US 20070009900A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2007/0009900 A1**

Doucette-Stamm et al. (43) **Pub. Date: Jan. 11, 2007**

(54) NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

(75) Inventors: **Lynn Doucette-Stamm**, Framingham, MA (US); **David Bush**, Somerville, MA (US); **Qiandong Zeng**, Waltham, MA (US); **Timothy Opperman**, Somerville, MA (US); **Chad Eric Houseweart**, Waltham, MA (US)

Correspondence Address:
HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
530 VIRGINIA ROAD
P.O. BOX 9133
CONCORD, MA 01742-9133 (US)

(73) Assignee: **Aventis Pasteur Ltd**, Toronto (CA)

(21) Appl. No.: **11/027,802**

(22) Filed: **Dec. 30, 2004**

Related U.S. Application Data

(63) Continuation of application No. 10/640,833, filed on Aug. 14, 2003, which is a continuation of application No. 09/583,110, filed on May 26, 2000, now Pat. No. 6,699,703, which is a continuation-in-part of application No. 09/107,433, filed on Jun. 30, 1998, now Pat. No. 6,800,744.

(60) Provisional application No. 60/085,131, filed on May 12, 1998. Provisional application No. 60/051,553, filed on Jul. 2, 1997.

Publication Classification

(51) **Int. Cl.**
C12Q 1/68 (2006.01)
C07H 21/04 (2006.01)
C12N 1/21 (2006.01)
C12N 15/74 (2006.01)

(52) **U.S. Cl.** **435/6**; 435/252.3; 435/320.1; 536/23.7

(57) ABSTRACT

The invention provides isolated polypeptide and nucleic acid sequences derived from *Streptococcus pneumoniae* that are useful in diagnosis and therapy of pathological conditions; antibodies against the polypeptides; and methods for the production of the polypeptides. The invention also provides methods for the detection, prevention and treatment of pathological conditions resulting from bacterial infection.

NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 10/640,833, filed Aug. 14, 2003, which is a continuation of U.S. application Ser. No. 09/583,110 (now U.S. Pat. No. 6,699,703) filed May 26, 2000, which is a continuation-in-part of U.S. application Ser. No. 09/107,433 (now U.S. Pat. No. 6,800,744), filed Jun. 30, 1998, which claims the benefit of U.S. Application No. 60/085,131, filed May 12, 1998 and of U.S. Application No. 60/051,553, filed Jul. 2, 1997. The entire teachings of the above applications are incorporated herein by reference.

INCORPORATION BY REFERENCE OF MATERIAL ON COMPACT DISK

[0002] This application incorporates by reference the Sequence Listing contained on the two compact disks (Copy 1 and Copy 2), filed concurrently herewith, containing the following file:

[0003] File name: 3687.1000-011SequenceList.txt; created Dec. 27, 2004, 8,134 KB in size.

[0004] This application also incorporates by reference Table 2 contained on the two compact disks (Copy 1 and Copy 2), filed concurrently herewith, containing the following file:

[0005] File name: 521614_1.txt; created Dec. 23, 2004, 351 KB in size.

TABLES FILED ON CD

The patent application contains tables filed on compact disc. These tables have been included at the end of the specification.

FIELD OF THE INVENTION

[0006] The invention relates to isolated nucleic acids and polypeptides derived from *Streptococcus pneumoniae* that are useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions, as well as materials and methods for the diagnosis, prevention, and amelioration of pathological conditions resulting from bacterial infection.

BACKGROUND OF THE INVENTION

[0007] *Streptococcus pneumoniae* (*S. pneumoniae*) is a common, spherical, gram-positive bacterium. Worldwide it is a leading cause of illness among children, the elderly, and individuals with debilitating medical conditions (Breiman, R. F. et al., 1994, JAMA 271: 1831). *S. pneumoniae* is estimated to be the causal agent in 3,000 cases of meningitis, 50,000 cases of bacteremia, 500,000 cases of pneumonia, and 7,000,000 cases of otitis media annually in the United States alone (Reichler, M. R. et al., 1992, J. Infect. Dis. 166: 1346; Stool, S. E. and Field, M. J., 1989 Pediatr. Infect. Dis J. 8: S11). In the United States alone, 40,000 deaths result annually from *S. pneumoniae* infections (Williams, W. W. et al., 1988 Ann. Intern. Med. 108: 616) with a death rate approaching 30% from bacteremia (Butler, J. C. et al., 1993, JAMA 270: 1826). Pneumococcal pneumonia is a serious problem among the elderly of industrialized nations (Käyhty, H. and Eskola, J., 1996 Emerg. Infect. Dis. 2: 289) and is a leading cause of death among children in developing nations (Käyhty, H. and Eskola, J., 1996 Emerg. Infect. Dis. 2: 289; Stansfield, S. K., 1987 Pediatr. Infect. Dis. 6: 622).

[0008] Vaccines against *S. pneumoniae* have been available for a number of years. There are a large number of serotypes based on the polysaccharide capsule (van Dam, J. E., Fleer, A., and Snippe, H., 1990 Antonie van Leeuwenhoek 58: 1) although only a fraction of the serotypes seem to be associated with infections (Martin, D. R. and Brett, M. S., 1996 N. Z. Med. J. 109: 288). A multivalent vaccine against capsular polysaccharides of 23 serotypes (Smart, L. E., Dougall, A. J. and Gridwood, R. W., 1987 J. Infect. 14: 209) has provided protection for some groups but not for several groups at risk for pneumococcal infections, such as infants and the elderly (Makel, P. H. et al., 1980 Lancet 2: 547; Sankilampi, U., 1996 J. Infect. Dis. 173: 387). Conjugated pneumococcal capsular polysaccharide vaccines have somewhat improved efficacy, but are costly and, therefore, are not likely to be in widespread use (Käyhty, H. and Eskola, J., 1996 Emerg. Infect. Dis. 2: 289).

[0009] At one time, *S. pneumoniae* strains were uniformly susceptible to penicillin. The report of a penicillin-resistant strain of (Hansman, D. and Bullen, M. M., 1967 Lancet 1: 264) was followed rapidly by many reports indicating the worldwide emergence of penicillin-resistant and penicillin non-susceptible strains (Klugman, K. P., 1990 Clin. Microbiol. Rev. 3: 171). *S. pneumoniae* strains which are resistant to multiple antibiotics (including penicillin) have also been observed recently within the United States (Welby, P. L., 1994 Pediatr. Infect. Dis. J. 13: 281; Ducin, J. S. et al., 1995 Pediatr. Infect. Dis. J. 14: 745; Butler, J. C., 1996 J. Infect. Dis. 174: 986) as well as internationally (Boswell, T. C. et al., 1996; J. Infect. 33: 17; Catchpole, C., Fraise, A., and Wise, R., 1996 Microb. Drug Resist. 2: 431; Tarasi, A. et al., 1997 Microb. Drug Resist. 3: 105).

[0010] A high incidence of morbidity is associated with invasive *S. pneumoniae* infections (Williams, W. W. et al., 1988 Ann. Intern. Med. 108: 616). Because of the incomplete effectiveness of currently available vaccines and antibiotics, the identification of new targets for antimicrobial therapies, including, but not limited to, the design of vaccines and antibiotics, which may help prevent infection or that may be useful in fighting existing infections, is highly desirable.

SUMMARY OF THE INVENTION

[0011] The present invention fulfills the need for diagnostic tools and therapeutics by providing bacterial-specific compositions and methods for detecting, treating, and preventing bacterial infection, in particular *S. pneumoniae* infection.

[0012] The present invention encompasses isolated polypeptides and nucleic acids derived from *S. pneumoniae* that are useful as reagents for diagnosis of bacterial infection, components of effective antibacterial vaccines, and/or as targets for antibacterial drugs, including anti-*S. pneumoniae* drugs. The nucleic acids and peptides of the present invention also have utility for diagnostics and therapeutics for *S. pneumoniae* and other *Streptococcus* species. They can also be used to detect the presence of *S. pneumoniae* and other *Streptococcus* species in a sample; and in screening compounds for the ability to interfere with the *S. pneumoniae* life cycle or to inhibit *S. pneumoniae* infection. More specifically, this invention features compositions of nucleic acids corresponding to entire coding sequences of *S. pneumoniae* proteins, including surface or secreted proteins or parts thereof, nucleic acids capable of binding mRNA from *S. pneumoniae* proteins to block protein translation, and methods for producing *S. pneumoniae* proteins or parts thereof using peptide synthesis and recombinant DNA techniques. This invention also features antibodies and nucleic acids useful as probes to detect *S. pneumoniae* infection. In addition, vaccine compositions and methods for the protection or treatment of infection by *S. pneumoniae* are within the scope of this invention.

[0013] The nucleotide sequences provided in SEQ ID NO: 1-SEQ ID NO: 2661, a fragment thereof, or a nucleotide sequence at least 99.5% identical to a sequence contained within SEQ ID NO: 1-SEQ ID NO: 2661 may be "provided" in a variety of medias to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid molecule, which contains a nucleotide sequence of the present invention, i.e., the nucleotide sequence provided in SEQ ID NO: 1-SEQ ID NO: 2661, a fragment thereof, or a nucleotide sequence at least 99.5% identical to a sequence contained within SEQ ID NO: 1-SEQ ID NO: 2661. Uses for and methods for providing nucleotide sequences in a variety of media is well known in the art (see e.g., EPO Publication No. EP 0 756 006).

[0014] In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any media which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage media, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A person skilled in the art can readily appreciate how any of the presently known computer readable media can be used to create a manufacture comprising computer readable media having recorded thereon a nucleotide sequence of the present invention.

[0015] As used herein, "recorded" refers to a process for storing information on computer readable media. A person skilled in the art can readily adopt any of the presently known methods for recording information on computer readable media to generate manufactures comprising the nucleotide sequence information of the present invention.

[0016] A variety of data storage structures are available to a person skilled in the art for creating a computer readable media having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable media. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A person skilled in the art can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable media having recorded thereon the nucleotide sequence information of the present invention.

[0017] By providing the nucleotide sequence of SEQ ID NO: 1-SEQ ID NO: 2661, a fragment thereof, or a nucleotide sequence at least 99.5% identical to a sequence contained within SEQ ID NO: 1-SEQ ID NO: 2661 in computer readable form, a person skilled in the art can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a person skilled in the art to access sequence information provided in a computer readable media. Examples of such computer software include programs of the "Staden Package", "DNA Star", "MacVector", GCG "Wisconsin Package" (Genetics Computer Group, Madison, Wis.) and "NCBI toolbox" (National Center for Biotechnology Information).

[0018] Computer algorithms enable the identification of *S. pneumoniae* open reading frames (ORFs) within SEQ ID NO: 1-SEQ ID NO: 2661 which contain homology to ORFs or proteins from other organisms. Examples of such similarity-search algorithms include the BLAST [Altschul et al., J. Mol. Biol. 215: 403-410 (1990)] and Smith-Waterman [Smith and Waterman (1981) Advances in Applied Mathematics, 2: 482-489] search algorithms. These algorithms are utilized on computer systems as exemplified below. The ORFs so identified represent protein encoding fragments within the *S. pneumoniae* genome and are useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

[0019] The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify commercially important fragments of the *S. pneumoniae* genome. As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A person skilled in the art can readily appreciate that any one of the currently available computer-based systems is suitable for use in the present invention. The computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

[0020] As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the *S. pneumoniae* genome which are similar to, or "match", a particular target sequence or target motif. A variety of known algorithms are known in the art and have been disclosed publicly, and a variety of commercially available software for conducting homology-based similarity searches are available and can be used in the computer-based systems of the present invention. Examples of such software include, but is not limited to, FASTA (GCG Wisconsin Package), Bic_SW (Compugen Bioccelerator, BLASTN2, BLASTP2 and BLASTX2 (NCBI) and Motifs (GCG). BLASTN2, A person skilled in the art can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

[0021] As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A person skilled in the art can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that many genes are longer than 500 amino acids, or 1.5 kb in length, and that commercially important fragments of the *S. pneumoniae* genome, such as sequence fragments involved in gene expression and protein processing, will often be shorter than 30 nucleotides.

[0022] As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a specific functional domain or three-dimensional configuration which is formed upon the folding of the target polypeptide. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymatic active sites, membrane spanning regions, and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

[0023] A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the *S. pneumoniae* genome possessing varying degrees of homology to the target sequence or target motif. Such presentation provides a person skilled in the art with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

[0024] A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments of the *S. pneumoniae* genome. In the present examples, implementing software which implement the BLASTP2 and bic_SW algorithms (Altschul et al., J. Mol. Biol. 215: 403-410 (1990); Compugen Biocellerator) was used to identify open reading frames within the *S. pneumoniae* genome. A person skilled in the art can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention.

[0025] The invention features *S. pneumoniae* polypeptides, preferably a substantially pure preparation of an *S. pneumoniae* polypeptide, or a recombinant *S. pneumoniae* polypeptide. In preferred embodiments: the polypeptide has biological activity; the polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% identical to an amino acid sequence of the invention contained in the Sequence Listing, preferably it has about 65% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing, and most preferably it has about 92% to about 99% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide has an amino acid sequence essentially the same as an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acid residues in length; the polypeptide includes at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acid residues of the invention contained in the Sequence Listing. In yet another preferred embodiment, the amino acid sequence which differs in sequence identity by about 7% to about 8% from the *S. pneumoniae* amino acid sequences of the invention contained in the Sequence Listing is also encompassed by the invention.

[0026] In preferred embodiments: the *S. pneumoniae* polypeptide is encoded by a nucleic acid of the invention contained in the Sequence Listing, or by a nucleic acid having at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a nucleic acid of the invention contained in the Sequence Listing.

[0027] In a preferred embodiment, the subject *S. pneumoniae* polypeptide differs in amino acid sequence at 1, 2, 3, 5, 10 or more residues from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that the *S. pneumoniae* polypeptide exhibits an *S. pneumoniae* biological activity, e.g., the *S. pneumoniae* polypeptide retains a biological activity of a naturally occurring *S. pneumoniae* enzyme.

[0028] In preferred embodiments, the polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

[0029] In yet other preferred embodiments, the *S. pneumoniae* polypeptide is a recombinant fusion protein having a first *S. pneumoniae* polypeptide portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to *S. pneumoniae*. The second polypeptide portion can be, e.g., any of glutathione-S-transferase, a DNA binding domain, or a polymerase activating domain. In preferred embodiment the fusion protein can be used in a two-hybrid assay.

[0030] Polypeptides of the invention include those which arise as a result of alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events.

[0031] In a preferred embodiment, the encoded *S. pneumoniae* polypeptide differs (e.g., by amino acid substitution, addition or deletion of at least one amino acid residue) in amino acid sequence at 1, 2, 3, 5, 10 or more residues, from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that: the *S. pneumoniae* encoded polypeptide exhibits a *S. pneumoniae* biological activity, e.g., the encoded *S. pneumoniae* enzyme retains a biological activity of a naturally occurring *S. pneumoniae*.

[0032] In preferred embodiments, the encoded polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

[0033] The *S. pneumoniae* strain, 14453, from which genomic sequences have been sequenced, has been deposited on Jun. 26, 1997 in the American Type Culture Collection and assigned the ATCC designation # 55987.

[0034] Included in the invention are: allelic variations; natural mutants; induced mutants; proteins encoded by DNA that hybridize under high or low stringency conditions to a nucleic acid which encodes a polypeptide of the invention contained in the Sequence Listing (for definitions of high and low stringency see Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1-6.3.6, hereby incorporated by reference); and, polypeptides specifically bound by antisera to *S. pneumoniae* polypeptides, especially by antisera to an active site or binding domain of *S. pneumoniae* polypeptide. The invention also includes fragments, preferably biologically active fragments. These and other polypeptides are also referred to herein as *S. pneumoniae* polypeptide analogs or variants.

[0035] The invention further provides nucleic acids, e.g., RNA or DNA, encoding a polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense single strands.

[0036] In preferred embodiments, the subject *S. pneumoniae* nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the *S. pneumoniae* gene sequence, e.g., to render the *S. pneumoniae* gene sequence suitable for expression in a recombinant host cell.

[0037] In yet a further preferred embodiment, the nucleic acid which encodes an *S. pneumoniae* polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 8 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 12 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 20 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 40 consecutive nucleotides of the invention contained in the Sequence Listing.

[0038] In another aspect, the invention provides a substantially pure nucleic acid having a nucleotide sequence which encodes an *S. pneumoniae* polypeptide. In preferred embodiments: the encoded polypeptide has biological activity; the encoded polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous to an amino acid sequence of the invention contained in the Sequence Listing; the encoded polypeptide has an amino acid sequence essentially the same as an amino acid sequence of the invention contained in the Sequence Listing; the encoded polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the encoded polypeptide comprises at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acids of the invention contained in the Sequence Listing.

[0039] In another aspect, the invention encompasses: a vector including a nucleic acid which encodes an *S. pneumoniae* polypeptide or an *S. pneumoniae* polypeptide variant as described herein; a host cell transfected with the vector; and a method of producing a recombinant *S. pneumoniae* polypeptide or *S. pneumoniae* polypeptide variant; including culturing the cell, e.g., in a cell culture medium, and isolating an *S. pneumoniae* polypeptide or an *S. pneumoniae* polypeptide variant, e.g., from the cell or from the cell culture medium.

[0040] In another series of embodiments, the invention provides isolated nucleic acids comprising sequences at least about 8 nucleotides in length, more preferably at least about 12 nucleotides in length, and most preferably at least about 15-20 nucleotides in length, that correspond to a subsequence of any one of SEQ ID NO: 1-SEQ ID NO: 2661 or complements thereof. Alternatively, the nucleic acids comprise sequences contained within any ORF (open reading frame), including a complete protein-coding sequence, of which any of SEQ ID NO: 1-SEQ ID NO: 2661 forms a part. The invention encompasses sequence-conservative variants and function-conservative variants of these sequences. The nucleic acids may be DNA, RNA, DNA/RNA duplexes, protein-nucleic acid (PNA), or derivatives thereof.

[0041] In another aspect, the invention features, a purified recombinant nucleic acid having at least 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a sequence of the invention contained in the Sequence Listing.

[0042] In another aspect, the invention features nucleic acids capable of binding mRNA of *S. pneumoniae*. Such nucleic acid is capable of acting as antisense nucleic acid to control the translation of mRNA of *S. pneumoniae*. A further aspect features a nucleic acid which is capable of binding specifically to an *S. pneumoniae* nucleic acid. These nucleic acids are also referred to herein as complements and have utility as probes and as capture reagents.

[0043] In another aspect, the invention features an expression system comprising an open reading frame corresponding to *S. pneumoniae* nucleic acid. The nucleic acid further comprises a control sequence compatible with an intended host. The expression system is useful for making polypeptides corresponding to *S. pneumoniae* nucleic acid.

[0044] In another aspect, the invention features a cell transformed with the expression system to produce *S. pneumoniae* polypeptides.

[0045] In yet another embodiment, the invention encompasses reagents for detecting bacterial infection, including *S. pneumoniae* infection, which comprise at least one *S. pneumoniae-derived* nucleic acid defined by any one of SEQ ID NO: 1-SEQ ID NO: 2661, or sequence-conservative or function-conservative variants thereof. Alternatively, the diagnostic reagents comprise polypeptide sequences that are contained within any open reading frames (ORFs), including complete protein-coding sequences, contained within any of SEQ ID NO: 1-SEQ ID NO: 2661, or polypeptide sequences contained within any of SEQ ID NO: 2662-SEQ ID NO: 5322, or polypeptides of which any of the above sequences forms a part, or antibodies directed against any of the above peptide sequences or function-conservative variants and/or fragments thereof.

[0046] The invention further provides antibodies, preferably monoclonal antibodies, which specifically bind to the polypeptides of the invention. Methods are also provided for producing antibodies in a host animal. The methods of the invention comprise immunizing an animal with at least one *S. pneumoniae-derived* immunogenic component, wherein the immunogenic component comprises one or more of the polypeptides encoded by any one of SEQ ID NO: 1-SEQ ID NO: 2661 or sequence-conservative or function-conservative variants thereof; or polypeptides that are contained within any ORFs, including complete protein-coding sequences, of which any of SEQ ID NO: 1-SEQ ID NO: 2661 forms a part; or polypeptide sequences contained within any of SEQ ID NO: 2662-SEQ ID NO: 5322; or polypeptides of which any of SEQ ID NO: 2662-SEQ ID NO: 5322 forms a part. Host animals include any warm blooded animal, including without limitation mammals and birds. Such antibodies have utility as reagents for immunoassays to evaluate the abundance and distribution of *S. pneumoniae*-specific antigens.

[0047] In yet another aspect, the invention provides a method for detecting bacterial antigenic components in a sample, which comprises the steps of: (i) contacting a sample suspected to contain a bacterial antigenic component with a bacterial-specific antibody, under conditions in which a stable antigen-antibody complex can form between the antibody and bacterial antigenic components in the sample; and (ii) detecting any antigen-antibody complex formed in step (i), wherein detection of an antigen-antibody complex indicates the presence of at least one bacterial antigenic component in the sample. In different embodiments of this method, the antibodies used are directed against a sequence encoded by any of SEQ ID NO: 1-SEQ ID NO: 2661 or sequence-conservative or function-conservative variants thereof, or against a polypeptide sequence contained in any of SEQ ID NO: 2662-SEQ ID NO: 5322 or function-conservative variants thereof.

[0048] In yet another aspect, the invention provides a method for detecting antibacterial-specific antibodies in a sample, which comprises: (i) contacting a sample suspected to contain antibacterial-specific antibodies with a *S. pneumoniae* antigenic component, under conditions in which a stable antigen-antibody complex can form between the *S. pneumoniae* antigenic component and antibacterial antibodies in the sample; and (ii) detecting any antigen-antibody complex formed in step (i), wherein detection of an antigen-antibody complex indicates the presence of antibacterial antibodies in the sample. In different embodiments of this method, the antigenic component is encoded by a sequence contained in any of SEQ ID NO: 1-SEQ ID NO: 2661 or sequence-conservative and function-conservative variants thereof, or is a polypeptide sequence contained in any of SEQ ID NO: 2662-SEQ ID NO: 5322 or function-conservative variants thereof.

[0049] In another aspect, the invention features a method of generating vaccines for immunizing an individual against *S. pneumoniae*. The method includes: immunizing a subject with an *S. pneumoniae* polypeptide, e.g., a surface or secreted polypeptide, or active portion thereof, and a pharmaceutically acceptable carrier. Such vaccines have therapeutic and prophylactic utilities.

[0050] In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *S. pneumoniae* polypeptide. The method includes: contacting the candidate compound with an *S. pneumoniae* polypeptide and determining if the compound binds or otherwise interacts with an *S. pneumoniae* polypeptide. Compounds which bind *S. pneumoniae* are candidates as activators or inhibitors of the bacterial life cycle These assays can be performed in vitro or in vivo.

[0051] In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *S. pneumoniae* nucleic acid, e.g., DNA or RNA. The method includes: contacting the candidate compound with an *S. pneumoniae* nucleic acid and determining if the compound binds or otherwise interacts with an *S. pneumoniae* polypeptide. Compounds which bind *S. pneumoniae* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be performed in vitro or in vivo.

DETAILED DESCRIPTION OF THE INVENTION

[0052] The sequences of the present invention include the specific nucleic acid and amino acid sequences set forth in the Sequence Listing that forms a part of the present specification, and which are designated SEQ ID NO: 1-SEQ ID NO: 5322. Use of the terms "SEQ ID NO: 1-SEQ ID NO: 2661", "SEQ ID NO: 2662-SEQ ID NO: 5322", "the sequences depicted in Table 2", etc., is intended, for convenience, to refer to each individual SEQ ID NO individually, and is not intended to refer to the genus of these sequences. In other words, it is a shorthand for listing all of these sequences individually. The invention encompasses each sequence individually, as well as any combination thereof.

Definitions

[0053] "Nucleic acid" or "polynucleotide" as used herein refers to purine- and pyrimidine-containing polymers of any length, either polyribonucleotides or polydeoxyribonucleotides or mixed polyribo-polydeoxyribo nucleotides. This includes single- and double-stranded molecules, i.e., DNA-

DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases.

[0054] A nucleic acid or polypeptide sequence that is "derived from" a designated sequence refers to a sequence that corresponds to a region of the designated sequence. For nucleic acid sequences, this encompasses sequences that are homologous or complementary to the sequence, as well as "sequence-conservative variants" and "function-conservative variants." For polypeptide sequences, this encompasses "function-conservative variants." Sequence-conservative variants are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position. Function-conservative variants are those in which a given amino acid residue in a polypeptide has been changed without altering the overall conformation and function of the native polypeptide, including, but not limited to, replacement of an amino acid with one having similar physico-chemical properties (such as, for example, acidic, basic, hydrophobic, and the like). "Function-conservative" variants also include any polypeptides that have the ability to elicit antibodies specific to a designated polypeptide.

[0055] An "*S. pneumoniae-derived*" nucleic acid or polypeptide sequence may or may not be present in other bacterial species, and may or may not be present in all *S. pneumoniae* strains. This term is intended to refer to the source from which the sequence was originally isolated. Thus, a *S. pneumoniae-derived* polypeptide, as used herein, may be used, e.g., as a target to screen for a broad spectrum antibacterial agent, to search for homologous proteins in other species of bacteria or in eukaryotic organisms such as fungi and humans, etc.

[0056] A purified or isolated polypeptide or a substantially pure preparation of a polypeptide are used interchangeably herein and, as used herein, mean a polypeptide that has been separated from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10, or 100 mg of the polypeptide.

[0057] A purified preparation of cells refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

[0058] A purified or isolated or a substantially pure nucleic acid, e.g., a substantially pure DNA, (are terms used interchangeably herein) is a nucleic acid which is one or both of the following: not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally-occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *S. pneumoniae* DNA sequence.

[0059] A "contig" as used herein is a nucleic acid representing a continuous stretch of genomic sequence of an organism.

[0060] An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a polypeptide. This region usually represents the total coding region for the polypeptide and can be determined from a stop to stop codon or from a start to stop codon.

[0061] As used herein, a "coding sequence" is a nucleic acid which is transcribed into messenger RNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop codon at the three prime terminus. A coding sequence can include but is not limited to messenger RNA, synthetic DNA, and recombinant nucleic acid sequences.

[0062] A "complement" of a nucleic acid as used herein refers to an anti-parallel or antisense sequence that participates in Watson-Crick base-pairing with the original sequence.

[0063] A "gene product" is a protein or structural RNA which is specifically encoded by a gene.

[0064] As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest. Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are readily discernable to one of ordinary skill in the art using routine experimentation.

[0065] "Homologous" refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared× 100. For example, if 6 of 10 of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give maximum homology.

[0066] Nucleic acids are hybridizable to each other when at least one strand of a nucleic acid can anneal to the other nucleic acid under defined stringency conditions. Stringency of hybridization is determined by: (a) the temperature at which hybridization and/or washing is performed; and (b) the ionic strength and polarity of the hybridization and washing solutions. Hybridization requires that the two nucleic acids contain complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. Typically, hybridization of two sequences at high stingency (such as, for example, in a solution of 0.5×SSC, at 65° C.) requires that the sequences be essentially completely homologous. Conditions of intermediate stringency (such as, for example, 2×SSC at 65° C.) and low stringency (such as, for example 2×SSC at 55° C.), require correspondingly less overall complementarity between the hybridizing sequences. (1×SSC is 0.15 M NaCl, 0.015 M Na citrate).

[0067] The terms peptides, proteins, and polypeptides are used interchangeably herein.

[0068] As used herein, the term "surface protein" refers to all surface accessible proteins, e.g. inner and outer membrane proteins, proteins adhering to the cell wall, and secreted proteins.

[0069] A polypeptide has *S. pneumoniae* biological activity if it has one, two and preferably more of the following properties: (1) if when expressed in the course of an *S. pneumoniae* infection, it can promote, or mediate the attachment of *S. pneumoniae* to a cell; (2) it has an enzymatic activity, structural or regulatory function characteristic of an *S. pneumoniae* protein; (3) or the gene which encodes it can rescue a lethal mutation in an *S. pneumoniae* gene. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the above-listed properties.

[0070] A biologically active fragment or analog is one having an in vivo or in vitro activity which is characteristic of the *S. pneumoniae* polypeptides of the invention contained in the Sequence Listing, or of other naturally occurring *S. pneumoniae* polypeptides, e.g., one or more of the biological activities described herein. Especially preferred are fragments which exist in vivo, e.g., fragments which arise from post transcriptional processing or which arise from translation of alternatively spliced RNA's. Fragments include those expressed in native or endogenous cells as well as those made in expression systems, e.g., in CHO cells. Because peptides such as *S. pneumoniae* polypeptides often exhibit a range of physiological properties and because such properties may be attributable to different portions of the molecule, a useful *S. pneumoniae* fragment or *S. pneumoniae* analog is one which exhibits a biological activity in any biological assay for *S. pneumoniae* activity. Most preferably the fragment or analog possesses 10%, preferably 40%, more preferably 60%, 70%, 80% or 90% or greater of the activity of *S. pneumoniae*, in any in vivo or in vitro assay.

[0071] Analogs can differ from naturally occurring *S. pneumoniae* polypeptides in amino acid sequence or in ways that do not involve sequence, or both. Non-sequence modifications include changes in acetylation, methylation, phosphorylation, carboxylation, or glycosylation. Preferred analogs include *S. pneumoniae* polypeptides (or biologically active fragments thereof) whose sequences differ from the wild-type sequence by one or more conservative amino acid substitutions or by one or more non-conservative amino acid substitutions, deletions, or insertions which do not substantially diminish the biological activity of the *S. pneumoniae* polypeptide. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics, e.g., substitutions within the following groups: valine, glycine; glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Other conservative substitutions can be made in view of the table below.

TABLE 1

CONSERVATIVE AMINO ACID REPLACEMENTS

For Amino Acid	Code	Replace with any of
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, β-Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

[0072] Other analogs within the invention are those with modifications which increase peptide stability; such analogs may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: analogs that include residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids; and cyclic analogs.

[0073] As used herein, the term "fragment", as applied to an *S. pneumoniae* analog, will ordinarily be at least about 20 residues, more typically at least about 40 residues, preferably at least about 60 residues in length. Fragments of *S. pneumoniae* polypeptides can be generated by methods known to those skilled in the art. The ability of a candidate fragment to exhibit a biological activity of *S. pneumoniae* polypeptide can be assessed by methods known to those skilled in the art as described herein. Also included are *S. pneumoniae* polypeptides containing residues that are not required for biological activity of the peptide or that result from alternative mRNA splicing or alternative protein processing events.

[0074] An "immunogenic component" as used herein is a moiety, such as an *S. pneumoniae* polypeptide, analog or fragment thereof, that is capable of eliciting a humoral and/or cellular immune response in a host animal.

[0075] An "antigenic component" as used herein is a moiety, such as an *S. pneumoniae* polypeptide, analog or fragment thereof, that is capable of binding to a specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

[0076] The term "antibody" as used herein is intended to include fragments thereof which are specifically reactive with *S. pneumoniae* polypeptides.

[0077] As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

[0078] Misexpression, as used herein, refers to a non-wild type pattern of gene expression. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-translational modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

[0079] As used herein, "host cells" and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refers to cells which can become or have been used as recipients for a recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood by individuals skilled in the art that the progeny of a single parental cell may not necessarily be completely identical in genomic or total DNA compliment to the original parent, due to accident or deliberate mutation.

[0080] As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site, terminators, and in some cases operators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences.

[0081] As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

[0082] The "metabolism" of a substance, as used herein, means any aspect of the expression, function, action, or regulation of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modifications of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modification, the substance induces in other substances. The metabolism of a substance also includes changes in the distribution of the substance. The metabolism of a substance includes changes the substance induces in the distribution of other substances.

[0083] A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including without limitation plasma, serum, cerebrospinal fluid, lymph, tears, saliva and tissue sections) or from in vitro cell culture constituents, as well as samples from the environment.

[0084] Technical and scientific terms used herein have the meanings commonly understood by one of ordinary skill in the art to which the present invention pertains, unless otherwise defined. Reference is made herein to various methodologies known to those of skill in the art. Publications and other materials setting forth such known methodologies to which reference is made are incorporated herein by reference in their entireties as though set forth in full. The practice of the invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, Fritsch, and Maniatis, *Molecular Cloning; Laboratory Manual* 2nd ed. (1989); *DNA Cloning*, Volumes I and II (D. N Glover ed. 1985); *Oligonucleotide Synthesis* (M. J. Gait ed, 1984); *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); the series, *Methods in Enzymology* (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, eds.); *PCR-A Practical Approach* (McPherson, Quirke, and Taylor, eds., 1991); *Immunology,* 2d Edition, 1989, Roitt et al., C. V. Mosby Company, and New York; *Advanced Immunology,* 2d Edition, 1991, Male et al., Grower Medical Publishing, New York.; *DNA Cloning: A Practical Approach*, Volumes I and II, 1985 (D. N. Glover ed.); *Oligonucleotide Synthesis,* 1984, (M. L. Gait ed); *Transcription and Translation,* 1984 (Hames and Higgins eds.); *Animal Cell Culture,* 1986 (R. I. Freshney ed.); *Immobilized Cells and Enzymes,* 1986 (IRL Press); Perbal, 1984, *A Practical Guide to Molecular Cloning*; and *Gene Transfer Vectors for Mammalian Cells,* 1987 (J. H. Miller and M. P. Calos eds., Cold Spring Harbor Laboratory).

[0085] Any suitable materials and/or methods known to those of skill can be utilized in carrying out the present invention: however preferred materials and/or methods are described. Materials, reagents and the like to which reference is made in the following description and examples are obtainable from commercial sources, unless otherwise noted.

S. pneumoniae Genomic Sequence

[0086] This invention provides nucleotide sequences of the genome of *S. pneumoniae* which thus comprises a DNA sequence library of *S. pneumoniae* genomic DNA. The detailed description that follows provides nucleotide sequences of *S. pneumoniae*, and also describes how the sequences were obtained and how ORFs and protein-coding sequences were identified. Also described are methods of using the disclosed *S. pneumoniae* sequences in methods including diagnostic and therapeutic applications. Furthermore, the library can be used as a database for identification and comparison of medically important sequences in this and other strains of *S. pneumoniae.*

[0087] To determine the genomic sequence of *S. pneumoniae*, DNA was isolated from strain 14453 of *S. pneumoniae* and mechanically sheared by nebulization to a median size of 2 kb. Following size fractionation by gel electrophoresis, the fragments were blunt-ended, ligated to adapter oligonucleotides, and cloned into each of 20 different pMPX vectors (Rice et al., abstracts of Meeting of Genome Mapping and Sequencing, Cold Spring Harbor, N.Y., May 11-May 15, 1994, p. 225) and the PUC19 vector to construct a series of "shotgun" subclone libraries.

[0088] DNA sequencing was achieved using two sequencing methods. The first method used multiplex sequencing procedures essentially as disclosed in Church et al., 1988, *Science* 240: 185; U.S. Pat. Nos. 4,942,124 and 5,149,625). DNA was extracted from pooled cultures and subjected to chemical or enzymatic sequencing. Sequencing reactions were resolved by electrophoresis, and the products were transferred and covalently bound to nylon membranes. Finally, the membranes were sequentially hybridized with a series of labelled oligonucleotides complimentary to "tag" sequences present in the different shotgun cloning vectors. In this manner, a large number of sequences could be obtained from a single set of sequencing reactions. The remainder of the sequencing was performed on ABI377 automated DNA sequencers. The cloning and sequencing procedures are described in more detail in the Exemplification.

[0089] Individual sequence reads were assembled using PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, January 1996, p. 157). The average contig length was about 3-4 kb.

[0090] A variety of approaches are used to order the contigs so as to obtain a continuous sequence representing the entire *S. pneumoniae* genome. Synthetic oligonucleotides are designed that are complementary to sequences at the end of each contig. These oligonucleotides may be hybridized to libaries of *S. pneumoniae* genomic DNA in, for example, lambda phage vectors or plasmid vectors to identify clones that contain sequences corresponding to the junctional regions between individual contigs. Such clones are then used to isolate template DNA and the same oligonucleotides are used as primers in polymerase chain reaction (PCR) to amplify junctional fragments, the nucleotide sequence of which is then determined.

[0091] The *S. pneumoniae* sequences were analyzed for the presence of open reading frames (ORFs) comprising at least 180 nucleotides. As a result of the initial analysis of ORFs based on stop-to-stop codon reads, it should be understood that these ORFs may not correspond to the ORF of a naturally-occurring *S. pneumoniae* polypeptide. These ORFs may contain start codons which indicate the initiation of protein synthesis of a naturally-occurring *S. pneumoniae* polypeptide. Such start codons within the ORFs provided herein can be identified by those of ordinary skill in the relevant art, and the resulting ORF and the encoded *S. pneumoniae* polypeptide is within the scope of this invention. For example, within the ORFs a codon such as AUG or GUG (encoding methionine or valine) which is part of the initiation signal for protein synthesis can be identified and the portion of an ORF to corresponding to a naturally-occurring *S. pneumoniae* polypeptide can be recognized.

[0092] The second analysis of the ORFs included identifying the start codons and the predicted coding regions. These ORFs provided in this invention were defined by one or more of the following methods: evaluating the coding potential of such sequences with the program GENEMARK™ (Borodovsky and McIninch, 1993, *Comp.* 17: 123), distinguishing the coding from noncoding regions using the program Glimmer (Fraser et al, *Nature,* 1997), determining codon usage (Staden et al., *Nucleic Acid Research* 10: 141), and each predicted ORF amino acid sequence was compared with all protein sequences found in current GENBANK, SWISS-PROT, and PIR databases using the BLAST algorithm. BLAST identifies local alignments occurring by chance between the ORF sequence and the sequence in the databank (Altschal et al., 1990, L Mol. Biol. 215: 403-410). Homologous ORFs (probabilities less than 10^{-5} by chance) and ORF's that are probably non-homologous (probabilities greater than 10^{-5} by chance) but have good codon usage were identified. Both homologous, sequences and non-homologous sequences with good codon usage are likely to encode proteins and are encompassed by the invention.

S. pneumoniae Nucleic Acids

[0093] The nucleic acids of this invention may be obtained directly from the DNA of the above referenced *S. pneumoniae* strain by using the polymerase chain reaction (PCR). See "*PCR, A Practical Approach*" (McPherson, Quirke, and Taylor, eds., IRL Press, Oxford, UK, 1991) for details about the PCR. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, the authenticity of amplified products can be verified by conventional sequencing methods. Clones carrying the desired sequences described in this invention may also be obtained by screening the libraries by means of the PCR or by hybridization of synthetic oligonucleotide probes to filter lifts of the library colonies or plaques as known in the art (see, e.g., Sambrook et al., *Molecular Cloning, A Laboratory Manual* 2nd edition, 1989, Cold Spring Harbor Press, NY).

[0094] It is also possible to obtain nucleic acids encoding *S. pneumoniae* polypeptides from a cDNA library in accordance with protocols herein described. A cDNA encoding an *S. pneumoniae* polypeptide can be obtained by isolating total mRNA from an appropriate strain. Double stranded cDNAs can then be prepared from the total mRNA. Subsequently, the cDNAs can be inserted into a suitable plasmid or viral (e.g., bacteriophage) vector using any one of a number of known techniques. Genes encoding *S. pneumoniae* polypeptides can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. The nucleic acids of the invention can be DNA or RNA. Preferred nucleic acids of the invention are contained in the Sequence Listing.

[0095] The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura et al. U.S. Pat. No. 4,598,049; Caruthers et al. U.S. Pat. No. 4,458,066; and Itakura U.S. Pat. Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

[0096] Nucleic acids isolated or synthesized in accordance with features of the present invention are useful, by way of example, without limitation, as probes, primers, capture ligands, antisense genes and for developing expression systems for the synthesis of proteins and peptides corresponding to such sequences. As probes, primers, capture ligands and antisense agents, the nucleic acid normally consists of all or part (approximately twenty or more nucleotides for specificity as well as the ability to form stable hybridization products) of the nucleic acids of the invention contained in the Sequence Listing. These uses are described in further detail below.

[0097] Probes

[0098] A nucleic acid isolated or synthesized in accordance with the sequence of the invention contained in the Sequence Listing can be used as a probe to specifically detect *S. pneumoniae*. With the sequence information set forth in the present application, sequences of twenty or more nucleotides are identified which provide the desired inclusivity and exclusivity with respect to *S. pneumoniae*, and extraneous nucleic acids likely to be encountered during hybridization conditions. More preferably, the sequence will comprise at least twenty to thirty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules.

[0099] Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques. Individuals skilled in the art will readily recognize that the nucleic acids, for use as probes, can be provided with a label to facilitate detection of a hybridization product.

[0100] Nucleic acid isolated and synthesized in accordance with the sequence of the invention contained in the Sequence Listing can also be useful as probes to detect homologous regions (especially homologous genes) of other *Streptococcus* species using appropriate stringency hybridization conditions as described herein.

[0101] Capture Ligand

[0102] For use as a capture ligand, the nucleic acid selected in the manner described above with respect to probes, can be readily associated with a support. The manner in which nucleic acid is associated with supports is well known. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing have utility to separate *S. pneumoniae* nucleic acid from the nucleic acid of each other and other organisms. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing can also have utility to separate other *Streptococcus* species from each other and from other organisms. Preferably, the sequence will comprise at least twenty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules. Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques.

[0103] Primers

[0104] Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility as primers for the amplification of *S. pneumoniae* nucleic acid. These nucleic acids may also have utility as primers for the amplification of nucleic acids in other *Streptococcus* species. With respect to polymerase chain reaction (PCR) techniques, nucleic acid sequences of $\geqq$10-15 nucleotides of the invention contained in the Sequence Listing have utility in conjunction with suitable enzymes and reagents to create copies of *S. pneumoniae* nucleic acid. More preferably, the sequence will comprise twenty or more nucleotides to convey stability to the hybridization product formed between the primer and the intended target molecules. Binding conditions of primers greater than 100 nucleotides are more difficult to control to obtain specificity. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, amplified products can be checked by conventional sequencing methods.

[0105] The copies can be used in diagnostic assays to detect specific sequences, including genes from *S. pneumoniae* and/or other *Streptococcus* species. The copies can also be incorporated into cloning and expression vectors to generate polypeptides corresponding to the nucleic acid synthesized by PCR, as is described in greater detail herein.

[0106] Antisense

[0107] Nucleic acid or nucleic acid-hybridizing derivatives isolated or synthesized in accordance with the sequences described herein have utility as antisense agents to prevent the expression of *S. pneumoniae* genes. These sequences also have utility as antisense agents to prevent expression of genes of other *Streptococcus* species.

[0108] In one embodiment, nucleic acid or derivatives corresponding to *S. pneumoniae* nucleic acids is loaded into a suitable carrier such as a liposome or bacteriophage for introduction into bacterial cells. For example, a nucleic acid having twenty or more nucleotides is capable of binding to bacteria nucleic acid or bacteria messenger RNA. Preferably, the antisense nucleic acid is comprised of 20 or more nucleotides to provide necessary stability of a hybridization product of non-naturally occurring nucleic acid and bacterial nucleic acid and/or bacterial messenger RNA. Nucleic acid having a sequence greater than 1000 nucleotides in length is difficult to synthesize but can be generated by recombinant DNA techniques. Methods for loading antisense nucleic acid in liposomes is known in the art as exemplified by U.S. Pat. No. 4,241,046 issued Dec. 23, 1980 to Papahadjopoulos et al.

[0109] The present invention encompasses isolated polypeptides and nucleic acids derived from *S. pneumoniae* that are useful as reagents for diagnosis of bacterial infection, components of effective antibacterial vaccines, and/or as targets for antibacterial drugs, including anti-*S. pneumoniae* drugs.

Expression of *S. pneumoniae* Nucleic Acids

[0110] Table 2 provides a list of open reading frames (ORFs) in both strands. An ORF is a region of nucleic acid which encodes a polypeptide. This region normally represents a complete coding sequence or a total sequence and was determined from an initial analysis of stop to stop codons followed by the prediction of start codons. The first column lists the ORF designation. The second and third columns list the SEQ ID numbers for the nucleic acid and amino acid sequences corresponding to each ORF, respectively. The fourth and fifth columns list the length of the nucleic acid ORF and the length of the amino acid ORF, respectively. Most of the nucleotide sequences corresponding to each ORF begin at the first nucleotide of the start codon and end at the nucleotide immediately preceding the next downstream stop codon in the same reading frame. It will be recognized by one skilled in the art that the natural translation initiation sites will correspond to ATG, GTG, or TTG codons located within the ORFs. The natural initiation sites depend not only on the sequence of a start codon but also on the context of the DNA sequence adjacent to the start codon. Usually, a recognizable ribosome binding site is found within 20 nucleotides upstream from the initiation codon. In some cases where genes are translationally coupled and coordinately expressed together in "operons", ribosome binding sites are not present, but the initiation codon of a downstream gene may occur very close to, or overlap, the stop codon of the an upstream gene in the same operon. The correct start codons can be generally identified rapidly and efficiently because only a few codons need be tested. It is recognized that the translational machinery in bacteria initiates most polypeptide chains with the amino acid methionine. In some cases, polypeptides are post-translationally modified, resulting in an N-terminal amino acid other than methionine in vivo. The sixth and seventh columns provide metrics for assessing the likelihood of the homology match (determined by the BLASTP2 algorithm), as is known in the art, to the genes indicated in the description field. Specifically, the sixth column represents the "Score" for the match (a higher score is a better match), and the seventh column represents the "P-value" for the match (the probability that such a match could have occurred by chance; the lower the value, the more likely the match is valid). If a BLASTP2 score of less than 46 was obtained, no value is reported in the table the "P-value". The description field provides, where available, the accession number (AC) or the Swissprot accession number (SP), the locus name (LN), Superfamily Classification (CL), the Organism (OR), Source of variant (SR), E.C. number (EC), the gene name (GN), the product name (PN), the Function Description (FN), the Map Position (MP), Left End (LE), Right End (RE), Coding Direction (DI), the Database from which the sequence originates (DB), and the description (DE) or notes (NT) for each ORF. This information allows one of ordinary skill in the art to determine a potential use and function for each identified coding sequence and, as a result, allows the use of the polypeptides of the present invention for commercial and industrial purposes.

[0111] Using the information provided in SEQ ID NO: 1-SEQ ID NO: 2661 and in Table 2 together with routine cloning and sequencing methods, one of ordinary skill in the art will be able to clone and sequence all the nucleic acid fragments of interest including open reading frames (ORFs) encoding a large variety proteins of *S. pneumoniae.*

[0112] Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility to generate polypeptides. The nucleic acid of the invention exemplified in SEQ ID NO: 1-SEQ ID NO: 2661 and in Table 2 or fragments of said nucleic acid encoding active portions of *S. pneumoniae* polypeptides can be cloned into suitable vectors or used to isolate nucleic acid. The isolated nucleic acid is combined with suitable DNA linkers and cloned into a suitable vector.

[0113] The function of a specific gene or operon can be ascertained by expression in a bacterial strain under conditions where the activity of the gene product(s) specified by the gene or operon in question can be specifically measured. Alternatively, a gene product may be produced in large quantities in an expressing strain for use as an antigen, an industrial reagent, for structural studies, etc. This expression can be accomplished in a mutant strain which lacks the activity of the gene to be tested, or in a strain that does not produce the same gene product(s). This includes, but is not limited to, Eucaryotic species such as the yeast *Saccharomyces cerevisiae, Methanobacterium* strains or other Archaea, and Eubacteria such as *E. coli, B. subtilis, S. aureus, S. pneumonia* or *Pseudomonas putida*. In some cases the expression host will utilize the natural *S. pneumoniae* promoter whereas in others, it will be necessary to drive the gene with a promoter sequence derived from the expressing organism (e.g., an *E. coli* beta-galactosidase promoter for expression in *E. coli*).

[0114] To express a gene product using the natural *S. pneumoniae* promoter, a procedure such as the following can be used. A restriction fragment containing the gene of interest, together with its associated natural promoter element and regulatory sequences (identified using the DNA sequence data) is cloned into an appropriate recombinant plasmid containing an origin of replication that functions in the host organism and an appropriate selectable marker. This can be accomplished by a number of procedures known to those skilled in the art. It is most preferably done by cutting the plasmid and the fragment to be cloned with the same restriction enzyme to produce compatible ends that can be ligated to join the two pieces together. The recombinant plasmid is introduced into the host organism by, for example, electroporation and cells containing the recombinant plasmid are identified by selection for the marker on the plasmid. Expression of the desired gene product is detected using an assay specific for that gene product.

[0115] In the case of a gene that requires a different promoter, the body of the gene (coding sequence) is specifically excised and cloned into an appropriate expression plasmid. This subcloning can be done by several methods, but is most easily accomplished by PCR amplification of a specific fragment and ligation into an expression plasmid after treating the PCR product with a restriction enzyme or exonuclease to create suitable ends for cloning.

[0116] A suitable host cell for expression of a gene can be any procaryotic or eucaryotic cell. For example, an *S.*

pneumoniae polypeptide can be expressed in bacterial cells such as *E. coli* or *B. subtilis*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cell (CHO). Other suitable host cells are known to those skilled in the art.

[0117] Expression in eucaryotic cells such as mammalian, yeast, or insect cells can lead to partial or complete glycosylation and/or formation of relevant inter- or intra-chain disulfide bonds of a recombinant peptide product. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari. et al., (1987) *Embo J.* 6: 229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30: 933-943), pJRY88 (Schultz et al., (1987) *Gene* 54: 113-123), and pYES2 (Invitrogen Corporation, San Diego, Calif.). Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9 cells) include the pAc series (Smith et al., (1983) *Mol. Cell Biol.* 3: 2156-2165) and the pVL series (Lucklow, V. A., and Summers, M. D., (1989) *Virology* 170: 31-39). Generally, COS cells (Gluzman, Y., (1981) *Cell* 23: 175-182) are used in conjunction with such vectors as pCDM 8 (Aruffo, A. and Seed, B., (1987) *Proc. Natl. Acad. Sci. USA* 84: 8573-8577) for transient amplification/expression in mammalian cells, while CHO ($dhfr^-$ Chinese Hamster Ovary) cells are used with vectors such as pMT2PC (Kaufman et al. (1987), *EMBO J.* 6: 187-195) for stable amplification/expression in mammalian cells. Vector DNA can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming host cells can be found in Sambrook et al. (*Molecular Cloning: A Laboratory Manual,* 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

[0118] Expression in procaryotes is most often carried out in *E. coli* with either fusion or non-fusion inducible expression vectors. Fusion vectors usually add a number of NH_2 terminal amino acids to the expressed target gene. These NH_2 terminal amino acids often are referred to as a reporter group or an affinity purification group. Such reporter groups usually serve two purposes: 1) to increase the solubility of the target recombinant protein; and 2) to aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the reporter group and the target recombinant protein to enable separation of the target recombinant protein from the reporter group subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) which fuse glutathione S-transferase, maltose E binding protein, or protein A, respectively, to the target recombinant protein. A preferred reporter group is poly(His), which may be fused to the amino or carboxy terminus of the protein and which renders the recombinant fusion protein easily purifiable by metal chelate chromatography.

[0119] Inducible non-fusion expression vectors include pTrc (Amann et al., (1988) Gene 69: 301-315) and pET11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) 60-89). While target gene expression relies on host RNA polymerase transcription from the hybrid trp-lac fusion promoter in pTrc, expression of target genes inserted into pET11d relies on transcription from the T7 gn10-lac 0 fusion promoter mediated by coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 under the transcriptional control of the lacUV 5 promoter.

[0120] For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding an *S. pneumoniae* polypeptide can be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the peptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. Polypeptides of the invention can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such polypeptides. Additionally, in many situations, polypeptides can be produced by chemical cleavage of a native protein (e.g., tryptic digestion) and the cleavage products can then be purified by standard techniques.

[0121] In the case of membrane bound proteins, these can be isolated from a host cell by contacting a membrane-associated protein fraction with a detergent forming a solubilized complex, where the membrane-associated protein is no longer entirely embedded in the membrane fraction and is solubilized at least to an extent which allows it to be chromatographically isolated from the membrane fraction. Several different criteria are used for choosing a detergent suitable for solubilizing these complexes. For example, one property considered is the ability of the detergent to solubilize the *S. pneumoniae* protein within the membrane fraction at minimal denaturation of the membrane-associated protein allowing for the activity or functionality of the membrane-associated protein to return upon reconstitution of the protein. Another property considered when selecting the detergent is the critical micelle concentration (CMC) of the detergent in that the detergent of choice preferably has a high CMC value allowing for ease of removal after reconstitution. A third property considered when selecting a detergent is the hydrophobicity of the detergent. Typically, membrane-associated proteins are very hydrophobic and therefore detergents which are also hydrophobic, e.g., the triton series, would be useful for solubilizing the hydrophobic proteins. Another property important to a detergent can be the capability of the detergent to remove the *S. pneumoniae* protein with minimal protein-protein interaction facilitating further purification. A fifth property of the detergent which should be considered is the charge of the detergent. For example, if it is desired to use ion exchange resins in the purification process then preferably detergent should be an uncharged detergent. Chromatographic techniques which can be used in the final purification step are known in the art and include hydrophobic interaction, lectin affinity, ion exchange, dye affinity and immunoaffinity.

[0122] One strategy to maximize recombinant *S. pneumoniae* peptide expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) 119-128). Another strategy would be to alter the nucleic acid encoding an *S. pneumoniae* peptide to be inserted into an expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins (Wada et al., (1992) *Nuc. Acids Res.* 20: 2111-2118). Such alteration of nucleic acids of the invention can be carried out by standard DNA synthesis techniques.

[0123] The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See, e.g., Itakura et al. U.S. Pat. No. 4,598,049; Caruthers et al. U.S. Pat. No. 4,458,066; and Itakura U.S. Pat. Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

[0124] The present invention provides a library of *S. pneumoniae-derived* nucleic acid sequences. The libraries provide probes, primers, and markers which can be used as markers in epidemiological studies. The present invention also provides a library of *S. pneumoniae-derived* nucleic acid sequences which comprise or encode targets for therapeutic drugs.

[0125] Nucleic acids comprising any of the sequences disclosed herein or sub-sequences thereof can be prepared by standard methods using the nucleic acid sequence information provided in SEQ ID NO: 1-SEQ ID NO: 2661. For example, DNA can be chemically synthesized using, e.g., the phosphoramidite solid support method of Matteucci et al., 1981*, J. Am. Chem. Soc.* 103: 3185, the method of Yoo et al., 1989*, J. Biol. Chem.* 764: 17078, or other well known methods. This can be done by sequentially linking a series of oligonucleotide cassettes comprising pairs of synthetic oligonucleotides, as described below.

[0126] Of course, due to the degeneracy of the genetic code, many different nucleotide sequences can encode polypeptides having the amino acid sequences defined by SEQ ID NO: 2662-SEQ ID NO: 5322 or sub-sequences thereof. The codons can be selected for optimal expression in prokaryotic or eukaryotic systems. Such degenerate variants are also encompassed by this invention.

[0127] Insertion of nucleic acids (typically DNAs) encoding the polypeptides of the invention into a vector is easily accomplished when the termini of both the DNAs and the vector comprise compatible restriction sites. If this cannot be done, it may be necessary to modify the termini of the DNAs and/or vector by digesting back single-stranded DNA overhangs generated by restriction endonuclease cleavage to produce blunt ends, or to achieve the same result by filling in the single-stranded termini with an appropriate DNA polymerase.

[0128] Alternatively, any site desired may be produced, e.g., by ligating nucleotide sequences (linkers) onto the termini. Such linkers may comprise specific oligonucleotide sequences that define desired restriction sites. Restriction sites can also be generated by the use of the polymerase chain reaction (PCR). See, e.g., Saiki et al., 1988*, Science* 239: 48. The cleaved vector and the DNA fragments may also be modified if required by homopolymeric tailing.

[0129] In certain embodiments, the invention encompasses isolated nucleic acid fragments comprising all or part of the individual nucleic acid sequences disclosed herein. The fragments are at least about 8 nucleotides in length, preferably at least about 12 nucleotides in length, and most preferably at least about 15-20 nucleotides in length.

[0130] The nucleic acids may be isolated directly from cells. Alternatively, the polymerase chain reaction (PCR) method can be used to produce the nucleic acids of the invention, using either chemically synthesized strands or genomic material as templates. Primers used for PCR can be synthesized using the sequence information provided herein and can further be designed to introduce appropriate new restriction sites, if desirable, to facilitate incorporation into a given vector for recombinant expression.

[0131] The nucleic acids of the present invention may be flanked by natural *S. pneumoniae* regulatory sequences, or may be associated with heterologous sequences, including promoters, enhancers, response elements, signal sequences, polyadenylation sequences, introns, 5'- and 3'-noncoding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, “caps”, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.). Nucleic acids may contain one or more additional covalently linked moieties, such as, for example, proteins (e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (e.g., acridine, psoralen, etc.), chelators (e.g., metals, radioactive metals, iron, oxidative metals, etc.), and alkylators. PNAs are also included. The nucleic acid may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the nucleic acid sequences of the present invention may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

[0132] The invention also provides nucleic acid vectors comprising the disclosed *S. pneumoniae*-derived sequences or derivatives or fragments thereof. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts, and may be used for gene therapy as well as for simple cloning or protein expression.

[0133] The encoded *S. pneumoniae* polypeptides may be expressed by using many known vectors, such as pUC plasmids, pET plasmids (Novagen, Inc., Madison, Wis.), or pRSET or pREP (Invitrogen, San Diego, Calif.), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. The particular choice of vector/host is not critical to the practice of the invention.

[0134] Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, e.g. antibiotic resistance, and one or more expression cassettes. The inserted *S. pneumoniae* coding sequences may be synthesized by standard methods, isolated from natural sources, or prepared as hybrids, etc. Ligation of the *S. pneumoniae* coding sequences to transcriptional regulatory elements and/or to other amino acid coding sequences may be achieved by known methods. Suitable host cells may be transformed/transfected/infected as appropriate by any suitable method including electroporation, $CaCl_2$ mediated DNA uptake, fungal infection, microinjection, microprojectile, or other established methods.

[0135] Appropriate host cells include bacteria, archebacteria, fungi, especially yeast, and plant and animal cells, especially mammalian cells. Of particular interest are *S. pneumoniae, E. coli, B. Subtilis, Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Schizosaccharomyces pombi*, SF9 cells, C129 cells, 293 cells, *Neurospora*, and CHO cells, COS cells, HeLa cells, and immortalized mammalian myeloid and lymphoid cell lines. Preferred replication systems include M13, ColE1, SV40, baculovirus, lambda, adenovirus, and the like. A large number of transcription initiation and termination regulatory regions have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. Examples of these regions, methods of isolation, manner of manipulation, etc. are known in the art. Under appropriate expression conditions, host cells can be used as a source of recombinantly produced *S. pneumoniae*-derived peptides and polypeptides.

[0136] Advantageously, vectors may also include a transcription regulatory element (i.e., a promoter) operably linked to the *S. pneumoniae* portion. The promoter may optionally contain operator portions and/or ribosome binding sites. Non-limiting examples of bacterial promoters compatible with *E. coli* include: b-lactamase (penicillinase) promoter; lactose promoter; tryptophan (trp) promoter; araBAD (arabinose) operon promoter; lambda-derived P_1 promoter and N gene ribosome binding site; and the hybrid tac promoter derived from sequences of the trp and lac Uv5 promoters. Non-limiting examples of yeast promoters include 3-phosphoglycerate kinase promoter, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter, galactokinase (GAL1) promoter, galactoepimerase promoter, and alcohol dehydrogenase (ADH) promoter. Suitable promoters for mammalian cells include without limitation viral promoters such as that from Simian Virus 40 (SV40), Rous sarcoma virus (RSV), adenovirus (ADV), and bovine papilloma virus (BPV). Mammalian cells may also require terminator sequences, polyA addition sequences and enhancer sequences to increase expression. Sequences which cause amplification of the gene may also be desirable. Furthermore, sequences that facilitate secretion of the recombinant product from cells, including, but not limited to, bacteria, yeast, and animal cells, such as secretory signal sequences and/or prohormone pro region sequences, may also be included. These sequences are well described in the art.

[0137] Nucleic acids encoding wild-type or variant *S. pneumoniae*-derived polypeptides may also be introduced into cells by recombination events. For example, such a sequence can be introduced into a cell, and thereby effect homologous recombination at the site of an endogenous gene or a sequence with substantial identity to the gene. Other recombination-based methods such as nonhomologous recombinations or deletion of endogenous genes by homologous recombination may also be used.

[0138] The nucleic acids of the present invention find use as templates for the recombinant production of *S. pneumoniae*-derived peptides or polypeptides.

Identification and Use of *S. pneumoniae* Nucleic Acid Sequences

[0139] The disclosed *S. pneumoniae* polypeptide and nucleic acid sequences, or other sequences that are contained within ORFs, including complete protein-coding sequences, of which any of the disclosed *S. pneumoniae*-specific sequences forms a part, are useful as target components for diagnosis and/or treatment of *S. pneumoniae*-caused infection.

[0140] It will be understood that the sequence of an entire protein-coding sequence of which each disclosed nucleic acid sequence forms a part can be isolated and identified based on each disclosed sequence. This can be achieved, for example, by using an isolated nucleic acid encoding the disclosed sequence, or fragments thereof, to prime a sequencing reaction with genomic *S. pneumoniae* DNA as template; this is followed by sequencing the amplified product. The isolated nucleic acid encoding the disclosed sequence, or fragments thereof, can also be hybridized to *S. pneumoniae* genomic libraries to identify clones containing additional complete segments of the protein-coding sequence of which the shorter sequence forms a part. Then, the entire protein-coding sequence, or fragments thereof, or nucleic acids encoding all or part of the sequence, or sequence-conservative or function-conservative variants thereof, may be employed in practicing the present invention.

[0141] Preferred sequences are those that are useful in diagnostic and/or therapeutic applications. Diagnostic applications include without limitation nucleic-acid-based and antibody-based methods for detecting bacterial infection. Therapeutic applications include without limitation vaccines, passive immunotherapy, and drug treatments directed against gene products that are both unique to bacteria and essential for growth and/or replication of bacteria.

Identification of Nucleic Acids Encoding Vaccine Components and Targets for Agents Effective Against *S. pneumoniae*

[0142] The disclosed *S. pneumoniae* genome sequence includes segments that direct the synthesis of ribonucleic acids and polypeptides, as well as origins of replication, promoters, other types of regulatory sequences, and intergenic nucleic acids. The invention encompasses nucleic acids encoding immunogenic components of vaccines and targets for agents effective against *S. pneumoniae*. Identification of said immunogenic components involved in the determination of the function of the disclosed sequences, which can be achieved using a variety of approaches. Non-limiting examples of these approaches are described briefly below.

[0143] Homology to Known Sequences:

[0144] Computer-assisted comparison of the disclosed *S. pneumoniae* sequences with previously reported sequences present in publicly available databases is useful for identifying functional *S. pneumoniae* nucleic acid and polypeptide sequences. It will be understood that protein-coding sequences, for example, may be compared as a whole, and that a high degree of sequence homology between two proteins (such as, for example, >80-90%) at the amino acid level indicates that the two proteins also possess some degree of functional homology, such as, for example, among enzymes involved in metabolism, DNA synthesis, or cell wall synthesis, and proteins involved in transport, cell division, etc. In addition, many structural features of particular protein classes have been identified and correlate with specific consensus sequences, such as, for example, binding domains for nucleotides, DNA, metal ions, and other small molecules; sites for covalent modifications such as phosphorylation, acylation, and the like; sites of protein:protein interactions, etc. These consensus sequences may be quite short and thus may represent only a fraction of the entire protein-coding sequence. Identification of such a feature in an *S. pneumoniae* sequence is therefore useful in determining the function of the encoded protein and identifying useful targets of antibacterial drugs.

[0145] Of particular relevance to the present invention are structural features that are common to secretory, transmembrane, and surface proteins, including secretion signal peptides and hydrophobic transmembrane domains. *S. pneumoniae* proteins identified as containing putative signal sequences and/or transmembrane domains are useful as immunogenic components of vaccines.

[0146] Targets for therapeutic drugs according to the invention include, but are not limited to, polypeptides of the invention, whether unique to *S. pneumoniae* or not, that are essential for growth and/or viability of *S. pneumoniae* under at least one growth condition. Polypeptides essential for growth and/or viability can be determined by examining the effect of deleting and/or disrupting the genes, i.e., by so-called gene "knockout". Alternatively, genetic footprinting can be used (Smith et al., 1995*, Proc. Natl. Acad. Sci. USA* 92: 5479-6433; Published International Application WO 94/26933; U.S. Pat. No. 5,612,180). Still other methods for assessing essentiality includes the ability to isolate conditional lethal mutations in the specific gene (e.g., temperature sensitive mutations). Other useful targets for therapeutic drugs, which include polypeptides that are not essential for growth or viability per se but lead to loss of viability of the cell, can be used to target therapeutic agents to cells.

[0147] Strain-Specific Sequences:

[0148] Because of the evolutionary relationship between different *S. pneumoniae* strains, it is believed that the presently disclosed *S. pneumoniae* sequences are useful for identifying, and/or discriminating between, previously known and new *S. pneumoniae* strains. It is believed that other *S. pneumoniae* strains will exhibit at least 70% sequence homology with the presently disclosed sequence. Systematic and routine analyses of DNA sequences derived from samples containing *S. pneumoniae* strains, and comparison with the present sequence allows for the identification of sequences that can be used to discriminate between strains, as well as those that are common to all *S. pneumoniae* strains. In one embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that discriminate between different strains of *S. pneumoniae*. Strain-specific components can also be identified functionally by their ability to elicit or react with antibodies that selectively recognize one or more *S. pneumoniae* strains.

[0149] In another embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that are common to all *S. pneumoniae* strains but are not found in other bacterial species.

S. pneumoniae Polypeptides

[0150] This invention encompasses isolated *S. pneumoniae* polypeptides encoded by the disclosed *S. pneumoniae* genomic sequences, including the polypeptides of the invention contained in the Sequence Listing. Polypeptides of the invention are preferably at least 5 amino acid residues in length. Using the DNA sequence information provided herein, the amino acid sequences of the polypeptides encompassed by the invention can be deduced using methods well-known in the art. It will be understood that the sequence of an entire nucleic acid encoding an *S. pneumoniae* polypeptide can be isolated and identified based on an ORF that encodes only a fragment of the cognate protein-coding region. This can be achieved, for example, by using the isolated nucleic acid encoding the ORF, or fragments thereof, to prime a polymerase chain reaction with genomic *S. pneumoniae* DNA as template; this is followed by sequencing the amplified product.

[0151] The polypeptides of the present invention, including function-conservative variants of the disclosed ORFs, may be isolated from wild-type or mutant *S. pneumoniae* cells, or from heterologous organisms or cells (including, but not limited to, bacteria, fungi, insect, plant, and mammalian cells) including *S. pneumoniae* into which a *S. pneumoniae*-derived protein-coding sequence has been introduced and expressed. Furthermore, the polypeptides may be part of recombinant fusion proteins.

[0152] *S. pneumoniae* polypeptides of the invention can be chemically synthesized using commercially automated procedures such as those referenced herein, including, without limitation, exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis. The polypeptides are preferably prepared by solid phase peptide synthesis as described by Merrifield, 1963*, J. Am. Chem. Soc.* 85: 2149. The synthesis is carried out with amino acids that are protected at the alpha-amino terminus. Trifunctional amino acids with labile side-chains are also protected with suitable groups to prevent undesired chemical reactions from occurring during the assembly of the polypeptides. The alpha-amino protecting group is selectively removed to allow subsequent reaction to take place at the amino-terminus. The conditions for the removal of the alpha-amino protecting group do not remove the side-chain protecting groups.

[0153] The alpha-amino protecting groups are those known to be useful in the art of stepwise polypeptide synthesis. Included are acyl type protecting groups, e.g., formyl, trifluoroacetyl, acetyl, aromatic urethane type protecting groups, e.g., benzyloxycarbonyl (Cbz), substituted benzyloxycarbonyl and 9-fluorenylmethyloxycarbonyl (Fmoc), aliphatic urethane protecting groups, e.g., t-butyloxycarbonyl (Boc), isopropyloxycarbonyl, cyclohexyloxycarbonyl, and alkyl type protecting groups, e.g., benzyl, triphenylmethyl. The preferred protecting group is Boc. The side-chain protecting groups for Tyr include tetrahydropyranyl, tert-butyl, trityl, benzyl, Cbz, 4-Br-Cbz and 2,6-dichlorobenzyl. The preferred side-chain protecting group for Tyr is 2,6-dichlorobenzyl. The side-chain protecting groups for Asp include benzyl, 2,6-dichlorobenzyl, methyl, ethyl and cyclohexyl. The preferred side-chain protecting group for Asp is cyclohexyl. The side-chain protecting groups for Thr and Ser include acetyl, benzoyl, trityl, tetrahydropyranyl, benzyl, 2,6-dichlorobenzyl and Cbz. The preferred protecting group for Thr and Ser is benzyl. The side-chain protecting groups for Arg include nitro, Tos, Cbz, adamantyloxycarbonyl and Boc. The preferred protecting group for Arg is Tos. The side-chain amino group of Lys may be protected with Cbz, 2-Cl-Cbz, Tos or Boc. The 2-Cl-Cbz group is the preferred protecting group for Lys.

[0154] The side-chain protecting groups selected must remain intact during coupling and not be removed during the deprotection of the amino-terminus protecting group or during coupling conditions. The side-chain protecting groups must also be removable upon the completion of synthesis, using reaction conditions that will not alter the finished polypeptide.

[0155] Solid phase synthesis is usually carried out from the carboxy-terminus by coupling the alpha-amino protected (side-chain protected) amino acid to a suitable solid support. An ester linkage is formed when the attachment is made to a chloromethyl or hydroxymethyl resin, and the resulting polypeptide will have a free carboxyl group at the C-terminus. Alternatively, when a benzhydrylamine or p-methylbenzhydrylamine resin is used, an amide bond is formed and the resulting polypeptide will have a carboxamide group at the C-terminus. These resins are commercially available, and their preparation was described by Stewart et al., 1984, *Solid Phase Peptide Synthesis* (2nd Edition), Pierce Chemical Co., Rockford, Ill.

[0156] The C-terminal amino acid, protected at the side chain if necessary and at the alpha-amino group, is coupled to the benzhydrylamine resin using various activating agents including dicyclohexylcarbodiimide (DCC), N,N'-diisopropyl-carbodiimide and carbonyldiimidazole. Following the attachment to the resin support, the alpha-amino protecting group is removed using trifluoroacetic acid (TFA) or HCl in dioxane at a temperature between 0 and 25° C. Dimethylsulfide is added to the TFA after the introduction of methionine (Met) to suppress possible S-alkylation. After removal of the alpha-amino protecting group, the remaining protected amino acids are coupled stepwise in the required order to obtain the desired sequence.

[0157] Various activating agents can be used for the coupling reactions including DCC, N,N'-diisopropyl-carbodiimide, benzotriazol-1-yl-oxy-tris-(dimethylamino)-phosphonium hexa-fluorophosphate (BOP) and DCC-hydroxybenzotriazole (HOBt). Each protected amino acid is used in excess (>2.0 equivalents), and the couplings are usually carried out in N-methylpyrrolidone (NMP) or in DMF, CH_2Cl_2 or mixtures thereof. The extent of completion of the coupling reaction is monitored at each stage, e.g., by the ninhydrin reaction as described by Kaiser et al., 1970, *Anal. Biochem.* 34: 595. In cases where incomplete coupling is found, the coupling reaction is repeated. The coupling reactions can be performed automatically with commercially available instruments.

[0158] After the entire assembly of the desired polypeptide, the polypeptide-resin is cleaved with a reagent such as liquid HF for 1-2 hours at 0° C., which cleaves the polypeptide from the resin and removes all side-chain protecting groups. A scavenger such as anisole is usually used with the liquid HF to prevent cations formed during the cleavage from alkylating the amino acid residues present in the polypeptide. The polypeptide-resin may be deprotected with TFA/dithioethane prior to cleavage if desired.

[0159] Side-chain to side-chain cyclization on the solid support requires the use of an orthogonal protection scheme which enables selective cleavage of the side-chain functions of acidic amino acids (e.g., Asp) and the basic amino acids (e.g., Lys). The 9-fluorenylmethyl (Fm) protecting group for the side-chain of Asp and the 9-fluorenylmethyloxycarbonyl (Fmoc) protecting group for the side-chain of Lys can be used for this purpose. In these cases, the side-chain protecting groups of the Boc-protected polypeptide-resin are selectively removed with piperidine in DMF. Cyclization is achieved on the solid support using various activating agents including DCC, DCC/HOBt or BOP. The HF reaction is carried out on the cyclized polypeptide-resin as described above.

[0160] Methods for polypeptide purification are well-known in the art, including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide in a recombinant system in which the *S. pneumoniae* protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence. The polypeptide can then be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against a *S. pneumoniae* protein or against peptides derived therefrom can be used as purification reagents. Other purification methods are possible.

[0161] The present invention also encompasses derivatives and homologues of *S. pneumoniae*-encoded polypeptides. For some purposes, nucleic acid sequences encoding the peptides may be altered by substitutions, additions, or deletions that provide for functionally equivalent molecules, i.e., function-conservative variants. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of similar properties, such as, for example, positively charged amino acids (arginine, lysine, and histidine); negatively charged amino acids (aspartate and glutamate); polar neutral amino acids; and non-polar amino acids. The isolated polypeptides may be modified by, for example, phosphorylation, sulfation, acylation, or other protein modifications. They may also be modified with a label capable of providing a detectable signal, either directly or indirectly, including, but not limited to, radioisotopes and fluorescent compounds.

Agents.

[0162] To identify *S. pneumoniae*-derived polypeptides for use in the present invention, essentially the complete genomic sequence of a virulent, methicillin-resistant isolate of *Streptococcus pneumoniae* isolate was analyzed. While, in very rare instances, a nucleic acid sequencing error may be revealed, resolving a rare sequencing error is well within the art, and such an occurrence will not prevent one skilled in the art from practicing the invention.

[0163] Also encompassed are any *S. pneumoniae* polypeptide sequences that are contained within the open reading frames (ORFs), including complete protein-coding sequences, of which any of SEQ ID NO: 2662-SEQ ID NO: 5322 forms a part. Table 2, which is appended herewith and which forms part of the present specification, provides a putative identification of the particular function of a polypeptide which is encoded by each ORF. As a result, one skilled in the art can use the polypeptides of the present invention for commercial and industrial purposes consistent with the type of putative identification of the polypeptide.

[0164] The present invention provides a library of *S. Pneumoniae*-derived polypeptide sequences, and a corresponding library of nucleic acid sequences encoding the polypeptides, wherein the polypeptides themselves, or polypeptides contained within ORFs of which they form a part, comprise sequences that are contemplated for use as components of vaccines. Non-limiting examples of such sequences are listed by SEQ ID NO in Table 2, which is appended herewith and which forms part of the present specification.

[0165] The present invention also provides a library of *S. pneumoniae*-derived polypeptide sequences, and a corresponding library of nucleic acid sequences encoding the polypeptides, wherein the polypeptides themselves, or polypeptides contained within ORFs of which they form a part, comprise sequences lacking homology to any known prokaryotic or eukaryotic sequences. Such libraries provide probes, primers, and markers which can be used to diagnose *S. pneumoniae* infection, including use as markers in epidemiological studies. Non-limiting examples of such sequences are listed by SEQ ID NO in Table 2, which is appended.

[0166] The present invention also provides a library of *S. pneumoniae*-derived polypeptide sequences, and a corresponding library of nucleic acid sequences encoding the polypeptides, wherein the polypeptides themselves, or polypeptides contained within ORFs of which they form a part, comprise targets for therapeutic drugs.

Specific Example: Determination of Candidate Protein Antigens for Antibody and Vaccine Development

[0167] The selection of candidate protein antigens for vaccine development can be derived from the nucleic acids encoding *S. pneumoniae* polypeptides. First, the ORF's can be analyzed for homology to other known exported or membrane proteins and analyzed using the discriminant analysis described by Klein, et al. (Klein, P., Kanehsia, M., and DeLisi, C. (1985) *Biochimica et Biophysica Acta* 815, 468-476) for predicting exported and membrane proteins.

[0168] Homology searches can be performed using the BLAST algorithm contained in the Wisconsin Sequence Analysis Package (Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711) to compare each predicted ORF amino acid sequence with all sequences found in the current GenBank, SWISS-PROT and PIR databases. BLAST searches for local alignments between the ORF and the databank sequences and reports a probability score which indicates the probability of finding this sequence by chance in the database. ORF's with significant homology (e.g. probabilities lower than 1×10^{-6} that the homology is only due to random chance) to membrane or exported proteins represent protein antigens for vaccine development. Possible functions can be provided to *S. pneumoniae* genes based on sequence homology to genes cloned in other organisms.

[0169] Discriminant analysis (Klein, et al. supra) can be used to examine the ORF amino acid sequences. This algorithm uses the intrinsic information contained in the ORF amino acid sequence and compares it to information derived from the properties of known membrane and exported proteins. This comparison predicts which proteins will be exported, membrane associated or cytoplasmic. ORF amino acid sequences identified as exported or membrane associated by this algorithm are likely protein antigens for vaccine development.

Production of Fragments and Analogs of *S. pneumoniae* Nucleic Acids and Polypeptides

[0170] Based on the discovery of the *S. pneumoniae* gene products of the invention provided in the Sequence Listing, one skilled in the art can alter the disclosed structure (of *S. pneumoniae* genes), e.g., by producing fragments or analogs, and test the newly produced structures for activity. Examples of techniques known to those skilled in the relevant art which allow the production and testing of fragments and analogs are discussed below. These, or analogous methods can be used to make and screen libraries of polypeptides, e.g., libraries of random peptides or libraries of fragments or analogs of cellular proteins for the ability to bind *S. pneumoniae* polypeptides. Such screens are useful for the identification of inhibitors of *S. pneumoniae*.

[0171] Generation of Fragments

[0172] Fragments of a protein can be produced in several ways, e.g., recombinantly, by proteolytic digestion, or by chemical synthesis. Internal or terminal fragments of a polypeptide can be generated by removing one or more nucleotides from one end (for a terminal fragment) or both ends (for an internal fragment) of a nucleic acid which encodes the polypeptide. Expression of the mutagenized DNA produces polypeptide fragments. Digestion with "end-nibbling" endonucleases can thus generate DNA's which encode an array of fragments. DNA's which encode fragments of a protein can also be generated by random shearing, restriction digestion or a combination of the above-discussed methods.

[0173] Fragments can also be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, peptides of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or divided into overlapping fragments of a desired length.

Alteration of Nucleic Acids and Polypeptides: Random Methods

[0174] Amino acid sequence variants of a protein can be prepared by random mutagenesis of DNA which encodes a protein or a particular domain or region of a protein. Useful methods include PCR mutagenesis and saturation mutagenesis. A library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences. (Methods for screening proteins in a library of variants are elsewhere herein).

[0175] PCR Mutagenesis

[0176] In PCR mutagenesis, reduced Taq polymerase fidelity is used to introduce random mutations into a cloned fragment of DNA (Leung et al., 1989, *Technique* 1: 11-15). The DNA region to be mutagenized is amplified using the polymerase chain reaction (PCR) under conditions that reduce the fidelity of DNA synthesis by Taq DNA polymerase, e.g., by using a dGTP/dATP ratio of five and adding Mn^{2+} to the PCR reaction. The pool of amplified DNA fragments are inserted into appropriate cloning vectors to provide random mutant libraries.

[0177] Saturation Mutagenesis

[0178] Saturation mutagenesis allows for the rapid introduction of a large number of single base substitutions into cloned DNA fragments (Mayers et al., 1985, *Science* 229: 242). This technique includes generation of mutations, e.g., by chemical treatment or irradiation of single-stranded DNA in vitro, and synthesis of a complimentary DNA strand. The mutation frequency can be modulated by modulating the severity of the treatment, and essentially all possible base substitutions can be obtained. Because this procedure does not involve a genetic selection for mutant fragments both neutral substitutions, as well as those that alter function, are obtained. The distribution of point mutations is not biased toward conserved sequence elements.

[0179] Degenerate Oligonucleotides

[0180] A library of homologs can also be generated from a set of degenerate oligonucleotide sequences. Chemical synthesis of a degenerate sequences can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The synthesis of degenerate oligonucleotides is known in the art (see for example, Narang, S A (1983) *Tetrahedron* 39: 3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. A G Walton, Amsterdam: Elsevier pp 273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53: 323; Itakura et al. (1984) *Science* 198: 1056; Ike et al. (1983) *Nucleic Acid Res.* 11: 477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249: 386-390; Roberts et al. (1992) *PNAS* 89: 2429-2433; Devlin et al. (1990) *Science* 249: 404-406; Cwirla et al. (1990) *PNAS* 87: 6378-6382; as well as U.S. Pat. Nos. 5,223,409, 5,198,346, and 5,096,815).

Alteration of Nucleic Acids and Polypeptides: Methods for Directed Mutagenesis

[0181] Non-random or directed, mutagenesis techniques can be used to provide specific sequences or mutations in specific regions. These techniques can be used to create variants which include, e.g., deletions, insertions, or substitutions, of residues of the known amino acid sequence of a protein. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conserved amino acids and then with more radical choices depending upon results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

[0182] Alanine Scanning Mutagenesis

[0183] Alanine scanning mutagenesis is a useful method for identification of certain residues or regions of the desired protein that are preferred locations or domains for mutagenesis, Cunningham and Wells (*Science* 244: 1081-1085, 1989). In alanine scanning, a residue or group of target residues are identified (e.g., charged residues such as Arg, Asp, His, Lys, and Glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine). Replacement of an amino acid can affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions are then refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation per se need not be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis may be conducted at the target codon or region and the expressed desired protein subunit variants are screened for the optimal combination of desired activity.

[0184] Oligonucleotide-Mediated Mutagenesis

[0185] Oligonucleotide-mediated mutagenesis is a useful method for preparing substitution, deletion, and insertion variants of DNA, see, e.g., Adelman et al., (*DNA* 2: 183, 1983). Briefly, the desired DNA is altered by hybridizing an oligonucleotide encoding a mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the desired protein. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the desired protein DNA. Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et al. (*Proc. Natl. Acad. Sci. USA,* 75: 5765 [1978]).

[0186] Cassette Mutagenesis

[0187] Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (*Gene,* 34: 315 [1985]). The starting material is a plasmid (or other vector) which includes the protein subunit DNA to be mutated. The codon(s) in the protein subunit DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the desired protein subunit DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are comparable with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated desired protein subunit DNA sequence.

[0188] Combinatorial Mutagenesis

[0189] Combinatorial mutagenesis can also be used to generate mutants (Ladner et al., WO 88/06630). In this method, the amino acid sequences for a group of homologs or other related proteins are aligned, preferably to promote the highest homology possible. All of the amino acids which appear at a given position of the aligned sequences can be selected to create a degenerate set of combinatorial sequences. The variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential sequences are expressible as individual peptides, or alternatively, as a set of larger fusion proteins containing the set of degenerate sequences.

Other Modifications of *S. pneumoniae* Nucleic Acids and Polypeptides

[0190] It is possible to modify the structure of an *S. pneumoniae* polypeptide for such purposes as increasing solubility, enhancing stability (e.g., shelf life ex vivo and resistance to proteolytic degradation in vivo). A modified *S. pneumoniae* protein or peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition as described herein.

[0191] An *S. pneumoniae* peptide can also be modified by substitution of cysteine residues preferably with alanine, serine, threonine, leucine or glutamic acid residues to minimize dimerization via disulfide linkages. In addition, amino acid side chains of fragments of the protein of the invention can be chemically modified. Another modification is cyclization of the peptide.

[0192] In order to enhance stability and/or reactivity, an *S. pneumoniae* polypeptide can be modified to incorporate one or more polymorphisms in the amino acid sequence of the protein resulting from any natural allelic variation. Additionally, D-amino acids, non-natural amino acids, or non-amino acid analogs can be substituted or added to produce a modified protein within the scope of this invention. Furthermore, an *S. pneumoniae* polypeptide can be modified using polyethylene glycol (PEG) according to the method of A. Sehon and co-workers (Wie et al., supra) to produce a protein conjugated with PEG. In addition, PEG can be added during chemical synthesis of the protein. Other modifications of *S. pneumoniae* proteins include reduction/alkylation (Tarr, *Methods of Protein Microcharacterization*, J. E. Silver ed., Humana Press, Clifton N.J. 155-194 (1986)); acylation (Tarr, supra); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds, *Selected Methods in Cellular Immunology*, WH Freeman, San Francisco, Calif. (1980), U.S. Pat. No. 4,939,239; or mild formalin treatment (Marsh, (1971) *Int. Arch. of Allergy and Appl. Immunol,* 41: 199-215).

[0193] To facilitate purification and potentially increase solubility of an *S. pneumoniae* protein or peptide, it is possible to add an amino acid fusion moiety to the peptide backbone. For example, hexa-histidine can be added to the protein for purification by immobilized metal ion affinity chromatography (Hochuli, E. et al., (1988) *Bio/Technology,* 6: 1321-1325). In addition, to facilitate isolation of peptides free of irrelevant sequences, specific endoprotease cleavage sites can be introduced between the sequences of the fusion moiety and the peptide.

[0194] To potentially aid proper antigen processing of epitopes within an *S. pneumoniae* polypeptide, canonical protease sensitive sites can be engineered between regions, each comprising at least one epitope via recombinant or synthetic methods. For example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a protein or fragment during recombinant construction thereof. The resulting peptide can be rendered sensitive to cleavage by cathepsin and/or other trypsin-like enzymes which would generate portions of the protein containing one or more epitopes. In addition, such charged amino acid residues can result in an increase in the solubility of the peptide.

Primary Methods for Screening Polypeptides and Analogs

[0195] Various techniques are known in the art for screening generated mutant gene products. Techniques for screening large gene libraries often include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the genes under conditions in which detection of a desired activity, e.g., in this case, binding to *S. pneumoniae* polypeptide or an interacting protein, facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the techniques described below is amenable to high through-put analysis for screening large numbers of sequences created, e.g., by random mutagenesis techniques.

[0196] Two Hybrid Systems

[0197] Two hybrid assays such as the system described above (as with the other screening methods described herein), can be used to identify polypeptides, e.g., fragments or analogs of a naturally-occurring *S. pneumoniae* polypeptide, e.g., of cellular proteins, or of randomly generated polypeptides which bind to an *S. pneumoniae* protein. (The *S. pneumoniae* domain is used as the bait protein and the library of variants are expressed as prey fusion proteins.) In an analogous fashion, a two hybrid assay (as with the other screening methods described herein), can be used to find polypeptides which bind a *S. pneumoniae* polypeptide.

[0198] Display Libraries

[0199] In one approach to screening assays, the candidate peptides are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind an appropriate receptor protein via the displayed product is detected in a "panning assay". For example, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9: 1370-1371; and Goward et al. (1992) *TIBS* 18: 136-140). In a similar fashion, a detectably labeled ligand can be used to score for potentially functional peptide homologs. Fluorescently labeled ligands, e.g., receptors, can be used to detect homologs which retain ligand-binding activity. The use of fluorescently labeled ligands, allows cells to be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, to be separated by a fluorescence-activated cell sorter.

[0200] A gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at concentrations well over 10^{13} phage per milliliter, a large number ofphage can be screened at one time. Second, since each infectious phage displays a gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd., and f1 are most often used in phage display libraries. Either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle. Foreign epitopes can be expressed at the NH_2-terminal end of pIII and phage bearing such epitopes recovered from a large excess of phage lacking this epitope (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267: 16007-16010; Griffiths et al. (1993) *EMBO J.* 12: 725-734; Clackson et al. (1991) *Nature* 352: 624-628; and Barbas et al. (1992) *PNAS* 89: 4457-4461).

[0201] A common approach uses the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner (Charbit et al. (1986) *EMBO* 5, 3029-3037). Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are available for binding to ligands, e.g., to antibodies, and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA (Schorr et al. (1991) *Vaccines* 91, pp. 387-392), PhoE (Agterberg, et al. (1990) *Gene* 88, 37-45), and PAL (Fuchs et al. (1991) *Bio/Tech* 9, 1369-1372), as well as large bacterial surface structures have served as vehicles for peptide display. Peptides can be fused to pilin, a protein which polymerizes to form the pilus-a conduit for interbacterial exchange of genetic information (Thiry et al. (1989) *Appl. Environ. Microbiol.* 55, 984-993). Because of its role in interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular environment. Another large surface structure used for peptide display is the bacterial motive organ, the flagellum. Fusion of peptides to the subunit protein flagellin offers a dense array of many peptide copies on the host cells (Kuwajima et al. (1988) *Bio/Tech.* 6, 1080-1083). Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus* protein A and the outer membrane IgA protease of *Neisseria* (Hansson et al. (1992) *J. Bacteriol.* 174, 4239-4245 and Klauser et al. (1990) *EMBO J.* 9, 1991-1999).

[0202] In the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures the particle and the DNA within. An alternative scheme uses the DNA-binding protein LacI to form a link between peptide and DNA (Cull et al. (1992) *PNAS USA* 89: 1865-1869). This system uses a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3'-end. Under the controlled induction by arabinose, a LacI-peptide fusion protein is produced. This fusion retains the natural ability of LacI to bind to a short DNA sequence known as LacO operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI-peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stablely associated with the DNA sequence that directed its synthesis. The cells of the library are gently lysed and the peptide-DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands. As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and selected on a monoclonal antibody raised against the opioid peptide dynorphin B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B. (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89-1869).

[0203] This scheme, sometimes referred to as peptides-on-plasmids, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of the fusion protein, resulting in the display of the library members as peptides having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein. (Cwirla, et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 6378-6382) A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells. The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the periplasmic compartment, remaining anchored in the membrane by their C-terminal hydrophobic domains, with the N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries (Gallop et al. (1994) *J. Med. Chem.* 37(9): 1233-1251). These particular biases are not a factor in the Lac display system.

[0204] The number of small peptides available in recombinant random libraries is enormous. Libraries of 10^7-10^9 independent clones are routinely prepared. Libraries as large as 10^{11} recombinants have been created, but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this limitation, an in vitro system based on the display of nascent peptides in polysome complexes has recently been developed. This display library method has the potential of producing libraries 3-6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. Furthermore, the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

[0205] In one application of this method (Gallop et al. (1994) *J. Med. Chem.* 37(9): 1233-1251), a molecular DNA library encoding 10^{12} decapeptides was constructed and the library expressed in an *E. coli* S30 in vitro coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent peptides still linked to their encoding RNA. The polysomes are sufficiently robust to be affinity purified on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the phage display system. Following several rounds of screening, cDNA from the enriched pool of polysomes was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides on individual clones for binding activity in a phage ELISA, or for binding specificity in a completion phage ELISA (Barret, et al. (1992) *Anal. Biochem* 204, 357-364). To identify the sequences of the active peptides one sequences the DNA produced by the phagemid host.

Secondary Screening of Polypeptides and Analogs

[0206] The high through-put assays described above can be followed by secondary screens in order to identify further biological activities which will, e.g., allow one skilled in the art to differentiate agonists from antagonists. The type of a secondary screen used will depend on the desired activity that needs to be tested. For example, an assay can be developed in which the ability to inhibit an interaction between a protein of interest and its respective ligand can be used to identify antagonists from a group of peptide fragments isolated though one of the primary screens described above.

[0207] Therefore, methods for generating fragments and analogs and testing them for activity are known in the art. Once the core sequence of interest is identified, it is routine for one skilled in the art to obtain analogs and fragments.

Peptide Mimetics of *S. pneumoniae* Polypeptides

[0208] The invention also provides for reduction of the protein binding domains of the subject *S. pneumoniae* polypeptides to generate mimetics, e.g. peptide or non-peptide agents. The peptide mimetics are able to disrupt binding of a polypeptide to its counter ligand, e.g., in the case of an *S. pneumoniae* polypeptide binding to a naturally occurring ligand. The critical residues of a subject *S. pneumoniae* polypeptide which are involved in molecular recognition of a polypeptide can be determined and used to generate *S. pneumoniae*-derived peptidomimetics which competitively or noncompetitively inhibit binding of the *S. pneumoniae* polypeptide with an interacting polypeptide (see, for example, European patent applications EP-412, 762A and EP-B31,080A).

[0209] For example, scanning mutagenesis can be used to map the amino acid residues of a particular *S. pneumoniae* polypeptide involved in binding an interacting polypeptide, peptidomimetic compounds (e.g. diazepine or isoquinoline derivatives) can be generated which mimic those residues in binding to an interacting polypeptide, and which therefore can inhibit binding of an *S. pneumoniae* polypeptide to an interacting polypeptide and thereby interfere with the function of *S. pneumoniae* polypeptide. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in Peptides: *Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gama lactam rings (Garvey et al. in Peptides: *Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29: 295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, Ill., 1985), b-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26: 647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1: 1231), and b-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126: 419; and et al. (1986) *Biochem Biophys Res Commun* 134: 71).

Vaccine Formulations for *S. pneumoniae* Nucleic Acids and Polypeptides

[0210] This invention also features vaccine compositions for protection against infection by *S. pneumoniae* or for treatment of *S. pneumoniae* infection, a gram-negative spiral microaerophilic bacterium. In one embodiment, the vaccine compositions contain one or more immunogenic components such as a surface protein from *S. pneumoniae*, or portion thereof, and a pharmaceutically acceptable carrier. Nucleic acids within the scope of the invention are exemplified by the nucleic acids of the invention contained in the Sequence Listing which encode *S. pneumoniae* surface proteins. Any nucleic acid encoding an immunogenic *S. pneumoniae* protein, or portion thereof, which is capable of expression in a cell, can be used in the present invention. These vaccines have therapeutic and prophylactic utilities.

[0211] One aspect of the invention provides a vaccine composition for protection against infection by *S. pneumoniae* which contains at least one immunogenic fragment of an *S. pneumoniae* protein and a pharmaceutically acceptable carrier. Preferred fragments include peptides of at least about 10 amino acid residues in length, preferably about 10-20 amino acid residues in length, and more preferably about 12-16 amino acid residues in length.

[0212] Immunogenic components of the invention can be obtained, for example, by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding the full-length *S. pneumoniae* protein. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry.

[0213] In one embodiment, immunogenic components are identified by the ability of the peptide to stimulate T cells. Peptides which stimulate T cells, as determined by, for example, T cell proliferation or cytokine secretion are defined herein as comprising at least one T cell epitope. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to the protein allergen which is responsible for the clinical symptoms of allergy. These T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of an antigen presenting cell, thereby stimulating the T cell subpopulation with the relevant T cell receptor for the epitope. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune cells to the site of antigen/T cell interaction, and activation of the B cell cascade, leading to the production of antibodies. A T cell epitope is the basic element, or smallest unit of recognition by a T cell receptor, where the epitope comprises amino acids essential to receptor recognition (e.g., approximately 6 or 7 amino acid residues). Amino acid sequences which mimic those of the T cell epitopes are within the scope of this invention.

[0214] Screening immunogenic components can be accomplished using one or more of several different assays. For example, in vitro, peptide T cell stimulatory activity is assayed by contacting a peptide known or suspected of being immunogenic with an antigen presenting cell which presents appropriate MHC molecules in a T cell culture. Presentation of an immunogenic *S. pneumoniae* peptide in association with appropriate MHC molecules to T cells in conjunction with the necessary co-stimulation has the effect of transmitting a signal to the T cell that induces the production of increased levels of cytokines, particularly of interleukin-2 and interleukin-4. The culture supernatant can be obtained and assayed for interleukin-2 or other known cytokines. For example, any one of several conventional assays for interleukin-2 can be employed, such as the assay described in *Proc. Natl. Acad. Sci USA,* 86: 1333 (1989) the pertinent portions of which are incorporated herein by reference. A kit for an assay for the production of interferon is also available from Genzyme Corporation (Cambridge, Mass.).

[0215] Alternatively, a common assay for T cell proliferation entails measuring tritiated thymidine incorporation. The proliferation of T cells can be measured in vitro by determining the amount of ^{3}H-labeled thymidine incorporated into the replicating DNA of cultured cells. Therefore, the rate of DNA synthesis and, in turn, the rate of cell division can be quantified.

[0216] Vaccine compositions of the invention containing immunogenic components (e.g., *S. pneumoniae* polypeptide or fragment thereof or nucleic acid encoding an *S. pneumoniae* polypeptide or fragment thereof) preferably include a pharmaceutically acceptable carrier. The term “pharmaceutically acceptable carrier” refers to a carrier that does not cause an allergic reaction or other untoward effect in patients to whom it is administered. Suitable pharmaceutically acceptable carriers include, for example, one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the antibody. For vaccines of the invention containing *S. pneumoniae* polypeptides, the polypeptide is co-administered with a suitable adjuvant.

[0217] It will be apparent to those of skill in the art that the therapeutically effective amount of DNA or protein of this invention will depend, inter alia, upon the administration schedule, the unit dose of antibody administered, whether the protein or DNA is administered in combination with other therapeutic agents, the immune status and health of the patient, and the therapeutic activity of the particular protein or DNA.

[0218] Vaccine compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Methods for intramuscular immunization are described by Wolff et al. (1990) *Science* 247: 1465-1468 and by Sedegah et al. (1994) *Immunology* 91: 9866-9870. Other modes of administration include oral and pulmonary formulations, suppositories, and transdermal applications. Oral immunization is preferred over parenteral methods for inducing protection against infection by *S. pneumoniae*. Cain et. al. (1993) *Vaccine* 11: 637-642. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like.

[0219] The vaccine compositions of the invention can include an adjuvant, including, but not limited to aluminum hydroxide; N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP); N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP); N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphos-phoryloxy)-ethylamine (CGP 19835A, referred to a MTP-PE); RIBI, which contains three components from bacteria; monophosphoryl lipid A; trehalose dimycoloate; cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion; and cholera toxin. Others which may be used are non-toxic derivatives of cholera toxin, including its B subunit, and/or conjugates or genetically engineered fusions of the *S. pneumoniae* polypeptide with cholera toxin or its B subunit, procholeragenoid, fungal polysaccharides, including schizophyllan, muramyl dipeptide, muramyl dipeptide derivatives, phorbol esters, labile toxin of *E. coli*, non-*S. pneumoniae* bacterial lysates, block polymers or saponins.

[0220] Other suitable delivery methods include biodegradable microcapsules or immuno-stimulating complexes (ISCOMs), cochleates, or liposomes, genetically engineered attenuated live vectors such as viruses or bacteria, and recombinant (chimeric) virus-like particles, e.g., bluetongue. The amount of adjuvant employed will depend on the type of adjuvant used. For example, when the mucosal adjuvant is cholera toxin, it is suitably used in an amount of 5 mg to 50 mg, for example 10 mg to 35 mg. When used in the form of microcapsules, the amount used will depend on the amount employed in the matrix of the microcapsule to achieve the desired dosage. The determination of this amount is within the skill of a person of ordinary skill in the art.

[0221] Carrier systems in humans may include enteric release capsules protecting the antigen from the acidic environment of the stomach, and including *S. pneumoniae* polypeptide in an insoluble form as fusion proteins. Suitable carriers for the vaccines of the invention are enteric coated capsules and polylactide-glycolide microspheres. Suitable diluents are 0.2 N NaHCO3 and/or saline.

[0222] Vaccines of the invention can be administered as a primary prophylactic agent in adults or in children, as a secondary prevention, after successful eradication of *S. pneumoniae* in an infected host, or as a therapeutic agent in the aim to induce an immune response in a susceptible host to prevent infection by *S. pneumoniae*. The vaccines of the invention are administered in amounts readily determined by persons of ordinary skill in the art. Thus, for adults a suitable dosage will be in the range of 10 mg to 10 g, preferably 10 mg to 100 mg. A suitable dosage for adults will also be in the range of 5 mg to 500 mg. Similar dosage ranges will be applicable for children. Those skilled in the art will recognize that the optimal dose may be more or less depending upon the patient's body weight, disease, the route of administration, and other factors. Those skilled in the art will also recognize that appropriate dosage levels can be obtained based on results with known oral vaccines such as, for example, a vaccine based on an *E. coli* lysate (6 mg dose daily up to total of 540 mg) and with an enterotoxigenic *E. coli* purified antigen (4 doses of 1 mg) (Schulman et al., *J. Urol.* 150: 917-921 (1993); Boedecker et al., *American Gastroenterological Assoc.* 999: A-222 (1993)). The number of doses will depend upon the disease, the formulation, and efficacy data from clinical trials. Without intending any limitation as to the course of treatment, the treatment can be administered over 3 to 8 doses for a primary immunization schedule over 1 month (Boedeker, *American Gastroenterological Assoc.* 888: A-222 (1993)).

[0223] In a preferred embodiment, a vaccine composition of the invention can be based on a killed whole *E. coli* preparation with an immunogenic fragment of an *S. pneumoniae* protein of the invention expressed on its surface or it can be based on an *E. coli* lysate, wherein the killed *E. coli* acts as a carrier or an adjuvant.

[0224] It will be apparent to those skilled in the art that some of the vaccine compositions of the invention are useful only for preventing *S. pneumoniae* infection, some are useful only for treating *S. pneumoniae* infection, and some are useful for both preventing and treating *S. pneumoniae* infection. In a preferred embodiment, the vaccine composition of the invention provides protection against *S. pneumoniae* infection by stimulating humoral and/or cell-mediated immunity against *S. pneumoniae*. It should be understood that amelioration of any of the symptoms of *S. pneumoniae* infection is a desirable clinical goal, including a lessening of the dosage of medication used to treat *S. pneumoniae-caused* disease, or an increase in the production of antibodies in the serum or mucous of patients.

Antibodies Reactive with *S. pneumoniae* Polypeptides

[0225] The invention also includes antibodies specifically reactive with the subject *S. pneumoniae* polypeptide. Anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, Antibodies: A Laboratory Manual ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of the subject *S. pneumoniae* polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies.

[0226] In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of the *S. pneumoniae* polypeptides of the invention, e.g. antigenic determinants of a polypeptide of the invention contained in the Sequence Listing, or a closely related human or non-human mammalian homolog (e.g., 90% homologous, more preferably at least 95% homologous). In yet a further preferred embodiment of the invention, the anti-*S. pneumoniae* antibodies do not substantially cross react (i.e., react specifically) with a protein which is for example, less than 80% percent homologous to a sequence of the invention contained in the Sequence Listing. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is less than 10 percent, more preferably less than 5 percent, and even more preferably less than 1 percent, of the binding affinity for a protein of the invention contained in the Sequence Listing. In a most preferred embodiment, there is no cross-reactivity between bacterial and mammalian antigens.

[0227] The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with *S. pneumoniae* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, $F(ab')_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab')_2$ fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the invention is further intended to include bispecific and chimeric molecules having an anti-*S. pneumoniae* portion.

[0228] Both monoclonal and polyclonal antibodies (Ab) directed against *S. pneumoniae* polypeptides or *S. pneumoniae* polypeptide variants, and antibody fragments such as Fab' and $F(ab')_2$, can be used to block the action of *S. pneumoniae* polypeptide and allow the study of the role of a particular *S. pneumoniae* polypeptide of the invention in aberrant or unwanted intracellular signaling, as well as the normal cellular function of the *S. pneumoniae* and by microinjection of anti-*S. pneumoniae* polypeptide antibodies of the present invention.

[0229] Antibodies which specifically bind *S. pneumoniae* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of *S. pneumoniae* antigens. Anti *S. pneumoniae* polypeptide antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *S. pneumoniae* levels in tissue or bodily fluid as part of a clinical testing procedure. Likewise, the ability to monitor *S. pneumoniae* polypeptide levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of an *S. pneumoniae* polypeptide can be measured in cells found in bodily fluid, such as in urine samples or can be measured in tissue, such as produced by gastric biopsy. Diagnostic assays using anti-*S. pneumoniae* antibodies can include, for example, immunoassays designed to aid in early diagnosis of *S. pneumoniae* infections. The present invention can also be used as a method of detecting antibodies contained in samples from individuals infected by this bacterium using specific *S. pneumoniae* antigens.

[0230] Another application of anti-*S. pneumoniae* polypeptide antibodies of the invention is in the immunological screening of cDNA libraries constructed in expression vectors such as lgt11, lgt18-23, lZAP, and lORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, lgt11 will produce fusion proteins whose amino termini consist of β-galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a subject *S. pneumoniae* polypeptide can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*S. pneumoniae* polypeptide antibodies. Phage, scored by this assay, can then be isolated from the infected plate. Thus, the presence of *S. pneumoniae* gene homologs can be detected and cloned from other species, and alternate isoforms (including splicing variants) can be detected and cloned.

Kits Containing Nucleic Acids, Polypeptides or Antibodies of the Invention

[0231] The nucleic acid, polypeptides and antibodies of the invention can be combined with other reagents and articles to form kits. Kits for diagnostic purposes typically comprise the nucleic acid, polypeptides or antibodies in vials or other suitable vessels. Kits typically comprise other reagents for performing hybridization reactions, polymerase chain reactions (PCR), or for reconstitution of lyophilized components, such as aqueous media, salts, buffers, and the like. Kits may also comprise reagents for sample processing such as detergents, chaotropic salts and the like. Kits may also comprise immobilization means such as particles, supports, wells, dipsticks and the like. Kits may also comprise labeling means such as dyes, developing reagents, radioisotopes, fluorescent agents, luminescent or chemiluminescent agents, enzymes, intercalating agents and the like. With the nucleic acid and amino acid sequence information provided herein, individuals skilled in art can readily assemble kits to serve their particular purpose. Kits further can include instructions for use.

Bio Chips and Microarrays

[0232] The nucleic acid sequence of the present invention may be used to detect *S. pneumoniae* or other species of *Streptococcus* acid sequence using bio chip technology. Bio chips containing arrays of nucleic acid sequence can also be used to measure expression of genes of *S. pneumoniae* or other species of *Streptococcus*. For example, to diagnose a patient with a *S. pneumoniae* or other *Streptococcus* infection, a sample from a human or animal can be used as a probe on a bio chip containing an array of nucleic acid sequence from the present invention. In addition, a sample from a disease state can be compared to a sample from a non-disease state which would help identify a gene that is up-regulated or expressed in the disease state. This would provide valuable insight as to the mechanism by which the disease manifests. Changes in gene expression can also be used to identify critical pathways involved in drug transport or metabolism, and may enable the identification of novel targets involved in virulence or host cell interactions involved in maintenance of an infection. Procedures using such techniques have been described by Brown et al., 1995, *Science* 270: 467-470.

[0233] Bio chips can also be used to monitor the genetic changes of potential therapeutic compounds including, deletions, insertions or mismatches. Once the therapeutic is added to the patient, changes to the genetic sequence can be evaluated for its efficacy. In addition, the nucleic acid sequence of the present invention can be used to determine essential genes in cell cycling. As described in Iyer et al., 1999 (*Science,* 283: 83-87) genes essential in the cell cycle can be identified using bio chips. Furthermore, the present invention provides nucleic acid sequence which can be used with bio chip technology to understand regulatory networks in bacteria, measure the response to environmental signals or drugs as in drug screening, and study virulence induction. (Mons et al., 1998*, Nature Biotechnology,* 16: 45-48. Patents teaching this technology include U.S. Pat. Nos. 5,445,934, 5,744,305, and 5800992.

Drug Screening Assays Using *S. pneumoniae* Polypeptides

[0234] By making available purified and recombinant *S. pneumoniae* polypeptides, the present invention provides assays which can be used to screen for drugs which are either agonists or antagonists of the normal cellular function, in this case, of the subject *S. pneumoniae* polypeptides, or of their role in intracellular signaling. Such inhibitors or potentiators may be useful as new therapeutic agents to combat *S. pneumoniae* infections in humans. A variety of assay formats will suffice and, in light of the present inventions, will be comprehended by the skilled artisan.

[0235] In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or change in enzymatic properties of the molecular target. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with an isolated and purified *S. pneumoniae* polypeptide.

[0236] Screening assays can be constructed in vitro with a purified *S. pneumoniae* polypeptide or fragment thereof, such as an *S. pneumoniae* polypeptide having enzymatic activity, such that the activity of the polypeptide produces a detectable reaction product. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. Suitable products include those with distinctive absorption, fluorescence, or chemiluminescence properties, for example, because detection may be easily automated. A variety of synthetic or naturally occurring compounds can be tested in the assay to identify those which inhibit or potentiate the activity of the *S. pneumoniae* polypeptide. Some of these active compounds may directly, or with chemical alterations to promote membrane permeability or solubility, also inhibit or potentiate the same activity (e.g., enzymatic activity) in whole, live *S. pneumoniae* cells.

Overexpression Assays

[0237] Overexpression assays are based on the premise that overproduction of a protein would lead to a higher level of resistance to compounds that selectively interfere with the function of that protein. Overexpression assays may be used to identify compounds that interfere with the function of virtually any type of protein, including without limitation enzymes, receptors, DNA- or RNA-binding proteins, or any proteins that are directly or indirectly involved in regulating cell growth.

[0238] Typically, two bacterial strains are constructed. One contains a single copy of the gene of interest, and a second contains several copies of the same gene. Identification of useful inhibitory compounds of this type of assay is based on a comparison of the activity of a test compound in inhibiting growth and/or viability of the two strains. The method involves constructing a nucleic acid vector that directs high level expression of a particular target nucleic acid. The vectors are then transformed into host cells in single or multiple copies to produce strains that express low to moderate and high levels of protein encoding by the target sequence (strain A and B, respectively). Nucleic acid comprising sequences encoding the target gene can, of course, be directly integrated into the host cell.

[0239] Large numbers of compounds (or crude substances which may contain active compounds) are screened for their effect on the growth of the two strains. Agents which interfere with an unrelated target equally inhibit the growth of both strains. Agents which interfere with the function of the target at high concentration should inhibit the growth of both strains. It should be possible, however, to titrate out the inhibitory effect of the compound in the overexpressing strain. That is, if the compound is affecting the particular target that is being tested, it should be possible to inhibit the growth of strain A at a concentration of the compound that allows strain B to grow.

[0240] Alternatively, a bacterial strain is constructed that contains the gene of interest under the control of an inducible promoter. Identification of useful inhibitory agents using this type of assay is based on a comparison of the activity of a test compound in inhibiting growth and/or viability of this strain under both inducing and non-inducing conditions. The method involves constructing a nucleic acid vector that directs high-level expression of a particular target nucleic acid. The vector is then transformed into host cells that are grown under both non-inducing and inducing conditions (conditions A and B, respectively).

[0241] Large numbers of compounds (or crude substances which may contain active compounds) are screened for their effect on growth under these two conditions. Agents that interfere with the function of the target should inhibit growth under both conditions. It should be possible, however, to titrate out the inhibitory effect of the compound in the overexpressing strain. That is, if the compound is affecting the particular target that is being tested, it should be possible to inhibit growth under condition A at a concentration that allows the strain to grow under condition B.

[0242] Ligand-Binding Assays

[0243] Many of the targets according to the invention have functions that have not yet been identified. Ligand-binding assays are useful to identify inhibitor compounds that interfere with the function of a particular target, even when that function is unknown. These assays are designed to detect binding of test compounds to particular targets. The detection may involve direct measurement of binding. Alternatively, indirect indications of binding may involve stabilization of protein structure or disruption of a biological function. Non-limiting examples of useful ligand-binding assays are detailed below.

[0244] A useful method for the detection and isolation of binding proteins is the Biomolecular Interaction Assay (BIAcore) system developed by Pharmacia Biosensor and described in the manufacturer's protocol (LKB Pharmacia, Sweden). The BIAcore system uses an affinity purified anti-GST antibody to immobilize GST-fusion proteins onto a sensor chip. The sensor utilizes surface plasmon resonance which is an optical phenomenon that detects changes in refractive indices. In accordance with the practice of the invention, a protein of interest is coated onto a chip and test compounds are passed over the chip. Binding is detected by a change in the refractive index (surface plasmon resonance).

[0245] A different type of ligand-binding assay involves scintillation proximity assays (SPA, described in U.S. Pat. No. 4,568,649).

[0246] Another type of ligand binding assay, also undergoing development, is based on the fact that proteins containing mitochondrial targeting signals are imported into isolated mitochondria in vitro (Hurt et al., 1985, *Embo J.* 4: 2061-2068; Eilers and Schatz, *Nature,* 1986, 322: 228-231). In a mitochondrial import assay, expression vectors are constructed in which nucleic acids encoding particular target proteins are inserted downstream of sequences encoding mitochondrial import signals. The chimeric proteins are synthesized and tested for their ability to be imported into isolated mitochondria in the absence and presence of test compounds. A test compound that binds to the target protein should inhibit its uptake into isolated mitochondria in vitro.

[0247] Another ligand-binding assay is the yeast two-hybrid system (Fields and Song, 1989, *Nature* 340: 245-246). The yeast two-hybrid system takes advantage of the properties of the GAL4 protein of the yeast *Saccharomyces cerevisiae*. The GAL4 protein is a transcriptional activator required for the expression of genes encoding enzymes of galactose utilization. This protein consists of two separable and functionally essential domains: an N-terminal domain which binds to specific DNA sequences (UAS_G); and a C-terminal domain containing acidic regions, which is necessary to activate transcription. The native GAL4 protein, containing both domains, is a potent activator of transcription when yeast are grown on galactose media. The N-terminal domain binds to DNA in a sequence-specific manner but is unable to activate transcription. The C-terminal domain contains the activating regions but cannot activate transcription because it fails to be localized to UAS_G. In the two-hybrid system, a system of two hybrid proteins containing parts of GAL4: (1) a GAL4 DNA-binding domain fused to a protein 'X' and (2) a GAL4 activation region fused to a protein 'Y'. If X and Y can form a protein-protein complex and reconstitute proximity of the GAL4 domains, transcription of a gene regulated by UAS_G occurs. Creation of two hybrid proteins, each containing one of the interacting proteins X and Y, allows the activation region of UAS_G to be brought to its normal site of action.

[0248] The binding assay described in Fodor et al., 1991, *Science* 251: 767-773, which involves testing the binding affinity of test compounds for a plurality of defined polymers synthesized on a solid substrate, may also be useful.

[0249] Compounds which bind to the polypeptides of the invention are potentially useful as antibacterial agents for use in therapeutic compositions.

[0250] Pharmaceutical formulations suitable for antibacterial therapy comprise the antibacterial agent in conjunction with one or more biologically acceptable carriers. Suitable biologically acceptable carriers include, but are not limited to, phosphate-buffered saline, saline, deionized water, or the like. Preferred biologically acceptable carriers are physiologically or pharmaceutically acceptable carriers.

[0251] The antibacterial compositions include an antibacterial effective amount of active agent. Antibacterial effective amounts are those quantities of the antibacterial agents of the present invention that afford prophylactic protection against bacterial infections or which result in amelioration or cure of an existing bacterial infection. This antibacterial effective amount will depend upon the agent, the location and nature of the infection, and the particular host. The amount can be determined by experimentation known in the art, such as by establishing a matrix of dosages and frequencies and comparing a group of experimental units or subjects to each point in the matrix.

[0252] The antibacterial active agents or compositions can be formed into dosage unit forms, such as for example, creams, ointments, lotions, powders, liquids, tablets, capsules, suppositories, sprays, aerosols or the like. If the antibacterial composition is formulated into a dosage unit form, the dosage unit form may contain an antibacterial effective amount of active agent. Alternatively, the dosage unit form may include less than such an amount if multiple dosage unit forms or multiple dosages are to be used to administer a total dosage of the active agent. Dosage unit forms can include, in addition, one or more excipient(s), diluent(s), disintegrant(s), lubricant(s), plasticizer(s), colorant(s), dosage vehicle(s), absorption enhancer(s), stabilizer(s), bactericide(s), or the like.

[0253] For general information concerning formulations, see, e.g., Gilman et al. (eds.), 1990*, Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 8th ed., Pergamon Press; and *Remington's Pharmaceutical Sciences,* 17th ed., 1990, Mack Publishing Co., Easton, Pa.; Avis et al. (eds.), 1993*, Pharmaceutical Dosage Forms: Parenteral Medications*, Dekker, New York; Lieberman et al (eds.), 1990*, Pharmaceutical Dosage Forms: Disperse Systems*, Dekker, New York.

[0254] The antibacterial agents and compositions of the present invention are useful for preventing or treating *S. pneumoniae* infections. Infection prevention methods incorporate a prophylactically effective amount of an antibacterial agent or composition. A prophylactically effective amount is an amount effective to prevent *S. pneumoniae* infection and will depend upon the specific bacterial strain, the agent, and the host. These amounts can be determined experimentally by methods known in the art and as described above.

[0255] *S. pneumoniae* infection treatment methods incorporate a therapeutically effective amount of an antibacterial agent or composition. A therapeutically effective amount is an amount sufficient to ameliorate or eliminate the infection. The prophylactically and/or therapeutically effective amounts can be administered in one administration or over repeated administrations. Therapeutic administration can be followed by prophylactic administration, once the initial bacterial infection has been resolved.

[0256] The antibacterial agents and compositions can be administered topically or systemically. Topical application is typically achieved by administration of creams, ointments, lotions, or sprays as described above. Systemic administration includes both oral and parental routes. Parental routes include, without limitation, subcutaneous, intramuscular, intraperitoneal, intravenous, transdermal, inhalation and intranasal administration.

Exemplification

I. Cloning and Sequencing of *S. pneumoniae* DNA

[0257] *S. pneumoniae* chromosomal DNA was isolated according to a basic DNA protocol outlined in Schleif R. F. and Wensink P. C., Practical Methods in Molecular Biology, p. 98, Springer-Verlag, NY., 1981, with minor modifications. Briefly, cells were pelleted, resuspended in TE (10 mM Tris, 1 mM EDTA, pH 7.6) and GES lysis buffer (5.1 M guanidium thiocyanate, 0.1 M EDTA, pH 8.0, 0.5% N-laurylsarcosine) was added. Suspension was chilled and ammonium acetate (NH4Ac) was added to final concentration of 2.0 M. DNA was extracted, first with chloroform, then with phenol-chloroform, and reextracted with chloroform. DNA was precipitated with isopropanol, washed twice with 70% EtOH, dried and resuspended in TE.

[0258] Following isolation whole genomic *S. pneumoniae* DNA was nebulized (Bodenteich et al., Automated DNA Sequencing and Analysis (J. C. Venter, ed.), Academic Press, 1994) to a median size of 2000 bp. After nebulization, the DNA was concentrated and separated on a standard 1% agarose gel. Several fractions, corresponding to approximate sizes 1000-1500 bp, 1500-2000 bp, 2000-2500 bp, 2500-3000 bp, were excised from the gel and purified by the GeneClean procedure (Bio101, Inc.).

[0259] The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The healed DNA was then ligated to unique BstXI-linker adapters (5' GTCTTCACCACGGGG and 5' GTGGTGAAGAC in 100-1000 fold molar excess). These linkers are complimentary to the BstXI-cut pMPX vectors, while the overhang is not self-complimentary. Therefore, the linkers will not concatemerize nor will the cut-vector religate itself easily. The linker-adopted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean. The linker-adopted inserts were then ligated to each of 20 pMPX vectors to construct a series of "shotgun" subclone libraries. Blunt ended vector was used for cloning into the PUC19 vector. The vectors contain an out-of-frame lacZ gene at the cloning site which becomes in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

[0260] All subsequent steps were based either on the multiplex DNA sequencing protocols outlined in Church G. M. and Kieffer-Higgins S., Science 240: 185-188, 1988 or by ABI377 automated DNA sequencing methods. Only major modifications to the protocols are highlighted. Briefly, each of the 20 vectors was then transformed into DHSa competent cells (Gibco/BRL, DH5a transformation protocol). The libraries were assessed by plating onto antibiotic plates containing ampicillin, methicillin and IPTG/Xgal. The plates were incubated overnight at 37° C. Successful transformants were then used for plating of clones and pooling into the multiplex pools. The clones were picked and pooled into 40 ml growth medium cultures. The cultures were grown overnight at 37° C. DNA was purified using the Qiagen Midi-prep kits and Tip-100 columns (Qiagen, Inc.). In this manner, 100 mg of DNA was obtained per pool.

[0261] These purified DNA samples were then sequenced either using the multiplex DNA sequencing based on chemical degradation methods (Church G. M. and Kieffer-Higgins S., Science 240: 185-188, 1988) or by Sequithrem (Epicenter Technologies) dideoxy sequencing protocols or by ABI dye-terminator chemistry. For the multiplex portion the sequencing reactions were electrophoresed and transferred onto nylon membranes by direct transfer electrophoresis from 40 cm gels (Richterich P. and Church G. M., Methods in Enzymology 218: 187-222, 1993). The DNA was covalently bound to the membranes by exposure to ultraviolet light, and hybridized with labeled oligonucleotides complimentary to tag sequences on the vectors (Church, supra). The membranes were washed to rinse off non-specifically bound probe, and exposed to X-ray film to visualize individual sequence ladders. After autoradiography, the hybridized probe was removed by incubation at 65° C., and the hybridization cycle repeated with another tag sequence until the membrane had been probed 41 times. Thus, each gel produced a large number of films, each containing new sequencing information. Whenever a new blot was processed, it was initially probed for an internal standard sequence added to each of the pools. Digital images of the films were generated using a laser-scanning densitometer (Molecular Dynamics, Sunnyvale, Calif.). The digitized images were processed on computer workstations (VaxStation 4000's) using the program REPLICA™ (Church et al., Automated DNA Sequencing and Analysis (J. C. Venter, ed.), Academic Press, 1994). Image processing included lane straightening, contrast adjustment to smooth out intensity differences, and resolution enhancement by iterative gaussian deconvolution. The sequences were then converted to an SCF format so that processing and assembly could proceed on UNIX machines. The ABI dye terminator sequence reads were run on ABI377 machines and the data was directly transferred to UNIX machines following lane tracking of the gels. All multiplex and ABI reads were assembled using PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, January 1996, p. 157) with default parameters and not using quality scores. The initial assembly was done at 7fold coverage and yielded 511 contigs. Short read length fragments of 200 bp or less found on the ends of contigs facing in the appropriate direction were used to extend off the end of the contigs. These reads were then resequenced with primers using ABI technology to give sequences with a read length of 500 or more bases. This allowed end extensions to be performed without ordering new primers. In addition, missing mates (sequences from clones that only gave one strand reads) were identified and sequenced with ABI technology to allow the identification of additional overlapping contigs.

[0262] End-sequencing of randomly picked genomic lambda was also performed. Sequencing on a both sides was done for all lambda sequences. The lambda library backbone helped to verify the integrity of the assembly and allowed closure of some of the physical gaps.

[0263] To identify *S. pneumoniae* polypeptides the complete genomic sequence of *S. pneumoniae* were analyzed essentially as follows: First, all possible stop-to-stop open reading frames (ORFs) greater than 180 nucleotides in all six reading frames were translated into amino acid sequences. Second, the identified ORFs were analyzed for homology to known (archeabacter, prokaryotic and eukaryotic) protein sequences. Third, the predicted coding regions of the sequences and start codons were evaluated with the programs GENEMARK™ (Borodovsky and McIninch, 1993, Comp. Chem. 17: 123) and Glimmer (Fraser et al, *Nature,* 1997).

Identification, Cloning and Expression of *S. pneumoniae* Nucleic Acids

[0264] Expression and purification of the *S. pneumoniae* polypeptides of the invention can be performed essentially as outlined below.

[0265] To facilitate the cloning, expression and purification of membrane and secreted proteins from *S. pneumoniae*, a gene expression system, such as the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli*, is selected. Also, a DNA sequence encoding a peptide tag, the His-Tag, is fused to the 3' end of DNA sequences of interest in order to facilitate purification of the recombinant protein products. The 3' end is selected for fusion in order to avoid alteration of any 5' terminal signal sequence.

[0266] PCR Amplification and Cloning of Nucleic Acids Containing ORF's Encoding Enzymes

[0267] Nucleic acids chosen (for example, from the nucleic acids set forth in SEQ ID NO: 1-SEQ ID NO: 2661) for cloning from the 14453 strain of *S. pneumoniae* are prepared for amplification cloning by polymerase chain reaction (PCR). Synthetic oligonucleotide primers specific for the 5' and 3' ends of open reading frames (ORFs) are designed and purchased from GibcoBRL Life Technologies (Gaithersburg, Md., USA). All forward primers (specific for the 5' end of the sequence) are designed to include an NcoI cloning site at the extreme 5' terminus. These primers are designed to permit initiation of protein translation at a methionine residue followed by a valine residue and the coding sequence for the remainder of the native *S. pneumoniae* DNA sequence. All reverse primers (specific for the 3' end of any *S. pneumoniae* ORF) include a EcoRI site at the extreme 5' terminus to permit cloning of each *S. pneumoniae* sequence into the reading frame of the pET-28b. The pET-28b vector provides sequence encoding an additional 20 carboxy-terminal amino acids including six histidine residues (at the extreme C-terminus), which comprise the His-Tag.

[0268] Genomic DNA prepared from strain 14453 of *S. pneumoniae* is used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). To amplify a DNA sequence containing an *S. pneumoniae* ORF, genomic DNA (50 nanograms) is introduced into a reaction vial containing 2 mM $MgCl_2$, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined *S. pneumoniae* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, N.J., USA) in a final volume of 100 microliters.

[0269] Upon completion of thermal cycling reactions, each sample of amplified DNA is washed and purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, Md., USA). All amplified DNA samples are subjected to digestion with the restriction endonucleases, e.g., NcoI and EcoRI (New England BioLabs, Beverly, Mass., USA) (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). DNA samples are then subjected to electrophoresis on 1.0% NuSeive (FMC BioProducts, Rockland, Me. USA) agarose gels. DNA is visualized by exposure to ethidium bromide and long wave uv irradiation. DNA contained in slices isolated from the agarose gel is purified using the Bio 101 GeneClean Kit protocol (Bio 101 Vista, Calif., USA).

Cloning of *S. pneumoniae* Nucleic Acids into an Expression Vector

[0270] The pET-28b vector is prepared for cloning by digestion with endonucleases, e.g., NcoI and EcoRI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). The pET-28a vector, which encodes a His-Tag that can be fused to the 5' end of an inserted gene, is prepared by digestion with appropriate restriction endonucleases.

[0271] Following digestion, DNA inserts are cloned (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) into the previously digested pET-28b expression vector. Products of the ligation reaction are then used to transform the BL21 strain of *E. coli* (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) as described below.

[0272] Transformation of Competent Bacteria with Recombinant Plasmids

[0273] Competent bacteria, *E coli* strain BL21 or *E. coli* strain BL21 (DE3), are transformed with recombinant pET expression plasmids carrying the cloned *S. pneumoniae* sequences according to standard methods (Current Protocols in Molecular, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). Briefly, 1 microliter of ligation reaction is mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which, samples are incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4 and 20, mM glucose) at 37° C. with shaking for 1 hour. Samples are then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 are then picked and analyzed to evaluate cloned inserts as described below.

[0274] Identification Of Recombinant Expression Vectors with *S. pneumoniae* Nucleic Acids

[0275] Individual BL21 clones transformed with recombinant pET-28b *S. pneumoniae* ORFs are analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *S. pneumoniae* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verifies the integration of the *S. pneumoniae* sequences in the expression vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994).

[0276] Isolation and Preparation of Nucleic Acids from Transformants

[0277] Individual clones of recombinant pET-28b vectors carrying properly cloned *S. pneumoniae* ORFs are picked and incubated in 5 mls of LB broth plus 25 microgram/ml kanamycin sulfate overnight. The following day plasmid DNA is isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, Calif., USA).

[0278] Expression Of Recombinant *S. pneumoniae* Sequences in *E. coli*

[0279] The pET vector can be propagated in any *E. coli* K-12 strain e.g. HMS 174, HB101, JM109, DH5, etc. for the purpose of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase is induced by addition of isopropyl-B-D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid, such as pET-28b, carrying its gene of interest. Strains used include: BL21(DE3) (Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Meth. Enzymol. 185, 60-89).

[0280] To express recombinant *S. pneumoniae* sequences, 50 nanograms of plasmid DNA isolated as described above is used to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression system kit). The lacZ gene (beta-galactosidase) is expressed in the pET-System as described for the *S. pneumoniae* recombinant constructions. Transformed cells are cultured in SOC medium for 1 hour, and the culture is then plated on LB plates containing 25 micrograms/ml kanamycin sulfate. The following day, bacterial colonies are pooled and grown in LB medium containing kanamycin sulfate (25 micrograms/ml) to an optical density at 600 nM of 0.5 to 1.0 O.D. units, at which point, 1 millimolar IPTG was added to the culture for 3 hours to induce gene expression of the *S. pneumoniae* recombinant DNA constructions.

[0281] After induction of gene expression with IPTG, bacteria are pelleted by centrifugation in a Sorvall RC-3B centrifuge at 3500×g for 15 minutes at 4° C. Pellets are resuspended in 50 milliliters of cold 10 mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.1 mM EDTA (STE buffer). Cells are then centrifuged at 2000×g for 20 min at 4° C. Wet pellets are weighed and frozen at −80° C. until ready for protein purification.

[0282] A variety of methodologies known in the art can be utilized to purify the isolated proteins. (Current Protocols in

Protein Science, John Wiley and Sons, Inc., J. E. Coligan et al., eds., 1995). For example, the frozen cells may be thawed, resupended in buffer and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International Corporation, Newton, Mass.). The resultant homogenate may be centrifuged to yield a clear supernatant (crude extract) and following filtration the crude extract may be fractionated over columns. Fractions may be monitored by absorbance at OD_{280} nm. and peak fractions may analyzed by SDS-PAGE.

[0283] The concentrations of purified protein preparations may be quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, S. J. 1986 Eur. J. Biochem. 157, 169-180). Protein concentrations are also measured by the method of Bradford, M. M. (1976) Anal. Biochem. 72, 248-254, and Lowry, O. H., Rosebrough, N., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. 193, pages 265-275, using bovine serum albumin as a standard.

[0284] SDS-polyacrylamide gels of various concentrations may be purchased from BioRad (Hercules, Calif., USA), and stained with Coomassie blue. Molecular weight markers may include rabbit skeletal muscle myosin (200 kDa), *E. coli* (-galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

Equivalents

[0285] While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0001	1	2662	111	333	552	2.20E−74	[GI:2804700]	87
ORF NAME	NTID	AAID	AALN	NTLN	SCORE	P-VALUE	DESCRIPTION	
SPX0001	1	2662	111	333	552	2.20E-74	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX0002	2	2663	173	519	885	3.60E-120	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX0003	3	2664	109	327	281	1.40E-34	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX0004	4	2665	190	570	948	7.40E-131	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX0005	5	2666	122	366			NO-HIT	6
SPX0006	6	2667	401	1203	815	7.40E-125	[LN:RS1__LEULA] [AC:P50889:P71450] [GN:RPS1] [OR:Leuconostoc lactis] [DE:40S RIBOSOMAL PROTEIN S1] [SP:P50889:P71450]	116
SPX0007	7	2668	79	237	89	2.90E-05	[LN:E64801] [AC:E64801] [PN:hypothetical protein b0663] [OR:Escherichia coli]	77
SPX0008	8	2669	77	231			NO-HIT	6
SPX0009	9	2670	215	645	832	6.30E-114	[LN:PCP__STRPY] [AC:Q01328] [GN:PCP] [OR:Streptococcus pyogenes] [EC:3.4.19.3] [DE:PEPTIDASE) (PYROGLUTAMYL-PEPTIDASE I) (PGP-I) (PYRASE)] [SP:Q01328]	149
SPX0010	10	2671	308	924	234	3.90E-61	[LN:A81402] [AC:A81402] [PN:probable integral membrane protein Cj0553 [imported]] [GN:Cj0553] [OR:Campylobacter jejuni]	119
SPX0011	11	2672	237	711	208	1.30E-37	[GI:6002226]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SC51A] [AC:AL121596] [PN:possible integral membrane protein] [GN:SCF51A.18c] [OR:Streptomyces coelicolor A3(2)]	
SPX0012	12	2673	235	705	213	1.30E-53	[GI:2804687] [LN:AF030359] [AC:AF030359] [PN:oligopeptide binding protein] [GN:aliA] [OR:Streptococcus pneumoniae]	114
SPX0013	13	2674	341	1023	1359	9.60E-184	[GI:6979306] [LN:AF164204] [AC:AF164204] [PN:branched-chain amino acid aminotransferase] [GN:bcaT] [FN:catalyzes the first reaction in the catabolism] [OR:Lactococcus lactis subsp. cremoris]	190
SPX0014	14	2675	824	2472	4143	0	[GI:6851038] [LN:SPPARCETP] [AC:Z67739] [PN:DNA topoisomerase IV] [GN:parC] [OR:Streptococcus pneumoniae]	105
SPX0015	15	2676	648	1944	3318	0	[GI:6851037] [LN:SPPARCETP] [AC:Z67739] [PN:DNA topoisomerase IV] [GN:parE] [OR:Streptococcus pneumoniae]	105
SPX0016	16	2677	214	642	1073	2.50E-150	[LN:YPAE_STRPN] [AC:Q54916] [OR:Streptococcus pneumoniae] [DE:HYPOTHETICAL 23.0 KD PROTEIN IN PARE 5'REGION (ORF2)] [SP:Q54916]	127
SPX0017	17	2678	410	1230	1044	8.20E-270	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX0018	18	2679	278	834	591	9.50E-139	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0019	19	2680	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0020	20	2681	34	102	124	5.30E-12	[LN:B30868]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	
SPX0021	21	2682	319	957	1002	3.80E-138	[LN:F70009] [AC:F70009] [PN:conserved hypothetical protein yufQ] [GN:yufQ] [CL:probable ribose ABC transporter rbsC-2] [OR:Bacillus subtilis]	141
SPX0022	22	2683	362	1086	385	3.70E-85	[LN:E70009] [AC:E70009] [PN:conserved hypothetical protein yufP] [GN:yufP] [OR:Bacillus subtilis]	97
SPX0023	23	2684	512	1536	1623	2.30E-217	[LN:D70009] [AC:D70009] [PN:probable ABC transporter yufO] [GN:yufO] [CL:Bacillus subtilis probable ABC transporter yufO:ATP-binding cassette homology] [OR:Bacillus subtilis]	174
SPX0024	24	2685	140	420	83	1.80E-06	[LN:G72493] [AC:G72493] [PN:hypothetical protein APE2590] [GN:APE2590] [OR:Aeropyrum pernix]	92
SPX0025	25	2686	364	1092	354	8.80E-82	[LN:YUFN_BACSU] [AC:O05252] [GN:YUFN] [OR:Bacillus subtilis] [DE:HYPOTHETICAL LIPOPROTEIN YUFN PRECURSOR]	117
SPX0026	26	2687	130	390	324	5.40E-40	[SP:O05252] [GI:6478520] [LN:AF187304] [AC:AF187304] [PN:cytidine deaminase] [GN:cdd] [OR:Bacillus caldolyticus]	100
SPX0027	27	2688	100	300	276	2.40E-33	[LN:D75424] [AC:D75424] [PN:deoxyribose-phosphate aldolase] [GN:DR1205] [CL:deoxyribose-phosphate aldolase] [OR:Deinococcus radiodurans]	136
SPX0028	28	2689	61	183			NO-HIT	6
SPX0029	29	2690	119	357	344	3.30E-42	[LN:DEOC_BACSU] [AC:P39121] [GN:DRA] [OR:Bacillus subtilis] [EC:4.1.2.4] [DE:(DEOXYRIBOALDOLASE)]	109

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P39121]	
SPX0030	30	2691	426	1278	921	9.70E-147	“[LN:PDP_BACST] [AC:P77836] [GN:PYN] [OR:Bacillus stearothermophilus] [EC:2.4.2.2] [DE:PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE, (PYNP)] [SP:P77836]”	144
SPX0031	31	2692	197	591	250	2.00E-56	[LN:S59955] [AC:S59955] [PN:hypothetical protein 202] [CL:hypothetical protein MJ0882] [OR:Staphylococcus aureus]	113
SPX0032	32	2693	73	219			NO-HIT	6
SPX0033	33	2694	100	300			NO-HIT	6
SPX0034	34	2695	317	951	345	3.80E-79	[GI:6960352] [LN:STYSTMF1] [AC:AF170176] [PN:96% identity over 316 amino acids with E. coli] [GN:coaA] [OR:Salmonella typhimurium LT2]	134
SPX0035	35	2696	79	237	98	8.40E-17	[LN:R3EC20] [AC:A30425:A02748:S40547:G64722:S07374] [PN:ribosomal protein S20/L26 [validated] :ribosomal protein L26:ribosomal protein S20] [GN:rpsT] [CL:Escherichia coli ribosomal protein S20] [OR:Escherichia coli]	214
SPX0036	36	2697	175	525	270	6.30E-50	[GI:6434823] [LN:SPU78969] [AC:U78969] [PN:FlaR] [GN:flaR] [OR:Streptococcus pyogenes]	86
SPX0037	37	2698	237	711	1042	4.10E-140	[LN:DEOD_STRTR] [AC:Q56037] [GN:DEOD] [OR:Streptococcus thermophilus] [EC:2.4.2.1] [DE:(PNP) (FRAGMENT)] [SP:Q56037]	116
SPX0038	38	2699	171	513	271	1.10E-30	[GI:6729344] [LN:AB029317] [AC:AB029317] [PN:intermedilysin] [GN:ily] [OR:Streptococcus intermedius] [SR:Streptococcus intermedius (strain:UNS46) DNA]	150
SPX0039	39	2700	192	576	117	2.80E-07	“[LN:PXN1_XENLA] [AC:P49263]	137

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:PXN1] [OR:Xenopus laevis] [SR:,African clawed frog] [DE:PENTRAXIN FUSION PROTEIN PRECURSOR] [SP:P49263]”	
SPX0040	40	2701	136	408	92	0.00026	“[LN:CP23_CHICK] [AC:P23614] [OR:Gallus gallus] [SR:,Chicken] [DE:23 KD CORTICAL CYTOSKELETON-ASSOCIATED PROTEIN (CAP-23)] [SP:P23614]”	135
SPX0041	41	2702	270	810	395	1.40E-105	[LN:DEOD_BACSU] [AC:P46354] [GN:DEOD:PNP] [OR:Bacillus subtilis] [EC:2.4.2.1] [DE:(PNP)] [SP:P46354]	100
SPX0042	42	2703	181	543			NO-HIT	6
SPX0043	43	2704	404	1212	736	2.30E-199	“[LN:DEOB_LACLC] [AC:O32808] [GN:DEOB] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [EC:5.4.2.7] [DE:PHOSPHOPENTOMUTASE, (PHOSPHODEOXYRIBOMUTASE)] [SP:O32808]”	181
SPX0044	44	2705	247	741	206	1.40E-52	[LN:G69180] [AC:G69180] [PN:ribose 5-phosphate isomerase] [GN:MTH608] [CL:Haemophilus influenzae ribose-5-phosphate isomerase] [OR:Methanobacterium thermoautotrophicum]	168
SPX0045	45	2706	62	186			NO-HIT	6
SPX0046	46	2707	291	873	386	5.80E-99	[LN:YFOL_STRTR] [AC:P96051] [OR:Streptococcus thermophilus] [DE:(ORF1091)] [SP:P96051]	86
SPX0047	47	2708	307	921	1049	3.00E-140	[LN:FOLD_STRTR] [AC:P96050] [GN:FOLD] [OR:Streptococcus thermophilus] [EC:1.5.1.5:3.5.4.9] [DE:(EC 3.5.4.9)]] [SP:P96050]	121
SPX0048	48	2709	72	216	99	5.40E-08	[LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Aeropyrum pernix]	
SPX0049	49	2710	245	735	796	7.00E-105	[LN:H69334] [AC:H69334] [PN:glutamine transport protein glnQ] [GN:glnQ] [CL:inner membrane protein malK:ATP-binding cassette homology] [OR:Archaeoglobus fulgidus]	162
SPX0050	50	2711	128	384	88	5.60E-05	[LN:D72757] [AC:D72757] [PN:hypothetical protein APE0049] [GN:APE0049] [OR:Aeropyrum pernix]	92
SPX0051	51	2712	228	684	314	1.10E-52	"[LN:H69278] [AC:H69278] [PN:glutamine ABC transporter, permease protein (glnP) homolog] [CL:histidine permease protein M] [OR:Archaeoglobus fulgidus]"	151
SPX0052	52	2713	77	231	89	8.60E-09	[LN:G69865] [AC:G69865] [PN:hypothetical protein ykuJ] [GN:ykuJ] [OR:Bacillus subtilis]	87
SPX0053	53	2714	753	2259	2641	0	[GI:4103470] [LN:AF023421] [AC:AF023421] [PN:ClpE] [GN:clpE] [OR:Lactococcus lactis]	84
SPX0054	54	2715	83	249	116	9.50E-10	[GI:4098132] [LN:MBU73653] [AC:U73653] [PN:63 kDa protein] [OR:Mycobacterium bovis]	83
SPX0055	55	2716	61	183			NO-HIT	6
SPX0056	56	2717	70	210			NO-HIT	6
SPX0057	57	2718	114	342	477	2.00E-73	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX0058	58	2719	116	348	569	6.90E-75	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96
SPX0059	59	2720	141	423	101	5.60E-07	[LN:A70315] [AC:A70315] [PN:AP4A hydrolase] [GN:apfA] [CL:Methanococcus jannaschii mutator protein mutT:mutT domain homology]	147

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Aquifex aeolicus]	
SPX0060	60	2721	100	300			NO-HIT	6
SPX0061	61	2722	198	594	597	3.80E-95	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0062	62	2723	88	264	308	1.70E-38	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0063	63	2724	34	102	126	2.60E-12	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0064	64	2725	117	351	210	5.10E-24	"[LN:JC1151] [AC:JC1151] [PN:hypothetical protein, 20.3K] [OR:Agrobacterium tumefaciens]"	89
SPX0065	65	2726	102	306	154	5.60E-19	[GI:722339] [LN:AXU22323] [AC:U22323] [PN:unknown] [OR:Acetobacter xylinus] [SR:Acetobacter xylinum]	100
SPX0066	66	2727	89	267			NO-HIT	6
SPX0067	67	2728	576	1728	285	5.20E-53	[LN:G70002] [AC:G70002] [PN:hypothetical protein ytwP] [GN:ytwP] [OR:Bacillus subtilis]	87
SPX0068	68	2729	649	1947	3322	0	[GI:1490397] [LN:SPGYRBORF] [AC:Z67740] [PN:DNA gyrase] [GN:gyrB] [OR:Streptococcus pneumoniae]	95
SPX0069	69	2730	191	573	709	2.40E-93	[GI:1052803] [LN:SPGYRBG] [AC:X83917] [GN:orflgyrb] [OR:Streptococcus pneumoniae]	81
SPX0070	70	2731	185	555	118	9.00E-28	"[LN:D70177] [AC:D70177] [PN:4-methyl-5(b-hydroxyethyl)-thiazole monophosphate biosynthesis protein (thiJ) homolog] [CL:signal transduction protein DJ-1] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	211
SPX0071	71	2732	417	1251	398	1.00E-50	[LN:S22738]	86

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:S22738:S22728] [PN:hypothetical protein] [OR:Streptococcus salivarius]	
SPX0072	72	2733	817	2451	203	3.80E-66	[LN:DING_BACSU] [AC:P54394] [GN:DING] [OR:Bacillus subtilis] [DE:PROBABLE ATP-DEPENDENT HELICASE DING HOMOLOG] [SP:P54394]	122
SPX0073	73	2734	310	930	189	6.20E-42	[LN:G69979] [AC:G69979] [PN:proteinase homolog yrrN] [GN:yrrN] [OR:Bacillus subtilis]	85
SPX0074	74	2735	96	288			NO-HIT	6
SPX0075	75	2736	445	1335	2202	4.50E-301	“[LN:CIAH_STRPN] [AC:Q54955] [GN:CIAH] [OR:Streptococcus pneumoniae] [EC:2.7.3.-] [DE:SENSOR PROTEIN CIAH,] [SP:Q54955]”	120
SPX0076	76	2737	225	675	1130	1.70E-152	[LN:CIAR_STRPN] [AC:Q54954] [GN:CIAR] [OR:Streptococcus pneumoniae] [DE:TRANSCRIPTIONAL REGULATORY PROTEIN CIAR] [SP:Q54954]	124
SPX0077	77	2738	849	2547	1653	0	[GI:5524752] [LN:STH007700] [AC:AJ007700] [PN:aminopeptidase N] [GN:pepN] [OR:Streptococcus thermophilus]	105
SPX0078	78	2739	147	441	176	1.80E-20	[LN:T35570] [AC:T35570] [PN:hypothetical protein SC6G4.19c SC6G4.19c] [GN:SC6G4.19c] [OR:Streptomyces coelicolor]	113
SPX0079	79	2740	292	876	80	8.00E-10	[LN:PPSA_PYRFU] [AC:P42850:Q59672] [GN:PPSA] [OR:Pyrococcus furiosus] [EC:2.7.9.2] [DE:DIKINASE) (PEP SYNTHASE)] [SP:P42850:Q59672]	131
SPX0080	80	2741	243	729			NO-HIT	6
SPX0081	81	2742	160	480	113	1.60E-07	[LN:C72513] [AC:C72513] [PN:hypothetical protein APE2080]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:APE2080] [OR:Aeropyrum pernix]	
SPX0082	82	2743	76	228			NO-HIT	6
SPX0083	83	2744	233	699	231	3.20E-28	[LN:GNO_GLUOX] [AC:P50199] [GN:GNO] [OR:Gluconobacter oxydans] [EC:1.1.1.69] [DE:REDUCTASE)] [SP:P50199]	104
SPX0084	84	2745	89	267			NO-HIT	6
SPX0085	85	2746	309	927	210	2.40E-41	[LN:YDHF_ECOLI] [AC:P76187] [GN:YDHF] [OR:Escherichia coli] [DE:HYPOTHETICAL OXIDOREDUCTASE IN SODC-NEMA INTERGENIC REGION] [SP:P76187]	135
SPX0086	86	2747	272	816	114	3.70E-09	[GI:3043880] [LN:LLU95841] [AC:U95841] [PN:transmembrane protein Tmp6] [OR:Lactococcus lactis]	94
SPX0087	87	2748	72	216	146	1.50E-15	[GI:3582220] [LN:AE001272] [AC:AE001272] [PN:conserved hypothetical protein] [GN:ORF00047] [OR:Lactococcus lactis]	114
SPX0088	88	2749	79	237			NO-HIT	6
SPX0089	89	2750	82	246	269	2.80E-32	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0090	90	2751	680	2040	734	6.10E-239	[LN:SYM_BACST] [AC:P23920] [GN:METS] [OR:Bacillus stearothermophilus] [EC:6.1.1.10] [DE:(METRS)] [SP:P23920]	108
SPX0091	91	2752	420	1260	287	2.30E-62	[LN:E69858] [AC:E69858] [PN:conserved hypothetical protein yknZ] [GN:yknZ] [OR:Bacillus subtilis]	97
SPX0092	92	2753	234	702	663	9.00E-86	[GI:2822199] [LN:SCU96166] [AC:U96166:AF227987] [PN:ATP-binding cassette protein] [GN:tptC]	120

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus cristatus]	
SPX0093	93	2754	400	1200	144	5.60E-36	[LN:C69858] [AC:C69858] [PN:conserved hypothetical protein yknX] [GN:yknX] [OR:Bacillus subtilis]	97
SPX0094	94	2755	449	1347	730	1.70E-192	“[LN:S41386] [AC:S41386] [PN:glutathione reductase (NADPH),] [CL:dihydrolipoamide dehydrogenase:dihydrolipoamide dehydrogenase homology] [OR:Streptococcus thermophilus] [EC:1.6.4.2]”	182
SPX0095	95	2756	179	537	218	6.90E-30	[LN:H72334] [AC:H72334] [PN:bioY protein] [GN:TM0799] [OR:Thermotoga maritima]	78
SPX0096	96	2757	112	336			NO-HIT	6
SPX0097	97	2758	80	240	171	1.50E-18	[LN:Y352__TREPA] [AC:O83371] [GN:TP0352] [OR:Treponema pallidum] [DE:HYPOTHETICAL PROTEIN TP0352] [SP:O83371]	108
SPX0098	98	2759	107	321			NO-HIT	6
SPX0099	99	2760	113	339	109	1.40E-12	[GI:4098081] [LN:LLU73336] [AC:U73336] [PN:anaerobic ribonucleotide reductase] [GN:nrdD] [OR:Lactococcus lactis subsp. cremoris]	128
SPX0100	100	2761	240	720	372	1.30E-85	[LN:TRMD__BACSU] [AC:O31741] [GN:TRMD] [OR:Bacillus subtilis] [EC:2.1.1.31] [DE:METHYLTRANSFERASE) (TRNA [GM37] METHYLTRANSFERASE)] [SP:O31741]	142
SPX0101	101	2762	173	519	292	5.20E-61	[LN:RIMM__BACSU] [AC:O31740] [GN:RIMM] [OR:Bacillus subtilis] [DE:PROBABLE 16S RRNA PROCESSING PROTEIN RIMM] [SP:O31740]	119
SPX0102	102	2763	267	801			NO-HIT	6
SPX0103	103	2764	80	240	141	9.70E-15	[LN:C69880] [AC:C69880] [PN:conserved hypothetical protein ylqC] [GN:ylqC]	147

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[CL:Bacillus conserved hypothetical protein ylqC] [OR:Bacillus subtilis]	
SPX0104	104	2765	91	273	315	5.10E-39	[LN:C47154] [AC:C47154:S11366:G69700] [PN:ribosomal protein S16 (BS17) rpsP:ribosomal protein BS17] [GN:rpsP] [CL:Escherichia coli ribosomal protein S16] [OR:Bacillus subtilis]	176
SPX0105	105	2766	70	210			NO-HIT	6
SPX0106	106	2767	135	405			NO-HIT	6
SPX0107	107	2768	71	213			NO-HIT	6
SPX0108	108	2769	268	804	243	3.60E-38	[LN:T41399] [AC:T41399] [PN:probable cyclophilin-related peptidyl prolyl cis-trans isomerase] [GN:SPCC553.04] [OR:Schizosaccharomyces pombe]	140
SPX0109	109	2770	120	360	83	5.110E-05	[GI:1773206] [LN:ECU82664] [AC:U82664] [OR:Escherichia coli]	60
SPX0110	110	2771	514	1542	1502	3.70E-258	[LN:D69813] [AC:D69813] [PN:ABC transporter (ATP-binding protein) homolog yfmM] [GN:yfmM] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0111	111	2772	161	483	88	2.20E-11	[GI:7110140] [LN:AF155139] [AC:AF155139:AF019976] [PN:VanZF] [GN:vanZF] [OR:Paenibacillus popilliae]	100
SPX0112	112	2773	362	1086	452	1.30E-135	[LN:YLON_BACSU] [AC:O34617] [GN:YLON] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 41.6 KD PROTEIN IN FMT-SPOVM INTERGENIC REGION] [SP:O34617]	137
SPX0113	113	2774	177	531	152	3.20E-25	[LN:F70023] [AC:F70023] [PN:hypothetical protein yutD] [GN:yutD] [CL:Bacillus subtilis hypothetical protein yutD] [OR:Bacillus subtilis]	136
SPX0114	114	2775	202	606	1034	4.90E-138	[GI:5758312] [LN:AF162664] [AC:AF162664] [PN:manganese co-factored superoxide dismutase] [GN:sodA] [FN:removes superoxide]	152

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX0115	115	2776	346	1038	952	2.60E-127	[GI:2765186] [LN:SASODA] [AC:Y12224] [PN:hypothetical protein] [OR:Streptococcus agalactiae]	92
SPX0116	116	2777	312	936	1616	3.20E-220	[GI:5578893] [LN:SPN131985] [AC:AJ131985] [PN:dihydroorotate dehydrogenase] [GN:pyrDA] [OR:Streptococcus pneumoniae]	116
SPX0117	117	2778	397	1191	1399	3.90E-187	[LN:METK__STAAU] [AC:P50307] [GN:METK] [OR:Staphylococcus aureus] [EC:2.5.1.6] [DE:ADENOSYLTRANSFERASE) (ADOMET SYNTHETASE)] [SP:P50307]	135
SPX0118	118	2779	448	1344	972	1.00E-140	[LN:YQFR__BACSU] [AC:P54475] [GN:YQFR] [OR:Bacillus subtilis] [DE:PROBABLE RNA HELICASE IN CCCA-SODA INTERGENIC REGION] [SP:P54475]	130
SPX0119	119	2780	272	816			NO-HIT	6
SPX0120	120	2781	88	264			NO-HIT	6
SPX0121	121	2782	727	2181	735	4.60E-182	[GI:4098489] [LN:SMU78600] [AC:U78600] [PN:putative ptsG protein] [OR:Streptococcus mutans]	91
SPX0122	122	2783	312	936	330	2.80E-63	[LN:G69627] [AC:G69627] [PN:cell-division protein ftsX] [GN:ftsX] [OR:Bacillus subtilis]	88
SPX0123	123	2784	69	207			NO-HIT	6
SPX0124	124	2785	231	693	798	9.90E-106	[LN:D69627] [AC:D69627] [PN:cell-division ATP-binding protein ftsE] [GN:ftsE] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	176
SPX0125	125	2786	75	225	94	1.10E-07	[LN:A71007] [AC:A71007] [PN:hypothetical protein PH1351] [GN:PH1351] [OR:Pyrococcus horikoshii]	95
SPX0126	126	2787	307	921	936	6.80E-124	[LN:RF2__BACSU]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P28367:O34444] [GN:PRFB] [OR:Bacillus subtilis] [DE:PEPTIDE CHAIN RELEASE FACTOR 2 (RF-2)] [SP:P28367:O34444]	
SPX0127	127	2788	219	657	279	3.40E-41	[LN:H72290] [AC:H72290] [PN:conserved hypothetical protein] [GN:TM1140] [OR:Thermotoga maritima]	96
SPX0128	128	2789	237	711	711	7.60E-93	[LN:T35757] [AC:T35757] [PN:probable branched chain amino acid transport ATP-binding protein] [GN:SC7H2.26] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Streptomyces coelicolor]	212
SPX0129	129	2790	67	201	109	6.60E-10	[LN:G72485] [AC:G72485] [PN:hypothetical protein APE2527] [GN:APE2527] [OR:Aeropyrum pernix]	92
SPX0130	130	2791	255	765	387	4.20E-90	“[LN:F72290] [AC:F72290] [PN:branched chain amino acid ABC transporter, ATP-binding protein] [GN:TM1138] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Thermotoga maritima]”	206
SPX0131	131	2792	319	957	334	1.10E-63	[LN:E81303] [AC:E81303] [PN:probable branched-chain amino-acid ABC transport system permease protein Cj1016c [imported]] [GN:livM:Cj1016c] [OR:Campylobacter jejuni]	164
SPX0132	132	2793	293	879	325	1.30E-90	“[LN:D72290] [AC:D72290] [PN:branched chain amino acid ABC transporter, permease protein] [GN:TM1136] [CL:leucine transport protein livH] [OR:Thermotoga maritima]”	163
SPX0133	133	2794	387	1161	193	5.70E-39	[LN:H81303] [AC:H81303] [PN:branched-chain amino-acid ABC transport system periplasmic binding protein Cj1019c [imported]] [GN:livJ:Cj1019c] [OR:Campylobacter jejuni]	166
SPX0134	134	2795	150	450			NO-HIT	6
SPX0135	135	2796	83	249	116	2.40E-11	[LN:D69874] [AC:D69874] [PN:conserved hypothetical protein ylbG] [GN:ylbG] [OR:Bacillus subtilis]	97
SPX0136	136	2797	197	591	933	8.10E-124	[LN:CLPP_STRSL]	119

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P36398] [GN:CLPP] [OR:Streptococcus salivarius] [EC:3.4.21.92] [DE:(ENDOPEPTIDASE CLP)] [SP:P36398]	
SPX0137	137	2798	217	651	961	5.20E-129	[LN:UPP__STRSL] [AC:P36399] [GN:UPP] [OR:Streptococcus salivarius] [EC:2.4.2.9] [DE:PYROPHOSPHORYLASE) (UPRTASE)] [SP:P36399]	124
SPX0138	138	2799	54	162			O-HIT	6
SPX0139	139	2800	156	468	265	4.40E-49	[LN:CME2__BACSU] [AC:P32393] [GN:COMEB:COME2] [OR:Bacillus subtilis] [DE:COME OPERON PROTEIN 2] [SP:P32393]	106
SPX0140	140	2801	186	558	114	1.10E-07	[LN:G70325] [AC:G70325] [PN:transcription regulator TetR/AcrR family] [GN:acrR3] [OR:Aquifex aeolicus]	102
SPX0141	141	2802	282	846	813	2.60E-107	[GI:4580621] [LN:AF118389] [AC:AF118389] [PN:unknown] [OR:Streptococcus suis]	77
SPX0142	142	2803	392	1176	330	1.60E-83	[GI:2293312] [LN:AF008220] [AC:AF008220] [PN:YtfP] [GN:ytfP] [OR:Bacillus subtilis]	83
SPX0143	143	2804	152	456	106	6.50E-12	[LN:G70031] [AC:G70031] [PN:mutator MutT protein homolog yvcI] [GN:yvcI] [CL:mutT domain homology] [OR:Bacillus subtilis]	121
SPX0144	144	2805	173	519	876	2.10E-115	[GI:6179679] [LN:SPN239034] [AC:AJ239034] [PN:regulator of pmrA] [GN:mta] [FN:regulator of pmrA expression] [OR:Streptococcus pneumoniae]	137
SPX0145	145	2806	490	1470	805	5.40E-151	[LN:F69825]	144

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:F69825] [PN:sodium-dependent transporter homolog yhdH] [GN:yhdH] [CL:gamma-aminobutyric acid transporter] [OR:Bacillus subtilis]	
SPX0146	146	2807	315	945	1165	3.00E-155	[LN:MANA__STRMU] [AC:Q59935] [GN:PMI] [OR:Streptococcus mutans] [EC:5.3.1.8] [DE:(PMI) (PHOSPHOHEXOMUTASE)] [SP:Q59935]	118
SPX0147	147	2808	134	402			NO-HIT	6
SPX0148	148	2809	110	330	270	1.80E-43	[LN:JE0396] [AC:JE0396] [PN:phospho-beta-galactosidase II] [CL:Agrobacterium beta-glucosidase] [OR:Lactobacillus gasseri]	121
SPX0149	149	2810	89	267			NO-HIT	6
SPX0150	150	2811	218	654	154	1.00E-20	[GI:722339] [LN:AXU22323] [AC:U22323] [PN:unknown] [OR:Acetobacter xylinus] [SR:Acetobacter xylinum]	100
SPX0151	151	2812	116	348			NO-HIT	6
SPX0152	152	2813	592	1776	3033	0	“[LN:POXB__STRPN] [AC:Q54970] [GN:SPXB] [OR:Streptococcus pneumoniae] [EC:1.2.3.3] [DE:PYRUVATE OXIDASE, (PYRUVIC OXIDASE) (POX)] [SP:Q54970]”	141
SPX0153	153	2814	67	201	149	2.30E-14	[GI:1513069] [LN:LMU15554] [AC:U15554] [PN:P-type adenosine triphosphatase] [GN:ctpA] [FN:involved in cation transport] [OR:Listeria monocytogenes]	147
SPX0154	154	2815	733	2199	648	8.80E-202	[GI:1513069] [LN:LMU15554] [AC:U15554] [PN:P-type adenosine triphosphatase] [GN:ctpA] [FN:involved in cation transport] [OR:Listeria monocytogenes]	147
SPX0155	155	2816	85	255	107	2.40E-09	[LN:H72624] [AC:H72624] [PN:hypothetical protein APE1456]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:APE1456] [OR:Aeropyrum pernix]	
SPX0156	156	2817	61	183			NO-HIT	6
SPX0157	157	2818	124	372	70	6.10E-06	“[LN:H69267] [AC:H69267] [PN:cytochrome-c oxidase, chain II AF0144] [CL:cytochrome-c oxidase chain II:cytochrome-c oxidase chain II homology] [OR:Archaeoglobus fulgidus] [EC:1.9.3.1]”	183
SPX0158	158	2819	142	426	240	1.00E-28	[LN:COPY_ENTHR] [AC:Q47839] [GN:COPY] [OR:Enterococcus hirae] [DE:COPAB ATPASES METAL-FIST TYPE REPRESSOR] [SP:Q47839]	118
SPX0159	159	2820	269	807	308	4.00E-99	[GI:6707002] [LN:AF109218] [AC:AF109218:U96108] [PN:ThiD] [GN:thiD] [OR:Staphylococcus carnosus]	96
SPX0160	160	2821	211	633	245	8.10E-55	[LN:D75087] [AC:D75087] [PN:thiamin phosphate pyrophosphorylase (thie) PAB1645] [GN:PAB1645] [CL:probable thiamin-phosphate pyrophosphorylase:thiamin-phosphate pyrophosphorylase homology] [OR:Pyrococcus abyssi]	210
SPX0161	161	2822	269	807	228	2.70E-27	[LN:THIM_ECOLI] [AC:P76423] [GN:THIM] [OR:Escherichia coli] [EC:2.7.1.50] [DE:HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE)] [SP:P76423]	144
SPX0162	162	2823	175	525			NO-HIT	6
SPX0163	163	2824	231	693	384	2.00E-48	[LN:TENA_BACSU] [AC:P25052] [GN:TENA] [OR:Bacillus subtilis] [DE:TRANSCRIPTIONAL ACTIVATOR TENA] [SP:P25052]	108
SPX0164	164	2825	217	651	112	1.10E-13	[GI:1296823] [LN:LHPEPIGN] [AC:Z56283] [GN:orf2] [OR:Lactobacillus helveticus]	78
SPX0165	165	2826	462	1386	279	4.50E-59	[GI:1296822] [LN:LHPEPIGN] [AC:Z56283] [GN:orf1]	78

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Lactobacillus helveticus]	
SPX0166	166	2827	187	561	203	2.20E-27	[LN:A69859] [AC:A69859] [PN:hypothetical protein ykoE] [GN:ykoE] [OR:Bacillus subtilis]	87
SPX0167	167	2828	75	225			NO-HIT	6
SPX0168	168	2829	69	207			NO-HIT	6
SPX0169	169	2830	210	630	520	4.10E-72	[GI:6707004] [LN:AF109218] [AC:AF109218:U96108] [PN:ThiE] [GN:thiE] [OR:Staphylococcus carnosus]	96
SPX0170	170	2831	268	804	523	4.60E-67	[GI:6707003] [LN:AF109218] [AC:AF109218:U96108] [PN:ThiM] [GN:thiM] [OR:Staphylococcus carnosus]	96
SPX0171	171	2832	137	411	114	2.00E-08	[LN:H71057] [AC:H71057] [PN:hypothetical protein PH1158] [GN:PH1158] [OR:Pyrococcus horikoshii]	95
SPX0172	172	2833	82	246			NO-HIT	6
SPX0173	173	2834	75	225	203	1.90E-23	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0174	174	2835	223	669	127	6.60E-15	“[LN:PT18_YEAST] [AC:P25362] [GN:PET18:HIT2:YCR020C:YCR20C] [OR:Saccharomyces cerevisiae] [SR:,Baker's yeast] [DE:PET18 PROTEIN] [SP:P25362]”	141
SPX0175	175	2836	379	1137	1379	5.20E-216	[GI:2239174] [LN:SILCT] [AC:Y07622] [PN:lactate oxidase] [GN:lctO] [FN:lactate utilisation] [OR:Streptococcus iniae]	116
SPX0176	176	2837	497	1491	1157	1.60E-228	“[LN:SYK_BACSU] [AC:P37477] [GN:LYSS] [OR:Bacillus subtilis] [EC:6.1.1.6] [DE:LYSYL-TRNA SYNTHETASE, (LYSINE--TRNA LIGASE) (LYSRS)]	144

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P37477]”	
SPX0177	177	2838	227	681	351	2.10E-70	“[LN:D71849] [AC:D71849] [PN:amino acid ABC transporter, permease protein] [GN:jhp1096] [CL:histidine permease protein M] [OR:Helicobacter pylori] [SR:strain J99, , strain J99] [SR:strain J99,]”	195
SPX0178	178	2839	127	381	279	2.10E-35	“[LN:B64666] [AC:B64666] [PN:glutamine ABC transporter, permease protein] [CL:histidine permease protein M] [OR:Helicobacter pylori]”	133
SPX0179	179	2840	96	288	249	4.10E-30	“[LN:B64666] [AC:B64666] [PN:glutamine ABC transporter, permease protein] [CL:histidine permease protein M] [OR:Helicobacter pylori]”	133
SPX0180	180	2841	71	213	92	5.90E-05	[GI:4163988] [LN:AF082511] [AC:AF082511] [PN:putative ATP-binding protein MglA] [GN:mglA] [OR:Treponema denticola]	114
SPX0181	181	2842	169	507	531	5.00E-70	[LN:H81391] [AC:H81391] [PN:amino-acid ABC transporter ATP-binding protein Cj0469 [imported]] [GN:Cj0469] [OR:Campylobacter jejuni]	131
SPX0182	182	2843	67	201	72	3.30E-07	[LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061] [OR:Aeropyrum pernix]	92
SPX0183	183	2844	81	243	132	5.50E-12	[GI:4204972] [LN:LLU60994] [AC:U60994] [FN:transport systems for basic amino acids and] [OR:Leuconostoc lactis]	111
SPX0184	184	2845	185	555	531	1.40E-67	[LN:C81373] [AC:C81373] [PN:probable amino-acid transporter periplasmic solute-binding protein Cj0982c [imported]] [GN:Cj0982c] [OR:Campylobacter jejuni]	153
SPX0185	185	2846	216	648	353	6.60E-51	[LN:YYBJ_BACSU] [AC:P37494] [GN:YYBJ] [OR:Bacillus subtilis] [DE:INTERGENIC REGION]	95

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P37494]	
SPX0186	186	2847	235	705			NO-HIT	6
SPX0187	187	2848	68	204			NO-HIT	6
SPX0188	188	2849	269	807			NO-HIT	6
SPX0189	189	2850	255	765			NO-HIT	6
SPX0190	190	2851	117	351			NO-HIT	6
SPX0191	191	2852	211	633	1057	1.30E-140	[GI:4009484] [LN:AF068902] [AC:AF068902] [PN:orotate phosphoribosyltransferase PyrE] [GN:pyrE] [OR:Streptococcus pneumoniae]	124
SPX0192	192	2853	234	702	1173	1.80E-157	[GI:4009483] [LN:AF068902] [AC:AF068902] [PN:orotidine-5'-decarboxylase PyrF] [GN:pyrF] [OR:Streptococcus pneumoniae]	117
SPX0193	193	2854	67	201			NO-HIT	6
SPX0194	194	2855	400	1200	1946	3.00E-251	[GI:4009482] [LN:AF068902] [AC:AF068902] [PN:cell division protein DivIB] [GN:divIB] [OR:Streptococcus pneumoniae]	114
SPX0195	195	2856	148	444	110	2.20E-07	“[GI:6473880] [LN:AB027890] [AC:AB027890] [PN:Hypothetical protein] [GN:SPBC3D5.14C] [OR:Schizosaccharomyces pombe] [SR:Schizosaccharomyces pombe (strain:968 h90) DNA, clone:TA46]]	180
SPX0196	196	2857	353	1059	1780	1.40E-242	[GI:4009481] [LN:AF068902] [AC:AF068902] [PN:undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc] [GN:murG] [OR:Streptococcus pneumoniae]	131
SPX0197	197	2858	451	1353	2206	1.20E-297	[GI:4009480] [LN:AF068902] [AC:AF068902] [PN:D-glutamic acid adding enzyme MurD] [GN:murD] [OR:Streptococcus pneumoniae]	120
SPX0198	198	2859	214	642	1062	1.20E-142	[GI:4009479] [LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0199	199	2860	674	2022	2263	0	[GI:4009478]	83

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0200	200	2861	135	405	86	1.80E-05	[LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae] [GI:7576923] [LN:AF242367] [AC:AF242367] [PN:lactococcin 972] [GN:lclA] [OR:Lactococcus lactis subsp. lactis]	109
SPX0201	201	2862	73	219			NO-HIT	6
SPX0202	202	2863	85	255	89	2.50E-05	[LN:T20916] [AC:T20916] [PN:hypothetical protein F14F8.3] [GN:F14F8.3] [OR:Caenorhabditis elegans]	98
SPX0203	203	2864	621	1863	1706	0	[LN:TYPA__BACSU] [AC:O07631] [GN:TYPA] [OR:Bacillus subtilis] [DE:GTP-BINDING PROTEIN TYPA/BIPA HOMOLOG] [SP:O07631]	115
SPX0204	204	2865	242	726	154	3.00E-34	[GI:7328274] [LN:SAY14816] [AC:Y14816] [PN:hypothetical protein] [GN:ORF231] [OR:Staphylococcus aureus]	103
SPX0205	205	2866	66	198			NO-HIT	6
SPX0206	206	2867	127	381	174	3.80E-28	[LN:YQHL__BACSU] [AC:P54510] [GN:YQHL] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION] [SP:P54510]	141
SPX0207	207	2868	74	222	158	1.80E-17	[LN:T44786] [AC:T44786] [PN:hypothetical protein 1 [imported]] [OR:Bacillus megaterium]	87
SPX0208	208	2869	303	909	1241	5.30E-168	[LN:T44638] [AC:T44638] [PN:capsular polysaccharide biosynthesis protein cpsY [imported]] [GN:cpsY] [CL:probable transcription regulator lsyR] [OR:Streptococcus agalactiae]	172
SPX0209	209	2870	252	756	476	5.30E-66	[LN:YQJQ__BACSU] [AC:P54554] [GN:YQJQ] [OR:Bacillus subtilis] [EC:1.-.-.-]	103

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:(EC 1.-.-.-)] [SP:P54554]	
SPX0210	210	2871	310	930	368	9.50E-98	[LN:YQJK_BACSU] [AC:P54548] [GN:YQJK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 34.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION] [SP:P54548]	137
SPX0211	211	2872	208	624			NO-HIT	6
SPX0212	212	2873	413	1239	792	1.40E-103	[LN:B69888] [AC:B69888] [PN:GTP-binding protein proteinase modulator homolog ynbA] [GN:ynbA] [CL:GTP-binding protein hflX:translation elongation factor Tu homology] [OR:Bacillus subtilis]	187
SPX0213	213	2874	312	936	317	7.10E-70	[LN:G69657] [AC:G69657] [PN:tRNA isopentenylpyrophosphate transferase miaA] [GN:miaA] [CL:delta(2)-isopentenylpyrophosphate transferase] [OR:Bacillus subtilis]	159
SPX0214	214	2875	63	189			NO-HIT	6
SPX0215	215	2876	66	198	129	6.10E-12	“[LN:TYSY_LACLA] [AC:P19368] [GN:THYA] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.1.1.45] [DE:THYMIDYLATE SYNTHASE, (TS)] [SP:P19368]]	160
SPX0216	216	2877	221	663	318	1.20E-39	[GI:7328278] [LN:SAY14816] [AC:Y14816] [PN:hypothetical protein] [GN:ORF242] [OR:Staphylococcus aureus]	103
SPX0217	217	2878	212	636	691	4.80E-91	[LN:B69997] [AC:B69997] [PN:conserved hypothetical protein ytmQ] [GN:ytmQ] [CL:hypothetical protein HI0340] [OR:Bacillus subtilis]	130
SPX0218	218	2879	89	267			NO-HIT	6
SPX0219	219	2880	103	309	215	6.90E-25	[LN:YLXS_BACSU] [AC:P32726] [GN:YLXS] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 17.6 KD PROTEIN IN NUSA 5'REGION (P15A) (ORF1)] [SP:P32726]	137
SPX0220	220	2881	379	1137	715	6.00E-122	[GI:2634032]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:BSUB0009] [AC:Z99112:AL009126] [GN:nusA] [FN:transcription termination] [OR:Bacillus subtilis]	
SPX0221	221	2882	98	294	220	1.70E-25	[LN:YLXR_BACSU] [AC:P32728] [GN:YLXR] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 10.4 KD PROTEIN IN NUSA-INFB INTERGENIC REGION (ORF3)] [SP:P32728]	144
SPX0222	222	2883	100	300	259	3.80E-31	“[LN:YLXQ_ENTFC] [AC:P55768] [OR:Enterococcus faecium] [SR:,Streptococcus faecium] [DE:PROBABLE RIBOSOMAL PROTEIN IN INFB 5'REGION] [SP:P55768]]	144
SPX0223	223	2884	165	495			NO-HIT	6
SPX0224	224	2885	931	2793	2768	0	[GI:3947714] [LN:SAAJ3164] [AC:AJ003164] [PN:initiation factor IF2] [GN:infB] [FN:translation initiation factor] [OR:Streptococcus agalactiae]	142
SPX0225	225	2886	82	246			NO-HIT	6
SPX0226	226	2887	117	351	478	2.00E-61	[GI:3947715] [LN:SAAJ3164] [AC:AJ003164] [PN:ribosome binding factor A] [GN:rbfA] [FN:ribosome maturation] [OR:Streptococcus agalactiae]	136
SPX0227	227	2888	260	780	217	3.30E-40	[LN:T35040] [AC:T35040] [PN:hypothetical protein SC4G2.05 SC4G2.05] [GN:SC4G2.05] [OR:Streptomyces coelicolor]	110
SPX0228	228	2889	84	252			NO-HIT	6
SPX0229	229	2890	74	222			NO-HIT	6
SPX0230	230	2891	65	195	202	2.70E-24	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0231	231	2892	78	234	102	4.60E-08	[LN:A69271] [AC:A69271] [PN:hypothetical protein AF0169] [CL:Archaeoglobus fulgidus hypothetical protein AF0169] [OR:Archaeoglobus fulgidus]	140
SPX0232	232	2893	445	1335	123	1.30E-41	[LN:B69271]	84

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:B69271] [PN:hypothetical protein AF0170] [OR:Archaeoglobus fulgidus]	
SPX0233	233	2894	85	255			NO-HIT	6
SPX0234	234	2895	132	396			NO-HIT	6
SPX0235	235	2896	69	207			NO-HIT	6
SPX0236	236	2897	150	450			NO-HIT	6
SPX0237	237	2898	186	558	148	2.10E-18	[LN:YP20_BACLI] [AC:P05332] [GN:P20] [OR:Bacillus licheniformis] [DE:HYPOTHETICAL P20 PROTEIN] [SP:P05332]	106
SPX0238	238	2899	321	963			NO-HIT	6
SPX0239	239	2900	884	2652	1594	0	"[LN:SYV_BACSU] [AC:Q05873] [GN:VALS] [OR:Bacillus subtilis] [EC:6.1.1.9] [DE:VALYL-TRNA SYNTHETASE, (VALINE--TRNA LIGASE) (VALRS)] [SP:Q05873]]	144
SPX0240	240	2901	252	756	642	7.80E-85	"[LN:B71947] [AC:B71947] [PN:hypothetical protein jhp0330] [GN:jhp0330] [OR:Helicobacter pylori] [SR:strain J99, , strain J99] [SR:strain J99,]]	145
SPX0241	241	2902	268	804	91	1.80E-05	"[LN:Y4LH_RHISN] [AC:P55548] [GN:Y4LH] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 22.4 KD PROTEIN Y4LH] [SP:P55548]]	128
SPX0242	242	2903	95	285	387	3.10E-49	[GI:1490399] [LN:SPPARCETP] [AC:Z67739] [PN:DNA transposase] [OR:Streptococcus pneumoniae]	90
SPX0243	243	2904	102	306	323	2.70E-40	[GI:1490399] [LN:SPPARCETP] [AC:Z67739] [PN:DNA transposase] [OR:Streptococcus pneumoniae]	90
SPX0244	244	2905	91	273	315	2.50E-39	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0245	245	2906	92	276	135	3.80E-13	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0246	246	2907	422	1266			NO-HIT	6
SPX0247	247	2908	399	1197	109	9.90E-10	[GI:6689196] [LN:SCE20] [AC:AL136058] [PN:putative helicase] [GN:SCE20.37] [OR:Streptomyces coelicolor A3(2)]	109
SPX0248	248	2909	98	294			NO-HIT	6
SPX0249	249	2910	280	840	489	2.30E-81	[LN:LICT__BACSU] [AC:P39805] [GN:LICT:N15A] [OR:Bacillus subtilis] [DE:TRANSCRIPTION ANTITERMINATOR LICT] [SP:P39805]	116
SPX0250	250	2911	613	1839	534	8.10E-147	“[LN:PTBA__BACSU] [AC:P40739:Q45661] [GN:BGLP:N17C] [OR:Bacillus subtilis] [EC:2.7.1.69] [DE:ENZYME II, ABC COMPONENT), (EII-BGL)] [SP:P40739:Q45661]]	149
SPX0251	251	2912	472	1416	1046	3.70E-228	“[LN:ABGA__CLOLO] [AC:Q46130] [GN:ABGA] [OR:Clostridium longisporum] [EC:3.2.1.86] [DE:6-PHOSPHO-BETA-GLUCOSIDASE,] [SP:Q46130]]	127
SPX0252	252	2913	349	1047	656	3.10E-156	[LN:SYFA__BACSU] [AC:P17921:P94539] [GN:PHES] [OR:Bacillus subtilis] [EC:6.1.1.20] [DE:-TRNA LIGASE ALPHA CHAIN) (PHERS)] [SP:P17921:P94539]	139
SPX0253	253	2914	331	993			NO-HIT	6
SPX0254	254	2915	170	510	333	1.80E-41	“[LN:YPIP__LACDL] [AC:P46543] [OR:Lactobacillus delbrueckii] [SR:,subsplactis] [DE:HYPOTHETICAL 19.8 KD PROTEIN IN PEPI 3'REGION] [SP:P46543]]	141
SPX0255	255	2916	801	2403	312	2.20E-132	“[LN:SYFB__SYNY3] [AC:P74296] [GN:PHET:SLL1553] [OR:Synechocystis sp]	154

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SR:,strain PCC 6803] [EC:6.1.1.20] [DE:TRNA LIGASE BETA CHAIN) (PHERS)] [SP:P74296]]	
SPX0256	256	2917	296	888			NO-HIT	6
SPX0257	257	2918	149	447	437	4.90E-55	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX0258	258	2919	89	267	245	2.00E-28	[LN:Y659_HAEIN] [AC:P44030] [GN:HI0659] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI0659] [SP:P44030]	112
SPX0259	259	2920	99	297			NO-HIT	6
SPX0260	260	2921	750	2250	1309	6.60E-297	“[LN:C81326] [AC:C81326] [PN:5-methyltetrahydropteroyltriglutamate-- homocystei methyltransferase, Cj1201 [imported]] [GN:metE:Cj1201] [OR:Campylobacter jejuni] [EC:2.1.1.14]]	175
SPX0261	261	2922	289	867	261	3.90E-77	“[LN:D81326] [AC:D81326] [PN:5,10-methylenetetrahydrofolate reductase, Cj1202 [imported]] [GN:metF:Cj1202] [OR:Campylobacter jejuni] [EC:1.7.99.5]]	147
SPX0262	262	2923	90	270			NO-HIT	6
SPX0263	263	2924	738	2214	1323	3.80E-298	[LN:PNPA_BACSU] [AC:P50849] [GN:PNPA:COMR] [OR:Bacillus subtilis] [EC:2.7.7.8] [DE:PHOSPHORYLASE) (PNPASE) (VEGETATIVE PROTEIN 15) (VEG15)] [SP:P50849]	151
SPX0264	264	2925	206	618	493	1.20E-63	[GI:6899995] [LN:CST130879] [AC:AJ130879] [PN:serine acetyltransferase] [GN:cysE] [OR:Clostridium sticklandii]	110
SPX0265	265	2926	295	885	113	3.80E-07	[LN:E69786] [AC:E69786] [PN:ribosomal-protein-alanine N-acetyltransfer homolog ydiD] [GN:ydiD] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimI] [OR:Bacillus subtilis]	190
SPX0266	266	2927	152	456			NO-HIT	6
SPX0267	267	2928	448	1344	570	9.00E-167	[GI:6899996]	112

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:CST130879] [AC:AJ130879] [PN:cysteinyl tRNA synthetase] [GN:cysRS] [OR:Clostridium sticklandii]	
SPX0268	268	2929	129	387	316	6.30E-39	[LN:C69742] [AC:C69742] [PN:conserved hypothetical protein yazC] [GN:yazC] [OR:Bacillus subtilis]	97
SPX0269	269	2930	295	885	224	3.10E-30	[LN:LRPR_STREQ] [AC:Q54087] [GN:LRP] [OR:Streptococcus equisimilis] [DE:LEUCINE RICH PROTEIN] [SP:Q54087]	105
SPX0270	270	2931	88	264			NO-HIT	6
SPX0271	271	2932	200	600			NO-HIT	6
SPX0272	272	2933	52	156	87	3.30E-06	[LN:D75542] [AC:D75542] [PN:hypothetical protein] [GN:DR0254] [OR:Deinococcus radiodurans]	90
SPX0273	273	2934	82	246	247	1.10E-29	[LN:F81737] [AC:F81737] [PN:hypothetical protein TC0129 [imported]] [GN:TC0129] [OR:Chlamydia muridarum:Chlamydia trachomatis MoPn]	131
SPX0274	274	2935	78	234	143	4.90E-15	[LN:F71245] [AC:F71245] [PN:hypothetical protein PHS004] [GN:PHS004] [OR:Pyrococcus horikoshii]	95
SPX0275	275	2936	72	216			NO-HIT	6
SPX0276	276	2937	63	189	68	2.40E-07	[LN:C64571] [AC:C64571] [PN:hypothetical protein HP0411] [CL:Helicobacter pylori hypothetical protein HP0411] [OR:Helicobacter pylori]	134
SPX0277	277	2938	312	936	194	2.60E-36	“[GI:773349] [LN:BSU20445] [AC:U20445] [PN:BirA protein] [GN:birA] [FN:biotin protein ligase, biotin operon repressor] [OR:Bacillus subtilis]]	142
SPX0278	278	2939	287	861	255	1.70E-56	[LN:MSMR_STRMU] [AC:Q00753] [GN:MSMR] [OR:Streptococcus mutans]	110

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:MSM OPERON REGULATORY PROTEIN] [SP:Q00753]	
SPX0279	279	2940	721	2163	2528	0	“[LN:AGAL__STRMU] [AC:P27756] [GN:AGA] [OR:Streptococcus mutans] [EC:3.2.1.22] [DE:ALPHA-GALACTOSIDASE, (MELIBIASE)] [SP:P27756]]	128
SPX0280	280	2941	420	1260	1106	2.40E-180	[LN:MSME__STRMU] [AC:Q00749] [GN:MSME] [OR:Streptococcus mutans] [DE:MULTIPLE SUGAR-BINDING PROTEIN PRECURSOR] [SP:Q00749]	121
SPX0281	281	2942	82	246			NO-HIT	6
SPX0282	282	2943	289	867	1188	1.40E-166	[LN:MSMF__STRMU] [AC:Q00750] [GN:MSMF] [OR:Streptococcus mutans] [DE:MULTIPLE SUGAR-BINDING TRANSPORT SYSTEM PERMEASE PROTEIN MSMF] [SP:Q00750]	142
SPX0283	283	2944	278	834	1174	3.40E-164	[LN:MSMG__STRMU] [AC:Q00751] [GN:MSMG] [OR:Streptococcus mutans] [DE:MULTIPLE SUGAR-BINDING TRANSPORT SYSTEM PERMEASE PROTEIN MSMG] [SP:Q00751]	142
SPX0284	284	2945	482	1446	2154	6.70E-294	“[LN:A27626] [AC:A27626] [PN:sucrose phosphorylase,] [CL:gtfA protein] [OR:Streptococcus mutans] [EC:2.4.1.7]]	110
SPX0285	285	2946	398	1194			NO-HIT	6
SPX0286	286	2947	134	402			NO-HIT	6
SPX0287	287	2948	662	1986	3341	0	[LN:AMIA__STRPN] [AC:P18791:P18792] [GN:AMIA] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE-BINDING PROTEIN AMIA PRECURSOR] [SP:P18791:P18792]	142
SPX0288	288	2949	499	1497	2538	0	[LN:AMIC__STRPN] [AC:P18793] [GN:AMIC] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE TRANSPORT PERMEASE PROTEIN AMIC] [SP:P18793]	129
SPX0289	289	2950	309	927	1564	1.10E-217	[LN:AMID__STRPN] [AC:P18794]	129

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:AMID] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE TRANSPORT PERMEASE PROTEIN AMID] [SP:P18794]	
SPX0290	290	2951	356	1068	1806	2.10E-245	[LN:AMIE__STRPN] [AC:P18765] [GN:AMIE] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE TRANSPORT ATP-BINDING PROTEIN AMIE] [SP:P18765]	132
SPX0291	291	2952	113	339			NO-HIT	6
SPX0292	292	2953	309	927	1534	5.40E-208	[LN:AMIF__STRPN] [AC:P18766] [GN:AMIF] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE TRANSPORT ATP-BINDING PROTEIN AMIF] [SP:P18766]	132
SPX0293	293	2954	419	1257	2039	1.30E-284	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX0294	294	2955	96	288	274	1.10E-34	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX0295	295	2956	173	519	840	6.50E-114	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX0296	296	2957	237	711	471	2.40E-60	[LN:TRER__BACSU] [AC:P39796] [GN:TRER] [OR:Bacillus subtilis] [DE:TREHALOSE OPERON TRANSCRIPTIONAL REPRESSOR] [SP:P39796]	120
SPX0297	297	2958	656	1968	383	7.20E-135	[LN:PTTB__ECOLI] [AC:P36672] [GN:TREB] [OR:Escherichia coli] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EII-TRE)] [SP:P36672]	114
SPX0298	298	2959	542	1626	1617	2.00E-272	[GI:2208998] [LN:SSU35633] [AC:U35633]	102

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0299	299	2960	60	180	125	2.30E-12	[PN:dextran glucosidase DexS] [GN:dexS] [OR:Streptococcus suis] [LN:T30285] [AC:T30285]	79
SPX0300	300	2961	84	252	104	4.40E-09	[PN:hypothetical protein] [OR:Streptococcus pneumoniae] [GI:6899236] [LN:AE002123] [AC:AE002123:AF222894]	116
SPX0301	301	2962	265	795	476	2.60E-84	[PN:conserved hypothetical] [GN:UU265] [OR:Ureaplasma urealyticum] [GI:2462097] [LN:BCBCTLGLR] [AC:Y10927]	95
SPX0302	302	2963	297	891	273	9.10E-35	[PN:glutamate racemase] [GN:bcglr] [OR:Bacillus cereus] [LN:C69986] [AC:C69986]	165
SPX0303	303	2964	74	222	158	1.60E-15	[PN:conserved hypothetical protein ysnA] [GN:ysnA] [CL:Methanococcus jannaschii conserved hypothetical protein MJ0226] [OR:Bacillus subtilis] [LN:YGGV_ECOLI] [AC:P52061] [GN:YGGV] [OR:Escherichia coli]	143
SPX0304	304	2965	174	522	229	8.10E-27	[DE:HYPOTHETICAL 21.0 KD PROTEIN IN GSHB-ANSB INTERGENIC REGION (O197)] [SP:P52061] [LN:YSNB_BACSU] [AC:P94559] [GN:YSNB] [OR:Bacillus subtilis]	136
SPX0305	305	2966	154	462	75	8.30E-08	[DE:HYPOTHETICAL 19.2 KD PROTEIN IN RPH-ILVB INTERGENIC REGION] [SP:P94559] [LN:A69866] [AC:A69866]	136
SPX0306	306	2967	247	741	61	0.0006	[PN:hypothetical protein ykuL] [GN:ykuL] [CL:Bacillus subtilis hypothetical protein ykuL] [OR:Bacillus subtilis] [LN:A75153] [AC:A75153]	164
SPX0307	307	2968	243	729	233	1.60E-40	[PN:integrase/recombinase xerd PAB0255] [GN:xerD-like:PAB0255] [CL:probable site-specific integrase/recombinase XerC] [OR:Pyrococcus abyssi] [LN:YPUG_BACSU]	145

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P35154] [GN:YPUG] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 29.6 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX7)] [SP:P35154]	
SPX0308	308	2969	190	570	216	2.00E-28	[LN:Y214_MYCPN] [AC:P75477] [OR:Mycoplasma pneumoniae] [DE:HYPOTHETICAL PROTEIN MG214 HOMOLOG] [SP:P75477]	106
SPX0309	309	2970	241	723	660	5.60E-86	[LN:RLUB_BACSU] [AC:P35159] [GN:RLUB] [OR:Bacillus subtilis] [EC:4.2.1.70] [DE:(PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE)] [SP:P35159]	138
SPX0310	310	2971	81	243	166	1.10E-32	[LN:G72251] [AC:G72251] [PN:conserved hypothetical protein] [GN:TM1462] [CL:conserved hypothetical protein HI1000] [OR:Thermotoga maritima]	139
SPX0311	311	2972	175	525			NO-HIT	6
SPX0312	312	2973	322	966	218	2.70E-57	[LN:E69763] [AC:E69763] [PN:probable ferrichrome ABC transporter yclQ] [GN:yclQ] [CL:iron(III) dicitrate transport protein] [OR:Bacillus subtilis]	146
SPX0313	313	2974	251	753	702	2.90E-92	[LN:D69763] [AC:D69763] [PN:ferrichrome ABC transporter (ATP-binding p) homolog yclP] [GN:yclP] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	194
SPX0314	314	2975	67	201	89	6.60E-07	[LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061] [OR:Aeropyrum pernix]	92
SPX0315	315	2976	319	957	474	8.60E-64	[LN:C69763] [AC:C69763] [PN:ferrichrome ABC transporter (permease) homolog yclO] [GN:yclO] [CL:ferrichrome ABC transporter] [OR:Bacillus subtilis]	146
SPX0316	316	2977	320	960	625	2.80E-84	[LN:B69763] [AC:B69763] [PN:ferrichrome ABC transporter (permease) homolog yclN]	153

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0317	317	2978	136	408	288	3.50E-36	[GN:yclN] [CL:vitamin B12 transport protein btuC] [OR:Bacillus subtilis] [GI:6707010] [LN:AF109295] [AC:AF109295:AF022796]	98
SPX0318	318	2979	81	243	164	6.20E-18	[PN:GalE] [GN:galE] [OR:Staphylococcus carnosus] [GI:3703059] [LN:AF082009] [AC:AF082009]	105
SPX0319	319	2980	354	1062	388	8.30E-143	[PN:UDP-galactose-4-epimerase] [GN:galE] [OR:Lactococcus lactis] “[LN:PEPA_LACLC] [AC:Q48677] [GN:PEPA] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [EC:3.4.11.7] [DE:GLUTAMYL-AMINOPEPTIDASE,] [SP:Q48677]]	162
SPX0320	320	2981	99	297			NO-HIT	6
SPX0321	321	2982	63	189			NO-HIT	6
SPX0322	322	2983	148	444	87	0.00059	[GI:2791904] [LN:SSK3MECA1] [AC:Y13052] [GN:ORF145] [OR:Staphylococcus sciuri]	78
SPX0323	323	2984	255	765			NO-HIT	6
SPX0324	324	2985	243	729	1211	8.40E-163	[GI:5579394] [LN:AF162655] [AC:AF162655] [PN:choline transporter] [GN:proV] [FN:ATPase] [OR:Streptococcus pneumoniae]	117
SPX0325	325	2986	29	87			NO-HIT	6
SPX0326	326	2987	311	933	1434	2.40E-200	[LN:T44634] [AC:T44634] [PN:choline transporter [imported]] [GN:proWX] [OR:Streptococcus pneumoniae]	100
SPX0327	327	2988	210	630	1042	3.30E-137	[LN:T44634] [AC:T44634] [PN:choline transporter [imported]] [GN:proWX] [OR:Streptococcus pneumoniae]	100
SPX0328	328	2989	270	810	125	1.90E-15	[LN:YA7B_HAEIN]	130

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Q57425:P96338]	
							[GN:HI1077.1]	
							[OR:Haemophilus influenzae]	
							[DE:HYPOTHETICAL PROTEIN HI1077.1]	
							[SP:Q57425:P96338]	
SPX0329	329	2990	179	537	150	8.10E-17	[GI:5881867]	139
							[LN:SC5G9]	
							[AC:AL117385]	
							[PN:putative tetR-family transcriptional regulator]	
							[GN:SC5G9.19c]	
							[OR:Streptomyces coelicolor A3(2)]	
SPX0330	330	2991	300	900	202	5.60E-39	[GI:1944409]	151
							[LN:D87026]	
							[AC:D87026:D28136]	
							[PN:membrane protein]	
							[OR:Bacillus stearothermophilus]	
							[SR:Bacillus stearothermophilus (strain:TRBE14) DNA]	
SPX0331	331	2992	118	354	282	3.60E-34	[LN:B69970]	144
							[AC:B69970]	
							[PN:transcription regulator MerR family homolog yraB]	
							[GN:yraB]	
							[CL:transcription repressor glnR]	
							[OR:Bacillus subtilis]	
SPX0332	332	2993	346	1038	403	3.80E-82	[LN:T44975]	146
							[AC:T44975]	
							[PN:dehydrogenase [imported]]	
							[CL:alcohol dehydrogenase:long-chain alcohol dehydrogenase homology]	
							[OR:Haloferax volcanii]	
SPX0333	333	2994	336	1008	729	4.00E-144	[GI:4097439]	90
							[LN:STU61402]	
							[AC:U61402]	
							[PN:GalR]	
							[GN:galR]	
							[OR:Streptococcus thermophilus]	
SPX0334	334	2995	393	1179	1674	3.10E-227	[GI:4097440]	125
							[LN:STU61402]	
							[AC:U61402]	
							[PN:galactokinase]	
							[GN:galK]	
							[FN:galactose metabolism]	
							[OR:Streptococcus thermophilus]	
SPX0335	335	2996	494	1482	1612	1.90E-230	[GI:4097441]	139
							[LN:STU61402]	
							[AC:U61402]	
							[PN:gal-1-P uridylyltransferase]	
							[GN:galT]	
							[FN:galactose metabolism]	
							[OR:Streptococcus thermophilus]	
SPX0336	336	2997	135	405	467	1.10E-60	[GI:1196924]	81
							[LN:STRDPN2A]	

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:M14339] [PN:unknown] [OR:Streptococcus pneumoniae]	
SPX0337	337	2998	296	888	1461	6.00E-199	[LN:MT21__STRPN] [AC:P04043] [GN:DPNM] [OR:Streptococcus pneumoniae] [EC:2.1.1.72] [DE:METHYLTRANSFERASE DPNII 1) (M.DPNII 1)] [SP:P04043]	137
SPX0338	338	2999	257	771	1366	2.60E-183	[LN:MT22__STRPN] [AC:P09358] [GN:DPNA] [OR:Streptococcus pneumoniae] [EC:2.1.1.72] [DE:METHYLTRANSFERASE DPNII 2) (M.DPNII 2)] [SP:P09358]	137
SPX0339	339	3000	289	867	1492	5.50E-202	[LN:T2D2__STRPN] [AC:P09357] [GN:DPNB] [OR:Streptococcus pneumoniae] [EC:3.1.21.4] [DE:(R.DPNII)] [SP:P09357]	108
SPX0340	340	3001	119	357	462	1.20E-61	[GI:6978345] [LN:STRDPN2A] [AC:M14339] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX0341	341	3002	283	849	436	8.70E-87	[LN:YUNK__BACSU] [AC:O32140] [GN:YUNK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 44.9 KD PROTEIN IN HOM-MRGA INTERGENIC REGION] [SP:O32140]	136
SPX0342	342	3003	194	582	786	3.40E-105	[GI:7160242] [LN:SPN272465] [AC:AJ272465] [PN:xanthine phosphoribosyltransferase] [GN:xpt] [OR:Streptococcus pneumoniae]	120
SPX0343	343	3004	58	174	79	7.70E-11	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0344	344	3005	80	240	113	5.90E-10	[LN:C75408] [AC:C75408] [PN:hypothetical protein] [GN:DR1341] [OR:Deinococcus radiodurans]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0345	345	3006	276	828	1432	8.60E-192	“[LN:EXOA_STRPN] [AC:P21998] [GN:EXOA] [OR:Streptococcus pneumoniae] [EC:3.1.11.2] [DE:EXODEOXYRIBONUCLEASE,] [SP:P21998]]	122
SPX0346	346	3007	76	228			NO-HIT	6
SPX0347	347	3008	90	270			NO-HIT	6
SPX0348	348	3009	177	531	132	9.60E-19	[LN:F69972] [AC:F69972] [PN:probable membrane protein yrbG] [GN:yrbG] [CL:probable membrane protein ycaP] [OR:Bacillus subtilis]	128
SPX0349	349	3010	79	237	129	5.90E-21	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX0350	350	3011	582	1746	685	1.90E-170	[LN:A69829] [AC:A69829] [PN:ABC transporter (ATP-binding protein) homolog yheI] [GN:yheI] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0351	351	3012	294	882	364	6.80E-49	[LN:H69828] [AC:H69828] [PN:ABC transporter (ATP-binding protein) homolog yheH] [GN:yheH] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0352	352	3013	69	207	82	0.00022	[LN:H69828] [AC:H69828] [PN:ABC transporter (ATP-binding protein) homolog yheH] [GN:yheH] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0353	353	3014	264	792	383	3.90E-79	[LN:H69828] [AC:H69828] [PN:ABC transporter (ATP-binding protein) homolog yheH] [GN:yheH] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0354	354	3015	232	696	632	1.10E-83	[GI:6601354] [LN:AF164515] [AC:AF164515] [PN:putative glycosyltransferase Cps7F] [GN:cps7F] [OR:Streptococcus suis]	115

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0355	355	3016	409	1227	973	4.70E-214	[GI:6601355] [LN:AF164515] [AC:AF164515] [PN:Cps7G] [GN:cps7G] [OR:Streptococcus suis]	86
SPX0356	356	3017	63	189			NO-HIT	6
SPX0357	357	3018	89	267			NO-HIT	6
SPX0358	358	3019	709	2127	107	4.60E-08	[GI:2952545] [LN:AF051898] [AC:AF051898] [PN:coronin binding protein] [GN:DB10] [OR:Dictyostelium discoideum]	109
SPX0359	359	3020	82	246			NO-HIT	6
SPX0360	360	3021	69	207			NO-HIT	6
SPX0361	361	3022	95	285			NO-HIT	6
SPX0362	362	3023	96	288	289	6.60E-34	[GI:43589] [LN:HILIC3] [AC:X57315] [PN:UDP-galactose-4-epimerase] [GN:galE] [OR:Haemophilus influenzae]	103
SPX0363	363	3024	66	198			NO-HIT	6
SPX0364	364	3025	220	660			NO-HIT	6
SPX0365	365	3026	356	1068	286	2.20E-30	“[LN:Y4FP_RHISN] [AC:P55454] [GN:Y4FP] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:PROBABLE ABC TRANSPORTER PERIPLASMIC BINDING PROTEIN Y4FP PRECURSOR] [SP:P55454]]	162
SPX0366	366	3027	337	1011	395	3.30E-90	“[LN:A70180] [AC:A70180] [PN:spermidine/putrescine ABC transporter, ATP-binding protein (potA) homolog] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete] ”””	237
SPX0367	367	3028	564	1692	433	1.50E-93	“[LN:Y4FN_RHISN] [AC:P55452] [GN:Y4FN] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:PROBABLE ABC TRANSPORTER PERMEASE PROTEIN Y4FN] [SP:P55452]”	141
SPX0368	368	3029	237	711	192	4.80E-28	[LN:Y647_HAEIN] [AC:Q57424:O05028] [GN:HI0647] [OR:Haemophilus influenzae]	126

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:HYPOTHETICAL PROTEIN HI0647] [SP:Q57424:O05028]	
SPX0369	369	3030	103	309			NO-HIT	6
SPX0370	370	3031	334	1002	144	3.70E-25	[LN:CCPA_BACME] [AC:P46828] [GN:CCPA] [OR:Bacillus megaterium] [DE:GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN)] [SP:P46828]	145
SPX0371	371	3032	66	198			NO-HIT	6
SPX0372	372	3033	69	207			NO-HIT	6
SPX0373	373	3034	488	1464	1100	2.50E-173	“[LN:TRPE_LACLA] [AC:Q02001] [GN:TRPE] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.3.27] [DE:ANTHRANILATE SYNTHASE COMPONENT I,] [SP:Q02001]”	168
SPX0374	374	3035	67	201	108	6.30E-10	[LN:C72489] [AC:C72489] [PN:hypothetical protein APE2554] [GN:APE2554] [OR:Aeropyrum pernix]	92
SPX0375	375	3036	189	567	440	8.80E-75	“[LN:TRPG_LACLA] [AC:Q02003] [GN:TRPG] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.3.27] [DE:TRANSFERASE)] [SP:Q02003]”	146
SPX0376	376	3037	335	1005	1008	9.70E-134	“[LN:TRPD_LACLA] [AC:Q02000] [GN:TRPD] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.4.2.18] [DE:ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE,] [SP:Q02000]”	173
SPX0377	377	3038	256	768	446	6.20E-107	“[LN:TRPC_LACLA] [AC:Q01999] [GN:TRPC] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.1.48] [DE:INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE, (IGPS)] [SP:Q01999]”	178
SPX0378	378	3039	210	630	225	3.50E-49	“[LN:TRPF_LACLA] [AC:Q02002]	185

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:TRPF] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:5.3.1.24] [DE:N-(5'-PHOSPHORIBOSYL)ANTHRANILATE ISOMERASE, (PRAI)] [SP:Q02002]”	
SPX0379	379	3040	408	1224	930	9.00E-225	“[LN:TRPB_LACLA] [AC:Q01998] [GN:TRPB] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.2.1.20] [DE:TRYPTOPHAN SYNTHASE BETA CHAIN,] [SP:Q01998]”	165
SPX0380	380	3041	278	834	1137	5.00E-153	[GI:5231181] [LN:AF157817] [AC:AF157817] [PN:tryptophan synthase alpha chain] [GN:trpA] [OR:Streptococcus pneumoniae]	117
SPX0381	381	3042	68	204			NO-HIT	6
SPX0382	382	3043	244	732			NO-HIT	6
SPX0383	383	3044	384	1152	92	2.70E-13	[LN:H70940] [AC:H70940] [PN:probable helix-turn-helix motif at aa 18-39] [GN:Rv2017] [OR:Mycobacterium tuberculosis]	116
SPX0384	384	3045	220	660	109	1.00E-14	[LN:B72220] [AC:B72220] [PN:type IV prepilin peptidase] [GN:TM1696] [CL:type IV prepilin peptidase] [OR:Thermotoga maritima]	124
SPX0385	385	3046	149	447	92	0.00011	[LN:E72312] [AC:E72312] [PN:conserved hypothetical protein] [GN:TM0968] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimJ] [OR:Thermotoga maritima]	169
SPX0386	386	3047	104	312	138	1.30E-12	“[GI:6332767] [LN:AB033763] [AC:AB033763:AB014419:AB014429:AB014439] [PN:hypothetical protein] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:NCTC10442) DNA, clone_lib:Lambda das]”	194
SPX0387	387	3048	68	204	98	8.10E-08	[GI:2772940] [LN:AF034574] [AC:AF034574] [PN:putative cruciform DNA binding protein] [GN:Gv1]	116

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Glomus versiforme]	
SPX0388	388	3049	203	609	515	8.70E-64	[GI:727436] [LN:LLU23376] [AC:U23376] [OR:Lactococcus lactis]	61
SPX0389	389	3050	191	573			NO-HIT	6
SPX0390	390	3051	513	1539	193	6.90E-35	“[GI:6782400] [LN:STC133440] [AC:AJ133440] [PN:multigene regulator protein Mgc, putative] [GN:mgc] [FN:gene regulatory function, putative] [OR:Streptococcus dysgalactiae]”	171
SPX0391	391	3052	334	1002	185	3.80E-46	[LN:SCRR_STRMU] [AC:Q54430] [GN:SCRR] [OR:Streptococcus mutans] [DE:SUCROSE OPERON REPRESSOR (SCR OPERON REGULATORY PROTEIN)] [SP:Q54430]	137
SPX0392	392	3053	315	945	376	2.50E-89	[LN:A69653] [AC:A69653] [PN:transmembrane lipoprotein lplB] [GN:lplB] [CL:inner membrane protein ugpA] [OR:Bacillus subtilis]	125
SPX0393	393	3054	306	918	450	1.20E-73	[LN:LPLC_BACSU] [AC:P39129] [GN:LPLC] [OR:Bacillus subtilis] [DE:LPLC PROTEIN] [SP:P39129]	90
SPX0394	394	3055	539	1617	167	2.90E-15	[GI:4056657] [LN:AF098273] [AC:AF098273] [PN:peripheral protein] [GN:glucuronic acid catabolism operon] [OR:Bacillus stearothermophilus]	136
SPX0395	395	3056	141	423			NO-HIT	6
SPX0396	396	3057	440	1320	305	2.30E-76	“[LN:RAFD_ECOLI] [AC:P16553] [GN:RAFD] [OR:Escherichia coli] [EC:3.2.1.26] [DE:RAFFINOSE INVERTASE, (INVERTASE)] [SP:P16553]”	125
SPX0397	397	3058	424	1272	722	7.10E-167	[LN:D69981] [AC:D69981] [PN:conserved hypothetical protein yrvN] [GN:yrvN] [CL:Haemophilus influenzae conserved hypothetical protein HI1590]	163

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis]	
SPX0398	398	3059	74	222			NO-HIT	6
SPX0399	399	3060	60	180			NO-HIT	6
SPX0400	400	3061	84	252			NO-HIT	6
SPX0401	401	3062	151	453	92	4.90E-09	[LN:T13557] [AC:T13557] [PN:hypothetical protein 17] [CL:Archaeoglobus fulgidus conserved hypothetical protein AF1072] [OR:Bacillus phage phi-105]	146
SPX0402	402	3063	157	471			NO-HIT	6
SPX0403	403	3064	363	1089			NO-HIT	6
SPX0404	404	3065	345	1035	130	8.00E-15	[GI:758793] [LN:ECU23723] [AC:U23723] [FN:unknown] [OR:Escherichia coli]	72
SPX0405	405	3066	748	2244	68	0.00048	[GI:758794] [LN:ECU23723] [AC:U23723] [FN:unknown] [OR:Escherichia coli]	72
SPX0406	406	3067	340	1020			NO-HIT	6
SPX0407	407	3068	317	951	395	1.40E-81	"[LN:T43740] [AC:T43740] [PN:probable ribosomal protein L11 methyltransferase, [imported] :hypothetical protein 35] [CL:ribosomal protein L11 methyltransferase:bioC homology] [OR:Listeria monocytogenes] [EC:2.1.1.-]"	215
SPX0408	408	3069	248	744	461	1.90E-64	[LN:T43741] [AC:T43741] [PN:conserved hypothetical protein orf29 [imported]] [OR:Listeria monocytogenes]	104
SPX0409	409	3070	599	1797	300	2.10E-57	"[LN:PEPF__BACSU] [AC:O31605] [GN:YJBG] [OR:Bacillus subtilis] [EC:3.4.24.-] [DE:OLIGOENDOPEPTIDASE F HOMOLOG,] [SP:O31605]"	123
SPX0410	410	3071	226	678	312	1.80E-58	[LN:D69780] [AC:D69780] [PN:hypothetical protein ydfF] [GN:ydfF] [OR:Bacillus subtilis]	87
SPX0411	411	3072	217	651	394	1.40E-54	[LN:G69803] [AC:G69803] [PN:ABC transporter (ATP-binding protein) homolog yfiL] [GN:yfiL] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology]	188

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis]	
SPX0412	412	3073	255	765	346	4.20E-44	“[LN:NODJ_RHILV] [AC:P06755] [GN:NODJ] [OR:Rhizobium leguminosarum] [SR:,biovar viciae] [DE:NODULATION PROTEIN J] [SP:P06755]”	126
SPX0413	413	3074	264	792			NO-HIT	6
SPX0414	414	3075	210	630	257	3.90E-39	“[LN:A69428] [AC:A69428] [PN:glycerol uptake facilitator, MIP channel (glpF) homolog] [CL:nodulin-26] [OR:Archaeoglobus fulgidus]”	130
SPX0415	415	3076	60	180			NO-HIT	6
SPX0416	416	3077	191	573	467	1.10E-60	[LN:D69868] [AC:D69868] [PN:conserved hypothetical protein ykvM] [GN:ykvM] [CL:hypothetical protein ykvM] [OR:Bacillus subtilis]	128
SPX0417	417	3078	105	315	193	1.80E-43	[GI:5640117] [LN:LMO133006] [AC:AJ133006] [PN:thioredoxin] [GN:trxA] [FN:thiol:disulfide interchange] [OR:Listeria monocytogenes]	129
SPX0418	418	3079	101	303			NO-HIT	6
SPX0419	419	3080	67	201			NO-HIT	6
SPX0420	420	3081	75	225			NO-HIT	6
SPX0421	421	3082	318	954	433	2.70E-56	[LN:YQJA_BACSU] [AC:P54538] [GN:YQJA] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 37.1 KD PROTEIN IN BMRU-ANSR INTERGENIC REGION] [SP:P54538]	137
SPX0422	422	3083	287	861	745	1.50E-100	[LN:YGJU_HAEIN] [AC:P45246] [GN:HI1545] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL SYMPORTER HI1545] [SP:P45246]	114
SPX0423	423	3084	75	225	153	4.00E-15	[LN:YGJU_ECOLI] [AC:P42602] [GN:YGJU] [OR:Escherichia coli] [DE:HYPOTHETICAL 43.5 KD PROTEIN IN EBGC-UXAA INTERGENIC REGION (O414)] [SP:P42602]	143
SPX0424	424	3085	140	420	163	2.90E-17	“[GI:2828366]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:AB010789] [AC:AB010789] [OR:Lactococcus lactis] [SR:Lactococcus lactis (sub_species:lactis, strain:01-7) DNA]”	
SPX0425	425	3086	66	198	102	1.80E-07	[LN:A69867] [AC:A69867] [PN:conserved hypothetical protein ykuT] [GN:ykuT] [CL:Escherichia coli hypothetical 30.9K protein (sbm-fba intergenic region)] [OR:Bacillus subtilis]	174
SPX0426	426	3087	59	177			NO-HIT	6
SPX0427	427	3088	303	909	270	1.70E-35	[GI:5123526] [LN:AF036485] [AC:AF036485:AF036486:AF036487:U93364] [PN:hypothetical protein] [OR:Plasmid pNZ4000]	112
SPX0428	428	3089	176	528	326	9.70E-41	[LN:YQEG_BACSU] [AC:P54452] [GN:YQEG] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 20.1 KD PROTEIN IN NUCB-AROD INTERGENIC REGION] [SP:P54452]	137
SPX0429	429	3090	369	1107	1037	1.30E-164	[LN:YQEH_BACSU] [AC:P54453] [GN:YQEH] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 41.0 KD PROTEIN IN NUCB-AROD INTERGENIC REGION] [SP:P54453]	137
SPX0430	430	3091	104	312	211	5.20E-24	[LN:YQEI_BACSU] [AC:P54454] [GN:YQEI] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 10.8 KD PROTEIN IN AROD-COMER INTERGENIC REGION] [SP:P54454]	138
SPX0431	431	3092	205	615	491	1.30E-63	[LN:YQEJ_BACSU] [AC:P54455] [GN:YQEJ] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 22.2 KD PROTEIN IN AROD-COMER INTERGENIC REGION] [SP:P54455]	138
SPX0432	432	3093	198	594	375	2.90E-47	[LN:YQEK_BACSU] [AC:P54456] [GN:YQEK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 21.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION] [SP:P54456]	138
SPX0433	433	3094	167	501	110	4.90E-10	[GI:7160132] [LN:SC5C11] [AC:AL158060] [PN:putative isochorismatase.]	119

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:SC5C11.12] [OR:Streptomyces coelicolor A3(2)]	
SPX0434	434	3095	118	354	267	7.00E-32	[LN:YQEL__BACSU] [AC:P54457] [GN:YQEL] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 13.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION] [SP:P54457]	138
SPX0435	435	3096	247	741	287	4.60E-55	[LN:YQEM__BACSU] [AC:P54458] [GN:YQEM] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 28.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION] [SP:P54458]	138
SPX0436	436	3097	188	564			NO-HIT	6
SPX0437	437	3098	366	1098	222	9.30E-38	[LN:B69875] [AC:B69875] [PN:conserved hypothetical protein ylbM] [GN:ylbM] [OR:Bacillus subtilis]	97
SPX0438	438	3099	63	189			NO-HIT	6
SPX0439	439	3100	916	2748	86	1.00E-10	[GI:1890605] [LN:RMEXPGNS] [AC:Z79692] [PN:ORF25] [OR:Sinorhizobium meliloti]	77
SPX0440	440	3101	537	1611	1547	2.30E-211	[GI:6984124] [LN:AF228345] [AC:AF228345] [PN:unknown] [OR:Listeria monocytogenes]	81
SPX0441	441	3102	209	627	614	1.40E-80	[LN:B69878] [AC:B69878] [PN:guanylate kinase homolog yloD] [GN:yloD] [CL:guanylate kinase:guanylate kinase homology] [OR:Bacillus subtilis]	139
SPX0442	442	3103	105	315	91	1.00E-09	[LN:C69878] [AC:C69878] [PN:hypothetical protein yloH] [GN:yloH] [OR:Bacillus subtilis]	87
SPX0443	443	3104	799	2397	953	2.70E-237	[LN:PRIA__BACSU] [AC:P94461:O34941] [GN:PRIA] [OR:Bacillus subtilis] [DE:PRIMOSOMAL PROTEIN N'(REPLICATION FACTOR Y)] [SP:P94461:O34941]	136
SPX0444	444	3105	312	936	434	1.70E-64	[GI:1772500] [LN:BSPRIADFS]	105

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Y10304] [PN:Met-tRNAi formyl transferase] [GN:fmt] [OR:Bacillus subtilis]	
SPX0445	445	3106	176	528	248	1.40E-39	[GI:1772501] [LN:BSPRIADFS] [AC:Y10304] [GN:sun] [OR:Bacillus subtilis]	71
SPX0446	446	3107	139	417	130	1.30E-09	[GI:4210751] [LN:LLA132604] [AC:AJ132604] [PN:sunL protein] [GN:sunL] [OR:Lactococcus lactis]	93
SPX0447	447	3108	233	699	606	2.70E-80	[GI:4210751] [LN:LLA132604] [AC:AJ132604] [PN:sunL protein] [GN:sunL] [OR:Lactococcus lactis]	93
SPX0448	448	3109	247	741	371	1.50E-85	[GI:4210752] [LN:LLA132604] [AC:AJ132604] [PN:pppL protein] [GN:pppL] [FN:putative phosphoprotein phosphatase] [OR:Lactococcus lactis]	134
SPX0449	449	3110	660	1980	985	1.70E-128	[GI:4210753] [LN:LLA132604] [AC:AJ132604] [PN:hypothetical protein] [OR:Lactococcus lactis]	91
SPX0450	450	3111	68	204			NO-HIT	6
SPX0451	451	3112	74	222			NO-HIT	6
SPX0452	452	3113	170	510	125	1.80E-09	[LN:YJJP_ECOLI] [AC:P39402] [GN:YJJP] [OR:Escherichia coli] [DE:HYPOTHETICAL 30.5 KD PROTEIN IN DNAT-BGLJ INTERGENIC REGION (F277)] [SP:P39402]	143
SPX0453	453	3114	112	336	94	4.60E-07	[LN:G81321] [AC:G81321] [PN:probable integral membrane protein Cj1165c [imported]] [GN:Cj1165c] [OR:Campylobacter jejuni]	121
SPX0454	454	3115	67	201	217	3.80E-24	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0455	455	3116	130	390	362	7.00E-46	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX0456	456	3117	96	288			NO-HIT	6
SPX0457	457	3118	391	1173	775	1.30E-101	[GI:4096796] [LN:SCU40157] [AC:U40157] [OR:Staphylococcus carnosus]	67
SPX0458	458	3119	425	1275	354	8.70E-92	[LN:MVAA_PSEMV] [AC:P13702] [GN:MVAA] [OR:Pseudomonas mevalonii] [EC:1.1.1.88] [DE:REDUCTASE)] [SP:P13702]	106
SPX0459	459	3120	322	966	1108	1.00E-150	[LN:SCRR_STRMU] [AC:Q54430] [GN:SCRR] [OR:Streptococcus mutans] [DE:SUCROSE OPERON REPRESSOR (SCR OPERON REGULATORY PROTEIN)] [SP:Q54430]	137
SPX0460	460	3121	485	1455	1691	3.60E-229	"[LN:S68598] [AC:S68598] [PN:sucrose-6-phosphate hydrolase ScrB] [GN:scrB] [CL:sucrose-6-phosphate hydrolase] [OR:Streptococcus sobrinus] [SR:strain 6715, , strain 6715] [SR:strain 6715,]"	189
SPX0461	461	3122	77	231			NO-HIT	6
SPX0462	462	3123	66	198			NO-HIT	6
SPX0463	463	3124	628	1884	717	6.70E-255	"[LN:S68599] [AC:S68599] [PN:phosphotransferase system enzyme II,, sucrose-specific:sucrose-specific enzyme II:sucrose-specific enzyme II] [GN:scrA] [OR:Streptococcus sobrinus]"	177
SPX0464	464	3125	296	888	701	3.80E-158	"[LN:SCRK_STRMU] [AC:Q07211] [GN:SCRK] [OR:Streptococcus mutans] [EC:2.7.1.4] [DE:FRUCTOKINASE,] [SP:Q07211]"	109
SPX0465	465	3126	71	213			NO-HIT	6
SPX0466	466	3127	200	600	103	2.90E-07	[LN:YDJZ_ECOLI] [AC:P76221] [GN:YDJZ] [OR:Escherichia coli] [DE:HYPOTHETICAL 26.2 KD PROTEIN IN XTHA-GDHA INTERGENIC REGION]	136

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0467	467	3128	75	225	200	1.30E-21	[SP:P76221] [GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX0468	468	3129	84	252			NO-HIT	6
SPX0469	469	3130	76	228	187	3.50E-21	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0470	470	3131	298	894	460	1.20E-78	[GI:3688825] [LN:AF084104] [AC:AF084104] [PN:NatA] [GN:natA] [OR:Bacillus firmus]	81
SPX0471	471	3132	400	1200	200	4.70E-36	[GI:1150487] [LN:LSLASAMPT] [AC:Z54312] [PN:unknown] [GN:orf414] [OR:Lactobacillus sakei]	89
SPX0472	472	3133	56	168			NO-HIT	6
SPX0473	473	3134	283	849			NO-HIT	6
SPX0474	474	3135	241	723	460	1.50E-58	“[LN:G71842] [AC:G71842] [PN:probable ABC transporter, ATP-binding protein] [GN:jhp1141] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Helicobacter pylori] [SR:strain J99, , strain J99] [SR:strain J99,]”	238
SPX0475	475	3136	70	210	95	4.40E-07	[LN:E72756] [AC:E72756] [PN:hypothetical protein APE0042] [GN:APE0042] [OR:Aeropyrum pernix]	92
SPX0476	476	3137	122	366	256	3.70E-30	[GI:5822822] [LN:AB024564] [AC:AB024564] [PN:YHCF] [GN:yhcF] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA]	129
SPX0477	477	3138	158	474	405	8.00E-51	[LN:E72221] [AC:E72221] [PN:conserved hypothetical protein] [GN:TM1707] [CL:conserved hypothetical protein HI0943] [OR:Thermotoga maritima]	139

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0478	478	3139	390	1170			NO-HIT	6
SPX0479	479	3140	299	897	300	2.40E-58	[LN:DNAI_BACSU] [AC:P06567] [GN:DNAI] [OR:Bacillus subtilis] [DE:PRIMOSOMAL PROTEIN DNAI] [SP:P06567]	101
SPX0480	480	3141	238	714	239	3.90E-35	"[LN:CHRR_PSESP] [AC:P96977] [GN:CHRR] [OR:Pseudomonas sp] [SR:,strain G-1] [DE:CR(VI) REDUCTASE] [SP:P96977]"	110
SPX0481	481	3142	437	1311	1986	1.10E-270	"[GI:6681650] [LN:AB016077] [AC:AB016077] [PN:phosphoglycerate dehydrogenase] [GN:pgdA] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:MT8148) DNA, clone:pYT570]"	174
SPX0482	482	3143	115	345			NO-HIT	6
SPX0483	483	3144	118	354			NO-HIT	6
SPX0484	484	3145	64	192			NO-HIT	6
SPX0485	485	3146	284	852	103	3.00E-11	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX0486	486	3147	131	393			NO-HIT	6
SPX0487	487	3148	264	792			NO-HIT	6
SPX0488	488	3149	226	678			NO-HIT	6
SPX0489	489	3150	211	633	282	3.20E-42	[LN:YYBJ_BACSU] [AC:P37494] [GN:YYBJ] [OR:Bacillus subtilis] [DE:INTERGENIC REGION] [SP:P37494]	95
SPX0490	490	3151	96	288			NO-HIT	6
SPX0491	491	3152	838	2514	989	0	[LN:SECA_LISMO] [AC:P47847] [GN:SECA] [OR:Listeria monocytogenes] [DE:PREPROTEIN TRANSLOCASE SECA SUBUNIT] [SP:P47847]	118
SPX0492	492	3153	344	1032	156	7.30E-47	[GI:7380801] [LN:NMA7Z2491] [AC:AL162758:AL157959]	136

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:phospho-2-dehydr-3-deoxyheptonate aldolase] [GN:aroG] [OR:Neisseria meningitidis]	
SPX0493	493	3154	215	645	178	3.50E-34	[LN:AROF_ECOLI] [AC:P00888] [GN:AROF] [OR:Escherichia coli] [EC:4.1.2.15] [DE:SYNTHETASE) (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE)] [SP:P00888]	157
SPX0494	494	3155	111	333			NO-HIT	6
SPX0495	495	3156	89	267	132	2.60E-14	[GI:7380801] [LN:NMA7Z2491] [AC:AL162758:AL157959] [PN:phospho-2-dehydr-3-deoxyheptonate aldolase] [GN:aroG] [OR:Neisseria meningitidis]	136
SPX0496	496	3157	121	363	177	1.70E-30	“[LN:H69772] [AC:H69772] [PN:holo-[acyl-carrier-protein] synthase,] [GN:ydcB] [CL:holo-ACP synthase] [OR:Bacillus subtilis] [EC:2.7.8.7]”	137
SPX0497	497	3158	368	1104	1852	1.10E-252	[GI:5759209] [LN:AF171873] [AC:AF171873] [PN:alanine racemase] [GN:alr] [OR:Streptococcus pneumoniae]	101
SPX0498	498	3159	672	2016	3357	0	“[LN:RECG_STRPN] [AC:Q54900] [GN:RECG:MMSA] [OR:Streptococcus pneumoniae] [EC:3.6.1.-] [DE:ATP-DEPENDENT DNA HELICASE RECG,] [SP:Q54900]”	137
SPX0499	499	3160	89	267	263	2.70E-32	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0500	500	3161	68	204			NO-HIT	6
SPX0501	501	3162	327	981	447	1.40E-92	[GI:2353697] [LN:AF001926] [AC:AF001926] [PN:xylan esterase 1] [GN:axe1] [OR:Thermoanaerobacterium sp. 'JW/SL YS485']	117
SPX0502	502	3163	206	618	355	1.20E-44	[LN:T30285]	79

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	
SPX0503	503	3164	168	504	172	1.10E-15	[GI:4101572] [LN:AF004842] [AC:AF004842] [PN:major royal jelly protein MRJP5] [GN:MRJP5] [OR:Apis mellifera] [SR:honeybee]	122
SPX0504	504	3165	76	228	371	3.40E-50	[LN:T30286] [AC:T30286:S26297] [PN:hypothetical protein 76] [OR:Streptococcus pneumoniae]	89
SPX0505	505	3166	938	2814	4378	0	"[LN:NANA_STRPN] [AC:Q59959:Q54722] [GN:NANA] [OR:Streptococcus pneumoniae] [EC:3.2.1.18] [DE:SIALIDASE A PRECURSOR, (NEURAMINIDASE A)] [SP:Q59959:Q54722]"	155
SPX0506	506	3167	72	216			NO-HIT	6
SPX0507	507	3168	151	453	773	6.10E-103	[GI:1163111] [LN:SPU43526] [AC:U43526] [OR:Streptococcus pneumoniae]	68
SPX0508	508	3169	446	1338	2307	0	[GI:1163112] [LN:SPU43526] [AC:U43526] [OR:Streptococcus pneumoniae]	68
SPX0509	509	3170	295	885	1542	2.00E-218	[GI:1163113] [LN:SPU43526] [AC:U43526] [OR:Streptococcus pneumoniae]	68
SPX0510	510	3171	278	834	1375	3.80E-192	[GI:1163114] [LN:SPU43526] [AC:U43526] [OR:Streptococcus pneumoniae]	68
SPX0511	511	3172	254	762	1250	1.40E-167	"[LN:NANB_STRPN] [AC:Q54727] [GN:NANB] [OR:Streptococcus pneumoniae] [EC:3.2.1.18] [DE:SIALIDASE B PRECURSOR, (NEURAMINIDASE B)] [SP:Q54727]"	141
SPX0512	512	3173	126	378	333	1.60E-40	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX0513	513	3174	211	633	762	7.90E-101	[GI:517210]	65

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	
SPX0514	514	3175	467	1401	2363	0	"[LN:NANB_STRPN] [AC:Q54727] [GN:NANB] [OR:Streptococcus pneumoniae] [EC:3.2.1.18] [DE:SIALIDASE B PRECURSOR, (NEURAMINIDASE B)] [SP:Q54727]"	141
SPX0515	515	3176	368	1104	1838	1.30E-249	[LN:YJHC_STRPN] [AC:Q54728] [OR:Streptococcus pneumoniae] [DE:HYPOTHETICAL PROTEIN IN NANB 3'REGION (ORF-5) (FRAGMENT)] [SP:Q54728]	131
SPX0516	516	3177	130	390			NO-HIT	6
SPX0517	517	3178	233	699	331	1.90E-64	"[LN:C70180] [AC:C70180] [PN:conserved hypothetical protein BB0644] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	125
SPX0518	518	3179	506	1518	839	2.60E-178	"[LN:D70180] [AC:D70180] [PN:phosphotransferase system enzyme II,, glucose-specific, factor II] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete] [EC:2.7.1.69]"	167
SPX0519	519	3180	443	1329	110	1.90E-09	[GI:6137033] [LN:SCF11] [AC:AL132662] [PN:putative sugar transporter sugar binding] [GN:SCF11.11] [OR:Streptomyces coelicolor A3(2)]	132
SPX0520	520	3181	296	888	211	1.60E-50	"[LN:E72357] [AC:E72357] [PN:sugar ABC transporter, permease protein] [GN:TM0596] [CL:inner membrane protein malF] [OR:Thermotoga maritima]"	140
SPX0521	521	3182	280	840	434	1.10E-65	"[LN:F72379] [AC:F72379] [PN:sugar ABC transporter, permease protein] [GN:TM0430] [CL:maltose transport protein malG] [OR:Thermotoga maritima]"	143
SPX0522	522	3183	151	453	109	2.60E-13	[LN:YJGK_ECOLI] [AC:P39335] [GN:YJGK] [OR:Escherichia coli] [DE:HYPOTHETICAL 17.3 KD PROTEIN IN PYRL-ARGI INTERGENIC REGION (O153B)]	144

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P39335]	
SPX0523	523	3184	118	354			NO-HIT	6
SPX0524	524	3185	66	198			NO-HIT	6
SPX0525	525	3186	215	645			NO-HIT	6
SPX0526	526	3187	220	660	830	1.30E-109	[GI:2385360] [LN:CTSIALIDA] [AC:Y08695] [PN:putative acylneuraminate lyase] [OR:Clostridium tertium]	100
SPX0527	527	3188	100	300	68	0.00016	[GI:42131] [LN:ECNPL] [AC:X03345] [OR:Escherichia coli]	55
SPX0528	528	3189	295	885	423	2.50E-53	[LN:YNGA_CLOPE] [AC:P26832] [OR:Clostridium perfringens] [DE:HYPOTHETICAL PROTEIN IN NAGH 5'REGION (ORFA) (FRAGMENT)] [SP:P26832]	129
SPX0529	529	3190	283	849	170	2.10E-25	[LN:Y143_HAEIN] [AC:P44540] [GN:HI0143] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI0143] [SP:P44540]	112
SPX0530	530	3191	681	2043	3365	0	[GI:4009463] [LN:AF068901] [AC:AF068901] [PN:penicillin-binding protein 2b] [GN:pbp2b] [OR:Streptococcus pneumoniae]	116
SPX0531	531	3192	119	357	524	5.00E-69	[GI:49383] [LN:SPPBP2BH] [AC:Z21808] [PN:internal region of the penicillin-binding] [FN:penicillin-resistantance] [OR:Streptococcus pneumoniae]	143
SPX0532	532	3193	171	513	625	4.80E-83	[GI:4009464] [LN:AF068901] [AC:AF068901] [PN:RecM] [GN:recM] [OR:Streptococcus pneumoniae]	90
SPX0533	533	3194	348	1044	1766	1.30E-239	[GI:4009465] [LN:AF068901] [AC:AF068901] [PN:D-Ala-D-Ala ligase] [GN:ddl] [OR:Streptococcus pneumoniae]	103
SPX0534	534	3195	458	1374	2278	0	[GI:4009466] [LN:AF068901]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AF068901] [PN:D-Ala-D-Ala adding enzyme] [GN:murF] [OR:Streptococcus pneumoniae]	
SPX0535	535	3196	449	1347	403	3.50E-76	[GI:5822769] [LN:AB024553] [AC:AB024553] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA]	109
SPX0536	536	3197	192	576	954	1.80E-127	[GI:4009467] [LN:AF068901] [AC:AF068901] [PN:MutT] [GN:mutT] [OR:Streptococcus pneumoniae]	90
SPX0537	537	3198	199	597	1003	1.80E-140	[GI:4009468] [LN:AF068901] [AC:AF068901] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0538	538	3199	87	261			NO-HIT	6
SPX0539	539	3200	458	1374	2266	0	[GI:4009469] [LN:AF068901] [AC:AF068901] [PN:cell division protein FtsA] [GN:ftsA] [OR:Streptococcus pneumoniae]	112
SPX0540	540	3201	419	1257	2094	4.10E-280	[GI:4009470] [LN:AF068901] [AC:AF068901] [PN:cell division protein FtsZ] [GN:ftsZ] [OR:Streptococcus pneumoniae]	112
SPX0541	541	3202	79	237	178	2.40E-19	[LN:A71218] [AC:A71218] [PN:hypothetical protein PH0004] [GN:PH0004] [OR:Pyrococcus horikoshii]	95
SPX0542	542	3203	224	672	1101	6.20E-148	[GI:4009471] [LN:AF068901] [AC:AF068901] [PN:YlmE] [GN:ylmE] [OR:Streptococcus pneumoniae]	90
SPX0543	543	3204	180	540	920	1.10E-122	[GI:4009472] [LN:AF068901] [AC:AF068901] [PN:YlmF] [GN:ylmF] [OR:Streptococcus pneumoniae]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0544	544	3205	88	264	431	1.50E-58	[GI:4009473] [LN:AF068901] [AC:AF068901] [PN:YlmG] [GN:ylmG] [OR:Streptococcus pneumoniae]	90
SPX0545	545	3206	262	786	1318	2.10E-179	[GI:4009474] [LN:AF068901] [AC:AF068901] [PN:YlmH] [GN:ylmH] [OR:Streptococcus pneumoniae]	90
SPX0546	546	3207	263	789	1297	1.90E-171	[GI:4009475] [LN:AF068901] [AC:AF068901] [PN:cell division protein DivIVA] [GN:divIVA] [OR:Streptococcus pneumoniae]	116
SPX0547	547	3208	931	2793	4840	0	[GI:4009476] [LN:AF068901] [AC:AF068901] [PN:isoleucine-tRNA synthetase] [GN:ileS] [OR:Streptococcus pneumoniae]	112
SPX0548	548	3209	62	186	82	0.00016	"[LN:A56034] [AC:A56034] [PN:insulin activator factor] [CL:human insulin activator factor] [OR:Homo sapiens] [SR:, man]"	120
SPX0549	549	3210	87	261	192	2.20E-20	"[LN:A56034] [AC:A56034] [PN:insulin activator factor] [CL:human insulin activator factor] [OR:Homo sapiens] [SR:, man]"	120
SPX0550	550	3211	63	189			NO-HIT	6
SPX0551	551	3212	72	216			NO-HIT	6
SPX0552	552	3213	231	693	1201	1.00E-160	[GI:5578891] [LN:SPN131985] [AC:AJ131985] [PN:phosphoglyceromutase] [GN:gpmA] [OR:Streptococcus pneumoniae]	107
SPX0553	553	3214	804	2412	232	2.70E-95	[GI:5912520] [LN:SCF12] [AC:AL117669] [PN:putative large secreted protein] [GN:SCF12.20c] [OR:Streptomyces coelicolor A3(2)]	124

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0554	554	3215	237	711	642	7.20E-84	“[LN:B69477] [AC:B69477] [PN:ABC transporter, ATP-binding protein homolog] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Archaeoglobus fulgidus]”	179
SPX0555	555	3216	925	2775	106	1.10E-16	[LN:C69477] [AC:C69477] [PN:hypothetical protein AF1820] [OR:Archaeoglobus fulgidus]	84
SPX0556	556	3217	173	519	878	1.10E-117	[GI:4193357] [LN:AF055088] [AC:AF055088:AF055087] [PN:putative hydrophobic transmembrane protein] [GN:psaD] [OR:Streptococcus pneumoniae]	137
SPX0557	557	3218	310	930	1579	2.30E-209	[GI:3258602] [LN:SPU40786] [AC:U40786] [PN:surface antigen A variant precursor] [GN:psaA] [FN:putative fimbrial adhesin] [OR:Streptococcus pneumoniae]	150
SPX0558	558	3219	283	849	1171	8.00E-165	[LN:P29K_STRGC] [AC:P42361] [OR:Streptococcus gordonii challis] [DE:29 KD MEMBRANE PROTEIN IN PSAA 5'REGION (ORF1)] [SP:P42361]	127
SPX0559	559	3220	252	756	1208	2.80E-164	[GI:4193356] [LN:AF055088] [AC:AF055088:AF055087] [PN:ATP-binding cassette] [GN:psaB] [OR:Streptococcus pneumoniae]	115
SPX0560	560	3221	651	1953	1477	0	[GI:5139244] [LN:AF116532] [AC:AF116532] [PN:endopeptidase O] [GN:pepO] [OR:Streptococcus parasanguinis]	104
SPX0561	561	3222	210	630	181	2.90E-36	[LN:YQGX_BACSU] [AC:P54501] [GN:YQGX] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 23.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION] [SP:P54501]	138
SPX0562	562	3223	67	201	131	2.00E-13	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0563	563	3224	81	243			NO-HIT	6
SPX0564	564	3225	741	2223	1601	0	[LN:RELA_STREQ] [AC:Q54089] [GN:RELA:REL] [OR:Streptococcus equisimilis] [EC:2.7.6.5] [DE:PROTEIN)] [SP:Q54089]	111
SPX0565	565	3226	81	243	70	0.0001	[LN:Q3ECS7] [AC:A30374:Q90796] [PN:hypothetical 77K protein (spoT 3'region)] [CL:Escherichia coli hypothetical 77K protein (spoT 3'region)] [OR:Escherichia coli]	163
SPX0566	566	3227	148	444	585	2.40E-76	[LN:S39974] [AC:S39974] [PN:hypothetical protein] [CL:conserved hypothetical protein HI0670] [OR:Streptococcus equisimilis]	123
SPX0567	567	3228	76	228			NO-HIT	6
SPX0568	568	3229	141	423			NO-HIT	6
SPX0569	569	3230	236	708			NO-HIT	6
SPX0570	570	3231	140	420	433	7.90E-56	[GI:6694218] [LN:AF182402] [AC:AF182402] [PN:metalloregulator RmtA] [GN:rmtA] [OR:Streptococcus gordonii]	105
SPX0571	571	3232	80	240	307	4.00E-38	[GI:6694218] [LN:AF182402] [AC:AF182402] [PN:metalloregulator RmtA] [GN:rmtA] [OR:Streptococcus gordonii]	105
SPX0572	572	3233	281	843	125	2.00E-13	[LN:G75297] [AC:G75297] [PN:conserved hypothetical protein] [GN:DR2233] [CL:probable phosphoesterase MJ0912:phosphoesterase core homology] [OR:Deinococcus radiodurans]	167
SPX0573	573	3234	65	195			NO-HIT	6
SPX0574	574	3235	105	315	142	7.90E-15	[LN:D70063] [AC:D70063] [PN:hypothetical protein ywnA] [GN:ywnA] [OR:Bacillus subtilis]	87
SPX0575	575	3236	73	219	129	2.40E-12	[LN:D70063] [AC:D70063] [PN:hypothetical protein ywnA] [GN:ywnA] [OR:Bacillus subtilis]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0576	576	3237	101	303			NO-HIT	6
SPX0577	577	3238	358	1074			NO-HIT	6
SPX0578	578	3239	226	678	1146	1.70E-157	[GI:5830520] [LN:SPAJ6391] [AC:AJ006391] [PN:response regulator] [GN:rr01] [OR:Streptococcus pneumoniae]	104
SPX0579	579	3240	325	975	1628	3.20E-226	[GI:5830521] [LN:SPAJ6391] [AC:AJ006391] [PN:histidine kinase] [GN:hk01] [OR:Streptococcus pneumoniae]	102
SPX0580	580	3241	661	1983	1149	2.20E-274	[LN:SYT1_BACSU] [AC:P18255:P06570] [GN:THRS:THRSV] [OR:Bacillus subtilis] [EC:6.1.1.3] [DE:(THRRS)] [SP:P18255:P06570]	118
SPX0581	581	3242	130	390			NO-HIT	6
SPX0582	582	3243	97	291			NO-HIT	6
SPX0583	583	3244	70	210			NO-HIT	6
SPX0584	584	3245	210	630	371	1.30E-60	[LN:E70063] [AC:E70063] [PN:hypothetical protein ywnB] [GN:ywnB] [OR:Bacillus subtilis]	87
SPX0585	585	3246	90	270	311	1.10E-38	[LN:F69700] [AC:F69700:S11365:S70690] [PN:ribosomal protein S15 (rpsO):ribosomal protein BS18] [GN:rpsO] [CL:Escherichia coli ribosomal protein S15:eubacterial ribosomal protein S15 homology] [OR:Bacillus subtilis]	214
SPX0586	586	3247	75	225			NO-HIT	6
SPX0587	587	3248	159	477	243	1.30E-38	[GI:1916729] [LN:AF134905] [AC:AF134905:U76550] [PN:CadD] [GN:cadD] [OR:Staphylococcus aureus]	94
SPX0588	588	3249	250	750	165	1.10E-23	[LN:E69826] [AC:E69826] [PN:1-acylglycerol-3-phosphate O-acyltransfera homolog yhdO] [GN:yhdO] [OR:Bacillus subtilis]	117
SPX0589	589	3250	779	2337	414	1.80E-126	"[LN:CTPE_MYCTU] [AC:O08365] [GN:CTPE:RV0908:MTCY21C12.02]	160

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Mycobacterium tuberculosis] [EC:3.6.1.-] [DE:PROBABLE CATION-TRANSPORTING ATPASE E,] [SP:O08365]"	
SPX0590	590	3251	77	231			NO-HIT	6
SPX0591	591	3252	488	1464	122	9.60E-23	[GI:1480429] [LN:BSU18943] [AC:U18943:X99465] [GN:MtlR] [OR:Bacillus stearothermophilus]	88
SPX0592	592	3253	143	429	60	0.00018	[GI:2317740] [LN:AF013987] [AC:AF013987] [PN:nitrogen regulatory IIA protein] [GN:ptsN] [OR:Vibrio cholerae]	108
SPX0593	593	3254	151	453	223	5.60E-24	"[LN:HRSA_ECOLI] [AC:P54745] [GN:HRSA] [OR:Escherichia coli] [EC:2.7.1.69] [DE:HRSA PROTEIN,] [SP:P54745]"	106
SPX0594	594	3255	104	312	236	1.30E-27	"[LN:PTWB_ECOLI] [AC:P32673] [GN:FRWB] [OR:Escherichia coli] [EC:2.7.1.69] [DE:II, B COMPONENT),] [SP:P32673]"	110
SPX0595	595	3256	369	1107	398	8.00E-60	"[LN:PTFB_XANCP] [AC:P23355] [GN:FRUA] [OR:Xanthomonas campestris] [SR:,pvcampestris] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EII-FRU)] [SP:P23355]"	141
SPX0596	596	3257	232	696	560	6.50E-74	"[LN:ALSE_ECOLI] [AC:P32719] [GN:ALSE] [OR:Escherichia coli] [EC:5.1.3.-] [DE:D-ALLULOSE-6-PHOSPHATE 3-EPIMERASE,] [SP:P32719]"	127
SPX0597	597	3258	657	1971	1371	2.00E-292	"[LN:TKT_STRPN] [AC:P22976] [GN:RECP] [OR:Streptococcus pneumoniae] [EC:2.2.1.1]	127

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:PROBABLE TRANSKETOLASE, (TK)] [SP:P22976]”	
SPX0598	598	3259	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0599	599	3260	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0600	600	3261	278	834	587	1.60E-140	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0601	601	3262	177	531			NO-HIT	6
SPX0602	602	3263	139	417	138	1.40E-10	[LN:T35180] [AC:T35180] [PN:hypothetical protein SC5A7.31] [GN:SC5A7.31] [CL:Streptomyces coelicolor hypothetical protein SC5A7.31] [OR:Streptomyces coelicolor]	160
SPX0603	603	3264	226	678	612	1.10E-79	[GI:1813343] [LN:D78182] [AC:D78182] [GN:ORF3] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	114
SPX0604	604	3265	266	798	344	5.40E-120	[GI:1813344] [LN:D78182] [AC:D78182] [GN:ORF4] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	114
SPX0605	605	3266	368	1104	335	4.50E-71	[LN:YURR_BACSU] [AC:O32159] [GN:YURR] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 39.4 KD OXIDOREDUCTASE IN HOM-MRGA INTERGENIC REGION] [SP:O32159]	143
SPX0606	606	3267	64	192			NO-HIT	6
SPX0607	607	3268	340	1020	1167	7.10E-156	[LN:GALE_BACSU] [AC:P55180] [GN:GALE] [OR:Bacillus subtilis] [EC:5.1.3.2] [DE:GALACTOSE 4-EPIMERASE)] [SP:P55180]	113
SPX0608	608	3269	318	954	441	7.80E-87	[LN:CSBB_BACSU] [AC:Q45539]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:CSBB] [OR:Bacillus subtilis] [DE:CSBB PROTEIN] [SP:Q45539]	
SPX0609	609	3270	222	666			NO-HIT	6
SPX0610	610	3271	69	207	55	0.00029	"[LN:FER_MOOTH] [AC:P00203] [OR:Moorella thermoacetica] [SR:,Clostridium thermoaceticum] [DE:FERREDOXIN] [SP:P00203]"	117
SPX0611	611	3272	159	477			NO-HIT	6
SPX0612	612	3273	224	672	413	6.00E-74	[LN:KCY_BACSU] [AC:P38493] [GN:CMK:JOFC] [OR:Bacillus subtilis] [EC:2.7.4.14] [DE:(CMP KINASE)] [SP:P38493]	107
SPX0613	613	3274	113	339	432	5.60E-55	[LN:PHNA_ECOLI] [AC:P16680] [GN:PHNA] [OR:Escherichia coli] [DE:PHNA PROTEIN] [SP:P16680]	89
SPX0614	614	3275	210	630			NO-HIT	6
SPX0615	615	3276	76	228	192	3.60E-23	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX0616	616	3277	384	1152	68	6.80E-05	[LN:C81266] [AC:C81266] [PN:probable efflux protein Cj1687 [imported]] [GN:Cj1687] [OR:Campylobacter jejuni]	108
SPX0617	617	3278	250	750	309	4.30E-69	[LN:TRUA_BACSU] [AC:P70973] [GN:TRUA] [OR:Bacillus subtilis] [EC:4.2.1.70] [DE:I) (PSEUDOURIDINE SYNTHASE I) (URACIL HYDROLYASE)] [SP:P70973]	141
SPX0618	618	3279	258	774	155	3.70E-39	[LN:THID_BACSU] [AC:P39610] [GN:THID:IPA-52R] [OR:Bacillus subtilis] [EC:2.7.4.7] [DE:(HMP-P KINASE)] [SP:P39610]	113

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0619	619	3280	154	462	69	5.80E-12	[LN:G75153] [AC:G75153] [PN:hypothetical protein PAB2090] [GN:PAB2090] [OR:Pyrococcus abyssi]	93
SPX0620	620	3281	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX0621	621	3282	91	273	288	2.30E-35	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0622	622	3283	198	594	583	9.30E-95	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX0623	623	3284	159	477			NO-HIT	6
SPX0624	624	3285	361	1083	1016	1.20E-136	[GI:2323341] [LN:AF014460] [AC:AF014460] [PN:PepQ] [FN:hydrolysis of Leu-Pro] [OR:Streptococcus mutans]	103
SPX0625	625	3286	261	783	147	5.10E-36	[LN:T31439] [AC:T31439] [PN:probable cobyric acid synthase CobQ] [OR:Heliobacillus mobilis]	91
SPX0626	626	3287	448	1344	137	1.30E-46	[LN:T31440] [AC:T31440] [PN:UDP-N-acetylmuramyl tripeptide synthetase homolog murC] [OR:Heliobacillus mobilis]	110
SPX0627	627	3288	439	1317	797	4.80E-136	[LN:YKGC_ECOLI] [AC:P77212] [GN:YKGC] [OR:Escherichia coli] [DE:INTERGENIC REGION] [SP:P77212]	94
SPX0628	628	3289	291	873	711	3.70E-137	[GI:7107009] [LN:AF168363] [AC:AF168363] [PN:oxalate:formate antiporter] [OR:Lactococcus lactis]	96
SPX0629	629	3290	82	246	193	5.00E-21	[GI:7107009] [LN:AF168363] [AC:AF168363] [PN:oxalate:formate antiporter] [OR:Lactococcus lactis]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0630	630	3291	525	1575	1329	1.80E-184	[GI:4409804] [LN:AF091502] [AC:AF091502] [PN:autoaggregation-mediating protein] [GN:aggH] [OR:Lactobacillus reuteri]	116
SPX0631	631	3292	263	789	547	3.10E-74	[LN:CODY__BACSU] [AC:P39779] [GN:CODY] [OR:Bacillus subtilis] [DE:CODY PROTEIN (VEGETATIVE PROTEIN 286B) (VEG286B)] [SP:P39779]	126
SPX0632	632	3293	69	207	174	1.30E-18	[LN:C70008] [AC:C70008] [PN:pyrazinamidase/nicotinamidase homolog yueJ] [GN:yueJ] [CL:hypothetical protein b1011] [OR:Bacillus subtilis]	136
SPX0633	633	3294	152	456	253	1.50E-30	[LN:C70008] [AC:C70008] [PN:pyrazinamidase/nicotinamidase homolog yueJ] [GN:yueJ] [CL:hypothetical protein b1011] [OR:Bacillus subtilis]	136
SPX0634	634	3295	78	234	328	1.70E-40	[GI:600072] [LN:SEDEXB] [AC:X72832] [PN:ABC transporter] [GN:abc] [OR:Streptococcus equisimilis]	96
SPX0635	635	3296	52	156	113	2.50E-10	[LN:E72756] [AC:E72756] [PN:hypothetical protein APE0042] [GN:APE0042] [OR:Aeropyrum pernix]	92
SPX0636	636	3297	285	855	731	1.70E-136	[LN:MSMK__STRMU] [AC:Q00752] [GN:MSMK] [OR:Streptococcus mutans] [DE:MULTIPLE SUGAR-BINDING TRANSPORT ATP-BINDING PROTEIN MSMK] [SP:Q00752]	138
SPX0637	637	3298	253	759	485	3.90E-95	[LN:Y095__HAEIN] [AC:Q57060:O05007] [GN:HI0095] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI0095] [SP:Q57060:O05007]	126
SPX0638	638	3299	177	531	603	2.90E-79	“[LN:B69587] [AC:B69587] [PN:adenine phosphoribosyltransferase,:AMP pyrophosphorylase:transphosphoribosidase]	194

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:apt] [CL:adenine phosphoribosyltransferase] [OR:Bacillus subtilis] [EC:2.4.2.7]”	
SPX0639	639	3300	315	945	567	6.90E-116	[LN:C72324] [AC:C72324] [PN:homoserine O-succinyltransferase] [GN:TM0881] [CL:homoserine succinyltransferase] [OR:Thermotoga maritima]	134
SPX0640	640	3301	226	678	357	5.50E-90	[GI:1813342] [LN:D78182] [AC:D78182] [GN:ORF2] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	114
SPX0641	641	3302	125	375			NO-HIT	6
SPX0642	642	3303	253	759	864	1.30E-113	“[LN:TPIS_LACDE] [AC:O32757] [GN:TPIA:TPI] [OR:Lactobacillus delbrueckii] [SR:,subspbulgaricus] [EC:5.3.1.1] [DE:TRIOSEPHOSPHATE ISOMERASE, (TIM)] [SP:O32757]”	159
SPX0643	643	3304	502	1506	2728	0	“[GI:4218526] [LN:SPAJ9639] [AC:AJ009639] [PN:1,4-beta-N-acetylmuramidase] [GN:lytC] [FN:lysis of cell wall peptidoglycan] [OR:Streptococcus pneumoniae]”	153
SPX0644	644	3305	173	519	800	2.60E-105	[GI:3513547] [LN:AF055720] [AC:AF055720] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0645	645	3306	213	639			NO-HIT	6
SPX0646	646	3307	169	507	882	2.00E-118	[GI:2196662] [LN:HSZ84379] [AC:Z84379] [PN:dihydrofolate reductase] [GN:dfr] [FN:trimethoprim resistance] [OR:Streptococcus pneumoniae]	135
SPX0647	647	3308	69	207			NO-HIT	6
SPX0648	648	3309	411	1233	1251	4.80E-206	[GI:7546983] [LN:AF236863] [AC:AF236863] [PN:protease ClpX]	93

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:clpX] [OR:Lactococcus lactis]	
SPX0649	649	3310	196	588	807	9.70E-107	[GI:7546984] [LN:AF236863] [AC:AF236863] [PN:hypothetical GTP-binding protein] [OR:Lactococcus lactis]	102
SPX0650	650	3311	127	381	392	2.30E-49	"[LN:ALDR_LACLA] [AC:O34133] [GN:ALDR] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:PUTATIVE REGULATOR ALDR] [SP:O34133]"	143
SPX0651	651	3312	297	891	337	2.00E-91	[GI:7328270] [LN:SAY14324] [AC:Y14324] [PN:hypothetical protein] [OR:Staphylococcus aureus]	91
SPX0652	652	3313	326	978	268	1.30E-62	[LN:YVCK_BACSU] [AC:O06974] [GN:YVCK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 34.7 KD PROTEIN IN CRH-TRXB INTERGENIC REGION] [SP:O06974]	136
SPX0653	653	3314	304	912	424	5.50E-81	[GI:7328272] [LN:SAY14324] [AC:Y14324] [OR:Staphylococcus aureus]	65
SPX0654	654	3315	323	969	422	5.00E-99	[LN:B70015] [AC:B70015] [PN:thioredoxin reductase homolog yumC] [GN:yumC] [CL:thioredoxin reductase:thioredoxin reductase homology] [OR:Bacillus subtilis]	154
SPX0655	655	3316	272	816	419	8.50E-81	[LN:H69744] [AC:H69744] [PN:conserved hypothetical protein ybbP] [GN:ybbP] [CL:hypothetical protein ybbP] [OR:Bacillus subtilis]	128
SPX0656	656	3317	260	780	104	2.70E-10	[LN:A69745] [AC:A69745] [PN:hypothetical protein ybbR] [GN:ybbR] [OR:Bacillus subtilis]	87
SPX0657	657	3318	64	192			NO-HIT	6
SPX0658	658	3319	451	1353	655	2.00E-180	[GI:3892895] [LN:SAARGFEMD] [AC:Y15477]	107

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:phosphoglucosamine-mutase] [GN:glmM] [OR:Staphylococcus aureus]	
SPX0659	659	3320	121	363			NO-HIT	6
SPX0660	660	3321	283	849	239	6.50E-25	[LN:DEGV_BACSU] [AC:P32436] [GN:DEGV] [OR:Bacillus subtilis] [DE:DEGV PROTEIN] [SP:P32436]	90
SPX0661	661	3322	256	768	614	6.50E-89	"[LN:DAPB_BACSU] [AC:P42976] [GN:DAPB] [OR:Bacillus subtilis] [EC:1.3.1.26] [DE:DIHYDRODIPICOLINATE REDUCTASE,] [SP:P42976]"	124
SPX0662	662	3323	400	1200	601	8.40E-92	"[LN:PAPS_BACSU] [AC:P42977] [GN:PAPS] [OR:Bacillus subtilis] [EC:2.7.7.19] [DE:POLY(A) POLYMERASE, (PAP)] [SP:P42977]"	119
SPX0663	663	3324	624	1872	669	1.80E-162	[LN:A69814] [AC:A69814] [PN:ABC transporter (ATP-binding protein) homolog yfmR] [GN:yfmR] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0664	664	3325	144	432			NO-HIT	6
SPX0665	665	3326	394	1182	629	1.90E-82	[LN:YEAB_BACSU] [AC:P46348:O05001] [GN:YEAB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 31.8 KD PROTEIN IN GABP-GUAA INTERGENIC REGION (ORFX)] [SP:P46348:O05001]	158
SPX0666	666	3327	899	2697	601	1.20E-189	"[LN:S77052] [AC:S77052] [PN:cation-transporting ATPase, pacL-1:protein sll0672:protein sll0672] [GN:pacL-1] [CL:Na+/K+-transporting ATPase alpha chain:ATPase nucleotide-binding domain homology] [OR:Synechocystis sp.]"	218
SPX0667	667	3328	97	291			NO-HIT	6
SPX0668	668	3329	83	249	73	0.00016	[LN:H72624] [AC:H72624] [PN:hypothetical protein APE1456] [GN:APE1456] [OR:Aeropyrum pernix]	92
SPX0669	669	3330	264	792	713	2.30E-105	[LN:YGHU_ECOLI]	135

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Q46845] [GN:YGHU] [OR:Escherichia coli] [DE:HYPOTHETICAL 34.2 KD PROTEIN IN GSP-HYBG INTERGENIC REGION] [SP:Q46845]	
SPX0670	670	3331	152	456	368	3.30E-46	“[LN:DEF_CLOBE] [AC:O08450] [GN:DEF:FMS] [OR:Clostridium beijerinckii] [SR:,Clostridium MP] [EC:3.5.1.31] [DE:DEFORMYLASE)] [SP:O08450]”	136
SPX0671	671	3332	532	1596			NO-HIT	6
SPX0672	672	3333	60	180			NO-HIT	6
SPX0673	673	3334	243	729	473	1.20E-60	[GI:4580623] [LN:AF118389] [AC:AF118389] [PN:unknown] [OR:Streptococcus suis]	77
SPX0674	674	3335	210	630	111	2.10E-12	[GI:6562870] [LN:SCM1] [AC:AL133422] [PN:putative secreted protein.] [GN:SCM1.21] [OR:Streptomyces coelicolor A3(2)]	116
SPX0675	675	3336	163	489	68	0.00021	[LN:YPMB_BACSU] [AC:P54396] [GN:YPMB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 17.9 KD PROTEIN IN DING-ASPB INTERGENIC REGION] [SP:P54396]	137
SPX0676	676	3337	396	1188	1400	2.80E-187	[GI:6465901] [LN:AF035157] [AC:AF035157] [PN:aspartate aminotransferase] [GN:aspC] [OR:Lactococcus lactis]	106
SPX0677	677	3338	124	372	178	1.90E-27	[LN:T03486] [AC:T03486] [PN:conserved hypothetical protein] [OR:Rhodobacter capsulatus]	87
SPX0678	678	3339	448	1344	703	1.60E-180	[LN:SYN_BACSU] [AC:P39772] [GN:ASNS] [OR:Bacillus subtilis] [EC:6.1.1.22] [DE:(ASNRS)] [SP:P39772]	98
SPX0679	679	3340	129	387			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0680	680	3341	97	291	236	7.30E-39	[LN:RS6_BACSU] [AC:P21468] [GN:RPSF] [OR:Bacillus subtilis] [DE:30S RIBOSOMAL PROTEIN S6 (BS9)] [SP:P21468]	107
SPX0681	681	3342	157	471	481	5.20E-71	[GI:6716352] [LN:AF145054] [AC:AF145054:AF001793:AF118440:U89246] [PN:ORF9] [GN:orf9] [OR:Streptococcus thermophilus bacteriophage 7201]	136
SPX0682	682	3343	87	261	134	1.30E-13	[LN:RS18_BACST] [AC:P10806] [GN:RPSR] [OR:Bacillus stearothermophilus] [DE:30S RIBOSOMAL PROTEIN S18 (BS21)] [SP:P10806]	120
SPX0683	683	3344	112	336			NO-HIT	6
SPX0684	684	3345	467	1401	270	7.50E-47	[GI:4678225] [LN:AC007135] [AC:AC007135:AE002093] [PN:cyclophilin-like protein] [GN:At2g36130] [OR:Arabidopsis thaliana] [SR:thale cress]	137
SPX0685	685	3346	84	252	172	6.10E-19	[LN:YABR_BACSU] [AC:P37560] [GN:YABR] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 14.2 KD PROTEIN IN DIVIC-SPOIIE INTERGENIC REGION] [SP:P37560]	140
SPX0686	686	3347	87	261	96	1.40E-15	[LN:YABB_BACSU] [AC:P37543] [GN:YABB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 28.3 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION] [SP:P37543]	137
SPX0687	687	3348	67	201	109	3.00E-09	[LN:YABB_BACSU] [AC:P37543] [GN:YABB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 28.3 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION] [SP:P37543]	137
SPX0688	688	3349	129	387	267	1.40E-39	[LN:YABB_BACSU] [AC:P37543] [GN:YABB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 28.3 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION] [SP:P37543]	137

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0689	689	3350	93	279	224	1.50E-26	[LN:A69742] [AC:A69742] [PN:conserved hypothetical protein yazA] [GN:yazA] [CL:hypothetical protein 312] [OR:Bacillus subtilis]	127
SPX0690	690	3351	60	180	143	5.60E-14	[GI:1743856] [LN:SGU57759] [AC:U57759] [PN:intrageneric coaggregation-relevant adhesin] [OR:Streptococcus gordonii]	115
SPX0691	691	3352	222	666	955	2.60E-126	[GI:1743856] [LN:SGU57759] [AC:U57759] [PN:intrageneric coaggregation-relevant adhesin] [OR:Streptococcus gordonii]	115
SPX0692	692	3353	171	513			NO-HIT	6
SPX0693	693	3354	105	315	140	1.80E-21	[GI:2952534] [LN:AF051356] [AC:AF051356] [PN:unknown] [OR:Streptococcus mutans]	79
SPX0694	694	3355	95	285			NO-HIT	6
SPX0695	695	3356	65	195	74	0.00038	[GI:2772940] [LN:AF034574] [AC:AF034574] [PN:putative cruciform DNA binding protein] [GN:Gv1] [OR:Glomus versiforme]	116
SPX0696	696	3357	504	1512	172	1.50E-35	[LN:MURE_BACSU] [AC:Q03523] [GN:MURE] [OR:Bacillus subtilis] [EC:6.3.2.13] [DE:DIAMINOPIMELATE-ADDING ENZYME)] [SP:Q03523]	122
SPX0697	697	3358	192	576	172	9.40E-16	[LN:MURE_BACSU] [AC:Q03523] [GN:MURE] [OR:Bacillus subtilis] [EC:6.3.2.13] [DE:DIAMINOPIMELATE-ADDING ENZYME)] [SP:Q03523]	122
SPX0698	698	3359	541	1623	187	2.90E-75	[LN:G69992] [AC:G69992] [PN:spore cortex protein homolog ytgP] [GN:ytgP] [OR:Bacillus subtilis]	95
SPX0699	699	3360	653	1959	3335	0	[LN:ALIB_STRPN] [AC:Q51933]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0700	700	3361	185	555	282	2.00E-38	[GN:ALIB] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE-BINDING PROTEIN ALIB PRECURSOR] [SP:Q51933] [GI:1125685] [LN:SADNAS55] [AC:X87104]	104
SPX0701	701	3362	180	540	179	9.60E-18	[GN:mdr] [FN:multiple grug resistance] [OR:Staphylococcus aureus] [LN:S58356] [AC:S66651:S58356] [PN:pepT protein] [GN:pepT] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology]	166
SPX0702	702	3363	182	546	444	4.80E-55	[OR:Staphylococcus epidermidis] [GI:1262136] [LN:SAPBP4GEN] [AC:X91786] [PN:ATP-binding cassette transporter A] [GN:abcA]	116
SPX0703	703	3364	365	1095	681	2.80E-123	[OR:Staphylococcus aureus] [LN:A69847] [AC:A69847] [PN:cystathionine gamma-synthase homolog yjcI] [GN:yjcI] [CL:O-succinylhomoserine (thiol)-lyase]	143
SPX0704	704	3365	389	1167	582	1.60E-99	[OR:Bacillus subtilis] “[LN:PATB_BACSU] [AC:Q08432] [GN:PATB] [OR:Bacillus subtilis] [EC:2.6.1.-] [DE:PUTATIVE AMINOTRANSFERASE B,] [SP:Q08432]”	121
SPX0705	705	3366	1033	3099	552	1.00E-129	[GI:1769947] [LN:BCX98455] [AC:X98455] [GN:SNF] [OR:Bacillus cereus]	68
SPX0706	706	3367	206	618			NO-HIT	6
SPX0707	707	3368	445	1335	491	2.00E-154	[LN:MURC_BACSU] [AC:P40778] [GN:MURC] [OR:Bacillus subtilis] [EC:6.3.2.8] [DE:ACETYLMURANOYL-L-ALANINE SYNTHETASE)] [SP:P40778]	127
SPX0708	708	3369	155	465			NO-HIT	6
SPX0709	709	3370	119	357	56	7.70E-05	“[LN:A64491]	192

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:A64491] [PN:N-terminal acetyltransferase complex, subunit ARD1 homolog] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimI] [OR:Methanococcus jannaschii]”	
SPX0710	710	3371	71	213			NO-HIT	6
SPX0711	711	3372	502	1506	185	2.10E-43	[LN:E69979] [AC:E69979] [PN:folate metabolism homolog yrrL] [GN:yrrL] [CL:yceG protein] [OR:Bacillus subtilis]	110
SPX0712	712	3373	161	483	271	3.60E-55	[LN:GREA_BACSU] [AC:P80240] [GN:GREA] [OR:Bacillus subtilis] [DE:GREA) (GENERAL STRESS PROTEIN 20M) (GSP20M)] [SP:P80240]	121
SPX0713	713	3374	61	183	121	2.50E-12	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0714	714	3375	163	489	152	1.50E-24	[LN:S31638] [AC:S31638] [PN:hypothetical protein] [OR:Lactobacillus curvatus]	77
SPX0715	715	3376	66	198			NO-HIT	6
SPX0716	716	3377	70	210	89	0.0001	[GI:4966270] [LN:CELK09H11] [AC:U97002] [GN:K09H11.1] [OR:Caenorhabditis elegans]	81
SPX0717	717	3378	148	444	419	7.50E-56	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX0718	718	3379	164	492	484	1.40E-63	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX0719	719	3380	67	201	331	3.00E-44	“[LN:S49404] [AC:S49404:S38206] [PN:H+-transporting ATP synthase, chain c] [GN:atpC] [CL:H+-transporting ATP synthase lipid-binding protein] [OR:Streptococcus pneumoniae] [EC:3.6.1.34]”	185
SPX0720	720	3381	239	717	1205	5.00E-168	“[LN:ATP6_STRPN] [AC:Q59954] [GN:ATPB:ATPA] [OR:Streptococcus pneumoniae]	139

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0721	721	3382	165	495	769	2.90E-100	[EC:3.6.1.34] [DE:ATP SYNTHASE A CHAIN, (PROTEIN 6)] [SP:Q59954]" "[LN:ATPF_STRPN] [AC:Q59952:Q59955] [GN:ATPF:ATPB] [OR:Streptococcus pneumoniae]	153
SPX0722	722	3383	179	537	509	8.40E-66	[EC:3.6.1.34] [DE:ATP SYNTHASE B CHAIN, (SUBUNIT I)] [SP:Q59952:Q59955]" [GI:4100654] [LN:AF001955] [AC:AF001955] [PN:proton-translocating ATPase delta subunit] [GN:uncH] [OR:Streptococcus sanguinis]	126
SPX0723	723	3384	502	1506	2347	0	[GI:4100655] [LN:AF001955] [AC:AF001955] [PN:proton-translocating ATPase alpha subunit] [GN:uncA] [OR:Streptococcus sanguinis]	126
SPX0724	724	3385	293	879	1011	1.70E-169	[GI:4100656] [LN:AF001955] [AC:AF001955] [PN:proton-translocating ATPase gamma subunit] [OR:Streptococcus sanguinis]	116
SPX0725	725	3386	469	1407	2287	0	[GI:4100657] [LN:AF001955] [AC:AF001955] [PN:proton-translocating ATPase beta subunit] [GN:uncD] [OR:Streptococcus sanguinis]	125
SPX0726	726	3387	140	420	581	1.40E-76	[GI:4100658] [LN:AF001955] [AC:AF001955] [PN:proton-translocating ATPase epsilon subunit] [GN:uncC] [OR:Streptococcus sanguinis]	128
SPX0727	727	3388	251	753	175	2.70E-42	[GI:6746427] [LN:AF179847] [AC:AF179847] [PN:putative transposase] [OR:Lactococcus lactis]	90
SPX0728	728	3389	65	195			NO-HIT	6
SPX0729	729	3390	182	546	107	1.20E-15	[GI:3849798] [LN:U91581] [AC:U91581:U04057] [PN:putative transposase] [GN:tpase]	118

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Lactococcus lactis subsp. lactis]	
SPX0730	730	3391	338	1014	444	1.50E-117	[GI:3242228] [LN:LLCADHE] [AC:AJ001007] [GN:orfB] [OR:Lactococcus lactis]	73
SPX0731	731	3392	78	234			NO-HIT	6
SPX0732	732	3393	389	1167	1141	1.60E-158	[GI:4098497] [LN:SMU78604] [AC:U78604] [PN:putative membrane protein] [OR:Streptococcus mutans]	95
SPX0733	733	3394	410	1230	457	6.20E-71	“[GI:473901] [LN:LACALS] [AC:L16975] [OR:Lactococcus lactis] [SR:Lactococcus lactis (strain DSM 20384, sub__species lactis) DNA]”	128
SPX0734	734	3395	214	642	351	1.70E-45	“[LN:A72357] [AC:A72357] [PN:amino acid ABC transporter, permease protein] [GN:TM0592] [CL:histidine permease protein M] [OR:Thermotoga maritima]”	146
SPX0735	735	3396	210	630	367	2.90E-65	[GI:1649037] [LN:STU73111] [AC:U73111] [PN:glutamine transport ATP-binding protein GLNQ] [OR:Salmonella typhimurium]	116
SPX0736	736	3397	279	837	1411	4.70E-188	[GI:5929889] [LN:AF165218] [AC:AF165218] [PN:AatB] [GN:aatB] [OR:Streptococcus pneumoniae]	90
SPX0737	737	3398	116	348	578	1.80E-76	[GI:5929888] [LN:AF165218] [AC:AF165218] [PN:Bta] [GN:bta] [OR:Streptococcus pneumoniae]	88
SPX0738	738	3399	573	1719	2903	0	[GI:5929887] [LN:AF165218] [AC:AF165218] [PN:Pgm] [GN:pgm] [OR:Streptococcus pneumoniae]	88
SPX0739	739	3400	109	327	129	8.30E-12	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX0740	740	3401	100	300	245	5.10E-30	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX0741	741	3402	95	285			NO-HIT	6
SPX0742	742	3403	85	255			NO-HIT	6
SPX0743	743	3404	71	213			NO-HIT	6
SPX0744	744	3405	371	1113			NO-HIT	6
SPX0745	745	3406	174	522	72	1.10E-05	“[LN:I38170] [AC:I38170] [PN:gene hr44 protein] [GN:hr44] [OR:Homo sapiens] [SR:, man]”	87
SPX0746	746	3407	290	870	335	2.40E-72	[LN:B72254] [AC:B72254] [PN:glycerol uptake facilitator protein] [GN:TM1429] [CL:glycerol facilitator protein] [OR:Thermotoga maritima]	135
SPX0747	747	3408	399	1197	2022	6.40E-273	[LN:EFTU_STROR] [AC:P33170] [GN:TUF] [OR:Streptococcus oralis] [DE:ELONGATION FACTOR TU (EF-TU)] [SP:P33170]	108
SPX0748	748	3409	128	384	104	1.80E-06	[LN:H71023] [AC:H71023] [PN:hypothetical protein PH1485] [GN:PH1485] [OR:Pyrococcus horikoshii]	95
SPX0749	749	3410	138	414	99	6.40E-06	[LN:S31840] [AC:S31840] [PN:probable transposase] [OR:Bacillus stearothermophilus]	82
SPX0750	750	3411	62	186			NO-HIT	6
SPX0751	751	3412	218	654	163	3.90E-21	[GI:722339] [LN:AXU22323] [AC:U22323] [PN:unknown] [OR:Acetobacter xylinus] [SR:Acetobacter xylinum]	100
SPX0752	752	3413	361	1083	272	2.20E-72	[LN:C69813] [AC:C69813] [PN:RNA helicase homolog yfmL] [GN:yfmL] [CL:unassigned DEAD/H box helicases:DEAD/H box helicase homology] [OR:Bacillus subtilis]	153

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0753	753	3414	354	1062	197	3.10E-44	[LN:YGJR_ECOLI] [AC:P42599:P42600:P76661] [GN:YGJR] [OR:Escherichia coli] [DE:HYPOTHETICAL 36.2 KD PROTEIN IN EBGC-UXAA INTERGENIC REGION] [SP:P42599:P42600:P76661]	164
SPX0754	754	3415	62	186	77	0.00028	"[LN:S67482] [AC:S67482:S52150] [PN:serine O-acetyltransferase,, cytosolic:serine acetyltransferase:serine acetyltransferase] [OR:Arabidopsis thaliana]"	152
SPX0755	755	3416	126	378			NO-HIT	6
SPX0756	756	3417	481	1443	485	5.60E-59	"[LN:G69849] [AC:G69849] [PN:endo-1,4-beta-xylanase homolog yjeA] [GN:yjeA] [CL:nodB homology] [OR:Bacillus subtilis]"	118
SPX0757	757	3418	256	768	357	5.70E-83	[LN:C70040] [AC:C70040] [PN:plant-metabolite dehydrogenase homolog yvgN] [GN:yvgN] [CL:aldehyde reductase] [OR:Bacillus subtilis]	129
SPX0758	758	3419	77	231			NO-HIT	6
SPX0759	759	3420	254	762			NO-HIT	6
SPX0760	760	3421	306	918	1187	3.50E-159	[LN:SYGA_BACSU] [AC:P54380] [GN:GLYQ] [OR:Bacillus subtilis] [EC:6.1.1.14] [DE:ALPHA CHAIN) (GLYRS)] [SP:P54380]	112
SPX0761	761	3422	679	2037	1334	6.50E-179	[LN:SYGB_BACSU] [AC:P54381] [GN:GLYS] [OR:Bacillus subtilis] [EC:6.1.1.14] [DE:BETA CHAIN) (GLYRS)] [SP:P54381]	111
SPX0762	762	3423	86	258	218	1.20E-24	[LN:E69894] [AC:E69894] [PN:hypothetical protein ynzC] [GN:ynzC] [OR:Bacillus subtilis]	87
SPX0763	763	3424	415	1245	97	5.10E-08	[LN:YIEF_ECOLI] [AC:P31465] [GN:YIEF] [OR:Escherichia coli] [DE:HYPOTHETICAL 20.4 KD PROTEIN IN TNAB-BGLB INTERGENIC REGION] [SP:P31465]	136

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0764	764	3425	79	237			NO-HIT	6
SPX0765	765	3426	202	606	110	2.80E-13	[LN:YIEF_ECOLI] [AC:P31465] [GN:YIEF] [OR:Escherichia coli] [DE:HYPOTHETICAL 20.4 KD PROTEIN IN TNAB-BGLB INTERGENIC REGION] [SP:P31465]	136
SPX0766	766	3427	308	924	172	3.10E-41	[LN:APBE_TREPA] [AC:O83774] [GN:APBE:TP0796] [OR:Treponema pallidum] [DE:THIAMINE BIOSYNTHESIS LIPOPROTEIN APBE PRECURSOR] [SP:O83774]	134
SPX0767	767	3428	91	273			NO-HIT	6
SPX0768	768	3429	460	1380	2342	0	[GI:4416519] [LN:AF014458] [AC:AF014458] [PN:NADH oxidase] [OR:Streptococcus pneumoniae]	88
SPX0769	769	3430	298	894	1411	3.40E-191	[LN:YG47_HAEIN] [AC:P45293] [GN:HI1647] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI1647] [SP:P45293]	112
SPX0770	770	3431	78	234	125	3.00E-12	[LN:F71007] [AC:F71007] [PN:hypothetical protein PH1356] [GN:PH1356] [CL:Pyrococcus horikoshii hypothetical protein PH1356] [OR:Pyrococcus horikoshii]	150
SPX0771	771	3432	63	189	157	1.20E-15	[LN:F72782] [AC:F72782] [PN:hypothetical protein APE0247] [GN:APE0247] [OR:Aeropyrum pernix]	92
SPX0772	772	3433	52	156	159	4.30E-16	[LN:F72782] [AC:F72782] [PN:hypothetical protein APE0247] [GN:APE0247] [OR:Aeropyrum pernix]	92
SPX0773	773	3434	194	582	720	2.40E-95	[LN:YG48_HAEIN] [AC:P45294] [GN:HI1648] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI1648] [SP:P45294]	112
SPX0774	774	3435	216	648	183	5.00E-34	[LN:YA37_TREPA] [AC:O84000] [GN:TP1037]	108

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Treponema pallidum] [DE:HYPOTHETICAL PROTEIN TP1037] [SP:O84000]	
SPX0775	775	3436	151	453			NO-HIT	6
SPX0776	776	3437	165	495	94	2.80E-12	[LN:G81269] [AC:G81269] [PN:probable acetyltransferase Cj1715 [imported]] [GN:Cj1715] [OR:Campylobacter jejuni]	111
SPX0777	777	3438	177	531	167	1.80E-29	[GI:3694956] [LN:AF091508] [AC:AF091508] [PN:O-6-methylguanine DNA methyltransferase] [GN:ogt] [OR:Salmonella muenster]	119
SPX0778	778	3439	119	357	272	2.10E-32	[LN:B70021] [AC:B70021] [PN:arsenate reductase homolog yusI] [GN:yusI] [CL:hypothetical protein yjbD] [OR:Bacillus subtilis]	124
SPX0779	779	3440	126	378			NO-HIT	6
SPX0780	780	3441	267	801	284	2.10E-54	[LN:YXEN_BACSU] [AC:P54953] [GN:YXEN:LP9F] [OR:Bacillus subtilis] [DE:INTERGENIC REGION] [SP:P54953]	100
SPX0781	781	3442	248	744	468	5.80E-82	[GI:666983] [LN:BSPAAT] [AC:X77636] [PN:putative ATP binding subunit] [GN:ORF3] [OR:Bacillus subtilis]	102
SPX0782	782	3443	79	237			NO-HIT	6
SPX0783	783	3444	74	222			NO-HIT	6
SPX0784	784	3445	266	798	542	4.60E-98	“[LN:TRXB_LISMO] [AC:O32823] [GN:TRXB] [OR:Listeria monocytogenes] [EC:1.6.4.5] [DE:THIOREDOXIN REDUCTASE,] [SP:O32823]”	120
SPX0785	785	3446	254	762	322	1.30E-58	“[LN:YACO_BACSU] [AC:Q06753] [GN:YACO] [OR:Bacillus subtilis] [EC:2.1.1.-] [DE:HYPOTHETICAL TRNA/RRNA METHYLTRANSFERASE YACO,] [SP:Q06753]”	139

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0786	786	3447	204	612	394	4.10E-93	“[GI:806487] [LN:LACLPAGAP] [AC:L36907] [FN:unknown] [OR:Lactococcus lactis] [SR:Lactococcus lactis (individual__isolate LM0230, sub__specie]”	140
SPX0787	787	3448	84	252			NO-HIT	6
SPX0788	788	3449	161	483	136	6.00E-14	[LN:H70069] [AC:H70069:JC7099] [PN:poly-gamma-glutamic synthesis PgsA protein] [GN:ywtB:pgsA] [OR:Bacillus subtilis]	116
SPX0789	789	3450	126	378			NO-HIT	6
SPX0790	790	3451	276	828	242	8.80E-42	[LN:F64819] [AC:F64819] [PN:hypothetical protein b0822] [CL:Methanobacterium thermoautotrophicum conserved hypothetical protein MTH1071] [OR:Escherichia coli]	158
SPX0791	791	3452	212	636			NO-HIT	6
SPX0792	792	3453	242	726	1196	1.80E-160	[GI:5001711] [LN:AF112358] [AC:AF112358] [PN:C3-degrading proteinase] [GN:cppA] [OR:Streptococcus pneumoniae]	109
SPX0793	793	3454	312	936	117	4.00E-22	[GI:6469268] [LN:SCC75A] [AC:AL133220] [PN:possible secreted esterase] [GN:SCC75A.29c] [OR:Streptomyces coelicolor A3(2)]	121
SPX0794	794	3455	299	897	487	1.00E-64	[LN:YICL__ECOLI] [AC:P31437] [GN:YICL] [OR:Escherichia coli] [DE:HYPOTHETICAL 33.1 KD PROTEIN IN SELC-NLPA INTERGENIC REGION] [SP:P31437]	136
SPX0795	795	3456	233	699	294	7.80E-49	[LN:D70044] [AC:D70044] [PN:transcription regulator GntR family homolog yvoA] [GN:yvoA] [CL:transcription regulator GntR] [OR:Bacillus subtilis]	144
SPX0796	796	3457	521	1563	1721	3.40E-294	[GI:4321715] [LN:AF058326] [AC:AF058326] [PN:GMP synthase] [GN:guaA] [OR:Lactococcus lactis]	92
SPX0797	797	3458	62	186			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0798	798	3459	96	288	204	2.90E-23	"[LN:JC1151] [AC:JC1151] [PN:hypothetical protein, 20.3K] [OR:Agrobacterium tumefaciens]"	89
SPX0799	799	3460	65	195			NO-HIT	6
SPX0800	800	3461	130	390	91	0.00062	"[GI:5824139] [LN:POL245436] [AC:AJ245436:J04618:J04619:S50571:X52935:X65936] [PN:hypothetical protein, 57.8 kD] [OR:Pseudomonas putida]"	137
SPX0801	801	3462	143	429	84	9.20E-08	"[GI:6009407] [LN:AB024946] [AC:AB024946] [GN:orf31] [OR:Escherichia coli] [SR:Escherichia coli (sub__species:enteropathogenic, strain:B171]"	140
SPX0802	802	3463	150	450	94	3.20E-15	"[LN:Y4HP__RHISN] [AC:P50360] [GN:Y4HP] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 61.7 KD PROTEIN Y4HP] [SP:P50360]"	128
SPX0803	803	3464	78	234	87	0.00033	"[GI:5824139] [LN:POL245436] [AC:AJ245436:J04618:J04619:S50571:X52935:X65936] [PN:hypothetical protein, 57.8 kD] [OR:Pseudomonas putida]"	137
SPX0804	804	3465	455	1365	253	2.70E-62	"[LN:D69159] [AC:D69159] [PN:methyl coenzyme M reductase system, component A2 homolog] [GN:MTH454] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Methanobacterium thermoautotrophicum]"	217
SPX0805	805	3466	74	222			NO-HIT	6
SPX0806	806	3467	232	696	88	1.90E-07	[LN:E72202] [AC:E72202] [PN:conserved hypothetical protein] [GN:TM1868] [OR:Thermotoga maritima]	96
SPX0807	807	3468	195	585			NO-HIT	6
SPX0808	808	3469	582	1746	554	3.00E-83	[GI:4097162] [LN:PMU46488] [AC:U46488] [PN:NrpB] [GN:nrpB] [OR:Proteus mirabilis]	81
SPX0809	809	3470	587	1761	526	2.60E-78	[GI:4097161] [LN:PMU46488] [AC:U46488]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0810	810	3471	159	477	131	3.30E-12	[PN:NrpA] [GN:nrpA] [OR:Proteus mirabilis] [LN:SOXS_ECOLI] [AC:P22539] [GN:SOXS] [OR:Escherichia coli] [DE:REGULATORY PROTEIN SOXS] [SP:P22539]	100
SPX0811	811	3472	69	207			NO-HIT	6
SPX0812	812	3473	397	1191	889	2.70E-126	[GI:2995646] [LN:AF051092] [AC:AF051092] [PN:DNA modification methyltransferase M.XbaI] [GN:xbaIM] [FN:recognizes ds DNA sequence TCTAGA; thought to] [OR:Xanthomonas campestris]	177
SPX0813	813	3474	63	189	128	1.30E-17	[GI:2995645] [LN:AF051092] [AC:AF051092] [PN:restriction endonuclease R.XbaI] [GN:xbaIR] [FN:recognizes ds DNA sequence TCTAGA; cleaves both] [OR:Xanthomonas campestris]	169
SPX0814	814	3475	61	183	199	4.80E-22	[GI:2995645] [LN:AF051092] [AC:AF051092] [PN:restriction endonuclease R.XbaI] [GN:xbaIR] [FN:recognizes ds DNA sequence TCTAGA; cleaves both] [OR:Xanthomonas campestris]	169
SPX0815	815	3476	71	213			NO-HIT	6
SPX0816	816	3477	429	1287	681	4.70E-147	[LN:H69979] [AC:H69979] [PN:proteinase homolog yrrO] [GN:yrrO] [CL:collagenase prtC] [OR:Bacillus subtilis]	107
SPX0817	817	3478	243	729	251	6.20E-58	[LN:H69979] [AC:H69979] [PN:proteinase homolog yrrO] [GN:yrrO] [CL:collagenase prtC] [OR:Bacillus subtilis]	107
SPX0818	818	3479	241	723	225	4.40E-26	[LN:YCBN_BACSU] [AC:P42246] [GN:YCBN] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 31.7 KD PROTEIN IN GLTP-CWLJ INTERGENIC REGION (ORF13)] [SP:P42246]	145

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0819	819	3480	243	729			NO-HIT	6
SPX0820	820	3481	80	240	152	2.50E-16	[GI:1536960] [LN:SOORFS] [AC:Z79691] [GN:yorfE] [FN:putative transcription regulator] [OR:Streptococcus pneumoniae]	115
SPX0821	821	3482	163	489			NO-HIT	6
SPX0822	822	3483	487	1461	981	1.70E-215	[LN:D70008] [AC:D70008] [PN:nicotinate phosphoribosyltransferase homolog yueK] [GN:yueK] [OR:Bacillus subtilis]	111
SPX0823	823	3484	275	825	923	1.70E-122	[LN:NADE_ECOLI] [AC:P18843:P78235] [GN:NADE:EFG:NTRL] [OR:Escherichia coli] [EC:6.3.5.1] [DE:PROTEIN)] [SP:P18843:P78235]	121
SPX0824	824	3485	66	198			NO-HIT	6
SPX0825	825	3486	184	552	135	1.80E-13	[GI:7635982] [LN:SCE6] [AC:AL353832] [PN:putative acetyltransferase.] [GN:SCE6.13] [OR:Streptomyces coelicolor A3(2)]	117
SPX0826	826	3487	250	750	66	0.00023	[GI:4185565] [LN:AF115379] [AC:AF115379] [PN:surface protein Pls] [GN:pls] [OR:Staphylococcus aureus]	101
SPX0827	827	3488	130	390	306	1.70E-36	[LN:A41971] [AC:A41971:A60282:A33134] [PN:surface protein pspA precursor:pneumococcal surface protein A] [GN:pspA] [CL:cpl repeat homology] [OR:Streptococcus pneumoniae]	169
SPX0828	828	3489	347	1041	1229	5.30E-166	[LN:QUEA_BACSU] [AC:O32054] [GN:QUEA] [OR:Bacillus subtilis] [EC:5.-.-.-] [DE:(QUEUOSINE BIOSYNTHESIS PROTEIN QUEA)] [SP:O32054]	128
SPX0829	829	3490	236	708	336	3.00E-66	[LN:NAGB_BACSU] [AC:O35000] [GN:NAGB] [OR:Bacillus subtilis]	139

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[EC:5.3.1.10] [DE:PHOSPHATE DEAMINASE) (GNPDA) (GLCN6P DEAMINASE)] [SP:O35000]	
SPX0830	830	3491	67	201	248	1.00E-29	[LN:T43742] [AC:T43742] [PN:ribosomal protein S21 [imported]] [OR:Listeria monocytogenes]	89
SPX0831	831	3492	313	939	1224	7.60E-165	[GI:4884536] [LN:AB027460] [AC:AB027460] [PN:Hpr kinase] [OR:Streptococcus bovis] [SR:Streptococcus bovis (strain:JB1) DNA]	123
SPX0832	832	3493	192	576	608	1.30E-84	"[LN:T11569] [AC:T11569] [PN:prolipoprotein diacylglyceryl transferase,] [GN:lgt] [CL:prolipoprotein diacylglyceryl transferase] [OR:Streptococcus mutans] [EC:2.4.99.-]"	169
SPX0833	833	3494	128	384	333	1.50E-41	[LN:T11570] [AC:T11570] [PN:hypothetical protein 1] [OR:Streptococcus mutans]	77
SPX0834	834	3495	110	330			NO-HIT	6
SPX0835	835	3496	130	390	164	5.70E-17	[LN:T11571] [AC:T11571] [PN:hypothetical protein 2] [OR:Streptococcus mutans]	77
SPX0836	836	3497	377	1131	596	4.30E-126	"[LN:B69640] [AC:B69640] [PN:coproporphyrinogen oxidase, III, oxygen-independent hemN] [GN:hemN] [CL:oxygen-independent coproporphyrinogen oxidase] [OR:Bacillus subtilis] [EC:1.3.3.3]"	184
SPX0837	837	3498	166	498			NO-HIT	6
SPX0838	838	3499	246	738	96	7.20E-11	[GI:4704640] [LN:AF110462] [AC:AF110462] [PN:oleoyl-ACP thioesterase] [OR:Elaeis guineensis]	92
SPX0839	839	3500	203	609	466	3.10E-59	[LN:H70023] [AC:H70023] [PN:N-acetyl-glucosamine catabolism homolog yutF] [GN:yutF] [CL:nagD protein] [OR:Bacillus subtilis]	124
SPX0840	840	3501	217	651			NO-HIT	6
SPX0841	841	3502	108	324			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0842	842	3503	134	402	447	5.60E-58	[LN:D69843] [AC:D69843] [PN:conserved hypothetical protein yjbD] [GN:yjbD] [CL:hypothetical protein yjbD] [OR:Bacillus subtilis]	128
SPX0843	843	3504	93	279	216	2.40E-24	[LN:C69864] [AC:C69864] [PN:hypothetical protein yktA] [GN:yktA] [OR:Bacillus subtilis]	87
SPX0844	844	3505	283	849	207	1.30E-46	[LN:SUHB_BACSU] [AC:Q45499] [GN:SUHB] [OR:Bacillus subtilis] [DE:EXTRAGENIC SUPPRESSOR PROTEIN SUHB HOMOLOG] [SP:Q45499]	120
SPX0845	845	3506	113	339			NO-HIT	6
SPX0846	846	3507	305	915	143	8.30E-21	[LN:YEBU_ECOLI] [AC:P76273:O07980] [GN:YEBU] [OR:Escherichia coli] [DE:HYPOTHETICAL 53.2 KD PROTEIN IN PRC-PRPA INTERGENIC REGION] [SP:P76273:O07980]	149
SPX0847	847	3508	76	228			NO-HIT	6
SPX0848	848	3509	56	168			NO-HIT	6
SPX0849	849	3510	78	234			NO-HIT	6
SPX0850	850	3511	293	879	465	1.40E-80	[LN:YQGG_BACSU] [AC:P46338] [GN:YQGG] [OR:Bacillus subtilis] [DE:REGION PRECURSOR (ORF108)] [SP:P46338]	103
SPX0851	851	3512	306	918	403	1.00E-99	[LN:YQGH_BACSU] [AC:P46339] [GN:YQGH] [OR:Bacillus subtilis] [DE:REGION (ORF72)] [SP:P46339]	92
SPX0852	852	3513	295	885	842	2.80E-114	[LN:YQGI_BACSU] [AC:P46340] [GN:YQGI] [OR:Bacillus subtilis] [DE:REGION (ORF73)] [SP:P46340]	92
SPX0853	853	3514	268	804	796	1.10E-105	[LN:PSTB_METJA] [AC:Q58418] [GN:PSTB:MJ1012] [OR:Methanococcus jannaschii] [DE:PROBABLE PHOSPHATE TRANSPORT ATP-BINDING PROTEIN PSTB]	145

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:Q58418]	
SPX0854	854	3515	253	759	763	7.10E-100	[LN:T43868] [AC:T43868] [PN:phosphate transport system peripheral membrane protein B [imported]] [GN:pstB] [CL:inner membrane protein malK:ATP-binding cassette homology] [OR:Pseudomonas putida]	193
SPX0855	855	3516	218	654	212	3.70E-40	[GI:4530451] [LN:AF118229] [AC:AF118229] [PN:PhoU] [GN:phoU] [OR:Streptococcus pneumoniae]	90
SPX0856	856	3517	272	816	190	2.40E-30	“[LN:B72357] [AC:B72357] [PN:amino acid ABC transporter, periplasmic amino acid-binding protein] [GN:TM0593] [CL:lysine-arginine-ornithine-binding protein] [OR:Thermotoga maritima]”	181
SPX0857	857	3518	424	1272	778	4.20E-129	[GI:4580622] [LN:AF118389] [AC:AF118389] [PN:unknown] [OR:Streptococcus suis]	77
SPX0858	858	3519	57	171			NO-HIT	6
SPX0859	859	3520	171	513	340	3.10E-77	[GI:1524117] [LN:LLALDB] [AC:X82620] [PN:alpha-acetolactate decarboxylase] [GN:aldB] [OR:Lactococcus lactis]	108
SPX0860	860	3521	67	201	265	2.40E-33	[GI:1808671] [LN:SGCSHAG] [AC:X65164:S52427] [PN:putative alpha-acetolactate decarboxylase] [GN:aldB] [OR:Streptococcus gordonii]	129
SPX0861	861	3522	243	729	106	1.50E-25	[LN:A69830] [AC:A69830] [PN:hypothetical protein yhfC] [GN:yhfC] [OR:Bacillus subtilis]	87
SPX0862	862	3523	302	906	212	5.60E-55	[LN:MURB_BACSU] [AC:P18579:P16669:P37581] [GN:MURB] [OR:Bacillus subtilis] [EC:1.1.1.158] [DE:ACETYLMURAMATE DEHYDROGENASE)] [SP:P18579:P16669:P37581]	150
SPX0863	863	3524	379	1137	792	1.80E-114	“[LN:A70180]	237

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:A70180] [PN:spermidine/putrescine ABC transporter, ATP-binding protein (potA) homolog] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	
SPX0864	864	3525	109	327			NO-HIT	6
SPX0865	865	3526	211	633	324	4.20E-44	"[LN:H70179] [AC:H70179] [PN:spermidine/putrescine ABC transporter, permease protein (potB) homolog] [CL:spermidine/putrescine transport system permease protein potH] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	224
SPX0866	866	3527	258	774	346	7.00E-66	"[LN:G70179] [AC:G70179] [PN:spermidine/putrescine ABC transporter, permease protein (potC) homolog] [CL:spermidine/putrescine transport system permease protein potI] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	224
SPX0867	867	3528	226	678	173	5.20E-48	[LN:POTD_ECOLI] [AC:P23861] [GN:POTD] [OR:Escherichia coli] [DE:SPERMIDINE/PUTRESCINE-BINDING PERIPLASMIC PROTEIN PRECURSOR (SPBP)] [SP:P23861]	143
SPX0868	868	3529	125	375	197	1.30E-20	[LN:POTD_ECOLI] [AC:P23861] [GN:POTD] [OR:Escherichia coli] [DE:SPERMIDINE/PUTRESCINE-BINDING PERIPLASMIC PROTEIN PRECURSOR (SPBP)] [SP:P23861]	143
SPX0869	869	3530	80	240			NO-HIT	6
SPX0870	870	3531	73	219			NO-HIT	6
SPX0871	871	3532	163	489	159	4.70E-15	[LN:G70079] [AC:G70079] [PN:hypothetical protein yxjI] [GN:yxjI] [CL:Bacillus subtilis hypothetical protein yxjI] [OR:Bacillus subtilis]	136
SPX0872	872	3533	89	267	332	2.40E-39	"[LN:SYA_BACSU] [AC:O34526] [GN:ALAS] [OR:Bacillus subtilis] [EC:6.1.1.7] [DE:ALANYL-TRNA SYNTHETASE, (ALANINE--TRNA LIGASE) (ALARS)] [SP:O34526]"	146
SPX0873	873	3534	787	2361	570	1.20E-202	"[LN:SYA_BACSU] [AC:O34526] [GN:ALAS] [OR:Bacillus subtilis] [EC:6.1.1.7]	146

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:ALANYL-TRNA SYNTHETASE, (ALANINE--TRNA LIGASE) (ALARS)] [SP:O34526]”	
SPX0874	874	3535	485	1455	1726	2.50E-257	[GI:2760119] [LN:AB000830] [AC:AB000830] [PN:alpha-amylase precursor] [OR:Streptococcus bovis] [SR:Streptococcus bovis (strain:148) DNA]	136
SPX0875	875	3536	64	192			NO-HIT	6
SPX0876	876	3537	250	750	489	1.10E-62	“[LN:H75077] [AC:H75077] [PN:abc transporter, ATP-binding protein PAB1696] [GN:PAB1696] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Pyrococcus abyssi]”	187
SPX0877	877	3538	545	1635			NO-HIT	6
SPX0878	878	3539	388	1164	468	2.20E-96	[LN:YWBD_BACSU] [AC:P39587] [GN:YWBD:IPA-19D] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 44.4 KD PROTEIN IN EPR-GALK INTERGENIC REGION] [SP:P39587]	144
SPX0879	879	3540	226	678	165	6.60E-27	[GI:5919207] [LN:AF184963] [AC:AF184963] [PN:3-dehydroquinase] [GN:aroD] [OR:Salmonella enteritidis]	100
SPX0880	880	3541	285	855	680	5.70E-89	[GI:3821433] [LN:SPN232281] [AC:AJ232281] [PN:shikimate dehydrogenase] [GN:aroE] [OR:Streptococcus pneumoniae]	110
SPX0881	881	3542	356	1068	439	3.10E-84	“[LN:AROB_SYNY3] [AC:P73997] [GN:AROB:SLR2130] [OR:Synechocystis sp] [SR:,strain PCC 6803] [EC:4.6.1.3] [DE:3-DEHYDROQUINATE SYNTHASE,] [SP:P73997]”	148
SPX0882	882	3543	389	1167	790	1.40E-142	[LN:AROC_BACSU] [AC:P31104] [GN:AROF] [OR:Bacillus subtilis] [EC:4.6.1.4] [DE:PHOSPHOLYASE) (VEGETATIVE PROTEIN 216) (VEG216)] [SP:P31104]	138
SPX0883	883	3544	368	1104	402	2.40E-127	“[LN:TYRA_LACLA]	165

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P43901] [GN:TYRA] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:1.3.1.12] [DE:PREPHENATE DEHYDROGENASE, (PDH)] [SP:P43901]”	
SPX0884	884	3545	113	339	152	7.80E-16	[GI:3688819] [LN:AF084104] [AC:AF084104] [PN:hypothetical protein] [OR:Bacillus firmus]	87
SPX0885	885	3546	428	1284	2094	1.40E-284	[GI:5616525] [LN:AF169483] [AC:AF169483] [PN:5-enolpyruvylshikimate-3-phosphate synthase] [GN:aroA] [OR:Streptococcus pneumoniae]	129
SPX0886	886	3547	159	477	262	3.60E-40	“[LN:AROK_LACLA] [AC:P43906] [GN:AROK] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.1.71] [DE:SHIKIMATE KINASE, (SK)] [SP:P43906]”	156
SPX0887	887	3548	282	846	707	3.20E-98	“[LN:PHEA_LACLA] [AC:P43909] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.2.1.51] [DE:PREPHENATE DEHYDRATASE, (PDT)] [SP:P43909]”	153
SPX0888	888	3549	231	693	91	7.60E-09	[GI:7160813] [LN:EFA276231] [AC:AJ276231] [PN:PSR protein] [GN:psr] [FN:unknown] [OR:Enterococcus faecalis]	107
SPX0889	889	3550	200	600	268	9.10E-46	[GI:7160813] [LN:EFA276231] [AC:AJ276231] [PN:PSR protein] [GN:psr] [FN:unknown] [OR:Enterococcus faecalis]	107
SPX0890	890	3551	282	846	172	1.10E-24	[LN:LICD_HAEIN] [AC:P14184] [GN:LICD]	95

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Haemophilus influenzae] [DE:LICD PROTEIN] [SP:P14184]	
SPX0891	891	3552	134	402			NO-HIT	6
SPX0892	892	3553	386	1158	244	1.70E-28	[LN:F70441] [AC:F70441] [PN:capsular polysaccharide biosynthsis protein] [GN:cap] [OR:Aquifex aeolicus]	103
SPX0893	893	3554	329	987	327	4.10E-51	[GI:3320393] [LN:AF030373] [AC:AF030373] [PN:galactosyl transferase] [GN:cps23FU] [OR:Streptococcus pneumoniae]	111
SPX0894	894	3555	478	1434	96	0.00022	[GI:5931973] [LN:AF125164] [AC:AF125164] [PN:putative polymerase] [GN:wzy] [OR:Bacteroides fragilis]	100
SPX0895	895	3556	151	453			NO-HIT	6
SPX0896	896	3557	202	606			NO-HIT	6
SPX0897	897	3558	99	297			NO-HIT	6
SPX0898	898	3559	185	555	349	1.10E-42	[GI:5360696] [LN:AB022909] [AC:AB022909] [PN:negative regulator of genetic competence] [GN:mecA] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	164
SPX0899	899	3560	128	384			NO-HIT	6
SPX0900	900	3561	71	213	127	1.70E-11	[GI:5360696] [LN:AB022909] [AC:AB022909] [PN:negative regulator of genetic competence] [GN:mecA] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	164
SPX0901	901	3562	429	1287	1425	2.60E-191	“[LN:DHOM_LACLA] [AC:P52985] [GN:HOM] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:1.1.1.3] [DE:HOMOSERINE DEHYDROGENASE, (HDH)] [SP:P52985]”	163
SPX0902	902	3563	290	870	1439	4.70E-194	“[LN:KHSE_STRPN] [AC:P72535] [GN:THRB]	124

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae] [EC:2.7.1.39] [DE:HOMOSERINE KINASE, (HK)] [SP:P72535]”	
SPX0903	903	3564	142	426			NO-HIT	6
SPX0904	904	3565	75	225	192	8.30E-23	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0905	905	3566	86	258	366	6.80E-46	[LN:PMSR__STRPN] [AC:P35593] [GN:MSRA:EXP3] [OR:Streptococcus pneumoniae] [DE:(EXPORTED PROTEIN 3)] [SP:P35593]	110
SPX0906	906	3567	256	768	1236	9.90E-167	[LN:PMSR__STRPN] [AC:P35593] [GN:MSRA:EXP3] [OR:Streptococcus pneumoniae] [DE:(EXPORTED PROTEIN 3)] [SP:P35593]	110
SPX0907	907	3568	103	309	123	6.70E-10	“[LN:E72396] [AC:E72396] [PN:ABC transporter, ATP-binding protein] [GN:TM0287] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Thermotoga maritima]”	180
SPX0908	908	3569	488	1464	472	3.40E-92	“[LN:E72396] [AC:E72396] [PN:ABC transporter, ATP-binding protein] [GN:TM0287] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Thermotoga maritima]”	180
SPX0909	909	3570	62	186			NO-HIT	6
SPX0910	910	3571	584	1752	722	7.90E-145	[LN:Y08A__MYCTU] [AC:Q11047] [GN:MTCY50.10] [OR:Mycobacterium tuberculosis] [DE:HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.10] [SP:Q11047]	148
SPX0911	911	3572	86	258	325	5.10E-41	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0912	912	3573	129	387	95	8.00E-06	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0913	913	3574	64	192			NO-HIT	6
SPX0914	914	3575	74	222			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0915	915	3576	69	207			NO-HIT	6
SPX0916	916	3577	65	195			NO-HIT	6
SPX0917	917	3578	99	297			NO-HIT	6
SPX0918	918	3579	295	885	447	1.10E-56	[LN:A70039] [AC:A70039] [PN:ABC transporter (ATP-binding protein) homolog yvfR] [GN:yvfR] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX0919	919	3580	246	738	289	1.00E-39	[GI:6759480] [LN:BCE243712] [AC:AJ243712] [PN:YVFS protein] [GN:yvfS] [OR:Bacillus cereus]	90
SPX0920	920	3581	366	1098	1715	5.40E-238	[GI:5830547] [LN:SPAJ6400] [AC:AJ006400] [PN:histidine kinase] [GN:hk11] [OR:Streptococcus pneumoniae]	102
SPX0921	921	3582	200	600	932	2.00E-123	[GI:5830548] [LN:SPAJ6400] [AC:AJ006400] [PN:response regulator] [GN:rr11] [OR:Streptococcus pneumoniae]	104
SPX0922	922	3583	337	1011	1328	1.10E-176	[GI:7328454] [LN:AB028599] [AC:AB028599] [PN:catabolite control protein A] [GN:ccpA] [OR:Streptococcus bovis] [SR:Streptococcus bovis (strain:JB1) DNA]	151
SPX0923	923	3584	131	393	270	4.60E-32	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0924	924	3585	113	339	89	5.70E-05	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX0925	925	3586	321	963	295	8.70E-56	[LN:G75283] [AC:G75283] [PN:L-asparaginase] [GN:DR2353] [CL:asparaginase] [OR:Deinococcus radiodurans]	102
SPX0926	926	3587	463	1389	213	9.90E-43	[LN:C69862] [AC:C69862]	178

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:conserved hypothetical protein ykrA] [GN:ykrA] [CL:Methanobacterium thermoautotrophicum conserved hypothetical protein MTH1071] [OR:Bacillus subtilis]	
SPX0927	927	3588	151	453	127	6.20E-14	[LN:A69220] [AC:A69220] [PN:conserved hypothetical protein MTH898] [GN:MTH898] [CL:Escherichia coli ybdQ protein] [OR:Methanobacterium thermoautotrophicum]	155
SPX0928	928	3589	97	291			NO-HIT	6
SPX0929	929	3590	239	717	819	4.70E-109	“[LN:YFBQ_HAEIN] [AC:P71348] [GN:HI0286] [OR:Haemophilus influenzae] [EC:2.6.1.-] [DE:PROBABLE AMINOTRANSFERASE HI0286,] [SP:P71348]”	133
SPX0930	930	3591	111	333			NO-HIT	6
SPX0931	931	3592	125	375	385	1.70E-48	“[LN:YFBQ_HAEIN] [AC:P71348] [GN:HI0286] [OR:Haemophilus influenzae] [EC:2.6.1.-] [DE:PROBABLE AMINOTRANSFERASE HI0286,] [SP:P71348]”	133
SPX0932	932	3593	45	135	170	7.60E-20	[LN:RL34_BACST] [AC:P23376] [GN:RPMH] [OR:Bacillus stearothermophilus] [DE:50S RIBOSOMAL PROTEIN L34] [SP:P23376]	113
SPX0933	933	3594	67	201			NO-HIT	6
SPX0934	934	3595	218	654	116	9.60E-07	[GI:546643] [LN:S70345] [AC:S70345] [PN:SpaA] [GN:SpaA] [OR:Streptococcus sobrinus] [SR:Streptococcus sobrinus MUCOB 263]	121
SPX0935	935	3596	75	225	180	6.20E-19	[LN:YABD_BACSU] [AC:P37545] [GN:YABD] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 29.2 KD PROTEIN IN METS-KSGA INTERGENIC REGION] [SP:P37545]	137
SPX0936	936	3597	181	543	469	3.60E-60	[LN:YABD_BACSU] [AC:P37545] [GN:YABD] [OR:Bacillus subtilis]	137

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0937	937	3598	187	561	235	2.70E-52	[DE:HYPOTHETICAL 29.2 KD PROTEIN IN METS-KSGA INTERGENIC REGION] [SP:P37545] [LN:YABF_BACSU] [AC:P37547] [GN:YABF] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 20.7 KD PROTEIN IN METS-KSGA INTERGENIC REGION] [SP:P37547]	137
SPX0938	938	3599	304	912			NO-HIT	6
SPX0939	939	3600	99	297	149	3.30E-15	[GI:7576923] [LN:AF242367] [AC:AF242367] [PN:lactococcin 972] [GN:lclA] [OR:Lactococcus lactis subsp. lactis]	109
SPX0940	940	3601	269	807	65	5.50E-08	[GI:3355783] [LN:LLJ002203] [AC:AJ002203] [PN:putative immunity] [GN:orf2] [OR:Lactococcus lactis]	98
SPX0941	941	3602	450	1350	357	1.30E-89	[GI:4009478] [LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0942	942	3603	214	642	667	2.50E-87	[GI:4009479] [LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0943	943	3604	103	309			NO-HIT	6
SPX0944	944	3605	169	507			NO-HIT	6
SPX0945	945	3606	291	873	781	3.80E-104	[LN:KSGA_BACSU] [AC:P37468] [GN:KSGA] [OR:Bacillus subtilis] [EC:2.1.1.-] [DE:DIMETHYLTRANSFERASE)] [SP:P37468]	111
SPX0946	946	3607	293	879	394	5.10E-100	[LN:A69879] [AC:A69879] [PN:conserved hypothetical protein yloQ] [GN:yloQ] [CL:conserved hypothetical protein HI1714] [OR:Bacillus subtilis]	140
SPX0947	947	3608	301	903	460	1.20E-69	"[LN:RPE_RHOCA] [AC:P51012] [GN:CBBE] [OR:Rhodobacter capsulatus]	154

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SR:,Rhodopseudomonas capsulata] [EC:5.1.3.1] [DE:EPIMERASE) (PPE) (R5P3E)] [SP:P51012]”	
SPX0948	948	3609	221	663	143	3.50E-18	[LN:C69879] [AC:C69879] [PN:hypothetical protein yloS] [GN:yloS] [OR:Bacillus subtilis]	87
SPX0949	949	3610	405	1215	91	5.10E-07	[LN:E81528] [AC:E81528] [PN:conserved hypothetical protein CP0874 [imported]] [GN:CP0874] [OR:Chlamydophila pneumoniae:Chlamydia pneumoniae]	140
SPX0950	950	3611	314	942	453	1.00E-88	[LN:G69818] [AC:G69818] [PN:CMP-binding factor homolog yhaM] [GN:yhaM] [OR:Bacillus subtilis]	93
SPX0951	951	3612	276	828	498	2.80E-98	[GI:2706406] [LN:LLAJ642] [AC:AJ222642] [GN:purR] [FN:activator of purine biosynthetic genes] [OR:Lactococcus lactis]	117
SPX0952	952	3613	70	210	83	0.00046	“[LN:TABA_PSESZ] [AC:P31851] [GN:TABA] [OR:Pseudomonas syringae] [SR:,pvtabaci] [DE:TABA PROTEIN] [SP:P31851]”	110
SPX0953	953	3614	128	384	199	3.60E-21	“[LN:TABA_PSESZ] [AC:P31851] [GN:TABA] [OR:Pseudomonas syringae] [SR:,pvtabaci] [DE:TABA PROTEIN] [SP:P31851]”	110
SPX0954	954	3615	100	300			NO-HIT	6
SPX0955	955	3616	335	1005	371	8.10E-63	“[LN:TABA_PSESZ] [AC:P31851] [GN:TABA] [OR:Pseudomonas syringae] [SR:,pvtabaci] [DE:TABA PROTEIN] [SP:P31851]”	110
SPX0956	956	3617	265	795	1121	8.40E-152	[LN:PFLA_STRMU] [AC:O68575] [GN:ACT:PFLC]	106

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus mutans] [EC:1.97.1.4] [DE:ENZYME)] [SP:O68575]	
SPX0957	957	3618	309	927	140	6.90E-32	[GI:6117974] [LN:AF139908] [AC:AF139908] [PN:membrane protein homolog] [OR:Listeria monocytogenes]	98
SPX0958	958	3619	93	279	108	9.80E-09	[LN:T35660] [AC:T35660] [PN:probable acylphosphatase] [GN:SC7A1.20] [OR:Streptomyces coelicolor]	96
SPX0959	959	3620	253	759	217	1.10E-43	[LN:G69984] [AC:G69984] [PN:rRNA methylase homolog ysgA] [GN:ysgA] [CL:conserved hypothetical protein HI0860] [OR:Bacillus subtilis]	132
SPX0960	960	3621	228	684	433	1.10E-57	[GI:3171165] [LN:AF064763] [AC:AF064763] [PN:putative membrane spanning protein] [OR:Lactococcus lactis subsp. cremoris]	120
SPX0961	961	3622	72	216			NO-HIT	6
SPX0962	962	3623	331	993	1387	1.70E-186	“[LN:ASNA_HAEIN] [AC:P44338] [GN:ASNA:HI0564] [OR:Haemophilus influenzae] [EC:6.3.1.1] [DE:ASPARTATE--AMMONIA LIGASE, (ASPARAGINE SYNTHETASE A)] [SP:P44338]”	157
SPX0963	963	3624	180	540	279	1.30E-33	[LN:E69874] [AC:E69874] [PN:conserved hypothetical protein ylbH] [GN:ylbH] [CL:Escherichia coli hypothetical 21.7K protein (ftsY-nikA intergenic region)] [OR:Bacillus subtilis]	176
SPX0964	964	3625	123	369	184	3.00E-25	[LN:G81347] [AC:G81347] [PN:3-deoxy-D-manno-octulosonic-acid transferase Cj0767c [imported]] [GN:kdtB:Cj0767c] [OR:Campylobacter jejuni]	136
SPX0965	965	3626	239	717	184	7.20E-17	[LN:A69875] [AC:A69875] [PN:hypothetical protein ylbL] [GN:ylbL] [OR:Bacillus subtilis]	87
SPX0966	966	3627	78	234	101	1.90E-16	[LN:A69875]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:A69875] [PN:hypothetical protein ylbL] [GN:ylbL] [OR:Bacillus subtilis]	
SPX0967	967	3628	95	285			NO-HIT	6
SPX0968	968	3629	428	1284	620	3.20E-133	[LN:MURA__ACICA] [AC:P33986] [GN:MURA:MURZ] [OR:Acinetobacter calcoaceticus] [EC:2.5.1.7] [DE:TRANSFERASE) (EPT)] [SP:P33986]	124
SPX0969	969	3630	75	225	320	4.60E-41	[LN:EPUA__STRPN] [AC:Q03159] [GN:EPUA] [OR:Streptococcus pneumoniae] [DE:EPUA PROTEIN] [SP:Q03159]	97
SPX0970	970	3631	275	825	1392	2.50E-182	“[LN:NUCE__STRPN] [AC:Q03158] [GN:ENDA] [OR:Streptococcus pneumoniae] [EC:3.1.30.-] [DE:DNA-ENTRY NUCLEASE (COMPETENCE-SPECIFIC NUCLEASE),] [SP:Q03158]”	151
SPX0971	971	3632	328	984	1168	4.20E-160	[GI:2952527] [LN:AF051356] [AC:AF051356] [PN:putative hemolysin] [GN:hlyX] [OR:Streptococcus mutans]	100
SPX0972	972	3633	95	285	266	6.50E-33	[GI:2952527] [LN:AF051356] [AC:AF051356] [PN:putative hemolysin] [GN:hlyX] [OR:Streptococcus mutans]	100
SPX0973	973	3634	1204	3612	1921	0	[LN:T44375] [AC:T44375] [PN:rpoB protein [imported]] [CL:DNA-directed RNA polymerase beta chain] [OR:Bacillus halodurans]	121
SPX0974	974	3635	116	348			NO-HIT	6
SPX0975	975	3636	57	171			NO-HIT	6
SPX0976	976	3637	1226	3678	4009	0	[LN:RPOC__STRPY] [AC:P95816] [GN:RPOC] [OR:Streptococcus pyogenes] [EC:2.7.7.6] [DE:BETA'CHAIN) (RNA POLYMERASE BETA'SUBUNIT) (FRAGMENT)]	150

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P95816]	
SPX0977	977	3638	131	393			NO-HIT	6
SPX0978	978	3639	138	414	192	4.30E-48	[GI:1655704] [LN:XLNM23] [AC:X97899] [PN:NM23/nucleoside diphosphate kinase] [OR:Xenopus laevis] [SR:African clawed frog]	121
SPX0979	979	3640	93	279			NO-HIT	6
SPX0980	980	3641	211	633	472	1.50E-87	[GI:4009479] [LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0981	981	3642	672	2016	277	4.50E-70	[GI:4009478] [LN:AF068902] [AC:AF068902] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX0982	982	3643	65	195			NO-HIT	6
SPX0983	983	3644	288	864	110	2.30E-09	[LN:JC6007] [AC:JC6007] [PN:transcription activator plcR] [GN:plcR] [CL:Bacillus thuringiensis transcription activator plcR] [OR:Bacillus thuringiensis]	152
SPX0984	984	3645	137	411			NO-HIT	6
SPX0985	985	3646	159	477	166	2.20E-36	[LN:C69786] [AC:C69786] [PN:conserved hypothetical protein ydiB] [GN:ydiB] [CL:hypothetical protein HI0065] [OR:Bacillus subtilis]	130
SPX0986	986	3647	173	519	149	7.60E-28	[LN:D72360] [AC:D72360] [PN:conserved hypothetical protein] [GN:TM0577] [OR:Thermotoga maritima]	96
SPX0987	987	3648	339	1017	182	6.50E-46	[LN:LYTR_BACSU] [AC:Q02115] [GN:LYTR] [OR:Bacillus subtilis] [DE:MEMBRANE-BOUND PROTEIN LYTR] [SP:Q02115]	105
SPX0988	988	3649	419	1257	2015	1.70E-272	[LN:CINA_STRPN] [AC:P54184:Q54853] [GN:CINA:EXP10] [OR:Streptococcus pneumoniae] [DE:PUTATIVE COMPETENCE-DAMAGE PROTEIN (EXPORTED PROTEIN 10)] [SP:P54184:Q54853]	161

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX0989	989	3650	389	1167	1928	6.10E-256	[LN:RECA__STRPN] [AC:P30758] [GN:RECA] [OR:Streptococcus pneumoniae] [DE:RECA PROTEIN] [SP:P30758]	97
SPX0990	990	3651	457	1371	2252	0	[GI:2398825] [LN:SPCINREC] [AC:Z34303] [PN:DinF protein] [GN:dinF] [OR:Streptococcus pneumoniae]	96
SPX0991	991	3652	319	957	1786	1.50E-240	[LN:ALYS__STRPN] [AC:P06653] [GN:LYTA] [OR:Streptococcus pneumoniae] [EC:3.5.1.28] [DE:HYDROLASE) (MUCOPEPTIDE AMINOHYDROLASE) (CELL WALL HYDROLASE)] [SP:P06653]	160
SPX0992	992	3653	79	237			NO-HIT	6
SPX0993	993	3654	81	243			NO-HIT	6
SPX0994	994	3655	120	360			NO-HIT	6
SPX0995	995	3656	142	426			NO-HIT	6
SPX0996	996	3657	75	225			NO-HIT	6
SPX0997	997	3658	332	996	676	4.50E-129	[GI:2398827] [LN:SPCINREC] [AC:Z34303] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	94
SPX0998	998	3659	82	246			NO-HIT	6
SPX0999	999	3660	90	270			NO-HIT	6
SPX1000	1000	3661	87	261			NO-HIT	6
SPX1001	1001	3662	101	303	296	7.00E-36	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1002	1002	3663	44	132	108	5.70E-09	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1003	1003	3664	153	459			NO-HIT	6
SPX1004	1004	3665	203	609			NO-HIT	6
SPX1005	1005	3666	137	411			NO-HIT	6
SPX1006	1006	3667	472	1416	2426	0	[LN:A28568] [AC:A28568:S12829] [PN:pneumolysin] [CL:dipeptide transport protein] [OR:Streptococcus pneumoniae]	110
SPX1007	1007	3668	239	717	915	1.30E-119	[LN:A64963] [AC:A64963]	119

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1008	1008	3669	150	450	104	9.20E-07	[PN:conserved hypothetical protein b1983] [CL:hypothetical protein MG332] [OR:Escherichia coli] [LN:C72329] [AC:C72329]	93
SPX1009	1009	3670	291	873	309	1.20E-35	[PN:hypothetical protein TM0816] [GN:TM0816] [OR:Thermotoga maritima] [GI:6759558] [LN:SC7A8] [AC:AL137187]	116
SPX1010	1010	3671	344	1032	335	7.80E-76	[PN:putative ABC transporter] [GN:SC7A8.02] [OR:Streptomyces coelicolor A3(2)] [GI:6759559] [LN:SC7A8] [AC:AL137187] [PN:putative ABC transporter] [GN:SC7A8.03] [OR:Streptomyces coelicolor A3(2)]	116
SPX1011	1011	3672	121	363			NO-HIT	6
SPX1012	1012	3673	168	504	130	2.10E-10	[LN:Y374_METJA] [AC:Q57819] [GN:MJ0374] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ0374] [SP:Q57819]	114
SPX1013	1013	3674	150	450			NO-HIT	6
SPX1014	1014	3675	225	675			NO-HIT	6
SPX1015	1015	3676	74	222	125	8.00E-13	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX1016	1016	3677	100	300			NO-HIT	6
SPX1017	1017	3678	106	318	276	6.70E-34	[LN:G69998] [AC:G69998] [PN:thioredoxin H1 homolog ytpP] [GN:ytpP] [CL:thioredoxin:thioredoxin homology] [OR:Bacillus subtilis]	127
SPX1018	1018	3679	126	378			NO-HIT	6
SPX1019	1019	3680	209	627	217	5.30E-44	[LN:A69999] [AC:A69999] [PN:phenylalanyl-tRNA synthetase (beta subunit) homolog ytpR] [GN:ytpR] [CL:Mycoplasma genitalium hypothetical protein MG449] [OR:Bacillus subtilis]	172
SPX1020	1020	3681	254	762	354	4.20E-67	[GI:2425123] [LN:AF019986]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AF019986] [PN:PksB] [GN:pksB] [OR:Dictyostelium discoideum]	
SPX1021	1021	3682	132	396	227	1.20E-34	[GI:6716352] [LN:AF145054] [AC:AF145054:AF001793:AF118440:U89246] [PN:ORF9] [GN:orf9] [OR:Streptococcus thermophilus bacteriophage 7201]	136
SPX1022	1022	3683	95	285	460	1.90E-58	[GI:4566772] [LN:AF117741] [AC:AF117741] [PN:cochaperonin GroES] [GN:groES] [OR:Streptococcus pneumoniae]	105
SPX1023	1023	3684	203	609	344	8.20E-43	[GI:1196510] [LN:MSGTCWPA] [AC:M15467] [PN:unknown protein] [OR:Mycobacterium tuberculosis] [SR:Mycobacterium tuberculosis (strain Erdman) DNA]	143
SPX1024	1024	3685	541	1623	2611	0	[GI:4566773] [LN:AF117741] [AC:AF117741] [PN:chaperonin GroEL] [GN:groEL] [OR:Streptococcus pneumoniae]	103
SPX1025	1025	3686	70	210	70	3.30E-06	[LN:C71647] [AC:C71647] [PN:hypothetical protein RP851] [GN:RP851] [OR:Rickettsia prowazekii]	93
SPX1026	1026	3687	84	252	139	1.80E-13	[LN:F72338] [AC:F72338] [PN:conserved hypothetical protein] [GN:TM0731] [OR:Thermotoga maritima]	96
SPX1027	1027	3688	64	192			NO-HIT	6
SPX1028	1028	3689	258	774	129	6.70E-09	[LN:T33885] [AC:T33885] [PN:hypothetical protein H14E04.1] [GN:H14E04.1] [CL:24-sterol C-methyltransferase:bioC homology] [OR:Caenorhabditis elegans]	149
SPX1029	1029	3690	113	339	133	1.50E-11	[LN:AB025228] [AC:AB025228] [GN:rgg] [OR:Streptococcus oralis] [SR:Streptococcus oralis (strain:ATCC10557) DNA]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1030	1030	3691	116	348	90	8.90E-06	[LN:AB025228] [AC:AB025228] [GN:rgg] [OR:Streptococcus oralis] [SR:Streptococcus oralis (strain:ATCC10557) DNA]	111
SPX1031	1031	3692	178	534	294	4.90E-36	[LN:F69815] [AC:F69815] [PN:hypothetical protein ygaC] [GN:ygaC] [CL:Bacillus subtilis hypothetical protein ygaC] [OR:Bacillus subtilis]	136
SPX1032	1032	3693	60	180			NO-HIT	6
SPX1033	1033	3694	287	861	226	3.30E-32	[LN:H69800] [AC:H69800] [PN:hypothetical protein yfhG] [GN:yfhG] [OR:Bacillus subtilis]	87
SPX1034	1034	3695	452	1356	606	1.00E-145	[LN:F69806] [AC:F69806] [PN:RNA methyltransferase homolog yfjO] [GN:yfjO] [CL:hypothetical protein HI0333] [OR:Bacillus subtilis]	129
SPX1035	1035	3696	82	246			NO-HIT	6
SPX1036	1036	3697	50	150	82	7.60E-06	[LN:G71244] [AC:G71244] [PN:hypothetical protein PH0217] [GN:PH0217] [OR:Pyrococcus horikoshii]	95
SPX1037	1037	3698	62	186			NO-HIT	6
SPX1038	1038	3699	87	261	136	3.90E-14	[LN:G81516] [AC:G81516] [PN:hypothetical protein CP0988 [imported]] [GN:CP0988] [OR:Chlamydophila pneumoniae:Chlamydia pneumoniae]	130
SPX1039	1039	3700	76	228	190	1.00E-21	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1040	1040	3701	79	237			NO-HIT	6
SPX1041	1041	3702	94	282	301	2.00E-35	[LN:DHAS_STRMU] [AC:P10539] [GN:ASD] [OR:Streptococcus mutans] [EC:1.2.1.11] [DE:DEHYDROGENASE)] [SP:P10539]	108
SPX1042	1042	3703	92	276	325	3.80E-40	[LN:DHAS_STRMU] [AC:P10539] [GN:ASD]	108

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus mutans] [EC:1.2.1.11] [DE:DEHYDROGENASE)] [SP:P10539]	
SPX1043	1043	3704	222	666	590	1.20E-108	[LN:DHAS_STRMU] [AC:P10539] [GN:ASD] [OR:Streptococcus mutans] [EC:1.2.1.11] [DE:DEHYDROGENASE)] [SP:P10539]	108
SPX1044	1044	3705	312	936	192	2.60E-55	“[LN:B72246] [AC:B72246] [PN:dihydrodipicolinate synthase, TM1521 [similarity]] [GN:TM1521] [CL:Pseudomonas 5-dehydro-4-deoxyglucarate dehydratase] [OR:Thermotoga maritima] [EC:4.2.1.52]”	187
SPX1045	1045	3706	458	1374	1312	1.10E-249	[GI:6448626] [LN:SAG251564] [AC:AJ251564] [PN:thiophene degradation protein F] [GN:thdF] [FN:putative thiophene and furan oxidation gene] [OR:Streptococcus agalactiae]	167
SPX1046	1046	3707	75	225	127	1.30E-12	“[LN:A43397] [AC:A43397:S35225] [PN:4-oxalocrotonate tautomerase, xylH] [GN:xylH] [CL:4-oxalocrotonate tautomerase] [OR:Pseudomonas putida] [EC:5.3.2.-]”	153
SPX1047	1047	3708	164	492	724	7.00E-97	“[LN:KITH_STRGC] [AC:P47848] [GN:TDK] [OR:Streptococcus gordonii challis] [EC:2.7.1.21] [DE:THYMIDINE KINASE,] [SP:P47848]”	123
SPX1048	1048	3709	281	843	642	2.20E-103	[LN:RF1_BACSU] [AC:P45872] [GN:PRFA] [OR:Bacillus subtilis] [DE:PEPTIDE CHAIN RELEASE FACTOR 1 (RF-1)] [SP:P45872]	114
SPX1049	1049	3710	143	429	350	2.60E-43	[LN:RF1_BACSU] [AC:P45872] [GN:PRFA] [OR:Bacillus subtilis] [DE:PEPTIDE CHAIN RELEASE FACTOR 1 (RF-1)]	114

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P45872]	
SPX1050	1050	3711	280	840	183	8.00E-41	[LN:HEMK__BACSU] [AC:P45873] [GN:YWKE] [OR:Bacillus subtilis] [DE:HEMK PROTEIN HOMOLOG] [SP:P45873]	98
SPX1051	1051	3712	201	603	244	1.80E-32	[GI:6015811] [LN:SSU18930] [AC:Y18930] [PN:hypothetical protein] [GN:ORF-c09__003] [OR:Sulfolobus solfataricus]	110
SPX1052	1052	3713	144	432	97	1.30E-05	[LN:PHNO__ECOLI] [AC:P16691] [GN:PHNO] [OR:Escherichia coli] [DE:PHNO PROTEIN] [SP:P16691]	89
SPX1053	1053	3714	419	1257	858	7.90E-176	[LN:GLYA__BACSU] [AC:P39148] [GN:GLYA:GLYC:IPC-34D] [OR:Bacillus subtilis] [EC:2.1.2.1] [DE:(SHMT)] [SP:P39148]	110
SPX1054	1054	3715	336	1008	130	3.90E-11	[GI:6899993] [LN:CST130879] [AC:AJ130879] [PN:hypothetical protein] [FN:unknown] [OR:Clostridium sticklandii]	109
SPX1055	1055	3716	205	615	243	1.60E-27	[LN:F69900] [AC:F69900] [PN:transposon-related protein homolog yocA] [GN:yocA] [OR:Bacillus subtilis]	101
SPX1056	1056	3717	325	975	104	7.60E-08	[LN:YPUA__BACSU] [AC:P31847:P37951] [GN:YPUA] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 31.3 KD PROTEIN IN LYSA-PPIB INTERGENIC REGION (ORFX19)] [SP:P31847:P37951]	160
SPX1057	1057	3718	548	1644	767	5.90E-131	[LN:E69793] [AC:E69793] [PN:RNA methyltransferase homolog yefA] [GN:yefA] [CL:hypothetical protein HI0333] [OR:Bacillus subtilis]	129
SPX1058	1058	3719	98	294			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1059	1059	3720	83	249			NO-HIT	6
SPX1060	1060	3721	65	195	81	3.20E-08	[LN:C69931] [AC:C69931] [PN:transcription regulator homolog yozG] [GN:yozG] [OR:Bacillus subtilis]	98
SPX1061	1061	3722	80	240			NO-HIT	6
SPX1062	1062	3723	345	1035	276	3.80E-70	[GI:2276374] [LN:CDU02617] [AC:U02617] [PN:DtxR/iron regulated lipoprotein precursor] [GN:irp1] [FN:iron transport] [OR:Corynebacterium diphtheriae]	148
SPX1063	1063	3724	212	636			NO-HIT	6
SPX1064	1064	3725	336	1008	643	1.40E-86	[LN:T44797] [AC:T44797] [PN:iron transport membrane protein irp1B [imported]] [GN:irp1B] [CL:vitamin B12 transport protein btuC] [OR:Corynebacterium diphtheriae]	161
SPX1065	1065	3726	336	1008	484	1.10E-82	[LN:T44798] [AC:T44798] [PN:iron transport membrane protein irp1C [imported]] [GN:irp1C] [CL:vitamin B12 transport protein btuC] [OR:Corynebacterium diphtheriae]	161
SPX1066	1066	3727	74	222			NO-HIT	6
SPX1067	1067	3728	138	414			NO-HIT	6
SPX1068	1068	3729	265	795	694	6.10E-91	[LN:G70022] [AC:G70022] [PN:iron(III) dicitrate transport permease homolog yusV] [GN:yusV] [CL:inner membrane protein malK:ATP-binding cassette homology] [OR:Bacillus subtilis]	176
SPX1069	1069	3730	87	261	151	1.30E-14	[GI:2467226] [LN:LLLPK214] [AC:X92946:Y10522] [PN:transposase] [GN:tnpA] [OR:Lactococcus lactis]	96
SPX1070	1070	3731	152	456	137	1.10E-10	“[LN:T4BB_BACCO] [AC:Q07606] [GN:BCGIB] [OR:Bacillus coagulans] [EC:3.1.21.-] [DE:RESTRICTION ENZYME BGCI BETA SUBUNIT,] [SP:Q07606]”	133
SPX1071	1071	3732	92	276			NO-HIT	6
SPX1072	1072	3733	83	249			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1073	1073	3734	70	210			NO-HIT	6
SPX1074	1074	3735	560	1680	138	3.40E-35	[GI:6681569] [LN:AB014436] [AC:AB014436] [PN:cassette chromosome recombinase B] [GN:ccrB] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:85/2082) DNA]	164
SPX1075	1075	3736	204	612			NO-HIT	6
SPX1076	1076	3737	63	189			NO-HIT	6
SPX1077	1077	3738	253	759			NO-HIT	6
SPX1078	1078	3739	285	855	70	4.40E-09	[GI:7020551] [LN:AK000453] [AC:AK000453] [OR:Homo sapiens] [SR:Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA t]	129
SPX1079	1079	3740	295	885	192	4.90E-36	[LN:F69795] [AC:F69795] [PN:conserved hypothetical protein yerQ] [GN:yerQ] [OR:Bacillus subtilis]	97
SPX1080	1080	3741	93	279	124	1.30E-19	“[LN:JS0673] [AC:JS0673] [PN:neopullulanase,] [CL:neopullulanase:alpha-amylase core homology] [OR:Bacillus sp.] [EC:3.2.1.135]”	127
SPX1081	1081	3742	485	1455	683	3.10E-176	“[LN:F70033] [AC:F70033] [PN:glucan 1,4-alpha-maltohydrolase homolog yvdF] [GN:yvdF] [CL:neopullulanase:alpha-amylase core homology] [OR:Bacillus subtilis]”	156
SPX1082	1082	3743	101	303			NO-HIT	6
SPX1083	1083	3744	143	429	249	9.00E-30	[GI:6102604] [LN:AF144880] [AC:AF144880] [PN:aminoglycoside 6'-N-acetyltransferase] [GN:aac(6')-Iy] [OR:Salmonella enteritidis]	127
SPX1084	1084	3745	190	570			NO-HIT	6
SPX1085	1085	3746	71	213			NO-HIT	6
SPX1086	1086	3747	94	282			NO-HIT	6
SPX1087	1087	3748	65	195	184	8.80E-20	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1088	1088	3749	410	1230	419	1.50E-74	[LN:FTSW_ENTHR] [AC:Q47866] [GN:FTSW]	114

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Enterococcus hirae] [DE:PROBABLE CELL DIVISION PROTEIN FTSW] [SP:Q47866]	
SPX1089	1089	3750	899	2697	418	1.60E-169	[GI:144985] [LN:CORPEPC] [AC:M25819] [PN:phosphoenolpyruvate carboxylase] [OR:Corynebacterium glutamicum] [SR:Corynebacterium glutamicum (clone: pTG1200.) (tissue library: ATC]	176
SPX1090	1090	3751	102	306			NO-HIT	6
SPX1091	1091	3752	210	630	105	2.60E-05	[LN:H71693] [AC:H71693] [PN:hypothetical protein RP367] [GN:RP367] [OR:Rickettsia prowazekii]	93
SPX1092	1092	3753	223	669	163	6.00E-23	[LN:A71694] [AC:A71694] [PN:hypothetical protein RP368] [GN:RP368] [OR:Rickettsia prowazekii]	93
SPX1093	1093	3754	253	759	365	1.80E-45	[GI:1661179] [LN:SMU75471] [AC:U75471] [PN:high affinity branched chain amino acid] [GN:livG] [OR:Streptococcus mutans]	119
SPX1094	1094	3755	64	192	82	2.10E-05	[LN:A71007] [AC:A71007] [PN:hypothetical protein PH1351] [GN:PH1351] [OR:Pyrococcus horikoshii]	95
SPX1095	1095	3756	592	1776	694	1.40E-161	“[LN:PRIM_LACLA] [AC:Q04505] [GN:DNAG:DNAE] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.7.-] [DE:DNA PRIMASE,] [SP:Q04505]”	150
SPX1096	1096	3757	305	915	1452	2.30E-194	[LN:RPOD_STRPN] [AC:O08388] [GN:RPOD] [OR:Streptococcus pneumoniae] [DE:RNA POLYMERASE SIGMA FACTOR RPOD] [SP:O08388]	117
SPX1097	1097	3758	66	198	290	2.40E-35	[LN:RPOD_STRPN] [AC:O08388] [GN:RPOD] [OR:Streptococcus pneumoniae] [DE:RNA POLYMERASE SIGMA FACTOR RPOD]	117

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:O08388]	
SPX1098	1098	3759	110	330	556	5.10E-73	[GI:2108332] [LN:SPDNAGCPO] [AC:Y11463] [OR:Streptococcus pneumoniae]	69
SPX1099	1099	3760	69	207			NO-HIT	6
SPX1100	1100	3761	267	801	1357	1.40E-186	[GI:2108333] [LN:SPDNAGCPO] [AC:Y11463] [GN:cpoA] [OR:Streptococcus pneumoniae]	79
SPX1101	1101	3762	66	198	318	4.00E-38	[GI:2108333] [LN:SPDNAGCPO] [AC:Y11463] [GN:cpoA] [OR:Streptococcus pneumoniae]	79
SPX1102	1102	3763	442	1326	1475	1.70E-200	[GI:2108334] [LN:SPDNAGCPO] [AC:Y11463] [OR:Streptococcus pneumoniae]	69
SPX1103	1103	3764	59	177	151	3.70E-16	[LN:B72287] [AC:B72287] [PN:conserved hypothetical protein] [GN:TM1156] [OR:Thermotoga maritima]	96
SPX1104	1104	3765	392	1176	1021	5.10E-184	[LN:OBG_BACSU] [AC:P20964] [GN:OBG] [OR:Bacillus subtilis] [DE:SPO0B-ASSOCIATED GTP-BINDING PROTEIN] [SP:P20964]	112
SPX1105	1105	3766	420	1260	424	9.00E-142	[LN:MURA_BACSU] [AC:P19670:Q03225] [GN:MURA:MURZ] [OR:Bacillus subtilis] [EC:2.5.1.7] [DE:ENOLPYRUVYL TRANSFERASE) (EPT)] [SP:P19670:Q03225]	140
SPX1106	1106	3767	176	528	156	1.00E-17	[LN:C69895] [AC:C69895] [PN:conserved hypothetical protein yoaA] [GN:yoaA] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimJ] [OR:Bacillus subtilis]	170
SPX1107	1107	3768	426	1278	425	4.20E-85	[LN:A69998] [AC:A69998] [PN:hypothetical protein ytoI] [GN:ytoI] [OR:Bacillus subtilis]	87
SPX1108	1108	3769	287	861	345	3.30E-65	"[LN:T44405]	158

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:T44405] [PN:methionyl aminopeptidase, map [imported]] [CL:Escherichia coli methionyl aminopeptidase] [OR:Bacillus halodurans] [EC:3.4.11.18]”	
SPX1109	1109	3770	181	543	307	2.50E-38	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1110	1110	3771	75	225	104	5.80E-08	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1111	1111	3772	130	390	384	4.80E-49	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1112	1112	3773	764	2292	1351	7.70E-282	[LN:E69794] [AC:E69794] [PN:ATP-dependent DNA helicase homolog yerF] [GN:yerF] [CL:helicase II] [OR:Bacillus subtilis]	118
SPX1113	1113	3774	240	720	488	2.80E-62	[LN:RADC_BACSU] [AC:Q02170] [GN:YSXA] [OR:Bacillus subtilis] [DE:DNA REPAIR PROTEIN RADC HOMOLOG (ORFB)] [SP:Q02170]	116
SPX1114	1114	3775	230	690	237	6.50E-27	“[LN:YDP3_LACLA] [AC:P22347] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:HYPOTHETICAL 18.7 KD PROTEIN IN PEPX 3'REGION (ORF3)] [SP:P22347]”	162
SPX1115	1115	3776	214	642	339	2.60E-68	[LN:A69787] [AC:A69787] [PN:hypothetical protein ydiH] [GN:ydiH] [OR:Bacillus subtilis]	87
SPX1116	1116	3777	84	252			NO-HIT	6
SPX1117	1117	3778	116	348			NO-HIT	6
SPX1118	1118	3779	265	795	288	2.40E-63	[LN:E69981] [AC:E69981] [PN:iron-sulfur cofactor synthesis protein yrvO] [GN:yrvO] [CL:nitrogen fixation protein nifS] [OR:Bacillus subtilis]	141
SPX1119	1119	3780	145	435	149	4.90E-18	[GI:2289093] [LN:CAU76387]	102

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:U76387] [PN:PRPP synthetase] [GN:prs] [OR:Corynebacterium ammoniagenes]	
SPX1120	1120	3781	183	549	547	6.60E-75	[GI:2289093] [LN:CAU76387] [AC:U76387] [PN:PRPP synthetase] [GN:prs] [OR:Corynebacterium ammoniagenes]	102
SPX1121	1121	3782	190	570	164	1.20E-26	[LN:C69844] [AC:C69844] [PN:hypothetical protein yjbK] [GN:yjbK] [CL:Bacillus subtilis hypothetical protein yjbK] [OR:Bacillus subtilis]	136
SPX1122	1122	3783	168	504	338	6.80E-68	[LN:E69844] [AC:E69844] [PN:GTP pyrophosphokinase homolog yjbM] [GN:yjbM] [CL:GTP pyrophosphokinase related protein] [OR:Bacillus subtilis]	139
SPX1123	1123	3784	71	213	90	6.20E-06	[LN:E69844] [AC:E69844] [PN:GTP pyrophosphokinase homolog yjbM] [GN:yjbM] [CL:GTP pyrophosphokinase related protein] [OR:Bacillus subtilis]	139
SPX1124	1124	3785	263	789	309	4.70E-67	[LN:F69844] [AC:F69844] [PN:conserved hypothetical protein yjbN] [GN:yjbN] [CL:conserved hypothetical protein HI0072] [OR:Bacillus subtilis]	140
SPX1125	1125	3786	299	897	286	3.40E-66	[LN:YJBO_BACSU] [AC:O31613] [GN:YJBO] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 31.5 KD PROTEIN IN MECA-TENA INTERGENIC REGION] [SP:O31613]	137
SPX1126	1126	3787	325	975	867	5.10E-115	[LN:PTA_BACSU] [AC:P39646] [GN:PTA:IPA-88D] [OR:Bacillus subtilis] [EC:2.3.1.8] [DE:(PHOSPHOTRANSACETYLASE) (VEGETATIVE PROTEIN 43) (VEG43)] [SP:P39646]	152
SPX1127	1127	3788	79	237			NO-HIT	6
SPX1128	1128	3789	199	597	472	3.10E-122	[GI:663279] [LN:STRCOMAA]	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	
SPX1129	1129	3790	109	327	280	2.00E-34	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1130	1130	3791	173	519	875	3.10E-119	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX1131	1131	3792	108	324	190	3.90E-21	[LN:YAAK_BACSU] [AC:P24281] [GN:YAAK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 11.8 KD PROTEIN IN DNAZ-RECR INTERGENIC REGION] [SP:P24281]	137
SPX1132	1132	3793	231	693	89	1.90E-05	[LN:E69896] [AC:E69896] [PN:hypothetical protein yoaK] [GN:yoaK] [OR:Bacillus subtilis]	87
SPX1133	1133	3794	105	315	360	3.20E-44	[LN:RL21_BACSU] [AC:P26908] [GN:RPLU] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L21 (BL20)] [SP:P26908]	110
SPX1134	1134	3795	147	441			NO-HIT	6
SPX1135	1135	3796	115	345	152	1.30E-22	[LN:YSXB_BACSU] [AC:P26942:Q45629] [GN:YSXB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 12.3 KD PROTEIN IN RPLU-RPMA INTERGENIC REGION (ORF X)] [SP:P26942:Q45629]	159
SPX1136	1136	3797	98	294	372	1.10E-46	[LN:RL27_BACSU] [AC:P05657] [GN:RPMA] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L27 (BL30) (BL24)] [SP:P05657]	117
SPX1137	1137	3798	55	165	77	4.70E-09	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1138	1138	3799	306	918	808	1.30E-107	[GI:2289231] [LN:SAU92073] [AC:U92073] [PN:macrolide-efflux protein] [GN:mreA] [FN:resistance to 14- and 15-membered macrolides] [OR:Streptococcus agalactiae]	158
SPX1139	1139	3800	283	849	531	2.80E-71	[LN:YPJC_BACSU] [AC:P42978] [GN:YPJC:JOJC] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 23.6 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION] [SP:P42978]	142
SPX1140	1140	3801	280	840	252	9.50E-36	[GI:671632] [LN:SADIRED] [AC:Z16422] [PN:unknown] [GN:ORF2] [OR:Staphylococcus aureus]	86
SPX1141	1141	3802	92	276	397	9.60E-50	[GI:5209334] [LN:L40356] [AC:L40356] [PN:histone-like DNA-binding protein] [GN:hlpA] [OR:Streptococcus gordonii]	112
SPX1142	1142	3803	634	1902	905	2.00E-202	[LN:YDIF_BACSU] [AC:O05519] [GN:YDIF] [OR:Bacillus subtilis] [DE:HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YDIF] [SP:O05519]	131
SPX1143	1143	3804	298	894	204	8.40E-33	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX1144	1144	3805	393	1179	113	5.30E-22	[LN:H72265] [AC:H72265] [PN:hypothetical protein TM1336] [GN:TM1336] [OR:Thermotoga maritima]	93
SPX1145	1145	3806	656	1968	521	1.40E-217	“[LN:F69794] [AC:F69794] [PN:DNA ligase (NAD+),] [GN:yerG] [CL:polydeoxyribonucleotide synthase (NAD+)] [OR:Bacillus subtilis] [EC:6.5.1.2]”	140

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1146	1146	3807	760	2280	620	2.20E-109	[GI:3089609] [LN:AF060205] [AC:AF060205] [PN:pullulanase] [GN:pul] [OR:Thermus sp. IM6501]	90
SPX1147	1147	3808	475	1425	1818	2.70E-245	[LN:GAPN_STRMU] [AC:Q59931] [GN:GAPN] [OR:Streptococcus mutans] [EC:1.2.1.9] [DE:DEHYDROGENASE)] [SP:Q59931]	108
SPX1148	1148	3809	643	1929	658	6.60E-176	"[LN:S18599] [AC:S18599] [PN:1,4-alpha-glucan branching enzyme,:glycogen branching enzyme] [CL:1,4-alpha-glucan branching enzyme] [OR:Bacillus stearothermophilus] [EC:2.4.1.18]"	177
SPX1149	1149	3810	381	1143	187	6.00E-19	[LN:CCE277601] [AC:AJ277601] [PN:ADP-glucose pyrophosphorylase] [GN:glgD] [OR:Clostridium cellulolyticum]	105
SPX1150	1150	3811	380	1140	324	4.80E-53	[LN:GLGD_BACST] [AC:O08327] [GN:GLGD] [OR:Bacillus stearothermophilus] [DE:GLYCOGEN BIOSYNTHESIS PROTEIN GLGD] [SP:O08327]	122
SPX1151	1151	3812	478	1434	813	6.80E-159	[LN:GLGA_BACSU] [AC:P39125] [GN:GLGA] [OR:Bacillus subtilis] [EC:2.4.1.21] [DE:SYNTHASE)] [SP:P39125]	101
SPX1152	1152	3813	124	372	214	7.80E-24	[LN:B69517] [AC:B69517] [PN:phosphoserine phosphatase (serB) homolog] [OR:Archaeoglobus fulgidus]	97
SPX1153	1153	3814	110	330	213	6.30E-24	"[LN:A64499] [AC:A64499] [PN:phosphoserine phosphatase,] [CL:phosphoserine phosphatase] [OR:Methanococcus jannaschii] [EC:3.1.3.3]"	131
SPX1154	1154	3815	372	1116	1670	1.30E-226	[GI:7380122] [LN:NMA5Z2491] [AC:AL162756:AL157959]	125

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:hypothetical protein NMA1473] [GN:NMA1473] [OR:Neisseria meningitidis]	
SPX1155	1155	3816	149	447			NO-HIT	6
SPX1156	1156	3817	435	1305	2119	6.30E-286	[GI:5263171] [LN:AB029313] [AC:AB029313] [PN:enolase] [GN:eno] [OR:Streptococcus intermedius] [SR:Streptococcus intermedius (strain:ATCC 27335) DNA]	148
SPX1157	1157	3818	156	468	74	6.80E-09	"[LN:T45116] [AC:T45116] [PN:phosphopyruvate hydratase, [imported]] [OR:Schizosaccharomyces pombe] [EC:4.2.1.11]"	113
SPX1158	1158	3819	44	132	75	2.00E-06	[LN:B72477] [AC:B72477] [PN:hypothetical protein APE2459] [GN:APE2459] [OR:Aeropyrum pernix]	92
SPX1159	1159	3820	169	507	174	3.60E-16	[GI:6782411] [LN:STH243106] [AC:AJ243106] [PN:integrase] [GN:int] [OR:Streptococcus thermophilus]	97
SPX1160	1160	3821	113	339	556	6.20E-74	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX1161	1161	3822	116	348	573	1.80E-75	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96
SPX1162	1162	3823	1092	3276	315	2.90E-101	[LN:T30307] [AC:T30307] [PN:rexB protein] [GN:rexB] [OR:Lactococcus lactis]	75
SPX1163	1163	3824	1217	3651	517	1.80E-246	[LN:T30308] [AC:T30308] [PN:rexA protein] [OR:Lactococcus lactis]	65
SPX1164	1164	3825	79	237			NO-HIT	6
SPX1165	1165	3826	64	192	205	2.00E-24	[LN:T30285] [AC:T30285] [PN:hypothetical protein]	79

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX1166	1166	3827	393	1179			NO-HIT	6
SPX1167	1167	3828	111	333			NO-HIT	6
SPX1168	1168	3829	1966	5898	7485	0	[GI:1213494] [LN:SPU47687] [AC:U47687] [PN:immunoglobulin A1 protease] [GN:iga] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae strain=R6]	149
SPX1169	1169	3830	76	228	138	8.00E-12	[GI:6911257] [LN:AF221126] [AC:AF221126] [PN:putative zinc metalloprotease] [GN:zmpB] [OR:Streptococcus pneumoniae]	115
SPX1170	1170	3831	284	852	707	2.50E-96	[LN:F69880] [AC:F69880] [PN:conserved hypothetical protein ylqF] [GN:ylqF] [CL:conserved hypothetical protein MG442] [OR:Bacillus subtilis]	139
SPX1171	1171	3832	260	780	615	3.40E-79	[LN:C69693] [AC:C69693] [PN:ribonuclease H rnh] [GN:rnh] [CL:ribonuclease HII] [OR:Bacillus subtilis]	101
SPX1172	1172	3833	517	1551	261	1.40E-51	[LN:YADQ_ECOLI] [AC:P37019:P77394] [GN:YADQ] [OR:Escherichia coli] [DE:HYPOTHETICAL 46.0 KD PROTEIN IN HEML-PFS INTERGENIC REGION] [SP:P37019:P77394]	149
SPX1173	1173	3834	357	1071	1529	9.60E-206	[GI:3152920] [LN:AF065141] [AC:AF065141] [PN:unknown] [OR:Streptococcus mutans]	79
SPX1174	1174	3835	330	990	577	6.60E-108	[LN:G69830] [AC:G69830] [PN:lipoate-protein ligase homolog yhfJ] [GN:yhfJ] [CL:lipoate-protein ligase] [OR:Bacillus subtilis]	125
SPX1175	1175	3836	568	1704	334	6.80E-104	[LN:I40794] [AC:I40794] [PN:dihydrolipoamide dehydrogenase] [OR:Clostridium magnum] [EC:1.8.1.4]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1176	1176	3837	376	1128	366	8.20E-70	"[LN:I40793] [AC:I40793] [PN:dihydrolipoamide S-acetyltransferase,] [CL:dihydrolipoamide acetyltransferase:lipoyl/biotin-binding homology] [OR:Clostridium magnum] [EC:2.3.1.12]"	177
SPX1177	1177	3838	331	993	550	5.00E-108	"[LN:I40791] [AC:I40791] [PN:acetoin dehydrogenase (TPP-dependent), beta chain] [CL:pyruvate dehydrogenase (lipoamide) beta chain] [OR:Clostridium magnum] [EC:1.-.-.-]"	168
SPX1178	1178	3839	146	438			NO-HIT	6
SPX1179	1179	3840	104	312	136	6.90E-13	[LN:G72548] [AC:G72548] [PN:hypothetical protein APE1675] [GN:APE1675] [OR:Aeropyrum pernix]	92
SPX1180	1180	3841	323	969	554	6.20E-107	"[LN:I40790] [AC:I40790] [PN:acetoin dehydrogenase (TPP-dependent), alpha chain] [CL:pyruvate dehydrogenase (lipoamide) alpha chain:thiamin pyrophosphate-binding domain homology] [OR:Clostridium magnum] [EC:1.-.-.-]"	216
SPX1181	1181	3842	454	1362	482	2.40E-117	[LN:F69906] [AC:F69906] [PN:conserved hypothetical protein yojI] [GN:yojI] [CL:conserved hypothetical protein HI1612] [OR:Bacillus subtilis]	140
SPX1182	1182	3843	87	261			NO-HIT	6
SPX1183	1183	3844	423	1269	351	1.70E-86	"[LN:PYRC_LACLE] [AC:P48795] [GN:PYRC] [OR:Lactobacillus leichmannii] [EC:3.5.2.3] [DE:DIHYDROOROTASE, (DHOASE)] [SP:P48795]"	125
SPX1184	1184	3845	155	465	845	7.70E-113	"[LN:MUTX_STRPN] [AC:P41354] [GN:MUTX] [OR:Streptococcus pneumoniae] [EC:3.6.1.-] [DE:(8-OXO-DGTPASE), (DGTP PYROPHOSPHOHYDROLASE)] [SP:P41354]"	144
SPX1185	1185	3846	222	666	1123	8.50E-151	"[LN:UNG_STRPN] [AC:P23379] [GN:UNG] [OR:Streptococcus pneumoniae] [EC:3.2.2.-]	127

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:URACIL-DNA GLYCOSYLASE, (UDG)] [SP:P23379]”	
SPX1186	1186	3847	201	603			NO-HIT	6
SPX1187	1187	3848	238	714	260	5.40E-57	[LN:E69814] [AC:E69814] [PN:conserved hypothetical protein yfnB] [GN:yfnB] [CL:Alcaligenes eutrophus phosphoglycolate phosphatase] [OR:Bacillus subtilis]	153
SPX1188	1188	3849	76	228	66	1.60E-05	[LN:H75623] [AC:H75623] [PN:conserved hypothetical protein] [GN:DRB0052] [OR:Deinococcus radiodurans]	101
SPX1189	1189	3850	852	2556	412	3.10E-104	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX1190	1190	3851	830	2490	397	3.60E-100	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX1191	1191	3852	411	1233	1763	4.80E-237	[GI:4958916] [LN:AB027569] [AC:AB027569] [PN:phosphoenolpyruvate-protein phosphotransferase] [GN:ptsI] [OR:Streptococcus bovis] [SR:Streptococcus bovis (strain:JB1) DNA]	169
SPX1192	1192	3853	46	138	144	1.50E-15	[LN:YKXH_BACSU] [AC:P08876] [GN:YKXH] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 7.4 KD PROTEIN IN PTSX OPERON (PROTEIN K)] [SP:P08876]	132
SPX1193	1193	3854	88	264	402	2.30E-50	[LN:PTHP_STRMU] [AC:P45596] [GN:PTSH] [OR:Streptococcus mutans] [DE:PHOSPHOCARRIER PROTEIN HPR (HISTIDINE-CONTAINING PROTEIN)] [SP:P45596]	138
SPX1194	1194	3855	73	219	233	1.50E-27	“[LN:NRDH_LACLC] [AC:Q48708] [GN:NRDH] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [DE:GLUTAREDOXIN-LIKE PROTEIN NRDH] [SP:Q48708]”	154
SPX1195	1195	3856	720	2160	1042	2.20E-255	[GI:3077613] [LN:CANRDFGEN]	125

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Y09572] [PN:ribonucleotide reductase subunit R1E] [GN:nrdE] [OR:Corynebacterium ammoniagenes]	
SPX1196	1196	3857	63	189			NO-HIT	6
SPX1197	1197	3858	321	963	818	2.90E-116	[GI:3068719] [LN:AF050168] [AC:AF050168] [PN:ribonucleoside diphosphate reductase small] [GN:nrdF] [OR:Corynebacterium ammoniagenes]	132
SPX1198	1198	3859	254	762	371	2.50E-80	"[LN:LACR_LACLA] [AC:P18816] [GN:LACR] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:LACTOSE PHOSPHOTRANSFERASE SYSTEM REPRESSOR] [SP:P18816]"	163
SPX1199	1199	3860	84	252			NO-HIT	6
SPX1200	1200	3861	472	1416	2452	0	"[LN:LACG_LACLA] [AC:P11546] [GN:LACG] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:3.2.1.85] [DE:GALACTOHYDROLASE) (PGALASE) (P-BETA-GAL) (PBG)] [SP:P11546]"	180
SPX1201	1201	3862	565	1695	1628	4.50E-299	"[LN:PTLB_LACLA] [AC:P23531] [GN:LACE] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EII-LAC)] [SP:P23531]"	157
SPX1202	1202	3863	106	318	454	2.90E-58	"[LN:PTLA_LACLA] [AC:P23532] [GN:LACF] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EIII-LAC)] [SP:P23532]"	158
SPX1203	1203	3864	279	837	487	1.40E-75	[LN:LACT_LACCA] [AC:P24401] [GN:LACT] [OR:Lactobacillus casei] [DE:TRANSCRIPTION ANTITERMINATOR LACT] [SP:P24401]	113
SPX1204	1204	3865	270	810	1232	1.10E-164	"[LN:LACD_LACLA]	182

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P26593] [GN:LACD] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.2.40] [DE:ALDOLASE) (D-TAGATOSE-1,6-BISPHOSPHATE ALDOLASE)] [SP:P26593]"	
SPX1205	1205	3866	273	819	810	9.10E-118	"[LN:LACC_LACLA] [AC:P23391] [GN:LACC] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.1.144] [DE:TAGATOSE-6-PHOSPHATE KINASE, (PHOSPHOTAGATOKINASE)] [SP:P23391]"	185
SPX1206	1206	3867	163	489			NO-HIT	6
SPX1207	1207	3868	172	516	810	5.00E-108	"[LN:LACB_LACLA] [AC:P23495] [GN:LACB] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:5.3.1.26] [DE:GALACTOSE-6-PHOSPHATE ISOMERASE LACB SUBUNIT,] [SP:P23495]"	179
SPX1208	1208	3869	142	426	612	2.90E-80	"[LN:LACA_LACLA] [AC:P23494] [GN:LACA] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:5.3.1.26] [DE:GALACTOSE-6-PHOSPHATE ISOMERASE LACA SUBUNIT,] [SP:P23494]"	179
SPX1209	1209	3870	97	291	57	0.00024	[GI:712785] [LN:S74218] [AC:S74218] [GN:E9] [OR:reindeer papillomavirus]	72
SPX1210	1210	3871	159	477	507	1.70E-65	"[LN:LAXP_LACLA] [AC:P23496] [GN:LACX] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:LACX PROTEIN, PLASMID] [SP:P23496]"""	141
SPX1211	1211	3872	207	621	374	1.20E-65	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1212	1212	3873	107	321	193	4.90E-21	[LN:S52544] [AC:S52544]	71

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:ISL2 protein] [OR:Lactobacillus helveticus]	
SPX1213	1213	3874	101	303			NO-HIT	6
SPX1214	1214	3875	148	444	88	1.30E-10	“[GI:4512375] [LN:AB011837] [AC:AB011837] [PN:phosphotransferase system (PTS)] [GN:fruA] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA, clone_lib:lambda no.9]”	181
SPX1215	1215	3876	431	1293	2159	2.90E-302	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1216	1216	3877	63	189	120	2.10E-12	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1217	1217	3878	417	1251	1395	2.00E-200	“[LN:THD1_LACLA] [AC:Q02145] [GN:ILVA] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.2.1.16] [DE:DEAMINASE)] [SP:Q02145]”	144
SPX1218	1218	3879	119	357			NO-HIT	6
SPX1219	1219	3880	78	234			NO-HIT	6
SPX1220	1220	3881	89	267			NO-HIT	6
SPX1221	1221	3882	341	1023	1329	9.50E-179	“[LN:ILVC_LACLA] [AC:Q02138] [GN:ILVC] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:1.1.1.86] [DE:ISOMEROREDUCTASE) (ALPHA-KETO-BETA-HYDROXYLACIL REDUCTOISOMERASE)] [SP:Q02138]”	199
SPX1222	1222	3883	167	501	477	4.10E-62	“[LN:ILVN_LACLA] [AC:Q02140] [GN:ILVN] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.3.18] [DE:(ACETOHYDROXY-ACID SYNTHASE SMALL SUBUNIT) (ALS)] [SP:Q02140]”	182
SPX1223	1223	3884	567	1701	1207	2.30E-266	“““[LN:ILVB_LACLA] [AC:Q02137] [GN:ILVB] [OR:Lactococcus lactis]	182

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SR:,subsplactis:Streptococcus lactis] [EC:4.1.3.18] [DE:(ACETOHYDROXY-ACID SYNTHASE LARGE SUBUNIT) (ALS)] [SP:Q02137]”	
SPX1224	1224	3885	679	2037	565	7.80E-167	[LN:E69879] [AC:E69879] [PN:conserved hypothetical protein yloV] [GN:yloV] [CL:Mycoplasma genitalium hypothetical protein MG369] [OR:Bacillus subtilis]	151
SPX1225	1225	3886	63	189	222	4.40E-26	[LN:RL28_BACSU] [AC:P37807] [GN:RPMB] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L28] [SP:P37807]	103
SPX1226	1226	3887	82	246	123	5.10E-19	[GI:6760462] [LN:AF224467] [AC:AF224467] [PN:putative glycosyl transferase] [OR:Haemophilus ducreyi]	100
SPX1227	1227	3888	66	198	157	8.80E-16	[GI:6760462] [LN:AF224467] [AC:AF224467] [PN:putative glycosyl transferase] [OR:Haemophilus ducreyi]	100
SPX1228	1228	3889	515	1545	771	2.70E-216	[GI:3256222] [LN:SAY14370] [AC:Y14370] [PN:peptide chain release factor 3] [GN:RF3] [OR:Staphylococcus aureus]	110
SPX1229	1229	3890	101	303	190	4.90E-21	[LN:A69795] [AC:A69795] [PN:conserved hypothetical protein yerL] [GN:yerL] [CL:probable glu-tRNA amidotransferase C chain] [OR:Bacillus subtilis]	145
SPX1230	1230	3891	489	1467	1333	1.60E-186	[LN:B69795] [AC:B69795] [PN:amidase homolog yerM] [GN:yerM] [CL:indoleacetamide hydrolase] [OR:Bacillus subtilis]	113
SPX1231	1231	3892	213	639	414	4.30E-94	[LN:T44293] [AC:T44293] [PN:hypothetical protein yerN [imported]] [CL:PET112 protein] [OR:Bacillus halodurans]	110
SPX1232	1232	3893	269	807	744	4.60E-97	[LN:C69795]	114

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:C69795] [PN:pet112-like protein homolog yerN] [GN:yerN] [CL:PET112 protein] [OR:Bacillus subtilis]	
SPX1233	1233	3894	348	1044	719	1.70E-145	“[LN:H69789] [AC:H69789] [PN:L-iditol 2-dehydrogenase, homolog ydjL:sorbitol dehydrogenase homolog] [GN:ydjL] [CL:alcohol dehydrogenase:long-chain alcohol dehydrogenase homology] [OR:Bacillus subtilis] [EC:1.1.1.14]”	216
SPX1234	1234	3895	67	201	138	4.80E-14	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1235	1235	3896	187	561	430	7.10E-54	[LN:EFP__BACSU] [AC:P49778] [GN:EFP] [OR:Bacillus subtilis] [DE:ELONGATION FACTOR P (EF-P)] [SP:P49778]	102
SPX1236	1236	3897	130	390	139	8.30E-19	[LN:YQHY__BACSU] [AC:P54519] [GN:YQHY] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 14.7 KD PROTEIN IN ACCC-FOLD INTERGENIC REGION] [SP:P54519]	137
SPX1237	1237	3898	141	423	105	1.00E-12	[GI:6580774] [LN:AF088897] [AC:AF088897:AF034613:AF086792:U63733] [PN:N-utilization substance protein B] [GN:nusB] [OR:Zymomonas mobilis]	137
SPX1238	1238	3899	190	570	948	7.40E-131	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX1239	1239	3900	261	783	1075	1.10E-148	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1240	1240	3901	278	834	80	9.90E-07	[LN:E69787] [AC:E69787] [PN:hypothetical protein ydiL] [GN:ydiL] [OR:Bacillus subtilis]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1241	1241	3902	84	252			NO-HIT	6
SPX1242	1242	3903	256	768	653	4.10E-90	[GI:285621] [LN:BACPK] [AC:D13095] [PN:undefined open reading frame] [OR:Bacillus stearothermophilus] [SR:Bacillus stearothermophilus (strain:NCA1503) DNA]	155
SPX1243	1243	3904	288	864	602	2.20E-92	"[LN:ACCD_SYNP7] [AC:Q54776] [GN:ACCD] [OR:Synechococcus sp] [SR:,strain PCC 7942:Anacystis nidulans R2] [EC:6.4.1.2] [DE:(EC 6.4.1.2) (ACCASE BETA CHAIN)] [SP:Q54776]"	168
SPX1244	1244	3905	456	1368	1101	1.50E-181	[LN:A69581] [AC:A69581] [PN:acetyl-CoA carboxylase (biotin carboxylase subunit) accC] [GN:accC] [CL:biotin carboxylase:biotin carboxylase homology] [OR:Bacillus subtilis]	170
SPX1245	1245	3906	141	423	299	1.60E-49	"[LN:D70065] [AC:D70065] [PN:(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase, ywpB] [GN:ywpB] [CL:(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase] [OR:Bacillus subtilis] [EC:4.2.1.-]"	201
SPX1246	1246	3907	162	486	248	1.10E-34	"[LN:C75558] [AC:C75558] [PN:acetyl-CoA carboxylase, bitoin carboxyl carrier protein] [GN:DR0118] [CL:biotin carboxyl carrier protein:lipoyl/biotin-binding homology] [OR:Deinococcus radiodurans]"	195
SPX1247	1247	3908	93	279	134	1.30E-13	[LN:FAB2_MYCTU] [AC:Q10525] [GN:RV2246:MTCY427.27] [OR:Mycobacterium tuberculosis] [EC:2.3.1.41] [DE:(BETA-KETOACYL-ACP SYNTHASE 2) (KAS 2)] [SP:Q10525]	152
SPX1248	1248	3909	360	1080	477	2.60E-113	"[LN:S77464] [AC:S77464] [PN:3-oxoacyl-[acyl-carrier-protein] synthase, beta chain:beta ketoacyl-acyl carrier protein synthase:protein sll1069:beta ketoacyl-acyl carrier protein synthase:protein sll1069] [GN:fabF] [OR:Synechocystis sp.]"	237

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1249	1249	3910	244	732	482	2.10E-76	[LN:FABG_CUPLA] [AC:P28643] [GN:CLKR27] [OR:Cuphea lanceolata] [EC:1.1.1.100] [DE:(3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)] [SP:P28643]	138
SPX1250	1250	3911	74	222	81	9.50E-05	[LN:H72482] [AC:H72482] [PN:hypothetical protein APE2504] [GN:APE2504] [OR:Aeropyrum pernix]	92
SPX1251	1251	3912	307	921	374	1.00E-85	"[LN:B41856] [AC:B41856:A42147:S20443:A64853] [PN:[acyl-carrier-protein] S-malonyltransferase,] [GN:fabD:tfpA] [CL:[acyl-carrier-protein] S-malonyltransferase:[acyl-carrier-protein] S-malonyltransferase homology] [OR:Escherichia coli] [EC:2.3.1.39]"	249
SPX1252	1252	3913	325	975	681	5.90E-101	[LN:A72335] [AC:A72335] [PN:conserved hypothetical protein] [GN:TM0800] [OR:Thermotoga maritima]	96
SPX1253	1253	3914	75	225	80	1.80E-12	[LN:ACP_BACSU] [AC:P80643:P51832] [GN:ACPA:ACPP] [OR:Bacillus subtilis] [DE:ACYL CARRIER PROTEIN (ACP)] [SP:P80643:P51832]	122
SPX1254	1254	3915	189	567	202	6.90E-38	"[LN:S75457] [AC:S75457] [PN:beta-ketoacyl-acyl carrier protein synthase III:protein slr1511:protein slr1511] [GN:fabH] [CL:3-oxoacyl-[acyl-carrier-protein] synthase III] [OR:Synechocystis sp.] [SR:PCC 6803, , PCC 6803] [SR:PCC 6803,]"	236
SPX1255	1255	3916	140	420	269	1.30E-46	[GI:7416010] [LN:AB025973] [AC:AB025973] [PN:3-oxoacyl-[acyl-carrier-protein] synthase III] [GN:accS] [OR:Lactobacillus plantarum] [SR:Lactobacillus plantarum (strain:L137) DNA]	177

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1256	1256	3917	66	198			NO-HIT	6
SPX1257	1257	3918	145	435	139	5.00E-12	[GI:4139249] [LN:AF110185] [AC:AF110185] [PN:unknown] [OR:Burkholderia pseudomallei]	84
SPX1258	1258	3919	66	198			NO-HIT	6
SPX1259	1259	3920	98	294	138	7.10E-13	[GI:3253198] [LN:AF029714] [AC:AF029714:Z71175] [PN:PhaB] [GN:phaB] [OR:Pseudomonas putida]	91
SPX1260	1260	3921	76	228	156	9.00E-18	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX1261	1261	3922	455	1365	1240	6.10E-166	[LN:A69763] [AC:A69763] [PN:homoserine dehydrogenase homolog yclM] [GN:yclM] [CL:aspartate kinase:aspartate kinase homology] [OR:Bacillus subtilis]	147
SPX1262	1262	3923	124	372	493	2.80E-64	[GI:5669858] [LN:AF130465] [AC:AF130465] [PN:unknown] [GN:manO] [OR:Streptococcus salivarius]	93
SPX1263	1263	3924	375	1125	564	8.90E-141	“[LN:SYS_BACSU] [AC:P37464] [GN:SERS] [OR:Bacillus subtilis] [EC:6.1.1.11] [DE:SERYL-TRNA SYNTHETASE, (SERINE--TRNA LIGASE) (SERRS)] [SP:P37464]”	145
SPX1264	1264	3925	91	273	101	6.30E-08	“[LN:SYS_MYCGE] [AC:P47251] [GN:SERS:MG005] [OR:Mycoplasma genitalium] [EC:6.1.1.11] [DE:SERYL-TRNA SYNTHETASE, (SERINE--TRNA LIGASE) (SERRS)] [SP:P47251]”	155
SPX1265	1265	3926	304	912			NO-HIT	6
SPX1266	1266	3927	69	207			NO-HIT	6
SPX1267	1267	3928	183	549	217	2.40E-57	[GI:7379894] [LN:NMA4Z2491] [AC:AL162755:AL157959] [PN:hypothetical protein NMA1203]	125

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:NMA1203] [OR:Neisseria meningitidis]	
SPX1268	1268	3929	441	1323	611	6.80E-118	“[LN:D71327] [AC:D71327] [PN:probable D-alanine glycine permease (dagA)] [GN:TP0414] [CL:sodium-dependent D-alanine/glycine transport protein] [OR:Treponema pallidum subsp. pallidum] [SR:, syphilis spirochete]”	210
SPX1269	1269	3930	779	2337	671	5.70E-155	[LN:MUS2_BACSU] [AC:P94545] [GN:MUTS2] [OR:Bacillus subtilis] [DE:MUTS2 PROTEIN] [SP:P94545]	92
SPX1270	1270	3931	251	753	175	7.10E-43	[GI:6746427] [LN:AF179847] [AC:AF179847] [PN:putative transposase] [OR:Lactococcus lactis]	90
SPX1271	1271	3932	65	195			NO-HIT	6
SPX1272	1272	3933	182	546	107	1.20E-15	[GI:3849798] [LN:U91581] [AC:U91581:U04057] [PN:putative transposase] [GN:tpase] [OR:Lactococcus lactis subsp. lactis]	118
SPX1273	1273	3934	183	549	95	4.30E-13	[LN:B69985] [AC:B69985] [PN:hypothetical protein yshB] [GN:yshB] [OR:Bacillus subtilis]	87
SPX1274	1274	3935	104	312			NO-HIT	6
SPX1275	1275	3936	294	882	1444	8.00E-194	“[LN:RNH2_STRPN] [AC:O07874] [GN:RNHB] [OR:Streptococcus pneumoniae] [EC:3.1.26.4] [DE:RIBONUCLEASE HII, (RNASE HII)] [SP:O07874]”	130
SPX1276	1276	3937	205	615	1065	6.70E-145	“[LN:LEP_STRPN] [AC:O07344] [GN:LEPB:SPI] [OR:Streptococcus pneumoniae] [EC:3.4.21.89] [DE:SIGNAL PEPTIDASE I, (SPASE I) (LEADER PEPTIDASE I)] [SP:O07344]”	155
SPX1277	1277	3938	93	279			NO-HIT	6
SPX1278	1278	3939	789	2367	738	5.80E-165	[LN:A69979] [AC:A69979]	103

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:conjugation transfer protein homolog yrrC] [GN:yrrC] [OR:Bacillus subtilis]	
SPX1279	1279	3940	428	1284	1691	1.40E-223	[GI:3549287] [LN:AF073922] [AC:AF073922] [PN:RopA] [GN:ropA] [OR:Streptococcus pyogenes]	88
SPX1280	1280	3941	379	1137	1350	2.20E-182	“[LN:MTLD__STRMU] [AC:Q02418] [GN:MTLD] [OR:Streptococcus mutans] [EC:1.1.1.17] [DE:MANNITOL-1-PHOSPHATE 5-DEHYDROGENASE,] [SP:Q02418]”	134
SPX1281	1281	3942	146	438	544	5.30E-71	[LN:PTMA__STRMU] [AC:Q02420] [GN:MTLF] [OR:Streptococcus mutans] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EIII-MTL)] [SP:Q02420]	119
SPX1282	1282	3943	652	1956	392	1.00E-47	[LN:YMTF__STRMU] [AC:Q02425] [OR:Streptococcus mutans] [DE:HYPOTHETICAL PROTEIN IN MTLF 5'REGION (ORFX) (FRAGMENT)] [SP:Q02425]	126
SPX1283	1283	3944	596	1788	1099	4.90E-194	[LN:PTMB__BACST] [AC:P50852] [GN:MTLA] [OR:Bacillus stearothermophilus] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EII-MTL)] [SP:P50852]	125
SPX1284	1284	3945	67	201			NO-HIT	6
SPX1285	1285	3946	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX1286	1286	3947	278	834	582	1.50E-140	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX1287	1287	3948	341	1023	1142	5.00E-219	[GI:1914875] [LN:SPZ82002] [AC:Z82002] [PN:PCPC] [GN:pcpC]	101

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[FN:unknown] [OR:Streptococcus pneumoniae]	
SPX1288	1288	3949	98	294	194	3.40E-25	[GI:4586910] [LN:AB017447] [AC:AB017447] [PN:protective antigen SpaA.1] [OR:Erysipelothrix rhusiopathiae] [SR:Erysipelothrix rhusiopathiae (strain:Fujisawa) DNA]	161
SPX1289	1289	3950	219	657	104	1.50E-08	[LN:S25140] [AC:S25140] [PN:serine proteinase homolog] [CL:staphylococcal serine proteinase] [OR:Enterococcus faecalis]	119
SPX1290	1290	3951	80	240	296	8.20E-35	[GI:1914874] [LN:SPZ82002] [AC:Z82002] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX1291	1291	3952	91	273	146	3.20E-14	[GI:4584089] [LN:BAJ10128] [AC:AJ010128] [PN:DNA alkylation repair enzyme] [GN:alkD] [FN:complements DNA alkylation repair deficiency] [OR:Bacillus cereus]	155
SPX1292	1292	3953	85	255	102	2.90E-08	[GI:4584089] [LN:BAJ10128] [AC:AJ010128] [PN:DNA alkylation repair enzyme] [GN:alkD] [FN:complements DNA alkylation repair deficiency] [OR:Bacillus cereus]	155
SPX1293	1293	3954	211	633	1040	1.50E-139	[GI:5830527] [LN:SPAJ6393] [AC:AJ006393] [PN:response regulator] [GN:rr03] [OR:Streptococcus pneumoniae]	104
SPX1294	1294	3955	332	996	1603	3.60E-220	[GI:5830526] [LN:SPAJ6393] [AC:AJ006393] [PN:histidine kinase] [GN:hk03] [OR:Streptococcus pneumoniae]	102
SPX1295	1295	3956	233	699	104	1.50E-11	[LN:G70045] [AC:G70045] [PN:hypothetical protein yvqF] [GN:yvqF] [OR:Bacillus subtilis]	87
SPX1296	1296	3957	337	1011	151	6.20E-42	[LN:D75084]	162

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:D75084] [PN:carotenoid biosynthetic gene erwcrts related PAB1662] [GN:PAB1662] [CL:carotenoid biosynthesis protein homolog] [OR:Pyrococcus abyssi]	
SPX1297	1297	3958	336	1008	67	0.00053	“[LN:KIME_METJA] [AC:Q58487] [GN:MJ1087] [OR:Methanococcus jannaschii] [EC:2.7.1.36] [DE:MEVALONATE KINASE, (MK)] [SP:Q58487]”	126
SPX1298	1298	3959	345	1035	141	5.90E-28	[GI:6625790] [LN:AF203779] [AC:AF203779] [PN:diphosphomevalonate decarboxylase MVD1] [GN:MVD1] [OR:Candida albicans]	116
SPX1299	1299	3960	293	879	134	1.60E-20	[LN:F72474] [AC:F72474] [PN:probable mevalonate kinase APE2439] [GN:APE2439] [OR:Aeropyrum pernix]	98
SPX1300	1300	3961	110	330	193	8.50E-21	[GI:1914875] [LN:SPZ82002] [AC:Z82002] [PN:PCPC] [GN:pcpC] [FN:unknown] [OR:Streptococcus pneumoniae]	101
SPX1301	1301	3962	292	876			NO-HIT	6
SPX1302	1302	3963	411	1233	97	4.90E-10	[GI:3582221] [LN:AE001272] [AC:AE001272] [PN:conserved hypothetical protein] [GN:ORF00049] [OR:Lactococcus lactis]	114
SPX1303	1303	3964	333	999	402	1.50E-124	[GI:1914875] [LN:SPZ82002] [AC:Z82002] [PN:PCPC] [GN:pcpC] [FN:unknown] [OR:Streptococcus pneumoniae]	101
SPX1304	1304	3965	371	1113	289	1.70E-89	[GI:1914875] [LN:SPZ82002] [AC:Z82002] [PN:PCPC] [GN:pcpC] [FN:unknown]	101

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX1305	1305	3966	230	690	212	7.30E-56	[GI:3599371] [LN:AF082668] [AC:AF082668] [PN:CsrR] [GN:csrR] [FN:negative regulator of hyaluronic acid capsule] [OR:Streptococcus pyogenes]	139
SPX1306	1306	3967	458	1374	801	5.10E-226	[GI:3033358] [LN:LLU74322] [AC:U74322] [PN:6-phosphogluconate dehydrogenase] [OR:Lactococcus lactis]	100
SPX1307	1307	3968	465	1395	109	1.10E-06	[GI:1684847] [LN:HSU77718] [AC:U77718] [PN:pinin] [OR:Homo sapiens] [SR:human]	78
SPX1308	1308	3969	386	1158	678	1.20E-133	[LN:YPSC__BACSU] [AC:P50840] [GN:YPSC] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 43.5 KD PROTEIN IN COTD-KDUD INTERGENIC REGION PRECURSOR] [SP:P50840]	147
SPX1309	1309	3970	110	330	166	6.70E-18	[LN:YPSB__BACSU] [AC:P50839] [GN:YPSB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 11.6 KD PROTEIN IN COTD-KDUD INTERGENIC REGION] [SP:P50839]	137
SPX1310	1310	3971	176	528	147	8.20E-19	[LN:YPSA__BACSU] [AC:P50838] [GN:YPSA] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 21.1 KD PROTEIN IN COTD-KDUD INTERGENIC REGION] [SP:P50838]	137
SPX1311	1311	3972	199	597	1043	3.40E-141	[LN:YPOA__STRPN] [AC:P38034] [OR:Streptococcus pneumoniae] [DE:HYPOTHETICAL 23.1 KD PROTEIN IN PONA 5'REGION] [SP:P38034]	120
SPX1312	1312	3973	720	2160	3734	0	[GI:6563337] [LN:AF210745] [AC:AF210745] [PN:penicillin-binding protein 1A] [GN:pbp1a] [OR:Streptococcus pneumoniae]	116
SPX1313	1313	3974	68	204			NO-HIT	6
SPX1314	1314	3975	1768	5304	1762	9.10E-232	[GI:1658320]	77

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SPDEXCAP] [AC:Z47210] [GN:orf] [OR:Streptococcus pneumoniae]	
SPX1315	1315	3976	661	1983	3378	0	LN:ALIA__STRPN] [AC:P35592:Q54782:O54620:O52228] [GN:ALIA:EXP1:PLPA] [OR:Streptococcus pneumoniae] [DE:OLIGOPEPTIDE-BINDING PROTEIN ALIA PRECURSOR (EXPORTED PROTEIN 1)] [SP:P35592:Q54782:O54620:O52228]	201
SPX1316	1316	3977	78	234	87	0.00034	“[GI:5824139] [LN:POL245436] [AC:AJ245436:J04618:J04619:S50571:X52935:X65936] [PN:hypothetical protein, 57.8 kD] [OR:Pseudomonas putida]”	137
SPX1317	1317	3978	150	450	94	3.20E-15	“[LN:Y4HP__RHISN] [AC:P50360] [GN:Y4HP] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 61.7 KD PROTEIN Y4HP] [SP:P50360]”	128
SPX1318	1318	3979	58	174			NO-HIT	6
SPX1319	1319	3980	86	258	80	8.90E-06	“[LN:Y4HP__RHISN] [AC:P50360] [GN:Y4HP] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 61.7 KD PROTEIN Y4HP] [SP:P50360]”	128
SPX1320	1320	3981	143	429	84	9.20E-08	“[GI:6009407] [LN:AB024946] [AC:AB024946] [GN:orf31] [OR:Escherichia coli] [SR:Escherichia coli (sub__species:enteropathogenic, strain:B171]”	140
SPX1321	1321	3982	284	852	1480	1.00E-198	[GI:2804705] [LN:AF030362] [AC:AF030362] [PN:dTDP-L-rhamnose synthase] [GN:cpsO] [OR:Streptococcus pneumoniae]	110
SPX1322	1322	3983	350	1050	1853	2.90E-249	[GI:4200435] [LN:AF026471] [AC:AF026471] [PN:Cps2N] [GN:cps2N] [OR:Streptococcus pneumoniae]	92
SPX1323	1323	3984	199	597	1027	4.50E-137	“[GI:3907611] [LN:AF094575]	131

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AF094575] [PN:dTDP-4-keto-6-deoxyglucose-3,5-epimerase] [GN:cps19aM] [OR:Streptococcus pneumoniae]”	
SPX1324	1324	3985	290	870	1430	7.20E-193	[GI:4406249] [LN:AF105113] [AC:AF105113] [PN:glucose-1-phosphate thymidylyl transferase] [GN:cps19AL] [OR:Streptococcus pneumoniae]	131
SPX1325	1325	3986	476	1428	2017	3.10E-287	[GI:3907608] [LN:AF094575] [AC:AF094575] [PN:putative repeat unit transporter Cps19aJ] [GN:cps19aJ] [OR:Streptococcus pneumoniae]	129
SPX1326	1326	3987	95	285			NO-HIT	6
SPX1327	1327	3988	448	1344	148	6.90E-31	[GI:3907607] [LN:AF094575] [AC:AF094575] [PN:polysaccharide polymerase Cps19aI] [GN:cps19aI] [OR:Streptococcus pneumoniae]	122
SPX1328	1328	3989	329	987	178	2.00E-31	[LN:T00087] [AC:T00087] [PN:rhamnosyltransferase] [GN:rgpBc] [OR:Streptococcus mutans]	86
SPX1329	1329	3990	242	726	93	1.90E-06	[LN:T35395] [AC:T35395] [PN:probable transferase] [GN:SC6A5.04] [OR:Streptomyces coelicolor]	92
SPX1330	1330	3991	186	558	154	3.00E-14	[GI:2209209] [LN:AF004325] [AC:AF004325] [PN:unknown] [GN:cps19bG] [OR:Streptococcus pneumoniae]	96
SPX1331	1331	3992	315	945	96	6.50E-16	[GI:3132871] [LN:AF019375] [AC:AF019375] [PN:UDP-glucose:(glucosyl) LPS] [GN:waaR] [OR:Escherichia coli]	104
SPX1332	1332	3993	456	1368	1141	0	[GI:3907603] [LN:AF094575] [AC:AF094575] [PN:glucosyl-1-phosphate transferase Cps19aE] [GN:cps19aE]	129

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1333	1333	3994	227	681	1055	4.60E-145	[OR:Streptococcus pneumoniae] [GI:3907602] [LN:AF094575] [AC:AF094575] [PN:Cps19aD] [GN:cps19aD] [FN:chain length regulation and export]	136
SPX1334	1334	3995	232	696	1061	2.00E-141	[OR:Streptococcus pneumoniae] [GI:3907601] [LN:AF094575] [AC:AF094575] [PN:Cps19aC] [GN:cps19aC] [FN:chain length regulation and export]	136
SPX1335	1335	3996	260	780	1195	2.90E-162	[OR:Streptococcus pneumoniae] [GI:3907600] [LN:AF094575] [AC:AF094575] [PN:Cps19aB] [GN:cps19aB]	96
SPX1336	1336	3997	485	1455	2238	0	[OR:Streptococcus pneumoniae] [GI:3550627] [LN:SPAJ6986] [AC:AJ006986] [GN:cap33fA]	83
SPX1337	1337	3998	358	1074	1690	9.20E-228	[OR:Streptococcus pneumoniae] [LN:TRA2__STRPN] [AC:Q54513] [OR:Streptococcus pneumoniae] [DE:TRANSPOSASE FOR INSERTION SEQUENCE IS1202] [SP:Q54513]	116
SPX1338	1338	3999	120	360	581	6.60E-75	[LN:TRA2__STRPN] [AC:Q54513] [OR:Streptococcus pneumoniae] [DE:TRANSPOSASE FOR INSERTION SEQUENCE IS1202] [SP:Q54513]	116
SPX1339	1339	4000	317	951	267	3.80E-61	""""[GI:6332767] [LN:AB033763] [AC:AB033763:AB014419:AB014429:AB014439] [PN:hypothetical protein] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:NCTC10442) DNA, clone__lib:Lambda das]""""	194
SPX1340	1340	4001	179	537	551	1.10E-70	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX1341	1341	4002	161	483	550	2.00E-80	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3]	107

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus agalactiae]	
SPX1342	1342	4003	278	834	587	1.60E-140	[LN:A33595] [AC:A33595:A30868] [PN:probable transposase] [CL:transposase IS3] [OR:Streptococcus agalactiae]	107
SPX1343	1343	4004	537	1611	2499	0	“[GI:3320387] [LN:AF030373] [AC:AF030373] [PN:alpha, 1-6-glucosidase] [GN:dexB] [OR:Streptococcus pneumoniae]”	110
SPX1344	1344	4005	495	1485			NO-HIT	6
SPX1345	1345	4006	161	483	288	9.40E-46	[LN:D75280] [AC:D75280] [PN:conserved hypothetical protein] [GN:DR2387] [CL:conserved hypothetical protein HI0491] [OR:Deinococcus radiodurans]	143
SPX1346	1346	4007	702	2106	2049	0	[GI:871784] [LN:BOVCLPAB] [AC:L34677] [PN:Clp-like ATP-dependent protease binding subunit] [OR:Bos taurus] [SR:Bos taurus calf thymus cDNA to mRNA]	147
SPX1347	1347	4008	80	240			NO-HIT	6
SPX1348	1348	4009	327	981	1619	8.40E-229	[GI:4009489] [LN:AF068903] [AC:AF068903] [PN:undecaprenyl-phosphate-UDP-MurNAc-pentapeptide] [GN:mraY] [OR:Streptococcus pneumoniae]	132
SPX1349	1349	4010	751	2253	3812	0	[LN:PBPX__STRPN] [AC:P14677] [GN:PBPX] [OR:Streptococcus pneumoniae] [DE:PENICILLIN-BINDING PROTEIN 2X (PBP-2X) (PBP2X)] [SP:P14677]	131
SPX1350	1350	4011	106	318	505	1.40E-65	[GI:4009487] [LN:AF068903] [AC:AF068903] [PN:YllD] [GN:yllD] [OR:Streptococcus pneumoniae]	90
SPX1351	1351	4012	317	951	1604	2.70E-217	[GI:4009486] [LN:AF068903] [AC:AF068903] [PN:YllC] [GN:yllC] [OR:Streptococcus pneumoniae]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1352	1352	4013	65	195	323	1.20E-40	[GI:1536960] [LN:SOORFS] [AC:Z79691] [GN:yorfE] [FN:putative transcription regulator] [OR:Streptococcus pneumoniae]	115
SPX1353	1353	4014	142	426	291	1.40E-36	[GI:1524349] [LN:SOORFS] [AC:Z79691] [PN:OrfC] [GN:yorfC] [OR:Streptococcus pneumoniae]	87
SPX1354	1354	4015	94	282			NO-HIT	6
SPX1355	1355	4016	85	255	178	5.70E-20	[GI:1536959] [LN:SOORFS] [AC:Z79691] [PN:OrfB] [GN:yorfb] [OR:Streptococcus pneumoniae]	87
SPX1356	1356	4017	81	243	391	2.00E-51	[GI:1536959] [LN:SOORFS] [AC:Z79691] [PN:OrfB] [GN:yorfb] [OR:Streptococcus pneumoniae]	87
SPX1357	1357	4018	112	336	219	4.90E-27	[GI:1536959] [LN:SOORFS] [AC:Z79691] [PN:OrfB] [GN:yorfb] [OR:Streptococcus pneumoniae]	87
SPX1358	1358	4019	64	192	111	4.30E-10	[GI:1536959] [LN:SOORFS] [AC:Z79691] [PN:OrfB] [GN:yorfb] [OR:Streptococcus pneumoniae]	87
SPX1359	1359	4020	334	1002	1703	2.00E-229	"[GI:1536958] [LN:SOORFS] [AC:Z79691] [PN:RegR] [GN:regR] [FN:putative transcription regulator, member GalR] [OR:Streptococcus pneumoniae]"	139
SPX1360	1360	4021	634	1902	1094	3.90E-148	[GI:1524347] [LN:SOORFS] [AC:Z79691] [PN:OrfA] [GN:yorfA] [OR:Streptococcus pneumoniae]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1361	1361	4022	98	294	96	7.80E-07	[LN:D70408] [AC:D70408] [PN:conserved hypothetical protein aq_1254] [GN:aq_1254] [CL:yajC protein] [OR:Aquifex aeolicus]	120
SPX1362	1362	4023	273	819	268	3.50E-51	"[LN:PTPD_ECOLI] [AC:P42911] [GN:AGAD] [OR:Escherichia coli] [DE:ENZYME II, D COMPONENT)] [SP:P42911]"	102
SPX1363	1363	4024	260	780	170	4.30E-27	[GI:1732200] [LN:VFU65015] [AC:U65015] [PN:PTS permease for mannose subunit IIPMan] [GN:manY] [OR:Vibrio furnissii]	115
SPX1364	1364	4025	164	492	286	3.40E-44	"[LN:PTPV_ECOLI] [AC:P42904:P76669] [GN:AGAV] [OR:Escherichia coli] [EC:2.7.1.69] [DE:ENZYME II, B COMPONENT 2),] [SP:P42904:P76669]"	133
SPX1365	1365	4026	336	1008	475	2.70E-95	[GI:5869507] [LN:AB019619] [AC:AB019619] [PN:unsaturated glucuronyl hydrolase] [GN:ugl] [OR:Bacillus sp. GL1] [SR:Bacillus sp. GL1 (strain:GL1) DNA]	148
SPX1366	1366	4027	146	438			NO-HIT	6
SPX1367	1367	4028	145	435	129	3.30E-17	[GI:5669855] [LN:AF130465] [AC:AF130465] [PN:mannose-specific phosphotransferase system] [GN:manL] [OR:Streptococcus salivarius]	128
SPX1368	1368	4029	218	654	436	1.40E-55	[LN:YJGU_ECOLI] [AC:P39345] [GN:YJGU] [OR:Escherichia coli] [EC:1.-.-.-] [DE:(EC 1.-.-.-)] [SP:P39345]	102
SPX1369	1369	4030	125	375			NO-HIT	6
SPX1370	1370	4031	98	294	169	2.50E-19	[LN:YJGU_ECOLI] [AC:P39345] [GN:YJGU]	102

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Escherichia coli] [EC:1.-.-.-] [DE:(EC 1.-.-.-)] [SP:P39345]	
SPX1371	1371	4032	214	642			NO-HIT	6
SPX1372	1372	4033	334	1002	317	1.00E-46	[LN:G72422] [AC:G72422] [PN:2-keto-3-deoxygluconate kinase] [GN:TM0067] [CL:ribokinase] [OR:Thermotoga maritima]	112
SPX1373	1373	4034	210	630	271	6.10E-46	[LN:F72422] [AC:F72422] [PN:2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase] [GN:TM0066] [CL:2-dehydro-3-deoxyphosphogluconate aldolase] [OR:Thermotoga maritima]	190
SPX1374	1374	4035	118	354	264	1.50E-50	[GI:1841332] [LN:D64071] [AC:D64071] [PN:putative protein of unknown function] [OR:Actinobacillus actinomycetemcomitans] [SR:Actinobacillus actinomycetemcomitans (strain:Y4) DNA]	178
SPX1375	1375	4036	1079	3237	4863	0	"[LN:HYSA_STRPN] [AC:Q54873:Q54874] [OR:Streptococcus pneumoniae] [EC:4.2.2.1] [DE:HYALURONATE LYASE PRECURSOR, (HYALURONIDASE) (HYASE)] [SP:Q54873:Q54874]"	156
SPX1376	1376	4037	159	477	482	1.90E-62	[GI:2598550] [LN:LLAJ109] [AC:AJ000109] [PN:gluthatione peroxidase] [GN:gpo] [OR:Lactococcus lactis]	100
SPX1377	1377	4038	738	2214	181	4.50E-33	[GI:6686567] [LN:AAC252161] [AC:AJ252161] [PN:putative alpha-glucosidase] [GN:glcA] [OR:Alicyclobacillus acidocaldarius]	120
SPX1378	1378	4039	61	183			NO-HIT	6
SPX1379	1379	4040	449	1347	231	3.40E-63	"[LN:PTCC_BACST] [AC:Q45400] [GN:CELB] [OR:Bacillus stearothermophilus] [DE:PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT)] [SP:Q45400]"	157
SPX1380	1380	4041	169	507			NO-HIT	6
SPX1381	1381	4042	105	315	234	2.00E-27	[LN:PTCA_BACSU]	121

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P46319] [GN:CELC:LICA] [OR:Bacillus subtilis] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EIII-CEL)] [SP:P46319]	
SPX1382	1382	4043	660	1980	138	9.20E-51	[LN:CELR__BACSU] [AC:P46321] [GN:CELR:LICR] [OR:Bacillus subtilis] [DE:PUTATIVE CEL OPERON REGULATOR] [SP:P46321]	112
SPX1383	1383	4044	114	342	84	2.40E-08	[LN:PTCB__BACST] [AC:Q45399] [GN:CELA] [OR:Bacillus stearothermophilus] [EC:2.7.1.69] [DE:(EC 2.7.1.69)] [SP:Q45399]	115
SPX1384	1384	4045	80	240	142	1.20E-15	[LN:Y110__METJA] [AC:Q57574] [GN:MJ0110] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ0110] [SP:Q57574]	114
SPX1385	1385	4046	493	1479	797	2.50E-225	[LN:JE0395] [AC:JE0395] [PN:phospho-beta-galactosidase I] [CL:Agrobacterium beta-glucosidase] [OR:Lactobacillus gasseri]	120
SPX1386	1386	4047	122	366			NO-HIT	6
SPX1387	1387	4048	69	207	359	2.90E-46	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX1388	1388	4049	131	393	456	1.30E-58	[LN:RS9__BACST] [AC:P07842] [GN:RPSI] [OR:Bacillus stearothermophilus] [DE:30S RIBOSOMAL PROTEIN S9 (BS10)] [SP:P07842]	118
SPX1389	1389	4050	149	447	487	2.30E-62	[LN:RL13__STACA] [AC:Q00990] [GN:RPLM] [OR:Staphylococcus carnosus] [DE:50S RIBOSOMAL PROTEIN L13] [SP:Q00990]	109
SPX1390	1390	4051	122	366			NO-HIT	6
SPX1391	1391	4052	106	318			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1392	1392	4053	131	393	630	8.40E-84	"[LN:SULD__STRPN] [AC:P22291:O33697] [GN:SULD] [OR:Streptococcus pneumoniae] [EC:4.1.2.25:2.7.6.3] [DE:HYDROXYMETHYL-7,8-DIHYDROPTERIN PYROPHOSPHOKINASE) (PPPK)]] [SP:P22291:O33697]"	181
SPX1393	1393	4054	170	510	834	1.40E-111	"[LN:A36704] [AC:A36704] [PN:bifunctional folate biosynthesis enzyme sulD:6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase:7, 8-dihydro-6-hydroxymethylpterin pyrophosphokinase] [GN:sulD] [OR:Streptococcus pneumoniae]"	220
SPX1394	1394	4055	185	555	938	4.90E-127	"[LN:GCH1__STRPN] [AC:P51595] [GN:SULC] [OR:Streptococcus pneumoniae] [EC:3.5.4.16] [DE:GTP CYCLOHYDROLASE I, (GTP-CH-I)] [SP:P51595]"	133
SPX1395	1395	4056	441	1323	2219	0	146	
SPX1396	1396	4057	327	981	1485	7.00E-201	[LN:DHPS__STRPN] [AC:P05382] [GN:SULA] [OR:Streptococcus pneumoniae] [EC:2.5.1.15] [DE:PYROPHOSPHORYLASE)] [SP:P05382]	117
SPX1397	1397	4058	236	708	143	3.00E-13	[GI:1402529] [LN:D78257] [AC:D78257] [PN:ORF8] [GN:orf8] [OR:Enterococcus faecalis] [SR:Enterococcus faecalis plasmid:pYI17 DNA]	128
SPX1398	1398	4059	62	186			NO-HIT	6
SPX1399	1399	4060	494	1482	447	1.70E-116	[GI:6624219] [LN:AB023289] [AC:AB023289] [GN:ORFX] [OR:Pseudomonas fluorescens] [SR:Pseudomonas fluorescens (strain:No.33) DNA]	127
SPX1400	1400	4061	271	813	199	8.70E-36	[LN:F69841] [AC:F69841] [PN:conserved hypothetical protein yitU] [GN:yitU] [CL:Methanobacterium thermoautotrophicum conserved hypothetical protein MTH1071] [OR:Bacillus subtilis]	178
SPX1401	1401	4062	311	933	861	2.70E-156	[GI:7379448] [LN:NMA3Z2491]	124

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AL162754:AL157959] [PN:putative alcohol dehydrogenase] [GN:adhA] [OR:Neisseria meningitidis]	
SPX1402	1402	4063	333	999	795	2.40E-188	[GI:5669855] [LN:AF130465] [AC:AF130465] [PN:mannose-specific phosphotransferase system] [GN:manL] [OR:Streptococcus salivarius]	128
SPX1403	1403	4064	268	804	974	1.10E-144	[GI:5669856] [LN:AF130465] [AC:AF130465] [PN:mannose-specific phosphotransferase system] [GN:manM] [OR:Streptococcus salivarius]	128
SPX1404	1404	4065	304	912	1157	3.60E-158	[GI:5669857] [LN:AF130465] [AC:AF130465] [PN:mannose-specific phosphotransferase system] [GN:manN] [OR:Streptococcus salivarius]	128
SPX1405	1405	4066	445	1335	1900	4.00E-256	"[LN:PEPC_STRTR] [AC:Q56115] [GN:PEPC] [OR:Streptococcus thermophilus] [EC:3.4.22.-] [DE:AMINOPEPTIDASE C,] [SP:Q56115]"	120
SPX1406	1406	4067	241	723	223	1.50E-59	[GI:7328274] [LN:SAY14816] [AC:Y14816] [PN:hypothetical protein] [GN:ORF231] [OR:Staphylococcus aureus]	103
SPX1407	1407	4068	137	411	328	2.80E-41	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1408	1408	4069	77	231			NO-HIT	6
SPX1409	1409	4070	77	231	154	1.90E-17	[LN:B69770] [AC:B69770] [PN:conserved hypothetical protein ydaS] [GN:ydaS] [OR:Bacillus subtilis]	97
SPX1410	1410	4071	414	1242	1792	6.20E-241	[GI:4731022] [LN:AF102860] [AC:AF102860] [PN:aminopeptidase PepS] [GN:pepS]	107

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus thermophilus]	
SPX1411	1411	4072	60	180			NO-HIT	6
SPX1412	1412	4073	77	231			NO-HIT	6
SPX1413	1413	4074	93	279	110	1.90E-19	"[LN:B71883] [AC:B71883] [PN:hypothetical protein jhp0831] [GN:jhp0831] [CL:conserved hypothetical protein HI0711] [OR:Helicobacter pylori] [SR:strain J99, , strain J99] [SR:strain J99,]"	188
SPX1414	1414	4075	88	264	80	0.00012	[LN:DINJ_ECOLI] [AC:Q47150] [GN:DINJ] [OR:Escherichia coli] [DE:DNA-DAMAGE-INDUCIBLE PROTEIN J] [SP:Q47150]	107
SPX1415	1415	4076	1464	4392	1689	0	"[GI:143342] [LN:BACPOLC] [AC:M22996] [GN:polC] [OR:Bacillus subtilis] [SR:B.subtilis SB19 DNA, clone PRO10]"	109
SPX1416	1416	4077	694	2082	2833	0	[LN:T44380] [AC:T44380] [PN:translation elongation factor EF-G fus [imported]] [GN:fus] [CL:translation elongation factor G:translation elongation factor Tu homology] [OR:Bacillus halodurans]	191
SPX1417	1417	4078	119	357	86	5.20E-10	[LN:YHA2_EIKCO] [AC:P35649] [OR:Eikenella corrodens] [DE:HYPOTHETICAL 66.3 KD PROTEIN IN HAG2 5'REGION] [SP:P35649]	115
SPX1418	1418	4079	150	450	88	1.30E-08	[LN:YHA2_EIKCO] [AC:P35649] [OR:Eikenella corrodens] [DE:HYPOTHETICAL 66.3 KD PROTEIN IN HAG2 5'REGION] [SP:P35649]	115
SPX1419	1419	4080	144	432	173	1.90E-17	[GI:7576785] [LN:AF214675] [AC:AF214675] [PN:unknown in vivo-induced protein 131-19] [OR:Pseudomonas aeruginosa]	112
SPX1420	1420	4081	157	471	633	1.10E-83	[LN:F69699] [AC:F69699:S11357] [PN:ribosomal protein S7 (rpsG):ribosomal protein BS7] [GN:rpsG] [CL:Escherichia coli ribosomal protein S7] [OR:Bacillus subtilis]	161

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1421	1421	4082	138	414	699	1.70E-91	[LN:RS12__STRPN] [AC:P30891] [GN:RPSL:STR] [OR:Streptococcus pneumoniae] [DE:30S RIBOSOMAL PROTEIN S12] [SP:P30891]	114
SPX1422	1422	4083	1205	3615	1478	0	[GI:1665738] [LN:D78258] [AC:D78258] [PN:alkaline amylopullulanase] [OR:Bacillus sp.] [SR:Bacillus sp. (strain:KSM-1378) DNA]	125
SPX1423	1423	4084	308	924	655	1.70E-115	"[GI:2735505] [LN:SCU96107] [AC:U96107] [PN:N5,N10-methylenetetrahydromethanopterin] [OR:Staphylococcus carnosus]"	114
SPX1424	1424	4085	603	1809	963	3.70E-194	"[LN:B69633] [AC:B69633] [PN:glutamine--fructose-6-phosphate transaminase (isomerizing), glmS] [GN:glmS] [CL:glutamine--fructose-6-phosphate aminotransferase (isomerizing)] [OR:Bacillus subtilis] [EC:2.6.1.16]"	210
SPX1425	1425	4086	460	1380	851	2.00E-207	[LN:D69785] [AC:D69785] [PN:beta-glucosidase homolog ydhP] [GN:ydhP] [CL:Agrobacterium beta-glucosidase] [OR:Bacillus subtilis]	127
SPX1426	1426	4087	347	1041	298	1.30E-82	"[LN:G69682] [AC:G69682] [PN:proline--tRNA ligase, proS:prolyl-tRNA synthetase] [GN:proS] [CL:proline--tRNA ligase] [OR:Bacillus subtilis] [EC:6.1.1.15]"	153
SPX1427	1427	4088	336	1008	691	5.20E-109	"[LN:G69682] [AC:G69682] [PN:proline--tRNA ligase, proS:prolyl-tRNA synthetase] [GN:proS] [CL:proline--tRNA ligase] [OR:Bacillus subtilis] [EC:6.1.1.15]"	153
SPX1428	1428	4089	420	1260	436	8.50E-136	[GI:5714510] [LN:AF152237] [AC:AF152237] [PN:Eep] [GN:eep] [FN:determinant for enhanced expression of]	129

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Enterococcus faecalis]	
SPX1429	1429	4090	268	804	202	1.20E-51	[LN:CDSA_BACSU] [AC:O31752] [GN:CDSA] [OR:Bacillus subtilis] [EC:2.7.7.41] [DE:SYNTHASE)] [SP:O31752]	101
SPX1430	1430	4091	259	777	490	2.80E-90	[LN:UPPS_BACSU] [AC:O31751] [GN:UPPS] [OR:Bacillus subtilis] [EC:2.5.1.31] [DE:(DI-TRANS-POLY-CIS-DECAPRENYLCISTRANSFERASE)] [SP:O31751]	136
SPX1431	1431	4092	191	573	127	1.50E-11	[LN:S70841] [AC:S70841] [PN:hypothetical transmembrane protein (sipS 3'region)] [CL:Bradyrhizobium japonicum hypothetical transmembrane protein (sipS 3'region)] [OR:Bradyrhizobium japonicum]	192
SPX1432	1432	4093	333	999	1045	1.50E-140	[LN:H81105] [AC:H81105] [PN:Holliday junction DNA helicase RuvB NMB1243 [imported]] [GN:NMB1243] [OR:Neisseria meningitidis]	124
SPX1433	1433	4094	35	105	76	0.00015	[LN:H72613] [AC:H72613] [PN:hypothetical protein APE1371] [GN:APE1371] [OR:Aeropyrum pernix]	92
SPX1434	1434	4095	139	417	103	1.50E-06	[GI:5616248] [LN:AF158628] [AC:AF158628] [PN:hypothetical protein] [OR:Prochlorococcus PCC9511]	95
SPX1435	1435	4096	232	696	482	1.50E-73	“[LN:YIW2_YEAST] [AC:P40586] [GN:YIR042C] [OR:Saccharomyces cerevisiae] [SR:,Baker's yeast] [DE:HYPOTHETICAL 27.4 KD PROTEIN IN HYR1 3'REGION] [SP:P40586]”	155
SPX1436	1436	4097	64	192			NO-HIT	6
SPX1437	1437	4098	834	2502	944	0	“LN:C70131] [AC:C70131] [PN:leucine--tRNA ligase, leuS:leucyl-tRNA synthetase:leucyl-tRNA synthetase] [CL:leucine--tRNA ligase] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete] [EC:6.1.1.4]”	199

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1438	1438	4099	363	1089	998	1.00E-137	“[LN:GLDA_BACST] [AC:P32816] [GN:GLDA:GLD] [OR:Bacillus stearothermophilus] [EC:1.1.1.6] [DE:GLYCEROL DEHYDROGENASE, (GLDH)] [SP:P32816]”	137
SPX1439	1439	4100	223	669	356	8.20E-53	[GI:2582658] [LN:CBAJ2527] [AC:AJ002527] [PN:OrfX] [GN:orfX] [FN:putative transaldolase (37.4% identity to talC] [OR:Clostridium beijerinckii]	142
SPX1440	1440	4101	816	2448	992	1.80E-275	[LN:PFLF_ECOLI] [AC:P75793] [GN:YBIW] [OR:Escherichia coli] [EC:2.3.1.54] [DE:LYASE 3)] [SP:P75793]	99
SPX1441	1441	4102	427	1281	337	4.90E-89	“[LN:PTCC_BACST] [AC:Q45400] [GN:CELB] [OR:Bacillus stearothermophilus] [DE:PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT)] [SP:Q45400]”	157
SPX1442	1442	4103	103	309	194	3.50E-22	[LN:A69785] [AC:A69785] [PN:cellobiose phosphotransferase system enzym homolog ydhM] [GN:ydhM] [CL:phosphotransferase system enzyme II cellobiose-specific factor IIB] [OR:Bacillus subtilis]	189
SPX1443	1443	4104	107	321	186	5.70E-21	[LN:PTCA_BACSU] [AC:P46319] [GN:CELC:LICA] [OR:Bacillus subtilis] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EIII-CEL)] [SP:P46319]	121
SPX1444	1444	4105	327	981	265	2.80E-49	[LN:SORC_KLEPN] [AC:P37078] [GN:SORC] [OR:Klebsiella pneumoniae] [DE:SORBITOL OPERON REGULATOR (SOR OPERON ACTIVATOR)] [SP:P37078]	130
SPX1445	1445	4106	249	747	252	6.10E-57	[LN:YCIT_ECOLI] [AC:P76034] [GN:YCIT] [OR:Escherichia coli]	145

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN OSMB-RNB INTERGENIC REGION] [SP:P76034]	
SPX1446	1446	4107	259	777	266	4.40E-55	"[LN:PFLE_ECOLI] [AC:P75794] [GN:YBIY] [OR:Escherichia coli] [EC:1.97.1.4] [DE:PUTATIVE PYRUVATE FORMATE-LYASE 3 ACTIVATING ENZYME,] [SP:P75794]"	145
SPX1447	1447	4108	83	249	105	3.60E-08	[GI:7331182] [LN:AF235048] [AC:AF235048] [PN:AgpT] [GN:agpT] [OR:Sinorhizobium meliloti]	88
SPX1448	1448	4109	65	195	177	4.60E-18	[GI:6127225] [LN:SHU75349] [AC:U75349] [PN:periplasmic-iron-binding protein BitA] [GN:bit operon] [OR:Brachyspira hyodysenteriae]	129
SPX1449	1449	4110	123	369	242	1.10E-26	[GI:6127225] [LN:SHU75349] [AC:U75349] [PN:periplasmic-iron-binding protein BitA] [GN:bit operon] [OR:Brachyspira hyodysenteriae]	129
SPX1450	1450	4111	318	954	799	1.70E-113	[GI:2766194] [LN:SHU75349] [AC:U75349] [PN:putative ABC transporter BitD] [GN:bit operon] [OR:Brachyspira hyodysenteriae]	121
SPX1451	1451	4112	209	627	283	2.70E-40	[GI:2766195] [LN:SHU75349] [AC:U75349] [PN:putative permease BitE] [GN:bit operon] [OR:Brachyspira hyodysenteriae]	114
SPX1452	1452	4113	257	771	295	1.50E-68	[GI:2766196] [LN:SHU75349] [AC:U75349] [PN:putative permease BitF] [GN:bit operon] [OR:Brachyspira hyodysenteriae]	114
SPX1453	1453	4114	66	198	164	3.20E-17	[GI:2766196] [LN:SHU75349] [AC:U75349] [PN:putative permease BitF] [GN:bit operon]	114

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Brachyspira hyodysenteriae]	
SPX1454	1454	4115	250	750	137	3.20E-10	[LN:PMGY__TREPA] [AC:P96121] [GN:GPM:PGM:TP0168] [OR:Treponema pallidum] [EC:5.4.2.1] [DE:(BPG-DEPENDENT PGAM)] [SP:P96121]	122
SPX1455	1455	4116	446	1338	855	6.60E-196	[GI:7380543] [LN:NMA6Z2491] [AC:AL162757:AL157959] [PN:conserved hypothetical protein] [GN:NMA1908] [OR:Neisseria meningitidis]	127
SPX1456	1456	4117	120	360	123	9.30E-21	[LN:E64494] [AC:E64494] [PN:hypothetical protein MJ1558] [OR:Methanococcus jannaschii]	86
SPX1457	1457	4118	129	387	282	2.20E-58	[LN:T44411] [AC:T44411] [PN:ribosomal protein L17 [imported]] [GN:rplQ] [CL:Escherichia coli ribosomal protein L17] [OR:Bacillus halodurans]	140
SPX1458	1458	4119	312	936	963	6.70E-130	[LN:RPOA__BACSU] [AC:P20429] [GN:RPOA] [OR:Bacillus subtilis] [EC:2.7.7.6] [DE:ALPHA CHAIN) (RNA POLYMERASE ALPHA SUBUNIT)] [SP:P20429]	134
SPX1459	1459	4120	134	402			NO-HIT	6
SPX1460	1460	4121	141	423	494	1.80E-63	[LN:RS11__BACSU] [AC:P04969] [GN:RPSK] [OR:Bacillus subtilis] [DE:30S RIBOSOMAL PROTEIN S11 (BS11)] [SP:P04969]	110
SPX1461	1461	4122	122	366	433	1.70E-55	[GI:1044989] [LN:BACRPLP] [AC:L47971] [PN:ribosomal protein S13] [GN:rpsM] [OR:Bacillus subtilis]	97
SPX1462	1462	4123	56	168	237	1.10E-28	[LN:S17988] [AC:S17988] [PN:translation initiation factor IF-1] [GN:infA] [CL:translation initiation factor IF-1] [OR:Lactococcus lactis subsp. lactis]	151

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1463	1463	4124	213	639	560	8.10E-95	"[LN:KAD_LACLA] [AC:P27143] [GN:ADK] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.4.3] [DE:ADENYLATE KINASE, (ATP-AMP TRANSPHOSPHORYLASE)] [SP:P27143]"	177
SPX1464	1464	4125	437	1311	927	4.30E-126	"[LN:SECY_LACLA] [AC:P27148] [GN:SECY] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:PREPROTEIN TRANSLOCASE SECY SUBUNIT] [SP:P27148]"	155
SPX1465	1465	4126	147	441	614	7.90E-79	[LN:RL15_STAAU] [AC:O06445] [GN:RPLO] [OR:Staphylococcus aureus] [DE:50S RIBOSOMAL PROTEIN L15] [SP:O06445]	107
SPX1466	1466	4127	61	183	185	6.90E-21	[LN:RL30_STAAU] [AC:O06444] [GN:RPMD] [OR:Staphylococcus aureus] [DE:50S RIBOSOMAL PROTEIN L30] [SP:O06444]	107
SPX1467	1467	4128	165	495	621	8.70E-81	[LN:RS5_BACST] [AC:P02357] [GN:RPSE] [OR:Bacillus stearothermophilus] [DE:30S RIBOSOMAL PROTEIN S5 (BS5)] [SP:P02357]	117
SPX1468	1468	4129	119	357	232	2.60E-53	[LN:RL18_BACSU] [AC:P46899:P70969] [GN:RPLR] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L18] [SP:P46899:P70969]	117
SPX1469	1469	4130	175	525	478	2.00E-66	[LN:B69695] [AC:B69695] [PN:ribosomal protein L6 (BL8) rplF] [GN:rplF] [CL:Escherichia coli ribosomal protein L6] [OR:Bacillus subtilis]	136
SPX1470	1470	4131	133	399	521	2.60E-67	[GI:1044978] [LN:BACRPLP] [AC:L47971] [PN:ribosomal protein S8] [GN:rpsH]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis]	
SPX1471	1471	4132	79	237			NO-HIT	6
SPX1472	1472	4133	90	270	325	8.80E-41	[LN:F69835] [AC:F69835] [PN:ribosomal protein S14 homolog yhzA] [GN:yhzA] [CL:Escherichia coli ribosomal protein S14] [OR:Bacillus subtilis]	140
SPX1473	1473	4134	181	543	728	6.40E-96	[LN:T44395] [AC:T44395] [PN:ribosomal protein L5 [imported]] [GN:rplE] [CL:Escherichia coli ribosomal protein L5] [OR:Bacillus halodurans]	138
SPX1474	1474	4135	102	306	508	2.80E-65	[GI:4927751] [LN:AF126059] [AC:AF126059] [PN:RpL24] [GN:rplX] [OR:Streptococcus pneumoniae]	91
SPX1475	1475	4136	123	369	597	2.20E-78	[GI:4927750] [LN:AF126059] [AC:AF126059] [PN:RpL14] [GN:rplN] [OR:Streptococcus pneumoniae]	91
SPX1476	1476	4137	110	330	100	4.40E-07	[LN:E71186] [AC:E71186] [PN:hypothetical protein PH1769] [GN:PH1769] [CL:Pyrococcus horikoshii hypothetical protein PH1769] [OR:Pyrococcus horikoshii]	150
SPX1477	1477	4138	87	261	430	3.10E-55	[GI:4927749] [LN:AF126059] [AC:AF126059] [PN:RpS17] [GN:rpsQ] [OR:Streptococcus pneumoniae]	91
SPX1478	1478	4139	133	399			NO-HIT	6
SPX1479	1479	4140	69	207	324	8.90E-40	[GI:4927748] [LN:AF126059] [AC:AF126059] [PN:RpL29] [GN:rpmC] [OR:Streptococcus pneumoniae]	91
SPX1480	1480	4141	138	414	708	1.30E-94	[GI:4927747] [LN:AF126059] [AC:AF126059] [PN:RpL16] [GN:rplP]	91

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX1481	1481	4142	209	627	1063	2.10E-142	[GI:4927746] [LN:AF126059] [AC:AF126059] [PN:RpS3] [GN:rpsC] [OR:Streptococcus pneumoniae]	90
SPX1482	1482	4143	75	225	356	1.60E-44	[GI:4927745] [LN:AF126059] [AC:AF126059] [PN:RpL22] [GN:rplV] [OR:Streptococcus pneumoniae]	91
SPX1483	1483	4144	82	246			NO-HIT	6
SPX1484	1484	4145	94	282	492	5.20E-64	[GI:4927744] [LN:AF126059] [AC:AF126059] [PN:RpS19] [GN:rpsS] [OR:Streptococcus pneumoniae]	91
SPX1485	1485	4146	278	834	1144	2.80E-152	[LN:RL2_BACST] [AC:P04257] [GN:RPLB] [OR:Bacillus stearothermophilus] [DE:50S RIBOSOMAL PROTEIN L2] [SP:P04257]	111
SPX1486	1486	4147	99	297	162	1.30E-30	[GI:3273331] [LN:AB015722] [AC:AB015722] [PN:ribosomal protrein L23] [OR:Bacillus stearothermophilus] [SR:Bacillus stearothermophilus DNA]	138
SPX1487	1487	4148	208	624	696	6.80E-91	[LN:RL4_BACST] [AC:P28601] [GN:RPLD] [OR:Bacillus stearothermophilus] [DE:50S RIBOSOMAL PROTEIN L4] [SP:P28601]	111
SPX1488	1488	4149	209	627	496	6.60E-110	[LN:RL3_BACSU] [AC:P42920] [GN:RPLC] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L3 (BL3)] [SP:P42920]	107
SPX1489	1489	4150	103	309	494	3.40E-64	[LN:RS10_STRMU] [AC:P48853] [GN:RPSJ] [OR:Streptococcus mutans] [DE:30S RIBOSOMAL PROTEIN S10] [SP:P48853]	106

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1490	1490	4151	206	618	71	0.00018	[GI:7302797] [LN:AE003803] [AC:AE003803:AE002787] [GN:CG4798] [OR:Drosophila melanogaster] [SR:fruit fly]	105
SPX1491	1491	4152	197	591	843	3.60E-112	[GI:4098082] [LN:LLU73336] [AC:U73336] [PN:anaerobic ribonucleotide reductase activator] [GN:nrdG] [FN:activation of anaerobic] [OR:Lactococcus lactis subsp. cremoris]	167
SPX1492	1492	4153	69	207			NO-HIT	6
SPX1493	1493	4154	169	507	187	4.50E-19	[LN:G75479] [AC:G75479] [PN:hypothetical protein] [GN:DR0763] [OR:Deinococcus radiodurans]	90
SPX1494	1494	4155	155	465			NO-HIT	6
SPX1495	1495	4156	738	2214	1123	1.10E-231	[GI:4098081] [LN:LLU73336] [AC:U73336] [PN:anaerobic ribonucleotide reductase] [GN:nrdD] [OR:Lactococcus lactis subsp. cremoris]	128
SPX1496	1496	4157	519	1557			NO-HIT	6
SPX1497	1497	4158	492	1476	519	3.70E-97	[LN:YWAP_STRMU] [AC:P34001] [OR:Streptococcus mutans] [DE:HYPOTHETICAL PROTEIN IN WAPA 3'REGION (FRAGMENT)] [SP:P34001]	119
SPX1498	1498	4159	186	558	93	0.00012	[LN:A26892] [AC:A26892] [PN:Mopa box protein] [OR:Mus musculus] [SR:, house mouse]”	84
SPX1499	1499	4160	70	210			NO-HIT	6
SPX1500	1500	4161	130	390	66	7.30E-09	[GI:1119198] [LN:BACCOMC] [AC:M30805] [PN:unknown protein] [OR:Bacillus subtilis] [SR:Bacillus subtilis (strain IS75) DNA]	122
SPX1501	1501	4162	161	483	119	5.10E-16	[LN:FOLC_LACCA] [AC:P15925] [GN:FGS] [OR:Lactobacillus casei] [EC:6.3.2.17] [DE:SYNTHETASE) (FPGS)]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P15925]	
SPX1502	1502	4163	164	492	251	2.00E-38	[LN:FOLC_BACSU] [AC:Q05865] [GN:FOLC] [OR:Bacillus subtilis] [EC:6.3.2.17] [DE:SYNTHETASE) (FPGS)] [SP:Q05865]	110
SPX1503	1503	4164	102	306	87	5.90E-11	[LN:A69982] [AC:A69982] [PN:hypothetical protein yrzB] [GN:yrzB] [OR:Bacillus subtilis]	87
SPX1504	1504	4165	140	420	327	2.00E-40	[LN:D69979] [AC:D69979] [PN:conserved hypothetical protein yrrK] [GN:yrrK] [CL:Haemophilus influenzae conserved hypothetical protein HI0305] [OR:Bacillus subtilis]	163
SPX1505	1505	4166	89	267			NO-HIT	6
SPX1506	1506	4167	190	570			NO-HIT	6
SPX1507	1507	4168	133	399	315	1.80E-38	[GI:6650536] [LN:AF103794] [AC:AF103794] [PN:unknown] [OR:Listeria monocytogenes]	81
SPX1508	1508	4169	126	378	333	1.60E-40	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX1509	1509	4170	131	393	340	3.70E-41	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX1510	1510	4171	296	888	1035	8.50E-139	[GI:6318592] [LN:AF146529] [AC:AF146529] [PN:aromatic amino acid aminotransferase] [GN:araT] [FN:catalyzes the last reaction in the biosynthesis] [OR:Lactococcus lactis subsp. cremoris]	185
SPX1511	1511	4172	62	186			NO-HIT	6
SPX1512	1512	4173	257	771	297	2.20E-40	[LN:YQXN_BACSU] [AC:P42095] [GN:YQXN:YQFI] [OR:Bacillus subtilis] [DE:(ORF3)] [SP:P42095]	89
SPX1513	1513	4174	331	993	623	6.80E-120	[LN:H69679] [AC:H69679]	148

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:involved in fatty acid/phospholipid synthesis plsX] [GN:plsX] [CL:phospholipid synthesis protein] [OR:Bacillus subtilis]	
SPX1514	1514	4175	112	336			NO-HIT	6
SPX1515	1515	4176	78	234	129	1.40E-17	[LN:C72349] [AC:C72349] [PN:acyl carrier protein] [GN:TM0662] [CL:acyl carrier protein:acyl carrier protein homology] [OR:Thermotoga maritima]	142
SPX1516	1516	4177	98	294	320	4.80E-40	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1517	1517	4178	77	231	64	3.00E-06	[GI:2695624] [LN:STU93029] [AC:U93029] [PN:amphipathic pore-forming peptide precursor] [GN:thmA] [OR:Streptococcus thermophilus]	128
SPX1518	1518	4179	69	207			NO-HIT	6
SPX1519	1519	4180	720	2160	3571	0	[LN:COMA_STRPN] [AC:Q03727] [GN:COMA] [OR:Streptococcus pneumoniae] [DE:TRANSPORT ATP-BINDING PROTEIN COMA] [SP:Q03727]	119
SPX1520	1520	4181	450	1350	2184	8.20E-289	[LN:COMB_STRPN] [AC:P36498] [GN:COMB] [OR:Streptococcus pneumoniae] [DE:TRANSPORT PROTEIN COMB] [SP:P36498]	107
SPX1521	1521	4182	250	750	1189	1.40E-160	[LN:PUR7_STRPN] [AC:Q07296] [GN:PURC] [OR:Streptococcus pneumoniae] [EC:6.3.2.6] [DE:(SAICAR SYNTHETASE)] [SP:Q07296]	117
SPX1522	1522	4183	254	762	127	2.80E-08	“““[LN:C69492] [AC:C69492] [PN:phosphoribosylformylglycinamidine synthase, component II:formylglycinamide ribotide amidotransferase:phosphoribosylformylglycinamidine synthetase] [OR:Archaeoglobus fulgidus] [EC:6.3.5.3]”””	217
SPX1523	1523	4184	95	285	155	3.80E-15	[GI:4928281] [LN:AF132127] [AC:AF132127]	110

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:glucose-6-phosphate isomerase] [GN:gpi] [OR:Streptococcus mutans]	
SPX1524	1524	4185	457	1371	1764	1.20E-238	[GI:4928281] [LN:AF132127] [AC:AF132127] [PN:glucose-6-phosphate isomerase] [GN:gpi] [OR:Streptococcus mutans]	110
SPX1525	1525	4186	246	738	246	3.10E-28	"[LN:YDP3__LACLA] [AC:P22347] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:HYPOTHETICAL 18.7 KD PROTEIN IN PEPX 3'REGION (ORF3)] [SP:P22347]"	162
SPX1526	1526	4187	604	1812	1001	2.10E-222	[GI:4104142] [LN:AF033015] [AC:AF033015] [PN:ABC transporter homolog Z] [OR:Listeria monocytogenes]	99
SPX1527	1527	4188	72	216			NO-HIT	6
SPX1528	1528	4189	108	324	105	1.60E-07	[GI:6015958] [LN:SSU18930] [AC:Y18930] [PN:hypothetical protein] [GN:ORF-c22__037] [OR:Sulfolobus solfataricus]	110
SPX1529	1529	4190	196	588	455	7.30E-58	[GI:4104141] [LN:AF033015] [AC:AF033015] [PN:ABC transporter homolog Y] [OR:Listeria monocytogenes]	99
SPX1530	1530	4191	361	1083	422	2.80E-66	[GI:4104140] [LN:AF033015] [AC:AF033015] [PN:ABC transporter homolog X] [OR:Listeria monocytogenes]	99
SPX1531	1531	4192	655	1965	3290	0	[LN:HEXA__STRPN] [AC:P10564] [GN:HEXA] [OR:Streptococcus pneumoniae] [DE:DNA MISMATCH REPAIR PROTEIN HEXA] [SP:P10564]	117
SPX1532	1532	4193	182	546	859	6.40E-114	[LN:HEXA__STRPN] [AC:P10564] [GN:HEXA] [OR:Streptococcus pneumoniae] [DE:DNA MISMATCH REPAIR PROTEIN HEXA] [SP:P10564]	117
SPX1533	1533	4194	149	447	740	7.50E-99	[LN:ARGR__STRPN]	112

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Q54870] [GN:ARGR] [OR:Streptococcus pneumoniae] [DE:PROBABLE ARGININE REPRESSOR] [SP:Q54870]	
SPX1534	1534	4195	564	1692	365	3.20E-102	"[LN:SYRC_YEAST] [AC:Q05506] [GN:YDR341C:D9651.10] [OR:Saccharomyces cerevisiae] [SR:,Baker's yeast] [EC:6.1.1.19] [DE:-TRNA LIGASE) (ARGRS)] [SP:Q05506]"	154
SPX1535	1535	4196	150	450	398	6.20E-51	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1536	1536	4197	68	204			NO-HIT	6
SPX1537	1537	4198	95	285	102	6.20E-17	"[LN:S74709] [AC:S74709] [PN:hypothetical protein sll1188] [OR:Synechocystis sp.] [SR:PCC 6803, , PCC 6803] [SR:PCC 6803,]"	124
SPX1538	1538	4199	236	708	1209	4.90E-164	[GI:5830529] [LN:SPAJ6394] [AC:AJ006394] [PN:response regulator] [GN:phoP] [OR:Streptococcus pneumoniae]	104
SPX1539	1539	4200	444	1332	2196	4.10E-300	[GI:5830530] [LN:SPAJ6394] [AC:AJ006394] [PN:histidine kinase] [GN:phoR] [OR:Streptococcus pneumoniae]	102
SPX1540	1540	4201	292	876	1435	1.70E-188	[GI:4530447] [LN:AF118229] [AC:AF118229] [PN:phosphate binding protein PstS] [GN:pstS] [OR:Streptococcus pneumoniae]	116
SPX1541	1541	4202	272	816	1333	2.80E-185	[GI:4530448] [LN:AF118229] [AC:AF118229] [PN:transmembrane protein PstC] [GN:pstC] [OR:Streptococcus pneumoniae]	112
SPX1542	1542	4203	272	816	1312	1.10E-182	[GI:4530449] [LN:AF118229]	112

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AF118229] [PN:transmembrane protein PstA] [GN:PstA] [OR:Streptococcus pneumoniae]	
SPX1543	1543	4204	251	753	1284	2.00E-174	[GI:4530450] [LN:AF118229] [AC:AF118229] [PN:ATP-binding cassette protein PstB] [GN:pstB] [OR:Streptococcus pneumoniae]	119
SPX1544	1544	4205	217	651	1067	8.20E-144	[GI:4530451] [LN:AF118229] [AC:AF118229] [PN:PhoU] [GN:phoU] [OR:Streptococcus pneumoniae]	90
SPX1545	1545	4206	142	426	131	2.90E-10	[GI:5822769] [LN:AB024553] [AC:AB024553] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA]	109
SPX1546	1546	4207	73	219			NO-HIT	6
SPX1547	1547	4208	153	459			NO-HIT	6
SPX1548	1548	4209	339	1017	921	6.40E-122	[LN:GPDA_BACSU] [AC:P46919] [GN:GPSA:GLYC] [OR:Bacillus subtilis] [EC:1.1.1.94] [DE:DEPENDENT DIHYDROXYACETONE-PHOSPHATE REDUCTASE)] [SP:P46919]	144
SPX1549	1549	4210	300	900	1481	1.60E-199	[GI:3550619] [LN:SPAJ4869] [AC:AJ004869] [PN:UTP-glucose-1-phosphate uridylyltransferase] [GN:galU] [FN:synthesis of UDP-glucose] [OR:Streptococcus pneumoniae]	159
SPX1550	1550	4211	225	675			NO-HIT	6
SPX1551	1551	4212	227	681	520	1.90E-70	[LN:SUB400707] [AC:AJ400707] [PN:hypothetical protein] [OR:Streptococcus uberis]	80
SPX1552	1552	4213	182	546	236	1.30E-55	[GI:3192049] [LN:AB001562] [AC:AB001562] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:Xc) DNA]	108
SPX1553	1553	4214	377	1131	622	4.50E-119	[LN:G69866] [AC:G69866] [PN:hippurate hydrolase homolog ykuR]	119

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:ykuR] [CL:hippurate hydrolase] [OR:Bacillus subtilis]	
SPX1554	1554	4215	233	699	578	1.10E-86	“[LN:H72245] [AC:H72245] [PN:2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase-related protein] [GN:TM1519] [OR:Thermotoga maritima]”	146
SPX1555	1555	4216	153	459	75	5.50E-08	[GI:6103625] [LN:AF172095] [AC:AF172095] [PN:unknown] [OR:Picea rubens]	71
SPX1556	1556	4217	72	216			NO-HIT	6
SPX1557	1557	4218	278	834	214	2.60E-30	[LN:YDED_BACSU] [AC:P96661] [GN:YDED] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 35.3 KD PROTEIN IN CSPC-NAP INTERGENIC REGION] [SP:P96661]	136
SPX1558	1558	4219	821	2463	4147	0	[GI:6165962] [LN:AF101781] [AC:AF101781] [PN:penicillin-binding protein 1b] [GN:pbp1b] [OR:Streptococcus pneumoniae]	116
SPX1559	1559	4220	419	1257	697	1.90E-174	[LN:SYY1_BACSU] [AC:P22326] [GN:TYRS] [OR:Bacillus subtilis] [EC:6.1.1.1] [DE:(TYRRS 1)] [SP:P22326]	100
SPX1560	1560	4221	691	2073	218	1.60E-63	“[LN:COPA_ENTHR] [AC:P32113:Q47841] [GN:COPA] [OR:Enterococcus hirae] [EC:3.6.1.36] [DE:COPPER/POTASSIUM-TRANSPORTING ATPASE A,] [SP:P32113:Q47841]”	148
SPX1561	1561	4222	65	195			NO-HIT	6
SPX1562	1562	4223	88	264	109	2.90E-08	[LN:RRMA_ECOLI] [AC:P36999] [GN:RRMA] [OR:Escherichia coli] [EC:2.1.1.51] [DE:METHYLTRANSFERASE)] [SP:P36999]	109
SPX1563	1563	4224	197	591	192	1.90E-25	[LN:YXJB_BACSU] [AC:P42313]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:YXJB:N15I] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 31.5 KD PROTEIN IN KATB 3'REGION] [SP:P42313]	
SPX1564	1564	4225	57	171			NO-HIT	6
SPX1565	1565	4226	191	573			NO-HIT	6
SPX1566	1566	4227	753	2259	473	5.80E-165	“[LN:PHSG__BACSU] [AC:P39123] [GN:GLGP] [OR:Bacillus subtilis] [EC:2.4.1.1] [DE:GLYCOGEN PHOSPHORYLASE,] [SP:P39123]”	116
SPX1567	1567	4228	506	1518	2687	0	[LN:MALQ__STRPN] [AC:P29851] [GN:MALM] [OR:Streptococcus pneumoniae] [EC:2.4.1.25] [DE:(DISPROPORTIONATING ENZYME) (D-ENZYME)] [SP:P29851]	137
SPX1568	1568	4229	424	1272	2167	3.30E-288	[LN:MALX__STRPN] [AC:P29850] [GN:MALX] [OR:Streptococcus pneumoniae] [DE:MALTOSE/MALTODEXTRIN-BINDING PROTEIN PRECURSOR] [SP:P29850]	131
SPX1569	1569	4230	431	1293	2176	0	[LN:MALC__STRPN] [AC:Q04698] [GN:MALC] [OR:Streptococcus pneumoniae] [DE:MALTODEXTRIN TRANSPORT SYSTEM PERMEASE PROTEIN MALC] [SP:Q04698]	136
SPX1570	1570	4231	281	843	1265	2.40E-177	[LN:MALD__STRPN] [AC:Q04699] [GN:MALD] [OR:Streptococcus pneumoniae] [DE:MALTODEXTRIN TRANSPORT SYSTEM PERMEASE PROTEIN MALD] [SP:Q04699]	136
SPX1571	1571	4232	267	801	540	1.70E-145	[LN:MALA__STRPN] [AC:Q08510] [GN:MALA] [OR:Streptococcus pneumoniae] [DE:MALA PROTEIN] [SP:Q08510]	97
SPX1572	1572	4233	329	987	1668	7.30E-227	[GI:2656094] [LN:STRMALR] [AC:L21856] [PN:repressor protein] [GN:malR] [FN:maltose operon transcriptional repressor]	146

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae]	
SPX1573	1573	4234	114	342			NO-HIT	6
SPX1574	1574	4235	127	381	378	8.30E-48	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1575	1575	4236	314	942	463	6.90E-73	[GI:1620924] [LN:BS168NPRB] [AC:Z79580] [OR:Bacillus subtilis]	62
SPX1576	1576	4237	588	1764	701	1.30E-229	[LN:SYD_BACSU] [AC:O32038] [GN:ASPS] [OR:Bacillus subtilis] [EC:6.1.1.12] [DE:(ASPRS)] [SP:O32038]	98
SPX1577	1577	4238	144	432			NO-HIT	6
SPX1578	1578	4239	64	192			NO-HIT	6
SPX1579	1579	4240	77	231			NO-HIT	6
SPX1580	1580	4241	214	642	120	9.80E-14	[GI:1402529] [LN:D78257] [AC:D78257] [PN:ORF8] [GN:orf8] [OR:Enterococcus faecalis] [SR:Enterococcus faecalis plasmid:pYI17 DNA]	128
SPX1581	1581	4242	247	741			NO-HIT	6
SPX1582	1582	4243	67	201			NO-HIT	6
SPX1583	1583	4244	155	465	78	0.00011	[LN:T07291] [AC:T07291] [PN:hypothetical protein 42c] [OR:chloroplast Chlorella vulgaris]	89
SPX1584	1584	4245	62	186			NO-HIT	6
SPX1585	1585	4246	64	192			NO-HIT	6
SPX1586	1586	4247	63	189	200	7.00E-24	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1587	1587	4248	430	1290	1740	7.50E-236	[LN:SYH_STREQ] [AC:P30053] [GN:HISS] [OR:Streptococcus equisimilis] [EC:6.1.1.21] [DE:(HISRS)] [SP:P30053]	106
SPX1588	1588	4249	307	921			NO-HIT	6
SPX1589	1589	4250	416	1248	334	2.10E-57	[GI:3582221] [LN:AE001272] [AC:AE001272]	114

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:conserved hypothetical protein] [GN:ORF00049] [OR:Lactococcus lactis]	
SPX1590	1590	4251	133	399			NO-HIT	6
SPX1591	1591	4252	284	852	301	1.70E-34	[LN:RGG_STRGC] [AC:P49330] [GN:RGG] [OR:Streptococcus gordonii challis] [DE:RGG PROTEIN] [SP:P49330]	100
SPX1592	1592	4253	114	342	227	6.60E-27	[GI:2258088] [LN:AB000631] [AC:AB000631] [OR:Streptococcus mutans] [SR:Streptococcus mutans DNA]	96
SPX1593	1593	4254	568	1704	994	1.30E-207	“[LN:S76895] [AC:S76895] [PN:hypothetical protein] [CL:dihydroxy-acid dehydratase] [OR:Synechocystis sp.] [SR:PCC 6803, , PCC 6803] [SR:PCC 6803,]”	148
SPX1594	1594	4255	311	933	272	2.40E-45	“[LN:TKTC_METJA] [AC:Q58092] [GN:MJ0679] [OR:Methanococcus jannaschii] [EC:2.2.1.1] [DE:PUTATIVE TRANSKETOLASE C-TERMINAL SECTION, (TK)] [SP:Q58092]”	149
SPX1595	1595	4256	286	858	347	7.70E-66	“[LN:TKTN_METJA] [AC:Q58094] [GN:MJ0681] [OR:Methanococcus jannaschii] [EC:2.2.1.1] [DE:PUTATIVE TRANSKETOLASE N-TERMINAL SECTION, (TK)] [SP:Q58094]”	149
SPX1596	1596	4257	449	1347	535	9.90E-132	[LN:T37066] [AC:T37066] [PN:probable integral membrane protein] [GN:SCJ21.17c] [OR:Streptomyces coelicolor]	107
SPX1597	1597	4258	95	285			NO-HIT	6
SPX1598	1598	4259	677	2031	197	8.80E-31	“[GI:4512373] [LN:AB011837] [AC:AB011837] [GN:yjdC] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA, clone_lib:lambda no.9]”	144
SPX1599	1599	4260	336	1008	133	1.30E-18	[GI:6689167] [LN:SCE20]	118

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AL136058] [PN:putative membrane protein] [GN:SCE20.08c] [OR:Streptomyces coelicolor A3(2)]	
SPX1600	1600	4261	67	201			NO-HIT	6
SPX1601	1601	4262	61	183	240	3.00E-28	“[LN:RL32__LACLC] [AC:O34101] [GN:RPMF] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [DE:50S RIBOSOMAL PROTEIN L32] [SP:O34101]”	149
SPX1602	1602	4263	642	1926	1320	0	[GI:1914872] [LN:SPZ82001] [AC:Z82001] [PN:PCPA] [GN:pcpA] [OR:Streptococcus pneumoniae]	88
SPX1603	1603	4264	76	228			NO-HIT	6
SPX1604	1604	4265	192	576	429	4.10E-54	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1605	1605	4266	126	378	303	4.10E-37	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1606	1606	4267	109	327	495	7.70E-65	[GI:1914871] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX1607	1607	4268	627	1881	95	0.00073	[LN:T27355] [AC:T27355] [PN:hypothetical protein Y70D2A.2] [GN:Y70D2A.2] [OR:Caenorhabditis elegans]	100
SPX1608	1608	4269	290	870	108	2.70E-22	“[LN:F70203] [AC:F70203] [PN:xylose operon regulatory protein (xylR-2) homolog] [CL:glucose kinase:glucose kinase homology] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]”	181
SPX1609	1609	4270	887	2661	345	6.30E-65	[LN:YBGG__ECOLI] [AC:P54746:P75753] [GN:YBGG] [OR:Escherichia coli] [DE:HYPOTHETICAL 100.0 KD PROTEIN IN HRSA-CYDA INTERGENIC REGION] [SP:P54746:P75753]	151
SPX1610	1610	4271	435	1305	409	6.20E-106	[LN:T37125]	101

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:T37125] [PN:hypothetical protein SCJ4.42c] [GN:SCJ4.42c] [OR:Streptomyces coelicolor]	
SPX1611	1611	4272	695	2085	408	9.50E-105	[GI:5759293] [LN:AF175722] [AC:AF175722] [PN:immunoreactive 89kD antigen PG87] [OR:Porphyromonas gingivalis]	108
SPX1612	1612	4273	560	1680	249	2.20E-63	"[GI:4096756] [LN:SSU39394] [AC:U39394] [PN:alpha-1,3/4-fucosidase precursor] [FN:alpha-fucosidase specific for alpha-1,3 and] [OR:Streptomyces sp.] [SR:Streptomyces sp]"	170
SPX1613	1613	4274	410	1230	942	1.30E-244	[LN:SAGP_STRPY] [AC:P16962] [GN:SAGP] [OR:Streptococcus pyogenes] [DE:STREPTOCOCCAL ACID GLYCOPROTEIN] [SP:P16962]	114
SPX1614	1614	4275	339	1017	897	3.80E-184	[GI:2764612] [LN:LSAJ1330] [AC:AJ001330] [PN:ornithine transcarbamoylase] [GN:arcB] [OR:Lactobacillus sakei]	108
SPX1615	1615	4276	316	948	465	3.40E-129	[GI:2894540] [LN:EFAJ3331] [AC:AJ223331] [PN:carbamate kinase] [GN:arcC] [FN:synthesis of ATP from carbamylphosphate] [OR:Enterococcus faecium]	143
SPX1616	1616	4277	504	1512	426	4.50E-106	[GI:2697115] [LN:AF008219] [AC:AF008219] [PN:unknown] [OR:Borrelia afzelii]	75
SPX1617	1617	4278	444	1332	277	8.20E-67	[LN:S43914] [AC:S43914] [PN:hypothetical protein 1] [CL:peptidase V] [OR:Bacillus stearothermophilus]	101
SPX1618	1618	4279	64	192			NO-HIT	6
SPX1619	1619	4280	79	237			NO-HIT	6
SPX1620	1620	4281	89	267			NO-HIT	6
SPX1621	1621	4282	51	153			NO-HIT	6
SPX1622	1622	4283	72	216	176	4.20E-20	[LN:T30285]	79

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	
SPX1623	1623	4284	55	165			NO-HIT	6
SPX1624	1624	4285	300	900	560	2.70E-78	[GI:6714460] [LN:ATAC008261] [AC:AC008261] [GN:T4P13.3] [OR:Arabidopsis thaliana] [SR:thale cress]	98
SPX1625	1625	4286	64	192			NO-HIT	6
SPX1626	1626	4287	384	1152	840	1.50E-119	“[LN:ADH2_ECOLI] [AC:P37686] [GN:YIAY] [OR:Escherichia coli] [EC:1.1.1.1] [DE:PROBABLE ALCOHOL DEHYDROGENASE,] [SP:P37686]”	123
SPX1627	1627	4288	589	1767	874	7.80E-287	[GI:6015981] [LN:AF137263] [AC:AF137263] [PN:L-fucose isomerase] [GN:fucI] [OR:Bacteroides thetaiotaomicron]	108
SPX1628	1628	4289	243	729	201	8.40E-20	“[LN:B70645] [AC:B70645] [PN:L-fuculose-phosphate aldolase,] [GN:fucA] [CL:L-ribulose-phosphate 4-epimerase] [OR:Mycobacterium tuberculosis] [EC:4.1.2.17]”	155
SPX1629	1629	4290	1029	3087			NO-HIT	6
SPX1630	1630	4291	738	2214	1709	3.00E-243	[GI:4567098] [LN:AF130985] [AC:AF130985] [PN:alpha-galactosidase AgaN] [GN:agaN] [OR:Bacillus stearothermophilus]	113
SPX1631	1631	4292	624	1872	112	5.60E-07	[LN:T36462] [AC:T36462] [PN:hypothetical protein SCF85.02] [GN:SCF85.02] [OR:Streptomyces coelicolor]	101
SPX1632	1632	4293	440	1320	174	1.10E-50	[LN:T36467] [AC:T36467] [PN:probable glycosyl hydrolase] [GN:SCF85.07] [CL:alpha-L-fucosidase] [OR:Streptomyces coelicolor]	123
SPX1633	1633	4294	62	186			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1634	1634	4295	295	885	372	3.60E-50	[LN:YURM_BACSU] [AC:O32154] [GN:YURM] [OR:Bacillus subtilis] [DE:HYPOTHETICAL ABC TRANSPORTER PERMEASE PROTEIN YURM] [SP:O32154]	128
SPX1635	1635	4296	310	930	163	7.70E-27	"[LN:E72357] [AC:E72357] [PN:sugar ABC transporter, permease protein] [GN:TM0596] [CL:inner membrane protein malF] [OR:Thermotoga maritima]"	140
SPX1636	1636	4297	431	1293	84	5.60E-07	[GI:1524333] [LN:SCMALREFG] [AC:Y07706] [PN:putative maltose-binding pootein] [GN:malE] [OR:Streptomyces coelicolor]	116
SPX1637	1637	4298	468	1404	210	2.00E-81	"[LN:E70014] [AC:E70014] [PN:rhamnulokinase, yulC] [GN:yulC] [CL:rhamnulokinase] [OR:Bacillus subtilis] [EC:2.7.1.5]"	117
SPX1638	1638	4299	59	177			NO-HIT	6
SPX1639	1639	4300	258	774	123	7.20E-21	[LN:SRLR_ECOLI] [AC:P15082:P77030] [GN:SRLR:GUTR] [OR:Escherichia coli] [DE:GLUCITOL OPERON REPRESSOR] [SP:P15082:P77030]	121
SPX1640	1640	4301	73	219			NO-HIT	6
SPX1641	1641	4302	69	207			NO-HIT	6
SPX1642	1642	4303	502	1506	2158	8.80E-289	[LN:T46756] [AC:T46756] [PN:Zn-binding lipoprotein adcA [imported]] [GN:adcA] [OR:Streptococcus pneumoniae]	107
SPX1643	1643	4304	269	807	1261	4.10E-177	[LN:T46755] [AC:T46755] [PN:membrane protein adcB [imported]] [GN:adcB] [OR:Streptococcus pneumoniae]	101
SPX1644	1644	4305	235	705	1247	1.40E-168	[LN:T46754] [AC:T46754] [PN:AdcC protein [imported]] [GN:adcC] [OR:Streptococcus pneumoniae]	92
SPX1645	1645	4306	64	192	89	1.20E-06	[LN:G72536]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:G72536] [PN:hypothetical protein APE1580] [GN:APE1580] [OR:Aeropyrum pernix]	
SPX1646	1646	4307	167	501	728	1.10E-95	[LN:T46753] [AC:T46753] [PN:repressor protein adcR [imported]] [GN:adcR] [OR:Streptococcus pneumoniae]	102
SPX1647	1647	4308	428	1284	612	1.10E-162	[GI:3403204] [LN:AF050517] [AC:AF050517:S78492] [PN:unknown] [GN:dlt4] [OR:Streptococcus mutans]	96
SPX1648	1648	4309	80	240	360	8.20E-46	[GI:2981430] [LN:AF049357] [AC:AF049357:S78492] [PN:Glg3] [GN:glg3] [OR:Streptococcus mutans]	93
SPX1649	1649	4310	415	1245	1236	2.30E-227	[GI:2952530] [LN:AF051356] [AC:AF051356] [PN:integral membrane protein] [GN:dltB] [OR:Streptococcus mutans]	107
SPX1650	1650	4311	517	1551	2154	1.90E-292	[LN:DLTA_STRMU] [AC:Q53526:O68576] [GN:DLTA] [OR:Streptococcus mutans] [EC:6.3.2.-] [DE:ALANYL CARRIER PROTEIN LIGASE) (DCL)] [SP:Q53526:O68576]	144
SPX1651	1651	4312	66	198			NO-HIT	6
SPX1652	1652	4313	174	522	215	3.50E-24	[LN:H69812] [AC:H69812] [PN:conserved hypothetical protein yfmI] [GN:yfmI] [OR:Bacillus subtilis]	97
SPX1653	1653	4314	237	711	97	2.60E-14	[LN:H69812] [AC:H69812] [PN:conserved hypothetical protein yfmI] [GN:yfmI] [OR:Bacillus subtilis]	97
SPX1654	1654	4315	62	186			NO-HIT	6
SPX1655	1655	4316	63	189			NO-HIT	6
SPX1656	1656	4317	331	993	1712	2.40E-235	[GI:2804734] [LN:AF030367] [AC:AF030367]	100

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:maturase-related protein] [OR:Streptococcus pneumoniae]	
SPX1657	1657	4318	94	282	470	6.40E-60	[GI:2804734] [LN:AF030367] [AC:AF030367] [PN:maturase-related protein] [OR:Streptococcus pneumoniae]	100
SPX1658	1658	4319	82	246			NO-HIT	6
SPX1659	1659	4320	66	198			NO-HIT	6
SPX1660	1660	4321	196	588			NO-HIT	6
SPX1661	1661	4322	68	204	344	4.40E-42	[GI:1217989] [LN:SPU12567] [AC:U12567] [PN:ORF3] [OR:Streptococcus pneumoniae]	78
SPX1662	1662	4323	253	759	968	2.20E-159	[LN:GLPF_STRPN] [AC:P52281] [GN:GLPF] [OR:Streptococcus pneumoniae] [DE:GLYCEROL UPTAKE FACILITATOR PROTEIN] [SP:P52281]	120
SPX1663	1663	4324	609	1827	3058	0	[GI:3551774] [LN:SPU94770] [AC:U94770] [PN:alpha-glycerophosphate oxidase] [GN:glpO] [FN:oxidizes alpha-glycerophosphate to] [OR:Streptococcus pneumoniae]	154
SPX1664	1664	4325	100	300	302	1.90E-38	“[LN:S67936] [AC:S67936] [PN:glycerol-3-phosphate dehydrogenase homolog GlpD] [GN:glpD] [OR:Streptococcus pneumoniae] [SR:strain P13, , strain P13] [SR:strain P13,]”	166
SPX1665	1665	4326	503	1509	1251	2.90E-274	“[LN:GLPK_ENTFA] [AC:O34154] [GN:GLPK] [OR:Enterococcus faecalis] [SR:,Streptococcus faecalis] [EC:2.7.1.30] [DE:(GLYCEROKINASE) (GK)] [SP:O34154]”	147
SPX1666	1666	4327	476	1428			NO-HIT	6
SPX1667	1667	4328	66	198			NO-HIT	6
SPX1668	1668	4329	291	873	1063	2.60E-140	[GI:4033717] [LN:SPU49397] [AC:U49397] [PN:unknown] [OR:Streptococcus pyogenes]	79

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1669	1669	4330	327	981	1303	4.70E-175	[GI:4033718] [LN:SPU49397] [AC:U49397] [PN:unknown] [OR:Streptococcus pyogenes]	79
SPX1670	1670	4331	63	189			NO-HIT	6
SPX1671	1671	4332	222	666	235	4.10E-25	[GI:558538] [LN:SUHSAPI] [AC:D38490] [PN:sperm-activating peptide I precursor] [OR:Hemicentrotus pulcherrimus] [SR:Hemicentrotus pulcherrimus female ovary accessory cell (library]	179
SPX1672	1672	4333	451	1353	500	6.50E-65	[GI:4838563] [LN:AF145055] [AC:AF145055] [PN:surface protein C PspC] [OR:Streptococcus pneumoniae]	98
SPX1673	1673	4334	183	549	213	1.70E-21	[GI:7293488] [LN:AE003509] [AC:AE003509:AE002593] [GN:CG15040] [OR:Drosophila melanogaster] [SR:fruit fly]	106
SPX1674	1674	4335	176	528	112	1.00E-08	[GI:406446] [LN:MGU02192] [AC:U02192] [OR:Mycoplasma genitalium]	64
SPX1675	1675	4336	182	546	107	1.20E-15	[GI:3849798] [LN:U91581] [AC:U91581:U04057] [PN:putative transposase] [GN:tpase] [OR:Lactococcus lactis subsp. lactis]	118
SPX1676	1676	4337	251	753	175	7.10E-43	[GI:6746427] [LN:AF179847] [AC:AF179847] [PN:putative transposase] [OR:Lactococcus lactis]	90
SPX1677	1677	4338	65	195			NO-HIT	6
SPX1678	1678	4339	55	165	183	1.50E-19	[GI:2576331] [LN:SPSPSA2] [AC:AJ002054] [PN:SpsA protein] [FN:IgA binding protein] [OR:Streptococcus pneumoniae]	112
SPX1679	1679	4340	419	1257	2039	1.30E-284	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1680	1680	4341	232	696	1296	1.20E-171	[GI:4097980] [LN:SPU72655] [AC:U72655] [PN:surface protein C] [GN:pspC] [OR:Streptococcus pneumoniae]	101
SPX1681	1681	4342	110	330			NO-HIT	6
SPX1682	1682	4343	105	315	98	7.50E-05	[LN:T14867] [AC:T14867] [PN:interaptin] [GN:abpD] [OR:Dictyostelium discoideum]	79
SPX1683	1683	4344	217	651	599	1.00E-117	[GI:6469845] [LN:AF068645] [AC:AF068645] [PN:unknown] [GN:pspC] [OR:Streptococcus pneumoniae]	93
SPX1684	1684	4345	73	219	66	0.001	"[GI:340613] [LN:LEIKPMURF2] [AC:L07545] [GN:MURF2] [OR:Kinetoplast Leishmania tarentolae] [SR:Kinetoplast Leishmania tarentolae (strain UC, organelle Kinetoplas]"	163
SPX1685	1685	4346	105	315	94	1.40E-05	"[LN:YHU3_YEAST] [AC:P38844] [GN:YHR143W] [OR:Saccharomyces cerevisiae] [SR:,Baker's yeast] [DE:PRECURSOR] [SP:P38844]"	119
SPX1686	1686	4347	276	828	697	7.60E-123	[GI:2576333] [LN:SPSPSA47] [AC:AJ002055] [PN:SpsA protein] [FN:IgA binding protein] [OR:Streptococcus pneumoniae]	113
SPX1687	1687	4348	176	528	382	1.80E-50	[LN:F81147] [AC:F81147] [PN:conserved hypothetical protein NMB0883 [imported]] [GN:NMB0883] [OR:Neisseria meningitidis]	119
SPX1688	1688	4349	447	1341	1245	0	[GI:5830533] [LN:SPAJ6395] [AC:AJ006395] [PN:histidine kinase] [GN:hk06] [OR:Streptococcus pneumoniae]	102
SPX1689	1689	4350	218	654	1100	3.60E-150	[GI:5830532] [LN:SPAJ6395]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AJ006395] [PN:response regulator] [GN:rr06] [OR:Streptococcus pneumoniae]	
SPX1690	1690	4351	75	225			NO-HIT	6
SPX1691	1691	4352	811	2433	1102	3.30E-266	[GI:4103472] [LN:AF023422] [AC:AF023422] [PN:ClpC] [GN:clpC] [OR:Lactococcus lactis]	84
SPX1692	1692	4353	174	522	323	4.90E-44	[LN:LLA249133] [AC:AJ249133] [PN:CtsR protein] [GN:ctsR] [FN:transcriptional regulator] [OR:Lactococcus lactis]	111
SPX1693	1693	4354	243	729	255	1.80E-47	“[LN:B72369] [AC:B72369] [PN:ABC transporter, ATP-binding protein] [GN:TM0483] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Thermotoga maritima]”	180
SPX1694	1694	4355	336	1008	561	9.10E-87	[LN:C72369] [AC:C72369] [PN:hypothetical protein TM0484] [GN:TM0484] [OR:Thermotoga maritima]	93
SPX1695	1695	4356	136	408			NO-HIT	6
SPX1696	1696	4357	250	750	440	1.10E-58	“[LN:D72369] [AC:D72369] [PN:ABC transporter, permease protein, cysTW family] [GN:TM0485] [CL:Synechococcus nitrate transport protein nrtB] [OR:Thermotoga maritima]”	165
SPX1697	1697	4358	97	291	118	4.00E-14	[LN:YVI2_CLOPE] [AC:Q46213] [OR:Clostridium perfringens] [DE:HYPOTHETICAL 10.7 KD PROTEIN IN VIRR 5'REGION (ORF2)] [SP:Q46213]	126
SPX1698	1698	4359	284	852	319	8.30E-35	[LN:S57721] [AC:S57721] [PN:cspB protein] [CL:cpl repeat homology] [OR:Clostridium acetobutylicum]	98
SPX1699	1699	4360	138	414	102	7.10E-06	[GI:1340128] [LN:SA1234] [AC:X97985] [OR:Staphylococcus aureus]	63
SPX1700	1700	4361	89	267	72	3.10E-05	[LN:VEG_BACSU]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P37466] [GN:VEG] [OR:Bacillus subtilis] [DE:VEG PROTEIN] [SP:P37466]	
SPX1701	1701	4362	451	1353	827	1.60E-171	“[LN:DNAC__BACSU] [AC:P37469] [GN:DNAC] [OR:Bacillus subtilis] [EC:3.6.1.-] [DE:REPLICATIVE DNA HELICASE,] [SP:P37469]”	118
SPX1702	1702	4363	151	453	193	1.20E-44	[LN:RL9__BACST] [AC:P02417] [GN:RPLI] [OR:Bacillus stearothermophilus] [DE:50S RIBOSOMAL PROTEIN L9 (BL17)] [SP:P02417]	118
SPX1703	1703	4364	658	1974	301	6.60E-68	[LN:YYBT__BACSU] [AC:P37484] [GN:YYBT] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION] [SP:P37484]	137
SPX1704	1704	4365	183	549	358	5.10E-68	[LN:LLA249134] [AC:AJ249134] [PN:hypothetical protein] [OR:Lactococcus lactis]	78
SPX1705	1705	4366	221	663	143	2.90E-23	[LN:CMF3__BACSU] [AC:P39147] [GN:COMFC:COMF3] [OR:Bacillus subtilis] [DE:COMF OPERON PROTEIN 3] [SP:P39147]	106
SPX1706	1706	4367	433	1299	269	2.00E-31	[LN:CMF1__BACSU] [AC:P39145] [GN:COMFA:COMF1] [OR:Bacillus subtilis] [DE:COMF OPERON PROTEIN 1] [SP:P39145]	106
SPX1707	1707	4368	212	636	475	1.00E-60	[LN:YVYE__BACSU] [AC:P32437:P96500] [GN:YVYE:YVHK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 24.8 KD PROTEIN IN DEGS-TAGO INTERGENIC REGION] [SP:P32437:P96500]	156
SPX1708	1708	4369	309	927	1289	7.30E-171	[GI:6567187] [LN:AB028865] [AC:AB028865] [PN:O-acetylserine lyase]	141

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:cysM] [OR:Streptococcus suis] [SR:Streptococcus suis (strain:SMR) DNA]	
SPX1709	1709	4370	347	1041	498	1.10E-82	[LN:EFTS__BACSU] [AC:P80700:O31748] [GN:TSF] [OR:Bacillus subtilis] [DE:ELONGATION FACTOR TS (EF-TS)] [SP:P80700:O31748]	119
SPX1710	1710	4371	129	387			NO-HIT	6
SPX1711	1711	4372	288	864	1013	1.60E-135	[LN:RS2__PEDAC] [AC:P49668] [GN:RPSB] [OR:Pediococcus acidilactici] [DE:30S RIBOSOMAL PROTEIN S2] [SP:P49668]	108
SPX1712	1712	4373	75	225			NO-HIT	6
SPX1713	1713	4374	393	1179	459	3.30E-76	[LN:JN0097] [AC:JN0097] [PN:secreted 45K protein precursor] [OR:Lactococcus lactis]	83
SPX1714	1714	4375	80	240			NO-HIT	6
SPX1715	1715	4376	165	495	108	7.90E-08	[LN:MRED__BACSU] [AC:Q01467] [GN:MRED:RODB] [OR:Bacillus subtilis] [DE:ROD SHAPE-DETERMINING PROTEIN MRED] [SP:Q01467]	117
SPX1716	1716	4377	273	819	141	1.60E-17	[LN:MREC__BACSU] [AC:Q01466] [GN:MREC] [OR:Bacillus subtilis] [DE:ROD SHAPE-DETERMINING PROTEIN MREC] [SP:Q01466]	112
SPX1717	1717	4378	265	795	694	1.30E-95	[LN:F69742] [AC:F69742] [PN:hypothetical protein ybaF] [GN:ybaF] [OR:Bacillus subtilis]	87
SPX1718	1718	4379	57	171			NO-HIT	6
SPX1719	1719	4380	280	840	623	4.20E-94	[LN:E69742] [AC:E69742] [PN:ABC transporter (ATP-binding protein) homolog ybaE] [GN:ybaE] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX1720	1720	4381	116	348	127	7.40E-11	[LN:C71234] [AC:C71234] [PN:hypothetical protein PH0133] [GN:PH0133]	95

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1721	1721	4382	262	786	853	1.50E-116	[OR:Pyrococcus horikoshii] [GI:3426368] [LN:AF082738] [AC:AF082738] [PN:ABC transporter ATP-binding protein] [GN:stpA]	119
SPX1722	1722	4383	75	225	106	1.30E-09	[OR:Streptococcus pyogenes] [LN:A71007] [AC:A71007] [PN:hypothetical protein PH1351] [GN:PH1351]	95
SPX1723	1723	4384	182	546	531	6.40E-72	[OR:Pyrococcus horikoshii] [GI:3426367] [LN:AF082738] [AC:AF082738] [PN:phosphotidylglycerophosphate synthase] [GN:pgsA]	121
SPX1724	1724	4385	277	831	100	1.50E-21	[OR:Streptococcus pyogenes] [GI:3426366] [LN:AF082738] [AC:AF082738] [PN:unknown]	81
SPX1725	1725	4386	428	1284	403	1.10E-102	[OR:Streptococcus pyogenes] [GI:3426365] [LN:AF082738] [AC:AF082738] [PN:unknown]	81
SPX1726	1726	4387	417	1251	694	2.70E-103	[OR:Streptococcus pyogenes] [GI:3426364] [LN:AF082738] [AC:AF082738] [PN:unknown]	81
SPX1727	1727	4388	123	369	148	1.20E-26	[OR:Streptococcus pyogenes] [LN:JC4754] [AC:JC4754] [PN:hypothetical 13.6k protein] [GN:recF]	89
SPX1728	1728	4389	366	1098	1127	1.20E-163	[OR:Lactococcus lactis] [LN:RECF_STRPY] [AC:P49999] [GN:RECF] [OR:Streptococcus pyogenes] [DE:RECF PROTEIN] [SP:P49999]	95
SPX1729	1729	4390	493	1479	2283	0	[LN:IMDH_STRPY] [AC:P50099] [GN:GUAB] [OR:Streptococcus pyogenes] [EC:1.1.1.205] [DE:DEHYDROGENASE) (IMPDH) (IMPD)]	127

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1730	1730	4391	49	147	99	1.30E-07	[SP:P50099] [LN:F71456] [AC:F71456] [PN:hypothetical protein PH0308] [GN:PH0308] [OR:Pyrococcus horikoshii]	95
SPX1731	1731	4392	342	1026	672	2.00E-153	[LN:SYW_CLOLO] [AC:Q46127] [GN:TRPS:TRSA] [OR:Clostridium longisporum] [EC:6.1.1.2] [DE:(TRPRS)] [SP:Q46127]	108
SPX1732	1732	4393	541	1623	1364	9.40E-252	[LN:E69861] [AC:E69861] [PN:ABC transporter (ATP-binding protein) homolog ykpA] [GN:ykpA] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188
SPX1733	1733	4394	851	2553	503	1.40E-88	[GI:3043878] [LN:LLU95840] [AC:U95840] [PN:transmembrane protein Tmp5] [OR:Lactococcus lactis]	94
SPX1734	1734	4395	105	315	181	5.50E-19	[LN:YHGE_BACSU] [AC:P32399] [GN:YHGE] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB)] [SP:P32399]	144
SPX1735	1735	4396	113	339	138	1.10E-11	[LN:YHGE_BACSU] [AC:P32399] [GN:YHGE] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB)] [SP:P32399]	144
SPX1736	1736	4397	256	768	307	2.50E-46	[GI:3043882] [LN:LLU95842] [AC:U95842] [PN:transmembrane protein Tmp7] [OR:Lactococcus lactis]	94
SPX1737	1737	4398	181	543	98	5.30E-09	[GI:6899263] [LN:AE002125] [AC:AE002125:AF222894] [PN:unique hypothetical] [GN:UU290] [OR:Ureaplasma urealyticum]	113
SPX1738	1738	4399	251	753	1294	2.50E-176	[GI:1613769] [LN:SPU33315] [AC:U33315]	102

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1739	1739	4400	442	1326	2246	0	[PN:response regulator] [GN:comE] [OR:Streptococcus pneumoniae] [GI:1613768] [LN:SPU33315] [AC:U33315]	108
SPX1740	1740	4401	160	480	781	4.50E-104	[PN:histidine protein kinase] [GN:comD] [OR:Streptococcus pneumoniae] [GI:2109449] [LN:SPDNAARG] [AC:AF000658]	83
SPX1741	1741	4402	398	1194	1963	8.90E-259	[FN:unknown] [OR:Streptococcus pneumoniae] [GI:2109443] [LN:SPDNAARG] [AC:AF000658]	112
SPX1742	1742	4403	253	759	1225	5.80E-163	[PN:putative serine protease] [GN:sphtra] [OR:Streptococcus pneumoniae] [GI:2109444] [LN:SPDNAARG] [AC:AF000658]	107
SPX1743	1743	4404	454	1362	2287	0	[PN:SPSpoJ] [GN:spspoJ] [FN:unknown] [OR:Streptococcus pneumoniae] [LN:DNAA_STRPN] [AC:O08397] [GN:DNAA] [OR:Streptococcus pneumoniae] [DE:CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA] [SP:O08397]	131
SPX1744	1744	4405	379	1137	1857	2.90E-250	“[LN:DP3B_STRPN] [AC:O06672] [GN:DNAN] [OR:Streptococcus pneumoniae] [EC:2.7.7.7] [DE:DNA POLYMERASE III, BETA CHAIN,] [SP:O06672]”	131
SPX1745	1745	4406	65	195	339	1.60E-43	[GI:2109447] [LN:SPDNAARG] [AC:AF000658] [FN:unknown]	83
SPX1746	1746	4407	375	1125	710	9.10E-176	[OR:Streptococcus pneumoniae] [LN:YYAF_BACSU] [AC:P37518] [GN:YYAF] [OR:Bacillus subtilis] [DE:REGION]	84

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P37518]	
SPX1747	1747	4408	190	570	368	3.70E-62	[LN:SP5C__BACSU] [AC:P37470] [GN:SPOVC:PTH] [OR:Bacillus subtilis] [EC:3.1.1.29] [DE:SPORULATION PROTEIN C)] [SP:P37470]	119
SPX1748	1748	4409	1170	3510	2088	0	[GI:3511015] [LN:AF054624] [AC:AF054624] [PN:transcription-repair coupling factor] [GN:mfd] [OR:Lactobacillus sakei]	116
SPX1749	1749	4410	99	297			NO-HIT	6
SPX1750	1750	4411	89	267	275	3.50E-33	[LN:YABO__BACSU] [AC:P37557] [GN:YABO] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 9.7 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION] [SP:P37557]	136
SPX1751	1751	4412	123	369	112	1.50E-08	[GI:4090866] [LN:AF023181] [AC:AF023181] [PN:DivIC homolog] [GN:divL] [OR:Listeria monocytogenes]	97
SPX1752	1752	4413	74	222			NO-HIT	6
SPX1753	1753	4414	446	1338			NO-HIT	6
SPX1754	1754	4415	426	1278	128	6.70E-28	[LN:D72358] [AC:D72358] [PN:conserved hypothetical protein] [GN:TM0579] [CL:hypothetical protein HI0404] [OR:Thermotoga maritima]	129
SPX1755	1755	4416	181	543	374	1.10E-80	“[LN:HPRT__LACLA] [AC:Q02522] [GN:HPT] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.4.2.8] [DE:(HGPRTASE)] [SP:Q02522]”	142
SPX1756	1756	4417	653	1959	3277	0	[GI:5030426] [LN:AF061748] [AC:AF061748] [PN:cell division protein FtsH] [GN:ftsH] [OR:Streptococcus pneumoniae]	112
SPX1757	1757	4418	208	624	181	9.80E-25	[LN:A71115]	95

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:A71115] [PN:hypothetical protein PH0688] [GN:PH0688] [OR:Pyrococcus horikoshii]	
SPX1758	1758	4419	36	108			NO-HIT	6
SPX1759	1759	4420	160	480	844	9.40E-114	[GI:5739312] [LN:AF161700] [AC:AF161700] [PN:ComX1] [GN:comX1] [FN:transcriptional regulator of] [OR:Streptococcus pneumoniae]	126
SPX1760	1760	4421	82	246			NO-HIT	6
SPX1761	1761	4422	50	150	82	7.60E-06	[LN:G71244] [AC:G71244] [PN:hypothetical protein PH0217] [GN:PH0217] [OR:Pyrococcus horikoshii]	95
SPX1762	1762	4423	62	186			NO-HIT	6
SPX1763	1763	4424	87	261	136	3.90E-14	[LN:G81516] [AC:G81516] [PN:hypothetical protein CP0988 [imported]] [GN:CP0988] [OR:Chlamydophila pneumoniae:Chlamydia pneumoniae]	130
SPX1764	1764	4425	136	408			NO-HIT	6
SPX1765	1765	4426	52	156			NO-HIT	6
SPX1766	1766	4427	52	156	87	3.30E-06	[LN:D75542] [AC:D75542] [PN:hypothetical protein] [GN:DR0254] [OR:Deinococcus radiodurans]	90
SPX1767	1767	4428	82	246	247	1.10E-29	[LN:F81737] [AC:F81737] [PN:hypothetical protein TC0129 [imported]] [GN:TC0129] [OR:Chlamydia muridarum:Chlamydia trachomatis MoPn]	131
SPX1768	1768	4429	78	234	143	4.90E-15	[LN:F71245] [AC:F71245] [PN:hypothetical protein PHS004] [GN:PHS004] [OR:Pyrococcus horikoshii]	95
SPX1769	1769	4430	69	207	340	7.40E-44	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX1770	1770	4431	82	246	256	1.40E-36	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase]	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	
SPX1771	1771	4432	47	141	119	1.30E-09	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1772	1772	4433	91	273	387	3.30E-50	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX1773	1773	4434	81	243			NO-HIT	6
SPX1774	1774	4435	443	1329	1739	2.60E-235	“[LN:A42280] [AC:S65968:A42280:H69683] [PN:adenylosuccinate synthase, purA:IMP--aspartate ligase] [GN:purA] [CL:adenylosuccinate synthase] [OR:Bacillus subtilis] [EC:6.3.4.4]”	175
SPX1775	1775	4436	141	423	365	1.90E-46	[LN:YAAJ_BACSU] [AC:P21335] [GN:YAAJ] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 17.8 KD PROTEIN IN SERS-DNAH INTERGENIC REGION] [SP:P21335]	137
SPX1776	1776	4437	148	444	205	2.80E-38	[GI:2765131] [LN:LLABIKORF] [AC:Y11901] [PN:dUTPase] [OR:Lactococcus lactis]	76
SPX1777	1777	4438	186	558			NO-HIT	6
SPX1778	1778	4439	68	204			NO-HIT	6
SPX1779	1779	4440	455	1365	1167	1.60E-201	[LN:RADA_BACSU] [AC:P37572] [GN:RADA:SMS] [OR:Bacillus subtilis] [DE:DNA REPAIR PROTEIN RADA HOMOLOG (DNA REPAIR PROTEIN SMS HOMOLOG)] [SP:P37572]	146
SPX1780	1780	4441	207	621			NO-HIT	6
SPX1781	1781	4442	166	498	268	2.80E-38	[LN:Y023_MYCTU] [AC:Q10612] [GN:MTCY373.03] [OR:Mycobacterium tuberculosis] [DE:HYPOTHETICAL 18.2 KD PROTEIN CY373.03] [SP:Q10612]	130
SPX1782	1782	4443	264	792			NO-HIT	6
SPX1783	1783	4444	323	969	607	9.40E-148	[GI:532204] [LN:LISTMS]	118

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:M92842] [GN:prs] [OR:Listeria monocytogenes] [SR:Listeria monocytogenes (strain L028) DNA]	
SPX1784	1784	4445	160	480	493	3.90E-65	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX1785	1785	4446	113	339			NO-HIT	6
SPX1786	1786	4447	122	366			NO-HIT	6
SPX1787	1787	4448	131	393			NO-HIT	6
SPX1788	1788	4449	118	354			NO-HIT	6
SPX1789	1789	4450	890	2670	4412	0	“[LN:DPO1_STRPN] [AC:P13252] [GN:POLA] [OR:Streptococcus pneumoniae] [EC:2.7.7.7] [DE:DNA POLYMERASE I, (POL I)] [SP:P13252]”	125
SPX1790	1790	4451	124	372	181	3.30E-32	[LN:B69892] [AC:B69892] [PN:conserved hypothetical protein yneT] [GN:yneT] [CL:hypothetical protein yneT] [OR:Bacillus subtilis]	128
SPX1791	1791	4452	137	411	328	2.80E-41	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1792	1792	4453	105	315	83	0.00036	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1793	1793	4454	239	717	398	3.00E-50	[LN:YOR3_BACCE] [AC:O31352] [OR:Bacillus cereus] [DE:HYPOTHETICAL PROTEIN (ORF3) (FRAGMENT)] [SP:O31352]	104
SPX1794	1794	4455	171	513	212	5.40E-30	[LN:YF05_METTH] [AC:O27549] [GN:MTH1505] [OR:Methanobacterium thermoautotrophicum] [DE:HYPOTHETICAL PROTEIN MTH1505] [SP:O27549]	128
SPX1795	1795	4456	126	378			NO-HIT	6
SPX1796	1796	4457	167	501	456	2.50E-57	[LN:RL10_BACSU] [AC:P42923] [GN:RPLJ]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis] [DE:(VEGETATIVE PROTEIN 300) (VEG300)] [SP:P42923]	
SPX1797	1797	4458	123	369	255	1.30E-46	"[LN:RL7_MICLU] [AC:P02395] [GN:RPLL] [OR:Micrococcus luteus] [SR:,Micrococcus lysodeikticus] [DE:50S RIBOSOMAL PROTEIN L7/L12 (MA1/MA2)] [SP:P02395]"	150
SPX1798	1798	4459	115	345	98	1.30E-05	[LN:T40374] [AC:T40374] [PN:hypothetical protein SPBC3D6.14c] [OR:Schizosaccharomyces pombe]	92
SPX1799	1799	4460	196	588			NO-HIT	6
SPX1800	1800	4461	505	1515	214	1.60E-47	[LN:T35180] [AC:T35180] [PN:hypothetical protein SC5A7.31] [GN:SC5A7.31] [CL:Streptomyces coelicolor hypothetical protein SC5A7.31] [OR:Streptomyces coelicolor]	160
SPX1801	1801	4462	198	594	136	6.60E-20	[GI:290801] [LN:FRNVALAB] [AC:L17003] [GN:valA] [OR:Francisella tularensis var. novicida] [SR:Francisella novicida (strain U112) DNA]	133
SPX1802	1802	4463	97	291			NO-HIT	6
SPX1803	1803	4464	247	741	357	1.50E-55	[GI:7288062] [LN:SCD40A] [AC:AL161691] [PN:putative ABC-transporter ATP-binding protein] [GN:SCD40A.12c] [OR:Streptomyces coelicolor A3(2)]	139
SPX1804	1804	4465	60	180			NO-HIT	6
SPX1805	1805	4466	129	387	90	6.90E-06	[LN:F72598] [AC:F72598] [PN:hypothetical protein APE1254] [GN:APE1254] [OR:Aeropyrum pernix]	92
SPX1806	1806	4467	440	1320			NO-HIT	6
SPX1807	1807	4468	60	180			NO-HIT	6
SPX1808	1808	4469	74	222			NO-HIT	6
SPX1809	1809	4470	45	135	109	3.10E-09	"[LN:Y4PE_RHISN] [AC:P55614] [GN:Y4PE,Y4SA] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 15.5 KD PROTEIN Y4PE/Y4SA] [SP:P55614]"	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1810	1810	4471	88	264	301	4.50E-37	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1811	1811	4472	233	699	331	2.60E-64	"[LN:C70180] [AC:C70180] [PN:conserved hypothetical protein BB0644] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]"	125
SPX1812	1812	4473	308	924	98	4.30E-21	[LN:I40867] [AC:I40867] [PN:hypothetical protein 2] [OR:Clostridium perfringens]	80
SPX1813	1813	4474	543	1629	102	3.10E-21	[LN:B71130] [AC:B71130] [PN:probable oligopeptide binding protein APPA] [GN:PH0807] [CL:dipeptide transport protein] [OR:Pyrococcus horikoshii]	143
SPX1814	1814	4475	317	951	335	2.30E-78	"[LN:D75202] [AC:D75202] [PN:dipeptide abc transporter, dipeptide-binding protein PAB0092] [GN:dppB-1:PAB0092] [CL:transmembrane protein dppB] [OR:Pyrococcus abyssi]"	166
SPX1815	1815	4476	296	888	642	5.70E-87	[LN:APPC_BACSU] [AC:P42063] [GN:APPC] [OR:Bacillus subtilis] [DE:OLIGOPEPTIDE TRANSPORT PERMEASE PROTEIN APPC] [SP:P42063]	122
SPX1816	1816	4477	204	612			NO-HIT	6
SPX1817	1817	4478	661	1983	644	1.60E-146	[LN:YLIA_ECOLI] [AC:P75796] [GN:YLIA] [OR:Escherichia coli] [DE:HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YLIA] [SP:P75796]	130
SPX1818	1818	4479	66	198			NO-HIT	6
SPX1819	1819	4480	470	1410			NO-HIT	6
SPX1820	1820	4481	163	489	450	8.60E-58	[LN:YUTG_BACSU] [AC:O32124] [GN:YUTG] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 18.7 KD PROTEIN IN HOM-MRGA INTERGENIC REGION] [SP:O32124]	136
SPX1821	1821	4482	136	408	190	1.30E-38	"[LN:CDD_BACSU] [AC:P19079] [GN:CDD] [OR:Bacillus subtilis]	142

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[EC:3.5.4.5] [DE:CYTIDINE DEAMINASE, (CYTIDINE AMINOHYDROLASE) (CDA)] [SP:P19079]"	
SPX1822	1822	4483	95	285			NO-HIT	6
SPX1823	1823	4484	274	822	771	1.10E-101	[GI:2385360] [LN:CTSIALIDA] [AC:Y08695] [PN:putative acylneuraminate lyase] [OR:Clostridium tertium]	100
SPX1824	1824	4485	68	204	173	1.10E-18	[LN:S43901] [AC:S43901:S27537] [PN:hypothetical protein A] [OR:Clostridium perfringens]	87
SPX1825	1825	4486	173	519	145	3.80E-20	[GI:2668605] [LN:AF015453] [AC:AF015453] [PN:unknown] [OR:Lactobacillus rhamnosus]	82
SPX1826	1826	4487	214	642	509	7.20E-64	[GI:6318176] [LN:BSP250862] [AC:AJ250862] [PN:MrsT protein] [GN:mrsT] [FN:putative ABC-transporter] [OR:Bacillus sp. HIL-Y85/54728]	131
SPX1827	1827	4488	63	189			NO-HIT	6
SPX1828	1828	4489	449	1347	731	5.00E-191	[GI:3702805] [LN:AF056335] [AC:AF056335] [PN:NADP-specific glutamate dehydrogenase] [GN:gdhA] [OR:Bacillus licheniformis]	121
SPX1829	1829	4490	138	414	99	6.40E-06	[LN:S31840] [AC:S31840] [PN:probable transposase] [OR:Bacillus stearothermophilus]	82
SPX1830	1830	4491	57	171	94	3.30E-07	"[LN:Y4PE_RHISN] [AC:P55614] [GN:Y4PE,Y4SA] [OR:Rhizobium sp] [SR:,strain NGR234] [DE:HYPOTHETICAL 15.5 KD PROTEIN Y4PE/Y4SA] [SP:P55614]"	138
SPX1831	1831	4492	218	654	154	1.00E-20	[GI:722339] [LN:AXU22323] [AC:U22323] [PN:unknown] [OR:Acetobacter xylinus] [SR:Acetobacter xylinum]	100
SPX1832	1832	4493	73	219			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1833	1833	4494	95	285	146	8.10E-34	[GI:4193373] [LN:AF072894] [AC:AF072894] [PN:ribosomal protein L31] [GN:rpmE] [OR:Listeria monocytogenes]	105
SPX1834	1834	4495	312	936	397	7.80E-97	[LN:F69999] [AC:F69999] [PN:conserved hypothetical protein ytqI] [GN:ytqI] [CL:Mycoplasma conserved hypothetical protein MG190] [OR:Bacillus subtilis]	150
SPX1835	1835	4496	145	435	271	2.00E-32	[LN:FLAV_BACSU] [AC:O34737] [GN:YKUN] [OR:Bacillus subtilis] [DE:PROBABLE FLAVODOXIN 1] [SP:O34737]	99
SPX1836	1836	4497	89	267	117	2.80E-11	[LN:Y246_METJA] [AC:Q57696] [GN:MJ0246] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ0246] [SP:Q57696]	114
SPX1837	1837	4498	125	375	169	5.60E-18	[LN:C75108] [AC:C75108] [PN:crcb protein PAB1925] [GN:PAB1925] [CL:hypothetical protein MJ1523] [OR:Pyrococcus abyssi]	118
SPX1838	1838	4499	110	330	93	1.20E-10	[LN:B71026] [AC:B71026] [PN:hypothetical protein PH1502] [GN:PH1502] [CL:hypothetical protein MJ1523] [OR:Pyrococcus horikoshii]	128
SPX1839	1839	4500	116	348	537	4.70E-71	[LN:RL19_STRTR] [AC:O34031] [GN:RPLS] [OR:Streptococcus thermophilus] [DE:50S RIBOSOMAL PROTEIN L19] [SP:O34031]	112
SPX1840	1840	4501	269	807	354	2.70E-85	[LN:YIDA_ECOLI] [AC:P09997:P76737] [GN:YIDA] [OR:Escherichia coli] [DE:HYPOTHETICAL 29.7 KD PROTEIN IN IBPA-GYRB INTERGENIC REGION] [SP:P09997:P76737]	150
SPX1841	1841	4502	448	1344	561	2.90E-126	[LN:YWFO_BACSU] [AC:P39651]	130

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:YWFO:IPA-93D] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 51.0 KD PROTEIN IN PTA 3'REGION] [SP:P39651]	
SPX1842	1842	4503	126	378			NO-HIT	6
SPX1843	1843	4504	111	333	499	3.70E-65	[GI:1850606] [LN:SMU88582] [AC:U88582] [PN:YlxM] [GN:ylxM] [FN:unknown] [OR:Streptococcus mutans]	97
SPX1844	1844	4505	185	555			NO-HIT	6
SPX1845	1845	4506	524	1572	1993	5.90E-290	[LN:SR54_STRMU] [AC:Q54431:P96469] [GN:FFH] [OR:Streptococcus mutans] [DE:SIGNAL RECOGNITION PARTICLE PROTEIN (FIFTY-FOUR HOMOLOG)] [SP:Q54431:P96469]	150
SPX1846	1846	4507	163	489			NO-HIT	6
SPX1847	1847	4508	257	771	500	1.10E-87	[LN:PYRP_BACCL] [AC:P41006] [GN:PYRP] [OR:Bacillus caldolyticus] [DE:URACIL PERMEASE (URACIL TRANSPORTER)] [SP:P41006]	118
SPX1848	1848	4509	147	441	231	5.00E-42	[LN:LLA132624] [AC:AJ132624] [PN:uracil transporter] [GN:pyrP] [OR:Lactococcus lactis]	86
SPX1849	1849	4510	238	714	353	3.70E-81	[LN:GIDB_BACSU] [AC:P25813] [GN:GIDB] [OR:Bacillus subtilis] [DE:GLUCOSE INHIBITED DIVISION PROTEIN B] [SP:P25813]	114
SPX1850	1850	4511	187	561	261	1.60E-57	[GI:1519287] [LN:LMU66186] [AC:U66186] [PN:LemA] [GN:lemA] [OR:Listeria monocytogenes]	86
SPX1851	1851	4512	303	909	757	1.70E-165	[LN:HTPX_STRGC] [AC:O30795] [GN:HTPX] [OR:Streptococcus gordonii challis] [DE:PUTATIVE HEAT SHOCK PROTEIN HTPX] [SP:O30795]	123
SPX1852	1852	4513	520	1560	197	1.40E-47	[LN:YHES_ECOLI]	130

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P45535] [GN:YHES] [OR:Escherichia coli] [DE:HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YHES] [SP:P45535]	
SPX1853	1853	4514	92	276			NO-HIT	6
SPX1854	1854	4515	122	366			NO-HIT	6
SPX1855	1855	4516	181	543	64	6.40E-07	[LN:C69875] [AC:C69875] [PN:hypothetical protein ylbN] [GN:ylbN] [OR:Bacillus subtilis]	87
SPX1856	1856	4517	62	186			NO-HIT	6
SPX1857	1857	4518	210	630	446	2.10E-57	[LN:END3__BACSU] [AC:P39788] [GN:NTH:JOOB] [OR:Bacillus subtilis] [EC:4.2.99.18] [DE:APYRIMIDINIC SITE) LYASE)] [SP:P39788]	122
SPX1858	1858	4519	174	522	626	4.30E-82	[LN:LLA132624] [AC:AJ132624] [PN:pyrimidine regulatory protein] [GN:pyrR] [FN:Regulates expression of the pyrimidine] [OR:Lactococcus lactis]	141
SPX1859	1859	4520	308	924	942	6.40E-134	[LN:LLA132624] [AC:AJ132624] [PN:aspartate transcarbamoylase] [GN:pyrB] [OR:Lactococcus lactis]	95
SPX1860	1860	4521	360	1080	1492	2.60E-202	[LN:LLA132624] [AC:AJ132624] [PN:carbamoyl phosphate synthetase small subunit] [GN:carA] [OR:Lactococcus lactis]	112
SPX1861	1861	4522	1059	3177	4000	0	[GI:2598551] [LN:LLAJ109] [AC:AJ000109] [PN:carbamoylphosphate synthetase] [GN:carB] [OR:Lactococcus lactis]	108
SPX1862	1862	4523	70	210			NO-HIT	6
SPX1863	1863	4524	281	843	1425	1.90E-195	[GI:5001693] [LN:AF106539] [AC:AF106539] [PN:LicD2] [GN:licD2] [OR:Streptococcus pneumoniae]	92
SPX1864	1864	4525	268	804	1419	6.40E-194	[GI:5001692]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:AF106539] [AC:AF106539] [PN:LicD1] [GN:licD1] [OR:Streptococcus pneumoniae]	
SPX1865	1865	4526	496	1488	2483	0	[GI:5001691] [LN:AF106539] [AC:AF106539] [PN:unknown] [OR:Streptococcus pneumoniae]	83
SPX1866	1866	4527	236	708	96	7.30E-17	[LN:S60902] [AC:S60902:S49238:S44071] [PN:CDP-ribitol pyrophosphorylase] [OR:Haemophilus influenzae]	100
SPX1867	1867	4528	347	1041	66	6.20E-05	[LN:YJJN_ECOLI] [AC:P39400] [GN:YJJN] [OR:Escherichia coli] [DE:INTERGENIC REGION] [SP:P39400]	94
SPX1868	1868	4529	290	870	1361	2.10E-181	[GI:2708632] [LN:AF036951] [AC:AF036951] [PN:choline kinase] [GN:pck] [OR:Streptococcus pneumoniae]	99
SPX1869	1869	4530	293	879	156	1.70E-15	[LN:LICB_HAEIN] [AC:P14182:Q57357:O05075] [GN:LICB:HI1538] [OR:Haemophilus influenzae] [DE:LICB PROTEIN] [SP:P14182:Q57357:O05075]	130
SPX1870	1870	4531	230	690	345	4.10E-54	[LN:D64128] [AC:D64128] [PN:lic-1 protein C] [GN:licC] [OR:Haemophilus influenzae]	82
SPX1871	1871	4532	283	849	458	8.00E-59	[LN:C72399] [AC:C72399] [PN:DNA processing chain A] [GN:TM0250] [OR:Thermotoga maritima]	88
SPX1872	1872	4533	171	513	309	5.10E-41	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1873	1873	4534	152	456	107	1.30E-07	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1874	1874	4535	69	207			NO-HIT	6
SPX1875	1875	4536	348	1044	179	2.10E-23	[LN:Y678_METJA] [AC:Q58091] [GN:MJ0678] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ0678] [SP:Q58091]	114
SPX1876	1876	4537	702	2106	1065	0	[LN:TOP1_BACSU] [AC:P39814] [GN:TOPA:TOPI] [OR:Bacillus subtilis] [EC:5.99.1.2] [DE:(UNTWISTING ENZYME) (SWIVELASE)] [SP:P39814]	128
SPX1877	1877	4538	120	360	208	4.20E-25	[LN:YBAN_ECOLI] [AC:P45808:P77478] [GN:YBAN] [OR:Escherichia coli] [DE:HYPOTHETICAL 14.8 KD PROTEIN IN PRIC-APT INTERGENIC REGION] [SP:P45808:P77478]	149
SPX1878	1878	4539	211	633	181	2.30E-34	[GI:4680703] [LN:AF132966] [AC:AF132966] [PN:CGI-32 protein] [OR:Homo sapiens] [SR:human]	89
SPX1879	1879	4540	219	657	142	3.40E-22	[LN:YYAQ_BACSU] [AC:P37507] [GN:YYAQ] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 13.9 KD PROTEIN IN COTF-TETB INTERGENIC REGION] [SP:P37507]	137
SPX1880	1880	4541	211	633	596	1.30E-93	[GI:2565151] [LN:LLU92974] [AC:U92974:M90760:M90761] [PN:LeuA] [GN:leuA] [OR:Lactococcus lactis]	96
SPX1881	1881	4542	70	210			NO-HIT	6
SPX1882	1882	4543	145	435	308	6.00E-39	“[LN:LEU1_LACLA] [AC:Q02141] [GN:LEUA] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.1.3.12] [DE:SYNTHASE) (ALPHA-IPM SYNTHETASE)] [SP:Q02141]”	166
SPX1883	1883	4544	346	1038	608	4.50E-153	“[LN:LEU3_LACLA] [AC:Q02143] [GN:LEUB]	151

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:1.1.1.85] [DE:(IMDH) (3-IPM-DH)] [SP:Q02143]"	
SPX1884	1884	4545	90	270	129	8.30E-13	[LN:G69983] [AC:G69983] [PN:hypothetical protein ysdA] [GN:ysdA] [OR:Bacillus subtilis]	87
SPX1885	1885	4546	120	360	373	9.40E-47	"[LN:LEUD_LACLA] [AC:Q02144] [GN:LEUD] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:4.2.1.33] [DE:(ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM ISOMERASE)] [SP:Q02144]"	183
SPX1886	1886	4547	79	237			NO-HIT	6
SPX1887	1887	4548	89	267			NO-HIT	6
SPX1888	1888	4549	645	1935	138	1.60E-33	"[LN:MCRB_ECOLI] [AC:P15005] [GN:MCRB:RGLB] [OR:Escherichia coli] [EC:3.1.21.-] [DE:5-METHYLCYTOSINE-SPECIFIC RESTRICTION ENZYME B,] [SP:P15005]"	145
SPX1889	1889	4550	448	1344	96	4.40E-09	[LN:D81431] [AC:D81431] [PN:hypothetical protein Cj0140 [imported]] [GN:Cj0140] [OR:Campylobacter jejuni]	105
SPX1890	1890	4551	329	987	913	1.40E-171	"[LN:C70015] [AC:C70015] [PN:probable GMP reductase, yumD:guanosine monophosphate reductase] [GN:yumD] [CL:GMP reductase:IMP dehydrogenase amino-terminal homology:IMP dehydrogenase catalytic homology] [OR:Bacillus subtilis] [EC:1.6.6.8]"	237
SPX1891	1891	4552	78	234			NO-HIT	6
SPX1892	1892	4553	67	201			NO-HIT	6
SPX1893	1893	4554	233	699	538	1.70E-69	"[LN:B69693] [AC:B69693:JC4821] [PN:ribonuclease III,:RNase D:RNase O] [GN:rncS:srb] [CL:ribonuclease III:double-stranded RNA-binding repeat homology] [OR:Bacillus subtilis] [EC:3.1.26.3]"	188
SPX1894	1894	4555	737	2211	1258	7.10E-169	[LN:G69708] [AC:G69708:JC4819:PC4029]	189

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:chromosome segregation SMC protein:minichromosome stabilizing protein SMC] [GN:smc] [CL:conserved hypothetical P115 protein] [OR:Bacillus subtilis]	
SPX1895	1895	4556	217	651	139	2.70E-09	[GI:2246532] [LN:U93872] [AC:U93872] [OR:Kaposi's sarcoma-associated herpesvirus] [SR:Kaposi's sarcoma-associated herpesvirus - Human herpesvirus 8]	148
SPX1896	1896	4557	102	306	95	0.00014	[LN:P115_MYCHR] [AC:P41508] [OR:Mycoplasma hyorhinis] [DE:P115 PROTEIN] [SP:P41508]	83
SPX1897	1897	4558	196	588	428	2.90E-69	[LN:G69708] [AC:G69708:JC4819:PC4029] [PN:chromosome segregation SMC protein:minichromosome stabilizing protein SMC] [GN:smc] [CL:conserved hypothetical P115 protein] [OR:Bacillus subtilis]	189
SPX1898	1898	4559	265	795	206	4.40E-31	“[GI:4062428] [LN:D90722] [AC:D90722:AB001340] [PN:Hypothetical 30.2 kd protein in idh-deoR] [OR:Escherichia coli] [SR:Escherichia coli(strain:K12) DNA, clone:Kohara clone #209]”	178
SPX1899	1899	4560	273	819	220	3.10E-42	[LN:YIDA_ECOLI] [AC:P09997:P76737] [GN:YIDA] [OR:Escherichia coli] [DE:HYPOTHETICAL 29.7 KD PROTEIN IN IBPA-GYRB INTERGENIC REGION] [SP:P09997:P76737]	150
SPX1900	1900	4561	430	1290	675	1.30E-113	[GI:2633967] [LN:BSUB0009] [AC:Z99112:AL009126] [PN:signal recognition particle (docking protein)] [GN:ftsY] [FN:involved in secretion of extracellular proteins] [OR:Bacillus subtilis]	184
SPX1901	1901	4562	419	1257	2039	1.30E-284	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1902	1902	4563	85	255	350	2.40E-45	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1903	1903	4564	109	327	288	1.40E-35	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX1904	1904	4565	173	519	858	5.70E-116	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX1905	1905	4566	496	1488	779	6.50E-166	[LN:G6PD_BACSU] [AC:P54547] [GN:ZWF] [OR:Bacillus subtilis] [EC:1.1.1.49] [DE:PROTEIN 11) (VEG11)] [SP:P54547]	110
SPX1906	1906	4567	247	741	772	4.80E-102	[LN:H69334] [AC:H69334] [PN:glutamine transport protein glnQ] [GN:glnQ] [CL:inner membrane protein malK:ATP-binding cassette homology] [OR:Archaeoglobus fulgidus]	162
SPX1907	1907	4568	493	1479	508	2.50E-80	[GI:6560693] [LN:AF141644] [AC:AF141644] [PN:putative integral membrane protein] [FN:putative inner membrane component of a] [OR:Lactococcus lactis]	148
SPX1908	1908	4569	237	711	91	0.00087	“[LN:S76167] [AC:S76167] [PN:hypothetical protein] [OR:Synechocystis sp.] [SR:PCC 6803, , PCC 6803] [SR:PCC 6803,]”	116
SPX1909	1909	4570	90	270	80	4.20E-05	[LN:T34651] [AC:T34651] [PN:probable transmembrane protein] [GN:SC1A9.02] [OR:Streptomyces coelicolor]	102
SPX1910	1910	4571	663	1989	3350	0	[LN:UVRB_STRPN] [AC:Q54986] [GN:UVRB:UVS402] [OR:Streptococcus pneumoniae] [DE:EXCINUCLEASE ABC SUBUNIT B] [SP:Q54986]	118
SPX1911	1911	4572	188	564	110	3.20E-15	[LN:G75474] [AC:G75474] [PN:probable acetyltransferase]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:DR0796] [OR:Deinococcus radiodurans]	
SPX1912	1912	4573	130	390	65	1.50E-06	[LN:T39482] [AC:T39482] [PN:N-acetyltransferase] [GN:SPBC15D4.06] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimI] [OR:Schizosaccharomyces pombe]	169
SPX1913	1913	4574	156	468	110	3.50E-09	[LN:G70031] [AC:G70031] [PN:mutator MutT protein homolog yvcI] [GN:yvcI] [CL:mutT domain homology] [OR:Bacillus subtilis]	121
SPX1914	1914	4575	172	516	206	7.90E-37	"[LN:F72234] [AC:F72234] [PN:transcription regulator, biotin repressor family] [GN:TM1602] [OR:Thermotoga maritima]"	116
SPX1915	1915	4576	194	582			NO-HIT	6
SPX1916	1916	4577	167	501	81	4.70E-09	[GI:806536] [LN:BAMALAMYA] [AC:Z22520] [PN:membrane protein] [OR:Bacillus acidopullulyticus]	92
SPX1917	1917	4578	184	552	465	2.50E-59	[LN:DFP__STRMU] [AC:Q54433] [GN:DFP] [OR:Streptococcus mutans] [DE:DNA/PANTOTHENATE METABOLISM FLAVOPROTEIN HOMOLOG (FRAGMENT)] [SP:Q54433]	138
SPX1918	1918	4579	253	759	120	2.90E-19	[LN:D69029] [AC:D69029] [PN:pantothenate metabolism flavoprotein dfp homolog MTH1216:probable aspartate 1-decarboxylase activase] [GN:MTH1216] [CL:pantothenate metabolism flavoprotein dfp] [OR:Methanobacterium thermoautotrophicum]	230
SPX1919	1919	4580	181	543			NO-HIT	6
SPX1920	1920	4581	557	1671	2475	0	[LN:FTHS__STRMU] [AC:Q59925:Q59926] [GN:FHS] [OR:Streptococcus mutans] [EC:6.3.4.3] [DE:SYNTHETASE) (FHS) (FTHFS)] [SP:Q59925:Q59926]	132
SPX1921	1921	4582	80	240	145	9.20E-16	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1922	1922	4583	392	1176	989	4.20E-132	[GI:6782414]	166

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SPN271596] [AC:AJ271596] [PN:A/G specific adenine glycosylase] [GN:mutY] [FN:antimutator prevents C to A transversions] [OR:Streptococcus pneumoniae]	
SPX1923	1923	4584	235	705	1194	9.00E-161	[GI:5830523] [LN:SPAJ6392] [AC:AJ006392] [PN:response regulator] [GN:rr02] [OR:Streptococcus pneumoniae]	104
SPX1924	1924	4585	452	1356	2249	0	[GI:5830524] [LN:SPAJ6392] [AC:AJ006392] [PN:histidine kinase] [GN:hk02] [OR:Streptococcus pneumoniae]	102
SPX1925	1925	4586	270	810	1379	5.30E-185	[GI:6689278] [LN:SPN012049] [AC:AJ012049] [PN:VicX protein] [GN:vicX] [FN:unknown] [OR:Streptococcus pneumoniae]	112
SPX1926	1926	4587	81	243			NO-HIT	6
SPX1927	1927	4588	88	264	93	1.90E-06	[LN:D70886] [AC:D70886] [PN:hypothetical protein Rv2866] [GN:Rv2866] [OR:Mycobacterium tuberculosis]	100
SPX1928	1928	4589	258	774	162	3.00E-20	[LN:C72692] [AC:C72692] [PN:probable potassium channel APE0955] [GN:APE0955] [OR:Aeropyrum pernix]	98
SPX1929	1929	4590	170	510			NO-HIT	6
SPX1930	1930	4591	333	999	1652	9.40E-223	[GI:2275101] [LN:SPR6LDH] [AC:AJ000336] [PN:L-lactate dehydrogenase] [GN:ldh] [FN:conversion of pyruvate to lactate] [OR:Streptococcus pneumoniae]	146
SPX1931	1931	4592	316	948	1617	3.30E-219	“[LN:GYRA_STRPN] [AC:P72524:Q54716:P72536] [GN:GYRA] [OR:Streptococcus pneumoniae] [EC:5.99.1.3] [DE:DNA GYRASE SUBUNIT A,]	150

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P72524:Q54716:P72536]”	
SPX1932	1932	4593	119	357	214	7.70E-24	“[LN:GYRA_STRPN] [AC:P72524:Q54716:P72536] [GN:GYRA] [OR:Streptococcus pneumoniae] [EC:5.99.1.3] [DE:DNA GYRASE SUBUNIT A,] [SP:P72524:Q54716:P72536]”	150
SPX1933	1933	4594	507	1521	2300	0	[GI:4138535] [LN:SPN5815] [AC:AJ005815] [PN:DNA gyrase subunit A] [GN:gyrA] [OR:Streptococcus pneumoniae]	105
SPX1934	1934	4595	248	744	299	5.50E-36	[GI:488339] [LN:SYNGIP3124] [AC:M77279] [PN:alpha-amylase] [OR:unidentified cloning vector] [SR:Cloning vector (sub_species Cloning vector pGIP3124) DNA]	153
SPX1935	1935	4596	266	798	184	6.20E-26	[GI:4433636] [LN:AF029224] [AC:AF029224:AF029225] [PN:NirC] [GN:nirC] [FN:putative nitrite transporter] [OR:Staphylococcus carnosus]	132
SPX1936	1936	4597	372	1116	426	6.30E-115	“[LN:T44655] [AC:T44655] [PN:O-acetylhomoserine (thiol)-lyase, [imported] :O-acetylhomoserine sulfhydrylase] [CL:O-succinylhomoserine (thiol)-lyase] [OR:Leptospira meyeri] [EC:4.2.99.10]”	186
SPX1937	1937	4598	425	1275	101	4.30E-18	[GI:6899348] [LN:AE002133] [AC:AE002133:AF222894] [PN:conserved hypothetical] [GN:UU367] [OR:Ureaplasma urealyticum]	116
SPX1938	1938	4599	293	879	661	5.10E-86	[LN:TRUB_BACSU] [AC:P32732] [GN:TRUB] [OR:Bacillus subtilis] [EC:4.2.1.70] [DE:HYDROLYASE)] [SP:P32732]	103
SPX1939	1939	4600	101	303			NO-HIT	6
SPX1940	1940	4601	239	717	707	3.50E-93	[LN:G69728] [AC:G69728]	99

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:uridine kinase udk] [GN:udk] [CL:uridine kinase] [OR:Bacillus subtilis]	
SPX1941	1941	4602	447	1341	633	1.60E-115	[LN:EX7L_BACSU] [AC:P54521] [GN:YQIB] [OR:Bacillus subtilis] [EC:3.1.11.6] [DE:VII LARGE SUBUNIT)] [SP:P54521]	110
SPX1942	1942	4603	71	213	113	3.50E-10	[LN:EX7S_ECOLI] [AC:P22938] [GN:XSEB] [OR:Escherichia coli] [EC:3.1.11.6] [DE:SMALL SUBUNIT)] [SP:P22938]	105
SPX1943	1943	4604	292	876	535	1.30E-78	“[LN:ISPA_MICLU] [AC:O66126] [GN:FPS] [OR:Micrococcus luteus] [SR:,Micrococcus lysodeikticus] [EC:2.5.1.10] [DE:(FPP SYNTHASE)] [SP:O66126]”	140
SPX1944	1944	4605	246	738	353	1.20E-79	[LN:YQXC_BACSU] [AC:P19672] [GN:YQXC:YQIF] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 29.7 KD PROTEIN IN FOLD-AHRC INTERGENIC REGION] [SP:P19672]	142
SPX1945	1945	4606	144	432	182	4.10E-26	[GI:4127534] [LN:BSAJ10954] [AC:AJ010954] [PN:arginine repressor] [GN:argR] [FN:ADN binding protein] [OR:Bacillus stearothermophilus]	133
SPX1946	1946	4607	298	894	317	4.20E-76	[LN:RECN_BACSU] [AC:P17894:P19671] [GN:RECN] [OR:Bacillus subtilis] [DE:DNA REPAIR PROTEIN RECN (RECOMBINATION PROTEIN N)] [SP:P17894:P19671]	141
SPX1947	1947	4608	261	783	395	3.20E-68	[LN:RECN_BACSU] [AC:P17894:P19671] [GN:RECN] [OR:Bacillus subtilis] [DE:DNA REPAIR PROTEIN RECN (RECOMBINATION PROTEIN N)]	141

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P17894:P19671]	
SPX1948	1948	4609	243	729	97	4.90E-15	[GI:2352096] [LN:U97022] [AC:U97022] [OR:Fervidobacterium islandicum]	69
SPX1949	1949	4610	345	1035	1284	9.40E-173	[LN:LEPA_BACSU] [AC:P37949] [GN:LEPA] [OR:Bacillus subtilis] [DE:GTP-BINDING PROTEIN LEPA] [SP:P37949]	102
SPX1950	1950	4611	304	912	1158	1.80E-156	[LN:LEPA_BACSU] [AC:P37949] [GN:LEPA] [OR:Bacillus subtilis] [DE:GTP-BINDING PROTEIN LEPA] [SP:P37949]	102
SPX1951	1951	4612	111	333	95	7.80E-05	“[LN:T04991] [AC:T04991] [PN:hypothetical protein T16L1.230] [OR:Arabidopsis thaliana] [SR:, mouse-ear cress]”	110
SPX1952	1952	4613	112	336	294	3.40E-35	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX1953	1953	4614	96	288			NO-HIT	6
SPX1954	1954	4615	230	690	70	3.10E-05	[LN:G75468] [AC:G75468] [PN:hypothetical protein] [GN:DR0857] [OR:Deinococcus radiodurans]	90
SPX1955	1955	4616	60	180			NO-HIT	6
SPX1956	1956	4617	82	246			NO-HIT	6
SPX1957	1957	4618	70	210			NO-HIT	6
SPX1958	1958	4619	68	204	138	3.20E-14	“[LN:LAFX_LACJO] [AC:Q48509] [GN:LAFX] [OR:Lactobacillus johnsonii] [DE:BACTERIOCIN LACTACIN F, SUBUNIT LAFX PRECURSOR] [SP:Q48509]”	132
SPX1959	1959	4620	85	255	114	8.90E-21	[GI:5441255] [LN:AB029612] [AC:AB029612] [PN:gassericin T1] [GN:gatA] [OR:Lactobacillus gasseri] [SR:Lactobacillus gasseri (strain:SBT2055) DNA]	144
SPX1960	1960	4621	56	168	79	0.00078	[GI:6751696] [LN:ATAC018908]	98

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AC018908] [GN:T7P1.21] [OR:Arabidopsis thaliana] [SR:thale cress]	
SPX1961	1961	4622	182	546	379	2.90E-48	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1962	1962	4623	133	399	364	8.90E-50	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX1963	1963	4624	55	165			NO-HIT	6
SPX1964	1964	4625	68	204			NO-HIT	6
SPX1965	1965	4626	69	207			NO-HIT	6
SPX1966	1966	4627	68	204	55	0.00018	[GI:6457574] [LN:AF200347] [AC:AF200347] [PN:lactocin 705 beta-subunit precursor] [OR:Lactobacillus casei]	106
SPX1967	1967	4628	73	219			NO-HIT	6
SPX1968	1968	4629	718	2154	2269	0	[LN:COMA_STRPN] [AC:Q03727] [GN:COMA] [OR:Streptococcus pneumoniae] [DE:TRANSPORT ATP-BINDING PROTEIN COMA] [SP:Q03727]	119
SPX1969	1969	4630	83	249	80	0.00032	[LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061] [OR:Aeropyrum pernix]	92
SPX1970	1970	4631	454	1362	697	7.40E-90	[GI:1698422] [LN:SGU40139] [AC:U40139] [PN:ComB] [GN:comB] [OR:Streptococcus gordonii] [SR:Streptococcus gordonii strain=Challis]	129
SPX1971	1971	4632	67	201			NO-HIT	6
SPX1972	1972	4633	110	330			NO-HIT	6
SPX1973	1973	4634	447	1341	2062	1.10E-288	[GI:5830551] [LN:SPAJ6401] [AC:AJ006401] [PN:histidine kinase] [GN:hk13] [OR:Streptococcus pneumoniae]	102
SPX1974	1974	4635	215	645	972	3.30E-133	[GI:5830550] [LN:SPAJ6401] [AC:AJ006401]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1975	1975	4636	61	183	239	4.40E-28	[PN:response regulator] [GN:rr13] [OR:Streptococcus pneumoniae] [GI:5830550] [LN:SPAJ6401] [AC:AJ006401]	104
SPX1976	1976	4637	113	339	69	0.000063	[PN:response regulator] [GN:rr13] [OR:Streptococcus pneumoniae] [GI:1495671] [LN:LPATOVGNS] [AC:X94434] [PN:response regulator PlnC] [GN:plnC] [OR:Lactobacillus plantarum]	107
SPX1977	1977	4638	110	330			NO-HIT	6
SPX1978	1978	4639	69	207			NO-HIT	6
SPX1979	1979	4640	324	972	150	4.10E-24	[LN:ECSB_BACSU] [AC:P55340] [GN:ECSB:PRST] [OR:Bacillus subtilis] [DE:PROTEIN ECSB] [SP:P55340]	95
SPX1980	1980	4641	74	222			NO-HIT	6
SPX1981	1981	4642	115	345	333	4.00E-42	[LN:ECSA_BACSU] [AC:P55339] [GN:ECSA:PRST] [OR:Bacillus subtilis] [DE:ABC-TYPE TRANSPORTER ATP-BINDING PROTEIN ECSA] [SP:P55339]	128
SPX1982	1982	4643	126	378	391	2.60E-49	[LN:ECSA_BACSU] [AC:P55339] [GN:ECSA:PRST] [OR:Bacillus subtilis] [DE:ABC-TYPE TRANSPORTER ATP-BINDING PROTEIN ECSA] [SP:P55339]	128
SPX1983	1983	4644	137	411	334	8.80E-42	[LN:HIT_BACSU] [AC:O07513] [GN:HIT] [OR:Bacillus subtilis] [DE:HIT PROTEIN] [SP:O07513]	87
SPX1984	1984	4645	96	288			NO-HIT	6
SPX1985	1985	4646	85	255			NO-HIT	6
SPX1986	1986	4647	379	1137	1885	4.90E-254	[LN:DNAJ_STRPN] [AC:P95830] [GN:DNAJ] [OR:Streptococcus pneumoniae] [DE:DNAJ PROTEIN (FRAGMENT)] [SP:P95830]	108

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX1987	1987	4648	47	141	125	7.20E-10	[GI:5305335] [LN:AF071081] [AC:AF071081] [PN:proline-rich mucin homolog] [OR:Mycobacterium tuberculosis]	104
SPX1988	1988	4649	119	357			NO-HIT	6
SPX1989	1989	4650	86	258			NO-HIT	6
SPX1990	1990	4651	608	1824	3008	0	[LN:DNAK__STRPN] [AC:P95829:O66035] [GN:DNAK] [OR:Streptococcus pneumoniae] [DE:DNAK PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70)] [SP:P95829:O66035]	143
SPX1991	1991	4652	120	360			NO-HIT	6
SPX1992	1992	4653	218	654			NO-HIT	6
SPX1993	1993	4654	183	549	346	3.60E-74	"[LN:GRPE__LACLA] [AC:P42369] [GN:GRPE] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:GRPE PROTEIN] [SP:P42369]"	132
SPX1994	1994	4655	356	1068	1712	6.90E-235	[GI:4566769] [LN:AF117740] [AC:AF117740] [PN:heat shock transcription repressor HrcA] [GN:hrcA] [OR:Streptococcus pneumoniae]	125
SPX1995	1995	4656	170	510			NO-HIT	6
SPX1996	1996	4657	778	2334	570	9.20E-238	[LN:I41291] [AC:I41291] [PN:EcoA type I restriction-modification enzyme R subunit] [OR:Escherichia coli]	104
SPX1997	1997	4658	488	1464	846	1.80E-144	[LN:I41293] [AC:I41293] [PN:EcoE type I restriction modification enzyme M subunit] [CL:site-specific methyltransferase (adenine-specific) EcoK] [OR:Escherichia coli]	165
SPX1998	1998	4659	517	1551	323	1.60E-37	[GI:6899439] [LN:AE002141] [AC:AE002141:AF222894] [PN:type I restriction enzyme S protein (fragment)] [GN:hsdS-5] [OR:Ureaplasma urealyticum]	141
SPX1999	1999	4660	430	1290	328	1.40E-38	[GI:6899439] [LN:AE002141] [AC:AE002141:AF222894] [PN:type I restriction enzyme S protein (fragment)] [GN:hsdS-5] [OR:Ureaplasma urealyticum]	141

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2000	2000	4661	74	222			NO-HIT	6
SPX2001	2001	4662	449	1347	1265	6.40E-246	[GI:1815634] [LN:SAU61271] [AC:U61271] [PN:glutamine synthetase type 1] [GN:glnA] [OR:Streptococcus agalactiae]	111
SPX2002	2002	4663	119	357	244	7.30E-29	[LN:GLNR_BACCE] [AC:P19083] [GN:GLNR] [OR:Bacillus cereus] [DE:REGULATORY PROTEIN GLNR] [SP:P19083]	99
SPX2003	2003	4664	176	528	100	1.00E-06	[LN:H69815] [AC:H69815] [PN:hypothetical protein ygaE] [GN:ygaE] [OR:Bacillus subtilis]	87
SPX2004	2004	4665	126	378	137	1.10E-11	[LN:S72776] [AC:S72776] [PN:B1496_F1_41 protein] [OR:Mycobacterium leprae]	74
SPX2005	2005	4666	399	1197	622	6.20E-166	[GI:4490614] [LN:SAU133520] [AC:AJ133520] [PN:phosphoglycerate kinase] [GN:pgk] [OR:Staphylococcus aureus]	106
SPX2006	2006	4667	1647	4941	207	3.10E-63	[GI:4204919] [LN:APU59168] [AC:U59168] [PN:endo-beta-N-acetylglucosaminidase] [OR:Arthrobacter protophormiae]	109
SPX2007	2007	4668	171	513			NO-HIT	6
SPX2008	2008	4669	544	1632	249	3.20E-57	“[LN:H71283] [AC:H71283] [PN:conserved hypothetical integral membrane protein TP0771] [GN:TP0771] [OR:Treponema pallidum subsp. pallidum] [SR:, syphilis spirochete]”	165
SPX2009	2009	4670	536	1608	2303	0	[LN:LAJ10153] [AC:AJ010153] [PN:CTP synthetase] [GN:pyrG] [OR:Lactococcus lactis subsp. cremoris]	97
SPX2010	2010	4671	44	132	137	6.60E-15	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2011	2011	4672	196	588	169	1.20E-25	“[LN:RPOE_BACSU]	135

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:P12464] [GN:RPOE] [OR:Bacillus subtilis] [EC:2.7.7.6] [DE:DNA-DIRECTED RNA POLYMERASE DELTA SUBUNIT,] [SP:P12464]"	
SPX2012	2012	4673	150	450			NO-HIT	6
SPX2013	2013	4674	130	390			NO-HIT	6
SPX2014	2014	4675	217	651	179	5.00E-19	[LN:S32217] [AC:S32217] [PN:hypothetical protein 2] [OR:Bacillus megaterium]	76
SPX2015	2015	4676	188	564	258	3.20E-32	[LN:YPAA_BACSU] [AC:P50726] [GN:YPAA] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 20.5 KD PROTEIN IN SERA-FER INTERGENIC REGION] [SP:P50726]	136
SPX2016	2016	4677	73	219			NO-HIT	6
SPX2017	2017	4678	156	468	243	6.60E-44	[GI:1381681] [LN:BSU58864] [AC:U58864] [PN:CspR] [GN:cspR] [OR:Bacillus subtilis] [SR:Bacillus subtilis strain=JH642]	117
SPX2018	2018	4679	277	831	83	4.70E-12	[LN:Y181_MYCPN] [AC:Q50292] [OR:Mycoplasma pneumoniae] [DE:HYPOTHETICAL PROTEIN MG181 HOMOLOG (GT9_ORF434)] [SP:Q50292]	119
SPX2019	2019	4680	416	1248	343	1.40E-50	[LN:G64435] [AC:G64435] [PN:cobalt transport ATP-binding protein O homolog] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Methanococcus jannaschii]	181
SPX2020	2020	4681	128	384	230	1.10E-25	[LN:G71192] [AC:G71192] [PN:probable cobalt transport ATP-binding protein] [GN:PH1815] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Pyrococcus horikoshii]	189
SPX2021	2021	4682	71	213			NO-HIT	6
SPX2022	2022	4683	86	258			NO-HIT	6
SPX2023	2023	4684	183	549	494	1.60E-73	[GI:6165407] [LN:LLA012388] [AC:AJ012388] [PN:hypothetical protein] [OR:Lactococcus lactis]	91
SPX2024	2024	4685	292	876	546	2.00E-98	[LN:E64608]	91

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:E64608] [PN:conserved hypothetical protein HP0709] [OR:Helicobacter pylori]	
SPX2025	2025	4686	450	1350	258	1.20E-60	[LN:F69354] [AC:F69354] [PN:TRK potassium uptake system protein (trkA-2) homolog] [CL:Methanococcus jannaschii TRK system potassium uptake protein A] [OR:Archaeoglobus fulgidus]	177
SPX2026	2026	4687	479	1437	232	7.20E-73	[LN:G69354] [AC:G69354] [PN:TRK potassium uptake system protein (trkH) homolog] [CL:potassium uptake protein trkG] [OR:Archaeoglobus fulgidus]	142
SPX2027	2027	4688	560	1680	777	1.20E-171	[LN:PTLB_STRMU] [AC:P50976] [GN:LACE] [OR:Streptococcus mutans] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EII-LAC)] [SP:P50976]	118
SPX2028	2028	4689	471	1413	643	5.80E-164	“[GI:153755] [LN:STRPBGSL] [AC:M19454] [OR:Lactococcus lactis subsp. cremoris] [SR:S.lactis (strain Z268) DNA, clone X25]”	122
SPX2029	2029	4690	65	195			NO-HIT	6
SPX2030	2030	4691	115	345	298	8.20E-37	[LN:PTLA_LACCA] [AC:P11502] [GN:LACF] [OR:Lactobacillus casei] [EC:2.7.1.69] [DE:(EC 2.7.1.69) (EIII-LAC)] [SP:P11502]	118
SPX2031	2031	4692	647	1941			NO-HIT	6
SPX2032	2032	4693	441	1323	395	1.00E-89	“[LN:PTCC_BACST] [AC:Q45400] [GN:CELB] [OR:Bacillus stearothermophilus] [DE:PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT)] [SP:Q45400]”	157
SPX2033	2033	4694	408	1224	147	2.50E-21	[LN:XYLR_BACSU] [AC:P16557] [GN:XYLR] [OR:Bacillus subtilis] [DE:XYLOSE REPRESSOR] [SP:P16557]	94
SPX2034	2034	4695	73	219			NO-HIT	6
SPX2035	2035	4696	58	174			NO-HIT	6
SPX2036	2036	4697	62	186			NO-HIT	6
SPX2037	2037	4698	95	285	167	2.40E-18	[GI:2707293]	77

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:AF036720] [AC:AF036720] [PN:unknown] [OR:Lactococcus lactis]	
SPX2038	2038	4699	173	519	879	6.50E-119	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX2039	2039	4700	256	768	1235	1.10E-172	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX2040	2040	4701	421	1263	948	4.80E-188	[GI:2687821] [LN:STIS1193] [AC:Y13713] [PN:transposase] [OR:Streptococcus thermophilus]	87
SPX2041	2041	4702	86	258	149	1.10E-14	[GI:2198546] [LN:SPCPS14E] [AC:X85787] [GN:tasA] [OR:Streptococcus pneumoniae]	78
SPX2042	2042	4703	775	2325	3774	0	[GI:3168596] [LN:AB014686] [AC:AB014686] [PN:pyruvate formate-lyase] [GN:pfl] [OR:Streptococcus bovis] [SR:Streptococcus bovis (strain:JB-1) DNA]	145
SPX2043	2043	4704	92	276	133	7.70E-20	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2044	2044	4705	78	234			NO-HIT	6
SPX2045	2045	4706	357	1071	491	2.70E-83	[GI:7380303] [LN:NMA5Z2491] [AC:AL162756:AL157959] [PN:impB/mucB/samB family protein] [GN:NMA1661] [OR:Neisseria meningitidis]	126
SPX2046	2046	4707	63	189			NO-HIT	6
SPX2047	2047	4708	282	846	289	6.60E-52	[GI:5739401] [LN:AF169967] [AC:AF169967] [PN:BacA] [GN:bacA]	91

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Flavobacterium johnsoniae]	
SPX2048	2048	4709	217	651	139	3.00E-12	[LN:YIS1_STRCO] [AC:P19780] [GN:SC3C8.10] [OR:Streptomyces coelicolor] [DE:INSERTION ELEMENT IS110 HYPOTHETICAL 43.6 KD PROTEIN] [SP:P19780]	140
SPX2049	2049	4710	62	186			NO-HIT	6
SPX2050	2050	4711	631	1893	95	3.20E-08	[LN:H64496] [AC:H64496] [PN:hypothetical protein MJ1577] [OR:Methanococcus jannaschii]	86
SPX2051	2051	4712	522	1566	331	9.10E-53	“[LN:A72357] [AC:A72357] [PN:amino acid ABC transporter, permease protein] [GN:TM0592] [CL:histidine permease protein M] [OR:Thermotoga maritima]”	146
SPX2052	2052	4713	247	741	388	9.60E-89	[LN:F81363] [AC:F81363] [PN:probable glutamine transport ATP-binding protein Cj0902 [imported]] [GN:glnQ:Cj0902] [OR:Campylobacter jejuni]	138
SPX2053	2053	4714	72	216			NO-HIT	6
SPX2054	2054	4715	77	231			NO-HIT	6
SPX2055	2055	4716	66	198	254	1.20E-29	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX2056	2056	4717	128	384	511	2.70E-66	[GI:5739312] [LN:AF161700] [AC:AF161700] [PN:ComX1] [GN:comX1] [FN:transcriptional regulator of] [OR:Streptococcus pneumoniae]	126
SPX2057	2057	4718	184	552	442	1.30E-63	[LN:NUSG_BACSU] [AC:Q06795] [GN:NUSG] [OR:Bacillus subtilis] [DE:TRANSCRIPTION ANTITERMINATION PROTEIN NUSG] [SP:Q06795]	120
SPX2058	2058	4719	59	177			NO-HIT	6
SPX2059	2059	4720	732	2196	3714	0	[GI:6165960] [LN:AF101780] [AC:AF101780] [PN:penicillin-binding protein 2a] [GN:pbp2a] [OR:Streptococcus pneumoniae]	116

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2060	2060	4721	292	876	229	4.10E-34	[LN:YHCT__BACSU] [AC:P54604] [GN:YHCT] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 33.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION] [SP:P54604]	137
SPX2061	2061	4722	360	1080	1061	2.30E-215	"[LN:A42963] [AC:A42963:B42963:JH0750] [PN:glyceraldehyde-3-phosphate dehydrogenase,;plasmin receptor] [CL:glyceraldehyde-3-phosphate dehydrogenase] [OR:Streptococcus sp.] [EC:1.2.1.12]"	186
SPX2062	2062	4723	92	276			NO-HIT	6
SPX2063	2063	4724	317	951	124	1.80E-15	[LN:Y797__METJA] [AC:Q58207] [GN:MJ0797] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ0797] [SP:Q58207]	114
SPX2064	2064	4725	113	339	552	2.20E-73	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX2065	2065	4726	87	261			NO-HIT	6
SPX2066	2066	4727	116	348	577	5.80E-76	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96
SPX2067	2067	4728	300	900	482	2.20E-71	[LN:B70375] [AC:B70375] [PN:quinolinate phosphoribosyl transferase] [GN:nadC] [CL:nicotinate-nucleotide pyrophosphorylase (carboxylating)] [OR:Aquifex aeolicus]	160
SPX2068	2068	4729	437	1311	448	2.70E-105	[LN:C81402] [AC:C81402] [PN:probable integral membrane protein Cj0555 [imported]] [GN:Cj0555] [OR:Campylobacter jejuni]	119
SPX2069	2069	4730	62	186			NO-HIT	6
SPX2070	2070	4731	65	195	111	7.10E-09	"[LN:H81018] [AC:H81018] [PN:iron(III) ABC transporter, ATP-binding protein NMB1993 [imported]] [GN:NMB1993] [OR:Neisseria meningitidis]"	137
SPX2071	2071	4732	243	729	352	3.20E-72	"[GI:4512387] [LN:AB011838] [AC:AB011838]	143

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:ydhQ]	
							[OR:Bacillus halodurans]	
							[SR:Bacillus halodurans (strain:C-125) DNA, clone_lib:lambda no.]”	
SPX2072	2072	4733	470	1410	808	8.30E-181	[LN:D69785]	127
							[AC:D69785]	
							[PN:beta-glucosidase homolog ydhP]	
							[GN:ydhP]	
							[CL:Agrobacterium beta-glucosidase]	
							[OR:Bacillus subtilis]	
SPX2073	2073	4734	432	1296	330	7.00E-64	“[LN:PTCC_BACSU]	152
							[AC:P46317]	
							[GN:CELB:LICC]	
							[OR:Bacillus subtilis]	
							[DE:PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT)]	
							[SP:P46317]”	
SPX2074	2074	4735	129	387			NO-HIT	6
SPX2075	2075	4736	103	309	191	2.70E-30	[LN:PTCB_BACST]	115
							[AC:Q45399]	
							[GN:CELA]	
							[OR:Bacillus stearothermophilus]	
							[EC:2.7.1.69]	
							[DE:(EC 2.7.1.69)]	
							[SP:Q45399]	
SPX2076	2076	4737	103	309	210	6.00E-24	[LN:PTCA_BACSU]	121
							[AC:P46319]	
							[GN:CELC:LICA]	
							[OR:Bacillus subtilis]	
							[EC:2.7.1.69]	
							[DE:(EC 2.7.1.69) (EIII-CEL)]	
							[SP:P46319]	
SPX2077	2077	4738	75	225			NO-HIT	6
SPX2078	2078	4739	891	2673	2272	0	“[LN:ADH2_ENTHI]	176
							[AC:Q24803:Q27649]	
							[GN:ADH2]	
							[OR:Entamoeba histolytica]	
							[EC:1.1.1.1:1.2.1.10]	
							[DE:(EC 1.1.1.1) (ADH); ALCETALDEHYDE DEHYDROGENASE, (ACDH)]]	
							[SP:Q24803:Q27649]”	
SPX2079	2079	4740	137	411	168	4.10E-18	[GI:6010051]	113
							[LN:ECA270205]	
							[AC:AJ270205]	
							[PN:putative phosphatidylinositol-4-phosphate]	
							[OR:Entodinium caudatum]	
SPX2080	2080	4741	143	429	91	7.10E-26	“““[LN:E69808]	161
							[AC:E69808]	
							[PN:protein-tyrosine phosphatase homolog yfkJ]	
							[GN:yfkJ]	
							[CL:protein-tyrosine-phosphatase, low molecular weight]	
							[OR:Bacillus subtilis]”	
SPX2081	2081	4742	100	300	177	9.30E-20	[GI:1402532]	130

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:D78257] [AC:D78257] [PN:ORF11] [GN:orf11] [OR:Enterococcus faecalis] [SR:Enterococcus faecalis plasmid:pYI17 DNA]	
SPX2082	2082	4743	659	1977	1671	0	“[LN:TKT_STRPN] [AC:P22976] [GN:RECP] [OR:Streptococcus pneumoniae] [EC:2.2.1.1] [DE:PROBABLE TRANSKETOLASE, (TK)] [SP:P22976]”	127
SPX2083	2083	4744	81	243	171	1.00E-17	“[LN:TKT_STRPN] [AC:P22976] [GN:RECP] [OR:Streptococcus pneumoniae] [EC:2.2.1.1] [DE:PROBABLE TRANSKETOLASE, (TK)] [SP:P22976]”	127
SPX2084	2084	4745	102	306	277	1.50E-33	“[LN:TKT_STRPN] [AC:P22976] [GN:RECP] [OR:Streptococcus pneumoniae] [EC:2.2.1.1] [DE:PROBABLE TRANSKETOLASE, (TK)] [SP:P22976]”	127
SPX2085	2085	4746	245	735	546	8.30E-96	[LN:YJFR_ECOLI] [AC:P39300] [GN:YJFR] [OR:Escherichia coli] [DE:HYPOTHETICAL 40.1 KD PROTEIN IN AIDB-SGAT INTERGENIC REGION] [SP:P39300]	136
SPX2086	2086	4747	136	408	430	3.30E-54	[LN:YJFR_ECOLI] [AC:P39300] [GN:YJFR] [OR:Escherichia coli] [DE:HYPOTHETICAL 40.1 KD PROTEIN IN AIDB-SGAT INTERGENIC REGION] [SP:P39300]	136
SPX2087	2087	4748	558	1674	178	6.30E-20	“[GI:4512373] [LN:AB011837] [AC:AB011837] [GN:yjdC] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA, clone_lib:lambda no.9]”	144
SPX2088	2088	4749	235	705	489	3.20E-93	[GI:5616307] [LN:AF160811] [AC:AF160811] [PN:L-ribulose 5-phosphate 4-epimerase] [GN:araD]	123

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus stearothermophilus]	
SPX2089	2089	4750	288	864	777	3.20E-106	“[LN:SGBU_HAEIN] [AC:P44990] [GN:SGBU:HI1026] [OR:Haemophilus influenzae] [EC:5.-.-.-] [DE:PUTATIVE HEXULOSE-6-PHOSPHATE ISOMERASE, (HUMPI)] [SP:P44990]”	153
SPX2090	2090	4751	222	666	349	1.00E-70	[LN:SGAH_ECOLI] [AC:P39304] [GN:SGAH] [OR:Escherichia coli] [EC:4.1.2.-] [DE:3-HEXULOSE 6-PHOSPHATE FORMALDEHYDE LYASE)] [SP:P39304]	132
SPX2091	2091	4752	162	486	266	1.10E-45	[LN:PTXA_ECOLI] [AC:P39303] [GN:SGAA] [OR:Escherichia coli] [EC:2.7.1.69] [DE:(EC 2.7.1.69)] [SP:P39303]	104
SPX2092	2092	4753	94	282	111	6.80E-22	[LN:PTXB_ECOLI] [AC:P39302] [GN:SGAB] [OR:Escherichia coli] [EC:2.7.1.69] [DE:(EC 2.7.1.69)] [SP:P39302]	104
SPX2093	2093	4754	509	1527	275	5.90E-99	[LN:SGAT_ECOLI] [AC:P39301] [GN:SGAT] [OR:Escherichia coli] [DE:PUTATIVE TRANSPORT PROTEIN SGAT] [SP:P39301]	108
SPX2094	2094	4755	208	624	210	1.10E-38	“[GI:6681651] [LN:AB016077] [AC:AB016077] [PN:sakacin A production response regulator] [GN:sapR] [OR:Streptococcus mutans] [SR:Streptococcus mutans (strain:MT8148) DNA, clone:pYT570]”	183
SPX2095	2095	4756	329	987	145	1.90E-24	[LN:JAG_BACSU] [AC:Q01620] [GN:JAG] [OR:Bacillus subtilis] [DE:JAG PROTEIN (SPOIIIJ ASSOCIATED PROTEIN)] [SP:Q01620]	116
SPX2096	2096	4757	277	831	357	1.00E-51	[LN:SP3J_BACSU] [AC:Q01625]	122

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:SPOIIIJ] [OR:Bacillus subtilis] [DE:STAGE III SPORULATION PROTEIN J PRECURSOR] [SP:Q01625]	
SPX2097	2097	4758	124	372	173	3.10E-31	“[GI:5672645] [LN:AB013492] [AC:AB013492] [GN:rnpA] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA, clone:ALBAC001]”	137
SPX2098	2098	4759	397	1191	822	1.30E-149	“[LN:ACKA_BACSU] [AC:P37877] [GN:ACKA] [OR:Bacillus subtilis] [EC:2.7.2.1] [DE:ACETATE KINASE, (ACETOKINASE)] [SP:P37877]”	122
SPX2099	2099	4760	318	954	297	4.70E-44	[LN:YTXK_BACSU] [AC:P37876] [GN:YTXK] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 37.4 KD PROTEIN IN ACKA-SSPA INTERGENIC REGION] [SP:P37876]	137
SPX2100	2100	4761	196	588	132	2.90E-10	[LN:H70323] [AC:H70323] [PN:hypothetical protein aq_262] [GN:aq_262] [OR:Aquifex aeolicus]	90
SPX2101	2101	4762	138	414			NO-HIT	6
SPX2102	2102	4763	169	507	157	1.60E-30	[GI:3287181] [LN:LLC3107] [AC:Y15043] [GN:orf150] [OR:Lactococcus lactis subsp. cremoris]	89
SPX2103	2103	4764	101	303	123	9.50E-12	[GI:3287182] [LN:LLC3107] [AC:Y15043] [PN:hypothetical protein] [GN:orf128] [OR:Lactococcus lactis subsp. cremoris]	115
SPX2104	2104	4765	135	405	626	3.00E-81	[GI:3211751] [LN:AF052207] [AC:AF052207] [PN:competence protein] [GN:cglD] [OR:Streptococcus pneumoniae]	104
SPX2105	2105	4766	109	327	512	5.80E-66	[GI:3211750] [LN:AF052207] [AC:AF052207] [PN:competence protein]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2106	2106	4767	348	1044	1440	3.30E-199	[GN:cglC] [OR:Streptococcus pneumoniae] [GI:3211749] [LN:AF052207] [AC:AF052207] [PN:competence protein]	104
SPX2107	2107	4768	314	942	1592	4.30E-217	[GN:cglB] [OR:Streptococcus pneumoniae] [GI:3211748] [LN:AF052207] [AC:AF052207] [PN:competence protein]	104
SPX2108	2108	4769	124	372	468	9.20E-61	[GN:cglA] [OR:Streptococcus pneumoniae] [GI:2058543] [LN:SGU81957] [AC:U81957]	100
SPX2109	2109	4770	353	1059	639	1.60E-125	[PN:putative DNA binding protein] [OR:Streptococcus gordonii] [LN:T36961] [AC:T36961] [PN:probable zinc-containing dehydrogenase] [GN:SCJ1.28c] [CL:alcohol dehydrogenase:long-chain alcohol dehydrogenase homology] [OR:Streptomyces coelicolor]	179
SPX2110	2110	4771	142	426			NO-HIT	6
SPX2111	2111	4772	384	1152	819	3.90E-108	[GI:6683552] [LN:AB024532] [AC:AB024532] [GN:SA8A11-1] [OR:Enterococcus seriolicida] [SR:Enterococcus seriolicida DNA]	118
SPX2112	2112	4773	606	1818	173	2.40E-27	[LN:C69975] [AC:C69975] [PN:acyltransferase homolog yrhL] [GN:yrhL] [OR:Bacillus subtilis]	90
SPX2113	2113	4774	381	1143	1469	3.70E-200	[LN:TGT_BACSU] [AC:O32053] [GN:TGT] [OR:Bacillus subtilis] [EC:2.4.2.29] [DE:TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME)] [SP:O32053]	134
SPX2114	2114	4775	286	858	90	9.00E-05	[LN:S32215] [AC:S32215] [PN:hypothetical protein 1] [OR:Bacillus megaterium]	76
SPX2115	2115	4776	215	645	769	1.20E-104	[LN:PCP_STRPY] [AC:Q01328]	149

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:PCP] [OR:Streptococcus pyogenes] [EC:3.4.19.3] [DE:PEPTIDASE) (PYROGLUTAMYL-PEPTIDASE I) (PGP-I) (PYRASE)] [SP:Q01328]	
SPX2116	2116	4777	90	270			NO-HIT	6
SPX2117	2117	4778	120	360	221	1.90E-43	“[LN:YTRP_LACLA] [AC:Q02009] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:HYPOTHETICAL 13.3 KD PROTEIN IN TRPE 5'REGION] [SP:Q02009]”	155
SPX2118	2118	4779	142	426	163	3.50E-17	[LN:D69783] [AC:D69783] [PN:transcription regulator MarR family homolog ydgJ] [GN:ydgJ] [CL:regulatory protein mprA] [OR:Bacillus subtilis]	139
SPX2119	2119	4780	69	207	107	2.10E-10	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX2120	2120	4781	371	1113	382	8.00E-79	[GI:517204] [LN:SPU09352] [AC:U09352] [OR:Streptococcus pyogenes]	65
SPX2121	2121	4782	207	621	153	7.60E-14	[LN:A57362] [AC:A57362] [PN:gyrb protein] [GN:gyrb] [OR:Streptococcus pneumoniae]	81
SPX2122	2122	4783	427	1281	216	1.10E-45	[LN:H72331] [AC:H72331] [PN:conserved hypothetical protein] [GN:TM0815] [CL:conserved hypothetical protein HI1612] [OR:Thermotoga maritima]	139
SPX2123	2123	4784	495	1485	412	4.80E-66	[LN:B64532] [AC:B64532] [PN:threonine synthase] [CL:threonine synthase] [OR:Helicobacter pylori]	96
SPX2124	2124	4785	426	1278	70	0.00014	“[LN:T05142] [AC:T05142:S60128:S45018] [PN:gamma-glutamylcysteine synthetase:protein F7H19.290] [OR:Arabidopsis thaliana] [SR:, mouse-ear cress] [EC:6.3.2.2]”	158
SPX2125	2125	4786	59	177	100	1.30E-08	[LN:C64671]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:C64671] [PN:hypothetical protein HP1211] [OR:Helicobacter pylori]	
SPX2126	2126	4787	61	183	81	9.30E-05	[LN:E72688] [AC:E72688] [PN:hypothetical protein APE0925] [GN:APE0925] [CL:Aeropyrum pernix hypothetical protein APE0925] [OR:Aeropyrum pernix]	143
SPX2127	2127	4788	78	234	143	4.90E-15	[LN:F71245] [AC:F71245] [PN:hypothetical protein PHS004] [GN:PHS004] [OR:Pyrococcus horikoshii]	95
SPX2128	2128	4789	82	246	247	1.10E-29	[LN:F81737] [AC:F81737] [PN:hypothetical protein TC0129 [imported]] [GN:TC0129] [OR:Chlamydia muridarum:Chlamydia trachomatis MoPn]	131
SPX2129	2129	4790	52	156	80	3.00E-05	[LN:D75542] [AC:D75542] [PN:hypothetical protein] [GN:DR0254] [OR:Deinococcus radiodurans]	90
SPX2130	2130	4791	105	315			NO-HIT	6
SPX2131	2131	4792	146	438	486	1.80E-62	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX2132	2132	4793	354	1062	824	6.40E-132	[GI:1915907] [LN:LLPEPPGEN] [AC:Y08842] [PN:aminopeptidase P] [GN:pepP] [OR:Lactococcus lactis]	95
SPX2133	2133	4794	944	2832	2029	0	[LN:UVRA_BACSU] [AC:O34863] [GN:UVRA] [OR:Bacillus subtilis] [DE:EXCINUCLEASE ABC SUBUNIT A] [SP:O34863]	104
SPX2134	2134	4795	315	945	236	1.50E-26	[LN:A75272] [AC:A75272] [PN:probable transport protein] [GN:DR2463] [CL:magnesium and cobalt transport protein] [OR:Deinococcus radiodurans]	140
SPX2135	2135	4796	68	204			NO-HIT	6
SPX2136	2136	4797	213	639			NO-HIT	6
SPX2137	2137	4798	51	153	123	1.60E-12	[GI:1914870]	81

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	
SPX2138	2138	4799	344	1032	86	1.10E-19	[LN:E75325] [AC:E75325] [PN:probable mccF protein] [GN:DR2000] [OR:Deinococcus radiodurans]	91
SPX2139	2139	4800	226	678	66	3.80E-08	[GI:6470197] [LN:AF188935] [AC:AF188935] [PN:pXO2-46] [OR:Bacillus anthracis]	77
SPX2140	2140	4801	188	564	465	3.30E-63	[LN:3MGA_HAEIN] [AC:P44321] [GN:TAG:HI0654] [OR:Haemophilus influenzae] [EC:3.2.2.20] [DE:GLYCOSIDASE) (TAG)] [SP:P44321]	121
SPX2141	2141	4802	186	558			NO-HIT	6
SPX2142	2142	4803	198	594	235	1.00E-38	[LN:RUVA_BACSU] [AC:O05392] [GN:RUVA] [OR:Bacillus subtilis] [DE:PROBABLE HOLLIDAY JUNCTION DNA HELICASE RUVA] [SP:O05392]	122
SPX2143	2143	4804	70	210			NO-HIT	6
SPX2144	2144	4805	367	1101	888	6.00E-128	"[LN:RIBD_ACTPL] [AC:P50853] [GN:RIBD:RIBG] [OR:Actinobacillus pleuropneumoniae] [SR:,Haemophilus pleuropneumoniae] [EC:3.5.4.-] [DE:RIBOFLAVIN-SPECIFIC DEAMINASE,] [SP:P50853]"	177
SPX2145	2145	4806	212	636	601	1.70E-81	"[LN:RISA_ACTPL] [AC:P50854] [GN:RIBE:RIBB] [OR:Actinobacillus pleuropneumoniae] [SR:,Haemophilus pleuropneumoniae] [EC:2.5.1.9] [DE:RIBOFLAVIN SYNTHASE ALPHA CHAIN,] [SP:P50854]"	179
SPX2146	2146	4807	412	1236	1266	1.80E-170	"[LN:GCH2_ACTPL] [AC:P50855] [GN:RIBA] [OR:Actinobacillus pleuropneumoniae] [SR:,Haemophilus pleuropneumoniae]	153

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[EC:3.5.4.25] [DE:SYNTHASE)]] [SP:P50855]”	
SPX2147	2147	4808	156	468	546	1.30E-70	“[LN:RISB_ACTPL] [AC:P50856] [GN:RIBH] [OR:Actinobacillus pleuropneumoniae] [SR:,Haemophilus pleuropneumoniae] [EC:2.5.1.9] [DE:(LUMAZINE SYNTHASE) (RIBOFLAVIN SYNTHASE BETA CHAIN)] [SP:P50856]”	194
SPX2148	2148	4809	650	1950	3286	0	[LN:HEXB_STRPN] [AC:P14160] [GN:HEXB] [OR:Streptococcus pneumoniae] [DE:DNA MISMATCH REPAIR PROTEIN HEXB] [SP:P14160]	117
SPX2149	2149	4810	84	252			NO-HIT	6
SPX2150	2150	4811	97	291			NO-HIT	6
SPX2151	2151	4812	61	183	122	1.10E-12	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2152	2152	4813	174	522	231	9.30E-27	“[GI:1196936] [LN:STRHEXB] [AC:M29686] [PN:unknown protein] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (clone: pSP(8,41).) DNA]”	144
SPX2153	2153	4814	153	459			NO-HIT	6
SPX2154	2154	4815	126	378	333	1.60E-40	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX2155	2155	4816	211	633	762	7.90E-101	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX2156	2156	4817	201	603	499	7.70E-65	[LN:C81084] [AC:C81084] [PN:hypothetical protein NMB1426 [imported]] [GN:NMB1426] [OR:Neisseria meningitidis]	109
SPX2157	2157	4818	64	192			NO-HIT	6
SPX2158	2158	4819	218	654	461	2.50E-67	[LN:C81084] [AC:C81084] [PN:hypothetical protein NMB1426 [imported]] [GN:NMB1426] [OR:Neisseria meningitidis]	109
SPX2159	2159	4820	201	603	353	3.50E-53	[GI:6782392]	125

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:SDY18363] [AC:Y18363] [PN:ribonucleotide reductase-like (Nrd-like)] [GN:nrd] [OR:Streptococcus dysgalactiae]	
SPX2160	2160	4821	194	582			NO-HIT	6
SPX2161	2161	4822	429	1287	2176	6.20E-298	[GI:5830536] [LN:SPAJ6396] [AC:AJ006396] [PN:response regulator] [GN:rr07] [OR:Streptococcus pneumoniae]	104
SPX2162	2162	4823	568	1704	1183	7.00E-160	[GI:5830535] [LN:SPAJ6396] [AC:AJ006396] [PN:histidine kinase] [GN:hk07] [OR:Streptococcus pneumoniae]	102
SPX2163	2163	4824	205	615			NO-HIT	6
SPX2164	2164	4825	190	570			NO-HIT	6
SPX2165	2165	4826	231	693	395	1.90E-87	[GI:6165406] [LN:LLA012388] [AC:AJ012388] [PN:hypothetical protein] [OR:Lactococcus lactis]	91
SPX2166	2166	4827	354	1062	616	4.50E-130	[GI:6165405] [LN:LLA012388] [AC:AJ012388] [PN:hypothetical protein] [OR:Lactococcus lactis]	91
SPX2167	2167	4828	73	219	94	8.50E-07	[LN:G72536] [AC:G72536] [PN:hypothetical protein APE1580] [GN:APE1580] [OR:Aeropyrum pernix]	92
SPX2168	2168	4829	458	1374	219	7.10E-53	[LN:E75327] [AC:E75327] [PN:ArgE/DapE/Acy1 family protein] [GN:DR2017] [OR:Deinococcus radiodurans]	99
SPX2169	2169	4830	285	855	319	1.50E-63	[GI:6165404] [LN:LLA012388] [AC:AJ012388] [PN:hypothetical protein] [OR:Lactococcus lactis]	91
SPX2170	2170	4831	66	198			NO-HIT	6
SPX2171	2171	4832	184	552	109	1.10E-12	[LN:YXEM_BACSU] [AC:P54952] [GN:YXEM:LP9E] [OR:Bacillus subtilis]	110

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[DE:INTERGENIC REGION PRECURSOR] [SP:P54952]	
SPX2172	2172	4833	62	186			NO-HIT	6
SPX2173	2173	4834	225	675	129	5.60E-10	[LN:G75494] [AC:G75494] [PN:AzlC family protein] [GN:DR0633] [CL:hypothetical protein b2682] [OR:Deinococcus radiodurans]	121
SPX2174	2174	4835	137	411	328	2.80E-41	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2175	2175	4836	80	240			NO-HIT	6
SPX2176	2176	4837	75	225			NO-HIT	6
SPX2177	2177	4838	410	1230	80	2.70E-10	[GI:3582221] [LN:AE001272] [AC:AE001272] [PN:conserved hypothetical protein] [GN:ORF00049] [OR:Lactococcus lactis]	114
SPX2178	2178	4839	222	666	71	1.80E-05	[LN:E69787] [AC:E69787] [PN:hypothetical protein ydiL] [GN:ydiL] [OR:Bacillus subtilis]	87
SPX2179	2179	4840	66	198			NO-HIT	6
SPX2180	2180	4841	296	888	587	4.20E-77	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX2181	2181	4842	69	207			NO-HIT	6
SPX2182	2182	4843	206	618	479	1.40E-72	[LN:UDG5_ECOLI] [AC:Q47329] [GN:KFID] [OR:Escherichia coli] [EC:1.1.1.22] [DE:(UDP-GLCDH) (UDPGDH)] [SP:Q47329]	111
SPX2183	2183	4844	252	756	86	3.90E-08	[GI:4689219] [LN:AF115779] [AC:AF115779] [PN:unknown] [GN:mitC] [OR:Streptomyces lavendulae]	92
SPX2184	2184	4845	74	222			NO-HIT	6
SPX2185	2185	4846	115	345			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2186	2186	4847	81	243			NO-HIT	6
SPX2187	2187	4848	539	1617	294	1.70E-34	"[LN:D72267] [AC:D72267] [PN:ABC transporter, ATP-binding protein] [GN:TM1328] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Thermotoga maritima]"	180
SPX2188	2188	4849	318	954	137	5.90E-20	[LN:B75096] [AC:B75096] [PN:glycosyl transferase PAB0772] [GN:PAB0772] [CL:Neisseria meningitidis glycosyl transferase A] [OR:Pyrococcus abyssi]	144
SPX2189	2189	4850	175	525	179	6.60E-17	[GI:1276880] [LN:STU40830] [AC:U40830] [PN:EpsG] [GN:epsG] [OR:Streptococcus thermophilus]	90
SPX2190	2190	4851	188	564			NO-HIT	6
SPX2191	2191	4852	107	321	187	3.70E-21	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96
SPX2192	2192	4853	88	264	196	2.60E-23	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX2193	2193	4854	355	1065	893	8.00E-121	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX2194	2194	4855	129	387	490	2.90E-65	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX2195	2195	4856	71	213	119	2.60E-10	[GI:2198546] [LN:SPCPS14E] [AC:X85787] [GN:tasA] [OR:Streptococcus pneumoniae]	78
SPX2196	2196	4857	337	1011	913	6.50E-132	[LN:YDIE_BACSU] [AC:O05518] [GN:YDIE] [OR:Bacillus subtilis]	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2197	2197	4858	146	438	92	1.00E-09	[DE:HYPOTHETICAL 36.8 KD PROTEIN IN PHOB-GROES INTERGENIC REGION] [SP:O05518] [LN:E69786] [AC:E69786] [PN:ribosomal-protein-alanine N-acetyltransfer homolog ydiD] [GN:ydiD] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimI] [OR:Bacillus subtilis]	190
SPX2198	2198	4859	228	684	303	1.40E-53	[GI:3341437] [LN:EFY17797] [AC:Y17797] [PN:hypothetical protein] [GN:ydiC] [OR:Enterococcus faecalis]	101
SPX2199	2199	4860	62	186			NO-HIT	6
SPX2200	2200	4861	77	231			NO-HIT	6
SPX2201	2201	4862	78	234	78	2.50E-10	[LN:D69871] [AC:D69871] [PN:hypothetical protein ykzG] [GN:ykzG] [OR:Bacillus subtilis]	87
SPX2202	2202	4863	560	1680	1841	6.00E-251	[LN:YKQC__BACSU] [AC:Q45493] [GN:YKQC] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 61.5 KD PROTEIN IN ADEC-PDHA INTERGENIC REGION] [SP:Q45493]	137
SPX2203	2203	4864	121	363			NO-HIT	6
SPX2204	2204	4865	638	1914	1698	2.50E-282	“[LN:GIDA__LACLC] [AC:O32806] [GN:GIDA] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [DE:GLUCOSE INHIBITED DIVISION PROTEIN A (FRAGMENT)] [SP:O32806]”	171
SPX2205	2205	4866	158	474			NO-HIT	6
SPX2206	2206	4867	161	483	73	2.50E-06	[LN:F75354] [AC:F75354] [PN:MutT/nudix family protein] [GN:DR1776] [OR:Deinococcus radiodurans]	95
SPX2207	2207	4868	392	1176	578	1.10E-170	[LN:TRMU__BACSU] [AC:O35020] [GN:TRMU] [OR:Bacillus subtilis] [EC:2.1.1.61] [DE:(EC 2.1.1.61)] [SP:O35020]	105
SPX2208	2208	4869	60	180			NO-HIT	6
SPX2209	2209	4870	66	198			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2210	2210	4871	551	1653	1324	7.00E-210	[LN:A41971] [AC:A41971:A60282:A33134] [PN:surface protein pspA precursor:pneumococcal surface protein A] [GN:pspA] [CL:cpl repeat homology] [OR:Streptococcus pneumoniae]	169
SPX2211	2211	4872	79	237	98	1.60E-06	[GI:311109] [LN:YSCISCLRP] [AC:L16900] [PN:intrastrand crosslink recognition protein] [GN:Ixr1] [OR:Saccharomyces cerevisiae] [SR:Saccharomyces cerevisiae (strain) DNA]	169
SPX2212	2212	4873	64	192	77	0.00063	[LN:B72392] [AC:B72392] [PN:hypothetical protein] [GN:TM0315] [OR:Thermotoga maritima]	86
SPX2213	2213	4874	103	309	87	0.0009	[GI:7292943] [LN:AE003494] [AC:AE003494:AE002593] [GN:CG11075] [OR:Drosophila melanogaster] [SR:fruit fly]	106
SPX2214	2214	4875	106	318			NO-HIT	6
SPX2215	2215	4876	103	309	459	8.20E-56	[GI:6752385] [LN:AF071807] [AC:AF071807] [PN:PspA] [GN:pspA] [OR:Streptococcus pneumoniae]	90
SPX2216	2216	4877	93	279			NO-HIT	6
SPX2217	2217	4878	100	300			NO-HIT	6
SPX2218	2218	4879	71	213			NO-HIT	6
SPX2219	2219	4880	253	759			NO-HIT	6
SPX2220	2220	4881	129	387			NO-HIT	6
SPX2221	2221	4882	301	903			NO-HIT	6
SPX2222	2222	4883	61	183			NO-HIT	6
SPX2223	2223	4884	125	375			NO-HIT	6
SPX2224	2224	4885	62	186			NO-HIT	6
SPX2225	2225	4886	227	681			NO-HIT	6
SPX2226	2226	4887	76	228			NO-HIT	6
SPX2227	2227	4888	319	957	1739	3.00E-234	“[LN:ALYS_BPHB3] [AC:P32762] [GN:HBL] [OR:Streptococcus pneumoniae phage HB-3] [EC:3.5.1.28] [DE:LYTIC AMIDASE, (N-ACETYLMURAMOYL-L-ALANINE AMIDASE)] [SP:P32762]”	162
SPX2228	2228	4889	111	333			NO-HIT	6

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2229	2229	4890	139	417	97	1.20E-08	[LN:VG14_BPB03] [AC:Q37895] [GN:14] [OR:Bacteriophage B103] [DE:LYSIS PROTEIN (LATE PROTEIN GP14)] [SP:Q37895]	110
SPX2230	2230	4891	117	351			NO-HIT	6
SPX2231	2231	4892	68	204			NO-HIT	6
SPX2232	2232	4893	2120	6360	154	1.40E-22	[GI:4530154] [LN:AF085222] [AC:AF085222] [PN:putative tail-host specificity protein] [OR:Streptococcus thermophilus bacteriophage DT1]	134
SPX2233	2233	4894	79	237			NO-HIT	6
SPX2234	2234	4895	117	351			NO-HIT	6
SPX2235	2235	4896	1218	3654	236	2.10E-30	[GI:2392838] [LN:AF011378] [AC:AF011378] [PN:unknown] [OR:Bacteriophage sk1]	76
SPX2236	2236	4897	187	561	106	2.00E-07	[GI:5305335] [LN:AF071081] [AC:AF071081] [PN:proline-rich mucin homolog] [OR:Mycobacterium tuberculosis]	104
SPX2237	2237	4898	70	210			NO-HIT	6
SPX2238	2238	4899	73	219			NO-HIT	6
SPX2239	2239	4900	127	381			NO-HIT	6
SPX2240	2240	4901	138	414			NO-HIT	6
SPX2241	2241	4902	123	369	77	0.00049	[LN:T13522] [AC:T13522] [PN:hypothetical protein 33] [CL:Bacillus phage phi-105 hypothetical protein 33] [OR:Bacillus phage phi-105]	132
SPX2242	2242	4903	172	516	71	7.90E-06	[LN:T42287] [AC:T42287] [PN:hypothetical protein] [OR:phage SPP1]	65
SPX2243	2243	4904	61	183			NO-HIT	6
SPX2244	2244	4905	113	339			NO-HIT	6
SPX2245	2245	4906	104	312			NO-HIT	6
SPX2246	2246	4907	63	189			NO-HIT	6
SPX2247	2247	4908	61	183	73	0.0003	[LN:F75258] [AC:F75258] [PN:hypothetical protein] [GN:DR2560] [OR:Deinococcus radiodurans]	90
SPX2248	2248	4909	282	846	226	2.60E-48	[GI:1369939] [LN:BTP9011] [AC:X84706]	92

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[PN:major head protein] [GN:mhp] [OR:Bacteriophage B1]	
SPX2249	2249	4910	195	585			NO-HIT	6
SPX2250	2250	4911	86	258			NO-HIT	6
SPX2251	2251	4912	76	228			NO-HIT	6
SPX2252	2252	4913	468	1404	80	1.20E-06	[LN:T13317] [AC:T13317] [PN:hypothetical protein 28] [OR:Streptococcus phage phi-O1205]	87
SPX2253	2253	4914	490	1470	154	3.40E-39	[LN:T13620] [AC:T13620] [PN:hypothetical protein gp502] [OR:Streptococcus phage phi-Sfi11]	90
SPX2254	2254	4915	433	1299	330	6.80E-48	[LN:TERL__BPSPP] [AC:P54308] [GN:2] [OR:Bacteriophage SPP1] [DE:TERMINASE LARGE SUBUNIT (G2P)] [SP:P54308]	105
SPX2255	2255	4916	160	480	215	2.30E-32	[GI:4680607] [LN:AF125198] [AC:AF125198] [PN:terminase small subunit] [GN:TS1] [OR:bacteriophage phi-FC1]	105
SPX2256	2256	4917	135	405	154	4.60E-16	[GI:4530181] [LN:AF085222] [AC:AF085222] [PN:unknown] [OR:Streptococcus thermophilus bacteriophage DT1]	103
SPX2257	2257	4918	124	372			NO-HIT	6
SPX2258	2258	4919	146	438	115	1.00E-14	[GI:5001708] [LN:AF109874] [AC:AF109874] [PN:unknown] [OR:Bacteriophage Tuc2009]	80
SPX2259	2259	4920	153	459			NO-HIT	6
SPX2260	2260	4921	87	261	109	9.50E-12	[GI:4530179] [LN:AF085222] [AC:AF085222] [PN:unknown] [OR:Streptococcus thermophilus bacteriophage DT1]	103
SPX2261	2261	4922	232	696	203	1.70E-24	[LN:T13308] [AC:T13308] [PN:hypothetical protein 19] [OR:Streptococcus phage phi-O1205]	87
SPX2262	2262	4923	450	1350	358	1.30E-46	[GI:2689558] [LN:U93688] [AC:U93688]	63

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Staphylococcus aureus]	
SPX2263	2263	4924	274	822	470	7.10E-87	[LN:T13301] [AC:T13301] [PN:hypothetical protein 12] [OR:Streptococcus phage phi-O1205]	87
SPX2264	2264	4925	78	234			NO-HIT	6
SPX2265	2265	4926	66	198			NO-HIT	6
SPX2266	2266	4927	169	507			NO-HIT	6
SPX2267	2267	4928	276	828			NO-HIT	6
SPX2268	2268	4929	106	318			NO-HIT	6
SPX2269	2269	4930	398	1194			NO-HIT	6
SPX2270	2270	4931	163	489			NO-HIT	6
SPX2271	2271	4932	120	360			NO-HIT	6
SPX2272	2272	4933	161	483	256	2.10E-45	[GI:2352435] [LN:AF004379] [AC:AF004379] [OR:Streptococcus thermophilus bacteriophage Sfi21]	92
SPX2273	2273	4934	258	774			NO-HIT	6
SPX2274	2274	4935	77	231			NO-HIT	6
SPX2275	2275	4936	114	342			NO-HIT	6
SPX2276	2276	4937	63	189			NO-HIT	6
SPX2277	2277	4938	87	261			NO-HIT	6
SPX2278	2278	4939	91	273			NO-HIT	6
SPX2279	2279	4940	80	240			NO-HIT	6
SPX2280	2280	4941	64	192			NO-HIT	6
SPX2281	2281	4942	68	204			NO-HIT	6
SPX2282	2282	4943	67	201			NO-HIT	6
SPX2283	2283	4944	72	216	101	3.70E-09	[LN:Y272__METJA] [AC:Q57720] [GN:MJ0272] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL TRANSCRIPTIONAL REGULATOR MJ0272] [SP:Q57720]	132
SPX2284	2284	4945	68	204			NO-HIT	6
SPX2285	2285	4946	230	690			NO-HIT	6
SPX2286	2286	4947	264	792	112	4.00E-15	[LN:T13264] [AC:T13264] [PN:repressor protein] [OR:Lactococcus lactis phage BK5-T]	82
SPX2287	2287	4948	317	951	102	5.00E-09	[LN:F59095] [AC:F59095] [PN:hypothetical protein pXO1-38] [GN:pXO1-38] [OR:Bacillus anthracis]	94
SPX2288	2288	4949	58	174			NO-HIT	6
SPX2289	2289	4950	376	1128	212	3.50E-43	[GI:2689564] [LN:U93688] [AC:U93688] [PN:integrase] [GN:int]	87

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Staphylococcus aureus]	
SPX2290	2290	4951	62	186			NO-HIT	6
SPX2291	2291	4952	353	1059	81	0.00071	[LN:AF147045] [AC:AF147045] [PN:cytochrome c oxidase subunit 1] [GN:COI] [OR:Mitochondrion Dolichoderus lutosus] [SR:Dolichoderus lutosus]	138
SPX2292	2292	4953	64	192			NO-HIT	6
SPX2293	2293	4954	71	213			NO-HIT	6
SPX2294	2294	4955	81	243			NO-HIT	6
SPX2295	2295	4956	149	447			NO-HIT	6
SPX2296	2296	4957	60	180			NO-HIT	6
SPX2297	2297	4958	117	351	137	3.30E-14	[GI:4760910] [LN:AF099088] [AC:AF099088] [PN:EntT] [GN:entT] [OR:Enterococcus faecium]	86
SPX2298	2298	4959	288	864	117	4.70E-08	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX2299	2299	4960	125	375			NO-HIT	6
SPX2300	2300	4961	76	228			NO-HIT	6
SPX2301	2301	4962	127	381			NO-HIT	6
SPX2302	2302	4963	75	225			NO-HIT	6
SPX2303	2303	4964	70	210			NO-HIT	6
SPX2304	2304	4965	68	204			NO-HIT	6
SPX2305	2305	4966	288	864	117	4.70E-08	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX2306	2306	4967	173	519	840	6.50E-114	[GI:663278] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX2307	2307	4968	122	366			NO-HIT	6
SPX2308	2308	4969	96	288	279	1.60E-35	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase]	138

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	
SPX2309	2309	4970	256	768	1235	1.10E-172	[GI:663279] [LN:STRCOMAA] [AC:M36180:L15190] [PN:transposase] [OR:Streptococcus pneumoniae] [SR:Streptococcus pneumoniae (strain RX1) DNA]	138
SPX2310	2310	4971	157	471	375	6.90E-56	[LN:ASSY_BACSU] [AC:O34347] [GN:ARGG] [OR:Bacillus subtilis] [EC:6.3.4.5] [DE:LIGASE)] [SP:O34347]	98
SPX2311	2311	4972	266	798	200	6.10E-32	“[LN:B72357] [AC:B72357] [PN:amino acid ABC transporter, periplasmic amino acid-binding protein] [GN:TM0593] [CL:lysine-arginine-ornithine-binding protein] [OR:Thermotoga maritima]”	181
SPX2312	2312	4973	214	642	420	1.80E-51	[LN:F81408] [AC:F81408] [PN:ABC-type transmembrane transport protein Cj0607 [imported]] [GN:Cj0607] [OR:Campylobacter jejuni]	125
SPX2313	2313	4974	59	177	80	1.80E-07	[LN:F72598] [AC:F72598] [PN:hypothetical protein APE1254] [GN:APE1254] [OR:Aeropyrum pernix]	92
SPX2314	2314	4975	439	1317			NO-HIT	6
SPX2315	2315	4976	371	1113			NO-HIT	6
SPX2316	2316	4977	72	216			NO-HIT	6
SPX2317	2317	4978	84	252			NO-HIT	6
SPX2318	2318	4979	288	864	117	4.70E-08	[GI:4102023] [LN:AF007761] [AC:AF007761] [PN:MutR] [GN:mutR] [FN:positive transcriptional regulator of mutA] [OR:Streptococcus mutans]	134
SPX2319	2319	4980	196	588	300	1.00E-34	[GI:1619598] [LN:LGAPFA] [AC:Y08498] [PN:aggregation promoting protein] [GN:apfA] [OR:Lactobacillus gasseri]	108
SPX2320	2320	4981	97	291			NO-HIT	6
SPX2321	2321	4982	224	672	276	8.70E-69	[LN:SDHB_BACSU]	115

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:O34635] [GN:YLOW] [OR:Bacillus subtilis] [EC:4.2.1.13] [DE:DEAMINASE) (SDH) (L-SD)] [SP:O34635]	
SPX2322	2322	4983	291	873	877	9.00E-119	[LN:SDHA_BACSU] [AC:O34607] [GN:YLPA] [OR:Bacillus subtilis] [EC:4.2.1.13] [DE:DEAMINASE) (SDH) (L-SD)] [SP:O34607]	115
SPX2323	2323	4984	178	534			NO-HIT	6
SPX2324	2324	4985	211	633	123	1.50E-15	[GI:7576264] [LN:ECH277403] [AC:AJ277403] [PN:IndB protein] [GN:indB] [OR:Erwinia chrysanthemi]	95
SPX2325	2325	4986	72	216			NO-HIT	6
SPX2326	2326	4987	617	1851	1006	3.00E-136	[GI:6601348] [LN:AF155805] [AC:AF155805] [PN:Cps9E] [GN:cps9E] [OR:Streptococcus suis]	86
SPX2327	2327	4988	84	252	135	6.70E-14	[LN:S28486] [AC:S28486] [PN:hypothetical protein 2] [OR:Vibrio cholerae]	72
SPX2328	2328	4989	392	1176	123	7.20E-25	[LN:T44514] [AC:T44514] [PN:hypothetical protein 6P [imported]] [OR:Plesiomonas shigelloides]	93
SPX2329	2329	4990	102	306			NO-HIT	6
SPX2330	2330	4991	249	747	142	5.40E-29	[LN:YYCB_BACSU] [AC:P37482] [GN:YYCB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 43.2 KD PROTEIN IN DNAC-RPLI INTERGENIC REGION] [SP:P37482]	137
SPX2331	2331	4992	143	429	143	9.10E-15	[LN:PQ0016] [AC:PQ0016] [PN:hypothetical 9K protein] [OR:Lactobacillus confusus]	80
SPX2332	2332	4993	109	327	102	4.10E-07	[LN:D75250] [AC:D75250] [PN:conserved hypothetical protein] [GN:DR2629]	100

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Deinococcus radiodurans]	
SPX2333	2333	4994	198	594			NO-HIT	6
SPX2334	2334	4995	304	912			NO-HIT	6
SPX2335	2335	4996	123	369			NO-HIT	6
SPX2336	2336	4997	355	1065	118	3.00E-07	[LN:T18283] [AC:T18283] [PN:hypothetical protein G5] [OR:Dictyostelium discoideum]	82
SPX2337	2337	4998	318	954			NO-HIT	6
SPX2338	2338	4999	329	987	925	1.60E-165	[GI:2565150] [LN:LLU92974] [AC:U92974:M90760:M90761] [PN:unknown] [OR:Lactococcus lactis]	89
SPX2339	2339	5000	69	207			NO-HIT	6
SPX2340	2340	5001	515	1545			NO-HIT	6
SPX2341	2341	5002	134	402			NO-HIT	6
SPX2342	2342	5003	308	924	389	1.50E-61	[GI:1483212] [LN:ATCELD] [AC:Z77855] [PN:sugar-binding transport protein] [OR:Anaerocellum thermophilum]	104
SPX2343	2343	5004	320	960	601	9.60E-95	[GI:1483211] [LN:ATCELD] [AC:Z77855] [PN:sugar-binding transport protein] [OR:Anaerocellum thermophilum]	104
SPX2344	2344	5005	70	210			NO-HIT	6
SPX2345	2345	5006	120	360			NO-HIT	6
SPX2346	2346	5007	116	348	577	5.80E-76	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96
SPX2347	2347	5008	177	531	299	3.70E-73	[LN:A37146] [AC:A37146:A44901:S11354:C69699:I39962] [PN:ribosomal protein S4:ribosomal protein BS4 (rpsD)] [GN:rpsD] [CL:Escherichia coli ribosomal protein S4] [OR:Bacillus subtilis]	182
SPX2348	2348	5009	359	1077	1709	2.00E-231	[GI:5830539] [LN:SPAJ6397] [AC:AJ006397] [PN:histidine kinase] [GN:hk08] [OR:Streptococcus pneumoniae]	102
SPX2349	2349	5010	233	699	1196	1.40E-161	[GI:5830538] [LN:SPAJ6397] [AC:AJ006397] [PN:response regulator]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:rr08] [OR:Streptococcus pneumoniae]	
SPX2350	2350	5011	124	372			NO-HIT	6
SPX2351	2351	5012	891	2673	106	0.00022	[GI:7293562] [LN:AE003511] [AC:AE003511:AE002593] [GN:CG7874] [OR:Drosophila melanogaster] [SR:fruit fly]	105
SPX2352	2352	5013	190	570			NO-HIT	6
SPX2353	2353	5014	137	411			NO-HIT	6
SPX2354	2354	5015	179	537	199	2.70E-20	[GI:3550634] [LN:SPAJ6986] [AC:AJ006986] [PN:glycosyl transferase] [GN:cap33fH] [FN:synthesis of capsular polysaccharide] [OR:Streptococcus pneumoniae]	151
SPX2355	2355	5016	222	666	346	1.00E-42	[LN:YKQB__BACSU] [AC:P39760] [GN:YKQB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 24.3 KD PROTEIN IN KINC-ADEC INTERGENIC REGION (ORF4)] [SP:P39760]	144
SPX2356	2356	5017	460	1380	339	7.90E-94	[LN:G75367] [AC:G75367] [PN:potassium uptake protein KtrB] [GN:DR1668] [OR:Deinococcus radiodurans]	99
SPX2357	2357	5018	62	186	162	1.60E-17	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2358	2358	5019	255	765	75	8.50E-08	[LN:D72516] [AC:D72516] [PN:probable uridine phosphorylase APE2105] [GN:APE2105] [CL:purine-nucleoside phosphorylase pnp] [OR:Aeropyrum pernix]	143
SPX2359	2359	5020	63	189	91	8.80E-06	“[GI:6332762] [LN:AB033763] [AC:AB033763:AB014419:AB014429:AB014439] [PN:hypothetical protein] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:NCTC10442) DNA, clone__lib:Lambda das]”	194
SPX2360	2360	5021	99	297	153	6.00E-20	[LN:YQHL__BACSU] [AC:P54510] [GN:YQHL] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION]	141

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P54510]	
SPX2361	2361	5022	121	363			NO-HIT	6
SPX2362	2362	5023	98	294	194	1.30E-21	"[GI:6332750] [LN:AB033763] [AC:AB033763:AB014419:AB014429:AB014439] [PN:hypothetical protein] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:NCTC10442) DNA, clone_lib:Lambda das]"	194
SPX2363	2363	5024	82	246	167	8.40E-17	[LN:A75356] [AC:A75356] [PN:conserved hypothetical protein] [GN:DR1764] [OR:Deinococcus radiodurans]	100
SPX2364	2364	5025	132	396	175	2.80E-26	[LN:T44495] [AC:T44495] [PN:hypothetical protein YFIE [imported]] [OR:Bacillus halodurans]	90
SPX2365	2365	5026	346	1038	219	2.60E-44	[GI:6562808] [LN:SC4A7] [AC:AL133423] [PN:putative aldose 1-epimerase] [GN:SC4A7.35] [OR:Streptomyces coelicolor A3(2)]	119
SPX2366	2366	5027	153	459			NO-HIT	6
SPX2367	2367	5028	389	1167	473	3.90E-95	[LN:AGAS_ECOLI] [AC:P42907] [GN:AGAS] [OR:Escherichia coli] [DE:AGAS PROTEIN] [SP:P42907]	89
SPX2368	2368	5029	83	249	77	6.20E-05	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX2369	2369	5030	135	405	103	3.80E-13	[GI:5669855] [LN:AF130465] [AC:AF130465] [PN:mannose-specific phosphotransferase system] [GN:manL] [OR:Streptococcus salivarius]	128
SPX2370	2370	5031	277	831	144	4.90E-34	"[LN:PTPD_ECOLI] [AC:P42911] [GN:AGAD] [OR:Escherichia coli] [DE:ENZYME II, D COMPONENT)] [SP:P42911]"	102
SPX2371	2371	5032	302	906	128	6.60E-16	[GI:1732200] [LN:VFU65015] [AC:U65015]	115

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2372	2372	5033	159	477	298	2.50E-36	[PN:PTS permease for mannose subunit IIPMan] [GN:manY] [OR:Vibrio furnissii] [GI:6690421] [LN:AF129168] [AC:AF129168] [PN:EIIB sorbose-PTS homolog] [GN:sorB] [OR:Lactobacillus casei]	105
SPX2373	2373	5034	509	1527	265	3.30E-90	"[LN:JC5618] [AC:JC5618] [PN:beta-galactosidase,:lactase] [GN:bgaC] [CL:beta-galactosidase bga] [OR:Bacillus circulans] [EC:3.2.1.23]"	134
SPX2374	2374	5035	131	393	358	5.70E-44	"[LN:JC5618] [AC:JC5618] [PN:beta-galactosidase,:lactase] [GN:bgaC] [CL:beta-galactosidase bga] [OR:Bacillus circulans] [EC:3.2.1.23]"	134
SPX2375	2375	5036	239	717	142	7.80E-15	[LN:F69750] [AC:F69750] [PN:transcription regulator GntR family homolog ybgA] [GN:ybgA] [CL:transcription regulator GntR] [OR:Bacillus subtilis]	144
SPX2376	2376	5037	1304	3912	6324	0	"[LN:STRH_STRPN] [AC:P49610] [GN:STRH] [OR:Streptococcus pneumoniae] [EC:3.2.1.52] [DE:BETA-N-ACETYLHEXOSAMINIDASE PRECURSOR,] [SP:P49610]"	139
SPX2377	2377	5038	130	390			NO-HIT	6
SPX2378	2378	5039	433	1299	1678	1.40E-227	[LN:PUR8_BACSU] [AC:P12047] [GN:PURB:PURE] [OR:Bacillus subtilis] [EC:4.3.2.2] [DE:(GLUTAMYL-TRNA SYNTHETASE REGULATORY FACTOR)] [SP:P12047]	140
SPX2379	2379	5040	76	228			NO-HIT	6
SPX2380	2380	5041	384	1152	286	1.70E-65	[GI:5051462] [LN:NME242842] [AC:AJ242842] [PN:putative phosphoribosylaminoimidazole] [GN:purK]	155

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[FN:purine nucleotide synthesis] [OR:Neisseria meningitidis]	
SPX2381	2381	5042	155	465	515	5.30E-66	[GI:3892884] [LN:LLJ000883] [AC:AJ000883] [GN:purE] [OR:Lactococcus lactis]	75
SPX2382	2382	5043	421	1263	589	2.10E-157	[GI:3892883] [LN:LLJ000883] [AC:AJ000883] [GN:purD] [OR:Lactococcus lactis]	75
SPX2383	2383	5044	227	681			NO-HIT	6
SPX2384	2384	5045	523	1569	881	2.30E-213	[LN:PUR9_BACSU] [AC:P12048] [GN:PURH:PURHJ] [OR:Bacillus subtilis] [EC:2.1.2.3:3.5.4.10] [DE:(IMP SYNTHETASE) (ATIC)]] [SP:P12048]	130
SPX2385	2385	5046	170	510	136	9.10E-15	“[LN:VANZ_ENTFC] [AC:Q06242] [GN:VANZ] [OR:Enterococcus faecium] [SR:,Streptococcus faecium] [DE:VANZ PROTEIN] [SP:Q06242]”	123
SPX2386	2386	5047	116	348			NO-HIT	6
SPX2387	2387	5048	182	546	553	2.20E-72	[GI:6446399] [LN:SPU70775] [AC:U70775] [PN:phosphoribosylglycinamide formyltransferase] [OR:Streptococcus pyogenes]	115
SPX2388	2388	5049	341	1023	1084	7.20E-148	[GI:3150047] [LN:AF016634] [AC:AF016634] [PN:phosphoribosylformylglycinamide cyclo-ligase] [GN:pur5] [OR:Lactococcus lactis subsp. cremoris]	140
SPX2389	2389	5050	481	1443	1727	7.50E-248	[GI:4097534] [LN:LLU64311] [AC:U64311] [PN:phosphoribosylpyrophosphate amidotransferase] [GN:purF] [OR:Lactococcus lactis]	122
SPX2390	2390	5051	1243	3729	149	3.00E-49	“[LN:C69492] [AC:C69492] [PN:phosphoribosylformylglycinamidine synthase, component II:formylglycinamide ribotide amidotransferase:phosphoribosylformylglycinamidine synthetase] [OR:Archaeoglobus fulgidus]	217

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[EC:6.3.5.3]”	
SPX2391	2391	5052	45	135			NO-HIT	6
SPX2392	2392	5053	204	612	875	9.90E-117	[LN:PUR7__STRPN] [AC:Q07296] [GN:PURC] [OR:Streptococcus pneumoniae] [EC:6.3.2.6] [DE:(SAICAR SYNTHETASE)] [SP:Q07296]	117
SPX2393	2393	5054	124	372	175	1.90E-17	“[LN:H75412] [AC:H75412] [PN:spermidine/putrescine ABC transporter, periplasmic spermidine/putrescine-binding protein] [GN:DR1305] [OR:Deinococcus radiodurans]”	160
SPX2394	2394	5055	431	1293	2159	2.90E-302	[GI:2804700] [LN:AF030361] [AC:AF030361] [PN:transposase] [OR:Streptococcus pneumoniae]	87
SPX2395	2395	5056	86	258			NO-HIT	6
SPX2396	2396	5057	166	498	300	8.70E-37	[LN:YEBR__ECOLI] [AC:P76270:O07976:O07978] [GN:YEBR] [OR:Escherichia coli] [DE:HYPOTHETICAL 20.3 KD PROTEIN IN PRC-PPHA INTERGENIC REGION] [SP:P76270:O07976:O07978]	163
SPX2397	2397	5058	552	1656	749	7.10E-135	“[LN:S13786] [AC:S13786:S00745:S66049:B69618] [PN:DNA-directed DNA polymerase, III chain dnaX:DNA polymerase III (gamma and tau subunits) dnaX] [GN:dnaX:dnaZX] [CL:DNA-directed DNA polymerase III gamma chain] [OR:Bacillus subtilis] [EC:2.7.7.7]”	245
SPX2398	2398	5059	65	195			NO-HIT	6
SPX2399	2399	5060	257	771	948	1.70E-125	[LN:V296__BACSU] [AC:P80866] [GN:YURI] [OR:Bacillus subtilis] [DE:VEGETATIVE PROTEIN 296 (VEG296)] [SP:P80866]	109
SPX2400	2400	5061	420	1260	583	1.90E-87	[LN:G70019] [AC:G70019] [PN:conserved hypothetical protein yurX] [GN:yurX] [CL:Methanobacterium thermoautotrophicum ABC transporter Ycf24] [OR:Bacillus subtilis]	161
SPX2401	2401	5062	409	1227	622	8.60E-181	[LN:F70019] [AC:F70019] [PN:nifS protein homolog yurW] [GN:yurW]	105

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[CL:nifS protein] [OR:Bacillus subtilis]	
SPX2402	2402	5063	157	471	229	1.90E-43	[LN:E70019] [AC:E70019] [PN:nitrogen fixation protein nifU homolog yurV] [GN:yurV] [CL:Yeast nitrogen fixation protein:nitrogen fixation protein homology] [OR:Bacillus subtilis]	177
SPX2403	2403	5064	470	1410	1721	1.20E-229	[LN:D70019] [AC:D70019] [PN:conserved hypothetical protein yurU] [GN:yurU] [CL:Methanobacterium thermoautotrophicum ABC transporter Ycf24] [OR:Bacillus subtilis]	161
SPX2404	2404	5065	65	195			NO-HIT	6
SPX2405	2405	5066	62	186	112	8.10E-11	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2406	2406	5067	414	1242	2059	6.00E-276	“[GI:1620468] [LN:SPDACAO] [AC:X99400] [PN:D,D-carboxypeptidase] [GN:dacA] [OR:Streptococcus pneumoniae]”	105
SPX2407	2407	5068	370	1110	1807	5.00E-255	[GI:1620467] [LN:SPDACAO] [AC:X99400] [PN:membrane protein] [OR:Streptococcus pneumoniae]	89
SPX2408	2408	5069	73	219			NO-HIT	6
SPX2409	2409	5070	67	201			NO-HIT	6
SPX2410	2410	5071	247	741	258	7.40E-41	[LN:B69627] [AC:B69627] [PN:transcription repressor of fructose operon fruR] [GN:fruR] [CL:regulatory protein gutR] [OR:Bacillus subtilis]	138
SPX2411	2411	5072	304	912	725	2.80E-94	[LN:A69627] [AC:A69627] [PN:fructose 1-phosphate kinase fruB] [GN:fruB] [CL:6-phosphofructokinase 2] [OR:Bacillus subtilis]	123
SPX2412	2412	5073	76	228			NO-HIT	6
SPX2413	2413	5074	651	1953	371	4.60E-168	“[LN:H69626] [AC:H69626] [PN:PTS fructose-specific enzyme IIBC component fruA] [GN:fruA] [CL:phosphotransferase system enzyme II, fructose-specific:phosphotransferase system mannitol-specific enzyme II factor III homology]	246

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis]”	
SPX2414	2414	5075	738	2214	934	4.10E-198	[LN:SP3E_BACSU] [AC:P21458:P21459] [GN:SPOIIIE] [OR:Bacillus subtilis] [DE:STAGE III SPORULATION PROTEIN E] [SP:P21458:P21459]	126
SPX2415	2415	5076	84	252	82	0.00017	[GI:2257458] [LN:AB000222] [AC:AB000222] [GN:epr] [FN:glycylglycine endopeptidase resistance] [OR:Staphylococcus capitis] [SR:Staphylococcus capitis DNA]	153
SPX2416	2416	5077	95	285	124	7.00E-12	“[LN:A70132] [AC:A70132] [PN:cell division protein homolog] [OR:Borrelia burgdorferi] [SR:, Lyme disease spirochete]”	117
SPX2417	2417	5078	72	216			NO-HIT	6
SPX2418	2418	5079	220	660	102	0.00013	“[LN:T06029] [AC:T06029] [PN:hypothetical protein T28I19.100] [GN:T28I19.100] [OR:Arabidopsis thaliana] [SR:, mouse-ear cress]”	127
SPX2419	2419	5080	401	1203	698	8.70E-120	“[LN:NIFS_LACDE] [AC:P31672] [OR:Lactobacillus delbrueckii] [SR:,subspbulgaricus] [DE:NIFS PROTEIN HOMOLOG (FRAGMENT)] [SP:P31672]”	131
SPX2420	2420	5081	405	1215	982	1.40E-130	[LN:THII_BACSU] [AC:O34595] [GN:THII] [OR:Bacillus subtilis] [DE:PROBABLE THIAMINE BIOSYNTHESIS PROTEIN THII] [SP:O34595]	121
SPX2421	2421	5082	210	630	115	3.40E-15	[LN:G75475] [AC:G75475] [PN:conserved hypothetical protein] [GN:DR0779] [OR:Deinococcus radiodurans]	100
SPX2422	2422	5083	116	348			NO-HIT	6
SPX2423	2423	5084	112	336			NO-HIT	6
SPX2424	2424	5085	144	432			NO-HIT	6
SPX2425	2425	5086	61	183			NO-HIT	6
SPX2426	2426	5087	389	1167	82	2.50E-06	[LN:T35924] [AC:T35924] [PN:hypothetical protein SC9B5.04]	101

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:SC9B5.04] [OR:Streptomyces coelicolor]	
SPX2427	2427	5088	498	1494	137	2.50E-32	[GI:1841495] [LN:SEHSDRMS] [AC:Y11005] [PN:StySKI methylase] [GN:hsdM] [OR:Salmonella enterica]	95
SPX2428	2428	5089	374	1122	109	3.30E-14	[LN:YC18_METJA] [AC:Q58615] [GN:MJ1218] [OR:Methanococcus jannaschii] [DE:HYPOTHETICAL PROTEIN MJ1218] [SP:Q58615]	114
SPX2429	2429	5090	79	237			NO-HIT	6
SPX2430	2430	5091	138	414			NO-HIT	6
SPX2431	2431	5092	276	828	153	1.90E-22	[LN:A75153] [AC:A75153] [PN:integrase/recombinase xerd PAB0255] [GN:xerD-like:PAB0255] [CL:probable site-specific integrase/recombinase XerC] [OR:Pyrococcus abyssi]	164
SPX2432	2432	5093	369	1107	148	4.50E-19	[GI:3057063] [LN:AF013165] [AC:AF013165] [PN:HsdS] [GN:hsds] [OR:Lactococcus lactis]	84
SPX2433	2433	5094	1117	3351	153	6.50E-30	“[LN:T1R_ECOLI] [AC:P08956] [GN:HSDR:HSR] [OR:Escherichia coli] [EC:3.1.21.3] [DE:TYPE I RESTRICTION ENZYME ECOK I R PROTEIN,] [SP:P08956]”	139
SPX2434	2434	5095	157	471	132	1.60E-24	[LN:AHRC_BACSU] [AC:P17893] [GN:AHRC] [OR:Bacillus subtilis] [DE:ARGININE HYDROXIMATE RESISTANCE PROTEIN] [SP:P17893]	117
SPX2435	2435	5096	758	2274	1298	1.20E-249	“[LN:PEPX_LACLC] [AC:P22093] [GN:PEPX] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [EC:3.4.14.11] [DE:PEPTIDASE) (X-PROLYL-DIPEPTIDYL AMINOPEPTIDASE) (X-PDAP)] [SP:P22093]”	195
SPX2436	2436	5097	1034	3102	563	5.40E-170	“[LN:DP3A_BACSU]	125

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:O34623] [GN:DNAE] [OR:Bacillus subtilis] [EC:2.7.7.7] [DE:DNA POLYMERASE III, ALPHA CHAIN,] [SP:O34623]”	
SPX2437	2437	5098	63	189			NO-HIT	6
SPX2438	2438	5099	336	1008	1356	2.50E-182	“[LN:K6PF_LACLA] [AC:Q07636] [GN:PFKA:PFK] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [EC:2.7.1.11] [DE:(PHOSPHOHEXOKINASE)] [SP:Q07636]”	157
SPX2439	2439	5100	236	708	1007	8.20E-134	[GI:6708108] [LN:AF172173] [AC:AF172173] [PN:pyruvate kinase] [GN:pyk] [FN:conversion of phosphoenolpyruvate to pyruvate] [OR:Streptococcus thermophilus]	153
SPX2440	2440	5101	305	915	1099	9.10E-150	[GI:6708108] [LN:AF172173] [AC:AF172173] [PN:pyruvate kinase] [GN:pyk] [FN:conversion of phosphoenolpyruvate to pyruvate] [OR:Streptococcus thermophilus]	153
SPX2441	2441	5102	74	222	137	1.60E-13	[LN:G71171] [AC:G71171] [PN:hypothetical protein PH0571] [GN:PH0571] [OR:Pyrococcus horikoshii]	95
SPX2442	2442	5103	67	201	73	3.00E-08	[LN:VINT_BPL54] [AC:P20709] [GN:INT] [OR:Bacteriophage L54a] [DE:INTEGRASE] [SP:P20709]	87
SPX2443	2443	5104	101	303	95	3.10E-05	[GI:763050] [LN:BTU21935] [AC:U21935] [PN:repressor protein] [OR:Bacteriophage T270]	84
SPX2444	2444	5105	85	255	89	1.90E-08	[GI:509672] [LN:TU2CIRPRSR] [AC:L26219] [PN:repressor protein] [GN:cI]	128

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacteriophage Tuc2009] [SR:Bacteriophage Tuc2009 DNA]	
SPX2445	2445	5106	291	873	393	1.90E-60	[GI:4580015] [LN:AF049342] [AC:AF049342] [PN:unknown] [OR:Treponema denticola]	78
SPX2446	2446	5107	128	384	364	3.60E-46	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX2447	2447	5108	104	312			NO-HIT	6
SPX2448	2448	5109	239	717			NO-HIT	6
SPX2449	2449	5110	132	396			NO-HIT	6
SPX2450	2450	5111	272	816	184	2.20E-58	[LN:D69759] [AC:D69759] [PN:hypothetical protein ycgQ] [GN:ycgQ] [CL:Bacillus subtilis hypothetical protein ycgQ] [OR:Bacillus subtilis]	136
SPX2451	2451	5112	302	906	369	7.00E-86	[LN:E69759] [AC:E69759] [PN:hypothetical protein ycgR] [GN:ycgR] [OR:Bacillus subtilis]	87
SPX2452	2452	5113	710	2130	644	1.20E-222	[LN:G69773] [AC:G69773] [PN:conserved hypothetical protein ydcI] [GN:ydcI] [CL:hypothetical protein ydcI] [OR:Bacillus subtilis]	128
SPX2453	2453	5114	150	450	256	7.80E-44	[GI:7007441] [LN:AB031213] [AC:AB031213] [PN:YdcK] [GN:ydcK] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA]	129
SPX2454	2454	5115	85	255	90	7.80E-11	[LN:E70043] [AC:E70043] [PN:hypothetical protein yvlC] [GN:yvlC] [OR:Bacillus subtilis]	87
SPX2455	2455	5116	64	192			NO-HIT	6
SPX2456	2456	5117	253	759	486	8.70E-81	[LN:A70001] [AC:A70001] [PN:ABC transporter (ATP-binding protein) homolog ytsC] [GN:ytsC] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Bacillus subtilis]	188

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2457	2457	5118	75	225	76	3.20E-05	[LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061] [OR:Aeropyrum pernix]	92
SPX2458	2458	5119	663	1989	249	1.70E-50	[LN:B70001] [AC:B70001] [PN:ABC transporter (permease) homolog ytsD] [GN:ytsD] [OR:Bacillus subtilis]	101
SPX2459	2459	5120	197	591	230	2.10E-25	[LN:SPBC1683] [AC:AL355920] [PN:hypothetical protein] [GN:SPBC1683.10c] [OR:Schizosaccharomyces pombe] [SR:fission yeast]	121
SPX2460	2460	5121	365	1095	1143	8.40E-154	[GI:517210] [LN:SPU11799] [AC:U11799] [OR:Streptococcus pyogenes]	65
SPX2461	2461	5122	61	183			NO-HIT	6
SPX2462	2462	5123	492	1476	585	3.60E-138	"[LN:DCLY__BACSU] [AC:P21885:P26934] [GN:CAD] [OR:Bacillus subtilis] [EC:4.1.1.18] [DE:LYSINE DECARBOXYLASE, (LDC)] [SP:P21885:P26934]"	134
SPX2463	2463	5124	79	237			NO-HIT	6
SPX2464	2464	5125	176	528	83	1.20E-07	"[GI:6009430] [LN:AB024946] [AC:AB024946] [GN:orf54] [OR:Escherichia coli] [SR:Escherichia coli (sub__species:enteropathogenic, strain:B171]"	140
SPX2465	2465	5126	287	861	439	5.40E-79	[LN:SPEE__BACSU] [AC:P70998] [GN:SPEE] [OR:Bacillus subtilis] [EC:2.5.1.16] [DE:(SPDSY)] [SP:P70998]	99
SPX2466	2466	5127	420	1260	1090	2.30E-187	[LN:C81435] [AC:C81435] [PN:hypothetical protein Cj0172c [imported]] [GN:Cj0172c] [OR:Campylobacter jejuni]	107
SPX2467	2467	5128	376	1128	411	7.60E-103	[LN:E75398] [AC:E75398] [PN:carboxynorspermidine decarboxylase]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:DR1410] [OR:Deinococcus radiodurans]	
SPX2468	2468	5129	362	1086	530	3.80E-94	[GI:5712716] [LN:AF153708] [AC:AF153708] [PN:unknown] [OR:Pseudomonas sp. BG33R]	80
SPX2469	2469	5130	292	876	585	1.20E-110	[GI:5262946] [LN:LES19104] [AC:Y19104] [PN:beta-alanine synthase] [OR:Lycopersicon esculentum] [SR:tomato]	106
SPX2470	2470	5131	270	810	209	6.80E-43	[LN:YXEH_BACSU] [AC:P54947] [GN:YXEH:IP1B] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 30.2 KD PROTEIN IN IDH-DEOR INTERGENIC REGION] [SP:P54947]	141
SPX2471	2471	5132	74	222			NO-HIT	6
SPX2472	2472	5133	224	672	90	2.30E-09	[LN:E69787] [AC:E69787] [PN:hypothetical protein ydiL] [GN:ydiL] [OR:Bacillus subtilis]	87
SPX2473	2473	5134	303	909	229	5.80E-46	[LN:T44638] [AC:T44638] [PN:capsular polysaccharide biosynthesis protein cpsY [imported]] [GN:cpsY] [CL:probable transcription regulator lsyR] [OR:Streptococcus agalactiae]	172
SPX2474	2474	5135	154	462	197	1.80E-31	[LN:LSPA_BACSU] [AC:Q45479] [GN:LSPA:LSP] [OR:Bacillus subtilis] [EC:3.4.23.36] [DE:PEPTIDASE) (SIGNAL PEPTIDASE II) (SPASE II)] [SP:Q45479]	140
SPX2475	2475	5136	296	888	888	1.80E-117	[LN:YLYB_BACSU] [AC:Q45480:O31732] [GN:YLYB] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 33.7 KD PROTEIN IN LSP-PYRR INTERGENIC REGION (ORF-X)] [SP:Q45480:O31732]	158
SPX2476	2476	5137	628	1884	461	1.70E-56	[GI:1914872] [LN:SPZ82001] [AC:Z82001] [PN:PCPA] [GN:pcpA] [OR:Streptococcus pneumoniae]	88

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2477	2477	5138	377	1131	1387	4.40E-185	[LN:F81125] [AC:F81125] [PN:glutamate 5-kinase NMB1069 [imported]] [GN:NMB1069] [OR:Neisseria meningitidis]	107
SPX2478	2478	5139	421	1263	1530	1.10E-204	[GI:7413448] [LN:AE002457] [AC:AE002457:AE002098] [PN:gamma-glutamyl phosphate reductase] [GN:NMB1068] [OR:Neisseria meningitidis]	130
SPX2479	2479	5140	266	798	396	3.30E-67	"[LN:PROC_ARATH] [AC:P54904] [GN:PROC1] [OR:Arabidopsis thaliana] [SR:,Mouse-ear cress] [EC:1.5.1.2] [DE:PYRROLINE-5-CARBOXYLATE REDUCTASE, (P5CR) (P5C REDUCTASE)] [SP:P54904]"	176
SPX2480	2480	5141	213	639	528	4.30E-68	"[LN:KTHY_BACSU] [AC:P37537] [GN:TMK] [OR:Bacillus subtilis] [EC:2.7.4.9] [DE:THYMIDYLATE KINASE, (DTMP KINASE)] [SP:P37537]"	125
SPX2481	2481	5142	84	252			NO-HIT	6
SPX2482	2482	5143	297	891	244	4.80E-37	"[LN:HOLB_BACSU] [AC:P37540] [GN:HOLB] [OR:Bacillus subtilis] [EC:2.7.7.7] [DE:DNA POLYMERASE III, DELTA'SUBUNIT,] [SP:P37540]"	128
SPX2483	2483	5144	106	318	103	9.60E-09	[LN:YABA_BACSU] [AC:P37542] [GN:YABA] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 14.1 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION] [SP:P37542]	137
SPX2484	2484	5145	290	870	618	1.60E-84	[LN:YABC_BACSU] [AC:P37544] [GN:YABC] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 33.0 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION] [SP:P37544]	137
SPX2485	2485	5146	182	546	362	2.50E-61	[LN:S62019] [AC:S62019] [PN:hypothetical protein YDR540c:hypothetical protein D3703.8] [GN:YDR540c]	177

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[CL:Saccharomyces hypothetical protein YDR540c] [OR:Saccharomyces cerevisiae]	
SPX2486	2486	5147	168	504	82	2.90E-09	[GI:6690333] [LN:AF117259] [AC:AF117259] [PN:replication protein] [GN:repX] [OR:Staphylococcus aureus]	102
SPX2487	2487	5148	150	450	400	5.20E-51	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX2488	2488	5149	134	402	370	1.00E-54	[LN:S52544] [AC:S52544] [PN:ISL2 protein] [OR:Lactobacillus helveticus]	71
SPX2489	2489	5150	91	273			NO-HIT	6
SPX2490	2490	5151	445	1335	585	1.40E-196	[LN:GID_BACSU] [AC:P39815] [GN:GID] [OR:Bacillus subtilis] [DE:GID PROTEIN] [SP:P39815]	87
SPX2491	2491	5152	246	738	793	2.20E-104	[LN:F69708] [AC:F69708] [PN:uridylate kinase smbA] [GN:smbA] [CL:uridine 5'-monophosphate kinase] [OR:Bacillus subtilis]	120
SPX2492	2492	5153	146	438			NO-HIT	6
SPX2493	2493	5154	186	558	545	2.50E-69	[LN:G69626] [AC:G69626] [PN:ribosome recycling factor frr] [GN:frr] [CL:ribosome releasing factor] [OR:Bacillus subtilis]	121
SPX2494	2494	5155	139	417			NO-HIT	6
SPX2495	2495	5156	285	855	180	3.70E-45	[GI:2145404] [LN:BSY09476] [AC:Y09476] [PN:YitL] [OR:Bacillus subtilis]	71
SPX2496	2496	5157	72	216	134	1.00E-13	[LN:A69931] [AC:A69931] [PN:hypothetical protein yozE] [GN:yozE] [OR:Bacillus subtilis]	87
SPX2497	2497	5158	323	969	956	1.30E-128	[LN:PHOL_BACSU] [AC:P46343] [GN:PHOH]	95

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis] [DE:PHOH-LIKE PROTEIN] [SP:P46343]	
SPX2498	2498	5159	65	195			NO-HIT	6
SPX2499	2499	5160	167	501	87	0.00031	[LN:E75272] [AC:E75272] [PN:hypothetical protein] [GN:DR2441] [OR:Deinococcus radiodurans]	90
SPX2500	2500	5161	92	276	253	1.30E-28	"[LN:DHA_BACSU] [AC:Q08352] [GN:ALD:SPOVN] [OR:Bacillus subtilis] [EC:1.4.1.1] [DE:ALANINE DEHYDROGENASE, (STAGE V SPORULATION PROTEIN N)] [SP:Q08352]"	151
SPX2501	2501	5162	65	195	168	5.10E-17	"[LN:JE0388] [AC:JE0388] [PN:alanine dehydrogenase,] [CL:alanine dehydrogenase:alanine dehydrogenase homology] [OR:Enterobacter aerogenes] [EC:1.4.1.1]"	152
SPX2502	2502	5163	73	219	239	3.30E-27	"[LN:S74638] [AC:S74638] [PN:alanine dehydrogenase:hypothetical protein sll1682:hypothetical protein sll1682] [CL:alanine dehydrogenase:alanine dehydrogenase homology] [OR:Synechocystis sp.] [SR:PCC 6803, , PCC 6803] [SR:PCC 6803,]"	233
SPX2503	2503	5164	142	426	405	2.90E-56	"[LN:DHA_BACSU] [AC:Q08352] [GN:ALD:SPOVN] [OR:Bacillus subtilis] [EC:1.4.1.1] [DE:ALANINE DEHYDROGENASE, (STAGE V SPORULATION PROTEIN N)] [SP:Q08352]"	151
SPX2504	2504	5165	196	588	147	8.00E-20	[LN:C69895] [AC:C69895] [PN:conserved hypothetical protein yoaA] [GN:yoaA] [CL:Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimJ] [OR:Bacillus subtilis]	170
SPX2505	2505	5166	217	651	1062	6.40E-138	[GI:3211753] [LN:AF052208] [AC:AF052208] [PN:competence protein] [GN:celA] [OR:Streptococcus pneumoniae]	104
SPX2506	2506	5167	459	1377	2213	0	[GI:3211754] [LN:AF052208]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:AF052208] [PN:competence protein] [GN:celB] [OR:Streptococcus pneumoniae]	
SPX2507	2507	5168	306	918	1215	1.90E-162	[GI:3211754] [LN:AF052208] [AC:AF052208] [PN:competence protein] [GN:celB] [OR:Streptococcus pneumoniae]	104
SPX2508	2508	5169	196	588	164	1.60E-26	[LN:YYBJ_BACSU] [AC:P37494] [GN:YYBJ] [OR:Bacillus subtilis] [DE:INTERGENIC REGION] [SP:P37494]	95
SPX2509	2509	5170	392	1176			NO-HIT	6
SPX2510	2510	5171	88	264			NO-HIT	6
SPX2511	2511	5172	196	588	589	4.10E-77	[LN:IF3_LISMO] [AC:O53084] [GN:INFC] [OR:Listeria monocytogenes] [DE:TRANSLATION INITIATION FACTOR IF-3] [SP:O53084]	116
SPX2512	2512	5173	67	201	235	1.10E-28	[LN:R5BS35] [AC:S05347] [PN:ribosomal protein L35] [GN:rpmI] [CL:Escherichia coli ribosomal protein L35] [OR:Bacillus stearothermophilus]	137
SPX2513	2513	5174	120	360	476	9.30E-63	[LN:RL20_BACSU] [AC:P55873] [GN:RPLT] [OR:Bacillus subtilis] [DE:50S RIBOSOMAL PROTEIN L20] [SP:P55873]	103
SPX2514	2514	5175	127	381	130	6.20E-24	[LN:LGUL_HAEIN] [AC:P44638] [GN:GLOA:HI0323] [OR:Haemophilus influenzae] [EC:4.4.1.5] [DE:(S-D-LACTOYLGLUTATHIONE METHYLGLYOXAL LYASE)] [SP:P44638]	147
SPX2515	2515	5176	190	570	354	2.10E-55	[LN:YLXD_BACCL] [AC:P46536] [OR:Bacillus caldolyticus] [DE:HYPOTHETICAL 27.6 KD PROTEIN IN PYRAB-PYRD INTERGENIC REGION (ORF2)] [SP:P46536]	139
SPX2516	2516	5177	94	282	172	7.70E-23	[LN:YLXD_BACCL] [AC:P46536]	139

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus caldolyticus] [DE:HYPOTHETICAL 27.6 KD PROTEIN IN PYRAB-PYRD INTERGENIC REGION (ORF2)] [SP:P46536]	
SPX2517	2517	5178	313	939	563	8.20E-147	“[LN:PYDB_LACLC] [AC:P54322] [GN:PYRDB] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [EC:1.3.3.1] [DE:(DHODEHASE B) (DHODB)] [SP:P54322]”	159
SPX2518	2518	5179	722	2166	2390	0	[GI:4218533] [LN:SPN010312] [AC:AJ010312] [PN:endo-beta-N-acetylglucosaminidase] [GN:lytB] [FN:cell wall degradation and cell separation] [OR:Streptococcus pneumoniae]	167
SPX2519	2519	5180	561	1683	2767	0	[GI:6175915] [LN:AF181976] [AC:AF181976] [PN:adherence and virulence protein A] [GN:pavA] [FN:adhesin] [OR:Streptococcus pneumoniae]	132
SPX2520	2520	5181	77	231			NO-HIT	6
SPX2521	2521	5182	166	498	705	5.20E-92	[GI:2749950] [LN:AF000954] [AC:AF000954] [OR:Streptococcus mutans]	66
SPX2522	2522	5183	132	396	507	6.90E-71	[LN:A36933] [AC:A36933] [PN:diacylglycerol kinase homolog] [CL:Bacillus subtilis diacylglycerol kinase dgkA] [OR:Streptococcus mutans]	134
SPX2523	2523	5184	300	900	1502	1.40E-204	[GI:5305399] [LN:AF072811] [AC:AF072811] [PN:GTPase Era] [GN:era] [OR:Streptococcus pneumoniae]	95
SPX2524	2524	5185	275	825	892	1.90E-119	[LN:FPG_STRMU] [AC:P55045] [GN:MUTM:FPG] [OR:Streptococcus mutans] [EC:3.2.2.23] [DE:GLYCOSYLASE)] [SP:P55045]	110
SPX2525	2525	5186	61	183			NO-HIT	6
SPX2526	2526	5187	203	609	167	2.40E-32	“[LN:Y553_SYNY3]	140

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[AC:Q55515] [GN:SLR0553] [OR:Synechocystis sp] [SR:,strain PCC 6803] [DE:HYPOTHETICAL 22.5 KD PROTEIN SLR0553] [SP:Q55515]"	
SPX2527	2527	5188	67	201			NO-HIT	6
SPX2528	2528	5189	400	1200	1989	9.60E-280	[GI:3820455] [LN:SPN7367] [AC:AJ007367] [PN:multi-drug resistance efflux pump] [GN:pmrA] [OR:Streptococcus pneumoniae]	118
SPX2529	2529	5190	50	150	177	8.70E-20	"[LN:RL33_LACLA] [AC:P27167] [GN:RPMG] [OR:Lactococcus lactis] [SR:,subsplactis:Streptococcus lactis] [DE:50S RIBOSOMAL PROTEIN L33] [SP:P27167]"	145
SPX2530	2530	5191	78	234	82	8.20E-06	[LN:A70028] [AC:A70028] [PN:hypothetical protein yvaL] [GN:yvaL] [CL:protein-export protein secG] [OR:Bacillus subtilis]	120
SPX2531	2531	5192	785	2355	564	4.00E-180	[LN:G70027] [AC:G70027] [PN:conserved hypothetical protein yvaJ] [GN:yvaJ] [CL:virulence-associated protein vacB homolog] [OR:Bacillus subtilis]	144
SPX2532	2532	5193	156	468	793	1.30E-104	[GI:3211758] [LN:AF052209] [AC:AF052209] [PN:VacB homolog] [OR:Streptococcus pneumoniae]	88
SPX2533	2533	5194	287	861	1499	8.30E-203	[GI:4883699] [LN:AF079807] [AC:AF079807] [PN:tellurite resistance protein TehB] [GN:tehB] [OR:Streptococcus pneumoniae]	119
SPX2534	2534	5195	129	387	326	5.30E-41	[GI:3211758] [LN:AF052209] [AC:AF052209] [PN:VacB homolog] [OR:Streptococcus pneumoniae]	88
SPX2535	2535	5196	74	222			NO-HIT	6
SPX2536	2536	5197	318	954	800	5.10E-182	[GI:3211759]	104

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[LN:AF052209] [AC:AF052209] [PN:competence protein] [GN:coiA] [OR:Streptococcus pneumoniae]	
SPX2537	2537	5198	601	1803	2159	2.10E-292	"[LN:PEPB_STRAG] [AC:Q53778] [GN:PEPB] [OR:Streptococcus agalactiae] [EC:3.4.24.-] [DE:GROUP B OLIGOPEPTIDASE PEPB,] [SP:Q53778]"	129
SPX2538	2538	5199	238	714	680	4.90E-90	[GI:1771204] [LN:LLLVSFPEP] [AC:X99710] [PN:methyltransferase] [OR:Lactococcus lactis]	86
SPX2539	2539	5200	314	942	189	3.10E-32	[LN:PRTM_LACPA] [AC:Q02473] [GN:PRTM] [OR:Lactobacillus paracasei] [DE:PROTEASE MATURATION PROTEIN PRECURSOR] [SP:Q02473]	121
SPX2540	2540	5201	114	342	567	4.00E-74	[GI:1490399] [LN:SPPARCETP] [AC:Z67739] [PN:DNA transposase] [OR:Streptococcus pneumoniae]	90
SPX2541	2541	5202	191	573	422	2.30E-53	[LN:B30868] [AC:B30868] [PN:hypothetical protein 1] [OR:Streptococcus agalactiae]	81
SPX2542	2542	5203	91	273			NO-HIT	6
SPX2543	2543	5204	64	192			NO-HIT	6
SPX2544	2544	5205	207	621	86	4.90E-12	"[LN:T46083] [AC:T46083] [PN:hypothetical protein T20E23.120] [CL:Aquifex aeolicus phosphoglycerate mutase:phosphoglycerate mutase homology] [OR:Arabidopsis thaliana] [SR:, mouse-ear cress]"	190
SPX2545	2545	5206	161	483	210	1.70E-35	"[LN:EBSC_ENTFA] [AC:P36922] [OR:Enterococcus faecalis] [SR:,Streptococcus faecalis] [DE:EBSC PROTEIN] [SP:P36922]"	115
SPX2546	2546	5207	281	843	128	6.60E-17	[LN:B72411] [AC:B72411] [PN:conserved hypothetical protein] [GN:TM0164]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Thermotoga maritima]	
SPX2547	2547	5208	263	789	102	2.10E-09	[GI:1865711] [LN:BPPLYBA] [AC:Y11477] [PN:endolysin] [GN:plyBa] [FN:cell wall hydrolase] [OR:Bacteriophage Bastille]	116
SPX2548	2548	5209	460	1380	1180	1.90E-156	[LN:S66080] [AC:S66080:I40018:C69629:S05371:S18903] [PN:UDP-N-acetylglucosamine pyrophosphorylase gcaD:cell division protein tms26:tms protein] [GN:gcaD:tms26] [CL:N-acetylglucosamine-1-phosphate uridyltransferase] [OR:Bacillus subtilis]	237
SPX2549	2549	5210	182	546	167	2.40E-45	[LN:YQKG_BACSU] [AC:P54570] [GN:YQKG] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 21.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION] [SP:P54570]	137
SPX2550	2550	5211	105	315			NO-HIT	6
SPX2551	2551	5212	231	693	477	1.60E-61	[LN:PFS_ECOLI] [AC:P24247] [GN:PFS] [OR:Escherichia coli] [EC:3.2.2.16:3.2.2.9] [DE:(EC 3.2.2.9)]] [SP:P24247]	110
SPX2552	2552	5213	164	492	124	1.90E-12	[GI:2769573] [LN:LLCPJW565] [AC:Y12736] [OR:Lactococcus lactis subsp. cremoris]	79
SPX2553	2553	5214	197	591	137	1.00E-08	"[LN:DPO3_BACSU] [AC:P13267] [GN:POLC:DNAF:MUTI] [OR:Bacillus subtilis] [EC:2.7.7.7] [DE:DNA POLYMERASE III, ALPHA CHAIN POLC-TYPE, (POLIII)] [SP:P13267]"	154
SPX2554	2554	5215	257	771	105	1.40E-11	[LN:E69827] [AC:E69827] [PN:glycerophosphodiester phosphodiesterase homolog yhdW] [GN:yhdW] [CL:glycerophosphodiester phosphodiesterase] [OR:Bacillus subtilis]	159
SPX2555	2555	5216	116	348	577	5.80E-76	[GI:4200438] [LN:AF026471] [AC:AF026471] [PN:putative transposase] [OR:Streptococcus pneumoniae]	96

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2556	2556	5217	47	141	186	5.40E-22	[GI:5019553] [LN:SPN239004] [AC:AJ239004] [PN:putative transposase] [OR:Streptococcus pneumoniae]	97
SPX2557	2557	5218	67	201			NO-HIT	6
SPX2558	2558	5219	211	633			NO-HIT	6
SPX2559	2559	5220	100	300	84	0.00082	[GI:1707287] [LN:BBU80959] [AC:U80959:L78251] [PN:putative outer membrane protein] [GN:ospFi] [OR:Borrelia burgdorferi] [SR:Lyme disease spirochete]	148
SPX2560	2560	5221	236	708	267	1.40E-53	[LN:YE54_HAEIN] [AC:P44202] [GN:HI1454] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL CYTOCHROME C-TYPE BIOGENESIS PROTEIN HI1454] [SP:P44202]	141
SPX2561	2561	5222	186	558	239	6.40E-28	[LN:YE53_HAEIN] [AC:Q57127:O05062] [GN:HI1453] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI1453 PRECURSOR] [SP:Q57127:O05062]	136
SPX2562	2562	5223	464	1392	481	1.60E-122	[LN:D69814] [AC:D69814] [PN:metabolite transporter homolog yfnA] [GN:yfnA] [CL:arginine permease] [OR:Bacillus subtilis]	120
SPX2563	2563	5224	308	924	985	7.10E-129	[LN:T46757] [AC:T46757] [PN:lipoprotein lmb [validated]] [GN:lmb] [OR:Streptococcus agalactiae]	95
SPX2564	2564	5225	839	2517	409	3.50E-104	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX2565	2565	5226	1040	3120	400	1.50E-93	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX2566	2566	5227	63	189			NO-HIT	6
SPX2567	2567	5228	63	189	240	1.40E-26	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2568	2568	5229	67	201	104	3.80E-14	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX2569	2569	5230	161	483	152	4.70E-20	[LN:T46758] [AC:T46758] [PN:hypothetical protein [imported]] [OR:Streptococcus agalactiae]	90
SPX2570	2570	5231	80	240			NO-HIT	6
SPX2571	2571	5232	86	258			NO-HIT	6
SPX2572	2572	5233	145	435	107	6.20E-05	[GI:7228476] [LN:AF163151] [AC:AF163151] [PN:dentin sialophosphoprotein precursor] [GN:DSPP] [OR:Homo sapiens] [SR:human]	121
SPX2573	2573	5234	408	1224	961	7.20E-190	"[LN:PEPT__LACLC] [AC:P42020] [GN:PEPT] [OR:Lactococcus lactis] [SR:,subspcremoris:Streptococcus cremoris] [EC:3.4.11.-] [DE:PEPTIDASE T, (AMINOTRIPEPTIDASE) (TRIPEPTIDASE)] [SP:P42020]"	185
SPX2574	2574	5235	365	1095	215	4.70E-41	[LN:HEMZ__BACSU] [AC:P32396] [GN:HEMH:HEMF] [OR:Bacillus subtilis] [EC:4.99.1.1] [DE:SYNTHETASE)] [SP:P32396]	108
SPX2575	2575	5236	126	378	149	2.80E-25	[GI:6136300] [LN:AF065159] [AC:AF065159] [PN:MscL] [GN:mscL] [OR:Bradyrhizobium japonicum]	90
SPX2576	2576	5237	87	261			NO-HIT	6
SPX2577	2577	5238	129	387	180	3.60E-29	"[LN:T44787] [AC:T44787] [PN:glucokinase, [imported]] [CL:glucose kinase:glucose kinase homology] [OR:Bacillus megaterium] [EC:2.7.1.2]"	136
SPX2578	2578	5239	88	264	61	0.0009	[GI:2897104] [LN:AF020798] [AC:AF020798] [PN:putative host cell surface-exposed lipoprotein] [OR:Streptococcus thermophilus bacteriophage TP-J34]	145

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2579	2579	5240	273	819	279	1.60E-30	“[GI:4218526] [LN:SPAJ9639] [AC:AJ009639] [PN:1,4-beta-N-acetylmuramidase] [GN:lytC] [FN:lysis of cell wall peptidoglycan] [OR:Streptococcus pneumoniae]”	153
SPX2580	2580	5241	510	1530	289	6.20E-30	“[GI:4218526] [LN:SPAJ9639] [AC:AJ009639] [PN:1,4-beta-N-acetylmuramidase] [GN:lytC] [FN:lysis of cell wall peptidoglycan] [OR:Streptococcus pneumoniae]”	153
SPX2581	2581	5242	574	1722	339	6.40E-76	“[LN:A71951] [AC:A71951] [PN:p-aminobenzoate synthetase] [GN:pabB] [OR:Helicobacter pylori] [SR:strain J99, , strain J99] [SR:strain J99,]”	140
SPX2582	2582	5243	1964	5892	7027	0	[GI:6911257] [LN:AF221126] [AC:AF221126] [PN:putative zinc metalloprotease] [GN:zmpB] [OR:Streptococcus pneumoniae]	115
SPX2583	2583	5244	78	234			NO-HIT	6
SPX2584	2584	5245	108	324	71	7.10E-06	[GI:501027] [LN:TBU01849] [AC:U01849] [OR:Kinetoplast Trypanosoma brucei] [SR:Trypanosoma brucei]	97
SPX2585	2585	5246	188	564	392	1.60E-49	[LN:YWLG_BACSU] [AC:P39157] [GN:YWLG:IPC-33D] [OR:Bacillus subtilis] [DE:HYPOTHETICAL 19.4 KD PROTEIN IN SPOIIR-GLYC INTERGENIC REGION] [SP:P39157]	147
SPX2586	2586	5247	378	1134	1862	7.80E-253	[GI:6911256] [LN:AF221126] [AC:AF221126] [PN:putative histidine kinase] [GN:zmpS] [OR:Streptococcus pneumoniae]	111
SPX2587	2587	5248	63	189			NO-HIT	6
SPX2588	2588	5249	186	558	898	5.90E-119	[GI:6911256] [LN:AF221126] [AC:AF221126] [PN:putative histidine kinase]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[GN:zmpS] [OR:Streptococcus pneumoniae]	
SPX2589	2589	5250	246	738	1237	2.00E-167	[GI:6911255] [LN:AF221126] [AC:AF221126] [PN:putative response regulator] [GN:zmpR] [OR:Streptococcus pneumoniae]	113
SPX2590	2590	5251	371	1113	1110	1.40E-147	[LN:YE55_HAEIN] [AC:P45213] [GN:HI1455] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI1455] [SP:P45213]	112
SPX2591	2591	5252	198	594	335	9.20E-41	[LN:YE53_HAEIN] [AC:Q57127:O05062] [GN:HI1453] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL PROTEIN HI1453 PRECURSOR] [SP:Q57127:O05062]	136
SPX2592	2592	5253	237	711	139	3.50E-49	[LN:YE54_HAEIN] [AC:P44202] [GN:HI1454] [OR:Haemophilus influenzae] [DE:HYPOTHETICAL CYTOCHROME C-TYPE BIOGENESIS PROTEIN HI1454] [SP:P44202]	141
SPX2593	2593	5254	287	861	312	4.30E-47	[GI:3608389] [LN:AF071085] [AC:AF071085] [PN:Orfde2] [OR:Enterococcus faecalis]	79
SPX2594	2594	5255	685	2055	564	1.00E-82	[LN:E70040] [AC:E70040] [PN:conserved hypothetical protein yvgP] [GN:yvgP] [CL:hypothetical protein yvgP] [OR:Bacillus subtilis]	128
SPX2595	2595	5256	64	192			NO-HIT	6
SPX2596	2596	5257	85	255			NO-HIT	6
SPX2597	2597	5258	63	189			NO-HIT	6
SPX2598	2598	5259	188	564	221	1.00E-42	[LN:E69999] [AC:E69999] [PN:hypothetical protein ytqB] [GN:ytqB] [OR:Bacillus subtilis]	87
SPX2599	2599	5260	149	447	329	6.50E-47	[LN:D69999] [AC:D69999] [PN:conserved hypothetical protein ytqA] [GN:ytqA] [CL:Methanococcus jannaschii conserved hypothetical protein MJ0486]	165

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[OR:Bacillus subtilis]	
SPX2600	2600	5261	84	252	209	2.10E-24	[LN:D69999] [AC:D69999] [PN:conserved hypothetical protein ytqA] [GN:ytqA] [CL:Methanococcus jannaschii conserved hypothetical protein MJ0486] [OR:Bacillus subtilis]	165
SPX2601	2601	5262	146	438	210	2.60E-23	[LN:D69999] [AC:D69999] [PN:conserved hypothetical protein ytqA] [GN:ytqA] [CL:Methanococcus jannaschii conserved hypothetical protein MJ0486] [OR:Bacillus subtilis]	165
SPX2602	2602	5263	120	360			NO-HIT	6
SPX2603	2603	5264	82	246	182	1.20E-19	[GI:1022726] [LN:SHU35635] [AC:U35635] [PN:unknown] [OR:Staphylococcus haemolyticus] [SR:Staphylococcus haemolyticus strain=Y176]	129
SPX2604	2604	5265	208	624	156	1.10E-37	"[GI:6332767] [LN:AB033763] [AC:AB033763:AB014419:AB014429:AB014439] [PN:hypothetical protein] [OR:Staphylococcus aureus] [SR:Staphylococcus aureus (strain:NCTC10442) DNA, clone_lib:Lambda das]"	194
SPX2605	2605	5266	2173	6519	309	2.70E-130	[GI:6468240] [LN:SCF81] [AC:AL133171] [PN:putative secreted beta-galactosidase] [GN:SCF81.25c] [OR:Streptomyces coelicolor A3(2)]	129
SPX2606	2606	5267	62	186			NO-HIT	6
SPX2607	2607	5268	64	192			NO-HIT	6
SPX2608	2608	5269	144	432	184	6.10E-21	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2609	2609	5270	262	786	92	2.70E-11	"[LN:SGCC_ECOLI] [AC:P39365] [GN:SGCC] [OR:Escherichia coli] [DE:PUTATIVE PHOSPHOTRANSFERASE ENZYME II, C COMPONENT SGCC] [SP:P39365]"	134
SPX2610	2610	5271	231	693	231	3.70E-25	"[GI:1736815] [LN:D90848] [AC:D90848:AB001340] [PN:PTS system, Galactitol-specific IIC component] [GN:gatC] [OR:Escherichia coli]	201

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SR:Escherichia coli (strain:K12) DNA, clone__lib:Kohara lambda minise]”	
SPX2611	2611	5272	102	306			NO-HIT	6
SPX2612	2612	5273	159	477	95	5.10E-13	“[GI:4512375] [LN:AB011837] [AC:AB011837] [PN:phosphotransferase system (PTS)] [GN:fruA] [OR:Bacillus halodurans] [SR:Bacillus halodurans (strain:C-125) DNA, clone__lib:lambda no.9]”	181
SPX2613	2613	5274	2139	6417	10556	0	[GI:5726291] [LN:AF127143] [AC:AF127143] [PN:cell wall-associated serine proteinase precursor] [GN:prtA] [OR:Streptococcus pneumoniae]	134
SPX2614	2614	5275	68	204	136	2.70E-14	[LN:T30285] [AC:T30285] [PN:hypothetical protein] [OR:Streptococcus pneumoniae]	79
SPX2615	2615	5276	66	198			NO-HIT	6
SPX2616	2616	5277	262	786	107	2.30E-18	[GI:7635999] [LN:SCE6] [AC:AL353832] [PN:putative integral membrane transport protein.] [GN:SCE6.32c] [OR:Streptomyces coelicolor A3(2)]	136
SPX2617	2617	5278	287	861	113	2.80E-10	[GI:7636000] [LN:SCE6] [AC:AL353832] [PN:putative integral membrane protein.] [GN:SCE6.33c] [OR:Streptomyces coelicolor A3(2)]	126
SPX2618	2618	5279	318	954	601	1.60E-83	“[LN:G75548] [AC:G75548] [PN:ABC transporter, ATP-binding protein] [GN:DR0205] [CL:unassigned ATP-binding cassette proteins:ATP-binding cassette homology] [OR:Deinococcus radiodurans]”	184
SPX2619	2619	5280	83	249			NO-HIT	6
SPX2620	2620	5281	132	396			NO-HIT	6
SPX2621	2621	5282	158	474	123	3.70E-11	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX2622	2622	5283	230	690	765	9.20E-100	[LN:RL1__BACST] [AC:P04447] [GN:RPLA] [OR:Bacillus stearothermophilus] [DE:50S RIBOSOMAL PROTEIN L1]	111

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P04447]	
SPX2623	2623	5284	142	426	596	4.50E-77	[LN:S38871] [AC:S38871] [PN:ribosomal protein L11] [GN:rplK] [CL:Escherichia coli ribosomal protein L11] [OR:Staphylococcus carnosus]	133
SPX2624	2624	5285	75	225			NO-HIT	6
SPX2625	2625	5286	239	717	133	3.00E-20	[GI:2529473] [LN:AF006665] [AC:AF006665] [PN:YokZ] [GN:yokZ] [FN:unknown] [OR:Bacillus subtilis]	96
SPX2626	2626	5287	60	180			NO-HIT	6
SPX2627	2627	5288	170	510	103	7.80E-07	[LN:A81062] [AC:A81062] [PN:conserved hypothetical protein NMB1619 [imported]] [GN:NMB1619] [OR:Neisseria meningitidis]	119
SPX2628	2628	5289	72	216	187	1.80E-19	[LN:B71121] [AC:B71121] [PN:hypothetical protein PH0737] [GN:PH0737] [CL:thermophilic aminopeptidase I alpha chain] [OR:Pyrococcus horikoshii]	142
SPX2629	2629	5290	199	597	306	4.10E-72	[LN:C69830] [AC:C69830] [PN:glucanase homolog yhfE] [GN:yhfE] [CL:thermophilic aminopeptidase I alpha chain] [OR:Bacillus subtilis]	131
SPX2630	2630	5291	114	342	145	6.70E-14	[GI:2315995] [LN:SAU87144] [AC:U87144] [PN:branched-chain amino acid carrier protein] [OR:Staphylococcus aureus]	112
SPX2631	2631	5292	447	1341	391	6.60E-90	[LN:BRNQ_BACSU] [AC:P94499:O07082] [GN:BRNQ] [OR:Bacillus subtilis] [DE:CHAIN AMINO ACID UPTAKE CARRIER)] [SP:P94499:O07082]	124
SPX2632	2632	5293	63	189	88	3.10E-06	“[LN:YC61_SYNY3] [AC:P73801] [GN:SLR1261] [OR:Synechocystis sp] [SR:,strain PCC 6803] [DE:HYPOTHETICAL 19.1 KD PROTEIN SLR1261]	140

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P73801]”	
SPX2633	2633	5294	80	240	100	1.90E-17	[GI:1914870] [LN:SPZ82001] [AC:Z82001] [PN:unknown] [OR:Streptococcus pneumoniae]	81
SPX2634	2634	5295	193	579	548	7.70E-95	[GI:3212185] [LN:U32707] [AC:U32707:L42023] [PN:H. influenzae predicted coding region HI0220.2] [GN:HI0220.2] [OR:Haemophilus influenzae Rd]	140
SPX2635	2635	5296	467	1401	646	4.20E-200	[GI:2160707] [LN:LLU78036] [AC:U78036] [PN:dipeptidase] [OR:Lactococcus lactis]	79
SPX2636	2636	5297	202	606	113	2.60E-17	“[LN:NOX_THETH] [AC:Q60049:Q53306] [GN:NOX] [OR:Thermus aquaticus] [SR:,subspthermophilus] [EC:1.6.99.3] [DE:(NADH:OXYGEN OXIDOREDUCTASE)] [SP:Q60049:Q53306]”	158
SPX2637	2637	5298	80	240			NO-HIT	6
SPX2638	2638	5299	267	801	141	2.90E-22	[GI:2108229] [LN:LFU97348] [AC:U97348] [PN:basic surface protein] [FN:L-cystine transporter] [OR:Lactobacillus fermentum]	121
SPX2639	2639	5300	278	834	106	6.20E-12	[LN:G75297] [AC:G75297] [PN:conserved hypothetical protein] [GN:DR2233] [CL:probable phosphoesterase MJ0912:phosphoesterase core homology] [OR:Deinococcus radiodurans]	167
SPX2640	2640	5301	615	1845	1882	5.00E-256	[LN:UVRC_BACSU] [AC:P14951] [GN:UVRC] [OR:Bacillus subtilis] [DE:EXCINUCLEASE ABC SUBUNIT C] [SP:P14951]	104
SPX2641	2641	5302	241	723			NO-HIT	6
SPX2642	2642	5303	411	1233	95	3.90E-07	[LN:SPN250764] [AC:AJ250764] [PN:MurM protein] [GN:murM] [FN:serine/alanine adding enzyme]	120

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2643	2643	5304	407	1221	2088	3.50E-283	[OR:Streptococcus pneumoniae] [LN:SPN277484] [AC:AJ277484] [PN:beta-lactam resistance factor] [GN:fibA] [FN:putative role in peptidoglycan crosslinking]	152
SPX2644	2644	5305	260	780	482	7.10E-100	[OR:Streptococcus pneumoniae] [GI:7453516] [LN:AF157484] [AC:AF157484] [PN:tributyrin esterase] [GN:estA]	113
SPX2645	2645	5306	554	1662	581	1.80E-136	[OR:Lactococcus lactis subsp. lactis] [LN:H69884] [AC:H69884] [PN:conserved hypothetical protein ymfA] [GN:ymfA] [CL:conserved hypothetical protein MG139]	139
SPX2646	2646	5307	647	1941	529	3.20E-120	[OR:Bacillus subtilis] [LN:H69980] [AC:H69980] [PN:single-strand DNA-specific exonuclease homolog yrvE] [GN:yrvE]	113
SPX2647	2647	5308	741	2223	529	8.40E-150	[OR:Bacillus subtilis] [LN:H69980] [AC:H69980] [PN:single-strand DNA-specific exonuclease homolog yrvE] [GN:yrvE]	113
SPX2648	2648	5309	253	759	527	8.90E-103	[OR:Bacillus subtilis] [LN:GLNQ_BACST] [AC:P27675] [GN:GLNQ] [OR:Bacillus stearothermophilus] [DE:GLUTAMINE TRANSPORT ATP-BINDING PROTEIN GLNQ]	132
SPX2649	2649	5310	65	195	69	0.00063	[SP:P27675] [LN:E72756] [AC:E72756] [PN:hypothetical protein APE0042] [GN:APE0042]	92
SPX2650	2650	5311	73	219	92	9.10E-12	[OR:Aeropyrum pernix] [LN:G72510] [AC:G72510] [PN:hypothetical protein APE2061] [GN:APE2061]	92
SPX2651	2651	5312	265	795	422	2.40E-66	[OR:Aeropyrum pernix] [LN:PEB1_CAMJE] [AC:P45678] [GN:PEB1A] [OR:Campylobacter jejuni] [DE:MAJOR CELL-BINDING FACTOR PRECURSOR (CBF1) (PEB1)]	131

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
							[SP:P45678]	
SPX2652	2652	5313	226	678	400	3.80E-74	[LN:G81365] [AC:G81365] [PN:probable ABC-type amino-acid transporter permease protein Cj0920c [imported]] [GN:Cj0920c] [OR:Campylobacter jejuni]	144
SPX2653	2653	5314	220	660	234	2.00E-50	[LN:F69633] [AC:F69633] [PN:glutamine ABC transporter (membrane protein) glnP] [GN:glnP] [CL:histidine permease protein M] [OR:Bacillus subtilis]	145
SPX2654	2654	5315	401	1203	78	7.00E-08	[GI:1255667] [LN:PSEORF1] [AC:D84146] [PN:reductase] [GN:pahA] [OR:Pseudomonas aeruginosa] [SR:Pseudomonas aeruginosa (strain:PaK1) DNA]	136
SPX2655	2655	5316	294	882	1498	3.00E-201	“[LN:ALF_STRPN] [AC:O65944] [GN:FBA] [OR:Streptococcus pneumoniae] [EC:4.1.2.13] [DE:FRUCTOSE-BISPHOSPHATE ALDOLASE,] [SP:O65944]”	130
SPX2656	2656	5317	443	1329	2174	1.90E-297	[GI:4583524] [LN:AF140356] [AC:AF140356] [PN:VncS] [GN:vncS] [FN:putative histidine kinase/phosphatase] [OR:Streptococcus pneumoniae]	133
SPX2657	2657	5318	219	657	1089	1.00E-146	[GI:4583523] [LN:AF140356] [AC:AF140356] [PN:VncR] [GN:vncR] [FN:putative response regulator] [OR:Streptococcus pneumoniae]	123
SPX2658	2658	5319	460	1380	2273	0	[GI:5712669] [LN:AF140784] [AC:AF140784] [PN:Vexp3] [GN:vex3] [OR:Streptococcus pneumoniae]	91
SPX2659	2659	5320	216	648	1066	2.30E-141	[GI:5712668] [LN:AF140784] [AC:AF140784] [PN:Vexp2]	91

-continued

ORF NAME	NT ID	AA ID	AA LN	NT LN	SCORE	P-VALUE	DESCRIPTION	
SPX2660	2660	5321	426	1278	2070	2.40E-280	[GN:vex2] [OR:Streptococcus pneumoniae] [GI:5712667] [LN:AF140784] [AC:AF140784] [PN:Vexp1] [GN:vex1] [OR:Streptococcus pneumoniae]	91
SPX2661	2661	5322	85	255			NO-HIT	6

[0286]

SEQUENCE LISTING

The patent application contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US20070009900A1). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

What is claimed is:

1. An isolated nucleic acid encoding a *S. pneumoniae* polypeptide comprising SEQ ID NO: 5200.

2. An isolated nucleic acid sequence comprising SEQ ID NO: 2539.

3. An isolated nucleic acid sequence selected from the group consisting of:

a) SEQ ID NO: 2539,

b) a nucleic acid sequence having at least about 90% identity to SEQ ID NO: 2539,

c) a complement of SEQ ID NO: 2539,

d) a complement of a nucleic acid sequence having at least about 90% identity to SEQ ID NO: 2539, and

e) an RNA of a), b), c) or d), wherein U is substituted for T.

4. An isolated nucleic acid comprising a nucleotide sequence that hybridizes under high stringency conditions to at least one member selected from the group consisting of:

a) SEQ ID NO: 2539,

b) a complement of SEQ ID NO: 2539,

c) an RNA of a) or b), wherein U is substituted for T.

5. An isolated nucleic acid comprising at least about 40 consecutive nucleotides that hybridizes under high stringency conditions to SEQ ID NO: 2539.

6. A recombinant expression vector comprising the nucleic acid of claim 1 operably linked to a transcription regulatory element.

7. A recombinant expression vector comprising the nucleic acid of claim 2 operably linked to a transcription regulatory element.

8. A cell comprising the recombinant expression vector of claim 6.

9. A cell comprising the recombinant expression vector of claim 7.

* * * * *