(12) STANDARD PATENT (11) Application No. **AU 2013209513 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Fused antigen vaccines and compositions against Streptococcus pneumoniae

(51) International Patent Classification(s)
A61K 39/09 (2006.01)

(21) Application No: **2013209513** (22) Date of Filing: **2013.01.18**

(87) WIPO No: **WO13/109995**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/589,267	**2012.01.20**	**US**

(43) Publication Date: **2013.07.25**
(44) Accepted Journal Date: **2017.10.05**

(71) Applicant(s)
Genocea Biosciences, Inc.;Children's Medical Center Corporation

(72) Inventor(s)
Flechtner, Jessica;Malley, Richard

(74) Agent / Attorney
FB Rice Pty Ltd, L 23 44 Market St, Sydney, NSW, 2000, AU

(56) Related Art
WO 2011/008548 A1

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

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 July 2013 (25.07.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/109995 A9

(51) International Patent Classification:
A61K 39/09 (2006.01)

(21) International Application Number:
PCT/US2013/022309

(22) International Filing Date:
18 January 2013 (18.01.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/589,267 20 January 2012 (20.01.2012) US

(71) Applicants: **GENOCEA BIOSCIENCES, INC.** [US/US]; Cambridge Discovery Park, 100 Acorn Park Drive, 5th Floor, Cambridge, MA 02140 (US). **CHILDREN'S MEDICAL CENTER CORPORATION** [US/US]; 55 Shattuck Street, Boston, MA 02115 (US).

(72) Inventors: **FLECHTNER, Jessica**; 30 Old Mill Road, Maynard, MA 01754 (US). **MALLEY, Richard**; 15 Bayview Avenue, Beverly, MA 01915 (US).

(74) Agents: **MEDINA, Rolando** et al.; Choate, Hall & Stewart LLP, Two International Place, Boston, MA 02110 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(48) Date of publication of this corrected version:
12 September 2014

(15) Information about Correction:
see Notice of 12 September 2014

(54) Title: FUSED ANTIGEN VACCINES AND COMPOSITIONS AGAINST STREPTOCOCCUS PNEUMONIAE

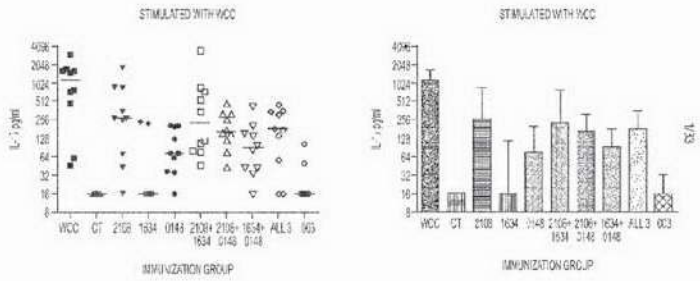


FIG. 1

(57) Abstract: Streptococcus pneumoniae is a major health concern, especially in very young, elderly, or immunocompromised patients. The present disclosure provides, inter alia, certain highly effective vaccines and pharmaceutical compositions in Streptococcus pneumoniae that contain fusion proteins. The antigens may be used therapeutically or prophylactically.

WO 2013/109995 A9

FUSED ANTIGEN VACCINES AND COMPOSITIONS AGAINST STREPTOCOCCUS PNEUMONIAE

Cross Reference to Related Applications

[0001] The present application claims the benefit of U.S. provisional application Serial No. 61/589,267, filed on January 20, 2012, the contents of which are herein incorporated by reference in their entirety.

Government Support

[0002] This work was made with Government support under Grant AI066013 awarded by the National Institutes of Health. Therefore, the U.S. Government has certain rights in this invention.

I. Background

[0003] Pneumococcal disease continues to be a leading cause of sickness and death in the United States and throughout the world. Each year, millions of cases of pneumonia, meningitis, bacteremia, and otitis media are attributed to infection with the pathogen *Streptococcus pneumoniae*. *S. pneumoniae* is a Gram-positive encapsulated coccus that colonizes the nasopharynx in about 5-10% of healthy adults and 20-40% of healthy children. Normal colonization becomes pathogenic when *S. pneumoniae* is carried into the Eustachian tubes, nasal sinuses, lungs, bloodstream, meninges, joint spaces, bones or peritoneal cavity. *S. pneumoniae* has several virulence factors that enable the organism to evade the immune system. Examples include a polysaccharide capsule that prevents phagocytosis by host immune cells, proteases that inhibit complement-mediated opsonization, and proteins that cause lysis of host cells. In the polysaccharide capsule, the presence of complex polysaccharides forms the basis for classifying pneumococci into different serotypes. To date, over 93 serotypes of *S. pneumoniae* have been identified.

[0004] Various pharmaceutical compositions have been used to harness an immune response against infection by *S. pneumoniae*. A polyvalent pneumococcal vaccine, PPV-23, was developed for preventing pneumonia and other invasive diseases due to *S. pneumoniae* in the adult and aging populations. The vaccine contains capsular polysaccharides (CPs) from 23 serotypes of *S. pneumoniae*. As T independent antigens, these CPs induce only short-lived antibody responses, necessitating repeated doses, which increases the risk of immunological tolerance. The antibodies raised against the *S. pneumoniae* capsular

polysaccharides, termed anticapsular antibodies, are recognized as protective in adult and immunocompetent individuals. However, children under 2 years of age and immunocompromised individuals, including the elderly, do not respond well to T independent antigens and, therefore, are not afforded optimal protection by PPV-23. Another *S. pneumoniae* vaccine, Prevnar, includes bacterial polysaccharides from 7 *S. pneumoniae* serovars conjugated to the diphtheria toxoid protein. This vaccine induces both B and T cell responses. However, because it only protects against 7 pneumococcal serotypes, serotype replacement can render Prevnar ineffective against non-vaccine serotypes. Serotype replacement has already been demonstrated in several clinical trials and epidemiologic studies, necessitating development of different formulations of these vaccines. An example is the recently introduced Prevnar 13, directed to 13 pneumococcal serotypes. Furthermore, the two Prevnar formulations are expensive to manufacture, greatly limiting their availability in the developing world. PPV-23, which consists of 23 purified but unconjugated polysaccharides, has broader coverage, but does not provide protection to children under the age of 2 years, a population which is at the highest risk for pneumococcal disease.

[0005] Thus, there remains a need to design more effective pharmaceutical compositions than the current strategies offer. In particular, such compositions need to incorporate novel or specific antigens that elicit an immune response against *S. pneumoniae*.

II. Summary

[0006] *Streptococcus pneumoniae* is a major health concern, especially in very young, elderly, or immunocompromised patients. While DNA and protein sequence information for *S. pneumoniae* have been known for some time, and researchers have long attempted to produce vaccines against *S. pneumoniae*, a major problem was how to elicit an immune response that is both long-lived and effective in all age groups. Certain of the *S. pneumoniae* antigens used in the instant application were initially identified by screening immune cells from mice either immunized with unencapsulated killed whole *S. pneumoniae* or infected with live *S. pneumoniae,* or from healthy human donors for T cell specific immune responses. Accordingly, the present disclosure provides, *inter alia*, certain highly effective vaccines against *Streptococcus pneumoniae*. The vaccines may be used therapeutically or prophylactically. The present disclosure also provides specific antigens and methods for using the antigens to elicit an immune response against *S. pneumoniae*.

[0007] In some embodiments, two, three, four, or more polypeptides from Table 1 or immunogenic fragments or variants thereof are covalently bound to each other, e.g. as a fusion protein.

[0008] In some embodiments, the vaccine formulation comprises or consists of one or more fusion proteins comprising (or consisting of) one or all or fragments of SP0148, SP2108 and SP1912 or variants thereof, such as those shown in Table 2.

[0009] In certain aspects, the present disclosure provides a vaccine formulation comprising or consisting of a pharmaceutically acceptable carrier and a fusion protein comprising or consisting of two or more immunogenic polypeptides, wherein the immunogenic polypeptides have an amino acid sequence at least 70, 80 or 90% homologous to SEQ ID NO: 6, SEQ ID NO: 2 or SEQ ID NO: 9, or an immunogenic fragment thereof.

[0010] In certain embodiments, the present disclosure provides fusion proteins comprising or consisting of a first immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 6, or an immunogenic fragment thereof and a second immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 9, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 6, or an immunogenic fragment thereof, and the second immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 9, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide is located at the amino terminus of the fusion protein and the second immunogenic polypeptide is located at the carboxyl terminus of the fusion protein.

[0011] In some embodiments, the fusion protein further comprises or consists of a third immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 2, or an immunogenic fragment thereof. In some embodiments, the third immunogenic polypeptide consists of or comprises SEQ ID NO: 2, or an immunogenic fragment thereof. In some embodiments, the third immunogenic polypeptide is located at the amino terminus of the fusion protein and the first immunogenic polypeptide is located at the carboxyl terminus of the fusion protein. In some embodiments, the first immunogenic polypeptide is located at the amino terminus of the fusion protein and the second immunogenic polypeptide is located at the carboxyl terminus of the fusion protein. In some embodiments, the third immunogenic polypeptide is located at the amino terminus of

the fusion protein and the second immunogenic polypeptide is located at the carboxyl terminus of the fusion protein. In some embodiments, the first immunogenic polypeptide is located at the amino terminus of the fusion protein and the third immunogenic polypeptide is located at the carboxyl terminus of the fusion protein.

[0012] In certain embodiments, the present disclosure provides fusion proteins comprising or consisting of a first immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 2, or an immunogenic fragment thereof and a second immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 9, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 2, or an immunogenic fragment thereof, and the second immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 9, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide is located at the amino terminus of the fusion protein and the second immunogenic polypeptide is located at the carboxyl terminus of the fusion protein.

[0013] In further embodiments, the present disclosure provides fusion proteins comprising or consisting of a first immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 2, or an immunogenic fragment thereof and a second immunogenic polypeptide having an amino acid sequence at least 90% homologous to the amino acid sequence of SEQ ID NO: 6, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 2, or an immunogenic fragment thereof, and the second immunogenic polypeptide consists of or comprises the amino acid sequence of SEQ ID NO: 6, or an immunogenic fragment thereof. In some embodiments, the first immunogenic polypeptide is located at the amino terminus of the fusion protein and the second immunogenic polypeptide is located at the carboxyl terminus of the fusion protein. In some embodiments, the second immunogenic polypeptide is located at the amino terminus of the fusion protein and the first immunogenic polypeptide is located at the carboxyl terminus of the fusion protein.

[0013A] In some embodiments, the present disclosure provides a fusion protein comprising: a first polypeptide comprising an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 6 or to the amino acid sequence of SEQ ID NO: 5; and a second polypeptide comprising an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 9.

[0014] In some embodiments, the present disclosure provides vaccine formulations comprising or consisting of one or more fusion proteins described herein and a pharmaceutically acceptable carrier.

[0015] In further embodiments, the present disclosure provides methods of treating a subject suffering from or susceptible to *S. pneumoniae* infection, comprising administering to the subject an effective amount of a vaccine formulation described herein.

[0015A] Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

[0015B] Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present disclosure as it existed before the priority date of each of the appended claims.

III. Brief Description of the Drawings

[0016] The present teachings described herein will be more fully understood from the following description of various illustrative embodiments, when read together with the accompanying drawings. It should be understood that the drawings described below are for illustration purposes only and are not intended to limit the scope of the present teachings in any way.

[0017] FIG. 1 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then stimulated with killed, unencapsulated whole cell *S. pneumoniae*, as described in Example 1. The left panel shows the data in scatter format, and the right panel shows the average and standard deviation for each sample. Immunization group “All 3” represents animals immunized with a combination of SP2108, SP0148, and SP1634.

[0018] FIG. 2 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then stimulated with a combination of three proteins (SP2108, SP0148, and SP1634), as described in Example 1.

[0019] FIG. 3 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated protein(s) and cholera toxin

adjuvant, then challenged with intranasal administration of *S. pneumoniae*, as described in Example 1. 003 represents a control unrelated antigen.

[0020] FIG. 4 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in BALB/c mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then challenged with intranasal administration of *S. pneumoniae*, as described in Example 2.

[0021] FIG. 5 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated proteins and cholera toxin adjuvant, then stimulated with the protein of immunization (left panel) or killed, unencapsulated whole cell *S. pneumoniae* (right panel), as described in Example 3.

[0022] FIG. 6 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and cholera toxin adjuvant, then challenged with intranasal administration of *S. pneumoniae*, as described in Example 4.

[0023] FIG. 7 shows survival of mice that were immunized with the indicated proteins and the adjuvant alum, then underwent aspiration challenge with *S. pneumoniae* as described in Example 5.

[0024] FIG. 8 shows survival of mice that were immunized with the indicated proteins and the adjuvant alum, then underwent aspiration challenge with *S. pneumoniae* as described in Example 6.

[0025] FIG. 9 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and cholera toxin adjuvant, then challenged with intranasal administration of *S. pneumoniae*, as described in Example 7.

[0026] FIG. 10 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated proteins and alum, then stimulated with the proteins indicated at upper left, as described in Example 8.

[0027] FIG. 11 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), then challenged with intranasal administration of *S. pneumoniae*, as described in Example 9.

[0028] FIG. 12 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), then challenged with intranasal administration of *S. pneumoniae*, in two pooled studies as described in Example 10.

[0029] FIG. 13 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCB), then challenged with intranasal administration of *S. pneumoniae*, as described in Example 11.

[0030] FIG. 14 shows survival of mice that were injected with antibodies or sera specific to the indicated proteins, then underwent aspiration challenge with *S. pneumoniae*, as described in Example 12.

[0031] FIG. 15 shows the percent of animals protected from sepsis in six separate aspiration challenge studies, two of which are described in more detail in Examples 6 and 12.

[0032] FIG. 16, Panels A-C show the mean IL-17A response generated by blood samples from mice that were immunized with the indicated proteins and alum, then stimulated with the proteins indicated at upper right, as described in Example 13.

[0033] FIG. 17, Panels A and B show the concentration of IL-17A generated by blood samples from mice immunized with the indicated fusion proteins and alum, or with alum alone. Blood samples were stimulated with the protein of immunization or a control protein, as indicated at upper right and described in Example 14.

[0034] FIG. 18 shows the number of IL-17A spot forming units (SFU) obtained from splenocytes of mice immunized with the indicated fusions proteins and alum, or with alum alone. Splenocytes were stimulated with overlapping peptides (OLP), killed unencapsulated whole cell *S. pneumoniae* (WCA), or media, as indicated at upper right and described in Example 15.

[0035] FIG. 19, Panels A and B show the mean IL-17A response generated by blood samples from mice immunized at the indicated doses with the indicated proteins and alum, or with alum alone. Blood samples were stimulated with overlapping peptides (OLP) or killed unencapsulated whole cell *S. pneumoniae* (WCA), as indicated at upper left and described in Example 16.

[0036] FIG. 20 shows the concentration of IL-17A generated by blood samples from mice immunized with the indicated proteins and alum, or with alum alone. Blood samples were stimulated with killed unencapsulated whole cell *S. pneumoniae* as described in Example 17.

[0037] FIG. 21, Panels A-D show the number of *S. pneumoniae* colonies (CFU) obtained from nasal washes of mice immunized with the indicated proteins and alum, or alum alone, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 18.

[0038] FIG. 22, Panels A and B show the concentration of IL-17A generated by blood samples from mice immunized with the indicated combinations of proteins and alum, killed unencapsulated whole cell *S. pneumoniae* and alum (WCB), or alum alone. Blood samples were stimulated with overlapping peptides (OLP) or killed unencapsulated whole cell *S. pneumoniae* (WCA), as indicated above each graph and described in Example 20.

[0039] FIG. 23, Panels A and B show the number of *S. pneumoniae* colonies (CFU) obtained from nasal washes of mice immunized with the indicated combinations of proteins and alum, killed unencapsulated whole cell *S. pneumoniae* and alum (WCB), or alum alone, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 21.

[0040] FIG. 24, Panels A and B show hours to moribund for mice immunized with the indicated combinations of proteins and alum, or alum alone, then challenged by intravenous administration of *S. pneumoniae*, as described in Example 22.

IV. Detailed Description

[0041] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described herein.

A. Fusion Proteins

[0042] This application describes *S. pneumoniae* fusion proteins. The fusion proteins include one, two or more polypeptides that elicit (e.g., primarily elicit) a T cell response, or that elicit both a T cell and a B cell response. In certain embodiments, the fusion protein comprises one, two or more of the polypeptides or genes listed in Table 1. In some embodiments, the fusion protein comprises two of the polypeptides or genes listed in Table 1. In some embodiments, the fusion protein comprises three of the polypeptides or genes listed in Table 1.

[0043] In some embodiments, a fusion protein comprises two or more immunogenic polypeptides having an amino acid sequence comprising SEQ ID NOs: 1-11, shown in Table 1, or immunogenic fragments thereof. In some embodiments, a fusion protein comprises two immunogenic polypeptides having an amino acid sequence comprising SEQ ID NOs: 1-11, shown in Table 1, or immunogenic fragments thereof. In some embodiments, a fusion protein comprises three immunogenic polypeptides having an amino acid sequence comprising SEQ ID NOs: 1-11, shown in Table 1, or immunogenic fragments thereof. In some embodiments, at least one immunogenic polypeptide is SP2108 (SEQ ID NO: 6). In

some embodiments, at least one immunogenic polypeptide is SP0148 (SEQ ID NO: 2). In some embodiments, at least one immunogenic polypeptide is SP1912 (SEQ ID NO: 9).

[0044] In some embodiments, a fusion protein comprises one or more polypeptides homologous to the polypeptides listed in Table 1 (for example, SP1912, SP1912L, SP0148 with or without a signal sequence, SP2108 with or without a signal sequence). Individual strains of *S. pneumoniae* contain numerous mutations relative to each other, and some of these result in different protein sequences between the different strains. One of skill in the art may readily substitute an amino acid sequence, or a portion thereof, with the homologous amino acid sequence from a different *S. pneumoniae* strain. In certain aspects, this application provides immunogenic polypeptides with at least 90%, 95%, 97%, 98%, 99%, or 99.5% identity to the polypeptides listed in Table 1 or an immunogenic fragment thereof. Serotypic variation may be used to design such variants of the polypeptides listed in Table 1.

[0045] In some embodiments, fusion proteins described herein comprise two or more fragments of immunogenic peptides listed in Table 1 (for example, fragments of SP1912, SP1912L, SP0148 with or without a signal sequence, SP2108 with or without a signal sequence). In some embodiments, fusion proteins described herein comprise truncation mutants that are close in size to the polypeptides listed in Table 1. For example, they may lack at most one, two three, four, five, ten, or twenty amino acids from one or both termini (referring to component polypeptides in a fusion protein). In certain embodiments, a fragment is a truncated fragment of any of SEQ ID NOs: 1-11 lacking 1-5, 1-10, or 1-20 amino acid residues from the N-terminus, C-terminus, or both, of any one of SEQ ID NOs: 1-11. In certain embodiments, a fragment is a truncated fragment of any of SEQ ID NOs: 1-11 lacking 1-10 amino acid residues from the N-terminus, C-terminus, or both, of any one of SEQ ID NOs: 1-11. For instance, a fragment may lack 10 amino acid residues at both the N-terminus and C-terminus of any one of SEQ ID NOs:1-11, resulting in a protein lacking 20 amino acid residues. Internal deletions, e.g., of 1-10, 11-20, 21-30, or 31-40 amino acids, are also contemplated.

Table 1. Immunogenic polypeptides

Locus tag name and description	**Protein SEQ ID No.**	**DNA SEQ ID No.**	**DNA GenBank Accession No.** (from March 30, 2010)

SP0148 lacking signal sequence	1	37	-
SP0148 including signal sequence	2	38	NC_003028.3|:145,513-146,343*
SP0148 consensus lacking signal sequence	3	-	-
SP0148 consensus including signal sequence	4	-	-
SP2108 lacking signal sequence	5	39	-
SP2108 including signal sequence	6	-	NC_003028.3|:2020750-2022021
SP2108 consensus lacking signal sequence	7	-	-
SP2108 consensus including signal sequence	8	-	-
SP1912	9	40	NC_003028.3|:1824672-1824971
SP1912L	10	41	-
SP1912 consensus	11	-	-
SP0641	12		
SP0641N	13	42	
SP0641M	14		-
SP0641N consensus	15	-	-
SP0641M consensus	16	-	-
SP0882	17	-	NC_003028.3|:831804-832628
SP0882N	18	43	-
SP0882 with exogenous signal sequence	19	44	-
SP0882N with exogenous signal sequence	20	45	-
SP0882 consensus	21	-	-
SP0882N consensus	22	-	-
SP0882 consensus with exogenous	23	-	-

leader			
SP0882N consensus with exogenous leader	24	-	-
SP1634	25	-	NC_003028.3|:1534348-1535421
SP0314	26	-	NC_003028.3|:287483-290683

*NB: The database sequence incorrectly lists TTG (encoding Leu) at nucleotide positions 541-543. The correct sequence, as shown in SEQ ID NO: 38, has TTC at that codon and encodes Phe. The database sequence further does not include a C-terminal Glu found in certain isolates.

[0046] Particular examples of fusion proteins are provided in Table 2.

Table 2. Immunogenic fusion proteins

Locus tag names	Protein SEQ ID No.	DNA SEQ ID No.
SP2108/0148	47	56
SP0148/2108	48	57
SP2108/1912	49	58
SP0148/1912	50	59
SP2108/1912/0148	51	60
SP0148/1912/2108	52	61
SP2108/0148/1912	53	62
SP0148/2108/1912	54	63

[0047] In some embodiments, a fusion protein comprises an N-terminal polypeptide and a C-terminal polypeptide. In some embodiments, one or both of the N-terminal polypeptide and the C-terminal polypeptide comprise an immunogenic polypeptide, for example, a polypeptide having an amino acid sequence comprising one of SEQ ID NOs: 1-11 or an immunogenic fragment or variant thereof.

[0048] In some embodiments, the N-terminal polypeptide and the C-terminal polypeptide are directly bound to each other. In some embodiments, the N-terminal polypeptide and the C-terminal polypeptide are linked via a linker peptide. The length and/or amino acids of a linker, when present, can be adjusted to obtain a more flexible or rigid linker. Exemplary peptide linkers are shown as SEQ ID NOs: 69-71. A linker can generally be from 1-40, such as 10-30 and specifically 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids in length.

[0049] In some embodiments, a fusion protein comprises an immunogenic fragment of a fusion protein shown in Table 2. In some embodiments, a fusion protein is or includes an immunogenic fragment of any of SEQ ID NOs: 47-54. For example, a fusion protein may lack at most one, two three, four, five, ten, or twenty amino acids from the N-terminus, C-terminus, or both, of any one of SEQ ID NOs: 47-54. In some such embodiments, the same number of residues is removed from the N-terminus and the C-terminus, while in other embodiments, a different number of residues is removed from the N-terminus compared to the C-terminus. Advantageously, in some embodiments, the N-terminal polypeptide is not truncated so as not to prevent lipidation.

[0050] In certain aspects, this application provides fusion proteins with at least 90%, 95%, 97%, 98%, 99%, or 99.5% sequence identity to a fusion protein listed in Table 2. In certain embodiments, a fusion protein is or includes an amino acid sequence having at least 90%, 95%, 97%, 98%, 99%, or 99.5% identity to any one of SEQ ID NOs: 47-54.

[0051] In some embodiments, the fusion proteins include variants or fragments of the polypeptides or genes listed in Table 1. In some embodiments, a fragment included in a fusion protein described herein is close in size to a full-length polypeptide or a polypeptide listed in Table 1. For example, they may lack at most one, two, three, four, five, ten, twenty, or thirty amino acids from one or both termini. In certain embodiments, the fragment is 100-500 amino acids in length, or 150-450, or 200-400, or 250-350 amino acids in length. In some embodiments, the fragment is 100-200, 150-250, 200-300, 250-350, 300-400, 350-450, or 400-500 amino acids in length. In certain embodiments, the fragments result from processing, or partial processing, of signal sequences by an expression host, e.g. *E. coli*, an insect cell line (e.g., the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line. The fragments described above or sub-fragments thereof (e.g., fragments of 8-50, 8-30, or 8-20 amino acid residues) preferably have one of the biological activities described below, such as increasing the amount of IL-17 released by at

least 1.5 fold or 2 fold or more (e.g., either as an absolute measure or relative to an immunologically inactive protein).

[0052] The DNA and protein sequence of each gene and polypeptide may be found by searching for the Locus Tag in the publicly available database, Entrez Gene (on the NCBI NIH web site on the World Wide Web, at www.ncbi.nlm.nih.gov/sites/entrez?db=gene), in the *Streptococcus pneumoniae* TIGR4 genome, and the indicated sequences are also included in this application.

[0053] Certain polypeptides of Table 1, and variants thereof, are described in greater detail below.

1. SP0148 (SEQ ID NO: 2) and variants thereof

[0054] The protein SP0148 is named "ABC transporter, substrate-binding protein". Proteins of this class are typically extracellular proteins that interact transiently with a transmembrane protein complex. Such complexes use energy generated by ATP hydrolysis to translocate specific substrates across a cell membrane. SP0148 is a 276 or 277 (depending on the isolate) amino acid protein that contains a conserved PBPb (periplasmic binding protein) domain, spanning amino acids 40-246, which is typical of membrane-bound transport complexes. In addition, SP0148 has a bacterial extracellular solute-binding proteins family 3 domain which is largely co-extensive with the PBPb domain and extends from amino acid 40 to 244. In some embodiments, a fusion protein comprises a truncation mutant of SP0148 comprising or lacking one or more of said domains and motifs.

[0055] In some embodiments, a fusion protein includes a polypeptide containing at least 20 consecutive amino acid residues selected from SP0148. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 250, 225, 200, 175, 150, 125, or 100 consecutive amino acids from SP0148.

[0056] Endogenous SP0148 comprises a signal sequence that directs its secretion and potential lipidation. In some embodiments, the signal sequence of the polypeptide of SEQ ID NO: 2 is partially or fully processed by an expression host, e.g. *E. coli*. In some embodiments, a variant of SP0148 that lacks the signal sequence (SEQ ID NO: 1) is used. The polypeptide of SEQ ID NO: 1 is encoded by the nucleic acid of SEQ ID NO: 37, although other nucleic acid sequences (including codon-optimized sequences) may be used. SEQ ID NO: 38 encodes the full length sequence of SP0148 used in the screens herein.

[0057] Variants of the amino acid sequence and nucleotide sequence of SP0148 may be found in U.S. Patent Application Publication No. 2005/0020813, U.S. Patent Nos. 7,378,514 and 7,504,110, and European Patent Application No. EP1572868 and EP1855717.

[0058] Consensus sequences illustrating combinations of SP0148 sequences from different *S. pneumoniae* serotypes are provided as SEQ ID NOs: 3 and 4. Accordingly, in certain embodiments, a fusion protein includes a polypeptide having an amino acid sequence comprising, or consisting of, either of SEQ ID NOs: 3-4, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising one of SEQ ID NOs: 1 or 2).

[0059] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP0148. Exemplary immunogenic fragments of SP0148 include ALGLVAAGV (SEQ ID NO: 74), ELTGYEIEV (SEQ ID NO: 75), AVNNLSYTK (SEQ ID NO: 76), TYLPAEADI (SEQ ID NO: 77), RYNMAVNNL (SEQ ID NO: 78), DFQQIMVRL (SEQ ID NO: 79), EHTDNPTIL (SEQ ID NO: 80), APIAQNPNV (SEQ ID NO: 81), LPSDQQPYV (SEQ ID NO: 82), YVYPLLAQG (SEQ ID NO: 83), QGLDNLKVI (SEQ ID NO: 84), KYLYAAPI (SEQ ID NO: 85), GELTGYEI (SEQ ID NO: 86), NPNVLVVKK (SEQ ID NO: 87), KLSKQFFGD (SEQ ID NO: 88), GSPRPFIYE (SEQ ID NO: 89), AVNNLSYTK (SEQ ID NO: 90), KIFDKIGVE (SEQ ID NO: 91), MVRLSDGQF (SEQ ID NO: 92), YVYPLLAQG (SEQ ID NO: 93), VVQATTSAK (SEQ ID NO: 94), TLEKLSKQF (SEQ ID NO: 95), VAAGVLAAC (SEQ ID NO: 96), LDNLKVIEL (SEQ ID NO: 97), and NMAVNNLSY (SEQ ID NO: 98).

2. SP2108 (SEQ ID NO: 6) and variants thereof

[0060] The polypeptide SP2108 is 423 amino acids in length and is alternatively known as MalX, maltose/maltodextrin ABC transporter, or maltose/maltodextrin-binding protein. Much of the protein (amino acids 3-423) is classified as a MalE (Maltose-binding periplasmic) domain. In addition, SP2108 contains a signal sequence that directs its secretion and potential lipidation. In some embodiments, the signal sequence of the polypeptide of SEQ ID NO: 6 is partially or fully processed by an expression host, e.g. *E. coli*. In some embodiments, a fusion protein comprises a truncation mutant of SP2108 comprising one or more of said domains and motifs.

[0061] In some embodiments, a fusion protein includes an SP2108 variant that lacks the signal sequence. This variant is represented by polypeptide sequence SEQ ID NO: 5 and

may be encoded by, for example, a nucleic acid according to SEQ ID NO: 39, although due to degeneracy in the genetic code, other DNA sequences (including codon-optimized sequences) may be used.

[0062] In some embodiments, a fusion protein includes a polypeptide containing at least 20 consecutive amino acid residues selected from SP2108. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 400, 350, 300, 250, 200, 150, or 100 consecutive amino acids from SP2108.

[0063] Consensus sequences illustrating combinations of SP2108 sequences from different serotypes are provided as SEQ ID NOs: 7 and 8. Thus, in certain embodiments, a fusion protein contains a polypeptide having an amino acid sequence comprising, or consisting of, either of SEQ ID NOs: 7-8, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising one of SEQ ID NOs: 5 or 6).

[0064] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP2108. Exemplary immunogenic fragments of SP2108 include AIIDGPWKA (SEQ ID NO: 99), VMMAPYDRV (SEQ ID NO: 100), SIAGINYAK (SEQ ID NO: 101), VWDPAKNML (SEQ ID NO: 102), QPLPNISQM (SEQ ID NO: 103), APYDRVGSL (SEQ ID NO: 104), APAVIESLV (SEQ ID NO: 105), FYYTYGLLA (SEQ ID NO: 106), SKYAFAGE (SEQ ID NO: 107), TEGAGNLI (SEQ ID NO: 108), LADWTNFYY (SEQ ID NO: 109), SLVMYYNKD (SEQ ID NO: 110), KEAGVKVTL (SEQ ID NO: 111), KSTAVLGTV (SEQ ID NO: 112), GAKTDDTTK (SEQ ID NO: 113), SQKFVDFLV (SEQ ID NO: 114), QAFKDAKVN (SEQ ID NO: 115), AVIESLVMY (SEQ ID NO: 116), DAKTAANDA (SEQ ID NO: 117), YGVATIPTL (SEQ ID NO: 118), KTAAIIDGP (SEQ ID NO: 119), KAYEKEAGV (SEQ ID NO: 120), AGNGAYVFG (SEQ ID NO: 121), and AWVIPQAVK (SEQ ID NO: 122).

3. SP1912 (SEQ ID NO: 9) and variants thereof

[0065] SP1912 is a hypothetical protein of 99 amino acids. While the protein function is not definitively known, sequence analysis suggests it is a putative thioredoxin.

[0066] In some embodiments, a fusion protein includes a polypeptide containing at least 20 consecutive amino acid residues selected from SP1912. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 90, 75, 60, 45 or 30 consecutive amino acids from SP1912.

[0067] In some embodiments, a fusion protein includes an SP1912 variant that comprises an exogenous lipidation sequence. In some embodiments, a signal sequence directs lipidation. Thus, the lipidation signal may be, e.g., the signal sequence of SP2108 (SEQ ID NO: 67) or SP0148, or an *E. coli* signal sequence. The exemplary variant SP1912L, comprising the signal sequence of the *E. coli* gene RlpB (SEQ ID NO: 68) is represented by polypeptide sequence SEQ ID NO: 10. SP1912 (SEQ ID NO: 9) and SP1912L (SEQ ID NO: 10) may be encoded, respectively, by nucleic acids according to SEQ ID NO: 40 and 41, although due to degeneracy in the genetic code, other DNA sequences (including codon-optimized sequences) may be used. In some embodiments, the lipidation sequence is provided by fusing a full-length polypeptide such as SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) to SP1912. Exemplary fusions are SEQ ID NOs: 49-54.

[0068] Consensus sequences illustrating combinations of SP1912 sequences from different serotypes are provided as SEQ ID NO: 11. Thus, in certain embodiments, formulation fusion protein comprises a polypeptide having an amino acid sequence comprising, or consisting of, SEQ ID NO: 11, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising SEQ ID NO: 9).

[0069] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP1912. Exemplary immunogenic fragments of SP1912 include KMWMAGLALLGIGSL (SEQ ID NO: 123), LLGIGSLALATKKVA (SEQ ID NO: 124), MAGLALLGIGSLALA (SEQ ID NO: 125), WMAGLALLGIGSLAL (SEQ ID NO: 126), GLALLGIGSLALATK (SEQ ID NO: 127), LALLGIGSLALATKK (SEQ ID NO: 128), FSDMGEIATLYVQVY (SEQ ID NO: 129), KAKKMWMAGLALLGI (SEQ ID NO: 130), ALLGIGSLALATKKVAKKMWMAGLALLGIG (SEQ ID NO: 131), SDMGEIATLYVQVYE (SEQ ID NO: 132), DMGEIATLYVQVYES (SEQ ID NO: 133), AGLALLGIGSLALAT (SEQ ID NO: 134), MGEIATLYVQVYESS (SEQ ID NO: 135), KKMWMAGLALLGIGS (SEQ ID NO: 136), GMKAKKMWMAGLALL (SEQ ID NO: 137), MKAKKMWMAGLALLG (SEQ ID NO: 138), HFSDMGEIATLYVQV (SEQ ID NO: 139), MNGMKAKKMWMAGLA (SEQ ID NO: 140), MWMAGLALLGIGSLA (SEQ ID NO: 141), DHFSDMGEIATLYVQ (SEQ ID NO: 142), RDHFSDMGEIATLYV (SEQ ID NO: 143), and NGMKAKKMWMAGLAL (SEQ ID NO: 144).

4. Lipidated Polypeptides

[0070] In certain embodiments, a fusion protein described herein contains at least one lipidated polypeptide. In some embodiments, the fusion protein is lipidated. In certain embodiments, the fusion protein is lipidated on the N-terminal peptide. Conjugation to the lipid moiety may be direct or indirect (e.g., via a linker). The lipid moiety may be synthetic or naturally produced. In certain embodiments, a polypeptide from Table 1 or 2 may be chemically conjugated to a lipid moiety. In certain embodiments, a construct may comprise a gene or polypeptide from Table 1, or 2, or an immunogenic fragment or variant thereof, and a lipidation sequence including a lipobox motif. A canonical lipobox motif is shown as SEQ ID NO: 66. A lipidation sequence may be N-terminal or C-terminal to the protein, and may be embedded in a signal or other sequence, or in a fusion protein. Exemplary lipidation sequences include the signal sequence of SP2108 (SEQ ID NO: 67) and the signal sequence of the *E. coli* gene RlpB (SEQ ID NO: 68). A signal sequence may be, for example, an *E. coli* or *S. pneumoniae* signal sequence. Exemplary *E. coli* signal sequences include the mlpA signal sequence (Lin, J.J. *et al.,* "An *Escherichia coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein" Proc Natl Acad Sci U S A. 1978 Oct;75(10):4891-5), the lamB signal sequence (Emr, S.D. *et al.* "Mutations altering the cellular localization of the phage lambda receptor, an *Escherichia coli* outer membrane protein", Proc Natl Acad Sci U S A. 1978 Dec;75(12):5802-6), or the MBP signal sequence (Bassford, P.J., "Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm" J Bacteriol. 1979 Jul;139(1):19-31). Lpp is an exemplary *E. coli* signal sequence that directs lipidation (Cullen, P.A. *et al.* "Construction and evaluation of a plasmid vector for the expression of recombinant lipoproteins in *Escherichia coli*" Plasmid. 2003 Jan;49(1):18-29.) *E. coli* signal sequences that direct lipidation are also described in Legrain, M. *et al.* ("Production of lipidated meningococcal transferrin binding protein 2 in Escherichia coli" Protein Expr Purif. 1995 Oct;6(5):570-8), e.g. the signal sequence of the gene RlpB (SEQ ID NO: 68). Numerous *S. pneumoniae* signal sequences are known in the art. One such signal sequence is SEQ ID NO: 67.

5. Tagged Polypeptides

[0071] In some embodiments, one or more polypeptides included in a fusion protein described herein may comprise a tag. A tag may be N-terminal or C-terminal. For instance, tags may be added to a nucleic acid or polypeptide to facilitate purification, detection, solubility, or confer other desirable characteristics on the protein or nucleic acid. For

instance, a purification tag may be a peptide, oligopeptide, or polypeptide that may be used in affinity purification. Examples include His, GST, TAP, FLAG, myc, HA, MBP, VSV-G, thioredoxin, V5, avidin, streptavidin, BCCP, Calmodulin, Nus, S tags, lipoprotein D, and β galactosidase. Particular exemplary His tags include HHHHHH (SEQ ID NO:64) and MSYYHHHHHH (SEQ ID NO: 65). In other embodiments, the polypeptide is free of tags such as protein purification tags, and is purified by a method not relying on affinity for a purification tag. In some embodiments, the fused portion is short. This, in some instances, the fusion protein comprises no more than 1, 2, 3, 4, 5, 10, or 20 additional amino acids on one or both termini of the polypeptide of Table 1 or 2.

B. Nucleic Acids

[0072] In certain embodiments, this application provides nucleic acids encoding one or more of the polypeptides and/or fusion proteins described herein, such as DNA, RNA, or an analog thereof. The underlying DNA sequences for the polypeptides described above may be modified in ways that do not affect the sequence of the protein product, and such sequences are included in the invention. For instance, the DNA sequence may be codon-optimized to improve expression in a host such as *E. coli*, an insect cell line (e.g., using the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line.

[0073] In certain embodiments, this application provides polynucleotides (such as DNA, RNA, or an analog thereof) that are at least 70%, 80%, 90%, 95%, 97%, 98%, 99%, or 100% identical to polynucleotide sequences of Table 1 or 2, or a variant or portion of said sequence. In certain embodiments, the polynucleotide is 600-2000, 800-1800, 1000-1600, 1200-1400 nucleotides in length. In some embodiments, the polynucleotide is 600-1600, 800-1800, 1000-2000, 2000-3000, or 3000-4000 nucleotides in length. The nucleic acids may be used, for example, for recombinant production of a polypeptide of Table 1 or a fusion protein of Table 2, or immunogenic fragments thereof.

[0074] Polynucleotides encoding peptides of Tables 1 or 2 or a fragment thereof can be cloned into any of a variety of expression vectors, under the control of a variety of regulatory elements, and fusions can be created with other peptides of Tables 1 or 2 or with other sequences of interest. Methods of cloning nucleic acids are routine and conventional in the art. For general references describing methods of molecular biology which are mentioned

in this application, e.g., isolating, cloning, modifying, labeling, manipulating, sequencing and otherwise treating or analyzing nucleic acids and/or proteins, see, e.g., Sambrook, J. et al. (1989). Molecular Cloning, a Laboratory Manual. Cold Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Ausubel, F. M. et al. (1995). Current Protocols in Molecular Biology, N.Y., John Wiley & Sons; Davis et al. (1986), Basic Methods in Molecular Biology, Elsevir Sciences Publishing,, Inc., New York; Hames et al. (1985), Nucleic Acid Hybridization, IL Press; Dracopoli, N.C. et al. Current Protocols in Human Genetics, John Wiley & Sons, Inc.; and Coligan, J. E., et al. Current Protocols in Protein Science, John Wiley & Sons, Inc.

C. Immunogenic compositions

[0075] The present disclosure also provides immunogenic compositions (e.g., vaccine compositions) of, or including, one or more fusion proteins described herein. In some embodiments, an immunogenic composition further includes a pharmaceutically acceptable carrier. The *S. pneumoniae* antigens described herein were identified by screening immune cells from mice either immunized with unencapsulated killed whole *S. pneumoniae* or infected with live *S. pneumoniae*, or from healthy human donors that elicit a T cell specific immune response. The human donors have almost certainly been exposed to *S. pneumoniae* at some point during their lifetimes, because *S. pneumoniae* is a very common disease and colonizing pathogen. Thus, the present disclosure contemplates immunogenic compositions of fusion proteins including *S. pneumoniae* antigens that elicit a strong immune response in all age groups that is long-lived in immunized or infected mice or humans for counteracting infection by *S. pneumoniae*.

[0076] In certain embodiments, an immunogenic composition described herein (e.g., a vaccine composition) induces a T_H17 cell response at least 1.5-fold greater than that induced by a control unrelated antigen (such as the HSV-2 protein ICP47 with the gene name US12) after contacting T_H17 cells. In some embodiments, an immunogenic composition described herein inhibits infection by *S. pneumoniae* in an uninfected subject. In certain embodiments, an immunogenic composition described herein reduces occurrence or duration of *S. pneumoniae* nasopharyngeal colonization in an individual colonized or infected by *S. pneumoniae*. In some embodiments, an immunogenic composition described herein inhibits development of sepsis in an individual infected by *S. pneumoniae*. In some embodiments, an immunogenic composition described herein inhibits development of invasive diseases such as

pneumonia, meningitis, otitis media, sinusitis or infection of other sites or organs with *S. pneumoniae.*

[0077] In some embodiments, an immunogenic composition described herein may also comprise portions of said Streptococcus polypeptides, including fusion proteins, for example deletion mutants, truncation mutants, oligonucleotides, and peptide fragments. In some embodiments, the portions of said polypeptides are immunogenic. The immunogenicity of a portion of a protein is readily determined using the same assays that are used to determine the immunogenicity of the full-length protein. In some embodiments, the portion of the polypeptide has substantially the same immunogenicity as the full-length proteins. In some embodiments, the immunogenicity is no more than 10%, 20%, 30%, 40%, or 50% less than that of the full-length protein (e.g., polypeptides of Tables 1, 2 and 3). The peptide fragments may be, for example, linear, circular, or branched.

[0078] Typically, polypeptides and fusion proteins described herein, and fragments and variants thereof, are immunogenic. These polypeptides may be immunogenic in mammals, for example mice, guinea pigs, or humans. An immunogenic polypeptide is typically one capable of raising a significant immune response in an assay or in a subject. The immune response may be innate, humoral, cell-mediated, or mucosal (combining elements of innate, humoral and cell-mediated immunity). For instance, an immunogenic polypeptide may increase the amount of IL-17 produced by T cells. Alternatively or additionally, an immunogenic polypeptide may (i) induce production of antibodies, e.g., neutralizing antibodies, that bind to the polypeptide and/or the whole bacteria, (ii) induce T_H17 immunity, (iii) activate the $CD4^+$ T cell response, for example by increasing $CD4^+$ T cells and/or increasing localization of $CD4^+$ T cells to the site of infection or reinfection, (iv) activate the $CD8^+$ CTL response, for example by increasing $CD8^+$ T cells and/or increasing localization of $CD8^+$ T cells to the site of infection or reinfection, (v) induce T_H1 immunity, (vi) induce anti-microbial peptides, and/or (vii) activate innate immunity. In some embodiments, an immunogenic polypeptide causes the production of a detectable amount of antibody specific to that antigen.

[0079] In certain embodiments, polypeptides described herein have less than 20%, 30%, 40%, 50%, 60% or 70% identity to human autoantigens and/or gut commensal bacteria (e.g., certain *Bacteroides*, *Clostridium*, *Fusobacterium*, *Eubacterium*, *Ruminococcus*, *Peptococcus*, *Peptostreptococcus*, *Bifidobacterium*, *Escherichia* and *Lactobacillus* species).

Examples of human autoantigens include insulin, proliferating cell nuclear antigen, cytochrome P450, and myelin basic protein.

[0080] A polypeptide described herein (e.g., a polypeptide included in a fusion protein described herein) may comprise one or more immunogenic portions and one or more non-immunogenic portions. The immunogenic portions may be identified by various methods, including protein microarrays, ELISPOT/ELISA techniques, and/or specific assays on different deletion mutants (e.g., fragments) of the polypeptide in question. Immunogenic portions may also be identified by computer algorithms. Some such algorithms, like EpiMatrix (produced by EpiVax), use a computational matrix approach. Other computational tools for identifying antigenic epitopes include PEPVAC (Promiscuous EPitope-based VACcine, hosted by Dana Farber Cancer Institute on the world wide web at immunax.dfci.harvard.edu/PEPVAC), MHCPred (which uses a partial least squares approach and is hosted by The Jenner Institute on the world wide web at www.jenner.ac.uk/MHCPred), and Immune Epitope Database algorithms on the world wide web at tools.immuneepitope.org. An immunogenic fragment of a polypeptide described herein comprises at least one immunogenic portion, as measured experimentally or identified by algorithm (for example, the SYFPEITHI algorithm found at www.syfpeithi.de).

D. Multi-component Immunogenic Compositions

[0081] In certain embodiments, an immunogenic composition described herein (e.g., a vaccine composition) includes a fusion protein described herein and additionally one or more, or two or more, known *S. pneumoniae* antigens. In some instances, the known *S. pneumoniae* antigens are predominantly antibody targets shown in Table 3. In some instances, the known *S. pneumoniae* antigens are polysaccharides. In some instances, the known *S. pneumoniae* antigens protect from *S. pneumoniae* colonization, or from *S. pneumoniae*-induced sepsis, pneumonia, meningitis, otitis media, sinusitis, or infection of other sites or organs by *S. pneumoniae*.

Table 3. Immunogenic polypeptides for vaccine formulations

PspA	27	46	
PR+NRB from PspA with coiled-coil	28		
CD2	29		
PR+NRB from PspA w/o coiled-coil	30		
PR only with coiled-coil	31		
PR only w/o coiled-coil	32		
H70 (PR+NRB from PspA aa 290-410)	33		
Non-proline Block (NPB)	34		
Non-proline Block (NPB)	35		
Non-proline Block (NPB)	36		
L460D	55		

[0082] One appropriate art-recognized class of *S. pneumoniae* antigen is Pneumococcal surface protein A (PspA) (SEQ ID NO: 27) and derivatives of PspA. Derivatives of PspA include proline-rich segments with the non-proline block (PR+NPB, also referred to as PRN and further described below as well as in Daniels, C.C. *et al.* (2010) Infection and Immunity 78:2163-72) and related constructs comprising all or a fragment of the proline-rich region of PspA (e.g., regions containing one or more of the sequences PAPAP, PKP, PKEPEQ and PEKP and optionally including a non-proline block). In some embodiments, fragments or variants of PspA comprise proline-rich segments with the non-proline block and 10, 20 30, 40 or more additional amino acids of PspA sequence. H70 (SEQ ID NO: 33) is one exemplary sequence which includes the proline-rich region and non-proline-block encompassing amino acids 290-410 PspA. An example of the non-proline-block has the exemplary sequence EKSADQQAEEDYARRSEEEYNRLTQQQ (SEQ ID NO: 34), which generally has no proline residues in an otherwise proline-rich area of the non-coiled region of PspA. Other embodiments of non-proline block (NPB) sequences include SEQ ID NOs: 35 and 36 and PspA and its derivatives can include genes expressing similar proline-rich structures (i.e. PKP, PKEPEQ and PEKP), with or without the NPB. The amino acids at either end of the NPB mark the boundaries of the proline-rich region. In one example, the amino-terminal boundary to the PR-region is DLKKAVNE (SEQ ID NO: 72), and the carboxy-terminal boundary is (K/G)TGW(K/G)QENGMW (SEQ ID NO: 73).

Peptides containing the NPB are particularly immunogenic, suggesting that the NPB may be an important epitope. Exemplary immunogenic PspA polypeptide derivatives containing the coiled-coil structure include SEQ ID NOs: 28 and 31. Particular embodiments of the immunogenic PspA polypeptide derivatives lacking the coiled-coil structure have the amino acid sequences shown as SEQ ID NOs: 29, 30 and 32. Immunogenic PspA polypeptides SEQ ID NO: 28-30 include both PR and NPB sequences (PR+NPB). Immunogenic PspA polypeptides of SEQ ID NOs: 31 and 32 include only a PR sequence (PR only) and lack the NPB.

[0083] Another appropriate art-recognized class of *S. pneumoniae* antigen is the pneumolysoids. Pneumolysoids have homology to the *S. pneumoniae* protein pneumolysin (PLY), but have reduced toxicity compared to pneumolysin. Pneumolysoids can be naturally occurring or engineered derivatives of pneumolysin. In some embodiments, a pneumolysoid has at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to pneumolysin. In some embodiments, the pneumolysoid demonstrates less than 1/2, 1/5, 1/10, 1/20, 1/50, 1/100, 1/200, 1/500, or 1/1000 the toxicity of pneumolysin in an assay for one or both of hemolytic activity towards erythrocytes and inhibition of polymorphonuclear leukocytes. Both assays are described in Saunders F.K. *et al.* ("Pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae*, does not require a thiol group for in vitro activity" Infect Immun. 1989 Aug;57(8):2547-52). Exemplary pneumolysoids include PdT (a triple mutant further described in Berry, A.M. *et al.* (1995) Infection and Immunity 63:1969-74); Pd-A and Pd-B (Paton J.C. *et al.* "Purification and immunogenicity of genetically obtained pneumolysin toxoids and their conjugation to *Streptococcus pneumoniae* type 19F polysaccharide" Infect Immun. 1991 Jul;59(7):2297-304); rPd2 and rPd3 (Ferreira *et al.* "DNA vaccines based on genetically detoxified derivatives of pneumolysin fail to protect mice against challenge with *Streptococcus pneumoniae*" FEMS Immunol Med Microbiol (2006) 46: 291-297); Ply8, Δ6PLY, L460D (see, e.g., US 2009/0285846 and L. Mitchell, Protective Immune Responses to Streptococcus pneumoniae Pneumolysoids, ASM2011 conference abstract, 2011); or a variant thereof. In some embodiments, the pneumolysin has a mutation in the catalytic center, such as at amino acid 428 or 433 or the vicinity.

[0084] Other appropriate *S. pneumoniae* antigens include Choline-binding protein A (CbpA) and derivatives thereof (AD Ogunniyi *et al.*, "Protection against Streptococcus pneumoniae elicited by immunization with pneumolysin and CbpA," Infect Immun. 2001 Oct;69(10):5997-6003); Pneumococcal surface adhesin A (PsaA); caseinolytic protease;

sortase A (SrtA); pilus 1 RrgA adhesin; PpmA; PrtA; PavA; LytA; Stk-PR; PcsB; RrgB and derivatives thereof. CpbA derivatives include fusion proteins described in WO 2012/134975. Such fusion proteins may comprise one or more copies of the R2 domain, $R2_1$ and/or $R2_2$ subdomains of CpbA, or active variants and fragments thereof, or any combination thereof. Such fusion proteins may further comprise a pneumolysoid. The construct YLN, for example, comprises CpbA polypeptides YPT and NEEK, and pneumolysoid L460D (SEQ ID NO: 55).

[0085] In some cases, the antigen is at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identical to the corresponding wild-type *S. pneumoniae* protein. Sequences of the above-mentioned polypeptides, and nucleic acids that encode them, are known; see, for example, the S. *pneumoniae* ATCC 700669 complete genome sequence under GenBank accession number FM211187.1 and linked polypeptide sequences therein.

[0086] In some embodiments, an immunogenic composition (e.g., a vaccine composition) contains one or more fusion proteins described herein in combination with one or more, e.g., two, three, four, or more polypeptides from Table 1 or 3 or immunogenic fragments or variants thereof in a mixture. In some embodiments, the mixture contains both full-length polypeptides and fragments resulting from processing, or partial processing, of signal sequences by an expression host, e.g. *E. coli*, an insect cell line (e.g., the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line. In some embodiments, an immunogenic composition contains one or more fusion proteins of any of SEQ ID NOs: 47-54 in the absence of any other antigens. In some embodiments, an immunogenic composition contains one or more fusion proteins of any of SEQ ID NOs: 47-54 in combination with one or more additional proteins of any of SEQ ID NOs: 1-36, in the absence of other antigens.

[0087] In some embodiments, polypeptides described in Tables 1 and 3 may be used without modification in conjunction with the fusion proteins described herein. In certain embodiments, when smaller related polypeptides are used, such as fragments or the like, and their molecular weight is less than about 5000 daltons, e.g., 1500 to 5000 daltons, modification may be useful in eliciting the desired immune response. For example, the smaller polypeptides can be conjugated to an appropriate immunogenic carrier such as tetanus toxoid, keyhole limpet hemocyanin or the like.

[0088] In some embodiments, polypeptides described herein may be conjugated to *S. pneumoniae* polysaccharides. The conjugated polysaccharides may be, for example, as described in US Patent 5,623,057, US Patent 5,371,197, or PCT/US2011/023526.

[0089] Some embodiments of immunogenic compositions described herein include an immunogenic polypeptide, including fusion proteins (e.g., a polypeptide of Table 1, 2 or 3) that contains a membrane translocating sequence (MTS), to facilitate introduction of the polypeptide into the mammalian cell and subsequent stimulation of the cell-mediated immune response. Exemplary membrane translocating sequences include hydrophobic region in the signal sequence of Kaposi fibroblast growth factor, the MTS of α-synuclein, β-synuclein, or γ-synuclein, the third helix of the Antennapedia homeodomain, SN50, integrin β3 h-region, HIV Tat, pAntp, PR-39, abaecin, apidaecin, Bac5, Bac7, *P. berghei* CS protein, and those MTSs described in US Patents 6,248,558, 6,432,680 and 6,248,558.

[0090] In certain embodiments, an antigen (e.g., a polypeptide of Table 1, 2 or 3) is covalently bound to another molecule. This may, for example, increase the half-life, solubility, bioabailability, or immunogenicity of the antigen. Molecules that may be covalently bound to the antigen include a carbohydrate, biotin, poly(ethylene glycol) (PEG), polysialic acid, N-propionylated polysialic acid, nucleic acids, polysaccharides, and PLGA. There are many different types of PEG, ranging from molecular weights of below 300 g/mol to over 10,000,000 g/mol. PEG chains can be linear, branched, or with comb or star geometries. In some embodiments, the naturally produced form of a protein is covalently bound to a moeity that stimulates the immune system. An example of such a moeity is a lipid moeity. In some instances, lipid moieties are recognized by a Toll-like receptor (TLR) such as TLR-2 or TLR-4, and activate the innate immune system.

[0091] In some embodiments, a fusion protein and one or more additional components described herein are mixed together using known methods to form a multi-component immunogenic composition. In some embodiments, a fusion protein and one or more additional components described herein are nano-encapsulated using known methods. In some embodiments, a fusion protein and one or more additional components described herein are molded into nano- or micro-particles using known methods. In some embodiments, a fusion protein and one or more additional components described herein are conjugated through a covalent bond using known methods to form a multi-component immunogenic composition. In some embodiments, a fusion protein and one or more additional components described herein are joined non-covalently using known methods to

form a multi-component immunogenic composition. Additional methods of combining a fusion protein and one or more additional components are described in, e.g., PCT/US12/37412 and PCT/US09/44956.

E. Additional Immunogenic Composition Components

[0092] In certain embodiments, an immunogenic composition described herein (e.g, a vaccine composition) comprises one or more of the following: an adjuvant, stabilizer, buffer, surfactant, controlled release component, salt, preservative, and/or an antibody specific to a polypeptide included in an immunogenic composition.

1. Adjuvants

[0093] In some embodiments, an immunogenic composition may include an adjuvant. Adjuvants can be broadly separated into two classes, based on their principal mechanisms of action: vaccine delivery systems and immunostimulatory adjuvants (see, *e.g.*, Singh *et al.*, *Curr. HIV Res.* 1:309-20, 2003). In most vaccine formulations, the adjuvant provides a signal to the immune system so that it generates a response to the antigen, and the antigen is required for driving the specificity of the response to the pathogen. Vaccine delivery systems are often particulate formulations, *e.g.*, emulsions, microparticles, immune-stimulating complexes (ISCOMs), nanoparticles, which may be, for example, particles and/or matrices, and liposomes. In contrast, immunostimulatory adjuvants are sometimes derived from pathogens and can represent pathogen associated molecular patterns (PAMP), *e.g.*, lipopolysaccharides (LPS), monophosphoryl lipid (MPL), or CpG-containing DNA, which activate cells of the innate immune system.

[0094] Alternatively, adjuvants may be classified as organic and inorganic. Inorganic adjuvants include alum salts such as aluminum phosphate, amorphous aluminum hydroxyphosphate sulfate, and aluminum hydroxide, which are commonly used in human vaccines. Organic adjuvants comprise organic molecules including macromolecules. An example of an organic adjuvant is cholera toxin.

[0095] Adjuvants may also be classified by the response they induce. In some embodiments, the adjuvant induces the activation of T_H1 cells or T_H2 cells. In other embodiments, the adjuvant induces the activation of B cells. In yet other embodiments, the adjuvant induces the activation of antigen-presenting cells. These categories are not mutually exclusive; in some cases, an adjuvant activates more than one type of cell.

[0096] In certain embodiments, the adjuvant induces the activation of T_H17 cells. It may promote the $CD4^+$ or $CD8^+$ T cells to secrete IL-17. In some embodiments, an adjuvant that induces the activation of T_H17 cells is one that produces at least a 2-fold, and in some cases a 10-fold, experimental sample to control ratio in the following assay. In the assay, an experimenter compares the IL-17 levels secreted by two populations of cells: (1) cells from animals immunized with the adjuvant and a polypeptide known to induce T_H17 activation, and (2) cells from animals treated with the adjuvant and an irrelevant (control) polypeptide. An adjuvant that induces the activation of T_H17 cells may cause the cells of population (1) to produce more than 2-fold, or more than 10-fold more IL-17 than the cells of population (2). IL-17 may be measured, for example, by ELISA or ELISPOT. Certain toxins, such as cholera toxin and labile toxin (produced by enterotoxigenic *E. coli*, or ETEC), activate a T_H17 response. Thus, in some embodiments, the adjuvant is a toxin. Cholera toxin was successfully used in the mouse model to induce protective immunity in conjunction with certain polypeptides from Table 1 (see Examples 1-2). One form of labile toxin is produced by Intercell. Mutant derivates of labile toxin that are active as adjuvants but significantly less toxic can be used as well. Exemplary detoxified mutant derivatives of labile toxin include mutants lacking ADP-ribosyltransferase activity. Particular detoxified mutant derivatives of labile toxin include LTK7 (Douce *et al.*, "Mutants of *Escherichia coli* heat-labile toxin lacking ADP-ribosyltransferase activity act as nontoxic, mucosal adjuvants" PNAS Vol. 92, pp. 1644-1648, February 1995) and LTK63 (Williams *et al.*, "Innate Imprinting by the Modified Heat-Labile Toxin of *Escherichia coli* (LTK63) Provides Generic Protection against Lung Infectious Disease" The Journal of Immunology, 2004, 173: 7435-7443), LT-G192 (Douce *et al.* "Genetically detoxified mutants of heat-labile toxin from *Escherichia coli* are able to act as oral adjuvants" Infect Immun. 1999 Sep;67(9):4400-6), and LTR72 ("Mucosal adjuvanticity and immunogenicity of LTR72, a novel mutant of Escherichia coli heat-labile enterotoxin with partial knockout of ADP-ribosyltransferase activity." J Exp Med. 1998 Apr 6;187(7):1123-32).

[0097] In some embodiments, the adjuvant comprises a VLP (virus-like particle). One such adjuvant platform, Alphavirus replicons, induces the activation of T_H17 cells using alphavirus and is produced by Alphavax. In certain embodiments of the Alphavirus replicon system, alphavirus may be engineered to express an antigen of interest, a cytokine of interest (for example, IL-17 or a cytokine that stimulates IL-17 production), or both, and may be produced in a helper cell line. More detailed information may be found in U.S. Patent Nos.

5,643,576 and 6,783,939. In some embodiments, a vaccine formulation is administered to a patient in combination with a nucleic acid encoding a cytokine.

[0098] Certain classes of adjuvants activate toll-like receptors (TLRs) in order to activate a T_H17 response. TLRs are well known proteins that may be found on leukocyte membranes, and recognize foreign antigens (including microbial antigens). Administering a known TLR ligand together with an antigen of interest (for instance, as a fusion protein) can promote the development of an immune response specific to the antigen of interest. One exemplary adjuvant that activates TLRs comprises Monophosphoryl Lipid A (MPL). Traditionally, MPL has been produced as a detoxified lipopolysaccharide (LPS) endotoxin obtained from gram negative bacteria, such as *S. minnesota*. In particular, sequential acid and base hydrolysis of LPS produces an immunoactive lipid A fraction (which is MPL), and lacks the saccharide groups and all but one of the phosphates present in LPS. A number of synthetic TLR agonists (in particular, TLR-4 agonists) are disclosed in Evans JT *et al.* "Enhancement of antigen-specific immunity via the TLR-4 ligands MPL adjuvant and Ribi.529." Expert Rev Vaccines 2003 Apr;2(2):219-29. Like MPL adjuvants, these synthetic compounds activate the innate immune system via TLR. Another type of TLR agonist is a synthetic phospholipid dimer, for example E6020 (Ishizaka ST *et al.* "E6020: a synthetic Toll-like receptor 4 agonist as a vaccine adjuvant" Expert Rev. Vaccines. 2007 Oct; 6(5):773-84.). Various TLR agonists (including TLR-4 agonists) have been produced and/or sold by, for example, the Infectious Disease Research Institute (IRDI), Corixa, Esai, Avanti Polar Lipids, Inc., and Sigma Aldrich. Another exemplary adjuvant that activates TLRs comprises a mixture of MPL, Trehalose Dicoynomycolate (TDM), and dioctadecyldimethylammonium bromide (DDA). Another TLR-activating adjuvant is R848 (resiquimod).

[0099] In some embodiments, the adjuvant is or comprises a saponin. Typically, the saponin is a triterpene glycoside, such as those isolated from the bark of the *Quillaja saponaria* tree. A saponin extract from a biological source can be further fractionated (e.g., by chromatography) to isolate the portions of the extract with the best adjuvant activity and with acceptable toxicity. Typical fractions of extract from *Quillaja saponaria* tree used as adjuvants are known as fractions A and C.

[0100] A particular form of saponins that may be used in vaccine formulations described herein is immunostimulating complexes (ISCOMs). ISCOMs are an art-recognized class of adjuvants, that generally comprise *Quillaja* saponin fractions and lipids (e.g.,

cholesterol and phospholipids such as phosphatidyl choline). In certain embodiments, an ISCOM is assembled together with a polypeptide or nucleic acid of interest. However, different saponin fractions may be used in different ratios. In addition, the different saponin fractions may either exist together in the same particles or have substantially only one fraction per particle (such that the indicated ratio of fractions A and C are generated by mixing together particles with the different fractions). In this context, "substantially" refers to less than 20%, 15%, 10%, 5%, 4%, 3%, 2% or even 1%. Such adjuvants may comprise fraction A and fraction C mixed into a ratio of 70-95 A: 30-5 C, such as 70 A : 30 C to 75 A : 5 C, 75 A : 5 C to 80 A : 20 C, 80 A : 20 C to 85 A : 15 C, 85 A : 15 C to 90 A : 10 C, 90 A : 10 C to 95 A : 5 C, or 95 A : 5 C to 99 A : 1 C.

[0101] In certain embodiments, combinations of adjuvants are used. Three exemplary combinations of adjuvants are MPL and alum, E6020 and alum, and MPL and an ISCOM.

[0102] Adjuvants may be covalently bound to antigens. In some embodiments, the adjuvant may comprise a protein which induces inflammatory responses through activation of antigen-presenting cells (APCs). In some embodiments, one or more of these proteins can be recombinantly fused with an antigen of choice, such that the resultant fusion molecule promotes dendritic cell maturation, activates dendritic cells to produce cytokines and chemokines, and ultimately, enhances presentation of the antigen to T cells and initiation of T cell responses (see Wu et al., Cancer Res 2005; 65(11), pp 4947-4954). In certain embodiments, a polypeptide, including a fusion protein, described herein is presented in the context of the trivalent conjugate system, comprising a fusion protein of *S. pneumoniae* Pneumococcal surface adhesin A (PsaA) with the pneumolysoid PdT and a cell wall polysaccharide (PsaA:PdT-CPs), described in Lu *et al.* ("Protection against Pneumococcal colonization and fatal pneumonia by a trivalent conjugate of a fusion protein with the cell wall polysaccharide." Infect Immun. 2009 May;77(5):2076-83). PdT carries three amino acid substitutions (W433F, D385N, and C428G) which render the molecule nontoxic but do not interfere with its TLR-4-mediated inflammatory properties. Conjugation of a polysaccharide to the fusion of a polypeptide to the TLR-4-agonist PdT enhances immunological response to the polypeptide. In some embodiments, one or more polypeptides described herein are used in place of PsaA in the trivalent conjugate. The trivalent conjugate system typically includes alum and is usually administered parenterally. Other exemplary adjuvants that may be covalently bound to antigens comprise polysaccharides, pneumolysin, synthetic peptides, lipopeptides, and nucleic acids.

[0103] Typically, the same adjuvant or mixture of adjuvants is present in each dose of a vaccine. Optionally, however, an adjuvant may be administered with the first dose of vaccine and not with subsequent doses (i.e., booster shots). Alternatively, a strong adjuvant may be administered with the first dose of vaccine and a weaker adjuvant or lower dose of the strong adjuvant may be administered with subsequent doses. The adjuvant can be administered before the administration of the antigen, concurrent with the administration of the antigen or after the administration of the antigen to a subject (sometimes within 1, 2, 6, or 12 hours, and sometimes within 1, 2, or 5 days). Certain adjuvants are appropriate for human patients, non-human animals, or both.

2. Other components

[0104] In addition to the antigens and the adjuvants described above, a vaccine formulation or immunogenic composition may include one or more additional components.

[0105] In certain embodiments, the vaccine formulation or immunogenic composition may include one or more stabilizers such as sugars (such as sucrose, glucose, or fructose), phosphate (such as sodium phosphate dibasic, potassium phosphate monobasic, dibasic potassium phosphate, or monosodium phosphate), glutamate (such as monosodium L-glutamate), gelatin (such as processed gelatin, hydrolyzed gelatin, or porcine gelatin), amino acids (such as arginine, asparagine, histidine, L-histidine, alanine, valine, leucine, isoleucine, serine, threonine, lysine, phenylalanine, tyrosine, and the alkyl esters thereof), inosine, or sodium borate.

[0106] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more buffers such as a mixture of sodium bicarbonate and ascorbic acid. In some embodiments, the vaccine formulation may be administered in saline, such as phosphate buffered saline (PBS), or distilled water.

[0107] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more surfactants such as polysorbate 80 (Tween 80), Triton X-100, Polyethylene glycol tert-octylphenyl ether t-Octylphenoxypolyethoxyethanol 4-(1,1,3,3-Tetramethylbutyl)phenyl-polyethylene glycol (TRITON X-100); Polyoxyethylenesorbitan monolaurate Polyethylene glycol sorbitan monolaurate (TWEEN 20); and 4-(1,1,3,3-Tetramethylbutyl)phenol polymer with formaldehyde and oxirane (TYLOXAPOL). A surfactant can be ionic or nonionic.

[0108] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more salts such as sodium chloride, ammonium chloride, calcium chloride, or potassium chloride.

[0109] In certain embodiments, a preservative is included in the vaccine or immunogenic composition. In other embodiments, no preservative is used. A preservative is most often used in multi-dose vaccine vials, and is less often needed in single-dose vaccine vials. In certain embodiments, the preservative is 2-phenoxyethanol, methyl and propyl parabens, benzyl alcohol, and/or sorbic acid.

[0110] In certain embodiments, the vaccine formulation or immunogenic composition is a controlled release formulation.

F. Use of Vaccines

[0111] The *S. pneumoniae* vaccines described herein may be used for prophylactic and/or therapeutic treatment of *S. pneumoniae*. Accordingly, this application provides a method for treating a subject suffering from or susceptible to *S. pneumoniae* infection, comprising administering an effective amount of any of the vaccine formulations described herein. In some aspects, the method inhibits *S. pneumoniae* colonization in an individual. In some aspects, the method inhibits *S. pneumoniae* symptoms, invasive disease or sequelae, such as sepsis, pneumonia, meningitis, otitis media, sinusitis or infection of other sites or organs with *S. pneumoniae*. The subject receiving the vaccination may be a male or a female, and may be a child or adult. In some embodiments, the subject being treated is a human. In other embodiments, the subject is a non-human animal.

1. Prophylactic use

[0112] In prophylactic embodiments, the vaccine is administered to a subject to induce an immune response that can help protect against the establishment of *S. pneumoniae*, for example by protecting against colonization, the first and necessary step in disease. Thus, in some aspects, the method inhibits infection by *S. pneumoniae* in a non-colonized or uninfected subject. In another aspect, the method may reduce the duration of colonization in an individual who is already colonized.

[0113] In some embodiments, the vaccine compositions of the invention confer protective immunity, allowing a vaccinated individual to exhibit delayed onset of symptoms or sequelae, or reduced severity of symptoms or sequelae, as the result of his or her exposure

to the vaccine. In certain embodiments, the reduction in severity of symptoms or sequelae is at least 25%, 40%, 50%, 60%, 70%, 80% or even 90%. In particular embodiments, vaccinated individuals may display no symptoms or sequelae upon contact with *S. pneumoniae*, do not become colonized by *S. pneumoniae*, or both. Protective immunity is typically achieved by one or more of the following mechanisms: mucosal, humoral, or cellular immunity. Mucosal immunity is primarily the result of secretory IgA (sIGA) antibodies on mucosal surfaces of the respiratory, gastrointestinal, and genitourinary tracts. The sIGA antibodies are generated after a series of events mediated by antigen-processing cells, B and T lymphocytes, that result in sIGA production by B lymphocytes on mucosa-lined tissues of the body. Humoral immunity is typically the result of IgG antibodies and IgM antibodies in serum. Cellular immunity can be achieved through cytotoxic T lymphocytes or through delayed-type hypersensitivity that involves macrophages and T lymphocytes, as well as other mechanisms involving T cells without a requirement for antibodies. In particular, cellular immunity may be mediated by T_H1 or T_H17 cells.

[0114] Essentially any individual has a certain risk of becoming infected with *S. pneumoniae*. However, certain sub-populations have an increased risk of infection. In some embodiments, a vaccine formulation as described herein (e.g., a composition comprising one or more polypeptides from Table 1, 2 or 3, or nucleic acids encoding the polypeptides, or antibodies reactive with the polypeptides) is administered to patients that are immunocompromised.

[0115] An immunocompromising condition arising from a medical treatment is likely to expose the individual in question to a higher risk of infection with *S. pneumoniae*. It is possible to treat an infection prophylactically in an individual having the immunocompromised condition before or during treatments known to compromise immune function. By prophylactically treating with an antigenic composition (e.g., two or more antigens from Table 1 or 2, or an antigen from Table 3, or nucleic acids encoding the antigens), or with antibodies reactive to two or more antigens from Table 1 or 2, or to an antigen from Table 3, before or during a treatment known to compromise immune function, it is possible to prevent a subsequent *S. pneumoniae* infection or to reduce the risk of the individual contracting an infection due to the immunocompromised condition. Should the individual contract an *S. pneumoniae* infection e.g., following a treatment leading to an immunocompromised condition it is also possible to treat the infection by administering to the individual an antigen composition.

[0116] The following groups are at increased risk of pneumococcal disease or its complications, and therefore it is advantageous for subjects falling into one or more of these groups to receive a vaccine formulation described herein: children, especially those from 1 month to 5 years old or 2 months to 2 years old; children who are at least 2 years of age with asplenia, splenic dysfunction or sickle-cell disease; children who are at least 2 years of age with nephrotic syndrome, chronic cerebrospinal fluid leak, HIV infection or other conditions associated with immunosuppression.

[0117] In another embodiment, at least one dose of the pneumococcal antigen composition is given to adults in the following groups at increased risk of pneumococcal disease or its complications: all persons 65 years of age; adults with asplenia, splenic dysfunction or sickle-cell disease; adults with the following conditions: chronic cardiorespiratory disease, cirrhosis, alcoholism, chronic renal disease, nephrotic syndrome, diabetes mellitus, chronic cerebrospinal fluid leak, HIV infection, AIDS and other conditions associated with immunosuppression (Hodgkin's disease, lymphoma, multiple myeloma, immunosuppression for organ transplantation), individuals with cochlear implants; individuals with long-term health problems such as heart disease and lung disease, as well as individuals who are taking any drug or treatment that lowers the body's resistance to infection, such as long-term steroids, certain cancer drugs, radiation therapy; Alaskan natives and certain Native American populations.

2. Therapeutic use

[0118] In therapeutic applications, the vaccine may be administered to a patient suffering from *S. pneumoniae* infection, in an amount sufficient to treat the patient. Treating the patient, in this case, refers to reducing *S. pneumoniae* symptoms and/or bacterial load and/or sequelae bin an infected individual. In some embodiments, treating the patient refers to reducing the duration of symptoms or sequelae, or reducing the intensity of symptoms or sequelae. In some embodiments, the vaccine reduces transmissibility of *S. pneumoniae* from the vaccinated patient. In certain embodiments, the reductions described above are at least 25%, 30%, 40%, 50%, 60%, 70%, 80% or even 90%.

[0119] In therapeutic embodiments, the vaccine is administered to an individual post-infection. The vaccine may be administered shortly after infection, e.g. before symptoms or sequelae manifest, or may be administered during or after manifestation of symptoms or sequelae.

[0120] A therapeutic *S. pneumoniae* vaccine can reduce the intensity and/or duration of the various symptoms or sequelae of *S. pneumoniae* infection. Symptoms or sequelae of *S. pneumoniae* infection can take many forms. In some cases, an infected patient develops pneumonia, acute sinusitis, otitis media (ear infection), meningitis, bacteremia, sepsis, osteomyelitis, septic arthritis, endocarditis, peritonitis, pericarditis, cellulitis, or brain abscess.

[0121] Sepsis is a rare but life-threatening complication of *S. pneumoniae* infection, where the bacterium invades the bloodstream and systemic inflammation results. Typically, fever is observed and white blood cell count increases. A further description of sepsis is found in Goldstein, B. *et al.* "International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics." Pediatr Crit Care Med. Jan 2005;6(1):2-8.

3. Assaying vaccination efficacy

[0122] The efficacy of vaccination with the vaccines disclosed herein may be determined in a number of ways, in addition to the clinical outcomes described above. First, one may assay IL-17 levels (particularly IL-17A) by stimulating T cells derived from the subject after vaccination. The IL-17 levels may be compared to IL-17 levels in the same subject before vaccination. Increased IL-17 (e.g., IL-17A) levels, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, would indicate an increased response to the vaccine. Alternatively (or in combination), one may assay neutrophils in the presence of T cells or antibodies from the patient for pneumococcal killing. Increased pneumococcal killing, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, would indicate an increased response to the vaccine. In addition, one may measure T_H17 cell activation, where increased T_H17 cell activation, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, correlates with an increased response to the vaccine. One may also measure levels of an antibody specific to the vaccine, where increased levels of the specific antibody, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, are correlated with increased vaccine efficacy. In certain embodiments, two or more of these assays are used. For example, one may measure IL-17 levels and the levels of vaccine-specific antibody. Alternatively, one may follow epidemiological markers such as incidence of, severity of, or duration of pneumococcal infection in vaccinated individuals compared to unvaccinated individuals.

[0123] Vaccine efficacy may also be assayed in various model systems such as the mouse model. For instance, BALB/c or C57BL/6 strains of mice may be used. After

administering the test vaccine to a subject (as a single dose or multiple doses), the experimenter administers a challenge dose of *S. pneumoniae.* In some cases, a challenge dose administered intranasally is sufficient to cause *S. pneumoniae* colonization (especially nasal colonization) in an unvaccinated animal, and in some cases a challenge dose administered via aspiration is sufficient to cause sepsis and a high rate of lethality in unvaccinated animals. One can then measure the reduction in colonization or the reduction in lethality in vaccinated animals. Examples 1-2 and 4 show the efficacy of polypeptides of Table 1 in inhibiting *S. pneumoniae* nasal colonization following intranasal challenge in the mouse model. Examples 5 and 6 show the efficacy of polypeptides of Table 1 in protecting against sepsis and death following infection with *S. pneumoniae* via aspiration in the mouse model.

G. Use of Immunogenic Compositions

1. Defense against S. pneumoniae infection

[0124] The immunogenic compositions of the present disclosure are designed to elicit an immune response against *S. pneumoniae.* Compositions described herein (e.g., ones comprising one or more polypeptides, including fusion proteins, of Table 1, 2 or 3, or nucleic acids encoding the polypeptides) may stimulate an antibody response or a cell-mediated immune response, or both, in the mammal to which it is administered. In some embodiments, the composition stimulates a T_H1-biased $CD4^+$ T cell response, a T_H17-biased $CD4^+$ T cell response and/or a $CD8^+$ T cell response. In some embodiments, the composition stimulates an antibody response. In some embodiments, the composition stimulates a T_H1-biased $CD4^+$ T cell response, T_H17-biased $CD4^+$ T cell response and/or a $CD8^+$ T cell response, and an antibody response.

[0125] In certain embodiments, the composition (e.g., one comprising one or more polypeptides of Table 1, 2 or 3, or nucleic acids encoding the polypeptides, or antibodies reactive with the peptides) includes a cytokine or nucleotide coding region encoding a cytokine such as IL-17, to provide additional stimulation to the immune system of the mammal. In certain embodiments, the composition comprises a cytokine such as IL-17.

[0126] While not wishing to be bound by theory, in some embodiments a T_H17 cell response is desirable in mounting an immune response to the compositions disclosed herein, e.g., ones comprising one or more polypeptides of Table 1, 2 or 3. In certain embodiments,

an active T_H17 response is beneficial in clearing a pneumococcal infection. For instance, mice lacking the IL-17A receptor show decreased whole cell vaccine-based protection from a pneumococcal challenge (Lu *et al.*, "Interleukin-17A mediates acquired immunity to pneumococcal colonization." PLoS Pathog. 2008 Sep 19;4(9)).

[0127] Thus, herein is provided a method of increasing IL-17 production by administering the compositions described herein (e.g., ones comprising one or more polypeptides of Table 1, 2 or 3) to a subject. Furthermore, this application provides a method of activating T_H17 cells by administering said compositions to a subject. In certain embodiments, increased IL-17A levels result in increased pneumococcal killing by neutrophils or neutrophil-like cells, for instance by inducing recruitment and activation of neutrophils of neutrophil-like cells. In certain embodiments, this pneumococcal killing is independent of antibodies and complement. However, specific antibody production and complement activation may be useful additional mechanisms that contribute to clearing of a pneumococcal infection.

[0128] Immunogenic compositions containing immunogenic polypeptides or polynucleotides encoding immunogenic polypeptides together with a pharmaceutical carrier are also provided.

[0129] In some instances, the immunogenic composition comprises one or more nucleic acids encoding one or more polypeptides of SEQ ID NOs: 1-26 and 47-54, such as one or more polypeptides selected from SEQ ID NOs: 47-54. In some embodiments these nucleic acids are expressed in the immunized individual, producing the encoded *S. pneumoniae* antigens, and the *S. pneumoniae* antigens so produced can produce an immunostimulatory effect in the immunized individual.

[0130] Such a nucleic acid-containing immunostimulatory composition may comprise, for example, an origin of replication, and a promoter that drives expression of one or more nucleic acids encoding one or more polypeptides of SEQ ID NOs: 47-54. Such a composition may also comprise a bacterial plasmid vector into which is inserted a promoter (sometimes a strong viral promoter), one or more nucleic acids encoding one or more polypeptides of SEQ ID NOs: 47-54, and a polyadenylation/transcriptional termination sequence. In some instances, the nucleic acid is DNA.

H. Doses and Routes of Administration

1. Dosage forms, amounts, and timing

[0131] The amount of antigen in each vaccine or immunogenic composition dose is selected as an effective amount, which induces a prophylactic or therapeutic response, as described above, in either a single dose or over multiple doses. Preferably, the dose is without significant adverse side effects in typical vaccines. Such amount will vary depending upon which specific antigen is employed. Generally, it is expected that a dose will comprise 1-1000 μg of each protein, in some instances 2-100 μg, for instance 4-40 μg. In some aspects, the vaccine formulation comprises 1-1000 μg of the polypeptide and 1-250 μg of the adjuvant. In some embodiments, the appropriate amount of antigen to be delivered will depend on the age, weight, and health (e.g. immunocompromised status) of a subject. When present, typically an adjuvant will be present in amounts from 1 μg – 250 μg per dose, for example 50-150 μg, 75-125 μg or 100 μg.

[0132] In some embodiments, only one dose of the vaccine is administered to achieve the results described above. In other embodiments, following an initial vaccination, subjects receive one or more boost vaccinations, for a total of two, three, four or five vaccinations. Advantageously, the number is three or fewer. A boost vaccination may be administered, for example, about 1 month, 2 months, 4 months, 6 months, or 12 months after the initial vaccination, such that one vaccination regimen involves administration at 0, 0.5-2 and 4-8 months. It may be advantageous to administer split doses of vaccines which may be administered by the same or different routes.

[0133] The vaccines and immunogenic compositions described herein may take on a variety of dosage forms. In certain embodiments, the composition is provided in solid or powdered (e.g., lyophilized) form; it also may be provided in solution form. In certain embodiments, a dosage form is provided as a dose of lyophilized composition and at least one separate sterile container of diluent.

[0134] In some embodiments, the composition will be administered in a dose escalation manner, such that successive administrations of the composition contain a higher concentration of composition than previous administrations. In some embodiments, the composition will be administered in a manner such that successive administrations of the composition contain a lower concentration of composition than previous administrations.

[0135] In therapeutic applications, compositions are administered to a patient suffering from a disease in an amount sufficient to treat the patient. Therapeutic applications of a composition described herein include reducing transmissibility, slowing disease progression, reducing bacterial viability or replication, or inhibiting the expression of proteins required for toxicity, such as by 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% or 10% of the levels at which they would occur in individuals who are not treated with the composition.

[0136] In prophylactic embodiments, compositions are administered to a human or other mammal to induce an immune response that can inhibit the establishment of an infectious disease or other condition. In some embodiments, a composition may partially block the bacterium from establishing an infection.

[0137] In some embodiments, the compositions are administered in combination with antibiotics. This co-administration is particularly appropriate when the pharmaceutical composition is administered to a patient who has recently been exposed (or is suspected of having been recently exposed) to *S. pneumoniae*. Many antibiotics are used to treat pneumococcal infections, including penicillin, amoxicillin, amoxicillin/clavulanate, cefuroxime, cefotaxime, ceftriaxone, and vancomycin. The appropriate antibiotic may be selected based on the type and severity of the infection, as well as any known antibiotic resistance of the infection (Jacobs MR "Drug-resistant *Streptococcus pneumoniae*: rational antibiotic choices" Am J Med. 1999 May 3;106(5A):19S-25S).

2. Routes of administration

[0138] The vaccine formulations and pharmaceutical compositions herein can be delivered by administration to an individual, typically by systemic administration (e.g., intravenous, intraperitoneal, intramuscular, intradermal, subcutaneous, subdermal, transdermal, intracranial, intranasal, mucosal, anal, vaginal, oral, buccal route or they can be inhaled) or they can be administered by topical application. In some embodiments, the route of administration is intramuscular. In other embodiments, the route of administration is subcutaneous. In yet other embodiments, the route of administration is mucosal. In certain embodiments, the route of administration is transdermal or intradermal.

[0139] Certain routes of administration are particularly appropriate for vaccine formulations and immunogenic compositions comprising specified adjuvants. In particular, transdermal administration is one suitable route of administration for *S. pneumoniae* vaccines comprising toxins (e.g. cholera toxin or labile toxin); in other embodiments, the

administration is intranasal. Vaccines formulated with Alphavirus replicons may be administered, for example, by the intramuscular or the subcutaneous route. Vaccines comprising Monophosphory Lipid A (MPL), Trehalose Dicoynomycolate (TDM), and dioctadecyldimethylammonium bromide (DDA) are suitable (*inter alia)* for intramuscular and subcutaneous administration. A vaccine comprising resiquimod may be administered topically or subcutaneously, for example.

3. Formulations

[0140] The vaccine formulation or immunogenic composition may be suitable for administration to a human patient, and vaccine or immunogenic composition preparation may conform to USFDA guidelines. In some embodiments, the vaccine formulation or immunogenic composition is suitable for administration to a non-human animal. In some embodiments, the vaccine or immunogenic composition is substantially free of either endotoxins or exotoxins. Endotoxins may include pyrogens, such as lipopolysaccharide (LPS) molecules. The vaccine or immunogenic composition may also be substantially free of inactive protein fragments which may cause a fever or other side effects. In some embodiments, the composition contains less than 1%, less than 0.1%, less than 0.01%, less than 0.001%, or less than 0.0001% of endotoxins, exotoxins, and/or inactive protein fragments. In some embodiments, the vaccine or immunogenic composition has lower levels of pyrogens than industrial water, tap water, or distilled water. Other vaccine or immunogenic composition components may be purified using methods known in the art, such as ion-exchange chromatography, ultrafiltration, or distillation. In other embodiments, the pyrogens may be inactivated or destroyed prior to administration to a patient. Raw materials for vaccines, such as water, buffers, salts and other chemicals may also be screened and depyrogenated. All materials in the vaccine may be sterile, and each lot of the vaccine may be tested for sterility. Thus, in certain embodiments the endotoxin levels in the vaccine fall below the levels set by the USFDA, for example 0.2 endotoxin (EU)/kg of product for an intrathecal injectable composition; 5 EU/kg of product for a non-intrathecal injectable composition, and 0.25-0.5 EU/mL for sterile water.

[0141] In certain embodiments, the preparation comprises less than 50%, 20%, 10%, or 5% (by dry weight) contaminating protein. In certain embodiments, the desired molecule is present in the substantial absence of other biological macromolecules, such as other proteins (particularly other proteins which may substantially mask, diminish, confuse or alter the characteristics of the component proteins either as purified preparations or in their

function in the subject reconstituted mixture). In certain embodiments, at least 80%, 90%, 95%, 99%, or 99.8% (by dry weight) of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). In some embodiments, the vaccine or immunogenic composition comprising purified subunit proteins contains less than 5%, 2%, 1%, 0.5%, 0.2%, 0.1% of protein from host cells in which the subunit proteins were expressed, relative to the amount of purified subunit. In some embodiments, the desired polypeptides are substantially free of nucleic acids and/or carbohydrates. For instance, in some embodiments, the vaccine or immunogenic composition contains less than 5%, less than 2%, less than 1%, less than 0.5%, less than 0.2%, or less than 0.1% host cell DNA and/or RNA. In certain embodiments, at least 80%, 90%, 95%, 99%, or 99.8% (by dry weight) of biological macromolecules of the same type are present in the preparation (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present).

[0142] It is preferred that the vaccine or immunogenic composition has low or no toxicity, within a reasonable risk-benefit ratio. In certain embodiments, the vaccine or immunogenic composition comprises ingredients at concentrations that are less than LD_{50} measurements for the animal being vaccinated. LD_{50} measurements may be obtained in mice or other experimental model systems, and extrapolated to humans and other animals. Methods for estimating the LD_{50} of compounds in humans and other animals are well-known in the art. A vaccine formulation or immunogenic composition, and any component within it, might have an LD_{50} value in rats of greater than 100 g/kg, greater than 50g/kg, greater than 20 g/kg, greater than 10 g/kg, greater than 5 g/kg, greater than 2 g/kg, greater than 1 g/kg, greater than 500 mg/kg, greater than 200 mg/kg, greater than 100 mg/kg, greater than 50 mg/kg, greater than 20 mg/kg, or greater than 10 mg/kg. A vaccine formulation or immunogenic composition that comprises a toxin such as botulinum toxin (which can be used as an adjuvant) should contain significantly less than the LD_{50} of botulinum toxin.

[0143] The formulations suitable for introduction of the vaccine formulations or pharmaceutical composition vary according to route of administration. Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, intranasal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the

blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampoules and vials.

[0144] Injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. In the case of adoptive transfer of therapeutic T cells, the cells can be administered intravenously or parenterally.

[0145] Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the polypeptides or packaged nucleic acids suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, tragacanth, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art. The pharmaceutical compositions can be encapsulated, e.g., in liposomes, or in a formulation that provides for slow release of the active ingredient.

[0146] The antigens, alone or in combination with other suitable components, can be made into aerosol formulations (e.g., they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. Aerosol formulations can be delivered orally or nasally.

[0147] Suitable formulations for vaginal or rectal administration include, for example, suppositories, which consist of the polypeptides or packaged nucleic acids with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a

combination of the polypeptides or packaged nucleic acids with a base, including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

I. Preparation and Storage of Vaccine Formulations and Immunogenic Compositions

[0148] The *S. pneumoniae* vaccines and immunogenic compositions described herein may be produced using a variety of techniques. For example, a polypeptide may be produced using recombinant DNA technology in a suitable host cell. A suitable host cell may be bacterial, yeast, mammalian, or other type of cell. The host cell may be modified to express an exogenous copy of one of the relevant polypeptide genes. Typically, the gene is operably linked to appropriate regulatory sequences such as a strong promoter and a polyadenylation sequence. In some embodiments, the promoter is inducible or repressible. Other regulatory sequences may provide for secretion or excretion of the polypeptide of interest or retention of the polypeptide of interest in the cytoplasm or in the membrane, depending on how one wishes to purify the polypeptide. The gene may be present on an extrachromosomal plasmid, or may be integrated into the host genome. One of skill in the art will recognize that it is not necessary to use a nucleic acid 100% identical to the naturally-occurring sequence. Rather, some alterations to these sequences are tolerated and may be desirable. For instance, the nucleic acid may be altered to take advantage of the degeneracy of the genetic code such that the encoded polypeptide remains the same. In some embodiments, the gene is codon-optimized to improve expression in a particular host. The nucleic acid may be produced, for example, by PCR or by chemical synthesis.

[0149] Once a recombinant cell line has been produced, a polypeptide may be isolated from it. The isolation may be accomplished, for example, by affinity purification techniques or by physical separation techniques (e.g., a size column).

[0150] In a further aspect of the present disclosure, there is provided a method of manufacture comprising mixing one or more polypeptides or an immunogenic fragment or variant thereof with a carrier and/or an adjuvant.

[0151] In some embodiments, antigens for inclusion the vaccine formulations and immunogenic compositions may be produced in cell culture. One method comprises providing one or more expression vectors and cloning nucleotides encoding one or more

polypeptides selected from polypeptides having an amino acid sequence of Table 1, 2 or 3, such as SEQ ID NO: 47-54, then expressing and isolating the polypeptides.

[0152] The immunogenic polypeptides described herein, and nucleic acid compositions that express the polypeptides, can be packaged in packs, dispenser devices, and kits for administering nucleic acid compositions to a mammal. For example, packs or dispenser devices that contain one or more unit dosage forms are provided. Typically, instructions for administration of the compounds will be provided with the packaging, along with a suitable indication on the label that the compound is suitable for treatment of an indicated condition, such as those disclosed herein.

[0153] The disclosure is further illustrated by the following examples. The examples are provided for illustrative purposes only. They are not to be construed as limiting the scope or content of the disclosure in any way.

V. Examples

Example 1

SP2108, SP0148 and SP1634 polypeptides

[0154] The SP2108 polypeptide (SEQ ID NO: 6), SP0148 polypeptide (SEQ ID NO: 2) and SP1634 polypeptide (SEQ ID NO: 25) were formulated as vaccine compositions using 4 μg of the polypeptide in combination with 1 μg cholera toxin adjuvant (CT). For combinations, 4 μg of each polypeptide was used. The compositions were administered intranasally to C57BL/6 mice three times, one week apart. The subjects were then allowed to rest for 3 weeks, and bled at that time for immunogenicity. For this assay, heparinized whole blood was collected from the retrograde orbital sinus. The total PBMC were stimulated with either killed, unencapsulated whole cell *S. pneumoniae* (WCC) or a combination of the three polypeptides in round bottomed tubes for three days. The supernatants were then harvested and evaluated by ELISA for IL-17 levels. Cholera toxin alone (CT) or an unrelated antigen from HSV (003) were used as negative controls. Results of the IL-17 immunogenicity assay are shown in FIGS. 1 and 2, where the left panels show data in scatter format, and the right panels show data as averages with standard deviations. The subjects were allowed to rest an additional 2 weeks, at which time they were challenged with intranasal administration of live, encapsulated *S. pneumoniae.* The subjects were sacrificed a week later, and the number of

colony-forming units (CFU) was counted from nasal washes. Results of the colonization assay are shown in FIG. 3.

Example 2

SP0148, SP0314, SP0882, and SP2108 polypeptides tested in the BALB/c mouse

[0155] To determine whether similar immune responses were seen across different mouse genotypes, vaccine compositions were administered to BALB/c mice. Vaccine compositions comprised the polypeptides SP0148 (SEQ ID NO: 2), SP0314 (SEQ ID NO: 26), SP0882 (SEQ ID NO: 17) or SP2108 (SEQ ID NO: 6), and cholera toxin adjuvant (CT). Using a protocol similar to that of Example 1, the mice were immunized, challenged intranasally with *S. pneumoniae*, and the number of CFU was recorded. The results of this colonization experiment are shown in FIG. 4.

Example 3

SP1912, SP2108 and SP0148 polypeptides: IL-17A immunogenicity assay

[0156] The polypeptides SP1912 (SEQ ID NO: 9), SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) were formulated as vaccine compositions with cholera toxin adjuvant (CT). The vaccine compositions were administered to mice two times, one week apart. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + CT (WCB), and the negative controls were CT alone or recombinant proteins without CT (with the exception of SP1912). Three weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured in duplicate with A) the protein of immunization, or B) the whole cell vaccine for six days. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 5. Each symbol in the graph represents responses from individual mice, and the line indicates the median response of the group.

Example 4

SP1912, SP2108 and SP0148 polypeptides: colonization assay

[0157] Animals were immunized with vaccine formulations comprising the polypeptides SP1912 (SEQ ID NO:9), SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) and cholera toxin adjuvant (CT) as described in Example 3, and then challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization (and one week after retroorbital blood collection). Seven days after challenge, animals were euthanized and

the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results are shown in FIG. 6 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. (*** = p-value <0.05).

Example 5

SP1912 polypeptide: aspiration challenge (sepsis assay)

[0158] Polypeptide SP1912 was evaluated for its ability to protect mice from sepsis. Groups of ten mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising either the SP1912 polypeptide (SEQ ID NO: 9) or pneumolysoid (PdT) adsorbed to alum. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + alum (WCB), and the negative control was alum alone. Three weeks after the final immunization, blood was collected for evaluation of IL-17A response and antibody levels, and then one week later, the mice underwent aspiration challenge with 10^7 live strain 0603 (type 6B) *S. pneumoniae.* Animals were monitored for survival for eight days. Results of the aspiration challenge are shown in FIG. 7 as survival curves for each immunized group.

Example 6

Pneumolysoid PdT, SP0148 and SP0641N polypeptides: aspiration challenge (sepsis assay)

[0159] Polypeptide SP0148 was evaluated for its ability to protect mice from sepsis when immunized singly or in combination with SP0641N and/or pneumolysoid (PdT). Groups of ten mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising polypeptide SP0148 (SEQ ID NO: 2), singly or in combination with polypeptide SP0641N (SEQ ID NO: 13) and/or PdT, adsorbed to alum. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + alum (WCB), and the negative control was alum alone. Three weeks after the final immunization, blood was collected for evaluation of IL-17 and antibody, and then one week later, the mice underwent aspiration challenge with 10^7 live strain 0603 (type 6B) *S. pneumoniae.* Animals were monitored for survival for eight days. The data are shown in FIG. 8 as survival curves for each immunized group.

Example 7

SP1912, SP2108 and SP0148 polypeptides: colonization assay

[0160] Additional studies were performed essentially as described in Example 4, for a total of four separate studies. Briefly, animals were immunized with vaccine formulations comprising the polypeptides SP1912 (SEQ ID NO: 9), SP2108 (SEQ ID NO: 6), SP0148 (SEQ ID NO: 2), or additionally SP2108 plus SP0148, and cholera toxin adjuvant (CT) as described in Example 3. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* plus CT (WCV), or CT alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Pooled results of four studies are shown in FIG. 9 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. (*** = p-value <0.05). N indicates the total number of animals evaluated. Percentages refer to the number of animals protected from colonization.

Example 8

SP1912 and SP0148 polypeptides: IL-17A immunogenicity assay

[0161] Groups of ten mice were subcutaneously immunized twice, two weeks apart with vaccine compositions comprising either SP1912 polypeptide (SEQ ID NO: 9), SP0148 polypeptide (SEQ ID NO: 2), or both adsorbed to alum. Control animals were immunized with alum alone. Three weeks after the last immunization, heparinized blood was collected by cardiac puncture and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with the protein(s) of immunization. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 10. Each symbol in the graph represents responses from individual mice, and the line indicates the median response of the group.

Example 9

SP1912 and SP0148 polypeptides: colonization assay

[0162] Animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2) at different doses plus and minus SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were

immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results are shown in FIG. 11 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The number of animals protected from colonization out of the number of animals in the group is indicated at the top of the figure.

Example 10

SP1912, SP0148, and SP2108 polypeptides: colonization assay

[0163] In two separate studies, animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2) and SP0148 plus SP1912 (SEQ ID NO: 9), or additionally with SP2108 (SEQ ID NO: 6), SP2108 plus SP0148, and SP2108 plus SP1912, adsorbed to alum. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Pooled results of the two studies are shown in FIG. 12 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The number of animals protected from colonization out of the number of animals in the group and corresponding percentage of animals protected from colonization are indicated at the top of the figure. (*$p<0.05$, **$p<0.01$, ***$p<0.001$ Dunn's Multiple Comparison Test compared to Alum control)

Example 11

Pneumolysoid L460D, PspA derivative CD2, SP1912, SP0148, and SP2108 polypeptides: colonization assay

[0164] Animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2), SP2108 (SEQ ID NO: 6), SP0148 plus SP2108, and SP0148 plus SP2108 in combination with SP1912 (SEQ ID NO: 9) or known *S. pneumoniae* antigens L460D plus CD2, adsorbed to alum. Two

separate studies were conducted. Control animals were immunized with alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results of the second study are shown in FIG. 13 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The number of animals protected from colonization out of the number of animals in the group is indicated at the top of the figure.

[0165] The chart below shows the absolute number and corresponding percentage of animals protected from colonization in the four studies described in Examples 10 and 11.

	1	2	3	4	# not colonized/ total	% not colonized
Alum	0/9	1/10	2/10	4/10	7/39	18%
WCV	9/9	9/10	9/10	9/10	36/39	92%
0148	8/18	4/10	5/9	1/10	18/47	38%
2108		8/10	6/10	4/10	18/30	60%
1912		4/10			4/10	40%
0148 + 2108		6/10	3/10	6/10	15/30	50%
0148 + 1912	13/18	3/10			16/28	57%
2108 + 1912		3/10			3/10	30%
0148 + 2108 + 1912		8/9	8/10	6/10	22/29	76%
0148 + 2108 + L460D + CD2			2/10	6/10	8/20	40%

Example 12

PspA, SP0148 and SP2108 passive antibody transfer and aspiration challenge (sepsis assay)

[0166] Groups of ten mice were injected with monoclonal antibodies specific for PspA, heat-inactivated rabbit sera specific for SP0148, SP2108, or combinations of these. Antibody and antisera concentrations and total injection volumes were adjusted with normal rabbit serum (NRS) and PBS. Control animals were injected with NRS, or serum against killed, unencapsulated whole cell *S. pneumoniae* (WCV). One day after injection, the mice underwent aspiration challenge with 10^6 live *S. pneumoniae* type WU-2 (ST-3). Animals

were monitored for survival for eight days. The data are shown in FIG. 14 as survival curves for each immunized group.

[0167] FIG. 15 shows the percent of animals protected from sepsis in six separate aspiration challenge studies, two of which are described in more detail in Examples 6 and 12.

Example 13

SP0148, SP2108, and SP1912 fusion proteins: IL-17A immunogenicity assay

[0168] Groups of four mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising the fusion proteins SP2108/0148/1912 (SEQ ID NO: 53), SP0148/2108 (SEQ ID NO: 48), SP0148/1912 (SEQ ID NO: 50) or the comparable combinations of the single antigens SP2108 (SEQ ID NO: 6), SP0148 (SEQ ID NO: 2) and SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with alum alone. Two weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with single (non-fusion) proteins of immunization. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 16, Panels A-C. Each bar represents the mean ± SD IL-17A response from each group of four mice in response to the indicated antigen.

Example 14

SP0148, SP2108, and SP1912 fusion proteins: IL-17A immunogenicity assay

[0169] Groups of four to eight mice were subcutaneously immunized three times, two weeks apart, with vaccine compositions comprising the fusion proteins SP0148/1912/2108 (SEQ ID NO: 52) or SP2108/1912 (SEQ ID NO: 49), adsorbed to alum. Control animals were immunized with alum alone. Three weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with the protein of immunization, a control protein, or media alone. The cell-free supernatants were harvested and IL-17A concentrations measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 17, Panels A and B. Each symbol represents the IL-17A response of an individual mouse, and the line indicates the median response of the group.

Example 15

SP2108 and SP1912 fusion proteins: Frequency of T cell response

[0170] Spleens were collected at the same time as peripheral blood from mice, shown in Example 14, that were immunized with the vaccine composition comprising fusion protein SP2108/1912 (SEQ ID NO: 49), or alum alone. Single cell suspensions were prepared from the spleens. Erythrocytes were lysed, and splenocytes were plated onto ELISpot plates coated with anti-IL-17A capture antibody at 200-800 cells per well and stimulated with overlapping peptides (OLP) spanning SP2108 or SP1912, killed unencapsulated whole-cell *S. pneumoniae* (WCA), or media alone. After two days of incubation, the cells were removed and IL-17A visualized by labeling with biotinylated anti-IL-17A antibody and a horseradish peroxidase substrate. Each colored spot represents a cell that secreted IL-17A. Results of the assay are shown in FIG. 18 as the number of spot forming units (SFU) per 10^6 splenocytes. Each symbol represents the spot frequency from an individual mouse, and the line indicates the median response of the group.

Example 16

SP0148, SP2108, SP1912 and fusion proteins: IL-17A dose response assay

[0171] Groups of four mice were subcutaneously immunized three times, two weeks apart, with vaccine compositions comprising 30, 100, or 300 pmol of the fusion proteins SP0148/1912/2108 (SEQ ID NO: 52), SP0148/2108/1912 (SEQ ID NO: 54), SP2108/0148/1912 (SEQ ID NO: 53), SP0148/1912 (SEQ ID NO: 50), SP2108/1912 (SEQ ID NO: 49), or indicated admixtures of SP0148 (SEQ ID NO: 2), SP2108 (SEQ ID NO: 6), and SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with alum alone. Two weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with overlapping peptides (OLP) spanning SP0148, SP1912 or SP2108, killed unencapsulated whole-cell *S. pneumoniae* (WCA), or media alone. The cell-free supernatants were harvested and IL-17A concentrations measured by ELISA. Results of the IL-17A dose response assays are shown in FIG. 19, Panels A and B. Each bar represents the mean specific IL-17A response of a group of mice with the media-induced (non-specific) background subtracted.

Example 17

SP0148, SP2108, SP1912, and fusion proteins: IL-17A immunogenicity assay

[0172] Groups of ten mice were subcutaneously immunized three times, two weeks apart, with vaccine compositions comprising the fusion proteins SP2108/1912 (SEQ ID NO: 49) or SP0148/1912/2108 (SEQ ID NO: 52), or indicated admixtures of SP0148 (SEQ ID NO: 2), SP2108 (SEQ ID NO: 6), and SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with alum alone. Three weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with killed unencapsulated whole-cell *S. pneumoniae* (WCA). The cell-free supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 20. Each symbol represents the IL-17A response of an individual mouse, and the line indicates the median response of the group.

Example 18

SP0148, SP2108, and SP1912 fusion proteins: intranasal challenge and colonization assay

[0173] Animals were subcutaneously immunized three times, two weeks apart at the indicated doses with vaccine formulations comprising the fusion proteins SP2108/1912 (SEQ ID NO: 49), SP0148/1912 (SEQ ID NO: 50), SP0148/2108/1912 (SEQ ID NO: 54), SP0148/1912/2108 (SEQ ID NO: 52), or SP2108/0148/1912 (SEQ ID NO: 53), or an admixture of SP0148 (SEQ ID NO: 2), SP2108 (SEQ ID NO: 6) and SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with alum alone. Immunized animals were challenged intranasally with 10^6 to 10^7 colony forming units (CFU) live strain 0603M20 *S. pneumoniae* twenty-five days after the last immunization. Ten days after challenge, animals were euthanized, the nasopharynx lavaged (nasal wash) and the fluid cultured on permissive media to evaluate the *S. pneumoniae* titers. Results of the colonization assays are shown in FIG. 21, panels A-D, as CFU per nasal wash. Each symbol represents the colonization density from an individual mouse, and the horizontal line represents the median of the group. In panel A, stars indicate statistical significance by Mann-Whitney U test; the absolute number of mice protected and corresponding median log reduction in bacterial colonization are also indicated. (* $p<0.05$; ** $p<0.005$; *** $p<0.001$; **** $p<0.0005$)

Example 19

SP0148, SP2108, and SP1912 fusion proteins: aspiration challenge and sepsis assay

[0174] Groups of ten mice are subcutaneously immunized three times, two weeks apart with vaccine compositions comprising the fusion proteins SP2108/0148 (SEQ ID NO: 47), SP0148/2108 (SEQ ID NO: 48), SP2108/1912 (SEQ ID NO: 49), SP0148/1912 (SEQ ID NO: 50), SP2108/1912/0148 (SEQ ID NO: 51), SP0148/1912/2108 (SEQ ID NO: 52), SP2108/0148/1912 (SEQ ID NO: 53), or SP0148/2108/1912 (SEQ ID NO: 54), adsorbed to alum. Control animals are immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Three weeks after the final immunization, blood is collected for evaluation of IL-17A and antibody, and then one week later, the mice undergo aspiration challenge with 10^7 live strain WU-2 (serotype 3) *S. pneumoniae.* Animals are monitored for survival for eight days. Results are reported as survival curves for each immunized group.

Example 20

Fusion protein SP2108/1912 plus B cell antigens: IL-17A immunogenicity assay

[0175] Groups of twenty mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising the fusion protein SP2108/1912 (SEQ ID NO: 49) plus either the B cell antigens YLN and H70 (SEQ ID NO: 33), or the B cell antigen PRN-L460D, adsorbed to alum. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* (WCB), or alum alone. Two to three weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparizined whole blood was diluted in media and then cultured for six days with overlapping peptides (OLP) spanning SP2108 or SP1912, or killed unencapsulated whole-cell *S. pneumoniae* (WCA). The cell-free supernatants were harvested and IL-17A concentrations measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 22, Panels A and B (duplicate experiments). Each symbol represents the IL-17A response of an individual mouse, and the line indicates the median response of the group.

Example 21

Fusion protein SP2108/1912 plus B cell antigens: intranasal challenge followed by colonization assay

[0176] Groups of twenty mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising killed, unencapsulated whole cell *S. pneumoniae* (WCB), fusion protein SP2108/1912 (SEQ ID NO: 49) plus either the B cell

antigens YLN and H70 (SEQ ID NO: 33), or the B cell antigen PRN-L460D, adsorbed to alum. Control animals were immunized with alum alone. Immunized animals were challenged intranasally with 10^7 live strain 0603M20 *S. pneumoniae* three to five weeks after the last immunization. Ten days after challenge, animals were euthanized and the nasopharynx lavaged (nasal wash) and the fluid cultured on permissive media to evaluate the *S. pneumoniae* titers. Results of the colonization assays are shown in FIG. 23, Panels A and B (duplicate experiments), as colony forming units of bacteria (CFU) per nasal wash. Each symbol represents the colonization density from an individual mouse, and the horizontal line represents the median of the group.

Example 22

Fusion protein SP2108/1912 plus B cell antigens: intravenous challenge followed by sepsis assay

[0177] Groups of twenty mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising B cell antigens YLN plus H70 (SEQ ID NO: 33), fusion protein SP2108/1912 (SEQ ID NO: 49) plus B cell antigens YLN and H70 (SEQ ID NO: 33), or fusion protein SP2108/1912 (SEQ ID NO: 49) plus B cell antigen PRN-L460D, adsorbed to alum. Control animals were immunized with alum alone. Immunized animals with injected intravenously with 3.6×10^6 live strain DBL6A *S. pneumoniae* three weeks after last immunization. Animals were monitored for survival for twenty-one days. Results are shown in FIG. 24, Panels A and B (duplicate experiments), as hours elapsed until each mouse reaches a moribund state or death. Each symbol represents the response of an individual mouse, and the line indicates the median response of the group.

Equivalents

[0178] It is to be understood that while the disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

SEQUENCES

SEQ ID NO: 1
SP0148 lacking signal sequence
MCSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSDKYDVKFEKTEWSGVFAGLDADR
YNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPT
ILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRI
KELYKDGTLEKLSKQFFGDTYLPAEADIKE

SEQ ID NO: 2
SP0148 including signal sequence (277 amino acids with N-terminal E)
MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSD
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKE

SEQ ID NO: 3
SP0148 consensus lacking signal sequence
MCSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSDKYDVKFE
Q S R N N X

KTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDD
I E

IGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADLQQIMVRLSDGQFDYKIFDKIG
F

VETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQ
Y S

FFGDTYLPAEADIK(E)

SEQ ID NO: 4
SP0148 consensus including signal sequence
MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELT
G L Q S R N

GYEIEVVRAIFKDSDKYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAP
N X I

IAQNPNVLVVKKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKAD
E

LQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELK
F Y S

SFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIK(E)

SEQ ID NO: 5
SP2108 lacking signal sequence
MCGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVM
MAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLA
KDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQ
DTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQK
FVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSG
QKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 6
SP2108 including signal sequence
>gi|14973620|gb|AAK76167.1| maltose/maltodextrin ABC transporter,
maltose/maltodextrin-binding protein (Streptococcus pneumoniae TIGR4)
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL
PNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 7
SP2108 consensus lacking signal sequence
MCGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLD
A I

KLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGA
I X T

PAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGL
A

LAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQE
G P A X H

GKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQK
A

FVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSA
S A S

VWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 8
SP2108 consensus including signal sequence
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVA
T T V A

KAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLS
I I X

DGAKTDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAF
T

AGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWY
A G P A X

EKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
H

AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDE
A S A

LTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQK
S

FGE

SEQ ID NO: 9
SP1912

MNGMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGV
IFEDGRHYTFVYENEDLVYEEEVL

SEQ ID NO: 10
SP1912L

MRYLATLLLSLAVLITAGCKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIFEDGRH
YTFVYENEDLVYEEEVL

SEQ ID NO: 11
SP1912 consensus
MNGMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGV
IF
H A L I L S

EDGRHYTFVYENEDLVYEEEVL
I

SEQ ID NO: 12
SP0641
>gi|14972117|gb|AAK74791.1| serine protease, subtilase family
[Streptococcus pneumoniae TIGR4]
MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKKTAVIKEK
EVVSKNPVIDNNTSNEEAKIKEENSNKSQGDYTDSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNT
KVLYTYDRIFNGSAIETTPDNLDKIKQIEGISSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRG
MVISNIDTGTDYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDYFDPHG
MHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIEDSIKHNVDVVSVSSGFTGTG
LVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVANNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFD
KVNIGGESFKYRNIGAFFDKSKITTNEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKK
AMDKGARAIMVVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDK
LEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRIDLLLKPDVSAPGKNIKSTLN
VINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKE
KSQYFASPRQQGAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSA
SAITTDSLTDRLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAKNKNKF
VESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLGGYDDDGKPKIPGTL
NKGIGGEHGIDKFNPAGVIQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGL
TPSPLVLRSAEEGLISIVNTNKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRG
REENAPESKDNQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLIT
KDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQGRNRKLDKDGNTI

YEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANQADEKGMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGK
EGSLKKDTTGVEHHHQENEESIKEKSSFTIDRNISTIRDFENKDLKKLIKKKFREVDDFTSETGKRMEEYDYKYD
DKGNIIAYDDGTDLEYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATKYDFHS
KTMTFDLYANINDIVDGLAFAGDMRLFVKDNDQKKAEIKIRMPEKIKETKSEYPYVSSYGNVIELGEGDLSKNKP
DNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYNPGKTDMLEGNGVYSKEDIAKIQKANPNLRALSETT
IYADSRNVEDGRSTQSVLMSALDGFNIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKFNV
EAIKEDGSMLFIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIYGDLHI
DNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAVGVHYQFLYDNVKPEVNIDPK
GNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYINGKEYTSFNDIKQIIDKTLNIKIVVKDFARNTTVKEFIL
NKDTGEVSELKPHRVTVTIQNGKEMSSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTF
IKPVFKKIEEKKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREEHSQKSDSTKDVTATVLDKNNISSKST
TNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD

SEQ ID NO: 13
SP0641N
MVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAKIKE
ENSNKSQGDYTDSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDNL
DKIKQIEGISSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGTDYRHKAMRIDDD
AKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGI
DGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIEDSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVV
ATGNYATSASSSSWDLVANNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKSK
ITTNEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIMVVNTVNYYNRDN
WTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDKLEQYYPIDMESFNSNKPNVGDE
KEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRIDLLLKPDVSAPGKNIKSTLNVINGKSTYG

SEQ ID NO: 14
SP0641M
MSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQ
QGAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD
RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAKNKNKFVESFIHFESV
EEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLGGYDDDGKPKIPGTLNKGIGGEHGI
DKFNPAGVIQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSA
EEGLISIVNTNKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD
NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLITKDTYHKVKDQ
YKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQGRNRKLDKDGNTIYEIKGAGDLR
GKIIEVIALDGSSNFTKIHRIKFANQADEKGMISYYLVDPDQDSSKYQ

SEQ ID NO: 15
SP641N consensus

MVVLADTSSSEDALNISDKEKVA-----
ENKEKHENIHSAMETSQDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAK - N S
VVDKET KD N I K TE TI EG A T TK R
L

IKEENSNKSQGDYTDSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGSAIETTP
DN
D- Q H Q S Q N G Q
NAH SA G RL G

LDKIKQIEGISSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGTDYRHKAMRIDD
DA
T I

KASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGID
GI

APNAQIFSYKMYSDAGSGFAGDETMFHAIEDSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATG
NY

ATSASSSSWDLVANNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKSKITTNE
DG
Q N

TKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIMVVNTVNYYNRDNWTELPAM
GY

EADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAP
DT
N

DKELYKEDIIVPAGSTSWGPRIDLLLKPDVSAPGKNIKSTLNVINGKSTYG

SEQ ID NO: 16
SP641M consensus

MSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQ
QG K
T

AGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTDRL
KL V

DETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAKNKNKFVESFIHFESVEEME
AL
Y R A

NSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPA
GV
S TD K ME

IQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIV
NT
R D D Q VH E T

NKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKDNQDPATKIRG
QF
K V G

EPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLITKDTYHKVKDQYKNETLFARDQK
EH

PEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSN
FT S A

KIHRIKFANQADEKGMISYYLVDPDQDSSKYQ
DH K A E

SEQ ID NO: 17
SP0882
>gi|14972356|gb|AAK75009.1| conserved hypothetical protein (Streptococcus pneumoniae TIGR4)

MNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNP
DISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAEFVMEVVKPFIDETYRTKADCQHTAMIGSSL
GGNITQFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFIYVGTEEADDTDKTLMDGNIKQAYID
SSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 18
SP0882N
MNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNP
DISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAEFVMEVVKPFI

SEQ ID NO: 19
SP0882 with exogenous signal sequence
MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDG
QNVFNSKESFIGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAEFVMEV
VKPFIDETYRTKADCQHTAMIGSSLGGNITQFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFI
YVGTEEADDTDKTLMDGNIKQAYIDSSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 20
SP0882N with exogenous signal sequence
MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDG
QNVFNSKESFIGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAEFVMEV
VKPFI

SEQ ID NO: 21
SP0882 consensus
MNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESF
I Y

IGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAE
Y H E E

FVMEVVKPFIDETYRTKADCQHTAMIGSSLGGNITQFIGLEYQDQIGCLGVFSSANWLHQ
EK

EAFNRYFECQKLSPDQRIFIYVGTEEADDTDKTLMDGNIKQAYIDSSLCYYHDLIAGGVH
I H R

LDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 22
SP0882N consensus
MNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESF
I Y

IGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNEYAAWKFQESPIPGQQFGGKGVEYAE
Y H E E

FVMEVVKPFI

SEQ ID NO: 23
SP0882 consensus with exogenous signal sequence
MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYE
T T V I

KDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNE
Y Y H

YAAWKFQESPIPGQQFGGKGVEYAEFVMEVVKPFIDETYRTKADCQHTAMIGSSLGGNIT
E E

QFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFIYVGTEEADDTDKTLM
EK I H

DGNIKQAYIDSSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW
R

SEQ ID NO: 24
SP0882N consensus with exogenous signal sequence
MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLKMKEHKLKVPYTGKERRVRILLPKDYE
T T V I

KDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNPDISRMIVVAIDNDGMGRMNE
Y Y H

YAAWKFQESPIPGQQFGGKGVEYAEFVMEVVKPFI

E E

SEQ ID NO: 25
SP1634
>gi|14973124|gb|AAK75714.1| hypothetical protein SP_1634 Streptococcus pneumoniae TIGR4
MANIFDYLKDVAYDSYYDLPLNELDILTLIEITYLSFDNLVSTLPQRLLDLAPQVPRDPTMLTSKNRLQLLDELA
QHKRFKNCKLSHFINDIDPELQKQFAAMTYRVSLDTYLIVFRGTDDSIIGWKEDFHLTYMKEIPAQKHALRYLKN
FFAHHPKQKVILAGHSKGGNLAIYAASQIEQSLQNQITAVYTFDAPGLHQELTQTAGYQRIMDRSKIFIPQGSII
GMMLEIPAHQIIVQSTALGGIAQHDTFSWQIEDKHFVQLDKTNSDSQQVDTTFKEWVATVPDEELQLYFDLFFGT
ILDAGISSINDLASLKALEYIHHLFVQAQSLTPEERETLGRLTQLLIDTRYQAWKNR

SEQ ID NO: 26
SP0314
>gi|14971788|gb|AAK74491.1| hyaluronidase Streptococcus pneumoniae TIGR4MQTKTKKLIVSLSSLVLSGFLLNHYMTIGAEETTTNTIQQSQKEVQYQQRDTKNLVENGDFGQTEDGSSP
WTGSKAQGWSAWVDQKNSADASTRVIEAKDGAITISSHEKLRAALHRMVPIEAKKKYKLRFKIKTDNKIGIAKVR
IIEESGKDKRLWNSATTSGTKDWQTIEADYSPTLDVDKIKLELFYETGTGTVSFKDIELVEVADQLSEDSQTDKQ
LEEKIDLPIGKKHVFSLADYTYKVENPDVASVKNGILEPLKEGTTNVIVSKDGKEVKKIPLKILASVKDAYTDRL
DDWNGIIAGNQYYDSKNEQMAKLNQELEGKVADSLSSISSQADRTYLWEKFSNYKTSANLTATYRKLEEMAKQVT
NPSSRYYQDETVVRTVRDSMEWMHKHVYNSEKSIVGNWWDYEIGTPRAINNTLSLMKEYFSDEEIKKYTDVIEKF
VPDPEHFRKTTDNPFKALGGNLVDMGRVKVIAGLLRKDDQEISSTIRSIEQVFKLVDQGEGFYQDGSYIDHTNVA
YTGAYGNVLIDGLSQLLPVIQKTKNPIDKDKMQTMYHWIDKSFAPLLVNGELMDMSRGRSISRANSEGHVAAVEV
LRGIHRIADMSEGETKQCLQSLVKTIVQSDSYYDVFKNLKTYKDISLMQSLLSDAGVASVPRPSYLSAFNKMDKT
AMYNAEKGFGFGLSLFSSRTLNYEHMNKENKRGWYTSDGMFYLYNGDLSHYSDGYWPTVNPYKMPGTTETDAKRA
DSDTGKVLPSAFVGTSKLDDANATATMDFTNWNQTLTAHKSWFMLKDKIAFLGSNIQNTSTDTAATTIDQRKLES
GNPYKVYVNDKEASLTEQEKDYPETQSVFLESFDSKKNIGYFFFKKSSISMSKALQKGAWKDINEGQSDKEVENE
FLTISQAHKQNRDSYGYMLIPNVDRATFNQMIKELESSLIENNETLQSVYDAKQGVWGIVKYDDSVSTISNQFQV
LKRGVYTIRKEGDEYKIAYYNPETQESAPDQEVFKKLEQAAQPQVQNSKEKEKSEEEKNHSDQKNLPQTGEGQSI
LASLGFLLLGAFYLFRRGKNN

SEQ ID NO: 27
PspA
MNKKKMILTSLASVAILGAGFVASSPTFVRAEEAPVANQSKAEKDYDAAVKKSEAAKKDYETAKKKAEDA
QKKYDEDQKKTEAKAEKERKASEKIAEATKEVQQAYLAYLQASNESQRKEADKKIKEATQRKDEAEAAFA
TIRTTIVVPEPSELAETKKKAEEATKEAEVAKKKSEEAAKEVEVEKNKILEQDAENEKKIDVLQNKVADL
EKGIAPYQNEVAELNKEIARLQSDLKDAEENNVEDYIKEGLEQAITNKKAELATTQQNIDKTQKDLEDAE
LELEKVLATLDPEGKTQDELDKEAAEAELNEKVEALQNQVAELEEELSKLEDNLKDAETNNVEDYIKEGL
EEAIATKKAELEKTQKELDAALNELGPDGDEEETPAPAPQPEKPAEEPENPAPAPKPEKSADQQAEEDYA
RRSEEEYNRLTQQQPPKAEKPAPAPQPEQPAPAPKIGWKQENGMWYFYNTDGSMATGWLQNNGSWYYLNS

NGAMATGWLQYNGSWYYLNANGAMATGWLQYNGSWYYLNANGAMATGWLQYNGSWYYLNANGDMATGWLQ
YNGSWYYLNANGDMATGWAKVHGSWYYLNANGSMATGWVKDGETWYYLEASGSMKANQWFQVSDKWYYVN
GLGSLSVNTTVDGYKVNANGEWV

SEQ ID NO: 28
Immunogenic PspA/PspC polypeptides including the coiled-coil structure (PR + NPB)
MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVA
KLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSG
HMHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMADLKKAVNE
PEKPAEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEEDYARRSEEEYNRLTQQ
QPPKAEKPAPAPVPKPEQPAPAPKTGWGQENGMWCRQACGRTRAPPPPPLRSGC

SEQ ID NO: 29
Immunogenic PspA/PspC polypeptide CD2
ADLKKAVNEPEKPAEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEEDYARR
SEEEYNRLTQQQPPKAEKPAPAPVPKPEQPAPAPKTGWGQENGMW

SEQ ID NO: 30
Immunogenic PspA/PspC polypeptides lacking the coiled-coil structure (PR + NPB)
MAKKAELEKTPEKPAEEPENPAPAPQPEKSADQQAEEDYARRSEEEYNRLTQQQPPKA

SEQ ID NO: 31
Immunogenic PspA/PspC polypeptides including the coiled-coil structure (PR only)
MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVA
KLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSG
HMHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMADLKKAVNE
PETPAPAPAPAPAPTPEAPAPAPAPKPAPAPKPAPAPKPAPAPKPAPAPKPAPAPKP
APAPAPAPKPEKPAEKPAPAPKPETPKTGWKQENGMWCRQACGRTRAPPPPPLRSG

SEQ ID NO: 32
Immunogenic PspA/PspC polypeptides lacking the coiled-coil structure (PR only)
DLKKAVNEPETPAPAPAPAPAPTPEAPAPAPAPAPKPAPAPKPAPAPKPAPAPKPA
PAPKPAPAPKPAPAPAPAPKPEKPAEKPAPAPKPETPKTGWKQENGMW

SEQ ID NO: 33
Immunogenic PspA/PspC polypeptide H70 (aa 290-410 of PspA)
YFKEGLEKTIAAKKAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPK

PEKPAEQPKPEKTDDQQAEEDYARRSEEEYNRLTQQQPPKAEKPAPAPKTGWKQENGM

SEQ ID NO: 34
Non-proline Block (NPB)
EKSADQQAEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 35
Non-proline Block (NPB)
DQQAEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 36
Non-proline Block (NPB)
MEKSADQQAEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 37
SP0148 lacking signal sequence (nucleotides)
ATGTGCTCAGGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCA
CCAAAGCCATTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAA
GATTCTGACAAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGT
TACAATATGGCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCC
CAAAATCCTAATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCG
ACGGAAGTCGTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACT
ATCCTTAACTATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATT
TTTGATAAAATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGC
GACCAACAACCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATC
AAAGAACTTTATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAA
GCTGATATTAAAGAGTAA

SEQ ID NO: 38
SP0148 including signal sequence (nucleotides)
ATGAAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT

TATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGAGTAA

SEQ ID NO: 39
SP2108 lacking signal sequence (nucleotides)
ATGTGCGGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTA
GACGAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTT
AAAACTGGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATG
ATGGCTCCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCT
AAAACAGACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAG
TCACTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCT
AAAGATAGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTAC
TATACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGT
CTTGCAAACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAA
GATACAGAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCT
TGGAAAGCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAA
GAATATGCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAA
TTTGTAGACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACT
GAGGCTCGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAA
CCACTGCCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGT
CAAAAAGATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAA
TAA

SEQ ID NO: 40
SP1912 (nucleotides)
ATGAATGGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
TCCGACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGA

SEQ ID NO: 41
SP1912L (nucleotides)
ATGAGATACCTGGCAACATTGTTGTTATCTCTGGCGGTGTTAATCACCGCCGGGTGCAAAAAAGTTGCAGATGAC
CGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATGGGGGAAATTGCG
ACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAGGATGGCCGTCAT
TATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGA

SEQ ID NO: 42
SP0641N (nucleotides)
ATGGTAGTCTTAGCAGACACATCTAGCTCTGAAGATGCTTTAAACATCTCTGATAAAGAAAAAGTAGCAGAAAAT
AAAGAGAAACATGAAAATATCCATAGTGCTATGGAAACTTCACAGGATTTTAAAGAGAAGAAAACAGCAGTCATT

AAGGAAAAAGAAGTTGTTAGTAAAAATCCTGTGATAGACAATAACACTAGCAATGAAGAAGCAAAAATCAAAGAA
GAAAATTCCAATAAATCCCAAGGAGATTATACGGACTCATTTGTGAATAAAAACACAGAAAATCCCAAAAAAGAA
GATAAAGTTGTCTATATTGCTGAATTTAAAGATAAAGAATCTGGAGAAAAAGCAATCAAGGAACTATCCAGTCTT
AAGAATACAAAAGTTTTATATACTTATGATAGAATTTTTAACGGTAGTGCCATAGAAACAACTCCAGATAACTTG
GACAAAATTAAACAAATAGAAGGTATTTCATCGGTTGAAAGGGCACAAAAAGTCCAACCCATGATGAATCATGCC
AGAAAGGAAATTGGAGTTGAGGAAGCTATTGATTACCTAAAGTCTATCAATGCTCCGTTTGGGAAAAATTTTGAT
GGTAGAGGTATGGTCATTTCAAATATCGATACTGGAACAGATTATAGACATAAGGCTATGAGAATCGATGATGAT
GCCAAAGCCTCAATGAGATTTAAAAAAGAAGACTTAAAAGGCACTGATAAAAATTATTGGTTGAGTGATAAAATC
CCTCATGCGTTCAATTATTATAATGGTGGCAAAATCACTGTAGAAAAATATGATGATGGAAGGGATTATTTTGAC
CCACATGGGATGCATATTGCAGGGATTCTTGCTGGAAATGATACTGAACAAGACATCAAAAACTTTAACGGCATA
GATGGAATTGCACCTAATGCACAAATTTTCTCTTACAAAATGTATTCTGACGCAGGATCTGGGTTTGCGGGTGAT
GAAACAATGTTTCATGCTATTGAAGATTCTATCAAACACAACGTTGATGTTGTTTCGGTATCATCTGGTTTTACA
GGAACAGGTCTTGTAGGTGAGAAATATTGGCAAGCTATTCGGGCATTAAGAAAAGCAGGCATTCCAATGGTTGTC
GCTACGGGTAACTATGCGACTTCTGCTTCAAGTTCTTCATGGGATTTAGTAGCAAATAATCATCTGAAAATGACC
GACACTGGAAATGTAACACGAACTGCAGCACATGAAGATGCGATAGCGGTCGCTTCTGCTAAAAATCAAACAGTT
GAGTTTGATAAAGTTAACATAGGTGGAGAAAGTTTTAAATACAGAAATATAGGGGCCTTTTTCGATAAGAGTAAA
ATCACAACAAATGAAGATGGAACAAAAGCTCCTAGTAAATTAAAATTTGTATATATAGGCAAGGGGCAAGACCAA
GATTTGATAGGTTTGGATCTTAGGGGCAAAATTGCAGTAATGGATAGAATTTATACAAAGGATTTAAAAAATGCT
TTTAAAAAAGCTATGGATAAGGGTGCACGCGCCATTATGGTTGTAAATACTGTAAATTACTACAATAGAGATAAT
TGGACAGAGCTTCCAGCTATGGGATATGAAGCGGATGAAGGTACTAAAAGTCAAGTGTTTTCAATTTCAGGAGAT
GATGGTGTAAAGCTATGGAACATGATTAATCCTGATAAAAAAACTGAAGTCAAAAGAAATAATAAAGAAGATTTT
AAAGATAAATTGGAGCAATACTATCCAATTGATATGGAAAGTTTTAATTCCAACAAACCGAATGTAGGTGACGAA
AAAGAGATTGACTTTAAGTTTGCACCTGACACAGACAAAGAACTCTATAAAGAAGATATCATCGTTCCAGCAGGA
TCTACATCTTGGGGGCCAAGAATAGATTTACTTTTAAAACCCGATGTTTCAGCACCTGGTAAAAATATTAAATCC
ACGCTTAATGTTATTAATGGCAAATCAACTTATGGC

SEQ ID NO: 43
SP0882N DNA
ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT
GTACGTATTCTTCTTCCTAAAGATTATGAGAAAGATACAGACCGTTCCTATCCTGTTGTATACTTTCATGACGGG
CAAAATGTTTTTAATAGCAAAGAGTCTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG
GATATCAGTCGCATGATTGTCGTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTCATGGAGGTG
GTCAAGCCTTTTATC

SEQ ID NO: 44
SP0882 with exogenous signal sequence (nucleotides)
ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT
GTACGTATTCTTCTTCCTAAAGATTATGAGAAAGATACAGACCGTTCCTATCCTGTTGTATACTTTCATGACGGG
CAAAATGTTTTTAATAGCAAAGAGTCTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG

GATATCAGTCGCATGATTGTCGTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTCATGGAGGTG
GTCAAGCCTTTTATCGATGAGACCTATCGTACAAAAGCAGACTGCCAGCATACGGCTATGATTGGTTCCTCACTA
GGAGGCAATATTACCCAGTTTATCGGTTTGGAATACCAAGACCAAATTGGTTGCTTGGGCGTTTTTTCATCTGCA
AACTGGCTCCACCAAGAAGCCTTTAACCGCTATTTCGAGTGCCAGAAACTATCGCCTGACCAGCGCATCTTCATC
TATGTAGGAACAGAAGAAGCAGATGATACAGACAAGACCTTGATGGATGGCAATATCAAACAAGCCTATATCGAC
TCGTCGCTTTGCTATTACCATGATTTGATAGCAGGGGGAGTACATCTGGATAATCTTGTGCTAAAAGTTCAGTCT
GGTGCCATCCATAGTGAAATCCCTTGGTCAGAAAATCTACCAGATTGTCTGAGATTTTTTGCAGAAAAATGGTAA

SEQ ID NO: 45
SP0882N with exogenous signal sequence (nucleotides)
ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT
GTACGTATTCTTCTTCCTAAAGATTATGAGAAAGATACAGACCGTTCCTATCCTGTTGTATACTTTCATGACGGG
CAAAATGTTTTTAATAGCAAAGAGTCTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG
GATATCAGTCGCATGATTGTCGTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTCATGGAGGTG
GTCAAGCCTTTTATC

SEQ ID NO: 46
PSPA (nucleotides)
TTGACAAATATTTACGGAGGAGGCTTATGCTTAATATAAGTATAGGCTAAAAATGATTATCAGAAAAGAG
GTAAATTTAGATGAATAAGAAAAAAATGATTTTAACAAGCCTAGCCAGCGTCGCTATCTTAGGGGCTGGT
TTTGTTGCGTCTTCGCCTACTTTTGTAAGAGCAGAAGAAGCTCCTGTAGCTAACCAGTCTAAAGCTGAGA
AAGACTATGATGCAGCAGTGAAAAAATCTGAAGCTGCTAAGAAAGATTACGAAACGGCTAAAAAGAAAGC
AGAAGACGCTCAGAAGAAATATGATGAGGATCAGAAGAAAACTGAGGCAAAAGCGGAAAAAGAAAGAAAA
GCTTCTGAAAAGATAGCTGAGGCAACAAAAGAAGTTCAACAAGCGTACCTAGCTTATCTACAAGCTAGCA
ACGAAAGTCAGAGAAAAGAGGCAGATAAGAAGATAAAAGAAGCTACGCAACGCAAAGATGAGGCGGAAGC
TGCATTTGCTACTATTCGAACAACAATTGTAGTTCCTGAACCAAGTGAGTTAGCTGAGACTAAGAAAAAA
GCAGAAGAGGCAACAAAAGAAGCAGAAGTAGCTAAGAAAAAATCTGAAGAGGCAGCTAAAGAGGTAGAAG
TAGAGAAAAATAAAATACTTGAACAAGATGCTGAAAACGAAAAGAAAATTGACGTACTTCAAAACAAAGT
CGCTGATTTAGAAAAAGGAATTGCTCCTTATCAAAACGAAGTCGCTGAATTAAATAAAGAAATTGCTAGA
CTTCAAAGCGATTTAAAAGATGCTGAAGAAAATAATGTAGAAGACTACATTAAAGAAGGTTTAGAGCAAG
CTATCACTAATAAAAAAGCTGAATTAGCTACAACTCAACAAAACATAGATAAAACTCAAAAAGATTTAGA
GGATGCTGAATTAGAACTTGAAAAAGTATTAGCTACATTAGACCCTGAAGGTAAAACTCAAGATGAATTA
GATAAAGAAGCTGCTGAAGCTGAGTTGAATGAAAAAGTTGAAGCTCTTCAAAACCAAGTTGCTGAATTAG
AAGAAGAACTTTCAAAACTTGAAGATAATCTTAAAGATGCTGAAACAAACAACGTTGAAGACTACATTAA
AGAAGGTTTAGAAGAAGCTATCGCGACTAAAAAAGCTGAATTGGAAAAAACTCAAAAAGAATTAGATGCA
GCTCTTAATGAGTTAGGCCCTGATGGAGATGAAGAAGAGACTCCAGCGCCGGCTCCTCAACCAGAAAAAC
CAGCTGAAGAGCCTGAGAATCCAGCTCCAGCACCAAAACCAGAGAAGTCAGCAGATCAACAAGCTGAAGA
AGACTATGCTCGTAGATCAGAAGAAGAATATAATCGCTTGACCCAACAGCAACCGCCAAAAGCAGAAAAA

CCAGCTCCTGCACCACAACCAGAGCAACCAGCTCCTGCACCAAAAATAGGTTGGAAACAAGAAAACGGTA
TGTGGTACTTCTACAATACTGATGGTTCAATGGCGACAGGTTGGCTACAAAACAACGGTTCATGGTACTA
CCTCAACAGCAATGGCGCTATGGCTACAGGTTGGCTCCAATACAATGGTTCATGGTATTACCTAAACGCT
AACGGCGCTATGGCGACAGGCTGGCTCCAATACAATGGCTCATGGTACTACCTCAACGCTAACGGCGCTA
TGGCGACAGGCTGGCTCCAATACAATGGCTCATGGTACTACCTCAACGCTAATGGTGATATGGCGACAGG
ATGGCTCCAATACAACGGTTCATGGTATTACCTCAACGCTAATGGTGATATGGCTACAGGTTGGGCTAAA
GTCCACGGTTCATGGTACTACCTCAACGCTAACGGTTCAATGGCAACAGGTTGGGTGAAAGATGGAGAAA
CCTGGTACTATCTTGAAGCATCAGGTTCTATGAAAGCAAACCAATGGTTCCAAGTATCAGATAAATGGTA
CTATGTCAATGGTTTAGGTTCCCTTTCAGTCAACACAACTGTAGATGGCTATAAAGTCAATGCCAATGGT
GAATGGGTTTAAGCCG

SEQ ID NO: 47
SP2108/SP0148 fusion protein (675 amino acids)
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL
PNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGESGGAKKEGEAASKKEIIVATNGSPKPF
IYEENGELTGYEIEVVRAIFKDSDKYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPN
VLVVKKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKI
GVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIK

SEQ ID NO: 48
SP0148/SP2108 fusion protein (674 amino acids)
MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSD
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKGSKTADKPADSGSSEVKELTVYVD
EGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAK
TDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYY
TYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPW
KAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTE
ARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 49
SP2108/SP1912 fusion protein (521 amino acids)
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY

AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL
PNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGENGMKAKKMWMAGLALLGIGSLALATKK
VADDRKLMKTQE ELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIFEDGRHYTFVYENEDLVYEEEVL

SEQ ID NO: 50
SP0148/SP1912 fusion protein (374 amino acids)
MKKIVKYSSLAALALVAAGVLAA**C**SGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSD
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKNGMKAKKMWMAGLALLGIGSLALA
TKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIFEDGRHYTFVYENEDLVYEEEVL

SEQ ID NO: 51
SP2108/SP1912/SP0148 fusion protein (773 amino acids)
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL
PNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGENGMKAKKMWMAGLALLGIGSLALATKK
VADDRKLMKTQE
ELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIFEDGRHYTFVYENEDLVYEEEVLSGGAKKEGEAASKKEI
IVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSDKYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKY
LYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSD
GQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFG
DTYLPAEADIK

SEQ ID NO: 52
SP0148/SP1912/SP2108 fusion protein (772 amino acids)
MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSD
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKNGMKAKKMWMAGLALLGIGSLALA
TKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIFEDGRHYTFVYENEDLVYEEEVLG
SKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAP
YDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDS
KYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQDTE
GAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFVD
FLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSGQKD
AKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 53
SP2108/SP0148/SP1912 fusion protein (773 amino acids)
MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL
PNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGESGGAKKEGEAASKKEIIVATNGSPKPF
IYEENGELTGYEIEVVRAIFKDSDKYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPN
VLVVKKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKI
GVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIK
NGMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVI
FEDGRHYTFVYENEDLVYEEEVL

SEQ ID NO: 54
SP0148/SP2108/SP1912 fusion protein (772 amino acids)
MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIYEENGELTGYEIEVVRAIFKDSD
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAEKYLYAAPIAQNPNVLVVKKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKGSKTADKPADSGSSEVKELTVYVD
EGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAPYDRVGSLGSDGQLSEVKLSDGAK
TDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYY
TYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTAAIIDGPW
KAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTE
ARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGEN
GMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLVGGVIF
EDGRHYTFVYENEDLVYEEEVL

SEQ ID NO: 55
L460D (471 amino acids)
MANKAVNDFILAMNYDKKKLLTHQGESIENRFIKEGNQLPDEFVVIERKKRSLSTNTSDISVTATNDSRLYPGAL
LVVDETLLENNPTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSSVRGAVNDLLAKWHQDYGQVNNVPARMQY
EKITAHSMEQLKVKFGSDFEKTGNSLDIDFNSVHSGEKQIQIVNFKQIYYTVSVDAVKNPGDVFQDTVTVEDLKQ
RGISAERPLVYISSVAYGRQVYLKLETTSKSDEVEAAFEALIKGVKVAPQTEWKQILDNTEVKAVILGGDPSSGA
RVVTGKVDMVEDLIQEGSRFTADHPGLPISYTTSFLRDNVVATFQNSTDYVETKVTAYRNGDLLLDHSGAYVAQY
YITWDELSYDHQGKEVLTPKAWDRNGQDLTAHFTTSIPLKGNVRNLSVKIRECTGLAWEWWRTVYEKTDLPLVRK
RTISIWGTTDYPQVEDKVEND

SEQ ID NO: 56
SP2108/SP0148 fusion (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGAT
AGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAATCAGGG
GGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCATTT
ATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGACAAA
TATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCT
GTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCTAAT
GTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTCGTT
CAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAACTAT
ACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAAATC
GGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAACCG
TACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTAT
AAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA

SEQ ID NO: 57

SP0148/SP2108 fusion (nucleotides)

ATGAAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTSCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT

AAAGGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGAC
GAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAA
ACTGGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATG
GCTCCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAA
ACAGACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCA
CTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAA
GATAGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTAT
ACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTT
GCAAACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGAT
ACAGAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGG
AAAGCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAA
TATGCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTT
GTAGACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAG
GCTCGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCA
CTGCCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAA
AAAGATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAA

SEQ ID NO: 58
SP2108/SP1912 fusion (nucleotides)
ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGAT
AGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAAATGGT
ATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAAAAA
GTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATG
GGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAG
GATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 59

SP0148/SP1912 fusion (nucleotides)

ATGAAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAAATGGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
TCCGACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 60

SP2108/SP1912/SP0148 fusion (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGAT
AGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAAATGGT
ATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAAAAA
GTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATG
GGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAG
GATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGCTCAGGGGGT

GCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCATTTATC
TATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGACAAATAT
GATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCTGTC
AACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCTAATGTC
CTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTCGTTCAA
GCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAACTATACT
AAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAAATCGGT
GTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAACCGTAC
GTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTATAAA
GATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA

SEQ ID NO: 61
SP0148/SP1912/SP2108 fusion (nucleotides)
ATGAAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAAATGGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
TCCGACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTAGGA
AGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAGGGA
TATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAAACTGGT
GATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCTCCA
TACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACAGAC
GACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTTGTT
ATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGATAGC
AAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACATAT
GGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCAAAC
GACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACAGAA
GGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAAGCT
CAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATATGCT
GCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTAGAC
TTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCTCGT

TCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTGCCA
AACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAAGAT
GCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAACACCACCAC
CACCACCACTGA

SEQ ID NO: 62
SP2108/SP0148/SP1912 fusion (nucleotides)
ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGAT
AGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAATCAGGG
GGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCATTT
ATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGACAAA
TATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCT
GTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCTAAT
GTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTCGTT
CAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAACTAT
ACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAAATC
GGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAACCG
TACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTAT
AAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA
AATGGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACG
AAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCC
GACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATT
TTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 63

SP0148/SP2108/SP1912 fusion (nucleotides)

ATGAAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGGAAGCAAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGAC
GAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAA
ACTGGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATG
GCTCCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAA
ACAGACGACACAACTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCA
CTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAA
GATAGCAAATACGCATTCGCTGGTGAAGATGGTAAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTAT
ACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTT
GCAAACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGAT
ACAGAAGGTGCTGGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGG
AAAGCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAA
TATGCTGCATTCGGTGGTGGTAAAGCTTGGGTCATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTT
GTAGACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAG
GCTCGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCA
CTGCCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAA
AAAGATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAAAT
GGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAA
AAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGAC
ATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTT
GAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 64

HHHHHH

SEQ ID NO: 65

MSYYHHHHHH

SEQ ID NO: 66

Canonical lipobox motif
[LIVMFESTAGPC]-[LVIAMFTG]-[IVMSTAGCP]-[AGS]-C

SEQ ID NO: 67
SP2108 signal sequence
MSSKFMKSAAVLGTATLASLLLVAC

SEQ ID NO: 68
E. coli RlpB signal sequence
MRYLATLLLSLAVLITAG[C]

SEQ ID NO: 69
Fusion protein linker (19-mer)
LGGGGSGGGGSGGGGSAAA

SEQ ID NO: 70
Fusion protein LR linker (20-mer)
LAEATAKEATAKEATAKATA

SEQ ID NO: 71
Fusion protein LC linker (14-mer)
GPKPHRIQSTPKGS

SEQ ID NO: 72
Amino-terminal boundary to the PR-region
DLKKAVNE

SEQ ID NO: 73
Carboxy-terminal boundary to the PR-region
(K/G)TGW(K/G)QENGMW

2013209513 15 Sep 2017

CLAIMS

We claim:

1. A fusion protein comprising:

a first polypeptide comprising an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 6 or to the amino acid sequence of SEQ ID NO: 5; and

a second polypeptide comprising an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 9.

2. The fusion protein of claim 1, wherein the first polypeptide comprises the amino acid sequence of SEQ ID NO: 6 or comprises the amino acid sequence of SEQ ID NO: 5, and the second polypeptide comprises the amino acid sequence of SEQ ID NO: 9.

3. The fusion protein of claim 1, wherein the first polypeptide consists of the amino acid sequence of SEQ ID NO: 6 or consists of the amino acid sequence of SEQ ID NO: 5.

4. The fusion protein of any one of claims 1 to 3, wherein the second polypeptide consists of the amino acid sequence of SEQ ID NO: 9.

5. The fusion protein of claim 1, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 49.

6. The fusion protein of claim 1, wherein the fusion protein consists of the amino acid sequence of SEQ ID NO: 49.

7. The fusion protein of any one of claims 1-6, further comprising a third polypeptide comprising an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2 or to the amino acid sequence of SEQ ID NO: 1.

8. The fusion protein of claim 7, wherein the third polypeptide comprises the amino acid sequence of SEQ ID NO: 2 or comprises the amino acid sequence of SEQ ID NO: 1.

9. The fusion protein of claim 7, wherein the third polypeptide consists of the amino acid sequence of SEQ ID NO: 2 or consists of the amino acid sequence of SEQ ID NO: 1.

10. The fusion protein of claim 7, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO:53.

11. The fusion protein of claim 7, wherein the fusion protein consists of the amino acid sequence of SEQ ID NO:53.

12. The fusion protein of claim 7, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 52.

13. The fusion protein of claim 7, wherein the fusion protein consists of the amino acid sequence of SEQ ID NO: 52.

14. The fusion protein of any one of claims 1 to 13, wherein the fusion protein is lipidated.

15. The fusion protein of claim 1, wherein the fusion protein comprises an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 49.

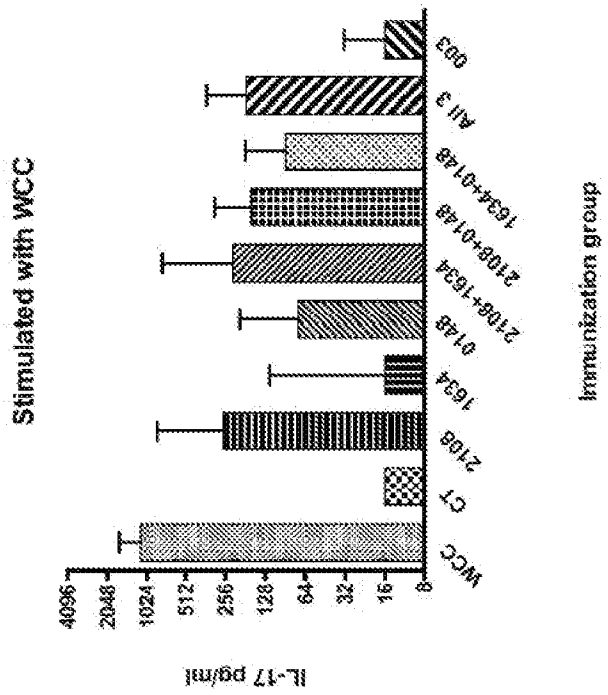
16. The fusion protein of claim 7, wherein the fusion protein comprises an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 53.

17. The fusion protein of claim 7, wherein the fusion protein comprises an amino acid sequence with at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 52.

18. An immunogenic composition comprising a pharmaceutically acceptable carrier and the fusion protein of any one of claims 1 to 17.

19. The immunogenic composition of claim 18, further comprising an adjuvant.

20. The immunogenic composition of claim 18 or 19, wherein the immunogenic composition comprises 1-1000 μg of the fusion protein.


Stimulated with WCC
IL-17 pg/ml
4096
2048
1024
512
256
128
64
32
16
0
WCC
CT
2108
1634
0148
2108+1634
2108+0148
1634+0148
All 3
003
Immunization group

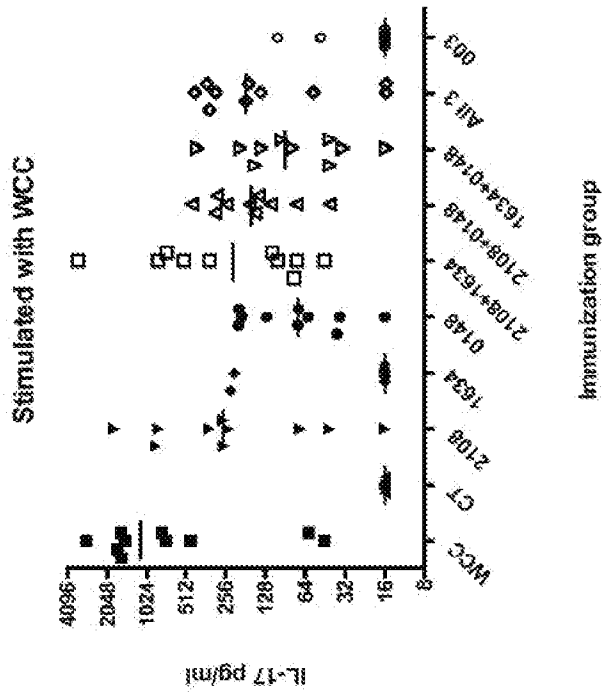

Stimulated with WCC
IL-17 pg/ml
4096
2048
1024
512
256
128
64
32
16
0
WCC
CT
2108
1634
0148
2108+1634
2108+0148
1634+0148
All 3
003
Immunization group

Figure 1

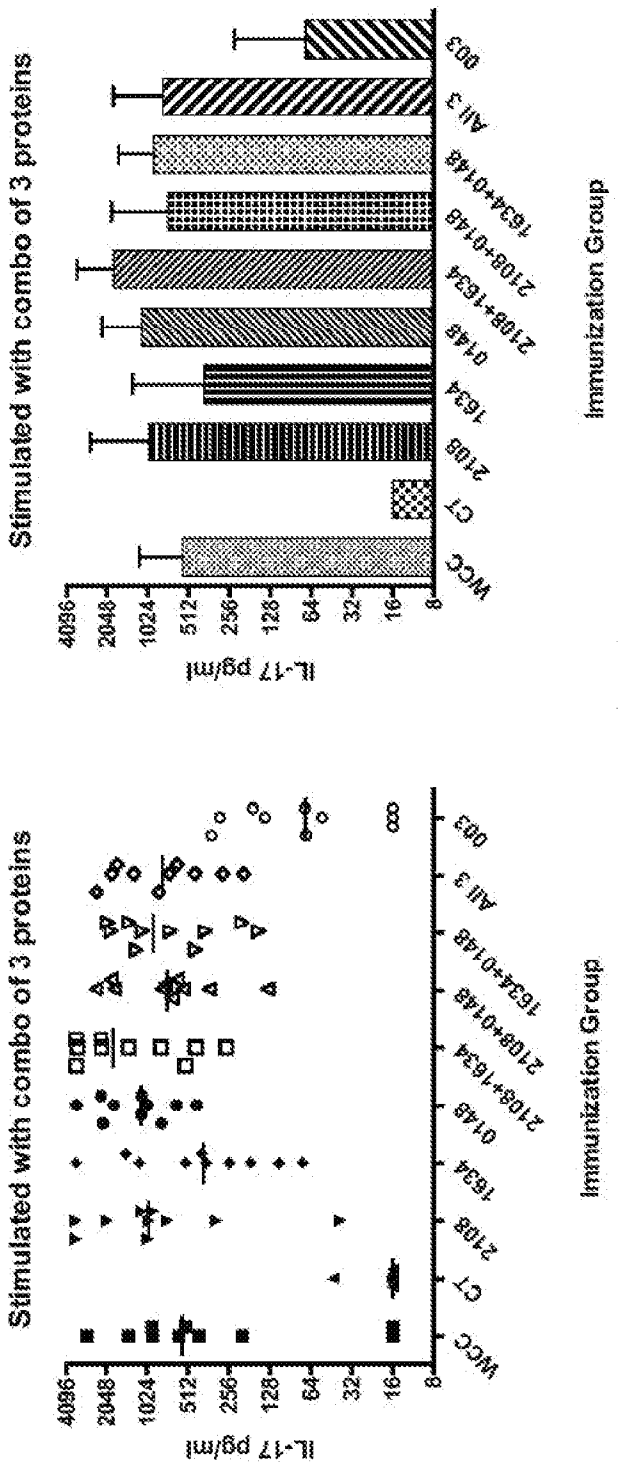

Stimulated with combo of 3 proteins
IL-17 pg/ml
4096
2048
1024
512
256
128
64
32
16
8
WCC
CT
2108
1634
0148
2108+1634
2108+0148
1634+0148
All 3
003
Immunization Group
Stimulated with combo of 3 proteins
IL-17 pg/ml
Immunization Group

Figure 2

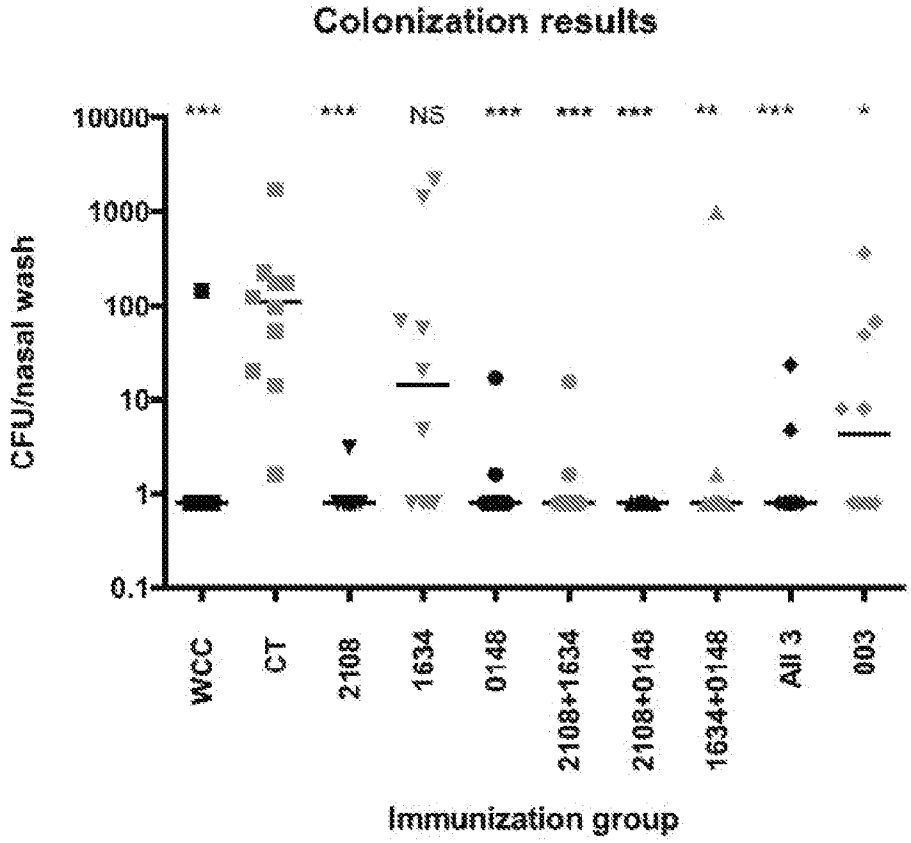

Colonization results
CFU/nasal wash
10000
1000
100
10
1
0.1
*** *** NS *** *** *** ** *** *
WCC
CT
2108
1634
0148
2108+1634
2108+0148
1634+0148
All 3
003
Immunization group

Figure 3

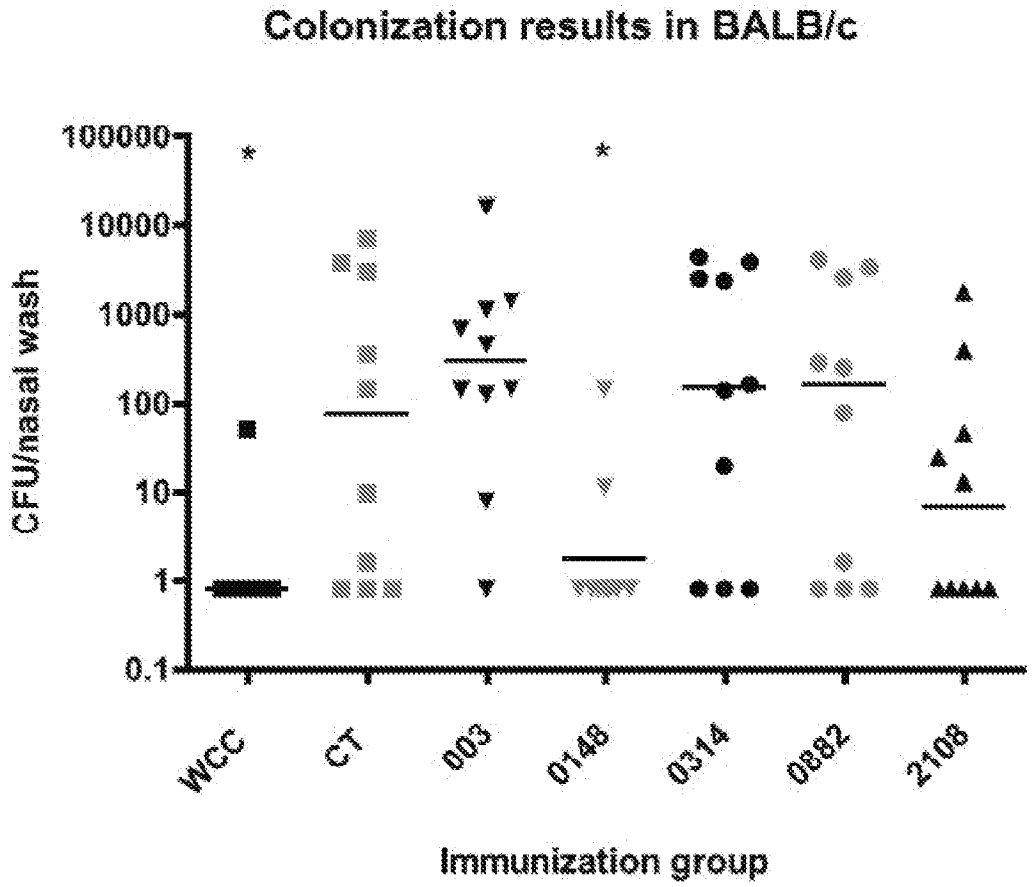

Colonization results in BALB/c
CFU/nasal wash
100000
10000
1000
100
10
1
0.1
WCC
CT
003
0148
0314
0882
2108
Immunization group

Figure 4

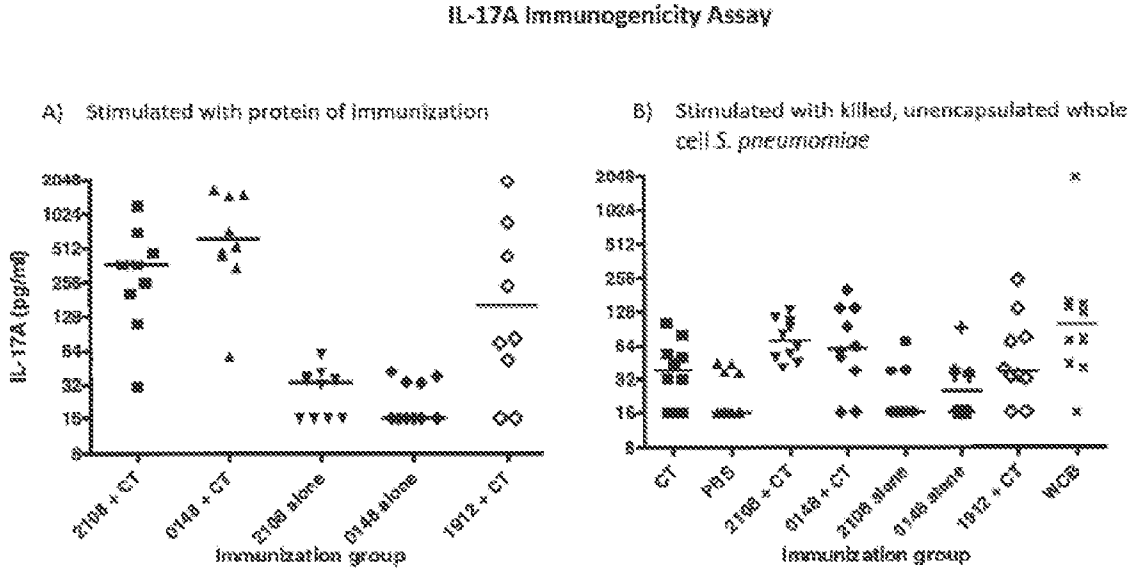

IL-17A Immunogenicity Assay
A) Stimulated with protein of immunization
B) Stimulated with killed, unencapsulated whole cell S. pneumoniae
IL-17A (pg/ml)
2048
1024
512
256
128
64
32
16
8
0
2108 + CT
0148 + CT
2108 alone
0148 alone
1912 + CT
Immunization group
CT
PBS
2108 + CT
0148 + CT
2108 alone
0148 alone
1912 + CT
WCSP
Immunization group

Figure 5

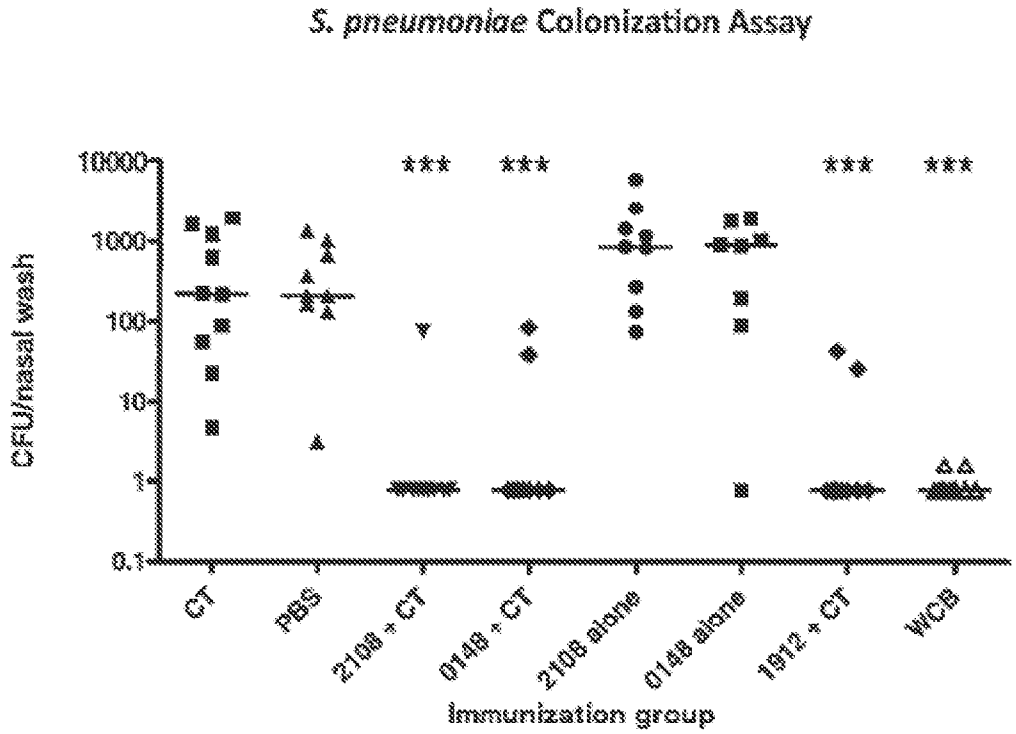

S. pneumoniae Colonization Assay
CFU/nasal wash
10000
1000
100
10
1
0.1
***
***
***
***
CT
PBS
2198 + CT
0148 + CT
2108 alone
0148 alone
1912 + CT
WCB
Immunization group

Figure 6

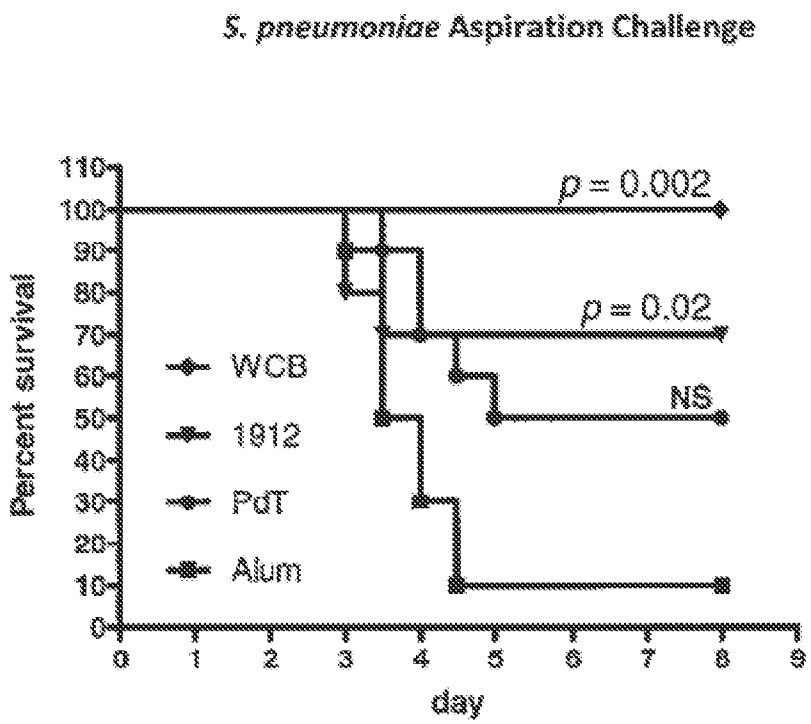

S. pneumoniae Aspiration Challenge
Percent survival
110
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
day
p = 0.002
p = 0.02
NS
WCB
1912
PdT
Alum

Figure 7

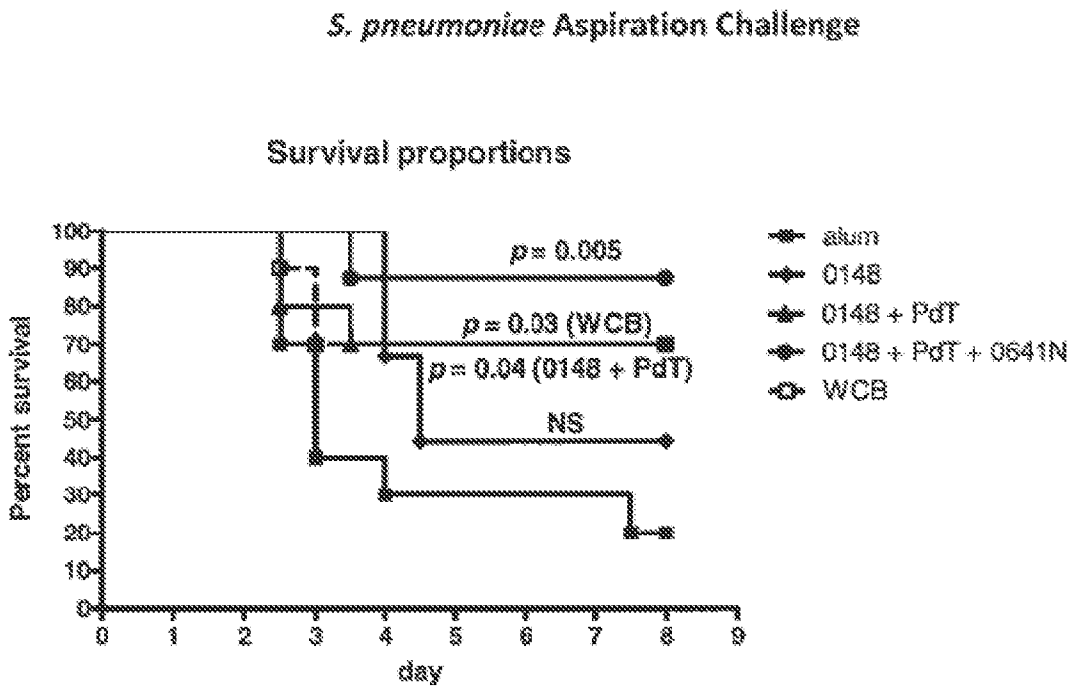

S. pneumoniae Aspiration Challenge
Survival proportions
Percent survival
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
day
p = 0.005
p = 0.03 (WCB)
p = 0.04 (0148 + PdT)
NS
alum
0148
0148 + PdT
0148 + PdT + 0641N
WCB

Figure 8

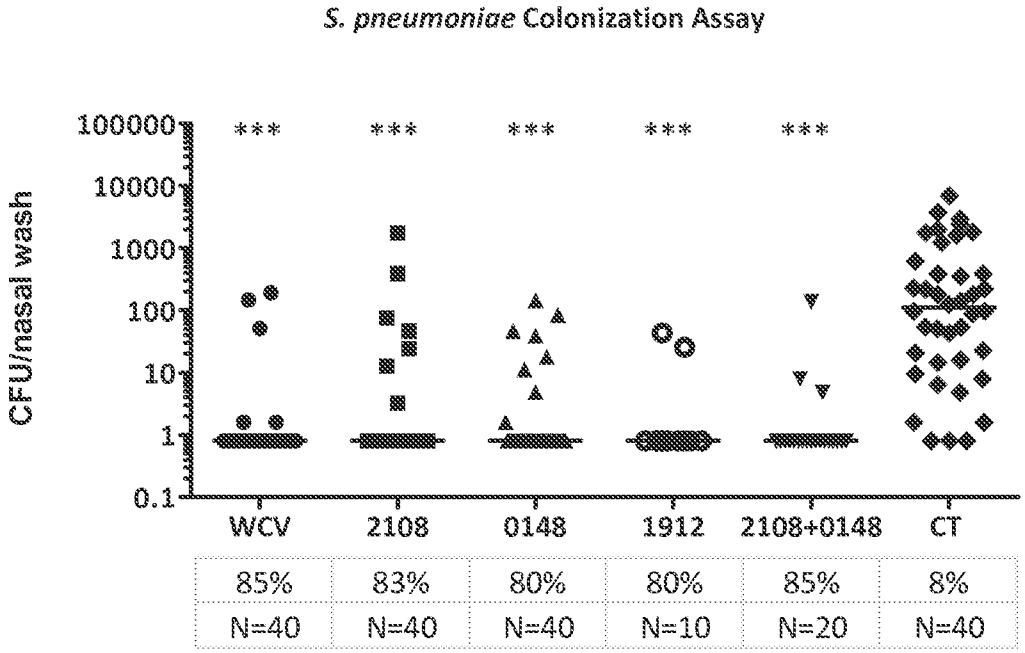

S. pneumoniae Colonization Assay
CFU/nasal wash
100000
10000
1000
100
10
1
0.1
***
***
***
***
***
WCV
2108
0148
1912
2108+0148
CT
85%
83%
80%
80%
85%
8%
N=40
N=40
N=40
N=10
N=20
N=40

Figure 9

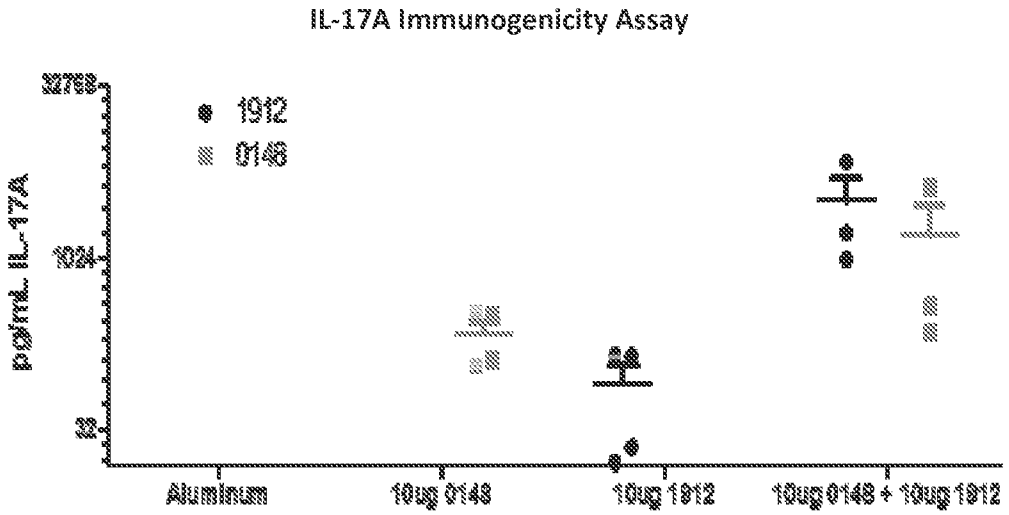

IL-17A Immunogenicity Assay
pg/mL IL-17A
2768
1024
32
1912
0148
Aluminum
10ug 0148
10ug 1912
10ug 0148 + 10ug 1912

Figure 10

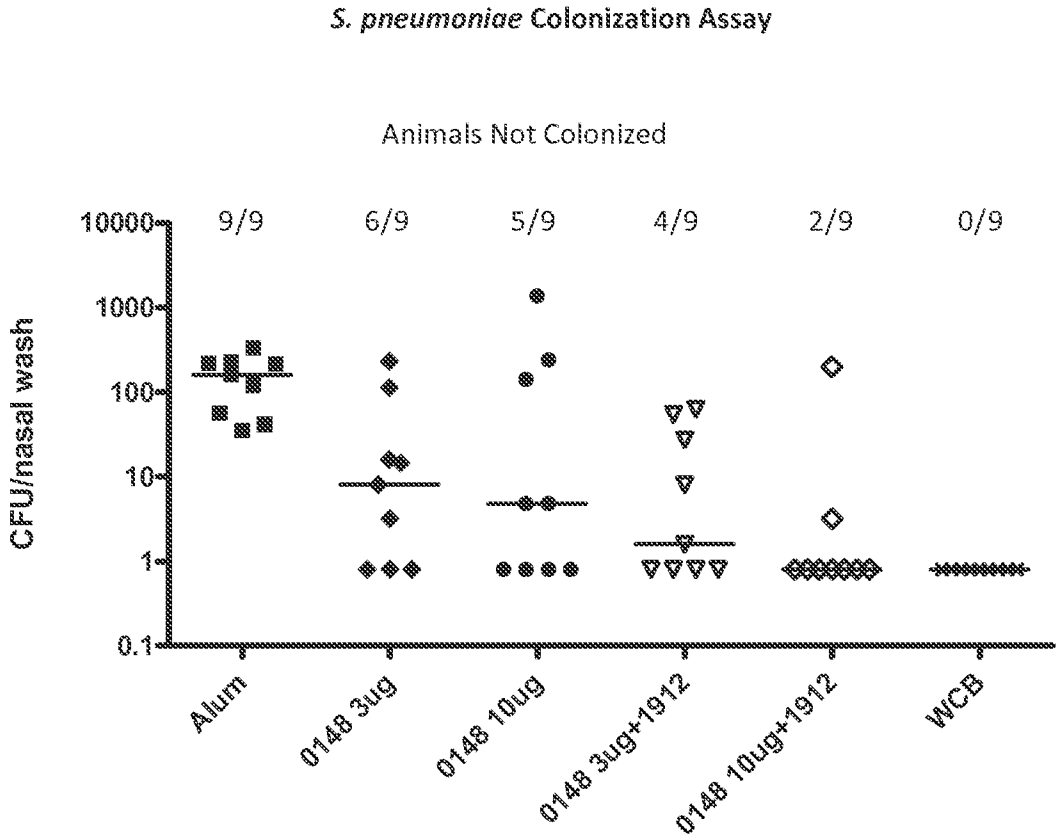

S. pneumoniae Colonization Assay
Animals Not Colonized
9/9
6/9
5/9
4/9
2/9
0/9
CFU/nasal wash
10000
1000
100
10
1
0.1
Alum
0148 3ug
0148 10ug
0148 3ug+1912
0148 10ug+1912
WCB

Figure 11

S. pneumoniae Colonization Assay: Pooled Data from 2 Vaccination Studies

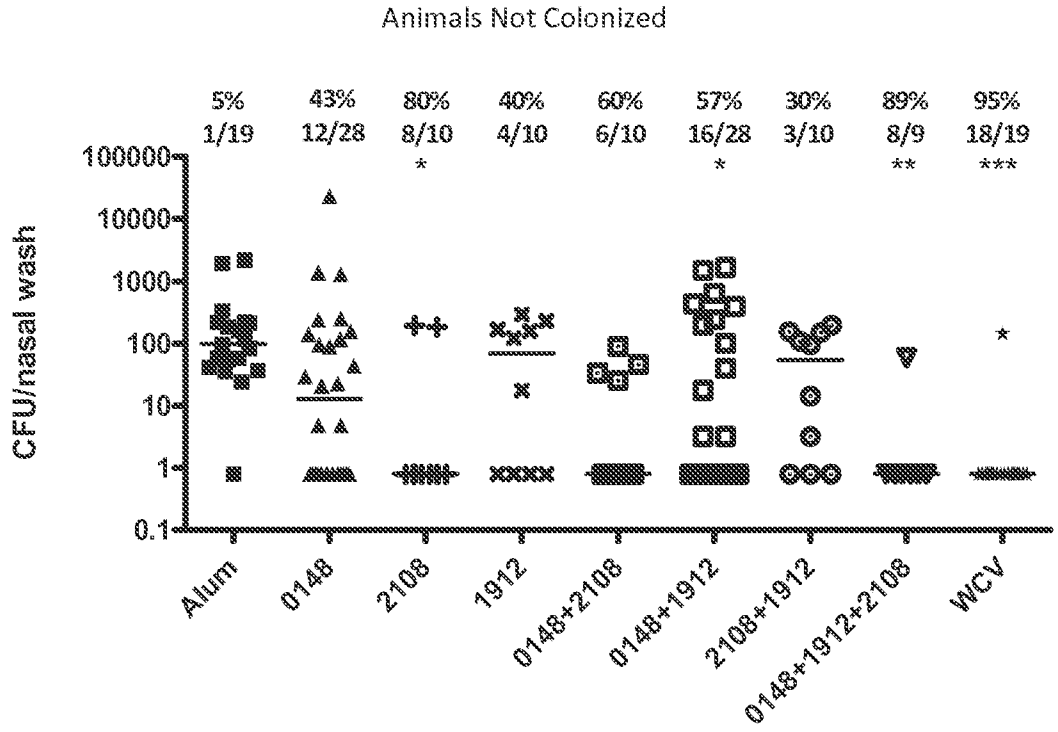


Figure 12

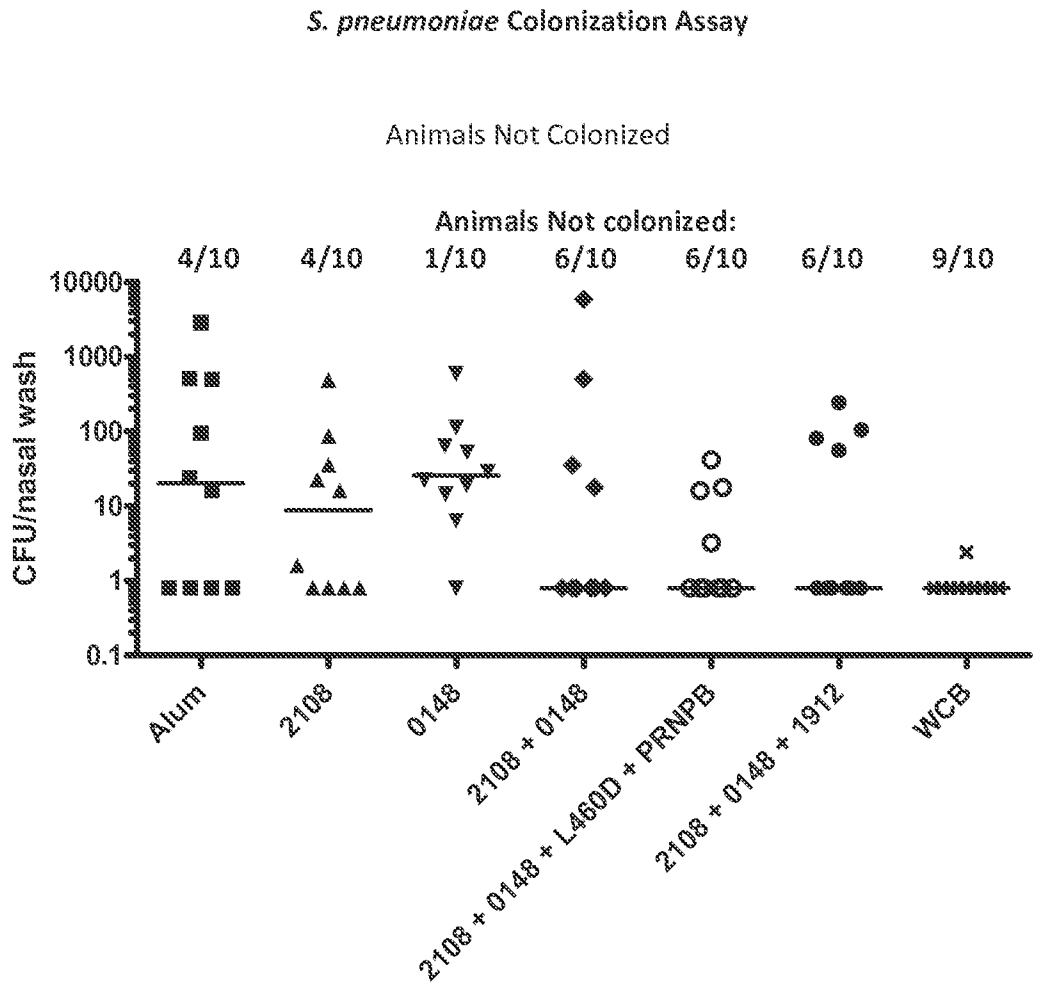

S. pneumoniae Colonization Assay
Animals Not Colonized
Animals Not colonized:
4/10
4/10
1/10
6/10
6/10
6/10
9/10
CFU/nasal wash
10000
1000
100
10
1
0.1
Alum
2108
0148
2108 + 0148
2108 + 0148 + L460D + PRNPB
2108 + 0148 + 1912
WCB

Figure 13

S. pneumoniae Aspiration Challenge: Passive Transfer Study

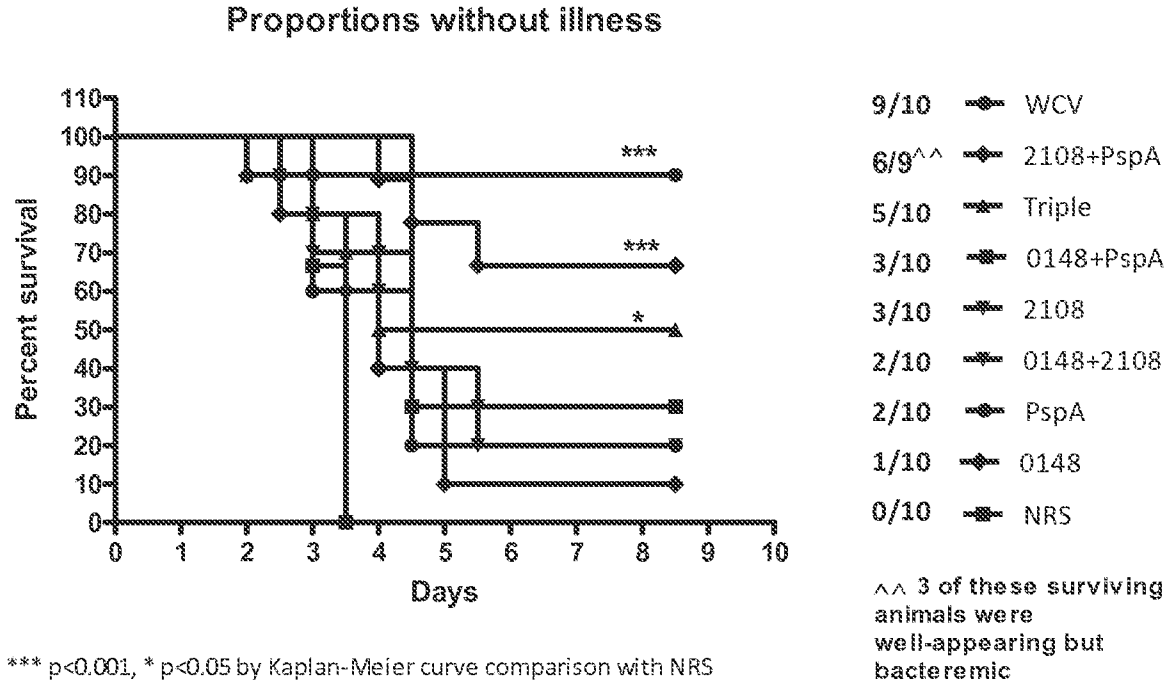


Figure 14

S. pneumoniae Aspiration Challenge: Summary of Studies

	Sepsis 1	Sepsis 2	Sepsis 3	Sepsis 4	PT1^	PT2*
alum or NRS	20%	10%	30%	20%	0%	20%
0148	44%	10%			10%	
1912		70%		20%		
2108					30%	
PdT		50%		20%		
PspA					10%	70%
0148 + 1912		40%	40%			
0148 + 2108					20%	
0148 + PdT	70%	17%	25%			
0148 + PspA					30%	70%
1912 + PdT				40%		
2108 + PspA					67%	
0148 + PdT + 0641N	90%					
0148 + PdT + 1912		60%	60%			
0148 + 2108 + PspA					50%	70%
WCB	70%	100%	100%	100%	90%	90%

°10ug protein dose

*PspA - purified IgG

^PspA - ascites

Figure 15

IL-17A Immunogenicity Assay: Fusion Proteins

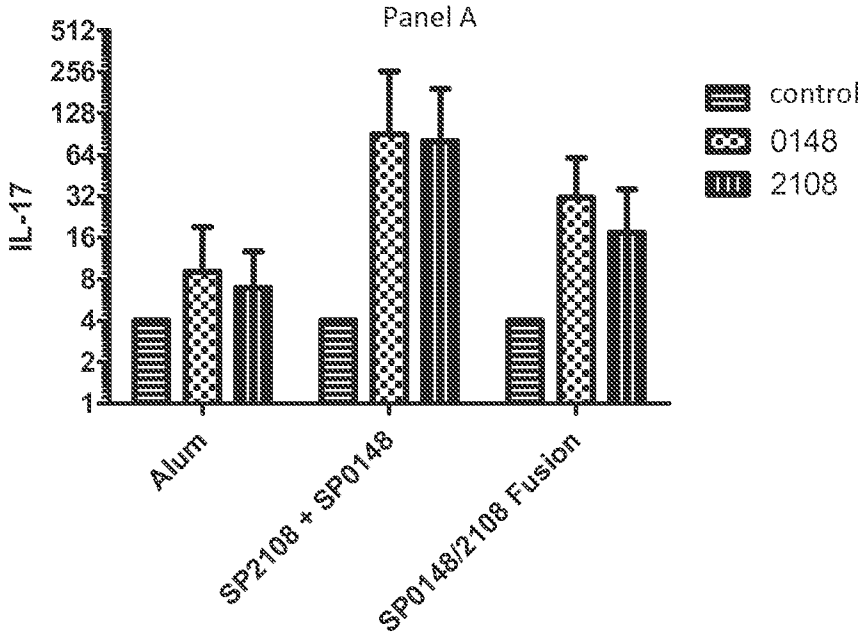


Figure 16A

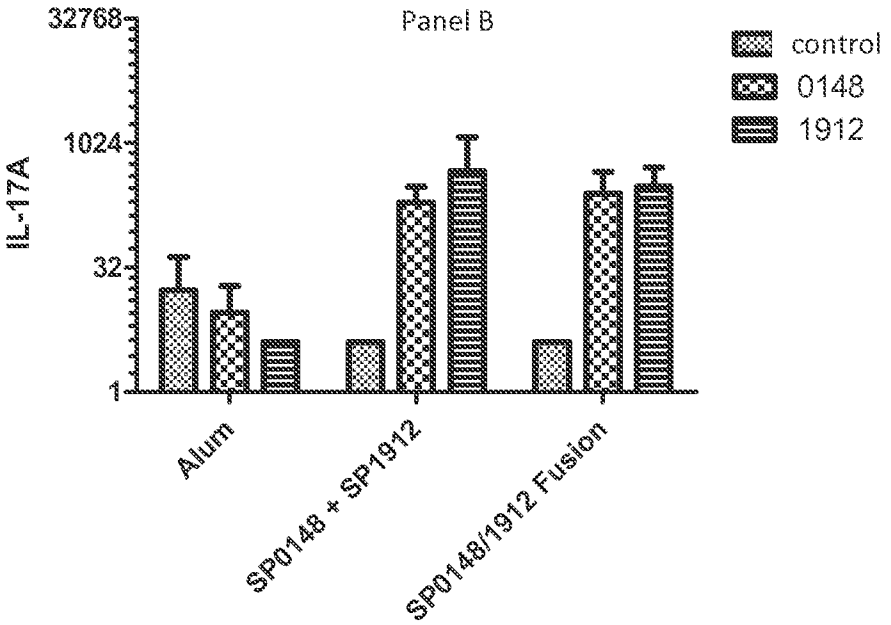


Figure 16B

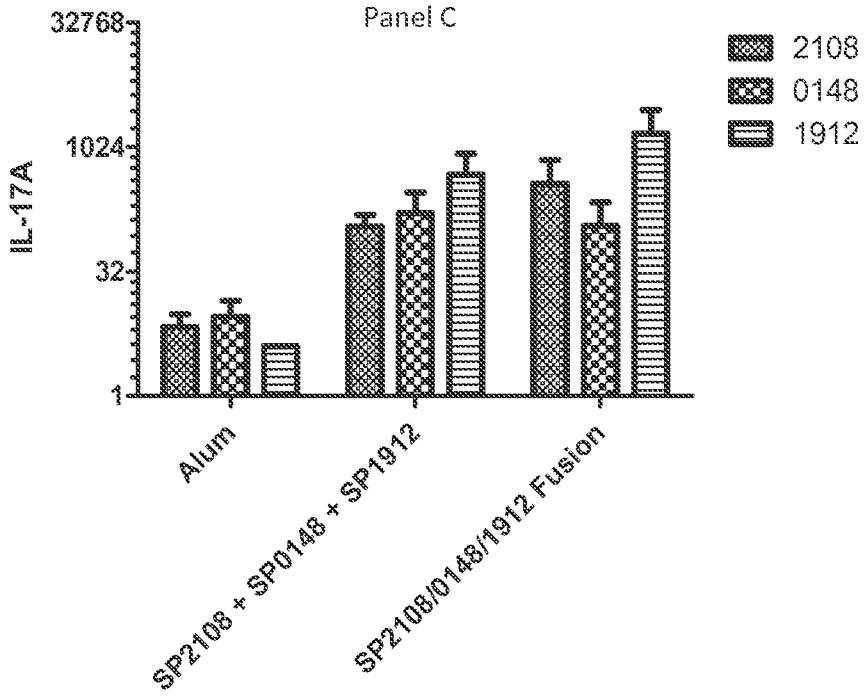

Panel C
32768
1024
32
1
IL-17A
2108
0148
1912
Alum
SP2108 + SP0148 + SP1912
SP2108/0148/1912 Fusion

Figure 16C

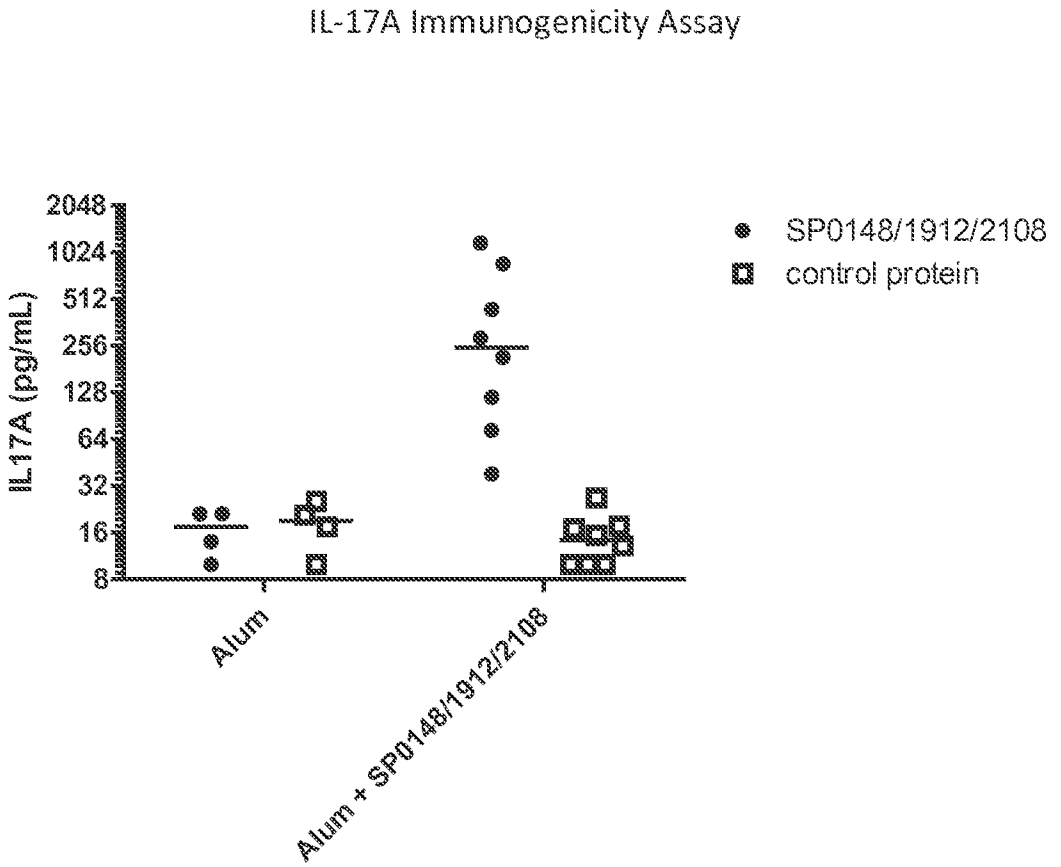


Figure 17, Panel A

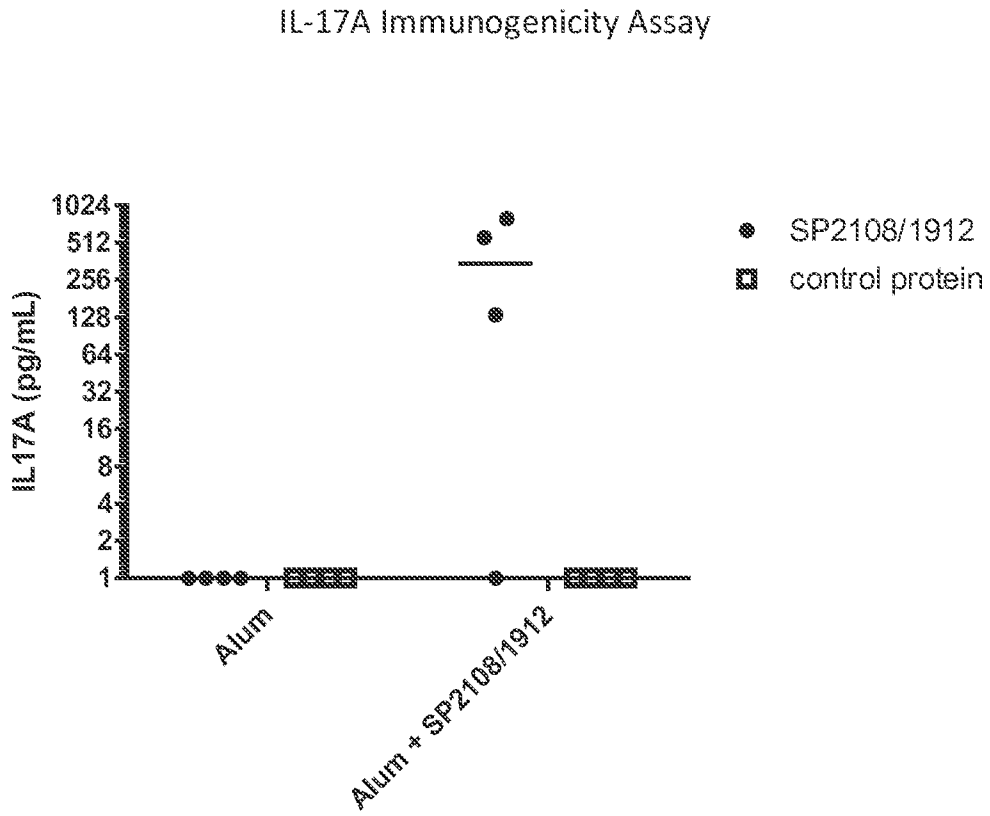


Figure 17, Panel B

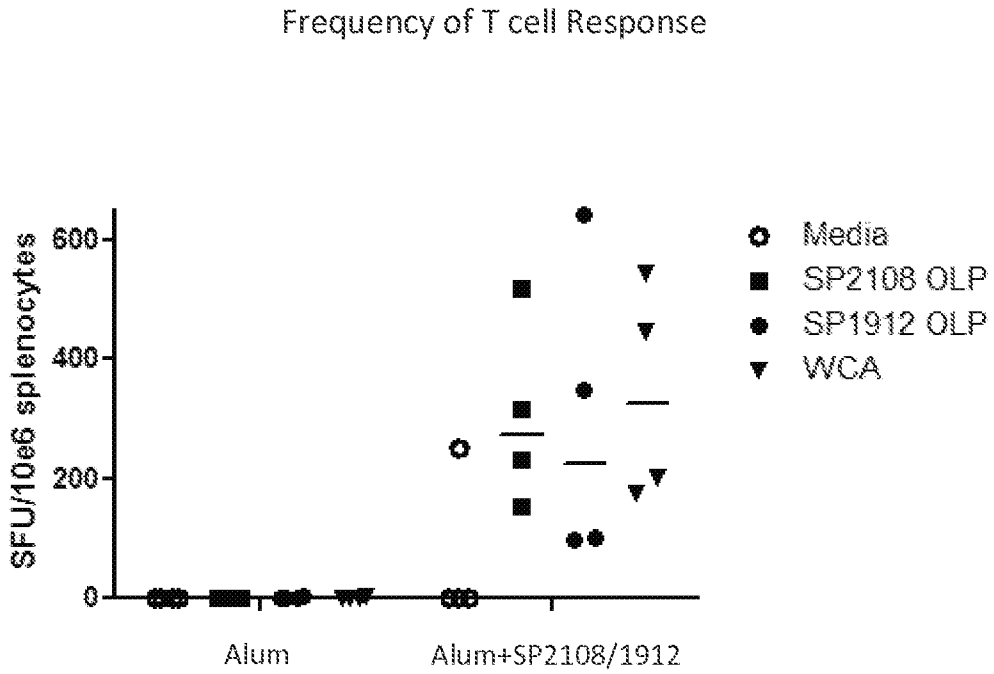

Frequency of T cell Response
SFU/10e6 splenocytes
600
400
200
0
Alum
Alum+SP2108/1912
Media
SP2108 OLP
SP1912 OLP
WCA

Figure 18

IL-17A Dose Response Assay

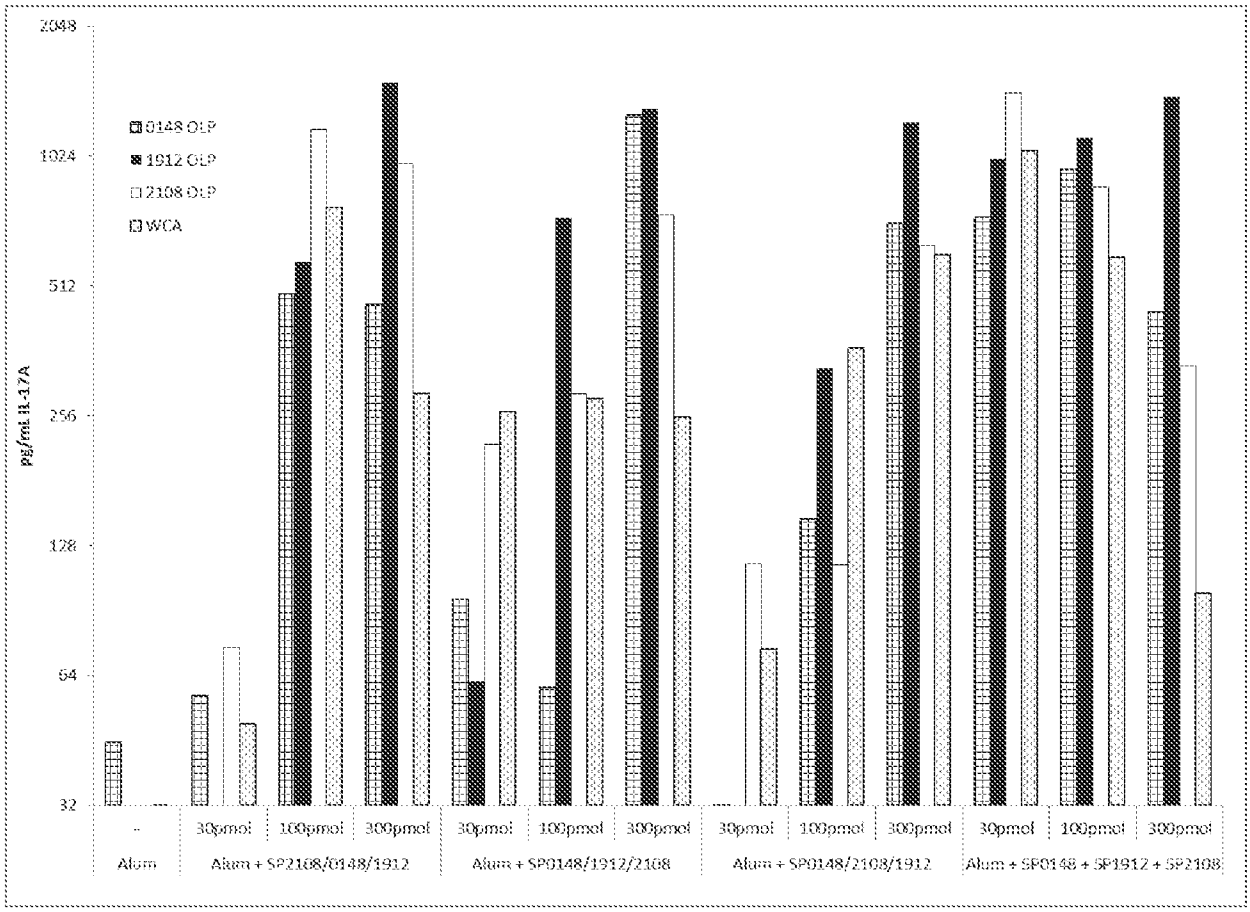


Figure 19, Panel A

IL-17A Dose Response Assay

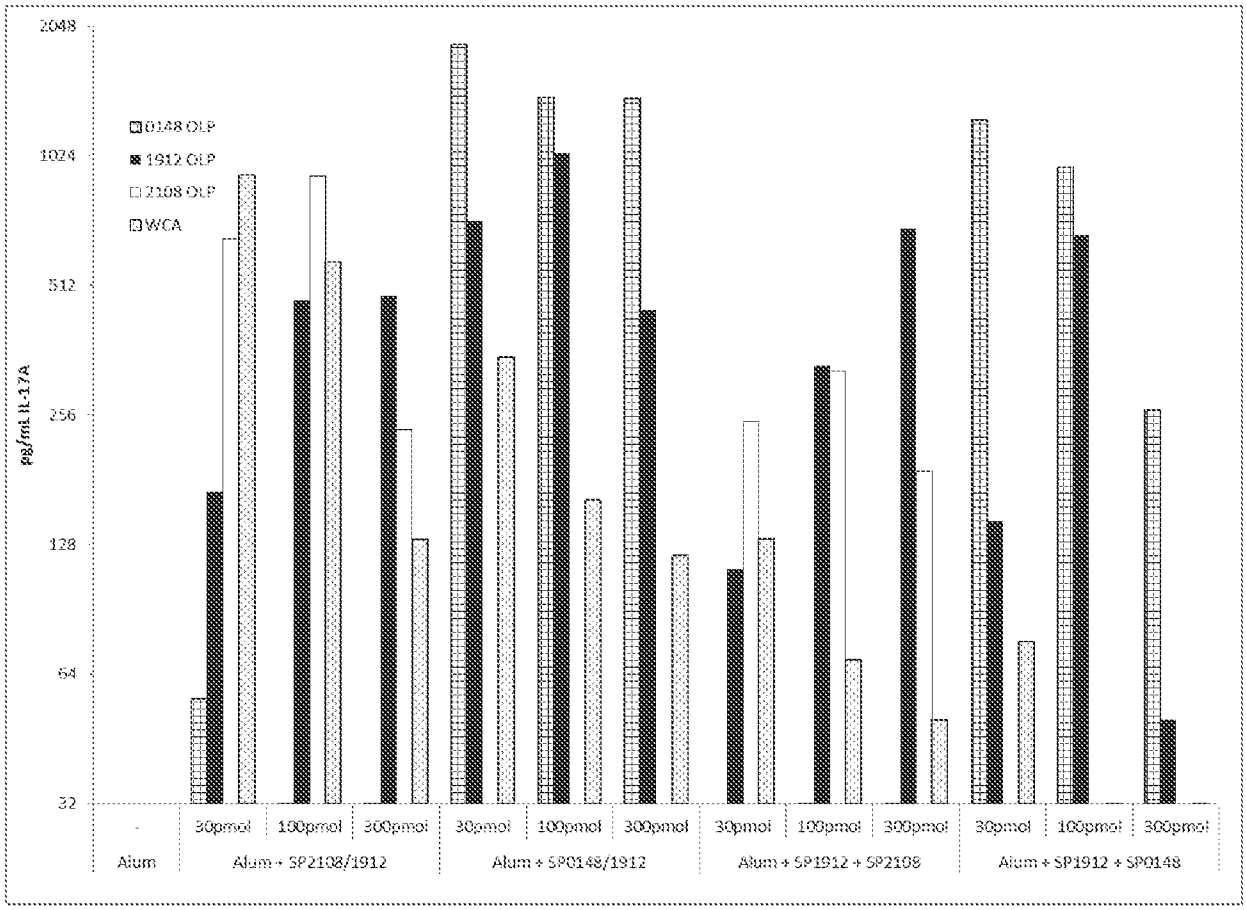


Figure 19, Panel B

IL-17A Immunogenicity Assay

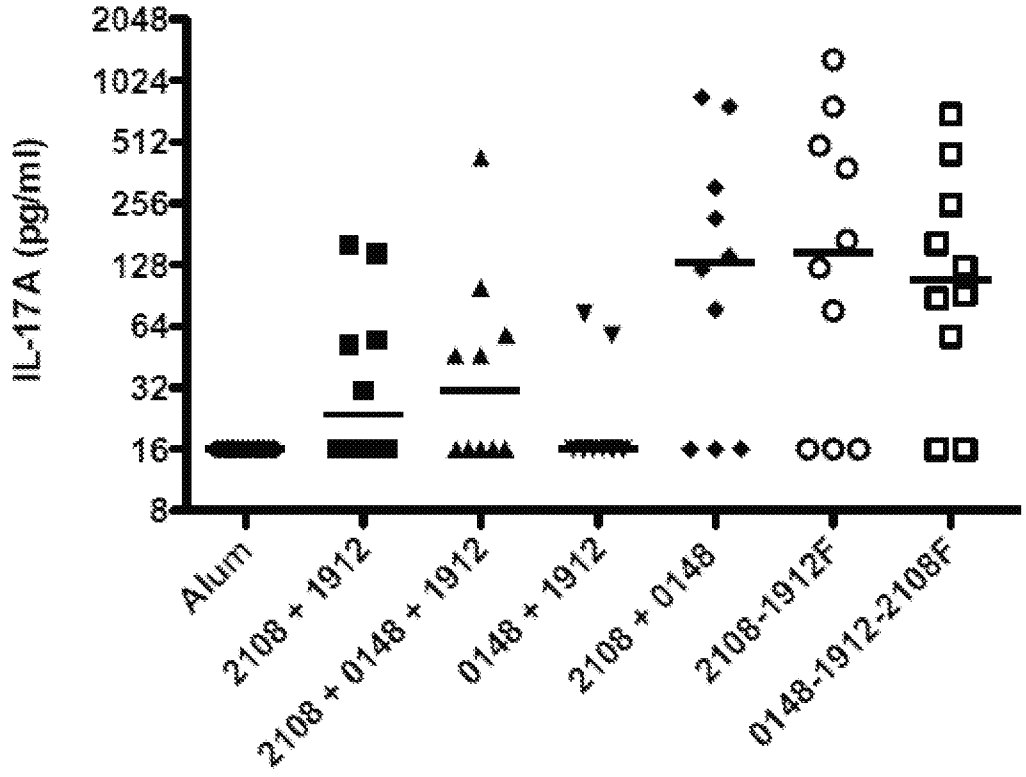
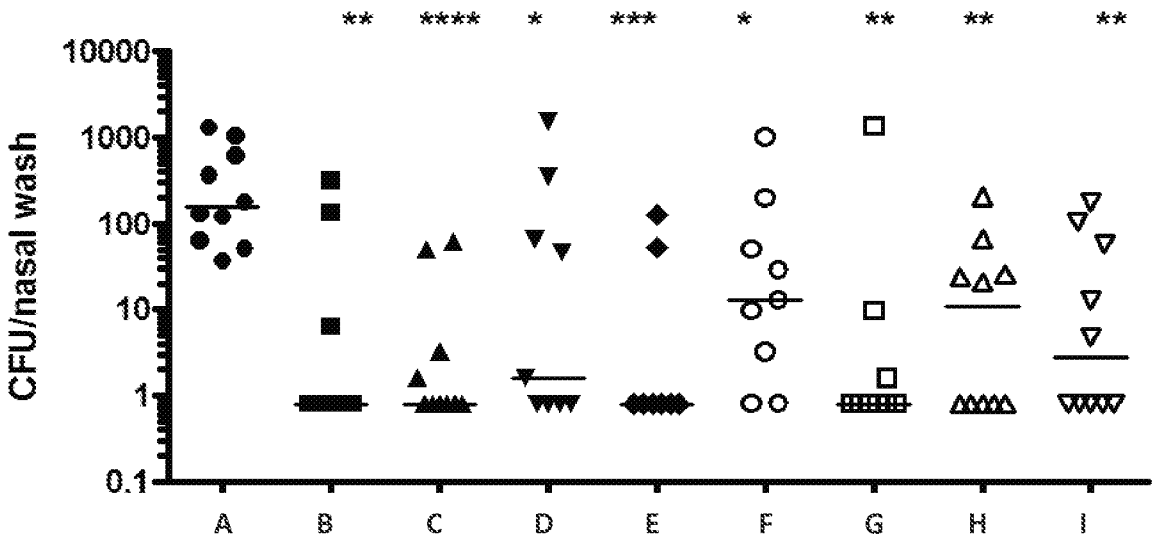


Figure 20

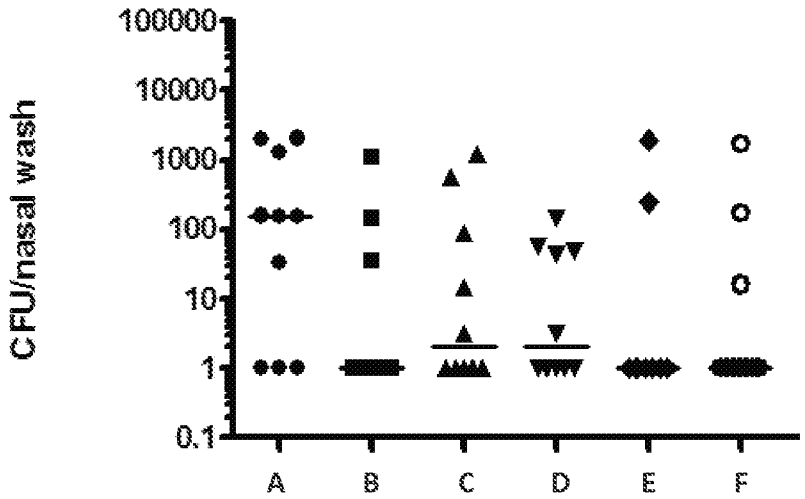
Intranasal Challenge: Colonization Assay



	A	B	C	D	E	F	G	H	I
immun group	alum	SP0148 + SP1912 + SP2108		SP2108 /1912F	SP0148 /1912F	SP0148 /2108/ 1912F	SP0148 /1912/ 2108F	SP2108/0148 /1912F	
dose		*1μg + 10μg* + 10μg	100 pmol each	100 pmol	100 pmol	100 pmol	100 pmol	100 pmol	20 pmol
# of mice protected	0	7/10	6/10	5/9	8/10	2/9	7/10	5/10	5/10
Log reduc (median)	0.00	2.28	2.28	1.98	2.28	1.07	2.28	1.57	1.89

Figure 21, Panel A

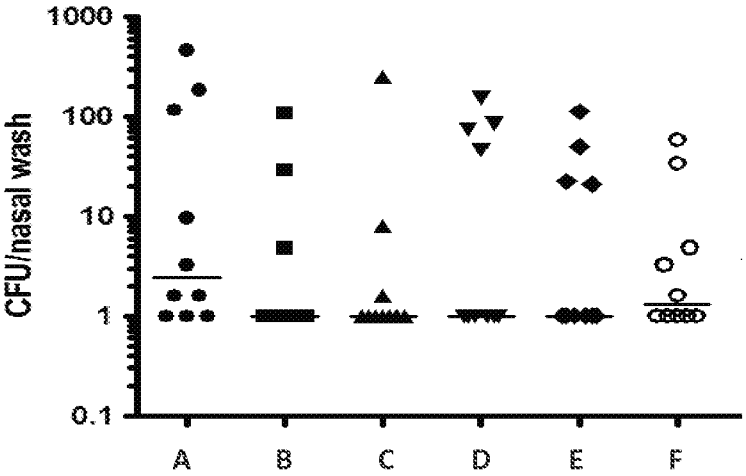
Intranasal Challenge: Colonization Assay



	A	B	C	D	E	F
immun group	alum	SP0148 + SP1912 + SP2108	SP0148/ 1912F	SP0148/ 1912F + SP2108	SP2108/ 1912F	SP2108/ 1912F + SP0148
# of mice protected	3/10	7/10	5/10	5/10	8/10	7/10
Log reduc (median)		1.1	0.8	1.0	1.2	1.1

Figure 21, Panel B

Intranasal Challenge: Colonization Assay



	A	B	C	D	E	F
Immun group	alum	SP0148 + SP1912 + SP2108	SP0148/ 1912F	SP0148/ 1912F + SP2108	SP2108/ 1912F	SP2108/ 1912F + SP0148
# of mice protected	3/10	7/10	7/10	6/10	6/10	5/10
Log reduc (median)		0.5	0.5	0.1	0.3	0.4

Figure 21, Panel C

Intranasal Challenge: Colonization Assay

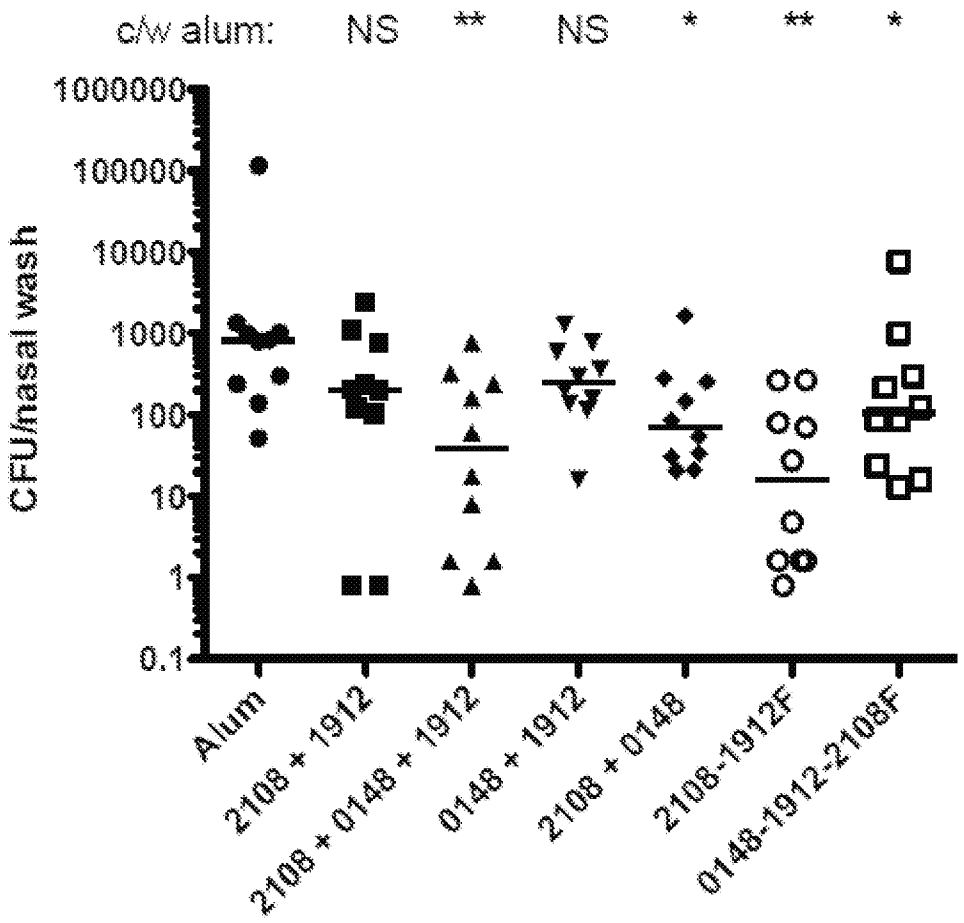


Figure 21, Panel D

IL-17A Immunogenicity Assay

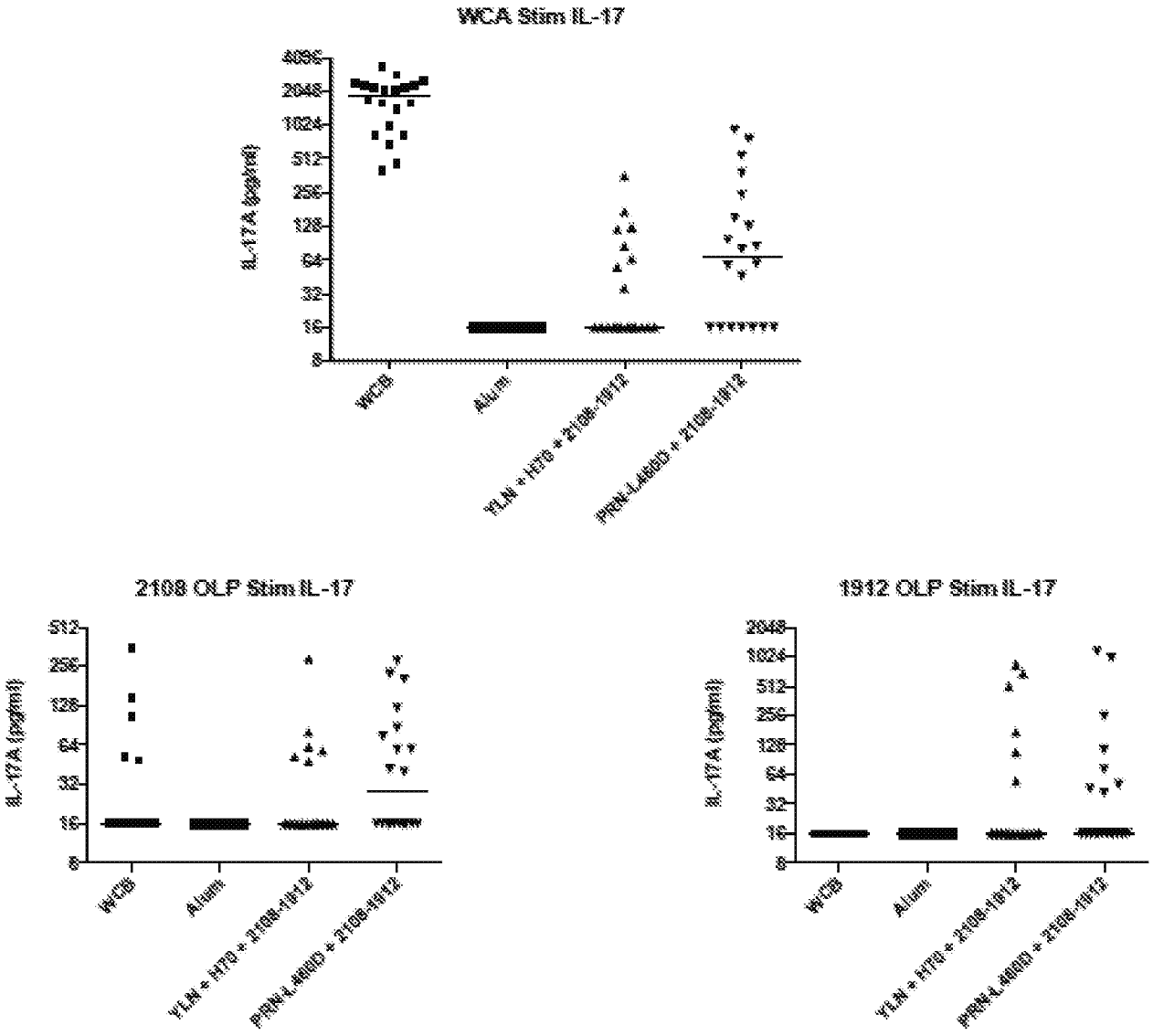


Figure 22, Panel A

IL-17A Immunogenicity Assay

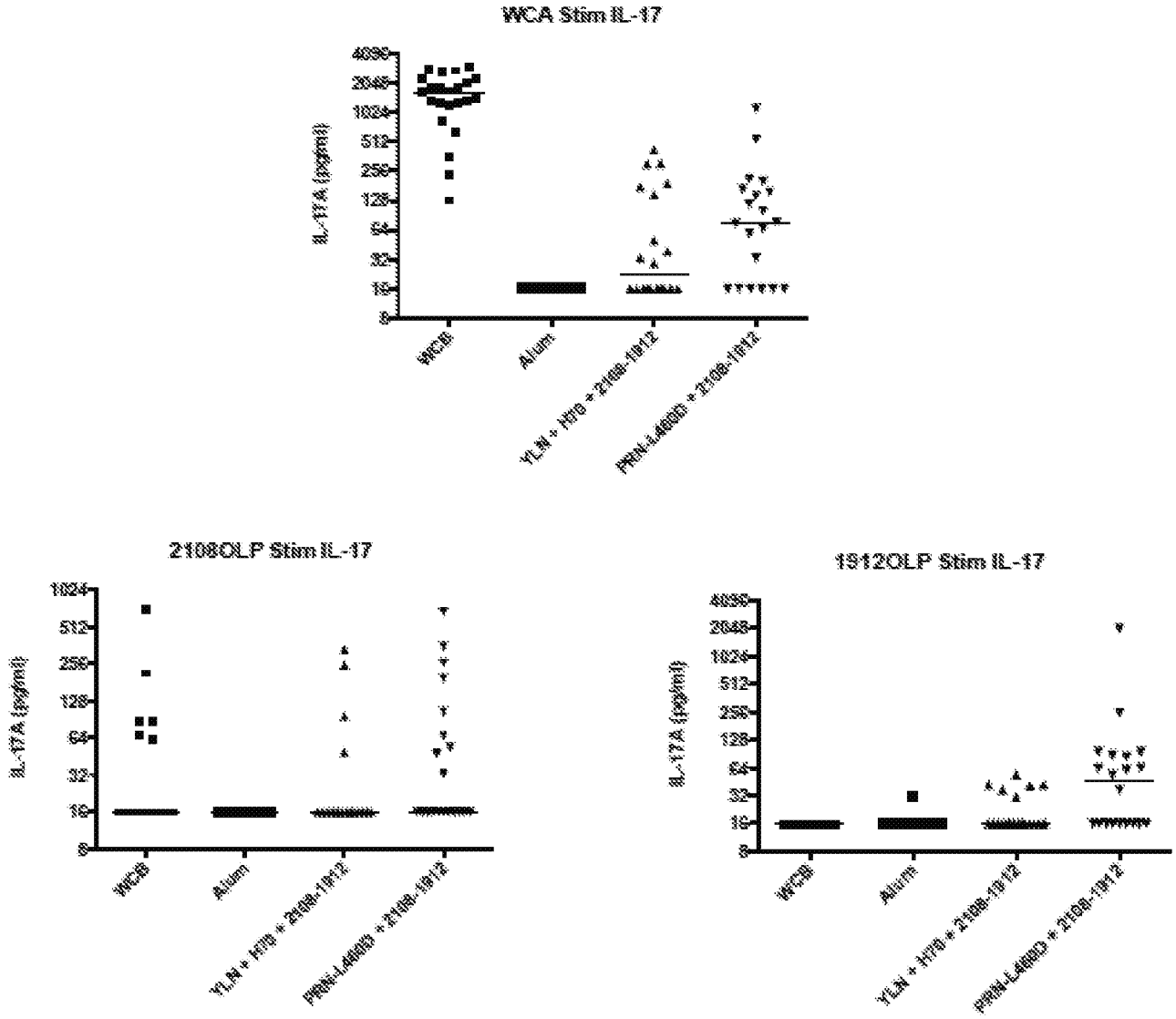


Figure 22, Panel B

Intranasal Challenge: Colonization Assay

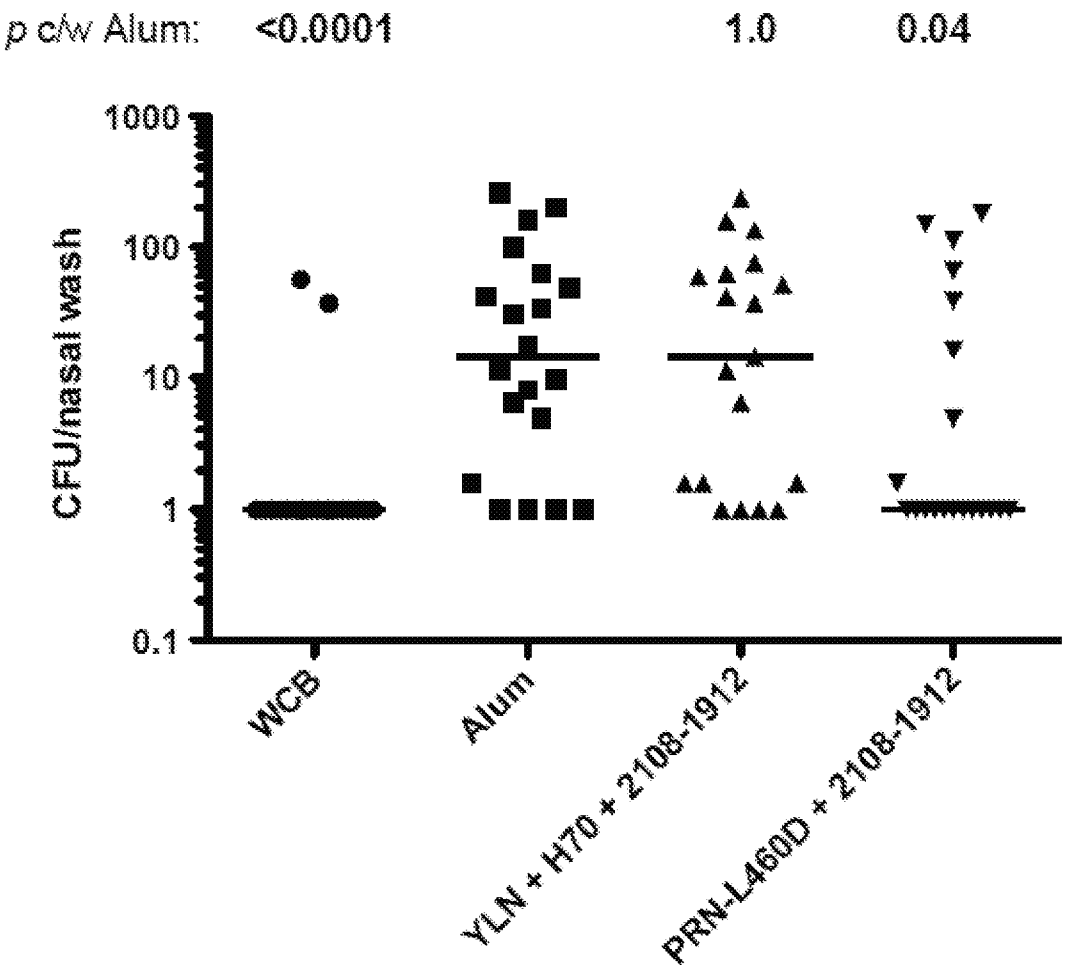


Figure 23, Panel A

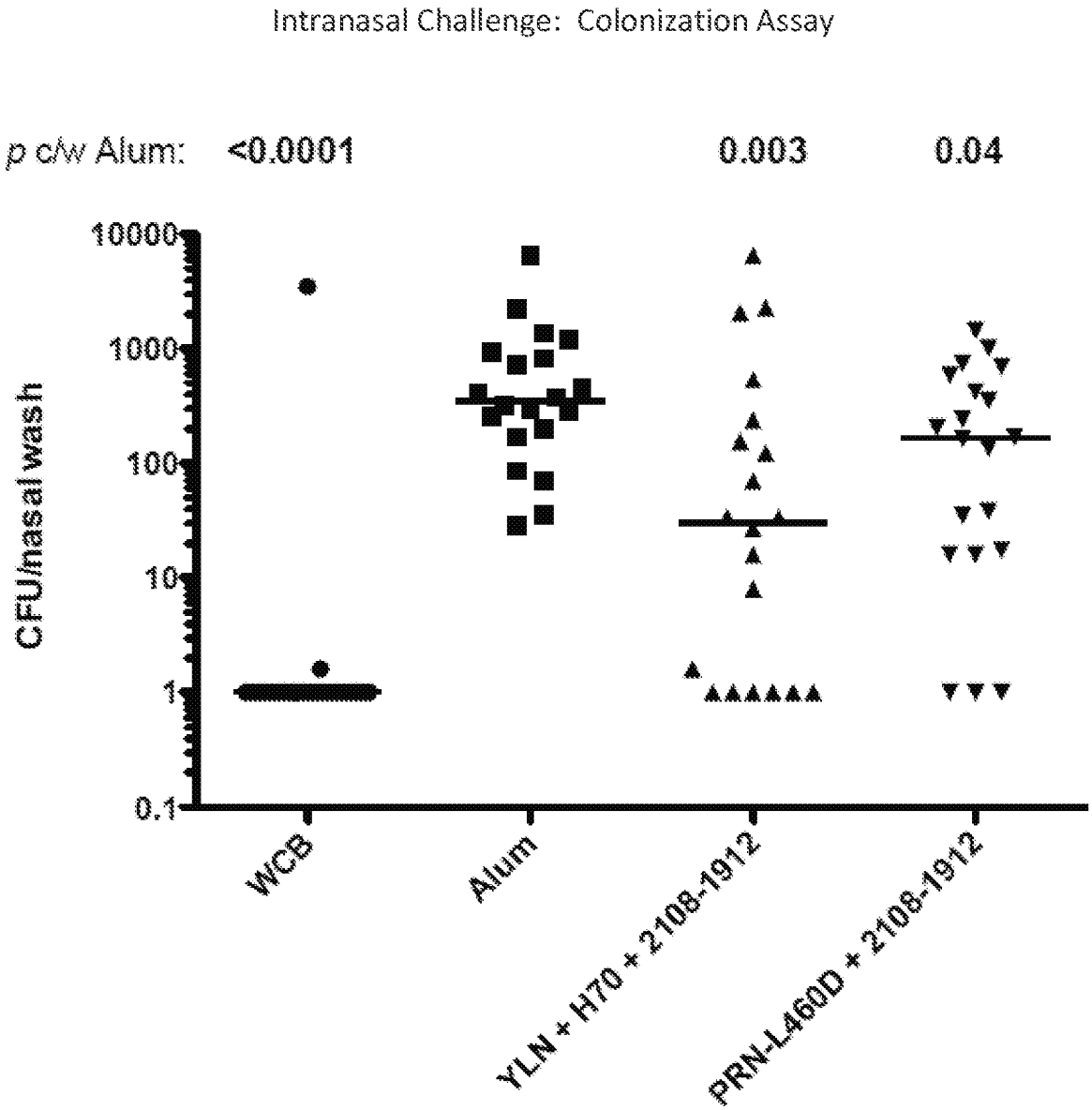
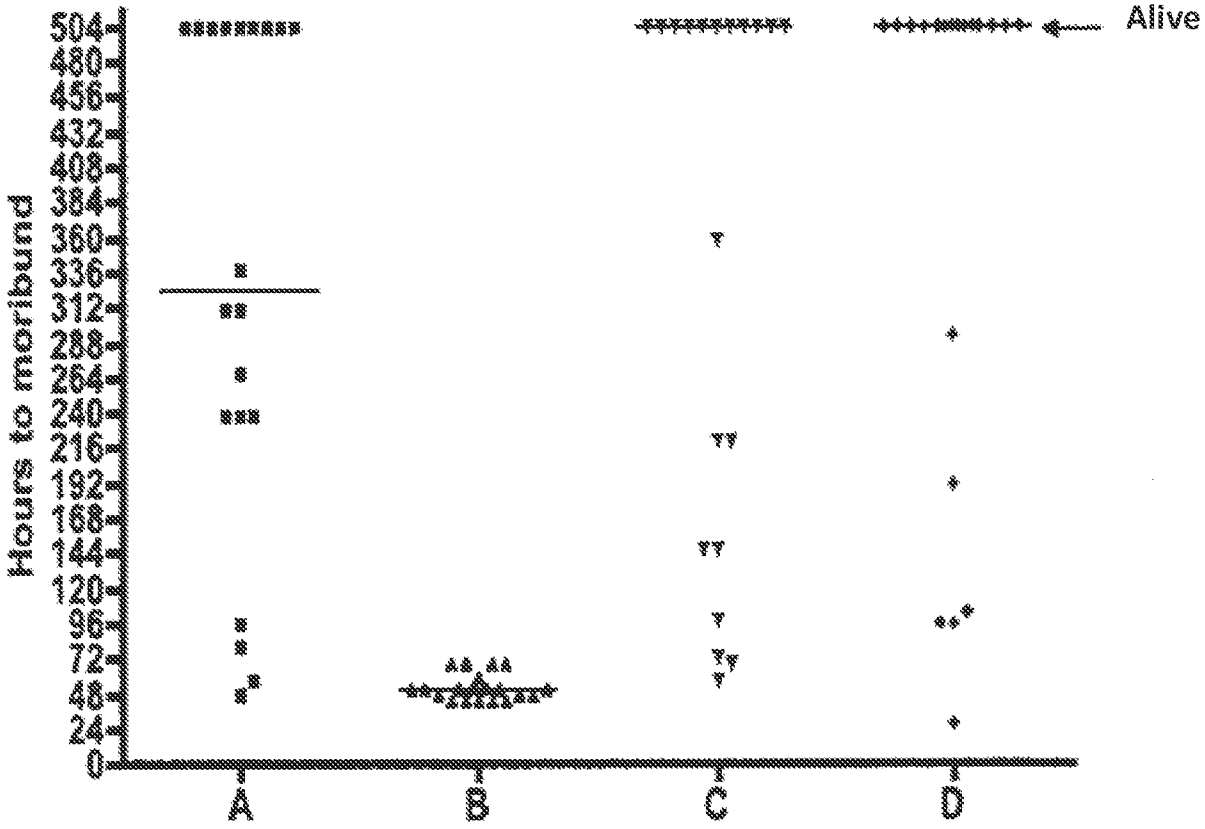


Figure 23, Panel B

Intravenous Challenge: Sepsis Assay



A = Alum

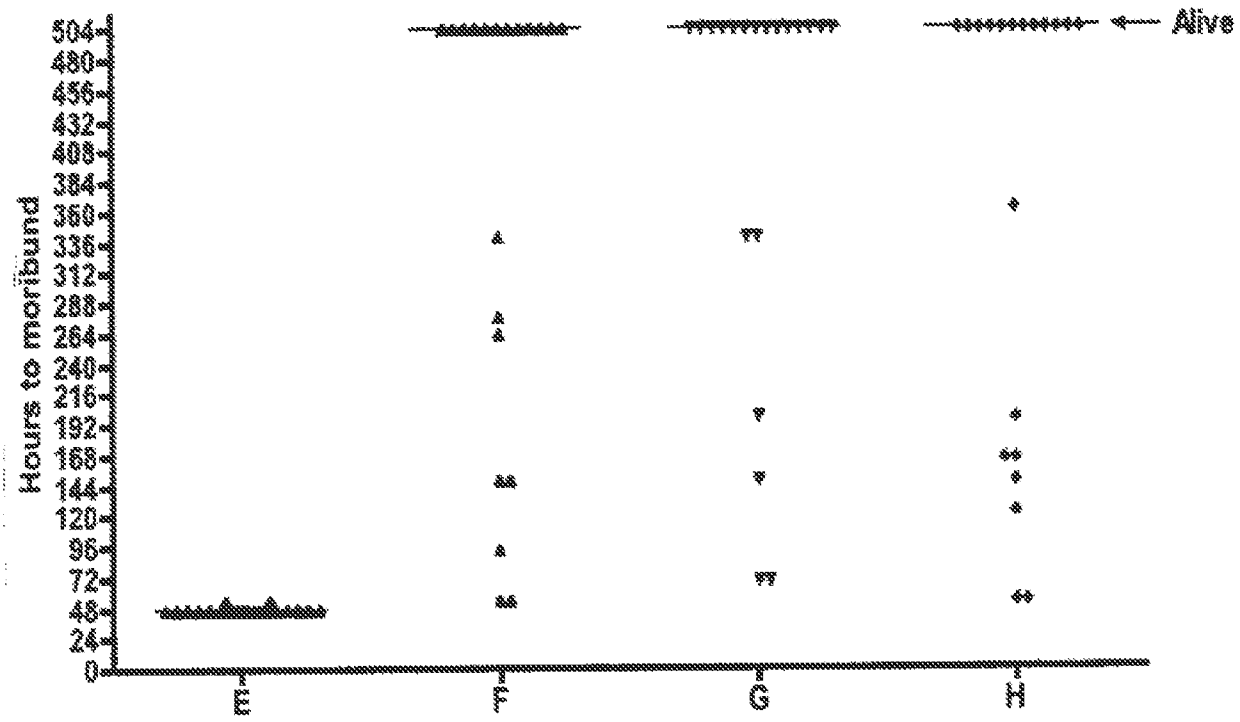
B = YLN + H70 + SP2108/1912

C = PRN-L460D + SP2108/1912

D = YLN + H70

Figure 24, Panel A

Intravenous Challenge: Sepsis Assay



E = Alum

F = YLN + H70 + SP2108/1912

G = PRN-L460D + SP2108/1912

H = YLN + H70

Figure 24, Panel B