



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁷ : C12N 15/31, C07K 14/21, A61K 38/16, C07K 16/12, G01N 33/53</p>	A1	<p>(11) International Publication Number: WO 00/09694</p> <p>(43) International Publication Date: 24 February 2000 (24.02.00)</p>
<p>(21) International Application Number: PCT/EP99/05828</p> <p>(22) International Filing Date: 11 August 1999 (11.08.99)</p> <p>(30) Priority Data: 9817824.7 14 August 1998 (14.08.98) GB</p> <p>(71) Applicant (for all designated States except US): SMITHK-LINE BEECHAM BIOLOGICALS S.A. [BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).</p> <p>(72) Inventor; and (75) Inventor/Applicant (for US only): THONNARD, Joelle [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE).</p> <p>(74) Agent: PRIVETT, Kathryn, Louise; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
<p>(54) Title: CLONING OF BASB023 ANTIGEN FROM MORAXELLA CATARRHALIS</p>		
<p>(57) Abstract</p> <p>The invention provides BASB023 polypeptides and polynucleotides encoding BASB023 polypeptides and methods for producing such polypeptides by recombinant techniques. Also provided are diagnostic, prophylactic and therapeutic uses.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakistan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CLONING OF BASB023 ANTIGEN FROM MORAXELLA CATARRHALIS

FIELD OF THE INVENTION

This invention relates to polynucleotides, (herein referred to as “BASB023
5 polynucleotide(s)”), polypeptides encoded by them (referred to herein as “BASB023” or
“BASB023 polypeptide(s)”), recombinant materials and methods for their production. In
another aspect, the invention relates to methods for using such polypeptides and
polynucleotides, including vaccines against bacterial infections. In a further aspect, the
invention relates to diagnostic assays for detecting infection of certain pathogens.

10

BACKGROUND OF THE INVENTION

Moraxella catarrhalis (also named *Branhamella catarrhalis*) is a Gram negative bacteria
frequently isolated from the human upper respiratory tract. It is responsible for several
pathologies the main ones being otitis media in infants and children, and pneumonia in
15 elderlies. It is also responsible of sinusitis, nosocomial infections and less frequently of
invasive diseases.

Otitis media is an important childhood disease both by the number of cases and its potential
sequelae. More than 3.5 millions cases are recorded every year in the United States, and it is
20 estimated that 80 % of the children have experienced at least one episode of otitis before
reaching the age of 3 (Klein, JO (1994) Clin.Inf.Dis 19:823). Left untreated, or becoming
chronic, this disease may lead to hearing losses that could be temporary (in the case of fluid
accumulation in the middle ear) or permanent (if the auditive nerve is damaged). In infants,
such hearing losses may be responsible for a delayed speech learning.

25

Three bacterial species are primarily isolated from the middle ear of children with otitis
media: *Streptococcus pneumoniae*, non typeable *Haemophilus influenza* (NTHi) and *M.*
catarrhalis. They are present in 60 to 90 % of the cases. A review of recent studies shows
that *S. pneumoniae* and NTHi represent both about 30 %, and *M. catarrhalis* about 15 % of

the otitis media cases (Murphy, TF (1996) Microbiol.Rev. 60:267). Other bacteria could be isolated from the middle ear (*H. influenza* type B, *S. pyogenes* etc) but at a much lower frequency (2 % of the cases or less).

5 Epidemiological data indicate that, for the pathogens found in the middle ear, the colonization of the upper respiratory tract is an absolute prerequisite for the development of an otitis; other are however also required to lead to the disease (Dickinson, DP et al. (1988) J. Infect.Dis. 158:205, Faden, HL et al. (1991) Ann.Otorhinol.Laryngol. 100:612). These are important to trigger the migration of the bacteria into the middle ear via the Eustachian
10 tubes, followed by the initiation of an inflammatory process. These factors are unknown todate. It has been postulated that a transient anomaly of the immune system following a viral infection, for example, could cause an inability to control the colonization of the respiratory tract (Faden, HL et al (1994) J. Infect.Dis. 169:1312). An alternative explanation is that the exposure to environmental factors allow a more important colonization of some
15 children, who subsequently become susceptible to the development of otitis media because of the sustained presence of middle ear pathogens (Murphy, TF (1996) Microbiol.Rev. 60:267).

The immune response to *M. catarrhalis* is poorly characterized. The analysis of strains
20 isolated sequentially from the nasopharynx of babies followed from 0 to 2 years of age, indicates that they get and eliminate frequently new strains. This indicates that an efficacious immune response against this bacteria is mounted by the colonized children (Faden, HL et al (1994) J. Infect.Dis. 169:1312).

25 In most adults tested, bactericidal antibodies have been identified (Chapman, AJ et al. (1985) J. Infect.Dis. 151:878). Strains of *M. catarrhalis* present variations in their capacity to resist serum bactericidal activity: in general, isolates from diseased individuals are more resistant than those who are simply colonized (Hol, C et al. (1993) Lancet 341:1281, Jordan, KL et al. (1990) Am.J.Med. 88 (suppl. 5A):28S). Serum resistance could therefore be

considered as a virulence factor of the bacteria. An opsonizing activity has been observed in the sera of children recovering from otitis media.

5 The antigens targetted by these different immune responses in humans have not been identified, with the exception of OMP B1, a 84 kDa protein which expression is regulated by iron, and that is recognized by the sera of patients with pneumonia (Sethi, S, et al. (1995) Infect.Immun. 63:1516) , and of UspA1 and UspA2 (Chen D. et al.(1999), Infect.Immun. 67:1310).

10 A few other membrane proteins present on the surface of *M. catarrhalis* have been characterized using biochemical method, or for their potential implication in the induction of a protective immunity (for review, see Murphy, TF (1996) Microbiol.Rev. 60:267). In a mouse pneumonia model, the presence of antibodies raised against some of them (UspA, CopB) favors a faster clearance of the pulmonary infection. Another polypeptide (OMP CD)
15 is highly conserved among *M. catarrhalis* strains, and presents homologies with a porin of *Pseudomonas aeruginosa*, which has been demonstrated efficacious against this bacterium in animal models.

The frequency of *Moraxella catarrhalis* infections has risen dramatically in the past few
20 decades. This has been attributed to the emergence of multiply antibiotic resistant strains and an increasing population of people with weakened immune systems. It is no longer uncommon to isolate *Moraxella catarrhalis* strains that are resistant to some or all of the standard antibiotics. This phenomenon has created an unmet medical need and demand for new anti-microbial agents, vaccines, drug screening methods, and diagnostic tests for this
25 organism.

SUMMARY OF THE INVENTION

- The present invention relates to BASB023, in particular BASB023 polypeptides and BASB023 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including prevention and treatment of microbial diseases, amongst others.
- 5 In a further aspect, the invention relates to diagnostic assays for detecting diseases associated with microbial infections and conditions associated with such infections, such as assays for detecting expression or activity of BASB023 polynucleotides or polypeptides.
- 10 Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

DESCRIPTION OF THE INVENTION

- 15 The invention relates to BASB023 polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of BASB023 of *Moraxella catarrhalis*, which is related by amino acid sequence homology to *Legionella adelaidensis* macrophage infectivity potentiator polypeptide. The invention relates especially to BASB023 having the nucleotide and amino acid sequences set out in
- 20 SEQ ID NO:1 or 3 and SEQ ID NO:2 or 4 respectively. It is understood that sequences recited in the Sequence Listing below as "DNA" represent an exemplification of one embodiment of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

25

Polypeptides

In one aspect of the invention there are provided polypeptides of *Moraxella catarrhalis* referred to herein as "BASB023" and "BASB023 polypeptides" as well as biologically,

diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

The present invention further provides for:

- 5 (a) an isolated polypeptide which comprises an amino acid sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, most preferably at least 97-99% or exact identity, to that of SEQ ID NO:2 or 4;
- (b) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more
10 preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1 or 3 over the entire length of SEQ ID NO:1 or 3 respectively; or
- (c) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90%
15 identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity, to the amino acid sequence of SEQ ID NO:2 or 4.

The BASB023 polypeptides provided in SEQ ID NO:2 or 4 are the BASB023 polypeptides from *Moraxella catarrhalis* strains Mc2931 (ATCC 43617).

- 20 The invention also provides an immunogenic fragment of a BASB023 polypeptide, that is, a contiguous portion of the BASB023 polypeptide which has the same or substantially the same immunogenic activity as the polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 4; That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the BASB023 polypeptide.
- 25 Such an immunogenic fragment may include, for example, the BASB023 polypeptide lacking an N-terminal leader sequence, and/or a transmembrane domain and/or a C-terminal anchor domain. In a preferred aspect the immunogenic fragment of BASB023 according to the invention comprises substantially all of the extracellular domain of a polypeptide which has at least 85% identity, preferably at least 90% identity, more

preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2 or 4 over the entire length of SEQ ID NO:2 or 4.

- 5 A fragment is a polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with BASB023 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.
- 10 Preferred fragments include, for example, truncation polypeptides having a portion of an amino acid sequence of SEQ ID NO:2 or 4 or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, are also preferred. Further preferred are fragments characterized by structural or functional
- 15 attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.
- 20 Further preferred fragments include an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO:2 or 4, or an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids truncated or
- 25 deleted from the amino acid sequence of SEQ ID NO:2 or 4.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these fragments

may be employed as intermediates for producing the full-length polypeptides of the invention.

5 Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

10 The polypeptides, or immunogenic fragments, of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

15 In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa.

20

25 Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or
5 assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

Fusion partners include protein D from *Haemophilus influenzae* and the non-structural
10 protein from influenzae virus, NS1 (hemagglutinin). Another fusion partner is the protein known as LytA. Preferably the C terminal portion of the molecule is used. LytA is derived from *Streptococcus pneumoniae* which synthesizes an N-acetyl-L-alanine amidase, amidase LytA, (coded by the *lytA* gene {Gene, 43 (1986) page 265-272}) an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The
15 C-terminal domain of the LytA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of E.coli C-LytA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LytA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible
20 to use the repeat portion of the LytA molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions,
25 whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

It is most preferred that a polypeptide of the invention is derived from *Moraxella catarrhalis*, however, it may preferably be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Polynucleotides

It is an object of the invention to provide polynucleotides that encode BASB023 polypeptides, particularly polynucleotides that encode the polypeptide herein designated BASB023.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding BASB023 polypeptides comprising a sequence set out in SEQ ID NO:1 or 3 which includes a full length gene, or a variant thereof.

The BASB023 polynucleotides provided in SEQ ID NO:1 or 3 are the BASB023 polynucleotides from *Moraxella catarrhalis* strains Mc2931 (ATCC 43617).

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing BASB023 polypeptides and polynucleotides, particularly *Moraxella catarrhalis* BASB023 polypeptides and polynucleotides, including, for example, unprocessed RNAs, ribozyme RNAs, mRNAs, cDNAs, genomic DNAs, B- and Z-DNAs. Further embodiments of the invention include biologically,

diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions comprising the same.

Another aspect of the invention relates to isolated polynucleotides, including at least one full
5 length gene, that encodes a BASB023 polypeptide having a deduced amino acid sequence of SEQ ID NO:2 or 4 and polynucleotides closely related thereto and variants thereof.

In another particularly preferred embodiment of the invention there is a BASB023 polypeptide from *Moraxella catarrhalis* comprising or consisting of an amino acid
10 sequence of SEQ ID NO:2 or 4 or a variant thereof.

Using the information provided herein, such as a polynucleotide sequence set out in SEQ ID NO:1 or 3, a polynucleotide of the invention encoding BASB023 polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and
15 sequencing chromosomal DNA fragments from bacteria using *Moraxella catarrhalis* Catlin cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as a polynucleotide sequence given in SEQ ID NO:1 or 3, typically a library of clones of chromosomal DNA of *Moraxella catarrhalis* Catlin in *E.coli* or some other suitable host is probed with a radiolabeled
20 oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both
25 directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By

Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence. Illustrative of the invention, each polynucleotide set out in SEQ ID NO:1 or 3 was discovered in a DNA library derived from *Moraxella catarrhalis*.

5

Moreover, each DNA sequence set out in SEQ ID NO:1 or 3 contains an open reading frame encoding a protein having about the number of amino acid residues set forth in SEQ ID NO:2 or 4 with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known to those skilled in the art.

10

The polynucleotide of SEQ ID NO:1, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 808 of SEQ ID NO:1, encodes the polypeptide of SEQ ID NO:2.

The polynucleotide of SEQ ID NO:3, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 808 of SEQ ID NO:3, encodes the polypeptide of SEQ ID NO:4.

In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of:

- 20 (a) a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1 or 3 over the entire length of SEQ ID NO:1 or 3 respectively; or
- 25 (b) a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or 100% exact, to the amino acid sequence of SEQ ID NO:2 or 4, over the entire length of SEQ ID NO:2 or 4 respectively.

A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *Moraxella catarrhalis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS
5 concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising the sequence of SEQ ID NO:1 or 3 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention provides a polynucleotide sequence identical over its entire length to a coding
10 sequence (open reading frame) in SEQ ID NO:1 or 3. Also provided by the invention is a coding sequence for a mature polypeptide or a fragment thereof, by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence. The polynucleotide of the invention may also contain at least one
15 non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, and polyadenylation signals. The polynucleotide sequence may also comprise additional coding sequence encoding
20 additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gents *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell* 37: 767 (1984), both of which may be useful in purifying
25 polypeptide sequence fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The nucleotide sequence encoding BASB023 polypeptide of SEQ ID NO:2 or 4 may be identical to the polypeptide encoding sequence contained in nucleotides 1 to 807 of SEQ ID NO:1 or 3 respectively. Alternatively it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID
5 NO:2 or 4.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *Moraxella*
10 *catarrhalis* BASB023 having an amino acid sequence set out in SEQ ID NO:2 or 4. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an integrated transposon sequence, or due to RNA editing or genomic DNA reorganization)
15 together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode variants of a polypeptide having a deduced amino acid sequence of SEQ ID NO:2 or 4. Fragments of polynucleotides of the invention may be used, for example, to synthesize full-
20 length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding BASB023 variants, that have the amino acid sequence of BASB023 polypeptide of SEQ ID NO:2 or 4 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these are
25 silent substitutions, additions and deletions, that do not alter the properties and activities of BASB023 polypeptide.

Further preferred embodiments of the invention are polynucleotides that are at least 85% identical over their entire length to a polynucleotide encoding BASB023 polypeptide having an amino acid sequence set out in SEQ ID NO:2 or 4, and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are
5 polynucleotides that comprise a region that is at least 90% identical over its entire length to a polynucleotide encoding BASB023 polypeptide and polynucleotides complementary thereto. In this regard, polynucleotides at least 95% identical over their entire length to the same are particularly preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99%
10 are particularly highly preferred, with at least 99% being the more preferred.

Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by a DNA of SEQ ID NO:1 or 3.
15

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to BASB023 polynucleotide sequences, such as those polynucleotides in SEQ ID NO:1 or 3.

20 The invention further relates to polynucleotides that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least
25 97% identity between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about

65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.

5

The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1 or 3 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in
10 SEQ ID NO:1 or 3 or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers fully described elsewhere herein.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for
15 instance, the polynucleotides of the invention, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding BASB023 and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the BASB023 gene. Such probes generally will
20 comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have less than 30 nucleotide residues or base pairs.

A coding region of a BASB023 gene may be isolated by screening using a DNA sequence
25 provided in SEQ ID NO:1 or 3 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA* 85: 8998-9002, 1988). Recent modifications of the technique, exemplified by the
5 Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific
10 oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the
15 product directly to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides and polypeptides of the invention may be employed, for example, as
20 research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ ID NOS:1 - 4 may be used in the processes herein as described, but preferably for
25 PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

10 For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

A precursor protein, having a mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleic acid that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

25 In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a proprotein, which is a precursor to a proprotein, having a leader sequence and one or

more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In accordance with an aspect of the invention, there is provided the use of a
5 polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles
10 (Wolff *et al.*, *Hum Mol Genet* (1992) 1: 363, Manthorpe *et al.*, *Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al.*, *J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, *Science* (1989) 243: 375), particle bombardment (Tang *et al.*,
15 *Nature* (1992) 356:152, Eisenbraun *et al.*, *DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS USA* (1984) 81: 5849).

Vectors, Host Cells, Expression Systems

20 The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

25 Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a polynucleotide or polynucleotides of the present

invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

- For recombinant production of the polypeptides of the invention, host cells can be
- 5 genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor
- 10 Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.
- 15 Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Neisseria meningitidis* and *Moraxella catarrhalis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*;
- 20 insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

- A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived
- 25 vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from

plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis

Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), Listeria, Salmonella, Shigella, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

5

Diagnostic, Prognostic, Serotyping and Mutation Assays

This invention is also related to the use of BASB023 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of BASB023 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a
10 diagnostic method for diagnosis of disease, staging of disease or response of an infectious organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the BASB023 gene or protein, may be detected at the nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

15

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to
20 analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence
25 selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled BASB023 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting

differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

- 10 In another embodiment, an array of oligonucleotides probes comprising BASB023 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see, for example, Chee *et al.*, *Science*, 274: 610 (1996)).

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:
(a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:1 or 3, or a fragment thereof ;

- 20 (b) a nucleotide sequence complementary to that of (a);
(c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or 4 or a fragment thereof; or
(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2 or 4.

25

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a Disease, among others.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferable, SEQ ID NO:1 or 3, which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

5

10

Cells from an organism carrying mutations or polymorphisms (allelic variations) in a polynucleotide and/or polypeptide of the invention may also be detected at the polynucleotide or polypeptide level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations in the RNA. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA, cDNA or genomic DNA may also be used for the same purpose, PCR. As an example, PCR primers complementary to a polynucleotide encoding BASB023 polypeptide can be used to identify and analyze mutations.

15

20

The invention further provides primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying BASB023 DNA and/or RNA isolated from a sample derived from an individual, such as a bodily material. The primers may be used to amplify a polynucleotide isolated from an infected individual, such that the polynucleotide may then be subject to various techniques for elucidation of the polynucleotide sequence. In this way, mutations in the polynucleotide sequence may be detected and used to diagnose and/or prognose the infection or its stage or course, or to serotype and/or classify the infectious agent.

25

- The invention further provides a process for diagnosing, disease, preferably bacterial infections, more preferably infections caused by *Moraxella catarrhalis*, comprising determining from a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence of SEQ ID NO:1 or 3.
- 5 Increased or decreased expression of a BASB023 polynucleotide can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting, spectrometry and other hybridization methods.
- 10 In addition, a diagnostic assay in accordance with the invention for detecting over-expression of BASB023 polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a BASB023 polypeptide, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include
- 15 radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

The polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are

20 particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as using hybridization or nucleic acid amplification, using a probes obtained or derived from a bodily sample, to determine the presence of a particular polynucleotide sequence or related sequence in

25 an individual. Such a presence may indicate the presence of a pathogen, particularly *Moraxella catarrhalis*, and may be useful in diagnosing and/or prognosing disease or a course of disease. A grid comprising a number of variants of the polynucleotide sequence of SEQ ID NO:1 or 3 are preferred. Also preferred is a comprising a number

of variants of a polynucleotide sequence encoding the polypeptide sequence of SEQ ID NO:2 or 4.

Antibodies

- 5 The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively.

In certain preferred embodiments of the invention there are provided antibodies against
10 BASB023 polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing
15 either or both, to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL*
20 *ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals,
25 may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR

amplified v-genes of lymphocytes from humans screened for possessing anti-BASB023 or from naive libraries (McCafferty, *et al.*, (1990), *Nature* 348, 552-554; Marks, *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature* 352: 628).

5

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

10 Thus, among others, antibodies against BASB023-polypeptide or BASB023-polynucleotide may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

15

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complementarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example
20 as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Antagonists and Agonists - Assays and Molecules

25 Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring BASB023 polypeptide and/or polynucleotide activity in the mixture, and comparing the BASB023 polypeptide and/or polynucleotide activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and BASB023 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide

in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively)
5 from suitably manipulated cells or tissues.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of BASB023 polypeptides or polynucleotides, particularly those compounds that are bacteriostatic and/or bactericidal.
10 The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising BASB023 polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a BASB023 agonist or antagonist. The
15 ability of the candidate molecule to agonize or antagonize the BASB023 polypeptide is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the effects of BASB023 polypeptide are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase
20 signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, signal transduction, or chemical channel activity may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in BASB023
25 polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for BASB023 agonists is a competitive assay that combines BASB023 and a potential agonist with BASB023-binding molecules, recombinant BASB023 binding molecules, natural substrates or ligands, or substrate or ligand mimetics,

under appropriate conditions for a competitive inhibition assay. BASB023 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of BASB023 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

5

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related
10 protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing BASB023-induced activities, thereby preventing the action or expression of BASB023 polypeptides and/or polynucleotides by excluding BASB023 polypeptides and/or polynucleotides from binding.

15 Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991);
20 *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of BASB023.

In a further aspect, the present invention relates to genetically engineered soluble fusion
25 proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by

incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host responsible for sequelae of infection. In particular, the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular gram positive and/or gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial BASB023 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

In accordance with yet another aspect of the invention, there are provided BASB023 agonists and antagonists, preferably bacteristatic or bactericidal agonists and antagonists.

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

5 In a further aspect, the present invention relates to mimotopes of the polypeptide of the invention. A mimotope is a peptide sequence, sufficiently similar to the native peptide (sequentially or structurally), which is capable of being recognised by antibodies which recognise the native peptide; or is capable of raising antibodies which recognise the native peptide when coupled to a suitable carrier.

10 Peptide mimotopes may be designed for a particular purpose by addition, deletion or substitution of elected amino acids. Thus, the peptides may be modified for the purposes of ease of conjugation to a protein carrier. For example, it may be desirable for some chemical conjugation methods to include a terminal cysteine. In addition it may be desirable for peptides conjugated to a protein carrier to include a hydrophobic terminus
15 distal from the conjugated terminus of the peptide, such that the free unconjugated end of the peptide remains associated with the surface of the carrier protein. Thereby presenting the peptide in a conformation which most closely resembles that of the peptide as found in the context of the whole native molecule. For example, the peptides may be altered to have an N-terminal cysteine and a C-terminal hydrophobic amidated
20 tail. Alternatively, the addition or substitution of a D-stereoisomer form of one or more of the amino acids may be performed to create a beneficial derivative, for example to enhance stability of the peptide.

25 Alternatively, peptide mimotopes may be identified using antibodies which are capable themselves of binding to the polypeptides of the present invention using techniques such as phage display technology (EP 0 552 267 B1). This technique, generates a large number of peptide sequences which mimic the structure of the native peptides and are, therefore, capable of binding to anti-native peptide antibodies, but may not necessarily themselves share significant sequence homology to the native polypeptide.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, preferably humans, which comprises
5 inoculating the individual with BASB023 polynucleotide and/or polypeptide, or a fragment or variant thereof, adequate to produce antibody and/ or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *Moraxella catarrhalis* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention
10 relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of BASB023 polynucleotide and/or polypeptide, or a fragment or a variant thereof, for expressing BASB023 polynucleotide and/or polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response, such as, to produce
15 antibody and/ or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a ribozyme, a
20 modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an
25 immunological response, induces an immunological response in such individual to a BASB023 polynucleotide and/or polypeptide encoded therefrom, wherein the composition comprises a recombinant BASB023 polynucleotide and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA which encodes and expresses an antigen of said BASB023 polynucleotide, polypeptide encoded therefrom, or other

polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+ T cells.

- 5 A BASB023 polypeptide or a fragment thereof may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as
- 10 lipoprotein D from *Haemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or
- 15 carboxy-terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science

20 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic

25 immunization experiments in animal models of infection with *Moraxella catarrhalis*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection,

for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Moraxella catarrhalis* infection, in mammals, particularly humans.

The invention also includes a vaccine formulation which comprises an immunogenic
5 recombinant polypeptide and/or polynucleotide of the invention together with a suitable carrier, such as a pharmaceutically acceptable carrier. Since the polypeptides and polynucleotides may be broken down in the stomach, each is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include
10 aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostatic compounds and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed
15 ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system
20 raises preferentially a TH1 type of response.

An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response
25 have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses.

In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

5

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

10

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF- γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2-type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

15

20

It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.

25

Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454). 3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose wherein the antigen will typically be present in a range 2-50µg per dose.

Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of *Quillaja Saponaria Molina*. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with an carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have

been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

Further adjuvants which are preferential stimulators of TH1 cell response include
5 immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential
10 stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

15 Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha
20 tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprilyn.

25 Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1µg - 200µg, such as 10-100µg, preferably 10µg - 50µg per dose. Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85

may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

5 Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

10

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

15

While the invention has been described with reference to certain BASB023 polypeptides and polynucleotides, it is to be understood that this covers fragments of the naturally occurring polypeptides and polynucleotides, and similar polypeptides and polynucleotides with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant polypeptides or polynucleotides.

20

Compositions, kits and administration

In a further aspect of the invention there are provided compositions comprising a BASB023 polynucleotide and/or a BASB023 polypeptide for administration to a cell or to a multicellular organism.

25

The invention also relates to compositions comprising a polynucleotide and/or a polypeptides discussed herein or their agonists or antagonists. The polypeptides and

polynucleotides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to an individual. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide and/or polynucleotide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides, polynucleotides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, agonist or antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one

or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides, polynucleotides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

- 5 The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic
- 10 acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, solutions, powders and the like.
- 15 For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course,
- 20 be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the

25 attending practitioner. Suitable dosages, however, are in the range of 0.1-100 $\mu\text{g}/\text{kg}$ of subject.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5

microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

5

Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

10

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of similar homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well known searching tools, such as the GCG program package.

20

Also provided by the invention are methods for the analysis of character sequences or strings, particularly genetic sequences or encoded protein sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, DNA, RNA and protein structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, codon usage analysis, nucleic acid base trimming, and sequencing chromatogram peak analysis.

25

A computer based method is provided for performing homology identification. This method comprises the steps of: providing a first polynucleotide sequence comprising the sequence of a polynucleotide of the invention in a computer readable medium; and comparing said first polynucleotide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

A computer based method is also provided for performing homology identification, said method comprising the steps of: providing a first polypeptide sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and comparing said first polypeptide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

20

DEFINITIONS

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993;

Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GAP program in the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990), and FASTA(Pearson and Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444-2448 (1988). The BLAST family of programs is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,
20 *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)

Gap Penalty: 8

Gap Length Penalty: 2

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default
25 parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

5

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide
 10 comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or
 100% identity to the reference sequence of SEQ ID NO:1, wherein said polynucleotide
 sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up
 to a certain integer number of nucleotide alterations as compared to the reference
 sequence, wherein said alterations are selected from the group consisting of at least one
 15 nucleotide deletion, substitution, including transition and transversion, or insertion, and
 wherein said alterations may occur at the 5' or 3' terminal positions of the reference
 nucleotide sequence or anywhere between those terminal positions, interspersed either
 individually among the nucleotides in the reference sequence or in one or more
 contiguous groups within the reference sequence, and wherein said number of nucleotide
 20 alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1
 by the integer defining the percent identity divided by 100 and then subtracting that
 product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

25

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides
 in SEQ ID NO:1, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for
 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \cdot is the symbol for
 the multiplication operator, and wherein any non-integer product of x_n and y is rounded

down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

5

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is it may be 100% identical, or it may include up to a certain integer number of nucleic acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such
 10 alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups
 15 within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of nucleic acids in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleic acids in SEQ ID NO:1, or:

$$20 \quad n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleic acid alterations, x_n is the total number of nucleic acids in SEQ ID NO:1, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n
 25 and y is rounded down to the nearest integer prior to subtracting it from x_n .

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may

be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and
5 wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in
10 SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

15 wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

20

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such
25 alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one

or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

5

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

10

"Individual(s)," when used herein with reference to an organism, means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

15

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

20

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double-stranded regions.

25

“Variant” refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Disease(s)" means any disease caused by or related to infection by a bacteria, including, for example, otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, chronic otitis media with hearing loss, fluid accumulation in the middle ear, auditive nerve damage, delayed speech learning, infection of the upper respiratory tract and inflammation of the middle ear.

EXAMPLES:

25

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1:**Discovery and confirmatory DNA sequencing of the BASB023 gene from *Moraxella catarrhalis* strain ATCC 43617.**

5 The BASB023 gene was first discovered in the Incyte PathoSeq data base containing unfinished genomic DNA sequences of the *Moraxella catarrhalis* strain ATCC 43617 (also referred to as strain MC2931). The translation of the BASB023 polynucleotide sequence showed significant similarity (43 % identity in a 212 amino acids overlap) to a *Legionella adelaidensis* macrophage infectivity potentiator protein.

10

The sequence of the BASB023 gene was further confirmed experimentally. For this purpose, genomic DNA was extracted from 10^{10} cells of the *M. catarrhalis* cells (strain ATCC 43617) using the QIAGEN genomic DNA extraction kit (Qiagen GmbH), and 1µg of this material was submitted to Polymerase Chain Reaction DNA amplification using primers MIP-P1 (5'- CAT AGA AAT CAT GCA GAG ACC -3') [SEQ ID NO:5] and MIP-P2 (5'GAG TCA ATG CCA ATA CGG C -3') [SEQ ID NO:6]. This PCR product was purified on a Biorobot 9600 (Qiagen GmbH) apparatus and subjected to DNA sequencing using the Big Dye Cycle Sequencing kit (Perkin-Elmer) and an ABI 377/PRISM DNA sequencer. DNA sequencing was performed on both strands with a redundancy of 2 and the full-length sequence was assembled using the Sequencher™ software (Applied Biosystems). The resulting DNA sequence and deduced polypeptide sequence are shown as SEQ ID NO:3 and SEQ ID NO:4 respectively. One nucleotide in SEQ ID NO:3 (at position 415) was found different from its counterpart in SEQ ID NO:1, as shown in Figure 2. As a consequence, the Proline at position 139 in SEQ ID NO:2 is replaced by a Serine in SEQ ID NO:4.

20

25

Example 2:**Variability analysis of the BASB023 gene among several *Moraxella catarrhalis* strains.**

2A: Restriction Fragment Length Analysis (RFLP).

Genomic DNA was extracted from 16 *M. catarrhalis* strains (presented in Table 1) as described below. *M. catarrhalis* was streaked for single colonies on BHI agar plates and
5 grown overnight at 37 °C. Three or four single colonies were picked and used to inoculate a ~1.5 ml BHI (Brain-heart infusion) broth seed culture which was grown overnight in a shaking incubator, ~300 rpm, at 37 °C. A 500ml erlenmeyer flask containing ~150 ml of BHI broth was inoculated with the seed culture and grown for
10 ~12-16 hours at 37 °C in a shaking incubator, ~175 rpm, to generate cell mass for DNA isolation. Cells were collected by centrifugation in a Sorvall GSA rotor at ~2000 X g for 15 minutes at room temperature. The supernatant was removed and the cell pellet suspended in ~5.0 ml of sterile water. An equal volume of lysis buffer (200 mM NaCl, 20 mM EDTA, 40 mM Tris-Hcl, pH 8.0, 0.5% (w/v) SDS, 0.5% (v/v) 2-mercaptoethanol, and 250 µg/ml of proteinase K) was added and the cells suspended by
15 gentle agitation and trituration. The cell suspension was then incubated ~12 hours at 50°C to lyse the bacteria and liberate chromosomal DNA. Proteinaceous material was precipitated by the addition of 5.0 ml of saturated NaCl (~6.0 M, in sterile water) and centrifugation at ~5,500xg in a Sorvall SS34 rotor at room temperature. Chromosomal DNA was precipitated from the cleared supernatant by the addition of two volumes of
20 100 % ethanol. Aggregated DNA was collected and washed using gentle agitation in a small volume of a 70 % ethanol solution. Purified chromosomal DNA was suspended in sterile water and allowed to dissolve/ disburse overnight at 4 °C by gentle rocking. The concentration of dissolved DNA was determined spectrophotometrically at 260 nm using an extinction coefficient of 1.0 O.D. unit ~50 µg/ml.

25 This material was next submitted to PCR amplification using the MIP-P1 and MIP-P2 oligonucleotides. The corresponding BASB023 gene amplicons were then subjected independantly to hydrolysis using restriction enzymes (*AluI*, *DraI*, *HinII*, *HphI*, *RsaI*, *Tsp509I*) and restriction products were separated by agarose or polyacrylamide gel electrophoresis using standard molecular biology procedures as described in “Molecular

Cloning, a Laboratory Manual, Second Edition, Eds: Sambrook, Fritsch & Maniatis, Cold Spring Harbor press 1989". The photographs of the resulting electrophoresis gels are displayed in Figure 1. For each strain, RFLP patterns corresponding to the 6 restriction enzymes were scored and combined. Groups of strains sharing identical
5 combination of RFLP patterns were then defined. Using this methodology, the strains tested in this study fell into 2 genomic groups (Group 1: Mc2911, Group 2: Mc2904, Mc 2905, Mc2907, Mc2908, Mc2910, Mc2912, Mc2926, Mc2931, Mc2956, Mc2960, Mc2969, Mc2969, Mc2975; Mc2906, Mc2909, Mc2913). These data support that the *Moraxella catarrhalis* population used in this study displays limited nucleotide
10 sequence diversity for the BASB023 gene.

Table 1: Features of the *Moraxella catarrhalis* strains used in this study

Strain	Isolated in:	from:
Mc2904	USA	Tympanocentesis
Mc2905	USA	Tympanocentesis
Mc2906	USA	Tympanocentesis
Mc2907	USA	Tympanocentesis
Mc2908	USA	Acute otitis Tympanocentesis
Mc2909	USA	Tympanocentesis
Mc2910	USA	Tympanocentesis
Mc2911	USA	Acute otitis Tympanocentesis
Mc2912	USA	Acute otitis Tympanocentesis
Mc2913	USA	Acute otitis Tympanocentesis
Mc2926	USA	Tympanocentesis
Mc2931 /ATCC 43617	USA	Transtracheal aspirate
Mc2956	Finland	Middle ear fluid
Mc2960	Finland	Middle ear fluid
Mc2969	Norway	Nasopharynx (Pharyngitis-Rhinitis)
Mc2975	Norway	Nasopharynx (Rhinitis)

Example 3: Construction of Plasmid to Express Recombinant BASB023

5

A: Cloning of BASB023.

Two oligonucleotide primers, MC-MIP-BADf-F (5'- ATT ACG CAG AGC CAT GGG CTG TCA AAA AGA CAC TTC AGC -3') [SEQ ID NO:7] and MC-MIP-BADf-RC (5'- ATT ACG CAG AGC TCT AGA CCT CGT TGA CCT TGC GAT TCT TGA ATG -3') [SEQ ID NO:8], were synthesized and used to amplify the MIP using the *M. catarrhalis* strain Mc2931 genomic DNA as a template. The *NcoI* and *XbaI* restriction sites engineered into the forward (MC-MIP-BADf-F) and reverse (MC-MIP-BADf-RC) amplification primers permitted directional cloning of the PCR product into the commercially available *E. coli* expression plasmid pBAD-gIII (Invitrogen, ampicillin resistant) such that a mature BASB023 protein could be expressed as a fusion protein, at the N-terminus with the pIII (bacteriophage fd) signal peptide and at the C-terminus with a stretch of 6 Histidine residues (affinity chromatography tag). The BASB023 PCR product was purified from the amplification reaction using silica gel-based spin columns (QiaGen) according to the manufacturers instructions. To produce the required *NcoI* and *XbaI* termini necessary for cloning, purified PCR product was sequentially digested to completion with *NcoI* and *XbaI* restriction enzymes as recommended by the manufacturer (Life Technologies). Following the first restriction digestion, the PCR product was purified via spin column as above to remove salts and eluted in sterile water prior to the second enzyme digestion. The digested DNA fragment was again purified using silica gel-based spin columns prior to ligation with the pBADgIII plasmid.

B: Production of Expression Vector.

To prepare the expression plasmid pBADgIII for ligation, it was similarly digested to completion with both *NcoI* and *XbaI* and then treated with calf intestinal phosphatase

30

(CIP, ~0.02 units / pmole of 5' end, Life Technologies) as directed by the manufacturer to prevent self ligation. An approximately 5-fold molar excess of the digested fragment to the prepared vector was used to program the ligation reaction. A standard ~20 μ l ligation reaction (~16°C, ~16 hours), using methods well known in the art. was performed using T4 DNA ligase (~2.0 units / reaction, Life Technologies). An aliquot of the ligation (~5 μ l) was used to transform electro-competent LMG194 cells according to methods well known in the art. Following a ~2-3 hour outgrowth period at 37°C in ~1.0 ml of LB broth, transformed cells were plated on LB agar plates ampicillin (100 μ g/ml). Plates were incubated overnight at 37°C for ~16 hours. Individual ApR colonies were picked with sterile toothpicks and used to "patch" inoculate fresh LB ApR plates as well as a ~1.0 ml LB ApR broth culture. Both the patch plates and the broth culture were incubated overnight at 37°C in either a standard incubator (plates) or a shaking water bath. A whole cell-based PCR analysis was employed to verify that transformants contained the BASB023 DNA insert. Here, the ~1.0 ml overnight LB Ap broth culture was transferred to a 1.5 ml polypropylene tube and the cells collected by centrifugation in a Beckmann microcentrifuge (~3 min., room temperature, ~12,000 X g). The cell pellet was suspended in ~200 μ l of sterile water and a ~10 μ l aliquot used to program a ~50 μ l final volume PCR reaction containing both BASB023 forward and reverse amplification primers. Final concentrations of the PCR reaction components were essentially the same as those specified in example 2 except ~5.0 units of Taq polymerase was used. The initial 95°C denaturation step was increased to 3 minutes to ensure thermal disruption of the bacterial cells and liberation of plasmid DNA. An ABI Model 9700 thermal cycler and a 32 cycles, three-step thermal amplification profile, i.e. 95°C, 45sec; 55-58°C, 45sec, 72°C, 1min., were used to amplify the BASB023 PCR fragment from the lysed transformant samples. Following thermal amplification, a ~20 μ l aliquot of the reaction was analyzed by agarose gel electrophoresis (0.8 % agarose in a Tris-acetate-EDTA (TAE) buffer). DNA fragments were visualized by UV illumination after gel electrophoresis and ethidium bromide staining. A DNA molecular size standard (1 Kb ladder, Life Technologies) was electrophoresed in parallel with the

test samples and was used to estimate the size of the PCR products. Transformants that produced the expected PCR product were identified as strains containing a BASB023 expression construct. Expression plasmid containing strains were then analyzed for the inducible expression of recombinant BASB023.

5

C: Expression Analysis of PCR-Positive Transformants.

For each PCR-positive transformant identified above, ~5.0 ml of LB broth containing ampicillin (100 µg/ml) was inoculated with cells from the patch plate and grown overnight at 37 °C with shaking (~250 rpm). An aliquot of the overnight seed culture
10 (~1.0 ml) was inoculated into a 125 ml erlenmeyer flask containing ~25 ml of LB Ap broth and grown at 37 °C with shaking (~250 rpm) until the culture turbidity reached O.D.600 of ~0.5, i.e. mid-log phase (usually about 1.5 - 2.0 hours). At this time approximately half of the culture (~12.5 ml) was transferred to a second 125 ml flask and expression of recombinant BASB023 protein was induced by the addition of
15 arabinose to a final concentration of 0.2%. Incubation of both the arabinose-induced and non-induced cultures continued for an additional ~4 hours at 37 °C with shaking. Samples (~1.0 ml) of both induced and non-induced cultures were removed after the induction period and the cells collected by centrifugation in a microcentrifuge at room temperature for ~3 minutes. Individual cell pellets were suspended in ~50µl of sterile
20 water, then mixed with an equal volume of 2X Laemelli SDS-PAGE sample buffer containing 2-mercaptoethanol, and placed in boiling water bath for ~3 min to denature protein. Equal volumes (~15µl) of both the crude arabinose-induced and the non-induced cell lysates were loaded onto duplicate 12% Tris/glycine polyacrylamide gel (1 mm thick Mini-gels, Novex). The induced and non-induced lysate samples were
25 electrophoresed together with prestained molecular weight markers (SeeBlue, Novex) under conventional conditions using a standard SDS/Tris/glycine running buffer (BioRad). Following electrophoresis, one gel was stained with commassie brilliant blue R250 (BioRad) and then destained to visualize novel BASB023 arabinose-inducible protein(s) migrating at about 37 kDa. The second gel was electroblotted onto a PVDF

membrane (0.45 micron pore size, Novex) for ~2 hrs at 4 °C using a BioRad Mini-Protean II blotting apparatus and Towbin's methanol (20 %) transfer buffer. Blocking of the membrane and antibody incubations were performed according to methods well known in the art. A monoclonal anti-RGS (His)₃ antibody, followed by a second rabbit anti-mouse antibody conjugated to HRP (QiaGen), was used to confirm the expression and identity of the BASB023 recombinant protein. Visualization of the anti-His antibody reactive pattern was achieved using either an ABT insoluble substrate or using Hyperfilm with the Amersham ECL chemiluminescence system.

10 **D: Sequence Confirmation.**

To further verify that the arabinose-inducible recombinant BASB023 protein being expressed is in the correct open reading frame and not a spurious molecule arising from a cloning artifact (i.e. a frame-shift), the DNA sequence of the cloned insert was determined. The DNA sequence for the *M.catarrhalis* BASB023 gene was obtained from one strand using conventional asymmetric PCR cycle sequencing methodologies (ABI Prism Dye-Terminator Cycle Sequencing, Perkin-Elmer). Sequencing reactions were programmed with undigested expression plasmid DNA (~0.5µg/rxn) as a template and appropriate pQE30 vector-specific and ORF-specific sequencing primers (~3.5 pmol/rxn). In addition to the template and sequencing primer, each sequencing reaction (~20µl) contained the four different dNTPs (i.e. A,G,C, and T) and the four corresponding ddNTPs (i.e. ddA, ddG, ddC, and ddT) terminator nucleotides; with each terminator being conjugated to one of the four fluorescent dyes, Joe, Tam, Rox, or Fam. Single strand sequencing elongation products were terminated at random positions along the template by the incorporation of the dye-labelled ddNTP terminators. Fluorescent dye-labelled termination products were purified using microcentrifuge size-exclusion chromatography columns (Princeton Genetics), dried under vacuum, suspended in a Template Resuspension Buffer (Perkin-Elmer) for capillary electrophoresis or deionized formamide for PAGE, denatured at 95°C for ~5 min, and analyzed by high resolution capillary electrophoresis (ABI 310 Automated DNA

Sequenator, Perkin-Elmer) or high resolution PAGE (ABI 377 Automated DNA Sequenator) as recommended by the manufacturer. DNA sequence data produced from individual reactions were collected and the relative fluorescent peak intensities analyzed automatically on a PowerMAC computer using ABI Sequence Analysis Software (Perkin-Elmer). Individually autoanalyzed DNA sequences were edited manually for accuracy before being merged into a consensus single strand sequence "string" using AutoAssembler software (Perkin-Elmer). Sequencing determined that the expression plasmid contained the correct sequence in the correct open reading frame.

10 **D: Sequence Confirmation.**

To further verify that the IPTG-inducible recombinant BASB023 protein being expressed is in the correct open reading frame and not a spurious molecule arising from a cloning artifact (i.e. a frame-shift), the DNA sequence of the cloned insert was determined. The DNA sequence for the *M. catarrhalis* BASB023 gene was obtained from one strand using conventional asymmetric PCR cycle sequencing methodologies (ABI Prism Dye-Terminator Cycle Sequencing, Perkin-Elmer). Sequencing reactions were programmed with undigested expression plasmid DNA (~0.5µg/rxn) as a template and appropriate pQE30 vector-specific and ORF-specific sequencing primers (~3.5 pmol/rxn). In addition to the template and sequencing primer, each sequencing reaction (~20µl) contained the four different dNTPs (i.e. A,G,C, and T) and the four corresponding ddNTPs (i.e. ddA, ddG, ddC, and ddT) terminator nucleotides; with each terminator being conjugated to one of the four fluorescent dyes, Joe, Tam, Rox, or Fam. Single strand sequencing elongation products were terminated at random positions along the template by the incorporation of the dye-labelled ddNTP terminators.

25 Fluorescent dye-labelled termination products were purified using microcentrifuge size-exclusion chromatography columns (Princeton Genetics), dried under vacuum, suspended in a Template Resuspension Buffer (Perkin-Elmer) for capillary electrophoresis or deionized formamide for PAGE, denatured at 95°C for ~5 min, and analyzed by high resolution capillary electrophoresis (ABI 310 Automated DNA

Sequenator, Perkin-Elmer) or high resolution PAGE (ABI 377 Automated DNA Sequenator) as recommended by the manufacturer. DNA sequence data produced from individual reactions were collected and the relative fluorescent peak intensities analyzed automatically on a PowerMAC computer using ABI Sequence Analysis Software (Perkin-Elmer). Individually autoanalyzed DNA sequences were edited manually for accuracy before being merged into a consensus single strand sequence "string" using AutoAssembler software (Perkin-Elmer). Sequencing determined that the expression plasmid contained the correct sequence in the correct open reading frame.

10

Example 4: Production of Recombinant BASB023

Bacterial strain

A recombinant expression strain of *E. coli* LMG194 containing a plasmid (pBADgIII) encoding BASB023 from *M. catarrhalis* was used to produce cell mass for purification of recombinant protein. The expression strain was cultivated on LB agar plates containing 100µg/ml ampicillin ("Ap") to ensure plasmid maintenance. For cryopreservation at -80° C, the strain was propagated in LB broth containing the same concentration of antibiotic then mixed with an equal volume of LB broth containing 30% (w/v) glycerol.

20

Media

The fermentation medium used for the production of recombinant protein consisted of 2X YT broth (Difco) containing 100µg/ml Ap. Antifoam was added to medium for the fermentor at 0.25 ml/L (Antifoam 204, Sigma). To induce expression of the BASB023 recombinant protein, arabinose was added to the fermentor (0.2%).

25

Fermentation

A 500-ml erlenmeyer seed flask, containing 50ml working volume, was inoculated with 0.3 ml of rapidly thawed frozen culture, or several colonies from a selective agar plate

culture, and incubated for approximately 12 hours at $37 \pm 1^\circ\text{C}$ on a shaking platform at 150rpm (Innova 2100, New Brunswick Scientific). This seed culture was then used to inoculate a 5-L working volume fermentor containing 2X YT broth and both Kn and Ap antibiotics. The fermentor (Bioflo 3000, New Brunswick Scientific) was operated at $37 \pm 1^\circ\text{C}$, 0.2 - 0.4 VVM air sparge, 250 rpm in Rushton impellers. The pH was not controlled in either the flask seed culture or the fermentor. During fermentation, the pH ranged 6.5 to 7.3 in the fermentor arabinose (2% stock, prepared in sterile water) was added to the fermentor when the culture reached mid-log of growth (~ 0.7 O.D.600 units). Cells were induced for 2 - 4 hours then harvested by centrifugation using either a 28RS Heraeus (Sepatech) or RC5C superspeed centrifuge (Sorvall Instruments). Cell paste was stored at -20°C until processed.

Purification

Chemicals and Materials

Imidazole, guanidine hydrochloride, Tris (hydroxymethyl), and EDTA (ethylene-diamine tetraacetic acid) biotechnology grade or better were all obtained from Ameresco Chemical, Solon, Ohio. Triton X-100 (t-Octylphenoxy polyethoxy-ethanol), sodium phosphate, monobasic, and Urea were reagent grade or better and obtained from Sigma Chemical Company, St. Louis, Missouri. Glacial acetic acid and hydrochloric acid were obtained from Mallincrodt Baker Inc., Phillipsburg, New Jersey. Methanol was obtained from Fisher Scientific, Fairlawn, New Jersey. Pefabloc®SC (4-(2-Aminoethyl)-benzenesulfonyl fluoride), Complete protease inhibitor cocktail tablets, and PMSF (phenylmethyl-sulfonyl fluoride) were obtained from Roche Diagnostics Corporation, Indianapolis, Indiana. Bestatin, Pepstatin A, and E-64 protease inhibitor were obtained from Calbiochem, LaJolla, California. Dulbecco's Phosphate Buffered Saline (1x PBS) was obtained from Quality Biological, Inc., Gaithersburg, Maryland. Dulbecco's Phosphate Buffered Saline (10x PBS) was obtained from BioWhittaker, Walkersville, Maryland. Penta-His Antibody, BSA free was obtained from QiaGen, Valencia, California. Peroxidase-conjugated AffiniPure Goat Anti-mouse IgG was

obtained from Jackson Immuno Research, West Grove, Penn. AEC single solution was obtained from Zymed, South San Francisco, California. All other chemicals were reagent grade or better.

Ni-NTA Superflow resin was obtained from QiaGen Inc., Valencia, California. Precast
5 Tris-Glycine 4-20% and 10-20% polyacrylamide gels, all running buffers and solutions, SeeBlue Pre-Stained Standards, MultiMark Multi-Colored Standards and PVDF transfer membranes were obtained from Novex, San Diego, California. SDS-PAGE Silver Stain kits were obtained from Daiichi Pure Chemicals Company Limited, Tokyo, Japan. Coomassie Stain Solution was obtained from Bio-Rad Laboratories, Hercules,
10 California. Acrodisc® PF 0.2 µm syringe filters were obtained from Pall Gelman Sciences, Ann Arbor, Michigan. GD/X 25mm disposable syringe filters were obtained from Whatman Inc., Clifton, New Jersey. Dialysis tubing 8,000 MWCO was obtained from BioDesign Inc. Od New York, Carmal New York. BCA Protein Assay Reagents and Snake Skin dialysis tubing 3,500 MWCO were obtained from Pierce Chemical Co.,
15 Rockford, Illinois.

Extraction Protocol

Cell paste was thawed at room temperature for 30 to 60 minutes. Five to six grams of material was weighed out into a 50 ml disposable centrifuge tube. To this five
20 mls/gram of Guanidine hydrochloride (Gu-HCl) buffer was added (6 M Guanidine hydrochloride, 100 mM Sodium phosphate, monobasic, 10 mM Tris and 0.05 % Triton X-100, pH 8.0). Cell paste was resuspended using a PRO300D proscientific homogenizer, at 3/4 power for one minute. The extraction mixture was then placed at room temperature with gentle agitation for 60 to 90 minutes. After 60 to 90 minutes the
25 extraction mixture was centrifuged at 15,800 x g for 15 minutes (Sorvall RC5C centrifuge, 11,500 rpm). The supernatant (S1) was decanted and saved for additional purification. The pellet (P1) was saved for analysis.

Binding of BASB023 to Nickel-NTA Resin

To the S1 three to four mls of Ni-NTA resin is added. This is then placed at room temperature with gentle agitation for one hour. After one hour the S1/Ni-NTA is packed into an XK16 Pharmacia column. The column is then washed with 1 M Gu-HCl buffer (1 M Guanidine hydrochloride, 100mM Sodium phosphate, monobasic, 10 mM Tris and 0.05% Triton X-100, pH 8.0). This is then followed by a wash with phosphate buffer (100mM Sodium phosphate, monobasic, 10 mM Tris and 0.05% Triton X-100, pH 6.3). The protein is then eluted from the column with a 250 mM imidazole buffer (250 mM imidazole, 100mM Sodium phosphate, monobasic, 10 mM Tris and 0.05% Triton X-100, pH 5.9).

10

Final Formulation

BASB023 was formulated by dialysis overnight against, three changes of 0.1 % Triton X-100 and 1x PBS, pH 7.4 (pH 12.0 for MC-D15), to remove residual Gu-HCl and imidazole. The purified protein was characterized and used to produced antibodies as described below.

15

Biochemical Characterizations

SDS-PAGE and Western Blot Analysis

The recombinant purified protein was resolved on 4-20 % polyacrylamide gels and electrophoretically transferred to PVDF membranes at 100 V for 1 hour as previously described (Thebaine et al. 1979, Proc. Natl. Acad. Sci. USA 76:4350-4354). The PVDF membranes were then pretreated with 25 ml of Dulbecco's phosphate buffered saline containing 5 % non-fat dry milk. All subsequent incubations were carried out using this pretreatment buffer.

25

PVDF membranes were incubated with 25 ml of a 1:500 dilution of preimmune serum or rabbit anti-His immune serum for 1 hour at room temperature. PVDF membranes were then washed twice with wash buffer (20 mM Tris buffer, pH 7.5, containing 150 mM sodium chloride and 0.05 % Tween-20). PVDF membranes were incubated with

25 ml of a 1:5000 dilution of peroxidase-labeled goat anti-rabbit IgG (Jackson ImmunoResearch Laboratories, West Grove, PA) for 30 minutes at room temperature. PVDF membranes were then washed 4 times with wash buffer, and were developed with 3-amino-9-ethylcarbazole and urea peroxide as supplied by Zymed (San Francisco, CA) for 10 minutes each.

The results of an SDS-PAGE (Figure 4) show a protein about 37 kDa purified to greater than 95 % and that is reactive to an anti-RGS(His) antibody by western blots (Figure 5) of the SDS-PAGE.

10

Protein Sequencing

Amino terminal amino acid sequencing of the purified protein was performed to confirm the production of the correct recombinant protein using well defined chemical protocols on Hewlett-Packard model G1000A sequencer with a model 1090 LC and a Hewlett-Packard model 241 sequencer with a model 1100 LC.

15

Example 5: Production of Antisera to Recombinant BASB023

Polyvalent antisera directed against the BASB023 protein were generated by vaccinating two rabbits with the purified recombinant BASB023 protein. Each animal is given a total of three immunizations intramuscularly (i.m.) of about 20 μ g BASB023 protein per injection (beginning with complete Freund's adjuvant and followed with incomplete Freund's adjuvant) at approximately 21 day intervals. Animals were bled prior to the first immunization ("pre-bleed") and on days 35 and 57.

25

Anti-BASB023 protein titres were measured by an ELISA using purified recombinant BASB023 protein (0.5 μ g/well). The titre is defined as the highest dilution equal or greater than 0.1 as calculated with the following equation: average OD of two test

samples of antisera - the average OD of two test samples of buffer. The titres after three immunizations were above 3000.

The antisera were used as the first antibody to identify the protein in a western blot as described in example 5 above. The western-blot showed the presence of anti-BASB023 antibody in the sera of immunized animals.

Example 6: Presence of Antibody to BASB023 in Human Convalescent Sera

Western blot analysis of purified recombinant BASB023 were performed as described in Example 4 above, except that a pool of human sera from children infected by *M. catarrhalis* is used as the first antibody preparation. Results show that antisera from naturally infected individuals react to the purified recombinant protein, as shown in Figure 6.

Deposited materials

5 A deposit containing a *Moraxella catarrhalis* Catlin strain has been deposited with the American Type Culture Collection (herein "ATCC") on June 21, 1997 and assigned deposit number 43617. The deposit was described as *Branhamella catarrhalis* (Frosch and Kolle) and is a freeze-dried, 1.5-2.9 kb insert library constructed from *M. catarrhalis* isolate obtained from a transtracheal aspirate of a coal miner with chronic bronchitis. The deposit is described in *Antimicrob. Agents Chemother.* 21: 506-508 (1982).

10

The *Moraxella catarrhalis* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains a full length BASB023 gene.

15

The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of any polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

20 The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The deposited strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a
25 deposit is required for enablement, such as that required under 35 U.S.C. §112.

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page <u>64</u> , line <u>4-25</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution AMERICAN TYPE CULTURE COLLECTION	
Address of depositary institution (including postal code and country) 10801 UNIVERSITY BLVD, MANASSAS, VIRGINIA 20110-2209, UNITED STATES OF AMERICA	
Date of deposit <u>21.06.97</u> (21 June 1997)	Accession Number <u>43617</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect of those designations where a European Patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European Patent or until the date on which the application has been refused or withdrawn, only by issue of such a sample to an expert nominated by the person requesting the sample.	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g. "Accession Number of Deposit")	

For receiving Office use only
<input checked="" type="checkbox"/> This sheet was received with the international application
11 AUG 1999
Authorized officer E. Speiser <i>ES</i>

For International Bureau use only
<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer

CLAIMS:

1. An isolated polypeptide comprising an amino acid sequence which has at least 85% identity to the amino acid sequence of: SEQ ID NO:4.
5
2. An isolated polypeptide as claimed in claim 1 in which the amino acid sequence has at least 95% identity to the amino acid sequence of: SEQ ID NO:4.
3. The polypeptide as claimed in claim 1 comprising the amino acid of: SEQ ID NO:4.
10
4. An isolated polypeptide of SEQ ID NO:4.
5. An isolated polypeptide of SEQ ID NO:2.
- 15 6. An immunogenic fragment of the polypeptide as claimed in any one of claims 1 to 5 in which the immunogenic activity of said immunogenic fragment is substantially the same as the polypeptide of SEQ ID NO:4.
7. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide
20 that has at least 85% identity to the amino acid sequence of SEQ ID NO: 4 over the entire length of SEQ ID NO: 4; or a nucleotide sequence complementary to said isolated polynucleotide.
8. An isolated polynucleotide comprising a nucleotide sequence that has at least 85%
25 identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO: 4 over the entire coding region; or a nucleotide sequence complementary to said isolated polynucleotide.

9. An isolated polynucleotide which comprises a nucleotide sequence which has at least 85% identity to that of SEQ ID NO: 3 over the entire length of SEQ ID NO: 3; or a nucleotide sequence complementary to said isolated polynucleotide.
- 5 10. The isolated polynucleotide as claimed in any one of claims 7 to 9 in which the identity is at least 95% to SEQ ID NO: 3.
11. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:4.
- 10 12. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:3.
13. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:4, obtainable by screening an appropriate library under stringent
- 15 hybridization conditions with a labeled probe having the sequence of SEQ ID NO:3 or a fragment thereof.
14. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2.
- 20 15. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1.
16. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2, obtainable by screening an appropriate library under stringent
- 25 hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or a fragment thereof.
17. An expression vector or a recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 7 - 16.

18. A host cell comprising the expression vector of claim 17 or a subcellular fraction or a membrane of said host cell expressing an isolated polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence of: SEQ ID NO:4.

5

19. A process for producing a polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence of SEQ ID NO:4 comprising culturing a host cell of claim 18 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

10

20. A process for expressing a polynucleotide of any one of claims 7 – 16 comprising transforming a host cell with the expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

15

21. A vaccine composition comprising an effective amount of the polypeptide of any one of claims 1 to 6 and a pharmaceutically acceptable carrier.

22. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 7 to 16 and a pharmaceutically effective carrier.

20

23. The vaccine composition according to either one of claims 21 or 22 wherein said composition comprises at least one other *Moraxella catarrhalis* antigen.

24. An antibody immunospecific for the polypeptide or immunological fragment as claimed in any one of claims 1 to 6.

25

25. A method of diagnosing a *Moraxella* infection, comprising identifying a polypeptide as claimed in any one of claims 1 - 6, or an antibody that is immunospecific for said

polypeptide, present within a biological sample from an animal suspected of having such an infection.

26. Use of a composition comprising an immunologically effective amount of a
5 polypeptide as claimed in any one of claims 1 – 6 in the preparation of a medicament for use in generating an immune response in an animal.

27. Use of a composition comprising an immunologically effective amount of a
10 polynucleotide as claimed in any one of claims 7 - 16 in the preparation of a medicament for use in generating an immune response in an animal.

28. A therapeutic composition useful in treating humans with *Moraxella catarrhalis*
disease comprising at least one antibody directed against the polypeptide of claims 1 – 6
and a suitable pharmaceutical carrier.

15

Fig. 1 : Photographs of the electrophoresis gels obtained in the RFLP analysis of BASB023

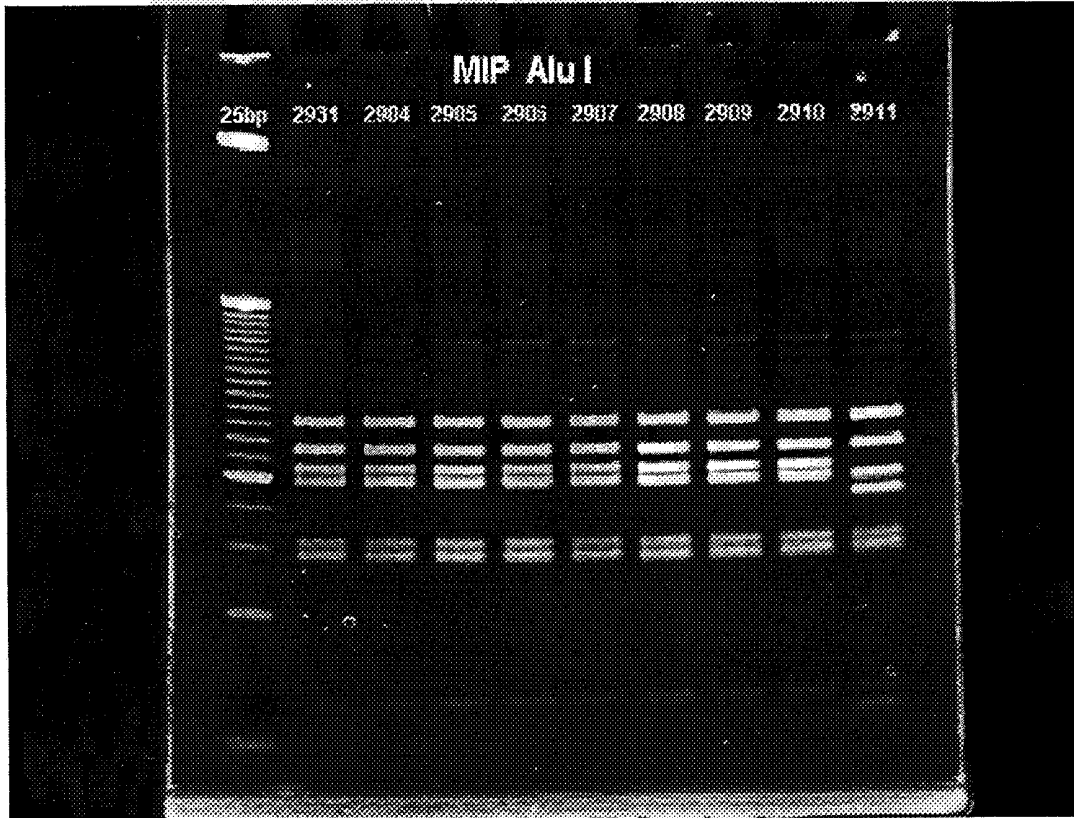


Fig. 1A

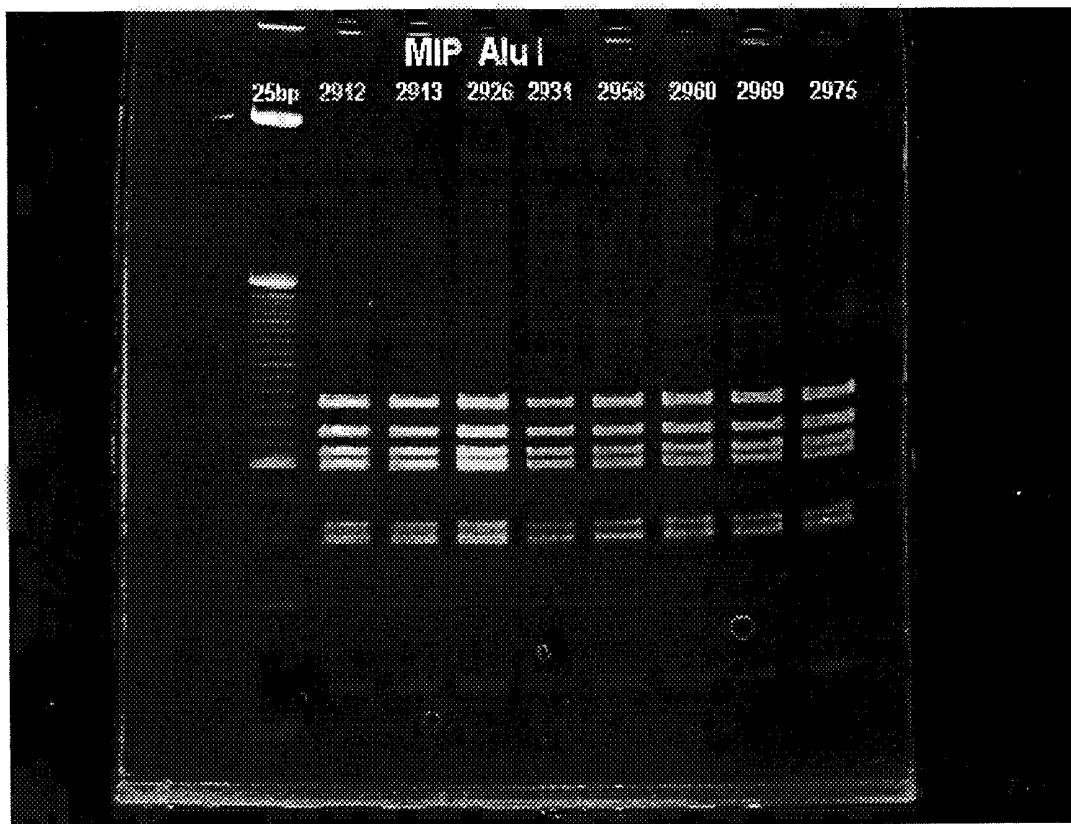


Fig. 1B

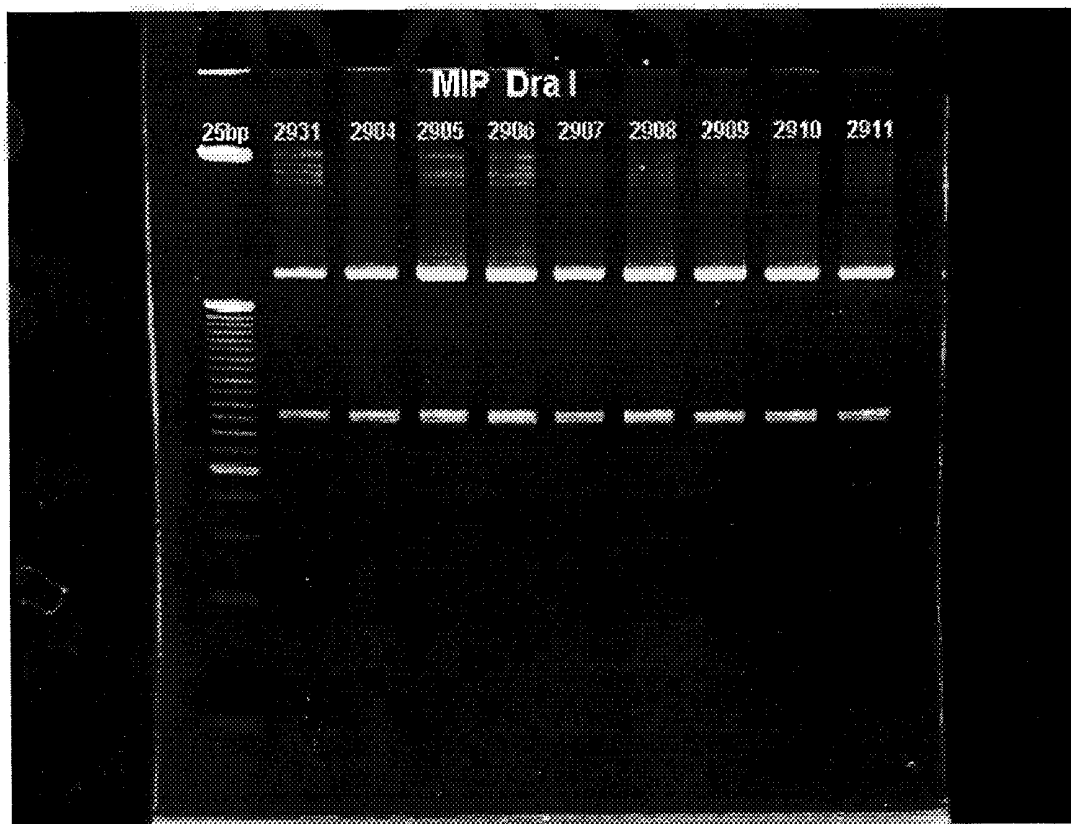


Fig. 1C



Fig. 1D

5/19

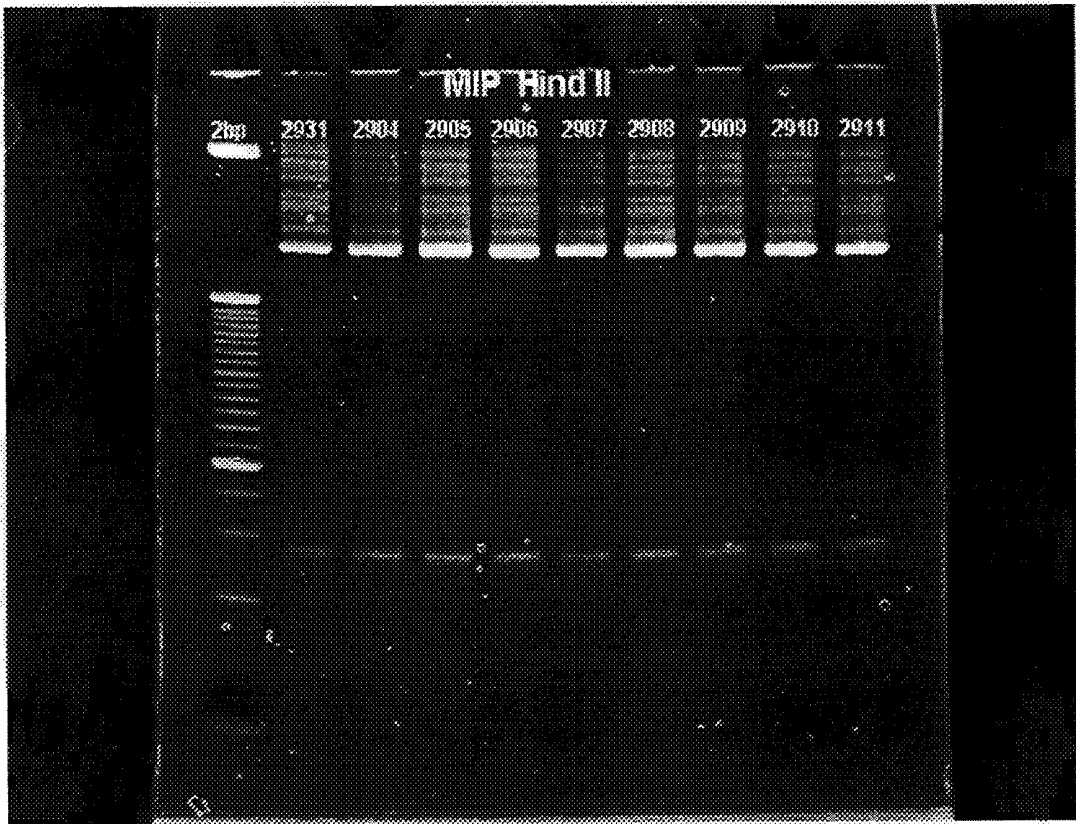


Fig. 1 E

6/19

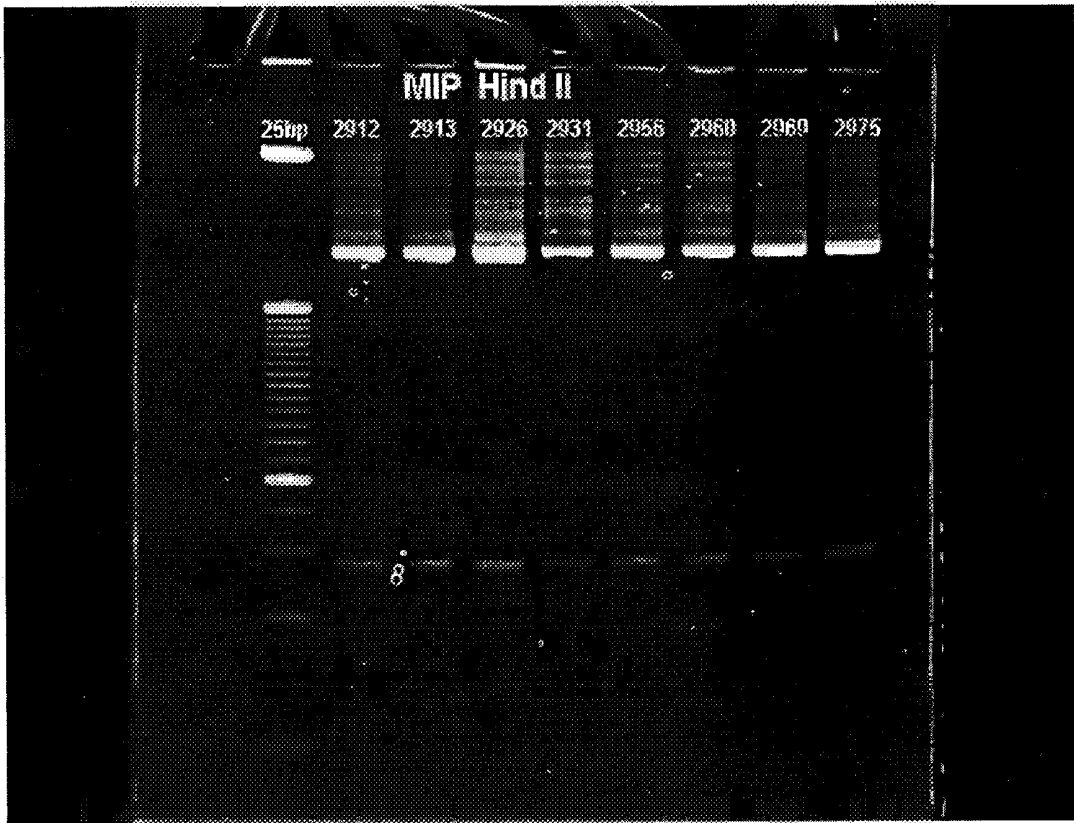


Fig. 1 F

7/19

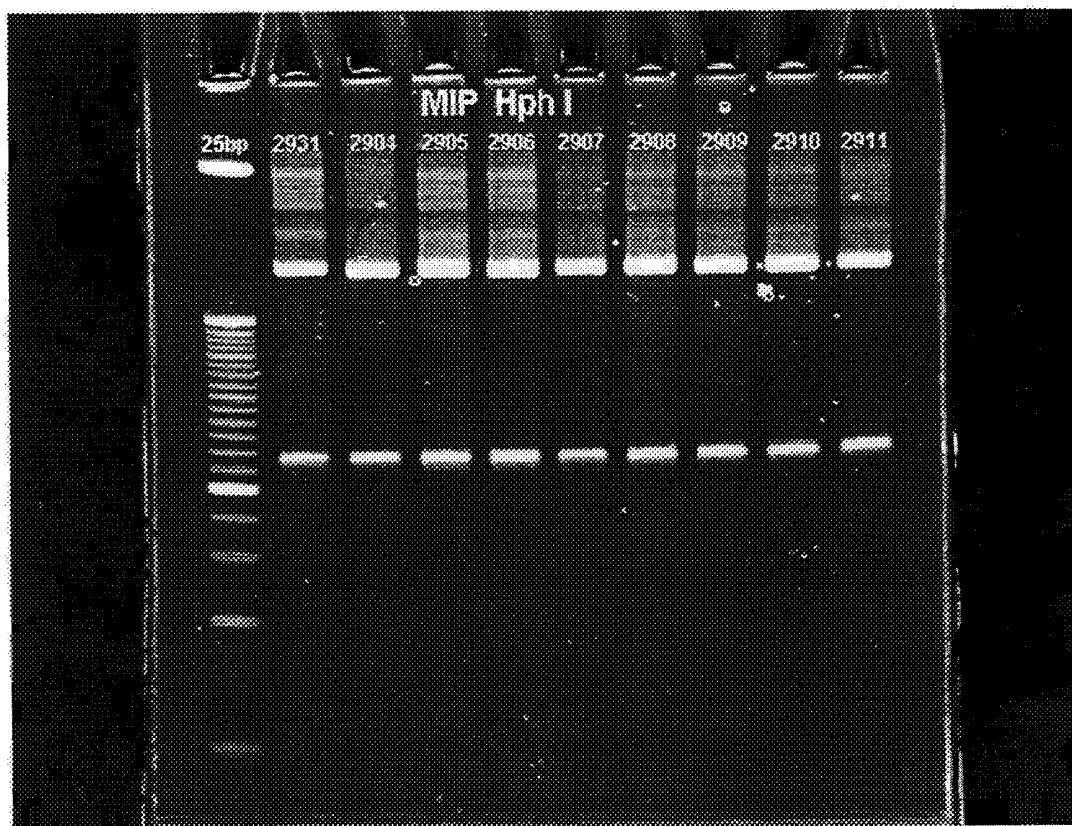


Fig. 1G

8/19

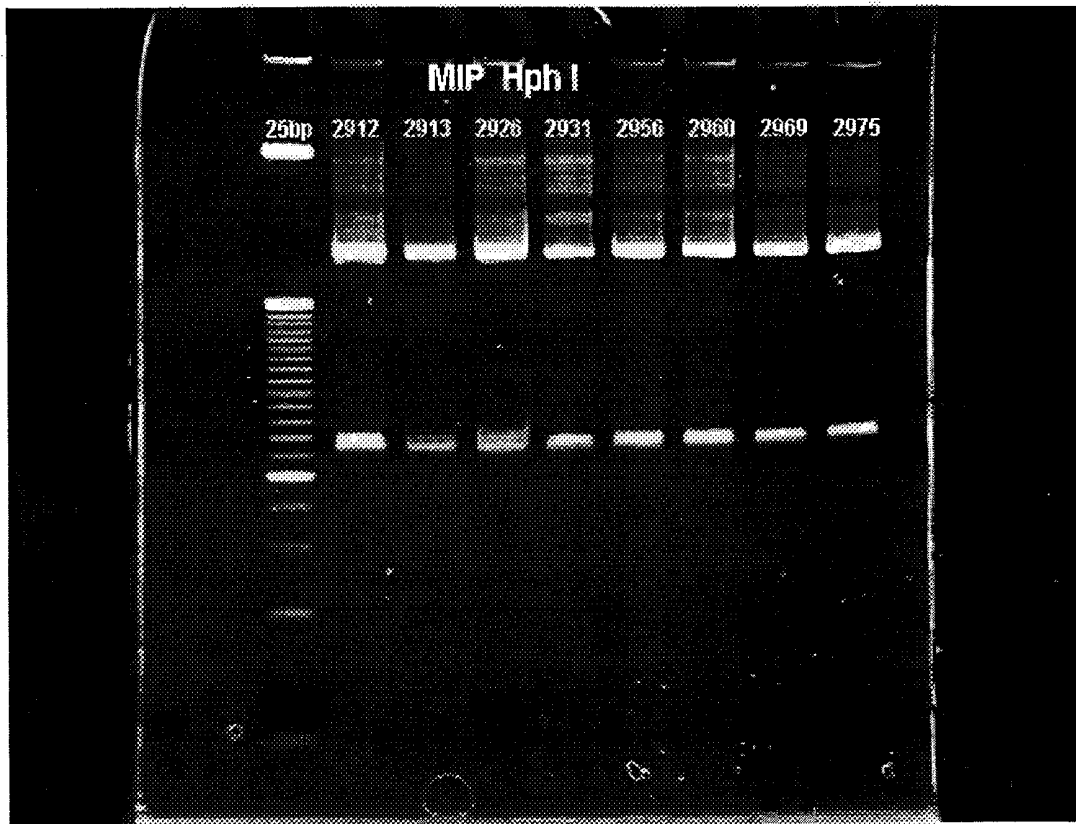


Fig. 1 H

9/19

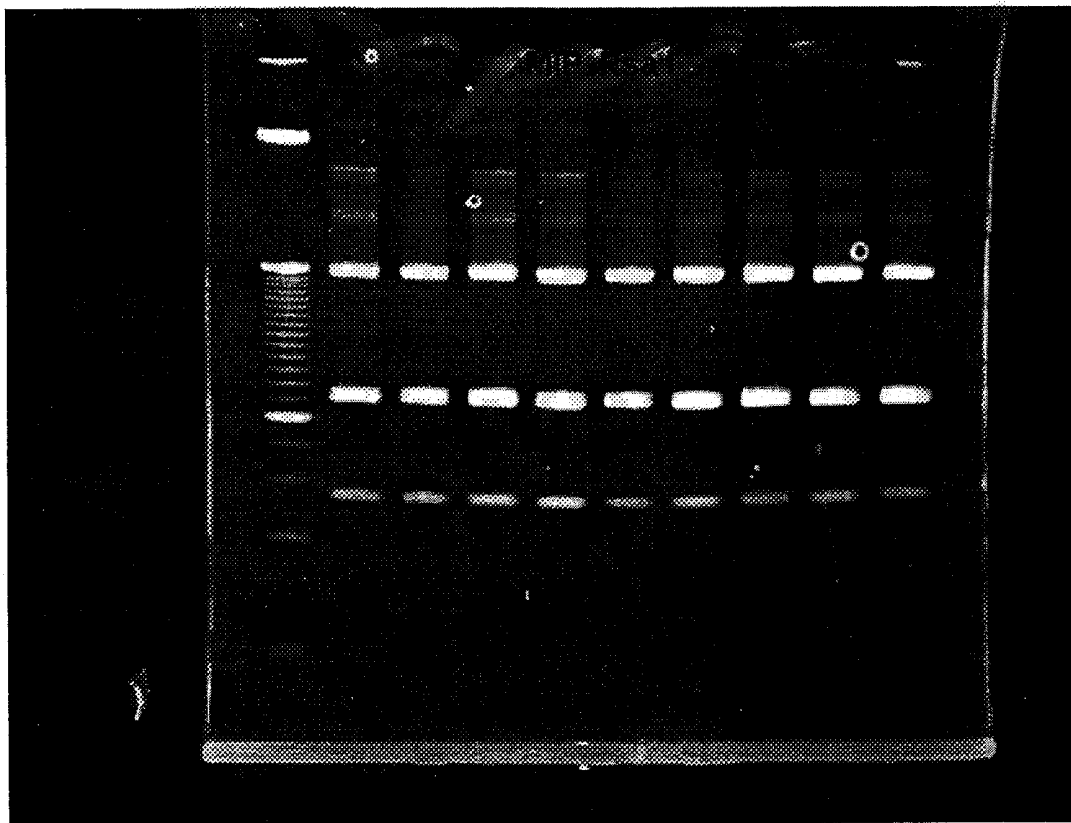


Fig. 1 I

10/19

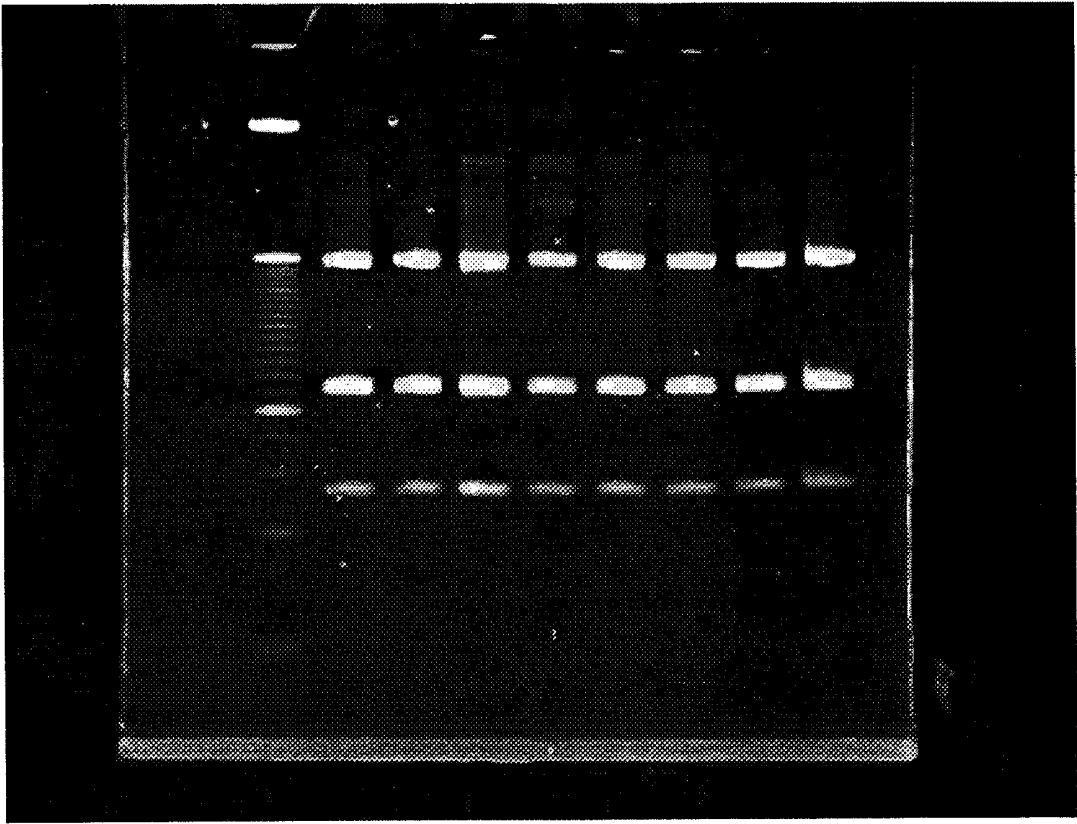


Fig. 1 J

11/19

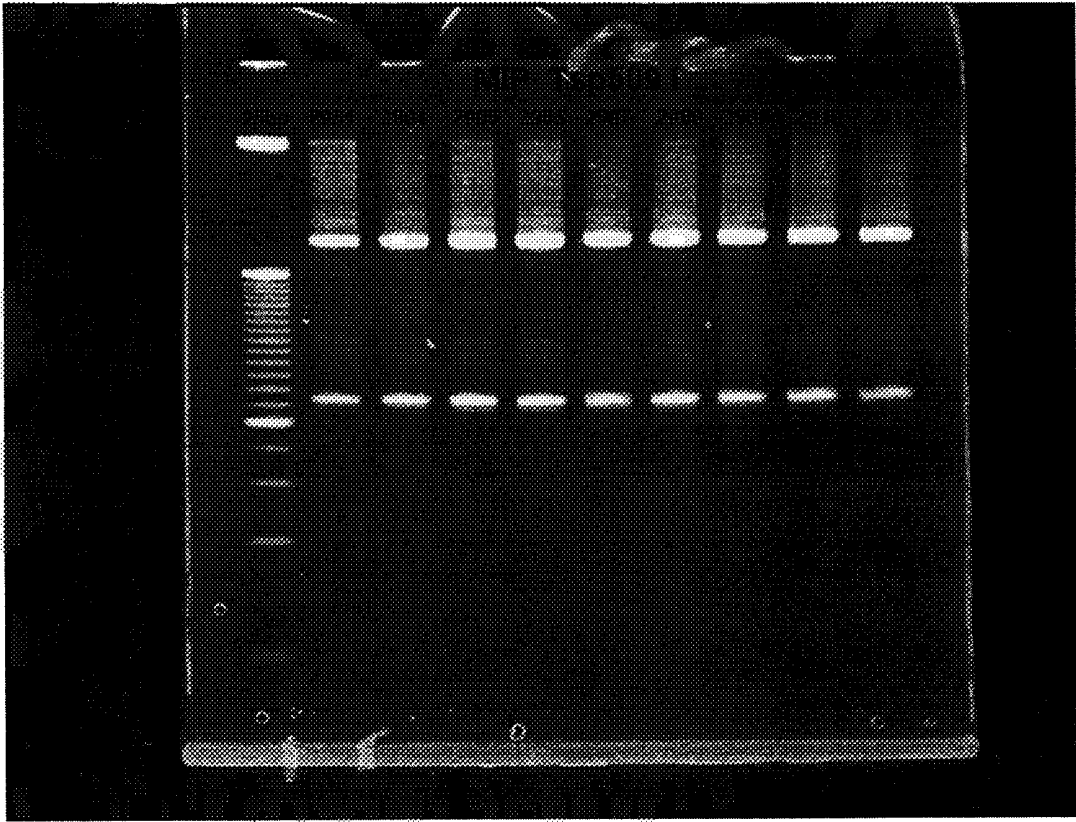


Fig. 1K

12/19

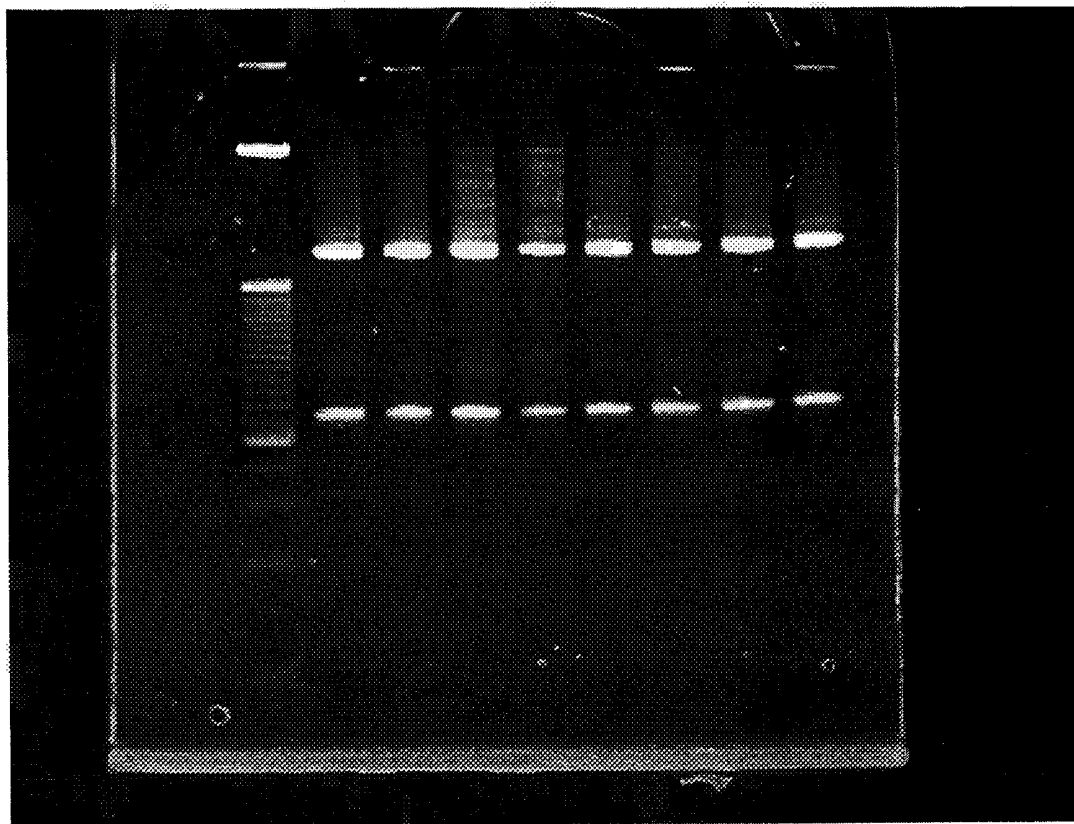


Fig. 1 L

Fig. 2 Alignment of the BASB023 polynucleotide sequences.
Identity to SeqID No:1 is indicated by a dot.

```

                *           20           *           40           *
Seqid1:ATGAAAAAACATGCACTGGTTGCCACCATGGCAGCAACATTGATTTGGT: 50
Seqid3:.....: 50

                60           *           80           *           100
Seqid1:AGGCTGTCAAAAAGACACTTCAGCTTCGCTGCCTAAAGCAGGCGAAAAGT:100
Seqid3:.....:100

                *           120           *           140           *
Seqid1:CAACCGTCGTTAGCGATAAAAAGCACAGAGATTGAACAAGTCAGCTATGTG:150
Seqid3:.....:150

                160           *           180           *           200
Seqid1:TTTGGCTATGACGCAGGCGAGTCTATGAAAAAATTGAAGAAAATTTAGA:200
Seqid3:.....:200

                *           220           *           240           *
Seqid1:CATCGATGTTTATATCAAAGCGTTTAAAGATGGCTATGCAGGTGTTGATA:250
Seqid3:.....:250

                260           *           280           *           300
Seqid1:GCGCTTTAACGAAAAAGCAGATTCAGACATTGGGTCAAGCATATGAAAAG:300
Seqid3:.....:300

                *           320           *           340           *
Seqid1:CGTAAAACCGAAGAAGCCATTCAAAAACAACAACAAGCTGCCGTTACCAA:350
Seqid3:.....:350
    
```

14/19

360 * 380 * 400
 Seqid1:TAAGGCTGATGGCGAAAAGTTTTGGCTGAAAATGCTAAAAAAGATGGCG:400
 Seqid3:.....:400

* 420 * 440 *
 Seqid1:TAAAGACAACACCACCAGGTCTTCAATATAAAGTTATTACTGAGGGTACG:450
 Seqid3:.....T.....:450

460 * 480 * 500
 Seqid1:GGCAAATCTCCCACGGCCAAAGATGGTGTATTATGCAGCTTATGAAGGCAG:500
 Seqid3:.....:500

* 520 * 540 *
 Seqid1:ATTGATTGATGGTACGGTATTTGATTCATCTGAAGGCGAGGCTGTGCCAT:550
 Seqid3:.....:550

560 * 580 * 600
 Seqid1:TTATGCTAAGTCAGGTCATTGAAGGCTGGTCTGAAGGCTTGCAACTGATG:600
 Seqid3:.....:600

* 620 * 640 *
 Seqid1:AAAGAAGGTGGCAAATATGAGCTGTATGTTCCAAGCCAAATGGCATATGG:650
 Seqid3:.....:650

660 * 680 * 700
 Seqid1:TGAACACGGTATGTACAATGCTGGCATTGGTCCAAACAGTTTTTGGTCT:700
 Seqid3:.....:700

* 720 * 740 *
 Seqid1:TTGTGATTGACCTAAAAAAGTTTCTGATGAAAAAGCCATAGCTGCCGAA:750
 Seqid3:.....:750

760 * 780 * 800
Seqid1:CAGCAGGCAATCATTGATGCTCAAATGCAAGCCATTCAAGAATCGCAAGG:800
Seqid3:.....:800

*
Seqid1:TCAACGATAG:810
Seqid3:.....:810

Fig. 3 Alignment of the BASB023 polypeptide sequences.
Identity to SeqID No:2 is indicated by a dot.

```

                *           20           *           40           *
Seqid2:MKKHALVATMAATLILVGCQKDTASALPKAGEKSTVVSDKSTEIEQVSYV: 50
Seqid4:.....: 50

                60           *           80           *           100
Seqid2:FGYDAGESMKKIEENLDIDVYIKAFKDG YAGVDSALTKKQIQTLGQAYEK:100
Seqid4:.....:100

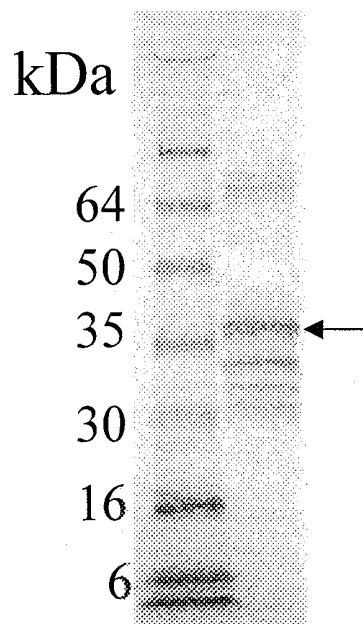
                *           120           *           140           *
Seqid2:RKTEEAIQKQQAAVTNKADGEKFLAENAKKDGVKTPPGLQYKVITEGT:150
Seqid4:.....S.....:150

                160           *           180           *           200
Seqid2:GKSPTAKDGVYAAAYEGR LIDGTVFDSSEGEAVPFMLSQVIEGWSEGLQLM:200
Seqid4:.....:200

                *           220           *           240           *
Seqid2:KEGGKYELYVPSQMAYGEHGMYNAGIGPNSVLV FVIDLKKVSDEKAIAAE:250
Seqid4:.....:250

                260
Seqid3:QQAII DAQMQA IQESGGQR :269
Seqid4:.....:269
    
```

Fig. 4 : Coomassie stained SDS-PAGE of recombinant purified BASB023



18/19

Fig. 5 : Western-blot of purified recombinant BASB023 protein probed with anti-His antibody

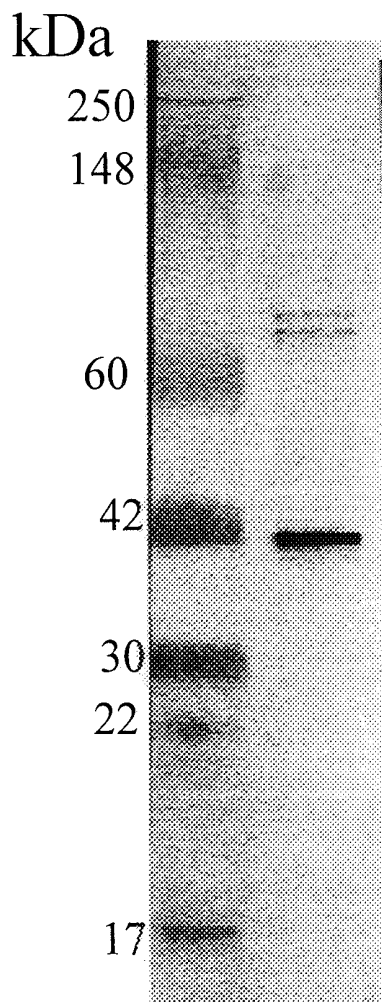
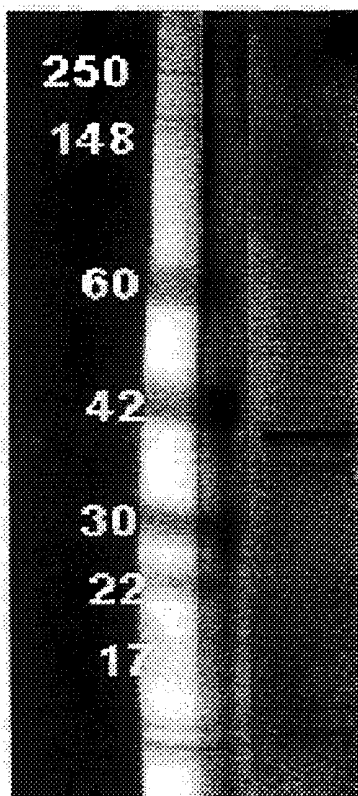


Fig. 6 : Western-blot of purified recombinant BASB023 protein probed with human convalescent sera.



SEQUENCE LISTING

<110> SmithKline Beecham Biologicals S.A.

<120> Novel Compounds

<130> BM45329

<160> 8

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 810

<212> DNA

<213> Bacteria

<400> 1

```

atgaaaaaac atgcactggt tgccaccatg gcagcaacat tgattttggt aggctgtcaa      60
aaagacactt cagcttcgct gcctaaagca ggcgaaaagt caaccgtcgt tagcgataaa      120
agcacagaga ttgaacaagt cagctatgtg tttggctatg acgcaggcga gtctatgaaa      180
aaaattgaag aaaatttaga catcgatggt tatatcaaag cgtttaaaga tggctatgca      240
gggtgtgata gcgctttaac gaaaaagcag attcagacat tgggtcaagc atatgaaaag      300
cgtaaaaccg aagaagccat tcaaaaacaa caacaagctg ccgttaccaaa taaggctgat      360
ggcgaaaagt ttttggctga aaatgctaaa aaagatggcg taaagacaac accaccaggt      420
cttcaatata aagttattac tgagggtagc ggcaaatctc ccacggccaa agatgggtgtt      480
tatgcagctt atgaaggcag attgattgat ggtacgggat ttgattcatc tgaaggcgag      540
gctgtgccat ttatgctaag tcaggtcatt gaaggctggt ctgaaggctt gcaactgatg      600
aaagaaggty gcaaatatga gctgtatggt ccaagccaaa tggcatatgg tgaacacggt      660
atgtacaatg ctggcattgg tccaaacagt gttttggtct ttgtgattga cctaaaaaaaa      720
gtttctgatg aaaaagccat agctgccgaa cagcaggcaa tcattgatgc tcaaatgcaa      780
gccattcaag aatcgcaagg tcaacgatag                                     810
    
```

<210> 2

<211> 269

<212> PRT

<213> Bacteria

<400> 2

```

Met Lys Lys His Ala Leu Val Ala Thr Met Ala Ala Thr Leu Ile Leu
 1                               10                15
Val Gly Cys Gln Lys Asp Thr Ser Ala Ser Leu Pro Lys Ala Gly Glu
      20                    25                30
Lys Ser Thr Val Val Ser Asp Lys Ser Thr Glu Ile Glu Gln Val Ser
      35                    40                45
Tyr Val Phe Gly Tyr Asp Ala Gly Glu Ser Met Lys Lys Ile Glu Glu
      50                    55                60
Asn Leu Asp Ile Asp Val Tyr Ile Lys Ala Phe Lys Asp Gly Tyr Ala
 65                    70                75                80
Gly Val Asp Ser Ala Leu Thr Lys Lys Gln Ile Gln Thr Leu Gly Gln
      85                    90                95
Ala Tyr Glu Lys Arg Lys Thr Glu Glu Ala Ile Gln Lys Gln Gln Gln
      100                   105                110
Ala Ala Val Thr Asn Lys Ala Asp Gly Glu Lys Phe Leu Ala Glu Asn
      115                   120                125
Ala Lys Lys Asp Gly Val Lys Thr Thr Pro Pro Gly Leu Gln Tyr Lys
      130                   135                140
Val Ile Thr Glu Gly Thr Gly Lys Ser Pro Thr Ala Lys Asp Gly Val
 145                   150                155                160
Tyr Ala Ala Tyr Glu Gly Arg Leu Ile Asp Gly Thr Val Phe Asp Ser
      165                   170                175
Ser Glu Gly Glu Ala Val Pro Phe Met Leu Ser Gln Val Ile Glu Gly
      180                   185                190
Trp Ser Glu Gly Leu Gln Leu Met Lys Glu Gly Gly Lys Tyr Glu Leu
      195                   200                205
Tyr Val Pro Ser Gln Met Ala Tyr Gly Glu His Gly Met Tyr Asn Ala
      210                   215                220
Gly Ile Gly Pro Asn Ser Val Leu Val Phe Val Ile Asp Leu Lys Lys
 225                   230                235                240
Val Ser Asp Glu Lys Ala Ile Ala Ala Glu Gln Gln Ala Ile Ile Asp
      245                   250                255
Ala Gln Met Gln Ala Ile Gln Glu Ser Gln Gly Gln Arg
      260                   265

```

<210> 3

<211> 810
 <212> DNA
 <213> Bacteria

<400> 3

```

atgaaaaaac atgcactggt tgccaccatg gcagcaacat tgattttggt aggctgtcaa      60
aaagacactt cagcttcgct gcctaaagca ggcgaaaagt caaccgtcgt tagcgataaa      120
agcacagaga ttgaacaagt cagctatgtg tttggctatg acgcaggcga gtctatgaaa      180
aaaattgaag aaaatTTaga catc gatggt tatatcaaag cgTTTaaaga tggctatgca      240
ggtgTTgata gcgctTTaac gaaaaagcag attcagacat tgggtcaagc atatgaaaag      300
cgtaaaaccg aagaagccat tcaaaaacaa caacaagctg ccgTTaccaa taaggctgat      360
ggcgaaaagt tTTTggctga aaatgctaaa aaagatggcg taaagacaac accatcaggt      420
cttcaatata aagTTattac tgagggtacg ggcaaatctc ccacggccaa agatggtgTT      480
tatgcagctt atgaaggcag attgattgat ggtacggtat ttgattcatc tgaaggcgag      540
gctgtgccat ttatgctaag tcaggtcatt gaaggctggt ctgaaggctt gcaactgatg      600
aaagaaggTg gcaaatatga gctgtatgTT ccaagccaaa tggcatatgg tgaacacggt      660
atgtacaatg ctggcattgg tccaaacagt gTTTTggtct ttgtgattga cctaaaaaaa      720
gtttctgatg aaaaagccat agctgccgaa cagcaggcaa tcattgatgc tcaaatgcaa      780
gccattcaag aatcgcaagg tcaacgatag                                     810
    
```

<210> 4
 <211> 269
 <212> PRT
 <213> Bacteria

<400> 4

```

Met Lys Lys His Ala Leu Val Ala Thr Met Ala Ala Thr Leu Ile Leu
 1           5           10           15
Val Gly Cys Gln Lys Asp Thr Ser Ala Ser Leu Pro Lys Ala Gly Glu
           20           25           30
Lys Ser Thr Val Val Ser Asp Lys Ser Thr Glu Ile Glu Gln Val Ser
           35           40           45
Tyr Val Phe Gly Tyr Asp Ala Gly Glu Ser Met Lys Lys Ile Glu Glu
           50           55           60
Asn Leu Asp Ile Asp Val Tyr Ile Lys Ala Phe Lys Asp Gly Tyr Ala
           65           70           75           80
Gly Val Asp Ser Ala Leu Thr Lys Lys Gln Ile Gln Thr Leu Gly Gln
           85           90           95
Ala Tyr Glu Lys Arg Lys Thr Glu Glu Ala Ile Gln Lys Gln Gln Gln
    
```

	100		105		110										
Ala	Ala	Val	Thr	Asn	Lys	Ala	Asp	Gly	Glu	Lys	Phe	Leu	Ala	Glu	Asn
	115		120		125										
Ala	Lys	Lys	Asp	Gly	Val	Lys	Thr	Thr	Pro	Ser	Gly	Leu	Gln	Tyr	Lys
	130		135		140										
Val	Ile	Thr	Glu	Gly	Thr	Gly	Lys	Ser	Pro	Thr	Ala	Lys	Asp	Gly	Val
145			150		155									160	
Tyr	Ala	Ala	Tyr	Glu	Gly	Arg	Leu	Ile	Asp	Gly	Thr	Val	Phe	Asp	Ser
			165		170									175	
Ser	Glu	Gly	Glu	Ala	Val	Pro	Phe	Met	Leu	Ser	Gln	Val	Ile	Glu	Gly
			180		185									190	
Trp	Ser	Glu	Gly	Leu	Gln	Leu	Met	Lys	Glu	Gly	Gly	Lys	Tyr	Glu	Leu
	195		200		205										
Tyr	Val	Pro	Ser	Gln	Met	Ala	Tyr	Gly	Glu	His	Gly	Met	Tyr	Asn	Ala
	210		215		220										
Gly	Ile	Gly	Pro	Asn	Ser	Val	Leu	Val	Phe	Val	Ile	Asp	Leu	Lys	Lys
225			230		235									240	
Val	Ser	Asp	Glu	Lys	Ala	Ile	Ala	Ala	Glu	Gln	Gln	Ala	Ile	Ile	Asp
			245		250									255	
Ala	Gln	Met	Gln	Ala	Ile	Gln	Glu	Ser	Gln	Gly	Gln	Arg			
	260		265												

<210> 5
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer sequence

<400> 5
 catagaaatc atgcagagac c

21

<210> 6
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Primer sequence

<400> 6

gagtcaatgc caatacggc

19

<210> 7

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 7

attacgcaga gccatgggct gtcaaaaaga cacttcagc

39

<210> 8

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 8

attacgcaga gctctagacc tcgttgacct tgcgattcct gaatg

45

SEQUENCE INFORMATION

BASB023 Polynucleotide and Polypeptide Sequences

SEQ ID NO:1

Moraxella catarrhalis BASB023 polynucleotide sequence from strain ATCC 43617

ATGAAAAACATGCACTGGTTGCCACCATGGCAGCAACATTGATTTTGGTAGGCTGTCAAAAAGACACTTCAGCTTCGCT
GCCTAAAGCAGGCGAAAAGTCAACCGTCGTTAGCGATAAAAAGCACAGAGATTGAACAAGTCAGCTATGTGTTTGGCTATG
ACGCAGGCGAGTCTATGAAAAAATGAAGAAAATTTAGACATCGATGTTTATATCAAAGCGTTTAAAGATGGCTATGCA
GGTGTGATAGCGCTTTAACGAAAAGCAGATTGACATTGGGTCAAGCATATGAAAAGCGTAAAACCGAAGAAGCCAT
TCAAAAACAACAAGCTGCCGTTACCAATAAGGCTGATGGCGAAAAGTTTGGCTGAAAATGCTAAAAAAGATGGCG

TAAAGACAACACCACCAGGTCTTCAATATAAAGTTATTACTGAGGGTACGGGCAAATCTCCACGGCCAAAGATGGTGTT
 TATGCAGCTTATGAAGGCAGATTGATTGATGGTACGGTATTTGATTTCATCTGAAGGCGAGGCTGTGCCATTTATGCTAAG
 TCAGGTCATTGAAGGCTGGTCTGAAGGCTTGCAACTGATGAAAGAAGGTGGCAAATATGAGCTGTATGTTCCAAGCCAAA
 TGGCATATGGTGAACACGGTATGTACAATGCTGGCATTGGTCCAAACAGTGTTTTGGTCTTTGTGATTGACCTAAAAAAA
 GTTTCTGATGAAAAAGCCATAGCTGCCGAACAGCAGGCAATCATTGATGCTCAAATGCAAGCCATTCAAGAATCGCAAGG
 TCAACGATAG

SEQ ID NO:2

Moraxella catarrhalis BASB023 polypeptide sequence deduced from the polynucleotide of
 SeQ ID NO:1

MKKHALVATMAATLILVGCQKDTASALPKAGEKSTVVSDKSTEIEQVSYVFGYDAGESMKKIEENLDIDVYIKAFKDG
 YAGVDSALTKKQIQTLGQAYEKRKTEEAIQKQQAAVTNKADGEKFLAENAKKDGVKTTTPGLQYKVIETEGTKSPTAKD
 GYAYAYEGRLLIDGTVFDSSEGEAVPFMLSQVIEGWSEGLQLMKEGGKYELYVPSQMAYGEHGMYNAGIGPNSVLV
 FVIDLKKVSDEKAIAAEQQAIIDAQMQA IQESQQR

SEQ ID NO:3

Moraxella catarrhalis BASB023 polynucleotide sequence from strain ATCC 43617

ATGAAAAACATGCACTGGTTGCCACCATGGCAGCAACATTGATTTTGGTAGGCTGTCAAAAAGACACTTCAGCTTCGCT
 GCCTAAAGCAGGCGAAAAGTCAACCGTCGTTAGCGATAAAAGCACAGAGATTGAACAAGTCAGCTATGTGTTGGCTATG
 ACGCAGGCGAGTCTATGAAAAAATTGAAGAAAATTTAGACATCGATGTTTATATCAAAGCGTTTAAAGATGGCTATGCA
 GGTGTTGATAGCGCTTTAACGAAAAAGCAGATTGAGACATTGGGTCAAGCATATGAAAAGCGTAAAACCGAAGAAGCCAT
 TCAAAAACAACAACAGCTGCCGTTACCAATAAGGCTGATGGCGAAAAGTTTTTGGCTGAAAATGCTAAAAAAGATGGCG
 TAAAGACAACACCATCAGGTCTTCAATATAAAGTTATTACTGAGGGTACGGGCAAATCTCCACGGCCAAAGATGGTGTT
 TATGCAGCTTATGAAGGCAGATTGATTGATGGTACGGTATTTGATTTCATCTGAAGGCGAGGCTGTGCCATTTATGCTAAG
 TCAGGTCATTGAAGGCTGGTCTGAAGGCTTGCAACTGATGAAAGAAGGTGGCAAATATGAGCTGTATGTTCCAAGCCAAA
 TGGCATATGGTGAACACGGTATGTACAATGCTGGCATTGGTCCAAACAGTGTTTTGGTCTTTGTGATTGACCTAAAAAAA
 GTTTCTGATGAAAAAGCCATAGCTGCCGAACAGCAGGCAATCATTGATGCTCAAATGCAAGCCATTCAAGAATCGCAAGG
 TCAACGATAG

SEQ ID NO:4

Moraxella catarrhalis BASB023 polypeptide sequence deduced from the polynucleotide of
 SeQ ID NO:3

MKKHALVATMAATLILVGCQKDTASALPKAGEKSTVVSDKSTEIEQVSYVFGYDAGESMKKIEENLDIDVYIKAFKDG
 YAGVDSALTKKQIQTLGQAYEKRKTEEAIQKQQAAVTNKADGEKFLAENAKKDGVKTTTPSGLQYKVIETEGTKSPTAKD
 GYAYAYEGRLLIDGTVFDSSEGEAVPFMLSQVIEGWSEGLQLMKEGGKYELYVPSQMAYGEHGMYNAGIGPNSVLV
 FVIDLKKVSDEKAIAAEQQAIIDAQMQA IQESQQR

SEQ ID NO:5

CAT AGA AAT CAT GCA GAG ACC

SEQ ID NO:6

GAG TCA ATG CCA ATA CGG C

SEQ ID NO:7

ATT ACG CAG AGC CAT GGG CTG TCA AAA AGA CAC TTC AGC

SEQ ID NO:8

ATT ACG CAG AGC TCT AGA CCT CGT TGA CCT TGC GAT TCT TGA ATG

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 99/05828

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/31 C07K14/21 A61K38/16 C07K16/12 G01N33/53		
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) IPC 7 C12N C07K A61K G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RADCLIFF, R.M. ET AL.: "Interspecies sequence differences in the Mip protein from the genus Legionella: implications for function and evolutionary relatedness." MOLECULAR MICROBIOLOGY, vol. 25, 1997, pages 1149-58, XP000863073 the whole document	1,2,13, 16-22, 26,27
A	WO 98 28333 A (COPE LESLIE D ;HANSEN ERIC J (US); AEBI CHRISTOPH (US); FISKE MICH) 2 July 1998 (1998-07-02) see whole document, particularly the claims.	
<input type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in 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 document 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 7 January 2000	Date of mailing of the international search report 14/01/2000	
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Smalt, R	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 99/05828

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9828333 A	02-07-1998	AU 5720198 A	17-07-1998
		EP 0948625 A	13-10-1999
