(11) Application No. AU 200023731 B2 (12) PATENT (10) Patent No. 776828 (19) AUSTRALIAN PATENT OFFICE (54)Streptococcus pneumoniae proteins and immunogenic fragments for vaccines (51) 6 International Patent Classification(s) A61K 039/09 A61P 031/04 A61K 039/40 CO7K 014/315 (21)Application No: 200023731 (22) Application Date: 1999.12.21 (87) WIPO No: WOOD/37105 Priority Data (30)Number Country (31) (32) Date (33)1998.12.21 60/113048 US (43)Publication Date : 2000.07.12 (43)Publication Journal Date: 2000.09.07 (44) Accepted Journal Date : 2004.09.23 (71)Applicant(s) MedImmune, Inc. (72)Inventor(s) Leslie S. Johnson; Scott Koenig; John E. Adamou Agent/Attorney (74) SPRUSON and FERGUSON, GPO Box 3898, SYDNEY 2001 NSW (56)Related Art

SEE REFERENCES OF WO 2000/037105 A3

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:		(1	1) International Publication Number:	WO 00/37105
A61K 39/09, 39/40, A61P 31/04 // C07K 14/315	A2	(4	3) International Publication Date:	29 June 2000 (29.06.00)
(21) International Application Number: PCT/US (22) International Filing Date: 21 December 1999 ((81) Designated States: 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, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE,	
(30) Priority Data: 60/113,048 21 December 1998 (21.12.9	21 December 1998 (21.12.98) US MMUNE, INC. (US/US); 35 West Watkins hersburg, MD 20878 (US). SON, Leslie, S.; 13545 Ambassador Drive,		SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MY, SD, SL, SZ, TZ, UG, ZW). Eursain patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, 1M), 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).	
(71) Applicant: MEDIMMUNE, INC. [US/US]; 35 West Mill Road, Gaithersburg, MD 20878 (US).				
Germantown, MD 20874 (US). KOENIG, Sco Ralston Road, Rockville, MD 20852 (US). Al John, E.; 20822 Shamrock Glen Circle, Germant 20874 (US).	tt; 107 DAMO	732 DU,	Published Without international search rep upon receipt of that report.	ort and to be republished
(74) Agents: GRANT, Alan, J. et al.; Carella, Byrne, Bain, Cecchi, Stewart & Olstein, 6 Becker Farm Road, NJ 07068 (US).				
(54) Title: STREPTOCOCCUS PNEUMONIAE PROTE	INS AN	ND	IMMUNOGENIC FRAGMENTS FOR VA	CCINES
(57) Abstract				
A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.				

STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

FIELD OF THE INVENTION

5

10

15

This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species with a polypeptide comprising at least one conserved histidine triad residue (HxxHxH) and at least one helix-forming polypeptide obtained from *Streptococcus pneumoniae* as a mechanism for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic *S. pneumoniae*. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

In a particular aspect, the present invention relates to the prevention and treatment of pneumococcal infections such as infections of the middle ear, nasopharynx, lung and bronchial areas, blood, CSF, and the like, that are caused by pneumococcal bacteria.

Background of the Invention

Streptococcus pneumoniae is a gram positive bacteria which is a major causative agent in invasive infections in animals and humans, such as sepsis, meningitis, otitis media and lobar pneumonia (Tuomanen et al. New Engl. J. Med. 322:1280-1284 (1995)). As part of the infective process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell et al., Micro. Path. 17:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell et al., Nature, 377:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila et al., J. Inf. Dis., 176:704-712 (1997)). A number of various other proteins have been suggested as being involved in the pathogenicity of S. pneumoniae. There remains a need for identifying polypeptides having epitopes in common from various strains of S. pneumoniae in order to utilize such polypeptides as vaccines to provide protection against a wide variety of S. pneumoniae.

Summary of the Invention

According to a first embodiment of the invention, there is provided a composition comprising a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

According to a second embodiment of the invention, there is provided a composition comprising an active fragment of a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 wherein said active fragment comprises at least two coiled coil regions and is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to said polypeptide when said composition is administered to a mammal.

II:\DavLib\LIBFF185145spec.doc:gcc

20

25

:··:

According to a third embodiment of the invention, there is provided a composition comprising a polypeptide comprising the sequence of amino acid residues 139-791 of SEQ ID NO:4.

According to a fourth embodiment of the invention, there is provided a method for preventing infection caused by *S. pneumoniae* comprising administering the vaccine comprising the composition in accordance with any one of the first to third embodiments of the present invention.

According to a fifth embodiment of the invention, there is provided an isolated polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 wherein said polypeptide binds to an antibody that binds to *S. pneumoniae*.

According to a sixth embodiment of the invention, there is provided a method for producing a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4 comprising expressing said polypeptide from a cell that comprises a polynucleotide containing a nucleotide sequence that encodes said polypeptide.

According to a seventh embodiment of the invention, there is provided a method for producing an active fragment of the polypeptide that comprises amino acids 20-838 of SEQ ID NO:4 comprising expressing said active fragment from a cell that contains a polynucleotide containing a nucleotide sequence that encodes said active fragment.

According to an eighth embodiment of the invention, there is provided an isolated polynucleotide comprising a nucleotide sequence having at least 90% identity to the sequence of SEQ ID NO:5.

According to a ninth embodiment of the invention, there is provided a vector comprising the isolated polynucleotide in accordance with the eighth embodiment of the present invention.

According to a tenth embodiment of the invention, there is provided a method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.

According to an eleventh embodiment of the invention, there is provided a method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.

[R:\LIBFF]99631spec.doc:GCC

According to a twelfth embodiment of the invention, there is provided a method for eliciting an immune response to *Streptococcus pneumoniae* in a mouse comprising administering to said mouse a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.

According to a thirteenth embodiment of the invention, there is provided a method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an active fragment of amino acids 20-838 of SEQ ID NO:4.

According to a fourteenth embodiment of the invention, there is provided a method for eliciting an immune response to *Streptococcus pneumoniae* in a mouse comprising administering to said mouse a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.

According to a fifteenth embodiment of the invention, there is provided a composition comprising a polypeptide having at least 90% identity to amino acids 21-480 of SEQ ID NO:6,

wherein said polypeptide is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

According to a sixteenth embodiment of the invention, there is provided a composition comprising an active fragment of the polypeptide comprising amino acids 21-480 of SEQ ID NO:6,

wherein said active fragment comprises at least about 10% and no more than about 85% of said polypeptide, wherein said active fragment comprises at least one histidine triad and at least one coiled coil region and wherein said active fragment is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

According to a seventeenth embodiment of the invention, there is provided a vaccine comprising a polypeptide containing the sequence of amino acids 21-480 of SEQ ID NO:6 in a pharmaceutically acceptable carrier in an amount effective to protect against pneumococcal infection when administered to a mammal.

According to an eighteenth embodiment of the invention, there is provided an isolated polypeptide comprising at least 90% identity to amino acids 21-480 of SEQ ID NO:6, wherein said polypeptide binds to an antibody that binds to S. pneumoniae.

[R:\LIBFF]99631spec.doc:GCC

According to a nineteenth embodiment of the invention, there is provided a composition comprising at least one antibody that binds to a polypeptide comprising an amino acid sequence at least 90% identical to a member selected from the group consisting of:

- (i) a sequence having at least 90% identity to amino acids 20-838 of SEQ ID NO:4 and
 - (ii) a sequence comprising amino acids 21-480 of SEQ ID NO:6,

wherein said antibody is in a pharmaceutically acceptable carrier.

According to a twentieth embodiment of the invention, there is provided a composition comprising at least one antibody that binds to an active fragment of a polypeptide comprising a member selected from the group consisting of:

- (i) amino acids 20-838 of SEQ ID NO:4 and
- (ii) amino acids 21-480 of SEQ ID NO:6,

5

25

wherein said active fragment comprises at least about 10% and no more than about 85% of said polypeptide, wherein said active fragment comprises at least one histidine triad and at least one coiled coil region and wherein said active fragment is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal

According to a twenty-first embodiment of the invention, there is provided a process for treating or preventing infection caused by *S. pneumoniae* comprising administering to a mammal afflicted with said infection, or at risk of said infection, an effective amount of the composition in accordance with the twentieth embodiment of the present invention.

According to a twenty-second embodiment of the invention, there is provided a process for treating or preventing infection caused by *S. pneumoniae* comprising administering to a mammal afflicted with said infection, or at risk of said infection, an effective amount of the composition in accordance with the twenty-first embodiment of the present invention.

According to a twenty-third embodiment of the invention, there is provided the use of a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4 for the manufacture of a medicament for eliciting an immune response to *Streptococcus pneumoniae* in a mammal.

According to a twenty-fourth embodiment of the invention, there is provided the use of a polypeptide comprising an active fragment of amino acids 20-838 of SEQ ID

[R:\LIBFF]99631spec.doc:GCC

NO:4 for the manufacture of a medicament for eliciting an immune response to Streptococcus pneumoniae in a mammal.

According to a twenty-fifth embodiment of the invention, there is provided the use of the polypeptide in accordance with the twenty-fourth embodiment of the present invention wherein said active fragment is a contiguous segment of amino acids 20-838 of SEQ ID NO:4 and at least 42 amino acid residues in length and not more than 356 residues in length.

In accordance with the present invention, there is provided vaccines and

[R:\LIBFF]99631spec.doc:GCC

vaccine compositions that include polypeptides obtained from *S. pneumoniae* and/or variants of said polypeptides and/or active fragments of such polypeptides.

The active fragments, as hereinafter defined, include a histidine triad residue(s) and/or coiled coil regions of such polypeptides.

The term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence from an alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The percent identity is determined as follows:

Percent Identity = [1- (C/R)] 100

15

20

10

5

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of the alignment between the Compared Sequence and the Reference Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have an aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, each being a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence in which the Percent Identity as calculated above is about

equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum Percent Identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

"Isolated" in the context of the present invention with respect to polypeptides and/or polynucleotides means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

10

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A-1C, respectively, report the results of three experiments using different preparations of SP36. The results demonstrate that active immunization with recombinant SP36 derived from pneumococcal strain Norway serotype 4 is able to protect mice from death in a model of pneumococcal sepsis using a heterologous strain, SJ2 (serotype 6B). In each of the three experiments shown, one hundred percent of the mice immunized

with SP36 survived for the 14-day observation period following challenge with approximately 500 cfu of pneumococci, while eighty to one hundred percent of sham-immunized mice (injected with PBS and adjuvant) died during the same period.

5

25

Figures 2A-2B show that passive administration of rabbit antiserum raised against Sp36 derived from Norway type 4 was able to protect mice in the pneumococcal sepsis model using two heterologous strains. Figure 2A shows that one hundred percent of the mice immunized with the SP36 antiserum survived the 21-day observation period after challenge with 172 CFU of strain SJ2 (serotype 6B). Eighty percent of the mice immunized with a control serum (rabbit anti-FimC) died by day 8, and ninety percent died by day 12. Figure 2B shows that 90 percent of the mice immunized with the Sp36 antiserum survived the 8-day observation after challenge with 862 CFU of strain EF6796 (serotype 6A). Ninety percent of the mice immunized with a control serum (collected before immunization) died by day 5.

Figure 3 is a western blot demonstrating the ability of antisera raised against recombinant Sp36 derived from strain Norway type 4 to react with Sp36 of heterologous strains. Total cell lysates were immunoblotted with mouse antisera to Sp36. A band representing Sp36 protein was detected in all 23 *S. pneumoniae* strains tested, which included isolates from each of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

Figure 4 is a Southern blot showing that the Sp36 gene from Norway type 4 hybridizes with genomic DNA from 24 other pneumococcal strains, indicating the presence of similar sequences in all these strains.

Figure 5 is a western blot showing the reactivity of patient sera with Sp36. Sp36 (either full-length, panel A; N-terminal half, panel B; or C-terminal half, panel C) was electrophoresed by SDS-PAGE and transferred to nitrocellulose. Patient sera collected soon after the onset of illness (acute serum, lanes A) or eight to 30 days later (convalescent serum, lanes C) were used to probe the blots. For patients 2, 3, and 5, convalescent serum reacted more strongly with Sp36 than did the corresponding acute serum.

Figure 6 is an amino acid alignment comparison of four related pneumococcal proteins, namely Sp36A (PhtA; SEQ ID NO:8), Sp36B (PhtB; SEQ ID NO:10), Sp36D (PhtD; SEQ ID NO:4), Sp36E (PhtE; SEQ ID NO:6), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity. Amino acid residues that match are boxed.

15

20

10

Figure 7 is a nucleotide alignment comparison of four related pneumococcal genes, namely Sp36A (PhtA; SEQ ID NO:9), Sp36B (PhtB; SEQ ID NO:11), Sp36D (PhtD; SEQ ID NO:5), Sp36E (PhtE; SEQ ID NO:7), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity.

Figure 8 shows the results of immunization of mice with PhtD recombinant protein, which leads to protection from lethal sepsis. C3H/HeJ (Panel A and B) or Balb/cByJ (Panel C) mice were immunized subcutaneously with PhtD protein (15 μ g in 50 μ l PBS emulsified in 50 μ l complete Freund's adjuvant (CFA)). The recombinant PhtD protein used in protection experiments consisted of 819 amino acid residues, starting with the cysteine

(residue 20). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 µg protein with incomplete Freund's adjuvant (IFA) was administered 3 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at week 7; and sera from each group was pooled for analysis of anti-PhtD antibody by ELISA. Mice were challenged at week 8 by an intraperitonial (i.p.) injection of approximately 550 CFU S. pneumoniae strain SJ2, serotype 6B (Panel A), 850 CFU of strain EF6796, serotype 6A (Panel B) or 450 CFU of strain EF5668, serotype 4 (Panel C). In preliminary experiments, the $\ensuremath{\text{LD}_{\text{50}}}$ for strain 10 SJ2 and EF6796 were determined to be approximately 10 CFU for both strains. The LD_{50} for strain EF5668 was determined to be < 5 CFU. Survival was determined in all groups over the course of 15 days following challenge. Data are presented as the percent survival for a total of 10 mice per experimental group. Two-sample Log-rank test was used for statistical analysis comparing recombinant Pht immunized mice to sham-immunized mice.

SUMMARY OF THE INVENTION

20

In accordance with one aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes at least one polypeptide that is at least 90% identical to (c) a polypeptide

sequence comprising amino acids 20-838 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 21-480 of SEQ ID NO:6 or an active fragment of the foregoing.

In accordance with another aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes an active fragment of a polypeptide that is at least 90% identical to (i) a polypeptide comprising amino acids 20-819 of SEQ ID NO:8 or (ii) a polypeptide

comprising amino acids 20-819 of SEQ ID NO:10.

10

The term "active fragment" means a fragment that includes one or more histidine triad residues and/or one or more coiled coil regions. A "histidine triad residue" is the portion of the polypeptide that has the sequence HxxHxH wherein H is histidine and x is an amino acid other than histidine

15

A coiled coil region is the region predicted by "Coils" algorithm: Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting Coiled Coils from Protein Sequences, *Science* 252:1162-1164.

20

In accordance with one embodiment, the active fragment includes both one or more histidine triad residues and at least one coiled coil region of the applicable polypeptide sequence. In accordance with another embodiment, the active fragment includes at least two histidine triad residues.

25

In another embodiment, the active fragment that includes at least one histidine triad residue or at least one coiled-coil region of the applicable polypeptide includes at least about ten percent of the applicable polypeptide and no more than about 85% of the applicable polypeptide.

The polypeptide of SEQ ID NO:4 includes five histidine triad residues, as follows: amino acids 83-88, 207-212, 315-320, 560-565 and 644-649.

The polypeptide of SEQ ID NO:6 includes five histidine triad residues, as follows: amino acids 83-88, 205-210, 309-314, 396-401 and 461-466.

In addition, the polypeptide of SEQ ID NO:4 includes two coiled-coil regions (amino acids 139-159 and amino acids 769-791) and the polypeptide of SEQ ID NO:6 includes one coiled-coil region (amino acids 139-172.)

The polypeptide of SEQ ID NO:8 includes the following regions:

HxxHxH: amino acids 82-87, 208-213, 328-333, 569-574 and 653-658.

Coiled-coils: amino acids 137-164, 425-453, 481-512 and 743-770.

In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by *S. pneumoniae*.

A vaccine, or vaccine composition, in accordance with the present invention may include one or more of the hereinabove described polypeptides or active fragments thereof. When employing more than one polypeptide or active fragment, such two or more polypeptides and/or active fragments may be used as a physical mixture or as a fusion of two or more polypeptides or active fragments. The fusion fragment or fusion polypeptide may be produced,

for example, by recombinant techniques or by the use of appropriate linkers for fusing previously prepared polypeptides or active fragments.

In an embodiment of the invention, there is provided (a) a polypeptide that is at least 95% identical or at least 97% identical or 100% identical to (i) a polypeptide sequence comprising amino acids 20-838 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 21-480 of SEQ ID NO:6; or (b) an active fragment of the polypeptide of (a).

In the case where the polypeptide is a variant of the polypeptide comprising the mature polypeptide of SEQ ID NO:4 or SEQ ID NO:6, or any of the active fragments of the invention, the variation in the polypeptide or fragment is generally in a portion thereof other than the histidine triad residues and the coiled-coil region, although variations in one or more of these regions may be made.

10

15

..... 20

25

In many cases, the variation in the polypeptide or active fragment is a conservative amino acid substitution, although other substitutions are within the scope of the invention.

In accordance with the present invention, a polypeptide variant includes variants in which one or more amino acids are substituted and/or deleted and/or inserted.

In another aspect, the invention relates to passive immunity vaccines formulated from antibodies against a polypeptide or active fragment of a polypeptide of the present invention. Such passive immunity vaccines can be utilized to prevent and/or treat pneumococcal infections in patients. In this manner, according to a further aspect of the invention, a vaccine can be

produced from a synthetic or recombinant polypeptide of the present invention or an antibody against such polypeptide.

In still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of diseases that are caused by pneumococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or treatment of infectious diseases that are caused by *S. pneumoniae*. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of otitis media, nasopharyngeal, bronchial infections, and the like in humans by utilizing a vaccine of the present invention.

Generally, vaccines are prepared as injectables, in the form of aqueous solutions or suspensions. Vaccines in an oil base are also well known such as for inhaling. Solid forms which are dissolved or suspended prior to use may also be formulated. Pharmaceutical carriers are generally added that are compatible with the active ingredients and acceptable for pharmaceutical use. Examples of such carriers include, but are not limited to, water, saline solutions, dextrose, or glycerol. Combinations of carriers may also be used.

Vaccine compositions may further incorporate additional substances to stabilize pH, or to function as adjuvants, wetting agents. or emulsifying agents, which can serve to improve the effectiveness of the vaccine.

20

25

Vaccines are generally formulated for parental administration and are injected either subcutaneously or intramuscularly. Such vaccines can also be formulated as suppositories or for oral administration, using methods known in the art.

The amount of vaccine sufficient to confer immunity to pathogenic bacteria is determined by methods well known to those skilled in the art. This quantity will be determined based upon the characteristics of the vaccine recipient and the level of immunity required. Typically, the amount of vaccine to be administered will be determined based upon the judgment of a skilled physician. Where vaccines are administered by subcutaneous or intramuscular injection, a range of 50 to 500 µg purified protein may be given.

The present invention is also directed to a vaccine in which a polypeptide or active fragment of the present invention is delivered or administered in the form of a polynucleotide encoding the polypeptide or active fragment, whereby the polypeptide or active fragment is produced *in vivo*. The polynucleotide may be included in a suitable expression vector and combined with a pharmaceutically acceptable carrier.

10

20

25

In addition, the polypeptides of the present invention can be used as immunogens to stimulate the production of antibodies for use in passive immunotherapy, for use as diagnostic reagents, and for use as reagents in other processes such as affinity chromatography.

In another aspect the present invention provides polynucleotides which encode the hereinabove described polypeptides and active fragments of the invention. The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand.

in accordance with another aspect of the present invention, there is

provided

10

20

25

(A) an isolated polynucleotide that is at least 90% identical to a polynucleotide sequence encoding (i) a polypeptide comprising amino acids 20-838 of SEQ ID NO:4 or (ii) a polypeptide comprising amino acids 21-480 of SEQ ID NO:6, or

(B) a fragment of the polynucleotide of (A) that encodes an active polypeptide fragment or

(C) a polynucleotide that is at least 90% identical to a polynucleotide sequence encoding an active fragment of (i) a polypeptide comprising amino acids 20-819 of SEQ ID NO:8 or (ii) a polypeptide comprising amino acids 20-819 of SEQ ID NO:10.

In specific embodiments, the polynucleotide is at least 95% identical, preferably at least 97% identical, and even 100% identical to such polynucleotide sequence.

The term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of polynucleotides. The variants of the polynucleotides may be a naturally occurring allelic variant of the polynucleotides or a non-naturally occurring variant of the polynucleotides. The variants include variants in which one or more bases are substituted, deleted or inserted. Complements to such coding polynucleotides may be utilized to isolate polynucleotides encoding the same or similar polypeptides. In particular, such procedures are useful to obtain native immunogenic portions of polypeptides from different serotypes of *S. pneumoniae*, which is especially

useful in the production of "chain" polypeptide vaccines containing multiple immunogenic segments.

SEQ ID NO:5 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:4 and SEQ ID NO:7 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:6. SEQ ID NO:9 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:8, and SEQ ID NO:11 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:10. As a result of the known degeneracy of the genetic code, other polynucleotides that encode the polypeptides of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10 should be apparent to those skilled in the art from the teachings herein.

The polynucleotides encoding the immunogenic polypeptides described above may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptides of the present invention. The marker sequence may be, for example, a hexa-histidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptides fused to the marker in the case of a bacterial host, or, for 20 example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

15

25

The present invention also relates to vectors which include polynucleotides encoding one or more of the polypeptides of the invention, host cells which are genetically engineered with vectors of the invention and the production of such immunogenic polypeptides by recombinant techniques in an isolated and substantially immunogenically pure form.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors comprising a polynucleotide encoding a polypeptide of the invention. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the polynucleotides which encode such polypeptides. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

10

20

25

Vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art.

Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the $\underline{\text{E. coli. lac}}$ or $\underline{\text{trp}}$, the phage lambda P_L promoter and other promoters known to control expression of genes in

prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

5

10

15

20

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in $\underline{\textbf{E}}$. $\underline{\textbf{coli}}$.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the proteins.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as <u>E. coli</u>, <u>Streptomyces</u>, <u>Salmonella typhimurium</u>; fungal cells, such as yeast; insect cells such as <u>Drosophila S2</u> and <u>Spodoptera Sf9</u>; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter,

operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen, Inc.), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

10

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and TRP. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

20

25

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to

produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

5

15

20

25

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of <u>E. coli</u> and <u>S. cerevisiae</u> TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with

translation initiation and termination sequences. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

5

15

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

25

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

5

25

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, a french press, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art. However, preferred are host cells which secrete the polypeptide of the invention and permit recovery of the polypeptide from the culture media.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV4O splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The polypeptides can be recovered and/or purified from recombinant cell cultures by well-known protein recovery and purification methods. Such methodology may include ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity

chromatography, hydroxylapatite chromatography and lectin chromatography.

Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. In this respect, chaperones may be used in such a refolding procedure. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides that are useful as immunogens in the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated.

Procedures for the isolation of the individually expressed polypeptides may be isolated by recombinant expression/isolation methods that are well-known in the art. Typical examples for such isolation may utilize an antibody to a conserved area of the protein or to a His tag or cleavable leader or tail that is expressed as part of the protein structure.

20

15

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a

sequence of the-present invention can be obtained by direct injection of the polypeptides into an animal.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The invention will be further described with respect to the following examples; however, the scope of the invention is not limited thereby:

20 Example 1

25

Active Protection with Anti-Sp36

A. Cloning, expression, and purification of SP36

The genomic DNA used as target for amplification was isolated from *S. pneumoniae* Norway strain (serotype 4), the same strain used for genomic sequencing. The complete sequence of the Sp36 gene (SEQ ID NO:9), and its predicted amino acid sequence (SEQ ID NO:8), are given in the Sequence Listing appended hereto. It was noted that the predicted amino acid

sequence included a hydrophobic leader sequence followed by a sequence (LSVC) similar to the consensus sequence for Type II signal peptidase (LxxC, in which both x's typically represent small amino acids). Primers (listed as SEQ ID NOS:1-3) were designed that would amplify the Sp36 gene and allow its cloning into pQE10 and expression as a histidine-tagged protein lacking the signal sequence for purification by nickel-affinity chromatography. Cloning of the fragment amplified by SEQ ID Nos 1 and 3 would result in a protein containing amino acids 21 through 819 of Sp36; cloning of the fragment amplified by SEQ ID Nos 2 and 3 would result in a protein containing amino acids 26 through 819 of Sp36 (amino acid numbers refer to SEQ ID NO:8).

B. Active Protection With Sp36 Vaccination

15

In each of the three experiments shown in Figures 1A-1C, C3H/HeJ mice (10/group) were immunized intraperitoneally (i.p.) with Sp36 protein (15 µg in 50 µl PBS emulsified in 50 µl complete Freund's adjuvant (CFA)). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 µg protein with incomplete Freund's adjuvant (IFA) was administered 4 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at weeks 3, 6, and 9; and sera from each group were pooled for analysis of anti-Sp36 antibody by ELISA. Mice were challenged at week 10 by an i.p. injection of approximately 500 CFU *S. pneumoniae* strain SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research Hospital, Memphis, TN). In preliminary experiments, the LD₅₀ of this strain was determined to be approximately 10 CFU. Mice were monitored for 14 days for survival.

The three experiments shown in Figures 1A-1C used slightly different

preparations of recombinant Sp36. The experiments shown in Figure 1A and 1B both used Sp36 containing amino acids 20-815, but different batches of protein were used in the two experiments. The experiment shown in Figure 1C used Sp36 containing amino acids 25-815.

5

10

20

In the experiment shown in Figure 1A, 9-week sera collected from the ten mice immunized with Sp36 (first batch) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimmunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (520 cfu of pneumococci) for 14 days. Eighty percent of sham-immunized mice were dead by day 4, and the remainder survived.

In the experiment shown in Figure 1B, 9-week sera collected from the ten mice immunized with Sp36 (second batch) had an endpoint ELISA titer of >1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimmunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, eighty percent were dead by day 4, and all died by day 9.

In the experiment shown in Figure 1C, 9-week sera collected from the ten mice immunized with Sp36 (containing amino acids 25- 815) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from sham-immunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, ninety percent died by day 4, and all died by day 12. These data demonstrate that immunization of mice with recombinant Sp36 proteins elicits a response capable of

protecting against systemic pneumococcal infection and death. This protection was not strain-specific: the recombinant pneumococcal protein was cloned from a serotype 4 strain, while the challenge was with a heterologous strain, SJ2 (serotype 6B).

5

Example 2

Passive Protection with Anti-Sp36 Antisera

A. Generation of Rabbit Immune Sera

10

Following collection of preimmune serum, a New Zealand White rabbit was immunized with 250 μg of Sp36 (containing amino acids 20-815) in CFA. The rabbit was given two boosts of 125 μg Sp36 in IFA on days 29 and 50 and bled on days 39 and 60. A second rabbit was immunized with a control antigen, *E. coli* FimC.

B. Passive Protection in Mice

C3H/HeJ mice (10 mice/group) were passively immunized by two i.p. injections of 100 µl of rabbit serum. The first injection was administered twenty-four hours before challenge with 172 cfu of *S. pneumoniae* strain SJ2, and the second injection was given four hours after challenge. Figure 2 shows the survival of mice after infection with two different strains of pneumococci.

25

Figure 2A shows that of mice injected with 172 cfu of strain SJ2 (Figure 2A), one hundred percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 21-day observation period.

Of the mice immunized with the control serum (anti-FimC), eighty percent

died by day 8, and ninety percent died by day 12. Figure 2B shows that of mice injected with 862 cfu of strain EF6796, ninety percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 8-day observation period. Of those given a control serum (collected from a rabbit before immunization), ninety percent died by day 8.

These data indicate that the protection against pneumococcal infection resulting from immunization with Sp36 is antibody-mediated, since mice can be protected by passive transfer of serum from a hyperimmunized rabbit. As seen in the mouse active challenge experiments described above, serum directed against recombinant Sp36 protein cloned from a serotype 4 strain was protective against challenge with heterologous strains.

Example 3

25

15 Conservation of Sp36 Among Strains of S. pneumoniae

A. Western blotting

The 23 pneumococcal strains used in this experiment were obtained from the American Type Culture Collection (Rockville, MD) and include one isolate each of the 23 serotypes in the multivalent pneumococcal vaccine. For total cell lysates, pneumococci were grown to mid-logarithmic phase (optical density at 620 nm, 0.4 to 0.6) in 2 ml Todd-Hewitt broth with 0.5% yeast extract (Difco, Detroit, ME) at 37°C. Bacteria were harvested by centrifugation and washed twice with water. Pellets were resuspended in 200 µl lysis buffer (0.01% sodium dodecyl sulfate, 0.15 M sodium citrate and 0.1% sodium deoxycholate) and incubated at 37°C for 30 min, then diluted in an equal volume 2x SSC (0.3 M sodium chloride, 0.03 M sodium citrate). Lysates were separated by SDS-PAGE, transferred to nitrocellulose

membranes (Bio-Rad Laboratories, Hercules, CA), and probed with antibody in a standard Western blotting procedure. Sera from ten C3H/HeJ mice immunized with Sp36 (as described in Example 1) were pooled and used at a dilution of 1:3000. Bound antibody was detected with peroxidase-conjugated sheep anti-mouse IgG using the chemiluminescence kit from Amersham, Inc. (Cambridge, MA).

The mouse anti-Sp36 sera detected two major bands with apparent molecular weights of 97 and 100 kDa in all 23 pneumococcal lysates tested (shown in Figure 3). The Sp36 signals obtained from *S. pneumoniae* serotypes 1, 5, 17F and 22F were lower, indicating either that the level of Sp36 expression is reduced in these strains, or that Sp36 in these strains is antigenically different.

These data show that Sp36 is antigenically conserved among strains of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

20 B. Southern blotting

15

Genomic DNA was prepared from each of the 23 pneumococcal strains listed in the previous section and also from strain SJ2. DNA was digested with *Pvull* and *BamHI*, electrophoresed in an agarose gel and transferred to a nylon membrane. A probe was prepared by amplifying the Sp36 gene from Norway type 4 DNA (as in Example 1) and labeling the amplified fragment with fluorescein by the random-priming method, using a kit from Amersham. Hybridization, washing, and exposure of film were carried out as in the protocol supplied by Amersham. Figure 4 shows that

the Sp36 probe hybridized with DNA from each of the 24 strains studied. The lane marked "M" contained DNA from lambda phage, digested with HindIII and labeled with fluorescein, as molecular weight markers.

5 Example 4

10

15

Immunogenicity of Sp36 in Humans

In order to determine whether Sp36 is immunogenic during human with culture-proven patients pneumococcal infection, from sera pneumococcal bacteremia were used in Western blots containing recombinant Sp36 protein. In the experiment shown in Figure 5, sera from five patients (indicated as 1 through 5) were diluted 1:3000 and used to probe blots containing full-length Sp36, the N-terminal half of Sp36 (preceding the proline-rich region), or the C-terminal half of Sp36 (following the proline-rich region). Lanes labeled A (acute) were probed with serum collected shortly after diagnosis of pneumococcal infection; lanes C (convalescent) were probed with serum collected either one month later (patients 1, 2, and 3) or eight days after the first serum collection (patients 4 and 5). For patients 2, 3 and 5, reactivity of the convalescent serum with Sp36 was stronger that that of the corresponding acute serum. The difference between the acute and convalescent sera was particularly evident for reactivity with the C-terminal half of the protein.

In additional experiments (not shown), convalescent sera from 23 patients with pneumococcal infections were tested individually for reactivity with full-length Sp36: 20 of the 23 sera were found to bind Sp36 on a Western blot.

These experiments indicate that Sp36 is recognized by the human

immune system and suggest that antibodies able to bind the Sp36 protein may be produced during natural *S. pneumoniae* infection in humans. Since the patients were infected with a variety of pneumococcal strains, these data also support the idea that Sp36 is antigenically conserved.

Example 5

5

10

15

Table 1 provides the percent identity between the various sequences.

Alignment of the predicted amino acid sequences of PhtA, PhtB, PhtD, and PhtE using the MEGALIGN program of Lasergene showed strong N-terminal homology with substantial divergence of the C-termini (Figure 6). The alignment of the nucleotide sequences of the same genes is shown in Figure 7. Amino acid and nucleotide sequences were compared using the identity weighting in a Lipman-Pearson pairwise alignment, in which the number of matching residues is divided by the total of matching residues plus the number of mismatched residues plus the number of residues in gaps. In the table below, the percent identity between each pair of sequences is shown at the intersection of the corresponding row and column.

20

25

:···: :···:

····:

Example 6

Active Protection with PhtD Vaccination.

Mice immunized with recombinant PhtD derived from strain N4 generated potent antibody titers (reciprocal endpoint titers ranging from 2,048,00 to 4,096,000). Mice immunized with PhtD were protected against death following intraperitoneal injection with either of three heterologous strains, SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research

Hospital; Memphis, TN), EF6796 (serotype 6A) or EF5668 (serotype 4; both strains provided by D. Briles, University of Alabama, Birmingham). In the experiment shown in Figure 8 (Panel A), all ten of the sham-immunized mice died within 10-days after challenge with virulent pneumococci (strain SJ2), while eighty percent of the PhtD-immunized mice survived the 15-day observation period. Immunization with PhtD also protected against a serotype 6A strain, EF6796 (Panel B) and a serotype 4 strain, EF5668 (Panel C). In the experiment shown in Figure 8 (Panel B), all ten of the sham-immunized mice died within 7-days after challenge with virulent pneumococci (strain EF6796), while ninety percent of the PhtD-immunized mice survived the 15-day observation period. In the experiment shown in Figure 8 (Panel C), all ten of the sham-immunized mice died within 6-days after challenge with virulent pneumoccoci (strain EF5668), while eight of nine mice immunized with PhtD survived the 15-day observation period.

15

Percent lo	lentity Between	Amino Acid Sed	auences	
	PhtA	PhtB	PhtD	PhtE
PhtA	**-	66.4	63.9	49.5
PhtB			87.2	49.5
PhtD				49.8
PhtE				
_		Nucleotide Segu	ences	
Percent Id	entity Between	Nocieotide Sequ	011000	
Percent Id	PhtA	PhtB	PhtD	PhtE
Percent Id	·	·		PhtE 47.9
	PhtA	PhtB	PhtD	
PhtA	PhtA	PhtB 58.3	PhtD 59.3	47.9

EDITORIAL NOTE

APPLICATION NUMBER - 2373 \ 1000

The following Sequence Listing pages \mid to \mid 8 are part of the description. The claims pages follow on pages 32 to 39 .

SEQUENCE LISTING

<110> Johnson, Leslie S. Koenig, Scott Adamou, John E.

<120> Streptococcus Pneumoniae and Immunogenic Fragments for Vaccines

<130> 469201-444

<140>

<141>

<150> 60/113,048 <151> 1998-12-21

<160> 11

<170> PatentIn Ver. 2.0

<210> 1

<211> 36

<212> DNA

<213> Artificial Sequence

<220:

<223> Description of Artificial Sequence:Forward primer used in amplification of the Sp36 gene sequence.

<400> 1

atcggatcct tcttacgagt tgggactgta tcaagc

36

<210> 2

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Forward primer used in amplification of the Sp36 gene sequence.

<400> 2

atcggatcca ctgtatcaag ctagaacggt taagg

35

<210> 3

<211> 40 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Reverse primer used in amplification of the Sp36 gene sequence.

<400> 3

agtcaagctt gtttattttt tccttactta cagatgaagg

40

<210> 4

<211> 83

<212> PRT

<213> Streptococcus pneumoniae

<400>

Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Leu Ser Val Cys Ser Tyr Glu Leu Gly Arg His Gln Ala Gly Gln Val \$20\$

Lys Lys Glu Ser Asn Arg Val Ser Tyr Ile Asp Gly Asp Gln Ala Gly

Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly 50 \$50\$

Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val 65 70 75 80

Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr 85 90 95

Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln 100 105 110

Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 115 120 125

Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140

Asp Asn Ile Arg Thr Lys Glu Glu Ile Lys Arg Gln Lys Gln Glu His 145 150 155 160

Ser His Asn His Gly Gly Gly Ser Asn Asp Gln Ala Val Val Ala Ala
165 170 175

2

-

Arg Ala Gln Gly Arg Tyr Thr Thr Asp Asp-Gly Tyr-Ile Phe Asn Ala 180 185 190

- Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly 195 200 205
- Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu 210 215 220
- Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser 225 230 235
- Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu 245 250 255
- Asn His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu 260 265 270
- Asn Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu 275 280 285
- Arg His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr 290 295
- Ser Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His 305 310 315 320
- Phe Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg 325 330 335
- Ile Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg 340 345 350
- Pro Glu Gln Pro Ser Pro Gln Ser Thr Pro Glu Pro Ser Pro Ser Pro 355 360 365
- Gln Pro Ala Pro Asn Pro Gln Pro Ala Pro Ser Asn Pro Ile Asp Glu 370 375 380
- Lys Leu Val Lys Glu Ala Val Arg Lys Val Gly Asp Gly Tyr Val Phe 385
- Glu Glu Asn Gly Val Ser Arg Tyr Ile Pro Ala Lys Asp Leu Ser Ala 405 $$ 410 $$ 415
- Glu Thr Ala Ala Gly Ile Asp Ser Lys Leu Ala Lys Gln Glu Ser Leu 420 425 430

Ser His Lys Leu Gly Ala Lys Lys Thr Asp Leu Pro Ser Ser Asp Arg

- Glu Phe Tyr Asn Lys Ala Tyr Asp Leu Leu Ala Arg Ile His Gln Asp 450 450 460
- Leu Leu Asp Asn Lys Gly Arg Gln Val Asp Phe Glu Ala Leu Asp Asn 465 470 475 480
- Leu Leu Glu Arg Leu Lys Asp Val Pro Ser Asp Lys Val Lys Leu Val
 485 490 495
- Asp Asp Ile Leu Ala Phe Leu Ala Pro Ile Arg His Pro Glu Arg Leu 500 505 510
- Ala Lys Leu Ala Gly Lys Tyr Thr Thr Glu Asp Gly Tyr Ile Phe Asp 530 535 540
- Pro Arg Asp Ile Thr Ser Asp Glu Gly Asp Ala Tyr Val Thr Pro His 545 550 555 560
- Met Thr His Ser His Trp Ile Lys Lys Asp Ser Leu Ser Glu Ala Glu 565 570 575
- Arg Ala Ala Ala Gln Ala Tyr Ala Lys Glu Lys Gly Leu Thr Pro Pro $580 \hspace{1.5cm} 585 \hspace{1.5cm} 590 \hspace{1.5cm}$
- Ser Thr Asp His Gln Asp Ser Gly Asn Thr Glu Ala Lys Gly Ala Glu 595 600 605
- Ala Ile Tyr Asn Arg Val Lys Ala Ala Lys Lys Val Pro Leu Asp Arg 610 615 620
- Met Pro Tyr Asn Leu Gln Tyr Thr Val Glu Val Lys Asn Gly Ser Leu 625 630 640
- Ile Ile Pro His Tyr Asp His Tyr His Asn Ile Lys Phe Glu Trp Phe 645 650 655
- Asp Glu Gly Leu Tyr Glu Ala Pro Lys Gly Tyr Thr Leu Glu Asp Leu 660 665 670
- Leu Ala Thr Val Lys Tyr Tyr Val Glu His Pro Asn Glu Arg Pro His 675 680 685

Ser Asp Asn Gly Phe Gly Asn Ala Ser Asp His Val Arg Lys Asn Lys 695 Val Asp Gln Asp Ser Lys Pro Asp Glu Asp Lys Glu His Asp Glu Val 705 710 715 Ser Glu Pro Thr His Pro Glu Ser Asp Glu Lys Glu Asn His Ala Gly 725 730 Leu Asn Pro Ser Ala Asp Asn Leu Tyr Lys Pro Ser Thr Asp Thr Glu 740 745 Glu Thr Glu Glu Glu Ala Glu Asp Thr Thr Asp Glu Ala Glu Ile Pro 755 760 Gln Val Glu Asn Ser Val Ile Asn Ala Lys Ile Ala Asp Ala Glu Ala 775 Leu Leu Glu Lys Val Thr Asp Pro Ser Ile Arg Gln Asn Ala Met Glu 790 795 Thr Leu Thr Gly Leu Lys Ser Ser Leu Leu Gly Thr Lys Asp Asn 805 810 Asn Thr Ile Ser Ala Glu Val Asp Ser Leu Leu Ala Leu Leu Lys Glu 820 825 Ser Gln Pro Ala Pro Ile

835

<210> 5
<211> 2531
<212> DNA
<213> Streptococcus pneumoniae

<400> 5

atgaaaatta ataaaaata tctagcaggt tcagtggaag tccttgcct aagtgttgt 60
tcctatgaac ttggtcgtca ccaagctggt caggttaaga aagagtctaa tcgagttct 120
tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180
aagagggagg ggatcaacgc cgaacaaatc gtcatcaaga ttacggatca aggttatgt 240
acctctcatg agaccatta tcattacta aatggcaagg tccttatga tgccatcatc 300
agtgaagagc tcctcatga agatccgaat tatcagtag aggattcaga cattgtcaat 360
gaaatcaaagg gtggttatgt tatcaaggta gatggaaaaat actatgtta ccttaaggat 420
gcagctcatt gcggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacac 480
agtcataatc acggggtgg ttctaacgat caagcagtag ttgcagccag agcccaagga 540
cgctatacaa cggatgatgg ttatatctt aatgcatct aatcattga gaacaccgag 540
gatgctata tcgtcctca cggcaccat taccattac ttcctaagaa tcctaagaa tgagttatca 660

```
gctagcgagt tagctgctgc agaagcctat tggaatggga agcagggatc tcgtccttct 720
tcaagttcta gttataatgc aaatccagct caaccaagat tgtcagagaa ccacaatctg 780
actgtcactc caacttatca tcaaaatcaa ggggaaaaca tttcaagcct tttacgtgaa 840
ttgtatgcta aaccettate agaacgccat gtggaatetg atggeettat tttegaccea 900
gegeaaatea caagtegaac egeeagaggt gtagetgtee eteatggtaa ecattaceac 960
tttatccctt atgaacaaat gtctgaattg gaaaaacgaa ttgctcgtat tattcccctt 1020
cgttatcgtt caaaccattg ggtaccagat tcaagaccag aacaaccaag tccacaatcg 1080
actocggaac ctagtecaag teegcaacet geaccaaate eteaaceage teeaageaat 1140
ccaattgatg agaaattggt caaagaagct gttcgaaaag taggcgatgg ttatgtcttt 1200
gaggagaatg gagtttctcg ttatatccca gccaaggatc tttcagcaga aacagcagca 1260
ggcattgata gcaaactggc caagcaggaa agtttatctc ataagctagg agctaagaaa 1320
actgacetee catetagtga tegagaattt tacaataagg ettatgactt actagcaaga 1380
attcaccaag atttacttga taataaaggt cgacaagttg attttgaggc tttggataac 1440
ctgttggaac gactcaagga tgtcccaagt gataaagtca agttagtgga tgatattctt 1500
geettettag etcegatteg teatecagaa egtttaggaa aaccaaatge geaaattace 1560
tacactgatg atgagattca agtagccaag ttggcaggca agtacacaac agaagacggt 1620
tatatetttg ateetegtga tataaeeagt gatgaggggg atgeetatgt aaeteeacat 1680
atgacccata gccactggat taaaaaagat agtttgtctg aagctgagag agcggcagcc 1740
caggettatg ctaaagagaa aggtttgace ceteettega cagaccatca ggattcagga 1800
aatactgagg caaaaggagc agaagctatc tacaaccgcg tgaaagcagc taagaaggtg 1860
ccacttgatc gtatgcctta caatcttcaa tatactgtag aagtcaaaaa cggtagttta 1920
atcatacctc attatgacca ttaccataac atcaaatttg agtggtttga cgaaggcctt 1980
tatgaggcac ctaaggggta tactettgag gatettttgg cgactgtcaa gtactatgtc 2040
gaacatccaa acgaacgtcc gcattcagat aatggttttg gtaacgctag cgaccatgtt 2100
cgtaaaaata aggtagacca agacagtaaa cctgatgaag ataaggaaca tgatgaagta 2160
agtgagccaa ctcaccctga atctgatgaa aaagagaatc acgctggttt aaatccttca 2220
gcagataatc tttataaacc aagcactgat acggaagaga cagaggaaga agctgaagat 2280
accacagatg aggetgaaat teeteaagta gagaattetg ttattaaege taagatagea 2340
gatgcggagg ccttgctaga aaaagtaaca gatcctagta ttagacaaaa tgctatggag 2400
acattgactg gtctaaaaag tagtcttctt ctcggaacga aagataataa cactatttca 2460
gcagaagtag atagtetett ggetttgtta aaagaaagte aaceggetee tatatagtaa 2520
aagcttaagc c
<210> 6
<211> 484
<212> PRT
<213> Streptococcus pneumoniae
<400> 6
Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val
                                     10
Ser Leu Ser Leu Cys Ala Tyr Ala Leu Asn Gln His Arg Ser Gln Glu
             20
Asn Lys Asp Asn Asn Arg Val Ser Tyr Val Asp Gly Ser Gln Ser Ser
         35
```

Gln Lys Ser Glu Asn Leu Thr Pro Asp Gln Val Ser Gln Lys Glu Gly
50 55 60

- Ile Gln Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val 65 70 75 80
- Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr 85 90 95
- Asp Ala Leu Phe Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln 100 105 110
- Leu Lys Asp Ala Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Ile Ile 115 120 125
- Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140
- Asp Asn Val Arg Thr Lys Asp Glu Ile Asn Arg Gln Lys Gln Glu His 145 150 155 160
- Val Lys Asp Asn Glu Lys Val Asn Ser Asn Val Ala Val Ala Arg Ser 165 170 175
- Gln Gly Arg Tyr Thr Thr Asn Asp Gly Tyr Val Phe Asn Pro Ala Asp 180 185 190
- Ile Ile Glu Asp Thr Gly Asn Ala Tyr Ile Val Pro His Gly Gly His
- Tyr His Tyr Ile Pro Lys Ser Asp Leu Ser Ala Ser Glu Leu Ala Ala 210 215 220
- Ala Lys Ala His Leu Ala Gly Lys Asn Met Gln Pro Ser Gln Leu Ser 225 230 240
- Tyr Ser Ser Thr Ala Ser Asp Asn Asn Thr Gln Ser Val Ala Lys Gly
- Ser Thr Ser Lys Pro Ala Asn Lys Ser Glu Asn Leu Gln Ser Leu Leu 260 265 270
- Lys Glu Leu Tyr Asp Ser Pro Ser Ala Gln Arg Tyr Ser Glu Ser Asp 275 280 285
- Gly Leu Val Phe Asp Pro Ala Lys Ile Ile Ser Arg Thr Pro Asn Gly
 290 295 300

325 330 335

Thr Gly Ser Thr Val Ser Thr Asn Ala Lys Pro Asn Glu Val Val Ser 340 345 350

Ser Leu Gly Ser Leu Ser Ser Asn Pro Ser Ser Leu Thr Thr Ser Lys 355 360 365

Glu Leu Ser Ser Ala Ser Asp Gly Tyr Ile Phe Asn Pro Lys Asp Ile 370 375 380

Val Glu Glu Thr Ala Thr Ala Tyr Ile Val Arg His Gly Asp His Phe 385 390 395 400

His Tyr Ile Pro Lys Ser Asn Gln Ile Gly Gln Pro Thr Leu Pro Asn
405
415

Asn Ser Leu Ala Thr Pro Ser Pro Ser Leu Pro Ile Asn Pro Gly Thr 420 425 430

Ser His Glu Lys His Glu Glu Asp Gly Tyr Gly Phe Asp Ala Asn Arg 435 440 445

Ile Ile Ala Glu Asp Glu Ser Gly Phe Val Met Ser His Gly Asp His 450 455 460

Asn His Tyr Phe Phe Lys Lys Asp Leu Thr Glu Glu Gln Ile Lys Val 465 470 480

Arg Lys Asn Ile

<210> 7

<211> 1455 <212> DNA

<213> Streptococcus pneumoniae

<400> 7

atgaaattta gtaaaaaata tatagcagct ggatcagctg ttatcgtatc cttgagtcta 60 tgtgcctatg cactaaacca gcatcgttcg caggaaaata aggacaataa tcgtgtctct 120 tatgtggatg gcagccagtc aagtcagaaa agtgaaaact tgacaccaga ccaggttagc 180 cagaaagaag gaattcaggc tgagcaaatt gtaatcaaaa ttacagatca gggctatgta 240 acgtcaccac gtgaccacta tcattactat aatgggaaag ttccttatga tgccctcttt 300

```
agtgaagaac tettgatgaa ggatecaaac tatcaactta aagacgetga tattgtcaat 360
gaagtcaagg gtggttatat catcaaggtc gatggaaaat attatgtcta cctgaaagat 420
gcagctcatg ctgataatgt tcgaactaaa gatgaaatca atcgtcaaaa acaagaacat 480
gtcaaagata atgagaaggt taactctaat gttgctgtag caaggtctca gggacgatat 540
acgacaaatg atggttatgt ctttaatcca gctgatatta tcgaagatac gggtaatgct 600
tatatcgttc ctcatggagg tcactatcac tacattccca aaagcgattt atctgctagt 660
gaattagcag cagctaaagc acatctggct ggaaaaaata tgcaaccgag tcagttaagc 720
tattetteaa eagetagtga caataacaeg caatetgtag caaaaggate aactageaag 780
ccagcaaata aatctgaaaa tctccagagt cttttgaagg aactctatga ttcacctagc 840
gcccaacgtt acagtgaatc agatggcctg gtctttgacc ctgctaagat tatcagtcgt 900
acaccaaatg gagttgcgat tccgcatggc gaccattacc actttattcc ttacagcaag 960
ctttctgcct tagaagaaaa gattgccaga atggtgccta tcagtggaac tggttctaca 1020
gtttctacaa atgcaaaacc taatgaagta gtgtctagtc taggcagtct ttcaagcaat 1080
cettettett taacgacaag taaggagete tetteageat etgatggtta tattttaat 1140
ccaaaagata tcgttgaaga aacggctaca gcttatattg taagacatgg tgatcatttc 1200
cattacattc caaaatcaaa tcaaattggg caaccgactc ttccaaacaa tagtctagca 1260
acaccttctc catctcttcc aatcaatcca ggaacttcac atgagaaaca tgaagaagat 1320
ggatacggat ttgatgctaa tcgtattatc gctgaagatg aatcaggttt tgtcatgagt 1380
cacggagacc acaatcatta tttcttcaag aaggacttga cagaagagca aattaaggtg 1440
cgcaaaaaca tttag
<210> 8
<211> 819
```

<212> PRT

<213> Streptococcus pneumoniae

<400> 8

Met Lys Ile Asn Lys Lys Tyr Leu Val Gly Ser Ala Ala Ala Leu Ile 10

Leu Ser Val Cys Ser Tyr Glu Leu Gly Leu Tyr Gln Ala Arg Thr Val

Lys Glu Asn Asn Arg Val Ser Tyr Ile Asp Gly Lys Gln Ala Thr Gln 40

Lys Thr Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly Ile 60

Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val Thr 70 75

Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr Asp 90

Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Lys Leu 100 105

Lys Asp Glu Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Val Ile Lys 120 125 Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala Asp 140 Asn Val Arg Thr Lys Glu Glu Ile Asn Arg Gln Lys Gln Glu His Ser 155 Gln His Arg Glu Gly Gly Thr Pro Arg Asn Asp Gly Ala Val Ala Leu 170 Ala Arg Ser Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His 200 205 Gly Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu 215 Leu Ala Ala Glu Ala Phe Leu Ser Gly Arg Gly Asn Leu Ser Asn 230 235 Ser Arg Thr Tyr Arg Arg Gln Asn Ser Asp Asn Thr Ser Arg Thr Asn 245 250 Trp Val Pro Ser Val Ser Asn Pro Gly Thr Thr Asn Thr Asn Thr Ser 260 265 Asn Asn Ser Asn Thr Asn Ser Gln Ala Ser Gln Ser Asn Asp Ile Asp 280 Ser Leu Leu Lys Gln Leu Tyr Lys Leu Pro Leu Ser Gln Arg His Val 295 Glu Ser Asp Gly Leu Val Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr 310 315 Ala Arg Gly Val Ala Val Pro His Gly Asp His Tyr His Phe Ile Pro 325 Tyr Ser Gln Met Ser Glu Leu Glu Glu Arg Ile Ala Arg Ile Ile Pro 340 345 Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro Glu Gln 360

Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Gly Pro Gln Pro Ala 375 Pro Asn Leu Lys Ile Asp Ser Asn Ser Ser Leu Val Ser Gln Leu Val 385 390 Arg Lys Val Gly Glu Gly Tyr Val Phe Glu Glu Lys Gly Ile Ser Arg 405 410 Tyr Val Phe Ala Lys Asp Leu Pro Ser Glu Thr Val Lys Asn Leu Glu 420 425 Ser Lys Leu Ser Lys Gln Glu Ser Val Ser His Thr Leu Thr Ala Lys 435 440 Lys Glu Asn Val Ala Pro Arg Asp Gln Glu Phe Tyr Asp Lys Ala Tyr 455 Asn Leu Leu Thr Glu Ala His Lys Ala Leu Phe Glu Asn Lys Gly Arg 470 475 Asn Ser Asp Phe Gln Ala Leu Asp Lys Leu Leu Glu Arg Leu Asn Asp 490 Glu Ser Thr Asn Lys Glu Lys Leu Val Asp Asp Leu Leu Ala Phe Leu 505 Ala Pro Ile Thr His Pro Glu Arg Leu Gly Lys Pro Asn Ser Gln Ile 520 Glu Tyr Thr Glu Asp Glu Val Arg Ile Ala Gln Leu Ala Asp Lys Tyr 535 Thr Thr Ser Asp Gly Tyr Ile Phe Asp Glu His Asp Ile Ile Ser Asp Glu Gly Asp Ala Tyr Val Thr Pro His Met Gly His Ser His Trp Ile 565 570 Gly Lys Asp Ser Leu Ser Asp Lys Glu Lys Val Ala Ala Gln Ala Tyr 585 Thr Lys Glu Lys Gly Ile Leu Pro Pro Ser Pro Asp Ala Asp Val Lys 600 Ala Asn Pro Thr Gly Asp Ser Ala Ala Ala Ile Tyr Asn Arg Val Lys 610 615 620

Gly Glu Lys Arg Ile Pro Leu Val Arg Leu Pro Tyr Met Val Glu His 625 630 635 640

Thr Val Glu Val Lys Asn Gly Asn Leu Ile Ile Pro His Lys Asp His 645 650 655

Tyr His Asn Ile Lys Phe Ala Trp Phe Asp Asp His Thr Tyr Lys Ala 660 665 670

Pro Asn Gly Tyr Thr Leu Glu Asp Leu Phe Ala Thr Ile Lys Tyr Tyr 675 680 685

Val Glu His Pro Asp Glu Arg Pro His Ser Asn Asp Gly Trp Gly Asn 690 695 700

Ala Ser Glu His Val Leu Gly Lys Lys Asp His Ser Glu Asp Pro Asn 705 710 715 720

Lys Asn Phe Lys Ala Asp Glu Glu Pro Val Glu Glu Thr Pro Ala Glu
725 730 735

Pro Glu Val Pro Gln Val Glu Thr Glu Lys Val Glu Ala Gln Leu Lys
740 745 750

Glu Ala Glu Val Leu Leu Ala Lys Val Thr Asp Ser Ser Leu Lys Ala 755 760 765

Asn Ala Thr Glu Thr Leu Ala Gly Leu Arg Asn Asn Leu Thr Leu Gln 770 785

Ile Met Asp Asn Asn Ser Ile Met Ala Glu Ala Glu Lys Leu Leu Ala 785 790 795 800

Leu Leu Lys Gly Ser Asn Pro Ser Ser Val Ser Lys Glu Lys Ile Asn 805 810 815

Lys Leu Asn

<210> 9

<211> 2451

<212> DNA

<213> Streptococcus pneumoniae

<400> 9

atgaaaatta ataagaaata ccttgttggt tctgcggcag ctttgatttt aagtgtttgt 60 tcttacgagt tgggactgta tcaagctaga acggttaagg aaaataatcg tgtttcctat 120

```
atagatggaa aacaagcgac gcaaaaaacg gagaatttga ctcctgatga ggttagcaag 180
cgtgaaggaa tcaatgctga gcaaatcgtc atcaagataa cagaccaagg ctatgtcact 240
tcacatggcg accactatca ttattacaat ggtaaggttc cttatgacgc tatcatcagt 300
gaagaattac tcatgaaaga tccaaactat aagctaaaag atgaggatat tgttaatgag 360
gtcaagggtg gatatgttat caaggtagat ggaaaatact atgtttacct taaggatgct 420
gcccacgcgg ataacgtccg tacaaaagag gaaatcaatc gacaaaaaca agagcatagt 480
caacatcgtg aaggtggaac tccaagaaac gatggtgctg ttgccttggc acgttcgcaa 540
ggacgctata ctacagatga tggttatatc tttaatgctt ctgatatcat agaggatact 600
ggtgatgctt atatcgttcc tcatggagat cattaccatt acattcctaa gaatgagtta 660
teagetageg agttggetge tgeagaagee tteetatetg gtegaggaaa tetgteaaat 720
tcaagaacct atcgccgaca aaatagcgat aacacttcaa gaacaaactg ggtaccttct 780
gtaagcaatc caggaactac aaatactaac acaagcaaca acagcaacac taacagtcaa 840
gcaagtcaaa gtaatgacat tgatagtete ttgaaacage tetacaaact geetttgagt 900
caacgacatg tagaatctga tggccttgtc tttgatccag cacaaatcac aagtcgaaca 960
gctagaggtg ttgcagtgcc acacggagat cattaccact tcatccctta ctctcaaatg 1020
totgaattgg aagaacgaat cgctcgtatt attocccttc gttatcgttc aaaccattgg 1080
gtaccagatt caaggccaga acaaccaagt ccacaaccga ctccggaacc tagtccaggc 1140
ccgcaacctg caccaaatct taaaatagac tcaaattctt ctttggttag tcagctggta 1200
cgaaaagttg gggaaggata tgtattcgaa gaaaagggca tctctcgtta tgtctttgcg 1260
aaagatttac catctgaaac tgttaaaaat cttgaaagca agttatcaaa acaagagagt 1320
gtttcacaca ctttaactgc taaaaaagaa aatgttgctc ctcgtgacca agaattttat 1380
gataaagcat ataatctgtt aactgaggct cataaagcct tgtttgnaaa taagggtcgt 1440
aattctgatt tccaagcctt agacaaatta ttagaacgct tgaatgatga atcgactaat 1500
aaagaaaaat tggtagatga tttattggca ttcctagcac caattaccca tccagagcga 1560
cttggcaaac caaattctca aattgagtat actgaagacg aagttcgtat tgctcaatta 1620
gctgataagt atacaacgtc agatggttac atttttgatg aacatgatat aatcagtgat 1680
gaaggagatg catatgtaac gcctcatatg ggccatagtc actggattgg aaaagatagc 1740
ctttctgata aggaaaaagt tgcagctcaa gcctatacta aagaaaaagg tatcctacct 1800
ccatctccag acgcagatgt taaagcaaat ccaactggag atagtgcagc agctatttac 1860
aatcgtgtga aaggggaaaa acgaattcca ctcgttcgac ttccatatat ggttgagcat 1920
acagttgagg ttaaaaacgg taatttgatt attcctcata aggatcatta ccataatatt 1980
aaatttgctt ggtttgatga tcacacatac aaagctccaa atggctatac cttggaagat 2040
ttgtttgcga cgattaagta ctacgtagaa caccctgacg aacgtccaca ttctaatgat 2100
ggatggggca atgccagtga gcatgtgtta ggcaagaaag accacagtga agatccaaat 2160
aagaacttca aagcggatga agagccagta gaggaaacac ctgctgagcc agaagtccct 2220
caagtagaga ctgaaaaagt agaagcccaa ctcaaagaag cagaagtttt gcttgcgaaa 2280
gtaacggatt ctagtctgaa agccaatgca acagaaactc tagctggttt acgaaataat 2340
ttgactcttc aaattatgga taacaatagt atcatggcag aagcagaaaa attacttgcg 2400
ttgttaaaag gaagtaatcc ttcatctgta agtaaggaaa aaataaacta a
<2105 10
<211> 819
<212> PRT
<213> Streptococcus pneumoniae
<400> 10
Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala
                                    10
```

Leu Ser Val Cys Ser Tyr Glu Leu Glý Arg Tỹr Gln Ala Gly Gln Asp 20 25 Lys Lys Glu Ser Asn Arg Val Ala Tyr Ile Asp Gly Asp Gln Ala Gly 40 Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly 55 Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 120 Lys Val Asn Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 135 Asp Asn Ile Arg Thr Lys Glu Glu Ile Lys Arg Gln Lys Gln Glu Arg 150 155 Ser His Asn His Asn Ser Arg Ala Asp Asn Ala Val Ala Ala Ala Arg 170 Ala Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala Ser 185 Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu Ala 215 Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser Ser 230 Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu Asn 245 250 His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu Asn

14

265

Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu Arg 280 His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr Ser 295 Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His Phe 310 Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg Ile 325 330 Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro 345 Glu Glu Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Ser Pro Gln 360 Pro Ala Pro Ser Asn Pro Ile Asp Gly Lys Leu Val Lys Glu Ala Val 375 Arg Lys Val Gly Asp Gly Tyr Val Phe Glu Glu Asn Gly Val Ser Arg 395 Tyr Ile Pro Ala Lys Asp Leu Ser Ala Glu Thr Ala Ala Gly Ile Asp 410 Ser Lys Leu Ala Lys Gln Glu Ser Leu Ser His Lys Leu Gly Thr Lys 425 Lys Thr Asp Leu Pro Ser Ser Asp Arg Glu Phe Tyr Asn Lys Ala Tyr 435 440 Asp Leu Leu Ala Arg Ile His Gln Asp Leu Leu Asp Asn Lys Gly Arg 455 Gln Val Asp Phe Glu Ala Leu Asp Asn Leu Leu Glu Arg Leu Lys Asp 470 Val Ser Ser Asp Lys Val Lys Leu Val Glu Asp Ile Leu Ala Phe Leu

Ala Pro Ile Arg His Pro Glu Arg Leu Gly Lys Pro Asn Ala Gln Ile 500 505 510

490

Thr Tyr Thr Asp Asp Glu Ile Gln Val Ala Lys Leu Ala Gly Lys Tyr 515 520 525

Thr Ala Glu Asp Gly Tyr Ile Phe Asp Pro Arg Asp Ile Thr Ser Asp 535 Glu Gly Asp Ala Tyr Val Thr Pro His Met Thr His Ser His Trp Ile 545 550 555 Lys Lys Asp Ser Leu Ser Glu Ala Glu Arg Ala Ala Gln Ala Tyr 565 570 Ala Glu Glu Lys Gly Leu Thr Pro Pro Ser Thr Asp His Gln Asp Ser 580 Gly Asn Thr Glu Ala Lys Gly Ala Glu Ala Ile Tyr Asn Arg Val Lys 600 Ala Ala Lys Lys Val Pro Leu Asp Arg Met Pro Tyr Asn Leu Gln Tyr 615 Thr Val Glu Val Lys Asn Gly Ser Leu Ile Ile Pro His Tyr Asp His 630 635 Tyr His Asn Ile Lys Phe Glu Trp Phe Asp Glu Gly Leu Tyr Glu Ala 645 650 Pro Lys Gly Tyr Thr Leu Glu Asp Leu Leu Ala Thr Val Lys Tyr Tyr 665 Val Glu His Pro Asn Glu Arg Pro His Ser Asp Asn Gly Phe Gly Asn Ala Ser Asp His Val Gln Arg Asn Lys Asn Gly Gln Ala Asp Thr Asn 695 Gln Thr Glu Lys Pro Ser Glu Glu Lys Pro Gln Thr Glu Lys Pro Glu 710 Glu Glu Thr Pro Arg Glu Glu Lys Pro Gln Ser Glu Lys Pro Glu Ser 725 Pro Lys Pro Thr Glu Glu Pro Glu Glu Ser Pro Glu Glu Ser Glu Glu 745 Pro Gln Val Glu Thr Glu Lys Val Glu Glu Lys Leu Arg Glu Ala Glu 760 Asp Leu Leu Gly Lys Ile Gln Asp Pro Ile Ile Lys Ser Asn Ala Lys 775 780

Glu Thr Leu Thr Gly Leu Lys Asn Asn Leu Leu Phe Gly Thr Gln Asp 785 790 795 800

Asn Asn Thr Ile Met Ala Glu Ala Glu Lys Leu Leu Ala Leu Leu Lys 805 810 815

Glu Ser Lys

<210> 11
<211> 2531

<211> 2531 <212> DNA <213> Streptococcus pneumoniae

<400> 11

atgaaaatta ataaaaaata totagcaggt toagtggcag toottgcoot aagtgtttgt 60 tcctatgagc ttggacgtta ccaagctggt caggataaga aagagtctaa tcgagttgct 120 tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180 aagagggagg ggatcaacgc cgaacaaatt gttatcaaga ttacggatca aggttatgtg 240 accteteatg gagaceatta teattaetat aatggeaagg tteettatga tgccateate 300 agtgaagage teeteatgaa agateegaat tateagttga aggatteaga cattgteaat 360 gaaatcaagg gtggttatgt cattaaggta aacggtaaat actatgttta ccttaaggat 420 gcrgctcatg cggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacgc 480 agtcataatc ataactcaag agcagataat gctgttgctg cagccagagc ccaaggacgt 540 tatacaacgg atgatgggta tatcttcaat gcatctgata tcattgagga cacgggtgat 600 gettatateg tteetcaegg egaceattac cattacatte ctaagaatga gttateaget 660 agegagttag etgetgeaga ageetattgg aatgggaage agggateteg teettettea 720 agttctagtt ataatgcaaa tccagctcaa ccaagattgt cagagaacca caatctgact 780 gtcactccaa cttatcatca aaatcaaggg gaaaacattt caagcctttt acgtgaattg 840 tatgctaaac ccttatcaga acgccatgtg gaatctgatg gccttatttt cgacccageg 900 caaatcacaa gtcgaaccgc cagaggtgta gctgtccctc atggtaacca ttaccacttt 960 atcccttatg aacaaatgtc tgaattggaa aaacgaattg ctcgtattat tccccttcgt 1020 tategtteaa accattgggt accagattea agaccagaag aaccaagtee acaaccgact 1080 ccagaaccta gtccaagtcc gcaaccagct ccaagcaatc caattgatgg gaaattggtc 1140 aaagaagctg ttcgaaaagt aggcgatggt tatgtctttg aggagaatgg agtttctcgt 1200 tatatcccag ccaaggatct ttcagcagaa acagcagcag gcattgatag caaactggcc 1260 aagcaggaaa gtttatctca taagctagga actaagaaaa ctgacctccc atctagtgat 1320 cgagaatttt acaataaggc ttatgactta ctagcaagaa ttcaccaaga tttacttgat 1380 aataaaggtc gacaagttga ttttgaggct ttggataacc tgttggaacg actcaaggat 1440 gtctcaagtg ataaagtcaa gttagtggaa gatattcttg ccttcttagc tccgattcgt 1500 catccagaac gtttaggaaa accaaatgcg caaattacct acactgatga tgagattcaa 1560 gtagccaagt tggcaggcaa gtacacagca gaagacggtt atatetttga teetegtgat 1620 ataaccagtg atgaggggga tgcctatgta actccacata tgacccatag ccactggatt 1680 aaaaaagata gtttgtctga agctgagaga gcggcagccc aggcttatgc traagagaaa 1740 ggtttgaccc ctccttcgac agaccatcag gattcaggaa atactgaggc aaaaggagca 1800 gaagetatet acaaccgmgt gaaagcaget aagaaggtge caettgateg tatgeettae 1860

aatetteaat atactgtaga agteaaaac ggtagtttaa teatacetea ttatgaceat 1920 taccataaca teaaatttga gtggtttgac gaaggeettt atgaggeace taaggggtat 1980 actettgagg atettttgge gactgteaag tactatgteg aacateeaaa egaaegteeg 2040 catteagata atggttttgg taaegetage gaccatgtte aaagaaacaa aaatggteaa 2100 getgatacea ateaaacgga aaaaceaage gaggagaaac etcagacaga aaaacetgag 2160 gaagaacaca cecgagaaga gaaacegeaa agegagaace etcagacaga aaaacetgag 2160 gaagaacacaa aagaaceaca agaggaatea gaggagaacete aggtegagae tgaaaaggtt 2280 gaagaaaaaa tegagaacete tgaagagge tgaagatta etcaggatea aacaatacea aagagacete cacaggatt aaaaataatat tactatttgg caccaggae 2400 aaggtagaage ttaagggega atttggeace caggacacaa atactatta ggeagaaget 2520 gaaaaaceta ttaagggega atttggeace caggacacaa atactatta ggeagaaget 2520 gaaaaaceta t

The claims defining the invention are as follows:

- 1. A composition comprising a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.
 - 2. The composition of claim 1 wherein said percent identity is at least 97%.
- 3. The composition of claim 1 wherein said polypeptide has the sequence of amino acids 20-838 of SEQ ID NO:4.
- 4. A composition comprising an active fragment of a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 wherein said active fragment comprises at least two coiled coil regions and is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to said polypeptide when said composition is administered to a mammal.
- The composition of claim 4 wherein said active fragment further comprises at least 5 histidine triad regions.
- 6. A composition comprising a polypeptide comprising the sequence of amino acid residues 139-791 of SEQ ID NO:4.
- 7. A vaccine comprising the composition of claim 1, 2 or 3 wherein said polypeptide is present in an amount effective to protect against pneumococcal infection when administered to a mammal.
- 8. A method for preventing infection caused by S. pneumoniae comprising administering the vaccine of claim 7.
- 9. An isolated polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 wherein said polypeptide binds to an antibody that binds to *S. pneumoniae*.
- 10. The isolated polypeptide of claim 9 wherein said percent identity is at least 97%.
- 11. The isolated polypeptide of claim 9 wherein said isolated polypeptide has the sequence of amino acids 20-838 of SEQ ID NO:4.
- 12. A method for producing a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4 comprising expressing said polypeptide from a cell that comprises a polynucleotide containing a nucleotide sequence that encodes said polypeptide.
 - 13. The method of claim 12 wherein said percent identity is at least 97%.

R:\LIBFF]99631spec.doc:GCC			

- 14. The method of claim 12 wherein said polypeptide comprises the amino acid sequence of SEQ ID NO:4.
- 15. The method of claim 12 wherein said polypeptide binds to an antibody specific for *Streptococcus pneumoniae*.
- 16. The method of claim 12 wherein the nucleotide sequence contained in said polynucleotide has at least 90% identity to the sequence of SEQ ID NO:5.
- 17. The method of claim 12 wherein said nucleotide sequence contained in said polynucleotide has at least 95% identity to the sequence of SEQ ID NO:5.
- 18. The method of claim 12 wherein said nucleotide sequence contained in said polynucleotide has at least 97% identity to the sequence of SEQ ID NO:5.

10

- 19. The method of claim 12 wherein said nucleotide sequence contained in said polynucleotide has the sequence of SEQ ID NO:5.
- 20. A method for producing an active fragment of the polypeptide that comprises amino acids 20-838 of SEQ ID NO:4 comprising expressing said active fragment from a cell that contains a polynucleotide containing a nucleotide sequence that encodes said active fragment.
- 21. The method of claim 20 wherein said active fragment is a contiguous segment of amino acids 20-838 of SEQ ID NO:4 and at least 42 amino acid residues in length and not more than 356 residues in length.
- 22. The method of claim 20 wherein said active fragment comprises at least one coiled coil region.
- 23. The method of claim 20 wherein said active fragment comprises at least two coiled coil regions.
- 24. The method of claim 20 wherein said active fragment comprises at least one histidine triad region.
- 25. The method of claim 20 wherein said active fragment comprises at least two histidine triad regions.
- 26. The method of claim 20 wherein said active fragment comprises at least three histidine triad regions.
- 27. An isolated polynucleotide comprising a nucleotide sequence having at least 90% identity to the sequence of SEQ ID NO:5.
- 28. The isolated polynucleotide of claim 27 wherein said nucleotide sequence has at least 95% identity to the sequence of SEQ ID NO:5.
- 29. The isolated polynucleotide of claim 27 wherein said nucleotide sequence has at least 97% identity to the sequence of SEQ ID NO:5.

[R\LIBFF]99631spec.doc:GCC

- 30. The isolated polynucleotide_of claim 27 wherein said nucleotide sequence comprises the sequence of SEQ ID NO:5.
 - 31. A vector comprising the isolated polynucleotide of claim 27, 28, 29, or 30.
 - 32. A vector comprising the polynucleotide used in 12, 16, 17, 18, 19 or 20.
 - 33. A recombinant cell genetically engineered with the vector of claim 31.
 - 34. A recombinant cell genetically engineered with the vector of claim 32.
- 35. A method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.
 - 36. The method of claim 35 wherein said percent identity is at least 97%.
- 37. The method of claim 35 wherein said polypeptide comprises the amino acid sequence of SEQ ID NO:4.
- 38. A method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an active fragment of amino acids 20-838 of SEQ ID NO:4.
- 39. The method of claim 38 wherein said active fragment is a contiguous segment of amino acids 20-838 of SEQ ID NO:4 and at least 42 amino acid residues in length and not more than 356 residues in length.
- 40. The method of claim 39 wherein said active fragment comprises at least one coiled coil region.
- 41. The method of claim 40 wherein said active fragment comprises at least two coiled coil regions.
- 42. The method of claim 39 wherein said active fragment comprises at least one histidine triad region.
- 43. The method of claim 39 wherein said active fragment comprises at least two histidine triad regions.
- 44. The method of claim 39 wherein said active fragment comprises at least three histidine triad regions.
- 45. The method of claim 35 wherein said polypeptide is administered in a pharmaceutically acceptable carrier.
- 46. The method of claim 39 wherein said active fragment is administered in a pharmaceutically acceptable carrier.
- 47. A method for eliciting an immune response to *Streptococcus pneumoniae* in a mouse comprising administering to said mouse a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4.

[I:\DayLib\LIBFF]85145spec.doc:gcc

48. A composition comprising a polypeptide having at least 90% identity to ____ amino acids 21-480 of SEQ ID NO:6,

wherein said polypeptide is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

- 49. The composition of claim 48 wherein said percent identity is at least 95%.
- 50. The composition of claim 48 wherein said percent identity is at least 97%.
- 51. The composition of claim 48 wherein said polypeptide has the sequence of amino acids 21-480 of SEQ ID NO:6.
- 52. A composition comprising an active fragment of the polypeptide comprising amino acids 21-480 of SEQ ID NO:6,

wherein said active fragment comprises at least about 10% and no more than about 85% of said polypeptide, wherein said active fragment comprises at least one histidine triad and at least one coiled coil region and wherein said active fragment is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

- 53. The composition of claim 52 wherein said active fragment further comprises at least 2 histidine triad regions.
- 54. The composition of claim 52 wherein said active fragment further comprises at least 2 coiled coil regions.
- 55. A vaccine comprising a polypeptide containing the sequence of amino acids 21-480 of SEQ ID NO:6 in a pharmaceutically acceptable carrier in an amount effective to protect against pneumococcal infection when administered to a mammal.
- 56. A process for preventing infection caused by S. pneumoniae comprising administering the vaccine of claim 55.
- 57. A process for preventing infection caused by *S. pneumoniae* comprising administering the composition of claim 52.
- 58. An isolated polypeptide comprising at least 90% identity to amino acids 21-480 of SEQ ID NO:6, wherein said polypeptide binds to an antibody that binds to S. pneumoniae.
- 59. The isolated polypeptide of claim 58 wherein said percent identity is at least 95%.
- 60. The isolated polypeptide of claim 58 wherein said percent identity is at least 97%.

[R:\LIBFF]99631spec.doc:GCC

35

- 61. The isolated polypeptide of claim .58 wherein said isolated polypeptide has the amino acid sequence of amino acids 21-480 of SEQ ID NO:6.
- 62. A composition comprising at least one antibody that binds to a polypeptide comprising an amino acid sequence at least 90% identical to a member selected from the group consisting of:
- (i) a sequence having at least 90% identity to amino acids 20-838 of SEQ ID NO:4 and $\,$
 - (ii) a sequence comprising amino acids 21-480 of SEQ ID NO:6, wherein said antibody is in a pharmaceutically acceptable carrier.
- 63. The composition of claim 62 wherein the antibody is the antibody against the polypeptide in (i).
- 64. The composition of claim 62 wherein the antibody is the antibody against the fragment in (ii).
- 65. A composition comprising at least one antibody that binds to an active fragment of a polypeptide comprising a member selected from the group consisting of:
 - (i) amino acids 20-838 of SEQ ID NO:4 and
 - (ii) amino acids 21-480 of SEQ ID NO:6,

10

30

wherein said active fragment comprises at least about 10% and no more than about 85% of said polypeptide, wherein said active fragment comprises at least one histidine triad and at least one coiled coil region and wherein said active fragment is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

- 66. A process for treating or preventing infection caused by *S. pneumoniae* comprising administering to a mammal afflicted with said infection, or at risk of said infection, an effective amount of the composition of claim 62.
- 67. A process for treating or preventing infection caused by S. pneumoniae comprising administering to a mammal afflicted with said infection, or at risk of said infection, an effective amount of the composition of claim 65.
- 68. A composition comprising a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae*, substantially as hereinbefore described with reference to any one of the examples.
- 69. A composition comprising an active fragment of a polypeptide at least 95% identical to amino acids 20-838 of SEQ ID NO:4 wherein said active fragment comprises

[R:\LiBFF]99631spec.doc:GCC		

at least two coiled coil regions and is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to said polypeptide, substantially as hereinbefore described with reference to any one of the examples.

- 70. A vaccine comprising the composition of claim 68 or 69 wherein said polypeptide is present in an amount effective to protect against pneumococcal infection when administered to a mammal.
- 71. A method for preventing infection caused by S. pneumoniae comprising administering the vaccine of claim 70.
- 72. A method for producing a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4, substantially as hereinbefore described with reference to any one of the examples.
- 73. An isolated polynucleotide comprising a nucleotide sequence having at least 90% identity to the sequence of SEQ ID NO:5, substantially as hereinbefore described with reference to any one of the examples.
- 74. A method for eliciting an immune response to *Streptococcus pneumoniae* in a mammal comprising administering to said mammal a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4, substantially as hereinbefore described with reference to any one of the examples.
- 75. A vaccine comprising a polypeptide containing the sequence of amino acids 21-480 of SEQ ID NO:6 in a pharmaceutically acceptable carrier, substantially as hereinbefore described with reference to any one of the examples.
- 76. A method for preventing infection caused by S. pneumoniae, comprising administering the vaccine of claim 75.
- 77. Use of a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4 for the manufacture of a medicament for eliciting an immune response to *Streptococcus pneumoniae* in a mammal.
 - 78. Use of a polypeptide of claim 77 wherein said percent identity is at least 97%.
- 79. Use of a polypeptide of claim 77 wherein said polypeptide comprises the amino acid sequence of residues 20-838 of SEQ ID NO:4.
- 80. Use of a polypeptide comprising an active fragment of amino acids 20-838 of SEQ ID NO:4 for the manufacture of a medicament for eliciting an immune response to *Streptococcus pneumoniae* in a mammal.
- 81. Use of the polypeptide of claim 80 wherein said active fragment is a contiguous segment of amino acids 20-838 of SEQ ID NO:4 and at least 42 amino acid residues in length and not more than 356 residues in length.

[R:\LIBFF]99631spec.doc:GCC

15

- 82. Use of the polypeptide of claim 81 wherein said active fragment comprises at least one coiled coil region.
- 83. Use of the polypeptide of claim 82 wherein said active fragment comprises at least two coiled coil regions.
- 84. Use of the polypeptide of claim 81 wherein said active fragment comprises at least one histidine triad region.
- 85. Use of the polypeptide of claim 81 wherein said active fragment comprises at least two histidine triad regions.
- 86. Use of the polypeptide of claim 81 wherein said active fragment comprises at least three histidine triad regions.
 - 87. Use of the polypeptide of claim 77 wherein said polypeptide is administered in a pharmaceutically acceptable carrier.
 - 88. Use of the polypeptide of claim 81 wherein said active fragment is administered in a pharmaceutically acceptable carrier.
 - 89. Use of a polypeptide comprising an amino acid sequence that has at least 95% identity to amino acids 20-838 of SEQ ID NO:4 for the manufacture of a medicament for eliciting an immune response to *Streptococcus pneumoniae* in a mouse.
 - 90. Use of an effective amount of a composition comprising a polypeptide having at least 90% identity to amino acids 21-480 of SEQ ID NO:6,

wherein said polypeptide is in a pharmaceutically acceptable carrier in an amount effective to elicit production of an antibody that binds to *S. pneumoniae* when said composition is administered to a mammal.

- 91. Use of a polypeptide containing the sequence of amino acids 21-480 of SEQ ID NO:6 in a pharmaceutically acceptable carrier in an amount effective to protect against pneumococcal infection when administered to a mammal, for the manufacture of a medicament for preventing infection caused by *S. pneumoniae*.
- 92. Use of the composition of claim 52 for the manufacture of a medicament for preventing infection caused by *S. pneumoniae*.
- 93. Use of the composition of claim 62 for the manufacture of a medicament for treating or preventing infection caused by *S. pneumoniae* of a mammal afflicted with said infection or at risk of said infection.
 - 94. Use of the composition of claim 65 for the manufacture of a medicament for treating or preventing infection caused by *S. pneumoniae* of a mammal afflicted with said infection or at risk of said infection.

[R:\LiBFF]99631spec.doc:GCC

10

15

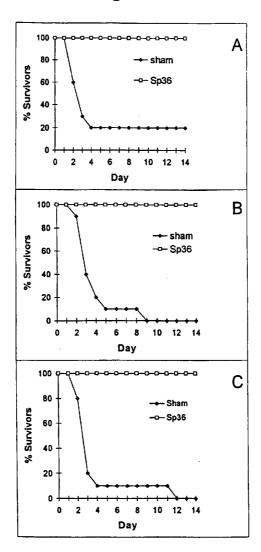
- 95. An isolated polynucleotide comprising a nucleotide sequence that encodes the polypeptide of SEQ ID NO:4.
- 96. An isolated polynucleotide comprising a nucleotide sequence that encodes the polypeptide of SEQ ID NO:6.
- 97. The isolated polynucleotide of claim 96 wherein said nucleotide sequence is the sequence of SEQ ID NO:7.

Dated 19 July, 2004 Med Immune, Inc.

Patent Attorneys for the Applicant/Nominated Person SPRUSON & FERGUSON

[R:\LIBFF]99631spec.doc:GCC

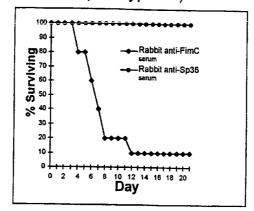
Figure 1



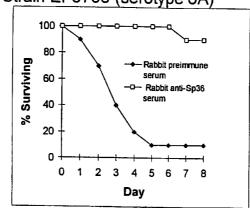
1/17

Figure 2

A. Strain SJ2 (serotype 6B)

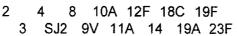


B. Strain EF6796 (serotype 6A)



2/17

FIG.3A



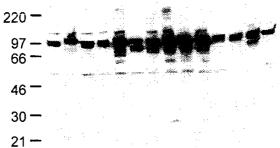
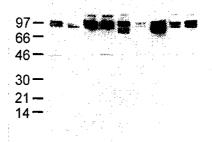


FIG.3B

1 7F 15N 20 33F 5 9N 17F 22F



3/17

F1G.4



4/17
SUBSTITUTE SHEET (RULE 26)

FIG.5A

FIG. 5B

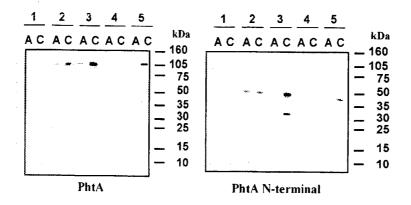
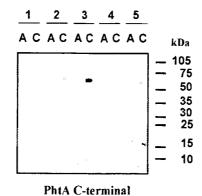


FIG.5C



5/17 SUBSTITUTE SHEET (RULE 26)

Figure 6(a)

CSYELGR	HOAGOXKKE	SNRVSYIDG	DOAGORAENL	TPDEVSKREG:	INAEQ Majority
	10	20	30	40	50
1 CSYELGR	HQAGQVKKE	SNRVSYIDG	DOAGORAENL	TPDEVSKREG:	NA E OI Phen PRO
1 icsyzlori	YID A G OIDIK K Z	SNRVAYIDG	DOAGOKAENL	TPDEVSKREG!	NA FO Phen are
1 CSYELGL	<u> </u>	NNRVSYIDG	KO ATO KTENL	TPDEVSKREG	NA E Q PhEA.PRO
1 CAYACHO	Ык ≳- [<u>О</u>] Σ и <u>(К</u>] ⊃	NM KAZ AIA (DC)	z fol z sio xi sis n r	T P D Q V S Q K 2 G	UQIA E O Phiz. PRO
IVIKITO	QGYVTSHGD	HYHYYNGKV	PYDAIISEEL	LHKDPNYQLKI	S D I V Majority
	60	70	80	30	100
SI IVIKITO	OCYVTSHGD	HYHYYNGKV	PYDAIISEPL	LHEDPNYQLE	
51 IVIKITO	QGYVTSHGD	HYHYYNGKV	PYDAIISEEL	LHKDPNYQLKI	S D I V Phts.pro
50 IVIKITO	QGYVTSHGD	HYHYYNGKV	PYDAIISEEL	LHKDPNYKLKE	ED T V Phea PPO
50 IVIKITO	OGYVTSHGD	HYHYYNGKV	PADVITABEST	LHKDPNYOLK	AD I V Phtz. PRO
NEVEGGY	VIKVDOKYY	VY	NVRTEBEIN	коковнянин:	GG Majority
	110	120	130	140	150
101 N ELLIK G G Y	VIKVDGKYY	VYLEDAAHAI	1	я Q х Q E н S н N н с	
101 N E I K G G Y 1	VIRVING KYY	VYLKDAAHAI	NIERTKERIK	ROKOEIRISHNHIS	S Phis.pro
100 NEVEGGY	VIKADGKAA	VYLKDAAHAI	NVRTEEEN	ROKOKKSOKRE	G G T P Phta PRO
100 NEVKOGY	I I K V D G K Y Y	VYLKDAAHAI	NVRTEDEIN	ROKOK HV KD NE	Pher.PRO
RNDXAVA	ARAQGRYT	TDDGYIPNAS	DIIBDTGDA	YIVPHODKYHY	IPKN Majority
	160	170	180	190	200
see chi plata vivi					
		TODUTIFHAS	COILEOTGOL	Y	
148 RADNA VA	A R A Q G R Y T	TDDGYIPNAS	DITEDTODA	Y I V P H G D H Y H 1	IPKN Phts.pro
146 RADNAVA 150 R N D G A V A I	AARAQORYT LARSQORYT	TDDGYIPNA! T <u>D</u> DGY <u>I</u> PN <u>A</u> !	DIIEDTODA DIIEDTODA	Y I V P H G D H Y H Y Y I V P H G D H Y H Y	IPKN Phts.pro
146 RADNAVA 150 R N D G A V A I	AARAQORYT LARSQORYT	TDDGYIPNA! T <u>D</u> DGY <u>I</u> PN <u>A</u> !	DIIEDTODA DIIEDTODA	Y I V P H G D H Y H Y	IPKN PhtB.pro
148 RAD NA VA 150 R N D G A V A 146 K V N S N V A	A A R A Q G R Y T LA RSQ G R Y T VA R SQ G R Y T	TDDGYIPHA: TDDGYIPHA: TNDGYVPHPI	DIIEDTODA DIIEDTODA DIIEDTONA	AIA B H C C H A H J A I A B H C C H A H J A I A B H C C H A H J	IPKN Pht8.pro
148 RAD NA VA 150 R N D G A V A 146 K V N S N V A	A A R A Q G R Y T LA RSQ G R Y T VA R SQ G R Y T	TDDGYIPHA: TDDGYIPHA: TNDGYVPHPI	DIIEDTODA DIIEDTODA DIIEDTONA	Y I V P H G D H Y H Y Y I V P H G D H Y H Y	IPKN Pht8.pro
146 RADHAVA 150 RNDGAVA 146 KVNSNVA ELSASEL	AARAQORYT LARSQORYT VARSQORYT AAAEAYLNG 210	TDDGYIPHA: TDDGYVPHP; TNDGYVPHP; 220	DIIEDTGDA DIIEDTGDA DIIEDTGNA	YIVPHGDHYHY YIVPHGDHYHY YIVPHGGHYHY SSSYNANPAOI 240	Y I P K N PhtB.pro Y I P K N PhtA.pro Y I P K S PhtB.pro P K L S E Majority 250 P K L S E PhtB.pro
146 RAID NA V A 150 R N D G A V A 146 K V N S N V A 159 E L S A S E L 3	AARAQ GRYT LARSQ GRYT VARS GRYT AAARAY LNG 210 AAARAY WN G	T D D G Y I P N A 9 T D D G Y I F N A 9 T N D G Y V P N P 1 K	DIIEDTGDA DIIEDTGDA DIIEDTGMA OGSRPSS QGSRPSS	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y 2 S S Y N A N P A O I S S S Y N A N P A Q I	FIPKN Phts.pro FIPKN Phts.PRO FIPKS Phts.PRO PRLSE Hajority 250 RLSE Phts.pro
146 RADDNA V A 150 R N D G A V A 146 K V N S N V A 146 K V N S N V A 146 K V N S N V A 150 K V A	A A R A Q G R Y T LA R S Q G R Y T VA R S Q G R Y T 210 A A B A Y W N G A A B A Y W N G A A B A Y W N G A A B A Y W N G	TDDGYIFNA: TDDGYVFNP: TNDGYVFNP: 220 K GGNLSHSRT	DIIRDTGDA DIIRDTGDA DIIRDTGNA	YIVPHGDHYHY YIVPHGGHYHY YIVPHGGHYHY SSSYNANPAOI 240 SSSYNANPAQI SSSYNANPAQI SSSYNANPAQI RTNMVPSVSNI	IPKN Pht8.pro IPKN Pht8.pro IPKN Pht8.pro PRLSE Hajority 250 RLSE PhtD.pro PRLSE Pht0.pro GTTN Pht8.pro
146 RAID NA V A 150 R N D G A V A 146 K V N S N V A 159 E L S A S E L 3	A A R A Q G R Y T LA R S Q G R Y T VA R S Q G R Y T 210 A A B A Y W N G A A B A Y W N G A A B A Y W N G A A B A Y W N G	TDDGYIFNA: TDDGYVFNP: TNDGYVFNP: 220 K GGNLSHSRT	DIIRDTGDA DIIRDTGDA DIIRDTGNA	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y 2 S S Y N A N P A O I S S S Y N A N P A Q I	IPKN Pht8.pro IPKN Pht8.pro IPKN Pht8.pro PRLSE Hajority 250 RLSE PhtD.pro PRLSE Pht0.pro GTTN Pht8.pro
146 RAJDINIA V A J 150 R N D G A V A I 146 K V N S N V A I E L S A S E L J 199 E L S A S E L J 200 E L S A S E L J 196 D L S A S E L J	A A A A A A A A A A A A A A A A A A A	TDDGYIFNA, TDDGYIFNA, TNDGYVVNPI 220 K	DIIRDTODA DIIRDTODA DIIRDTONA OGSRPSS OGSRPSS RROMSDNTS NMOPS	YIVPHGDHYHY YIVPHGGHYHY YIVPHGGHYHY SSSYNANPAOI 240 SSSYNANPAQI SSSYNANPAQI SSSYNANPAQI RTNMVPSVSNI	I P K N Phtb.pro I P K N Phtb.pro I P K N Phtt.pro I P K S Phtt.pro R L S E Phtb.pro R L S E Phtb.pro R L S E Phtb.pro G T T N Phta.pro
146 RAJDINIA V A J 150 R N D G A V A I 146 K V N S N V A I E L S A S E L J 199 E L S A S E L J 200 E L S A S E L J 196 D L S A S E L J	A A A A A A A A A A A A A A A A A A A	TDDGYIFNA, TDDGYIFNA, TNDGYVVNPI 220 K	DIIRDTODA DIIRDTODA DIIRDTONA OGSRPSS OGSRPSS RROMSDNTS NMOPS	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y S S S Y N A N P A O I S S S Y N A N P A Q F S S S Y N A N P A Q F R T N W V P S V S N E Q L S Y S S T A S D -	I P K N Phtb.pro I P K N Phtb.pro I P K N Phtt.pro I P K S Phtt.pro R L S E Phtb.pro R L S E Phtb.pro R L S E Phtb.pro G T T N Phta.pro
146 RAJDINIA V A. 150 RN D G A V A. 146 K V N S N V A. E L S A S E L. 199 E L S A S E L. 198 E L S A S E L. 200 E L S A S E L. 196 D L S A S E L. THN L T V T S	A A A A Q Q R Y T LA R S Q Q R Y Y VA R S Q Q R Y Y 210 210 A A A E A Y W N G A A E A Y W N G A A E A Y W N G A A E A Y W N G A A E A Y W N G A A E A Y W N G A W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G A W N G W N G W N G W N G A W N G W N G W N G W N G A W N G W N G W N G W N G A W N G W N G W N G W N G W N G A W N G W N G W N G W N G W N G W N G A W N G W N	TDDGYIFNA, TDDGYIFNA, TDDGYIFNA, TNDGYIFNA,	DITEDTODA DITEDTONA OGSRPSS 220 OGSRPSS RROMSDNTS NHOPS	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y S S S Y N A N P A O I S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P S S S Y N A N P A Q P V S S D G L V P D P A 290	I P K N PhtB.pro I P K N PhtB.pro I P K N PhtE.PRO I P K L S E Majority 250 R L S E PhtB.pro R L S E PhtB.pro G T T N PhtL.PRO - N N PhtL.PRO O I T S Majority
146 RAJD NIA V A. 150 RN D G A V A. 150 RN D G A V A. 146 KV N S N V A. E L S A S E L J. 198 E L S A S E L J. 200 E L S A S E L J. 196 D L S A S E L J. TH N L T V T E. 238 NH N L T V T E. 237 NH N L T V T E.	AAA AQ G R Y T LA R S Q G R Y T LA R S Q G R Y T T AAA E A Y L N G 210 AAA E A Y L N G AAA E A Y L N G AAA E A Y L L S G AA E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L L S G A E A Y L N G G E A Y L S G A E A Y L N G G E A Y L N G G E A Y L N G G E A Y L N G G E A Y L N G G E A Y L N G E A Y L N G E A Y L N G E A Y L N	TD D G Y I F N A: TD D G Y I F N A: TN D G Y V F N F J K	DIIEDTODA DIIEDTODA DIIEDTONA	YIVPHGDHYHY YIVPHGDHYHY YIVPHGGHYHY SSSYNANPAOS	I P K N Phtb.pro (I P K N Phtb.pro (I P K N Phtb.pro F R L S E Phtb.pro R L S E Phtb.pro G T T N Phtb.pro G T T N Phtb.pro G O I T S Majority 100 Q I T S Phtb.pro
146 RAJDINIA V A 1 150 RN D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 147 E L S A S E L 1 199 E L S A S E L 1 198 E L S A S E L 1 198 E L S A S E L 1 198 D L S A S E L 1 198 D L S A S E L 2 200 E L S A S E L 2 217 H N L T V T T T T T T T T T T T T T T T T T	A A A A Q Q R Y T L A R S Q Q R Y T L A R S Q Q R Y T A A A E A Y L N G 210 A A A E A Y L N G A A A E A Y L N G A A A E A Y L N G A A A E A Y L L S G A A A E A Y L L S G T Y H Q - N Q G T Y H Q - N Q G T Y H Q - N Q G T Y N S Q A S Q S	TDDGYIFNA; TDDGYIFNA; TDDGYIFNA; TNDGYIFNA;	DITEDTODA DITEDTONA OSRPSS 220 OSRPSS RROHSDNTS NHOPS YARPLSERH YARPLSERH YARPLSERH	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y S S Y N A N P A O I 240 S S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S S Y N A N	IPKN Pheb.pro (IPKN) Pheb.PRO FIPKN Pheb.PRO RLSE Majority 250 RLSE Pheb.PRO ORLSE Pheb.PRO OLTS Pheb.PRO OLTS Majority 300 QITS Pheb.PRO QITS Pheb.PRO
146 RAJDINIA V A 1 150 RN D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 147 E L S A S E L 1 199 E L S A S E L 1 198 E L S A S E L 1 198 E L S A S E L 1 198 D L S A S E L 1 198 D L S A S E L 2 200 E L S A S E L 2 217 H N L T V T T T T T T T T T T T T T T T T T	A A A A Q Q R Y T L A R S Q Q R Y T L A R S Q Q R Y T A A A E A Y L N G 210 A A A E A Y L N G A A A E A Y L N G A A A E A Y L N G A A A E A Y L L S G A A A E A Y L L S G T Y H Q - N Q G T Y H Q - N Q G T Y H Q - N Q G T Y N S Q A S Q S	TDDGYIFNA; TDDGYIFNA; TDDGYIFNA; TNDGYIFNA;	DITEDTODA DITEDTONA OSRPSS 220 OSRPSS RROHSDNTS NHOPS YARPLSERH YARPLSERH YARPLSERH	YIVPHGDHYHY YIVPHGDHYHY YIVPHGHHYHY SSSYNANPAOS 240 SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS THWVPSVSNIP QLSYSSTASD-	IPKN Pheb.pro (IPKN) Pheb.PRO FIPKN Pheb.PRO RLSE Majority 250 RLSE Pheb.PRO ORLSE Pheb.PRO OLTS Pheb.PRO OLTS Majority 300 QITS Pheb.PRO QITS Pheb.PRO
146 RAJD NIA V A 1 150 R N D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N T 1 146 K	AARAQORYT LARSQGRYT LARSQGRYT AAAEAYLHG 210 AAAEAYHNG	TD D G Y I F N A: TD D G Y I F N A: TN D G Y I F N P I 220 K	DITEDTODA DTTEDTODA DTTEDTONA QGSRPSS RRQWSDNTSQGSRPSS RRQWSDNTSNMQPS YARPLSERH YARPLSERH YARPLSERH YARPLSERH YARPLSERH YELFLSQRH YDSPSAQRY	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y S S Y M A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S Y N A N P A O I S S S Y N A N P A O I	I P K N Phtb.pro (I P K N Phtb.pro (I P K N Phtb.pro F I P K N Phtb.pro P R L S E Phtb.pro R L S E Phtb.pro G T T N Phtb.pro G T T N Phtb.pro G T S Phtb.pro G I T S Phtb.pro D I T S Phtb.pro
146 RAJD NIA V A 1 150 R N D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N T 1 146 K	AARAQORYT LARSQGRYT LARSQGRYT AAAEAYLHG 210 AAAEAYHNG	TD D G Y I F N A: TD D G Y I F N A: TN D G Y I F N P I 220 K	DIIEDTGDA DIIEDTGNA DIIEDTGNA 210QGSRPSS RRQNEDNTSQGSRPSS RRQNEDNTSNMQPS 210 210 210 24AFPLSERH YAKPLSERH YAKPLSERH YAKPLSERH YAKPLSERH YAKPLSERH YAKPLSERH	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S Y N A N P A O S S S S Y N A N P A O S S S S Y N A N P A O S S S	I P K N Phtb.pro (I P K N Phtb.pro (I P K N Phtb.pro F R L S E Phtb.pro R L S E Phtb.pro G T N Phtb.pro G T N Phtb.pro G T N Phtb.pro G T S Phtb.pro
146 RAID NIA V A 1 150 R N D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N T 1 146 K V T 1 146 K	AAAAQ ORYTLARSQ GRYTUARSQ GRYTOAAAAEAYLNG 210 AAAEAYLNG	TD D G Y I F N A: TD D G Y I F N A: TN D G Y I F N P J K	DITEDTODA DITEDTONA DITEDTONA 210QGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTSQGSRPSS RRQMSDNTS 210 YAKPLSERH YKLPLSQRH YKLPLSQRH YKLPLSQRH YBLSQRH YRLPLSQRH YBLSQRH YRLPLSQRH YBLSQRH YRLPLSQRH YBLSQRH YRLPLSQRH YBLSQRH YRLPLSQRH	YIVPHGDHYHY YIVPHGDHYHY YIVPHGDHYHY YIVPHGDHYHY SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS SSSYNANPAOS STAWVFSVSNIP OLEYPTA	I P K N Phtb.pro (I P K N Phtb.pro (I P K N Phtb.pro F R L S E Phtb.pro R L S E Phtb.pro G T T N Phtb.pro G T T N Phtb.pro G T T N Phtb.pro G T T S Phtb.pro Q I T S Phtb.pro R I I I S Phtb.pro
146 RAID NIA V A. 150 RN D G A V A. 150 RN D G A V A. 146 KV N S N V A. E L S A S E L J. 199 E L S A S E L J. 198 E L S A S E L J. 200 E L S A S E L J. 196 D L S A S E L J. 217 NH N L T V T S. 218 NH N L T V T S. 217 NH N L T V T S. 218 T N T S N N S & E Z. 210 T Q S V A K Q S. R T A R Q V A V. 287 R T A R Q V A V.	AAA A Q Q R Y T LARS Q G R Y T VARS Q G R Y T AAA E A Y L N G AAA E A Y W N G AAA E A Y W N G AAA E A Y W N G AAA E A Y W N G AAA E A Y L S G AAA E A Y W N G AAA E A Y L S G AAA E A Y W N	TDDGYIFNA. TDDGYIFNA. TDDGYVFMPJ K	DITEDTODA DITEDTONA OSRPSS 230OSRPSS RROMSDNTSNOSRPSS RROMSDNTSNOSPSS RROMSDN	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y S S Y N A N P A Q I S S Y N A N	I P K N Pheb.pro (I P K S E Pheb.pro (R L S E Pheb.pro (R I S P E O Pheb.pro (R P E O Pheb.pro
146 RAID NIA V A 1 150 R N D G A V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 146 K V N S N V A 1 150 E L S A S E L J 196 E L S A S E L J 196 D L	AARAQORYT LARSQGRYT LARSQGRYT AAAEAYLMG 210 AAAEAYMMG AAAEAYMMG AAAEAYMMG AAAEAYMG 250 7 Y H Q - N Q Q 27 Y H Q - N Q Q	TDDGYIFNA: TDDGYIFNA: TDDGYVFNPI 220 K	DIIEDTGDA DIIEDTGNA DIIEDTGNA 210QGSRPSS RRQNEDNTSNMQPS YAKPLSERH	Y I V P H G D H Y H Y Y I V P H G D H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y Y I V P H G G H Y H Y S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S S Y N A N P A Q I S Y R S N H M V P D S R Y R S N H M V P D S R Y R S N H M V P D S R Y R S N H M V P D S R Y R S N H M V P D S	I P K N Phtb.pro (I P K N Phtb.PRO (I P K N Phtb.PRO (I P K N Phtb.PRO PR L S E Phtb.PRO R L S E Phtb.PRO G T N Phtb.PRO G T N Phtb.PRO G T S Phtb.PRO G I T S Phtb.PRO F P E Q Phtb.PRO R P E Q Phtb.PRO R P E Q Phtb.PRO R P E Q Phtb.PRO

6/17

Figure 6(b)

PSPOPTPEPSPSPOPA	PN A P S N P	DXKLVKEAV	RKVGDGYVFE	ENGV HAJORITY
360	370	380	390	400
337 PSPQSTPEPSPSPQPA	PNPQPAPSNP	DEKLVKEAV	RKVGDGYVFE	E-N G V PhtD. PRO
116 PSPOPTPEPSPSPSPGP	APSNP	DCKLVKEAV	RKVGDGYVFE	E N G V PhtB.pro
150 PSPOPTPEPSPGPOPA	PNULKIUS	4 2 2 F A 2 G F A		- SGT Phts.PRO
				_
SRYVPAKDLSAETAAG	LDSKLARQESI	SHRLGARKT	DLPSSDREFY	N K A Y Hajority
410	420	430	440	450
JET SRYIPAKDL SAETAAG	IDSKLARQESI	SHKLGAKKT	DLPSSDREFY	N K A Y PheD. PRO
JBO SRYIPAKOL SARTAAG J96 <u>SRY</u> VFAKOLPSETVKN	LIZIS K LISIKO E SI	SHELGIER	DLPSSDREFY	H K A Y PhtB.pro
116 GSTVSTNAKPHEVVSS	LGS	LSSNPS	SLTTS	Phil.PRO
DLLARIHODLLDHKOR				PLAP Hajority
460	470	480	490	500
437 DLLARIHQDLLDNKGR 430 DLLARIHODLLDNKGR	Q V D F E A L D N L I	ERLED VIPIS D	KAKTABOIFY	FLAP PhtD.PRO
446 NL LTEAME ALPENKOR	N SD PQA L DEL L	BRLHDBSTH	KER L V D DLL A	F L A P Phta.pro
348		RELSS-		PhtE.PRO
IRHPERLGRPHAGITY	TDDEIOVAKLA	GKYTASDay	************	E G D A Weterton
510	520	530	540	a o o a mountly
			1	330
		G T V min win a v		
487 I R H P E R L G K P H A Q I T Y 480 I R H P E R L G K P H A Q I T Y	T D D E I Q V A K L A	GEYTARDGY	I F D P R D I T S D	E G D A Phts.pro
480 I R H P E R L G K P H A Q I T Y 496 I T H P E R L G K P H S Q I E Y	T D D E I Q V A K L A	GEYTANDGY DEYTTSDGY	TPDPRDITSD C2IICHBBCII	E G D A Phts.pro
480 IRHPERLOKPHAQITY	T D D E I Q V A K L A	GEYTANDGY DEYTTSDGY	I F D P R D I T S D	E G D A Phts.pro
480 IRHPERLOKPHAQITY	TODEIOVAKLA TEDEVRIAQLA	GRYTARDGY DRYTTSDGY ASDGY	IFOFROITSO IFORHOILISO IPNPRIJISE	E G D A Phts.pro E G D A Phts.pro T A T A Phtz.pro
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S Q I E Y 153	TODEIOVAKLA TEDEVRIAQLA	GRYTARDGY DRYTTSDGY ASDGY	IFOFROITSO IFORHOILISO IPNPRIJISE	E G D A Phts.pro E G D A Phts.pro T A T A Phtz.pro
480 THPERLORPHAQITY 496 THPERLORPHSQ TEV 153) YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS	T D D E I Q V A K L A TED EV R I A Q L A L S B A B R A A A Q A 570 L S B A B R A A A Q A	GRYTANDGY DRYTTSDGYASDGY YAKERGLTP 580 YAKERGLTP	IFDPRDITSDIFNDRDIVES IFNDRDIVEE PSTDHQDSGN PSTDHQDSGN	E G D A PhtB.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Majority 500 T E A K PhtD.PRO
480 TRHPERLGKPNAQITY 496 TTHPERLGKPNSQ TEV 153 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS	T D D E I Q V A K L A T E D E V R T L A Q L A 570 L S E A E R A A A Q A	GRYTANDGY DRYTTSDGYASDGY YAKEKGLTP S80 YAKEKGLTP YAKEKGLTP YAKEKGLTP	I F D P R D I T S D I F D R HD I I I S D I F N P R D I V R R P S T D H Q D S G H P S T D H Q D S G N	E G D A PhtB.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Majority 500 T E A K PhtD.PRO
480 T R H P E R L G K P N A Q I T Y 496 TTH P E R L G K P N S Q T E Y 153 Y V T P H N T H S H W I K K D S 560 517 Y V T P H N T H S H W I K K D S 530 Y V T P H H G H B H W I K K D S 546 Y V T P H H G H B H W I G K D S 546 Y V T P H H G H B H W I G K D S	T D D E I Q V A K L A T E D E V R T I A Q L A S T O E A E R A A A Q A S T O L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A L S D K E J R V A A Q A	GRYTANDGY DRYTTSDGYASDGY YAKERGLTP SEC YAKERGLTP YAKERGLTP YAKERGLTP YITKERGILTP	I F D P R D I T S D I F D R HD I I I S O I P N P R D I V X X P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N	E G D A Pht8.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Majority 600 T E A K PhtD.PRO T E A K PhtD.PRO T E A K PhtB.Pro P T G D PhtA.PRO
480 TRHPERLGKPNAQITY 496 TTHPERLGKPNSQ TEV 153 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS	T D D E I Q V A K L A T E D E V R T I A Q L A S T O E A E R A A A Q A S T O L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A L S D K E J R V A A Q A L S D K E J R V A A Q A	GRYTANDGY DRYTTSDGYASDGY YAKERGLTP SEC YAKERGLTP YAKERGLTP YAKERGLTP YITKERGILTP	I F D P R D I T S D I F D R HD I I I S O I P N P R D I V X X P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N	E G D A Pht8.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Majority 600 T E A K PhtD.PRO T E A K PhtD.PRO T E A K PhtB.Pro P T G D PhtA.PRO
480 T R H P E R L G K P N A Q I T Y 496 TTH P E R L G K P N S Q T E Y 153 Y V T P H N T H S H W I K K D S 560 517 Y V T P H N T H S H W I K K D S 530 Y V T P H H G H B H W I K K D S 546 Y V T P H H G H B H W I G K D S 546 Y V T P H H G H B H W I G K D S	T D D E I Q V A K L A TED EV R I A Q L A L S E A E R A A A Q A S 70 L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A S N Q I Q Q P T	GRYTARDOY DRYTTS DOY ASDGY YAKEKGLTP SECULTS YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP	I F D P R D I T S D I F D B N D I I S O I F N P R D I V S G P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N	E G D A Pht8.pro E G D A PhtA.PRO T A T A PhtA.PRO T E A K Majority 600 T E A K PhtB.pro F S A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.Pro
480 T R H P E R L G K P N A Q I T Y 496 T T H P E R L G K P N S Q I E Y 1537 Y V T P H H T H S H H I K K D S 537 Y V T P H H T H S H H I K K D S 530 Y V T P H H T H S H H I K K D S 546 Y V T P H H G H S H M I G K D S 572 Y I V R H G D M P H Y I P K -	T D D E I Q V A K L A TED EV R I A Q L A L S E A E R A A A Q A S 70 L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A S N Q I Q Q P T	GRYTARDOY DRYTTS DOY ASDGY YAKEKGLTP SECULTS YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP	I F D P R D I T S D I F D B N D I I S O I F N P R D I V S G P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N P S T D N Q D S G N	E G D A Pht8.pro E G D A PhtA.PRO T A T A PhtA.PRO T E A K Majority 600 T E A K PhtB.pro F S A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.Pro
480 TRHPERLOKPNAQITY 496 TTHPERLOKPNSQIEV 1533	T D D E I Q V A K L A T E D E V R I A Q L A 570 L S B A B R A A A Q A 570 L S B A B R A A A Q A L S D K B K V A A Q A S N Q I Q Q P T L D R K P Y N L Q Y T 620 L D R K P Y N L Q Y T	GRYTANDGY ORYTTS DGY ASDGY YAKERGLTP SEC YAREKGLTP YINEROLLTP YINEROLLTP YINEROLLTP VEVENOSLI 630 VEVENOSLI	I F D P R D I T S D I F D R MD I I S D I F M P R D I V R R P S T D H Q D S G M S 90 P S T D H Q D S G M P S T D H Q D S G M P S T D H Q D S G M P S F D A D V K A M S P S L P I M P G T I P H Y D M Y H M I 640 I P H Y D M Y H M I	E G D A Pht8.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Hajority 600 T E A K PhtB.PRO T E A K PhtB.PRO F T G D PhtA.PRO S H E K PhtB.PRO K F E W Majority 650 K F E W PhtB.PRO
496 ITH PERLOKPNAQITY 496 ITH PERLOKPNSQIEV 1533 YVTPHHTHSHMIKKDS 560 537 YVTPHHTHSHMIKKDS 530 YVTPHHTHSHMIKKDS 546 YVTPHHTHSHMIKKDS 546 YVTPHHGHSHMIGNDS 572 YIVRHGDHFHYIPK GARAIYNRVKAAKKVP 580 GARAIYNRVKAAKKVP	T D D E T Q V A K L A T E D E V R T A Q L A L S E A E R A A A Q A 570 L S E A E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S D X E R A A A Q A L S E A E R A A	GRYTARDOY ORYTTS DOY YAKEROLTP SSO YAKEROLTP YAKEROLTP YAKEROLTP YAKEROLTP YAKEROLTP YAKEROLTP YAKEROLTP YAKEROLTP YEVENOSLT GOO VEVENOSLT	I P D P R D I T S D I P D E N D I I S O I P N P R D I V S E P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N E S P S L P I N P G T I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I	E G D A PheB.pro E G D A PheA.PRO T A T A PheA.PRO T E A K Hajority 600 T E A K PheB.pro P T G D PheA.PRO S H E K PheB.pro F F E W Hajority 650 K F E W PheB.pro
480 I R H P B R L G K P N A Q I T Y 496 ITH P E R L G K P N S Q I E Y 1533 Y V T P H H T H S H M I K K D S 550 537 Y V T P H H T H S H M I K K D S 546 Y V T P H H T H S H M I K K D S 546 Y V T P H H G H S H M I G K D S 572 Y I V R H G D M P H Y I P K - G A K A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P	TDDEIQVALAGILA LSEAERAAAQA LSEAERAAAQAA LSEAERAAAQAAAQAA LSEAERAAAQAAAQAA LSEAERAAAQAAAQAA LSEAERAAAQAAAQAAAQAAAAQAAAAQAAAAQAAAAQAAA	GRYTARD GY ORYTARD GY YAKEKGLTP SEC YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YAKEKGLTP YEVENGSLT GOO VEVENGSLT VEVENGSLT VEVENGSLT VEVENGSLT VEVENGSLT	I F D P R D I T S D I F D B H D I I S O I F N P R D I V S G P S T D H Q D S G N 5 5 0 P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H R D H Y H N I I P H R D H Y R N I	E G D A PhtB.pro E G D A PhtP.PRO T A T A PhtP.PRO T E A K Majority 600 T E A K PhtD.PRO T E A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.pro K F E W PhtB.pro K F E W PhtD.PRO K F E W PhtD.PRO K F E W PhtD.PRO K F E W PhtA.PRO
496 I TH P ERL G K P N A Q I T Y 496 I TH P ERL G K P N S Q I E Y 153 YVT P H H T H S H M I K K D S 560 537 Y V T P H H T H S H M I K K D S 546 Y V T P H H T H S H M I K K D S 546 Y V T P H H G H S H M I G K D S 546 Y V T P H H G H S H M I G K D S 547 Y I V R H G D H F H Y I P R GAEAIYN R V K A A K K V P 590 GAEAIYN R V K A A K K V P 596 S A E A I Y N R V K A A K K V P 596 S A E A I Y N R V K A A K K V P 596 S A E A I Y N R V K G E K R I P 417 H E E D G Y G	T D D E I Q V A K L A T E D E V R I A G L A 570 L S B A B R A A A Q A L S B A B R A A A	GRYTANDOY ORYTTS DOY YAKERGLTP SSO YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YEVKNOSLI SSO VEVKNOSLI VEVKNOSLI	I P D P R D I T S D I P D B H D I T S D I P D B H D I V S B P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N E S P I N P G I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I S H G D H N H	E G D A PhtB.pro E G D A PhtA.PRO T A T A T A PhtA.PRO T E A K PhtD.PRO T E A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.PRO K F E W PhtB.pro
480 I R H P B R L G K P N A Q I T Y 496 ITH P E R L G K P N S Q I E Y 1533 Y V T P H H T H S H M I K K D S 550 537 Y V T P H H T H S H M I K K D S 546 Y V T P H H T H S H M I K K D S 546 Y V T P H H G H S H M I G K D S 572 Y I V R H G D M P H Y I P K - G A K A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P 580 G S S A A I Y N R V K A A K K V P	T D D E I Q V A K L A T E D E V R I A G L A 570 L S B A B R A A A Q A L S B A B R A A A	GRYTANDOY ORYTTS DOY YAKERGLTP SSO YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YEVKNOSLI SSO VEVKNOSLI VEVKNOSLI	I P D P R D I T S D I P D B H D I T S D I P D B H D I V S B P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N E S P I N P G I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I S H G D H N H	E G D A PhtB.pro E G D A PhtA.PRO T A T A T A PhtA.PRO T E A K PhtD.PRO T E A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.PRO K F E W PhtB.pro
496 I TH P ERL G K P N A Q I T Y 496 I TH P ERL G K P N S Q I E Y 153 YVT P H H T H S H M I K K D S 560 537 Y V T P H H T H S H M I K K D S 546 Y V T P H H T H S H M I K K D S 546 Y V T P H H G H S H M I G K D S 546 Y V T P H H G H S H M I G K D S 547 Y I V R H G D H F H Y I P R GAEAIYN R V K A A K K V P 590 GAEAIYN R V K A A K K V P 596 S A E A I Y N R V K A A K K V P 596 S A E A I Y N R V K A A K K V P 596 S A E A I Y N R V K G E K R I P 417 H E E D G Y G	T D D E I Q V A K L A T E D E V R I A G L A 570 L S B A B R A A A Q A L S B A B R A A A	GRYTANDOY ORYTTS DOY YAKERGLTP SSO YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YEVKNOSLI SSO VEVKNOSLI VEVKNOSLI	I P D P R D I T S D I P D B H D I T S D I P D B H D I V S B P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N P S T D H Q D S G N E S P I N P G I P H Y D H Y H N I I P H Y D H Y H N I I P H Y D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I P H K D H Y H N I I S H G D H N H	E G D A PhtB.pro E G D A PhtA.PRO T A T A T A PhtA.PRO T E A K PhtD.PRO T E A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.PRO K F E W PhtB.pro
480	T D D E I Q V A K L A TED E V R I A Q L A L S E A E R A A A Q A S 70 L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A L S E A E R A A A Q A L S D K E K V A A Q A S N Q I Q Q P T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L D R M P Y N L Q Y T L L A T V R Y Y V E R 670 L L A T V R Y Y V E R	GRYTARD GY GRYTARD GY ASDGY YAKEKGLTP SSO YAKEKGLTP YANDEKOLTP YANDEKOLTP YANDEKOLTP YANDEKOLTP YANDEKOLTP YANDEKOLTP YEVKNOSLI VEVKNOSLI VEVKNOSLI ZAEDES GPV PNERPHSDN	I F D P R D I T S D I F D R H D I I S O I F N P R D I V S R P S T D H Q D S G N P S T	E G D A Pht8.pro E G D A Pht2.pro T A T A Pht2.pro T E A K Hajority 600 T E A K Pht8.pro P T G D PhtA.pro P T G D PhtA.pro K F E W Majority 650 K F E W Majority 650 K F E W PhtB.pro K F E W PhtB.pro K F E W PhtA.pro PhtB.pro K F A W Majority 700 K N K X Majority 700 K N K X PhtB.pro
486	T D D E I Q V A K L A T E D E V R I A Q L A 570 L S B A B R A A A Q A L S B A B R A A	GRYTANDOY ORYTOSOOY YAKERGLTP SSO YAKERGLTP SSO YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YEVRNOSLI VEVRNOSLI VEVRNOSLI VEVRNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI PNERPHSDN	I P D P R D I T S D I P D B H D I I S O I P D B H D I I S O P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R	E G D A Phts.pro E G D A Phts.pro D A Phts.pro T E A K Phts.pro T E A K Phts.pro T E A K Phts.pro P T G D Phts.pro P T G D Phts.pro F E W Phts.pro 650 K F E W Phts.pro
480	T D D E I Q V A K L A T E D E V R I A Q L A 570 L S B A B R A A A Q A L S B A B R A A	GRYTANDOY ORYTOSOOY YAKERGLTP SSO YAKERGLTP SSO YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YAKERGLTP YEVRNOSLI VEVRNOSLI VEVRNOSLI VEVRNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI VEVKNOSLI PNERPHSDN	I P D P R D I T S D I P D B H D I I S O I P D B H D I I S O P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N P S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T D H O D S G N E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R E S T G N A S D H V R	E G D A Phts.pro E G D A Phts.pro D A Phts.pro T E A K Phts.pro T E A K Phts.pro T E A K Phts.pro P T G D Phts.pro P T G D Phts.pro F E W Phts.pro 650 K F E W Phts.pro

7/17

Figure 6(c)

x	XXXEEX	P E	<u></u>	Hajority
7	710	720 730	740	750
	RPSEEKPQTE	THPESDEKENHAG EKPEEET	LNPSADNLYKPSTPREEKPQSEKP	ESPKPT PhtB.pro
E E X E E T P X E	XEXPOVETER	KVEAKLXEAEXLL	XKVTDPSIKXNAX	ETLTGL Majority
7	760	770 780	790	800
	SEEPQVETER	VINAKIADA EALL KVEEKLREA EDLL KVEAQLKEA EVLL - PKKDLTEEQI	OKTOD PIT KSNAK	
		BKLLALLKESXPX 820 830	X X K K	Hajority
772 KNNLLFGTQ	DHNTIKABAZ	DSLLALLX ESQPA EKLLALLX ESK EKLLALLX GSNPS		Phed. Pro Pheb. pro Pheb. Pro Pheb. Pro

Decoration 'Decoration #2': Box residues that match the Consensus exactly

8/17

Figure 7(a)

	TECTATGAGETTGG	A. JTTATCAAG	T G G T C A G G T T		T C T A A Majority
	10	20	30	40	50
			 	<u></u>	
61	TCTTACGAGTTGGG				
1	TCCTATGAACTTGG				
64	CCTATGCACTAAA				
••					
	TCGTGTTTCTTATA	TAGATGGTGAT	CAGGCTGGTCX		A A A C T Majority
	60	70	AD.	90	100
					
108	TCGTGTTTCCTATA				
51	TCGAGTTGCTTATA TCGAGTTTCTTATA				
51 111	TCGTGTCTCTTATA				
111	.co.d.c.c				pilez
	TGACACCAGATGAG	GTTAGTAAGAG		ACGCTGAGC	A A A T T Majority
	110	120	110	140	150
	110	120	130	140	
	TGACTCCTGATGAG				
101	TGACACCAGATGAA				
101	TGACACCAGATGAA				
151	TGACACCAGACCAG	GTTAGCCAGAA	. U X X U U X X T T U	. AUGCTUAGE	A A A I I PRIE.SEQ
	GTCATCAAGATTAC	GGATCAAGGTTJ	. TGTGACCTCT		CATTA Majority
			ı	,	200
	160	170	180	190	
		AGACCAAGGCT			
151	GTTATCAAGATTAC	GGATCAAGGTTJ	.TGTGACCTCT	CATGGAGAC	CATTA phtB.seq
151 151	GTTATCAAGATTAC GTCATCAAGATTAC	G G A T C A A G G T T J G G A T C A A G G T T J	.TGTGACCTCT	. C	CATTA phtB.seq CATTA phtD.SEQ
151 151	GTTATCAAGATTAC	G G A T C A A G G T T J G G A T C A A G G T T J	.TGTGACCTCT	. C	CATTA phtB.seq CATTA phtD.SEQ
151 151	GTTATCAAGATTAC GTCATCAAGATTAC	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J	. T G T G A C C T C T . T G T G A C C T C T . T G T A A C G T C A	C	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ
151 151	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A	CATGGAGAC CATGGAGAC ACACGGTGAC	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Hajority
151 151 211	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A T G 210	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J G C A A G G T T C C T T 220	A T G T G A C C T C T A T G T G A C C T C T A T O T A A C G T C A E A T G A T G C C A T	CATGGAGAC CATGGAGAC ACACGGTGAC CATCAGTGA 240	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Hajority
151 151 211 258	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 210 TCATTATTACAATG	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J G C A A G G T T C C T T 220 G T A A G G T T C C T T	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A E A T G A T G C C A T 230 F A T G A C G C T A T	C	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Hajority 250 AGAAT phtA.SEQ
151 151 211 258 201	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A C A A T G T C A T T A T T A C T A T A A T G T C A T T A C T A T A A T G	G G A T C A A G G T T J G G A T C A G G G T T J A G A T C A G G G T T 230 G T A A G G T T C C T T G C A A G G T T C C T T	TATGACCTCA TGTGACCTCA TGTAACGTCA TATGATGCCAT 230 TATGACGCTAT TATGACGCTAT	CCATGGAGAC CCATGGAGAC ACACGGTGAC CCATCAGTGA 240 CCATCAGTGA	CATTA pht8.seq CATTA pht8.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAT phtA.SEQ AGAAC pht8.seq
151 151 211 258 201 201	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G T C A T T A T T A C A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J C C A A G G T T C C T T 220 G T A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A 230 FA T G A C G C T A T FA T G A T G C C A T	CCATGGAGAC CCATGGAGAC CCATCAGTGA 240 CCATCAGTGA CCATCAGTGA CCATCAGTGA	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq
151 151 211 258 201	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A C A A T G T C A T T A T T A C T A T A A T G T C A T T A C T A T A A T G	G G A T C A A G G T T J G G A T C A A G G T T J A G A T C A G G G C T J C C A A G G T T C C T T 220 G T A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A 230 FA T G A C G C T A T FA T G A T G C C A T	CCATGGAGAC CCATGGAGAC CCATCAGTGA 240 CCATCAGTGA CCATCAGTGA CCATCAGTGA	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq
151 151 211 258 201 201	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G T C A T T A T T A C A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G	GOATCAAGOTTA GOATCAAGOTTA GCAAGOTTCCTT GCAAGOTTCCTT GTAAGOTTCCTT GCAAGGTTCCTT GCAAGGTTCCTT	ATGTGACCTCT ATGTGACCTCT ATGTAACGTCA PATGATGCCAT PATGACGCTAT SATGATGCCAT FATGATGCCAT FATGATGCCCT	CCATGGAGAC CCATGAGTGA CCATCAGTGA ICATCAGTGA ICATCAGTGA ICATCAGTGA ICATCAGTGA ICATCAGTGA ICATCAGTGA	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.seq
151 151 211 258 201 201	GTTATCAAGATTAC GTCATCAAGATTAC TCATTACTATAATG TCATTATTACAATG TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG	G G A T C A A G G T T T A G G A T C A A G G T T T A A G G T T C C T T T C C T T A G G T T C C T T T C C T A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G G A A A G T T C C T T C C T T C C C A A A G T T C C T T C C C C C C C C C C C C C	TOTOLOGICA TOTOLOGICA TATGATOCCA TATGAT	C A T G G A G A C A C A C A G A C A C A C A C A G A C A C A C A C A G A C A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A C A T A T T G	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.seq
151 151 211 258 201 201 261	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 1210 TCATTATTACAATG TCATTACTATAATG TCATTACTATACAATG TCATTACTATACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG	G G A T C A A G G T T T A G G A T C A A G G T T C A G G G C T T C C T T C C T A G G G C T C C T T C C T A G G G C T C C C T C C C A A G G T T C C T T C C C A A G G T T C C T T C C C A A A G T T C C T T C C C C A A G G T C C C T C C C C C C C C C C C C C	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G C C A 2310 FA T G A T G C C A T FA T G A T G C C A T FA T G A T G C C C A FA T G A T G C C C C C 2 T T G A A G G A T T 280	CATGGAGAC ACACGGTGAC CATCAGTGA	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAC phtB.seq AGAGC phtB.seq AGAGC phtB.seq AGAAC phtB.SEQ TCAAT Majority 100
151 151 211 258 201 201 261	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T C A T G A A A G A T	G G A T C A A G G T T J G G A T C A A G G T T T J A G A T C A G G G C T J 220 G T A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T C C G A A T T A T C A C 270 C C C A A C T A T A A C	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A 230 FA T G A T G C C A T FA T G A T G C C A T FA T G A T G C C C A FA T G A T G C C C A FA T G A T G C C C C FA T G A T G C C C C FA T G A A G G A T T 280 G C T A A A A G A T G	C A T G G A G A C C A T G G A G A C C A T C A G T G A 240 C A T C A G T G A C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A T C A G T G A C C A G A T A T T G C A G A T A T T G	CATTA phth.seq CATTA phtc.SEQ CACTA phte.SEQ AGAGC Majority 250 AGAAT phth.SEQ AGAGC phth.seq
151 151 211 258 201 201 261	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C C T C A T G A A A G A T T C C T C A T G A A A G A T	G G A T C A A G G T T 7 G G A T C A A G G T T 7 A G A T C A G G G C T 7 G C A A G G T T C C T 7 G C A A G G T T C C C T 7 G C A A G G T T C C C T 7 G C A A G G T T C C T 7 C C A A A G T T C C T 7 C C G A A T T A T C A C C C G A A T T A T C A C	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G C C A T 230 F A T G A T G C C C T A T G A T G C C C T A T G A T G C C C T A T G A T G C C C T 280 5 T T G A A G A T T 5 T T G A A G A T T 5 T T G A A G G A T	CATGGAGAC CATGGAGAC CATGATGA 240 ICATCAGTGA ICAGATATTG	CATTA phtB.seq CATTA phtD.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT PhtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.SEQ TCAAT Majority 300 TTAAT phtA.SEQ TAAAT phtA.SEQ
151 151 211 258 201 201 261	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 210 TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT	G G A T C A A G G T T T A G G A T C A A G G T T C A G G G C T T A G G G G C T T A G G G G C T A G G G G C T A G G G G C T C C T T G C A A G G T T C C T T G C A A G G T C C C T T G G A A A G T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C C C A A G G T C C C T T C C C C C C C C C C C C	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A 230 FA T G A T G C C A T FA T G A T G C C A T FA T G A T G C C C A T A T G A T G C C C A T A T G A T G C C C C T T T G A A G G A T T C T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T	CATGGAGAC CATCAGTGA 240 CATCAGTGA CAGAGACATTG CAGAGACATTG	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.SEQ TCAAT Majority 100 TTAAT phtA.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ
151 151 211 258 201 201 261	G T T A T C A A G A T T A C G T C A T C A A G A T T A C G T A A T C A A A A T T A C T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C A T T A C T A T A A T G T C C T C A T G A A A G A T T C C T C A T G A A A G A T	G G A T C A A G G T T T A G G A T C A A G G T T C A G G G C T T A G G G G C T T A G G G G C T A G G G G C T A G G G G C T C C T T G C A A G G T T C C T T G C A A G G T C C C T T G G A A A G T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C T T C C C C A A G G T C C C T T C C C C C C C C C C C C	A T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A 230 FA T G A T G C C A T FA T G A T G C C A T FA T G A T G C C C A T A T G A T G C C C A T A T G A T G C C C C T T T G A A G G A T T C T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T T T G A A G G A T T	CATGGAGAC CATCAGTGA 240 CATCAGTGA CAGAGACATTG CAGAGACATTG	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.SEQ TCAAT Majority 100 TTAAT phtA.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ
151 151 211 258 201 201 261	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 210 TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT	G G A T C A A G G T T A A G G T T A A G G T T C A C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G G T T C C T A A G C C G A A T T A T T C A G C C G A A T T A T C A C C C G A A T T A T C A C C C C A A A C T A T A	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G C C A 230 F A T G A T G C C A T F A T G A T G C C C A F A T G A T G C C C A F A T G A T G C C C C 270 280 F A T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T	CATGGAGAC CATGAGAGA CATCAGTGA CAGATATTG CAGACATTG CAGACATTG CAGACATTG	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAGC phtB.seq AGAAGC phtB.SEQ AGAAC phtE.SEQ TCAAT Majority 300 TCAAT phtA.SEQ TCAAT phtB.seq TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ
151 151 211 258 201 201 261	GTTATCAAGATTAC GTCATCAAGATTAC GTCATCAAGATTAC TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT	G G A T C A A G G T T T A G G A T C A A G G T T C C T T A A G G T T C C T T A A G G T T C C T T A A G G T C C T T A A G G T T C C T T A A G C C C A A A C T A T A A C C C C A A A C T A T C A A C C C C	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G C C A T 230 F A T G A T G C C C T A T G A T G C C C T A T G A T G C C C T A T G A T G C C C T A T G A T G C C C T 280 G T T A A A G A T T G T T G A A G G A T T G T T G T T G T G T T G T T G T T G T G T T G T T G T T G T T T G T T T G T T T T	TCAT GGAGAC TCAT GAGAC TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAT CAGTGA TCAGATATTG TCAGATATTG TCAGACATTG TCAGACATTG TCAGACATTG	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAGC phtB.seq AGAAGC phtB.SEQ AGAAC phtE.SEQ TCAAT Majority 300 TCAAT phtA.SEQ TCAAT phtB.seq TCAAT phtB.SEQ TCAAT phtB.SEQ TCAAT phtB.SEQ
151 151 211 258 201 201 261 108 251 251 311	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT	G G A T C A A G G T T T A G G A T C A A G G T T C A A G G T T C A A G G T T C A A A G G T T C A A A G G T C C T T C A A A G G T C C T T C C C A A A G T T C C T T C C C C A A A C T A T A A C C C C	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A E A T G A T G C C A T F A T G A T G C C C A F A T G A T G C C C A F A T G A T G C C C A F A T G A T G C C C C A F A T G A T G C C C C A F A T G A T G C C C C A F A T G A A G G A T T F T G A A G G A T T F T G A A G G A T F T G A A G A T G G A F T G A A G A T G G A F T G A G G T A G A T G G A F T G A G G T A G A T G G A F T G A G G T A G A T G G A F T G A G G T A G A T G G A F T G T G T G T G T G T G T G T G T G T	CATGGAGAC ACACGGTGAC ACACGGTAC ACACGGTGAC ACACGGTAC ACACGACACAC ACACGGTAC ACACGACACACAC ACACGACACACACACACACACAC	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAC phtB.seq AGAGC phtB.SEQ AGAAC phtB.SEQ TCAAT Majority 300 TTAAT phtA.SEQ TCAAT phtB.seq
151 151 211 258 201 201 261 108 251 251 311	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 210 TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT	G G A T C A A G G T T T A G G A T C A A G G T T C C T T A G G T T C C T T A G G T T C C T T A G G T T C C T T A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T C T T C T T C T T C T T C T T C T T C T T C T T T C T T T T C T T T T C T T T T C T T T T C T T T T C T T T T C T T T T T C T T T T C T T T T C T T T T C T T T T T C T T T T T C T T T T T T C T T T T T C T	230 FATGATGACCTAT TOTAACGTCA TATGATGCCAT TATGATGATGCAT TATGATGATGCAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGATG	CATGGAGAC CATCAGTGA 240 CATCAGTGA CATCAG	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.SEQ TCAAT Majority 100 TTAAT phtA.SEQ TCAAT phtB.SEQ TTTA Majority 150 GTTTA phtA.SEQ
151 151 211 258 201 201 261 251 251 311	GTTATCAAGATTAC GTCATCAAGATTAC GTCATCAAGATTAC TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT TCATGAAAGGTTGG AAGGTCAAAGGGTGG GAAATCAAAGGGTTGG	G G A T C A A G G T T Y G G A T C A A G G T T T A G A T C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G G T T C C T T G C A A G T T C C T T C C G A A T T A T C A C C C G A A T T A T C A C C C G A A T T A T C A C C C G A A T T A T C A C C C C A A C T A T A T C A C T T A T G T T A T C A A T T A T G T T A T C A A T T A T G T T A T C A A T T A T G T T A T C A A T T A T G T T A T C A A T T A T G T T A T C A A T T A T G T C A T T T A T C A A T T A T G T T A T C A A T T A T G T C A T T T A T C A A T T A T G T C A T T A T C A A T T A T G T C A T T A T C A A T T A T G T C A T T A T C A A T T A T G T C A T T A T C A A T T A T G T C A T T A T C A A T T A T G T C A T T A T C A A	TATGATGATGTAT Z310 FATGATGCTAT FATGATGCCAT GTTGAAGGATT GTTGAAGGATT GTTGAAGGATT GTTGAAGGATT GTTGAAGGATT GTTAAAGATGGAT GGTTAAAGATGGAT GGTTAAACGAT	TCATTGAGAGACATGATGACATCATCAGTGAATCAGTGAATCAGTGAATCAGTGAATCAGTGAATCATTAGTGAATCAGTGAATCAGAGATATTGAGAGACATTGAAATACTATTAAAATACTAT	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAGC PhtB.seq AGAGC PhtB.seq AGAGC PhtB.SEQ TCAAT Hajority TCAAT PhtA.SEQ TCAAT PhtB.seq
151 151 211 258 201 201 261 108 251 251 311	GTTATCAAGATTAC GTCATCAAGATTAC GTAATCAAAATTAC TCATTACTATAATG 210 TCATTATTACAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCATTACTATAATG TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCCTCATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT TCTTGATGAAAGAT	G G A T C A A G G T T T A G G A T C A A G G T T C C T T A G G T C C T T A G G T C C T T A G G T C C T T A G G T C C T T A G G T C C T T C C C C A A G G T T C C T T C C C C A A C T T A T C A G C C C A A C T A T A T C A G C C C A A C T A T A T C A G C C C A A C T A T A T C A G C C C A A C T A T A T C A G C C C A A C T A T C A A G C C C C A A C T A T C A A G C C C C A A C T A T C A A G C C C C A A C T A T C A A G C C C C A A C T A T C A A G C C C C C C C C C C C C C C C C	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A E A T G A T G C C A T E A T G A T G C C C A E A T G A T G C C C A E A T G A T G C C C A E A T G A T G C C C C A E A T G A T G C C C C A E A T G A A G G A T T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A A G G A T E T T G A G G T A G A T G G A A G G T A G A T G G G G A A G G T A G A T G G G A A G G T A G A T G G G A A G G T A G A T G G G A A G G T A G A T G G G A A G G T A G A T G G G G T G G A A G G T A G A T G G G C G G T G G C G G T G G C G C G	CATGGAGAC CATGGAGAC CATGAGTGA CATCAGTGA	CATTA phtB.seq CATTA phtB.SEQ CACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAAC phtB.seq AGAAC phtB.SEQ TCAAT Majority 100 TTAAT phtA.SEQ TCAAT phtB.seq TTAA Majority 150 GTTTA phtA.SEQ GTTTA phtB.SEQ GTTTA phtB.SEQ

9/17

Figure 7(b)

	CCTTA	λG	G A	Ŧ	c	A G	c	c	λ 1	G	c	6	A T	A A	T	G T	· T	c s	G	c	λ.			Gλ	λ	6	A G	A	τ 1	· A	Hajor	ity
				360						371	•					38	0						390							400	•	
408	CCTTA	A G	Gλ	T	c	T G	C I	: c	λ	G	c c	G	A T	λ.	\ c	C I	C	c c	T	· c	λ.		. A 1	G A	G	G .	١.٨	À	T	: A	phtA.	SEO
351	CCTTA																															
351	CCTTA																															
411	CCTGA			. т (A G		re			٠.	•	•	^ -	• •				^ '		T			5 X	т	G ,		^	ŢÇ		phtE.	SEQ
	A T C G T	. c y	G A		: c	A G	6	A A	c x	Ť	A C	T	C A	Ť	A A	T C	A	T G	A C	0	G 1	rc	G.	A X	c	T.		-	۸ -		Hajos	ity
				410						42	,					43	0						440							450	· •	
458	ATCGA	. c .	A A			<u>, , , , , , , , , , , , , , , , , , , </u>	G	l G	c x	T	A G	T	C A	A (: 1	Ŧ	G	T G	A	G	G 1	G	6	Ä A	c	T (- c	<u> </u>	A G	: 1	phta.	SED
401	AACGT																														phtB.	
401	AACGT																														•	
461	ATCGT	CA	^ ^			^ ^		٠.	C ,	. т	6 1		^ ^	^ '	• •	TA		T G	A	-	•	-	-		-		• -	-	,	G	phtE.	SEQ
	GATGA	тx	хт	G	7	O T	T	3 C	ŦG	Ť	A G	C	C A	G J	Ť	c c	c.	۸.	0 0	; A	e c	<u> </u>	т :	A T	λ	c i	۸ ۸	С	G	, A	Major	ity
				460						470)					48	0						490							500)	
	AACGA																															
445	GATAA																															
445	AACGA																															
***																															•	
	TGATG	G T	T A	77	T	C T	T	Α ?	λ 7	6	C X	T	C T	0)	T	A T	· c	A T	Ţ) A	0	٠.	T	A C	G	G	3 T	G	λ :	٤٠	Kajor	ity
				510						520)					53	0						540							550	•	
	TGATE			_		==			-	_		_			_	_	_	. =	_		_	-			-	_	_				nhes	CPA
558	TGATG	GT	TA	т,	T	CT	т.		A T		CI	T	CT	G 1	L T	A T	C.	A T	y (3 %	G	; A	τ.	N C	T	G	3 T	G	A 7		pu.	360
492	TGATG	GG	T A	T A	T	C T	T	: A	A T	G	C A	Ť	C T	g J	T	A T	c.	A T	Ť	3 A	6 0	3 A	c	R C	G	G	3 7	G	A 1	C G	phtB.	800
492 495	TGATG	GG	T A T A	TA	T	C T	T	: A : A	A T	G	C A	T	C T	G 1	T	A T A T	C.	A T	T	3 A 3 A	G (; A	c .	A C	G	G 1	5 T	G	A 1	. G	phtB.	800 SEQ
492	TGATG TGATG	GG	T A T A	TA	T	C T	T	: A : A	A T	G	C A	T	C T	G 1	T	A T A T	C.	A T	T	3 A 3 A	G (; A	c .	A C	G	G 1	5 T	G	A 1	. G	phtB.	800 SEQ
492 495	TGATG	G T	T A T A	TA	T	CT	T	. A . A	A T A T	C	C A C A	T	C T C T	9 1	T	A T A T	C	A T A T	TO	3 A 3 A 3 A	G C	3 A 3 A 3 A	C	A C	0	G	3 T	G G A	A T	. G	phtB. phtD. phtE.	SEQ SEQ
492 495	TGATG TGATG	G T	T A T A C G	TA	T	CT	T	. A . A	A T A T	C	C A C A	T	C T C T	9 1	T	A T A T	C	A T A T	TO	3 A 3 A 3 A	G C	A	C	A C	0	G	3 T	G G A	A T	0 0	phtB. phtD. phtE. Major	SEQ SEQ
492 495 549	TGATG TGATG TGATG	G T G T A T	TA TA CG	T 7 T T T T T T T T T T T T T T T T T T	T	C T	TO	1 A T	A T	C 570	C A C A G A	.T	C T C T	0 J 0 J T 1	T	A T A T C C	C	A T A T T T	T C	: A : A	G G G G A G	A & A	C	A C	G	6 1	G T G T	G G A T	A T A T	600	phtB. phtD. phtE. Kajor	SEQ SEQ SEQ
492 495 549 608 542	TGATG TGATG TGATG CTTAT CTTAT	G T G T A T	TA TA CG	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T	C T C T	T C	T T	A T A T	C C	C A C A G A	T	C T C A	TI	TA	A T A T C C 58	C C T	A T A T T T T T T T T T T T T T T T T T	A C	: A : A : A : A	G G G A G	2 2 2	590	A C	GGAAA	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	GT	G O A T	A 1 A 1 G 3	0 0 0	phtB. phtD. phtE. Major phtA. phtA.	SEQ
492 495 549 608 542 545	TGATG TGATG TGATG CTTAT CTTAT	G G T A T A T A T	TA TA CG CG CG	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T	C T C T C T C T C T	CI	TOC	A T A T G G G G G	570 A	G A G A G A	T	C T C A	TITITI	TA	A T A T C C C C C C	CCTA	A T T T T T T T T T T T T T T T T T T T	A C		TTTTTT	AGA	590		GGGA	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	G T G T	G G A T T T T	A 2 G 3	600	phtB. phtD. phtE. Major phtA. phtB. phtB.	SEQ SEQ SEQ SEQ SEQ SEQ SEQ
492 495 549 608 542 545	TGATG TGATG TGATG CTTAT CTTAT CTTAT	G G T A T A T A T	TA TA CG CG CG	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		CT CT CT	CICI	TOOT	A T G G G G G G G G	570 A C C A	G A G A G A G A	T	C T C A C A C A C A C A C A C A C A C A	TITT	T . A . A . A . A . A . A . A . A . A .	A T C C C C C C C C C C C C C C C C C C	C C T	A T T T T T T T T T T T T T T T T T T T	A CA	: A : A : A : A	G G G G G G G G G G G G G G G G G G G	A	C		GGGA	G : G : A :	G T G T A A A A A A A A A A A	GOA T	A 1 G 2 G 2 G 2 G 2	600	phtB. phtD. phtE. Hajor phtA. phtB. phtB. phtB.	SEQ SEQ SEQ SEQ SEQ SEQ SEQ
492 495 549 608 542 545	TGATG TGATG TGATG CTTAT CTTAT	G G T A T A T A T	TA TA CG CG CG	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		CT CT CT	CICI	TOOT	A T G G G G G G G G	570 A C C A	G A G A G A G A	T	C T C A C A C A C A C A C A C A C A C A	TITT	T . A . A . A . A . A . A . A . A . A .	A T C C C C C C C C C C C C C C C C C C	C C T	A T T T T T T T T T T T T T T T T T T T	A CA	: A : A : A : A	G G G G G G G G G G G G G G G G G G G	A	C		GGGA	G : G : A :	G T G T A A A A A A A A A A A	GOA T	A 1 G 2 G 2 G 2 G 2	600	phtB. phtD. phtE. Hajor phtA. phtB. phtB. phtB.	SEQ SEQ SEQ SEQ SEQ SEQ SEQ
492 495 549 608 542 545	TGATG TGATG TGATG CTTAT CTTAT CTTAT	G G T A T A T A T	TA TA CG CG CG CG	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		CT CT CT	CICI	TOOT	A T G G G G G G G G G G G G G G G G G G	570 A C C A	CA CA GA GA GG	T	C T C A C A C A C A C A C A C A C A C A	TITT	T . A . A . A . A . A . A . A . A . A .	A T C C C C C C C C C C C C C C C C C C	C C T	A T T T T T T T T T T T T T T T T T T T	A CA	: A : A : A : A	G G G G G G G G G G G G G G G G G G G	ASA	C		GGGA	G : G : A :	G T G T A A A A A A A A A A A	GOA T	A 1 G 2 G 2 G 2 G 2	600	phtB. phtD. phtE. Major phtA. phtB. phtB. phtB.	SEQ SEQ SEQ SEQ SEQ SEQ SEQ
492 495 549 608 542 545 599	TGATG TGATG TGATG CTTAT CTTAT CTTAT CTTAT TTATC	G G T G T A T A T A T A G A G	TA TA CG CG CG CG CC CG	T 7 T T T T T T T T T T T T T T T T T T	T C	CTT CTT CTT A	C / C / C / C / C / C / C / C / C / C /	TCCTT	A T G G G G G G G G G G G G G G G G G G	570 A C C A C 620	GAAGGG TG	TOOT	CTT CA	TTTTCT	T A A A A A	ATT C SECCCT A C G A A	C C T A A A A A A A A A A A A A A A A A		A CA	: A : A : A : C : C	G G G A G T T T T T T T T T T T T T T T	AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG	C		G G G A A A A A G	GGGA	G T T A A A A A A A A T C	GOA T TTTC G	A T G J G J G J G J	60 G G G G G G G G G G G G G G G G G G G	phtB. phtD. phtE. Major phtA. phtB. phtB. phtB.	SEO
492 495 549 608 542 545 599	TGATG TGATG CTTAT CTTAT CTTAT CTTAT TTATC TTATC	G G T G T A T A T A T A G A G A G	TAATA CGCCGGCCGCCGCCTCTT	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T C C C C C C C C C C C C C C C C C C C	CTT CTT A	C C C C C C C C C C C C C C C C C C C	AAA T TOOT T	A T G G G G G G G G G G G G G G G G G G	570 A C C A	GAAA GAAAGG TG	TOOT		TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	T A A A A A A	A T C 58 C C C C C C A 63 A A	C C T A O A A A A A A A A A A A A A A A A A		ACC	: A : A : A : A : A : A : A : A : A : A	G G G A G T T T T T T T T T T T T T T T	A COOOL A CAR	C C C C T 640		G G G A A A A A G G G	G G G A G G A	GTT A AAAG T	GOA T TTTC G	A T G J G J G J G J G J G J G J G J G J G	65 G G G G G G A	phtB. phtD. phtE. Major phtA. phtB. phtB. phtB. phtB. phtB.	SEO
608 542 545 549 608 542 545 599	TGATG TGATG CTTAT CTTAT CTTAT CTTAT TTATC TTATC	G G T T A T A T A T A G A G A G	TAA CGCCGGCGCGCCTCTT	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		CTT	C C C C C C C C C C C C C C C C C C C	AAA T TOOT T TTT	A T G G G G G G G G G G G G G G G G G G	570 A C C A C 620 C C	CCA A GGGG T GGGTT	THE THE COLUMN	CTT A A A A G G G G G G G G G G G G G G G	TTTTCT	TT A A A G G G G	A T C	C C T A A A A A A A A A A A A A A A A A		ACC ACC ACC	: A A A	G G A C T T T T T T T T T T T T T T T T T T	A A C C C C C A C A C A C A C A C A C A	590 CCC T		GGG A AAAA G GGG	G G G A A	GGT A AAAG TAA	GOA T TTTC C CGC	A T G J G J G J G J G J G J G J G J G J G	G G G G G G A A	phtB. phtD. phtA. phtB. phtB. phtB. phtB. phtB. phtD. phtE. Hajor	SEO SEO SEO SEO SEO SEO SEO SEO SEO SEO
492 495 549 608 542 545 599	TGATG TGATG CTTAT CTTAT CTTAT CTTAT TTATC TTATC TTATC	G G T T A T A T A T A G A G A G	TAA CGCCGGCGCGCCTCTT	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		CTT	C C C C C C C C C C C C C C C C C C C	AAA T TOOT T TTT	A T G G G G G G G G G G G G G G G G G G	570 A C C A C 620 C C	CCA A GGGG T GGGTTTT	THE THE COLUMN	CTT A A A A G G G G G G G G G G G G G G G	TTTTCT	TT A A A G G G G	A T C	C C T A A A A A A A A A A A A A A A A A		ACC ACC ACC	: A A A	G G A C T T T T T T T T T T T T T T T T T T	A A C C C C C A C A C A C A C A C A C A	590 CCC T		GGG A AAAA G GGG	G G G A A	GGT A AAAG TAA	GOA T TTTC C CGC	A T G J G J G J G J G J G J G J G J G J G	G G G G G G A A	phtB. phtD. phtA. phtB. phtB. phtB. phtB. phtB. phtD. phtE. Hajor	SEO
608 542 545 549 608 542 545 599	TGATG TGATG CTTAT CTTAT CTTAT CTTAT TTATC TTATC	G G T A T A T A T A G A G G T G	TAA CG CGG CG CT CT CT CT	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T : C : C : C : C : C : C : C : C : C :	CTT CTT CA	C J C J C J C J C J C J C J C J C J C J	AAA T TCCT T TTTT	A T G G G G G G G G G G G G G G G G G G	570 A C C A C C C C C	GAA GAA GGGGGGTGTA	TOUT		T1 T1 T1 T1 C1 C1	T A A A G G G A	A T C 58 C C C C C A 63 A A A A	C C T A C A A A A A A A A A A A A A A A	A T T T T T T T C C C C C C C C C C C C	ACC ACC TO	: A	GGA 77	AAA C CCCC A AAAA AAAA AAAA AAAA AAAA	C C C C T 640		GGG A AAAA G GGGG	G G G A A G	3	GOA T TTTC C CGCC	A 1	G G G G G G A A A	phtB. phtD. phtA. phtB.	SEO
608 542 545 549 608 542 545 599	TGATG TGATG CTTAT	G G T A T A T A T A G A G G T G	TAA CG CG CG CT CTT CTT CTT CTT CTT CTT CTT	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T : C : C : C : C : C : C : C : C : C :	CTT	C J C J C J C J C J C J C J C J C J C J	AAA T TCCT T TTTT	GGGGGGGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	570 A C C A C C C C C	GAAAA GAAAA GTTTGG	TOUT		T1 T1 T1 T1 C1 C1	T A A A G G G A	A T C 58 C C C C C A 63 A A A A	C C T A A A A A A A A A A A A A A A A A	A T T T T T T T C C C C C C C C C C C C	ACC ACC TO	: A	GGA 77	AAA C CCCC A AAA C	C C C C T 640		GGG A AAAA G GGGG	G G G A A G	3	GOA T TTTC C CGCC	A 1	G G G G G G A A A	phtB.	SEO
608 542 545 549 608 542 545 599	TGATG TGATG CTTAT	G G T A T A T A G A G G G T A T A T A G A G	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T C C C C C C C C C C C C C C C C C C C	CCC T T A AAAA -	C C C C C C C C C C C C C C C C C C C	TCCT T TTTT G	AAA GGGGG AGGGAAA GAAAGAA	570 A C C A C C C C T 670 A	CCC GAAG G TTTTA CT	TTG T	CTT A AAAA G FFFFA T T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	T A A A G G G A T	AAA C 58 CCCCC A 63 AAAA C 68 G	C C T A O G G G G G G G G G G G G G G G G G G	AAA TT TTTTT C CCCA C	A A A A A A A A A A A A A A A A A A A	: A	GGA T 7 T 7 T 7 T 7 T 7 T 7 T 7 T 7 T 7 T	AAA C C C C C AAAA C C C AAAA C C C AAAA C C C AAAA C C C C C AAAA C C C C C C AAAA C C C C C C C C AAAA C	590 C T T T T T T T T T T T T T T T T T T		GGG A AAAA G GGGG T	GGGA GAAG	G G G G A A A A A A A A A A A A A A A A	GOA T TTTC C CGCC A A	A T G J G J G J G J G J G J G J G J G J G	G G G G G G G G G G G G G G G G G G G	phtB. Hajor phtA. phtB.	SEO
608 542 545 545 547 545 599 658 592 595 649	TGATG TGATG CTTAT CTTAT CTTAT TTATC TTATC TTATC TTATC TTATC	G G T A T A T A G A G G G T A T A T A G A G	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	T 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	T C C C C C C C C C C C C C C C C C C C	CCC T T A AAAA -	C C C J J C J C J C J C J C J C J C J C	TAA T TCCT T TTTT G	AAA G GGGG A G GAAA G AA	GGC C 570 A C C A C 620 C C T 670 A T	CCC G AAAG G TTTTA CT CCT	TTOOT OF TOO	CTT A CAAAA G FFFFFFFFFFFFFFFFFFFFFFFFFFFFF	T1111111111111111111111111111111111111	T A A A A G G G A T	AAA C 58 CCCC A 63 AAAA C 68 GC	C C T A O A A A O G G G G G G G G G G G G G G	AAA TT TTTT C CCCA C	A A A A A A A A A A A A A A A A A A A		GGA T T T T T T T T T T T T T T T T T T	AAAA COOO AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	590 C C C T T T T T T T T T T T T T T T T	ARA I ARAA I IIII A AAAA AAAA AAAA AAAA	GGG A AAAA G GGGG T TT	G G G G G A A G G A A G T A A T A T A T	G	GOA T TTTC C GGGG A AA	A 1 2 3 3 4 4 5 5 5 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5	G G G G G G A A T 700 T G	phtB.	SEQ
608 542 545 599 658 592 595 649	TGATG TGATG CTTAT CTTAT CTTAT TTATC TTATC TTATC TTATC TTATC	G G T A T A T A G A G G G T A T A T A G A G	TAA CCCCCC CT CCCCC - C TCCC	T 7 T T T T T T T T T T T T T T T T T T	C C C C C C C C C C C C C C C C C C C	CTT T AAAA -	C C C J J J C C C J C C C C C C C C C C	A T T T T T T T T T T T T T T T T T T T		GGC C-570 ACCA C-620 CCC T-670 ATT	CCC G AAAG G TTTTA C CCT	TTO T CO CO CO TO CO	CTT A CAAAA G FOGG T ATT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	T A A A A G G G A T	ATT C. SECCE A. G. AAAA C. G. G. C.	C.C.T.A.O.A.A.A.O.G.G.G.G.G.G.G.G.G.G.G.G.G.G	AAA TT TTTTT C CCCA C ACC	AAAA TOO AAAA		GGA T T T T T T T T T T T T T T T T T T	AAA C CCCC A TAAA C CCCC	590 CCCC T 640 T 7	ARA T RARA T TOOT S AGA	GGG A AAAA G GGGG T TTT	GGGA G	CCC A AAAG T CTTT T CTT	GOA T TTTC C CGCG A AAA	A A A A A A A A A A A A A A A A A A A	G G G G G G G G G G G G G G G G G G G	phtB. Hajor phtA. phtB.	SEO

10/17

Figure 7(c)

	CAA-ATCCAG	CTCAGTAC	C A A G	ATTGTCAGAG	AACCACAAT -	- C T Majority
	711		720	730	740	750
758	CAAGAACAAA	CIGGGTAC	CTTCTGTAA	GCAATCCAGG	AACTACAAAT	A C T phtA.SEO
677	CAA-ATCCAG					
680	CAA-ATCCAG	C T C A A C	C A A G	ATTGTCAGAG	AACCACAAT -	- C T phtD.SEQ
728	C A A C A G	CT - AGT	G	A C A A T A A C	A C G C A A T C	T G T phtE.SEQ
			1.	A - TCAAGCAA		
	760	·	770	780	790	800
	AACACAAGCA					
717	G A C T					
720 759	ACCANANG-G			A - T C A A A A A G C C A G C A A		
739	X C C X X X X C + C			~~~~~~		p
	CATTTCAAGT	CTTTTGCG	TGAATTGTA	TGCTAAACCT	TTATCAGAAC	GCC Majority
	810	•	B20	830	840	850
858	CATTGATAGT	CTCTTGAA	ACAGCTCTA	CAAACTGCCT	TTGAGTCAAC	GAC phta.SEO
756	CATTTCAAGC					
759	CATTTCAAGC	CTTTTACG	TGAATTGTA	TGCTAAACCC	TTATCAGAAC	G C C phtD.SEQ
801	TCTCCAGAGT	CITTTGAA	GGAACTCTA	TGATTCACCT	AGCGCCCAAC	GTT phtE.SEQ
	ATGTGGAATC	****		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		CG & Majority
	860		670	8 0	890	900
908	A T C T A C A A T C					
					AATCACAAGT	
806	ATGTGGAATC	TGATGGCC	T T A T T T T C G	A C C C A G C G C A	AATCACAAGT	CGA phtB.seq
806 809	ATGTGGAATC ATGTGGAATC	TGATGGCC TGATGGCC	T T A T T T T C G T T A T T T T C G	A C C C A G C G C A A C C C A G C G C A	A A T C A C A A G T	CGA phtB.seq
806 809	ATGTGGAATC	TGATGGCC TGATGGCC	T T A T T T T C G T T A T T T T C G	A C C C A G C G C A A C C C A G C G C A	A A T C A C A A G T	CGA phtB.seq
806 809	ATGTGGAATC ATGTGGAATC	TGATGGCC TGATGGCC AGATGGCC	T T A T T T T C G T T A T T T T C G T G G T C T T T G	A C C C A G C G C A A C C C A G C G C A A C C C T G C T A A	A A T C A C A A G T A A T C A C A A G T G A T T A T C A G T	CGA phtB.seq CGA phtD.SEQ CGT phtE.SEQ
806 809	ATGTGGAATC ATGTGGAATC ACAGTGAATC	T G A T G G C C T G A T G G C C A G A T G G C C	T T A T T T T C G T T A T T T T C G T G G T C T T T G	A C C C A G C G C A A C C C A G C G C A A C C C T G C T A A	A A T C A C A A G T A A T C A C A A G T G A T T A T C A G T	CGA phtB.seq CGA phtD.SEQ CGT phtE.SEQ
806 809 851	A T G T G G A A T C A T G T G G A A T C A C A G T G A A T C A C C G C C A G A G	T	TTATTTTCG TTATTTTCG TGGTCTTTG	ACCCAGCGCA ACCCAGCGCA ACCCTGCTAA GGTGACCATT	A A T C A C A A G T A A T C A C A A G T G A T T A T C A G T A C C A C T T T A T 940	CGA phtB.seq CGA phtB.SEQ CGT phtE.SEQ
806 809 851	ATGTGGAATCATGTGGAATCACAGTGAATCACAGTGAATCACAGCCAGAG	TOATGGCC TGATGGCC AGATGGCC GTGTTGCT GTGTTGCT	TTATTTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTGCCACAC	ACCCAGCGCA ACCCAGCGCA ACCCTGCTAA GGTGACCATT 930 GGAGATCATT	AATCACAAGT AATCACAAGT GATTATCAGT ACCACTTTAT 940 ACCACTTCAT	C G A phtB.seq C G A phtD.SEQ C G T phtE.SEQ C C C C Hajority 950
806 809 851	A T G T G G A A T C A T G T G G A A T C A C A G T G A A T C A C C G C C A G A G	TOATGGCC TGATGGCC AGATGGCC GTGTTGCT GTGTTGCA GTGTTGCA	TTATTTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTGCCACAC GTCCCTCAT	ACCCAGCGCA ACCCAGCGCA ACCCTGCTAA GGTGACCATT 930 GGAGATCATT GGTAACCATT	AATCACAAGT AATCACAAGT GATTATCAGT ACCACTTTAT 940 ACCACTTCAT ACCACTTCAT	C G A phtB.seq C G A phtD.SEQ C G T phtE.SEQ C C C Majority 950 C C C phtA.SEQ C C C phtB.seq
806 809 851 958 856 859	ATGTGGAATC ATGTGGAATC ACAGTGAATC ACCGCCAGAG 91 ACAGCTAGAG ACCGCCAGAG	T G A T G G C C T G A T G G C C A G A T G G C C G T G T T G C T G T G T T G C A G T G T A G C C G T G T A G C T	TTATTTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTGCCACAC GTCCCTCAT GTCCCTCAT	A C C C A G C G C A A C C C T G C T A G G T G A C C A T T 930 G G A G A T C A T T G G T A A C C A T T G G T A A C C A T T	AATCACAAGT AATCACAAGT GATTATCAGT ACCACTTTAT 940 ACCACTTCAT ACCACTTTAT ACCACTTTAT	C G A phtB.seq C G A phtD.SEQ C G T phtE.SEQ C C C Majority 950 C C C phtB.seq C C C phtB.seq C C C phtB.seq
806 809 851 958 856 859	ATGTGGAATC ATGTGGAATC ACAGCCAGAG 920 ACAGCTAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACAGCTAGAG	TOATGGCC TGATGGCC AGATGCC GTGTTGCT GTGTTGCA GTGTAGCT GTGTAGCT GAGTTGCG	TTATTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTCCCTCAT GTCCCTCAT GTCCCTCAT ATTCCGCAT	ACCCAGCGA ACCCTGCTA ACCCTGCTA GGTGACCATT 930 GGTGACCATT GGTAACCATT GGTAACCATT	AATCACAAGT AATCACAAGT GATTATCAT 940 ACCACTTTAT ACCACTTTAT ACCACTTTAT ACCACTTTAT	C C A phtB.seq C C A phtB.SEQ C C C phtB.SEQ
806 809 851 958 856 859	ATGTGGAATC ATGTGGAATC ACAGCCAGAG ACAGCTAGAG ACAGCTAGAG ACCGCCAGAG ACCGCCAGAG ACAGCCAAATG	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C G A T G T C T G A	TTATTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTGCCACAC GTCCCTCAT GTCCCTCAT ATTCCACAT ATTCCACAT	A C C C A G C G C A A C C C C A G C G C	AATCACAAGT AATCACAAGT ACCACTTTAT 940 ACCACTTTAT ACCACTTTAT ACCACTTTAT ACCACTTTAT	C C C PhtA.SEQ C C C PhtA.SEQ C C C PhtA.SEQ C C C PhtA.SEQ C C C PhtB.SEQ C C C C Hajority
806 809 851 958 856 859	ATGTGGAATC ATGTGGAATC ACAGCCAGAG 920 ACAGCTAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACAGCTAGAG	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C G A T G T C T G A	TTATTTCG TTATTTTCG TGGTCTTTG GTCCCTCAT 920 GTCCCTCAT GTCCCTCAT GTCCCTCAT ATTCCGCAT	ACCCAGCGA ACCCTGCTA ACCCTGCTA GGTGACCATT 930 GGTGACCATT GGTAACCATT GGTAACCATT	AATCACAAGT AATCACAAGT GATTATCAT 940 ACCACTTTAT ACCACTTTAT ACCACTTTAT ACCACTTTAT	C C A phtB.seq C C A phtB.SEQ C C C phtB.SEQ
958 851 958 856 859 901	ATGTGGAATC ATGTGGAATC ACAGTGAATC ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACACCAAATG TTATGAACAA TTATGAACAA TTACTCTAGA	T G A T G G C C T G A T G G C C G T G T T G C T	TTATTTCG TTATTTTCG TGGTCTTTT 920 GTGCCACAC GTCCCTCAT GTCCCTCAT ATTCCGCAT ATTCCGCAT ATTGGAAGA 970 ATTGGAAGA	A C C C A G C G C A A C C C C A G C G C	A A T C A C A A G T A A T C A C A A G T G A T A T C A G T G A C T T T A T A C C A C T T T A T A C C A C T T T A T A	C C A phth.seq C C A phth.SEQ C C C Majority 950 C C C phth.SEQ C C phth.SEQ C C C phth.SEQ C C C phth.SEQ C C C phth.SEQ C C C phth.SEQ
958 851 958 856 859 901	ATGTGGAATC ATGTGGAATC ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACACCAAATG TTATGAACAA TTATGAACAA	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C C G A G T T G C G A T G T C T G A A T G T C T G A	T T A T T T T C G T T A T T T T C G T G T T C T T T T T 920 G T G C C A C A C G T C C C T C A T A T T C C G C A T A T T C G G A A G A 370 A T T G G A A G A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G A A A A	A C C C A G C G C A A C C C A G C G C A A C C C A T T G C T A A C C A T T G C T A A C C A T T G C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C	A A T C A C A A G T A A T C A C A A G T A T A T A C A C A C T T T A T A C C A C T T T A T A	C C A phtB.seq C C C phtA.SEQ C C C phtA.SEQ C C C phtB.seq
958 851 958 856 859 901	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G A C C A G A G A C C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A	TTATTTCG TTATTTTCG TGGTCTTTTG GTCCTCAT 920 GTGCCACAC GTCCCTCAT GTCCCTCAT ATTCCGCAT ATTCGGAAGA ATTGGAAGA ATTGGAAAA	A C C C A G C G C A A C C C C A G C G C	A T C A C A A G T A A T C A C A A G T A T A C A C T T T A T A C C A C T T T A T A	C C A phth.seq C C A phth.SEQ C C C phth.SEQ
958 851 958 856 859 901	ATGTGGAATC ATGTGGAATC ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACACCAAATG TTATGAACAA TTATGAACAA	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G A	TTATTTCG TTATTTTCG TGGTCTTTTG GTCCTCAT 920 GTGCCACAC GTCCCTCAT GTCCCTCAT ATTCCGCAT ATTCGGAAGA ATTGGAAGA ATTGGAAAA	A C C C A G C G C A A C C C C A G C G C	A T C A C A A G T A A T C A C A A G T A T A C A C T T T A T A C C A C T T T A T A	C C A phth.seq C C A phth.SEQ C C C phth.SEQ
958 851 958 856 859 901	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G A C C A G A G A C C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C G A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G C A T G T C T G C	TTATTTCG TTATTTTCG TTGTTTTTG GTCCCTCAT 320 GTGCCACAC GTCCCTCAT ATTCCGTAT ATTCCGTAT ATTCGGAAAA ATTGGAAAAA ATTGGAAAAA CTTAGAAAAA	A C C C A G C G C A A C C C C A G C G C	A A T C A C A A G T A A T C A C A A G T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T C C T A T T A T T C C T T C T T C T C C T T C T T C T C	C C A phtB.seq C C C phtA.SEQ C C C phtA.SEQ C C C phtB.seq
958 851 958 856 859 901	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G A C C A C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C G A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G C A T G T C T G C T T C T G A A C C	T T A T T T T C G T T A T T T T C G T G G T C T T T T T	A C C C A G C G C A A C C C A G C G C A A C C C A T T 9300 G G A G A T C A T T G G T A A C C A T T G C C A C C A T T G C C A C C A T T G C C A C C A T T G C C A C C A T T C C C A C C A C C A C C A C C C A C C C A C C C A C C C A C C C A C	A T C A C A A G T A T C A C A A G T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T A C C A C T T T A T C C T A T T A T T C C T A T T T T C C T T T T T T T T T T T T T T	C C A phth.seq C C C phth.seq
958 851 958 859 901 1008 909 951	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G A C C A C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C G A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G C C T T C T G C	T T A T T T T C G T T A T T T T C G T G G T C T T T T 320 G T G C C A C A C G T C C C T C A T A T T C C G C A T A T T G G A A G A A T T G G A A A A A T T G G A A A A C T T A G A C A A A C T T A G A A A A A C T T A G A A A A A C T T A G A A A A A C T T A G A A A A A C T T A G A A A A A A C T T A G A A A A A A A A A A A A A A A A	A C C C A G C G C A A C C C A G C G C A A C C C A T T G C T A A A G A T T G C T A A A G A T T G C C C A G A T T G C C C A A A G A T T G C C C A G A T T G C C C A G A T T G C C C A G A T T G C C C A G A T T G C C C A G A T T G C C C A G A T T G C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C A G A T T C C C C C C C C C C C C C C C C C	A T C A C A A G T A T C A C A C T T T A T A C C A C T T T A T A	C C A phtB.seq C C C phtA.SEQ C C C phtA.SEQ C C C phtB.seq
958 856 859 901 1008 909 951	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G G A C C A G A G A C C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C T G T G T T G C T G T G T T G C A G T G T T G C A G T G T T G C A A T G T C T G A A T G T C T G A A T G T C T G C C T T T C T G C C T T T C T G C	T T A T T T T C G T T A T T T T C G T G G T C T T T T 920 G T C C C C C A T G T C C C C C A T G T C C C T C A T G T C C C T C A T G T C C C T C A T G T C C C T C A T A T T C C G C A T A T T G G A A G A A T T G G A A A A A T T G G A A A A C T T A G A A G A A T T G G A A A A C T T A G A A G A A T T G G A A A A C T T A G A A G A A T T G G G T A C 1020 A T T G G G T A C	A C C C A G C G C A A C C C C A G C G C A A C C C C A G C G C A G G T G A C C A T T G G T A A C C A T T G G C G A C C A T T A C G A A T T G C T A C G A A T T G C T A C G A A T T G C T A A G A T T G C C C A G A T T C A A G 1010 C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G C A G A T T C A A G	A T C A C A A G T A A T C A C A A G T T A T C A C A C T T T A T A C C A C T T T A T A	C C A phth.seq C C C A Kajority
958 855 958 856 859 901 1008 906 909 951	ATGTGGAATC ATGTGGAATC ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACCGCCAGAG ACACCAAATG TTATGAACAA TTATGATTATCG	T G A T G G C C T G A T G G C C G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G C T T C A A A C C G T C A A A C C	T T A T T T T C G T T A T T T T C G T G G T C T C T T T T 920 G T G C C A C A C G T C C C T C A T A T C C C C A A C A T T G G A A G A A T T G G A A A A A T T G G A A A A A T T G G A A A A A T T G G G T A C 1020 A T T G G G T A C 1020 A T T G G G T A C	A C C C A G C G C A A C C C A G C G C A A C C C A T T 9310 G A G A T C A C A T T G G T A A C C A T T G C C A C C A T T G C C A C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C	A A T C A C A A G T A A T C A C A A G T A T A T A T A C A C A C T T T A T A C C A C T T T A T A	C C A phth.seq C C C A phth.seq
958 856 859 901 1008 909 951	ATGT G G A AT C ATGT G G A AT C A C G C C A G A G G A C C A G A G A C C C A G A G	T G A T G G C C T G A T G G C C G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A G T G T T G C A A T G T C T G A A T G T C T G A A T G T C T G A A T G T C T G C T T C A A A C C G T C A A A C C	T T A T T T T C G T T A T T T T C G T G G T C T T T T	A C C C A G C G C A A C C C A G C G C A A C C C A T T 9310 G A G A T C A C A T T G G T A A C C A T T G C C A C C A T T G C C A C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C	A T C A C A A G T A T C A C A C T T T A T A C C A C T T T A T A	C C A phth. seq C C C phth. SEQ C C A phth. SEQ

11/17

Figure 7(d)

AGTCCACAATCGACTC	CGGAACCTA	TCCAAGTCC	GCAACCTGCAC	CAAA Hajority
1050	1070	1080	1090	1100
1108 AGTCCACAACCGACTC	CGGAACCTA	TCCAGGCCC	GCAACCTGCAG	CAAA phtA.SEQ
1006 AGTCCACAACCGACTC				phtB.seq
1009 A G T C C A C A A T C G A C T C			GCAACCTGCAC	CAAA phtD.SEQ
1013 GTTCTACAGTT	TCTA	CAAA	T G C A -	- A A A phtE.SEQ
TC-T-AAAGCTCCA	AGCAATCCA	TTGATG-GA	AATTGGTCAA	GAAG Majority
1110	1120	1130	1140	1150
1158 TCTTAAAATAGACTCA	AATTCT	CT	TTGGTTAG1	CAGC phtA.SEQ
			AATTGGTCAAJ	
1059 TCCTCAACCAGCTCCA				GAAG phtD.SEQ
1039 C C		·- T A A T G		phtE.SEQ
CTGTTCGAAAAGTAGG	CGATGGTTA	GTCTTTGAG	GAGAATGGAG1	TTCT Majority
1150	1170	1180	1190	1200
1196 TGGTACGAAAAGTTGG				
1088 CTGTTCGAAAAGTAGG				
1109 CTGTTCGAAAAGTAGG				
1046 A A G T A G -		GTCT	A G 7	CT phts.SEQ
CGTTATATCCCAGCCA	AGGATETTT	. A G C A G A A A C	AGCAGCAGGC	TTGA Majority
1210	1220	1230	1240	1250
1246 CGTTATGTCTTTGCGA				
1138 CGTTATATCCCAGCCA	AGGATOTTT	AGCAGAAAC	Y @ C Y @ C Y @ C Y	TTGA phtB.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA	AGGATOTTT	AGCAGAAAC	A G C A G C A G G C J	TTGA phtB.seq
1138 CGTTATATCCCAGCCA	AGGATOTTT	AGCAGAAAC	A G C A G C A G G C J	TTGA phtB.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA	AGGATCTTTC	A G C A G A A A C	Y G C Y G C Y G C C Y	TTGA phtB.seq TTGA phtB.SEQ
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062	AGGATCTTTC	A G C A G A A A C	Y G C Y G C Y G C C Y	TTGA phtB.seq TTGA phtB.SEQ
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062	AGGATETTT AGGATETTT CAGGAAAGTT 1270 CAAGAAAGTT	TTTCTCATA 1280	A G C A G C A G G C A G C C A G C A G C A G C C A G G C C A G G C C A G C T A G G A G C T A G G A G C T A A C T A C C C C T T T A A C T G C C C	TTGA phtB.seq TTGA phtD.SEQ phtB.SEQ TAAGA Majority 1300 TAAAA phtA.SEQ
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062	AGGATCTTTCAGGATCTTTCCCAGGAAAGTT	TTTCTCATA 1280 TTTCACACA TATCTCATA	A G C A G C A G C C A G C A G C A G C A G C A G C C A G C C A G C T A G C A G C T A G	TTGA phth.seq TTGA phtD.SEQ phtE.SEQ TAAGA Majority 1300 TAAAA phtA.SEQ AAGA phtB.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062 TAGCARACTGGCCAAG 1260 1296 AAGCAAGTTATCAAAA 1209 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG	AGGATCTTTCAGGATCTTTCCAGGAAAGTT	AGCAGAAAC TTTCTCATA 1280 TTTCACACACA TATCTCATA	A G C A G C A G C C A G C A G C A G C A G C A G C C A G C C A G C T A G C A G C T A G	TTGA phtB.seq TTGA phtB.SEQ TAAGA Majority 1300 TAAAA phtA.SEQ TAAGA phtB.seq TAAGA phtB.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062 TAGCARACTGGCCAAG 1260 1296 AAGCAAGTTATCAAAA 1209 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG	AGGATCTTTCAGGATCTTTCCCAGGAAAGTT	AGCAGAAAC TTTCTCATA 1280 TTTCACACACA TATCTCATA	A G C A G C A G C C A G C A G C A G C A G C A G C C A G C C A G C T A G C A G C T A G	TTGA phth.seq TTGA phtD.SEQ phtE.SEQ TAAGA Majority 1300 TAAAA phth.SEQ TAAGA phth.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062 TAGCARACTGGCCAAG 1260 1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT:	TTTCTCATA 1280 TTTCACACA TATCACACA TATCTCATA	AGCAGCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	TTGA phtB.seq LTTGA phtD.SEQ FAAGA Majority 1300 FAAGA phtB.SEQ FAAGA phtB.seq FAAGA phtB.seq FAAGA phtB.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062 TAG C A A A C T G G C C A A G 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1066	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT:	TTTCTCATA 1280 TTTCACACA TATCACACA TATCTCATA	AGCAGCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	TTGA phtB.seq LTTGA phtD.SEQ phtE.SEQ TAAGA Majority 1300 TAAGA phtB.seq TAAGA phtB.seq TAAGA phtB.seq TAAGA phtB.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT:	TTTCTCATA 1280 TTTCACACA TATCACACA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA	AGCAGCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	TTGA phtB.seq LTTGA phtD.SEQ
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062 T A G C A A A C T G G C C A A 1260 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G C C C A T C A A C T G A T C T T C C T T C 1310 1346 A A G A A A A T G T T G C T C C C A T C	A G G A T C T T T C A G G A A A G T T 1270 C A A G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T C A G G A A A G T G T A G T G A T C G G T A G T G A T C G G	TTTCTCATA 1280 TTTCTCACACA TATCTCATA TATCTCA	A G C A G C A G C C A G C C A G C A G C A G C A G C C C A G C C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C C A G C C C A G C C C C	TTGA phtB.seq 17TGA phtB.seq 17TGA phtB.SEQ FAAGA Majority 1300 FAAGA phtB.seq 1350
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: TAGTGATCGJ TAGTGATCGJ	TTTCTCATA 1280 TTTCACACA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA GAATTTTAC GAATTTTAC GAATTTTAC	AGCAGCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	TTGA phtB.seq LTTGA phtD.SEQ LTTGA phtD.SEQ LTTGA Majority 1300 LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAGAC Hajority 1350 LTGAC Hajority LTGAC phtB.seq LTGAC phtB.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062 T A G C A A A C T G G C C A A 1260 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G C C C A T C A A C T G A T C T T C C T T C 1310 1346 A A G A A A A T G T T G C T C C C A T C	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: TAGTGATCGJ TAGTGATCGJ	TTTCTCATA 1280 TTTCACACA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA GAATTTTAC GAATTTTAC GAATTTTAC	AGCAGCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	TTGA phtB.seq LTTGA phtD.SEQ LTTGA phtD.SEQ LTTGA Majority 1300 LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAAGA phtB.seq LAGAC Hajority 1350 LTGAC Hajority LTGAC phtB.seq LTGAC phtB.seq
1138 CGTTATATCCCAGCCA 1159 CGTTATATCCCAGCCA 1062	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: TAGTGATCGG TAGTGATCGGT TAGTGATCGGT	TTTCTCATA 1280 TTTCTCATA TATCTCATA CAATTTTAC GAATTTTAC GAATTTTAC GAATTTTAC	A G C A G C A G C C A G C C A G C C A G C C C C	TTGA pht8.seq 17TGA pht0.seq 17TGA pht0.seq 17TGA pht0.seq 1300 TAAAA phtA.seq 1AAGA pht0.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062	CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: TAGTGATCGG TAGTGATCGGT TAGTGATCGGT	TTTCTCATA 1280 TTTCTCATA TATCTCATA CAATTTTAC GAATTTTAC GAATTTTAC GAATTTTAC	A G C A G C A G C C A G C C A G C C A G C C C C	TTGA phtB.seq 17TGA phtD.SEQ 17TGA Majority 1300 17AAAA phtA.SEQ 17AAGA Majority 1300 17AAAA phtB.seq 17AAGA phtB.seq 17AAGA Majority 1350 17GAC Majority 1350 17GAC phtB.seq 17GAC phtB.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062 TAG C A A A C T G G C C A A G 1260 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1204 A A C T G A T C T T C C T T C 1310 1346 A A G A A A A T G T T G C T C C 1259 A A A C T G A C C T C C C A T C 1274 - A A G C A A T C C T T C T T C T T A C T A G C A A G A A T T C 1360 1196 C T G T T A A C T G A G G C T C	270 CAGGAAAGT: 1270 CAAGAAAGT: CAGGAAAGT: CAGGAAAGT: CAGGAAAGT: TAGTGATCGA TAGTGATCGA TAGTGATCGA TAGTGATCGA TAGTGATCGATAGTGATCGATAGTGATCGATAGTGATCGATAGTGATCGATAGTGATCGATAGTGATCGATAGTGATTGAT	1280 TTTCTCATA 1280 TTTCACACA TATCTCATA CAATTTTAC CAATTTTAC CAATTTTAC ACTTGATA 1380 CTTTGNAAA	A G C A G C A G C C A A G C A A G C A G C A G C A G C A G C A	TTGA phtB.seq LTTGA phtD.SEQ LTTGA Majority 1300 LAAAA phtB.SEQ LAGA Majority LAGA A phtB.seq LAGA C phtB.seq LTGAC Majority LTGAC phtB.SEQ LAGT Majority L100 LATTC PhtA.SEQ LATTC PhtA.SEQ LATTC PhtA.SEQ LATTC PhtA.SEQ
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1162 T A G C A A C T G G C C A A 1260 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G C C C A T C 1310 1346 A A G A A A A T G T T G C T C C 1238 A A A C T G A C C T C C C A T C 1259 A A A C T G A C C T C C C A T C 1274 - A A G C A A T C C T T C T T C T T A C T A G C A A G A A T T C 1288 T T A C T A G C A A G A T T C 1288 T T A C T A G C A A G A A T T C	A G G A T C T T T C A G G A A A G T T 1270 C A G G A A A G T T C A G G A A A G T T C A G G A A A G T T T A G T G A T C G G T A G T G T G G T A G T G T G G T A G T G T G G T A G T G T G T G T G T A G T G T G T G T G T A G T G T G T G T G T A G T G T G T G T G T A G T G T G T G T G T A G T G T G T G T G T A G T G T	1280 TTTCTCATA 1280 TTTCACACA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA CAATTTTAC 1330 GAATTTTAC CAATTTTAC ACTTGATA 1380 GTTTGAAAA ACTTGATAA	A G C A G C A G G C A G C A G C A G C A G C A G C C A G C A	TTGA pht8.seq 17 TGA pht8.seq 17 TGA pht8.SEQ FAAGA Majority 1300 FAAAA pht8.seq FAAGA Majority 1350 TAAGA pht8.seq TAAGA pht8.seq TAAGA pht8.seq TAAGA pht8.seq TAAGA pht8.seq TGAC Majority 1350 TAAT pht8.seq TGAC pht8.seq TGAC pht8.seq TGAC pht8.seq TGAC pht8.seq TGAC pht8.seq
1138 C G T T A T A T C C C A G C C A 1159 C G T T A T A T C C C A G C C A 1062 TAG C A A A C T G G C C A A G 1260 1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1204 A A C T G A T C T T C C T T C 1310 1346 A A G A A A A T G T T G C T C C 1259 A A A C T G A C C T C C C A T C 1274 - A A G C A A T C C T T C T T C T T A C T A G C A A G A A T T C 1360 1196 C T G T T A A C T G A G G C T C	200 A A A G T T T C A G G A A A G T T T C A G G A A A G T T T C A G G A A A G T T T C A G G A A A G T T T C A G G A A A G T T T C A G G A A A G T T T A G T G A T C C A G T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A G T G A T C G J T A C C A A G A T T T A C C A A G A	1280 TTTCTCATA 1280 TTTCACACA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA TATCTCATA CAATTTTAC GAATTTTAC GAATTTTAC GAATTTTAC ACTTGATAA 1280 CTTTGATAA ACTTGATAA	A G C A G C A G G C A G C A G C A G C A G C A G C C A G C A	TTGA phtB.seq TTGA phtB.SEQ TAAGA Majority 1300 TAAAA phtA.SEQ TAAGA Majority 1350 TAAAA phtB.seq TAAGA Majority 1350 TAAAT phtB.SEQ TAAGA Majority 1350 TAAT phtA.SEQ TGAC Majority 1350 TAAT phtB.Seq TGAC TGAC phtB.Seq TGAC TGAC phtB.Seq TAAGT phtB.Seq TAAGT phtB.Seq

12/17

Figure 7(e)

	2	. C	λ	Ţ	T	Τ.	Ţ	3)	-	G	c	T	T :		G	A	т :	<u>, , , , , , , , , , , , , , , , , , , </u>	<u> </u>	ς.		, T	T	G	۸.	٨	Ç	; <u>,</u>	Ç	TC	A	A	3 0	A	τ	C 1	2	τ	c ,		Ha	7051	LEY
									14	10								20							143							14	40							149	_		Ī
144	6. 7	· c	λ	Ŧ	Ŧ	T	c (: 7		G	С	c	T :	۲ ۸	. 6	$\overline{\lambda}$	c 7	\ A	. λ	Ŧ	r	Ī	Ŧ	A C	· 🖈	A 1	c 0	c	7	rc		<u>.</u>	F G		T			Ŧ	e 6	-	• nh	EA.S	
133	8 1	. с	ı A	T	T	т.	T (.,	C		С	T	T	r a	G	A	T,	\ A	С	c ·	r	: 1	T	G 6			c c	: 1	•		٠.				•			-					
135	9 7	G	A	Ŧ	Ŧ	7	r	3 3	. 0	Q	C	Ŧ	T :	rG	G	A	Ţ	\ A	c	c ·	7	ī	Ŧ	GG	A	λ	c	i A	¢ '	rc	λ		3 0	A	Ť	6 1	Ç	c	c a	L A	ph	ED.S	EC
110	7 -	-	-	-	-			•	-	-	•	-		-	-	-		-	-	-	- •	-	-		-			-	-		-	-		-	-		-	-		-	ph	EE.S	EQ
	9	Ŧ	c	1	Ŧ	<u> </u>		١. ٥	7	c	λ		3 7	7	A	G ·		0	λ	T (; ,	. 7	λ	T 7	c	7 :	r ç	c	c	r	c	Ŧ 1	r a	G	c ·	r c	: c	G	2 1	T	Ma	jori	Ev
									4								14	•							48							14	_	_	_				-	150	•	•	,
1496	ī c	Ŧ	λ	λ	T		, ,		7	Ā	λ.		. 7	7	6	G	. ,		λ	Ŧ	; ;	Ŧ	T	7 7	<u>+</u>	Ŧ	1 6	c	1	-	_	_	-	_	-	-	-	_	1 7		٠	CA.S	
1381	G	T	G	x	T	,,	١,	G	T	С	A i		37		×	c ·	2 7	G	λ	A C	: 2	·Ŧ	A.	T T	C	T 1	C	c	C 3	•		Ŧ 1		c				~					
1409	G	Ŧ	G	λ	T.	۸,		G	Ŧ	c	λ	A C	3 1	T	λ	G 1	C	G	λ	Ţ	3 2	Ť	y,	7	C	T	c	c	c a	T	C	7 7	٠,	G	c :	c	c	G	A T	T	ph	tD.S	EC
1101	-	-	-	•	-		• -	-	•	-	•		-	-	-	-	-	-	-		•	•	-		-				T					-			•	-		-		tE.S	
	c	G	т	c	λ,	rc	: c	: A	G	A	. 1	c (Ŧ	a (; ;	A	,				λ.		Ŧ.	a c	: 0				.	T 1		_	T 1			_	T C		w.	jori	
	_								51		_						15:		_						•		_	-	-	<u> </u>	Ė	_	_	Ť	• '	` `	-	-		•	•	jarı	cy
	_	_	_	_	_	_	_	_			_				_							_	_		.531			_		_		15								155			
1546		c	_	ç	۸.		: 0	٠.	9	٨	G (C 0		. =	I	T		c	۸	٠,		c	٨.		T	TC	7	C	λ,	\ Ā	T	T		G	T J	T	y	¢	7 0	, y	ph	tA.S	EC
1459	-	6	T	-		,,			6	۸.		~ 0	, ,	-	Ŧ			,	۸.			c		, ,	Ĩ.	8 0	: G	c	2)		7	T A	c	c	T	٠.	λ	c	T G	; A	ÞÞ	tE.s	eq
1115	-	ŭ	:	_	•			•	-	_	- :			-	•	•		•	•	• •		-	^ /	• •	T !	G C	: 0	-	, ,		T	T A		Ç	T A	, c	λ	С	ŦO				
																								_	_	_	_	_		_	-		_	•	•	•	•	-		-	-	tZ.S	_
	Ŧ	G	Y.	Ţ	G 2			T	Ŧ	<u>c</u> .	~	١.	Ŧ	Α	0	<u> </u>	4	λ	9 '	rī	. G	0	C /	0	9.	C A		0	T A	<u>. c</u>	λ	<u> </u>	. 0	C	<u> </u>	c	A	G	A T	G	Ha	jori	ty
	_							_	56								157							_	580							15								160			
1596	A	Ģ	A I	C (0 1		. 0	1	Ŧ	C	3 7	· A	T	Ŧ	G (7	C	λ	A :	7	۸	0	C 7	. 0	λ:	žλ	۸.	G	T A	T	λ	د ۲	Ā	c	0 7	¢	À	a .	A T	G	ph	tA.S	EQ
1488	Ī	G	۸.	Ŧ (3	. 0		I	I	c :		١ ۵	Ŧ	*	a (2 0	٨	٨	G :	7	G	3	C J	G	9 (CA		G	T A	C	X I	C Y	. a	c.	A C	,	X	G .	A C	0	ph	tB.s	80
1509	T	-	^ :	. ·			_	-	T				· Ŧ	^									C)		6 (E A		a	T A	C													
	_	_					_		_	-		-	_	_	-		_	_		_	_	-		-	•		•	-		_	- '	C A		С.	. 7	c	T	G .	N T	G	рh	te.s	20
	9	Ŧ	Ŧ.		,	1	7	Ŧ	Ŧ	T	, ,	T	c	c	T (: 0	Ŧ	G	λ:	. 1	7	λ	A C	: ¢	λ	3 T	G	A.	TG	A	a e	3 0	G	o.	۱ 1	, C	c	¢.	r a	T	Ha;	jori	τy
								1	61	0						1	62	٥						1	630	ŀ						164	LO			•			1	65	0		
1646	6	T	T		= ,	т	т	T	Ŧ	T	7 7	T	G	Ä	Ä	: 7	7	Ğ	λ 1		Ŧ	7	λ 1	· c	, ,	. T	0	<u>.</u>	ŧ o	•	1		1	a :	. 7		_	١,		÷		~	_
1538	G	Т	T	١.	r ,	T	c	T	Ŧ	T	3 8	T	c	C	T (: 0	T	٥.	A :	. 7	Ŧ	A.	A C	: c	A C	3 T	G	A :	r a	λ	G (GG	a	G i	1	0	c	c ·		Ŧ	phi	. B . e	·
1559	G	T	Ť	. :		T	c	Ŧ	Ŧ	T		T	¢	¢	T	: 0	T	G.	λ 1	۱ ۸	7	A.	A C	c		3 T	G	A S	r G	A	G	3 0	G	G	N 7	G	ċ	ċ	r a	Ŧ	phi	D. 57	EO
1127	Ģ	T	T	١.	r J	T	Ŧ	Ŧ	T	T	٠,	T	C	C		-	-	-		• •	•	-		•		-	-	-		-			-			-	-			•		E.S	
	G	Ŧ.	λ,		: 1	c	c	,	c.	λ 1	· A	. 7	G	A	c (: c	λ	т.	A · c	; c	c	A I	c 7		.,		Ŧ							T I	L a	7	•	T (· •	_	wa-	jori	
						٠			561								.67								580							169				_	_			70		,	-,
1696	~	-		_	- 7	_	_	=	_		_	Ţ	-	~	-		•		-	_	_	-	_	_			_			_				_		_	_	_					
1696 1588	Ğ	T				č	ċ	ì	c .				a	1			:	T :			-		-			I	Ξ		3 A				٠	T !		c	c :	T :	Ī	c			
1609	G	1	a i		. 1	c	č	λ	c .	λ 7		T	Ğ	A (2 0	: -	ĩ	T			č	2				T	T		•	•	^ ′		٠.									:B. sc :D. 51	
1143	-	-				-	-	-	_		-	-	-				-	_ :	-																							. E. SI	
																																									-		•
	Ξ.	3	<u>, , , , , , , , , , , , , , , , , , , </u>		; C	T	G	λ	9 2	A C	Α	0	c	0 (3 0	<u> </u>	٩	e e	: c	Α	G	<u>a</u> (: 7	7	<u> </u>	0	c	T J	\ <u>A</u>	λ	0 ,	١ ٥	λ	٨.	G	¢	T '	7	G	4	Ha j	ori	t
				_	_	_		17	1)						1	720	0						17	30						:	17	0						1	750)		
1746	τ	c .	۸ 7								Α	G	Ŧ	Ŧ	3 C	λ	G	c	C	λ	λ	G (: c	T.	À I	À	c	7 ,	\ A	7	ر د	<u>, , , , , , , , , , , , , , , , , , , </u>	λ.	, ,	G	0	T	7	c	÷	pht	.A. 51	20
1618						Ŧ	Œ	1	2 :			n	~		3 0		G	c	: c	λ	a	c (-				~ :						.,	à	G	+ -				-	. ac	7
	Ŧ	5 2	•			Ξ	Ξ	-			•	•	-				_							•		_	-									_	•			^	pac		
1659 1153	Ŧ	G 2	,,		C	T	G	y (3 ,	A G	λ	G	C	6 6	3 0	λ	G	c	: 0	λ	G	G	7	T.	A T	6	c ·	T	ì			G	λ	λ,	a	ā	T		G	À		D.SI	ΕQ

13/17

Figure 7(f)

CCCCTCCTTCGACAG	1 CATEAGGAT	TCAGGAAAA	ACTGAGGCAA	AAGGA Hajority
1760	1770	1780	1790	1800
1796 TACCTCCATCTCCAG	ACGCAGATGTT	AAAGCAAAT	CCAACTGGAG	ATAGT phta.SEO
1688 CCCCTCCTTCGACAG	ACCATCAGGAT	TCAGGAAAT	ACTGAGGCAA	AAGGA phtB.seg
1709 CCCCTCCTTCGACAG	ACCATCAGGAT	TCAGGAAAT	ACTGAGGCAA	AAGGA phtD.SEQ
1167				phtE.SEQ
GCAGAAGCTATCTAC	AACCGXGTGAA	AGCAGCTAA	GARGGTGCCA	C'T T G A Hajority
1810	1820	1830	1840	1850
1846 GCAGCAGCTATTAC	AATCGTGTGAA	AGGGGAAAA	ACGAATTCCA	CTCGT phta.SEQ
1738 GCAGAAGCTATCTAC	AACCGHGTGAA	AGCAGCTAA	GAAGGTGCCA	CTTGA phtB.seq
1759 GCAGAAGCTATCTAC		A G C A G C T A A	GAAGGTGCCA	
1167 T A C	AGCT			phtE.SEQ
TCGTATGCCTTACAA	TCTTCAATATA	CTGTAGAAG	<u> </u>	TAGTT Hajority
10,60	1870	1880	1890	1900
1896 TEGACTTECATATAT	GGTTGAGCATA	CAGTTGAGG	TTAAAAACGG	TAATT phtA.SEO
1788 TCGTATGCCTTACAA				
1809 T C G T A T G C C T T A C A A	TCTTCAATATA	CTOTAGAAG	TCX4444CGG	TAGTT phtD.SEQ
1174		TTGTAAGA-		phtE.SEQ
TAATCATACCTCATT	ATGATCATTAC	CATAACATT	AAATTTGAGT	G G T T T Majority
1910	1920	1930	1940	1950
1946 TGATTATTCCTCATA				
1838 TAATCATACCTCATT				
1859 TAATCATACCTCATT	ATGACCATTAC			
1104				
1186 C A T G	GTGATCATTTC			pheE.SEC
GACGAAGGCCTTTAT	GTGATCATTTC	CATTACATT		phtE.SEC
	GTGATCATTTC	CATTACATT		phtE.SEC
GACGAAGGCCTTTAT 1960 1996 GATGATCACACATAC	GTGATCATTTC GAGGCACCTAA 1970 AAAGCTCCAAA	CATTACATT GGGGTATAC 1980 TOGCTATAC	TCTTGAGGAT 1990 CTTGGAAGAT	C T T T T Majority 2000 T T G T T phth.SEO
950 GATGATCATATATATATATATATATATATATATATATATA	G T G A T C A T T T C G A G G C A C C T A A 1970 A A A G C T C C A A A G A G G C A C C T A A	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	TCTTGAGGAT 1990 CTTGGAAGAT	C T T T T Majority 2000 TT G T T phtA.SEO
GACGAAGGCCTTTAT 1960 1996 GATGATCACACATAC	GTGATCATTTC GAGGCACCTAA 1970 AAAGCTCCAAA GAGGCACCTAA GAGGCACCTAA	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T	C T T T T Majority 2000 T T G T T phtA.SEQ C T T T T phtB.seq C T T T T phtD.SEQ
950 GATGATCATATATATATATATATATATATATATATATATA	GTGATCATTTC GAGGCACCTAA 1970 AAAGCTCCAAA GAGGCACCTAA GAGGCACCTAA	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T	C T T T T Majority 2000 TT G T T phtA.SEO
950 GATGATCATATATATATATATATATATATATATATATATA	0 T G A T C A T T T C 0 A G G C A C C T A A 1970 A A A G C T C C A A A G A G G C A C C T A A G A G G C A C C T A A C A C C C A A A	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	TCTTGAGGAT 1990 CTTGGAAGAT TCTTGAGGAT	2000 TTGTT phta.szc CTTTT phta.szc CTTTT phta.szc CTTTT phtb.seq CTTTT phtb.szc
1950 1996 GATGATCACATAC 1888 GACGAAGGCCTTTAT 1909 GACGAAGGCCTTTAT 1909 GACGAAGGCCTTTAT 1210	0 T G A T C A T T T C 0 A G G C A C C T A A 1970 A A A G C T C C A A A G A G G C A C C T A A G A G G C A C C T A A C A C C C A A A	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	TCTTGAGGAT 1990 CTTGGAAGAT TCTTGAGGAT	2000 TTGTT phta.szc CTTTT phta.szc CTTTT phta.szc CTTTT phtb.seq CTTTT phtb.szc
GACGAAGGCCTTTAT 1950 1996 GATGATCACATAC 1888 GACGAAGGCCTTTAT 1909 GACGAAGGCCTTTAT 1210 GGCGACTGTCAAGTA 2010	0 T G A T C A T T T C 0 A G C C C C T A A 1970 A A G C T C C A A A G A G G C A C C T A A G A G G C A C C T A A C T A T G T C G A A C 2020	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG	TCTTGAGGAT 1990 CTTGGAAGAT TCTTGAGGAT TCTTGAGGAT AACGTCCGCA 2040	2000 TTGTT PhtA.SEC CTTTT PhtA.SEC CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SEC TTCAC Hajority
1950 1996 GATGATCACATAC 1888 GACGAAGGCCTTTAT 1909 GACGAAGGCCTTTAT 1909 GACGAAGGCCTTTAT 1210	0 T G A T C A T T T C 0 A G O C A C C T A A 1970 A A A G C T C C A A A A A G C A C C T A A C C A A A C T A T G T C G A A C 2020 C T A C G T A G A A C	CATTACATT GGGGTATAC 1940 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG	T C T T G A G O A T 1990 C T T G G A A G A T T C T T O A G G A T T C T T O A G G A T T C T T O A C G C A 2040 A A C G T C C A C A	2000 TTGTT PhtA.SEO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SEO TTCAG Majority 2050 TTCAG PhtA.SEO
O A C G A A G G C C T T T A T	0 T G A T C A T T T C 0 A G C C C C C A A A 1970 A A G C T C C A A A G A G C A C C T A A G A G C A C C T A A C T A T G T C G A A C 2020 C T A C G T A G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC 2010 ACCCAGACG ACCCAAACG	1990 CTTGGAAGAT TCTTGAAGAT TCTTGAAGAT TCTTGAAGAT TCTTGAAGAT AACGTCCGCA 2040 AACGTCCACA AACGTCCACA	2000 TTGTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGA O PhtB.SZO
2046 TGCGACGATTAAGTA 2938 GCGACGATTAA 2010	0 T G A T C A T T T C 0 A G C C C C C A A A 1970 A A G C T C C A A A G A G C A C C T A A G A G C A C C T A A C T A T G T C G A A C 2020 C T A C G T A G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC 2010 ACCCAGACG ACCCAAACG	1990 CTTGGAAGAT TCTTGAGGAT TCTTGAGGAT TCTTGAGGAT TCTTGAGGAT AACGTCCGCA 2040 AACGTCCGCA	2000 TTGTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGA O PhtB.SZO
O A C G A A G G C C T T T A T	G T G A T C A T T T C	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC 2010 ATCCAGACG ATCCAAACG ATCCAAACG ATCCAAACG	T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A	2000 TTGTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGA G Majority 2050 TTGA G PhtB.SZO TTGA G PhtB.SZO TTGA G PhtB.SZO
O A C G A A G G C C T T T A T	G T G A T C A T T T C	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC 2010 ATCCAGACG ATCCAAACG ATCCAAACG ATCCAAACG	T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A	2000 TTGTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGT PhtB.SZO TTGA G Majority 2050 TTGA G PhtB.SZO TTGA G PhtB.SZO TTGA G PhtB.SZO
O A C G A A G G C C T T T A T	G T G A T C A T T T C G A G G C A C C T A A 1970 AAA G C T C C A AA G A G G C A C C T A A A G G C A C C T A A C T A T G T C G A A C 2020 C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C C T A T G T C G A A C A C G C T A G C G A C 2070	CATTACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC 2010 ATCCAGACG	T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T T C T T G A G G A T 2040 A A C G T C C G C A A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A	2000 TTGTT PhtA.SZO CTTTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SZO TTCAO Majority 2050 TTCAO Majority TTCAO PhtB.seq TTCAO PhtB.seq TTCAO PhtB.seq TTCAO PhtB.SZO TCCAA PhtB.SZO
2010 2046 2056 2056 2056 2056 2056 2056 2056 205	G T G A T C A T T T C	C A T T A C A T T G G G G T A T A C 1940 T G G C T A T A C G G G G T A T A C A T C C A G A C G 2030 A C C C T G A C G A T C C A A A C G A T C C A A A C G A T C C A A A C G A T C C A A A C G A T C C A A C G A T C C A T G T T T T T X 2080 C A T G T T G T T T	T C T T G A G O A T 1990 C T T G G A A G A T T C T T O A G G A T T C T T O A G G A T T C T T O A G G A T A A C G T C C G C A G A G A	2000 TTGTT PhtA.SEQ CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.SEQ TTCAG Hajority 2050 TTCAG Hajority 2050 TTCAG PhtB.seq TTCAG PhtB.seq TTCAG PhtB.seq TTCAG PhtB.seq TCAA PhtB.seq TCAA PhtB.seq TCAA PhtB.seq
O A C G A A G G C C T T T A T	T G A T C A T T T C	C A T T A C A T T G G G G T A T A C 1940 T G G C T A T A C G G G G T A T A C A T C C A G A C G A T C C A A A C G A T C C A A A C G A T C C A A A C G A T C C A A A C G A T C C A A A C G A T C C A T T T T X 2080 C A T G T T T T C A A C G T T T C A A C T T T C A A C C C C C C C	T C T T G A G Q A T 1990 C T T G G A A G A T T C T T O A G G A T T C T T O A G G A T A A C O T C C O C A A C O T C C C C C A A C O T C C C C C A A C O T C C C C C C A A C O T C C C C C C C C C C C C C C C C C	2000 TTGTT PhtA.SZO CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.seq CTTTT PhtB.seq TTCAQ Majority 2050 TTCAQ PhtA.SZO TTCAQ PhtA.SZO TTCAQ PhtA.SZO TTCAA PhtA.SZO TTCAA PhtA.SZO TTCAA PhtA.SZO AAGAT Majority 2100 AAGAT Majority

14/17

Figure 7(g)

C-A A G C C A G T A A A C C T	***********	UNNACATUAL	CAAGTAAG - G	A G Majority
2110	2120	2130	2140	2150
2143 CA C	AGTGAAGAT	c	CAAATAAG	phtA.SEC
2038 CAAGCTGATA CC	AATCAA-ACGG	AAAAACI	CAAGCGAG-G	
2059 CAAGACAGTAAACCT		GGAACATGATI		A G C C phtD.SEQ
1284 CAATCCAG - GAACTT	C A C A T G A	G A A A C A T G A	• • • • • • • • •	phtE.SEQ
A - C T C A C - G A A				
2160	•			CCT- Majority
	2170	2180	2190	2200
2164		A A C	TTCAAA	
2079 ACCTCAGACAGAAAA. 2109 AACTCACCCTGAATC	RCCTGAGGAAG	X - X A C C		pes. Edd CTC pht8.seq
1314			COTTTALLT	
			OXTITUX-T	CT - phtE.SEQ
-AGCAGATAAACEGT:	TARGCCAG	A C	A - A C A	A Hajority
2210	2220	2230	2240	2250
2173 G C G G A T G A A	G A G C C A G			phtA.SED
2114 GAGAAGAGAAACCGC	- AAGCGAGAA.	CCAGAGTCTC		CAGA phtB.seq
2159 CAGCAGATAATCTTT		CTGATACGGA	AGAGACAGAC	GAA phed.SEQ
1339AATCOT	.TTATC			phtE.SEQ
G-AGCTGGAGGAAXC	CCAGATGAGT	AGAAGTXCCT	CAAGTAGAG	CTG Majority
2250	2270	2280	2290	2300
2189 TAGAGGAAAC	CCTGCTGAGC	AGAAGTCCCT	CAAGTAGAGA	CTC phra SPO
Z163 GGAACCAGAAGAATCA	CCAGAGGAAT	LOSSESSES	CAGGGGGAGA	C T C oben con
2209 GAAGCTGAAGATACCA	. C - A G A T G A G G	TGAAATTCCT	CAAGTAGAG	ATT phtD.SEQ
1351 G C T G A A G A	TGAAT	:		phte.seq
AAAAXGTTGAAGCXAA				
	XXXXXXXXXXX	CXGAGGTTTT		G T C Matoriry
•	ACTXAXAGAXO		-,	
2310	2320	2330	2340	2350
2310 2234 AAAAGTAGAAGCCCA	2320	2330	2340 GCTTGCGAAA	2350 G T A phtA.SEQ
2310 2234 AAAAAGTAGAAGCCCA 2233 AAAAGTTTGAAGAAA	2320 A C T C A A A G A A C A C T G A G A G A G	2330 CAGAAGTTTT	2340 GCTTGCGAAA	2350
2310 2234 AAAAGTAGAAGCCCA	2320 A C T C A A A G A A C A C T G A G A G A G	23JD CAGAAGTTTT CTGAAGATTT CGGAGGCCTT	2340 GCTTGCGAAA ACTTGGAAAA GCTAGAAAAA	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ
2214 AAAAAGTAGAAGCCCA 2213 AAAAAGTTGAAGAAA 2256 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACACTGAGAGAGAGAGAGAGAGAGAGAGA	Z3JD CAGAAGTTTT CTGAAGATTT CGGAGGCCTT	2340 G C T T G C G A A A A C T T G G A A A G C T A G A A A A A	2350 A G T A phtA.SEQ A T C phtB.seq A T C phtB.seq - T C phtE.SEQ
2310 2234 AAAAAGTAGAAGCCCA 2231 AAAAAGTTTGAAGAAAA 2258 CTGTTATTAACGCTAA	2320 ACTCAAAGAACACACTGAGAGAGAGAGAGAGAGAGAGAGA	Z3JD CAGAAGTTTT CTGAAGATTT CGGAGGCCTT	2340 G C T T G C G A A A A C T T G G A A A G C T A G A A A A A	2350 A G T A phtA.SEQ A T C phtB.seq A T C phtB.seq - T C phtE.SEQ
2214 AAAAAGTAGAAGCCCA 2213 AAAAAGTTGAAGAAA 2256 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACACTGAGAGAGAGAGAGAGAGAGAGAGA	Z3JD CAGAAGTTTT CTGAAGATTT CGGAGGCCTT	2340 G C T T G C G A A A A C T T G G A A A G C T A G A A A A A	2350 A G T A phtA.SEQ A T C phtB.seq A T C phtB.seq - T C phtE.SEQ
2214 AAAAAGTAGAAGCCCA 2213 AAAAAGTTAAAAAA 2215 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2330 CAGAAGTTTT CTGAAGATTTT CGGAGGCCTTAGGTTTT ACGGAGACTCC	2340 GCTTGCGAAA ACTTGGAAAA GCTAGAAAAA TXACTGGTTT	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A A Majority 2600
2214 AAAAAGTAGAAGCCA 2211 AAAAAGTAGAAGAAA 2213 AAAAAGTTGAAGAAAA 2258 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACACTGAGAACACACAATGCAAAATGCAAAAAAAA	Z330 CAGAAGTTTT CTGAAGATTTT CGGAGGCCTTAGGTTTT ACGGAGACACTC Z380 ACAGAAACTC	2340 G C T T G C G A A A A C T T G G A A A A A G C T A G A T T T T T T T T T T T T T T T T	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A Hajority 2400 A C G phtA.SEQ
2214 AAAAAGTAGAAGCCA 2213 AAAAAGTAGAAGAAA 2258 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACACTGAGAACACACAATGCAAAATGCAAAAAAAA	Z330 CAGAAGTTTT CTGAAGATTTT CGGAGGCCTTAGGTTTT ACGGAGACACTC Z380 ACAGAAACTC	2340 G C T T G C G A A A A C T T G G A A A A A G C T A G A T T T T T T T T T T T T T T T T	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A Hajority 2400 A C G phtA.SEQ
2214 AAAAAGTAGAAGCCA 2211 AAAAAGTAGAAGAAA 2213 AAAAAGTTGAAGAAAA 2258 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACACTGAGAACACAATGCAAATGCAAATGCAAATGCAAATGCAAATGCAAATGCAACACAAAAAATGCT	Z330 CAGAAGTTTT CTGAAGATTTT CGGAGGCCTTAGGTTTT ACGGAGACACTC Z380 ACAGAAACTC	2340 G C T T G C G A A A A C T T G G A A A A A G C T A G A T T T T T T T T T T T T T T T T	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A Hajority 2400 A C G phtA.SEQ
2214 AAAAAGTAGAAGCCA 2211 AAAAGTAGAAGCCA 2211 AAAAGTTAAAAA 2156 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACCAATAGCAAAACCAATGCAAATGCAAATGCAAATGCAAATGCAAATGCAAAAAAAA	2330 CAGAAGTTTT CTGAAGATTTC CGGAGGCCTTAGGTTTT ACGGAGACTC 2380 ACAGAAACTC AAAGAAACTC ATGGAGACACT ACGGAGACAC	2340 G C T T G C G A A A A C T T G G A A A A A G C T G G T T T 2390 T A G C T G G T T T T C A C A G G A T T T C A C A G G A T C T	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A A Bajority 2400 A C G phtA.SEQ A A A phtB.seq A A A phtB.seq A A A phtB.seq A A A phtB.SEQ phtE.SEQ
2214 AAAAAGTAGAAGCCA 2213 AAAAAGTAGAAGAAA 2213 AAAAGGTTAAAAGAAAA 2256 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAA ACTGAGAGAGAGATAGCAGATAGCAAATGCA AAGCCAATGCAAGGCAAATGCAAGTCCAATGCAAGGCAATGCAAGGCAAATGCAAGGCAAAAATGCAAGGAAAAATGCT	Z330 CAGAAGTTTT CTGAAGATTTC CGGAGGCCTTAGGTTTT ACGGAGACTC Z380 ACAGAAACTC AAAGAAACTC ATGGAGACCT TAATAATACT	2340 G C T T G C G A A A A C T T G G A A A A A G C T G G T T T T C G C A G T T T T C A C A C G T C T T T C A C A C G T C T T T C A C A C G T C T T T C A C A C G T C T T T C A C A C G T C T T T C A C C A C G T C T C A C A C A C A C A C A C A C A C	2350 G T A phtA.SEQ A T C phtB.seq C T A phtD.SEQ - T C phtE.SEQ A A A Majority 2400 A C C phtA.SEQ A A A phtB.seq
2214 AAAAAGTAGAAGCCA 2211 AAAAGTAGAAGCCA 2211 AAAAGTTGAAGAAA 2258 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAACACCAATAGCAAAAAAAAAAAAAAAAAAA	Z330 CAGAAGTTTT CTGAAGATTT CGGAGGCCTTAGGTTTT ACGGAGACACTC Z380 ACAGAAACTC AAAGAAACTC ATGGAGACAT ACGGAGACAT CTGAAGACTC ATGGAGACAT ACGGAGACAT	2340 G C T T G C G A A A A C T T G G A A A A A G C T G G T T T C G C A G G A T T T T C A C A G G A T T T T G A C T G G T C T T T G A C T G G T C T T T G A C T G G T C T T T G A C T G G T C T C A C A G A T T T G A C T G G T C T C A C A G A T T T G A C T G G T C T C A C A G A T T T G A C T G G T C T C A C A G A T T T T G A C T G G T C T C T C T C T C T C T C T C	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A A Bajority 2400 A C G phtA.SEQ A A A phtB.seq
2210 2214 AAAAAGTAGAAGCCCA 2211 AAAAGGTTGAAGACCCA 2213 AAAGGTTGAAGAAA 1165 ACGGATCCTAGTATXA 2360 2284 ACGGATCCTAGTATXA 2360 2284 ACGGATCCTAGTATATCA 2361 2310 AAATAATTTXCTTCTT 2410 2314 AAATAATTTGACTCTT	2320 ACTCAAAGAACACTGAGAGACACTGAGACAGATGCAATGCAATGCAATGCAAAGCCAATGCAACACAAAGAACAAAGAACAAATGCA	Z330 CAGAAGTTTT CTGAAGATTTC CGGAGGCCTTAGGTTTT ACGGAGACACTC Z380 ACAGAAACTC ATGGAGACACT ATGGAGACACT TAATAATACT Z430 TAACAATAGT	2340 G C T T G C G A A A A C T T G G A A A A A G C T A G A A A A A A C T G C T T T T C A C A G G A T T T T G A C T G G T C T T T G A C T G G T C T T T G A C T G G T C T T T G A C T G G T C T C A C A G A T C T C A C A G A T C T C A C A G A T C T C A C A C A C A C A C A C A C A	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A A Bajority 2400 A C G phtA.SEQ A A A phtB.seq A A A phtB.seq A A A phtB.seQ A A G Majority 2450 A A G phtA.SEQ A A G phtA.SEQ
2214 AAAAAGTAGAAGCCA 2213 AAAAGTTGAAGAAGCCA 2213 AAAAGGTTGAAGAAAA 1365	2320 ACTCAAAGAA ACTGAGAGAG ACTGAGAGAGAG AAXCCAATGCX 2370 AAGCCAATGCX AGTCCAATGCX GACAAAATGCT GAAACXAAGGA 2420 CAAATTATGGA	2330 CAGAAGTTT CTGAAGATTT CGGAGGCCTTAGGTTTT ACGGAGACCTC 2380 ACAGAAACTC AAAGAAACTC AAAGAAACT 2430 TAATAATAGT CAACAATAGT	2340 G C T T G C G A A A A C T T G G A A A A A G C T G G T T T 2190 T A G C T G G T T T C A C A G G A T T T G A C A G G A T T A T T T T G G C A G 2440 A T C A T G G C A G A T T A T G G C A G A T T A T G G C A G	2350 G T A phtA.SEQ A T C phtB.seq G T A phtD.SEQ - T C phtE.SEQ A A A Majority 2400 A C G phtA.SEQ A A A phtB.seq A A A phtB.seq A A A phtB.SEQ A A G phtA.SEQ A A G phtB.seq
2214 AAAAAGTAGAAGCCA 2213 AAAAAGTAGAAGCCA 2213 AAAAGGTTGAAGAAAA 2258 CTGTTATTAACGCTAA 1165	2320 ACTCAAAGAA ACTGAGAGAG ACTGAGAGAGAG AAXCCAATGCX 2370 AAGCCAATGCX AGTCCAATGCX GACAAAATGCT GAAACXAAGGA 2420 CAAATTATGGA	Z330 CAGAAGTTT CTGAAGATTT CTGAAGATTT CGGAGGCCTTAGGTTTT ACGGAGACTC Z380 ACAGAAACTC ATGGAGACTC ATGGAGACTT ATGGAGACAT ACGGAGACTT Z430 TAATAATACT TAATAACACT	2340 G C T T G C G A A A A C T T G G A A A A A G C T G G T T T T G G C A G A C T T T T G G C A G A C T T T T G G C A G A T T T T G G C A G A T T A T C A C A G A T T T T C A C A G A T T T T C A C A G A T T T T C A C C A G A T T T T C A C C A G A T T T T C A C C A G A T T T T C A C C A G A T T T T C A C C A G A T T T T C A C C A G A T T C A C C A G A T T T C A C C A G A T T C A C C A G A T T C A C C A C A C C A C C A C C C C	2350 G T A phtA.SEQ A T C phtB.seq C T A phtD.SEQ - T C phtE.SEQ A A A Bajority 2400 A C G phtA.SEQ A A A phtB.seq A A A phtB.seq A A A phtB.seq A A G Majority 2450 A A G phtA.SEQ A A G majority 2450 A A G phtA.SEQ A A C phtA.SEQ A A A D phtB.SEQ A A A D phtB.SEQ

15/17

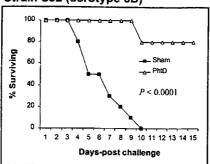
Figure 7(h)

	c			2	À.	A	λ	G	A	c	7		ī	7	G	G	c	Ŧ	T	7	G	T	Ţ	λ	λ	A	G	Ģ	A (Ģ.	۸.	G	T i	Α.	. :	7		Δ	A	G	c ·	r	_			_	С	Ŧ	τ	Majority
	_			_						2	4	50						_		2	47	0								24	80	•							2	49	,							2	50	0
2384	_	-	-		À		Ÿ		,	Ŧ	7	λ.	. 0	I	Ŧ	0	c	C	Ŧ	Ŧ	0	=					Ā					G .	Ŧ		. 1	c	-	-	:	Ξ	-			_	Ξ	=	c	Ŧ	Ŧ	phtA.SEC
2408	T				ı.	T	À	Ĝ	ī	c	Ī	Ĉ	i	ī	G	a	c	Ī	Ť	Ť	â	Ť	÷	À	ì	ì	Ä	G.	λ.			5	7			-	-	-	-	-				, ,		-	-	-	T	phtB.seq
1409	-	À		3 .	A	A	G	C	λ	C	T	-	-	T	G	A	C	•	-	-	-	•	-	-	-	À	G	λ.	A (3 2		3 (c ı		l J	T	Ŧ	λ	λ	Ģ	3 1	•			٠.,	-	٠	-	-	phtE.SEQ
	۸	٨				G	<u>c</u>	G	-	-	Ţ	C	Ŧ	٥	a	c	-	c	5	T	Ļ	-	G	•	¢	λ	λ	-	A J	<u>.</u>	. ,	١.	• •	٠.		Ţ	λ	Ŧ	c	c :	K 2	,	,	L A	. c	c	Ŧ	×	λ	Majority
										2	5,1	lo								2	52	0									30	,							2	540	1							2	55	0
2423	č	λ				=	-	-	-	-	I	2	7	0 0	Ξ	Ξ	-	=	-	7	<u>,</u>	<u>.</u>	0	-	-	:	-									Ŧ	Ä	<u>.</u>	G	Ģ,	. ,	. ,			7	Ĩ				phtA.SEQ
2441	î	Â			•	-	-	-	•	-	:	ċ	ċ	G	ä	ċ	î	c	ċ	Ť	À	-	-	•	-	-	-	- :	- :							T	•	T	G	G 1										phtB.seq phtD.SEQ
1440	-	-	•			G	C	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	С	À	λ.	A .	A A	١.	: ,	. :			-	-	-	T	A	G					_	-	•	•	•	phtz.SEQ
	A	^				<u>e</u>	T	A	x	X																																								Majority
2445			-		-	_	-	λ	A		•																																							phea. Seq
2463 2468					•	C '	•	••	T	•																																								phtB.seq
1455	•	•	•	•		-	_	-	٠																																									phtD.SEQ

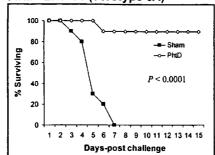
16/17

Figure 8

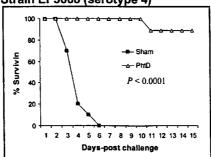
A. Strain SJ2 (serotype 6B)



B. Strain EF6796 (serotype 6A)



C. Strain EF5668 (serotype 4)



17/17