gcccggcacc 900

accacgccgc acaccgagtt cgaaggccag gtctacatcc tgtccaagga cgagggcggc 960

cggcacacgc cgttcttcaa caactaccgt ccgcagttct acttccgcac caccgacgtg 1020
accggtgtgg tgacactgcc ggagggcacc gagatggtga tgcccggtga caacaccaac 1080
atctcggtga agttgatcca gcccgtcgcc atggacgaag gtctgcgttt cgcgatccgc 1140
gagggtggcc gcaccgtggg cgccggccgg gtcaccaaga tcatcaagta g 1191

<210> 2
<211> 953
<212> DNA
<213> Legionella longbeachiae

<400> 2
gtttggtggt atagcaaaag catacgatca gatcgatgct gcaccagaag aaagagagcg 60
aggcatcaca atctcaacag cgcacgttga atacgaatca gcgaacagac actatgcgca 120
tgttgatccc aggacatgct gactatgtga agaacatgat tacaggtgct gcacaaatgg 180
acggagcgat acttgtagta tctgcagcag atggtcctat gcctcaaact agagaacata 240
ttctattgtc acgcagtagg ggtgccctac atagtagtgt ttatgaataa agcggacatg 300
gttgatgatc cagagctctt agagcttgta gaaatggaag tccgtgattt gttgagcagt 360
tatgattttc ctggtgatgt atccaattgt tgttggatct gcattgaaag cattagaagg 420
tgataccagt gatattgggg tgccagcaat agagaaattg gtagagacta tggattctta 480
tattccagag cctgtaagaa acattacaat cgttcttatt accgattgaa gatgtatttt 540
caatatctgg ccgaggaaca gtagtaacag gacgtatcga aagcggaatt atcaaagttg 600
gcgaagagat cgaaattgta ggaattcgtg aactgaaaga ctacctgtac tggtgttgag 660
atgttccgta agttattaga cgaaggacgt gcaggtgata acgttggtat attactccgt 720
ggaacaaaac gagacgaagt tgagcgtggt caggtatagc caaccaggta caattaagcc 780
tcacactaaa tttgaagcgg aagtatacgt gttatctaaa gaagaaggtg gacgacatac 840
cccgtttttc aatggatata gaccacagtt ttattttaga acacagacga acaggaactt 900
gtgatttacc atctggagtt gaaatggtaa tgcctggaga taacgtacag ttg 953

<210> 3
<211> 924
<212> DNA
<213> Fusobacterium nucleatum

<400> 3
agctaaaaaa gtagattttg accaaattga tgctgctcca gaagaaaaag aaagaggaat 60
tactattaat acagctcata ttgaatatga aacagcaaat agacactatg ctcacgttga 120
ctgtccagca tgctgactat gttaaaaata tgataactgg tgctgcacaa atggatggag 180

ctatacttgt tgtatcagct gctgatggtc ctatgcctca aacaagagaa catatcttac 240
tttctagaca agtgagtacc atatattatt gtttatttaa ataaatcaga tatggttgat 300
gatgaagaat tactagaatt ggtagaaatg gaagttagag aattattaac tgaatatgga 360
ttcccaggag atgacatcct gaattagagg ttcatcatta ggagctttaa atggtgaaga 420
aaaatggata gaaaagataa tggaacttat ggatgcagta gatagctata tccctactcc 480
tgaaagagca gtagatcaac cattctgagc caatagaaga cgttttcact atcacaggaa 540
gaggaacagt tgttacagga agagttgaaa gaggaatcat caaagttgga gaagaaattg 600
aaatagttgg aattaaacct acaactaaaa cactttacag gggttgaaat gtttagaaaa 660
cttcttgatc aaggtcaagc aggagataat atcggagtat tattaagagg aactaagaaa 720
gaagaagttg aaagaggaca agttcttgct aaaccagaag taccaccctc atacaaactt 780
taaaggtgaa gtctatgtat taactaaaga tgaaggagga agacacactc catttttcac 840
aggatacaga cctcaattct attttagaac tactgatatc actgtgcaga actttaccag 900
atggagtaga aatggttatg ccag 924

<210> 4
<211> 909
<212> DNA
<213> Moraxella catarrhalis

<400> 4
agaaaaagca cgtggtatta ctattaacac atctcacatc gagtatgaca cagaagctcg 60
tcactacgca catgtagact gcccaggcca cgctgactat gttaaaaaca tgatcacagg 120
tgctgcagat ggacggcgca atcttggttg tttctgcaac tgacggccca atgccacaaa 180
ctcgtgagca catcctactg tctcgtcagg ttggtgtacc atatatcatg gtattcatga 240
acaagtgcga catgagatga tgaagagcta ctagaattgg ttgaaatgga agttcgtgaa 300
cttctatctg actatgactt cccaggtgat gacaccccaa tcatcaaagg ttcagcacta 360
gaagcattga atggttctgc gtaaatatgg tgagcctgct gttctagaac tactagacac 420
gctagattca tacatcccag agcctgagcg tgacatcgat aagtcattcc taatgccaat 480
cgaagatgtc ttctcaatct caggtgtgta ctgttgtaac tggccgtgtt gaatcaggca 540
ttattaaagt tggtgatgaa attgaaatca tcggtatcaa accaactact aaaaccactt 600
gtactggtgt tgaaatgttc cgtaagctgc tgacgaggtc gtgcaggtga gaactgtggt 660
atcctactac gtggtactaa gcgtgaagaa gttcaacgtg gtcaagtatt ggctaagcca 720
ggttcaatca ccccacacac caagtttgat gcagaagata cggctatcaa aagaagaagg 780
tggtcgtcat accccattcc taaatggcta tcgtccacag ttctacttcc gtaccacaga 840
cgtaactggt gccatcaccc tacaagaagg cactgaaatg gtttgccagt gataacgttg 900
agatgagcg 909

<210> 5

<211> 982

<212> DNA

<213> Stenotrophomonas maltophilia

<400> 5

tgacaagatc ggtgccgagc gcttcggtgg cgagttcaag gactactcct cgatcgacgc 60
cgcgccggaa gaaaaggctc gtggtatcac gatctcgacc gcgcacgtcg aatacgaatc 120
cccgatctca ctacgcccac gttgactgcc cgggccacgc tgactacgtc aagaacatga 180
tcaccggtgc cgcccagatg gacggcgcga tcctggtgtg ctcggccgct gacggcccga 240
tgccgcagac ccgggcacat cctgctgtcg cgccaggtcg gcgtgccgta catcgtcgtg 300
ttcctgaaca aggccgacat ggtcgacgac gccgagctgc tcgagctggt cgagatggaa 360
gtgcgcgaac tgctgagcag tcgagttccc gggcgacgac accccgatca tcgccggttc 420
ggcccgcctg gcgctggaag gcgaccagag cgacatcggc gtgccggcca tcctgaagct 480
ggtcgacgcg ctggacagct ggattcgggc cggagcgtgc gatcgacaag ccgttcctga 540
tgccggtgga agacgtgttc tcgatctcgg gccgcggcac cgtggtgacc ggtcgtatcg 600
agcgcggcgt gatcaaggtt ggcgacgaaa tcgaaatcgt cggcatccgt ccggtgcaga 660
agaccaccgt ccggcgttga aatgttccgc aagctgctgg accagggtca ggcaggcgac 720
aacgctggcc tgctgctgcg cggcaccaag cgtgacgacg tcgagcgtgg ccaggtgctg 780
gccaagccgg gcacgacagc cgcacaccaa gttcgaaggc gaagtgtacg tcctgtcgaa 840
ggacgaaggc ggccgccaca ccccgttctt caacggctac cgtccgcagt tctacttccg 900
caccaccgac atcaccggcg cactcgctgc cggaaggcgt cgaaatggtg atgccgggtg 960
acaacgtcaa gatggtcgtc ac 982

<210> 6

<211> 984

<212> DNA

<213> Streptococcus sanguis

<400> 6

ctgtattggc acgtcgcttg ccttcagcag ttaaccaacc taaagactat gcgtctatcg 60
atgctgctcc agaagaacgc gaacgcggaa tcactatcaa cactgcgcac gttgagtatg 120
aaactgcagc gtcactacgc tcacatcgac gctccaggac acgcggacta cgttaaaaac 180
atgatcactg gtgccgctca aatggacgga gctatccttg tagtagcttc aactgacgga 240
ccaatgccac aaatgtgagc acatcttgct ttcacgtcag gttggtgtta aacacttgat 300
cgtcttcatg aacaaagttg acttggttga cgatgaagaa ttgcttgaat tggttgaaat 360
ggaaatccgt gacctcttgt cagaatacga cttcccaggt gacgatcttc cagttatcca 420

aggttcagct taaagctctt gaaggtgact ctaaatatga agacatcatc atggaattga 480
tggacactgt tgatgagtac atcccagaac cagaacgcga tactgacaag ccattgcttc 540
ttccagtcga agacgttctc aatcactggt cgtggtacag ttgcttcagg acgtatcgac 600
cgtggtatcg ttaaagtcaa cgacgaaatc gaaatcgttg gtatcaaaga agaaatccaa 660
aaagcagttg ttactggtgt tgaagttccg taaacagctt gacgaaggtc ttgcagggac 720
aacgtaggtg tgcttctccg tggtatccaa cgtgatgaaa tcgaacgtgg acaagttatc 780
gctaaaccag gttcaatcaa cccacacata attcaagggt gaagtttata tccttactaa 840
agaagaaggc ggacgtcaca ctccattctt caacaactac cgtccacagt tctacttccg 900
tacaactgac gttacaggtt caatcgaact tccacaggac tgaaatggta atgcctggtg 960
ataacgtaac aatcgacgtt gagt 984

<210> 7
<211> 964
<212> DNA
<213> Streptococcus mutans

<400> 7
ctgtactggc acgtcgtctt ccaagtgcag taaaccaacc aaaagattat tcatctattg 60
atgctgctcc tgaagaacgc gaacgcggta tcactatcaa tactgcgcac gttgagtatg 120
aaactgaaac gtcactatgc tcacattgac gctccaggac acgcggacta tgttaaaaac 180
atgattactg gtgctgccca aatggatggt gctatccttg tagtagcttc aactgatggt 240
ccaatgccac aaatgtgaac acattcttct ttcacgtcaa gttggtgtta aataccttat 300
tgtcttcatg aataaggttg atttggttga tgatgaagaa ctgcttgaat tggttgaaat 360
ggaaatccgt gatcttcttc aaatatgatt tcccaggtga tgatattcca gttattcaag 420
gttcagctct taaagctctt gaaggtgata ctgctcaaga agatatcatc atggaattaa 480
tgcatactgt tgatgactac attccgatca gaacgtgata ctgataagcc actccttctt 540
ccagtcgaag atgttttctc aatcactggt cgtggtactg ttgcttcagg acgtattgat 600
cgtggtactg ttaaagttaa cgatgaagtt gaatcttggt atccgtgatg acattcaaaa 660
agctgttgtt actggtgttg aaatgttccg taaacaattg gatgaaggta ttgcagggga 720
taatgttggt gttctccttc gtggtatcca acgtgataaa tcaacgtggt caagttcttg 780
ctaaaccagg ttcaattcac ccacatacta aattcaaagg tgaagtttat atccttacta 840
aagaagaagg tggacgtcat acaccattct tcaataacta tcgccacaat ctacttccgt 900
acaactgacg taactggttc aattgagttg ccagcaggta ctgaaatggt tatgcctggt 960
gata 964

<210> 8
<211> 870

<212> DNA

<213> Burkholderia cepacia

<400> 8

gttcggcggc gaagcgaaga agtacgacga aatcgacgcg gcgccggaag aaaaggcacg 60
cggcatcacg atcaacaccg cgcacatcga gtacgaaacg gcgaaccgcc actacgcgca 120
cgtcgacccc gggccacgcc gactacgtga agaacatgat cacgggtgca gcgcagatgg 180
acggcgcaat cctggtgtgc tcggccgcag acggcccgat gccgcaaacg cgtgagcaca 240
tcctgctggc gcgcggttgg cgttccgtac atcatcgtgt tcctgaacaa gtgcgacatg 300
gtggacgacg ctgaactgct cgagctggtc gagatggaag ttcgcgaact gctgtcgaag 360
tacgacttcc cgggcgacgc agccgatcat caagggttcg gcaaagctgg cgctggaagg 420
cgacaagggc gagctgggcg aaacggcgat catgaacctg gccgacgcgc tggacacgta 480
catcccgacg ccggagcgcg cggtgacgta cgttcctgat gccggtggaa gacgtgttct 540
cgatctcggg ccgcggtacg gtggtgacgg gtcgtgtgga gcgcggcgtg gtgaaggtcg 600
gtgaggaaat cgaaatcgtc ggtatcaagc cacgggaaga cgacctgcac gggcgttgaa 660
atgttccgca agctgctgga ccagggtcag gcaggcgaca acgtcggtat cctgctgcgc 720
ggcacgaagc gtgaagacgt ggagcgcggt caggttcggc gagccgggtt cgatcacgcc 780
gcacacgcac ttcacggccg aagtgtacgt gctgagcaag gacgaaggcg gccgtcacac 840
gccgttcttc aacaactacc gtccgcagtt 870

<210> 9

<211> 936

<212> DNA

<213> Enterococcus faecalis

<400> 9

gggaaagcac aaagctacga ttctatcgat aacgctccag aagaaaaaga acgtggaatc 60
acaatcaaca cttctcatat cgaatatgaa actgaaactc gtcactatgc acacgttgac 120
tgcccagcac gcggactacg ttaaaaacat gatcactggt gctgctcaaa tggacggagc 180
tatcttagta gtttctgctg ctgatggtcc tatgcctcaa acacgtgaac atatcttatt 240
atcacgtaac gttggtacca tacatcgttg tattcttaaa caaaatggat atggttgatg 300
acgaagaatt attagaatta gtagaaatgg aagttcgtga cttattatca gaatacgatt 360
tcccaggcga tgatgttccg tatcgcaggt tctgctttga aagctttaga aggcgacgag 420
tcttatgaag aaaaaatctt agaattaatg gctgcagttg acgaatatat cccaactcca 480
gaacgtgata ctgacaaacc attcagatcc agtcgaagac gtattctcaa tcactggacg 540
tggtactgtt gctacaggac gtgttgaacg tggtgaagtt cgcgttggtg acgaagttga 600
aatcgttggt attaaagacg aaacatctaa acaacgttac aggtgttgaa atgttccgta 660

aattattaga ctacgctgaa gcaggcgaca acatcggtgc tttattacgt ggtgtagcac 720
gtgaagatat cgaacgtgga caagtattag ctaaaccgct acatcactcc acacacaaaa 780
ttcaaagctg aagtatacgt attatcaaaa gaagaaggcg gacgtcacac tccattcttc 840
actaactacc gtcctcaatt ctacttccgt acaacagacg ttatggtgtg tagaattgcc 900
agaaggtact gaaatggtaa tgcctggtga taacgt 936

<210> 10
<211> 945
<212> DNA
<213> Corynebacterium diphtheriae

<400> 10
tgcacaatca tacgacatga ttgacaacgc accagaagaa aaagaacgtg gtattacaat 60
caatactgca cacatcgaat atcaaactga caaacgtcac tatgctcacg ttgactgccc 120
aggacactga ctatgttaaa aacatgatca ctggtgctgc gcaaatggac ggcggtatct 180
tagttgtatc tgcagctgat ggtccaatgc cacaaactcg tgaacacatt cttttatcac 240
gtaacgttgg tgtcagcttt agttgtattc ttaaacaaag ttgatatggt agacgacgaa 300
gaattattag aattagtaga aatggaagtt cgtgacttac tatctgaata cgacttccca 360
ggtgacgacg tacctgtaac gtggttcagc attaaaagct ttagaaggcg acgaaaaata 420
cgaagaaaaa atcttagaat taatgcaagc agttgatgac tacatcccaa ctccagaacg 480
tgattctgac aaaccattca tgatgcagtg aggacgtatt ctcaatcact ggtcgtggta 540
ctgttgctac aggccgtgtt gaacgtggtc aaatcaaagt tggtgaagaa gttgaaatca 600
tcggtttaca tgacacttct aaaacaactg tactgtgtag aaatgttccg taagttatta 660
gactacgctg aagctggtga caacattggt gctttattac gtggtgttgc tcgtgaagac 720
gtacaacgtg gtcaagtatt agctgctcct ggttcaatac acacatacaa aattcaaagc 780
ggaagtttac gttttatcta aagacgaagg tggacgtcac actccattct tcagtaacta 840
ccgcccacaa ttctatttcc gtactactga cgtaactggc gttttcaata ccagaaggta 900
ctgaaatggt tatgcctggt gacaacgttg aaatgacagt agaat 945

<210> 11
<211> 969
<212> DNA
<213> Neisseria meningitidis

<400> 11
cccgaagaaa aagcacgcgg tattaccatt aacacctcgc acgtggaata cgaaaccgaa 60
acccgccact acgcacacgt agactgcccg gggcacgccg actacgttaa aaacatgatt 120

accggtggca caaatggacg gtgcaatcct ggtatgttcc gcagccgacg gccctatgcc 180
gcaaacccgc gaacacatcc tgctggcccg tcaagtaggc gtaccttaca tcatcgtgtt 240
catgaacaaa tgcaatggtc gacgatgccg agctgttgga actggttgaa atggaaatcc 300
gcgacctgct gtccagctac gacttccccg gcgacgactg cccgattgta caaggttctg 360
cactgaaagc cttggaaggg agccgcttat gaagagaaaa tcttcgaact ggctgccgca 420
ttggacagct acatcccgac tcccgagcgt gccgtggaca aacctttcct gctgcctatc 480
gaagacgtgt tctccatttc cggccaggac agtagtaacc ggccgtgtag agcgcggtat 540
catccacgtt ggtgacgaga ttgaaatcgt cggtctgaaa gaaacccaaa aaaccacctg 600
taccggtgtt gaaatgttcc gcaaactgct gacgaggtca ggcgggcgac aacgtacgcg 660
tattgctgcg cggtaccaaa cgtgaagacg tggaacgcgg tcaggtattg gccaaacccg 720
ggtactatca ctcctcacac caagttcaaa gcagacggta cgactgagca aagaagaggg 780
cggccgccat actccgttct tcgccaacta ccgtccgcat tctacttccg taccaccgac 840
gtaaccggcg cggttacttt ggaagaagga gtggaaatgg taagcctggg agaacgtaac 900
catcaccgta gactgattgc gtctatcgct atggaagagc tgcgcttgcg tcgtgtcgcg 960
cgtgtccat 969

<210> 12
<211> 29
<212> DNA
<213> Haemophilus influenzae

<400> 12
tattatccgt acaggtgatg aagtagaaa 29

<210> 13
<211> 27
<212> DNA
<213> Haemophilus influenzae

<400> 13
aaagatacag cgaaaactac tgtaacg 27

<210> 14
<211> 24
<212> DNA
<213> haemophilus influenzae

<400> 14
catcggtgca ttattacgtg gtac 24

<210> 15
<211> 26
<212> DNA
<213> haemophilus influenzae

<400> 15
catactccat tcttcaaagg ttaccg 26

<210> 16
<211> 26
<212> DNA
<213> Haemophilus influenzae

<400> 16
tgactggtac aatcgaatta ccagaa 26

<210> 17
<211> 24
<212> DNA
<213> Haemophilus influenzae

<400> 17
ccgtaaatta cttgacgaag gtcg 24

<210> 18
<211> 22
<212> DNA
<213> Moraxella catarrhalis

<400> 18
tgagcgtgac atcgataagt ca 22

<210> 19
<211> 29

<212> DNA
<213> Moraxella catarrhalis

<400> 19
caggcattat taaagttggt gatgaaatt 29

<210> 20
<211> 28
<212> DNA
<213> Moraxella catarrhalis

<400> 20
aaaccaacta ctaaaaccac ttgtactg 28

<210> 21
<211> 21
<212> DNA
<213> Moraxella catarrhalis

<400> 21
cgtaagctgc tagacgaagg t 21

<210> 22
<211> 24
<212> DNA
<213> Moraxella catarrhalis

<400> 22
taccccattc ctaaatggct atcg 24

<210> 23
<211> 21
<212> DNA
<213> Moraxella catarrhalis

<400> 23
taactggtgc catcacccta c 21

<210> 24
<211> 22
<212> DNA
<213> Enterococcus faecalis

<400> 24
aaggcgacga gtcttatgaa ga 22

<210> 25
<211> 24
<212> DNA
<213> Enterococcus faecalis

<400> 25
aattaatggc tgcagttgac gaat 24

<210> 26
<211> 19
<212> DNA
<213> Enterococcus faecalis

<400> 26
gaagttcgcg ttggtgacg 19

<210> 27
<211> 28
<212> DNA
<213> Enterococcus faecalis

<400> 27
gacgaaacat ctaaaacaac tgttacag 28

<210> 28
<211> 23
<212> DNA
<213> Enterococcus faecalis

<400> 28
tggtgttgta gaattgccag aag 23

<210> 29
<211> 24
<212> DNA
<213> Enterococcus faecalis

<400> 29
taaaccagct acaatcactc caca 24

<210> 30
<211> 27
<212> DNA
<213> Staphylococcus aureus

<400> 30
gttgacatgg ttgacgatga agaatta 27

<210> 31
<211> 32
<212> DNA
<213> Staphylococcus aureus

<400> 31
cttagaatta atggaagctg tagatactta ca 32

<210> 32
<211> 28
<212> DNA
<213> Staphylococcus aureus

<400> 32
atcatcggtt tacatgacac atctaaaa 28

<210> 33

<211> 29
<212> DNA
<213> Staphylococcus aureus

<400> 33
caccacatac tgaatttaaa gcagaagta 29

<210> 34
<211> 26
<212> DNA
<213> Staphylococcus aureus

<400> 34
cattcttctc aaactatcgt ccacaa 26

<210> 35
<211> 24
<212> DNA
<213> Staphylococcus aureus

<400> 35
ctggtgttgt tcacttacca gaag 24

<210> 36
<211> 26
<212> DNA
<213> Staphylococcus aureus

<400> 36
tactgaaatg gtaatgcctg gtgata 26

<210> 37
<211> 20
<212> DNA
<213> Staphylococcus aureus

<400> 37

ccaatcgcga ttgaagacgg 20

<210> 38
<211> 27
<212> DNA
<213> Staphylococcus aureus

<400> 38
ctggttcagc attaaaagct ttagaag 27

<210> 39
<211> 25
<212> DNA
<213> Staphylococcus aureus

<400> 39
tgacaacatt ggtgcattat tacgt 25

<210> 40
<211> 24
<212> DNA
<213> Staphylococcus aureus

<400> 40
tgttacaggt gttgaaatgt tccg 24

<210> 41
<211> 26
<212> DNA
<213> Staphylococcus aureus

<400> 41
gtattatcaa aagacgaagg tggacg 26

<210> 42
<211> 26
<212> DNA

<213> Staphylococcus aureus

<400> 42
gcgatgctca atacgaagaa aaaatc 26

<210> 43
<211> 29
<212> DNA
<213> Staphylococcus epidermidis

<400> 43
tagacttaat gcaagcagtt gatgattac 29

<210> 44
<211> 25
<212> DNA
<213> Staphylococcus epidermidis

<400> 44
ccacacacaa aattcaaagc tgaag 25

<210> 45
<211> 25
<212> DNA
<213> Staphylococcus epidermidis

<400> 45
ctggtgttgt aaacttacca gaagg 25

<210> 46
<211> 21
<212> DNA
<213> Staphylococcus epidermidis

<400> 46
gaaatggtta tgcctggcga c 21

<210> 47
<211> 27
<212> DNA
<213> Staphylococcus epidermidis

<400> 47
ctggttctgc attaaaagca ttagaag 27

<210> 48
<211> 24
<212> DNA
<213> Staphylococcus epidermidis

<400> 48
tgacaacatc ggtgctttat tacg 24

<210> 49
<211> 25
<212> DNA
<213> Staphylococcus epidermidis

<400> 49
ctgttactgg tgtagaaatg ttccg 25

<210> 50
<211> 27
<212> DNA
<213> Staphylococcus epidermidis

<400> 50
cgtattatct aaagatgaag gtggacg 27

<210> 51
<211> 17
<212> DNA
<213> Pseudomonas aeruginosa

<400> 51
aagttcgagt gcgaagt 17

<210> 52
<211> 19
<212> DNA
<213> Pseudomonas aeruginosa

<400> 52
aaggcgtaga gatggtaat 19

<210> 53
<211> 17
<212> DNA
<213> Pseudomonas aeruginosa

<400> 53
ttcttcaagg gctaccg 17

<210> 54
<211> 23
<212> DNA
<213> Pseudomonas aeruginosa

<400> 54
atcatcaagg tccaggaaga agt 23

<210> 55
<211> 19
<212> DNA
<213> Pseudomonas aeruginosa

<400> 55
accaagacta cctgcaccg 19

<210> 56
<211> 20

<212> DNA
<213> Pseudomonas aeruginosa

<400> 56
tgtacgtgct gtccaaggaa 20

<210> 57
<211> 18
<212> DNA
<213> Pseudomonas aeruginosa

<400> 57
ccggtaactg cgaactgc 18

<210> 58
<211> 20
<212> DNA
<213> Bordetella pertussis

<400> 58
gtacattctg tccaaggaag 20

<210> 59
<211> 18
<212> DNA
<213> Bordetella pertussis

<400> 59
gacaacgtgg gtatcttg 18

<210> 60
<211> 17
<212> DNA
<213> Bordetella pertussis

<400> 60
gacaaggaaa tggtgct 17

<210> 61
<211> 21
<212> DNA
<213> Chlamydophila pneumoniae

<400> 61
aatttaagtc agctgtttac g 21

<210> 62
<211> 20
<212> DNA
<213> Chlamydophila pneumoniae

<400> 62
gaggtattgg aaagaacgat 20

<210> 63
<211> 20
<212> DNA
<213> Chlamydophila pneumoniae

<400> 63
ataacgttga gcttgatgtt
20

<210> 64
<211> 20
<212> DNA
<213> Haemophilus influenzae

<400> 64
acttcgaatc agaagtgtac 20

<210> 65
<211> 20
<212> DNA

<213> Haemophilus influenzae

<400> 65
caaacgtgaa gaaatcgaac 20

<210> 66
<211> 21
<212> DNA
<213> Haemophilus influenzae

<400> 66
cgataacatc aagatgacag t 21

<210> 67
<211> 20
<212> DNA
<213> Legionella pneumophila

<400> 67
agtttgaagc agaagtgtat 20

<210> 68
<211> 20
<212> DNA
<213> Legionella pneumophila

<400> 68
tgttattacg aggtacgaag 20

<210> 69
<211> 23
<212> DNA
<213> Legionella pneumophila

<400> 69
agataatgtg caattagttg tta 23

<210> 70
<211> 23
<212> DNA
<213> Mycoplasma pneumoniae

<400> 70
gaaatttaaa gcggaaatct atg 23

<210> 71
<211> 19
<212> DNA
<213> Mycoplasma pneumoniae

<400> 71
gaacgtggtc aagtgttag 19

<210> 72
<211> 20
<212> DNA
<213> Mycoplasma pneumoniae

<400> 72
gacaatacct cgattacagt 20

<210> 73
<211> 20
<212> DNA
<213> Mycobacterium tuberculosis

<400> 73
aacatctcgg tgaagttgat 20

<210> 74
<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 74
tgacaacacc aacatctc 18

<210> 75
<211> 16
<212> DNA
<213> Mycobacterium tuberculosis

<400> 75
acgaaggtct gcgttt 16

<210> 76
<211> 22
<212> DNA
<213> Staphylococcus aureus

<400> 76
tgaattcaaa gcagaagtat ac 22

<210> 77
<211> 20
<212> DNA
<213> Staphylococcus aureus

<400> 77
attattacgt ggtgttgctc 20

<210> 78
<211> 21
<212> DNA
<213> Staphylococcus aureus

<400> 78
gtgataacgt tgaaatgaca g 21

<210> 79
<211> 21

<212> DNA
<213> Streptococcus pneumoniae

<400> 79
aattcaaagg tgaagtctac a 21

<210> 80
<211> 19
<212> DNA
<213> Streptococcus pneumoniae

<400> 80
caacgtgatg aaatcgaac 19

<210> 81
<211> 18
<212> DNA
<213> Streptococcus pneumoniae

<400> 81
gacaatcgac gttgagtt 18

<210> 82
<211> 24
<212> DNA
<213> Streptococcus pyogenes

<400> 82
caaaggtgaa gtatatatcc tttc 24

<210> 83
<211> 18
<212> DNA
<213> Streptococcus pyogenes

<400> 83
caacgtgacg aaatcgaa 18

<210> 84
<211> 18
<212> DNA
<213> Streptococcus pyogenes

<400> 84
cgtgacaatc aacgttga 18

<210> 85
<211> 23
<212> DNA
<213> Pseudomonas aeruginosa

<400> 85
cacgttgact gccccggtca cgc 23

<210> 86
<211> 25
<212> DNA
<213> Pseudomonas aeruginosa

<400> 86
acgccttccg gcagttcgca gttac 25

<210> 87
<211> 22
<212> DNA
<213> Pseudomonas aeruginosa

<400> 87
atcggtcacg ttgaccatgg ca 22

<210> 88
<211> 21
<212> DNA
<213> Pseudomonas aeruginosa

<400> 88
gccgccttcg cggatcgcga a 21

<210> 89
<211> 24
<212> DNA
<213> Neisseria gonorrhoeae

<400> 89
caccctgact gctgctttga ctac 24

<210> 90
<211> 22
<212> DNA
<213> Neisseria gonorrhoeae

<400> 90
ccacctgtac cggcgttgaa at 22

<210> 91
<211> 26
<212> DNA
<213> Neisseria gonorrhoeae

<400> 91
gtattggcca aaccgggtac tatcac 26

<210> 92
<211> 17
<212> DNA
<213> Neisseria gonorrhoeae

<400> 92
gagggcggcc gccatac 17

<210> 93

<211> 27
<212> DNA
<213> Neisseria gonorrhoeae

<400> 93
ccattaacac ctcgcacgta gaatacg 27

<210> 94
<211> 26
<212> DNA
<213> Neisseria gonorrhoeae

<400> 94
attttagcta aaaaattcgg cggcgc 26

<210> 95
<211> 22
<212> DNA
<213> Helicobacter pylori

<400> 95
agcgatttca gcggtgcttt ct 22

<210> 96
<211> 25
<212> DNA
<213> Helicobacter pylori

<400> 96
gagactgaaa ccagacacta tgcgc 25

<210> 97
<211> 25
<212> DNA
<213> Helicobacter pylori

<400> 97

ggatggagcg atcttggttg tttct 25

<210> 98
<211> 26
<212> DNA
<213> Helicobacter pylori

<400> 98
gcgtgcctca cattgttgtt ttctta 26

<210> 99
<211> 33
<212> DNA
<213> Helicobacter pylori

<400> 99
atggtagatg accaagaatt gttagagctt gta 33

<210> 100
<211> 24
<212> DNA
<213> Helicobacter pylori

<400> 100
gttgagcgcg tatgaatttc ctgg 24

<210> 101
<211> 24
<212> DNA
<213> Helicobacter pylori

<400> 101
cctatcgtag cgggttcagc ttta 24

<210> 102
<211> 26
<212> DNA

<213> Helicobacter pylori

<400> 102
gtgaatgggg cgaaaaagtg cttaaa 26

<210> 103
<211> 28
<212> DNA
<213> Helicobacter pylori

<400> 103
tcatggctga agtggattcc tatatccc 28

<210> 104
<211> 21
<212> DNA
<213> Helicobacter pylori

<400> 104
tgggcgtgct tttgagagga a 21

<210> 105
<211> 28
<212> DNA
<213> Helicobacter pylori

<400> 105
atggttctat gcaaaccagg ttctatca 28

<210> 106
<211> 29
<212> DNA
<213> Helicobacter pylori

<400> 106
gttaggaacc aaatttgcga ttcgtgaag 29

<210> 107
<211> 23
<212> DNA
<213> Helicobacter hepaticus

<400> 107
gcagctattt ctgctgtgtt ggc 23

<210> 108
<211> 25
<212> DNA
<213> Helicobacter hepaticus

<400> 108
cggagcaatt cttgttgttt ctgcc 25

<210> 109
<211> 30
<212> DNA
<213> Helicobacter hepaticus

<400> 109
gatgctgaat tacttgagct tgtggaaatg 30

<210> 110
<211> 33
<212> DNA
<213> Helicobacter hepaticus

<400> 110
ttgcttagtc aatatgattt ccctggagat gat 33

<210> 111
<211> 34
<212> DNA
<213> Helicobacter hepaticus

<400> 111
taatgtaggt gaatggggtg aaaaagtatt aaaa 34

<210> 112
<211> 27
<212> DNA
<213> Helicobacter hepaticus

<400> 112
gtagaaagag gtgttgtgca agtaggc 27

<210> 113
<211> 34
<212> DNA
<213> Helicobacter hepaticus

<400> 113
ggtattttgt tgcgaggaac aaaaaaagaa gaag 34

<210> 114
<211> 23
<212> DNA
<213> Helicobacter hepaticus

<400> 114
atggttcttt gtaagcccgg ctc 23

<210> 115
<211> 25
<212> DNA
<213> Helicobacter hepaticus

<400> 115
ggaacttatt gctcctgtgg cactt 25

<210> 116
<211> 33

<212> DNA
<213> Campylobacter jejuni

<400> 116
agcgtggtat tactattgct acttctcata ttg 33

<210> 117
<211> 28
<212> DNA
<213> Campylobacter jejuni

<400> 117
cacattcttc tttctcgtca agtaggcg 28

<210> 118
<211> 34
<212> DNA
<213> Campylobacter jejuni

<400> 118
gatgatgctg aacttttaga gttagttgaa atgg 34

<210> 119
<211> 34
<212> DNA
<213> Campylobacter jejuni

<400> 119
ttattaagct cttatgattt cccaggcgat gata 34

<210> 120
<211> 31
<212> DNA
<213> Campylobacter jejuni

<400> 120
ggcgatgata cacctattat ttctggttct g 31

<210> 121
<211> 27
<212> DNA
<213> Campylobacter jejuni

<400> 121
gctcttgaag aagctaaagc tggacaa 27

<210> 122
<211> 32
<212> DNA
<213> Campylobacter jejuni

<400> 122
cttatggctg cagttgatag ctatattcca ac 32

<210> 123
<211> 32
<212> DNA
<213> Campylobacter jejuni

<400> 123
cgtggtacta aaaaagaaga agttatccgt gg 32

<210> 124
<211> 34
<212> DNA
<213> Campylobacter jejuni

<400> 124
gatgttacag gttcgattaa attagctgat ggtg 34

<210> 125
<211> 25
<212> DNA
<213> Campylobacter jejuni

<400> 125
ctgtaagctt gatcgctcca gtagc 25

<210> 126
<211> 22
<212> DNA
<213> Chlamydophila abortus

<400> 126
gcaattacac gcgcactgtc ag 22

<210> 127
<211> 27
<212> DNA
<213> Chlamydophila abortus

<400> 127
gttgaatacg aaacccctaa ccgtcac 27

<210> 128
<211> 21
<212> DNA
<213> Chlamydophila abortus

<400> 128
ccgctactga cggtgctatg c 21

<210> 129
<211> 22
<212> DNA
<213> Chlamydophila abortus

<400> 129
ctgttggcaa gacaggtagg gg 22

<210> 130

<211> 20
<212> DNA
<213> Chlamydophila abortus

<400> 130
aggatgccga gcttgtcgac 20

<210> 131
<211> 29
<212> DNA
<213> Chlamydophila abortus

<400> 131
gaagaaaaag gttacaaagg ctgccctat 29

<210> 132
<211> 33
<212> DNA
<213> Chlamydophila abortus

<400> 132
gaaatgttta gaaaagaact tccagaaggt cag 33

<210> 133
<211> 24
<212> DNA
<213> Chlamydophila abortus

<400> 133
atggtgattt gccaacctaa cagc 24

<210> 134
<211> 28
<212> DNA
<213> Chlamydophila abortus

<400> 134

gatatgattt ctcaagagga tgccgagc 28

<210> 135
<211> 29
<212> DNA
<213> Clostridium perfringens

<400> 135
agcagctata acaacagttt tagcacaag 29

<210> 136
<211> 33
<212> DNA
<213> Clostridium perfringens

<400> 136
ggtggagcag aagcattcaa atatgatgaa ata 33

<210> 137
<211> 29
<212> DNA
<213> Clostridium perfringens

<400> 137
gttgagtacg aaacagctaa cagacacta 29

<210> 138
<211> 30
<212> DNA
<213> Clostridium perfringens

<400> 138
ggagctatat tagtttgttc agcagctgat 30

<210> 139
<211> 34
<212> DNA

<213> Clostridium perfringens

<400> 139
cttattatca tcaagagttg gagttgacca catc 34

<210> 140
<211> 22
<212> DNA
<213> Clostridium perfringens

<400> 140
aacccaactg acgaagctgc aa 22

<210> 141
<211> 28
<212> DNA
<213> Clostridium perfringens

<400> 141
agaggagttc tacatgtagg agacgaag 28

<210> 142
<211> 30
<212> DNA
<213> Clostridium perfringens

<400> 142
ggtatccaaa gaactgatat cgaaagaggt 30

<210> 143
<211> 31
<212> DNA
<213> Clostridium perfringens

<400> 143
caagttttag ctcaagttgg aacaatcaac c 31

<210> 144
<211> 33
<212> DNA
<213> Clostridium perfringens

<400> 144
ggaagttgaa ttaatcacag aaatcgctat gga 33

<210> 145
<211> 28
<212> DNA
<213> Clostridium tetani

<400> 145
gggtttgcga aagcattcaa atatgacg 28

<210> 146
<211> 28
<212> DNA
<213> Clostridium tetani

<400> 146
atcctagtag taagtgcagc agatggtc 28

<210> 147
<211> 31
<212> DNA
<213> Clostridium tetani

<400> 147
cacatactat tagcatccag agttggagtt g 31

<210> 148
<211> 28
<212> DNA
<213> Clostridium tetani

<400> 148
gaccaagtag atgacgcaga gttaatcg 28

<210> 149
<211> 24
<212> DNA
<213> Clostridium tetani

<400> 149
acgcaccagt agtagtagga tccg 24

<210> 150
<211> 28
<212> DNA
<213> Clostridium tetani

<400> 150
gaaaatccag aagatgatgc agcaacac 28

<210> 151
<211> 33
<212> DNA
<213> Clostridium tetani

<400> 151
gacgaaatag aaatcgtagg attaagtgat gaa 33

<210> 152
<211> 33
<212> DNA
<213> Clostridium tetani

<400> 152
agcaagaaat cagtaatcac aggaatagaa atg 33

<210> 153
<211> 29

<212> DNA
<213> Clostridium tetani

<400> 153
attagcagca acaggatcag taaaaccac 29

<210> 154
<211> 35
<212> DNA
<213> Clostridium tetani

<400> 154
atgaaggtag aattaataac aagagtagca atgga 35

<210> 155
<211> 26
<212> DNA
<213> Listeria monocytogenes

<400> 155
ggctatgctg atgcacaagc ttatga 26

<210> 156
<211> 31
<212> DNA
<213> Staphylococcus saprophyticus

<400> 156
ggtgactctg tagcacaatc atatgacatg a 31

<210> 157
<211> 33
<212> DNA
<213> Staphylococcus saprophyticus

<400> 157
acttcacaca tcgagtatac tactgacaaa cgt 33

<210> 158
<211> 30
<212> DNA
<213> Staphylococcus saprophyticus

<400> 158
cggagctatc ttagtagtat ctgctgctga 30

<210> 159
<211> 26
<212> DNA
<213> Staphylococcus saprophyticus

<400> 159
acaaggtgac aaggaagctc aagacc 26

<210> 160
<211> 29
<212> DNA
<213> Staphylococcus saprophyticus

<400> 160
ttaatgcaag ctgttgatga cttcattcc 29

<210> 161
<211> 32
<212> DNA
<213> Staphylococcus saprophyticus

<400> 161
gtatgcaaga agaatcaagc aaaacaactg tt 32

<210> 162
<211> 26
<212> DNA
<213> Staphylococcus saprophyticus

<400> 162
ggtgtttcac gtgatgatgt acaacg 26

<210> 163
<211> 38
<212> DNA
<213> Staphylococcus saprophyticus

<400> 163
gaaatggatg ttgaattaat ttctccaatc gctattga 38

<210> 164
<211> 20
<212> DNA
<213> Vibrio vulnificus

<400> 164
tacggcggta cagctcgtga 20

<210> 165
<211> 33
<212> DNA
<213> Vibrio vulnificus

<400> 165
ggtatcacaa tctctacttc tcacgtagag tac 33

<210> 166
<211> 24
<212> DNA
<213> Vibrio vulnificus

<400> 166
tacgacactc catctcgcca ctac 24

<210> 167

<211> 31
<212> DNA
<213> Vibrio vulnificus

<400> 167
gatgatgaag agcttctaga gctagtagaa a 31

<210> 168
<211> 24
<212> DNA
<213> Vibrio vulnificus

<400> 168
ctagcagaag cgctagacag ctac 24

<210> 169
<211> 26
<212> DNA
<213> Vibrio vulnificus

<400> 169
gttgttgagc taatctctcc aatcgc 26

<210> 170
<211> 30
<212> DNA
<213> Vibrio parahaemolyticus

<400> 170
gctgctatct gtactactct agctaaagtg 30

<210> 171
<211> 22
<212> DNA
<213> Vibrio parahaemolyticus

<400> 171

ggcggtgaag cgaaagactt cg 22

<210> 172
<211> 28
<212> DNA
<213> Vibrio parahaemolyticus

<400> 172
ggtatcacaa tcgcaacttc tcacgtag 28

<210> 173
<211> 23
<212> DNA
<213> Vibrio parahaemolyticus

<400> 173
gtacgacact ccaagccgtc act 23

<210> 174
<211> 33
<212> DNA
<213> Vibrio parahaemolyticus

<400> 174
acgatgaaga gctactagaa ctagtagaaa tgg 33

<210> 175
<211> 33
<212> DNA
<213> Vibrio parahaemolyticus

<400> 175
ctagctgaag ctctagatac ttacatccct gag 33

<210> 176
<211> 20
<212> DNA

<213> Vibrio parahaemolyticus

<400> 176
agagcgtgcg gtagatcagc 20

<210> 177
<211> 25
<212> DNA
<213> Vibrio parahaemolyticus

<400> 177
gcgtggtatc ctaacagttg gtgac 25

<210> 178
<211> 27
<212> DNA
<213> Vibrio parahaemolyticus

<400> 178
gttgagctaa ttgcaccaat cgcaatg 27

<210> 179
<211> 34
<212> DNA
<213> Vibrio cholerae

<400> 179
acaatcaata cttctcacgt agagtacgat actc 34

<210> 180
<211> 26
<212> DNA
<213> Vibrio cholerae

<400> 180
tgtccaggac acgcggatta tgttaa 26

<210> 181
<211> 23
<212> DNA
<213> Vibrio cholerae

<400> 181
tagttgtagc ggcaactgac ggt 23

<210> 182
<211> 30
<212> DNA
<213> Vibrio cholerae

<400> 182
gcactggata cttatattcc agagccagag 30

<210> 183
<211> 32
<212> DNA
<213> Vibrio cholerae

<400> 183
acagtaaaaa cgacctgtac aggtgtagag at 32

<210> 184
<211> 24
<212> DNA
<213> Vibrio cholerae

<400> 184
gacgtaacag gcagcatcga gcta 24

<210> 185
<211> 28
<212> DNA
<213> Vibrio cholerae

<400> 185
gtgaagatgg ttgtagacct gattgcac 28

<210> 186
<211> 24
<212> DNA
<213> Acinetobacter baumanii

<400> 186
cgatcgcaac aatctgtgcg aaaa 24

<210> 187
<211> 32
<212> DNA
<213> Acinetobacter baumanii

<400> 187
gaaagattac tcacaaatcg actcagctcc ag 32

<210> 188
<211> 30
<212> DNA
<213> Acinetobacter baumanii

<400> 188
cacgtagaat acgactctcc aattcgtcac 30

<210> 189
<211> 27
<212> DNA
<213> Acinetobacter baumanii

<400> 189
gacgacgaag agcttcttga gcttgta 27

<210> 190
<211> 28

<212> DNA
<213> Acinetobacter baumanii

<400> 190
tggtcaatat ggcgagtctt cagttctt 28

<210> 191
<211> 19
<212> DNA
<213> Acinetobacter baumanii

<400> 191
aaccagagcg cgcaatcga 19

<210> 192
<211> 23
<212> DNA
<213> Acinetobacter baumanii

<400> 192
aagctggtat cgtgaaagtg ggc 23

<210> 193
<211> 25
<212> DNA
<213> Acinetobacter baumanii

<400> 193
tgcaatccag ttgaaagaag gcgtt 25

<210> 194
<211> 33
<212> DNA
<213> Acinetobacter baumanii

<400> 194
cgttgaaatg tcagtagaat tgatccaccc aat 33

<210> 195
<211> 25
<212> DNA
<213> Yersinia pestis

<400> 195
atggatggtg caatcctggt tgttg 25

<210> 196
<211> 19
<212> DNA
<213> Yersinia pestis

<400> 196
ttgggtcgtc aggttggcg 19

<210> 197
<211> 23
<212> DNA
<213> Yersinia pestis

<400> 197
ctggcgatga tctgccagtt gtt 23

<210> 198
<211> 30
<212> DNA
<213> Yersinia pestis

<400> 198
gctggttacc tggattctta tatcccagaa 30

<210> 199
<211> 33
<212> DNA
<213> Yersinia pestis

<400> 199
actgttaaat ctacttgtac tggcgttgaa atg 33

<210> 200
<211> 31
<212> DNA
<213> Yersinia pestis

<400> 200
aaccacacac tacctttgaa tcagaagttt a 31

<210> 201
<211> 33
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 201
ggtcatgctg actacgttaa aaatatgatc acc 33

<210> 202
<211> 24
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 202
atggatggag cgatcttggt tgtt 24

<210> 203
<211> 29
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 203
acgctgagtg ggaagctaaa attatcgag 29

<210> 204

<211> 28
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 204
ttggcagaag ctctggatag ctatattc 28

<210> 205
<211> 30
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 205
acgattaaaa caacttgtac tggcgttgaa 30

<210> 206
<211> 25
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 206
ggtactaagc gtgacgatgt tcagc 25

<210> 207
<211> 32
<212> DNA
<213> Yersinia pseudotuberculosis

<400> 207
gttgttaacc taattgctcc tatcgcaatg ga 32

<210> 208
<211> 21
<212> DNA
<213> Moraxella catarrhalis

<400> 208

agtttgatgc agaagtatac g 21

<210> 209
<211> 21
<212> DNA
<213> Moraxella catarrhalis

<400> 209
tatcctacta cgtggtacta a 21

<210> 210
<211> 18
<212> DNA
<213> Moraxella catarrhalis

<400> 210
ataacgttga gatgagcg 18

<210> 211
<211> 23
<212> DNA
<213> Streptococcus pyogenes

<400> 211
accttgttga tgacgaagag ttg 23

<210> 212
<211> 27
<212> DNA
<213> Streptococcus pyogenes

<400> 212
aattagttga gatggaaatt cgtgacc 27

<210> 213
<211> 26
<212> DNA

<213> Streptococcus pyogenes

<400> 213
gtattctcaa ttacaggtcg tggtac 26

<210> 214
<211> 21
<212> DNA
<213> Streptococcus pyogenes

<400> 214
gtggtactgt tcgtgtcaac g 21

<210> 215
<211> 23
<212> DNA
<213> Streptococcus pyogenes

<400> 215
aacgtaggta tccttcttcg tgg 23

<210> 216
<211> 21
<212> DNA
<213> Streptococcus pyogenes

<400> 216
cttcgtggtg ttcaacgtga c 21

<210> 217
<211> 22
<212> DNA
<213> Streptococcus pyogenes

<400> 217
ttcaacgtga cgaaatcgaa cg 22

<210> 218
<211> 31
<212> DNA
<213> Streptococcus pyogenes

<400> 218
gtgaagtata tatcctttct aaagacgaag g 31

<210> 219
<211> 22
<212> DNA
<213> Streptococcus pyogenes

<400> 219
tccagcaggt acagaaatgg tt 22

<210> 220
<211> 23
<212> DNA
<213> Streptococcus pyogenes

<400> 220
cgtgacaatc aacgttgagt tga 23

<210> 221
<211> 20
<212> DNA
<213> Streptococcus pneumoniae

<400> 221
cactggtgct gctcaaatgg 20

<210> 222
<211> 24
<212> DNA
<213> Streptococcus pneumoniae

<400> 222
caaggttcag cacttaaagc tctt 24

<210> 223
<211> 27
<212> DNA
<213> Streptococcus pneumoniae

<400> 223
tcttgaaggt gactctaaat acgaaga 27

<210> 224
<211> 26
<212> DNA
<213> Streptococcus pneumoniae

<400> 224
gttgatgagt atatcccaga accaga 26

<210> 225
<211> 22
<212> DNA
<213> Streptococcus pneumoniae

<400> 225
ttccagtcga ggacgtattc tc 22

<210> 226
<211> 21
<212> DNA
<213> Streptococcus pneumoniae

<400> 226
aatcactgga cgtggtacag t 21

<210> 227
<211> 24

<212> DNA
<213> Streptococcus mutans

<400> 227
ttcccaggtg atgatattcc agtt 24

<210> 228
<211> 28
<212> DNA
<213> Streptococcus mutans

<400> 228
gatactgctc aagaagatat catcatgg 28

<210> 229
<211> 23
<212> DNA
<213> Streptococcus mutans

<400> 229
cgttggtatc cgtgatgaca ttc 23

<210> 230
<211> 27
<212> DNA
<213> Streptococcus mutans

<400> 230
gtaaacaatt ggatgaaggt attgcag 27

<210> 231
<211> 24
<212> DNA
<213> Streptococcus mutans

<400> 231
ggataatgtt ggtgttctcc ttcg 24

<210> 232
<211> 28
<212> DNA
<213> Streptococcus mutans

<400> 232
ggttcaattc acccacatac taaattca 28

<210> 233
<211> 22
<212> DNA
<213> Streptococcus mutans

<400> 233
gttcaattga gttgccagca gg 22

<210> 234
<211> 29
<212> DNA
<213> Streptococcus mutans

<400> 234
tcatacacca ttcttcaata actatcgtc 29

<210> 235
<211> 23
<212> DNA
<213> Streptococcus mutans

<400> 235
tcaggacgta ttgatcgtgg tac 23

<210> 236
<211> 25
<212> DNA
<213> Streptococcus sanguis

<400> 236
ggtgttaaac acttgatcgt cttca 25

<210> 237
<211> 21
<212> DNA
<213> Streptococcus sanguis

<400> 237
ccgtaaacag cttgacgaag g 21

<210> 238
<211> 20
<212> DNA
<213> Streptococcus sanguis

<400> 238
tgtgcttctc cgtggtatcc 20

<210> 239
<211> 30
<212> DNA
<213> Streptococcus sanguis

<400> 239
acactaaatt caagggtgaa gtttatatcc 30

<210> 240
<211> 18
<212> DNA
<213> Streptococcus sanguis

<400> 240
aagaaggcgg acgtcaca 18

<210> 241

<211> 27
<212> DNA
<213> Staphylococcus aureus

<400> 241
gttgacatgg ttgacgatga agaatta 27

<210> 242
<211> 21
<212> DNA
<213> Neisseria meningitidis

<400> 242
cgcggttact ttggaagaag g 21

<210> 243
<211> 20
<212> DNA
<213> Neisseria meningitidis

<400> 243
aaactgctgg acgaaggtca 20

<210> 244
<211> 23
<212> DNA
<213> Neisseria meningitidis

<400> 244
tacgtactga gcaaagaaga ggg 23

<210> 245
<211> 19
<212> DNA
<213> Neisseria meningitidis

<400> 245

tcttcgccaa ctaccgtcc 19

<210> 246
<211> 19
<212> DNA
<213> Bordetella pertussis

<400> 246
caagccgggt tcgatcaac 19

<210> 247
<211> 23
<212> DNA
<213> Bordetella pertussis

<400> 247
gaggtgtaca ttctgtccaa gga 23

<210> 248
<211> 21
<212> DNA
<213> Bordetella pertussis

<400> 248
cgttcttcaa cggctatcgt c 21

<210> 249
<211> 20
<212> DNA
<213> Bordetella pertussis

<400> 249
gacaaggaaa tggtgctgcc 20

<210> 250
<211> 20
<212> DNA

<213> Bordetella pertussis

<400> 250
caacgtgggt atcttgctgc 20

<210> 251
<211> 18
<212> DNA
<213> Bordetella pertussis

<400> 251
tgactggccg tatcgagc 18

<210> 252
<211> 30
<212> DNA
<213> Enterococcus faecalis

<400> 252
tggatatggt tgatgacgaa gaattattag 30

<210> 253
<211> 29
<212> DNA
<213> Enterococcus faecalis

<400> 253
gtgacttatt atcagaatac gatttccca 29

<210> 254
<211> 25
<212> DNA
<213> Fusobacterium nucleatum

<400> 254
actgaatatg gattcccagg agatg 25

<210> 255
<211> 26
<212> DNA
<213> Fusobacterium nucleatum

<400> 255
acatccctgt aattagaggt tcatca 26

<210> 256
<211> 31
<212> DNA
<213> Fusobacterium nucleatum

<400> 256
ggtgaagaaa aatggataga aaagataatg g 31

<210> 257
<211> 29
<212> DNA
<213> Fusobacterium nucleatum

<400> 257
acttatggat gcagtagata gctatatcc 29

<210> 258
<211> 26
<212> DNA
<213> Fusobacterium nucleatum

<400> 258
gaaagagcag tagatcaacc attctt 26

<210> 259
<211> 28
<212> DNA
<213> Fusobacterium nucleatum

<400> 259
aggaatcatc aaagttggag aagaaatt 28

<210> 260
<211> 30
<212> DNA
<213> Fusobacterium nucleatum

<400> 260
taaacctaca actaaaacaa cttgtacagg 30

<210> 261
<211> 23
<212> DNA
<213> Fusobacterium nucleatum

<400> 261
cttcttgatc aaggtcaagc agg 23

<210> 262
<211> 30
<212> DNA
<213> Fusobacterium nucleatum

<400> 262
aagaggaact aagaaagaag aagttgaaag 30

<210> 263
<211> 24
<212> DNA
<213> Fusobacterium nucleatum

<400> 263
cttgctaaac caggaagtat ccac 24

<210> 264
<211> 21

<212> DNA
<213> Stenotrophomonas maltophilia

<400> 264
aatacgaatc cccgatccgt c 21

<210> 265
<211> 21
<212> DNA
<213> Stenotrophomonas maltophilia

<400> 265
actgctgagc aagtacgagt t 21

<210> 266
<211> 19
<212> DNA
<213> Stenotrophomonas maltophilia

<400> 266
atcctgaagc tggtcgacg 19

<210> 267
<211> 21
<212> DNA
<213> Stenotrophomonas maltophilia

<400> 267
gatcaaggtt ggcgacgaaa t 21

<210> 268
<211> 20
<212> DNA
<213> Stenotrophomonas maltophilia

<400> 268
tgtacgtcct gtcgaaggac 20

<210> 269
<211> 19
<212> DNA
<213> Burkholderia cepacia

<400> 269
acgccgatca tcaagggtt 19

<210> 270
<211> 19
<212> DNA
<213> Burkholderia cepacia

<400> 270
cgatcatgaa cctggccga 19

<210> 271
<211> 18
<212> DNA
<213> Burkholderia cepacia

<400> 271
acgtcggtat cctgctgc 18

<210> 272
<211> 18
<212> DNA
<213> Burkholderia cepacia

<400> 272
gaagccgggt tcgatcac 18

<210> 273
<211> 22
<212> DNA
<213> Burkholderia cepacia

<400> 273
cgttcttcaa caactaccgt cc 22

<210> 274
<211> 17
<212> DNA
<213> Burkholderia cepacia

<400> 274
gtggtgacgg gtcgtgt 17

<210> 275
<211> 17
<212> DNA
<213> Burkholderia cepacia

<400> 275
cgtacatccc gacgccg 17

<210> 276
<211> 22
<212> DNA
<213> Corynebacterium diphtheriae

<400> 276
aacaagtgcg acatggttga tg 22

<210> 277
<211> 22
<212> DNA
<213> Corynebacterium diphtheriae

<400> 277
aagaggctcc aatcatccac at 22

<210> 278

<211> 20
<212> DNA
<213> Corynebacterium diphtheriae

<400> 278
gaagtggacc cagtccatca 20

<210> 279
<211> 20
<212> DNA
<213> Corynebacterium diphtheriae

<400> 279
ctgaaggtca acgaggacgt 20

<210> 280
<211> 19
<212> DNA
<213> Corynebacterium diphtheriae

<400> 280
ttctcgacta caccgaggc 19

<210> 281
<211> 22
<212> DNA
<213> Corynebacterium diphtheriae

<400> 281
gttgttgtta agccaggcgc tt 22

<210> 282
<211> 21
<212> DNA
<213> Corynebacterium diphtheriae

<400> 282

gagttcgagg gctctgtcta c 21

<210> 283
<211> 23
<212> DNA
<213> Chlamydophila pneumoniae

<400> 283
cgcttctcac gttgaatacg aaa 23

<210> 284
<211> 21
<212> DNA
<213> Chlamydophila pneumoniae

<400> 284
acatatcttg ctagctcgcc a 21

<210> 285
<211> 29
<212> DNA
<213> Chlamydophila pneumoniae

<400> 285
tagatatgat ctctcaagaa gatgctgaa 29

<210> 286
<211> 24
<212> DNA
<213> Chlamydophila pneumoniae

<400> 286
aagaaaaagg ctacaaagga tgcc 24

<210> 287
<211> 30
<212> DNA

<213> Chlamydophila pneumoniae

<400> 287
ttgaaggtga tgcaaattat atcgaaaaag 30

<210> 288
<211> 23
<212> DNA
<213> Chlamydophila pneumoniae

<400> 288
ttcgagaact tatgcaagct gtg 23

<210> 289
<211> 30
<212> DNA
<213> Chlamydophila pneumoniae

<400> 289
tggaatcgtt aaagtttctg ataaagttca 30

<210> 290
<211> 28
<212> DNA
<213> Chlamydophila pneumoniae

<400> 290
ggagagacta aagaaacaat cgttactg 28

<210> 291
<211> 26
<212> DNA
<213> Chlamydophila pneumoniae

<400> 291
gaaaacgttg gtttactcct cagagg 26

<210> 292
<211> 24
<212> DNA
<213> Chlamydophila pneumoniae

<400> 292
ttgaaagagg tatggtggtt tgtc 24

<210> 293
<211> 26
<212> DNA
<213> Chlamydophila pneumoniae

<400> 293
cagctgttta cgttcttcag aaagaa 26

<210> 294
<211> 23
<212> DNA
<213> Chlamydophila pneumoniae

<400> 294
caggagtcgt aactcttcct gaa 23

<210> 295
<211> 27
<212> DNA
<213> Legionella pneumophila

<400> 295
atgaccctga gttattagag ttagtgg 27

<210> 296
<211> 26
<212> DNA
<213> Legionella pneumophila

<400> 296
gtgcgagatt tattaagcag ttacga 26

<210> 297
<211> 22
<212> DNA
<213> Legionella pneumophila

<400> 297
tgttggttca gctttgaaag ca 22

<210> 298
<211> 26
<212> DNA
<213> legionella pneumophila

<400> 298
ttggaaggtg aagacagtga tatagg 26

<210> 299
<211> 26
<212> DNA
<213> Legionella pneumophila

<400> 299
ttcatacatt cctgagccag ttagaa 26

<210> 300
<211> 19
<212> DNA
<213> Legionella pneumophila

<400> 300
tttctggacg cggaacagt 19

<210> 301
<211> 26

<212> DNA
<213> Legionella pneumophila

<400> 301
aagttggtga ggaagttgaa attgtt 26

<210> 302
<211> 20
<212> DNA
<213> Legionella pneumophila

<400> 302
gatgaaggtc gagctggtga 20

<210> 303
<211> 26
<212> DNA
<213> Legionella pneumophila

<400> 303
ggataccgtc cacaattcta tttcag 26

<210> 304
<211> 22
<212> DNA
<213> Legionella pneumophila

<400> 304
catcaagcca cacaccaagt tt 22

<210> 305
<211> 28
<212> DNA
<213> Legionella longbeachiae

<400> 305
aggtacaatt aagcctcaca ctaaattt 28

<210> 306
<211> 30
<212> DNA
<213> Legionella longbeachiae

<400> 306
cggaagtata cgtgttatct aaagaagaag 30

<210> 307
<211> 31
<212> DNA
<213> Legionella longbeachiae

<400> 307
atagaccaca gttttatttt agaacaacag a 31

<210> 308
<211> 26
<212> DNA
<213> Legionella longbeachiae

<400> 308
tatagcaaaa gcatacgatc agatcg 26

<210> 309
<211> 22
<212> DNA
<213> Legionella longbeachiae

<400> 309
gaggcatcac aatctcaaca gc 22

<210> 310
<211> 21
<212> DNA
<213> Legionella longbeachiae

<400> 310
tattggggtg ccagcaatag a 21

<210> 311
<211> 29
<212> DNA
<213> Legionella longbeachiae

<400> 311
attcttatat tccagagcct gtaagaaac 29

<210> 312
<211> 29
<212> DNA
<213> Legionella longbeachiae

<400> 312
attgacaaat cgttcttatt accgattga 29

<210> 313
<211> 25
<212> DNA
<213> Legionella longbeachiae

<400> 313
aattcgtgat actgcaaaga ctacc 25

<210> 314
<211> 26
<212> DNA
<213> Legionella longbeachiae

<400> 314
ggtatattac tccgtggaac aaaacg 26

<210> 315

<211> 23
<212> DNA
<213> Legionella longbeachiae

<400> 315
agttattaga cgaaggacgt gca 23

<210> 316
<211> 24
<212> DNA
<213> Haemophilus influenzae

<400> 316
tagtagtagc agcaacagat ggtc 24

<210> 317
<211> 25
<212> DNA
<213> Haemophilus influenzae

<400> 317
atcttattag gtcgccaagt aggtg 25

<210> 318
<211> 30
<212> DNA
<213> Haemophilus influenzae

<400> 318
ccatacatca tcgtattctt aaacaaatgc 30

<210> 319
<211> 21
<212> DNA
<213> Haemophilus influenzae

<400> 319

cgttaaacgg cgtagcagaa t 21

<210> 320
<211> 27
<212> DNA
<213> Haemophilus influenzae

<400> 320
aaatccttga gttagcaaac cacttag 27

<210> 321
<211> 20
<212> DNA
<213> Haemophilus influenzae

<400> 321
aaccagaacg tgcgattgac 20

<210> 322
<211> 23
<212> DNA
<213> Moraxella catarrhalis

<400> 322
cgactctgca cctgaagaaa aag 23

<210> 323
<211> 30
<212> DNA
<213> Moraxella catarrhalis

<400> 323
ggtattacta ttaacacatc tcacatcgag 30

<210> 324
<211> 24
<212> DNA

<213> Moraxella catarrhalis

<400> 324
tatgacacag aagctcgtca ctac 24

<210> 325
<211> 21
<212> DNA
<213> Moraxella catarrhalis

<400> 325
tggttgtttc tgcaactgac g 21

<210> 326
<211> 22
<212> DNA
<213> Moraxella catarrhalis

<400> 326
atcctactgt ctcgtcaggt tg 22

<210> 327
<211> 29
<212> DNA
<213> Moraxella catarrhalis

<400> 327
atatatcatg gtattcatga acaagtgcg 29

<210> 328
<211> 26
<212> DNA
<213> Moraxella catarrhalis

<400> 328
gcattgaatg gttctgacgg taaata 26

<210> 329
<211> 25
<212> DNA
<213> Moraxella catarrhalis

<400> 329
tgctgttcta gaactactag acacg 25

<210> 330
<211> 20
<212> DNA
<213> Mycoplasma pneumoniae

<400> 330
attaactccg ctcacgtgga 20

<210> 331
<211> 26
<212> DNA
<213> Mycoplasma pneumoniae

<400> 331
ttctagtagt ttcagcaact gacagt 26

<210> 332
<211> 24
<212> DNA
<213> Mycoplasma pneumoniae

<400> 332
ttcctaaaca agtgtgacat tgca 24

<210> 333
<211> 27
<212> DNA
<213> Mycoplasma pneumoniae

<400> 333
actgatgaag aagtgcaaga gttagta 27

<210> 334
<211> 22
<212> DNA
<213> Mycoplasma pneumoniae

<400> 334
ttacggcttt gatggcaaga ac 22

<210> 335
<211> 23
<212> DNA
<213> Mycoplasma pneumoniae

<400> 335
ttatggttct gcacttaaag cgc 23

<210> 336
<211> 25
<212> DNA
<213> Mycoplasma pneumoniae

<400> 336
tgatcctaag tgggaagcta agatc 25

<210> 337
<211> 30
<212> DNA
<213> Mycoplasma pneumoniae

<400> 337
gatttaatga atgcagttga tgaatggatt 30

<210> 338
<211> 24

<212> DNA
<213> Mycoplasma pneumoniae

<400> 338
gaacgtggtg aattgaaagt aggt 24

<210> 339
<211> 24
<212> DNA
<213> Mycoplasma pneumoniae

<400> 339
gtccaatccg taaagcagtt gtta 24

<210> 340
<211> 23
<212> DNA
<213> Mycoplasma pneumoniae

<400> 340
aaggaacttg attcagcaat ggc 23

<210> 341
<211> 22
<212> DNA
<213> Mycoplasma pneumoniae

<400> 341
gtggaccgta aagaagtgga ac 22

<210> 342
<211> 23
<212> DNA
<213> Mycoplasma pneumoniae

<400> 342
taaaccaggt tcgattaaac cgc 23

<210> 343
<211> 25
<212> DNA
<213> Mycoplasma pneumoniae

<400> 343
tactacagac gttactggtt cgatt 25

<210> 344
<211> 20
<212> DNA
<213> Mycoplasma pneumoniae

<400> 344
gctgctactc gttacgacca 20

<210> 345
<211> 15
<212> DNA
<213> Mycobacterium tuberculosis

<400> 345
acaacgcccc cgagg 15

<210> 346
<211> 14
<212> DNA
<213> Mycobacterium tuberculosis

<400> 346
acgcccctgg ccac 14

<210> 347
<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 347
gatggacggt gcgatcct 18

<210> 348
<211> 16
<212> DNA
<213> Mycobacterium tuberculosis

<400> 348
cccgatgccc cagacc 16

<210> 349
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 349
gagcacgttc tgctggc 17

<210> 350
<211> 19
<212> DNA
<213> Mycobacterium tuberculosis

<400> 350
atcctggtag cgctgaaca 19

<210> 351
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 351
gacgcagtgg acgacga 17

<210> 352

<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 352
ggctgcccag gaattcga 18

<210> 353
<211> 15
<212> DNA
<213> Mycobacterium tuberculosis

<400> 353
ccccggttgt gcggg 15

<210> 354
<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 354
tgggttgcct ctgtcgag 18

<210> 355
<211> 19
<212> DNA
<213> Mycobacterium tuberculosis

<400> 355
aggaactgat gaacgcggt 19

<210> 356
<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 356

cgagtcgatt ccggaccc 18

<210> 357
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 357
cattaccggc cgcggaa 17

<210> 358
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 358
gtcaccggac gtgtgga 17

<210> 359
<211> 21
<212> DNA
<213> Mycobacterium tuberculosis

<400> 359
acgaggaagt tgagatcgtc g 21

<210> 360
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 360
ttcgcccatc gaccacc 17

<210> 361
<211> 19
<212> DNA

<213> Mycobacterium tuberculosis

<400> 361
acaacgttgg tttgctgct 19

<210> 362
<211> 18
<212> DNA
<213> Mycobacterium tuberculosis

<400> 362
cgagttcgaa ggccaggt 18

<210> 363
<211> 17
<212> DNA
<213> Mycobacterium tuberculosis

<400> 363
gtggtgacac tgccgga 17

<210> 364
<211> 22
<212> DNA
<213> Mycobacterium tuberculosis

<400> 364
gacaaattcc ccgatctgaa cg 22

<210> 365
<211> 20
<212> DNA
<213> Mycobacterium tuberculosis

<400> 365
aacgagacga aggcattcga 20

<210> 366
<211> 17
<212> DNA
<213> Serratia marcescens

<400> 366
gttctgcacg tgctttc 17

<210> 367
<211> 18
<212> DNA
<213> Serratia marcescens

<400> 367
ggttggcgtt cctttcat 18

<210> 368
<211> 16
<212> DNA
<213> Serratia marcescens

<400> 368
acctgccagt gatccg 16

<210> 369
<211> 17
<212> DNA
<213> Serratia marcescens

<400> 369
tggacagcta catccca 17

<210> 370
<211> 20
<212> DNA
<213> Serratia marcescens

<400> 370
agacaccgtt aagtctacct 20

<210> 371
<211> 19
<212> DNA
<213> Serratia marcescens

<400> 371
taaaccaggc tccatcaag 19

<210> 372
<211> 19
<212> DNA
<213> Serratia marcescens

<400> 372
tcctgagcaa agatgaagg 19

<210> 373
<211> 17
<212> DNA
<213> Serratia marcescens

<400> 373
gaaggtggtc gtcacac 17

<210> 374
<211> 19
<212> DNA
<213> Serratia marcescens

<400> 374
gacaacgtga acatggttg 19

<210> 375
<211> 16

<212> DNA
<213> Serratia marcescens

<400> 375
acgtgaccgg taccat 16

<210> 376
<211> 17
<212> DNA
<213> Serratia marcescens

<400> 376
tcctggtagt agctgcg 17

<210> 377
<211> 19
<212> DNA
<213> Serratia marcescens

<400> 377
caaagttggc gaagaagtt 19

<210> 378
<211> 17
<212> DNA
<213> Klebsiella pneumoniae

<400> 378
atcctggttg ttgctgc 17

<210> 379
<211> 18
<212> DNA
<213> Klebsiella pneumoniae

<400> 379
gtaggcgttc cgtacatc 18

<210> 380
<211> 16
<212> DNA
<213> Klebsiella pneumoniae

<400> 380
acaccccgat cgttcg 16

<210> 381
<211> 20
<212> DNA
<213> Klebsiella pneumoniae

<400> 381
cctggatacc tatatcccgg 20

<210> 382
<211> 22
<212> DNA
<213> Klebsiella pneumoniae

<400> 382
atcaaagtag gtgaagaagt tg 22

<210> 383
<211> 16
<212> DNA
<213> Klebsiella pneumoniae

<400> 383
accgcgaaaa ccacct 16

<210> 384
<211> 16
<212> DNA
<213> Klebsiella pneumoniae

<400> 384
taagccgggc accatc 16

<210> 385
<211> 18
<212> DNA
<213> Klebsiella pneumoniae

<400> 385
tcctgtccaa agacgaag 18

<210> 386
<211> 16
<212> DNA
<213> Klebsiella pneumoniae

<400> 386
gaaggcggcc gtcata 16

<210> 387
<211> 19
<212> DNA
<213> Klebsiella pneumoniae

<400> 387
ggcgacaaca tcaaaatgg 19

<210> 388
<211> 17
<212> DNA
<213> MICRO-ORGANISM

<220>
<221> variation
<222> (3)..(3)
<223> r can be a or g

<220>
<221> variation
<222> (8)..(8)
<223> w can be a or t

<220>
<221> variation
<222> (9)..(9)
<223> y can be c or t

<220>
<221> variation
<222> (13)..(13)
<223> m can be a or c

<220>
<221> variation
<222> (15)..(15)
<223> h can be a or c or t

<400> 388
carttctwyt tcmghac 17

<210> 389
<211> 20
<212> DNA
<213> MICRO-ORGANISM

<220>
<221> variation
<222> (3)..(3)
<223> r can be a or g

<220>
<221> variation
<222> (6)..(6)

<223> h can be a or c or t

<220>
<221> variation
<222> (9)..(9)
<223> r can be a or g

<220>
<221> variation
<222> (12)..(12)
<223> y can be c or t

<220>
<221> variation
<222> (15)..(15)
<223> h can be a or c or t

<220>
<221> variation
<222> (18)..(18)
<223> k can be g or t

<400> 389
aargahgarg gygghcgkca 20

<210> 390
<211> 17
<212> DNA
<213> MICRO-ORGANISM

<220>
<221> variation
<222> (3)..(3)
<223> r can be a or g

<220>
<221> variation

<222> (9)..(9)
<223> n can be a or c or g or t

<220>
<221> misc_feature
<222> (9)..(9)
<223> n is a, c, g, or t

<220>
<221> variation
<222> (10)..(10)
<223> m can be a or c

<220>
<221> variation
<222> (12)..(12)
<223> r can be a or g

<220>
<221> variation
<222> (15)..(15)
<223> n can be a or g or c or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 390
garatggtnm trccngg 17

<210> 391
<211> 21
<212> DNA
<213> MICRO-ORGANISM

<400> 391

agtttgatgc agaagtatac g 21

<210> 392
<211> 21
<212> DNA
<213> MICRO-ORGANISM

<400> 392
tatcctacta cgtggtacta a 21

<210> 393
<211> 18
<212> DNA
<213> MICRO-ORGANISM

<400> 393
ataacgttga gatgagcg 18

<210> 394
<211> 20
<212> DNA
<213> artificial

<220>
<223> Primer directed at all tuf genes

<220>
<221> variation
<222> (12)..(12)
<223> y can be c or t

<220>
<221> variation
<222> (15)..(15)
<223> y can be c or t

<220>

<221> variation
<222> (18)..(18)
<223> r can be a or g

<400> 394
cacgttgacc ayggyaarac 20

<210> 395
<211> 20
<212> DNA
<213> artificial

<220>
<223> Primer directed at all tuf genes

<220>
<221> variation
<222> (3)..(3)
<223> d can be a or g or t

<220>
<221> variation
<222> (6)..(6)
<223> d can be a or g or t

<220>
<221> variation
<222> (9)..(9)
<223> r can be a or g

<220>
<221> variation
<222> (12)..(12)
<223> y can be c or t

<220>
<221> variation

<222> (15)..(15)
<223> w can be a or t

<220>
<221> variation
<222> (18)..(18)
<223> r can be a or g

<220>
<221> variation
<222> (18)..(18)
<223> r can be a or g

<400> 395
gtdcgdccrc cytcwcgrat 20

<210> 396
<211> 20
<212> DNA
<213> ARTIFICIAL

<220>
<223> Primer directed at all tuf genes

<220>
<221> misc_feature
<222> (1)..(1)
<223> k can be g or t

<220>
<221> misc_feature
<222> (2)..(2)
<223> y can be c or t

<220>
<221> misc_feature

<222> (5)..(5)
<223> i stands for deoxyinosine - a base that binds to a, g, c or t

<220>
<221> misc_feature
<222> (8)..(8)
<223> h can be a, c or t

<220>
<221> misc_feature
<222> (11)..(11)
<223> b can be g, c or t

<220>
<221> misc_feature
<222> (14)..(14)
<223> r can be a or g

<400> 396
kyacigghgt bgaratgttc 20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2006/001056

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12Q 1/04 (2006.01) ***C12Q 1/14*** (2006.01)
C12Q 1/10 (2006.01) ***C12Q 1/68*** (2006.01)

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)
SEE BELOW

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

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

Databases: DGENE, CAPLUS, USPATFULL, BIOSIS, GENBANK (SEQ ID NO: 394-396)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2001/023604 A (INFECTIO DIAGNOSTIC INC) 5 April 2001 Claim 1, 22; SEQ ID 561; p. 12-p. 19	15, 17, 18 and 21
A	US 6,489,110 B (OUDSHOORN, P. et al.) 3 December 2002 Whole document	15, 17, 18 and 21

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 September 2006	- 9 OCT 2006
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer **Sophina Calanni** Telephone No : (02) 6283 2038

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2006/001056

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Martineau, F. et al., 2001, Development of a PCR assay for identification of staphylococci at genus and species levels, *Journal of Clinical Microbiology*, **39**(7):2541-2547	15, 17, 18 and 21
A	Ke, D. et al., 1999, Development of a PCR assay for rapid detection of enterococci, *Journal of Clinical Microbiology*, **37**(11): 3497-3503	15, 17, 18 and 21
A	Ventura, M. et al., 2003, Analysis, characterization, and loci of the *tuf* genes in *Lactobacillus* and *Bifidobacterium* species and their direct application for species identification, *Applied and Environmental Microbiology*, **69**(11): 6908-6922	15, 17, 18 and 21
A	Luneberg, E. et al, 1993, Detection of *Mycoplasma pneumoniae* by polymerase chain reaction and nonradioactive hybridization in microtiter plates, *Journal of Clinical Microbiology*, **31**(5): 1088-1094	15, 17, 18 and 21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2006/001056

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: **15, 17, 18 and 21 (partially)**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claims 15 and 21 refer to primers selected from the group consisting of SEQ ID NO: 394-396 or *equivalents thereof*, provided said equivalents specifically amplify the *tuf* gene variable region or a part thereof. No meaningful international search could be carried out with respect to primers that are *equivalent to* the primers represented by SEQ ID NO: 394-396 because the scope of the subject matter claimed could not be determined. The present search is therefore limited to primers comprising SEQ ID NO: 394-396.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

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

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

See Supplemental Box

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention covered by claims Nos.: **15, 17, 18 and 21 (all partially),** as nominated by the applicant.

That is, a method for the detection or identification of at least one microorganism or for the simultaneous detection or identification of several microorganisms in a test sample comprising amplifying at least a portion of a *tuf* gene variable region with at least one primer pair, wherein at least one of the primers is selected from the group consisting of SEQ ID NO: 394-396. The search also covers a composition comprising at least one primer selected from the group consisting essentially of SEQ ID NO: 394-396.

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Supplemental Box I
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: III

More than one invention is claimed. Rule 13.1 of the PCT requires that an International Application should relate to one invention only, or if there is more than one invention, that the inclusion of those inventions in one International Application is only permitted if all inventions are so linked as to form a single general inventive concept.

Rule 13.2 of the PCT defines the method for determining whether the requirement of unity of invention is satisfied in respect of a group of inventions claimed in an International application. Unity of invention exists only when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding "special technical features". The expression "special technical features" is defined in Rule 13.2 as meaning those technical features that define a contribution that each of the inventions, considered as a whole, makes over the prior art. The determination is made on the contents of the claims as interpreted in light of the description and any drawings.

The problem addressed by the present application is the identification of microorganisms in a sample. The specification describes a method of identification based on hybridisation of nucleic acids isolated from the microorganism with one or more probes derived from the *tuf* gene variable regions. The specification discloses 382 probes (SEQ ID NO: 12-393) directed to the *tuf* gene variable region of 40 distinct species of bacteria. Three primers (SEQ ID NO: 394-396) that enable amplification of the *tuf* gene variable region are also disclosed.

The International Searching Authority has found that there are different inventions as follows:

- Invention 1: Claims 1-6, 11-14, 16-18, 23, 25 relating to a method for the identification of a microorganism by contacting nucleic acid molecules isolated from the microorganism with one or more probes derived from the *tuf* gene variable region, wherein the presence or absence of hybridisation is indicative of the presence or absence of one or more microorganisms.

  It is considered that the method described represents a "first special technical feature".

  Inventions 2-385: Claims 7-10, 15, 19-22, 24, 26-29. These claims relate to a composition comprising at least one probe or primer selected from the group consisting of SEQ ID NO: 12-396, a kit and microarray comprising said probes and/or primers, and various methods that make use of said probes and/or primers.

  It is considered that the probes and/or primers set forth in SEQ ID NO: 12-396 each represent a distinct "special technical feature".

It is noted that the common concept linking these inventions resides in the use of probes and/or primers directed to the variable region of the *tuf* gene. However, such primers and probes exist in the prior art, and have been used for the purpose of identifying different species of bacteria (see the citations listed below). Therefore, this common concept does not constitute a "special technical feature" as it makes no contribution over the prior art. Because the common feature does not satisfy the requirement for being a special technical feature it follows that it cannot provide the necessary technical relationship between the identified inventions. Therefore the claims do not satisfy the requirement of unity of invention *a posteriori*.

US 6489110 (Oudshoorn, P. et al.) 3 December 2002
Martineau, F. et al., 2001, *Journal of Clinical Microbiology*, **39**(7):2541-2547
Ke, D. et al., 1999, *Journal of Clinical Microbiology*, **37**(11): 3497-3503
Ventura, M. et al., 2003, *Applied and Environmental Microbiology*, **69**(11): 6908-6922
Luneberg, E. et al, 1993, *Journal of Clinical Microbiology*, **31**(5): 1088-1094

Supplemental Box II
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Supplemental Box I

In response to the request to pay additional fees the applicant has requested that the International Search Report is conducted on the basis of Claim 15. Claim 15 defines a method for the detection or identification of at least one microorganism or for the simultaneous detection or identification of several microorganisms in a test sample comprising amplifying at least a portion of a *tuf* gene variable region with at least one primer pair, wherein at least one of the primers is selected from the group consisting of SEQ ID NO: 394-396. The examiner notes that as indicated above SEQ ID NO: 394-396 may be considered separate inventions. However, it is considered that the search and examination of all three sequences will require only negligible additional search and examination effort over that for only one of the sequences. As such, the present search report covers the invention as defined in claims 15, 17, 18 and 21.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU2006/001056

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	0123604	AU	76360/00	BR	0014370	CA	2283458
		CA	2307010	CA	2388461	EP	1246935
US	6489110	AU	27176/99	CA	2318726	EP	1049809
		WO	9937804	ZA	9900362		

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX