



US 20120070881A1

(19) **United States**

(12) **Patent Application Publication**
Berka et al.

(10) **Pub. No.: US 2012/0070881 A1**

(43) **Pub. Date: Mar. 22, 2012**

(54) **BACILLUS LICHENIFORMIS
CHROMOSOME**

(75) Inventors: **Randy Berka**, Davis, CA (US);
Michael Rey, Davis, CA (US);
Preethi Ramaiya, Fairfield, CA
(US); **Jens Tonne Andersen**,
Naerum (DK); **Michael Dolberg**
Rasmussen, Vallensbaek (DK);
Peter Bjarke Olsen, Copenhagen
(DK)

(73) Assignees: **NOVOZYMES A/S**, Bagsvaerd
(DK); **NOVOZYMES, INC.**,
Davis, CA (US)

(21) Appl. No.: **13/305,556**

(22) Filed: **Nov. 28, 2011**

Related U.S. Application Data

(62) Division of application No. 12/972,306, filed on Dec. 17, 2010, which is a division of application No. 12/322,974, filed on Feb. 9, 2009, now Pat. No. 7,863,032, which is a division of application No. 10/983,128, filed on Nov. 5, 2004, now Pat. No. 7,494,798.

(60) Provisional application No. 60/572,403, filed on May 18, 2004, provisional application No. 60/561,059, filed on Apr. 8, 2004, provisional application No. 60/535,988, filed on Jan. 9, 2004.

Publication Classification

(51) **Int. Cl.**
C12N 1/21 (2006.01)
C12N 15/63 (2006.01)
C12N 15/60 (2006.01)

(52) **U.S. Cl.** **435/252.31**; 536/23.2; 435/320.1;
435/252.35; 435/252.33; 435/252.34

(57) **ABSTRACT**

The present invention relates to an isolated polynucleotide of the complete chromosome of *Bacillus licheniformis*. The present invention also relates to isolated genes of the chromosome of *Bacillus licheniformis* which encode biologically active substances and to nucleic acid constructs, vectors, and host cells comprising the genes as well as methods for producing biologically active substances encoded by the genes and to methods of using the isolated genes of the complete chromosome of *Bacillus licheniformis*.

BACILLUS LICHENIFORMIS CHROMOSOME

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a divisional of U.S. application Ser. No. 12/972,306, filed Dec. 17, 2010, which is a divisional of U.S. application Ser. No. 12/322,974, filed Feb. 9, 2009, now U.S. Pat. No. 7,863,032, which is a divisional of U.S. application Ser. No. 10/983,128, filed Nov. 5, 2004, now U.S. Pat. No. 7,494,798, which claims the benefit of U.S. Provisional Application No. 60/535,988, filed Jan. 9, 2004, U.S. Provisional Application No. 60/561,059, filed Apr. 8, 2004, and U.S. Provisional Application No. 60/572,403, filed May 18, 2004. The content of these applications is hereby incorporated by reference as if it was set forth in full below.

REFERENCE TO A SEQUENCE LISTING

[0002] This application contains a Sequence Listing in computer readable form, which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0003] 1. Field of the Invention

[0004] The present invention relates to an isolated polynucleotide molecule comprising the complete chromosome of *Bacillus licheniformis*. The present invention also relates to features (genes) of the complete chromosomal DNA molecule of *Bacillus licheniformis* which encode biologically active substances and to nucleic acid constructs, vectors, and host cells comprising the features as well as methods for producing biologically active substances encoded by the features and to methods of using the isolated features derived from the complete chromosomal DNA molecule of *Bacillus licheniformis*.

[0005] 2. Description of the Related Art

[0006] Microbes, which make up most of the earth's biomass, have evolved for some 3.8 billion years. They are found in virtually every environment, surviving and thriving in extremes of heat, cold, radiation, pressure, salt, acidity, and darkness. Often in these environments, no other forms of life are found and the only nutrients come from inorganic matter. The diversity and range of their environmental adaptations indicate that microbes long ago "solved" many problems for which scientists are still actively seeking solutions. The value in determining the complete genome sequence of microbes is that it provides a detailed blueprint for the organism revealing all of the biochemical pathways, substrates, intermediates, and end products as well as regulatory networks, and evolutionary relationships to other microbes. A complete manifest of proteins, both structural and catalytic, is encoded as a list of features in the DNA molecule comprising the genome, as well as their likely cellular location.

[0007] Knowledge about the enormous range of microbial capacities has broad and far-reaching implications for environmental, energy, health, and industrial applications, such as cleanup of toxic-waste, production of novel therapeutic and preventive agents (drugs and vaccines), energy generation and development of renewable energy sources, production of chemical catalysts, reagents, and enzymes to improve efficiency of industrial processes, management of environmental carbon, nitrogen and nutrient cycling, detection of disease-causing organisms and monitoring of the safety of food and

water supplies, use of genetically altered bacteria as living sensors (biosensors) to detect harmful chemicals in soil, air, or water, and understanding of specialized systems used by microbial cells to live in natural environments.

[0008] *Bacillus licheniformis* is a gram positive spore-forming bacterium that is widely distributed as a saprophytic organism in the environment. Unlike most other bacilli that are predominantly aerobic, *Bacillus licheniformis* is a facultative anaerobe which may allow it to grow in additional ecological niches. This species produces a diverse assortment of extracellular enzymes that are believed to contribute to the process of nutrient cycling in nature (Claus, D. and Berkeley, R. C. W., 1986, In *Bergey's Manual of Systematic Bacteriology*, Vol. 2., eds. Sneath, P. H. A. et al., Williams and Wilkins Co., Baltimore, Md., pp. 1105-1139). Certain *Bacillus licheniformis* isolates are capable of denitrification, however, the relevance of this characteristic to environmental denitrification may be small since the species generally persists in soil as endospores (Alexander, M., 1977, *Introduction to Soil Microbiology*. John Wiley and Sons, Inc., New York).

[0009] There are numerous industrial and agricultural uses for *Bacillus licheniformis* and its extracellular products. The species has been used for decades in the manufacture of industrial enzymes including several proteases, α -amylase, penicillinase, pentosanase, cycloglucosyltransferase, β -mannanase, and several pectinolytic enzymes, owing largely to its ability to secrete sizeable amounts of degradative enzymes. *Bacillus licheniformis* is also used to produce peptide antibiotics such as bacitracin and proticin, in addition to a number of specialty chemicals such as citric acid, inosine, inosinic acid, and poly- γ -glutamic acid. The proteases from *Bacillus licheniformis* are used in the detergent industry as well as for dehairing and batting of leather (Eveleigh, D. E., 1981, *Scientific American* 245, 155-178). Amylases from *Bacillus licheniformis* are deployed for the hydrolysis of starch, desizing of textiles, and sizing of paper (Erickson, R. J., 1976, In *Microbiology*, ed. Schlesinger, D. (Am. Soc. Microbiol., Washington, D.C.), pp. 406-419.). Certain strains of *Bacillus licheniformis* have shown efficacy to destroy fungal pathogens affecting maize, grasses, and vegetable crops (U.S. Pat. No. 5,589,381; U.S. Pat. No. 5,665,354). As an endospore-forming bacterium, the ability of the organism to survive under unfavorable environmental conditions may enhance its potential as a natural control agent.

[0010] *Bacillus licheniformis* can be differentiated from other bacilli on the basis of metabolic and physiological tests (Logan, N. A. and Berkeley, R. C. W., 1981, In *The Aerobic Endospore-Forming Bacteria Classification and Identification*, eds. Berkeley, R. C. W. and Goodfellow, M., Academic Press, Inc., London, pp. 106-140; O'Donnell, A. G., Norris, J. R., Berkeley, R. C. W., Claus, D., Kanero, T., Logan, N. A., and Nozaki, R., 1980, *Internat. J. Systematic Bacteriol.* 30: 448-459). However, biochemical and phenotypic characteristics may be ambiguous among closely related species. Lapidus et al. (Lapidus, A., Galleron, N., Andersen, J. T., Jørgensen, P. L., Ehrlich, S. D., and Sorokin, A., 2002, *FEMS Microbiol. Lett.* 209: 23-30) recently constructed a physical map of the *Bacillus licheniformis* chromosome using a PCR approach, and established a number of regions of co-linearity where gene content and organization were ostensibly conserved with the *Bacillus subtilis* chromosome.

[0011] It would be advantageous to the art to have available the complete primary structure of the chromosomal DNA molecule of the *Bacillus licheniformis* type strain ATCC

14580. With the complete chromosome data in hand, it should be possible to do comparative genomics and proteomics studies that can lead to improved industrial strains as well as to a better understanding of genome evolution among closely-related bacilli in the *subtilis-licheniformis* group.

[0012] It is an object of the present invention to provide an isolated polynucleotide with the sequence of the complete chromosome of *Bacillus licheniformis*.

SUMMARY OF THE INVENTION

[0013] The present invention relates to an isolated polynucleotide of the complete chromosomal DNA molecule of *Bacillus licheniformis* ATCC 14580 having the nucleotide sequence of SEQ ID NO: 1.

[0014] The present invention also relates to isolated features (genes) of the complete chromosomal DNA molecule of *Bacillus licheniformis* ATCC 14580 encoding biologically active substances, selected from the group consisting of:

[0015] (a) a gene comprising a nucleotide sequence having at least 60% identity with any of the polynucleotides of SEQ ID NOs: 2-4198; and

[0016] (b) a gene comprising a nucleotide sequence which hybridizes under at least medium stringency conditions with any of the polynucleotides of SEQ ID NOs: 2-4198, or a complementary strand thereof.

[0017] The present invention also relates to biologically active substances encoded by the isolated genes, and nucleic acid constructs, vectors, and host cells comprising the genes.

[0018] The present invention also relates to methods for producing such substances having biological activity comprising (a) cultivating a recombinant host cell comprising a nucleic acid construct comprising a gene encoding the biologically active substance under conditions suitable for production of the biologically active substance; and (b) recovering the biologically active substance.

[0019] The present invention also relates to methods for monitoring differential expression of a plurality of genes in a first bacterial cell relative to expression of the same genes in one or more second bacterial cells, comprising:

[0020] (a) adding a mixture of detection reporter-labeled nucleic acids isolated from the bacterial cells to a substrate containing an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198, under conditions where the detection reporter-labeled nucleic acids hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array, wherein the nucleic acids from the first bacterial cell and the one or more second bacterial cells are labeled with a first detection reporter and one or more different second detection reporters, respectively; and

[0021] (b) examining the array under conditions wherein the relative expression of the genes in the bacterial cells is determined by the observed detection signal of each spot on the array in which (i) the *Bacillus licheniformis* genes on the array that hybridize to the nucleic acids obtained from either the first or the one or more second bacterial cells produce a distinct first detection signal or one or more second detection signals, respectively, and (ii) the *Bacillus licheniformis* genes on the array that hybridize to the nucleic acids obtained from both the first and one or more second bacterial produce a distinct combined detection signal.

[0022] The present invention also relates to methods for isolating a gene encoding an enzyme, comprising:

[0023] (a) adding a mixture of labeled first nucleic acid probes, isolated from a microbial strain cultured on medium without an inducing substrate, and labeled second nucleic acid probes, isolated from the microbial strain cultured on medium with the inducing substrate, to an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198, under conditions where the labeled nucleic acid probes hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array, wherein the first nucleic acid probes are labeled with a first reporter and the second nucleic acid probes are labeled with a second reporter;

[0024] (b) examining the array under conditions wherein the relative expression of the genes of the microbial strain is determined by the observed hybridization reporter signal of each spot on the array in which (i) the *Bacillus licheniformis* genes on the array that hybridize to the first nucleic acid probes produce a distinct first hybridization reporter signal or the second nucleic acid probes produce a distinct second hybridization reporter signal, and (ii) the *Bacillus licheniformis* genes on the array that hybridize to both the first and second nucleic acid probes produce a distinct combined hybridization reporter signal; and

[0025] (c) isolating a gene from the microbial strain that encodes an enzyme that degrades or converts the substrate.

[0026] The present invention also relates to genes isolated by such methods and nucleic acid constructs, vectors, and host cells containing the genes.

DEFINITIONS

[0027] Biologically active substance: The term “biologically active substance” is defined herein as any substance which is encoded by a single gene or a series of genes (contiguous or non-contiguous) composing a biosynthetic or metabolic pathway or operon or may be the direct or indirect result of the product of a single gene or products of a series of genes of the *Bacillus licheniformis* chromosome. Such substances include, but are not limited to, biopolymers, metabolites, and cellular structures and components (e.g., ribosome, flagella, etc.). For purposes of the present invention, biological activity is determined according to procedures known in the art such as those described by Carpenter and Sabatini, 2004, *Nature* 5: 11-22; Sordie et al., 2003, *Proceedings of the National Academy of Sciences USA* 100: 11964-11969; Braun and LaBaer, 2003, *TRENDS in Biotechnology* 21: 383-388; and Kaberdin and McDowall, 2003, *Genome Research* 13: 1961-1965.

[0028] In the methods of the present invention, the biopolymer may be any biopolymer. The term “biopolymer” is defined herein as a chain (or polymer) of identical, similar, or dissimilar subunits (monomers). The biopolymer may be, but is not limited to, a nucleic acid, polyamine, polyol, polypeptide (or polyamide), or polysaccharide.

[0029] In a preferred aspect, the biopolymer is a polypeptide. The polypeptide may be any polypeptide having a biological activity of interest. The term “polypeptide” is not meant herein to refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and proteins. The term “polypeptide” also encompasses naturally occurring allelic variations.

[0030] In a preferred aspect, the polypeptide is an antibody, antigen, antimicrobial peptide, enzyme, growth factor, hormone, immunodilator, neurotransmitter, receptor, reporter protein, structural protein, transcription factor, and transporter.

[0031] In a more preferred aspect, the polypeptide is an oxidoreductase, transferase, hydrolase, lyase, isomerase, or ligase. In a most preferred aspect, the polypeptide is an alpha-glucosidase, aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, gluco-cerebrosidase, alpha-glucosidase, beta-glucosidase, invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phospholipase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, urokinase, or xylanase.

[0032] In another preferred aspect, the polypeptide is a collagen or gelatin.

[0033] In another preferred aspect, the biopolymer is a polysaccharide. The polysaccharide may be any polysaccharide, including, but not limited to, a mucopolysaccharide (e.g., heparin and hyaluronic acid) and a nitrogen-containing polysaccharide (e.g., chitin). In a more preferred aspect, the polysaccharide is hyaluronic acid (hyaluronan).

[0034] In the methods of the present invention, the metabolite may be any metabolite. The metabolite may be encoded by one or more genes, such as a biosynthetic or metabolic pathway. The term "metabolite" encompasses both primary and secondary metabolites. Primary metabolites are products of primary or general metabolism of a cell, which are concerned with energy metabolism, growth, and structure. Secondary metabolites are products of secondary metabolism (see, for example, R. B. Herbert, *The Biosynthesis of Secondary Metabolites*, Chapman and Hall, New York, 1981).

[0035] The primary metabolite may be, but is not limited to, an amino acid, fatty acid, nucleoside, nucleotide, sugar, triglyceride, or vitamin.

[0036] The secondary metabolite, may be, but is not limited to, an alkaloid, coumarin, flavonoid, polyketide, quinine, steroid, peptide, or terpene. In a preferred aspect, the secondary metabolite is an antibiotic, antifedant, attractant, bactericide, fungicide, hormone, insecticide, or rodenticide.

[0037] isolated biologically active substance: The term "isolated biologically active substance" is defined herein as a substance which is at least about 20% pure, preferably at least about 40% pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by SDS-PAGE, HPLC, capillary electrophoresis, or any other method used in the art.

[0038] Substantially pure biologically active substance or pure biologically active substance: The term "substantially pure biologically active substance" is defined herein as a biologically active substance preparation which contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5% by weight, more preferably at most 4%, at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other material with which it is natively associated. It is, therefore, preferred that the substantially pure biologically active substance is at least 92% pure, preferably at least 94% pure, more preferably at least 95% pure, more preferably at least 96% pure, more preferably at least

96% pure, more preferably at least 97% pure, even more preferably at least 98% pure, most preferably at least 99%, and even most preferably at least 99.5% pure by weight of the total material present in the preparation. The term "pure biologically active substance" is defined as a biologically active substance preparation which contains no other material with which it is natively associated.

[0039] The biologically active substances of the present invention are preferably in a substantially pure form. In particular, it is preferred that the biologically active substances are in "essentially pure form", i.e., that the biologically active substance preparation is essentially free of other material with which it is natively associated. This can be accomplished, for example, by preparing the biologically active substance by means of well-known recombinant methods or by classical purification methods.

[0040] Identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "identity".

[0041] For purposes of the present invention, the degree of identity between two amino acid sequences is determined by the Smith-Waterman Protein method for the Genematcher2, as implemented by Paracel Inc. (Pasadena, Calif.), or the BLASTP method as described by Altschul et al., 1990, *Journal of Molecular Biology* 215: 403-410.

[0042] For purposes of the present invention, the degree of identity between two nucleotide sequences is determined by the Smith Waterman nucleotide method for the Genematcher2 or BLASTN for the BlastMachine as implemented by Paracel inc.

[0043] Polypeptide Fragment: The term "polypeptide fragment" is defined herein as a polypeptide, which retains biological activity, having one or more amino acids deleted from the amino and/or carboxyl terminus of a polypeptide encoded by any of the genes of the present invention, i.e., polypeptides of SEQ ID NOs: 4199-8395. Preferably, a fragment contains at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97% of the amino acid residues of the mature encoded polypeptide product.

[0044] Subsequence: The term "subsequence" is defined herein as a polynucleotide comprising a nucleotide sequence of any of SEQ ID NOs: 2-4198 except that one or more nucleotides have been deleted from the 5' and/or 3' end. Preferably, a subsequence contains at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97% of the nucleotides of any of the genes of the present invention.

[0045] Allelic variant: The term "allelic variant" denotes herein any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

[0046] Substantially pure polynucleotide or pure polynucleotide: The term "substantially pure polynucleotide" as used herein refers to a polynucleotide preparation free of other extraneous or unwanted nucleotides and is in a form suitable for use within genetically engineered production systems. Thus, a substantially pure polynucleotide contains at most 10%, preferably at most 8%, more preferably at most

6%, more preferably at most 5%, more preferably at most 4%, more preferably at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polynucleotide material with which it is natively associated. A substantially pure polynucleotide may, however, include naturally occurring 5' and 3' untranslated regions, such as promoters and terminators. It is preferred that the substantially pure polynucleotide is at least 92% pure, preferably at least 94% pure, more preferably at least 95% pure, more preferably at least 96% pure, more preferably at least 96% pure, more preferably at least 97% pure, more preferably at least 98% pure, most preferably at least 99%, and even most preferably at least 99.5% pure by weight. The polynucleotides of the present invention are preferably in a substantially pure form. In particular, it is preferred that the polynucleotides disclosed herein are in "essentially pure form", i.e., that the polynucleotide preparation is essentially free of other polynucleotide material with which it is natively associated. The term "pure polynucleotide" is defined as a polynucleotide preparation which contains no other material with which it is natively associated.

[0047] Nucleic acid construct: The term "nucleic acid construct" as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

[0048] Control sequence: The term "control sequences" is defined herein to include all components, which are necessary or advantageous for the expression of a biologically active substance of the present invention. Each control sequence may be native or foreign to the polynucleotide encoding the substance. Such control sequences include, but are not limited to, a leader, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a biologically active substance.

[0049] Operably linked: The term "operably linked" as used herein refers to a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of the DNA sequence, such that the control sequence directs the expression of a biologically active substance.

[0050] Coding sequence: When used herein the term "coding sequence" is intended to cover a nucleotide sequence, which directly specifies the amino acid sequence of its protein product. The boundaries of the coding sequence are generally determined by an open reading frame, which usually begins with the ATG start codon or alternative start codons such as GTG and TTG.

[0051] Expression: The term "expression" includes any step involved in the production of a biologically active substance including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0052] Expression vector: The term "expression vector" herein covers a DNA molecule, linear or circular, that com-

prises a segment encoding a biologically active substance of the invention, and which is operably linked to additional segments that provide for its transcription.

[0053] Host cell: The term "host cell", as used herein, includes any cell type which is susceptible to transformation, transfection, conjugation, electroporation, etc. with a nucleic acid construct, plasmid, or vector.

DETAILED DESCRIPTION OF THE INVENTION

[0054] *Bacillus licheniformis* Chromosome and Features (Genes) Thereof

[0055] The present invention relates to an isolated polynucleotide of the complete chromosomal DNA molecule of *Bacillus licheniformis* ATCC 14580 having the nucleotide sequence of SEQ ID NO: 1. *Bacillus licheniformis* ATCC 14580, consists of a circular molecule of 4,222,336 base pairs with a mean G+C content of 46.2%. The chromosome contains 4208 predicted protein-coding genes (SEQ ID NOs: 2-4198) with an average size of 873 bp, 7 rRNA operons, and 72 tRNA genes. The deduced amino acid sequences of the 4208 predicted protein-coding genes are shown in SEQ ID NOs: 4199-8395. SEQ ID NO: 4210 corresponds to SEQ ID NO: 2, SEQ ID NO: 4211 corresponds to SEQ ID NO: 3, SEQ ID NO: 4212 corresponds to SEQ ID NO: 4, etc. The predicted functions of the 4208 gene products are shown in Table 1.

[0056] The *Bacillus licheniformis* chromosome possesses regions that are markedly co-linear with the chromosomes of *Bacillus subtilis* and *Bacillus halodurans*, and approximately 80% of the predicted genes have *Bacillus subtilis* orthologues.

[0057] The present invention also relates to isolated features (genes) of the complete chromosomal DNA molecule of *Bacillus licheniformis* ATCC 14580 encoding biologically active substances, selected from the group consisting of:

[0058] (a) a gene comprising a nucleotide sequence having at least 60% identity with any of the nucleotide sequences of SEQ ID NOs: 2-4198; and

[0059] (b) a gene comprising a nucleotide sequence which hybridizes under at least medium stringency conditions with any of the genes of SEQ ID NOs: 2-4198, or a complementary strand thereof.

[0060] In a first aspect, the present invention relates to isolated genes, which have a degree of identity to the nucleotide sequences of any of SEQ ID NOs: 2-4198 of at least about 60%, preferably at least about 65%, more preferably at least about 70%, more preferably at least about 75%, more preferably at least about 80%, more preferably at least about 85%, even more preferably at least about 90%, most preferably at least about 95%, and even most preferably at least about 97%, which encode biologically active substances having a particular biological activity (hereinafter "homologous biologically active substances").

[0061] In a second aspect, the present invention relates to isolated genes comprising nucleotide sequences which hybridize under very low stringency conditions, preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with any of (i) the genes of SEQ ID NOs: 2-4198, or subsequences thereof, or (ii) complementary strands thereof (J. Sambrook, E. F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, N.Y.).

Subsequences of SEQ ID NOs: 2-4198 may be at least 100 nucleotides or preferably at least 200 nucleotides. Moreover, the subsequences may encode fragments of a gene product which have biological activity. The biologically active substances may also be biologically active allelic variants of the biologically active substances.

[0062] The nucleotide sequences of SEQ ID NOs: 2-4198 or subsequences thereof, as well as the amino acid sequences of SEQ ID NOs: 4199-8395 or fragments thereof, may be used to design nucleic acid probes to identify and clone DNA encoding biologically active substances from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic DNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 14, preferably at least 25, more preferably at least 35 nucleotides in length, such as at least 70 nucleotides in length. It is preferred, however, that the nucleic acid probes are at least 100 nucleotides in length. For example, the nucleic acid probes may be at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length. Even longer probes may be used, e.g., nucleic acid probes which are at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin). Such probes are encompassed by the present invention.

[0063] A genomic DNA library prepared from such other organisms may, therefore, be screened for DNA which hybridizes with the probes described above and which encodes a biologically active substance. Genomic DNA from such other organisms may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA which is homologous with any of SEQ ID NOs: 2-4198 or subsequences thereof, the carrier material is used in a Southern blot.

[0064] For purposes of the present invention, hybridization indicates that a polynucleotide hybridizes to a labeled gene having the nucleotide sequence shown in any of SEQ ID NOs: 2-4198, complementary strands thereof, or subsequences thereof, under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using X-ray film.

[0065] In a preferred aspect, the nucleic acid probe is any of the genes of SEQ ID NOs: 2-4198, or subsequences thereof. In another preferred aspect, the nucleic acid probe is the mature coding region of any of the genes of SEQ ID NOs: 2-4198. In another preferred aspect, the nucleic acid probe is the gene of any of SEQ ID NOs: 2-4198 contained in *Bacillus licheniformis* ATCC 14580, in another preferred aspect, the nucleic acid probe is the mature coding region of any of the genes of SEQ ID NOs: 2-4198 contained in *Bacillus licheniformis* ATCC 14580.

[0066] For long probes of at least 100 nucleotides in length, very low to very high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm

DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures.

[0067] For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2×SSC, 0.2% SDS preferably at least at 45° C. (very low stringency), more preferably at least at 50° C. (low stringency), more preferably at least at 55° C. (medium stringency), more preferably at least at 60° C. (medium-high stringency), even more preferably at least at 65° C. (high stringency), and most preferably at least at 70° C. (very high stringency).

[0068] For short probes which are about 14 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at about 5° C. to about 10° C. below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1×Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures.

[0069] For short probes which are about 14 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6×SSC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6×SSC at 5° C. to 10° C. below the calculated T_m .

[0070] Under salt-containing hybridization conditions, the effective T_m is what controls the degree of identity required between the probe and the filter bound DNA for successful hybridization. The effective T_m may be determined using the formula below to determine the degree of identity required for two DNAs to hybridize under various stringency conditions.

$$\text{Effective } T_m = 81.5 + 16.6(\log M[\text{Na}^+]) + 0.41(\%G+C) - 0.72(\%\text{formamide})$$

[0071] The % G+C content of any of the genes of SEQ ID NOs: 2-4198 can easily be determined. For medium stringency, for example, the concentration of formamide is 35% and the Na^+ concentration for 5×SSPE is 0.75 M. Applying this formula to these values, the Effective T_m in ° C. can be calculated. Another relevant relationship is that a 1% mismatch of two DNAs lowers the T_m 1.4° C. To determine the degree of identity required for two DNAs to hybridize under medium stringency conditions at 42° C., the following formula is used:

$$\% \text{Homology} = 100 - [(\text{Effective } T_m - \text{Hybridization Temperature}) / 1.4]$$

[0072] Applying this formula, the degree of identity required for two DNAs to hybridize under medium stringency conditions at 42° C. can be calculated.

[0073] Similar calculations can be made under other stringency conditions, as defined herein.

[0074] The present invention also relates to isolated polynucleotides obtained by (a) hybridizing a population of DNA under very low, low, medium, medium-high, high, or very high stringency conditions with any of (i) the genes of SEQ ID NOs: 2-4198, or subsequences thereof, or (ii) complementary strands thereof; and (b) isolating the hybridizing polynucleotide from the population of DNA. In a preferred aspect, the hybridizing polynucleotide encodes a polypeptide of any of SEQ ID NOs: 2-4198, or homologous polypeptides thereof.

[0075] In a third aspect, the present invention relates to isolated polypeptides having amino acid sequences which have a degree of identity to any of SEQ ID NOs: 4199-8395 of at least about 60%, preferably at least about 65%, more preferably at least about 70%, more preferably at least about 75%, more preferably at least about 80%, more preferably at least about 85%, even more preferably at least about 90%, most preferably at least about 95%, and even most preferably at least about 97%, which have biological activity (hereinafter "homologous polypeptides"). In a preferred aspect, the homologous polypeptides have an amino acid sequence which differs by ten amino acids, preferably by five amino acids, more preferably by four amino acids, even more preferably by three amino acids, most preferably by two amino acids, and even most preferably by one amino acid from the amino acid sequences of SEQ ID NOs: 4199-8395.

[0076] The polypeptides of the present invention preferably comprise the amino acid sequence of any of SEQ ID NOs: 4199-8395 or an allelic variant thereof; or a fragment thereof that has biological activity. In a more preferred aspect, the polypeptides of the present invention comprise the amino acid sequence of any of SEQ ID NOs: 4199-8395. In another preferred aspect, the polypeptides of the present invention comprise the mature polypeptide region of any of SEQ ID NOs: 4199-8395, or an allelic variant thereof; or a fragment thereof that has biological activity. In another preferred aspect, the polypeptides of the present invention comprise the mature polypeptide region of any of SEQ ID NOs: 4199-8395. In another preferred aspect, the polypeptides of the present invention consist of the amino acid sequence of any of SEQ ID NOs: 4199-8395 or an allelic variant thereof; or a fragment thereof that has biological activity. In another preferred aspect, the polypeptides of the present invention consist of the amino acid sequence of any of SEQ ID NOs: 4199-8395. In another preferred aspect, the polypeptides consist of the mature polypeptide region of any of SEQ ID NOs: 4199-8395 or an allelic variant thereof; or a fragment thereof that has biological activity. In another preferred aspect, the polypeptides consist of the mature polypeptide region of any of SEQ ID NOs: 4199-8395.

[0077] In a fourth aspect, the present invention relates to isolated substances having biological activity which are encoded by polynucleotides which hybridize, as described above, under very low stringency conditions, preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with which hybridize under very low stringency conditions, preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with any of (i) the genes of SEQ ID NOs: 2-4198, or subsequences thereof, or (ii) complementary strands thereof. A subsequence of any of SEQ ID NOs: 2-4198 may be at least 100 nucleotides or preferably at least 200 nucleotides. Moreover, the subsequence may encode a fragment, e.g., a polypeptide fragment, which has biological activity.

Nucleic Acid Constructs

[0078] The present invention also relates to nucleic acid constructs comprising an isolated gene or isolated genes (e.g., operon) of the present invention operably linked to one or

more control sequences which direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

[0079] An isolated gene(s) of the present invention may be manipulated in a variety of ways to provide for production of a biologically active substance encoded directly or indirectly by the gene(s). Manipulation of the nucleotide sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleotide sequences utilizing recombinant DNA methods are well known in the art.

[0080] The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of the gene(s) encoding the biologically active substance. The promoter sequence contains transcriptional control sequences which mediate the expression of the biologically active substance. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides or biologically active substances either homologous or heterologous to the host cell.

[0081] Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli* lac operon, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis* xylA and xylB genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), as well as the tac promoter (DeBoer et al., 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74-94; and in Sambrook et al., 1989, *supra*.

[0082] The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the gene encoding the biologically active substance. Any terminator which is functional in the host cell of choice may be used in the present invention.

[0083] The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs

the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

[0084] Effective signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (nprT, nprS, nprM), and *Bacillus subtilis* prsA. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

[0085] The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for *Bacillus subtilis* alkaline protease (aprE) and *Bacillus subtilis* neutral protease (nprT).

[0086] Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

[0087] It may also be desirable to add regulatory sequences which allow the regulation of the expression of a biologically active substance relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the lac, lac, and trp operator systems. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleotide sequence encoding the biologically active substance would be operably linked with the regulatory sequence.

Expression Vectors

[0088] The present invention also relates to recombinant expression vectors comprising an isolated gene of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites. Alternatively, a gene of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0089] The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of a gene of the present invention. The choice of the vector will typically depend on the compatibility of the

vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

[0090] The vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

[0091] The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

[0092] Examples of bacterial selectable markers are the dal genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance.

[0093] The vectors of the present invention preferably contain an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

[0094] For integration into the host cell genome, the vector may rely on portions of the sequence of the gene or any other element of the vector for integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleotide sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleotide sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome (s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleotides, such as 100 to 10,000 base pairs, preferably 400 to 10,000 base pairs, and most preferably 800 to 10,000 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleotide sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0095] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication which functions in a cell. The term "origin of replication" or "plasmid replicator" is defined herein as a sequence that enables a plasmid or vector to replicate in vivo. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

[0096] More than one copy of a gene of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the gene can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with a gene of the present invention where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the gene of the present invention, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0097] The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g. Sambrook et al., 1989, *supra*).

Host Cells

[0098] The present invention also relates to recombinant host cells, comprising an isolated gene of the present invention, where the host cells are advantageously used in the recombinant production of a biologically active substance encoded by the gene. A vector comprising a gene of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the biologically active substance and its source.

[0099] The host cell may be any unicellular microorganism, e.g., a prokaryote.

[0100] Useful unicellular cells are bacterial cells such as gram positive bacteria including, but not limited to, a *Bacillus* cell, e.g., *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus fastidiosus*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentils*, *Bacillus licheniformis*, *Bacillus macerans*, *Bacillus megaterium*, *Bacillus methanolicus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis*; or a *Streptomyces* e.g., *Streptomyces lividans* and *Streptomyces murinus*, or gram negative bacteria such as *E. coli* and *Pseudomonas* sp. In a preferred aspect, the bacterial host cell is a *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, or *Bacillus subtilis* cell. In another preferred aspect, the *Bacillus* cell is an alkalophilic *Bacillus*.

[0101] The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), using competent cells (see, e.g., Young and Spizizen, 1961, *Journal of Bacteriology* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

Methods of Production

[0102] The present invention also relates to methods for producing a biologically active substance of the present invention comprising (a) cultivating a strain, which in its wild-type form is capable of producing the biologically active

substance, under conditions conducive for production of the biologically active substance; and (b) recovering the biologically active substance. Preferably, the strain is of the genus *Bacillus*, and more preferably *Bacillus licheniformis*.

[0103] The present invention also relates to methods for producing a biologically active substance of the present invention comprising (a) cultivating a host cell under conditions conducive for production of the biologically active substance; and (b) recovering the biologically active substance.

[0104] The present invention also relates to methods for producing a biologically active substance of the present invention comprising (a) cultivating a host cell under conditions conducive for production of the biologically active substance, wherein the host cell comprises a mutant polynucleotide comprising at least one mutation in the coding region of any of SEQ ID NOs: 2-4198, wherein the mutant polynucleotide encodes a biologically active substance which consists of SEQ ID NOs: 4199-8395, respectively, and (b) recovering the biologically active substance.

[0105] In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the biologically active substance using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the biologically active substance to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). If the biologically active substance is secreted into the nutrient medium, the biologically active substance can be recovered directly from the medium. If the biologically active substance is not secreted, it can be recovered from cell lysates.

[0106] The biologically active substances may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of an enzyme.

[0107] The resulting biologically active substances may be recovered by methods known in the art. For example, the biologically active substances may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

[0108] The biologically active substances of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

Plants

[0109] The present invention also relates to a transgenic plant, plant part, or plant cell which has been transformed with a gene encoding a biologically active substance of the

present invention so as to express and produce the biologically active substance in recoverable quantities. The biologically active substance may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the recombinant biologically active substance may be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

[0110] The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as *festuca*, *lolium*, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn).

[0111] Examples of dicot plants are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family *Brassicaceae*), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*.

[0112] Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers as well as the individual tissues comprising these parts, e.g., epidermis, mesophyll, parenchyme, vascular tissues, meristems. In the present context, also specific plant cell compartments, such as chloroplast, apoplast, mitochondria, vacuole, peroxisomes and cytoplasm are considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part. Likewise, plant parts such as specific tissues and cells isolated to facilitate the utilisation of the invention are also considered plant parts e.g. embryos, endosperms, aleurone and seeds coats.

[0113] Also included within the scope of the present invention are the progeny of such plants, plant parts, and plant cells.

[0114] The transgenic plant or plant cell expressing a biologically active substance of the present invention may be constructed in accordance with methods known in the art. Briefly, the plant or plant cell is constructed by incorporating one or more expression constructs encoding a biologically active substance of the present invention into the plant host genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

[0115] The expression construct is conveniently a nucleic acid construct which comprises a gene encoding a biologically active substance of the present invention operably linked with appropriate regulatory sequences required for expression of the nucleotide sequence in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

[0116] The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences is determined, for example, on the basis of when, where, and how the biologically active substance is desired to be expressed. For instance, the expression of the gene encoding a biologically active substance of the present invention may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague et al., 1988, *Plant Physiology* 86: 506.

[0117] For constitutive expression the 35S-CaMV, the maize ubiquitin 1 and the rice actin 1 promoter may be used (Franck et al., 1980, *Cell* 21: 285-294, Christensen A H, Sharrock R A and Quail, 1992, *Plant Mo. Biol.* 18: 675-689; Zhang W, McElroy D. and Wu R., 1991, *Plant Cell* 3: 1155-1165). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards and Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito et al., 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu et al., 1998, *Plant and Cell Physiology* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad et al., 1998, *Journal of Plant Physiology* 152: 708-711), a promoter from a seed oil body protein (Chen et al., 1998, *Plant and Cell Physiology* 39: 935-941), the storage protein napA promoter from *Brassica napus*, or any other seed specific promoter known in the art, e.g., as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the rbcS promoter from rice or tomato (Kyo-zuka et al., 1993, *Plant Physiology* 102: 991-1000), the chloro-*rella* virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Molecular Biology* 26: 85-93), or the aldP gene promoter from rice (Kagaya et al., 1995, *Molecular and General Genetics* 248: 668-674), or a wound inducible promoter such as the potato pint promoter (Xu et al., 1993, *Plant Molecular Biology* 22: 573-588). Likewise, the promoter may be inducible by abiotic treatments such as temperature, drought or alterations in salinity or induced by exogenously applied substances that activate the promoter, e.g., ethanol, oestrogens, plant hormones like ethylene, abscisic acid and gibberellic acid and heavy metals.

[0118] A promoter enhancer element may also be used to achieve higher expression of the enzyme in the plant. For instance, the promoter enhancer element may be an intron which is placed between the promoter and the nucleotide sequence encoding a biologically active substance of the present invention. For instance, Xu et al., 1993, supra disclose the use of the first intron of the rice actin 1 gene to enhance expression.

[0119] The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

[0120] The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including *Agrobacterium*-mediated transformation, virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation (Gasser et al., 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto et al., 1989, *Nature* 338: 274).

[0121] Presently, *Agrobacterium tumefaciens*-mediated gene transfer is the method of choice for generating transgenic dicots (for a review, see Hooykas and Schilperoot, 1992, *Plant Molecular Biology* 19: 15-38). However it can also be used for transforming monocots, although other transformation methods are generally preferred for these plants. Presently, the method of choice for generating transgenic monocots is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil et al., 1992, *Bio/Technology* 10: 667-674). An alternative method for transforma-

tion of monocots is based on protoplast transformation as described by Omirulleh et al., 1993, *Plant Molecular Biology* 21: 415-428.

[0122] Following transformation, the transformants having incorporated therein the expression construct are selected and regenerated into whole plants according to methods well-known in the art. Often the transformation procedure is designed for the selective elimination of selection genes either during regeneration or in the following generations by using, for example, co-transformation with two separate T-DNA constructs or site specific excision of the selection gene by a specific recombinase.

[0123] The present invention also relates to methods for producing a biologically active substance of the present invention comprising (a) cultivating a transgenic plant or a plant cell comprising a gene encoding a biologically active substance of the present invention under conditions conducive for production of the biologically active substance; and (b) recovering the biologically active substance.

Removal or Reduction of Biologically Active Substance

[0124] The present invention also relates to methods for producing a mutant of a parent cell, which comprises disrupting or deleting all or a portion of a gene encoding a biologically active substance of the present invention, which results in the mutant cell producing less of the biologically active substance than the parent cell when cultivated under the same conditions.

[0125] The mutant cell may be constructed by reducing or eliminating expression of a gene encoding or regulatory synthesis of a biologically active substance of the present invention using methods well known in the art, for example, insertions, disruptions, replacements, or deletions. The gene to be modified or inactivated may be, for example, the coding region or a part thereof essential for activity, or a regulatory element of the gene required for the expression of the coding region. An example of such a regulatory or control sequence may be a promoter sequence or a functional part thereof, i.e., a part that is sufficient for affecting expression of the gene. Other control sequences for possible modification include, but are not limited to, a leader, propeptide sequence, signal peptide sequence, transcription terminator, and transcriptional activator.

[0126] Modification or inactivation of the gene may be performed by subjecting the parent cell to mutagenesis and selecting for mutant cells in which expression of the gene has been reduced or eliminated. The mutagenesis, which may be specific or random, may be performed, for example, by use of a suitable physical or chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the mutagenesis may be performed by use of any combination of these mutagenizing agents.

[0127] Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues.

[0128] When such agents are used, the mutagenesis is typically performed by incubating the parent cell to be mutagenized in the presence of the mutagenizing agent of

choice under suitable conditions, and screening and/or selecting for mutant cells exhibiting reduced or no expression of the gene.

[0129] Modification or inactivation of the nucleotide sequence may be accomplished by introduction, substitution, or removal of one or more nucleotides in the gene or a regulatory element required for the transcription or translation thereof. For example, nucleotides may be inserted or removed so as to result in the introduction of a stop codon, the removal of the start codon, or a change in the open reading frame. Such modification or inactivation may be accomplished by site-directed mutagenesis or PCR generated mutagenesis in accordance with methods known in the art. Although, in principle, the modification may be performed *in vivo*, i.e., directly on the cell expressing the nucleotide sequence to be modified, it is preferred that the modification be performed *in vitro* as exemplified below.

[0130] An example of a convenient way to eliminate or reduce expression of a nucleotide sequence by a cell of choice is based on techniques of gene replacement, gene deletion, or gene disruption. For example, in the gene disruption method, a nucleic acid sequence corresponding to the endogenous nucleotide sequence is mutagenized *in vitro* to produce a defective nucleic acid sequence which is then transformed into the parent cell to produce a defective gene. By homologous recombination, the defective nucleic acid sequence replaces the endogenous nucleotide sequence. It may be desirable that the defective nucleotide sequence also encodes a marker that may be used for selection of transformants in which the nucleotide sequence has been modified or destroyed. In a particularly preferred aspect, the nucleotide sequence is disrupted with a selectable marker such as those described herein.

[0131] Alternatively, modification or inactivation of the nucleotide sequence may be performed by established anti-sense techniques using a sequence complementary to the nucleotide sequence. More specifically, expression of the nucleotide sequence by a cell may be reduced or eliminated by introducing a sequence complementary to the nucleic acid sequence of the gene that may be transcribed in the cell and is capable of hybridizing to the mRNA produced in the cell. Under conditions allowing the complementary anti-sense nucleotide sequence to hybridize to the mRNA, the amount of protein translated is thus reduced or eliminated.

[0132] The present invention further relates to a mutant cell of a parent cell which comprises a disruption or deletion of a nucleotide sequence encoding the biologically active substance or a control sequence thereof, which results in the mutant cell producing less of the biologically active substance than the parent cell.

[0133] The biologically active substance-deficient mutant cells so created are particularly useful as host cells for the expression of homologous and/or heterologous substances, such as polypeptides. Therefore, the present invention further relates to methods for producing a homologous or heterologous substance comprising (a) cultivating the mutant cell under conditions conducive for production of the substance; and (b) recovering the substance. The term "heterologous substances" is defined herein as substances which are not native to the host cell, a native substance in which modifications have been made to alter the native sequence, or a native substance whose expression is quantitatively altered as a result of a manipulation of the host cell by recombinant DNA techniques.

[0134] In a further aspect, the present invention relates to a method for producing a protein product essentially free of a biologically active substance by fermentation of a cell which produces both a biologically active substance of the present invention as well as the protein product of interest by adding an effective amount of an agent capable of inhibiting activity of the biologically active substance to the fermentation broth before, during, or after the fermentation has been completed, recovering the product of interest from the fermentation broth, and optionally subjecting the recovered product to further purification.

[0135] In accordance with this aspect of the invention, it is possible to remove at least 60%, preferably at least 75%, more preferably at least 85%, still more preferably at least 95%, and most preferably at least 99% of the biologically active substance. Complete removal of biologically active substance may be obtained by use of this method.

[0136] The methods used for cultivation and purification of the product of interest may be performed by methods known in the art.

[0137] The methods of the present invention for producing an essentially biologically active substance-free product is of particular interest in the production of prokaryotic polypeptides, in particular bacterial proteins such as enzymes. The enzyme may be selected from, e.g., an amylolytic enzyme, lipolytic enzyme, proteolytic enzyme, cellulolytic enzyme, oxidoreductase, or plant cell-wall degrading enzyme. Examples of such enzymes include an aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinolytic enzyme, peroxidase, phytase, phenoloxidase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transferase, transglutaminase, or xylanase. The biologically active substance-deficient cells may also be used to express heterologous proteins of pharmaceutical interest such as hormones, growth factors, receptors, and the like.

[0138] It will be understood that the term "prokaryotic polypeptides" includes not only native polypeptides, but also those polypeptides, e.g., enzymes, which have been modified by amino acid substitutions, deletions or additions, or other such modifications to enhance activity, thermostability, pH tolerance and the like.

[0139] In a further aspect, the present invention relates to a product of a protein or substance essentially free of a biologically active substance of the invention, produced by a method of the present invention.

Compositions

[0140] The present invention also relates to compositions comprising a biologically active substance of the present invention. Preferably, the compositions are enriched in the biologically active substance. The term "enriched" indicates that the biologically active substance of the composition has been increased, e.g., with an enrichment factor of 1.1.

[0141] The composition may comprise a biologically active substance of the invention as the major component, e.g., a mono-component composition. Alternatively, the composition may comprise multiple biologically active substances, for example, multiple enzymes, such as an aminopeptidase, amylase, carbohydrase, carboxypeptidase,

catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase, haloperoxidase, invertase, laccase, lipase, mannosidase, oxidase, pectinolytic enzyme, peptidoglutaminase, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, or xylanase.

[0142] The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition. For instance, the composition may be in the form of a granulate or a microgranulate. The biologically active substance to be included in the composition may be stabilized in accordance with methods known in the art.

Methods for Using the *Bacillus licheniformis* Chromosome

[0143] The present invention also relates to methods for using the *Bacillus licheniformis* chromosome.

[0144] The chromosome of *Bacillus licheniformis* serves as a reservoir of new genes/proteins that have likely environmental, energy, health, and industrial applications (e.g., enzymes, antibiotics, biochemicals). A clear extension of this is that the newly discovered molecules can be used as starting points for further improvements via well-established gene shuffling, directed evolution, and protein engineering methods. Additionally, regions or motifs (e.g., signal peptides, active sites, substrate-binding regions) from the newly discovered molecules may be employed to derive novel chimeras with industrially advantageous properties.

[0145] The genes encoded in the chromosome may be used for monitoring global gene expression during the life cycle of the organism or during industrial fermentations (e.g., implemented on DNA microarrays). By monitoring global gene expression, for example, improved processes for industrial fermentation can be implemented with greater efficiency and economy.

[0146] The chromosome is useful in comparative evolutionary and ecological studies. For example, dozens of *Bacillus licheniformis* isolates can be readily compared on a global scale by hybridization of their genomic DNAs to a microarray fabricated from the reference strain presented in this case (so-called comparative genomic hybridization). Using this method, one can compare various isolates to look for similarities/differences among geographical and environmental niches or among biocontrol strains versus saprophytic isolates.

[0147] The chromosome sequence may be used to construct the metabolic blueprint for *Bacillus licheniformis* that includes all catabolic and anabolic pathways, signaling pathways, regulatory networks, growth substrates, biochemical intermediates, end products, electron donors/acceptors and others. In doing so, it is possible to modify the metabolic machinery of the organism by deleting unwanted pathways and/or adding enzymes/pathways from other organisms to generate useful chemicals and intermediates.

[0148] The pathways and components that contribute to production of extracellular and surface proteins in *Bacillus licheniformis* can be extracted from the chromosomal sequence. This affords opportunities for improved production of extracellular proteins by genetic manipulation of the secretion machinery.

[0149] The chromosome data allows deduction of the essential genes for *Bacillus licheniformis* (either by comparison to related bacteria such as *Bacillus subtilis* or by systematic gene-by-gene knock outs). Thus it has become possible to

design custom-made strains which contain only the genes that are essential for production of specific proteins or metabolites (so-called cell factory concept).

[0150] The chromosome data may be used to construct interspecies hybrids between *Bacillus licheniformis* and other bacteria. Venter et al., 2003, *Proc. Nat. Acad. Sci. USA* 100, 15440-15445 have shown that it is possible to construct an entire virus genome from smaller DNA segments. Thus, segments of the *Bacillus licheniformis* chromosome may be employed to derive novel chromosomal segments or even entire chimeric chromosomes for specific applications.

[0151] In a preferred aspect, methods for using the *Bacillus licheniformis* chromosome include host improvement, e.g., secretion of a protein or metabolite, genome shuffling, construction of new genomes, metabolic engineering and pathway reconstruction, carrier for heterologous expression vectors, microarrays as described herein, identification of polypeptides in proteomics analyses, and comparative genomics with other *Bacillus* species or related organisms.

Methods for Isolating Genes

[0152] The present invention also relates to methods for isolating a gene encoding a biologically active substance from a microbial strain. The method comprises first the addition of a mixture of first labeled nucleic acid probes, isolated from a microbial strain cultured on medium without an inducing substrate, and a mixture of second labeled nucleic acid probes, isolated from the microbial strain cultured on medium with the inducing substrate, to an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198, under conditions where the labeled nucleic acid probes hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array. The first nucleic acid probes are labeled with a first reporter and the second nucleic acid probes are labeled with a second reporter. The array is then examined under conditions wherein the relative expression of the genes of the microbial strain is determined by the observed hybridization reporter signal of each spot on the array in which (i) the *Bacillus licheniformis* genes on the array that hybridize to the first nucleic acid probes produce a distinct first hybridization reporter signal or to the second nucleic acid probes produce a distinct second hybridization reporter signal, and (ii) the *Bacillus licheniformis* genes on the array that hybridize to both the first and second nucleic acid probes produce a distinct combined hybridization reporter signal. The probe is then sequenced to isolate from the microbial strain the corresponding gene that encodes an enzyme that degrades or converts the substrate.

[0153] Enzymes. The gene of interest may encode any enzyme including an oxidoreductase, transferase, hydrolase, lyase, isomerase, or ligase. In a preferred aspect, the enzyme is an acylase, alpha-glucosidase, amidase, aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, dextrinase, endoglucanase, esterase, galactanase, alpha-galactosidase, beta-galactosidase, glucoamylase, glucanase, glucocerebrosidase, alpha-glucosidase, beta-glucosidase, hemicellulase, invertase, laccase, lignase, lipase, lysin, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phosphatase, phospholi-

pase, phytase, polyphenoloxidase, proteolytic enzyme, pullulanase, ribonuclease, transglutaminase, urokinase, or xylanase.

[0154] Inducing Substrate. The inducing substrate may be any substrate that is subject to the action of an enzyme, i.e., that degrades or converts the substrate. In a preferred aspect, the inducing substrate is lignin or a lignin-containing material. In a more preferred aspect, the lignin-containing material is lignocellulose. In another preferred aspect, the inducing substrate is cellulose. In another preferred aspect, the inducing substrate is hemicellulose. In another preferred aspect, the inducing substrate is pectin. In another preferred aspect, the inducing substrate is a lipid. In another preferred aspect, the inducing substrate is phospholipid. In another preferred aspect, the inducing substrate is phytic acid. In another preferred aspect, the inducing substrate is protein. In another preferred aspect, the inducing substrate is a starch. In another preferred aspect, the inducing substrate is a medium that is low in nutrients such as amino acids, carbon, nitrogen, phosphate, or iron.

[0155] In a more preferred aspect, the protein substrate is blood, casein, egg, gelatin, gluten, milk protein, or soy protein. In another more preferred aspect, the lignin-containing material is hardwood thermomechanical pulp. In another more preferred aspect, the lignocellulose is corn stover. In another more preferred aspect, the lignocellulose is white poplar. In another more preferred aspect, the lignocellulose is rice straw. In another more preferred aspect, the lignocellulose is switch grass.

[0156] Microbial Strains. In the methods of the present invention, the microbial strain may be any microbial strain. The strain is cultured on a suitable nutrient medium with and without a substrate of interest. The strain cultured on medium without the substrate is used as a reference for identifying differences in expression of the same or similar complement of genes in the strain cultured on medium with substrate. The strain may be a wild-type, mutant, or recombinant strain.

[0157] In the methods of the present invention, the microbial strain is preferably a bacterium. In a more preferred aspect, the bacterium is a *Bacillus*, *Pseudomonas*, or *Streptomyces* strain or *E. coli*.

[0158] The *Bacillus* strain may be any *Bacillus* strain. In a preferred aspect, the *Bacillus* strain is *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus fastidiosus*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus macerans*, *Bacillus megaterium*, *Bacillus methanolicus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis*. It will be understood that the term "*Bacillus*" also encompasses relatives of *Bacillus* such as *Paenibacillus*, *Oceanobacillus*, and the like.

[0159] The *Pseudomonas* strain may be any *Pseudomonas* strain. In a preferred aspect, the *Pseudomonas* strain is *Pseudomonas acidovorans*, *Pseudomonas aeruginosa*, *Pseudomonas alcaligenes*, *Pseudomonas anguilliseptica*, *Pseudomonas abtimicrobica*, *Pseudomonas aurantiaca*, *Pseudomonas aureofaciens*, *Pseudomonas beijerinckii*, *Pseudomonas boreopolis*, *Pseudomonas chlororaphis*, *Pseudomonas citronellolis*, *Pseudomonas cocovenenans*, *Pseudomonas diminuta*, *Pseudomonas doudoroffii*, *Pseudomonas echinoides*, *Pseudomonas elongata*, *Pseudomonas fluorescens*, *Pseudomonas fragi*, *Pseudomonas halophobica*, *Pseudomonas huttiensis*, *Pseudomonas*

indigofera, *Pseudomonas lanceolata*, *Pseudomonas lemoignei*, *Pseudomonas lundensis*, *Pseudomonas mendocina*, *Pseudomonas mephitica*, *Pseudomonas mucidolens*, *Pseudomonas oleovorans*, *Pseudomonas phenazinium*, *Pseudomonas pictorium*, *Pseudomonas putida*, *Pseudomonas resinovorans*, *Pseudomonas saccharophila*, *Pseudomonas stamieri*, *Pseudomonas stutzeri*, *Pseudomonas taetrolensis*, or *Pseudomonas vesicularis*.

[0160] The *Streptomyces* strain may be any *Streptomyces* strain. In a preferred aspect, the *Streptomyces* strain is *Streptomyces lividans*. In another preferred aspect, the *Streptomyces* strain is *Streptomyces murinus*.

[0161] Microarrays. The term "an array of *Bacillus licheniformis* genes" is defined herein as a linear or two-dimensional array of preferably discrete elements of an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198 (e.g., synthetic oligonucleotides of, for example, 40-60 nucleotides), wherein each discrete element has a finite area, formed on the surface of a solid support. It shall be understood that the term "*Bacillus licheniformis* genes" encompasses nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198. The term "microarray" is defined herein as an array of *Bacillus licheniformis* gene elements having a density of discrete of *Bacillus licheniformis* gene elements of at least about 100/cm², and preferably at least about 1000/cm². The *Bacillus licheniformis* gene elements in a microarray have typical dimensions, e.g., diameters, in the range of between about 10 to about 250 μm, preferably in the range of between about 10 to about 200 μm, more preferably in the range of between about 20 to about 150 μm, even more preferably in the range of between about 20 to about 100 μm, most preferably in the range of between about 50 to about 100 μm, and even most preferably in the range of between about 80 to about 100 μm, and are separated from other gene elements in the microarray by about the same distance.

[0162] Methods and instruments for forming microarrays on the surface of a solid support are well known in the art. See, for example, U.S. Pat. No. 5,807,522; U.S. Pat. No. 5,700,637; and U.S. Pat. No. 5,770,151. The instrument may be an automated device such as described in U.S. Pat. No. 5,807,522.

[0163] The term "a substrate containing an array of *Bacillus licheniformis* genes" is defined herein as a solid support having deposited on the surface of the support one or more of a plurality of *Bacillus licheniformis* genes, as described herein, for use in detecting binding of labeled nucleic acids to the *Bacillus licheniformis* genes.

[0164] The substrate may, in one aspect, be a glass support (e.g., glass slide) having a hydrophilic or hydrophobic coating on the surface of the support, and an array of distinct random nucleic acid fragments bound to the coating, where each distinct random nucleic acid fragment is disposed at a separate, defined position.

[0165] Each microarray in the substrate preferably contains at least 10³ distinct *Bacillus licheniformis* in a surface area of less than about 5 or 6 cm². Each distinct *Bacillus licheniformis* gene (i) is disposed at a separate, defined position on the array, (ii) has a length of at least 50 bp, and (iii) is present in a defined amount between about 0.1 femtomoles and 100 nanomoles or higher if necessary.

[0166] For a hydrophilic coating, the glass slide is coated by placing a film of a polycationic polymer with a uniform thickness on the surface of the slide and drying the film to form a dried coating. The amount of polycationic polymer added should be sufficient to form at least a monolayer of polymers on the glass surface. The polymer film is bound to the surface via electrostatic binding between negative silyl-OH groups on the surface and charged cationic groups in the polymers. Such polycationic polymers include, but are not limited to, polylysine and polyarginine.

[0167] Another coating strategy employs reactive aldehydes to couple DNA to the slides (Scheda et al., 1996, *Proceedings of the National Academy of Science USA* 93: 10614-10619; Heller et al., 1997, *Proceedings of the National Academy of Science USA* 94: 2150-2155).

[0168] Alternatively, the surface may have a relatively hydrophobic character, i.e., one that causes aqueous medium deposited on the surface to bead. A variety of known hydrophobic polymers, such as polystyrene, polypropylene, or polyethylene, have desirable hydrophobic properties, as do glass and a variety of lubricant or other hydrophobic films that may be applied to the support surface. A support surface is "hydrophobic" if an aqueous droplet applied to the surface does not spread out substantially beyond the area size of the applied droplet, wherein the surface acts to prevent spreading of the droplet applied to the surface by hydrophobic interaction with the droplet.

[0169] In another aspect, the substrate may be a multi-cell substrate where each cell contains a microarray of *Bacillus licheniformis* and preferably an identical microarray, formed on a porous surface. For example, a 96-cell array may typically have array dimensions between about 12 and 244 mm in width and 8 and 400 mm in length, with the cells in the array having width and length dimension of 1/2 and 1/8 the array width and length dimensions, respectively, i.e., between about 1 and 20 in width and 1 and 50 mm in length.

[0170] The solid support may include a water-impermeable backing such as a glass slide or rigid polymer sheet, or other non-porous material. Formed on the surface of the backing is a water-permeable film, which is formed of porous material. Such porous materials include, but are not limited to, nitrocellulose membrane nylon, polypropylene, and polyvinylidene difluoride (PVDF) polymer. The thickness of the film is preferably between about 10 and 1000 μm. The film may be applied to the backing by spraying or coating, or by applying a preformed membrane to the backing.

[0171] Alternatively, the solid support may be simply a filter composed of nitrocellulose, nylon, polypropylene, or polyvinylidene difluoride (PVDF) polymer, or, for that matter, any material suitable for use.

[0172] The film surface may be partitioned into a desirable array of cells by water-impermeable grid lines typically at a distance of about 100 to 2000 μm above the film surface. The grid lines can be formed on the surface of the film by laying down an uncured flowable resin or elastomer solution in an array grid, allowing the material to infiltrate the porous film down to the backing, and then curing the grid lines to form the cell-array substrate.

[0173] The barrier material of the grid lines may be a flowable silicone, wax-based material, thermoset material (e.g., epoxy), or any other useful material. The grid lines may be applied to the solid support using a narrow syringe, printing techniques, heat-seal stamping, or any other useful method known in the art.

[0174] Each well preferably contains a microarray of distinct *Bacillus licheniformis* genes. "Distinct *Bacillus licheniformis* genes" as applied to the genes forming a microarray is defined herein as an array member which is distinct from other array members on the basis of a different *Bacillus licheniformis* gene sequence or oligo sequence thereof, and/or different concentrations of the same or distinct *Bacillus licheniformis* genes and/or different mixtures of distinct *Bacillus licheniformis* genes or different concentrations of *Bacillus licheniformis* genes. Thus an array of "distinct *Bacillus licheniformis* genes" may be an array containing, as its members, (i) distinct *Bacillus licheniformis* genes which may have a defined amount in each member, (ii) different, graded concentrations of a specific *Bacillus licheniformis* gene, and/or (iii) different-composition mixtures of two or more distinct *Bacillus licheniformis* genes.

[0175] It will be understood, however, that in the methods of the present invention, any type of substrate known in the art may be used.

[0176] The delivery of a known amount of a selected *Bacillus licheniformis* gene to a specific position on the support surface is preferably performed with a dispensing device equipped with one or more tips for insuring reproducible deposition and location of the *Bacillus licheniformis* genes and for preparing multiple arrays. Any dispensing device known in the art may be used in the methods of the present invention. See, for example, U.S. Pat. No. 5,807,522.

[0177] For liquid-dispensing on a hydrophilic surface, the liquid will have less of a tendency to bead, and the dispensed volume will be more sensitive to the total dwell time of the dispenser tip in the immediate vicinity of the support surface.

[0178] For liquid-dispensing on a hydrophobic surface, flow of fluid from the tip onto the support surface will continue from the dispenser onto the support surface until it forms a liquid bead. At a given bead size, i.e., volume, the tendency of liquid to flow onto the surface will be balanced by the hydrophobic surface interaction of the bead with the support surface, which acts to limit the total bead area on the surface, and by the surface tension of the droplet, which tends toward a given bead curvature. At this point, a given bead volume will have formed, and continued contact of the dispenser tip with the bead, as the dispenser tip is being withdrawn, will have little or no effect on bead volume.

[0179] The desired deposition volume, i.e., bead volume, formed is preferably in the range 2 pl (picoliters) to 2 nl (nanoliters), although volumes as high as 100 nl or more may be dispensed. It will be appreciated that the selected dispensed volume will depend on (i) the "footprint" of the dispenser tip(s), i.e., the size of the area spanned by the tip(s), (ii) the hydrophobicity of the support surface, and (iii) the time of contact with and rate of withdrawal of the tip(s) from the support surface. In addition, bead size may be reduced by increasing the viscosity of the medium, effectively reducing the flow time of liquid from the dispensing device onto the support surface. The drop size may be further constrained by depositing the drop in a hydrophilic region surrounded by a hydrophobic grid pattern on the support surface.

[0180] At a given tip size, bead volume can be reduced in a controlled fashion by increasing surface hydrophobicity, reducing time of contact of the tip with the surface, increasing rate of movement of the tip away from the surface, and/or increasing the viscosity of the medium. Once these param-

eters are fixed, a selected deposition volume in the desired picoliter to nanoliter range can be achieved in a repeatable fashion.

[0181] After depositing a liquid droplet of a *Bacillus licheniformis* gene sample at one selected location on a support, the tip may be moved to a corresponding position on a second support, the *Bacillus licheniformis* gene sample is deposited at that position, and this process is repeated until the random nucleic acid fragment sample has been deposited at a selected position on a plurality of supports.

[0182] This deposition process may then be repeated with another random nucleic acid fragment sample at another microarray position on each of the supports.

[0183] The diameter of each *Bacillus licheniformis* gene region is preferably between about 20-200 μm . The spacing between each region and its closest (non-diagonal) neighbor, measured from center-to-center, is preferably in the range of about 20-400 μm . Thus, for example, an array having a center-to-center spacing of about 250 μm contains about 40 regions/cm or 1,600 regions/cm². After formation of the array, the support is treated to evaporate the liquid of the droplet forming each region, to leave a desired array of dried, relatively flat *Bacillus licheniformis* gene or oligo thereof regions. This drying may be done by heating or under vacuum. The DNA can also be UV-crosslinked to the polymer coating.

[0184] Nucleic Acid Probes. In the methods of the present invention, the strains are cultivated in a nutrient medium with and without a substrate using methods well known in the art for isolation of nucleic acids to be used as probes. For example, the strains may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection).

[0185] The nucleic acid probes from the microbial strains cultured on medium with and without substrate may be any nucleic acid including genomic DNA, cDNA, and RNA, and may be isolated using standard methods known in the art.

[0186] The populations of isolated nucleic acid probes may be labeled with detection reporters such as colorimetric, radioactive for example, ³²P, ³³P, or ³⁵S), fluorescent reporters, or other reporters using methods known in the art (Chen et al., 1998, *Genomics* 51: 313-324; DeRisi et al., 1997, *Science* 278: 680-686; U.S. Pat. No. 5,770,367).

[0187] In a preferred aspect, the probes are labeled with fluorescent reporters. For example, the DNA probes may be labeled during reverse transcription from the respective RNA pools by incorporation of fluorophores as dye-labeled nucleotides (DeRisi et al., 1997, supra), e.g., Cy5-labeled deoxyuridine triphosphate, or the isolated cDNAs may be directly labeled with different fluorescent functional groups. Fluorescent-labeled nucleotides include, but are not limited to, fluorescein conjugated nucleotide analogs (green fluorescence), lissamine nucleotide analogs (red fluorescence). Fluorescent functional groups include, but are not limited to, Cy3 (a green fluorescent dye) and Cy5 (red fluorescent dye).

[0188] Array Hybridization. The labeled nucleic acids from the two strains cultivated with and without substrate are then added to an array of *Bacillus licheniformis* genes under

conditions where the nucleic acid pools from the two strains hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array. For purposes of the present invention, hybridization indicates that the labeled nucleic acids from the two strains hybridize to the *Bacillus licheniformis* genes under very low to very high stringency conditions.

[0189] A small volume of the labeled nucleic acids mixture is loaded onto the substrate. The solution will spread to cover the entire microarray. In the case of a multi-cell substrate, one or more solutions are loaded into each cell which stop at the barrier elements.

[0190] For nucleic acid probes of at least about 100 nucleotides in length, microarray hybridization conditions described by Eisen and Brown, 1999, *Methods of Enzymology* 303: 179-205, may be used. Hybridization is conducted under a cover slip at 65° C. in 3×SSC for 4-16 hours followed by post-hybridization at room temperature after removal of the cover slip in 2×SSC, 0.1% SDS by washing the array two or three times in the solution, followed by successive washes in 1×SSC for 2 minutes and 0.2×SSC wash for two or more minutes.

[0191] Conventional conditions of very low to very high stringency conditions may also be used. Very low to very high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures.

[0192] The carrier material is finally washed three times each for 15 minutes using 2×SSC, 0.2% SDS preferably at least at 45° C. (very low stringency), more preferably at least at 50° C. (low stringency), more preferably at least at 55° C. (medium stringency), more preferably at least at 60° C. (medium-high stringency), even more preferably at least at 65° C. (high stringency), and most preferably at least at 70° C. (very high stringency).

[0193] For shorter nucleic acid probes which are less than 50 nucleotides, microarray hybridization conditions described by Kane et al., 2000, *Nucleic Acids Research* 28: 4552-4557, may be used. Hybridization is conducted under a supported coverslip at 42° C. for 16-18 hours at high humidity in 50% formamide, 4.1×Denhardt's solution, 4.4×SSC, and 100 µg/ml of herring sperm DNA. Arrays are washed after removal of the coverslip in 4×SSC by immersion into 1×SSC, 0.1% SDS for 10 minutes, 0.1×SSC, 0.1% SDS twice for 10 minutes, and 0.1×SSC twice for 10 minutes.

[0194] For shorter nucleic acid probes which are about 50 nucleotides to about 100 nucleotides in length, conventional stringency conditions may be used. Such stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at 5° C. to 10° C. below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1×Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures.

[0195] The carrier material is finally washed once in 6×SSC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6×SSC at 5° C. to 10° C. below the calculated T_m .

[0196] The choice of hybridization conditions will depend on the degree of homology between the *Bacillus licheniformis* genes and the nucleic acid probes obtained from the strain cultured with and without inducing substrate. For example, where the nucleic acid probes and the *Bacillus licheniformis* genes are obtained from identical strains, high stringency conditions may be most suitable. Where the strains are from a genus or species different from which the *Bacillus licheniformis* genes were obtained, low or medium stringency conditions may be more suitable.

[0197] In a preferred aspect, the hybridization is conducted under low stringency conditions. In a more preferred aspect, the hybridization is conducted under medium stringency conditions. In a most preferred aspect, the hybridization is conducted under high stringency conditions.

[0198] The entire solid support is then reacted with detection reagents if needed and analyzed using standard colorimetric, radioactive, or fluorescent detection means. All processing and detection steps are performed simultaneously to all of the microarrays on the solid support ensuring uniform assay conditions for all of the microarrays on the solid support.

[0199] Detection. The most common detection method is laser-induced fluorescence detection using confocal optics (Cheung et al., 1998, *Nat. Genet.* 18: 225-230). The array is examined under fluorescence excitation conditions such that (i) the *Bacillus licheniformis* genes on the array that hybridize to the first nucleic acid probes obtained from the strain cultured without inducing substrate and to the second nucleic acid probes obtained from the strain cultured with inducing substrate produce a distinct first fluorescence emission color and a distinct second fluorescence emission color, respectively, and (ii) the *Bacillus licheniformis* genes on the array that hybridize to substantially equal numbers of nucleic acid probes obtained from the strain cultured without inducing substrate and from the strain cultured with inducing substrate produce a distinct combined fluorescence emission color; wherein the relative expression of the genes in the strains can be determined by the observed fluorescence emission color of each spot on the array.

[0200] The fluorescence excitation conditions are based on the selection of the fluorescence reporters. For example, Cy3 and Cy5 reporters are detected with solid state lasers operating at 532 nm and 632 nm, respectively.

[0201] However, other methods of detection well known in the art may be used such as standard photometric, colorimetric, or radioactive detection means, as described earlier.

[0202] Data Analysis. The data obtained from the scanned image may then be analyzed using any of the commercially available image analysis software. The software preferably identifies array elements, subtracts backgrounds, deconvolutes multi-color images, flags or removes artifacts, verifies that controls have performed properly, and normalizes the signals (Chen et al., 1997, *Journal of Biomedical Optics* 2: 364-374).

[0203] Several computational methods have been described for the analysis and interpretation of microarray-based expression profiles including cluster analysis (Eisen et al., 1998, *Proc. Nat. Acad. Sci. USA* 95: 14863-14868), parametric ordering of genes (Spellman et al., 1998, *Mol. Biol.*

Cell 9: 3273-3297), and supervised clustering methods based on representative hand-picked or computer-generated expression profiles (Chu et al., 1998, *Science* 282: 699-705). Preferred methods for evaluating the results of the microarrays employ statistical analysis to determine the significance of the differences in expression levels. In the methods of the present invention, the difference in the detected expression level is at least about 10% or greater, preferably at least about 20% or greater, more preferably at least about 50% or greater, even more preferably at least about 75% or greater; and most preferably at least about 100% or greater.

[0204] One such preferred system is the Significance Analysis of Microarrays (SAM) (Tusher et al., 2001, *Proc. Natl. Acad. Sci. USA* 98: 5116-5121). Statistical analysis allows the determination of significantly altered expression of levels of about 50% or even less. The PAM (or predictive analysis for microarrays) represents another approach for analyzing the results of the microarrays (Tibshirani et al., 2002, *Proc. Natl. Acad. Sci. USA* 99: 6567-6572).

[0205] Cluster algorithms may also be used to analyze microarray expression data. From the analysis of the expression profiles it is possible to identify co-regulated genes that perform common metabolic or biosynthetic functions. Hierarchical clustering has been employed in the analysis of microarray expression data in order to place genes into clusters based on sharing similar patterns of expression (Eisen et al., 1998, *supra*). This method yields a graphical display that resembles a kind of phylogenetic tree where the relatedness of the expression behavior of each gene to every other gene is depicted by branch lengths. The programs Cluster and Tree-View, both written by Michael Eisen (Eisen et al., 1998 *Proc. Nat. Acad. Sci. USA* 95: 14863-14868) are freely available. Genespring is a commercial program available for such analysis (Silicon Genetics, Redwood City, Calif.).

[0206] Self-organizing maps (SOMs), a non-hierarchical method, have also been used to analyze microarray expression data (Tamayo et al., 1999, *Proc. Natl. Acad. Sci. USA* 96: 2907-2912). This method involves selecting a geometry of nodes, where the number of nodes defines the number of clusters. Then, the number of genes analyzed and the number of experimental conditions that were used to provide the expression values of these genes are subjected to an iterative process (20,000-50,000 iterations) that maps the nodes and data points into multidimensional gene expression space. After the identification of significantly regulated genes, the expression level of each gene is normalized across experiments. As a result, the expression profile of the genome is highlighted in a manner that is relatively independent of each gene's expression magnitude. Software for the "GENECLUSTER" SOM program for microarray expression analysis can be obtained from the Whitehead/MIT Center for Genome Research. SOMs can also be constructed using the GeneSpring software package.

[0207] Isolation of Genes. Probes containing genes or portions thereof identified to be induced by the present of substrate in the medium are characterized by determining the sequence of the probe. Based on the sequence, the gene can then be isolated using methods well known in the art.

[0208] The techniques used to isolate or clone a gene include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the gene from such genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA frag-

ments with shared structural features. See, e.g., Innis et al., 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleic acid sequence-based amplification (NASBA) may be used. The gene may be cloned from the strain of interest, or another or related organism and thus, for example, may be an allelic or species variant of the gene.

Methods for Monitoring Differential Expression of a Plurality of Genes

[0209] The present invention also relates to methods for monitoring differential expression of a plurality of genes in a first bacterial cell relative to expression of the same genes in one or more second bacterial cells, comprising:

[0210] (a) adding a mixture of detection reporter-labeled nucleic acids isolated from the bacterial cells to a substrate containing an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198, under conditions where the detection reporter-labeled nucleic acids hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array, wherein the nucleic acids from the first bacterial cell and the one or more second bacterial cells are labeled with a first detection reporter and one or more different second detection reporters, respectively; and

[0211] (b) examining the array under conditions wherein the relative expression of the genes in the bacterial cells is determined by the observed detection signal of each spot on the array in which (i) the *Bacillus licheniformis* genes on the array that hybridize to the nucleic acids obtained from either the first or the one or more second bacterial cells produce a distinct first detection signal or one or more second detection signals, respectively, and (ii) the *Bacillus licheniformis* genes on the array that hybridize to the nucleic acids obtained from both the first and one or more second bacterial produce a distinct combined detection signal.

[0212] The methods of the present invention may be used to monitor global expression of a plurality of genes from a *Bacillus* cell, discover new genes, identify possible functions of unknown open reading frames, and monitor gene copy number variation and stability. For example, the global view of changes in expression of genes may be used to provide a picture of the way in which *Bacillus* cells adapt to changes in culture conditions, environmental stress, or other physiological provocation. Other possibilities for monitoring global expression include spore morphogenesis, recombination, metabolic or catabolic pathway engineering.

[0213] The methods of the present invention are particularly advantageous since one spot on an array equals one gene or open reading frame because extensive follow-up characterization is unnecessary since sequence information is available, and the *Bacillus licheniformis* microarrays can be organized based on function of the gene products.

[0214] Microarrays. Methods for preparing the microarrays are described herein.

[0215] Bacterial Cells. In the methods of the present invention, the two or more *Bacillus* cells may be any *Bacillus* cell where one of the cells is used as a reference for identifying differences in expression of the same or similar complement of genes in the other cell(s). In one aspect, the two or more cells are the same cell. For example, they may be compared under different growth conditions, e.g., oxygen limitation,

nutrition, and/or physiology. In another aspect, one or more cells are mutants of the reference cell. For example, the mutant(s) may have a different phenotype. In a further aspect, the two or more cells are of different species (e.g., *Bacillus clausii* and *Bacillus subtilis*). In another further aspect, the two or more cells are of different genera. In an even further aspect, one or more cells are transformants of the reference cell, wherein the one or more transformants exhibit a different property. For example, the transformants may have an improved phenotype relative to the reference cell and/or one of the other transformants. The term “phenotype” is defined herein as an observable or outward characteristic of a cell determined by its genotype and modulated by its environment. Such improved phenotypes may include, but are not limited to, improved secretion or production of a protein or compound, reduced or no secretion or production of a protein or compound, improved or reduced expression of a gene, desirable morphology, an altered growth rate under desired conditions, relief of over-expression mediated growth inhibition, or tolerance to low oxygen conditions.

[0216] The *Bacillus* cells may be any *Bacillus* cells, but preferably *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus fastidiosus*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus macerans*, *Bacillus megaterium*, *Bacillus methanolicus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis* cells.

[0217] In a preferred aspect, the *Bacillus* cells are *Bacillus alkalophilus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus amyloliquefaciens* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus brevis* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus cereus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus circulans* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus clausii* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus coagulans* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus fastidiosus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus firmus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus lautus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus lentus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus licheniformis* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus macerans* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus megaterium* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus methanolicus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus pumilus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus sphaericus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus stearothermophilus* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus subtilis* cells. In another preferred aspect, the *Bacillus* cells are *Bacillus thuringiensis* cells.

[0218] In a more preferred aspect, the *Bacillus* cells are *Bacillus licheniformis* cells. In a most preferred aspect, the *Bacillus licheniformis* cells are *Bacillus licheniformis* ATCC 14580 cells.

[0219] In another more preferred aspect, the *Bacillus* cells are *Bacillus clausii* cells. In another most preferred aspect, the *Bacillus clausii* cells are *Bacillus clausii* NUB 10309 cells.

[0220] It will be understood that the term “*Bacillus*” also encompasses relatives of *Bacillus* such as *Paenibacillus*, *Oceanobacillus*, and the like.

[0221] In the methods of the present invention, the cells are cultivated in a nutrient medium suitable for growth using methods well known in the art for isolation of the nucleic acids to be used as probes. For example, the cells may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection).

[0222] Nucleic Acid Probes. The nucleic acid probes from the two or more *Bacillus* cells may be any nucleic acid including genomic DNA, cDNA, and RNA, and may be isolated using standard methods known in the art, as described herein. The populations of isolated nucleic acid probes may be labeled with colorimetric, radioactive, fluorescent reporters, or other reporters using methods described herein.

[0223] In a preferred aspect, the probes are labeled with fluorescent reporters, e.g., Cy3 (a green fluorescent dye) and Cy5 (red fluorescent dye), as described herein.

[0224] Array Hybridization. The labeled nucleic acids from the two or more *Bacillus* cells are then added to a substrate containing an array of *Bacillus licheniformis* genes under conditions, as described herein, where the nucleic acid pools from the two or more *Bacillus* cells hybridize to complementary sequences of the *Bacillus licheniformis* genes on the array.

[0225] Detection and Data Analysis. The same methods as described herein are used for detection and data analysis.

Computer Readable Media and Computer-Based Systems

[0226] The *Bacillus licheniformis* chromosome and its genes described herein may be “provided” in a variety of media to facilitate their use. The term “provided” refers to a manufacture comprising an array of *Bacillus licheniformis* genes. Such manufactures provide the *Bacillus licheniformis* genes in a form which allows one skilled in the art to examine the manufacture using means not directly applicable to examining the chromosome or a subset thereof as it exists in nature or in purified form.

[0227] Thus, the present invention also relates to such a manufacture in the form of a computer readable medium comprising an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198.

[0228] In one application of this aspect, the *Bacillus licheniformis* genes of the present invention can be recorded on computer readable media. The term “computer readable media” is defined herein as any medium which can be read and accessed by a computer. Such computer readable media include, but are not limited to, magnetic storage media, e.g., floppy discs, hard disc storage medium, and magnetic tape; optical storage media, e.g., CD-ROM, DVD; electrical storage media, e.g., RAM and ROM; and hybrids of these categories, e.g., magnetic/optical storage media. One 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 medium having recorded thereon a nucleotide sequence of the present invention. Likewise, it will be clear to those of skill how additional computer readable media that may be developed also can be used to create analogous manufactures having recorded thereon a nucleotide sequence of the present invention.

[0229] As used herein, “recorded” refers to a process for storing information on computer readable medium. One skilled in the art can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

[0230] A variety of data storage structures are available for creating a computer readable medium 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 medium. 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 skilled artisan can readily adapt any number of data-processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

[0231] Various computer software programs are publicly available that allow a skilled artisan to access sequence information provided in a computer readable medium. Thus, by providing in computer readable form an array of *Bacillus licheniformis* genes selected from the group consisting of nucleotides SEQ ID NOs: 2-4198, complementary strands of SEQ ID NOs: 2-4198, or fragments of SEQ ID NOs: 2-4198, enables one skilled in the art to routinely access the provided sequence information for a wide variety of purposes.

[0232] Software utilizing the BLAST (Altschul et al., 1990, supra), BLAZE (Brutlag et al., 1993, *Comp. Chem.* 17: 203-207), GENEMARK (Lukashin and Borodovsky, 1998, *Nucleic Acids Research* 26: 1107-1115), GENSCAN (Burge and Karlin, 1997, *Journal of Molecular Biology* 268: 78-94), GLIMMER (Salzberg et al., 1998, *Nucleic Acids Research* 26: 544-548), and GRAIL (Xu et al., 1994, *Comput. Appl. Biosci.* 10: 613-623) search algorithms may be used to identify open reading frames (ORFs) within a genome of interest, which contain homology to ORFs or proteins from both *Bacillus licheniformis* and *Bacillus clausii* and from other organisms. Among the ORFs discussed herein are protein encoding fragments of the *Bacillus licheniformis* and *Bacillus clausii* genomes useful in producing commercially important proteins, such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

[0233] The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify, among other things, genes and gene products—many of which could be products themselves or used to genetically modify an industrial expression host through increased or decreased expression of a specific gene sequence(s).

[0234] The term “a computer-based system” is herein defined as 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. One skilled in the art can readily appreciate that any currently available computer-based system is suitable for use in the present invention.

[0235] As stated above, 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.

[0236] The term “data storage means” is defined herein as 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.

[0237] The term “search means” refers is defined herein as 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 present genomic sequences which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (Fuchs, 1991, *Comput. Appl. Biosci.* 7: 105-106), BLASTN and BLASTX National Center for Biotechnology information (NCBI). One 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.

[0238] The term “target sequence” is defined here as any DNA (genomic DNA, cDNA) or amino acid sequence of six or more nucleotides or two or more amino acids. One 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 searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

[0239] The term “a target structural motif” or “target motif” is defined herein as any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences, substrate and cofactor binding domains, transmembrane domains, and sites for post-translational modifications. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences), repeats, palindromes, dyad symmetries, and transcription and translation start and stop sites.

[0240] 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 *Bacillus licheniformis* or *Bacillus clausii* genomic sequences possessing varying degrees of homology to the target sequence or target motif. Such presentation provides one 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.

[0241] 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 *Bacillus licheniformis* and *Bacillus clausii* genomes. For example, implementing software which utilize the BLAST and BLAZE algorithms, described in Altschul et al., 1990, supra, may be used to identify open reading frames within the *Bacillus licheniformis* or *Bacillus clausii* genome or the genomes of other organisms. A skilled artisan 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. Suitable proprietary systems that may be known to those of skill also may be employed in this regard.

Codon Usage Tables

[0242] The present invention further relates to methods for preparing a synthetic gene, comprising (a) generating a codon usage table based on codons used in one or more open reading frames or portions thereof of SEQ ID NO: 1, (b) constructing a synthetic gene or portion thereof that contains in place of one or more native codons one or more preferred codons from the codon usage table, and (c) recovering the synthetic gene. In a preferred aspect, the codon usage table is Table 4 and/or Table 5.

[0243] The *Bacillus licheniformis* chromosomal sequence of SEQ ID NO: 1 or portions thereof can be used to generate codon usage tables to design synthetic genes for their efficient heterologous expression in *Bacillus licheniformis* host cells. The codon usage tables can be based on (1) the codon used in all the open reading frames, (2) selected open reading frames, (3) fragments of the open reading frames, or (4) fragments of selected open reading frames. With a codon usage table, synthetic genes can be designed with only the most preferred codon for each amino acid; with a number of common codons for each amino acid; or with the same or a similar statistical average of codon usages found in the table of choice.

[0244] The synthetic gene can be constructed using any method such as site-directed mutagenesis or PCR generated mutagenesis in accordance with methods known in the art. Although, in principle, the modification may be performed *in vivo*, i.e., directly on the cell expressing the nucleotide sequence to be modified, it is preferred that the modification is performed *in vitro*.

[0245] The synthetic gene can be further modified by operably linking the synthetic gene to one or more control sequences which direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences using the methods described herein. Nucleic acid constructs, recombinant expression vectors, and recombinant host cells comprising the synthetic gene can also be prepared using the methods described herein.

[0246] The present invention also relates to methods for producing a polypeptide encoded by such a synthetic gene comprising (a) cultivating a host cell comprising the synthetic gene under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

[0247] The present invention is further described by the following examples which should not be construed as limiting the scope of the invention.

EXAMPLES

Example 1

Shotgun DNA Sequencing and Genome Assembly

[0248] The genome of the type strain *Bacillus licheniformis* ATCC 14580 was sequenced by a combination of the whole genome shotgun method described by Wilson, R. K. and Mardis, E. R., 1997, In *Genome Analysis: A Laboratory Manual*, Vol. 1, eds. Birren, B., Green, E. D., Meyers, R. M., and Roskams, J. (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.), pp. 397-454, and fosmid end sequencing (Kim, U. J., Shizuya, H., de Jong, P. J., Birren, B. and Simon, M. I., 1992, *Nucleic Acids Res.* 20: 1083-1085; Longmire, J. L. and Brown, N. C., 2003, *Biotechniques* 35: 50-54; Zhao, S., Malek, J., Mahairas, G., Fu, L., Nierman, W., Venter, J. C., and Adams, M. D., 2000, *Genomics* 63: 321-332).

[0249] Genomic DNA of *Bacillus licheniformis* ATCC 14580 was isolated using the following method: A single colony was used to inoculate 20 ml of LB broth (Davis, R. W., Botstein, D., and Roth, J. R. 1980, *Advanced Bacterial Genetics*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y.) in a sterile 125 ml Erlenmeyer flask. The culture was incubated at 37° C. overnight with agitation at 240 rpm. The resulting cells were collected by centrifugation in a 45 ml Oak Ridge tube for 10 minutes at 6000×g, and the cell pellet was resuspended in 5 ml of Tris-glucose buffer (50 mM Tris-HCl, pH 8.0, 50 mM glucose, 10 mM EDTA). Lysozyme was added to a final concentration of 50 µg/ml and the suspension was incubated in a 37° C. water bath for 25 minutes. Next, 200 µl of 10% SDS was added and the tubes were gently inverted several times. Five milliliters of a second detergent mixture (1% Brij, 1% deoxycholate, 50 mM EDTA, pH 7.5) was added, and the tubes were inverted several times while incubating for 20 minutes at room temperature. An equal volume of phenol: chloroform (1:1 v/v) was added and the tubes were inverted gently at room temperature for 20-30 minutes. The tubes were centrifuged for 20 minutes at 12,000×g, 4° C. The top aqueous layer was carefully removed with a wide-bore pipette and placed in a clean 45 ml Oak Ridge tube. The phenol:chloroform extraction was repeated and 1/10 volume of 3 M sodium acetate pH 5.2 was added to the aqueous layer. Two volumes of cold ethanol were carefully layered on top and the DNA was spooled from the solution onto a sterile glass rod. Spooled DNA was carefully rinsed in 70% ethanol and resuspended in a suitable amount of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

[0250] Plasmid libraries were constructed using randomly-sheared and BamHI-digested genomic DNA that was enriched for 2-3 kb fragments by preparative agarose gel electrophoresis (Berka, R. M., Schneider, P., Golightly, E. J., Brown, S. H., Madden, M., Brown, K. M., Halkier, T., Mondorf, K., and Xu, F., 1997, *Appl. Environ. Microbiol.* 63: 3151-3157), Approximately 49,000 random clones were sequenced using dye-terminator chemistry (Applied Biosystems, Foster City, Calif.) with ABI 377 and ABI 3700 auto-

mated sequencers yielding approximately 6x coverage of the genome. A combination of methods was employed for gap closure including sequencing on fosmids (Kim, U. J., Shizuya, H., de Jong, P. J., Birren, B., and Simon, M. I., 1992, *Nucleic Acids Res.* 20: 1083-1085), primer walking on selected clones, and PCR-amplified DNA fragments. Fosmid libraries were constructed using a commercial kit from Epicentre (Madison, Wis.). Data from both ends of approximately 1975 fosmid clones with an average insert size of 40 kb were incorporated to aid in validation of the final assembly. In total, the number of input reads was 62,685 with 78.6% of these incorporated into the final assembly. Sequences were base called using TraceTuner 2.0 (Paracel, Inc., Pasadena, Calif.) and assembled using the Paracel Genome Assembler (Paracel, Inc., Pasadena, Calif.) with optimized parameters and the quality score set to >20. Phrap, Crossmatch, and Consed were used for sequence finishing (Gordon D., Abajian C., and Green P., 1998, *Genome Res.* 8: 195-202).

Example 2

Identification and Annotation of Open Reading Frames (ORFs)

[0251] Protein coding regions in the assembled genome sequence data were identified using EasyGene (Larsen, T. S., and Krogh, A., 2003, *BMC Bioinformatics* 4: 21), Glimmer (Delcher, A. L., Harmon, D., Kasif, S., White, O. and Salzberg, S. L., 1999, *Nucleic Acids Res.* 27, 4636-4641), and FrameD (Schiex, T., Gouzy, J., Moisan, A. and de Oliveira, Y., 2003, *Nucleic Acids Res.* 31, 3738-3741). Only EasyGene gene models with an R-value of less than 2 and log-odds score greater than -10 were used. Predicted proteins were compared to the non-redundant database PIR-NREF (Wu, C. H., Huang, H., Arminski, L., Castro-Alvear, J., Chen, Y., Hu, Z. Z., Ledley, R. S., Lewis, K. C., Mewes, H. W., Orcutt, B. C., 2002, *Nucleic Acids Research* 30: 35-37) and the *Bacillus subtilis* genome (SubtilList) using BLASTP with an E-value threshold of 10^{-5} . InterProScan was used to predict putative function. (Zdobnov, E. M. and Apweiler, R., 2001, *Bioinformatics* 17, 847-848). The InterPro analysis included comparison to Pfam (Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Sonnhammer, E. L., et al., 2004, *Nucleic Acids Res.* 32, D138-D141), TIGRFam (Haft, D. J., Selengut, J. D. and White, O., 2003, *Nucleic Acids Res.* 31: 371-373), Interpro (Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D., et al., 2001, *Nucleic Acids Res.* 29: 37-40), signal peptide prediction using SignalP (Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G., 1997, *Protein Engineering* 10: 1-6), and trans-membrane domain prediction using TMHMM (Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. L., 2000, *J. Mol. Biol.* 305, 567-580). These ORFs were assigned to functional categories based on the Cluster of Orthologous Groups (COG) database (www.ncbi.nlm.nih.gov/COG) with manual verification as described (Tatusov, R. L., Koonin, E. V. and Lipman, D. J., 1997, *Science* 278: 631-637; Koonin, E. V. and Galperin, M. Y., 2002, Sequence—Evolution—Function: Computational Approaches in Comparative Genomics (Kluwer, Boston)). Transfer RNA genes were identified using tRNAscan-SE (Lowe, T. M. and Eddy, S. R., 1997, *Nucleic Acids Res.* 25: 955-964).

Example 3

General Features of the *Bacillus licheniformis* Genome

[0252] The genome of *Bacillus licheniformis* ATCC 14580 was determined to consist of a circular molecule of 4,222,336

bp with an average GC content of 46.2% (Table 2). No plasmids were found during the genome analysis, and none were found by agarose gel electrophoresis.

[0253] The genome contains 4208 predicted protein-coding genes with an average size of 873 bp, seven rRNA operons, and 81 RNA genes. Using a combination of several gene-finding algorithms 4208 protein coding ORFs were predicted. These ORFs constitute 87% of the genome and have an average length of 873 bp. Approximately 48% of the ORFs are encoded on one DNA strand and 52% on the other strand. Among the protein coding ORFs, 3948 (94%) have significant similarity to proteins in PIR, and 3187 of these gene models contain Interpro motifs and 2895 contain protein motifs found in PFAM. The number of hypothetical and conserved hypothetical proteins in the *Bacillus licheniformis* genome with hits in the PIR database was 1318 (212 conserved hypothetical ORFs). Among the list of hypothetical and conserved hypothetical ORFs, 683 (52%) have protein motifs contained in PFAM (148 conserved hypothetical ORFs). There are 72 tRNA genes representing all 20 amino acids and 7 rRNA operons.

[0254] The likely origin of replication was identified by similarities to several features of *Bacillus subtilis* origin (Moriya, S., Fukuoka, T., Ogasawara, N., and Yoshikawa, H., 1988, *EMBO Journal* 7: 2911-2917; Ogasawara, N., Nakai, S., and Yoshikawa, H., 1994, *DNA Res.* 1, 1-14; Kadoya, R., Hassan, A. K., Kasahara, Y., Ogasawara, N., and Moriya, S., 2002, *Mol. Microbiol.* 45: 73-87; Tosato, V., Gjuracic, K., Vlahovick, K., Pongor, S., Danchin, A., and Bruschi, C. V., 2003, *FEMS Microbiol. Lett.* 218: 23-30). These included (a) colocalization of four genes (rpmH, dnaA, dnaN, and recF) found near the origin of the *Bacillus subtilis* chromosome, (b) GC nucleotide skew $[(G-C)/(G+C)]$ analysis, and (c) the presence of multiple dnaA-boxes (Pedersen, A. G., Jensen, L. J., Brunak, S., Staerfeldt, H. H., and Ussery, D. W., 2000, *Mol. Biol.* 299: 907-930; Christensen, B. B., Atlung, T., and Hansen, F. G., 1999, *J. Bacteriol.* 181: 2683-2688; Majka, J., Jakimowicz, D., Messer, W., Schrermpf, H., Lisowski, M., and Zakrzewska-Czerwińska, J., 1999, *Eur. J. Biochem.* 260: 325-335) and AT-rich sequences in the region immediately upstream of the dnaA gene. On the basis of these observations, a cytosine residue of the BstBI restriction site was assigned between the rpnH and dnaA genes to be the first nucleotide of the *Bacillus licheniformis* genome. The replication termination site was localized near 2.02 Mb by GC skew analysis. This region lies roughly opposite the origin of replication.

[0255] Unlike *Bacillus subtilis*, no apparent gene encoding a replication terminator protein (rtp) was found in *Bacillus licheniformis*. The *Bacillus halodurans* genome also lacks an rtp function (Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N. et al., 2000, *Nucleic Acids Res.* 28; 4317-4331), and it seems likely that *Bacillus subtilis* acquired the rip gene following its divergence from *Bacillus halodurans* and *Bacillus licheniformis*.

[0256] Transposable elements and prophages. The genome of *Bacillus licheniformis* ATCC 14580 was determined to contain nine identical copies of a 1285 bp insertion sequence element termed IS3Bli1 (Lapidus, A., Galleron, N., Andersen, J. T., Jørgensen, P. L., Ehrlich, S. D., and Sorokin, A., 2002, *FEMS Microbiol. Lett.* 209: 23-30). This sequence shares a number of features with other IS3 family elements including direct repeats of three to five bp, a ten bp left

inverted repeat, and a nine bp right inverted repeat. IS3Bli1 encodes two predicted overlapping ORFs, designated orfA and orfB in relative translational reading frames of 0 and -1. The presence of a “slippery heptamer” motif, AAAAAAG, before the stop codon in orfA suggests that programmed translational frameshifting occurs between these two ORFs, resulting in a single gene product (Farabaugh, P., 1996, *Microbiol. Rev.* 60: 103-134). The orfB gene product harbors the DD[35]E[7]K motif, a highly conserved pattern among insertion sequences. Eight of the IS3Bli1 elements lie in intergenic regions, and one interrupts the *comP* gene. In addition to these insertion sequences, the genome encodes a putative transposase that is most closely related ($E=1.8e-11$) to one identified in the *Thermoanaerobacter tengcongensis* genome (Bao, Q., Tian, Y., Li, W., Xu, Z.; Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., et al., 2002, *Genome Res.* 12: 689-700), however, similar genes are also found in the chromosomes of *Bacillus halodurans* (Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N. et al., 2000, *Nucleic Acids Res.* 28: 4317-4331), *Oceanobacillus iheyensis* (Takami, H., Takaki, Y., and Uchiyama, I., 2002, *Nucleic Acids Res.* 30: 3927-3935.), *Streptococcus agalactiae* (Takahashi, S., Detrick, S., Whiting, A. A., Blaschke-Bonkowsky, A. J., Aoyagi, Y., Adderson, E. E., and Bohnsack, J. F., 2002, *J. Infect. Dis.* 186: 1034-1038), and *Streptococcus pyogenes* (Smoot, J. C., Barbian, K. D., Van Gompel, J. J., Smoot, L. M., Chaussee, M. S., Sylva, G. L., Sturdevant, D. E., Ricklefs, S. M., Porcella, S. F., Parkins, L. D., et al., 2002, *Proc. Natl. Acad. Sci. U.S.A.* 99: 4668-4673).

[0257] The presence of several bacteriophage lysogens or prophage-like elements was revealed by Smith-Waterman comparisons to other bacterial genomes and by their AT-rich signatures. Prophage sequences, designated NZP1 and NZP3 (similar to PBSX and ϕ -105, respectively), were uncovered by noting the presence of nearby genes encoding the large subunit of terminase, a signature protein that is highly conserved among prophages (Casjens, S., 2003, *Mol. Microbiol.* 49: 277-300). A terminase gene was not observed in a third putative prophage, termed NZP2 (similarity to SPP1), however, its absence may be the result of genome deterioration during evolution. Regions were observed in which the GC content is less than 39% usually encoded proteins that have no *Bacillus subtilis* orthologue and share identity only to hypothetical and conserved hypothetical genes. Two of these AT-rich segments correspond to the NZP2 and NZP3 prophages.

[0258] An isochore plot also revealed the presence of a region with an atypically high (62%) G+C content. This segment contains two hypothetical ORFs whose sizes (3831 and 2865 bp) greatly exceed the size of an average gene in *Bacillus licheniformis*. The first protein encodes a protein of 1277 amino acids for which Interpro predicted 16 collagen triple helix repeats, and the amino acid pattern TGATGPT is repeated 75 times within the polypeptide. The second ORF is smaller, and encodes a protein with 11 collagen triple helix repeats, and the same TGATGPT motif recurs 56 times. Interestingly, the chromosomal region (19 kb) adjacent to these genes is clearly non-colinear with the *Bacillus subtilis* genome, and virtually all of the predicted ORFs are hypothetical or conserved hypothetical proteins. There are a number of bacterial proteins listed in PIR that contain collagen triple helix repeat regions including two from *Mesorhizobium loti* (accession numbers NF00607049 and NF00607035) and three from *Bacillus cereus* (accession numbers NF01692528,

NF01269899, and NF01694666). These putative orthologs share 53-76% amino acid sequence identity with their counterparts in *Bacillus licheniformis*, although their functions are unknown.

[0259] Extracellular enzymes. In the *Bacillus licheniformis* genome, 689 of the 4208 gene models have signal peptides as forecasted by SignalP (Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Henrik Nielsen, Jacob Engelbrecht, Søren Brunak and Gunnar von Heijne, 1997, *Protein Engineering* 10: 1-6). Of these, 309 have no trans-membrane domain as predicted with TMHMM (A. Krogh, B. Larsson, G. von Heijne, and E. L. L. Sonnhammer, 2000, *Journal of Molecular Biology* 305: 567-580) and 134 are hypothetical or conserved hypothetical genes. Based on a manual examination of the remaining 175 ORFs, at least 82 were determined to likely encode secreted proteins and enzymes. The sequence ID numbers for each of these genes are listed in Table 3.

[0260] Protein secretion, sporulation, and competence pathways. Kunst et al. (Kunst, F., Ogasawara, N., Mozser, I., Albertini, A. M., Alloni, G., Azebedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., and Borchert, S., 1997, *Nature* 390: 249-256) listed 18 genes that play a major role in the secretion of extracellular enzymes by *Bacillus subtilis* 168. This list includes several chaperonins, signal peptidases, components of the signal recognition particle and protein translocase complexes. Thus, it is reasonable to expect that the central features of the secretory apparatus are conserved in *Bacillus subtilis* and *Bacillus licheniformis*.

[0261] From the list of 139 sporulation genes tabulated by Kunst et al. (Kunst, F., Ogasawara, N., Mozser, I., Albertini, A. M., Alloni, G., Azebedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., and Borchert, S. et al., 1997, *Nature* 390: 249-256), all but six have obvious counterparts in *Bacillus licheniformis*. These six exceptions (spsABCEFG) comprise an operon involved in synthesis of spore coat polysaccharide in *Bacillus subtilis*. Additionally, the response regulator gene family (phrACEFGI) appears to have a low level of sequence conservation between *Bacillus subtilis* and *Bacillus licheniformis*.

[0262] Natural competence (the ability to take up and process exogenous DNA in specific growth conditions) is a feature of few *Bacillus licheniformis* strains (Gwinn, D. D. and Thorne, C. B., 1964, *J. Bacteriol.* 87: 519-526). The reasons for variability in competence phenotype have not been explored at the genetic level, but the genome data offered several possible explanations. Although the type strain genome encodes all of the late competence functions ascribed in *Bacillus subtilis* (e.g., *comC*, *comEFG* operons, *comK*, *mecA*), it lacks an obvious *comS* gene, and the *comP* gene is punctuated by an insertion sequence element, suggesting that the early stages of competence development have been pre-empted in *Bacillus licheniformis* ATCC 14580. Whether these early functions can be restored by introducing the corresponding genes from *Bacillus subtilis* is unknown. In addition to an apparent deficiency in DNA uptake, two Type I restriction-modification systems were discovered that may also contribute diminished transformation efficiencies. These are distinct from the *ydiOPS* genes of *Bacillus subtilis*, and could participate in degradation of improperly modified DNA from heterologous hosts used during construction of recombinant expression vectors. Lastly, the synthesis of a glutamyl polypeptide capsule has also been implicated as a potential barrier to transformation of *Bacillus licheniformis* strains

(Thorne, C. B. and Stull, H. B., 1966, *J. Bacteriol.* 91: 1012-1020). Six genes were predicted (ywtABDEF and ywsC orthologues) that may be involved in the synthesis of this capsular material.

[0263] Antibiotics and secondary metabolites. Bacitracin is a cyclic peptide antibiotic that is synthesized non-ribosomally in *Bacillus licheniformis* (Katz, E. and Demain, A. L., 1977, *Bacteriol. Rev.* 41: 449-474). While there is variation in the prevalence of bacitracin synthase genes in laboratory strains of this species, one study suggested that up to 50% may harbor the bac operon (Ishihara, H., Takoh, M., Nishibayashi, R., and Sato, A., 2002, *Curr. Microbiol.* 45: 18-23). The bac operon was determined not to be present in the type strain (ATCC 14580) genome. Seemingly, the only non-ribosomal peptide synthase operon encoded by the *Bacillus licheniformis* type strain genome is that which is responsible for lichenysin biosynthesis. Lichenysin structurally resembles surfactin from *Bacillus subtilis* (Grangemard, I., Wallach, J., Maget-Dana, R., and Peypoux, F., 2001, *Appl. Biochem. Biotechnol.* 90: 199-210), and their respective biosynthetic operons are highly similar. No *Bacillus licheniformis* counterparts were found for the pps (plipastatin synthase) and polyketide synthase (pks) operons of *Bacillus subtilis*. Collectively, these two regions represent sizeable portions (80 kb and 38 kb, respectively) of the chromosome in *Bacillus subtilis*, although they are reportedly dispensable (Westers, H., Dorenbos, R., van Dijk, J. M., Kable, J., Flanagan, T., Devine, K. M., Jude, F., Séror, S. J., Beekman, A. C., Darmon, E., 2003, *Mol. Biol. Evol.* 20: 2076-2090). Unexpectedly, a gene cluster was found encoding a lantibiotic and associated processing and transport functions. This peptide of 69 amino acids was designated as lichenicidin, and its closest known orthologue is mersacidin from *Bacillus* sp. strain HIL-Y85/54728 (Altena, K., Guder, A., Cramer, C., and Bierbaum, G., 2000, *Appl. Environ. Microbiol.* 66: 2565-2571). Lantibiotics are ribosomally synthesized peptides that are modified post-translationally so that the final molecules contain rare thioether amino acids such as lanthionine and/or methyl-lanthionine (Pag, U. and Sahl, H. G., 2002, *Curr. Pharm. Des.* 8: 815-833). These antimicrobial compounds have attracted much attention in recent years as models for the design of new antibiotics (Hoffmann, A., Pag, U., Wiedemann, I., and Sahl, H. G., 2002, *Farmaco*, 57: 685-691).

[0264] Essential Genes. The gene models were also compared to the list of essential genes in *Bacillus subtilis* (Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerch, S., Besieres, P., 2003, *Proc. Natl Acad. Sci. USA* 100: 4678-4683). All the essential genes in *Bacillus subtilis* have orthologues in *Bacillus licheniformis*, and most are present in a wide range of bacterial taxa (Pedersen, P. B., Bjørnvad, M. E., Rasmussen, M. D., and Petersen, J. N., 2002, *Reg. Toxicol. Pharmacol.* 36: 155-161).

Example 4

Comparison of *Bacillus licheniformis* Genome with Other Bacilli

[0265] VisualGenome software (Rational Genomics, San Francisco, Calif.) was used for GC-skew analysis and global homology comparisons of the *Bacillus licheniformis*, *Bacillus subtilis*, and *Bacillus halodurans* genomes with pre-computed BLAST results stored in a local database. In pairwise comparisons (E-score threshold of 10^{-5}) 66% (2771/4208) of

the predicted *Bacillus licheniformis* ORFs have orthologs in *Bacillus subtilis*, and 55% (2321/4208) of the gene models are represented by orthologous sequences in *Bacillus halodurans*. Using a reciprocal BLASTP analysis 1719 orthologs were found that are common to all three species (E-score threshold of 10^{-5}).

[0266] As noted by Lapidus et al. (Lapidus, A., Galleron, N., Andersen, J. T., Jørgensen, P. L., Ehrlich, S. D., and Sorokin, A., 2002, *FEMS Microbiol. Lett.* 209: 23-30), there are broad regions of colinearity between the genomes of *Bacillus licheniformis* and *Bacillus subtilis*. Less conservation of genome organization exists between *Bacillus licheniformis* and *Bacillus halodurans*, and inversion of one or more large genomic segments is evident. Clearly this supports previous findings (Xu, D. and Côté, J. C., 2003, *Internat. J. Syst. Evol. Microbiol.* 53: 695-704) that *Bacillus subtilis* and *Bacillus licheniformis* are phylogenetically and evolutionarily closer than either species is to *Bacillus halodurans*. However, a number of important differences were also observed, both in the numbers and locations of prophages and transposable elements and in a number of biochemical pathways, which distinguish *Bacillus licheniformis* from *Bacillus subtilis*, including a region of more than 80 kb that comprises a cluster of polyketide synthase genes that are absent in *Bacillus licheniformis*.

Example 5

Codon Usage Tables

[0267] The evolution of codon bias, the unequal usage of synonymous codons, is thought to be due to natural selection for the use of preferred codons that match the most abundant species of isoaccepting tRNAs, resulting in increased translational efficiency and accuracy. The practical applications for utilizing codon bias information include optimizing expression of heterologous and mutant genes (Jiang and Mannervik, 1999, *Protein Expression and Purification* 15: 92-98), site-directed mutagenesis to derive variant polypeptides from a given gene (Wong et al., 1995, *J. Immunol.* 154: 3351-3358; Kaji, H. et al., 1999, *J. Biochem.* 126: 769-775), design and synthesis of synthetic genes (Libertini and Di Donato, 1992, *Protein Engineering* 5: 821-825; Fang et al., 2000, *Biochem.* 39: 15399-15409), and fine-tuning or reducing of translation efficiency of specific genes by introduction of non-preferred codons (Crombie, T. et al., 1992, *J. Mol. Biol.* 228: 7-12; Carlini and Stephan, 2003, *Genetics* 163: 239-243).

[0268] A codon usage table (Table 4) was generated from SEQ ID NO: 1 with CUSP, a software component of the EMBOSS package (Rice, Longden, and Bleasby, 2000, EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16: 276-277) on all the predicted protein-coding genes of the *Bacillus licheniformis* chromosome. CUSP read the coding sequences and calculated the codon frequency table shown in Table 4.

[0269] A codon usage table (Table 5) was also generated based on the signal peptides of the 82 extracellular proteins described in Example 3.

TABLE 1

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2	Chromosomal replication initiator protein dnaA [<i>Bacillus subtilis</i>]	UniRef100_P05648	<i>Bacillus subtilis</i>	DnaA
3	DNA polymerase III, beta chain [<i>Bacillus subtilis</i>]	UniRef100_P05649	<i>Bacillus subtilis</i>	DnaN
4				
5	DNA replication and repair protein recF [<i>Bacillus subtilis</i>]	UniRef100_P05651	<i>Bacillus subtilis</i>	RecF
6				
7	DNA gyrase subunit B [<i>Bacillus subtilis</i>]	UniRef100_P05652	<i>Bacillus subtilis</i>	GyrB
8	DNA gyrase subunit A [<i>Bacillus subtilis</i>]	UniRef100_P05653	<i>Bacillus subtilis</i>	GyrA
9				YaaC
10	Inosine-5'-monophosphate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P21879	<i>Bacillus subtilis</i>	GuaB
11	D-alanyl-D-alanine carboxypeptidase precursor [<i>Bacillus subtilis</i>]	UniRef100_P08750	<i>Bacillus subtilis</i>	DacA
12	Pyridoxine biosynthesis protein pdx1 [<i>Bacillus subtilis</i>]	UniRef100_P37527	<i>Bacillus subtilis</i>	YaaD
13	Hypothetical UPF0030 protein yaaE [<i>Bacillus subtilis</i>]	UniRef100_P37528	<i>Bacillus subtilis</i>	YaaE
14	Seryl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P37464	<i>Bacillus subtilis</i>	SerS
15	Glycerate kinase [<i>Bacillus subtilis</i>]	UniRef100_P42100	<i>Bacillus subtilis</i>	YxaA
16	H ⁺ /gluconate symporter [<i>Vibrio vulnificus</i>]	UniRef100_Q7MHW6	<i>Vibrio vulnificus</i>	YuiF
17	Sugar diacid utilization regulator [<i>Vibrio vulnificus</i>]	UniRef100_Q8DBZ9	<i>Vibrio vulnificus</i>	YsfB
18	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HH43	<i>Bacillus thuringiensis</i>	
19	Hypothetical protein yaaF [<i>Bacillus subtilis</i>]	UniRef100_P37529	<i>Bacillus subtilis</i>	Dck
20	Hypothetical protein yaaG [<i>Bacillus subtilis</i>]	UniRef100_P37530	<i>Bacillus subtilis</i>	Dgk
21	Hypothetical protein yaaH [<i>Bacillus subtilis</i>]	UniRef100_P37531	<i>Bacillus subtilis</i>	YaaH
22	Hypothetical protein yaaI [<i>Bacillus subtilis</i>]	UniRef100_P37532	<i>Bacillus subtilis</i>	YaaI
23				YaaJ
24				DnaX
25	Hypothetical UPF0133 protein yaaK [<i>Bacillus subtilis</i>]	UniRef100_P24281	<i>Bacillus subtilis</i>	YaaK
26	Recombination protein recR [<i>Bacillus subtilis</i>]	UniRef100_P24277	<i>Bacillus subtilis</i>	RecR
27	Hypothetical protein yaaL [<i>Bacillus subtilis</i>]	UniRef100_P37533	<i>Bacillus subtilis</i>	
28	Sigma-K factor processing regulatory protein BOFA [<i>Bacillus subtilis</i>]	UniRef100_P24282	<i>Bacillus subtilis</i>	
29	CsfB protein [<i>Bacillus subtilis</i>]	UniRef100_P37534	<i>Bacillus subtilis</i>	
30	XpaC protein [<i>Bacillus subtilis</i>]	UniRef100_P37467	<i>Bacillus subtilis</i>	XpaC
31	Hypothetical protein yaaN [<i>Bacillus subtilis</i>]	UniRef100_P37535	<i>Bacillus subtilis</i>	YaaN
32				YaaO
33	Thymidylate kinase [<i>Bacillus subtilis</i>]	UniRef100_P37537	<i>Bacillus subtilis</i>	Trnk
34	Hypothetical protein yaaQ [<i>Bacillus subtilis</i>]	UniRef100_P37538	<i>Bacillus subtilis</i>	YaaQ
35				
36	DNA polymerase III, delta' subunit [<i>Bacillus subtilis</i>]	UniRef100_P37540	<i>Bacillus subtilis</i>	HolB
37	Hypothetical protein yaaT [<i>Bacillus subtilis</i>]	UniRef100_P37541	<i>Bacillus subtilis</i>	YaaT
38	Hypothetical protein yabA [<i>Bacillus subtilis</i>]	UniRef100_P37542	<i>Bacillus subtilis</i>	YabA
39	Hypothetical protein yabB [<i>Bacillus subtilis</i>]	UniRef100_P37543	<i>Bacillus subtilis</i>	YabB
40	Hypothetical UPF0213 protein yazA [<i>Bacillus subtilis</i>]	UniRef100_O31414	<i>Bacillus subtilis</i>	
41	Hypothetical UPF0011 protein yabC [<i>Bacillus subtilis</i>]	UniRef100_P37544	<i>Bacillus subtilis</i>	YabC

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
42	Transition state regulatory protein abrB [<i>Bacillus subtilis</i>]	UniRef100_P08874	<i>Bacillus subtilis</i>	
43	Methionyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P37465	<i>Bacillus subtilis</i>	MetS
44	Putative deoxyribonuclease yabD [<i>Bacillus subtilis</i>]	UniRef100_P37545	<i>Bacillus subtilis</i>	YabD
45				YabE
46	Hypothetical protein yabF [<i>Bacillus subtilis</i>]	UniRef100_P37547	<i>Bacillus subtilis</i>	RnmV
47	Dimethyladenosine transferase (EC 2.1.1.—) (S-adenosylmethionine-6-N',N'-adenosyl(rRNA) dimethyltransferase) [<i>Bacillus subtilis</i>]	UniRef100_P37468	<i>Bacillus subtilis</i>	KsgA
48	Hypothetical protein yabG [<i>Bacillus subtilis</i>]	UniRef100_P37548	<i>Bacillus subtilis</i>	YabG
49	Veg protein [<i>Bacillus subtilis</i>]	UniRef100_P37466	<i>Bacillus subtilis</i>	
50	SspF protein [<i>Bacillus subtilis</i>]	UniRef100_P37549	<i>Bacillus subtilis</i>	
51				IspE
52	Pur operon repressor [<i>Bacillus subtilis</i>]	UniRef100_P37551	<i>Bacillus subtilis</i>	PurR
53	UPF0076 protein yabJ [<i>Bacillus subtilis</i>]	UniRef100_P37552	<i>Bacillus subtilis</i>	YabJ
54	Stage V sporulation protein G [<i>Bacillus subtilis</i>]	UniRef100_P28015	<i>Bacillus subtilis</i>	
55	Bifunctional gcaD protein (TMS protein) [Includes: UDP-N-acetylglucosamine pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase (EC 2.3.1.157)] [<i>Bacillus subtilis</i>]	UniRef100_P14192	Includes: UDP-N-acetylglucosamine pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase (EC 2.3.1.157)	GcaD
56	Ribose-phosphate pyrophosphokinase [<i>Bacillus subtilis</i>]	UniRef100_P14193	<i>Bacillus subtilis</i>	Prs
57	General stress protein etc [<i>Bacillus subtilis</i>]	UniRef100_P14194	<i>Bacillus subtilis</i>	Ctc
58	Peptidyl-tRNA hydrolase [<i>Bacillus subtilis</i>]	UniRef100_P37470	<i>Bacillus subtilis</i>	SpoVC
59	Hypothetical protein yabK [<i>Bacillus subtilis</i>]	UniRef100_P37553	<i>Bacillus subtilis</i>	
60	Transcription-repair coupling factor [<i>Bacillus subtilis</i>]	UniRef100_P37474	<i>Bacillus subtilis</i>	Mfd
61	Stage V sporulation protein T [<i>Bacillus subtilis</i>]	UniRef100_P37554	<i>Bacillus subtilis</i>	SpoVT
62	Hypothetical protein yabM [<i>Bacillus subtilis</i>]	UniRef100_P37555	<i>Bacillus subtilis</i>	YabM
63	Hypothetical protein yabN [<i>Bacillus subtilis</i>]	UniRef100_P37556	<i>Bacillus subtilis</i>	YabN
64	Hypothetical protein yabO [<i>Bacillus subtilis</i>]	UniRef100_P37557	<i>Bacillus subtilis</i>	
65	Hypothetical protein yabP [<i>Bacillus subtilis</i>]	UniRef100_P37558	<i>Bacillus subtilis</i>	YabP
66	Hypothetical protein yabQ [<i>Bacillus subtilis</i>]	UniRef100_P37559	<i>Bacillus subtilis</i>	YabQ
67	Cell division protein divC [<i>Bacillus subtilis</i>]	UniRef100_P37471	<i>Bacillus subtilis</i>	DivC
68	Hypothetical protein yabR [<i>Bacillus subtilis</i>]	UniRef100_P37560	<i>Bacillus subtilis</i>	YabR
69	Stage II sporulation protein E [<i>Bacillus subtilis</i>]	UniRef100_P37475	<i>Bacillus subtilis</i>	SpoIIE
70	Hypothetical protein yabS [<i>Bacillus subtilis</i>]	UniRef100_P37561	<i>Bacillus subtilis</i>	YabS
71	Probable serine/threonine-protein kinase yabT [<i>Bacillus subtilis</i>]	UniRef100_P37562	<i>Bacillus subtilis</i>	YabT

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
72	Hypothetical UPF0072 protein yacA [<i>Bacillus subtilis</i>]	UniRef100_P37563	<i>Bacillus subtilis</i>	YacA
73	Hypoxanthine-guanine phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P37472	<i>Bacillus subtilis</i>	HprT
74	Cell division protein ftsH homolog [<i>Bacillus subtilis</i>]	UniRef100_P37476	<i>Bacillus subtilis</i>	FtsH
75	Putative 32 kDa replication protein [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9F985	<i>Bacillus stearothermophilus</i>	YacB
76	33 kDa chaperonin [<i>Bacillus subtilis</i>]	UniRef100_P37565	<i>Bacillus subtilis</i>	YacC
77				YacD
78				CysK
79	Para-aminobenzoate synthase component I [<i>Bacillus subtilis</i>]	UniRef100_P28820	<i>Bacillus subtilis</i>	PabB
80	Para-aminobenzoate/anthranilate synthase glutamine amidotransferase component II [Includes: Para-aminobenzoate synthase glutamine amidotransferase component II (EC 6.3.5.8) (ADC synthase); Anthranilate synthase component II (EC 4.1.3.27)] [<i>Bacillus subtilis</i>]	UniRef100_P28819	Includes: Para-aminobenzoate synthase glutamine amidotransferase component II (EC 6.3.5.8) (ADC synthase); Anthranilate synthase component II (EC 4.1.3.27)	PabA
81	Aminodeoxychorismate lyase [<i>Bacillus subtilis</i>]	UniRef100_P28821	<i>Bacillus subtilis</i>	PabC
82	Dihydropteroate synthase [<i>Bacillus subtilis</i>]	UniRef100_P28822	<i>Bacillus subtilis</i>	SuI
83	Dihydroneopterin aldolase [<i>Bacillus subtilis</i>]	UniRef100_P28823	<i>Bacillus subtilis</i>	FolB
84	2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase [<i>Bacillus subtilis</i>]	UniRef100_P29252	<i>Bacillus subtilis</i>	FolK
85	YazB protein [<i>Bacillus subtilis</i>]	UniRef100_O31417	<i>Bacillus subtilis</i>	
86	Probable tRNA-dihydrouridine synthase 1 [<i>Bacillus subtilis</i>]	UniRef100_P37567	<i>Bacillus subtilis</i>	YacF
87	Lysyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P37477	<i>Bacillus subtilis</i>	LysS
88	Transcriptional regulator ctsR [<i>Bacillus subtilis</i>]	UniRef100_P37568	<i>Bacillus subtilis</i>	CtsR
89	Hypothetical protein yacH [<i>Bacillus subtilis</i>]	UniRef100_P37569	<i>Bacillus subtilis</i>	McsA
90	Hypothetical ATP:guanido phosphotransferase yacI [<i>Bacillus subtilis</i>]	UniRef100_P37570	<i>Bacillus subtilis</i>	McsB
91	Negative regulator of genetic competence clpC/mecB [<i>Bacillus subtilis</i>]	UniRef100_P37571	<i>Bacillus subtilis</i>	ClpC
92				RadA
93	Hypothetical protein yacK [<i>Bacillus subtilis</i>]	UniRef100_P37573	<i>Bacillus subtilis</i>	YacK
94	Hypothetical protein yacL [<i>Bacillus subtilis</i>]	UniRef100_Q06754	<i>Bacillus subtilis</i>	YacL
95	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase [<i>Bacillus subtilis</i>]	UniRef100_Q06755	<i>Bacillus subtilis</i>	YacM
96	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase [<i>Bacillus subtilis</i>]	UniRef100_Q06756	<i>Bacillus subtilis</i>	YacN
97	Glutamyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P22250	<i>Bacillus subtilis</i>	Gltx
98	Serine acetyltransferase [<i>Bacillus subtilis</i>]	UniRef100_Q06750	<i>Bacillus subtilis</i>	CysE
99	CysteinyI-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_Q06752	<i>Bacillus subtilis</i>	CysS
100	YazC protein [<i>Bacillus subtilis</i>]	UniRef100_O31418	<i>Bacillus subtilis</i>	YazC

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
101	Hypothetical tRNA/rRNA methyltransferase yacO [<i>Bacillus subtilis</i>]	UniRef100_Q06753	<i>Bacillus subtilis</i>	YacO
102	Hypothetical protein yacP [<i>Bacillus subtilis</i>]	UniRef100_P37574	<i>Bacillus subtilis</i>	YacP
103	RNA polymerase sigma-H factor [<i>Bacillus subtilis</i>]	UniRef100_P17869	<i>Bacillus subtilis</i>	SigH
104	Preprotein translocase secE subunit [<i>Bacillus subtilis</i>]	UniRef100_Q06799	<i>Bacillus subtilis</i>	
105				NusG
106	50S ribosomal protein L11 [<i>Bacillus subtilis</i>]	UniRef100_Q06796	<i>Bacillus subtilis</i>	RplK
107	50S ribosomal protein L1 [<i>Bacillus subtilis</i>]	UniRef100_Q06797	<i>Bacillus subtilis</i>	RplA
108	50S ribosomal protein L10 [<i>Bacillus subtilis</i>]	UniRef100_P42923	<i>Bacillus subtilis</i>	RplJ
109	50S ribosomal protein L7/L12 [<i>Bacillus subtilis</i>]	UniRef100_P02394	<i>Bacillus subtilis</i>	RplL
110	Hypothetical protein yxbB [<i>Bacillus subtilis</i>]	UniRef100_P37872	<i>Bacillus subtilis</i>	YxbB
111	DNA-directed RNA polymerase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P37870	<i>Bacillus subtilis</i>	RpoB
112	DNA-directed RNA polymerase beta' chain [<i>Bacillus subtilis</i>]	UniRef100_P37871	<i>Bacillus subtilis</i>	RpoC
113	Putative ribosomal protein L7Ae-like [<i>Bacillus subtilis</i>]	UniRef100_P46350	<i>Bacillus subtilis</i>	
114	30S ribosomal protein S12 [<i>Bacillus subtilis</i>]	UniRef100_P21472	<i>Bacillus subtilis</i>	RpsL
115	30S ribosomal protein S7 [<i>Bacillus subtilis</i>]	UniRef100_P21469	<i>Bacillus subtilis</i>	RpsG
116	Elongation factor G [<i>Bacillus subtilis</i>]	UniRef100_P80868	<i>Bacillus subtilis</i>	FusA
117	Elongation factor Tu [<i>Bacillus subtilis</i>]	UniRef100_P33166	<i>Bacillus subtilis</i>	TufA
118	30S ribosomal protein S10 [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9L5	<i>Bacillus halodurans</i>	RpsJ
119	50S ribosomal protein L3 [<i>Bacillus subtilis</i>]	UniRef100_P42920	<i>Bacillus subtilis</i>	RplC
120	50S ribosomal protein L4 [<i>Bacillus subtilis</i>]	UniRef100_P42921	<i>Bacillus subtilis</i>	RplD
121	50S ribosomal protein L23 [<i>Bacillus subtilis</i>]	UniRef100_P42924	<i>Bacillus subtilis</i>	
122	50S ribosomal protein L2 [<i>Bacillus subtilis</i>]	UniRef100_P42919	<i>Bacillus subtilis</i>	RplB
123	30S ribosomal protein S19 [<i>Bacillus subtilis</i>]	UniRef100_P21476	<i>Bacillus subtilis</i>	
124	50S ribosomal protein L22 [<i>Bacillus subtilis</i>]	UniRef100_P42060	<i>Bacillus subtilis</i>	RplV
125	30S ribosomal protein S3 [<i>Bacillus subtilis</i>]	UniRef100_P21465	<i>Bacillus subtilis</i>	RpsC
126	50S ribosomal protein L16 [<i>Bacillus subtilis</i>]	UniRef100_P14577	<i>Bacillus subtilis</i>	RplP
127	50S ribosomal protein L29 [<i>Bacillus subtilis</i>]	UniRef100_P12873	<i>Bacillus subtilis</i>	
128	30S ribosomal protein S17 [<i>Bacillus subtilis</i>]	UniRef100_P12874	<i>Bacillus subtilis</i>	
129				
130	50S ribosomal protein L24 [<i>Bacillus subtilis</i>]	UniRef100_P12876	<i>Bacillus subtilis</i>	RplX
131	50S ribosomal protein L5 [<i>Bacillus subtilis</i>]	UniRef100_P12877	<i>Bacillus subtilis</i>	RplE
132	Ribosomal protein S14 [<i>Bacillus cereus</i> ZK]	UniRef100_Q63H77	<i>Bacillus cereus</i> ZK	
133	30S ribosomal protein S8 [<i>Bacillus subtilis</i>]	UniRef100_P12879	<i>Bacillus subtilis</i>	RpsH
134	50S ribosomal protein L6 [<i>Bacillus subtilis</i>]	UniRef100_P46898	<i>Bacillus subtilis</i>	RplF
135	50S ribosomal protein L18 [<i>Bacillus subtilis</i>]	UniRef100_P46899	<i>Bacillus subtilis</i>	RplR

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
136	30S ribosomal protein S5 [<i>Bacillus subtilis</i>]	UniRef100_P21467	<i>Bacillus subtilis</i>	RpsE
137	50S ribosomal protein L30 [<i>Bacillus subtilis</i>]	UniRef100_P19947	<i>Bacillus subtilis</i>	
138	50S ribosomal protein L15 [<i>Bacillus subtilis</i>]	UniRef100_P19946	<i>Bacillus subtilis</i>	RplO
139				SecY
140	Adenylate kinase [<i>Bacillus subtilis</i>]	UniRef100_P16304	<i>Bacillus subtilis</i>	Adk
141	Methionine aminopeptidase [<i>Bacillus subtilis</i>]	UniRef100_P19994	<i>Bacillus subtilis</i>	Map
142	C-125 initiation factor IF-I, RNA polymerase alpha subunit and ribosomal proteins, partial and complete cds [<i>Bacillus halodurans</i>]	UniRef100_O50629	<i>Bacillus halodurans</i>	
143	Translation initiation factor IF-1 [<i>Bacillus subtilis</i>]	UniRef100_P20458	<i>Bacillus subtilis</i>	
144				
145				RpsM
146	30S ribosomal protein S11 [<i>Bacillus subtilis</i>]	UniRef100_P04969	<i>Bacillus subtilis</i>	RpsK
147	DNA-directed RNA polymerase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P20429	<i>Bacillus subtilis</i>	RpoA
148	50S ribosomal protein L17 [<i>Bacillus subtilis</i>]	UniRef100_P20277	<i>Bacillus subtilis</i>	RplQ
149	Hypothetical ABC transporter ATP-binding protein ybxA [<i>Bacillus subtilis</i>]	UniRef100_P40735	<i>Bacillus subtilis</i>	YbxA
150	Hypothetical protein orf5 [<i>Bacillus subtilis</i>]	UniRef100_P70970	<i>Bacillus subtilis</i>	YbaE
151	YbaF protein [<i>Bacillus subtilis</i>]	UniRef100_P70972	<i>Bacillus subtilis</i>	YbaF
152	tRNA pseudouridine synthase A [<i>Bacillus subtilis</i>]	UniRef100_P70973	<i>Bacillus subtilis</i>	TruA
153	50S ribosomal protein L13 [<i>Bacillus subtilis</i>]	UniRef100_P70974	<i>Bacillus subtilis</i>	RplM
154	30S ribosomal protein S9 [<i>Bacillus subtilis</i>]	UniRef100_P21470	<i>Bacillus subtilis</i>	RpsI
155	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q737T6	<i>Bacillus cereus</i>	YizA
156	Hypothetical protein ybaK [<i>Bacillus subtilis</i>]	UniRef100_P50862	<i>Bacillus subtilis</i>	YbaK
157	Germination-specific N-acetylmuramoyl-L-alanine amidase [<i>Bacillus subtilis</i>]	UniRef100_P50864	<i>Bacillus subtilis</i>	CwlD
158				YbaL
159	Spore germination protein gerD precursor [<i>Bacillus subtilis</i>]	UniRef100_P16450	<i>Bacillus subtilis</i>	GerD
160	KinB signaling pathway activation protein [<i>Bacillus subtilis</i>]	UniRef100_P16449	<i>Bacillus subtilis</i>	KbaA
161	Hypothetical protein ybaN precursor [<i>Bacillus subtilis</i>]	UniRef100_P50865	<i>Bacillus subtilis</i>	YbaN
162	Penicillin-binding protein [<i>Bacillus subtilis</i>]	UniRef100_O31773	<i>Bacillus subtilis</i>	PbpX
163				
164	Hypothetical protein ybaS [<i>Bacillus subtilis</i>]	UniRef100_P55190	<i>Bacillus subtilis</i>	YbaS
165	Hypothetical protein yxaJ [<i>Bacillus subtilis</i>]	UniRef100_P42109	<i>Bacillus subtilis</i>	YxaJ
166	Phenazine biosynthetic protein [<i>Halobacterium</i> sp.]	UniRef100_Q9HHG6	<i>Halobacterium</i> sp.	YfhB
167	Hypothetical protein ybbC precursor [<i>Bacillus subtilis</i>]	UniRef100_P40407	<i>Bacillus subtilis</i>	YbbC
168	Hypothetical lipoprotein ybbD precursor [<i>Bacillus subtilis</i>]	UniRef100_P40406	<i>Bacillus subtilis</i>	YbbD
169	Hypothetical UPF0214 protein ybbE precursor [<i>Bacillus subtilis</i>]	UniRef100_O05213	<i>Bacillus subtilis</i>	YbbE
170	YbbF protein [<i>Bacillus subtilis</i>]	UniRef100_Q797S1	<i>Bacillus subtilis</i>	YbbF
171	Putative HTH-type transcriptional regulator ybbH [<i>Bacillus subtilis</i>]	UniRef100_Q45581	<i>Bacillus subtilis</i>	YbbH
172	Hypothetical protein ybbI [<i>Bacillus subtilis</i>]	UniRef100_Q45582	<i>Bacillus subtilis</i>	YbbI

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
173	Hypothetical protein ybbK [<i>Bacillus subtilis</i>]	UniRef100_Q45584	<i>Bacillus subtilis</i>	YbbK
174	Arginase [<i>Bacillus caldovelox</i>]	UniRef100_P53608	<i>Bacillus caldovelox</i>	RocF
175	RNA polymerase sigma factor sigW [<i>Bacillus subtilis</i>]	UniRef100_Q45585	<i>Bacillus subtilis</i>	SigW
176	YbbM protein [<i>Bacillus subtilis</i>]	UniRef100_Q45588	<i>Bacillus subtilis</i>	YbbM
177	YbbP protein [<i>Bacillus subtilis</i>]	UniRef100_Q45589	<i>Bacillus subtilis</i>	YbbP
178	YbbR protein [<i>Bacillus subtilis</i>]	UniRef100_O34659	<i>Bacillus subtilis</i>	YbbR
179	YbbT protein [<i>Bacillus subtilis</i>]	UniRef100_O34824	<i>Bacillus subtilis</i>	YbbT
180	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] [<i>Bacillus subtilis</i>]	UniRef100_P39754	isomerizing	GlnS
181				
182	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HH47	<i>Bacillus thuringiensis</i>	
183	UPI00003CBF92 UniRef100 entry	UniRef100_UPI00003CBF92		YtrB
184	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KF35	<i>Bacillus halodurans</i>	YtrA
185				
186	YbcL protein [<i>Bacillus subtilis</i>]	UniRef100_O34663	<i>Bacillus subtilis</i>	YbcL
187	BH0186 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KGB8	<i>Bacillus halodurans</i>	
188	ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KEY6	<i>Bacillus halodurans</i>	YvcC
189	Hypothetical protein ywbO [<i>Bacillus subtilis</i>]	UniRef100_P39598	<i>Bacillus subtilis</i>	YwbO
190				
191	BH0695 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KF04	<i>Bacillus halodurans</i>	
192	Aminoglycoside 6-adenylyltransferase [<i>Enterococcus faecium</i>]	UniRef100_Q6V4U6	<i>Enterococcus faecium</i>	AadK
193	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ERX2	<i>Oceanobacillus iheyensis</i>	
194	YfnB [<i>Bacillus subtilis</i>]	UniRef100_O06480	<i>Bacillus subtilis</i>	YfnB
195	Hypothetical transport protein ybxG [<i>Bacillus subtilis</i>]	UniRef100_P54425	<i>Bacillus subtilis</i>	YbxG
196	Mg(2+)/citrate complex secondary transporter [<i>Bacillus subtilis</i>]	UniRef100_P55069	<i>Bacillus subtilis</i>	CitM
197	YfiP protein [<i>Bacillus subtilis</i>]	UniRef100_O34439	<i>Bacillus subtilis</i>	YfiP
198				CitT
199	Sensor protein citS [<i>Bacillus subtilis</i>]	UniRef100_O34427	<i>Bacillus subtilis</i>	CitS
200	Transcriptional regulator [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RB37	<i>Thermoanaerobacter tengcongensis</i>	YwfK
201	Complete genome; segment 8/17 [<i>Photobacterium luminescens</i>]	UniRef100_Q7N4W7	<i>Photobacterium luminescens</i>	YwfE
202	Multidrug resistance protein B [<i>Bacillus cereus</i>]	UniRef100_Q81AF0	<i>Bacillus cereus</i>	YqjV
203	Sigma-G-dependent sporulation specific SASP protein [<i>Bacillus subtilis</i>]	UniRef100_P54379	<i>Bacillus subtilis</i>	
204	Hypothetical protein ybxH [<i>Bacillus subtilis</i>]	UniRef100_P54426	<i>Bacillus subtilis</i>	
205	YbyB protein [<i>Bacillus subtilis</i>]	UniRef100_O31441	<i>Bacillus subtilis</i>	
206	Hypothetical protein yyaL [<i>Bacillus subtilis</i>]	UniRef100_P37512	<i>Bacillus subtilis</i>	YyaL
207	Hypothetical protein yyaO [<i>Bacillus subtilis</i>]	UniRef100_P37509	<i>Bacillus subtilis</i>	
208	Threonyl-tRNA synthetase 2 [<i>Bacillus subtilis</i>]	UniRef100_P18256	<i>Bacillus subtilis</i>	ThrZ
209				
210				YttB
211	Hypothetical protein [<i>Acinetobacter</i> sp.]	UniRef100_Q6FDN0	<i>Acinetobacter</i> sp.	
212	Hypothetical UPF0053 protein yrkA [<i>Bacillus subtilis</i>]	UniRef100_P54428	<i>Bacillus subtilis</i>	YrkA
213				YdeG
214				YbaJ
215				
216				YqcW
217	YubF protein [<i>Bacillus subtilis</i>]	UniRef100_O32082	<i>Bacillus subtilis</i>	YubF
218				

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
219	Hypothetical protein [<i>Bacteroides fragilis</i>]	UniRef100_Q64UG6	<i>Bacteroides fragilis</i>	
220	Hypothetical transport protein ybhF [<i>Bacillus subtilis</i>]	UniRef100_O31448	<i>Bacillus subtilis</i>	YbhF
221	YbfI protein [<i>Bacillus subtilis</i>]	UniRef100_O31449	<i>Bacillus subtilis</i>	YbfI
222				
223	UPI00003CBC98 UniRef100 entry	UniRef100_UPI00003CBC98		
224				
225	Oxidoreductase [<i>Lactococcus lactis</i>]	UniRef100_Q9CFC5	<i>Lactococcus lactis</i>	
226	Methyl-accepting chemotaxis protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EST3	<i>Oceanobacillus iheyensis</i>	YvaQ
227	D-xylose-binding protein [<i>Thermoanaerobacter ethanolicus</i>]	UniRef100_O68456	<i>Thermoanaerobacter ethanolicus</i>	RbsB
228	Hypothetical protein OB0544 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESS4	<i>Oceanobacillus iheyensis</i>	YomI
229	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81IR4	<i>Bacillus cereus</i>	
230	Streptogramin B lactonase [<i>Staphylococcus cohnii</i>]	UniRef100_O87275	<i>Staphylococcus cohnii</i>	
231	Hypothetical protein ycbP [<i>Bacillus subtilis</i>]	UniRef100_P42248	<i>Bacillus subtilis</i>	YcbP
232	YbgF protein [<i>Bacillus subtilis</i>]	UniRef100_O31462	<i>Bacillus subtilis</i>	YbgF
233	YbgG protein [<i>Bacillus subtilis</i>]	UniRef100_O31463	<i>Bacillus subtilis</i>	YbgG
234				
235	PTS system, n-acetylglucosamine-specific enzyme II, ABC component [<i>Bacillus halodurans</i>]	UniRef100_Q9KF24	<i>Bacillus halodurans</i>	NagP
236	UPI00003CBDDB UniRef100 entry	UniRef100_UPI00003CBDDB		Mta
237	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HIW0	<i>Bacillus thuringiensis</i>	YdfS
238	Carbonic anhydrase [<i>Methanosarcina mazei</i>]	UniRef100_Q8PSJ1	<i>Methanosarcina mazei</i>	YtoA
239	Hypothetical protein yjgA [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HHU2	<i>Bacillus thuringiensis</i>	YjgA
240	Amino acid carrier protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGU4	<i>Bacillus thuringiensis</i>	YbgH
241	Probable glutaminase ybgJ [<i>Bacillus subtilis</i>]	UniRef100_O31465	<i>Bacillus subtilis</i>	YbgJ
242	Two-component sensor kinase ycbA [<i>Bacillus cereus</i>]	UniRef100_Q81BN8	<i>Bacillus cereus</i>	YcbA
243	Hypothetical sensory transduction protein ycbB [<i>Bacillus subtilis</i>]	UniRef100_P40759	<i>Bacillus subtilis</i>	YcbB
244				YoaA
245				YbxI
246				
247	Beta-lactamase precursor (EC 3.5.2.6) (Penicillinase) [Contains: Large exopenicillinase; Small exopenicillinase] [<i>Bacillus licheniformis</i>]	UniRef100_P00808	Contains: Large exopenicillinase; Small exopenicillinase	PenP
248	Alkaline phosphatase D precursor [<i>Bacillus subtilis</i>]	UniRef100_P42251	<i>Bacillus subtilis</i>	PhoD
249				
250	Hypothetical protein ycbT [<i>Bacillus subtilis</i>]	UniRef100_P42252	<i>Bacillus subtilis</i>	TatCD
251				YcbC
252	Probable aldehyde dehydrogenase ycbD [<i>Bacillus subtilis</i>]	UniRef100_P42236	<i>Bacillus subtilis</i>	YcbD
253	Probable glucarate transporter [<i>Bacillus subtilis</i>]	UniRef100_P42237	<i>Bacillus subtilis</i>	YcbE
254	Probable glucarate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P42238	<i>Bacillus subtilis</i>	YcbF
255	Hypothetical transcriptional regulator ycbG [<i>Bacillus subtilis</i>]	UniRef100_P42239	<i>Bacillus subtilis</i>	YcbG
256	Probable D-galactarate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P42240	<i>Bacillus subtilis</i>	YcbH
257	Hypothetical sensory transduction protein ycbL [<i>Bacillus subtilis</i>]	UniRef100_P42244	<i>Bacillus subtilis</i>	YcbL
258	Sensor histidine kinase [<i>Bacillus cereus</i> ZK]	UniRef100_Q633Q6	<i>Bacillus cereus</i> ZK	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
259	Hypothetical ABC transporter ATP-binding protein ycbN [<i>Bacillus subtilis</i>]	UniRef100_P42246	<i>Bacillus subtilis</i>	YcbN
260	Hypothetical protein ycbO [<i>Bacillus subtilis</i>]	UniRef100_P42247	<i>Bacillus subtilis</i>	YcbO
261				YcbO
262	YfnK [<i>Bacillus subtilis</i>]	UniRef100_O06490	<i>Bacillus subtilis</i>	YetN
263	Hypothetical protein ybdO [<i>Bacillus subtilis</i>]	UniRef100_O31437	<i>Bacillus subtilis</i>	YbdO
264	Hypothetical protein ycbJ [<i>Bacillus subtilis</i>]	UniRef100_P42242	<i>Bacillus subtilis</i>	YcbJ
265	Hypothetical protein ywhA [<i>Bacillus subtilis</i>]	UniRef100_P70993	<i>Bacillus subtilis</i>	YwhA
266				YdaB
267	AII0778 protein [<i>Anabaena</i> sp.]	UniRef100_Q8YYR8	<i>Anabaena</i> sp.	
268	BH1298 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDB5	<i>Bacillus halodurans</i>	YbdN
269	BH1299 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDB4	<i>Bacillus halodurans</i>	
270	Transcriptional activator tipA, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81XN3	<i>Bacillus anthracis</i>	Mta
271				
272	Transcriptional regulator, PadR family [<i>Bacillus cereus</i>]	UniRef100_Q81BZ0	<i>Bacillus cereus</i>	
273	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q639U7	<i>Bacillus cereus</i> ZK	
274	Tryptophan RNA-binding attenuator protein-inhibitory protein [<i>Bacillus subtilis</i>]	UniRef100_O31466	<i>Bacillus subtilis</i>	
275	Hypothetical transport protein ycbK [<i>Bacillus subtilis</i>]	UniRef100_P42243	<i>Bacillus subtilis</i>	YcbK
276				YezC
277	Hypothetical protein yccF [<i>Bacillus subtilis</i>]	UniRef100_O34478	<i>Bacillus subtilis</i>	YccF
278	SpaF [<i>Bacillus subtilis</i>]	UniRef100_Q45404	<i>Bacillus subtilis</i>	YhcH
279	SpaE [<i>Bacillus subtilis</i>]	UniRef100_O52853	<i>Bacillus subtilis</i>	
280	Putative SpaG [<i>Bacillus subtilis</i>]	UniRef100_Q93GG9	<i>Bacillus subtilis</i>	
281	Subtilin biosynthesis regulatory protein spaR [<i>Bacillus subtilis</i>]	UniRef100_P33112	<i>Bacillus subtilis</i>	YycF
282	Putative histidine kinase [<i>Bacillus subtilis</i>]	UniRef100_Q93GG7	<i>Bacillus subtilis</i>	ResE
283	UPI00003CA401 UniRef100 entry	UniRef100_UPI00003CA401		YtII
284	YusQ protein [<i>Bacillus subtilis</i>]	UniRef100_O32183	<i>Bacillus subtilis</i>	YusQ
285	Complete genome; segment 6/17 [<i>Photorhabdus luminescens</i>]	UniRef100_Q7N6K9	<i>Photorhabdus luminescens</i>	YusR
286	Methyltransferase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HK82	<i>Bacillus thuringiensis</i>	
287	Hypothetical lipoprotein ycdA precursor [<i>Bacillus subtilis</i>]	UniRef100_O34538	<i>Bacillus subtilis</i>	YcdA
288	UPI00003CB481 UniRef100 entry	UniRef100_UPI00003CB481		YcgA
289	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESQ2	<i>Oceanobacillus iheyensis</i>	YtnL
290	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ET83	<i>Oceanobacillus iheyensis</i>	YvbK
291				GatA
292	Oligoendopeptidase F, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81JJ8	<i>Bacillus anthracis</i>	YjbG
293				
294	UPI00003CC42D UniRef100 entry	UniRef100_UPI00003CC42D		PbpX
295	Hypothetical protein ysfD [<i>Bacillus subtilis</i>]	UniRef100_P94534	<i>Bacillus subtilis</i>	YsfD
296	Hypothetical protein ysfC [<i>Bacillus subtilis</i>]	UniRef100_P94535	<i>Bacillus subtilis</i>	YsfC
297	PTS system, cellobiose-specific enzyme II, B component [<i>Bacillus halodurans</i>]	UniRef100_Q9KEE3	<i>Bacillus halodurans</i>	
298	PTS system, cellobiose-specific enzyme II, A component [<i>Bacillus halodurans</i>]	UniRef100_Q9KEE2	<i>Bacillus halodurans</i>	LicA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
299	PTS system, cellobiose-specific enzyme II, C component [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EP43	<i>Oceanobacillus iheyensis</i>	YwbA
300	6-phospho-beta-glucosidase [<i>Bacillus halodurans</i>]	UniRef100_Q9KEE0	<i>Bacillus halodurans</i>	LicH
301	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KED8	<i>Bacillus halodurans</i>	YbgA
302				
303				YvbX
304	Chitinase precursor [<i>Streptomyces olivaceoviridis</i>]	UniRef100_Q9L3E8	<i>Streptomyces olivaceoviridis</i>	
305	Extracellular metalloprotease precursor [<i>Bacillus subtilis</i>]	UniRef100_P39790	<i>Bacillus subtilis</i>	Mpr
306	Glucose 1-dehydrogenase II [<i>Bacillus subtilis</i>]	UniRef100_P80869	<i>Bacillus subtilis</i>	YcdF
307	Hypothetical protein OB0244 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETL6	<i>Oceanobacillus iheyensis</i>	
308	Hypothetical protein ycdC [<i>Bacillus subtilis</i>]	UniRef100_O34772	<i>Bacillus subtilis</i>	YcdC
309				
310	Cell wall hydrolase cwIJ [<i>Bacillus subtilis</i>]	UniRef100_P42249	<i>Bacillus subtilis</i>	CwIJ
311	Hypothetical protein yceB [<i>Bacillus subtilis</i>]	UniRef100_O34504	<i>Bacillus subtilis</i>	YceB
312	Hypothetical protein yvcE [<i>Bacillus subtilis</i>]	UniRef100_P40767	<i>Bacillus subtilis</i>	YvcE
313	ABC transporter, permease [<i>Bacillus cereus</i> ZK]	UniRef100_Q637N2	<i>Bacillus cereus</i> ZK	YfiM
314	ABC transporter, ATP-binding protein [<i>Bacillus cereus</i>]	UniRef100_Q733P8	<i>Bacillus cereus</i>	YfiL
315	UPI00003CC482 UniRef100 entry	UniRef100_UPI00003CC482		YsiA
316	Hypothetical protein OB3113 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELV2	<i>Oceanobacillus iheyensis</i>	
317	BH3953 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5Y3	<i>Bacillus halodurans</i>	
318	BH3951 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5Y4	<i>Bacillus halodurans</i>	PadR
319	Stress response protein SCP2 [<i>Bacillus subtilis</i>]	UniRef100_P81100	<i>Bacillus subtilis</i>	YceC
320	General stress protein 16U [<i>Bacillus subtilis</i>]	UniRef100_P80875	<i>Bacillus subtilis</i>	YceD
321	Hypothetical protein yceE [<i>Bacillus subtilis</i>]	UniRef100_O34384	<i>Bacillus subtilis</i>	YceE
322	Hypothetical protein yceF [<i>Bacillus subtilis</i>]	UniRef100_O34447	<i>Bacillus subtilis</i>	YceF
323	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q72YD0	<i>Bacillus cereus</i>	
324	YceG [<i>Bacillus subtilis</i>]	UniRef100_O34809	<i>Bacillus subtilis</i>	YceG
325	Hypothetical protein yceH [<i>Bacillus subtilis</i>]	UniRef100_O34833	<i>Bacillus subtilis</i>	YceH
326	UPI00003CB694 UniRef100 entry	UniRef100_UPI00003CB694		CcdA
327	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELB4	<i>Oceanobacillus iheyensis</i>	YdiL
328				Mta
329	Nitrate transporter [<i>Bacillus subtilis</i>]	UniRef100_P42432	<i>Bacillus subtilis</i>	NasA
330	L-lactate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P13714	<i>Bacillus subtilis</i>	Ldh
331	L-lactate permease [<i>Bacillus subtilis</i>]	UniRef100_P55910	<i>Bacillus subtilis</i>	LctP
332	YcgF protein [<i>Bacillus subtilis</i>]	UniRef100_P94381	<i>Bacillus subtilis</i>	YcgF
333	Homologue of aromatic amino acids transport protein of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94383	<i>Bacillus subtilis</i>	YcgH
334	NH(3)-dependent NAD(+) synthetase [<i>Bacillus subtilis</i>]	UniRef100_P08164	<i>Bacillus subtilis</i>	NadE
335	Shikimate kinase [<i>Bacillus subtilis</i>]	UniRef100_P37944	<i>Bacillus subtilis</i>	AroK
336	YcgL protein [<i>Bacillus subtilis</i>]	UniRef100_P94389	<i>Bacillus subtilis</i>	YcgL
337	Proline dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_Q8RL79	<i>Bacillus subtilis</i>	YcgM
338	1-pyrroline-5-carboxylate dehydrogenase 2 [<i>Bacillus subtilis</i>]	UniRef100_P94391	<i>Bacillus subtilis</i>	YcgN

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
339	Homologue of proline permease of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94392	<i>Bacillus subtilis</i>	YcgO
340	Hypothetical protein ycgP [<i>Bacillus subtilis</i>]	UniRef100_P94393	<i>Bacillus subtilis</i>	YcgP
341	YcgQ protein [<i>Bacillus subtilis</i>]	UniRef100_P94394	<i>Bacillus subtilis</i>	YcgQ
342				YcgR
343	Cephalosporin-C deacetylase [<i>Bacillus subtilis</i>]	UniRef100_Q59233	<i>Bacillus subtilis</i>	Cah
344	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KF41	<i>Bacillus halodurans</i>	YdgG
345				
346				
347	Probable amino-acid ABC transporter permease protein yckA [<i>Bacillus subtilis</i>]	UniRef100_P42399	<i>Bacillus subtilis</i>	YckA
348	Probable ABC transporter extracellular binding protein yckB precursor [<i>Bacillus subtilis</i>]	UniRef100_P42400	<i>Bacillus subtilis</i>	YckB
349	NAD(P)H dehydrogenase (Quinone); possible modulator of drug activity B [<i>Bacillus cereus</i> ZK]	UniRef100_Q63FG6	<i>Bacillus cereus</i> ZK	YwrO
350	BH0315 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KG01	<i>Bacillus halodurans</i>	
351	Probable beta-glucosidase [<i>Bacillus subtilis</i>]	UniRef100_P42403	<i>Bacillus subtilis</i>	YckE
352	Nin [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70KK3	<i>Bacillus amyloliquefaciens</i>	Nin
353	DNA-entry nuclease [<i>Bacillus subtilis</i>]	UniRef100_P12667	<i>Bacillus subtilis</i>	NucA
354				YbdM
355	Methyl-accepting chemotaxis protein tlpC [<i>Bacillus subtilis</i>]	UniRef100_P39209	<i>Bacillus subtilis</i>	TlpC
356	Hypothetical protein yqcG [<i>Bacillus subtilis</i>]	UniRef100_P45942	<i>Bacillus subtilis</i>	YqcG
357	Hydantoin utilization protein A [<i>Bordetella bronchiseptica</i>]	UniRef100_Q7WCS3	<i>Bordetella bronchiseptica</i>	
358	Hydantoin utilization protein B [<i>Rhizobium loti</i>]	UniRef100_Q987J6	<i>Rhizobium loti</i>	
359	UPI000027D233 UniRef100 entry	UniRef100_UPI000027D233		YvfK
360	358aa long hypothetical transporter ATP-binding protein [<i>Aeropyrum permix</i>]	UniRef100_Q9YB65	<i>Aeropyrum permix</i>	MsmX
361	AttA2-like ABC transporter, permease protein [<i>Rhizobium meliloti</i>]	UniRef100_Q92ZH0	<i>Rhizobium meliloti</i>	YesP
362				YurM
363				SrfAA
364				SrfAB
365				SrfAC
366				SrfAD
367				YcxC
368				YcxD
369	4'-phosphopantetheinyl transferase sfp [<i>Bacillus subtilis</i>]	UniRef100_P39135	<i>Bacillus subtilis</i>	Sfp
370	Predicted esterase of alpha/beta hydrolase superfamily, YBBA <i>B. subtilis</i> ortholog [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HP7	<i>Clostridium acetobutylicum</i>	YbbA
371	Transcriptional regulator [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97LX8	<i>Clostridium acetobutylicum</i>	YdeE
372	YczE [<i>Bacillus subtilis</i>]	UniRef100_Q9F4F8	<i>Bacillus subtilis</i>	YczE
373	Hypothetical protein [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67MZ7	<i>Symbiobacterium thermophilum</i>	
374				YfiL
375				
376	Hypothetical protein [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67MZ9	<i>Symbiobacterium thermophilum</i>	
377				
378				
379	Methyl-accepting chemotaxis protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HJV7	<i>Bacillus thuringiensis</i>	TlpB
380	YckI [<i>Bacillus subtilis</i>]	UniRef100_Q9F4F9	<i>Bacillus subtilis</i>	YckI
381	YckJ [<i>Bacillus subtilis</i>]	UniRef100_Q9F4G0	<i>Bacillus subtilis</i>	YckJ

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
382				YckK
383				RocR
384	Ornithine aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P38021	<i>Bacillus subtilis</i>	RocD
385	Amino-acid permease rocE [<i>Bacillus subtilis</i>]	UniRef100_P39137	<i>Bacillus subtilis</i>	RocE
386	Arginase [<i>Bacillus subtilis</i>]	UniRef100_P39138	<i>Bacillus subtilis</i>	RocF
387	Homologue of als operon regulatory protein AlsR of <i>B. subtilis</i> [<i>Bacillus subtilis</i>]	UniRef100_P94403	<i>Bacillus subtilis</i>	YclA
388	Probable aromatic acid decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P94404	<i>Bacillus subtilis</i>	YclB
389	Hypothetical protein yclC [<i>Bacillus subtilis</i>]	UniRef100_P94405	<i>Bacillus subtilis</i>	YclC
390	YclD protein [<i>Bacillus subtilis</i>]	UniRef100_P94406	<i>Bacillus subtilis</i>	YclD
391				
392				
393				
394	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HM09	<i>Bacillus thuringiensis</i>	
395	Hypothetical protein OB2810 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMN2	<i>Oceanobacillus iheyensis</i>	
396				
397				YxiD
398				
399				YxiB
400	Sugar ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K7B8	<i>Bacillus halodurans</i>	RbsB
401	Two-component sensor histidine kinase [<i>Bacillus halodurans</i>]	UniRef100_Q9K7B9	<i>Bacillus halodurans</i>	YesM
402				YesN
403	Multiple sugar transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9K7C2	<i>Bacillus halodurans</i>	RbsB
404	L-arabinose transport ATP-binding protein araG [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9S472	<i>Bacillus stearothermophilus</i>	RbsA
405	L-arabinose membrane permease [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9S471	<i>Bacillus stearothermophilus</i>	RbsC
406	Transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K9F3	<i>Bacillus halodurans</i>	YbfB
407	YbfA protein [<i>Bacillus subtilis</i>]	UniRef100_Q31443	<i>Bacillus subtilis</i>	YbfA
408	Beta-D-galactosidase [<i>Bacillus circulans</i>]	UniRef100_Q45093	<i>Bacillus circulans</i>	LacA
409				Phy
410	Hypothetical transporter yclF [<i>Bacillus subtilis</i>]	UniRef100_P94408	<i>Bacillus subtilis</i>	YclF
411	YclG protein [<i>Bacillus subtilis</i>]	UniRef100_P94409	<i>Bacillus subtilis</i>	YclG
412				
413				GerKA
414	Spore germination protein KC precursor [<i>Bacillus subtilis</i>]	UniRef100_P49941	<i>Bacillus subtilis</i>	GerKC
415	Spore germination protein KB [<i>Bacillus subtilis</i>]	UniRef100_P49940	<i>Bacillus subtilis</i>	GerKB
416				Mta
417	Hypothetical protein yclH [<i>Bacillus subtilis</i>]	UniRef100_P94411	<i>Bacillus subtilis</i>	YclH
418	Hypothetical protein yclI [<i>Bacillus subtilis</i>]	UniRef100_P94412	<i>Bacillus subtilis</i>	YclI
419	Hypothetical sensory transduction protein yclJ [<i>Bacillus subtilis</i>]	UniRef100_P94413	<i>Bacillus subtilis</i>	YclJ
420	Hypothetical sensor-like histidine kinase yclK [<i>Bacillus subtilis</i>]	UniRef100_P94414	<i>Bacillus subtilis</i>	YclK
421	Methyl-accepting chemotaxis protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K632	<i>Bacillus halodurans</i>	TlpA
422	Probable aspartokinase [<i>Bacillus subtilis</i>]	UniRef100_P94417	<i>Bacillus subtilis</i>	YclM

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
423	Homologue of ferric anguibactin transport system permerase protein FatD of <i>V. anguillarum</i> [<i>Bacillus subtilis</i>]	UniRef100_P94418	<i>Bacillus subtilis</i>	YclN
424	Homologue of ferric anguibactin transport system permerase protein FatC of <i>V. anguillarum</i> [<i>Bacillus subtilis</i>]	UniRef100_P94419	<i>Bacillus subtilis</i>	YclO
425	Homologue of iron dicitrate transport ATP-binding protein FecE of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94420	<i>Bacillus subtilis</i>	YclP
426	Ferric anguibactin-binding protein precursor FatB of <i>V. anguillarum</i> [<i>Bacillus subtilis</i>]	UniRef100_P94421	<i>Bacillus subtilis</i>	YclQ
427	Homologue of multidrug resistance protein B, EmrB, of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94422	<i>Bacillus subtilis</i>	YcnB
428	YcnC protein [<i>Bacillus subtilis</i>]	UniRef100_P94423	<i>Bacillus subtilis</i>	YcnC
429	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q636R0	<i>Bacillus cereus</i> ZK	YqiQ
430	Hypothetical oxidoreductase ycnD [<i>Bacillus subtilis</i>]	UniRef100_P94424	<i>Bacillus subtilis</i>	YcnD
431	Hypothetical protein ycnE [<i>Bacillus subtilis</i>]	UniRef100_P94425	<i>Bacillus subtilis</i>	
432	YczG protein [<i>Bacillus subtilis</i>]	UniRef100_O31480	<i>Bacillus subtilis</i>	
433	Homologue of regulatory protein MocR of <i>R. meliloti</i> [<i>Bacillus subtilis</i>]	UniRef100_P94426	<i>Bacillus subtilis</i>	GabR
434	Probable 4-aminobutyrate aminotransferase (EC 2.6.1.19) ((S)-3-amino-2-methylpropionate transaminase) [<i>Bacillus subtilis</i>]	UniRef100_P94427	<i>Bacillus subtilis</i>	GabT
435	Cationic amino acid transporter [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESX7	<i>Oceanobacillus iheyensis</i>	YhdG
436	Homologue of succinate semialdehyde dehydrogenase GabD of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94428	<i>Bacillus subtilis</i>	GabD
437				YwfM
438				
439	YcnI protein [<i>Bacillus subtilis</i>]	UniRef100_P94431	<i>Bacillus subtilis</i>	YcnI
440	Homologue of copper export protein PcoD of <i>E. coli</i> [<i>Bacillus subtilis</i>]	UniRef100_P94432	<i>Bacillus subtilis</i>	YcnJ
441	YcnK protein [<i>Bacillus subtilis</i>]	UniRef100_P94433	<i>Bacillus subtilis</i>	YcnK
442	Assimilatory nitrate reductase electron transfer subunit [<i>Bacillus subtilis</i>]	UniRef100_P42433	<i>Bacillus subtilis</i>	NasB
443	Assimilatory nitrate reductase catalytic subunit [<i>Bacillus subtilis</i>]	UniRef100_P42434	<i>Bacillus subtilis</i>	NasC
444	Nitrite reductase [NAD(P)H] [<i>Bacillus subtilis</i>]	UniRef100_P42435	NAD(P)H	NasD
445	Assimilatory nitrite reductase [NAD(P)H] small subunit [<i>Bacillus subtilis</i>]	UniRef100_P42436	NAD(P)H	NasE
446	Uroporphyrin-III C-methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P42437	<i>Bacillus subtilis</i>	NasF
447	Hypothetical transcriptional regulator ydhC [<i>Bacillus subtilis</i>]	UniRef100_O05494	<i>Bacillus subtilis</i>	YdhC
448	Sodium-dependent transporter [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ENE3	<i>Oceanobacillus iheyensis</i>	YfIS
449	Hypothetical protein ycsD [<i>Bacillus subtilis</i>]	UniRef100_P42961	<i>Bacillus subtilis</i>	
450	UPI00003CC424 UniRef100 entry	UniRef100_UPI00003CC424		
451	Possible transcriptional antiterminator, bgIG family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63A16	<i>Bacillus cereus</i> ZK	MtlR
452	Putative sugar-specific PTS component EIIB [<i>Lactococcus raffinolactis</i>]	UniRef100_Q7X1N9	<i>Lactococcus raffinolactis</i>	
453	SgaT protein [<i>Mannheimia succiniciproducens</i> MBEL55E]	UniRef100_Q65WA2	<i>Mannheimia succiniciproducens</i> MBEL55E	
454	Transketolase, N-terminal subunit [<i>Streptococcus agalactiae</i>]	UniRef100_Q8E202	<i>Streptococcus agalactiae</i>	Tkt
455	Putative transketolase [<i>Salmonella typhimurium</i>]	UniRef100_Q8ZND5	<i>Salmonella typhimurium</i>	Dxs

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
456	Hypothetical protein ycsE [<i>Bacillus subtilis</i>]	UniRef100_P42962	<i>Bacillus subtilis</i>	YcsE
457	Hypothetical UPF0271 protein ycsF [<i>Bacillus subtilis</i>]	UniRef100_P42963	<i>Bacillus subtilis</i>	YcsF
458	Hypothetical protein ycsG [<i>Bacillus subtilis</i>]	UniRef100_P42964	<i>Bacillus subtilis</i>	YcsG
459	Hypothetical UPF0317 protein ycsI [<i>Bacillus subtilis</i>]	UniRef100_P42966	<i>Bacillus subtilis</i>	YcsI
460	Kinase A inhibitor [<i>Bacillus subtilis</i>]	UniRef100_P60495	<i>Bacillus subtilis</i>	KipI
461	KipI antagonist [<i>Bacillus subtilis</i>]	UniRef100_Q7WY77	<i>Bacillus subtilis</i>	KipA
462	HTH-type transcriptional regulator kipR [<i>Bacillus subtilis</i>]	UniRef100_P42968	<i>Bacillus subtilis</i>	KipR
463	Hypothetical protein ycsK [<i>Bacillus subtilis</i>]	UniRef100_P42969	<i>Bacillus subtilis</i>	YcsK
464	PTS system, mannitol-specific IIABC component [<i>Bacillus subtilis</i>]	UniRef100_P42956	<i>Bacillus subtilis</i>	MtlA
465	Mannitol-1-phosphate 5-dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P42957	<i>Bacillus subtilis</i>	MtlD
466	YdaA protein [<i>Bacillus subtilis</i>]	UniRef100_P96574	<i>Bacillus subtilis</i>	MtlR
467	General stress protein 39 [<i>Bacillus subtilis</i>]	UniRef100_P80873	<i>Bacillus subtilis</i>	YdaD
468	Hypothetical protein ydaE [<i>Bacillus subtilis</i>]	UniRef100_P96578	<i>Bacillus subtilis</i>	YdaE
469	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81U55	<i>Bacillus anthracis</i>	
470	General stress protein 26 [<i>Bacillus subtilis</i>]	UniRef100_P80238	<i>Bacillus subtilis</i>	YdaG
471	YdaH protein [<i>Bacillus subtilis</i>]	UniRef100_P96581	<i>Bacillus subtilis</i>	YdaH
472	Lin0463 protein [<i>Listeria innocua</i>]	UniRef100_Q92EJ6	<i>Listeria innocua</i>	YvhJ
473	YdzA protein [<i>Bacillus subtilis</i>]	UniRef100_O31485	<i>Bacillus subtilis</i>	
474	BH0424 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFFQ4	<i>Bacillus halodurans</i>	
475	HTH-type transcriptional regulator lrpC [<i>Bacillus subtilis</i>]	UniRef100_P96582	<i>Bacillus subtilis</i>	LrpC
476	PROBABLE DNA TOPOISOMERASE III [<i>Bacillus subtilis</i>]	UniRef100_P96583	<i>Bacillus subtilis</i>	TopB
477				
478				
479				
480				YdaO
481				YdaO
482	YdaP protein [<i>Bacillus subtilis</i>]	UniRef100_P96591	<i>Bacillus subtilis</i>	YdaP
483				
484	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
485	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
486				
487	Similar to ribosomal-protein-serine N-acetyltransferase [<i>Staphylococcus aureus</i>]	UniRef100_Q99WN5	<i>Staphylococcus aureus</i>	YdaF
488	Manganese transport protein mntH [<i>Bacillus subtilis</i>]	UniRef100_P96593	<i>Bacillus subtilis</i>	MntH
489				
490				AnsB
491				YojK
492	YdaT protein [<i>Bacillus subtilis</i>]	UniRef100_P96595	<i>Bacillus subtilis</i>	YdaT
493	Hypothetical protein ydbA [<i>Bacillus subtilis</i>]	UniRef100_P96596	<i>Bacillus subtilis</i>	YdbA
494	Na ⁺ /H ⁺ antiporter NhaC [<i>Bacillus cereus</i>]	UniRef100_Q81FX8	<i>Bacillus cereus</i>	NhaC
495	YdbB protein [<i>Bacillus subtilis</i>]	UniRef100_P96597	<i>Bacillus subtilis</i>	YdbB
496	Glucose starvation-inducible protein B [<i>Bacillus subtilis</i>]	UniRef100_P26907	<i>Bacillus subtilis</i>	GsiB
497	Hypothetical UPF0118 protein ydbI [<i>Bacillus subtilis</i>]	UniRef100_P96604	<i>Bacillus subtilis</i>	YdbI
498				GltT
499	YdbJ protein [<i>Bacillus subtilis</i>]	UniRef100_P96605	<i>Bacillus subtilis</i>	YdbJ
500	YdbK protein [<i>Bacillus subtilis</i>]	UniRef100_P96606	<i>Bacillus subtilis</i>	YdbK
501	Hypothetical protein ydbL [<i>Bacillus subtilis</i>]	UniRef100_P96607	<i>Bacillus subtilis</i>	YdbL

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
502 503 504	YdbM protein [<i>Bacillus subtilis</i>]	UniRef100_P96608	<i>Bacillus subtilis</i>	YdbM
505 506	YdbP protein [<i>Bacillus subtilis</i>] D-alanine--D-alanine ligase [<i>Bacillus subtilis</i>]	UniRef100_P96611 UniRef100_P96612	<i>Bacillus subtilis</i> <i>Bacillus subtilis</i>	YdbP Ddl
507 508	Esterase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESM0	<i>Oceanobacillus iheyensis</i>	MurF YvaK
509 510 511 512 513 514 515	YdbR protein [<i>Bacillus subtilis</i>] YdbS protein [<i>Bacillus subtilis</i>] YdbT protein [<i>Bacillus subtilis</i>] YdcA protein [<i>Bacillus subtilis</i>]	UniRef100_P96614 UniRef100_P96615 UniRef100_P96616 UniRef100_P96617	<i>Bacillus subtilis</i> <i>Bacillus subtilis</i> <i>Bacillus subtilis</i> <i>Bacillus subtilis</i>	YdbR YdbS YdbT YdcA YdcC Alr
516 517 518 519 520 521 522 523 524	YdcE protein [<i>Bacillus subtilis</i>]	UniRef100_P96622	<i>Bacillus subtilis</i>	YdcE RsbR RsbS RsbT RsbU RsbV RsbW SigB RsbX
525 526	YdcI protein [<i>Bacillus subtilis</i>] Transcriptional regulator, TetR family [<i>Bacillus thuringiensis</i>]	UniRef100_Q31489 UniRef100_Q6HGY5	<i>Bacillus subtilis</i> <i>Bacillus thuringiensis</i>	YdcI YxbF
527 528	Lin1189 protein [<i>Listeria innocua</i>]	UniRef100_Q92CI2	<i>Listeria innocua</i>	YdgH
529 530	Protein sprT-like [<i>Bacillus subtilis</i>] Possible transporter, EamA family [<i>Bacillus cereus</i> ZK]	UniRef100_P96628 UniRef100_Q638K5	<i>Bacillus subtilis</i> <i>Bacillus cereus</i> ZK	YdcK
531 532	Hypothetical protein ORF00034 [<i>Lactococcus lactis</i>]	UniRef100_O87235	<i>Lactococcus lactis</i>	
533	Delta5 acyl-lipid desaturase [<i>Bacillus cereus</i>]	UniRef100_Q81C02	<i>Bacillus cereus</i>	Des
534 535	Cold shock protein cspC [<i>Bacillus subtilis</i>]	UniRef100_P39158	<i>Bacillus subtilis</i>	YogA YyaS
536 537	Membrane protein, putative [<i>Bacillus cereus</i>] Transcriptional regulator, MarR family [<i>Bacillus cereus</i>]	UniRef100_Q734Y0 UniRef100_Q734X9	<i>Bacillus cereus</i> <i>Bacillus cereus</i>	YybA
538 539	Acetyltransferase, GNAT family [<i>Bacillus cereus</i>] Protease synthase and sporulation negative regulatory protein PAI 2 [<i>Bacillus cereus</i>]	UniRef100_Q734X8 UniRef100_Q734X7	<i>Bacillus cereus</i> <i>Bacillus cereus</i>	PaiA PaiB
540 541	BH0654 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KF33	<i>Bacillus halodurans</i>	RocF
542 543	YdeO protein [<i>Bacillus subtilis</i>] Transporter, LysE family [<i>Bacillus cereus</i>]	UniRef100_P96672 UniRef100_Q81DI7	<i>Bacillus subtilis</i> <i>Bacillus cereus</i>	YdeO YrhP
544 545				YwqM
546 547	Permease, putative [<i>Bacillus anthracis</i>] Putative cyclase [<i>Rhodopseudomonas palustris</i>]	UniRef100_Q81QY7 UniRef100_Q6N497	<i>Bacillus anthracis</i> <i>Rhodopseudomonas palustris</i>	YvqJ
548 549 550	Hypothetical transport protein ydgF [<i>Bacillus subtilis</i>]	UniRef100_P96704	<i>Bacillus subtilis</i>	YdgF
551	RNA polymerase sigma factor sigV [<i>Bacillus subtilis</i>]	UniRef100_O05404	<i>Bacillus subtilis</i>	YknW SigV

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
552	Putative anti-SigV factor [<i>Bacillus subtilis</i>]	UniRef100_O05403	<i>Bacillus subtilis</i>	YrhM
553	Hypothetical protein yrhL [<i>Bacillus subtilis</i>]	UniRef100_O05402	<i>Bacillus subtilis</i>	YrhL
554				YdgK
555				YwpD
556				LytT
557	Collagen adhesion protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q630P2	<i>Bacillus cereus</i> ZK	
558	Lin0929 protein [<i>Listeria innocua</i>]	UniRef100_Q92D88	<i>Listeria innocua</i>	
559				
560	Metabolite transport protein [<i>Bacillus subtilis</i>]	UniRef100_O34718	<i>Bacillus subtilis</i>	YdjK
561	Thiamine-monophosphate kinase [<i>Bacillus subtilis</i>]	UniRef100_O05514	<i>Bacillus subtilis</i>	ThiL
562	Hypothetical UPF0079 protein ydiB [<i>Bacillus subtilis</i>]	UniRef100_O05515	<i>Bacillus subtilis</i>	YdiB
563	YdiC protein [<i>Bacillus subtilis</i>]	UniRef100_O05516	<i>Bacillus subtilis</i>	YdiC
564	YdiD protein [<i>Bacillus subtilis</i>]	UniRef100_O05517	<i>Bacillus subtilis</i>	YdiD
565	Probable O-sialoglycoprotein endopeptidase [<i>Bacillus subtilis</i>]	UniRef100_O05518	<i>Bacillus subtilis</i>	Gep
566				YdiF
567				YdiF
568	Molybdenum cofactor biosynthesis protein C [<i>Bacillus subtilis</i>]	UniRef100_O05520	<i>Bacillus subtilis</i>	YdiG
569	Redox-sensing transcriptional repressor rex [<i>Bacillus subtilis</i>]	UniRef100_O05521	<i>Bacillus subtilis</i>	YdiH
570	YdiI [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9P5	<i>Bacillus halodurans</i>	
571				TatCY
572	Hypothetical lipoprotein ydiK precursor [<i>Bacillus subtilis</i>]	UniRef100_O05524	<i>Bacillus subtilis</i>	
573				YdiL
574	10 kDa chaperonin [<i>Bacillus subtilis</i>]	UniRef100_P28599	<i>Bacillus subtilis</i>	
575	60 kDa chaperonin [<i>Bacillus subtilis</i>]	UniRef100_P28598	<i>Bacillus subtilis</i>	GroEL
576				
577				
578	Hypothetical protein yoiD [Bacteriophage SPBc2]	UniRef100_O64030	Bacteriophage SPBc2	
579				
580				
581	YoaR [<i>Bacillus subtilis</i>]	UniRef100_O34611	<i>Bacillus subtilis</i>	YoaR
582	Hypothetical protein yfmQ [<i>Bacillus subtilis</i>]	UniRef100_O06475	<i>Bacillus subtilis</i>	YfmQ
583				
584	YoqW protein [Bacteriophage SPBc2]	UniRef100_O64131	Bacteriophage SPBc2	YoqW
585	Lin2076 protein [<i>Listeria innocua</i>]	UniRef100_Q92A46	<i>Listeria innocua</i>	YerO
586	PEP synthase [<i>Bacillus subtilis</i>]	UniRef100_O34309	<i>Bacillus subtilis</i>	Pps
587	Hypothetical protein yoaF [<i>Bacillus subtilis</i>]	UniRef100_O31829	<i>Bacillus subtilis</i>	
588	Hypothetical protein [<i>Staphylococcus aureus</i>]	UniRef100_Q6GK76	<i>Staphylococcus aureus</i>	
589	Short-chain dehydrogenase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97LM1	<i>Clostridium acetobutylicum</i>	DltE
590	Type B carboxylesterase [<i>Bacillus</i> sp. BP-7]	UniRef100_Q9L378	<i>Bacillus</i> sp. BP-7	PnbA
591	Inositol transport protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESX2	<i>Oceanobacillus iheyensis</i>	IolF
592	Phage shock protein A homolog [<i>Bacillus subtilis</i>]	UniRef100_P54617	<i>Bacillus subtilis</i>	PspA
593	YdjG protein [<i>Bacillus subtilis</i>]	UniRef100_O34434	<i>Bacillus subtilis</i>	YdjG
594	YdjH protein [<i>Bacillus subtilis</i>]	UniRef100_O35004	<i>Bacillus subtilis</i>	YdjH
595	YdjI protein [<i>Bacillus subtilis</i>]	UniRef100_O34789	<i>Bacillus subtilis</i>	YdjI
596	Putative oxidoreductase [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67S08	<i>Symbiobacterium thermophilum</i>	YtmO
597	YrhO [<i>Bacillus subtilis</i>]	UniRef100_O05405	<i>Bacillus subtilis</i>	YrhO
598				YrhP

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
599	Helix-turn-helix domain protein [<i>Bacillus cereus</i>]	UniRef100_Q73C00	<i>Bacillus cereus</i>	
600	Stage V sporulation protein E [<i>Bacillus halodurans</i>]	UniRef100_Q9K7T4	<i>Bacillus halodurans</i>	SpoVE
601	Stage V sporulation protein E [<i>Bacillus halodurans</i>]	UniRef100_Q9K7T3	<i>Bacillus halodurans</i>	FtsW
602				
603	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q72Z89	<i>Bacillus cereus</i>	
604	BH1889 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KBN6	<i>Bacillus halodurans</i>	YobV
605				YjeA
606				YjeA
607				TreA
608	Putative HTH-type transcriptional regulator yvdE [<i>Bacillus subtilis</i>]	UniRef100_O06987	<i>Bacillus subtilis</i>	YvdE
609	Hypothetical protein yvdF [<i>Bacillus subtilis</i>]	UniRef100_O06988	<i>Bacillus subtilis</i>	YvdF
610	Hypothetical protein yvdG [<i>Bacillus subtilis</i>]	UniRef100_O06989	<i>Bacillus subtilis</i>	YvdG
611	Hypothetical protein yvdH [<i>Bacillus subtilis</i>]	UniRef100_O06990	<i>Bacillus subtilis</i>	YvdH
612	Hypothetical protein yvdI [<i>Bacillus subtilis</i>]	UniRef100_O06991	<i>Bacillus subtilis</i>	YvdI
613	Hypothetical protein yvdJ [<i>Bacillus subtilis</i>]	UniRef100_O06992	<i>Bacillus subtilis</i>	YvdJ
614	Hypothetical glycosyl hydrolase yvdK [<i>Bacillus subtilis</i>]	UniRef100_O06993	<i>Bacillus subtilis</i>	YvdK
615	Oligo-1,6-glucosidase [<i>Bacillus subtilis</i>]	UniRef100_O06994	<i>Bacillus subtilis</i>	MallL
616	Putative beta-phosphoglucomutase [<i>Bacillus subtilis</i>]	UniRef100_O06995	<i>Bacillus subtilis</i>	PgcM
617	Tyrosyl-tRNA synthetase 2 [<i>Bacillus subtilis</i>]	UniRef100_P25151	<i>Bacillus subtilis</i>	TyrZ
618	Putative HTH-type transcriptional regulator ywaE [<i>Bacillus subtilis</i>]	UniRef100_P25150	<i>Bacillus subtilis</i>	YwaE
619				
620	Hypothetical protein SE2399 [<i>Staphylococcus epidermidis</i>]	UniRef100_Q8CQM7	<i>Staphylococcus epidermidis</i>	
621	Hypothetical protein ydjM [<i>Bacillus subtilis</i>]	UniRef100_P40775	<i>Bacillus subtilis</i>	YdjM
622	YdjN protein [<i>Bacillus subtilis</i>]	UniRef100_O34353	<i>Bacillus subtilis</i>	YdjN
623	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81AT6	<i>Bacillus cereus</i>	YeaA
624				
625	Signal peptidase I [<i>Bacillus cereus</i>]	UniRef100_Q73C25	<i>Bacillus cereus</i>	SipS
626				
627	Hypothetical protein yhfK [<i>Bacillus subtilis</i>]	UniRef100_O07609	<i>Bacillus subtilis</i>	YhfK
628	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9K766	<i>Bacillus halodurans</i>	YdeE
629	Spore coat protein A [<i>Bacillus subtilis</i>]	UniRef100_P07788	<i>Bacillus subtilis</i>	CotA
630				YkrP
631	Extracellular protein [<i>Lactobacillus plantarum</i>]	UniRef100_Q88T27	<i>Lactobacillus plantarum</i>	YcdA
632	Hypothetical UPF0018 protein yeaB [<i>Bacillus subtilis</i>]	UniRef100_P46348	<i>Bacillus subtilis</i>	YeaB
633	YeaC [<i>Bacillus subtilis</i>]	UniRef100_P94474	<i>Bacillus subtilis</i>	YeaC
634	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q739D8	<i>Bacillus cereus</i>	YeaD
635	YebA [<i>Bacillus subtilis</i>]	UniRef100_P94476	<i>Bacillus subtilis</i>	YebA
636	GMP synthase [glutamine-hydrolyzing] [<i>Bacillus subtilis</i>]	UniRef100_P29727	<i>Bacillus subtilis</i>	GuaA
637	Hypoxanthine/guanine permease [<i>Bacillus subtilis</i>]	UniRef100_O34987	<i>Bacillus subtilis</i>	PbuG
638				YebC
639	Hypothetical UPF0316 protein yebE [<i>Bacillus subtilis</i>]	UniRef100_O34624	<i>Bacillus subtilis</i>	YebE
640	YebG protein [<i>Bacillus subtilis</i>]	UniRef100_O34700	<i>Bacillus subtilis</i>	
641	Phosphoribosylaminoimidazole carboxylase catalytic subunit [<i>Bacillus subtilis</i>]	UniRef100_P12044	<i>Bacillus subtilis</i>	PurE

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
642				PurK
643	Adenylosuccinate lyase [<i>Bacillus subtilis</i>]	UniRef100_P12047	<i>Bacillus subtilis</i>	PurB
644	Phosphoribosylaminoimidazole-succinocarboxamide synthase [<i>Bacillus subtilis</i>]	UniRef100_P12046	<i>Bacillus subtilis</i>	PurC
645	Hypothetical UPF0062 protein yexA [<i>Bacillus subtilis</i>]	UniRef100_P12049	<i>Bacillus subtilis</i>	
646	Phosphoribosylformylglycinamide synthase I [<i>Bacillus subtilis</i>]	UniRef100_P12041	<i>Bacillus subtilis</i>	PurQ
647	Phosphoribosylformylglycinamide synthase II [<i>Bacillus subtilis</i>]	UniRef100_P12042	<i>Bacillus subtilis</i>	PurL
648	Amidophosphoribosyltransferase precursor [<i>Bacillus subtilis</i>]	UniRef100_P00497	<i>Bacillus subtilis</i>	PurF
649				PurM
650	Phosphoribosylglycinamide formyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P12040	<i>Bacillus subtilis</i>	PurN
651	Bifunctional purine biosynthesis protein purH [Includes: Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3) (AICAR transformylase); IMP cyclohydrolase (EC 3.5.4.10) (Inosinicase) (IMP synthetase) (ATIC)] [<i>Bacillus subtilis</i>]	UniRef100_P12048	Includes: Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3) (AICAR transformylase); IMP cyclohydrolase (EC 3.5.4.10) (Inosinicase) (IMP synthetase) (ATIC)	PurH
652	Phosphoribosylamine--glycine ligase [<i>Bacillus subtilis</i>]	UniRef100_P12039	<i>Bacillus subtilis</i>	PurD
653				YxbF
654	Putative cytochrome P450 yjiB [<i>Bacillus subtilis</i>]	UniRef100_O34374	<i>Bacillus subtilis</i>	YjiB
655				
656	Hypothetical lipoprotein yybP precursor [<i>Bacillus subtilis</i>]	UniRef100_P37488	<i>Bacillus subtilis</i>	YybP
657	Transposase [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RCM3	<i>Thermoanaerobacter tengcongensis</i>	
658	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81ZG4	<i>Bacillus anthracis</i>	
659	Putative adenine deaminase yerA [<i>Bacillus subtilis</i>]	UniRef100_O34909	<i>Bacillus subtilis</i>	YerA
660	YerB protein [<i>Bacillus subtilis</i>]	UniRef100_O34968	<i>Bacillus subtilis</i>	YerB
661	YecD [<i>Bacillus subtilis</i>]	UniRef100_Q7BVT7	<i>Bacillus subtilis</i>	YerC
662	PcrB protein homolog [<i>Bacillus subtilis</i>]	UniRef100_O34790	<i>Bacillus subtilis</i>	PcrB
663	ATP-dependent DNA helicase pcrA [<i>Bacillus subtilis</i>]	UniRef100_O34580	<i>Bacillus subtilis</i>	PcrA
664	DNA ligase [<i>Bacillus subtilis</i>]	UniRef100_O31498	<i>Bacillus subtilis</i>	LigA
665	YerH protein [<i>Bacillus subtilis</i>]	UniRef100_O34629	<i>Bacillus subtilis</i>	YerH
666	BH0586 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KF99	<i>Bacillus halodurans</i>	
667				
668				
669	BH0589 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KF96	<i>Bacillus halodurans</i>	
670	Phosphotriesterase homology protein [<i>Escherichia coli</i>]	UniRef100_P45548	<i>Escherichia coli</i>	
671	Similar to unknown protein YhfS of <i>Escherichia coli</i> [<i>Phototribadus luminescens</i>]	UniRef100_Q7N5F2	<i>Phototribadus luminescens</i>	Csd
672	Phosphopentomutase [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67S06	<i>Symbiobacterium thermophilum</i>	Drm
673	Putative alanine racemase [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67S05	<i>Symbiobacterium thermophilum</i>	
674	Glucosamine-6-phosphate deaminase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESL6	<i>Oceanobacillus iheyensis</i>	NagB
675	SapB protein [<i>Bacillus subtilis</i>]	UniRef100_Q45514	<i>Bacillus subtilis</i>	SapB
676				OpuE
677	Glutamyl-tRNA(Gln) amidotransferase subunit C [<i>Bacillus subtilis</i>]	UniRef100_O06492	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
678	Glutamyl-tRNA(Gln) amidotransferase subunit A [<i>Bacillus subtilis</i>]	UniRef100_O06491	<i>Bacillus subtilis</i>	GatA
679				GatB
680	Hypothetical protein [<i>Bacillus megaterium</i>]	UniRef100_Q848Y2	<i>Bacillus megaterium</i>	
681				YdhT
682				
683				
684	Putative HTH-type transcriptional regulator yerO [<i>Bacillus subtilis</i>]	UniRef100_O31500	<i>Bacillus subtilis</i>	YerO
685	Swarming motility protein swrC [<i>Bacillus subtilis</i>]	UniRef100_O31501	<i>Bacillus subtilis</i>	YerP
686				YjcK
687	Inosine-uridine preferring nucleoside hydrolase [<i>Bacillus cereus</i>]	UniRef100_Q81DM6	<i>Bacillus cereus</i>	
688	YerQ protein [<i>Bacillus subtilis</i>]	UniRef100_O31502	<i>Bacillus subtilis</i>	YerQ
689	Hypothetical RNA methyltransferase yefA [<i>Bacillus subtilis</i>]	UniRef100_O31503	<i>Bacillus subtilis</i>	YefA
690	Type I restriction-modification system specificity subunit [<i>Bacillus cereus</i>]	UniRef100_Q817S1	<i>Bacillus cereus</i>	
691	Type I restriction-modification system methylation subunit [<i>Bacillus cereus</i>]	UniRef100_Q817S2	<i>Bacillus cereus</i>	
692	Type IC specificity subunit [<i>Streptococcus thermophilus</i>]	UniRef100_Q9RNW0	<i>Streptococcus thermophilus</i>	
693	Type I restriction-modification system restriction subunit [<i>Bacillus cereus</i>]	UniRef100_Q817S4	<i>Bacillus cereus</i>	
694				
695	Beta-glucosides PTS, EIIBC.A [<i>Lactobacillus plantarum</i>]	UniRef100_Q88T54	<i>Lactobacillus plantarum</i>	BglP
696	6-phospho-beta-glucosidase [<i>Lactobacillus plantarum</i>]	UniRef100_Q88T55	<i>Lactobacillus plantarum</i>	BglH
697				LicT
698	Response regulator aspartate phosphatase K [<i>Bacillus subtilis</i>]	UniRef100_O34930	<i>Bacillus subtilis</i>	RapK
699				
700	Methyltransferase [<i>Bacillus cereus</i> ZK]	UniRef100_Q639N2	<i>Bacillus cereus</i> ZK	
701	Hypothetical UPF0082 protein BCE0595 [<i>Bacillus cereus</i>]	UniRef100_P62032	<i>Bacillus cereus</i>	YeeI
702	Putative HTH-type transcriptional regulator AF1627 [<i>Archaeoglobus fulgidus</i>]	UniRef100_O28646	<i>Archaeoglobus fulgidus</i>	
703				
704	YfmT [<i>Bacillus subtilis</i>]	UniRef100_O06478	<i>Bacillus subtilis</i>	YfmT
705	YfmS [<i>Bacillus subtilis</i>]	UniRef100_O06477	<i>Bacillus subtilis</i>	YfmS
706				YfS
707	YfmR [<i>Bacillus subtilis</i>]	UniRef100_O06476	<i>Bacillus subtilis</i>	YfmR
708				
709	YciA protein [<i>Bacillus subtilis</i>]	UniRef100_P94398	<i>Bacillus subtilis</i>	YciA
710	UPI00002BDF65 UniRef100 entry	UniRef100_UPI00002BDF65		YpdA
711	Ferrus ion transporter protein [<i>Klebsiella pneumoniae</i>]	UniRef100_Q6U5S9	<i>Klebsiella pneumoniae</i>	
712				
713	YciC protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70KK8	<i>Bacillus amyloliquefaciens</i>	YciC
714				BioW
715	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P53555	<i>Bacillus subtilis</i>	BioA
716	8-amino-7-oxononanoate synthase [<i>Bacillus subtilis</i>]	UniRef100_P53556	<i>Bacillus subtilis</i>	BioF
717	BioD protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70JZ0	<i>Bacillus amyloliquefaciens</i>	BioD
718	BioB protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70JZ1	<i>Bacillus amyloliquefaciens</i>	BioB
719	BioI protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70JZ2	<i>Bacillus amyloliquefaciens</i>	BioI
720	BH1501 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KCR8	<i>Bacillus halodurans</i>	
721				YfmP

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
722	Multidrug efflux protein yfmO [<i>Bacillus subtilis</i>]	UniRef100_O06473	<i>Bacillus subtilis</i>	YfmO
723	YfmM protein [<i>Bacillus subtilis</i>]	UniRef100_O34512	<i>Bacillus subtilis</i>	YfmM
724	YfmL protein [<i>Bacillus subtilis</i>]	UniRef100_O34750	<i>Bacillus subtilis</i>	YfmL
725	YfmJ protein [<i>Bacillus subtilis</i>]	UniRef100_O34812	<i>Bacillus subtilis</i>	YfmJ
726	Hypothetical protein yfmB [<i>Bacillus subtilis</i>]	UniRef100_O34626	<i>Bacillus subtilis</i>	YfmB
727	General stress protein 17M [<i>Bacillus subtilis</i>]	UniRef100_P80241	<i>Bacillus subtilis</i>	YfiT
728				YfiS
729				YfiS
730	Putative permease [<i>Clostridium tetani</i>]	UniRef100_Q895A0	<i>Clostridium tetani</i>	
731	Possible Zn-dependent hydrolase, beta-lactamase superfamily [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HKH5	<i>Bacillus thuringiensis</i>	YqgX
732	YfiN protein [<i>Bacillus subtilis</i>]	UniRef100_O34409	<i>Bacillus subtilis</i>	YfiN
733	Zink-carboxypeptidase [<i>Clostridium tetani</i>]	UniRef100_Q898E1	<i>Clostridium tetani</i>	
734	Nitric oxide synthase oxygenase [<i>Bacillus subtilis</i>]	UniRef100_O34453	<i>Bacillus subtilis</i>	YfiM
735	Membrane protein, putative [<i>Bacillus cereus</i>]	UniRef100_Q72YN4	<i>Bacillus cereus</i>	YvaZ
736	Transcriptional regulator, ArsR family [<i>Bacillus cereus</i> ZK]	UniRef100_Q632K6	<i>Bacillus cereus</i> ZK	
737	Putative acylphosphatase [<i>Bacillus subtilis</i>]	UniRef100_O35031	<i>Bacillus subtilis</i>	
738	YfiK protein [<i>Bacillus subtilis</i>]	UniRef100_O34542	<i>Bacillus subtilis</i>	YfiK
739				
740				
741	YfiG protein [<i>Bacillus subtilis</i>]	UniRef100_O34484	<i>Bacillus subtilis</i>	YfiG
742				YfiE
743	Hypothetical protein yfiB [<i>Bacillus subtilis</i>]	UniRef100_O34887	<i>Bacillus subtilis</i>	
744				YfiA
745	Probable PTS system, trehalose-specific IIBC component [<i>Bacillus subtilis</i>]	UniRef100_P39794	<i>Bacillus subtilis</i>	TreP
746	Alpha-glucosidase [<i>Bacillus</i> sp. DG0303]	UniRef100_Q9L872	<i>Bacillus</i> sp. DG0303	TreA
747	Trehalose operon transcriptional repressor [<i>Bacillus subtilis</i>]	UniRef100_P39796	<i>Bacillus subtilis</i>	TreR
748	Acetyltransferases [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RBZ9	<i>Thermoanaerobacter tengcongensis</i>	Yvfd
749				
750				
751	Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RBY7	<i>Thermoanaerobacter tengcongensis</i>	SpsC
752				YkuQ
753	UPI000029FB28 UniRef100 entry	UniRef100_UPI000029FB28		YtcB
754	Hypothetical protein MA2181 [<i>Methanosarcina acetivorans</i>]	UniRef100_Q8TNU7	<i>Methanosarcina acetivorans</i>	
755				
756	Beta 1,4 glucosyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q81GJ1	<i>Bacillus cereus</i>	YolJ
757	Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RC02	<i>Thermoanaerobacter tengcongensis</i>	SpsC
758	Predicted dehydrogenases and related proteins [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RC00	<i>Thermoanaerobacter tengcongensis</i>	YrbE
759	UDP-glucose:GDP-mannose dehydrogenase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CXB6	<i>Oceanobacillus iheyensis</i>	TuaD
760	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q635H5	<i>Bacillus cereus</i> ZK	YqkD

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
761	Hypothetical UPF0087 protein ydeP [<i>Bacillus subtilis</i>]	UniRef100_P96673	<i>Bacillus subtilis</i>	YdeP
762	Putative NAD(P)H nitroreductase yfkO [<i>Bacillus subtilis</i>]	UniRef100_O34475	<i>Bacillus subtilis</i>	YfkO
763				YfkN
764	General stress protein 18 [<i>Bacillus subtilis</i>]	UniRef100_P80876	<i>Bacillus subtilis</i>	YfkM
765	YfkK protein [<i>Bacillus subtilis</i>]	UniRef100_O35019	<i>Bacillus subtilis</i>	
766	Amino acid transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K5Q5	<i>Bacillus halodurans</i>	YfiA
767	YfkJ protein [<i>Bacillus subtilis</i>]	UniRef100_O35016	<i>Bacillus subtilis</i>	YfkJ
768	Hypothetical protein yfkI precursor [<i>Bacillus subtilis</i>]	UniRef100_O34418	<i>Bacillus subtilis</i>	YfkI
769	YfkH protein [<i>Bacillus subtilis</i>]	UniRef100_O34437	<i>Bacillus subtilis</i>	YfkH
770	YfkF protein [<i>Bacillus subtilis</i>]	UniRef100_O34929	<i>Bacillus subtilis</i>	YfkF
771	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELS1	<i>Oceanobacillus iheyensis</i>	
772	YfkE protein [<i>Bacillus subtilis</i>]	UniRef100_O34840	<i>Bacillus subtilis</i>	YfkE
773	YfkD protein [<i>Bacillus subtilis</i>]	UniRef100_O34579	<i>Bacillus subtilis</i>	YfkD
774	Thioredoxin-like oxidoreductases [<i>Bacillus cereus</i>]	UniRef100_Q81IC7	<i>Bacillus cereus</i>	YfkA
775	YljT protein [<i>Bacillus subtilis</i>]	UniRef100_O35041	<i>Bacillus subtilis</i>	
776				YfjS
777	UPI000029390C UniRef100 entry	UniRef100_UPI000029390C		AraM
778				
779	YfjQ protein [<i>Bacillus subtilis</i>]	UniRef100_O31543	<i>Bacillus subtilis</i>	YfjQ
780	YfjP protein [<i>Bacillus subtilis</i>]	UniRef100_O31544	<i>Bacillus subtilis</i>	YfjP
781				YfjO
782	YfjM protein [<i>Bacillus subtilis</i>]	UniRef100_O31547	<i>Bacillus subtilis</i>	YfjM
783				
784	Hypothetical protein yfjL [<i>Bacillus subtilis</i>]	UniRef100_P40773	<i>Bacillus subtilis</i>	YfjL
785	UPI00003CB259 UniRef100 entry	UniRef100_UPI00003CB259		YvkB
786	YdhE protein [<i>Bacillus subtilis</i>]	UniRef100_O05496	<i>Bacillus subtilis</i>	YdhE
787	Hypothetical protein yckD precursor [<i>Bacillus subtilis</i>]	UniRef100_P42402	<i>Bacillus subtilis</i>	
788	Hypothetical metabolite transport protein yceI [<i>Bacillus subtilis</i>]	UniRef100_O34691	<i>Bacillus subtilis</i>	YceI
789				
790	Levansucrase and sucrose synthesis operon antiterminator [<i>Bacillus subtilis</i>]	UniRef100_P15401	<i>Bacillus subtilis</i>	SacX SacY
791	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81NL1	<i>Bacillus anthracis</i>	YbcF
792	Hypothetical protein ybcD [<i>Bacillus cereus</i> ZK]	UniRef100_Q639F7	<i>Bacillus cereus</i> ZK	YbcD
793	Potential NADH-quinone oxidoreductase subunit 5 [<i>Bacillus subtilis</i>]	UniRef100_P39755	<i>Bacillus subtilis</i>	NdhF
794	YbcI protein [<i>Bacillus subtilis</i>]	UniRef100_O34380	<i>Bacillus subtilis</i>	YbcI
795				YraA
796	TPP-dependent acetoin dehydrogenase E1 alpha-subunit [<i>Bacillus anthracis</i>]	UniRef100_Q81PM6	<i>Bacillus anthracis</i>	AcoA
797	TPP-dependent acetoin dehydrogenase E1 beta-subunit [<i>Bacillus cereus</i>]	UniRef100_Q736U7	<i>Bacillus cereus</i>	AcoB
798	Dihydrolypoyllysine-residue acetyltransferase component of acetoin cleaving system [<i>Bacillus subtilis</i>]	UniRef100_O31550	<i>Bacillus subtilis</i>	AcoC
799	Dihydrolypoyl dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O34324	<i>Bacillus subtilis</i>	AcoL
800	Acetoin operon transcriptional activator, putative [<i>Bacillus cereus</i>]	UniRef100_Q736V6	<i>Bacillus cereus</i>	AcoR
801	Hypothetical UPF0060 protein yfjF [<i>Bacillus subtilis</i>]	UniRef100_O31553	<i>Bacillus subtilis</i>	YfjF
802	Maltose-6'-phosphate glucosidase [<i>Bacillus subtilis</i>]	UniRef100_P54716	<i>Bacillus subtilis</i>	MalA
803	HTH-type transcriptional regulator glvR [<i>Bacillus subtilis</i>]	UniRef100_P54717	<i>Bacillus subtilis</i>	YfiA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
804	PTS system, arbutin-like IIBC component [<i>Bacillus subtilis</i>]	UniRef100_P54715	<i>Bacillus subtilis</i>	MalP
805	UPI00003651CF UniRef100 entry	UniRef100_UPI00003651CF		
806	UPI000034AA1D UniRef100 entry	UniRef100_UPI000034AA1D		
807	Transcriptional regulator, MarR family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63812	<i>Bacillus cereus</i> ZK	YvaP
808				YfiD
809	Hypothetical protein yfiE [<i>Bacillus subtilis</i>]	UniRef100_P54721	<i>Bacillus subtilis</i>	YfiE
810	Xylosidase/arabinosidase [<i>Bacteroides thetaiotaomicron</i>]	UniRef100_Q8A036	<i>Bacteroides thetaiotaomicron</i>	XynB
811	Xylan beta-1,4-xylosidase [<i>Bacillus halodurans</i>]	UniRef100_Q9K6P5	<i>Bacillus halodurans</i>	XynB
812	Transcriptional regulator of AraC family [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97FW8	<i>Clostridium acetobutylicum</i>	YbfI
813				YtcQ
814	NAD(P)H dehydrogenase, quinone family [<i>Bacillus cereus</i> ZK]	UniRef100_Q638S8	<i>Bacillus cereus</i> ZK	
815	Mutator MutT protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K8B7	<i>Bacillus halodurans</i>	YjhB
816	YfiT protein [<i>Bacillus subtilis</i>]	UniRef100_O31562	<i>Bacillus subtilis</i>	YfiT
817	YfiX [<i>Bacillus subtilis</i>]	UniRef100_O52961	<i>Bacillus subtilis</i>	YfiX
818	Hypothetical protein yfhB [<i>Bacillus subtilis</i>]	UniRef100_O31570	<i>Bacillus subtilis</i>	YfhB
819	YfhC protein [<i>Bacillus subtilis</i>]	UniRef100_O31571	<i>Bacillus subtilis</i>	YfhC
820	Hypothetical protein yfhD [<i>Bacillus subtilis</i>]	UniRef100_O31572	<i>Bacillus subtilis</i>	
821				
822	BH0923 homolog [<i>Bacillus cereus</i>]	UniRef100_Q81IA0	<i>Bacillus cereus</i>	
823	Hypothetical UPF0105 protein yfhF [<i>Bacillus subtilis</i>]	UniRef100_O31574	<i>Bacillus subtilis</i>	YfhF
824	Regulatory protein recX [<i>Bacillus subtilis</i>]	UniRef100_O31575	<i>Bacillus subtilis</i>	YfhG
825	YfhH protein [<i>Bacillus subtilis</i>]	UniRef100_O31576	<i>Bacillus subtilis</i>	YfhH
826				
827	YfhJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31578	<i>Bacillus subtilis</i>	
828	CsbB protein [<i>Bacillus subtilis</i>]	UniRef100_Q45539	<i>Bacillus subtilis</i>	CsbB
829	Hypothetical protein SE1997 [<i>Staphylococcus epidermidis</i>]	UniRef100_Q8CR87	<i>Staphylococcus epidermidis</i>	
830	YfhO protein [<i>Bacillus subtilis</i>]	UniRef100_O31582	<i>Bacillus subtilis</i>	YfhO
831	YfhP protein [<i>Bacillus subtilis</i>]	UniRef100_O31583	<i>Bacillus subtilis</i>	YfhP
832	YfhQ protein [<i>Bacillus subtilis</i>]	UniRef100_O31584	<i>Bacillus subtilis</i>	YfhQ
833	YfhS protein [<i>Bacillus subtilis</i>]	UniRef100_O31585	<i>Bacillus subtilis</i>	
834	Unidentified dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P71079	<i>Bacillus subtilis</i>	FabL
835				
836	Hypothetical protein ygaB [<i>Bacillus subtilis</i>]	UniRef100_P71080	<i>Bacillus subtilis</i>	
837	YgaC protein [<i>Bacillus subtilis</i>]	UniRef100_Q796Z1	<i>Bacillus subtilis</i>	YgaC
838	Unidentified transporter-ATP binding [<i>Bacillus subtilis</i>]	UniRef100_P71082	<i>Bacillus subtilis</i>	YgaD
839	Oligopeptide ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K6T0	<i>Bacillus halodurans</i>	AppD
840	Oligopeptide ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K6T1	<i>Bacillus halodurans</i>	AppF
841	Dipeptide transporter protein DppA [<i>Bacillus firmus</i>]	UniRef100_P94310	<i>Bacillus firmus</i>	OppA
842	Dipeptide ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K6T3	<i>Bacillus halodurans</i>	AppB
843	Dipeptide transport system permease protein dppC [<i>Bacillus pseudofirmus</i>]	UniRef100_P94312	<i>Bacillus pseudofirmus</i>	AppC
844	Hypothetical 40.7 kd protein [<i>Bacillus subtilis</i>]	UniRef100_P71083	<i>Bacillus subtilis</i>	YgaE
845	Glutamate-1-semialdehyde 2,1-aminomutase 2 [<i>Bacillus subtilis</i>]	UniRef100_P71084	<i>Bacillus subtilis</i>	GsaB

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
846	YgaF protein [<i>Bacillus subtilis</i>]	UniRef100_Q796Y8	<i>Bacillus subtilis</i>	YgaF
847	Peroxide operon regulator [<i>Bacillus subtilis</i>]	UniRef100_P71086	<i>Bacillus subtilis</i>	PerR
848				
849	Hypothetical protein ygxA [<i>Bacillus subtilis</i>]	UniRef100_Q04385	<i>Bacillus subtilis</i>	YgxA
850				RapD
851				
852				
853				
854				
855				
856	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HH72	<i>Bacillus thuringiensis</i>	
857				YxiD
858	Hypothetical protein [uncultured archaeon GZfos18F2]	UniRef100_Q64DF2	uncultured archaeon GZfos18F2	
859				YxiD
860				RapE
861				
862	Putative membrane protein [<i>Streptomyces avermitilis</i>]	UniRef100_Q82NM0	<i>Streptomyces avermitilis</i>	
863	3-dehydroquinate dehydratase [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y9N4	<i>Listeria monocytogenes</i>	AroC
864				ThiC
865	Putative aliphatic sulfonates transport ATP-binding protein ssuB [<i>Bacillus subtilis</i>]	UniRef100_P97027	<i>Bacillus subtilis</i>	SsuB
866	Putative aliphatic sulfonates binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P40400	<i>Bacillus subtilis</i>	SsuA
867	Putative aliphatic sulfonates transport permease protein ssuC [<i>Bacillus subtilis</i>]	UniRef100_P40401	<i>Bacillus subtilis</i>	SsuC
868				SsuD
869				SsuD
870	Hypothetical lipoprotein ygaO precursor [<i>Bacillus subtilis</i>]	UniRef100_P97029	<i>Bacillus subtilis</i>	YgaO
871	DNA-binding protein [<i>Bacillus anthracis</i>]	UniRef100_Q81V18	<i>Bacillus anthracis</i>	
872				
873	ABC-type multidrug transport system, ATPase component [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RBH0	<i>Thermoanaerobacter tengcongensis</i>	YhaQ
874				
875				
876				
877				
878	30S ribosomal protein S14-2 [<i>Bacillus subtilis</i>]	UniRef100_O31587	<i>Bacillus subtilis</i>	
879	Hypothetical protein yhzB [<i>Bacillus subtilis</i>]	UniRef100_O31588	<i>Bacillus subtilis</i>	YhzB
880	Hypothetical 48.5 kd protein [<i>Bacillus subtilis</i>]	UniRef100_P97030	<i>Bacillus subtilis</i>	YhbA
881	Hypothetical 35.8 kd protein [<i>Bacillus subtilis</i>]	UniRef100_P97031	<i>Bacillus subtilis</i>	YhbB
882	CspR [<i>Bacillus subtilis</i>]	UniRef100_Q45512	<i>Bacillus subtilis</i>	CspR
883	Hypothetical 27 kd protein [<i>Bacillus subtilis</i>]	UniRef100_P97032	<i>Bacillus subtilis</i>	YhbD
884	Hypothetical Cytosolic Protein [<i>Bacillus cereus</i>]	UniRef100_Q813H1	<i>Bacillus cereus</i>	YhbE
885	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81PD2	<i>Bacillus anthracis</i>	YhbF
886				
887	PrkA protein [<i>Bacillus subtilis</i>]	UniRef100_P39134	<i>Bacillus subtilis</i>	PrkA
888	Stress response UPF0229 protein yhbH [<i>Bacillus subtilis</i>]	UniRef100_P45742	<i>Bacillus subtilis</i>	YhbH
889	YhbJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31593	<i>Bacillus subtilis</i>	YhbJ
890	Hypothetical transport protein yhcA [<i>Bacillus subtilis</i>]	UniRef100_P54585	<i>Bacillus subtilis</i>	YhcA

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
891	Hypothetical protein yhcB [<i>Bacillus subtilis</i>]	UniRef100_P54586	<i>Bacillus subtilis</i>	YhcB
892				YhcC
893				
894	Hypothetical ABC transporter ATP-binding protein yhcG [<i>Bacillus subtilis</i>]	UniRef100_P54591	<i>Bacillus subtilis</i>	
895	Hypothetical ABC transporter ATP-binding protein yhcH [<i>Bacillus subtilis</i>]	UniRef100_P54592	<i>Bacillus subtilis</i>	YhcH
896	Hypothetical protein yhcI [<i>Bacillus subtilis</i>]	UniRef100_P54593	<i>Bacillus subtilis</i>	YhcI
897	Cold shock protein cspB [<i>Bacillus subtilis</i>]	UniRef100_P32081	<i>Bacillus subtilis</i>	
898	ABC transporter, ATP-binding protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HP89	<i>Bacillus thuringiensis</i>	YusC
899	MW0417 protein [<i>Staphylococcus aureus</i>]	UniRef100_Q8NY20	<i>Staphylococcus aureus</i>	YusB
900	ABC transporter substrate-binding protein [<i>Bacillus cereus</i>]	UniRef100_Q81IN6	<i>Bacillus cereus</i>	YhcJ
901	Hypothetical symporter yhcL [<i>Bacillus subtilis</i>]	UniRef100_P54596	<i>Bacillus subtilis</i>	YhcL
902	Hypothetical protein yhcM [<i>Bacillus subtilis</i>]	UniRef100_P54597	<i>Bacillus subtilis</i>	YhcM
903	Acylamino-acid-releasing enzyme [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CXN6	<i>Oceanobacillus iheyensis</i>	YuxL
904	Lipoprotein yhcN precursor [<i>Bacillus subtilis</i>]	UniRef100_P54598	<i>Bacillus subtilis</i>	YhcN
905	Hypothetical protein yhcP [<i>Bacillus subtilis</i>]	UniRef100_P54600	<i>Bacillus subtilis</i>	YhcP
906	Hypothetical protein yhcQ [<i>Bacillus subtilis</i>]	UniRef100_P54601	<i>Bacillus subtilis</i>	YhcQ
907	Two-component sensor histidine kinase [<i>Bacillus subtilis</i>]	UniRef100_O31661	<i>Bacillus subtilis</i>	KinE
908	Hypothetical protein yhcR precursor [<i>Bacillus subtilis</i>]	UniRef100_P54602	<i>Bacillus subtilis</i>	YhcR
909	Hypothetical protein yhcS [<i>Bacillus subtilis</i>]	UniRef100_P54603	<i>Bacillus subtilis</i>	YhcS
910	Hypothetical pseudouridine synthase yhcT [<i>Bacillus subtilis</i>]	UniRef100_P54604	<i>Bacillus subtilis</i>	YhcT
911	Hypothetical protein yhcU [<i>Bacillus subtilis</i>]	UniRef100_P54605	<i>Bacillus subtilis</i>	YhcU
912	Hypothetical protein yhcV [<i>Bacillus subtilis</i>]	UniRef100_P54606	<i>Bacillus subtilis</i>	YhcV
913	Hypothetical protein yhcW [<i>Bacillus subtilis</i>]	UniRef100_P54607	<i>Bacillus subtilis</i>	YhcW
914	Hypothetical UPF0012 protein yhcX [<i>Bacillus subtilis</i>]	UniRef100_P54608	<i>Bacillus subtilis</i>	YhcX
915	ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KBA3	<i>Bacillus halodurans</i>	YdiF
916	Hypothetical transport protein yrhG [<i>Bacillus subtilis</i>]	UniRef100_O05399	<i>Bacillus subtilis</i>	YrhG
917	Lin0826 protein [<i>Listeria innocua</i>]	UniRef100_Q92DI7	<i>Listeria innocua</i>	YwkB
918	Alcohol dehydrogenase [<i>Bacillus cereus</i>]	UniRef100_Q81CT4	<i>Bacillus cereus</i>	YogA
919	Glycerol uptake operon antiterminator regulatory protein [<i>Bacillus subtilis</i>]	UniRef100_P30300	<i>Bacillus subtilis</i>	GlpP
920	Glycerol uptake facilitator protein [<i>Bacillus subtilis</i>]	UniRef100_P18156	<i>Bacillus subtilis</i>	GlpF
921	Glycerol kinase [<i>Bacillus subtilis</i>]	UniRef100_P18157	<i>Bacillus subtilis</i>	GlpK
922	Aerobic glycerol-3-phosphate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P18158	<i>Bacillus subtilis</i>	GlpD
923	Alpha-phosphoglucomutase [<i>Bacillus subtilis</i> subsp. <i>subtilis</i>]	UniRef100_Q68VA2	<i>Bacillus subtilis</i> subsp. <i>subtilis</i>	YhxB
924	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELS7	<i>Oceanobacillus iheyensis</i>	YcgB
925	Transcriptional regulator [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELS8	<i>Oceanobacillus iheyensis</i>	
926				YhcY

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
927	Hypothetical protein yhcZ [<i>Bacillus subtilis</i>]	UniRef100_O07528	<i>Bacillus subtilis</i>	YhcZ
928	Hypothetical protein yhdA [<i>Bacillus subtilis</i>]	UniRef100_O07529	<i>Bacillus subtilis</i>	YhdA
929				
930				
931	Hypothetical UPF0074 protein yhdE [<i>Bacillus subtilis</i>]	UniRef100_O07573	<i>Bacillus subtilis</i>	YhdE
932	Flavohepotein [<i>Bacillus halodurans</i>]	UniRef100_Q9RC40	<i>Bacillus halodurans</i>	Hmp
933	Stage V sporulation protein R [<i>Bacillus subtilis</i>]	UniRef100_P37875	<i>Bacillus subtilis</i>	SpoVR
934	Probable endopeptidase lytE precursor [<i>Bacillus subtilis</i>]	UniRef100_P54421	<i>Bacillus subtilis</i>	LytE
935				CitR
936				CitA
937	Glucose dehydrogenase-B [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9R3	<i>Bacillus halodurans</i>	
938	Hypothetical oxidoreductase yhdF [<i>Bacillus subtilis</i>]	UniRef100_O07575	<i>Bacillus subtilis</i>	YhdF
939	Hypothetical protein yhdH [<i>Bacillus subtilis</i>]	UniRef100_O07577	<i>Bacillus subtilis</i>	YhdH
940	2,4-diaminobutyrate decarboxylase [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB9	<i>Bacillus halodurans</i>	
941	YdeE protein [<i>Bacillus subtilis</i>]	UniRef100_P96662	<i>Bacillus subtilis</i>	YdeE
942	YdeL protein [<i>Bacillus subtilis</i>]	UniRef100_P96669	<i>Bacillus subtilis</i>	YdeL
943	BH1582 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KCI9	<i>Bacillus halodurans</i>	YhdJ
944	YhdK [<i>Bacillus subtilis</i> subsp. <i>spizizenii</i>]	UniRef100_Q7X2K9	<i>Bacillus subtilis</i> subsp. <i>spizizenii</i>	
945	YhdL [<i>Bacillus subtilis</i> subsp. <i>spizizenii</i>]	UniRef100_Q7X2L0	<i>Bacillus subtilis</i> subsp. <i>spizizenii</i>	YhdL
946	Hypothetical protein yhdM (RNA polymerase ECF (Extracytoplasmic function)-type sigma factor) [<i>Bacillus subtilis</i>]	UniRef100_O07582	<i>Bacillus subtilis</i>	SigM
947				YrkC
948	Hypothetical protein yhdO [<i>Bacillus subtilis</i>]	UniRef100_O07584	<i>Bacillus subtilis</i>	YhdO
949				
950	Acyl-CoA thioesterase 1 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97DR5	<i>Clostridium acetobutylicum</i>	
951	UPI00003CB259 UniRef100 entry	UniRef100_UPI00003CB259		YvkB
952				YhdP
953	HTH-type transcriptional regulator cueR [<i>Bacillus subtilis</i>]	UniRef100_O07586	<i>Bacillus subtilis</i>	YhdQ
954				YhdT
955				
956	BH3511 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K762	<i>Bacillus halodurans</i>	
957	Sporulation specific N-acetylmuramoyl-L-alanine amidase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CX69	<i>Oceanobacillus iheyensis</i>	CwIC
958	Protein creB homolog 1 [<i>Bacillus subtilis</i>]	UniRef100_O07590	<i>Bacillus subtilis</i>	YhdU
959	Protein creB homolog 2 [<i>Bacillus subtilis</i>]	UniRef100_O07591	<i>Bacillus subtilis</i>	YhdV
960				YhdW
961				
962	Hypothetical UPF0003 protein yhdY [<i>Bacillus subtilis</i>]	UniRef100_O07594	<i>Bacillus subtilis</i>	YhdY
963	NAD-dependent deacetylase [<i>Bacillus subtilis</i>]	UniRef100_O07595	<i>Bacillus subtilis</i>	YhdZ
964				
965	Hypothetical protein yheN [<i>Bacillus subtilis</i>]	UniRef100_O07596	<i>Bacillus subtilis</i>	YheN
966				Dat
967	Na(+)/H(+) antiporter [<i>Bacillus subtilis</i>]	UniRef100_O07553	<i>Bacillus subtilis</i>	NhaC

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
968	Hypothetical protein yoxA [<i>Bacillus subtilis</i>]	UniRef100_P39840	<i>Bacillus subtilis</i>	YoxA
969	Hypothetical protein ydhH [<i>Bacillus subtilis</i>]	UniRef100_O05500	<i>Bacillus subtilis</i>	YdhH
970	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63EB4	<i>Bacillus cereus</i> ZK	
971	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81GF4	<i>Bacillus cereus</i>	
972	Stress response protein nhaX [<i>Bacillus subtilis</i>]	UniRef100_O07552	<i>Bacillus subtilis</i>	NhaX
973	Hypothetical protein yheI [<i>Bacillus subtilis</i>]	UniRef100_O07550	<i>Bacillus subtilis</i>	YheI
974	Hypothetical protein yheH [<i>Bacillus subtilis</i>]	UniRef100_O07549	<i>Bacillus subtilis</i>	YheH
975	Hypothetical protein yheG [<i>Bacillus subtilis</i>]	UniRef100_O07548	<i>Bacillus subtilis</i>	YheG
976	Small, acid-soluble spore protein B [<i>Bacillus subtilis</i>]	UniRef100_P04832	<i>Bacillus subtilis</i>	
977	BH1139 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDS2	<i>Bacillus halodurans</i>	
978	Sugar ABC transporter ATP-binding protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CUH3	<i>Oceanobacillus iheyensis</i>	MsmX
979	Fis-type helix-turn-helix domain protein [<i>Bacillus cereus</i>]	UniRef100_Q73A84	<i>Bacillus cereus</i>	YxkF
980	Hypothetical protein yheE [<i>Bacillus subtilis</i>]	UniRef100_O07546	<i>Bacillus subtilis</i>	
981	Hypothetical protein yheD [<i>Bacillus subtilis</i>]	UniRef100_O07545	<i>Bacillus subtilis</i>	YheD
982	Hypothetical protein yheC [<i>Bacillus subtilis</i>]	UniRef100_O07544	<i>Bacillus subtilis</i>	YheC
983	Hypothetical protein yheB [<i>Bacillus subtilis</i>]	UniRef100_O07543	<i>Bacillus subtilis</i>	YheB
984	Hypothetical protein yheA [<i>Bacillus subtilis</i>]	UniRef100_O07542	<i>Bacillus subtilis</i>	YheA
985	Stress response protein yhaX [<i>Bacillus subtilis</i>]	UniRef100_O07539	<i>Bacillus subtilis</i>	YhaX
986				HemZ
987	YhaR protein [<i>Bacillus subtilis</i>]	UniRef100_O07533	<i>Bacillus subtilis</i>	YhaR
988	Response regulator aspartate phosphatase I [<i>Bacillus subtilis</i>]	UniRef100_P96649	<i>Bacillus subtilis</i>	RapI
989				
990	Hypothetical protein yhaQ [<i>Bacillus subtilis</i>]	UniRef100_O07524	<i>Bacillus subtilis</i>	YhaQ
991	Hypothetical protein yhaP [<i>Bacillus subtilis</i>]	UniRef100_O07523	<i>Bacillus subtilis</i>	YhaP
992				YhaO
993	Hypothetical protein yhaN [<i>Bacillus subtilis</i>]	UniRef100_O08455	<i>Bacillus subtilis</i>	YhaN
994				YhaM
995	Hypothetical protein yhaL [<i>Bacillus subtilis</i>]	UniRef100_O07520	<i>Bacillus subtilis</i>	
996	Foldase protein prsA precursor [<i>Bacillus subtilis</i>]	UniRef100_P24327	<i>Bacillus subtilis</i>	PrsA
997				
998	Hypothetical protein yhaK [<i>Bacillus subtilis</i>]	UniRef100_O07519	<i>Bacillus subtilis</i>	
999	Hypothetical protein yhaI [<i>Bacillus subtilis</i>]	UniRef100_O07517	<i>Bacillus subtilis</i>	YhaI
1000	Protease production regulatory protein hpr [<i>Bacillus subtilis</i>]	UniRef100_P11065	<i>Bacillus subtilis</i>	Hpr
1001	Hypothetical protein yhaH [<i>Bacillus subtilis</i>]	UniRef100_O07516	<i>Bacillus subtilis</i>	YhaH
1002	Probable tryptophan transport protein [<i>Bacillus subtilis</i>]	UniRef100_O07515	<i>Bacillus subtilis</i>	YhaG
1003	Phosphoserine aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P80862	<i>Bacillus subtilis</i>	SerC
1004	Hit protein [<i>Bacillus subtilis</i>]	UniRef100_O07513	<i>Bacillus subtilis</i>	Hit
1005				
1006	ABC-type transporter ATP-binding protein ecsA [<i>Bacillus subtilis</i>]	UniRef100_P55339	<i>Bacillus subtilis</i>	EcsA

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1007	Protein ecsB [<i>Bacillus subtilis</i>]	UniRef100_P55340	<i>Bacillus subtilis</i>	EcsB
1008	Protein ecsC [<i>Bacillus subtilis</i>]	UniRef100_P55341	<i>Bacillus subtilis</i>	EcsC
1009	YhaA protein [<i>Bacillus subtilis</i>]	UniRef100_O07598	<i>Bacillus subtilis</i>	YhaA
1010	Hypothetical protein yhfA [<i>Bacillus subtilis</i>]	UniRef100_O07599	<i>Bacillus subtilis</i>	YhfA
1011	Hypothetical protein yhgC [<i>Bacillus subtilis</i>]	UniRef100_P38049	<i>Bacillus subtilis</i>	YhgC
1012	Penicillin-binding protein 1F [<i>Bacillus subtilis</i>]	UniRef100_P38050	<i>Bacillus subtilis</i>	PbpF
1013	Uroporphyrinogen decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P32395	<i>Bacillus subtilis</i>	HemE
1014	Ferrochelatase [<i>Bacillus subtilis</i>]	UniRef100_P32396	<i>Bacillus subtilis</i>	HemH
1015	Protoporphyrinogen oxidase [<i>Bacillus subtilis</i>]	UniRef100_P32397	<i>Bacillus subtilis</i>	HemY
1016				YhgD
1017	Hypothetical protein yhgE [<i>Bacillus subtilis</i>]	UniRef100_P32399	<i>Bacillus subtilis</i>	YhgE
1018	3-oxoacyl-[acyl-carrier-protein] synthase III protein 2 [<i>Bacillus subtilis</i>]	UniRef100_O07600	acyl-carrier-protein	FabHB
1019	Hypothetical protein yhfE [<i>Bacillus subtilis</i>]	UniRef100_O07603	<i>Bacillus subtilis</i>	YhfE
1020				
1021	Hypothetical protein yhfG [<i>Bacillus subtilis</i>]	UniRef100_O07605	<i>Bacillus subtilis</i>	GltT
1022	Hypothetical protein yhfI [<i>Bacillus subtilis</i>]	UniRef100_O07607	<i>Bacillus subtilis</i>	YhfI
1023	Hypothetical protein yhfJ [<i>Bacillus subtilis</i>]	UniRef100_O07608	<i>Bacillus subtilis</i>	YhfJ
1024	Hypothetical protein yhfL [<i>Bacillus subtilis</i>]	UniRef100_O07610	<i>Bacillus subtilis</i>	YhfL
1025	Hypothetical protein yhfM precursor [<i>Bacillus subtilis</i>]	UniRef100_O07611	<i>Bacillus subtilis</i>	YhfM
1026	BH2909 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K8U3	<i>Bacillus halodurans</i>	
1027	Branched-chain amino acid transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K8U2	<i>Bacillus halodurans</i>	AziC
1028	BH2911 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K8U1	<i>Bacillus halodurans</i>	
1029	Putative metalloprotease yhfN [<i>Bacillus subtilis</i>]	UniRef100_P40769	<i>Bacillus subtilis</i>	YhfN
1030				AprE
1031	Transporter, drug/metabolite exporter family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63D40	<i>Bacillus cereus</i> ZK	YdeD
1032	Hypothetical protein yhfQ [<i>Bacillus subtilis</i>]	UniRef100_O07616	<i>Bacillus subtilis</i>	YhfQ
1033	YfmD protein [<i>Bacillus subtilis</i>]	UniRef100_O34933	<i>Bacillus subtilis</i>	YfmD
1034	YfmE protein [<i>Bacillus subtilis</i>]	UniRef100_O34832	<i>Bacillus subtilis</i>	YfmE
1035	Hypothetical protein yhfR [<i>Bacillus subtilis</i>]	UniRef100_O07617	<i>Bacillus subtilis</i>	YhfR
1036	Heme-based aerotactic transducer hemAT [<i>Bacillus subtilis</i>]	UniRef100_O07621	<i>Bacillus subtilis</i>	HemAT
1037	Rieske 2Fe–2S iron-sulfur protein, putative [<i>Bacillus cereus</i>]	UniRef100_Q73E94	<i>Bacillus cereus</i>	YhfW
1038				YhxC
1039	Hypothetical protein yhzC [<i>Bacillus subtilis</i>]	UniRef100_O31594	<i>Bacillus subtilis</i>	
1040				ComK
1041				YhjD
1042				YhjE
1043	Signal peptidase I V [<i>Bacillus subtilis</i>]	UniRef100_O07560	<i>Bacillus subtilis</i>	SipV
1044	Minor extracellular protease epr precursor [<i>Bacillus subtilis</i>]	UniRef100_P16396	<i>Bacillus subtilis</i>	Epr
1045	Putative permease [<i>Klebsiella pneumoniae</i>]	UniRef100_Q765R6	<i>Klebsiella pneumoniae</i>	YybO
1046	Hypothetical protein [<i>Enterococcus faecalis</i>]	UniRef100_Q82ZQ4	<i>Enterococcus faecalis</i>	PucR
1047	Putative allantoinase [<i>Staphylococcus xylosum</i>]	UniRef100_Q9EV52	<i>Staphylococcus xylosum</i>	PucH
1048	Peptidase, M20/M25/M40 family [<i>Enterococcus faecalis</i>]	UniRef100_Q82ZQ2	<i>Enterococcus faecalis</i>	YurH

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1049	Hypothetical protein STY0574 [<i>Salmonella typhi</i>]	UniRef100_Q8XFX7	<i>Salmonella typhi</i>	
1050	Ureidoglycolate dehydrogenase [<i>Enterococcus faecalis</i>]	UniRef100_Q838P9	<i>Enterococcus faecalis</i>	YjmC
1051	Ureidoglycolate dehydrogenase [<i>Enterococcus faecalis</i>]	UniRef100_Q838P9	<i>Enterococcus faecalis</i>	YjmC
1052	Hypothetical protein [<i>Enterococcus faecalis</i>]	UniRef100_Q838Q3	<i>Enterococcus faecalis</i>	SucD
1053	Hypothetical protein [<i>Enterococcus faecalis</i>]	UniRef100_Q838Q3	<i>Enterococcus faecalis</i>	
1054				
1055	Carbamate kinase [<i>Clostridium tetani</i>]	UniRef100_Q890W1	<i>Clostridium tetani</i>	
1056	Major facilitator family transporter [<i>Enterococcus faecalis</i>]	UniRef100_Q838Q1	<i>Enterococcus faecalis</i>	YcbE
1057	Hypothetical protein yjfA precursor [<i>Bacillus subtilis</i>]	UniRef100_O34554	<i>Bacillus subtilis</i>	
1058	Response regulator aspartate phosphatase G [<i>Bacillus subtilis</i>]	UniRef100_O32294	<i>Bacillus subtilis</i>	RapG
1059				
1060				
1061				
1062	LacI-family transcription regulator [<i>Bacillus subtilis</i>]	UniRef100_O34829	<i>Bacillus subtilis</i>	MsmR
1063	Multiple sugar-binding protein [<i>Bacillus subtilis</i>]	UniRef100_O34335	<i>Bacillus subtilis</i>	MsmE
1064	Sugar transporter [<i>Bacillus subtilis</i>]	UniRef100_O34706	<i>Bacillus subtilis</i>	AmyD
1065	Sugar transporter [<i>Bacillus subtilis</i>]	UniRef100_O34518	<i>Bacillus subtilis</i>	AmyC
1066	Alpha-galactosidase [<i>Bacillus subtilis</i>]	UniRef100_O34645	<i>Bacillus subtilis</i>	MelA
1067	Hypothetical protein yhjN [<i>Bacillus subtilis</i>]	UniRef100_O07568	<i>Bacillus subtilis</i>	YhjN
1068	Spore coat-associated protein JA [<i>Bacillus cereus</i> ZK]	UniRef100_Q63FK5	<i>Bacillus cereus</i> ZK	
1069	CotJB protein [<i>Bacillus subtilis</i>]	UniRef100_Q45537	<i>Bacillus subtilis</i>	
1070	CotJC protein [<i>Bacillus subtilis</i>]	UniRef100_Q45538	<i>Bacillus subtilis</i>	CotJC
1071	Long-chain fatty-acid-CoA ligase [<i>Bacillus halodurans</i>]	UniRef100_Q9KDT0	<i>Bacillus halodurans</i>	YngI
1072	Hypothetical protein yhjO [<i>Bacillus subtilis</i>]	UniRef100_O07569	<i>Bacillus subtilis</i>	YhjO
1073	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63FZ3	<i>Bacillus cereus</i> ZK	LytB
1074	Sensor histidine kinase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HNG3	<i>Bacillus thuringiensis</i>	PhoR
1075	Two-component response regulator [<i>Bacillus cereus</i>]	UniRef100_Q81136	<i>Bacillus cereus</i>	YelJ
1076				
1077	Putative molybdate binding protein, YvgL [<i>Bacillus subtilis</i>]	UniRef100_O32208	<i>Bacillus subtilis</i>	YhjR YvgL
1078	Putative molybdate transport protein, YvgM [<i>Bacillus subtilis</i>]	UniRef100_O32209	<i>Bacillus subtilis</i>	YvgM
1079	ATP-dependent nuclease subunit B [<i>Bacillus subtilis</i>]	UniRef100_P23477	<i>Bacillus subtilis</i>	AddB
1080	ATP-dependent nuclease subunit A [<i>Bacillus subtilis</i>]	UniRef100_P23478	<i>Bacillus subtilis</i>	AddA
1081	Exonuclease sbcD homolog [<i>Bacillus subtilis</i>]	UniRef100_P23479	<i>Bacillus subtilis</i>	SbcD
1082	Nuclease sbcCD subunit C [<i>Bacillus subtilis</i>]	UniRef100_O06714	<i>Bacillus subtilis</i>	YirY
1083	Probable spore germination protein gerPF [<i>Bacillus subtilis</i>]	UniRef100_O06716	<i>Bacillus subtilis</i>	
1084	Probable spore germination protein gerPE [<i>Bacillus subtilis</i>]	UniRef100_O06717	<i>Bacillus subtilis</i>	GerPE
1085	Probable spore germination protein gerPD [<i>Bacillus subtilis</i>]	UniRef100_O06718	<i>Bacillus subtilis</i>	
1086	Probable spore germination protein gerPC [<i>Bacillus subtilis</i>]	UniRef100_O06719	<i>Bacillus subtilis</i>	GerPC
1087	Probable spore germination protein gerPB [<i>Bacillus subtilis</i>]	UniRef100_O06720	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1088	Probable spore germination protein gerPA [<i>Bacillus subtilis</i>]	UniRef100_O06721	<i>Bacillus subtilis</i>	
1089	Hypothetical protein yitR [<i>Bacillus subtilis</i>]	UniRef100_O06753	<i>Bacillus subtilis</i>	
1090				
1091	Spore coat protein H [<i>Bacillus cereus</i>]	UniRef100_Q81EE9	<i>Bacillus cereus</i>	CotH
1092				CotG
1093	YisK [<i>Bacillus subtilis</i>]	UniRef100_O06724	<i>Bacillus subtilis</i>	YisK
1094	Hypothetical protein yloQ [<i>Bacillus cereus</i> ZK]	UniRef100_Q63CF2	<i>Bacillus cereus</i> ZK	YloQ
1095	YisL [<i>Bacillus subtilis</i>]	UniRef100_O06725	<i>Bacillus subtilis</i>	YisL
1096	Hypothetical protein yisN [<i>Bacillus subtilis</i>]	UniRef100_O06727	<i>Bacillus subtilis</i>	YisN
1097	Asparagine synthetase [glutamine-hydrolyzing] 3 [<i>Bacillus subtilis</i>]	UniRef100_O05272	glutamine-hydrolyzing	AsnO
1098				NrgA
1099	Nitrogen regulatory PII protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ERT8	<i>Oceanobacillus iheyensis</i>	
1100	YisQ [<i>Bacillus subtilis</i>]	UniRef100_O07940	<i>Bacillus subtilis</i>	YisQ
1101	Putative HTH-type transcriptional regulator yisR [<i>Bacillus subtilis</i>]	UniRef100_P40331	<i>Bacillus subtilis</i>	YisR
1102	Acetyltransferase, GNAT family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63C80	<i>Bacillus cereus</i> ZK	YokL
1103	HTH-type transcriptional regulator degA [<i>Bacillus subtilis</i>]	UniRef100_P37947	<i>Bacillus subtilis</i>	DegA
1104	Hypothetical oxidoreductase yisS [<i>Bacillus subtilis</i>]	UniRef100_P40332	<i>Bacillus subtilis</i>	YisS
1105	YisV protein [<i>Bacillus subtilis</i>]	UniRef100_Q796Q6	<i>Bacillus subtilis</i>	YisV
1106	Diaminobutyrate--pyruvate transaminase [<i>Bacillus halodurans</i>]	UniRef100_Q9K9M1	<i>Bacillus halodurans</i>	GabT
1107	L-2,4-diaminobutyrate decarboxylase [<i>Anabaena</i> sp.]	UniRef100_Q8YZR2	<i>Anabaena</i> sp.	
1108	AlI0394 protein [<i>Anabaena</i> sp.]	UniRef100_Q8YZR3	<i>Anabaena</i> sp.	
1109	BH2621 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K9M4	<i>Bacillus halodurans</i>	
1110	BH2620 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K9M5	<i>Bacillus halodurans</i>	
1111	BH2618 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K9M7	<i>Bacillus halodurans</i>	
1112	YitI protein [<i>Bacillus subtilis</i>]	UniRef100_O06744	<i>Bacillus subtilis</i>	YitI
1113	Glr2355 protein [<i>Gloeobacter violaceus</i>]	UniRef100_Q7NI29	<i>Gloeobacter violaceus</i>	YcdF
1114	BH0411 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFR6	<i>Bacillus halodurans</i>	YobV
1115	5-methyltetrahydrofolate S-homocysteine methyltransferase [<i>Bacillus halodurans</i>]	UniRef100_Q9KCE1	<i>Bacillus halodurans</i>	YitJ
1116	YitJ [<i>Bacillus subtilis</i>]	UniRef100_O06745	<i>Bacillus subtilis</i>	YitJ
1117	Hypothetical UPF0234 protein yitK [<i>Bacillus subtilis</i>]	UniRef100_O06746	<i>Bacillus subtilis</i>	YitK
1118	YitL protein [<i>Bacillus subtilis</i>]	UniRef100_O06747	<i>Bacillus subtilis</i>	YitL
1119				
1120				
1121	Hypothetical UPF0230 protein yitS [<i>Bacillus subtilis</i>]	UniRef100_P70945	<i>Bacillus subtilis</i>	YitS
1122	Hypothetical protein yitT [<i>Bacillus subtilis</i>]	UniRef100_P39803	<i>Bacillus subtilis</i>	YitT
1123	Intracellular proteinase inhibitor [<i>Bacillus subtilis</i>]	UniRef100_P39804	<i>Bacillus subtilis</i>	Ipi
1124	GMP reductase [<i>Bacillus subtilis</i>]	UniRef100_O05269	<i>Bacillus subtilis</i>	GuaC
1125				
1126				
1127	Putative orf protein [<i>Bacillus subtilis</i>]	UniRef100_P70947	<i>Bacillus subtilis</i>	YitU
1128	Putative orf protein [<i>Bacillus subtilis</i>]	UniRef100_P70948	<i>Bacillus subtilis</i>	YitV
1129				YitW
1130	N-acetyl-gamma-glutamyl-phosphate reductase [<i>Bacillus subtilis</i>]	UniRef100_P23715	<i>Bacillus subtilis</i>	ArgC
1131	Arginine biosynthesis bifunctional protein argJ [Includes: Glutamate N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase) (Ornithine	UniRef100_Q9ZJ14	Includes: Glutamate N-acetyltransferase (EC 2.3.1.35)	ArgJ

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
	transacetylase) (OATase); Amino-acid acetyltransferase (EC 2.3.1.1) (N- acetylglutamate synthase) (AGS)] [Contai		(Ornithine acetyltransferase) (Ornithine transacetylase) (OATase); Amino-acid acetyltransferase (EC 2.3.1.1) (N- acetylglutamate synthase) (AGS)	
1132	Acetylglutamate kinase [<i>Bacillus subtilis</i>]	UniRef100_P36840	<i>Bacillus subtilis</i>	ArgB
1133	Acetylmithine aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P36839	<i>Bacillus subtilis</i>	ArgD
1134	Carbamoyl-phosphate synthase, arginine-specific, small chain [<i>Bacillus subtilis</i>]	UniRef100_P36838	<i>Bacillus subtilis</i>	CarA
1135	Carbamoyl-phosphate synthase, arginine-specific, large chain [<i>Bacillus subtilis</i>]	UniRef100_P18185	<i>Bacillus subtilis</i>	CarB
1136	Ornithine carbamoyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P18186	<i>Bacillus subtilis</i>	ArgF
1137	Undecaprenyl-diphosphatase 1 [<i>Bacillus cereus</i>]	UniRef100_Q81HV4	<i>Bacillus cereus</i>	YubB
1138	YjzC protein [<i>Bacillus subtilis</i>]	UniRef100_O34585	<i>Bacillus subtilis</i>	
1139				
1140	Hypothetical protein yjaU [<i>Bacillus subtilis</i>]	UniRef100_O35001	<i>Bacillus subtilis</i>	YjaU
1141	ArgF and med genes, partial and complete cds [<i>Bacillus subtilis</i>]	UniRef100_O32435	<i>Bacillus subtilis</i>	YjaV
1142	Transcriptional activator protein med precursor [<i>Bacillus subtilis</i>]	UniRef100_O32436	<i>Bacillus subtilis</i>	Med
1143	ComG operon repressor [<i>Bacillus subtilis</i>]	UniRef100_O32437	<i>Bacillus subtilis</i>	
1144	Hypothetical protein yjzB [<i>Bacillus subtilis</i>]	UniRef100_O34891	<i>Bacillus subtilis</i>	
1145	3-oxoacyl-[acyl-carrier-protein] synthase III protein 1 [<i>Bacillus subtilis</i>]	UniRef100_O34746	acyl-carrier-protein <i>Bacillus subtilis</i>	FabHA
1146	Beta-ketoacyl-acyl carrier protein synthase II [<i>Bacillus subtilis</i>]	UniRef100_O34340	<i>Bacillus subtilis</i>	FabF
1147	YjaZ protein [<i>Bacillus subtilis</i>]	UniRef100_O31596	<i>Bacillus subtilis</i>	YjaZ
1148	Oligopeptide transport ATP-binding protein appD [<i>Bacillus subtilis</i>]	UniRef100_P42064	<i>Bacillus subtilis</i>	AppD
1149	Oligopeptide transport ATP-binding protein appF [<i>Bacillus subtilis</i>]	UniRef100_P42065	<i>Bacillus subtilis</i>	AppF
1150	Oligopeptide-binding protein appA precursor [<i>Bacillus subtilis</i>]	UniRef100_P42061	<i>Bacillus subtilis</i>	AppA
1151	Oligopeptide transport system permease protein appB [<i>Bacillus subtilis</i>]	UniRef100_P42062	<i>Bacillus subtilis</i>	AppB
1152	Oligopeptide transport system permease protein appC [<i>Bacillus subtilis</i>]	UniRef100_P42063	<i>Bacillus subtilis</i>	AppC
1153	Permease, putative [<i>Bacillus cereus</i>]	UniRef100_Q734Y3	<i>Bacillus cereus</i>	YvqJ
1154				YjbA
1155	Tryptophanyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P21656	<i>Bacillus subtilis</i>	TrpS
1156	Oligopeptide-binding protein oppA precursor [<i>Bacillus subtilis</i>]	UniRef100_P24141	<i>Bacillus subtilis</i>	OppA
1157	Oligopeptide transport system permease protein oppB [<i>Bacillus subtilis</i>]	UniRef100_P24138	<i>Bacillus subtilis</i>	OppB
1158	Oligopeptide transport system permease protein oppC [<i>Bacillus subtilis</i>]	UniRef100_P24139	<i>Bacillus subtilis</i>	OppC
1159	Oligopeptide transport ATP-binding protein oppD [<i>Bacillus subtilis</i>]	UniRef100_P24136	<i>Bacillus subtilis</i>	OppD

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1160	Oligopeptide transport ATP-binding protein oppF [<i>Bacillus subtilis</i>]	UniRef100_P24137	<i>Bacillus subtilis</i>	OppF
1161	YjbC protein [<i>Bacillus subtilis</i>]	UniRef100_O31601	<i>Bacillus subtilis</i>	YjbC
1162	Regulatory protein spx [<i>Bacillus subtilis</i>]	UniRef100_O31602	<i>Bacillus subtilis</i>	YjbD
1163	YjbE protein [<i>Bacillus subtilis</i>]	UniRef100_O31603	<i>Bacillus subtilis</i>	YjbE
1164	Adapter protein mecA 1 [<i>Bacillus subtilis</i>]	UniRef100_P37958	<i>Bacillus subtilis</i>	MecA
1165	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELH8	<i>Oceanobacillus iheyensis</i>	YfiP
1166	Hypothetical protein OB3248 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELH9	<i>Oceanobacillus iheyensis</i>	
1167	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELI0	<i>Oceanobacillus iheyensis</i>	
1168	Response regulator of citrate/malate metabolism [<i>Vibrio vulnificus</i>]	UniRef100_Q7ML23	<i>Vibrio vulnificus</i>	CitT
1169	Sensor protein citS [<i>Bacillus halodurans</i>]	UniRef100_Q9RC53	<i>Bacillus halodurans</i>	YufL
1170	YjbF protein [<i>Bacillus subtilis</i>]	UniRef100_O31604	<i>Bacillus subtilis</i>	YjbF
1171				YjbG
1172				
1173	YjbH protein [<i>Bacillus subtilis</i>]	UniRef100_O31606	<i>Bacillus subtilis</i>	YjbH
1174	YjbI protein [<i>Bacillus subtilis</i>]	UniRef100_O31607	<i>Bacillus subtilis</i>	YjbI
1175	YjbJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31608	<i>Bacillus subtilis</i>	YjbJ
1176	YjbK protein [<i>Bacillus subtilis</i>]	UniRef100_O31609	<i>Bacillus subtilis</i>	YjbK
1177	YjbL protein [<i>Bacillus subtilis</i>]	UniRef100_O31610	<i>Bacillus subtilis</i>	YjbL
1178	YjbM protein [<i>Bacillus subtilis</i>]	UniRef100_O31611	<i>Bacillus subtilis</i>	YjbM
1179				YjbN
1180	Hypothetical pseudouridine synthase yjbO [<i>Bacillus subtilis</i>]	UniRef100_O31613	<i>Bacillus subtilis</i>	YjbO
1181	YjbP protein [<i>Bacillus subtilis</i>]	UniRef100_O31614	<i>Bacillus subtilis</i>	YjbP
1182	YjbQ protein [<i>Bacillus subtilis</i>]	UniRef100_O31615	<i>Bacillus subtilis</i>	YjbQ
1183	Transcriptional activator tenA [<i>Bacillus subtilis</i>]	UniRef100_P25052	<i>Bacillus subtilis</i>	TenA
1184	Regulatory protein tenI [<i>Bacillus subtilis</i>]	UniRef100_P25053	<i>Bacillus subtilis</i>	TenI
1185	Glycine oxidase [<i>Bacillus subtilis</i>]	UniRef100_O31616	<i>Bacillus subtilis</i>	GoxB
1186	ThiS protein [<i>Bacillus subtilis</i>]	UniRef100_O31617	<i>Bacillus subtilis</i>	
1187	Thiazole biosynthesis protein thiG [<i>Bacillus subtilis</i>]	UniRef100_O31618	<i>Bacillus subtilis</i>	ThiG
1188	ThiF protein [<i>Bacillus subtilis</i>]	UniRef100_O31619	<i>Bacillus subtilis</i>	ThiF
1189	YjbV protein [<i>Bacillus subtilis</i>]	UniRef100_O31620	<i>Bacillus subtilis</i>	YjbV
1190	Enoyl-[acyl-carrier-protein] reductase [NADH] [<i>Bacillus subtilis</i>]	UniRef100_P54616	<i>Bacillus subtilis</i>	FabI
1191	YjbX protein [<i>Bacillus subtilis</i>]	UniRef100_O31622	<i>Bacillus subtilis</i>	YjbX
1192	Spore coat protein Z [<i>Bacillus subtilis</i>]	UniRef100_Q08312	<i>Bacillus subtilis</i>	CotZ
1193	Spore coat protein Y [<i>Bacillus subtilis</i>]	UniRef100_Q08311	<i>Bacillus subtilis</i>	CotY
1194	Spore coat protein X [<i>Bacillus subtilis</i>]	UniRef100_Q08313	<i>Bacillus subtilis</i>	CotX
1195	Spore coat protein W [<i>Bacillus subtilis</i>]	UniRef100_Q08310	<i>Bacillus subtilis</i>	CotW
1196	Spore coat protein V [<i>Bacillus subtilis</i>]	UniRef100_Q08309	<i>Bacillus subtilis</i>	CotV
1197	YjcA protein [<i>Bacillus subtilis</i>]	UniRef100_O31623	<i>Bacillus subtilis</i>	YjcA
1198				
1199				
1200	YjcC protein [<i>Bacillus subtilis</i>]	UniRef100_O31625	<i>Bacillus subtilis</i>	YjcD
1201				YngC
1202				GalE
1203				YngB
1204	YngB protein [<i>Bacillus subtilis</i>]	UniRef100_O31822	<i>Bacillus subtilis</i>	YngB
1205	YngA protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70JY6	<i>Bacillus amyloliquefaciens</i>	YngA
1206	YjcF protein [<i>Bacillus subtilis</i>]	UniRef100_O31628	<i>Bacillus subtilis</i>	YjcF
1207	YjcG protein [<i>Bacillus subtilis</i>]	UniRef100_O31629	<i>Bacillus subtilis</i>	YjcG
1208	YjcH protein [<i>Bacillus subtilis</i>]	UniRef100_O31630	<i>Bacillus subtilis</i>	YjcH
1209	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q739H9	<i>Bacillus cereus</i>	
1210	BH1889 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KBN6	<i>Bacillus halodurans</i>	YobV
1211	YjcI protein [<i>Bacillus subtilis</i>]	UniRef100_O31631	<i>Bacillus subtilis</i>	YjcI
1212	YjcJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31632	<i>Bacillus subtilis</i>	YjcJ
1213	YjcL protein [<i>Bacillus subtilis</i>]	UniRef100_O31634	<i>Bacillus subtilis</i>	YjcL

TABLE 1-continued

Predicted functions			
SEQ ID NO.	Description	UniRef Accession No.	Organism
1214	Transcriptional regulator, MarR/EmrR family [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97DR6	<i>Clostridium acetobutylicum</i>
1215	Penicillin-binding protein 4* [<i>Bacillus subtilis</i>]	UniRef100_P32959	<i>Bacillus subtilis</i>
1216			PbpE
1217			AbnA
1218	Maltose transacetylase [<i>Bacillus stearothermophilus</i>]	UniRef100_Q75TH6	<i>Bacillus stearothermophilus</i>
1219			Maa
1220	Putative HTH-type transcriptional regulator ywfK [<i>Bacillus subtilis</i>]	UniRef100_P39647	<i>Bacillus subtilis</i>
1221	Sulfite reductase [<i>Bacillus halodurans</i>]	UniRef100_Q9KF76	<i>Bacillus halodurans</i>
1222	Sulfite reductase [<i>Bacillus halodurans</i>]	UniRef100_Q9KF75	<i>Bacillus halodurans</i>
1223	Putative HTH-type transcriptional regulator yoaU [<i>Bacillus subtilis</i>]	UniRef100_O34701	<i>Bacillus subtilis</i>
1224	Hypothetical transport protein yoaV [<i>Bacillus subtilis</i>]	UniRef100_O34416	<i>Bacillus subtilis</i>
1225	Hypothetical protein VPA0302 [<i>Vibrio parahaemolyticus</i>]	UniRef100_Q87JF1	<i>Vibrio parahaemolyticus</i>
1226	Hypothetical protein yoeB precursor [<i>Bacillus subtilis</i>]	UniRef100_O34841	<i>Bacillus subtilis</i>
1227			YoeB
1228	Permease, general substrate transporter [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HMC3	<i>Bacillus thuringiensis</i>
1229	Putative HTH-type transcriptional regulator yxaF [<i>Bacillus subtilis</i>]	UniRef100_P42105	<i>Bacillus subtilis</i>
1230			YoeH
1231			LmrB
1232			YxaF
1233			YmzD
1234	Phage-like element PBSX protein xkdA [<i>Bacillus subtilis</i>]	UniRef100_P39780	<i>Bacillus subtilis</i>
1235	HTH-type transcriptional regulator xre [<i>Bacillus subtilis</i>]	UniRef100_P23789	<i>Bacillus subtilis</i>
1236			YeeF
1237			YjqB
1238	Phage-like element PBSX protein xkdB [<i>Bacillus subtilis</i>]	UniRef100_P39781	<i>Bacillus subtilis</i>
1239	Phage-like element PBSX protein xkdC [<i>Bacillus subtilis</i>]	UniRef100_P39782	<i>Bacillus subtilis</i>
1240	Phage-like element PBSX protein xkdD [<i>Bacillus subtilis</i>]	UniRef100_P39783	<i>Bacillus subtilis</i>
1241	Phage-like element PBSX protein xtrA [<i>Bacillus subtilis</i>]	UniRef100_P54344	<i>Bacillus subtilis</i>
1242	Positive control factor [<i>Bacillus subtilis</i>]	UniRef100_P39784	<i>Bacillus subtilis</i>
1243	PBSX phage terminase small subunit [<i>Bacillus subtilis</i>]	UniRef100_P39785	<i>Bacillus subtilis</i>
1244	PBSX phage terminase large subunit [<i>Bacillus subtilis</i>]	UniRef100_P39786	<i>Bacillus subtilis</i>
1245	Phage-like element PBSX protein xkdE [<i>Bacillus subtilis</i>]	UniRef100_P54325	<i>Bacillus subtilis</i>
1246	Phage-like element PBSX protein xkdF [<i>Bacillus subtilis</i>]	UniRef100_P54326	<i>Bacillus subtilis</i>
1247	Phage-like element PBSX protein xkdG [<i>Bacillus subtilis</i>]	UniRef100_P54327	<i>Bacillus subtilis</i>
1248	Hypothetical protein yqbG [<i>Bacillus subtilis</i>]	UniRef100_P45923	<i>Bacillus subtilis</i>
1249	Hypothetical protein yqbH [<i>Bacillus subtilis</i>]	UniRef100_P45924	<i>Bacillus subtilis</i>
1250	Phage-like element PBSX protein xkdI [<i>Bacillus subtilis</i>]	UniRef100_P54329	<i>Bacillus subtilis</i>
1251	Phage-like element PBSX protein xkdJ [<i>Bacillus subtilis</i>]	UniRef100_P54330	<i>Bacillus subtilis</i>
1252	Lin1277 protein [<i>Listeria innocua</i>]	UniRef100_Q92CB2	<i>Listeria innocua</i>
1253	Phage-like element PBSX protein xkdK [<i>Bacillus subtilis</i>]	UniRef100_P54331	<i>Bacillus subtilis</i>
			XkdK

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1254	Phage-like element PBSX protein xkdM [<i>Bacillus subtilis</i>]	UniRef100_P54332	<i>Bacillus subtilis</i>	XkdM
1255	Phage-like element PBSX protein xkdN [<i>Bacillus subtilis</i>]	UniRef100_P54333	<i>Bacillus subtilis</i>	XkdN
1256	Phage-like element PBSX protein xkdO [<i>Bacillus subtilis</i>]	UniRef100_P54334	<i>Bacillus subtilis</i>	XkdO
1257	Phage-like element PBSX protein xkdP [<i>Bacillus subtilis</i>]	UniRef100_P54335	<i>Bacillus subtilis</i>	YqbP
1258	Hypothetical protein yqbQ [<i>Bacillus subtilis</i>]	UniRef100_P45950	<i>Bacillus subtilis</i>	YqbQ
1259	Hypothetical protein yqbR [<i>Bacillus subtilis</i>]	UniRef100_P45933	<i>Bacillus subtilis</i>	YqbR
1260	Phage-like element PBSX protein xkdS [<i>Bacillus subtilis</i>]	UniRef100_P54338	<i>Bacillus subtilis</i>	XkdS
1261	Hypothetical protein yqbT [<i>Bacillus subtilis</i>]	UniRef100_P45935	<i>Bacillus subtilis</i>	YqbT
1262	Phage-like element PBSX protein xkdU [<i>Bacillus subtilis</i>]	UniRef100_P54340	<i>Bacillus subtilis</i>	XkdU
1263				
1264				XkdV
1265				YomR
1266				
1267				
1268				
1269	Regulatory protein [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9ZFL9	<i>Bacillus stearothermophilus</i>	BlyA YdhC
1270	Zinc-containing alcohol dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O35045	<i>Bacillus subtilis</i>	YjmD
1271	Mannonate dehydratase 1 [<i>Bacillus halodurans</i>]	UniRef100_Q9KZD8	<i>Bacillus halodurans</i>	UxuA
1272	D-mannonate oxidoreductase [<i>Bacillus halodurans</i>]	UniRef100_Q9KZD4	<i>Bacillus halodurans</i>	YjmF
1273	UPI00002F2634 UniRef100 entry	UniRef100_UPI00002F2634		YjmD
1274	Hexuronate transporter [<i>Bacillus subtilis</i>]	UniRef100_O34456	<i>Bacillus subtilis</i>	ExuT
1275	Stage II sporulation protein SB [<i>Bacillus subtilis</i>]	UniRef100_O34800	<i>Bacillus subtilis</i>	
1276	Stage II sporulation protein SA [<i>Bacillus subtilis</i>]	UniRef100_O34853	<i>Bacillus subtilis</i>	SpoIIISA
1277	UPI00003CC121 UniRef100 entry	UniRef100_UPI00003CC121		Pit
1278	Hypothetical UPF0111 protein ykaA [<i>Bacillus subtilis</i>]	UniRef100_O34454	<i>Bacillus subtilis</i>	YkaA
1279				Ggt
1280	YesL protein [<i>Bacillus subtilis</i>]	UniRef100_O31515	<i>Bacillus subtilis</i>	YesL
1281	YesM protein [<i>Bacillus subtilis</i>]	UniRef100_O31516	<i>Bacillus subtilis</i>	YesM
1282	YesN protein [<i>Bacillus subtilis</i>]	UniRef100_O31517	<i>Bacillus subtilis</i>	YesN
1283	YesO protein [<i>Bacillus subtilis</i>]	UniRef100_O31518	<i>Bacillus subtilis</i>	YesO
1284	Probable ABC transporter permease protein yesP [<i>Bacillus subtilis</i>]	UniRef100_O31519	<i>Bacillus subtilis</i>	YesP
1285	Probable ABC transporter permease protein yesQ [<i>Bacillus subtilis</i>]	UniRef100_O31520	<i>Bacillus subtilis</i>	YesQ
1286	YesR protein [<i>Bacillus subtilis</i>]	UniRef100_O31521	<i>Bacillus subtilis</i>	YesR
1287	YesS protein [<i>Bacillus subtilis</i>]	UniRef100_O31522	<i>Bacillus subtilis</i>	YesS
1288	YesT protein [<i>Bacillus subtilis</i>]	UniRef100_O31523	<i>Bacillus subtilis</i>	YesT
1289				YesU
1290	YesV protein [<i>Bacillus subtilis</i>]	UniRef100_O31525	<i>Bacillus subtilis</i>	YesV
1291	YesW protein [<i>Bacillus subtilis</i>]	UniRef100_O31526	<i>Bacillus subtilis</i>	YesW
1292				
1293				YesT
1294	YesX protein [<i>Bacillus subtilis</i>]	UniRef100_O31527	<i>Bacillus subtilis</i>	YesX
1295	Putative ion-channel protein [<i>Salmonella typhi</i>]	UniRef100_Q8Z4X6	<i>Salmonella typhi</i>	YccK
1296	YesY protein [<i>Bacillus subtilis</i>]	UniRef100_O31528	<i>Bacillus subtilis</i>	YesY
1297	YesZ protein [<i>Bacillus subtilis</i>]	UniRef100_O31529	<i>Bacillus subtilis</i>	YesZ
1298				YetA
1299	Lipoprotein lplA precursor [<i>Bacillus subtilis</i>]	UniRef100_P37966	<i>Bacillus subtilis</i>	LplA

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1300	LplB protein [<i>Bacillus subtilis</i>]	UniRef100_P39128	<i>Bacillus subtilis</i>	LplB
1301	LplC protein [<i>Bacillus subtilis</i>]	UniRef100_P39129	<i>Bacillus subtilis</i>	LplC
1302	YkbA protein [<i>Bacillus subtilis</i>]	UniRef100_O34739	<i>Bacillus subtilis</i>	YkbA
1303	YkcA protein [<i>Bacillus subtilis</i>]	UniRef100_O34689	<i>Bacillus subtilis</i>	YkcA
1304	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81CP9	<i>Bacillus cereus</i>	
1305	Probable serine protease do-like htrA [<i>Bacillus subtilis</i>]	UniRef100_O34358	<i>Bacillus subtilis</i>	HtrA
1306	Pyrroline-5-carboxylate reductase 3 [<i>Bacillus subtilis</i>]	UniRef100_Q00777	<i>Bacillus subtilis</i>	ProG
1307	D-aminopeptidase [<i>Bacillus subtilis</i>]	UniRef100_P26902	<i>Bacillus subtilis</i>	DppA
1308	Dipeptide transport system permease protein dppB [<i>Bacillus subtilis</i>]	UniRef100_P26903	<i>Bacillus subtilis</i>	DppB
1309	Dipeptide transport system permease protein dppC [<i>Bacillus subtilis</i>]	UniRef100_P26904	<i>Bacillus subtilis</i>	DppC
1310	Dipeptide transport ATP-binding protein dppD [<i>Bacillus subtilis</i>]	UniRef100_P26905	<i>Bacillus subtilis</i>	DppD
1311	Dipeptide-binding protein dppE precursor [<i>Bacillus subtilis</i>]	UniRef100_P26906	<i>Bacillus subtilis</i>	DppE
1312	Hypothetical protein ykfA [<i>Bacillus subtilis</i>]	UniRef100_O34851	<i>Bacillus subtilis</i>	YkfA
1313	Ykfb [<i>Bacillus subtilis</i>]	UniRef100_O34508	<i>Bacillus subtilis</i>	Ykfb
1314	Ykfc [<i>Bacillus subtilis</i>]	UniRef100_O35010	<i>Bacillus subtilis</i>	Ykfc
1315	Ykfd [<i>Bacillus subtilis</i>]	UniRef100_O34480	<i>Bacillus subtilis</i>	Ykfd
1316	BH1779 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KJBZ5	<i>Bacillus halodurans</i>	YkgA
1317	Putative acyl-CoA thioester hydrolase ykhA [<i>Bacillus subtilis</i>]	UniRef100_P49851	<i>Bacillus subtilis</i>	YkhA
1318				YkjA
1319	Pectate lyase 47 precursor [<i>Bacillus</i> sp. TS-47]	UniRef100_Q9AJM4	<i>Bacillus</i> sp. TS-47	Pel
1320	Transcriptional regulator, PadR family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HED5	<i>Bacillus thuringiensis</i>	
1321	Hypothetical protein OB0568 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESQ3	<i>Oceanobacillus iheyensis</i>	
1322	BH1312 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDA2	<i>Bacillus halodurans</i>	
1323	Hypothetical protein ykkaA [<i>Bacillus subtilis</i>]	UniRef100_P49854	<i>Bacillus subtilis</i>	YkkaA
1324	Hypothetical protein ykkC [<i>Bacillus subtilis</i>]	UniRef100_P49856	<i>Bacillus subtilis</i>	YkkC
1325				
1326	Ykke [<i>Bacillus subtilis</i>]	UniRef100_O34990	<i>Bacillus subtilis</i>	Ykke
1327	Glutamate 5-kinase 1 [<i>Bacillus subtilis</i>]	UniRef100_P39820	<i>Bacillus subtilis</i>	ProB
1328	Gamma-glutamyl phosphate reductase [<i>Bacillus subtilis</i>]	UniRef100_P39821	<i>Bacillus subtilis</i>	ProA
1329	Organic hydroperoxide resistance protein ohrA [<i>Bacillus subtilis</i>]	UniRef100_O34762	<i>Bacillus subtilis</i>	YkIA
1330	Organic hydroperoxide resistance transcriptional regulator [<i>Bacillus subtilis</i>]	UniRef100_O34777	<i>Bacillus subtilis</i>	YkmA
1331	Organic hydroperoxide resistance protein ohrB [<i>Bacillus subtilis</i>]	UniRef100_P80242	<i>Bacillus subtilis</i>	YkzA
1332				
1333	Guanine deaminase [<i>Bacillus subtilis</i>]	UniRef100_O34598	<i>Bacillus subtilis</i>	GuaD
1334	Phosphoglycerate mutase [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9ALU0	<i>Bacillus stearothermophilus</i>	YhfR
1335				
1336				
1337	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P80877	<i>Bacillus subtilis</i>	MetE
1338	Intracellular serine protease [<i>Bacillus</i> sp. WRD-2]	UniRef100_Q69DB4	<i>Bacillus</i> sp. WRD-2	IspA
1339				
1340	YkoK [<i>Bacillus subtilis</i>]	UniRef100_O34442	<i>Bacillus subtilis</i>	YkoK
1341				
1342	Integrase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETV2	<i>Oceanobacillus iheyensis</i>	YdcL
1343				YqaB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1344	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81UE0	<i>Bacillus anthracis</i>	
1345				YonS
1346				
1347	Putative HTH-type transcriptional regulator yqaE [<i>Bacillus subtilis</i>]	UniRef100_P45902	<i>Bacillus subtilis</i>	YqaE
1348	Transcriptional regulator [<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> ATCC 49256]	UniRef100_Q7P886	<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> ATCC 49256	
1349				
1350				
1351	Lin1236 protein [<i>Listeria innocua</i>]	UniRef100_Q92CD6	<i>Listeria innocua</i>	
1352				
1353				
1354				
1355				YqaJ
1356	35 protein [Bacteriophage SPP1]	UniRef100_Q38143	Bacteriophage SPP1	YqaK
1357				YqaL
1358				YqaM
1359	Hypothetical protein yqaO [<i>Bacillus subtilis</i>]	UniRef100_P45912	<i>Bacillus subtilis</i>	
1360	Hypothetical protein yopY [Bacteriophage SPBc2]	UniRef100_O64108	Bacteriophage SPBc2	
1361				
1362				
1363				
1364	Hypothetical protein CTC02137 [<i>Clostridium tetani</i>]	UniRef100_Q892G2	<i>Clostridium tetani</i>	
1365	Hypothetical protein MW1918 [<i>Staphylococcus aureus</i>]	UniRef100_Q8NVN5	<i>Staphylococcus aureus</i>	YqaN
1366				
1367	Single-strand binding protein 2 [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y4X1	<i>Listeria monocytogenes</i>	
1368				
1369				
1370	Hypothetical protein yqaQ [<i>Bacillus subtilis</i>]	UniRef100_P45948	<i>Bacillus subtilis</i>	YqaQ
1371				
1372	Hypothetical protein yqaS [<i>Bacillus subtilis</i>]	UniRef100_P45915	<i>Bacillus subtilis</i>	YqaS
1373	Hypothetical protein yqaT [<i>Bacillus subtilis</i>]	UniRef100_P45916	<i>Bacillus subtilis</i>	YqaT
1374	Hypothetical phage associated protein SpyM3_1326 [<i>Streptococcus pyogenes</i>]	UniRef100_Q8K6I0	<i>Streptococcus pyogenes</i>	
1375	Minor head structural component GP7 [Bacteriophage SPP1]	UniRef100_Q38442	Bacteriophage SPP1	
1376				
1377				
1378				
1379	Hypothetical protein CTC01553 [<i>Clostridium tetani</i>]	UniRef100_Q894J0	<i>Clostridium tetani</i>	
1380	Major capsid protein [Bacteriophage A118]	UniRef100_Q9T1B7	Bacteriophage A118	
1381	ORF28 [Bacteriophage phi-105]	UniRef100_Q9ZXF5	Bacteriophage phi-105	
1382				
1383	15 protein [Bacteriophage SPP1]	UniRef100_Q38584	Bacteriophage SPP1	
1384	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48446	Bacteriophage SPP1	
1385				
1386	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48448	Bacteriophage SPP1	
1387	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48449	Bacteriophage SPP1	
1388				
1389	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48453	Bacteriophage SPP1	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1390				
1391	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48455	Bacteriophage SPP1	XkdO
1392	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48459	Bacteriophage SPP1	
1393	Complete nucleotide sequence [Bacteriophage SPP1]	UniRef100_O48463	Bacteriophage SPP1	
1394				
1395				
1396	LycA [<i>Clostridium botulinum</i>]	UniRef100_Q6RI00	<i>Clostridium botulinum</i>	
1397	Hypothetical protein yrkC [<i>Bacillus subtilis</i>]	UniRef100_P54430	<i>Bacillus subtilis</i>	YrkC
1398				YhjR
1399	YdfS protein [<i>Bacillus subtilis</i>]	UniRef100_P96697	<i>Bacillus subtilis</i>	YdfS
1400	HTH-type transcriptional regulator tnrA [<i>Bacillus subtilis</i>]	UniRef100_Q45666	<i>Bacillus subtilis</i>	TnrA
1401	Hypothetical protein ykzB [<i>Bacillus subtilis</i>]	UniRef100_O34923	<i>Bacillus subtilis</i>	
1402				
1403	YkoM [<i>Bacillus subtilis</i>]	UniRef100_O34949	<i>Bacillus subtilis</i>	YkoM
1404	YkoU protein [<i>Bacillus subtilis</i>]	UniRef100_O34398	<i>Bacillus subtilis</i>	YkoU
1405	YkoV protein [<i>Bacillus subtilis</i>]	UniRef100_O34859	<i>Bacillus subtilis</i>	YkoV
1406	Signaling protein ykoW [<i>Bacillus subtilis</i>]	UniRef100_O34311	<i>Bacillus subtilis</i>	YkoW
1407	YkoX protein [<i>Bacillus subtilis</i>]	UniRef100_O34908	<i>Bacillus subtilis</i>	YkoX
1408	YkoY protein [<i>Bacillus subtilis</i>]	UniRef100_O34997	<i>Bacillus subtilis</i>	YkoY
1409	RNA polymerase sigma factor [<i>Bacillus subtilis</i>]	UniRef100_O31654	<i>Bacillus subtilis</i>	SigI
1410				YkrI
1411	Small, acid-soluble spore protein C3 [<i>Bacillus megaterium</i>]	UniRef100_P10572	<i>Bacillus megaterium</i>	
1412	YkrK protein [<i>Bacillus subtilis</i>]	UniRef100_O31656	<i>Bacillus subtilis</i>	YkrK
1413	Probable protease htpX homolog [<i>Bacillus subtilis</i>]	UniRef100_O31657	<i>Bacillus subtilis</i>	YkrL
1414	YkrM protein [<i>Bacillus subtilis</i>]	UniRef100_O31658	<i>Bacillus subtilis</i>	YkrM
1415	Penicillin-binding protein 3 [<i>Bacillus subtilis</i>]	UniRef100_P42971	<i>Bacillus subtilis</i>	PbpC
1416	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q73BI4	<i>Bacillus cereus</i>	
1417				
1418	YkrP protein [<i>Bacillus subtilis</i>]	UniRef100_O31660	<i>Bacillus subtilis</i>	YkrP
1419	Two-component sensor histidine kinase [<i>Bacillus subtilis</i>]	UniRef100_O31661	<i>Bacillus subtilis</i>	KinE
1420	Methylated-DNA--protein-cysteine methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P11742	<i>Bacillus subtilis</i>	Ogt
1421				
1422	Methylthioribose-1-phosphate isomerase [<i>Bacillus subtilis</i>]	UniRef100_O31662	<i>Bacillus subtilis</i>	YkrS
1423	Methylthioribose kinase [<i>Bacillus subtilis</i>]	UniRef100_O31663	<i>Bacillus subtilis</i>	YkrT
1424	YkrU protein [<i>Bacillus subtilis</i>]	UniRef100_O31664	<i>Bacillus subtilis</i>	YkrU
1425	Transaminase mtnE [<i>Bacillus subtilis</i>]	UniRef100_O31665	<i>Bacillus subtilis</i>	YkrV
1426	2,3-diketo-5-methylthiopentyl-1-phosphate enolase [<i>Bacillus subtilis</i>]	UniRef100_O31666	<i>Bacillus subtilis</i>	YkrW
1427	Methylthioribulose-1-phosphate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_O31668	<i>Bacillus subtilis</i>	YkrY
1428	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase [<i>Bacillus subtilis</i>]	UniRef100_O31669	<i>Bacillus subtilis</i>	YkrZ
1429	Metallothiol transferase fosB [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CXX5	<i>Oceanobacillus iheyensis</i>	YndN
1430	YkvA protein [<i>Bacillus subtilis</i>]	UniRef100_O31670	<i>Bacillus subtilis</i>	
1431	Stage 0 sporulation regulatory protein [<i>Bacillus subtilis</i>]	UniRef100_P05043	<i>Bacillus subtilis</i>	
1432	Two-component sensor histidine kinase [<i>Bacillus subtilis</i>]	UniRef100_O31671	<i>Bacillus subtilis</i>	KinD
1433				YkvE
1434	Chemotaxis motB protein [<i>Bacillus subtilis</i>]	UniRef100_P28612	<i>Bacillus subtilis</i>	MotB

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1435	Chemotaxis motA protein [<i>Bacillus subtilis</i>]	UniRef100_P28611	<i>Bacillus subtilis</i>	MotA
1436	ATP-dependent Clp protease-like [<i>Bacillus subtilis</i>]	UniRef100_O31673	<i>Bacillus subtilis</i>	ClpE
1437	YkvI protein [<i>Bacillus subtilis</i>]	UniRef100_O31674	<i>Bacillus subtilis</i>	YkvI
1438	YkvJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31675	<i>Bacillus subtilis</i>	YkvJ
1439	YkvK protein [<i>Bacillus subtilis</i>]	UniRef100_O31676	<i>Bacillus subtilis</i>	YkvK
1440	YkvL protein [<i>Bacillus subtilis</i>]	UniRef100_O31677	<i>Bacillus subtilis</i>	YkvL
1441	YkvM protein [<i>Bacillus subtilis</i>]	UniRef100_O31678	<i>Bacillus subtilis</i>	YkvM
1442	DNA integration/recombination protein [<i>Clostridium tetani</i>]	UniRef100_Q894H7	<i>Clostridium tetani</i>	CodV
1443	Integrase/recombinase [<i>Bacillus cereus</i> ZK]	UniRef100_Q633V7	<i>Bacillus cereus</i> ZK	RipX
1444				
1445				
1446				
1447	YoqV protein [Bacteriophage SPBc2]	UniRef100_O64130	Bacteriophage SPBc2	LigB
1448				
1449				
1450	UPI00003CC586 UniRef100 entry	UniRef100_UPI00003CC586		
1451				
1452				
1453				
1454				
1455	Prophage LambdaBa02, HNH endonuclease family protein [<i>Bacillus anthracis</i>]	UniRef100_Q81W86	<i>Bacillus anthracis</i>	
1456	Terminase small subunit [<i>Staphylococcus aureus</i>]	UniRef100_Q6GAL5	<i>Staphylococcus aureus</i>	
1457	Prophage LambdaBa02, terminase, large subunit, putative [<i>Bacillus anthracis</i>]	UniRef100_Q6HUD2	<i>Bacillus anthracis</i>	
1458	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81W89	<i>Bacillus anthracis</i>	
1459	ClpP family serine protease, possible phage related [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HW4	<i>Clostridium acetobutylicum</i>	ClpP
1460	Prophage LambdaBa02, major capsid protein, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81W91	<i>Bacillus anthracis</i>	
1461	Precursor polypeptide (AA -37 to 1647) precursor [unidentified bacterium]	UniRef100_Q03658	unidentified bacterium	
1462	Gp7 protein [Bacteriophage phi3626]	UniRef100_Q8SBP7	Bacteriophage phi3626	
1463	Uncharacterized phage related protein [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HW7	<i>Clostridium acetobutylicum</i>	
1464	Hypothetical protein CAC1887 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HW9	<i>Clostridium acetobutylicum</i>	
1465	Prophage LambdaBa02, major tail protein, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81W97	<i>Bacillus anthracis</i>	
1466				
1467				
1468	Prophage LambdaBa02, tape measure protein, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81WA0	<i>Bacillus anthracis</i>	YqbO
1469	Lin2382 protein [<i>Listeria innocua</i>]	UniRef100_Q928Z8	<i>Listeria innocua</i>	
1470	Protein gp18 [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y4Z4	<i>Listeria monocytogenes</i>	
1471				YclG
1472				XkdV
1473				XkdW
1474	YomP protein [Bacteriophage SPBc2]	UniRef100_O64052	Bacteriophage SPBc2	
1475	Glycerophosphoryl diester phosphodiesterase, putative [<i>Bacillus cereus</i>]	UniRef100_Q737E6	<i>Bacillus cereus</i>	GlpQ
1476	Protein bh1A [Bacteriophage SPBc2]	UniRef100_O64039	Bacteriophage SPBc2	
1477	ORF46 [Bacteriophage phi-105]	UniRef100_Q9ZXD7	Bacteriophage phi-105	XlyB
1478				
1479				
1480				

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1481	Transcriptional regulator, DeoR family [<i>Bacillus cereus</i>]	UniRef100_Q816D5	<i>Bacillus cereus</i>	
1482	Hypothetical protein yold [Bacteriophage SPBe2]	UniRef100_O64030	Bacteriophage SPBe2	
1483	DNA integration/recombination/inversion protein [<i>Bacillus cereus</i>]	UniRef100_Q81GD4	<i>Bacillus cereus</i>	YdcL
1484	YkvM protein [<i>Bacillus subtilis</i>]	UniRef100_O31678	<i>Bacillus subtilis</i>	
1485	Response regulator aspartate phosphatase H [<i>Bacillus subtilis</i>]	UniRef100_P40771	<i>Bacillus subtilis</i>	RapH
1486	YoaT [<i>Bacillus subtilis</i>]	UniRef100_O34535	<i>Bacillus subtilis</i>	YoaT
1487	YozG protein [<i>Bacillus subtilis</i>]	UniRef100_O31834	<i>Bacillus subtilis</i>	
1488	YoaS protein [<i>Bacillus subtilis</i>]	UniRef100_O31833	<i>Bacillus subtilis</i>	YoaS
1489				
1490				
1491				
1492				
1493				
1494	YkvS protein [<i>Bacillus subtilis</i>]	UniRef100_O31684	<i>Bacillus subtilis</i>	
1495	BH2327 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KAG0	<i>Bacillus halodurans</i>	
1496	YkvT protein [<i>Bacillus subtilis</i>]	UniRef100_O31685	<i>Bacillus subtilis</i>	YkvT
1497	YkvU protein [<i>Bacillus subtilis</i>]	UniRef100_O31686	<i>Bacillus subtilis</i>	YkvU
1498	YkvV protein [<i>Bacillus subtilis</i>]	UniRef100_O31687	<i>Bacillus subtilis</i>	YkvV
1499				YkvW
1500	YkvY protein [<i>Bacillus subtilis</i>]	UniRef100_O31689	<i>Bacillus subtilis</i>	YkvY
1501	Necrosis and ethylene inducing protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFT2	<i>Bacillus halodurans</i>	
1502	Putative HTH-type transcriptional regulator ykvZ [<i>Bacillus subtilis</i>]	UniRef100_O31690	<i>Bacillus subtilis</i>	YkvZ
1503	Transcription antiterminator [<i>Bacillus subtilis</i>]	UniRef100_O06710	<i>Bacillus subtilis</i>	GlcT
1504				PtsG
1505	Phosphocarrier protein HPr [<i>Bacillus subtilis</i>]	UniRef100_P08877	<i>Bacillus subtilis</i>	
1506	Phosphoenolpyruvate-protein phosphotransferase [<i>Bacillus subtilis</i>]	UniRef100_P08838	<i>Bacillus subtilis</i>	PtsI
1507	SplA [<i>Bacillus amyloliquefaciens</i>]	UniRef100_O54358	<i>Bacillus amyloliquefaciens</i>	
1508	Spore photoproduct lyase [<i>Bacillus amyloliquefaciens</i>]	UniRef100_O54359	<i>Bacillus amyloliquefaciens</i>	SplB
1509	Hypothetical protein orf1 [<i>Bacillus subtilis</i>]	UniRef100_O05187	<i>Bacillus subtilis</i>	YkwB
1510	Methyl-accepting chemotaxis protein mcpC [<i>Bacillus subtilis</i>]	UniRef100_P54576	<i>Bacillus subtilis</i>	McpC
1511	3-oxoacyl-(Acyl carrier protein) reductase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMP9	<i>Oceanobacillus iheyensis</i>	FabG
1512	Hypothetical oxidoreductase ykwC [<i>Bacillus subtilis</i>]	UniRef100_O34948	<i>Bacillus subtilis</i>	YkwC
1513	YkwD protein [<i>Bacillus subtilis</i>]	UniRef100_O31694	<i>Bacillus subtilis</i>	YkwD
1514	YkuA protein [<i>Bacillus subtilis</i>]	UniRef100_O31399	<i>Bacillus subtilis</i>	YkuA
1515	Sporulation kinase A [<i>Bacillus subtilis</i>]	UniRef100_P16497	<i>Bacillus subtilis</i>	KinA
1516	Putative aminotransferase A [<i>Bacillus subtilis</i>]	UniRef100_P16524	<i>Bacillus subtilis</i>	PatA
1517				
1518	Hypothetical protein yxaI [<i>Bacillus subtilis</i>]	UniRef100_P42108	<i>Bacillus subtilis</i>	YxaI
1519				YxiO
1520	Chemotaxis protein cheV [<i>Bacillus subtilis</i>]	UniRef100_P37599	<i>Bacillus subtilis</i>	CheV
1521	Hypothetical protein ykyB [<i>Bacillus subtilis</i>]	UniRef100_P42430	<i>Bacillus subtilis</i>	YkyB
1522				YkuC
1523	YkuD protein [<i>Bacillus subtilis</i>]	UniRef100_O34816	<i>Bacillus subtilis</i>	YkuD
1524				YkuE
1525	Hypothetical oxidoreductase ykuF [<i>Bacillus subtilis</i>]	UniRef100_O34717	<i>Bacillus subtilis</i>	YkuF
1526	YkuI protein [<i>Bacillus subtilis</i>]	UniRef100_O35014	<i>Bacillus subtilis</i>	YkuI
1527				
1528	YkuJ protein [<i>Bacillus subtilis</i>]	UniRef100_O34588	<i>Bacillus subtilis</i>	
1529	YkuK protein [<i>Bacillus subtilis</i>]	UniRef100_O34776	<i>Bacillus subtilis</i>	YkuK

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1530	Hypothetical protein ykzF [<i>Bacillus subtilis</i>]	UniRef100_O31697	<i>Bacillus subtilis</i>	
1531	YkuL protein [<i>Bacillus subtilis</i>]	UniRef100_O31698	<i>Bacillus subtilis</i>	YkuL
1532	Putative HTH-type transcriptional regulator ykuM [<i>Bacillus subtilis</i>]	UniRef100_O34827	<i>Bacillus subtilis</i>	CepC
1533	Probable flavodoxin 1 [<i>Bacillus subtilis</i>]	UniRef100_O34737	<i>Bacillus subtilis</i>	YkuN
1534	YkuO protein [<i>Bacillus subtilis</i>]	UniRef100_O34879	<i>Bacillus subtilis</i>	YkuO
1535	Probable flavodoxin 2 [<i>Bacillus subtilis</i>]	UniRef100_O34589	<i>Bacillus subtilis</i>	YkuP
1536	YkuQ protein [<i>Bacillus subtilis</i>]	UniRef100_O34981	<i>Bacillus subtilis</i>	YkuQ
1537	YkuR protein [<i>Bacillus subtilis</i>]	UniRef100_O34916	<i>Bacillus subtilis</i>	YkuR
1538	Hypothetical UPF0180 protein ykuS [<i>Bacillus subtilis</i>]	UniRef100_O34783	<i>Bacillus subtilis</i>	
1539	YkuU protein [<i>Bacillus subtilis</i>]	UniRef100_O34564	<i>Bacillus subtilis</i>	YkuU
1540	YkuV protein [<i>Bacillus subtilis</i>]	UniRef100_O31403	<i>Bacillus subtilis</i>	YkuV
1541	Repressor rok [<i>Bacillus subtilis</i>]	UniRef100_O34857	<i>Bacillus subtilis</i>	Rok
1542	YknT protein [<i>Bacillus subtilis</i>]	UniRef100_O31700	<i>Bacillus subtilis</i>	YknT
1543				MobA
1544	Molybdopterin biosynthesis protein MoeB [<i>Bacillus subtilis</i>]	UniRef100_O31702	<i>Bacillus subtilis</i>	MoeB
1545	Molybdopterin biosynthesis protein MoeA [<i>Bacillus subtilis</i>]	UniRef100_O31703	<i>Bacillus subtilis</i>	MoeA
1546	Molybdopterin-guanine dinucleotide biosynthesis protein B [<i>Bacillus subtilis</i>]	UniRef100_O31704	<i>Bacillus subtilis</i>	MobB
1547	Molybdopterin converting factor subunit 2 [<i>Bacillus subtilis</i>]	UniRef100_O31705	<i>Bacillus subtilis</i>	MoaE
1548	Molybdopterin converting factor, subunit 1 [<i>Bacillus subtilis</i>]	UniRef100_O31706	<i>Bacillus subtilis</i>	
1549				
1550	YknU protein [<i>Bacillus subtilis</i>]	UniRef100_O31707	<i>Bacillus subtilis</i>	YknU
1551	YknV protein [<i>Bacillus subtilis</i>]	UniRef100_O31708	<i>Bacillus subtilis</i>	YknV
1552	Hypothetical protein yknW [<i>Bacillus subtilis</i>]	UniRef100_O31709	<i>Bacillus subtilis</i>	YknW
1553	YknX protein [<i>Bacillus subtilis</i>]	UniRef100_O31710	<i>Bacillus subtilis</i>	YknX
1554	YknY protein [<i>Bacillus subtilis</i>]	UniRef100_O31711	<i>Bacillus subtilis</i>	YknY
1555	Hypothetical protein yknZ [<i>Bacillus subtilis</i>]	UniRef100_O31712	<i>Bacillus subtilis</i>	YknZ
1556				FruR
1557	1-phosphofructokinase [<i>Bacillus subtilis</i>]	UniRef100_O31714	<i>Bacillus subtilis</i>	FruK
1558	Phosphotransferase system (PTS) fructose-specific enzyme IIABC component [<i>Bacillus subtilis</i>]	UniRef100_P71012	<i>Bacillus subtilis</i>	FruA
1559	Signal peptidase I T [<i>Bacillus subtilis</i>]	UniRef100_P71013	<i>Bacillus subtilis</i>	SipT
1560	Hypothetical protein ykoA [<i>Bacillus subtilis</i>]	UniRef100_O31715	<i>Bacillus subtilis</i>	
1561				
1562	YkpA protein [<i>Bacillus subtilis</i>]	UniRef100_O31716	<i>Bacillus subtilis</i>	YkpA
1563	BH1921 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KKB5	<i>Bacillus halodurans</i>	
1564	Aminopeptidase ampS [<i>Bacillus subtilis</i>]	UniRef100_P39762	<i>Bacillus subtilis</i>	AmpS
1565				
1566	MreBH protein [<i>Bacillus subtilis</i>]	UniRef100_P39763	<i>Bacillus subtilis</i>	MreBH
1567				
1568	Sporulation kinase C [<i>Bacillus subtilis</i>]	UniRef100_P39764	<i>Bacillus subtilis</i>	KinC
1569	Hypothetical protein ykqB [<i>Bacillus subtilis</i>]	UniRef100_P39760	<i>Bacillus subtilis</i>	YkqB
1570	Adenine deaminase [<i>Bacillus subtilis</i>]	UniRef100_P39761	<i>Bacillus subtilis</i>	AdeC
1571				YkqC
1572	YkzG protein [<i>Bacillus subtilis</i>]	UniRef100_O31718	<i>Bacillus subtilis</i>	
1573	Hypothetical protein ykrA [<i>Bacillus subtilis</i>]	UniRef100_Q45494	<i>Bacillus subtilis</i>	YkrA
1574				YkrB
1575				
1576	Hypothetical protein ykyA [<i>Bacillus subtilis</i>]	UniRef100_P21884	<i>Bacillus subtilis</i>	YkyA
1577	Pyruvate dehydrogenase E1 component, alpha subunit [<i>Bacillus subtilis</i>]	UniRef100_P21881	<i>Bacillus subtilis</i>	PdhA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1578	Pyruvate dehydrogenase E1 component, beta subunit [<i>Bacillus subtilis</i>]	UniRef100_P21882	<i>Bacillus subtilis</i>	PdhB
1579	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex [<i>Bacillus subtilis</i>]	UniRef100_P21883	<i>Bacillus subtilis</i>	PdhC
1580	Dihydrolipoyl dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P21880	<i>Bacillus subtilis</i>	PdhD
1581	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
1582	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
1583				
1584				
1585	Arginine decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P21885	<i>Bacillus subtilis</i>	SpeA
1586	Hypothetical UPF0223 protein yktA [<i>Bacillus subtilis</i>]	UniRef100_Q45497	<i>Bacillus subtilis</i>	
1587	Hypothetical protein yktB [<i>Bacillus subtilis</i>]	UniRef100_Q45498	<i>Bacillus subtilis</i>	YktB
1588	YkzI protein [<i>Bacillus subtilis</i>]	UniRef100_O31719	<i>Bacillus subtilis</i>	
1589	Inositol-1-monophosphatase [<i>Bacillus subtilis</i>]	UniRef100_Q45499	<i>Bacillus subtilis</i>	YktC
1590	Hypothetical protein ykzC [<i>Bacillus subtilis</i>]	UniRef100_O31720	<i>Bacillus subtilis</i>	YkzC
1591	Hypothetical protein ylaA [<i>Bacillus subtilis</i>]	UniRef100_O07625	<i>Bacillus subtilis</i>	YlaA
1592	Hypothetical protein ylaB [<i>Bacillus subtilis</i>]	UniRef100_O07626	<i>Bacillus subtilis</i>	
1593	YlaC protein [<i>Bacillus subtilis</i>]	UniRef100_O07627	<i>Bacillus subtilis</i>	YlaC
1594	Hypothetical protein ylaD [<i>Bacillus subtilis</i>]	UniRef100_O07628	<i>Bacillus subtilis</i>	
1595	Hypothetical protein ylaF [<i>Bacillus subtilis</i>]	UniRef100_O07630	<i>Bacillus subtilis</i>	
1596	GTP-binding protein typA/bipA homolog [<i>Bacillus subtilis</i>]	UniRef100_O07631	<i>Bacillus subtilis</i>	YlaG
1597	YlaH protein [<i>Bacillus subtilis</i>]	UniRef100_O07632	<i>Bacillus subtilis</i>	YlaH
1598	YhzA homolog [<i>Bacillus subtilis</i>]	UniRef100_O07562	<i>Bacillus subtilis</i>	YhjH
1599	Hypothetical protein yhjG [<i>Bacillus subtilis</i>]	UniRef100_O07561	<i>Bacillus subtilis</i>	YhjG
1600				
1601	Hypothetical lipoprotein ylaJ precursor [<i>Bacillus subtilis</i>]	UniRef100_O07634	<i>Bacillus subtilis</i>	YlaJ
1602	YlaK protein [<i>Bacillus subtilis</i>]	UniRef100_O07635	<i>Bacillus subtilis</i>	YlaK
1603	UPI00003CB7B1 UniRef100 entry	UniRef100_UPI00003CB7B1		YlaL
1604	Probable glutaminase ylaM [<i>Bacillus subtilis</i>]	UniRef100_O07637	<i>Bacillus subtilis</i>	YlaM
1605	YlaN protein [<i>Bacillus subtilis</i>]	UniRef100_O07638	<i>Bacillus subtilis</i>	
1606	Hypothetical protein ylaO [<i>Bacillus subtilis</i>]	UniRef100_O07639	<i>Bacillus subtilis</i>	FtsW
1607				PycA
1608	Cytochrome AA3 controlling protein [<i>Bacillus subtilis</i>]	UniRef100_P12946	<i>Bacillus subtilis</i>	CtaA
1609	Protoheme IX farnesyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P24009	<i>Bacillus subtilis</i>	CtaB
1610	Cytochrome c oxidase polypeptide II precursor (EC 1.9.3.1) (Cytochrome aa3 subunit 2) (Caa-3605 subunit 2) (Oxidase aa(3) subunit 2) [<i>Bacillus subtilis</i>]	UniRef100_P24011	<i>Bacillus subtilis</i>	CtaC
1611	Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Cytochrome aa3 subunit 1) (Caa-3605 subunit 1) (Oxidase aa(3) subunit 1) [<i>Bacillus subtilis</i>]	UniRef100_P24010	<i>Bacillus subtilis</i>	CtaD
1612	Cytochrome c oxidase polypeptide III (EC 1.9.3.1) (Cytochrome aa3 subunit 3) (Caa-3605 subunit 3) (Oxidase aa(3) subunit 3) [<i>Bacillus subtilis</i>]	UniRef100_P24012	<i>Bacillus subtilis</i>	CtaE

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1613	Cytochrome c oxidase polypeptide IVB [<i>Bacillus subtilis</i>]	UniRef100_P24013	<i>Bacillus subtilis</i>	CtaF
1614	CtaG protein [<i>Bacillus subtilis</i>]	UniRef100_O34329	<i>Bacillus subtilis</i>	CtaG
1615	YlbA protein [<i>Bacillus subtilis</i>]	UniRef100_O34743	<i>Bacillus subtilis</i>	YlbA
1616	YlbB protein [<i>Bacillus subtilis</i>]	UniRef100_O34682	<i>Bacillus subtilis</i>	YlbB
1617	YlbC protein [<i>Bacillus subtilis</i>]	UniRef100_O34586	<i>Bacillus subtilis</i>	YlbC
1618	YlbD protein [<i>Bacillus subtilis</i>]	UniRef100_O34880	<i>Bacillus subtilis</i>	YlbD
1619	YlbE protein [<i>Bacillus subtilis</i>]	UniRef100_O34958	<i>Bacillus subtilis</i>	
1620	Regulatory protein ylbF [<i>Bacillus subtilis</i>]	UniRef100_O34412	<i>Bacillus subtilis</i>	YlbF
1621	Hypothetical UPF0298 protein ylbG [<i>Bacillus subtilis</i>]	UniRef100_O34658	<i>Bacillus subtilis</i>	
1622	YlbH protein [<i>Bacillus subtilis</i>]	UniRef100_O34331	<i>Bacillus subtilis</i>	YlbH
1623	Phosphopantetheine adenylyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34797	<i>Bacillus subtilis</i>	YlbI
1624				YlbJ
1625	YlbL protein [<i>Bacillus subtilis</i>]	UniRef100_O34470	<i>Bacillus subtilis</i>	YlbL
1626	YlbM protein [<i>Bacillus subtilis</i>]	UniRef100_O34513	<i>Bacillus subtilis</i>	YlbM
1627	YlbN protein [<i>Bacillus subtilis</i>]	UniRef100_O34445	<i>Bacillus subtilis</i>	YlbN
1628	50S ribosomal protein L32 [<i>Bacillus subtilis</i>]	UniRef100_O34687	<i>Bacillus subtilis</i>	
1629	Hypothetical protein ylbO [<i>Bacillus subtilis</i>]	UniRef100_O34549	<i>Bacillus subtilis</i>	YlbO
1630	YlbP protein [<i>Bacillus subtilis</i>]	UniRef100_O34468	<i>Bacillus subtilis</i>	YlbP
1631	Probable 2-dehydropanoate 2-reductase [<i>Bacillus subtilis</i>]	UniRef100_O34661	<i>Bacillus subtilis</i>	YlbQ
1632				YllA
1633	Protein mraZ [<i>Bacillus subtilis</i>]	UniRef100_P55343	<i>Bacillus subtilis</i>	YllB
1634	S-adenosyl-methyltransferase mraW [<i>Bacillus subtilis</i>]	UniRef100_Q07876	<i>Bacillus subtilis</i>	YlxA
1635				FtsL
1636	Penicillin-binding protein 2B [<i>Bacillus subtilis</i>]	UniRef100_Q07868	<i>Bacillus subtilis</i>	PbpB
1637	Stage V sporulation protein D [<i>Bacillus subtilis</i>]	UniRef100_Q03524	<i>Bacillus subtilis</i>	SpoVD
1638	UDP-N-acetylmuramoylalanine-D-glutamate-2,6-diaminopimelate ligase [<i>Bacillus subtilis</i>]	UniRef100_Q03523	<i>Bacillus subtilis</i>	MurE
1639	Phospho-N-acetylmuramoyl-pentapeptide-transferase [<i>Bacillus subtilis</i>]	UniRef100_Q03521	<i>Bacillus subtilis</i>	MraY
1640	UDP-N-acetylmuramoylalanine--D-glutamate ligase [<i>Bacillus subtilis</i>]	UniRef100_Q03522	<i>Bacillus subtilis</i>	MurD
1641	Stage V sporulation protein E [<i>Bacillus subtilis</i>]	UniRef100_P07373	<i>Bacillus subtilis</i>	SpoVE
1642	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase [<i>Bacillus subtilis</i>]	UniRef100_P37585	<i>Bacillus subtilis</i>	MurG
1643	UDP-N-acetylenolpyruvoylglucosamine reductase [<i>Bacillus subtilis</i>]	UniRef100_P18579	<i>Bacillus subtilis</i>	MurB
1644				DivIB
1645				YlxW
1646				YlxX
1647				Sbp
1648				FtsA
1649	Cell division protein ftsZ [<i>Bacillus subtilis</i>]	UniRef100_P17865	<i>Bacillus subtilis</i>	FtsZ
1650				Bpr
1651	Bacillopeptidase F precursor [<i>Bacillus subtilis</i>]	UniRef100_P16397	<i>Bacillus subtilis</i>	Bpr
1652	Sporulation sigma-E factor processing peptidase [<i>Bacillus subtilis</i>]	UniRef100_P13801	<i>Bacillus subtilis</i>	SpoIIGA
1653	RNA polymerase sigma-E factor precursor [<i>Bacillus subtilis</i>]	UniRef100_P06222	<i>Bacillus subtilis</i>	SigE
1654	RNA polymerase sigma-G factor [<i>Bacillus subtilis</i>]	UniRef100_P19940	<i>Bacillus subtilis</i>	SigG

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1655	YlmA protein [<i>Bacillus subtilis</i>]	UniRef100_O31723	<i>Bacillus subtilis</i>	YlmA
1656				
1657	YlmC protein [<i>Bacillus subtilis</i>]	UniRef100_O31725	<i>Bacillus subtilis</i>	
1658	Hypothetical UPF0124 protein ylmD [<i>Bacillus subtilis</i>]	UniRef100_O31726	<i>Bacillus subtilis</i>	YlmD
1659				YlmE
1660	YlmF protein [<i>Bacillus subtilis</i>]	UniRef100_O31728	<i>Bacillus subtilis</i>	YlmF
1661	YlmG protein [<i>Bacillus subtilis</i>]	UniRef100_O31729	<i>Bacillus subtilis</i>	
1662	Minicell-associated protein [<i>Bacillus subtilis</i>]	UniRef100_P71020	<i>Bacillus subtilis</i>	YlmH
1663	Minicell-associated protein DivIVA [<i>Bacillus subtilis</i>]	UniRef100_P71021	<i>Bacillus subtilis</i>	DivIVA
1664	Isoleucyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_Q45477	<i>Bacillus subtilis</i>	IleS
1665				YlyA
1666				LspA
1667	Hypothetical pseudouridine synthase ylyB [<i>Bacillus subtilis</i>]	UniRef100_Q45480	<i>Bacillus subtilis</i>	YlyB
1668	PyrR bifunctional protein [Includes: Pyrimidine operon regulatory protein; Uracil phosphoribosyltransferase (EC 2.4.2.9) (UPRTase)] [<i>Bacillus subtilis</i>]	UniRef100_P39765	Includes: Pyrimidine operon regulatory protein; Uracil phosphoribosyltransferase (EC 2.4.2.9) (UPRTase)	PyrR
1669	Uracil permease [<i>Bacillus subtilis</i>]	UniRef100_P39766	<i>Bacillus subtilis</i>	PyrP
1670	Aspartate carbamoyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P05654	<i>Bacillus subtilis</i>	PyrB
1671	Dihydroorotate [<i>Bacillus subtilis</i>]	UniRef100_P25995	<i>Bacillus subtilis</i>	PyrC
1672	Carbamoyl-phosphate synthase, pyrimidine-specific, small chain [<i>Bacillus subtilis</i>]	UniRef100_P25993	<i>Bacillus subtilis</i>	PyrAA
1673	Carbamoyl-phosphate synthase, pyrimidine-specific, large chain [<i>Bacillus subtilis</i>]	UniRef100_P25994	<i>Bacillus subtilis</i>	PyrAB
1674	Dihydroorotate dehydrogenase electron transfer subunit [<i>Bacillus subtilis</i>]	UniRef100_P25983	<i>Bacillus subtilis</i>	PyrK
1675	Dihydroorotate dehydrogenase, catalytic subunit [<i>Bacillus subtilis</i>]	UniRef100_P25996	<i>Bacillus subtilis</i>	PyrD
1676	Orotidine 5'-phosphate decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P25971	<i>Bacillus subtilis</i>	PyrF
1677	Orotate phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P25972	<i>Bacillus subtilis</i>	PyrE
1678				
1679				CysH
1680	YlnA protein [<i>Bacillus subtilis</i>]	UniRef100_O34734	<i>Bacillus subtilis</i>	CysP
1681	Sulfate adenyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34764	<i>Bacillus subtilis</i>	Sat
1682	Probable adenyl-sulfate kinase [<i>Bacillus subtilis</i>]	UniRef100_O34577	<i>Bacillus subtilis</i>	CysC
1683	Putative S-adenosyl L-methionine: uroporphyrinogen III methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34744	<i>Bacillus subtilis</i>	YlnD
1684	Sirohydrochlorin ferrochelatae [<i>Bacillus subtilis</i>]	UniRef100_O34632	<i>Bacillus subtilis</i>	YlnE
1685	YlnF protein [<i>Bacillus subtilis</i>]	UniRef100_O34813	<i>Bacillus subtilis</i>	YlnF
1686	Putative fibronectin-binding protein [<i>Bacillus subtilis</i>]	UniRef100_O34693	<i>Bacillus subtilis</i>	YloA
1687	YloB protein [<i>Bacillus subtilis</i>]	UniRef100_O34431	<i>Bacillus subtilis</i>	YloB
1688	YloC protein [<i>Bacillus subtilis</i>]	UniRef100_O34441	<i>Bacillus subtilis</i>	YloC
1689	Hypothetical UPF0296 protein ylzA [<i>Bacillus subtilis</i>]	UniRef100_Q7WY72	<i>Bacillus subtilis</i>	
1690	Guanylate kinase [<i>Bacillus subtilis</i>]	UniRef100_O34328	<i>Bacillus subtilis</i>	Gmk
1691	DNA-directed RNA polymerase omega chain [<i>Bacillus subtilis</i>]	UniRef100_O35011	<i>Bacillus subtilis</i>	
1692	YloI protein [<i>Bacillus subtilis</i>]	UniRef100_O35033	<i>Bacillus subtilis</i>	YloI

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1693	Primosomal protein N' [<i>Bacillus subtilis</i>]	UniRef100_P94461	<i>Bacillus subtilis</i>	PriA
1694	Peptide deformylase 1 [<i>Bacillus subtilis</i>]	UniRef100_P94462	<i>Bacillus subtilis</i>	Def
1695	Methionyl-tRNA formyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P94463	<i>Bacillus subtilis</i>	Fmt
1696	Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.—) (rRNA (cytosine-C(5)-)-methyltransferase) [<i>Bacillus subtilis</i>]	UniRef100_P94464	<i>Bacillus subtilis</i>	YloM
1697	Hypothetical UPF0063 protein yloN [<i>Bacillus subtilis</i>]	UniRef100_O34617	<i>Bacillus subtilis</i>	YloN
1698	Protein phosphatase [<i>Bacillus subtilis</i>]	UniRef100_O34779	<i>Bacillus subtilis</i>	PrpC
1699	Probable serine/threonine-protein kinase yloP [<i>Bacillus subtilis</i>]	UniRef100_O34507	<i>Bacillus subtilis</i>	PrkC
1700	Probable GTPase engC [<i>Bacillus subtilis</i>]	UniRef100_O34530	<i>Bacillus subtilis</i>	YloQ
1701	Ribulose-phosphate 3-epimerase [<i>Bacillus subtilis</i>]	UniRef100_O34557	<i>Bacillus subtilis</i>	Rpe
1702	YloS protein [<i>Bacillus subtilis</i>]	UniRef100_O34664	<i>Bacillus subtilis</i>	YloS
1703				
1704	50S ribosomal protein L28 [<i>Bacillus subtilis</i>]	UniRef100_P37807	<i>Bacillus subtilis</i>	
1705	Hypothetical protein yloU [<i>Bacillus subtilis</i>]	UniRef100_O34318	<i>Bacillus subtilis</i>	YloU
1706				YloV
1707	Probable L-serine dehydratase, beta chain [<i>Bacillus subtilis</i>]	UniRef100_O34635	<i>Bacillus subtilis</i>	SdaAB
1708	Probable L-serine dehydratase, alpha chain [<i>Bacillus subtilis</i>]	UniRef100_O34607	<i>Bacillus subtilis</i>	SdaAA
1709	ATP-dependent DNA helicase recG [<i>Bacillus subtilis</i>]	UniRef100_O34942	<i>Bacillus subtilis</i>	RecG
1710	Transcription factor fapR [<i>Bacillus subtilis</i>]	UniRef100_O34835	<i>Bacillus subtilis</i>	YlpC
1711	Fatty acid/phospholipid synthesis protein plsX [<i>Bacillus subtilis</i>]	UniRef100_P71018	<i>Bacillus subtilis</i>	PlsX
1712	Malonyl CoA-acyl carrier protein transacylase [<i>Bacillus subtilis</i>]	UniRef100_P71019	<i>Bacillus subtilis</i>	FabD
1713	3-oxoacyl-[acyl-carrier-protein] reductase [<i>Bacillus subtilis</i>]	UniRef100_P51831	acyl-carrier- protein <i>Bacillus subtilis</i>	FabG
1714	Acyl carrier protein [<i>Bacillus subtilis</i>]	UniRef100_P80643	<i>Bacillus subtilis</i>	
1715				Rnc
1716	Chromosome partition protein smc [<i>Bacillus subtilis</i>]	UniRef100_P51834	<i>Bacillus subtilis</i>	Smc
1717				FtsY
1718				
1719	Signal recognition particle protein [<i>Bacillus subtilis</i>]	UniRef100_P37105	<i>Bacillus subtilis</i>	Ffh
1720	30S ribosomal protein S16 [<i>Bacillus subtilis</i>]	UniRef100_P21474	<i>Bacillus subtilis</i>	
1721				
1722	YlqD protein [<i>Bacillus subtilis</i>]	UniRef100_O31739	<i>Bacillus subtilis</i>	YlqD
1723				RimM
1724				TmnD
1725	50S ribosomal protein L19 [<i>Bacillus subtilis</i>]	UniRef100_O31742	<i>Bacillus subtilis</i>	RplS
1726	YlqF protein [<i>Bacillus subtilis</i>]	UniRef100_O31743	<i>Bacillus subtilis</i>	YlqF
1727	Ribonuclease HII [<i>Bacillus subtilis</i>]	UniRef100_O31744	<i>Bacillus subtilis</i>	RnhB
1728	YlqG protein [<i>Bacillus subtilis</i>]	UniRef100_O31745	<i>Bacillus subtilis</i>	YlqG
1729	YlqH protein [<i>Bacillus subtilis</i>]	UniRef100_O34867	<i>Bacillus subtilis</i>	
1730	Succinyl-CoA synthetase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P80886	<i>Bacillus subtilis</i>	SucC
1731	Succinyl-CoA synthetase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P80865	<i>Bacillus subtilis</i>	SucD
1732				Smf
1733				TopA
1734	Protein gid [<i>Bacillus subtilis</i>]	UniRef100_P39815	<i>Bacillus subtilis</i>	Gid
1735	Tyrosine recombinase xerC [<i>Bacillus subtilis</i>]	UniRef100_P39776	<i>Bacillus subtilis</i>	CodV

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1736	ATP-dependent protease hslV precursor [<i>Bacillus subtilis</i>]	UniRef100_P39070	<i>Bacillus subtilis</i>	ClpQ
1737	ATP-dependent hsl protease ATP-binding subunit hslU [<i>Bacillus subtilis</i>]	UniRef100_P39778	<i>Bacillus subtilis</i>	ClpY
1738	GTP-sensing transcriptional pleiotropic repressor codY [<i>Bacillus subtilis</i>]	UniRef100_P39779	<i>Bacillus subtilis</i>	CodY
1739	Flagellar basal-body rod protein flgB [<i>Bacillus subtilis</i>]	UniRef100_P24500	<i>Bacillus subtilis</i>	FlgB
1740	Flagellar basal-body rod protein flgC [<i>Bacillus subtilis</i>]	UniRef100_P24501	<i>Bacillus subtilis</i>	FlgC
1741	Flagellar hook-basal body complex protein flhE [<i>Bacillus subtilis</i>]	UniRef100_P24502	<i>Bacillus subtilis</i>	FliE
1742	Flagellar M-ring protein [<i>Bacillus subtilis</i>]	UniRef100_P23447	<i>Bacillus subtilis</i>	FliF
1743	Flagellar motor switch protein flhG [<i>Bacillus subtilis</i>]	UniRef100_P23448	<i>Bacillus subtilis</i>	FliG
1744	Probable flagellar assembly protein flhH [<i>Bacillus subtilis</i>]	UniRef100_P23449	<i>Bacillus subtilis</i>	FliH
1745	Flagellum-specific ATP synthase [<i>Bacillus subtilis</i>]	UniRef100_P23445	<i>Bacillus subtilis</i>	FliI
1746	Flagellar flhJ protein [<i>Bacillus subtilis</i>]	UniRef100_P20487	<i>Bacillus subtilis</i>	FliJ
1747	FlaA locus 22.9 kDa protein [<i>Bacillus subtilis</i>]	UniRef100_P23454	<i>Bacillus subtilis</i>	YlxF
1748	Probable flagellar hook-length control protein [<i>Bacillus subtilis</i>]	UniRef100_P23451	<i>Bacillus subtilis</i>	FliK
1749	FlaA locus hypothetical protein ylxG [<i>Bacillus subtilis</i>]	UniRef100_P23455	<i>Bacillus subtilis</i>	YlxG
1750				FlgE
1751	BH2448 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KA42	<i>Bacillus halodurans</i>	
1752	Flagellar flhL protein [<i>Bacillus subtilis</i>]	UniRef100_P23452	<i>Bacillus subtilis</i>	FliL
1753	Flagellar motor switch protein flhM [<i>Bacillus subtilis</i>]	UniRef100_P23453	<i>Bacillus subtilis</i>	FliM
1754	Flagellar motor switch protein flhY [<i>Bacillus subtilis</i>]	UniRef100_P24073	<i>Bacillus subtilis</i>	FliY
1755	Chemotaxis protein cheY homolog [<i>Bacillus subtilis</i>]	UniRef100_P24072	<i>Bacillus subtilis</i>	CheY
1756	Flagellar biosynthetic protein flhZ precursor [<i>Bacillus subtilis</i>]	UniRef100_P35536	<i>Bacillus subtilis</i>	FliZ
1757	Flagellar biosynthetic protein flhP [<i>Bacillus subtilis</i>]	UniRef100_P35528	<i>Bacillus subtilis</i>	FliP
1758	Flagellar biosynthetic protein flhQ [<i>Bacillus subtilis</i>]	UniRef100_P35535	<i>Bacillus subtilis</i>	
1759	Flagellar biosynthetic protein flhR [<i>Bacillus subtilis</i>]	UniRef100_P35537	<i>Bacillus subtilis</i>	FliR
1760	Flagellar biosynthetic protein flhB [<i>Bacillus subtilis</i>]	UniRef100_P35538	<i>Bacillus subtilis</i>	FliB
1761	Flagellar biosynthesis protein flhA [<i>Bacillus subtilis</i>]	UniRef100_P35620	<i>Bacillus subtilis</i>	FliA
1762	Flagellar biosynthesis protein flhF [<i>Bacillus subtilis</i>]	UniRef100_Q01960	<i>Bacillus subtilis</i>	FliF
1763	Hypothetical protein ylxH [<i>Bacillus subtilis</i>]	UniRef100_P40742	<i>Bacillus subtilis</i>	YlxH
1764	Chemotaxis response regulator protein-glutamate methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_Q05522	<i>Bacillus subtilis</i>	CheB
1765	Chemotaxis protein cheA [<i>Bacillus subtilis</i>]	UniRef100_P29072	<i>Bacillus subtilis</i>	CheA
1766	Chemotaxis protein cheW [<i>Bacillus subtilis</i>]	UniRef100_P39802	<i>Bacillus subtilis</i>	CheW
1767	Chemotaxis protein cheC [<i>Bacillus subtilis</i>]	UniRef100_P40403	<i>Bacillus subtilis</i>	CheC
1768	Chemotaxis protein cheD [<i>Bacillus subtilis</i>]	UniRef100_P40404	<i>Bacillus subtilis</i>	CheD
1769	RNA polymerase sigma-D factor [<i>Bacillus subtilis</i>]	UniRef100_P10726	<i>Bacillus subtilis</i>	SigD
1770	Swarming motility protein swrB [<i>Bacillus subtilis</i>]	UniRef100_P40405	<i>Bacillus subtilis</i>	YlxL

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1771	30S ribosomal protein S2 [<i>Bacillus subtilis</i>]	UniRef100_P21464	<i>Bacillus subtilis</i>	RpsB
1772	Translation elongation factor Ts [<i>Bacillus cereus</i>]	UniRef100_Q65JJ8	<i>Bacillus cereus</i>	Tsf
1773				PyrH
1774				Frr
1775	Undecaprenyl pyrophosphate synthetase [<i>Bacillus subtilis</i>]	UniRef100_O31751	<i>Bacillus subtilis</i>	UppS
1776	Phosphatidate cytidyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O31752	<i>Bacillus subtilis</i>	CdsA
1777	1-deoxy-D-xylulose 5-phosphate reductoisomerase [<i>Bacillus subtilis</i>]	UniRef100_O31753	<i>Bacillus subtilis</i>	Dxr
1778	Hypothetical zinc metalloprotease yIuc [<i>Bacillus subtilis</i>]	UniRef100_O31754	<i>Bacillus subtilis</i>	YIuC
1779	Prolyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_O31755	<i>Bacillus subtilis</i>	ProS
1780	DNA polymerase III polC-type [<i>Bacillus subtilis</i>]	UniRef100_P13267	<i>Bacillus subtilis</i>	PolC
1781				
1782	Cellulose 1,4-beta-cellobiosidase precursor [<i>Paenibacillus</i> sp. BP-23]	UniRef100_Q8KKF7	<i>Paenibacillus</i> sp. BP-23	
1783	Endoglucanase B precursor [<i>Paenibacillus lautus</i>]	UniRef100_P23550	<i>Paenibacillus lautus</i>	
1784	Beta-mannosidase [<i>Thermotoga neapolitana</i>]	UniRef100_Q9RIK7	<i>Thermotoga neapolitana</i>	
1785	Hypothetical UPF0090 protein yIxs [<i>Bacillus subtilis</i>]	UniRef100_P32726	<i>Bacillus subtilis</i>	YIxs
1786	Transcription elongation protein nusA [<i>Bacillus subtilis</i>]	UniRef100_P32727	<i>Bacillus subtilis</i>	NusA
1787	Hypothetical protein yIxr [<i>Bacillus subtilis</i>]	UniRef100_P32728	<i>Bacillus subtilis</i>	
1788	Probable ribosomal protein yIxq [<i>Bacillus subtilis</i>]	UniRef100_P32729	<i>Bacillus subtilis</i>	
1789	Translation initiation factor IF-2 [<i>Bacillus subtilis</i>]	UniRef100_P17889	<i>Bacillus subtilis</i>	InfB
1790	Hypothetical protein yIxP [<i>Bacillus subtilis</i>]	UniRef100_P32730	<i>Bacillus subtilis</i>	
1791	Ribosome-binding factor A [<i>Bacillus subtilis</i>]	UniRef100_P32731	<i>Bacillus subtilis</i>	RbfA
1792				TruB
1793	Riboflavin biosynthesis protein ribC [Includes: Riboflavin kinase (EC 2.7.1.26) (Flavokinase); FMN adenylyltransferase (EC 2.7.7.2) (FAD pyrophosphorylase) (FAD synthetase)] [<i>Bacillus subtilis</i>]	UniRef100_P54575	Includes: Riboflavin kinase (EC 2.7.1.26) (Flavokinase); FMN adenylyltransferase (EC 2.7.7.2) (FAD pyrophosphorylase) (FAD synthetase)	RibC
1794	30S ribosomal protein S15 [<i>Bacillus subtilis</i>]	UniRef100_P21473	<i>Bacillus subtilis</i>	
1795	Polyribonucleotide nucleotidyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P50849	<i>Bacillus subtilis</i>	PnpA
1796	Hypothetical protein yIxY precursor [<i>Bacillus subtilis</i>]	UniRef100_P50850	<i>Bacillus subtilis</i>	YIxY
1797	Hypothetical zinc protease ymxG [<i>Bacillus subtilis</i>]	UniRef100_Q04805	<i>Bacillus subtilis</i>	MlpA
1798	Hypothetical protein ymxH [<i>Bacillus subtilis</i>]	UniRef100_Q04811	<i>Bacillus subtilis</i>	
1799	Dipicolinate synthase, A chain [<i>Bacillus subtilis</i>]	UniRef100_Q04809	<i>Bacillus subtilis</i>	SpoVFA
1800	Dipicolinate synthase, B chain [<i>Bacillus subtilis</i>]	UniRef100_Q04810	<i>Bacillus subtilis</i>	SpoVFB
1801	Aspartate-semialdehyde dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_Q04797	<i>Bacillus subtilis</i>	Asd

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1802	Aspartokinase 1 (EC 2.7.2.4) (Aspartokinase I) (Aspartate kinase 1) [Contains: Aspartokinase I alpha subunit; Aspartokinase I beta subunit] [<i>Bacillus subtilis</i>]	UniRef100_Q04795	Contains: Aspartokinase I alpha subunit; Aspartokinase I beta subunit	DapG
1803	Dihydrodipicolinate synthase [<i>Bacillus subtilis</i>]	UniRef100_Q04796	<i>Bacillus subtilis</i>	DapA
1804	YmfA protein [<i>Bacillus subtilis</i>]	UniRef100_O31760	<i>Bacillus subtilis</i>	YmfA
1805	Translocation-enhancing protein tepA [<i>Bacillus subtilis</i>]	UniRef100_Q99171	<i>Bacillus subtilis</i>	TepA
1806				
1807	DNA translocase ftsK [<i>Bacillus subtilis</i>]	UniRef100_P21458	<i>Bacillus subtilis</i>	SpoIIIE
1808	Hypothetical transcriptional regulator ymfC [<i>Bacillus subtilis</i>]	UniRef100_O31761	<i>Bacillus subtilis</i>	YmfC
1809	Multidrug resistance protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K7Q2	<i>Bacillus halodurans</i>	YitG
1810	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81WP6	<i>Bacillus anthracis</i>	YmfG
1811	YmfH protein [<i>Bacillus subtilis</i>]	UniRef100_O31766	<i>Bacillus subtilis</i>	YmfH
1812				
1813				
1814	YmfJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31768	<i>Bacillus subtilis</i>	
1815	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q636P2	<i>Bacillus cereus</i> ZK	YmfK
1816				
1817	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P46322	<i>Bacillus subtilis</i>	YmfM PgsA
1818	CinA-like protein [<i>Bacillus subtilis</i>]	UniRef100_P46323	<i>Bacillus subtilis</i>	CinA
1819	RecA protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q8GJG2	<i>Bacillus amyloliquefaciens</i>	RecA
1820	Hypothetical UPF0144 protein ymdA [<i>Bacillus subtilis</i>]	UniRef100_O31774	<i>Bacillus subtilis</i>	YmdA
1821	YmdB protein [<i>Bacillus subtilis</i>]	UniRef100_O31775	<i>Bacillus subtilis</i>	YmdB
1822	Stage V sporulation protein S [<i>Bacillus subtilis</i>]	UniRef100_P45693	<i>Bacillus subtilis</i>	
1823				
1824				
1825	L-threonine 3-dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O31776	<i>Bacillus subtilis</i>	Tdh
1826	2-amino-3-ketobutyrate coenzyme A ligase [<i>Bacillus subtilis</i>]	UniRef100_O31777	<i>Bacillus subtilis</i>	Kbl
1827	Hypothetical UPF0004 protein ymcB [<i>Bacillus subtilis</i>]	UniRef100_O31778	<i>Bacillus subtilis</i>	YmcB
1828	YmcA protein [<i>Bacillus subtilis</i>]	UniRef100_O31779	<i>Bacillus subtilis</i>	YmcA
1829	Spore coat protein E [<i>Bacillus subtilis</i>]	UniRef100_P14016	<i>Bacillus subtilis</i>	CotE
1830				MutS
1831	DNA mismatch repair protein mutL [<i>Bacillus subtilis</i>]	UniRef100_P49850	<i>Bacillus subtilis</i>	MutL
1832	YjcS protein [<i>Bacillus subtilis</i>]	UniRef100_O31641	<i>Bacillus subtilis</i>	
1833				YxiD
1834				
1835				
1836	AII1751 protein [<i>Anabaena</i> sp.]	UniRef100_Q8YW65	<i>Anabaena</i> sp.	YciC
1837	BH0367 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFP4	<i>Bacillus halodurans</i>	
1838	Phosphinothricin N-acetyltransferase [<i>Bacillus halodurans</i>]	UniRef100_Q9KFP5	<i>Bacillus halodurans</i>	YwnH
1839	UPI00003CC0D8 UniRef100 entry	UniRef100_UPI00003CC0D8		
1840				
1841	Putative L-amino acid oxidase [<i>Bacillus subtilis</i>]	UniRef100_O34363	<i>Bacillus subtilis</i>	YobN
1842				YoaK
1843	Na ⁺ /H ⁺ antiporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K5Q0	<i>Bacillus halodurans</i>	YvgP
1844	Aromatic hydrocarbon catabolism protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CV32	<i>Oceanobacillus iheyensis</i>	
1845	Hypothetical UPF0145 protein VP1283 [<i>Vibrio parahaemolyticus</i>]	UniRef100_Q87Q67	<i>Vibrio parahaemolyticus</i>	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1846	Hypothetical protein yqeD [<i>Bacillus subtilis</i>]	UniRef100_P54449	<i>Bacillus subtilis</i>	YqeD
1847	Penicillin-binding protein, putative [<i>Bacillus cereus</i>]	UniRef100_Q738U9	<i>Bacillus cereus</i>	PbpE
1848	Hypothetical glycosyl transferase [<i>Bacillus subtilis</i>]	UniRef100_O34539	<i>Bacillus subtilis</i>	YjiC
1849	Asparate-proton symporter [<i>Bacillus subtilis</i>]	UniRef100_O07002	<i>Bacillus subtilis</i>	YveA
1850	Spore coat protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KEV6	<i>Bacillus halodurans</i>	
1851				YdhD
1852				YezE
1853	Putative HTH-type transcriptional regulator yezE [<i>Bacillus subtilis</i>]	UniRef100_Q7WY76	<i>Bacillus subtilis</i>	
1854	Hypothetical protein yesE [<i>Bacillus subtilis</i>]	UniRef100_O31511	<i>Bacillus subtilis</i>	YesE
1855	YesF protein [<i>Bacillus subtilis</i>]	UniRef100_O31512	<i>Bacillus subtilis</i>	YesF
1856	UPI00003CBA3B UniRef100 entry	UniRef100_UPI00003CBA3B		
1857	YmaD protein [<i>Bacillus subtilis</i>]	UniRef100_O31790	<i>Bacillus subtilis</i>	YmaD
1858	Multidrug resistance protein ebrB [<i>Bacillus subtilis</i>]	UniRef100_O31791	<i>Bacillus subtilis</i>	EbrB
1859	Multidrug resistance protein ebrA [<i>Bacillus subtilis</i>]	UniRef100_O31792	<i>Bacillus subtilis</i>	EbrA
1860				
1861	Hypothetical protein ymaF [<i>Bacillus subtilis</i>]	UniRef100_O31794	<i>Bacillus subtilis</i>	YmaF
1862	tRNA delta(2)-isopentenylpyrophosphate transferase [<i>Bacillus subtilis</i>]	UniRef100_O31795	<i>Bacillus subtilis</i>	MiaA
1863	Hfq protein [<i>Bacillus subtilis</i>]	UniRef100_O31796	<i>Bacillus subtilis</i>	
1864	Hypothetical protein ymzA [<i>Bacillus subtilis</i>]	UniRef100_O31798	<i>Bacillus subtilis</i>	
1865	NrdI protein [<i>Bacillus subtilis</i>]	UniRef100_P50618	<i>Bacillus subtilis</i>	YmaA
1866	Ribonucleoside-diphosphate reductase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P50620	<i>Bacillus subtilis</i>	NrdE
1867	Ribonucleoside-diphosphate reductase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P50621	<i>Bacillus subtilis</i>	NrdF
1868	Hypothetical protein ymaB [<i>Bacillus subtilis</i>]	UniRef100_P50619	<i>Bacillus subtilis</i>	YmaB
1869	Blr6966 protein [<i>Bradyrhizobium japonicum</i>]	UniRef100_Q89EV4	<i>Bradyrhizobium japonicum</i>	YtmP
1870	Nitrogen fixation protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KfV2	<i>Bacillus halodurans</i>	YurV
1871	Transcription regulator Fur family-like protein [<i>Staphylococcus epidermidis</i>]	UniRef100_Q8CNQ7	<i>Staphylococcus epidermidis</i>	PerR
1872	UPI00003CB681 UniRef100 entry	UniRef100_UPI00003CB681		YdhC
1873	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81C60	<i>Bacillus cereus</i>	YjlA
1874				CwIC
1875	Membrane protein, putative [<i>Listeria monocytogenes</i>]	UniRef100_Q720L9	<i>Listeria monocytogenes</i>	
1876	Lin1174 protein [<i>Listeria innocua</i>]	UniRef100_Q92CJ7	<i>Listeria innocua</i>	
1877				
1878				
1879	Transcriptional regulator [<i>Aquifex aeolicus</i>]	UniRef100_O66635	<i>Aquifex aeolicus</i>	YdgC
1880	Hypothetical membrane-anchored protein [<i>Rhizobium meliloti</i>]	UniRef100_Q92VA1	<i>Rhizobium meliloti</i>	
1881	Cytosine permease [<i>Bacillus halodurans</i>]	UniRef100_Q9KBP3	<i>Bacillus halodurans</i>	YxLA
1882	AgaF [<i>Agrobacterium tumefaciens</i>]	UniRef100_O50265	<i>Agrobacterium tumefaciens</i>	
1883				
1884	Hydantoin utilization protein B [<i>Pseudomonas putida</i>]	UniRef100_Q88H51	<i>Pseudomonas putida</i>	
1885	Hypothetical protein SMb20139 [<i>Rhizobium meliloti</i>]	UniRef100_Q92X23	<i>Rhizobium meliloti</i>	
1886	BH2340 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KAE7	<i>Bacillus halodurans</i>	
1887				
1888				

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1889	Stage V sporulation protein K [<i>Bacillus subtilis</i>]	UniRef100_P27643	<i>Bacillus subtilis</i>	SpoVK
1890	YnbA [<i>Bacillus subtilis</i>]	UniRef100_P94478	<i>Bacillus subtilis</i>	YnbA
1891	YnbB [<i>Bacillus subtilis</i>]	UniRef100_P94479	<i>Bacillus subtilis</i>	YnbB
1892	HTH-type transcriptional regulator glnR [<i>Bacillus subtilis</i>]	UniRef100_P37582	<i>Bacillus subtilis</i>	GlnR
1893	Glutamine synthetase [<i>Bacillus subtilis</i>]	UniRef100_P12425	<i>Bacillus subtilis</i>	GlnA
1894				
1895	Hypothetical protein CAC3435 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97DN7	<i>Clostridium acetobutylicum</i>	
1896				
1897				
1898	Hypothetical protein CAC0350 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97M50	<i>Clostridium acetobutylicum</i>	
1899	Hypothetical HIT-like protein MJ0866 [<i>Methanococcus jannaschii</i>]	UniRef100_Q58276	<i>Methanococcus jannaschii</i>	Hit
1900	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q612B3	<i>Bacillus anthracis</i>	
1901	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81CM1	<i>Bacillus cereus</i>	YdjC
1902				
1903	Methyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q81CJ2	<i>Bacillus cereus</i>	
1904				YoaO
1905	Acetyltransferase, GNAT family [<i>Bacillus cereus</i>]	UniRef100_Q737B4	<i>Bacillus cereus</i>	
1906	Repressor rok [<i>Bacillus subtilis</i>]	UniRef100_O34857	<i>Bacillus subtilis</i>	Rok
1907				
1908	Hypothetical protein [Bacteriophage T5]	UniRef100_Q6QGK2	Bacteriophage T5	
1909	Hypothetical protein yvdT [<i>Bacillus subtilis</i>]	UniRef100_O07001	<i>Bacillus subtilis</i>	YvdT
1910	Hypothetical protein yvdS [<i>Bacillus subtilis</i>]	UniRef100_O32262	<i>Bacillus subtilis</i>	YvdS
1911	Hypothetical protein yvdR [<i>Bacillus subtilis</i>]	UniRef100_O06999	<i>Bacillus subtilis</i>	YvdR
1912	Spermidine N1-acetyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q72Y03	<i>Bacillus cereus</i>	YoaA
1913	Hypothetical Membrane Associated Protein [<i>Bacillus cereus</i>]	UniRef100_Q812L6	<i>Bacillus cereus</i>	
1914	Hypothetical protein yoaW precursor [<i>Bacillus subtilis</i>]	UniRef100_O34541	<i>Bacillus subtilis</i>	YoaW
1915	Thiol-disulfide oxidoreductase resA [<i>Bacillus halodurans</i>]	UniRef100_Q9KCJ4	<i>Bacillus halodurans</i>	ResA
1916	Manganese-containing catalase [<i>Bacillus halodurans</i>]	UniRef100_Q9KAU6	<i>Bacillus halodurans</i>	YdbD
1917	BH1562 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KCK9	<i>Bacillus halodurans</i>	
1918	Acetyltransferase, GNAT family [<i>Bacillus cereus</i>]	UniRef100_Q739K0	<i>Bacillus cereus</i>	YjcK
1919	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELR7	<i>Oceanobacillus iheyensis</i>	
1920				
1921	Transcriptional regulator, MarR family [<i>Bacillus cereus</i>]	UniRef100_Q81BM5	<i>Bacillus cereus</i>	YkvE
1922	Putative NAD(P)H nitroreductase ydfN [<i>Bacillus subtilis</i>]	UniRef100_P96692	<i>Bacillus subtilis</i>	YdfN
1923	YdfO protein [<i>Bacillus subtilis</i>]	UniRef100_P96693	<i>Bacillus subtilis</i>	YdfO
1924	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q630S7	<i>Bacillus cereus</i> ZK	YwrF
1925	BH1010 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KE48	<i>Bacillus halodurans</i>	YhcB
1926	ORF28 [<i>Staphylococcus</i> phage K]	UniRef100_Q6Y7T8	<i>Staphylococcus</i> phage K	
1927				
1928	Thymidylate synthase [Bacteriophage phi-3T]	UniRef100_P07606	Bacteriophage phi-3T	ThyA
1929	Sporulation-specific extracellular nuclease precursor [<i>Bacillus subtilis</i>]	UniRef100_P42983	<i>Bacillus subtilis</i>	NucB
1930	LexA repressor [<i>Bacillus subtilis</i>]	UniRef100_P31080	<i>Bacillus subtilis</i>	LexA
1931	YneA [<i>Bacillus subtilis</i>]	UniRef100_Q45056	<i>Bacillus subtilis</i>	YneA
1932	YneB [<i>Bacillus subtilis</i>]	UniRef100_Q45057	<i>Bacillus subtilis</i>	YneB
1933	Hypothetical UPF0291 protein ynzC [<i>Bacillus subtilis</i>]	UniRef100_O31818	<i>Bacillus subtilis</i>	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
1934	Transketolase [<i>Bacillus subtilis</i>]	UniRef100_P45694	<i>Bacillus subtilis</i>	Tkt
1935	Hypothetical protein yneE [<i>Bacillus subtilis</i>]	UniRef100_P45707	<i>Bacillus subtilis</i>	YneE
1936	Hypothetical UPF0154 protein yneF [<i>Bacillus subtilis</i>]	UniRef100_P45708	<i>Bacillus subtilis</i>	
1937				
1938	Cytochrome c-type biogenesis protein ccdA [<i>Bacillus subtilis</i>]	UniRef100_P45706	<i>Bacillus subtilis</i>	CcdA
1939	CcdB protein [<i>Bacillus subtilis</i>]	UniRef100_P45709	<i>Bacillus subtilis</i>	YneI
1940	CcdC protein [<i>Bacillus subtilis</i>]	UniRef100_P45710	<i>Bacillus subtilis</i>	YneJ
1941	Hypothetical protein yneK [<i>Bacillus subtilis</i>]	UniRef100_P45711	<i>Bacillus subtilis</i>	YneK
1942	Spore coat protein M [<i>Bacillus subtilis</i>]	UniRef100_Q45058	<i>Bacillus subtilis</i>	CotM
1943				
1944				
1945				CitB
1946	YneN protein [<i>Bacillus subtilis</i>]	UniRef100_O31820	<i>Bacillus subtilis</i>	YneN
1947				
1948				
1949	Small, acid-soluble spore protein tlp [<i>Bacillus subtilis</i>]	UniRef100_Q45060	<i>Bacillus subtilis</i>	
1950	YneP [<i>Bacillus subtilis</i>]	UniRef100_Q45061	<i>Bacillus subtilis</i>	YneP
1951	YneQ [<i>Bacillus subtilis</i>]	UniRef100_Q45062	<i>Bacillus subtilis</i>	
1952	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q815P1	<i>Bacillus cereus</i>	
1953	Conserved domain protein [<i>Bacillus cereus</i>]	UniRef100_Q72YR7	<i>Bacillus cereus</i>	
1954	YneR [<i>Bacillus subtilis</i>]	UniRef100_Q45063	<i>Bacillus subtilis</i>	
1955	Hypothetical UPF0078 protein yneS [<i>Bacillus subtilis</i>]	UniRef100_Q45064	<i>Bacillus subtilis</i>	YneS
1956	YneT [<i>Bacillus subtilis</i>]	UniRef100_Q45065	<i>Bacillus subtilis</i>	YneT
1957	Topoisomerase IV subunit B [<i>Bacillus subtilis</i>]	UniRef100_Q59192	<i>Bacillus subtilis</i>	ParE
1958	Topoisomerase IV subunit A [<i>Bacillus subtilis</i>]	UniRef100_Q45066	<i>Bacillus subtilis</i>	ParC
1959				AraR
1960	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMP2	<i>Oceanobacillus iheyensis</i>	XylB
1961	L-ribulose-5-phosphate 4-epimerase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMP3	<i>Oceanobacillus iheyensis</i>	AraD
1962	L-arabinose isomerase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMP4	<i>Oceanobacillus iheyensis</i>	AraA
1963				YwtG
1964				FabG
1965	Hypothetical protein ynfC [<i>Bacillus subtilis</i>]	UniRef100_Q45067	<i>Bacillus subtilis</i>	YnfC
1966	Amino acid carrier protein alsT [<i>Bacillus subtilis</i>]	UniRef100_Q45068	<i>Bacillus subtilis</i>	AlsT
1967				NarI
1968				NarJ
1969				NarH
1970				NarG
1971				
1972	Hypothetical protein yqfB [<i>Lactococcus lactis</i>]	UniRef100_Q9CF70	<i>Lactococcus lactis</i>	
1973				AlbA
1974				ArfM
1975				YwiC
1976	Transcriptional regulator of anaerobic genes [<i>Bacillus halodurans</i>]	UniRef100_Q9KG81	<i>Bacillus halodurans</i>	Fnr
1977	Nitrite extrusion protein [<i>Bacillus subtilis</i>]	UniRef100_P46907	<i>Bacillus subtilis</i>	NarK
1978				
1979	CAMP-binding domains-Catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8R5P4	<i>Thermoanaerobacter tengcongensis</i>	Fnr

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
1980	Putative nitric oxide reductase [<i>Staphylococcus aureus</i>]	UniRef100_Q6GK48	<i>Staphylococcus aureus</i>	
1981				YngL
1982				BglC
1983				
1984	Hypothetical protein ynfE [<i>Bacillus subtilis</i>]	UniRef100_Q45069	<i>Bacillus subtilis</i>	
1985	Hypothetical protein [<i>Bacillus megaterium</i>]	UniRef100_Q9ZF48	<i>Bacillus megaterium</i>	YkkB
1986	Hypothetical protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70K06	<i>Bacillus amyloliquefaciens</i>	
1987	Alkyl hydroperoxide reductase large subunit [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9W3	<i>Bacillus halodurans</i>	AhpF
1988	Methyltransferase [<i>Methanosarcina mazei</i>]	UniRef100_Q8PU82	<i>Methanosarcina mazei</i>	
1989	Similar to <i>B. subtilis</i> ywgB gene [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9W2	<i>Bacillus halodurans</i>	YwgB
1990	Hypothetical protein ywoF [<i>Bacillus subtilis</i>]	UniRef100_P94576	<i>Bacillus subtilis</i>	YwoF
1991	Branched-chain amino acid transport system carrier protein brnQ [<i>Bacillus subtilis</i>]	UniRef100_P94499	<i>Bacillus subtilis</i>	BrnQ
1992	NADP-dependent alcohol dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O06007	<i>Bacillus subtilis</i>	AdhA
1993	Transcriptional regulator, MerR family [<i>Listeria monocytogenes</i>]	UniRef100_Q721Z3	<i>Listeria monocytogenes</i>	YraB
1994	HPr-like protein crh [<i>Bacillus subtilis</i>]	UniRef100_O06976	<i>Bacillus subtilis</i>	
1995	BH2089 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9R4	<i>Bacillus halodurans</i>	YddR
1996				
1997	Enoyl-CoA hydratase/isomerase family protein [<i>Bacillus cereus</i>]	UniRef100_Q738L0	<i>Bacillus cereus</i>	YngF
1998	Hypothetical protein ysiB [<i>Bacillus subtilis</i>]	UniRef100_P94549	<i>Bacillus subtilis</i>	YsiB
1999	Methylmalonic acid semialdehyde dehydrogenase [<i>Bacillus cereus</i> ZK]	UniRef100_Q63BL0	<i>Bacillus cereus</i> ZK	MmsA
2000	3-hydroxyisobutyrate dehydrogenase [<i>Bacillus cereus</i> ZK]	UniRef100_Q63BL1	<i>Bacillus cereus</i> ZK	YkwC
2001	Acyl-CoA dehydrogenase [<i>Bacillus cereus</i>]	UniRef100_Q81DR7	<i>Bacillus cereus</i>	YusJ
2002	Mannose-6-phosphate isomerase [<i>Bacillus subtilis</i>]	UniRef100_O31646	<i>Bacillus subtilis</i>	ManA
2003	Phosphotransferase system (PTS) mannose-specific enzyme IIBCA component [<i>Bacillus subtilis</i>]	UniRef100_O31645	<i>Bacillus subtilis</i>	ManP
2004				
2005				
2006	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q72YT6	<i>Bacillus cereus</i>	
2007	Transcriptional regulator [<i>Bacillus subtilis</i>]	UniRef100_O31644	<i>Bacillus subtilis</i>	ManR
2008				
2009	UPI00003CC220 UniRef100 entry	UniRef100_UPI00003CC220		YtrB
2010	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KF35	<i>Bacillus halodurans</i>	YtrA
2011	Probable oxidoreductase [<i>Clostridium perfringens</i>]	UniRef100_Q8XP17	<i>Clostridium perfringens</i>	YjmF
2012	Mannonate dehydratase [<i>Clostridium perfringens</i>]	UniRef100_Q8XP15	<i>Clostridium perfringens</i>	UxuA
2013	Glucosidase [<i>Bacillus halodurans</i>]	UniRef100_Q9KEZ5	<i>Bacillus halodurans</i>	
2014	C4-dicarboxylate transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KEZ6	<i>Bacillus halodurans</i>	
2015	C4-dicarboxylate transport system permease small protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMM5	<i>Oceanobacillus iheyensis</i>	
2016	C4-dicarboxylate transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KEZ8	<i>Bacillus halodurans</i>	DetB
2017	Transcriptional regulator [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EL22	<i>Oceanobacillus iheyensis</i>	CepA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2018	Arsenate reductase [<i>Bacillus subtilis</i>]	UniRef100_P45947	<i>Bacillus subtilis</i>	ArsC
2019	YdfA protein [<i>Bacillus subtilis</i>]	UniRef100_P96678	<i>Bacillus subtilis</i>	YdfA
2020	YdeT protein [<i>Bacillus subtilis</i>]	UniRef100_P96677	<i>Bacillus subtilis</i>	YdeT
2021	YdeI [<i>Bacillus halodurans</i>]	UniRef100_Q9Z9R5	<i>Bacillus halodurans</i>	YdeI
2022	Putative secreted protein [<i>Streptomyces coelicolor</i>]	UniRef100_Q9S1Z5	<i>Streptomyces coelicolor</i>	
2023	Probable glucose uptake protein glcU [<i>Bacillus megaterium</i>]	UniRef100_P40419	<i>Bacillus megaterium</i>	GlcU
2024	YngK protein [<i>Bacillus subtilis</i>]	UniRef100_O35015	<i>Bacillus subtilis</i>	YngK
2025	YngD protein [<i>Bacillus subtilis</i>]	UniRef100_O31824	<i>Bacillus subtilis</i>	YngD
2026	Pyruvate formate-lyase-activating enzyme [<i>Bacillus cereus</i>]	UniRef100_Q73DZ6	<i>Bacillus cereus</i>	YkvL
2027	Formate acetyltransferase [<i>Bacillus anthracis</i>]	UniRef100_Q81YX1	<i>Bacillus anthracis</i>	
2028				DacC
2029	NADH dehydrogenase-like protein yjID [<i>Bacillus subtilis</i>]	UniRef100_P80861	<i>Bacillus subtilis</i>	YjID
2030	Hypothetical protein yjIC [<i>Bacillus subtilis</i>]	UniRef100_O34633	<i>Bacillus subtilis</i>	YjIC
2031				
2032	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81IJ8	<i>Bacillus cereus</i>	
2033	Hypothetical protein ykzH [<i>Bacillus subtilis</i>]	UniRef100_O31653	<i>Bacillus subtilis</i>	
2034	Acetyl-CoA synthetase [<i>Bacillus halodurans</i>]	UniRef100_Q9KDS4	<i>Bacillus halodurans</i>	AcsA
2035	YngE protein [<i>Bacillus subtilis</i>]	UniRef100_O31825	<i>Bacillus subtilis</i>	YngE
2036	Hydroxybutyryl-dehydratase [<i>Bacillus subtilis</i>]	UniRef100_Q9L7W1	<i>Bacillus subtilis</i>	YngF
2037	YngG protein [<i>Bacillus subtilis</i>]	UniRef100_O34873	<i>Bacillus subtilis</i>	YngG
2038	YngXX [<i>Bacillus subtilis</i>]	UniRef100_Q9R9I3	<i>Bacillus subtilis</i>	
2039	YngH [<i>Bacillus subtilis</i>]	UniRef100_Q9R9I4	<i>Bacillus subtilis</i>	YngH
2040	YngI [<i>Bacillus subtilis</i>]	UniRef100_Q9R9I5	<i>Bacillus subtilis</i>	YngI
2041	YngJ protein [<i>Bacillus subtilis</i>]	UniRef100_O34421	<i>Bacillus subtilis</i>	YngJ
2042	NAD(P)H oxidoreductase YRKL [<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> ATCC 49256]	UniRef100_Q7P6P0	<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> ATCC 49256	YrkL
2043	Transcriptional regulator, MarR family [<i>Fusobacterium nucleatum</i>]	UniRef100_Q8RE85	<i>Fusobacterium nucleatum</i>	
2044	Glutamate-5-semialdehyde dehydrogenase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HHC2	<i>Bacillus thuringiensis</i>	ProA
2045	Glutamate 5-kinase 2 [<i>Bacillus subtilis</i>]	UniRef100_O07509	<i>Bacillus subtilis</i>	ProJ
2046	Pyrraline-5-carboxylate reductase 1 [<i>Bacillus subtilis</i>]	UniRef100_P14383	<i>Bacillus subtilis</i>	ProH
2047	UPI00003CB6CD UniRef100 entry	UniRef100_UPI00003CB6CD		YetF
2048	Sodium-dependent phosphate transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KCT1	<i>Bacillus halodurans</i>	CysP
2049	Probable phosphoadenosine phosphosulfate reductase [<i>Bacillus subtilis</i>]	UniRef100_O06737	<i>Bacillus subtilis</i>	YitB
2050	Phosphosulfolactate synthase (EC 4.4.1.19) ((2R)-phospho-3-sulfolactate synthase) [<i>Bacillus subtilis</i>]	UniRef100_O06739	<i>Bacillus subtilis</i>	YitD
2051	YitE [<i>Bacillus subtilis</i>]	UniRef100_O06740	<i>Bacillus subtilis</i>	YitE
2052	YitF [<i>Bacillus subtilis</i>]	UniRef100_O06741	<i>Bacillus subtilis</i>	YitF
2053	YitG [<i>Bacillus subtilis</i>]	UniRef100_O06742	<i>Bacillus subtilis</i>	YitG
2054	Putative glycosyl transferase ykoT [<i>Bacillus subtilis</i>]	UniRef100_O34755	<i>Bacillus subtilis</i>	YkoT
2055	YkoR [<i>Bacillus subtilis</i>]	UniRef100_O34830	<i>Bacillus subtilis</i>	YkoS
2056	Glutamate synthase [NADPH] small chain [<i>Bacillus subtilis</i>]	UniRef100_O34399	NADPH	GltB
2057				GltA
2058	HTH-type transcriptional regulator gltC [<i>Bacillus subtilis</i>]	UniRef100_P20668	<i>Bacillus subtilis</i>	GltC
2059	AII7121 protein [<i>Anabaena</i> sp.]	UniRef100_Q8YL17	<i>Anabaena</i> sp.	
2060				

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2061	Lmo0606 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y9C6	<i>Listeria monocytogenes</i>	
2062	ABC transporter ATP-binding protein [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67MU2	<i>Symbiobacterium thermophilum</i>	YfiB
2063	ABC transporter ATP-binding protein [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67MU3	<i>Symbiobacterium thermophilum</i>	YfiC
2064	Cytochrome P450 109 [<i>Bacillus subtilis</i>]	UniRef100_P27632	<i>Bacillus subtilis</i>	YjiB
2065	Hypothetical oxidoreductase yoxD [<i>Bacillus subtilis</i>]	UniRef100_P14802	<i>Bacillus subtilis</i>	YoxD
2066				Pps
2067	Permease, general substrate transporter [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HMC3	<i>Bacillus thuringiensis</i>	LmrB
2068	Putative formate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O34323	<i>Bacillus subtilis</i>	YoaE
2069	Transcriptional regulatory protein [<i>Bradyrhizobium japonicum</i>]	UniRef100_Q89KD1	<i>Bradyrhizobium japonicum</i>	YcgE
2070	Drug resistance transporter, EmrB/QacA family [<i>Bacillus cereus</i>]	UniRef100_Q73615	<i>Bacillus cereus</i>	Mdr
2071	YndE protein [<i>Bacillus subtilis</i>]	UniRef100_O31809	<i>Bacillus subtilis</i>	YndE
2072	YndF protein [<i>Bacillus subtilis</i>]	UniRef100_O31810	<i>Bacillus subtilis</i>	YndF
2073				YndE
2074				YndE
2075	UPI00003CB22E UniRef100 entry	UniRef100_UPI00003CB22E		YndD
2076				
2077	Nucleotide binding protein expZ [<i>Bacillus subtilis</i>]	UniRef100_P39115	<i>Bacillus subtilis</i>	ExpZ
2078				TlpB
2079	CsaA protein [<i>Bacillus subtilis</i>]	UniRef100_P37584	<i>Bacillus subtilis</i>	CsaA
2080	Alkaline serine protease [<i>Bacillus subtilis</i>]	UniRef100_O31788	<i>Bacillus subtilis</i>	AprX
2081	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81SD4	<i>Bacillus anthracis</i>	YnzE
2082	Transcriptional regulator, TetR family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63D70	<i>Bacillus cereus</i> ZK	Yrhl
2083	Putative HTH-type transcriptional regulator yvmB [<i>Bacillus subtilis</i>]	UniRef100_P40762	<i>Bacillus subtilis</i>	YvmB
2084	Hydrolase [<i>Bacillus cereus</i>]	UniRef100_Q81D79	<i>Bacillus cereus</i>	YqeK
2085	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETG3	<i>Oceanobacillus iheyensis</i>	
2086	YndJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31813	<i>Bacillus subtilis</i>	YndJ
2087	YndH protein [<i>Bacillus subtilis</i>]	UniRef100_O31812	<i>Bacillus subtilis</i>	YndH
2088	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q73A97	<i>Bacillus cereus</i>	YndG
2089	UPI00003CBA97 UniRef100 entry	UniRef100_UPI00003CBA97		YobS
2090	UPI00003CBA98 UniRef100 entry	UniRef100_UPI00003CBA98		YobT
2091	DNA-binding protein YobU [<i>Bacillus subtilis</i>]	UniRef100_O34637	<i>Bacillus subtilis</i>	YobU
2092	Hypothetical protein [<i>Bacteroides fragilis</i>]	UniRef100_Q64RP1	<i>Bacteroides fragilis</i>	
2093	Transcription regulator [<i>Bacillus subtilis</i>]	UniRef100_O34920	<i>Bacillus subtilis</i>	YobV
2094				YobW
2095	YozA protein [<i>Bacillus subtilis</i>]	UniRef100_O31844	<i>Bacillus subtilis</i>	
2096	Possible metallo-beta-lactamase family protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q638G1	<i>Bacillus cereus</i> ZK	YmaE
2097				
2098				YfnA
2099				YocA
2100	UPI00003CBE6E UniRef100 entry	UniRef100_UPI00003CBE6E		YveM
2101	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q817C2	<i>Bacillus cereus</i>	
2102				
2103	Glycosyl transferase [<i>Bacillus halodurans</i>]	UniRef100_Q9K7I1	<i>Bacillus halodurans</i>	YtcC
2104	YozB protein [<i>Bacillus subtilis</i>]	UniRef100_O31845	<i>Bacillus subtilis</i>	YozB
2105	Lmo2079 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y5I3	<i>Listeria monocytogenes</i>	
2106	YocC [<i>Bacillus subtilis</i>]	UniRef100_O35042	<i>Bacillus subtilis</i>	YocC
2107	Na ⁺ /myo-inositol cotransporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KAR5	<i>Bacillus halodurans</i>	YegO
2108	YocH [<i>Bacillus subtilis</i>]	UniRef100_O34669	<i>Bacillus subtilis</i>	YocH

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2109	RecQ homolog [<i>Bacillus subtilis</i>]	UniRef100_O34748	<i>Bacillus subtilis</i>	YocI
2110	Hypothetical protein yqbC [<i>Bacillus subtilis</i>]	UniRef100_P45919	<i>Bacillus subtilis</i>	YqbC
2111	Hypothetical protein yjfb [<i>Bacillus subtilis</i>]	UniRef100_O34438	<i>Bacillus subtilis</i>	
2112	Hypothetical protein yyaQ [<i>Bacillus subtilis</i>]	UniRef100_P37507	<i>Bacillus subtilis</i>	YyaQ
2113	Hypothetical protein OB2103 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPJ8	<i>Oceanobacillus iheyensis</i>	
2114				
2115	Hypothetical protein yjgD [<i>Bacillus subtilis</i>]	UniRef100_O34681	<i>Bacillus subtilis</i>	YjgD
2116				YjgC
2117				YjgC
2118	Hypothetical protein OB3361 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EL70	<i>Oceanobacillus iheyensis</i>	
2119	Hypothetical protein ypfA [<i>Bacillus subtilis</i>]	UniRef100_P38491	<i>Bacillus subtilis</i>	YpfA
2120				
2121				
2122				YxiB
2123				
2124				
2125				
2126				YndF
2127				
2128	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q816C2	<i>Bacillus cereus</i>	
2129	BH0185 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KGB9	<i>Bacillus halodurans</i>	
2130				
2131				YisY
2132	Hypothetical protein yoqH [Bacteriophage SPBc2]	UniRef100_O64117	Bacteriophage SPBc2	YoqH
2133				
2134	BH0429 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFP9	<i>Bacillus halodurans</i>	YrhP
2135	30S ribosomal protein S14 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETX0	<i>Oceanobacillus iheyensis</i>	
2136	UPI00002DEBB5 UniRef100 entry	UniRef100_UPI00002DEBB5		MutT
2137	Pyruvate water dikinase [<i>Methanosarcina acetivorans</i>]	UniRef100_Q8TN35	<i>Methanosarcina acetivorans</i>	Pps
2138	Transcriptional regulator [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESJ8	<i>Oceanobacillus iheyensis</i>	YxbF
2139				
2140	YndM protein [<i>Bacillus subtilis</i>]	UniRef100_O31816	<i>Bacillus subtilis</i>	YndM
2141	Hypothetical protein yisT [<i>Bacillus subtilis</i>]	UniRef100_O07939	<i>Bacillus subtilis</i>	YisT
2142	Putative acyl carrier protein phosphodiesterase 1 [<i>Bacillus subtilis</i>]	UniRef100_O35022	<i>Bacillus subtilis</i>	YocJ
2143	General stress protein 16O [<i>Bacillus subtilis</i>]	UniRef100_P80872	<i>Bacillus subtilis</i>	YocK
2144				
2145				
2146	Aldehyde dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O34660	<i>Bacillus subtilis</i>	DhaS
2147	YjbB protein [<i>Bacillus subtilis</i>]	UniRef100_O31600	<i>Bacillus subtilis</i>	YjbB
2148	Aminoglycoside N6'-acetyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q81AT3	<i>Bacillus cereus</i>	YjcK
2149	Blasticidin S deaminase, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81Y61	<i>Bacillus anthracis</i>	
2150				
2151	Probable superoxide dismutase [Fe] [<i>Bacillus subtilis</i>]	UniRef100_O35023	Fe	SqhC SodF
2152	Stress response protein yvgO precursor [<i>Bacillus subtilis</i>]	UniRef100_O32211	<i>Bacillus subtilis</i>	YvgO
2153	Putative transporter [<i>Bacillus subtilis</i>]	UniRef100_O34383	<i>Bacillus subtilis</i>	YocR
2154	Putative transporter [<i>Bacillus subtilis</i>]	UniRef100_O34524	<i>Bacillus subtilis</i>	YocS
2155	Dihydrolipoylysine-residue succinyltransferase component of 2-	UniRef100_P16263	<i>Bacillus subtilis</i>	OdhB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
	oxoglutarate dehydrogenase complex [<i>Bacillus subtilis</i>]			
2156	2-oxoglutarate dehydrogenase E1 component [<i>Bacillus subtilis</i>]	UniRef100_P23129	<i>Bacillus subtilis</i>	OdhA
2157	YojO protein [<i>Bacillus subtilis</i>]	UniRef100_O31849	<i>Bacillus subtilis</i>	YojO
2158	YojN protein [<i>Bacillus subtilis</i>]	UniRef100_O31850	<i>Bacillus subtilis</i>	YojN
2159	Hypothetical superoxide dismutase-like protein yojM precursor [<i>Bacillus subtilis</i>]	UniRef100_O31851	<i>Bacillus subtilis</i>	YojM
2160	Hypothetical protein yojL precursor [<i>Bacillus subtilis</i>]	UniRef100_O31852	<i>Bacillus subtilis</i>	LytF
2161	Probable multidrug resistance protein norM (Na(+)/drug antiporter) [<i>Bacillus subtilis</i>]	UniRef100_O31855	<i>Bacillus subtilis</i>	YojI
2162	Hypothetical protein [<i>Bacillus cereus ZK</i>]	UniRef100_Q637Z8	<i>Bacillus cereus ZK</i>	YojG
2163	YojF protein [<i>Bacillus subtilis</i>]	UniRef100_O31858	<i>Bacillus subtilis</i>	YojF
2164	YojE [<i>Bacillus subtilis</i>]	UniRef100_O68260	<i>Bacillus subtilis</i>	
2165	YojE protein [<i>Bacillus subtilis</i>]	UniRef100_O31859	<i>Bacillus subtilis</i>	YojE
2166	Hypothetical protein yozR [<i>Bacillus subtilis</i>]	UniRef100_Q7WY67	<i>Bacillus subtilis</i>	YozR
2167	YoaJ [<i>Bacillus subtilis</i>]	UniRef100_O34918	<i>Bacillus subtilis</i>	YoaJ
2168				
2169	UPI00003CBE9B UniRef100 entry	UniRef100_UPI00003CBE9B		YwfD
2170	Hypothetical UPF0087 protein yodB [<i>Bacillus subtilis</i>]	UniRef100_O34844	<i>Bacillus subtilis</i>	YodB
2171	Putative NAD(P)H nitroreductase 12C [<i>Bacillus subtilis</i>]	UniRef100_P81102	<i>Bacillus subtilis</i>	YodC
2172	YolF [<i>Bacillus subtilis</i>]	UniRef100_O34842	<i>Bacillus subtilis</i>	YodD
2173	UPI0000315ACC UniRef100 entry	UniRef100_UPI0000315ACC		
2174	YodF protein [<i>Bacillus subtilis</i>]	UniRef100_O34745	<i>Bacillus subtilis</i>	YodF
2175	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
2176	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
2177	OrfRM1 [<i>Bacillus subtilis</i>]	UniRef100_O34666	<i>Bacillus subtilis</i>	CtpA
2178	YolB [<i>Bacillus subtilis</i>]	UniRef100_O34954	<i>Bacillus subtilis</i>	YodH
2179				
2180	Carboxypeptidase [<i>Bacillus subtilis</i>]	UniRef100_O34866	<i>Bacillus subtilis</i>	YodJ
2181	Purine nucleoside phosphorylase II [<i>Bacillus subtilis</i>]	UniRef100_O34925	<i>Bacillus subtilis</i>	DeoD
2182	Hypothetical Membrane Spanning Protein [<i>Bacillus cereus</i>]	UniRef100_Q813P0	<i>Bacillus cereus</i>	YcgR
2183	YcgQ protein [<i>Bacillus subtilis</i>]	UniRef100_P94394	<i>Bacillus subtilis</i>	YcgQ
2184				
2185	Hypothetical protein yodL [<i>Bacillus subtilis</i>]	UniRef100_O30472	<i>Bacillus subtilis</i>	YodL
2186				YodM
2187	Hypothetical protein yozD [<i>Bacillus subtilis</i>]	UniRef100_O31863	<i>Bacillus subtilis</i>	
2188	Hypothetical protein yodN [<i>Bacillus subtilis</i>]	UniRef100_O34414	<i>Bacillus subtilis</i>	YodN
2189				
2190	YokU [<i>Bacillus subtilis</i>]	UniRef100_O30470	<i>Bacillus subtilis</i>	
2191	Hypothetical UPF0069 protein yodO [<i>Bacillus subtilis</i>]	UniRef100_O34676	<i>Bacillus subtilis</i>	KamA
2192	YodP [<i>Bacillus subtilis</i>]	UniRef100_O34895	<i>Bacillus subtilis</i>	YodP
2193	Acetylornithine deacetylase [<i>Bacillus subtilis</i>]	UniRef100_O34984	<i>Bacillus subtilis</i>	YodQ
2194	Butirate-acetoacetate CoA transferase [<i>Bacillus subtilis</i>]	UniRef100_O34466	<i>Bacillus subtilis</i>	YodR
2195	Butyrate acetoacetate-CoA transferase [<i>Bacillus subtilis</i>]	UniRef100_O34317	<i>Bacillus subtilis</i>	YodS
2196	Probable aminotransferase yodT [<i>Bacillus subtilis</i>]	UniRef100_O34662	<i>Bacillus subtilis</i>	YodT
2197	Multidrug resistance protein; possible tetracycline resistance determinant [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HK46	<i>Bacillus thuringiensis</i>	YkuC
2198	Protein cgeE [<i>Bacillus subtilis</i>]	UniRef100_P42093	<i>Bacillus subtilis</i>	CgeE

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2199	Peptide methionine sulfoxide reductase msrB [<i>Bacillus subtilis</i>]	UniRef100_P54155	<i>Bacillus subtilis</i>	YppQ
2200				MsrA
2201	Putative HTH-type transcriptional regulator ypoP [<i>Bacillus subtilis</i>]	UniRef100_P54182	<i>Bacillus subtilis</i>	YpoP
2202				
2203	Hypothetical protein yhcK [<i>Bacillus subtilis</i>]	UniRef100_P54595	<i>Bacillus subtilis</i>	YhcK
2204	Hypothetical protein ypnP [<i>Bacillus subtilis</i>]	UniRef100_P54181	<i>Bacillus subtilis</i>	YpnP
2205	Hypothetical conserved protein [Thermus thermophilus]	UniRef100_Q746K9	<i>Thermus thermophilus</i>	
2206				
2207	Hypothetical protein ypmS [<i>Bacillus subtilis</i>]	UniRef100_P54179	<i>Bacillus subtilis</i>	YpmS
2208	Hypothetical protein ypmR [<i>Bacillus subtilis</i>]	UniRef100_P40766	<i>Bacillus subtilis</i>	YpmR
2209	Hypothetical protein ypmQ [<i>Bacillus subtilis</i>]	UniRef100_P54178	<i>Bacillus subtilis</i>	YpmQ
2210	DegV family protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63BU6	<i>Bacillus cereus</i> ZK	YviA
2211	Hypothetical protein ypmP [<i>Bacillus subtilis</i>]	UniRef100_P54177	<i>Bacillus subtilis</i>	
2212	Threonine dehydratase biosynthetic [<i>Bacillus subtilis</i>]	UniRef100_P37946	<i>Bacillus subtilis</i>	IlvA
2213	Putative sigma L-dependent transcriptional regulator yplP [<i>Bacillus subtilis</i>]	UniRef100_P54156	<i>Bacillus subtilis</i>	YplP
2214	Hemolysin III homolog [<i>Bacillus subtilis</i>]	UniRef100_P54175	<i>Bacillus subtilis</i>	YplQ
2215	Hypothetical protein ypkP [<i>Bacillus subtilis</i>]	UniRef100_P54174	<i>Bacillus subtilis</i>	YpkP
2216	Dihydrofolate reductase [<i>Bacillus subtilis</i>]	UniRef100_P11045	<i>Bacillus subtilis</i>	DfrA
2217	Hypothetical protein ypjQ [<i>Bacillus subtilis</i>]	UniRef100_P54173	<i>Bacillus subtilis</i>	YpjQ
2218				YpjP
2219				YpiP
2220	Hypothetical protein yphP [<i>Bacillus subtilis</i>]	UniRef100_P54170	<i>Bacillus subtilis</i>	YphP
2221	Dihydroxy-acid dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P51785	<i>Bacillus subtilis</i>	IlvD
2222				YpgR
2223	Hypothetical protein ypgQ [<i>Bacillus subtilis</i>]	UniRef100_P54168	<i>Bacillus subtilis</i>	YpgQ
2224	Glutathione peroxidase homolog bsaA [<i>Bacillus subtilis</i>]	UniRef100_P52035	<i>Bacillus subtilis</i>	BsaA
2225	UPI00003CBA0F UniRef100 entry	UniRef100_UPI00003CBA0F		
2226	Homoserine O-succinyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P54167	<i>Bacillus subtilis</i>	MetA
2227	Putative glycosyl transferase ypfP [<i>Bacillus subtilis</i>]	UniRef100_P54166	<i>Bacillus subtilis</i>	UgtP
2228				
2229	Cold shock protein cspD [<i>Bacillus subtilis</i>]	UniRef100_P51777	<i>Bacillus subtilis</i>	
2230	Regulatory protein degR [<i>Bacillus subtilis</i>]	UniRef100_P06563	<i>Bacillus subtilis</i>	
2231	Hypothetical protein ypzA [<i>Bacillus subtilis</i>]	UniRef100_O32007	<i>Bacillus subtilis</i>	
2232				
2233	Hypothetical protein ypeP [<i>Bacillus subtilis</i>]	UniRef100_P54164	<i>Bacillus subtilis</i>	YpeP
2234	Hypothetical protein ypdP [<i>Bacillus subtilis</i>]	UniRef100_P54163	<i>Bacillus subtilis</i>	YpdP
2235	14.7 kDa ribonuclease H-like protein [<i>Bacillus subtilis</i>]	UniRef100_P54162	<i>Bacillus subtilis</i>	YpdQ
2236	Probable 5'-3' exonuclease [<i>Bacillus subtilis</i>]	UniRef100_P54161	<i>Bacillus subtilis</i>	YpcP
2237				

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2238	Hypothetical protein ypbS [<i>Bacillus subtilis</i>]	UniRef100_P54160	<i>Bacillus subtilis</i>	
2239	Hypothetical protein ypbR [<i>Bacillus subtilis</i>]	UniRef100_P54159	<i>Bacillus subtilis</i>	YpbR
2240				
2241	Hypothetical protein ypbQ [<i>Bacillus subtilis</i>]	UniRef100_P54158	<i>Bacillus subtilis</i>	YpbQ
2242				BcsA
2243	Predicted acetyltransferase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97G03	<i>Clostridium acetobutylicum</i>	YokL
2244	Xanthine permease [<i>Bacillus subtilis</i>]	UniRef100_P42086	<i>Bacillus subtilis</i>	PbuX
2245	Xanthine phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P42085	<i>Bacillus subtilis</i>	Xpt
2246	Hypothetical metalloprotease ypwA [<i>Bacillus subtilis</i>]	UniRef100_P50848	<i>Bacillus subtilis</i>	YpwA
2247				YpvA
2248	Hypothetical protein yptA precursor [<i>Bacillus subtilis</i>]	UniRef100_P50841	<i>Bacillus subtilis</i>	
2249				
2250	Hypothetical UPF0020 protein ypsC precursor [<i>Bacillus subtilis</i>]	UniRef100_P50840	<i>Bacillus subtilis</i>	YpsC
2251	Hypothetical protein ypsB [<i>Bacillus subtilis</i>]	UniRef100_P50839	<i>Bacillus subtilis</i>	
2252	Hypothetical protein ypsA [<i>Bacillus subtilis</i>]	UniRef100_P50838	<i>Bacillus subtilis</i>	YpsA
2253	Spore coat protein D [<i>Bacillus subtilis</i>]	UniRef100_P07791	<i>Bacillus subtilis</i>	
2254				
2255				YprB
2256				YprA
2257	Putative PTS system IIA component ypqE [<i>Bacillus subtilis</i>]	UniRef100_P50829	<i>Bacillus subtilis</i>	YpqE
2258	Hypothetical protein ypqA precursor [<i>Bacillus subtilis</i>]	UniRef100_P50836	<i>Bacillus subtilis</i>	YpqA
2259	Hypothetical protein yppG [<i>Bacillus subtilis</i>]	UniRef100_P50835	<i>Bacillus subtilis</i>	YppG
2260				
2261				YppE
2262				
2263				
2264	Hypothetical protein yppC [<i>Bacillus subtilis</i>]	UniRef100_P39791	<i>Bacillus subtilis</i>	YppC
2265	Recombination protein U [<i>Bacillus subtilis</i>]	UniRef100_P39792	<i>Bacillus subtilis</i>	RecU
2266	Penicillin-binding protein 1A/1B (PBP1) [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.—) (Peptidoglycan TGase); Penicillin-sensitive transpeptidase (EC 3.4.—.—) (DD-transpeptidase)] [<i>Bacillus subtilis</i>]	UniRef100_P39793	Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.—) (Peptidoglycan TGase); Penicillin-sensitive transpeptidase (EC 3.4.—.—) (DD-transpeptidase)	PonA
2267	Hypothetical protein ypoC [<i>Bacillus subtilis</i>]	UniRef100_P39789	<i>Bacillus subtilis</i>	YpoC
2268	Probable endonuclease III (EC 4.2.99.18) (DNA-(apurinic or apyrimidinic site) lyase) [<i>Bacillus subtilis</i>]	UniRef100_P39788	<i>Bacillus subtilis</i>	Nth
2269	DNA replication protein dnaD [<i>Bacillus subtilis</i>]	UniRef100_P39787	<i>Bacillus subtilis</i>	DnaD
2270	Asparaginyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P39772	<i>Bacillus subtilis</i>	AsnS
2271	Aspartate aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P53001	<i>Bacillus subtilis</i>	AspB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2272	Hypothetical protein ypmB [<i>Bacillus subtilis</i>]	UniRef100_P54396	<i>Bacillus subtilis</i>	YpmB
2273	Hypothetical protein ypmA [<i>Bacillus subtilis</i>]	UniRef100_P54395	<i>Bacillus subtilis</i>	
2274	Probable ATP-dependent helicase dinG homolog [<i>Bacillus subtilis</i>]	UniRef100_P54394	<i>Bacillus subtilis</i>	DinG
2275	Aspartate 1-decarboxylase precursor [<i>Bacillus subtilis</i>]	UniRef100_P52999	<i>Bacillus subtilis</i>	PanD
2276	Pantoate--beta-alanine ligase [<i>Bacillus subtilis</i>]	UniRef100_P52998	<i>Bacillus subtilis</i>	PanC
2277	3-methyl-2-oxobutanoate hydroxymethyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P52996	<i>Bacillus subtilis</i>	PanB
2278	BirA bifunctional protein [Includes: Biotin operon repressor; Biotin--[acetyl-CoA-carboxylase] synthetase (EC 6.3.4.15) (Biotin--protein ligase)] [<i>Bacillus subtilis</i>]	UniRef100_P42975	Includes: Biotin operon repressor; Biotin--	BirA
2279	Poly(A) polymerase [<i>Bacillus subtilis</i>]	UniRef100_P42977	<i>Bacillus subtilis</i>	Cca
2280	Putative glycosyl transferase ypjH [<i>Bacillus subtilis</i>]	UniRef100_P42982	<i>Bacillus subtilis</i>	YpjH
2281	Hypothetical protein ypjG [<i>Bacillus subtilis</i>]	UniRef100_P42981	<i>Bacillus subtilis</i>	YpjG
2282	Methylglyoxal synthase [<i>Bacillus subtilis</i>]	UniRef100_P42980	<i>Bacillus subtilis</i>	MgsA
2283	Dihydrodipicolinate reductase [<i>Bacillus subtilis</i>]	UniRef100_P42976	<i>Bacillus subtilis</i>	DapB
2284	Hypothetical protein ypjD [<i>Bacillus subtilis</i>]	UniRef100_P42979	<i>Bacillus subtilis</i>	YpjD
2285				YpjC
2286	Hypothetical protein ypjB precursor [<i>Bacillus subtilis</i>]	UniRef100_P54393	<i>Bacillus subtilis</i>	YpjB
2287				YpjA
2288				QcrC
2289	Menaquinol-cytochrome c reductase cytochrome b subunit [<i>Bacillus subtilis</i>]	UniRef100_P46912	<i>Bacillus subtilis</i>	QcrB
2290	Menaquinol-cytochrome c reductase iron-sulfur subunit [<i>Bacillus subtilis</i>]	UniRef100_P46911	<i>Bacillus subtilis</i>	QcrA
2291	Hypothetical protein ypiF [<i>Bacillus subtilis</i>]	UniRef100_P54391	<i>Bacillus subtilis</i>	YpiF
2292	Hypothetical UPF0302 protein ypiB [<i>Bacillus subtilis</i>]	UniRef100_P54390	<i>Bacillus subtilis</i>	YpiB
2293	Hypothetical protein ypiA [<i>Bacillus subtilis</i>]	UniRef100_P54389	<i>Bacillus subtilis</i>	YpiA
2294	3-phosphoshikimate 1-carboxyvinyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P20691	<i>Bacillus subtilis</i>	AroE
2295	Prephenate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P20692	<i>Bacillus subtilis</i>	TyrA
2296				HisC
2297	Tryptophan synthase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P07601	<i>Bacillus subtilis</i>	TrpA
2298	Tryptophan synthase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P07600	<i>Bacillus subtilis</i>	TrpB
2299	N-(S-phosphoribosyl)anthranilate isomerase [<i>Bacillus subtilis</i>]	UniRef100_P20167	<i>Bacillus subtilis</i>	TrpF
2300	Indole-3-glycerol phosphate synthase [<i>Bacillus subtilis</i>]	UniRef100_P03964	<i>Bacillus subtilis</i>	TrpC
2301	Anthranilate phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P03947	<i>Bacillus subtilis</i>	TrpD
2302	Anthranilate synthase component I [<i>Bacillus subtilis</i>]	UniRef100_P03963	<i>Bacillus subtilis</i>	TrpE
2303	Chorismate mutase [<i>Bacillus subtilis</i>]	UniRef100_P19080	<i>Bacillus subtilis</i>	AroH
2304	3-dehydroquinate synthase [<i>Bacillus subtilis</i>]	UniRef100_P31102	<i>Bacillus subtilis</i>	AroB
2305	Chorismate synthase [<i>Bacillus subtilis</i>]	UniRef100_P31104	<i>Bacillus subtilis</i>	AroF
2306	Chemotaxis protein methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P31105	<i>Bacillus subtilis</i>	CheR

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2307	Nucleoside diphosphate kinase [<i>Bacillus subtilis</i>]	UniRef100_P31103	<i>Bacillus subtilis</i>	Ndk
2308	Heptaprenyl diphosphate synthase component II [<i>Bacillus subtilis</i>]	UniRef100_P31114	<i>Bacillus subtilis</i>	HepT
2309	Menaquinone biosynthesis methyltransferase ubiE [<i>Bacillus subtilis</i>]	UniRef100_P31113	<i>Bacillus subtilis</i>	MenH
2310	Heptaprenyl diphosphate synthase component I [<i>Bacillus subtilis</i>]	UniRef100_P31112	<i>Bacillus subtilis</i>	HepS
2311	Transcription attenuation protein mtrB [<i>Bacillus subtilis</i>]	UniRef100_P19466	<i>Bacillus subtilis</i>	
2312	GTP cyclohydrolase I [<i>Bacillus subtilis</i>]	UniRef100_P19465	<i>Bacillus subtilis</i>	MtrA
2313	DNA-binding protein HU 1 [<i>Bacillus subtilis</i>]	UniRef100_P08821	<i>Bacillus subtilis</i>	
2314	Stage IV sporulation protein A [<i>Bacillus subtilis</i>]	UniRef100_P35149	<i>Bacillus subtilis</i>	SpoIVA
2315	Hypothetical protein yphF [<i>Bacillus subtilis</i>]	UniRef100_P39911	<i>Bacillus subtilis</i>	YphF
2316	Hypothetical protein yphE [<i>Bacillus subtilis</i>]	UniRef100_P50744	<i>Bacillus subtilis</i>	
2317	Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H- dependent glycerol-3-phosphate dehydrogenase) (NAD(P)H-dependent dihydroxyacetone-phosphate reductase) [<i>Bacillus subtilis</i>]	UniRef100_P46919	NAD(P)+	GpsA
2318				YphC
2319	Hypothetical protein OBI798 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EQA7	<i>Oceanobacillus iheyensis</i>	
2320	Hypothetical protein yphB [<i>Bacillus subtilis</i>]	UniRef100_P50742	<i>Bacillus subtilis</i>	SeaA
2321	Hypothetical protein yphA [<i>Bacillus subtilis</i>]	UniRef100_P50741	<i>Bacillus subtilis</i>	YphA
2322				
2323				YpgA YpfD
2324	30S ribosomal protein S1 homolog [<i>Bacillus subtilis</i>]	UniRef100_P38494	<i>Bacillus subtilis</i>	
2325	Cytidylate kinase [<i>Bacillus subtilis</i>]	UniRef100_P38493	<i>Bacillus subtilis</i>	Cmk
2326	Hypothetical protein ypfB [<i>Bacillus subtilis</i>]	UniRef100_P38492	<i>Bacillus subtilis</i>	
2327	Sporulation protein ypeB [<i>Bacillus subtilis</i>]	UniRef100_P38490	<i>Bacillus subtilis</i>	YpeB
2328	Spore cortex-lytic enzyme precursor [<i>Bacillus subtilis</i>]	UniRef100_P50739	<i>Bacillus subtilis</i>	SleB
2329				YpdC
2330				YccC YpdA
2331	Hypothetical protein ypdA [<i>Bacillus subtilis</i>]	UniRef100_P50736	<i>Bacillus subtilis</i>	
2332	NAD-specific glutamate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P50735	<i>Bacillus subtilis</i>	GudB
2333	Adapter protein mecA 2 [<i>Bacillus subtilis</i>]	UniRef100_P50734	<i>Bacillus subtilis</i>	YpbH
2334	Hypothetical protein ypbG precursor [<i>Bacillus subtilis</i>]	UniRef100_P50733	<i>Bacillus subtilis</i>	YpbG
2335				YpbF YpbE
2336	Hypothetical protein ypbE [<i>Bacillus subtilis</i>]	UniRef100_P50731	<i>Bacillus subtilis</i>	
2337	Hypothetical protein ypbD [<i>Bacillus subtilis</i>]	UniRef100_P50730	<i>Bacillus subtilis</i>	YpbD
2338	ATP-dependent DNA helicase recQ [<i>Bacillus subtilis</i>]	UniRef100_P50729	<i>Bacillus subtilis</i>	RecQ
2339	Hypothetical protein ypbB [<i>Bacillus subtilis</i>]	UniRef100_P50728	<i>Bacillus subtilis</i>	YpbB
2340	Ferredoxin [<i>Bacillus subtilis</i>]	UniRef100_P50727	<i>Bacillus subtilis</i>	
2341	Hypothetical protein ypaA [<i>Bacillus subtilis</i>]	UniRef100_P50726	<i>Bacillus subtilis</i>	YpaA
2342				

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2343	D-3-phosphoglycerate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P35136	<i>Bacillus subtilis</i>	SerA
2344	BH1600 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KCH1	<i>Bacillus halodurans</i>	
2345	Sigma-X negative effector [<i>Bacillus subtilis</i>]	UniRef100_P35166	<i>Bacillus subtilis</i>	RsiX
2346	RNA polymerase sigma factor sigX [<i>Bacillus subtilis</i>]	UniRef100_P35165	<i>Bacillus subtilis</i>	SigX
2347	Transcriptional regulator [<i>Bacillus cereus</i>]	UniRef100_Q72XJ3	<i>Bacillus cereus</i>	LytR
2348	Endo-1,4-beta-xylanase [<i>Bacillus halodurans</i>]	UniRef100_Q9K630	<i>Bacillus halodurans</i>	YheN
2349				
2350	Alkaline phosphatase synthesis sensor protein phoR [<i>Clostridium tetani</i>]	UniRef100_Q898N3	<i>Clostridium tetani</i>	YclK
2351	Response regulators consisting of a CheY-like receiver domain and a HTH DNA-binding domain [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8R9H7	<i>Thermoanaerobacter tengcongensis</i>	YycF
2352	Sensor protein resE [<i>Bacillus subtilis</i>]	UniRef100_P35164	<i>Bacillus subtilis</i>	ResE
2353	Transcriptional regulatory protein resD [<i>Bacillus subtilis</i>]	UniRef100_P35163	<i>Bacillus subtilis</i>	ResD
2354	Protein resC [<i>Bacillus subtilis</i>]	UniRef100_P35162	<i>Bacillus subtilis</i>	ResC
2355	ResB protein [<i>Bacillus subtilis</i>]	UniRef100_P35161	<i>Bacillus subtilis</i>	ResB
2356	Thiol-disulfide oxidoreductase resA [<i>Bacillus subtilis</i>]	UniRef100_P35160	<i>Bacillus subtilis</i>	ResA
2357	Ribosomal large subunit pseudouridine synthase B [<i>Bacillus subtilis</i>]	UniRef100_P35159	<i>Bacillus subtilis</i>	RluB
2358	Spore maturation protein B [<i>Bacillus subtilis</i>]	UniRef100_P35158	<i>Bacillus subtilis</i>	SpmB
2359	Spore maturation protein A [<i>Bacillus subtilis</i>]	UniRef100_P35157	<i>Bacillus subtilis</i>	SpmA
2360	Penicillin-binding protein 5* precursor [<i>Bacillus subtilis</i>]	UniRef100_P35150	<i>Bacillus subtilis</i>	DacB
2361	Hypothetical protein ypuI [<i>Bacillus subtilis</i>]	UniRef100_P35156	<i>Bacillus subtilis</i>	YpuI
2362	Segregation and condensation protein B [<i>Bacillus subtilis</i>]	UniRef100_P35155	<i>Bacillus subtilis</i>	YpuH
2363	Segregation and condensation protein A [<i>Bacillus subtilis</i>]	UniRef100_P35154	<i>Bacillus subtilis</i>	YpuG
2364				
2365	Hypothetical protein ypuF [<i>Bacillus subtilis</i>]	UniRef100_P17617	<i>Bacillus subtilis</i>	YpuF
2366	RibT protein [<i>Bacillus subtilis</i>]	UniRef100_P17622	<i>Bacillus subtilis</i>	RibT
2367	6,7-dimethyl-8-ribityllumazine synthase [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q44681	<i>Bacillus amyloliquefaciens</i>	RibH
2368	Riboflavin biosynthesis protein ribA [Includes: GTP cyclohydrolase II (EC 3.5.4.25); 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBP synthase)] [<i>Bacillus subtilis</i>]	UniRef100_P17620	<i>Bacillus subtilis</i>	RibA
2369	Riboflavin synthase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P16440	<i>Bacillus subtilis</i>	RibE
2370	Riboflavin biosynthesis protein ribD [Includes: Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26) (Riboflavin-specific deaminase); 5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193) (HTP reductase)] [<i>Bacillus subtilis</i>]	UniRef100_P17618	Includes: Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26) (Riboflavin-specific deaminase); 5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193) (HTP reductase)	RibD

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2371	Hypothetical protein ypuD [<i>Bacillus subtilis</i>]	UniRef100_P17616	<i>Bacillus subtilis</i>	YpuD
2372	Putative serine/threonine protein phosphatase [<i>Helicobacter hepaticus</i>]	UniRef100_Q7VFC1	<i>Helicobacter hepaticus</i>	YjbP
2373	Stress response homolog Hsp [<i>Bacillus subtilis</i>]	UniRef100_Q9X3Z5	<i>Bacillus subtilis</i>	
2374	Response regulator aspartate phosphatase A [<i>Bacillus subtilis</i>]	UniRef100_Q00828	<i>Bacillus subtilis</i>	RapA
2375	Peptidyl-prolyl cis-trans isomerase B [<i>Bacillus subtilis</i>]	UniRef100_P35137	<i>Bacillus subtilis</i>	PpiB
2376				YpuA
2377	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
2378	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		YndL
2379				LysA
2380	Diaminopimelate decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P23630	<i>Bacillus subtilis</i>	
2381				SpoVAF
2382	Stage V sporulation protein AF [<i>Bacillus subtilis</i>]	UniRef100_P31845	<i>Bacillus subtilis</i>	
2383	Stage V sporulation protein AE [<i>Bacillus subtilis</i>]	UniRef100_P40870	<i>Bacillus subtilis</i>	SpoVAE
2384	Stage V sporulation protein AE [<i>Bacillus subtilis</i>]	UniRef100_P40870	<i>Bacillus subtilis</i>	SpoVAE
2385	Stage V sporulation protein AD [<i>Bacillus subtilis</i>]	UniRef100_P40869	<i>Bacillus subtilis</i>	SpoVAD
2386				SpoVAC
2387	Stage V sporulation protein AB [<i>Bacillus subtilis</i>]	UniRef100_P40867	<i>Bacillus subtilis</i>	SpoVAB
2388	Stage V sporulation protein AA [<i>Bacillus subtilis</i>]	UniRef100_P40866	<i>Bacillus subtilis</i>	SpoVAA
2389				SigF
2390				SpoIIAB
2391				SpoIIAA
2392	Penicillin-binding protein dacF precursor [<i>Bacillus subtilis</i>]	UniRef100_P38422	<i>Bacillus subtilis</i>	DacF
2393	Purine nucleoside phosphorylase I [<i>Bacillus subtilis</i>]	UniRef100_P46354	<i>Bacillus subtilis</i>	PunA
2394	Phosphopentomutase [<i>Bacillus subtilis</i>]	UniRef100_P46353	<i>Bacillus subtilis</i>	Drm
2395	Tyrosine recombinase xerD [<i>Bacillus subtilis</i>]	UniRef100_P46352	<i>Bacillus subtilis</i>	RipX
2396				
2397	Ferric uptake regulation protein [<i>Bacillus subtilis</i>]	UniRef100_P54574	<i>Bacillus subtilis</i>	Fur
2398	Stage II sporulation protein M [<i>Bacillus subtilis</i>]	UniRef100_P37873	<i>Bacillus subtilis</i>	SpoIIM
2399	Lmo2763 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q926Y4	<i>Listeria monocytogenes</i>	YdhO
2400	UPI00003CA2F1 UniRef100 entry	UniRef100_UPI00003CA2F1		
2401	Probable allantoin permease [<i>Bacillus subtilis</i>]	UniRef100_P94575	<i>Bacillus subtilis</i>	YwoE
2402				YdaF
2403	Hypothetical protein yqkK [<i>Bacillus subtilis</i>]	UniRef100_P54573	<i>Bacillus subtilis</i>	
2404				YybD
2405	Hypothetical protein [<i>Bacillus pumilus</i>]	UniRef100_Q93PN4	<i>Bacillus pumilus</i>	YqxK
2406	ADP-ribose pyrophosphatase [<i>Bacillus subtilis</i>]	UniRef100_P54570	<i>Bacillus subtilis</i>	NudF
2407				
2408	YdgC protein [<i>Bacillus subtilis</i>]	UniRef100_P96701	<i>Bacillus subtilis</i>	YdgC
2409	Hypothetical protein ydgD [<i>Bacillus subtilis</i>]	UniRef100_P96702	<i>Bacillus subtilis</i>	YdgD
2410	Hypothetical oxidoreductase yqkF [<i>Bacillus subtilis</i>]	UniRef100_P54569	<i>Bacillus subtilis</i>	YqkF
2411				
2412				YqkD
2413	Hypothetical protein yqkC [<i>Bacillus subtilis</i>]	UniRef100_P54566	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2414	Hypothetical protein yqkB [<i>Bacillus subtilis</i>]	UniRef100_P54565	<i>Bacillus subtilis</i>	YqkB
2415	UPI00003CBE2B UniRef100 entry	UniRef100_UPI00003CBE2B		
2416	Hypothetical UPF0157 protein yqkA [<i>Bacillus subtilis</i>]	UniRef100_P54564	<i>Bacillus subtilis</i>	YqkA
2417	Hypothetical protein yqjZ [<i>Bacillus subtilis</i>]	UniRef100_P54563	<i>Bacillus subtilis</i>	YqjZ
2418	Lipase precursor [<i>Bacillus subtilis</i>]	UniRef100_P37957	<i>Bacillus subtilis</i>	Lip
2419				
2420	Nickel ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K BX8	<i>Bacillus halodurans</i>	AppA
2421	Nickel ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K BX7	<i>Bacillus halodurans</i>	AppB
2422	Nickel ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K BX6	<i>Bacillus halodurans</i>	AppC
2423	Putative oligopeptide ABC transporter [<i>Clostridium tetani</i>]	UniRef100_Q895A4	<i>Clostridium tetani</i>	DppD
2424	Putative oligopeptide ABC transporter [<i>Clostridium tetani</i>]	UniRef100_Q895A5	<i>Clostridium tetani</i>	OppF
2425	DNA-damage-inducible protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDR1	<i>Bacillus halodurans</i>	YojI
2426	Fibronectin-binding protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63CW1	<i>Bacillus cereus</i> ZK	
2427	Hypothetical protein yolD [Bacteriophage SPBc2]	UniRef100_O64030	Bacteriophage SPBc2	YolD
2428				UvrX
2429	Hypothetical protein yqzH [<i>Bacillus subtilis</i>]	UniRef100_O32014	<i>Bacillus subtilis</i>	
2430	Hypothetical transport protein yqjV [<i>Bacillus subtilis</i>]	UniRef100_P54559	<i>Bacillus subtilis</i>	YqjV
2431	Hypothetical protein yqjT [<i>Bacillus subtilis</i>]	UniRef100_P54557	<i>Bacillus subtilis</i>	YqjT
2432				CoaA
2433	LacG [<i>Lactococcus lactis</i>]	UniRef100_Q9RAU9	<i>Lactococcus lactis</i>	
2434	LacF [<i>Lactococcus lactis</i>]	UniRef100_Q9RAV2	<i>Lactococcus lactis</i>	YdbJ
2435				
2436	Hypothetical protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70K07	<i>Bacillus amyloliquefaciens</i>	
2437	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q737G3	<i>Bacillus cereus</i>	
2438	YmaC protein [<i>Bacillus subtilis</i>]	UniRef100_O31789	<i>Bacillus subtilis</i>	YmaC
2439	Hypothetical oxidoreductase yqjQ [<i>Bacillus subtilis</i>]	UniRef100_P54554	<i>Bacillus subtilis</i>	YqjQ
2440				YqjP
2441				ProI
2442	Hypothetical protein yqjN [<i>Bacillus subtilis</i>]	UniRef100_P54551	<i>Bacillus subtilis</i>	YqjN
2443	Probable NADH-dependent flavin oxidoreductase yqjM [<i>Bacillus subtilis</i>]	UniRef100_P54550	<i>Bacillus subtilis</i>	YqjM
2444	Hypothetical protein yqjL [<i>Bacillus subtilis</i>]	UniRef100_P54549	<i>Bacillus subtilis</i>	YqjL
2445				
2446	Ribonuclease Z [<i>Bacillus subtilis</i>]	UniRef100_P54548	<i>Bacillus subtilis</i>	YqjK
2447	Glucose-6-phosphate 1-dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P54547	<i>Bacillus subtilis</i>	Zwf
2448	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELZ9	<i>Oceanobacillus iheyensis</i>	CitH
2449	Hypothetical protein OB3065 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELZ8	<i>Oceanobacillus iheyensis</i>	
2450	Immunogenic protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELZ7	<i>Oceanobacillus iheyensis</i>	SsuA
2451	YdhQ protein [<i>Bacillus subtilis</i>]	UniRef100_O05509	<i>Bacillus subtilis</i>	YdhQ
2452	Beta-glucosidase [<i>Bacillus halodurans</i>]	UniRef100_Q9K615	<i>Bacillus halodurans</i>	YdhP
2453	Putative cellobiose-specific enzyme IIC [<i>Bacillus pumilus</i>]	UniRef100_Q8KP28	<i>Bacillus pumilus</i>	YdhO
2454	YdhO protein [<i>Bacillus subtilis</i>]	UniRef100_O05507	<i>Bacillus subtilis</i>	YdhO
2455	YdhN protein [<i>Bacillus subtilis</i>]	UniRef100_O05506	<i>Bacillus subtilis</i>	YdhN

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2456	PTS system, cellobiose-specific enzyme II, B component [<i>Bacillus halodurans</i>]	UniRef100_Q9K613	<i>Bacillus halodurans</i>	
2457	Alkaline phosphatase [<i>Bacillus cereus</i> ZK]	UniRef100_Q639W1	<i>Bacillus cereus</i> ZK	PhoB
2458	Glucose 1-dehydrogenase A [<i>Bacillus megaterium</i>]	UniRef100_P10528	<i>Bacillus megaterium</i>	Gdh
2459	6-phosphogluconate dehydrogenase, decarboxylating II [<i>Bacillus subtilis</i>]	UniRef100_P80859	<i>Bacillus subtilis</i>	YqjI
2460	DNA polymerase IV 1 [<i>Bacillus subtilis</i>]	UniRef100_P54545	<i>Bacillus subtilis</i>	YqjH
2461	Hypothetical protein yqzJ [<i>Bacillus subtilis</i>]	UniRef100_Q7WY64	<i>Bacillus subtilis</i>	
2462				YqjG
2463	Hypothetical protein yqjE [<i>Bacillus subtilis</i>]	UniRef100_P54542	<i>Bacillus subtilis</i>	YqjE
2464	Methylmalonyl-CoA decarboxylase alpha subunit [<i>Bacillus halodurans</i>]	UniRef100_Q9K8P6	<i>Bacillus halodurans</i>	YqjD
2465	Hypothetical protein yqjA [<i>Bacillus subtilis</i>]	UniRef100_P54538	<i>Bacillus subtilis</i>	YqjA
2466	Probable amino-acid ABC transporter ATP-binding protein yqiZ [<i>Bacillus subtilis</i>]	UniRef100_P54537	<i>Bacillus subtilis</i>	YqiZ
2467	Probable amino-acid ABC transporter permease protein yqiY [<i>Bacillus subtilis</i>]	UniRef100_P54536	<i>Bacillus subtilis</i>	YqiY
2468	Probable amino-acid ABC transporter extracellular binding protein yqiX precursor [<i>Bacillus subtilis</i>]	UniRef100_P54535	<i>Bacillus subtilis</i>	YqiX
2469	Hypothetical protein yqiW [<i>Bacillus subtilis</i>]	UniRef100_P54534	<i>Bacillus subtilis</i>	YqiW
2470	Protein bmrU [<i>Bacillus subtilis</i>]	UniRef100_P39074	<i>Bacillus subtilis</i>	BmrU
2471	Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.168) (Dihydrolipoyllysine-residue (2-methylpropanoyl)transferase) [<i>Bacillus subtilis</i>]	UniRef100_P37942	<i>Bacillus subtilis</i>	BkdB
2472	2-oxoisovalerate dehydrogenase beta subunit [<i>Bacillus subtilis</i>]	UniRef100_P37941	<i>Bacillus subtilis</i>	BkdAB
2473	2-oxoisovalerate dehydrogenase alpha subunit [<i>Bacillus subtilis</i>]	UniRef100_P37940	<i>Bacillus subtilis</i>	BkdAA
2474	Dihydrolipoyl dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P54533	<i>Bacillus subtilis</i>	LpdV
2475	Probable butyrate kinase [<i>Bacillus subtilis</i>]	UniRef100_P54532	<i>Bacillus subtilis</i>	Buk
2476	Leucine dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P54531	<i>Bacillus subtilis</i>	Bed
2477	Probable phosphate butyryltransferase [<i>Bacillus subtilis</i>]	UniRef100_P54530	<i>Bacillus subtilis</i>	Ptb
2478	Putative sigma L-dependent transcriptional regulator yqiR [<i>Bacillus subtilis</i>]	UniRef100_P54529	<i>Bacillus subtilis</i>	BkdR
2479	Hypothetical protein yqzF [<i>Bacillus subtilis</i>]	UniRef100_O32015	<i>Bacillus subtilis</i>	
2480	Hypothetical protein yqiK [<i>Bacillus subtilis</i>]	UniRef100_P54527	<i>Bacillus subtilis</i>	YqiK
2481	Hypothetical protein yqiI [<i>Bacillus subtilis</i>]	UniRef100_P54525	<i>Bacillus subtilis</i>	YqiI
2482				YqiH
2483	Serine O-acetyltransferase [<i>Methanosarcina mazei</i>]	UniRef100_Q8PSY4	<i>Methanosarcina mazei</i>	CysE
2484				
2485	Stage 0 sporulation protein A [<i>Bacillus subtilis</i>]	UniRef100_P06534	<i>Bacillus subtilis</i>	Spo0A
2486				SpoIVB
2487	DNA repair protein recN [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q659H4	<i>Bacillus amyloliquefaciens</i>	RecN
2488	Arginine repressor [<i>Bacillus subtilis</i>]	UniRef100_P17893	<i>Bacillus subtilis</i>	AhrC

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2489	Hypothetical protein yqxC [<i>Bacillus subtilis</i>]	UniRef100_P19672	<i>Bacillus subtilis</i>	YqxC
2490				Dxs
2491	Geranyltranstransferase [<i>Bacillus subtilis</i>]	UniRef100_P54383	<i>Bacillus subtilis</i>	YqiD
2492	Probable exodeoxyribonuclease VII small subunit [<i>Bacillus subtilis</i>]	UniRef100_P54522	<i>Bacillus subtilis</i>	
2493	Probable exodeoxyribonuclease VII large subunit [<i>Bacillus subtilis</i>]	UniRef100_P54521	<i>Bacillus subtilis</i>	YqiB
2494	FoID bifunctional protein [Includes: Methylenetetrahydrofolate dehydrogenase (EC 1.5.1.5); Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)] [<i>Bacillus subtilis</i>]	UniRef100_P54382	Includes: Methylenetetrahydrofolate dehydrogenase (EC 1.5.1.5); Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	FoID
2495	N utilization substance protein B homolog [<i>Bacillus subtilis</i>]	UniRef100_P54520	<i>Bacillus subtilis</i>	NusB
2496	Hypothetical protein yqhY [<i>Bacillus subtilis</i>]	UniRef100_P54519	<i>Bacillus subtilis</i>	YqhY
2497				AccC
2498	Biotin carboxyl carrier protein of acetyl-CoA carboxylase [<i>Bacillus subtilis</i>]	UniRef100_P49786	<i>Bacillus subtilis</i>	AccB
2499	Stage III sporulation protein AH [<i>Bacillus subtilis</i>]	UniRef100_P49785	<i>Bacillus subtilis</i>	SpoIIIAH
2500	Stage III sporulation protein AG [<i>Bacillus subtilis</i>]	UniRef100_P49784	<i>Bacillus subtilis</i>	SpoIILAG
2501	Stage III sporulation protein AF [<i>Bacillus subtilis</i>]	UniRef100_P49783	<i>Bacillus subtilis</i>	SpoIIIAF
2502	Stage III sporulation protein AE [<i>Bacillus subtilis</i>]	UniRef100_P49782	<i>Bacillus subtilis</i>	SpoIIIAE
2503	Stage III sporulation protein AD [<i>Bacillus subtilis</i>]	UniRef100_P49781	<i>Bacillus subtilis</i>	SpoIIIDAD
2504	Stage III sporulation protein AC [<i>Bacillus subtilis</i>]	UniRef100_P49780	<i>Bacillus subtilis</i>	
2505	Stage III sporulation protein AB [<i>Bacillus subtilis</i>]	UniRef100_Q01368	<i>Bacillus subtilis</i>	SpoIIIAAB
2506				SpoIIIAA
2507	Hypothetical protein yqhV [<i>Bacillus subtilis</i>]	UniRef100_P49779	<i>Bacillus subtilis</i>	
2508	Elongation factor P [<i>Bacillus subtilis</i>]	UniRef100_P49778	<i>Bacillus subtilis</i>	Efp
2509	Putative peptidase yqHT [<i>Bacillus subtilis</i>]	UniRef100_P54518	<i>Bacillus subtilis</i>	YqhT
2510	3-dehydroquinate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P54517	<i>Bacillus subtilis</i>	YqhS
2511	Hypothetical protein yqHR [<i>Bacillus subtilis</i>]	UniRef100_P54516	<i>Bacillus subtilis</i>	YqhR
2512	Hypothetical protein yqhQ [<i>Bacillus subtilis</i>]	UniRef100_P54515	<i>Bacillus subtilis</i>	YqhQ
2513	Hypothetical protein yqhP [<i>Bacillus subtilis</i>]	UniRef100_P54514	<i>Bacillus subtilis</i>	YqhP
2514	Hypothetical protein yqhO [<i>Bacillus subtilis</i>]	UniRef100_P54513	<i>Bacillus subtilis</i>	YqhO
2515	Transcriptional regulator mntR [<i>Bacillus subtilis</i>]	UniRef100_P54512	<i>Bacillus subtilis</i>	MntR
2516	Hypothetical protein yqhM [<i>Bacillus subtilis</i>]	UniRef100_P54511	<i>Bacillus subtilis</i>	YqhM
2517	Hypothetical protein yqhL [<i>Bacillus subtilis</i>]	UniRef100_P54510	<i>Bacillus subtilis</i>	YqhL
2518	Glycine betaine-binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P46922	<i>Bacillus subtilis</i>	OpuAC
2519	Glycine betaine transport system permease protein opuAB [<i>Bacillus subtilis</i>]	UniRef100_P46921	<i>Bacillus subtilis</i>	OpuAB
2520	Glycine betaine transport ATP-binding protein opuAA [<i>Bacillus subtilis</i>]	UniRef100_P46920	<i>Bacillus subtilis</i>	OpuAA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2521	Probable glycine dehydrogenase [decarboxylating] subunit 2 [<i>Bacillus subtilis</i>]	UniRef100_P54377	decarboxylating	GcvPB
2522	Probable glycine dehydrogenase [decarboxylating] subunit 1 [<i>Bacillus subtilis</i>]	UniRef100_P54376	decarboxylating	GcvPA
2523	Aminomethyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P54378	<i>Bacillus subtilis</i>	GcvT
2524	Hypothetical helicase yqhH [<i>Bacillus subtilis</i>]	UniRef100_P54509	<i>Bacillus subtilis</i>	YqhH
2525	Hypothetical protein yqhG [<i>Bacillus subtilis</i>]	UniRef100_P54508	<i>Bacillus subtilis</i>	YqhG
2526				
2527				SinR
2528	Spore coat-associated protein N [<i>Bacillus subtilis</i>]	UniRef100_P54507	<i>Bacillus subtilis</i>	TasA
2529	Signal peptidase I W [<i>Bacillus subtilis</i>]	UniRef100_P54506	<i>Bacillus subtilis</i>	SipW
2530	Hypothetical protein yqxM [<i>Bacillus subtilis</i>]	UniRef100_P40949	<i>Bacillus subtilis</i>	YqxM
2531	YqzG protein [<i>Bacillus subtilis</i>]	UniRef100_O32019	<i>Bacillus subtilis</i>	YqzG
2532	YqzE protein [<i>Bacillus subtilis</i>]	UniRef100_O32020	<i>Bacillus subtilis</i>	
2533				ComGG
2534				ComGF
2535	ComG operon protein 5 precursor [<i>Bacillus subtilis</i>]	UniRef100_P25957	<i>Bacillus subtilis</i>	ComGE
2536	ComG operon protein 4 precursor [<i>Bacillus subtilis</i>]	UniRef100_P25956	<i>Bacillus subtilis</i>	ComGD
2537	ComG operon protein 3 precursor [<i>Bacillus subtilis</i>]	UniRef100_P25955	<i>Bacillus subtilis</i>	
2538				ComGB
2539	ComG operon protein 1 [<i>Bacillus subtilis</i>]	UniRef100_P25953	<i>Bacillus subtilis</i>	ComGA
2540	Hypothetical protein yqhA [<i>Bacillus subtilis</i>]	UniRef100_P54504	<i>Bacillus subtilis</i>	YqhA
2541	Hypothetical protein yqgZ [<i>Bacillus subtilis</i>]	UniRef100_P54503	<i>Bacillus subtilis</i>	YqgZ
2542	Hypothetical protein yqgY [<i>Bacillus subtilis</i>]	UniRef100_P54502	<i>Bacillus subtilis</i>	
2543	Hypothetical protein yqgX [<i>Bacillus subtilis</i>]	UniRef100_P54501	<i>Bacillus subtilis</i>	YqgX
2544	Hypothetical protein yqgW [<i>Bacillus subtilis</i>]	UniRef100_P54500	<i>Bacillus subtilis</i>	
2545	Hypothetical protein [<i>Staphylococcus epidermidis</i>]	UniRef100_Q8CSE8	<i>Staphylococcus epidermidis</i>	
2546				YqgU
2547	Hypothetical protein yqgT [<i>Bacillus subtilis</i>]	UniRef100_P54497	<i>Bacillus subtilis</i>	YqgT
2548	Ferrichrome-binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P37580	<i>Bacillus subtilis</i>	FhuD
2549	Hypothetical protein yqgS [<i>Bacillus subtilis</i>]	UniRef100_P54496	<i>Bacillus subtilis</i>	YqgS
2550	Glucokinase [<i>Bacillus subtilis</i>]	UniRef100_P54495	<i>Bacillus subtilis</i>	GlcK
2551	Hypothetical protein yqgQ [<i>Bacillus subtilis</i>]	UniRef100_P54494	<i>Bacillus subtilis</i>	
2552	Stage V sporulation protein AF [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EQ08	<i>Oceanobacillus iheyensis</i>	SpoVAF
2553	Hypothetical protein yqgP [<i>Bacillus subtilis</i>]	UniRef100_P54493	<i>Bacillus subtilis</i>	YqgP
2554	Hypothetical protein yqgO [<i>Bacillus subtilis</i>]	UniRef100_P54492	<i>Bacillus subtilis</i>	
2555	Hypothetical protein yqgN [<i>Bacillus subtilis</i>]	UniRef100_P54491	<i>Bacillus subtilis</i>	YqgN
2556				
2557				YqgM
2558	Hypothetical protein yqgL [<i>Bacillus subtilis</i>]	UniRef100_P54489	<i>Bacillus subtilis</i>	YqgL
2559				YqzD
2560	YqzC protein [<i>Bacillus subtilis</i>]	UniRef100_O32023	<i>Bacillus subtilis</i>	YqzC

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2561	Phosphate import ATP-binding protein pstB 1 [<i>Bacillus subtilis</i>]	UniRef100_P46342	<i>Bacillus subtilis</i>	PstBB
2562	Phosphate import ATP-binding protein pstB 2 [<i>Bacillus subtilis</i>]	UniRef100_P46341	<i>Bacillus subtilis</i>	PstBA
2563	Probable ABC transporter permease protein yqgI [<i>Bacillus subtilis</i>]	UniRef100_P46340	<i>Bacillus subtilis</i>	PstA
2564	Probable ABC transporter permease protein yqgH [<i>Bacillus subtilis</i>]	UniRef100_P46339	<i>Bacillus subtilis</i>	PstC
2565	Probable ABC transporter binding protein yqgG precursor [<i>Bacillus subtilis</i>]	UniRef100_P46338	<i>Bacillus subtilis</i>	PstS
2566	Hypothetical protein yqgF [<i>Bacillus subtilis</i>]	UniRef100_P54488	<i>Bacillus subtilis</i>	PbpA
2567	Hypothetical protein yqgE [<i>Bacillus subtilis</i>]	UniRef100_P54487	<i>Bacillus subtilis</i>	YqgE
2568				SodA
2569	Hypothetical protein yqgC [<i>Bacillus subtilis</i>]	UniRef100_P54486	<i>Bacillus subtilis</i>	YqgC
2570	Hypothetical protein yqgB [<i>Bacillus subtilis</i>]	UniRef100_P54485	<i>Bacillus subtilis</i>	YqgB
2571	Hypothetical protein yqfZ [<i>Bacillus subtilis</i>]	UniRef100_P54483	<i>Bacillus subtilis</i>	
2572	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (EC 1.17.4.3) (1- hydroxy-2-methyl-2-(E)-butenyl 4- diphosphate synthase) [<i>Bacillus subtilis</i>]	UniRef100_P54482	<i>Bacillus subtilis</i>	YqfY
2573	Hypothetical protein yqfX [<i>Bacillus subtilis</i>]	UniRef100_P54481	<i>Bacillus subtilis</i>	YqfX
2574	Putative nucleotidase yqfW [<i>Bacillus subtilis</i>]	UniRef100_P54480	<i>Bacillus subtilis</i>	YqfW
2575	Zinc-specific metalloregulatory protein [<i>Bacillus subtilis</i>]	UniRef100_P54479	<i>Bacillus subtilis</i>	Zur
2576	Metal (Zinc) transport protein [<i>Listeria innocua</i>]	UniRef100_Q926D9	<i>Listeria innocua</i>	YceA
2577				YcdI
2578	Hypothetical protein yqfU [<i>Bacillus subtilis</i>]	UniRef100_P54478	<i>Bacillus subtilis</i>	YqfU
2579				
2580	Probable endonuclease IV [<i>Bacillus subtilis</i>]	UniRef100_P54476	<i>Bacillus subtilis</i>	YqfS
2581	Probable RNA helicase yqfR [<i>Bacillus subtilis</i>]	UniRef100_P54475	<i>Bacillus subtilis</i>	YqfR
2582	Hypothetical protein yqfQ [<i>Bacillus subtilis</i>]	UniRef100_P54474	<i>Bacillus subtilis</i>	YqfQ
2583	4-hydroxy-3-methylbut-2-enyl diphosphate reductase [<i>Bacillus subtilis</i>]	UniRef100_P54473	<i>Bacillus subtilis</i>	YqfP
2584				YqfO
2585	Hypothetical protein yqfN [<i>Bacillus subtilis</i>]	UniRef100_P54471	<i>Bacillus subtilis</i>	YqfN
2586	YwqL protein [<i>Bacillus subtilis</i>]	UniRef100_P96724	<i>Bacillus subtilis</i>	YwqL
2587				
2588				
2589				
2590				
2591	Hypothetical protein CAC0336 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97M62	<i>Clostridium acetobutylicum</i>	
2592	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGF2	<i>Bacillus thuringiensis</i>	
2593				
2594	YwqJ protein [<i>Bacillus subtilis</i>]	UniRef100_P96722	<i>Bacillus subtilis</i>	YwqJ
2595	Hypothetical protein ywqI [<i>Bacillus subtilis</i>]	UniRef100_P96721	<i>Bacillus subtilis</i>	
2596				YwqH
2597	Cytochrome c-550 [<i>Bacillus subtilis</i>]	UniRef100_P24469	<i>Bacillus subtilis</i>	CccA
2598				SigA
2599	DNA primase [<i>Bacillus subtilis</i>]	UniRef100_P05096	<i>Bacillus subtilis</i>	DnaG

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2600	Hypothetical UPF0178 protein yqxD [<i>Bacillus subtilis</i>]	UniRef100_P17868	<i>Bacillus subtilis</i>	YqxD
2601	Hypothetical UPF0085 protein yqfL [<i>Bacillus subtilis</i>]	UniRef100_P54470	<i>Bacillus subtilis</i>	YqfL
2602	YqzB protein [<i>Bacillus subtilis</i>]	UniRef100_O34994	<i>Bacillus subtilis</i>	YqzB
2603	Glycyl-tRNA synthetase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P54381	<i>Bacillus subtilis</i>	GlyS
2604	Glycyl-tRNA synthetase alpha chain [<i>Bacillus subtilis</i>]	UniRef100_P54380	<i>Bacillus subtilis</i>	GlyQ
2605	DNA repair protein recO [<i>Bacillus subtilis</i>]	UniRef100_P42095	<i>Bacillus subtilis</i>	RecO
2606				
2607	GTP-binding protein era homolog [<i>Bacillus subtilis</i>]	UniRef100_P42182	<i>Bacillus subtilis</i>	Era
2608	Cytidine deaminase [<i>Bacillus subtilis</i>]	UniRef100_P19079	<i>Bacillus subtilis</i>	Cdd
2609				
2610	Hypothetical UPF0054 protein yqfG [<i>Bacillus subtilis</i>]	UniRef100_P46347	<i>Bacillus subtilis</i>	YqfG
2611	Hypothetical protein yqfF [<i>Bacillus subtilis</i>]	UniRef100_P46344	<i>Bacillus subtilis</i>	YqfF
2612	PhoH-like protein [<i>Bacillus subtilis</i>]	UniRef100_P46343	<i>Bacillus subtilis</i>	PhoH
2613				YqfD
2614	Hypothetical protein yqfC [<i>Bacillus subtilis</i>]	UniRef100_P54468	<i>Bacillus subtilis</i>	
2615	Hypothetical protein yqfB [<i>Bacillus subtilis</i>]	UniRef100_P54467	<i>Bacillus subtilis</i>	YqfB
2616	Hypothetical protein yqfA [<i>Bacillus subtilis</i>]	UniRef100_P54466	<i>Bacillus subtilis</i>	YqfA
2617				YqeZ
2618	Hypothetical protein yqeY [<i>Bacillus subtilis</i>]	UniRef100_P54464	<i>Bacillus subtilis</i>	YqeY
2619				
2620	Hypothetical protein yqeW [<i>Bacillus subtilis</i>]	UniRef100_P54463	<i>Bacillus subtilis</i>	YqeW
2621	Deoxyribose-phosphate aldolase [<i>Listeria innocua</i>]	UniRef100_Q92A19	<i>Listeria innocua</i>	Dra
2622	Hypothetical UPF0004 protein yqeV [<i>Bacillus subtilis</i>]	UniRef100_P54462	<i>Bacillus subtilis</i>	YqeV
2623	Hypothetical UPF0088 protein yqeU [<i>Bacillus subtilis</i>]	UniRef100_P54461	<i>Bacillus subtilis</i>	YqeU
2624	Ribosomal protein L11 methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P54460	<i>Bacillus subtilis</i>	YqeT
2625	Chaperone protein dnaJ [<i>Bacillus subtilis</i>]	UniRef100_P17631	<i>Bacillus subtilis</i>	DnaJ
2626				DnaK
2627				
2628				
2629	Heat-inducible transcription repressor hrcA [<i>Bacillus subtilis</i>]	UniRef100_P25499	<i>Bacillus subtilis</i>	HrcA
2630	Probable oxygen-independent coproporphyrinogen III oxidase [<i>Bacillus subtilis</i>]	UniRef100_P54304	<i>Bacillus subtilis</i>	HemN
2631				LepA
2632	Hypothetical protein yqxA [<i>Bacillus subtilis</i>]	UniRef100_P38425	<i>Bacillus subtilis</i>	YqxA
2633				SpolIP
2634	Germination protease precursor [<i>Bacillus subtilis</i>]	UniRef100_P22322	<i>Bacillus subtilis</i>	Gpr
2635	30S ribosomal protein S20 [<i>Bacillus subtilis</i>]	UniRef100_P21477	<i>Bacillus subtilis</i>	
2636	Hypothetical protein yqeN [<i>Bacillus subtilis</i>]	UniRef100_P54459	<i>Bacillus subtilis</i>	YqeN
2637				
2638	ComE operon protein 3 [<i>Bacillus subtilis</i>]	UniRef100_P39695	<i>Bacillus subtilis</i>	ComEC
2639	ComE operon protein 2 [<i>Bacillus subtilis</i>]	UniRef100_P32393	<i>Bacillus subtilis</i>	ComEB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2640	ComE operon protein 1 [<i>Bacillus subtilis</i>]	UniRef100_P39694	<i>Bacillus subtilis</i>	ComEA
2641	ComE operon protein 4 [<i>Bacillus subtilis</i>]	UniRef100_P39696	<i>Bacillus subtilis</i>	ComER
2642	Hypothetical protein yqeM [<i>Bacillus subtilis</i>]	UniRef100_P54458	<i>Bacillus subtilis</i>	YqeM
2643	Hypothetical protein yqeL [<i>Bacillus subtilis</i>]	UniRef100_P54457	<i>Bacillus subtilis</i>	YqeL
2644	Hypothetical protein yqeK [<i>Bacillus subtilis</i>]	UniRef100_P54456	<i>Bacillus subtilis</i>	YqeK
2645	Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18) (Deamido-NAD(+) pyrophosphorylase) (Deamido-NAD(+) diphosphorylase) [<i>Bacillus subtilis</i>]	UniRef100_P54455	<i>Bacillus subtilis</i>	YqeJ
2646	Hypothetical UPF0044 protein yqeI [<i>Bacillus subtilis</i>]	UniRef100_P54454	<i>Bacillus subtilis</i>	
2647	Shikimate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P54374	<i>Bacillus subtilis</i>	AroD
2648	Hypothetical protein yqeH [<i>Bacillus subtilis</i>]	UniRef100_P54453	<i>Bacillus subtilis</i>	YqeH
2649	Hypothetical protein yqeG [<i>Bacillus subtilis</i>]	UniRef100_P54452	<i>Bacillus subtilis</i>	YqeG
2650				
2651				
2652	Hypothetical lipoprotein yqeF precursor [<i>Bacillus subtilis</i>]	UniRef100_P54451	<i>Bacillus subtilis</i>	YqeF
2653	Acetyltransferase, GNAT family [<i>Bacillus anthracis</i>]	UniRef100_Q81KW8	<i>Bacillus anthracis</i>	YdfB
2654	Hypothetical protein yrhF [<i>Bacillus subtilis</i>]	UniRef100_O05398	<i>Bacillus subtilis</i>	YrhF
2655	Formate dehydrogenase chain A [<i>Bacillus subtilis</i>]	UniRef100_O05397	<i>Bacillus subtilis</i>	YrhE
2656	Hypothetical protein yrhD [<i>Bacillus subtilis</i>]	UniRef100_O05396	<i>Bacillus subtilis</i>	YrhD
2657				
2658	RNA polymerase sigma-K factor precursor [<i>Bacillus subtilis</i>]	UniRef100_P12254	<i>Bacillus subtilis</i>	SpoIIIC
2659				YcnB
2660	BH2157 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KAX9	<i>Bacillus halodurans</i>	YuaI
2661				
2662				
2663	Alanyl-tRNA synthetase family protein [<i>Bacillus anthracis</i>]	UniRef100_Q81Y73	<i>Bacillus anthracis</i>	AlaS
2664	METAL-ACTIVATED PYRIDOXAL ENZYME [<i>Brucella melitensis</i>]	UniRef100_Q8YCI2	<i>Brucella melitensis</i>	
2665	Probable translation initiation inhibitor [<i>Photobacterium profundum</i>]	UniRef100_Q6LKM3	<i>Photobacterium profundum</i>	YabJ
2666				YccC
2667	Putative threonine synthase [<i>Streptomyces avermitilis</i>]	UniRef100_Q82IF6	<i>Streptomyces avermitilis</i>	ThrC
2668				YabJ
2669	Hypothetical protein [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8RBA0	<i>Thermoanaerobacter tengcongensis</i>	
2670				
2671	UPI000032CE59 UniRef100 entry	UniRef100_UPI000032CE59		
2672	Multidrug-efflux transporter 1 regulator [<i>Bacillus subtilis</i>]	UniRef100_P39075	<i>Bacillus subtilis</i>	BmrR
2673	Metallo-beta-lactamase/rhodanese-like domain protein [<i>Bacillus anthracis</i>]	UniRef100_Q81Q95	<i>Bacillus anthracis</i>	YrkH
2674	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63B51	<i>Bacillus cereus</i> ZK	YumB
2675	NreC [<i>Staphylococcus carnosus</i>]	UniRef100_Q7WZY4	<i>Staphylococcus carnosus</i>	DegU
2676	Two-component sensor histidine kinase [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67JE7	<i>Symbiobacterium thermophilum</i>	DegS
2677	YdfQ protein [<i>Bacillus subtilis</i>]	UniRef100_P96695	<i>Bacillus subtilis</i>	YdfQ

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2678	Hypothetical Membrane Spanning Protein [<i>Bacillus cereus</i>]	UniRef100_Q813Y5	<i>Bacillus cereus</i>	YrkJ
2679	Hypothetical UPF0033 protein yrkI [<i>Bacillus subtilis</i>]	UniRef100_P54436	<i>Bacillus subtilis</i>	
2680	UPI00003CB3C6 UniRef100 entry	UniRef100_UPI00003CB3C6		YrkH
2681	Molybdopterin biosynthesis MoeB protein [<i>Bacillus cereus</i>]	UniRef100_Q81HL2	<i>Bacillus cereus</i>	YrkF
2682	Hypothetical protein yrkE [<i>Bacillus subtilis</i>]	UniRef100_P54432	<i>Bacillus subtilis</i>	YrkE
2683	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EN37	<i>Oceanobacillus iheyensis</i>	
2684	S-adenosylmethionine-dependent methyltransferase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97FB3	<i>Clostridium acetobutylicum</i>	YcgJ
2685	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81N81	<i>Bacillus anthracis</i>	
2686	YciB protein [<i>Bacillus subtilis</i>]	UniRef100_P94399	<i>Bacillus subtilis</i>	
2687	Acetylxyylan esterase related enzyme [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97LM8	<i>Clostridium acetobutylicum</i>	
2688	Hypothetical UPF0161 protein BCE4947 [<i>Bacillus cereus</i>]	UniRef100_P61464	<i>Bacillus cereus</i>	
2689	Delta-aminolevulinic acid dehydratase [<i>Bacillus halodurans</i>]	UniRef100_Q9K8G2	<i>Bacillus halodurans</i>	
2690	6-phospho-3-hexuloisomerase [<i>Bacillus methanolicus</i>]	UniRef100_Q6TV53	<i>Bacillus methanolicus</i>	HxlB
2691	Probable hexulose-6-phosphate synthase [<i>Bacillus subtilis</i>]	UniRef100_P42405	<i>Bacillus subtilis</i>	HxlA
2692	Transcriptional regulator [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q70KJ9	<i>Bacillus amyloliquefaciens</i>	HxlR
2693	Fatty acid desaturase [<i>Bacillus subtilis</i>]	UniRef100_O34653	<i>Bacillus subtilis</i>	Des
2694	Sensor kinase [<i>Bacillus subtilis</i>]	UniRef100_O34757	<i>Bacillus subtilis</i>	YocF
2695	Sensor regulator [<i>Bacillus subtilis</i>]	UniRef100_O34723	<i>Bacillus subtilis</i>	YocG
2696	UPI00003CC1E4 UniRef100 entry	UniRef100_UPI00003CC1E4		YcgT
2697	Nickel transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB8	<i>Bacillus halodurans</i>	AppA
2698	Nickel transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB7	<i>Bacillus halodurans</i>	AppB
2699	Nickel transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB6	<i>Bacillus halodurans</i>	AppC
2700	Oligopeptide ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB5	<i>Bacillus halodurans</i>	DppD
2701	Oligopeptide ABC transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB4	<i>Bacillus halodurans</i>	AppF
2702	UPI00003CB880 UniRef100 entry	UniRef100_UPI00003CB880		YdfL
2703	UPI00003CA374 UniRef100 entry	UniRef100_UPI00003CA374		YoeA
2704	PROBABLE TRANSCRIPTION REGULATORY PROTEIN [<i>Ralstonia solanacearum</i>]	UniRef100_Q8XS91	<i>Ralstonia solanacearum</i>	
2705	Short chain dehydrogenase family protein [<i>Enterococcus faecalis</i>]	UniRef100_Q83415	<i>Enterococcus faecalis</i>	YvaG
2706	Uncharacterized protein, containing predicted phosphatase domain [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97L24	<i>Clostridium acetobutylicum</i>	
2707	UPI000025758C UniRef100 entry	UniRef100_UPI000025758C		PnbA
2708	Cytochrome P450 [<i>Bacillus subtilis</i>]	UniRef100_O08469	<i>Bacillus subtilis</i>	CypA
2709	YtnM [<i>Bacillus subtilis</i>]	UniRef100_O34430	<i>Bacillus subtilis</i>	YtnM
2710	Hypothetical protein yndA precursor [<i>Bacillus subtilis</i>]	UniRef100_O31805	<i>Bacillus subtilis</i>	YndA
2711	YvaG protein [<i>Bacillus subtilis</i>]	UniRef100_O32229	<i>Bacillus subtilis</i>	YvaG
2712	Levanase precursor [<i>Bacillus subtilis</i>]	UniRef100_P05656	<i>Bacillus subtilis</i>	SacC
2713	PTS system, fructose-specific IID component [<i>Bacillus subtilis</i>]	UniRef100_P26382	<i>Bacillus subtilis</i>	LevG
2714	PTS system, fructose-specific IIC component [<i>Bacillus subtilis</i>]	UniRef100_P26381	<i>Bacillus subtilis</i>	LevF
2715	PTS system, fructose-specific IIB component [<i>Bacillus subtilis</i>]	UniRef100_P26380	<i>Bacillus subtilis</i>	LevE
2716	PTS system, fructose-specific IIA component [<i>Bacillus subtilis</i>]	UniRef100_P26379	<i>Bacillus subtilis</i>	LevD

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2717	Transcriptional regulatory protein levR [<i>Bacillus subtilis</i>]	UniRef100_P23914	<i>Bacillus subtilis</i>	LevR
2718	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q734Q5	<i>Bacillus cereus</i>	
2719	Hypothetical protein yrhK [<i>Bacillus subtilis</i>]	UniRef100_O05401	<i>Bacillus subtilis</i>	
2720	UPI00003CB785 UniRef100 entry	UniRef100_UPI00003CB785		
2721				AdaA
2722	Methylated-DNA--protein-cysteine S-methyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q732Y7	<i>Bacillus cereus</i>	AdaB
2723	Oxidoreductase, aldo/keto reductase family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HBJ5	<i>Bacillus thuringiensis</i>	YtbE
2724				YtbD
2725	Hypothetical UPF0087 protein ytcD [<i>Bacillus subtilis</i>]	UniRef100_O34533	<i>Bacillus subtilis</i>	YtcD
2726				
2727	Hypothetical protein yjjA [<i>Bacillus subtilis</i>]	UniRef100_O34394	<i>Bacillus subtilis</i>	YjjA
2728	UPI000028298B UniRef100 entry	UniRef100_UPI000028298B		
2729				YxeB
2730	Putative HTH-type transcriptional regulator yybE [<i>Bacillus subtilis</i>]	UniRef100_P37499	<i>Bacillus subtilis</i>	YybE
2731	Hypothetical transport protein yybF [<i>Bacillus subtilis</i>]	UniRef100_P37498	<i>Bacillus subtilis</i>	YybF
2732	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q732J0	<i>Bacillus cereus</i>	
2733	Probable bifunctional P-450:NADPH-P450 reductase 2 [Includes: Cytochrome P450 102 (EC 1.14.14.1); NADPH--cytochrome P450 reductase (EC 1.6.2.4)] [<i>Bacillus subtilis</i>]	UniRef100_O08336	Includes: Cytochrome P450 102 (EC 1.14.14.1); NADPH--cytochrome P450 reductase (EC 1.6.2.4)	YrhJ
2734	Regulatory protein [<i>Bacillus subtilis</i>]	UniRef100_O08335	<i>Bacillus subtilis</i>	YrhI
2735				WprA
2736	YrhH [<i>Bacillus subtilis</i>]	UniRef100_O05400	<i>Bacillus subtilis</i>	YrhH
2737				
2738				
2739				
2740	Cystathionine gamma-lyase [<i>Bacillus subtilis</i>]	UniRef100_O05394	<i>Bacillus subtilis</i>	YrhB
2741	Cysteine synthase [<i>Bacillus subtilis</i>]	UniRef100_O05393	<i>Bacillus subtilis</i>	YrhA
2742	MTA/SAH nucleosidase [<i>Bacillus subtilis</i>]	UniRef100_O32028	<i>Bacillus subtilis</i>	Mtn
2743	YrrT protein [<i>Bacillus subtilis</i>]	UniRef100_O32029	<i>Bacillus subtilis</i>	YrrT
2744	Hypothetical protein yrzA [<i>Bacillus subtilis</i>]	UniRef100_O32030	<i>Bacillus subtilis</i>	
2745				YrrS
2746	YrrR protein [<i>Bacillus subtilis</i>]	UniRef100_O32032	<i>Bacillus subtilis</i>	YrrR
2747	Transcription elongation factor greA [<i>Bacillus subtilis</i>]	UniRef100_P80240	<i>Bacillus subtilis</i>	GreA
2748	Uridine kinase [<i>Bacillus subtilis</i>]	UniRef100_O32033	<i>Bacillus subtilis</i>	Udk
2749	YrrO protein [<i>Bacillus subtilis</i>]	UniRef100_O32034	<i>Bacillus subtilis</i>	YrrO
2750	YrrN protein [<i>Bacillus subtilis</i>]	UniRef100_O32035	<i>Bacillus subtilis</i>	YrrN
2751	YrrM protein [<i>Bacillus subtilis</i>]	UniRef100_O32036	<i>Bacillus subtilis</i>	YrrM
2752	YrrL protein [<i>Bacillus subtilis</i>]	UniRef100_O34758	<i>Bacillus subtilis</i>	YrrL
2753	YrzB protein [<i>Bacillus subtilis</i>]	UniRef100_O34828	<i>Bacillus subtilis</i>	
2754	Putative Holliday junction resolvase [<i>Bacillus subtilis</i>]	UniRef100_O34634	<i>Bacillus subtilis</i>	YrrK
2755	Hypothetical UPF0297 protein yrzL [<i>Bacillus subtilis</i>]	UniRef100_Q7WY61	<i>Bacillus subtilis</i>	
2756	Alanyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_O34526	<i>Bacillus subtilis</i>	AlaS
2757	Hypothetical UPF0118 protein yrrI [<i>Bacillus subtilis</i>]	UniRef100_O34472	<i>Bacillus subtilis</i>	YrrI
2758				
2759	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q634F2	<i>Bacillus cereus</i> ZK	
2760	YrrD protein [<i>Bacillus subtilis</i>]	UniRef100_O34402	<i>Bacillus subtilis</i>	YrrD

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2761	YrrC protein [<i>Bacillus subtilis</i>]	UniRef100_O34481	<i>Bacillus subtilis</i>	YrrC
2762	YrrB protein [<i>Bacillus subtilis</i>]	UniRef100_O34452	<i>Bacillus subtilis</i>	YrrB
2763	Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O35020	<i>Bacillus subtilis</i>	TmU
2764	YrvO protein [<i>Bacillus subtilis</i>]	UniRef100_O34599	<i>Bacillus subtilis</i>	YrvO
2765	BH1259 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KDF4	<i>Bacillus halodurans</i>	
2766	YrvN protein [<i>Bacillus subtilis</i>]	UniRef100_O34528	<i>Bacillus subtilis</i>	YrvN
2767	YrvM protein [<i>Bacillus subtilis</i>]	UniRef100_O32037	<i>Bacillus subtilis</i>	YrvM
2768	Aspartyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_O32038	<i>Bacillus subtilis</i>	AspS
2769	Histidyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_O32039	<i>Bacillus subtilis</i>	HisS
2770				
2771				YrvJ
2772	Putative D-tyrosyl-tRNA(Tyr) deacylase-like protein [<i>Bacillus subtilis</i>]	UniRef100_O32042	<i>Bacillus subtilis</i>	YrvI
2773	GTP pyrophosphokinase (EC 2.7.6.5) (ATP:GTP 3'-pyrophosphotransferase) (ppGpp synthetase I) ((P)ppGpp synthetase) [<i>Bacillus subtilis</i>]	UniRef100_O54408	<i>Bacillus subtilis</i>	RelA
2774	Adenine phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34443	<i>Bacillus subtilis</i>	Apt
2775	YrvE protein [<i>Bacillus subtilis</i>]	UniRef100_O32044	<i>Bacillus subtilis</i>	YrvE
2776	YrvD protein [<i>Bacillus subtilis</i>]	UniRef100_O32045	<i>Bacillus subtilis</i>	
2777	SecDF protein [<i>Bacillus subtilis</i>]	UniRef100_O32047	<i>Bacillus subtilis</i>	SecDF
2778	YrzD protein [<i>Bacillus subtilis</i>]	UniRef100_O32049	<i>Bacillus subtilis</i>	
2779				SpoVB
2780	Hypothetical protein yrbG [<i>Bacillus subtilis</i>]	UniRef100_O32050	<i>Bacillus subtilis</i>	YrbG
2781	YrzE protein [<i>Bacillus subtilis</i>]	UniRef100_O32051	<i>Bacillus subtilis</i>	YrzE
2782	Hypothetical UPF0092 protein yrbF [<i>Bacillus subtilis</i>]	UniRef100_O32052	<i>Bacillus subtilis</i>	
2783	Queuine tRNA-ribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O32053	<i>Bacillus subtilis</i>	Tgt
2784	S-adenosylmethionine:tRNA ribosyltransferase-isomerase [<i>Bacillus subtilis</i>]	UniRef100_O32054	<i>Bacillus subtilis</i>	QueA
2785				
2786	Holliday junction DNA helicase ruvB [<i>Bacillus subtilis</i>]	UniRef100_O32055	<i>Bacillus subtilis</i>	RuvB
2787	Holliday junction DNA helicase ruvA [<i>Bacillus subtilis</i>]	UniRef100_O05392	<i>Bacillus subtilis</i>	RuvA
2788	BofC protein precursor [<i>Bacillus subtilis</i>]	UniRef100_O05391	<i>Bacillus subtilis</i>	BofC
2789	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ERL7	<i>Oceanobacillus iheyensis</i>	YrzF
2790				
2791	Small, acid-soluble spore protein H [<i>Bacillus halodurans</i>]	UniRef100_Q9KB75	<i>Bacillus halodurans</i>	
2792	Hypothetical protein yjoA [<i>Bacillus subtilis</i>]	UniRef100_O34334	<i>Bacillus subtilis</i>	YjoA
2793	YmaC protein [<i>Bacillus subtilis</i>]	UniRef100_O31789	<i>Bacillus subtilis</i>	YmaC
2794				
2795				
2796	Hypothetical UPF0082 protein yrbC [<i>Bacillus subtilis</i>]	UniRef100_P94447	<i>Bacillus subtilis</i>	YrbC
2797	Sporulation cortex protein coxA [<i>Bacillus subtilis</i>]	UniRef100_P94446	<i>Bacillus subtilis</i>	CoxA
2798	Morphogenetic protein associated with SpoVID [<i>Bacillus subtilis</i>]	UniRef100_O32062	<i>Bacillus subtilis</i>	SafA
2799	Quinolinolate synthetase A [<i>Bacillus subtilis</i>]	UniRef100_Q9KWZ1	<i>Bacillus subtilis</i>	NadA
2800	Probable nicotinate-nucleotide pyrophosphorylase [carboxylating] [<i>Bacillus subtilis</i>]	UniRef100_P39666	carboxylating	NadC
2801	L-aspartate oxidase [<i>Bacillus subtilis</i>]	UniRef100_P38032	<i>Bacillus subtilis</i>	NadB
2802	Probable cysteine desulfurase [<i>Bacillus subtilis</i>]	UniRef100_P38033	<i>Bacillus subtilis</i>	NifS

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
2803				YrxA
2804	Prephenate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P21203	<i>Bacillus subtilis</i>	PheA
2805	ACT domain protein pheB [<i>Bacillus subtilis</i>]	UniRef100_P21204	<i>Bacillus subtilis</i>	PheB
2806	Spo0B-associated GTP-binding protein [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q659J4	<i>Bacillus amyloliquefaciens</i>	Obg
2807	Sporulation initiation phosphotransferase B [<i>Bacillus subtilis</i>]	UniRef100_P06535	<i>Bacillus subtilis</i>	Spo0B
2808	50S ribosomal protein L27 [<i>Bacillus subtilis</i>]	UniRef100_P05657	<i>Bacillus subtilis</i>	
2809				
2810	50S ribosomal protein L21 [<i>Bacillus subtilis</i>]	UniRef100_P26908	<i>Bacillus subtilis</i>	RplU
2811	Stage IV sporulation protein FB [<i>Bacillus subtilis</i>]	UniRef100_P26937	<i>Bacillus subtilis</i>	SpoIVFB
2812	Stage IV sporulation protein FA [<i>Bacillus subtilis</i>]	UniRef100_P26936	<i>Bacillus subtilis</i>	SpoIVFA
2813	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q816V6	<i>Bacillus cereus</i>	YndB
2814	Transcriptional regulator, ArsR family [<i>Bacillus cereus</i> ZK]	UniRef100_Q632Y0	<i>Bacillus cereus</i> ZK	
2815	Septum site-determining protein minD [<i>Bacillus subtilis</i>]	UniRef100_Q01464	<i>Bacillus subtilis</i>	MinD
2816				MinC
2817	Rod shape-determining protein mreD [<i>Bacillus subtilis</i>]	UniRef100_Q01467	<i>Bacillus subtilis</i>	MreD
2818	Rod shape-determining protein mreC [<i>Bacillus subtilis</i>]	UniRef100_Q01466	<i>Bacillus subtilis</i>	MreC
2819	Rod shape-determining protein mreB [<i>Bacillus subtilis</i>]	UniRef100_Q01465	<i>Bacillus subtilis</i>	MreB
2820	DNA repair protein radC homolog [<i>Bacillus subtilis</i>]	UniRef100_Q02170	<i>Bacillus subtilis</i>	RadC
2821	Septum formation protein Maf [<i>Bacillus subtilis</i>]	UniRef100_Q02169	<i>Bacillus subtilis</i>	Maf
2822	Stage II sporulation protein B [<i>Bacillus subtilis</i>]	UniRef100_P37575	<i>Bacillus subtilis</i>	SpoIIB
2823	Type 4 prepilin-like proteins leader peptide processing enzyme (Late competence protein comC) [Includes: Leader peptidase (EC 3.4.23.43) (Prepilin peptidase); N-methyltransferase (EC 2.1.1.—)] [<i>Bacillus subtilis</i>]	UniRef100_P15378	Includes: Leader peptidase (EC 3.4.23.43) (Prepilin peptidase); N-methyltransferase (EC 2.1.1.—)	ComC
2824				FolC
2825	Valyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_Q05873	<i>Bacillus subtilis</i>	ValS
2826	Hypothetical protein OB2062 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPN1	<i>Oceanobacillus iheyensis</i>	
2827	Hypothetical protein ysxE [<i>Bacillus subtilis</i>]	UniRef100_P37964	<i>Bacillus subtilis</i>	YsxE
2828	Stage VI sporulation protein D [<i>Bacillus subtilis</i>]	UniRef100_P37963	<i>Bacillus subtilis</i>	SpoVID
2829	Glutamate-1-semialdehyde 2,1-aminomutase [<i>Bacillus subtilis</i>]	UniRef100_P30949	<i>Bacillus subtilis</i>	HemL
2830	Delta-aminolevulinic acid dehydratase [<i>Bacillus subtilis</i>]	UniRef100_P30950	<i>Bacillus subtilis</i>	HemB
2831	Uroporphyrinogen-III synthase [<i>Bacillus subtilis</i>]	UniRef100_P21248	<i>Bacillus subtilis</i>	HemD
2832	Porphobilinogen deaminase [<i>Bacillus subtilis</i>]	UniRef100_P16616	<i>Bacillus subtilis</i>	HemC
2833	Protein hemX [<i>Bacillus subtilis</i>]	UniRef100_P16645	<i>Bacillus subtilis</i>	HemX
2834	Glutamyl-tRNA reductase [<i>Bacillus subtilis</i>]	UniRef100_P16618	<i>Bacillus subtilis</i>	HemA
2835	Hypothetical protein ysxD [<i>Bacillus subtilis</i>]	UniRef100_P40736	<i>Bacillus subtilis</i>	YsxD
2836	Probable GTP-binding protein engB [<i>Bacillus subtilis</i>]	UniRef100_P38424	<i>Bacillus subtilis</i>	YsxC
2837	ATP-dependent protease La 1 [<i>Bacillus subtilis</i>]	UniRef100_P37945	<i>Bacillus subtilis</i>	LonA

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2838	ATP-dependent protease La homolog [<i>Bacillus subtilis</i>]	UniRef100_P42425	<i>Bacillus subtilis</i>	LonB
2839	ATP-dependent Clp protease ATP- binding subunit clpX [<i>Bacillus subtilis</i>]	UniRef100_P50866	<i>Bacillus subtilis</i>	ClpX
2840	Trigger factor [<i>Bacillus subtilis</i>]	UniRef100_P80698	<i>Bacillus subtilis</i>	Tig
2841	Hypothetical protein ysoA [<i>Bacillus subtilis</i>]	UniRef100_P94569	<i>Bacillus subtilis</i>	YsoA
2842	3-isopropylmalate dehydratase small subunit [<i>Bacillus subtilis</i>]	UniRef100_P94568	<i>Bacillus subtilis</i>	LeuD
2843	3-isopropylmalate dehydratase large subunit [<i>Bacillus subtilis</i>]	UniRef100_P80858	<i>Bacillus subtilis</i>	LeuC
2844				LeuB
2845	2-isopropylmalate synthase [<i>Bacillus subtilis</i>]	UniRef100_P94565	<i>Bacillus subtilis</i>	LeuA
2846	Ketol-acid reductoisomerase [<i>Bacillus subtilis</i>]	UniRef100_P37253	<i>Bacillus subtilis</i>	IlvC
2847	Acetolactate synthase small subunit [<i>Bacillus subtilis</i>]	UniRef100_P37252	<i>Bacillus subtilis</i>	IlvH
2848	Acetolactate synthase large subunit [<i>Bacillus subtilis</i>]	UniRef100_P37251	<i>Bacillus subtilis</i>	IlvB
2849	Branched-chain amino acid aminotransferase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HLF7	<i>Bacillus thuringiensis</i>	Dat
2850				RocG
2851				YxeD
2852	BH3337 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K7M4	<i>Bacillus halodurans</i>	
2853				YqaT
2854	Hypothetical protein yqaA [<i>Bacillus subtilis</i>]	UniRef100_P45917	<i>Bacillus subtilis</i>	
2855	Hypothetical protein yqaT [<i>Bacillus subtilis</i>]	UniRef100_P45916	<i>Bacillus subtilis</i>	
2856	Lin1266 protein [<i>Listeria innocua</i>]	UniRef100_Q92CC3	<i>Listeria innocua</i>	
2857	Lin1733 protein [<i>Listeria innocua</i>]	UniRef100_Q92B18	<i>Listeria innocua</i>	
2858				
2859				
2860				
2861				MtbP
2862	BH3535 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K738	<i>Bacillus halodurans</i>	
2863				YeeF
2864				YxiD
2865				RapI
2866				
2867				
2868				
2869	Hypothetical UPF0025 protein ysnB [<i>Bacillus subtilis</i>]	UniRef100_P94559	<i>Bacillus subtilis</i>	YsnB
2870	HAM1 protein homolog [<i>Bacillus subtilis</i>]	UniRef100_P94558	<i>Bacillus subtilis</i>	YsnA
2871				Rph
2872	Germination protein gerM [<i>Bacillus subtilis</i>]	UniRef100_P39072	<i>Bacillus subtilis</i>	GerM
2873				RacE
2874	Hypothetical protein ysmB [<i>Bacillus subtilis</i>]	UniRef100_P97247	<i>Bacillus subtilis</i>	YsmB
2875	Germination protein gerE [<i>Bacillus subtilis</i>]	UniRef100_P11470	<i>Bacillus subtilis</i>	
2876	Oxidoreductase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97TP7	<i>Clostridium acetobutylicum</i>	YqjQ
2877	Hypothetical protein ysmA [<i>Bacillus subtilis</i>]	UniRef100_Q6L874	<i>Bacillus subtilis</i>	YsmA
2878	Succinate dehydrogenase iron-sulfur protein [<i>Bacillus subtilis</i>]	UniRef100_P08066	<i>Bacillus subtilis</i>	SdhB
2879				SdhA
2880				
2881	Succinate dehydrogenase cytochrome B-558 subunit [<i>Bacillus subtilis</i>]	UniRef100_P08064	<i>Bacillus subtilis</i>	SdhC
2882	Hypothetical protein yslB [<i>Bacillus subtilis</i>]	UniRef100_P42955	<i>Bacillus subtilis</i>	YslB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2883	Aspartokinase 2 (EC 2.7.2.4) (Aspartokinase II) (Aspartate kinase 2) [Contains: Aspartokinase II alpha subunit; Aspartokinase II beta subunit] [<i>Bacillus subtilis</i>]	UniRef100_P08495	Contains: Aspartokinase II alpha subunit; Aspartokinase II beta subunit	LysC
2884	UvrABC system protein C [<i>Bacillus subtilis</i>]	UniRef100_P14951	<i>Bacillus subtilis</i>	UvrC
2885	Thioredoxin [<i>Bacillus subtilis</i>]	UniRef100_P14949	<i>Bacillus subtilis</i>	TrxA
2886	Electron transfer flavoprotein alpha-subunit [<i>Bacillus subtilis</i>]	UniRef100_P94551	<i>Bacillus subtilis</i>	EtfA
2887	Electron transfer flavoprotein beta-subunit [<i>Bacillus subtilis</i>]	UniRef100_P94550	<i>Bacillus subtilis</i>	EtfB
2888	Hypothetical protein ysiB [<i>Bacillus subtilis</i>]	UniRef100_P94549	<i>Bacillus subtilis</i>	YsiB
2889	Hypothetical protein ysiA [<i>Bacillus subtilis</i>]	UniRef100_P94548	<i>Bacillus subtilis</i>	YsiA
2890	Long-chain-fatty-acid--CoA ligase [<i>Bacillus subtilis</i>]	UniRef100_P94547	<i>Bacillus subtilis</i>	LcfA
2891	Hypothetical protein yshE [<i>Bacillus subtilis</i>]	UniRef100_P94546	<i>Bacillus subtilis</i>	YshE
2892	MutS2 protein [<i>Bacillus subtilis</i>]	UniRef100_P94545	<i>Bacillus subtilis</i>	MutSB
2893	Hypothetical protein yshC [<i>Bacillus subtilis</i>]	UniRef100_P94544	<i>Bacillus subtilis</i>	YshC
2894				YshB
2895				
2896	Ribonuclease HIII [<i>Bacillus subtilis</i>]	UniRef100_P94541	<i>Bacillus subtilis</i>	RnhC
2897				
2898				
2899				
2900				
2901				YxIF
2902				
2903	Phenylalanyl-tRNA synthetase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P17922	<i>Bacillus subtilis</i>	PheT
2904	Phenylalanyl-tRNA synthetase alpha chain [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q659J3	<i>Bacillus amyloliquefaciens</i>	PheS
2905	Hypothetical protein ysgA [<i>Bacillus subtilis</i>]	UniRef100_P94538	<i>Bacillus subtilis</i>	YsgA
2906	Small, acid-soluble spore protein I [<i>Bacillus subtilis</i>]	UniRef100_P94537	<i>Bacillus subtilis</i>	
2907	Carbon starvation protein A homolog [<i>Bacillus subtilis</i>]	UniRef100_P94532	<i>Bacillus subtilis</i>	CstA
2908	Alpha-N-arabinoxylanase [<i>Bacillus stearothermophilus</i>]	UniRef100_Q9XBQ3	<i>Bacillus stearothermophilus</i>	AbfA
2909	L-arabinose transport system permease protein araQ [<i>Bacillus subtilis</i>]	UniRef100_P94530	<i>Bacillus subtilis</i>	AraQ
2910	L-arabinose transport system permease protein araP [<i>Bacillus subtilis</i>]	UniRef100_P94529	<i>Bacillus subtilis</i>	AraP
2911	Probable arabinose-binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P94528	<i>Bacillus subtilis</i>	AraN
2912	Arabinose operon protein araM [<i>Bacillus subtilis</i>]	UniRef100_P94527	<i>Bacillus subtilis</i>	AraM
2913	L-ribulose-5-phosphate 4-epimerase [<i>Bacillus subtilis</i>]	UniRef100_P94525	<i>Bacillus subtilis</i>	AraD
2914	Ribulokinase [<i>Bacillus subtilis</i>]	UniRef100_P94524	<i>Bacillus subtilis</i>	AraB
2915	L-arabinose isomerase [<i>Bacillus subtilis</i>]	UniRef100_P94523	<i>Bacillus subtilis</i>	AraA
2916				AbnA
2917	Hypothetical protein ysdC [<i>Bacillus subtilis</i>]	UniRef100_P94521	<i>Bacillus subtilis</i>	YsdC
2918	Hypothetical protein ysdB [<i>Bacillus subtilis</i>]	UniRef100_P94520	<i>Bacillus subtilis</i>	YsdB
2919	Hypothetical protein ysdA [<i>Bacillus subtilis</i>]	UniRef100_P94519	<i>Bacillus subtilis</i>	
2920	50S ribosomal protein L20 [<i>Bacillus subtilis</i>]	UniRef100_P55873	<i>Bacillus subtilis</i>	RplT
2921	50S ribosomal protein L35 [<i>Bacillus subtilis</i>]	UniRef100_P55874	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
2922				InfC
2923	Antiholin-like protein IrgB [<i>Bacillus subtilis</i>]	UniRef100_P94516	<i>Bacillus subtilis</i>	YsbB
2924	Antiholin-like protein IrgA [<i>Bacillus subtilis</i>]	UniRef100_P94515	<i>Bacillus subtilis</i>	YsbA
2925				PhoD
2926	Sensory transduction protein lytT [<i>Bacillus subtilis</i>]	UniRef100_P94514	<i>Bacillus subtilis</i>	LytT
2927	Sensor protein lytS [<i>Bacillus subtilis</i>]	UniRef100_P94513	<i>Bacillus subtilis</i>	LytS
2928	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81N00	<i>Bacillus anthracis</i>	
2929	Hypothetical protein ysaA [<i>Bacillus subtilis</i>]	UniRef100_P94512	<i>Bacillus subtilis</i>	YsaA
2930	Threonyl-tRNA synthetase 1 [<i>Bacillus subtilis</i>]	UniRef100_P18255	<i>Bacillus subtilis</i>	ThrS
2931	Hypothetical protein ytxC [<i>Bacillus subtilis</i>]	UniRef100_P06569	<i>Bacillus subtilis</i>	YtxC
2932	Hypothetical UPF0043 protein ytxB [<i>Bacillus subtilis</i>]	UniRef100_P06568	<i>Bacillus subtilis</i>	YtxB
2933	Primosomal protein dnaI [<i>Bacillus subtilis</i>]	UniRef100_P06567	<i>Bacillus subtilis</i>	DnaI
2934	Replication initiation and membrane attachment protein [<i>Bacillus subtilis</i>]	UniRef100_P07908	<i>Bacillus subtilis</i>	DnaB
2935	Hypothetical UPF0168 protein ytcG [<i>Bacillus subtilis</i>]	UniRef100_Q45549	<i>Bacillus subtilis</i>	YtcG
2936				SpeD
2937				GapB
2938				PeIB
2939	Pectin lyase [<i>Bacillus subtilis</i>]	UniRef100_P94449	<i>Bacillus subtilis</i>	YtaG
2940	Dephospho-CoA kinase [<i>Bacillus subtilis</i>]	UniRef100_O34932	<i>Bacillus subtilis</i>	
2941				YtaF
2942	Formamidopyrimidine-DNA glycosylase [<i>Bacillus subtilis</i>]	UniRef100_O34403	<i>Bacillus subtilis</i>	MutM
2943	DNA polymerase I [<i>Bacillus subtilis</i>]	UniRef100_O34996	<i>Bacillus subtilis</i>	PoIA
2944	Alkaline phosphatase synthesis sensor protein phoR [<i>Bacillus subtilis</i>]	UniRef100_P23545	<i>Bacillus subtilis</i>	PhoR
2945	Alkaline phosphatase synthesis transcriptional regulatory protein phoP [<i>Bacillus subtilis</i>]	UniRef100_P13792	<i>Bacillus subtilis</i>	PhoP
2946	Malate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P49814	<i>Bacillus subtilis</i>	Mdh
2947	Isocitrate dehydrogenase [NADP] [<i>Bacillus subtilis</i>]	UniRef100_P39126	NADP	Icd
2948	Citrate synthase II [<i>Bacillus subtilis</i>]	UniRef100_P39120	<i>Bacillus subtilis</i>	CitZ
2949	YtwI [<i>Bacillus subtilis</i>]	UniRef100_O34811	<i>Bacillus subtilis</i>	YtwI
2950	Hypothetical UPF0118 protein ytvI [<i>Bacillus subtilis</i>]	UniRef100_O34991	<i>Bacillus subtilis</i>	YtvI
2951	YtzA protein [<i>Bacillus subtilis</i>]	UniRef100_O32064	<i>Bacillus subtilis</i>	YtzA
2952				Pyk
2953	6-phosphofructokinase [<i>Bacillus subtilis</i>]	UniRef100_O34529	<i>Bacillus subtilis</i>	PfkA
2954	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha [<i>Bacillus subtilis</i>]	UniRef100_O34847	<i>Bacillus subtilis</i>	AccA
2955	Acetyl-CoA carboxylase subunit [<i>Bacillus subtilis</i>]	UniRef100_O34571	<i>Bacillus subtilis</i>	AccD
2956				YtsJ
2957	DNA polymerase III alpha subunit [<i>Bacillus subtilis</i>]	UniRef100_O34623	<i>Bacillus subtilis</i>	DnaE
2958	Hypothetical Membrane Spanning Protein [<i>Bacillus cereus</i>]	UniRef100_Q812P3	<i>Bacillus cereus</i>	
2959	YtrI [<i>Bacillus subtilis</i>]	UniRef100_O34460	<i>Bacillus subtilis</i>	YtrI
2960	BH3172 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K835	<i>Bacillus halodurans</i>	
2961	YtqI [<i>Bacillus subtilis</i>]	UniRef100_O34600	<i>Bacillus subtilis</i>	YtqI
2962	YtPI [<i>Bacillus subtilis</i>]	UniRef100_O34922	<i>Bacillus subtilis</i>	
2963				YtoI
2964				PadR
2965	YtkL protein [<i>Bacillus subtilis</i>]	UniRef100_Q795U4	<i>Bacillus subtilis</i>	YtkL

TABLE 1-continued

		Predicted functions			
SEQ ID NO.	Description	UniRef Accession No.	Organism		<i>Bacillus subtilis</i> homolog (Gene Name)
2966					YtkK
2967					
2968	Argininosuccinate lyase [<i>Bacillus subtilis</i>]	UniRef100_O34858	<i>Bacillus subtilis</i>		ArgH
2969	Argininosuccinate synthase [<i>Bacillus subtilis</i>]	UniRef100_O34347	<i>Bacillus subtilis</i>		ArgG
2970	Molybdenum cofactor biosynthesis protein B [<i>Bacillus subtilis</i>]	UniRef100_O34457	<i>Bacillus subtilis</i>		MoaB
2971					AckA
2972	Hypothetical protein ytxK [<i>Bacillus subtilis</i>]	UniRef100_P37876	<i>Bacillus subtilis</i>		YtxK
2973	Probable thiol peroxidase [<i>Bacillus subtilis</i>]	UniRef100_P80864	<i>Bacillus subtilis</i>		Tpx
2974	YtfJ [<i>Bacillus subtilis</i>]	UniRef100_O34806	<i>Bacillus subtilis</i>		YtfJ
2975					YtfI
2976	YteJ [<i>Bacillus subtilis</i>]	UniRef100_O34424	<i>Bacillus subtilis</i>		YteJ
2977	Putative signal peptide peptidase sppA [<i>Bacillus subtilis</i>]	UniRef100_O34525	<i>Bacillus subtilis</i>		SppA
2978	Probable inorganic polyphosphate/ATP-NAD kinase 2 (EC 2.7.1.23) (Poly(P)/ATP NAD kinase 2) [<i>Bacillus subtilis</i>]	UniRef100_O34934	<i>Bacillus subtilis</i>		YtdI
2979	YhbJ protein [<i>Bacillus subtilis</i>]	UniRef100_O31593	<i>Bacillus subtilis</i>		YhbJ
2980	Multidrug resistance protein [<i>Staphylococcus epidermidis</i>]	UniRef100_Q8CQB1	<i>Staphylococcus epidermidis</i>		YubD
2981	Putative HTH-type transcriptional regulator yxaD [<i>Bacillus subtilis</i>]	UniRef100_P42103	<i>Bacillus subtilis</i>		YxaD
2982	YtcI [<i>Bacillus subtilis</i>]	UniRef100_O34613	<i>Bacillus subtilis</i>		YtcI
2983	Small, acid-soluble spore protein 1 [<i>Bacillus stearothermophilus</i>]	UniRef100_P06552	<i>Bacillus stearothermophilus</i>		
2984	Probable thiamine biosynthesis protein thil [<i>Bacillus subtilis</i>]	UniRef100_O34595	<i>Bacillus subtilis</i>		YtbJ
2985	NifS2 [<i>Bacillus subtilis</i>]	UniRef100_O34874	<i>Bacillus subtilis</i>		NifZ
2986	Branched-chain amino acid transport system carrier protein braB [<i>Bacillus subtilis</i>]	UniRef100_O34545	<i>Bacillus subtilis</i>		BraB
2987	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012		
2988	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069			
2989	Septation ring formation regulator ezrA [<i>Bacillus subtilis</i>]	UniRef100_O34894	<i>Bacillus subtilis</i>		EzrA
2990	Histidinol-phosphatase [<i>Bacillus subtilis</i>]	UniRef100_O34411	<i>Bacillus subtilis</i>		HisJ
2991	Probable HTH-type transcriptional regulator yttP [<i>Bacillus subtilis</i>]	UniRef100_O34970	<i>Bacillus subtilis</i>		YttP
2992	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPB0	<i>Oceanobacillus iheyensis</i>		
2993	YtrP [<i>Bacillus subtilis</i>]	UniRef100_O34325	<i>Bacillus subtilis</i>		YtrP
2994	30S ribosomal protein S4 [<i>Bacillus subtilis</i>]	UniRef100_P21466	<i>Bacillus subtilis</i>		RpsD
2995					
2996					
2997					YddR
2998	HTH-type transcriptional regulator lrpA [<i>Bacillus subtilis</i>]	UniRef100_P96652	<i>Bacillus subtilis</i>		LrpA
2999					
3000	Tyrosyl-tRNA synthetase 1 [<i>Bacillus subtilis</i>]	UniRef100_P22326	<i>Bacillus subtilis</i>		TyrS
3001	Acetyl-coenzyme A synthetase [<i>Bacillus subtilis</i>]	UniRef100_P39062	<i>Bacillus subtilis</i>		AcsA
3002	Acetoin utilization protein acuA [<i>Bacillus subtilis</i>]	UniRef100_P39065	<i>Bacillus subtilis</i>		AcuA
3003	Acetoin utilization acuB protein [<i>Bacillus subtilis</i>]	UniRef100_P39066	<i>Bacillus subtilis</i>		AcuB
3004	Acetoin utilization protein acuC [<i>Bacillus subtilis</i>]	UniRef100_P39067	<i>Bacillus subtilis</i>		AcuC
3005	Hypothetical protein ytxE [<i>Bacillus subtilis</i>]	UniRef100_P39064	<i>Bacillus subtilis</i>		YtxE

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3006	Hypothetical protein ytxD [<i>Bacillus subtilis</i>]	UniRef100_P39063	<i>Bacillus subtilis</i>	YtxD
3007	Catabolite control protein A [<i>Bacillus subtilis</i>]	UniRef100_P25144	<i>Bacillus subtilis</i>	CcpA
3008	AroA(G) protein [Includes: Phospho-2-dehydro-3-deoxyheptonate aldolase (EC 2.5.1.54) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase); Chorismate mutase (EC 5.4.99.5)] [<i>Bacillus subtilis</i>]	UniRef100_P39912	Includes: Phospho-2-dehydro-3-deoxyheptonate aldolase (EC 2.5.1.54) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase); Chorismate mutase (EC 5.4.99.5)	AroA
3009	Similar to hypothetical repeat containing protein [<i>Phototribadus luminescens</i>]	UniRef100_Q7N3B8	<i>Phototribadus luminescens</i>	
3010	Hypothetical protein ytxJ [<i>Bacillus subtilis</i>]	UniRef100_P39914	<i>Bacillus subtilis</i>	YtxJ
3011	Hypothetical protein ytxH [<i>Bacillus subtilis</i>]	UniRef100_P40780	<i>Bacillus subtilis</i>	YtxH
3012	Hypothetical protein ytxG [<i>Bacillus subtilis</i>]	UniRef100_P40779	<i>Bacillus subtilis</i>	YtxG
3013				MurC
3014	YtpT [<i>Bacillus subtilis</i>]	UniRef100_O34749	<i>Bacillus subtilis</i>	YtpT
3015	YtpR [<i>Bacillus subtilis</i>]	UniRef100_O34943	<i>Bacillus subtilis</i>	YtpR
3016	YtpQ [<i>Bacillus subtilis</i>]	UniRef100_O34496	<i>Bacillus subtilis</i>	YtpQ
3017	Putative thioredoxin [<i>Bacillus subtilis</i>]	UniRef100_O34357	<i>Bacillus subtilis</i>	YtpP
3018	YtoQ [<i>Bacillus subtilis</i>]	UniRef100_O34305	<i>Bacillus subtilis</i>	YtoQ
3019	YtoP [<i>Bacillus subtilis</i>]	UniRef100_O34924	<i>Bacillus subtilis</i>	YtoP
3020	YtzB protein [<i>Bacillus subtilis</i>]	UniRef100_O32065	<i>Bacillus subtilis</i>	YtzB
3021	Probable NAD-dependent malic enzyme 3 [<i>Bacillus subtilis</i>]	UniRef100_O34389	<i>Bacillus subtilis</i>	MalS
3022	YtnP [<i>Bacillus subtilis</i>]	UniRef100_O34760	<i>Bacillus subtilis</i>	YtnP
3023	tRNA (guanine-N(7)-methyltransferase (EC 2.1.1.33) (tRNA(m7G46)-methyltransferase) [<i>Bacillus subtilis</i>]	UniRef100_O34522	<i>Bacillus subtilis</i>	YtmQ
3024	YtzH protein [<i>Bacillus subtilis</i>]	UniRef100_O32066	<i>Bacillus subtilis</i>	
3025	YtmP [<i>Bacillus subtilis</i>]	UniRef100_O34935	<i>Bacillus subtilis</i>	YtmP
3026	AmyX protein [<i>Bacillus subtilis</i>]	UniRef100_O34587	<i>Bacillus subtilis</i>	AmyX
3027	YtlR [<i>Bacillus subtilis</i>]	UniRef100_O34799	<i>Bacillus subtilis</i>	YtlR
3028	YtlQ [<i>Bacillus subtilis</i>]	UniRef100_O34471	<i>Bacillus subtilis</i>	YtlQ
3029	Hypothetical UPF0097 protein ytlP [<i>Bacillus subtilis</i>]	UniRef100_O34570	<i>Bacillus subtilis</i>	YtlP
3030	Probable cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfhydrylase) (O-acetylserine (Thiol)-lyase) [<i>Bacillus subtilis</i>]	UniRef100_O34476	<i>Bacillus subtilis</i>	YtkP
3031	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q81BR8	<i>Bacillus cereus</i>	YncE
3032	Putative peptidase [<i>Bacillus subtilis</i>]	UniRef100_O34944	<i>Bacillus subtilis</i>	YtjP
3033	YtiP [<i>Bacillus subtilis</i>]	UniRef100_O34978	<i>Bacillus subtilis</i>	YtiP
3034	YtzE protein [<i>Bacillus subtilis</i>]	UniRef100_O32067	<i>Bacillus subtilis</i>	
3035	Ribosomal small subunit pseudouridine synthase A [<i>Bacillus cereus</i>]	UniRef100_Q816W1	<i>Bacillus cereus</i>	YtzF
3036	YtgP [<i>Bacillus subtilis</i>]	UniRef100_O34674	<i>Bacillus subtilis</i>	YtgP
3037	YthP [<i>Bacillus subtilis</i>]	UniRef100_O30505	<i>Bacillus subtilis</i>	YthP
3038				OpuD

TABLE 1-continued

Predicted functions			
SEQ ID NO.	Description	UniRef Accession No.	Organism
3039	Protein cse60 [<i>Bacillus subtilis</i>]	UniRef100_P94496	<i>Bacillus subtilis</i>
3040	Rhodanese-like domain protein [<i>Bacillus cereus</i>]	UniRef100_Q72YZ9	<i>Bacillus cereus</i>
3041			RapA
3042	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HI31	<i>Bacillus thuringiensis</i>
3043	YteU [<i>Bacillus subtilis</i>]	UniRef100_O34378	<i>Bacillus subtilis</i>
3044			YteU
3045			YteT
3046	YteR [<i>Bacillus subtilis</i>]	UniRef100_O34559	<i>Bacillus subtilis</i>
3047	Transmembrane lipoprotein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFJ5	<i>Bacillus halodurans</i>
3048	YtdP protein [<i>Bacillus subtilis</i>]	UniRef100_O32071	<i>Bacillus subtilis</i>
3049	YtcQ protein [<i>Bacillus subtilis</i>]	UniRef100_Q795R2	<i>Bacillus subtilis</i>
3050			YtdP
3051	Hypothetical protein ytbQ [<i>Bacillus subtilis</i>]	UniRef100_P53560	<i>Bacillus subtilis</i>
3052	YtaP [<i>Bacillus subtilis</i>]	UniRef100_O34973	<i>Bacillus subtilis</i>
3053	Amino acid/polyamine transporter; family I [<i>Methanococcus maripaludis</i>]	UniRef100_Q6LYX9	<i>Methanococcus maripaludis</i>
3054	Transcriptional regulator, LysR family [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97DX1	<i>Clostridium acetobutylicum</i>
3055	Prolyl endopeptidase [<i>Bacillus cereus</i>]	UniRef100_Q81C54	<i>Bacillus cereus</i>
3056	Leucyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P36430	<i>Bacillus subtilis</i>
3057			YycE
3058	YtvB [<i>Bacillus subtilis</i>]	UniRef100_O34546	<i>Bacillus subtilis</i>
3059	Lipoprotein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPK3	<i>Oceanobacillus iheyensis</i>
3060	YtaP [<i>Bacillus subtilis</i>]	UniRef100_O30500	<i>Bacillus subtilis</i>
3061	YtrF [<i>Bacillus subtilis</i>]	UniRef100_O35005	<i>Bacillus subtilis</i>
3062	Hypothetical ABC transporter ATP-binding protein ytrE [<i>Bacillus subtilis</i>]	UniRef100_O34392	<i>Bacillus subtilis</i>
3063	YtrC [<i>Bacillus subtilis</i>]	UniRef100_O34898	<i>Bacillus subtilis</i>
3064	Transporter [<i>Bacillus subtilis</i>]	UniRef100_O34641	<i>Bacillus subtilis</i>
3065	Transcription regulator [<i>Bacillus subtilis</i>]	UniRef100_O34712	<i>Bacillus subtilis</i>
3066	Hypothetical protein ytzC [<i>Bacillus subtilis</i>]	UniRef100_O32073	<i>Bacillus subtilis</i>
3067	YtqA [<i>Bacillus subtilis</i>]	UniRef100_O35008	<i>Bacillus subtilis</i>
3068			YtqA
3069	Proton glutamate symport protein [<i>Bacillus subtilis</i>]	UniRef100_P39817	<i>Bacillus subtilis</i>
3070	Hypothetical protein ytpB [<i>Bacillus subtilis</i>]	UniRef100_O34707	<i>Bacillus subtilis</i>
3071	Probable lysophospholipase [<i>Bacillus subtilis</i>]	UniRef100_O34705	<i>Bacillus subtilis</i>
3072	YtoA [<i>Bacillus subtilis</i>]	UniRef100_O34696	<i>Bacillus subtilis</i>
3073			YtpB
3074	Glycosyl transferase, group 1 family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HCB9	<i>Bacillus thuringiensis</i>
3075	Asparagine synthetase [glutamine-hydrolyzing] 1 [<i>Bacillus subtilis</i>]	UniRef100_P54420	glutamine-hydrolyzing
3076	S-adenosylmethionine synthetase [<i>Bacillus subtilis</i>]	UniRef100_P54419	<i>Bacillus subtilis</i>
3077	Phosphoenolpyruvate carboxykinase [ATP] [<i>Bacillus subtilis</i>]	UniRef100_P54418	ATP
3078	Sodium:dicarboxylate symporter [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EP16	<i>Oceanobacillus iheyensis</i>
3079	Hypothetical protein ytmB [<i>Bacillus subtilis</i>]	UniRef100_O34365	<i>Bacillus subtilis</i>
3080	Putative peptidase [<i>Bacillus subtilis</i>]	UniRef100_O34493	<i>Bacillus subtilis</i>
3081	ABC transporter substrate-binding protein [<i>Bacillus cereus</i>]	UniRef100_Q816P5	<i>Bacillus cereus</i>
3082	Putative transporter [<i>Bacillus subtilis</i>]	UniRef100_O34314	<i>Bacillus subtilis</i>
3083			YtmA
3084	YtkD [<i>Bacillus subtilis</i>]	UniRef100_O35013	<i>Bacillus subtilis</i>
3085	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HC91	<i>Bacillus thuringiensis</i>
			YtlA
			YtlC
			YtlD
			YtkD

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3086	Hypothetical protein ytkC [<i>Bacillus subtilis</i>]	UniRef100_O34883	<i>Bacillus subtilis</i>	YtkC
3087	General stress protein 20U [<i>Bacillus subtilis</i>]	UniRef100_P80879	<i>Bacillus subtilis</i>	Dps
3088	Hypothetical protein ytkA [<i>Bacillus subtilis</i>]	UniRef100_P40768	<i>Bacillus subtilis</i>	YtkA
3089	S-ribosylhomocysteinase [<i>Bacillus subtilis</i>]	UniRef100_O34667	<i>Bacillus subtilis</i>	LuxS
3090	Hypothetical UPF0161 protein ytlA [<i>Bacillus subtilis</i>]	UniRef100_O34601	<i>Bacillus subtilis</i>	
3091	YtiB [<i>Bacillus subtilis</i>]	UniRef100_O34872	<i>Bacillus subtilis</i>	YtiB
3092	Low-affinity zinc transport protein [<i>Bacillus cereus</i>]	UniRef100_Q81F90	<i>Bacillus cereus</i>	YciC
3093	High-affinity zinc uptake system protein znuA [<i>Bacillus cereus</i>]	UniRef100_Q81EF8	<i>Bacillus cereus</i>	YcdH
3094	50S ribosomal protein L31 type B [<i>Bacillus subtilis</i>]	UniRef100_O34967	<i>Bacillus subtilis</i>	
3095	YthA [<i>Bacillus subtilis</i>]	UniRef100_O34655	<i>Bacillus subtilis</i>	YthA
3096	YthB [<i>Bacillus subtilis</i>]	UniRef100_O34505	<i>Bacillus subtilis</i>	YthB
3097	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q737J1	<i>Bacillus cereus</i>	
3098				
3099	O-succinylbenzoate synthase (EC 4.2.1.—) (OSB synthase) (OSBS) (4-(2'-carboxyphenyl)-4-oxybutyric acid synthase) [<i>Bacillus subtilis</i>]	UniRef100_O34514	<i>Bacillus subtilis</i>	MenC
3100	O-succinylbenzoate--CoA ligase [<i>Bacillus subtilis</i>]	UniRef100_P23971	<i>Bacillus subtilis</i>	MenE
3101	Naphthoate synthase [<i>Bacillus subtilis</i>]	UniRef100_P23966	<i>Bacillus subtilis</i>	MenB
3102				YtxM
3103	Menaquinone biosynthesis protein menD [Includes: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 2.5.1.64) (SHCHC synthase); 2-oxoglutarate decarboxylase (EC 4.1.1.71) (Alpha-ketoglutarate decarboxylase) (KDC)] [<i>Bacillus subtilis</i>]	UniRef100_P23970	Includes: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 2.5.1.64) (SHCHC synthase); 2-oxoglutarate decarboxylase (EC 4.1.1.71) (Alpha-ketoglutarate decarboxylase) (KDC)	MenD
3104	Menaquinone-specific isochorismate synthase [<i>Bacillus subtilis</i>]	UniRef100_P23973	<i>Bacillus subtilis</i>	MenF
3105	Probable 1,4-dihydroxy-2-naphthoate octaprenyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P39582	<i>Bacillus subtilis</i>	MenA
3106	Hypothetical protein yteA [<i>Bacillus subtilis</i>]	UniRef100_P42408	<i>Bacillus subtilis</i>	YteA
3107	Glycogen phosphorylase [<i>Bacillus subtilis</i>]	UniRef100_P39123	<i>Bacillus subtilis</i>	GlgP
3108	Glycogen synthase [<i>Bacillus subtilis</i>]	UniRef100_P39125	<i>Bacillus subtilis</i>	GlgA
3109	Glycogen biosynthesis protein glgD [<i>Bacillus subtilis</i>]	UniRef100_P39124	<i>Bacillus subtilis</i>	GlgD
3110	Glucose-1-phosphate adenylyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P39122	<i>Bacillus subtilis</i>	GlgC
3111	1,4-alpha-glucan branching enzyme [<i>Bacillus subtilis</i>]	UniRef100_P39118	<i>Bacillus subtilis</i>	GlgB
3112				AraR
3113	YuaJ protein [<i>Bacillus subtilis</i>]	UniRef100_O32074	<i>Bacillus subtilis</i>	YuaJ
3114	BH4010 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5S8	<i>Bacillus halodurans</i>	YhcS
3115	BH4011 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5S7	<i>Bacillus halodurans</i>	
3116				RapD
3117				

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3118	Pyrrolidone-carboxylate peptidase [<i>Bacillus amyloliquefaciens</i>]	UniRef100_P46107	<i>Bacillus amyloliquefaciens</i>	Pcp
3119	BH0597 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KF88	<i>Bacillus halodurans</i>	YuaA
3120				YubG
3121				YxxF
3122	YuaE protein [<i>Bacillus subtilis</i>]	UniRef100_O32078	<i>Bacillus subtilis</i>	YuaE
3123	YuaD protein [<i>Bacillus subtilis</i>]	UniRef100_O32079	<i>Bacillus subtilis</i>	YuaD
3124	Alcohol dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P71017	<i>Bacillus subtilis</i>	GbsB
3125	Betaine aldehyde dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P71016	<i>Bacillus subtilis</i>	GbsA
3126	Hypothetical protein yuaC [<i>Bacillus subtilis</i>]	UniRef100_P71015	<i>Bacillus subtilis</i>	YuaC
3127	UPI00002D3D35 UniRef100 entry	UniRef100_UPI00002D3D35		OpuE
3128	Hypothetical protein yktD [<i>Bacillus subtilis</i>]	UniRef100_Q45500	<i>Bacillus subtilis</i>	YktD
3129	Alanine racemase 2 [<i>Bacillus subtilis</i>]	UniRef100_P94494	<i>Bacillus subtilis</i>	YncD
3130	Oxalate decarboxylase [<i>Bacillus cereus</i>]	UniRef100_Q81GZ6	<i>Bacillus cereus</i>	YoaN
3131	Hypothetical protein CAC0135 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97MQ7	<i>Clostridium acetobutylicum</i>	
3132	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGC9	<i>Bacillus thuringiensis</i>	
3133	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGC8	<i>Bacillus thuringiensis</i>	
3134				
3135	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGC6	<i>Bacillus thuringiensis</i>	
3136	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGC5	<i>Bacillus thuringiensis</i>	
3137	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETF5	<i>Oceanobacillus iheyensis</i>	
3138				
3139				
3140	YkoN [<i>Bacillus subtilis</i>]	UniRef100_O34625	<i>Bacillus subtilis</i>	YkoN
3141	Hypothetical protein ykoP [<i>Bacillus subtilis</i>]	UniRef100_O34495	<i>Bacillus subtilis</i>	YkoP
3142	Hypothetical UPF0151 protein ykoQ [<i>Bacillus subtilis</i>]	UniRef100_O35040	<i>Bacillus subtilis</i>	YkoQ
3143	Undecaprenyl-diphosphatase [<i>Bacillus subtilis</i>]	UniRef100_P94507	<i>Bacillus subtilis</i>	YubB
3144	Hypothetical UPF0118 protein yubA [<i>Bacillus subtilis</i>]	UniRef100_O32086	<i>Bacillus subtilis</i>	YubA
3145	Hypothetical oxidoreductase yulF [<i>Bacillus subtilis</i>]	UniRef100_O05265	<i>Bacillus subtilis</i>	YulF
3146	Lmo2256 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q929B9	<i>Listeria monocytogenes</i>	YraA
3147				McpA
3148				McpA
3149				McpA
3150	Protein-glutamine gamma-glutamyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P40746	<i>Bacillus subtilis</i>	Tgl
3151	2-nitropropane dioxygenase [<i>Bacillus subtilis</i>]	UniRef100_O05413	<i>Bacillus subtilis</i>	YrpB
3152	Hypothetical UPF0047 protein yugU [<i>Bacillus subtilis</i>]	UniRef100_O05243	<i>Bacillus subtilis</i>	YugU
3153	Hypothetical protein yugT [<i>Bacillus subtilis</i>]	UniRef100_O05242	<i>Bacillus subtilis</i>	YugT
3154	Transcriptional regulator, TetR family [<i>Bacillus cereus</i>]	UniRef100_Q81GX6	<i>Bacillus cereus</i>	YfiR
3155	Hypothetical protein yqeB [<i>Bacillus subtilis</i>]	UniRef100_P54447	<i>Bacillus subtilis</i>	YqeB
3156	Beta(1,4)-glucan glucanohydrolase [<i>Erwinia carotovora</i>]	UniRef100_Q6D3B7	<i>Erwinia carotovora</i>	
3157	Hypothetical UPF0053 protein yugS [<i>Bacillus subtilis</i>]	UniRef100_O05241	<i>Bacillus subtilis</i>	YugS

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3158	Hypothetical protein yugP [<i>Bacillus subtilis</i>]	UniRef100_O05248	<i>Bacillus subtilis</i>	YugP
3159	YugO protein [<i>Bacillus subtilis</i>]	UniRef100_Q795M8	<i>Bacillus subtilis</i>	
3160	Hypothetical protein yugN [<i>Bacillus subtilis</i>]	UniRef100_O05246	<i>Bacillus subtilis</i>	YugN
3161	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HIW1	<i>Bacillus thuringiensis</i>	YdfR
3162	YtaB protein [<i>Bacillus subtilis</i>]	UniRef100_O34694	<i>Bacillus subtilis</i>	YtaB
3163	Predicted acetyltransferase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97IT3	<i>Clostridium acetobutylicum</i>	YkkB
3164	Glucose-6-phosphate isomerase [<i>Bacillus subtilis</i>]	UniRef100_P80860	<i>Bacillus subtilis</i>	Pgi
3165	Probable NADH-dependent butanol dehydrogenase 1 [<i>Bacillus subtilis</i>]	UniRef100_O05239	<i>Bacillus subtilis</i>	YugJ
3166	YuzA protein [<i>Bacillus subtilis</i>]	UniRef100_O32087	<i>Bacillus subtilis</i>	
3167	General stress protein 13 [<i>Bacillus subtilis</i>]	UniRef100_P80870	<i>Bacillus subtilis</i>	YugI
3168	Alanine transaminase [<i>Bacillus subtilis</i>]	UniRef100_Q795M6	<i>Bacillus subtilis</i>	AlaT
3169	Transcriptional regulator [<i>Bacillus subtilis</i>]	UniRef100_O05236	<i>Bacillus subtilis</i>	AlaR
3170	Hypothetical protein yugF [<i>Bacillus subtilis</i>]	UniRef100_O05235	<i>Bacillus subtilis</i>	YugF
3171	Hypothetical protein yugE [<i>Bacillus subtilis</i>]	UniRef100_O05234	<i>Bacillus subtilis</i>	
3172	Hypothetical protein SMU.305 [<i>Streptococcus mutans</i>]	UniRef100_Q9X669	<i>Streptococcus mutans</i>	
3173	Putative aminotransferase B [<i>Bacillus subtilis</i>]	UniRef100_Q08432	<i>Bacillus subtilis</i>	PatB
3174				
3175	Kinase-associated lipoprotein B precursor [<i>Bacillus subtilis</i>]	UniRef100_Q08429	<i>Bacillus subtilis</i>	KapB
3176	Hypothetical protein yugB [<i>Bacillus subtilis</i>]	UniRef100_O05231	<i>Bacillus subtilis</i>	KapD
3177				YuxJ
3178				PbpD
3179	Hypothetical protein yuxK [<i>Bacillus subtilis</i>]	UniRef100_P40761	<i>Bacillus subtilis</i>	YuxK
3180	Hypothetical protein yufK [<i>Bacillus subtilis</i>]	UniRef100_O05249	<i>Bacillus subtilis</i>	YufK
3181	Hypothetical protein yufL [<i>Bacillus subtilis</i>]	UniRef100_O05250	<i>Bacillus subtilis</i>	YufL
3182	Hypothetical protein yufM [<i>Bacillus subtilis</i>]	UniRef100_O05251	<i>Bacillus subtilis</i>	YufM
3183				
3184	UPI00003CB938 UniRef100 entry	UniRef100_UPI00003CB938		PssA
3185	Hypothetical protein ybfM [<i>Bacillus subtilis</i>]	UniRef100_O31453	<i>Bacillus subtilis</i>	YbfM
3186	Phosphatidylserine decarboxylase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HDI5	<i>Bacillus thuringiensis</i>	Psd
3187	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
3188	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
3189	Na(+)-malate symporter [<i>Bacillus subtilis</i>]	UniRef100_O05256	<i>Bacillus subtilis</i>	MaeN
3190				
3191	Na(+)/H(+) antiporter subunit A [<i>Bacillus subtilis</i>]	UniRef100_Q9K2S2	<i>Bacillus subtilis</i>	MrpA
3192	Na(+)/H(+) antiporter subunit B [<i>Bacillus subtilis</i>]	UniRef100_O05259	<i>Bacillus subtilis</i>	MrpB
3193	Na(+)/H(+) antiporter subunit C [<i>Bacillus subtilis</i>]	UniRef100_O05260	<i>Bacillus subtilis</i>	MrpC
3194	Na(+)/H(+) antiporter subunit D [<i>Bacillus subtilis</i>]	UniRef100_O05229	<i>Bacillus subtilis</i>	MrpD
3195	Na(+)/H(+) antiporter subunit E [<i>Bacillus subtilis</i>]	UniRef100_Q7WY60	<i>Bacillus subtilis</i>	MrpE
3196	Na(+)/H(+) antiporter subunit F [<i>Bacillus subtilis</i>]	UniRef100_O05228	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3197	Na(+)/H(+) antiporter subunit G [<i>Bacillus subtilis</i>]	UniRef100_O05227	<i>Bacillus subtilis</i>	MrpG
3198				YuxO
3199				ComA
3200				ComP
3201	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
3202	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		ComP
3203				
3204				
3205	ComQ [<i>Bacillus subtilis</i>]	UniRef100_Q9K5L3	<i>Bacillus subtilis</i>	ComQ
3206				
3207	Hypothetical protein yuzC [<i>Bacillus subtilis</i>]	UniRef100_O32089	<i>Bacillus subtilis</i>	YuzC
3208	Hypothetical protein yuxH [<i>Bacillus subtilis</i>]	UniRef100_P14203	<i>Bacillus subtilis</i>	YuxH
3209	YueK protein [<i>Bacillus subtilis</i>]	UniRef100_O32090	<i>Bacillus subtilis</i>	YueK
3210	YueJ protein [<i>Bacillus subtilis</i>]	UniRef100_O32091	<i>Bacillus subtilis</i>	YueJ
3211				YueI
3212				
3213	Hypothetical protein yueG [<i>Bacillus subtilis</i>]	UniRef100_O32094	<i>Bacillus subtilis</i>	
3214				YueF
3215	RRF2 family protein [<i>Bacillus cereus</i>]	UniRef100_Q81EX1	<i>Bacillus cereus</i>	YwnA
3216	Probable lipase/esterase [<i>Rhodopirellula baltica</i>]	UniRef100_Q7UUA2	<i>Rhodopirellula baltica</i>	YuxL
3217	BH1896 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KBN0	<i>Bacillus halodurans</i>	
3218	YueE protein [<i>Bacillus subtilis</i>]	UniRef100_O32098	<i>Bacillus subtilis</i>	YueE
3219	YueD protein [<i>Bacillus subtilis</i>]	UniRef100_O32099	<i>Bacillus subtilis</i>	YueD
3220	Hypothetical protein yueC [<i>Bacillus subtilis</i>]	UniRef100_O32100	<i>Bacillus subtilis</i>	YueC
3221	YueB protein [<i>Bacillus subtilis</i>]	UniRef100_O32101	<i>Bacillus subtilis</i>	YueB
3222	YukA protein [<i>Bacillus subtilis</i>]	UniRef100_P71068	<i>Bacillus subtilis</i>	YukA
3223	YukC protein [<i>Bacillus subtilis</i>]	UniRef100_P71070	<i>Bacillus subtilis</i>	YukC
3224	YukD protein [<i>Bacillus subtilis</i>]	UniRef100_P71071	<i>Bacillus subtilis</i>	
3225	Lin0049 protein [<i>Listeria innocua</i>]	UniRef100_Q92FQ4	<i>Listeria innocua</i>	
3226				YflA
3227	YukF protein [<i>Bacillus subtilis</i>]	UniRef100_P71073	<i>Bacillus subtilis</i>	YukF
3228	Alanine dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_Q08352	<i>Bacillus subtilis</i>	Ald
3229				
3230	YuiH protein [<i>Bacillus subtilis</i>]	UniRef100_O32103	<i>Bacillus subtilis</i>	YuiH
3231	YuiG protein [<i>Bacillus subtilis</i>]	UniRef100_O32104	<i>Bacillus subtilis</i>	YuiG
3232	YuiF protein [<i>Bacillus subtilis</i>]	UniRef100_O32105	<i>Bacillus subtilis</i>	YuiF
3233	Probable cytosol aminopeptidase [<i>Bacillus subtilis</i>]	UniRef100_O32106	<i>Bacillus subtilis</i>	YuiE
3234	YuiD protein [<i>Bacillus subtilis</i>]	UniRef100_O32107	<i>Bacillus subtilis</i>	YuiD
3235	YuiC protein [<i>Bacillus subtilis</i>]	UniRef100_O32108	<i>Bacillus subtilis</i>	YuiC
3236	YuiB protein [<i>Bacillus subtilis</i>]	UniRef100_O32109	<i>Bacillus subtilis</i>	YuiB
3237				
3238				YumB
3239	Thioredoxine reductase [<i>Bacillus subtilis</i>]	UniRef100_O05268	<i>Bacillus subtilis</i>	YumC
3240				
3241	YdjO protein [<i>Bacillus subtilis</i>]	UniRef100_O34759	<i>Bacillus subtilis</i>	
3242				YxbD
3243	Hypothetical protein yutM [<i>Bacillus subtilis</i>]	UniRef100_O32113	<i>Bacillus subtilis</i>	YutM
3244	Diaminopimelate epimerase [<i>Bacillus subtilis</i>]	UniRef100_O32114	<i>Bacillus subtilis</i>	DapF
3245	YutK protein [<i>Bacillus subtilis</i>]	UniRef100_O32115	<i>Bacillus subtilis</i>	YutK
3246	YuzB protein [<i>Bacillus subtilis</i>]	UniRef100_O32116	<i>Bacillus subtilis</i>	
3247	YutJ protein [<i>Bacillus subtilis</i>]	UniRef100_O32117	<i>Bacillus subtilis</i>	YutJ
3248	YdhG protein [<i>Bacillus subtilis</i>]	UniRef100_O05499	<i>Bacillus subtilis</i>	YdhG
3249	Response regulator aspartate phosphatase [<i>Bacillus halodurans</i>]	UniRef100_Q9KBE1	<i>Bacillus halodurans</i>	RapI
3250	Phenolic acid decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_O07006	<i>Bacillus subtilis</i>	PadC

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3251	BH2266 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KAM1	<i>Bacillus halodurans</i>	
3252	YuzD protein [<i>Bacillus subtilis</i>]	UniRef100_O32118	<i>Bacillus subtilis</i>	YuzD
3253	YutI protein [<i>Bacillus subtilis</i>]	UniRef100_O32119	<i>Bacillus subtilis</i>	
3254	Probable peptidase yuxL [<i>Bacillus subtilis</i>]	UniRef100_P39839	<i>Bacillus subtilis</i>	YuxL
3255	Homoserine kinase [<i>Bacillus subtilis</i>]	UniRef100_P04948	<i>Bacillus subtilis</i>	ThrB
3256	Threonine synthase [<i>Bacillus subtilis</i>]	UniRef100_P04990	<i>Bacillus subtilis</i>	ThrC
3257	Homoserine dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P19582	<i>Bacillus subtilis</i>	Hom
3258	Glycerate dehydrogenase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ENW9	<i>Oceanobacillus iheyensis</i>	YvcT
3259	YutH protein [<i>Bacillus subtilis</i>]	UniRef100_O32123	<i>Bacillus subtilis</i>	YutH
3260	Hypothetical protein yutG [<i>Bacillus subtilis</i>]	UniRef100_O32124	<i>Bacillus subtilis</i>	YutG
3261	YutF protein [<i>Bacillus subtilis</i>]	UniRef100_O32125	<i>Bacillus subtilis</i>	YutF
3262	YutE protein [<i>Bacillus subtilis</i>]	UniRef100_O32126	<i>Bacillus subtilis</i>	YutE
3263	YutD protein [<i>Bacillus subtilis</i>]	UniRef100_O32127	<i>Bacillus subtilis</i>	
3264	YutC protein [<i>Bacillus subtilis</i>]	UniRef100_O32128	<i>Bacillus subtilis</i>	YutC
3265				LipA
3266	YunA protein [<i>Bacillus subtilis</i>]	UniRef100_O32130	<i>Bacillus subtilis</i>	YunA
3267				
3268	Sodium-dependent transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K7C5	<i>Bacillus halodurans</i>	YocR
3269	YunB protein [<i>Bacillus subtilis</i>]	UniRef100_O32131	<i>Bacillus subtilis</i>	YunB
3270	YunC protein [<i>Bacillus subtilis</i>]	UniRef100_O32132	<i>Bacillus subtilis</i>	YunC
3271	YunD protein [<i>Bacillus subtilis</i>]	UniRef100_O32133	<i>Bacillus subtilis</i>	YunD
3272	YunE protein [<i>Bacillus subtilis</i>]	UniRef100_O32134	<i>Bacillus subtilis</i>	YunE
3273	YunF protein [<i>Bacillus subtilis</i>]	UniRef100_O32135	<i>Bacillus subtilis</i>	YunF
3274				YmcC
3275	TetR family transcriptional regulator? [<i>Symbiobacterium thermophilum</i>]	UniRef100_Q67KA4	<i>Symbiobacterium thermophilum</i>	PksA
3276				
3277	Purine catabolism protein pucG [<i>Bacillus subtilis</i>]	UniRef100_O32148	<i>Bacillus subtilis</i>	YurG
3278	Allantoate amidohydrolase [<i>Bacillus subtilis</i>]	UniRef100_O32149	<i>Bacillus subtilis</i>	YurH
3279	Purine catabolism regulatory protein [<i>Bacillus subtilis</i>]	UniRef100_O32138	<i>Bacillus subtilis</i>	PucR
3280	Multidrug resistance protein B [<i>Bacillus cereus</i> ZK]	UniRef100_Q63FH7	<i>Bacillus cereus</i> ZK	Blt
3281	BH2308 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KAH9	<i>Bacillus halodurans</i>	YcgA
3282				TrpD
3283	Anthranilate phosphoribosyltransferase [<i>Pyrococcus furiosus</i>]	UniRef100_Q8U089	<i>Pyrococcus furiosus</i>	
3284	Extracellular ribonuclease precursor [<i>Bacillus subtilis</i>]	UniRef100_O32150	<i>Bacillus subtilis</i>	YurI
3285	BH1977 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KBF1	<i>Bacillus halodurans</i>	
3286				YurR
3287	Putative membrane protein [<i>Bordetella bronchiseptica</i>]	UniRef100_Q7WGW7	<i>Bordetella bronchiseptica</i>	
3288	UPI00003CB453 UniRef100 entry	UniRef100_UPI00003CB453		
3289	Response regulator aspartate phosphatase I [<i>Bacillus subtilis</i>]	UniRef100_P96649	<i>Bacillus subtilis</i>	RapI
3290	YurU protein [<i>Bacillus subtilis</i>]	UniRef100_O32162	<i>Bacillus subtilis</i>	YurU
3291	NifU-like protein [<i>Bacillus subtilis</i>]	UniRef100_O32163	<i>Bacillus subtilis</i>	YurV
3292	Probable cysteine desulfurase [<i>Bacillus subtilis</i>]	UniRef100_O32164	<i>Bacillus subtilis</i>	Csd
3293	YurX protein [<i>Bacillus subtilis</i>]	UniRef100_O32165	<i>Bacillus subtilis</i>	YurX
3294	Vegetative protein 296 [<i>Bacillus subtilis</i>]	UniRef100_P80866	<i>Bacillus subtilis</i>	YurY
3295	Lmo2575 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q8Y480	<i>Listeria monocytogenes</i>	CzcD
3296				
3297	BH3473 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K796	<i>Bacillus halodurans</i>	YurZ
3298	YusA protein [<i>Bacillus subtilis</i>]	UniRef100_O32167	<i>Bacillus subtilis</i>	YusA
3299	YusB protein [<i>Bacillus subtilis</i>]	UniRef100_O32168	<i>Bacillus subtilis</i>	YusB
3300	YusC protein [<i>Bacillus subtilis</i>]	UniRef100_O32169	<i>Bacillus subtilis</i>	YusC

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3301	Hypothetical protein yusD [<i>Bacillus subtilis</i>]	UniRef100_O32170	<i>Bacillus subtilis</i>	YusD
3302	YusE protein [<i>Bacillus subtilis</i>]	UniRef100_O32171	<i>Bacillus subtilis</i>	
3303	YusF protein [<i>Bacillus subtilis</i>]	UniRef100_O32172	<i>Bacillus subtilis</i>	YusF
3304				
3305	Glycine cleavage system H protein [<i>Bacillus subtilis</i>]	UniRef100_O32174	<i>Bacillus subtilis</i>	GevH
3306	Hypothetical protein yusI [<i>Bacillus subtilis</i>]	UniRef100_O32175	<i>Bacillus subtilis</i>	YusI
3307	YusJ protein [<i>Bacillus subtilis</i>]	UniRef100_O32176	<i>Bacillus subtilis</i>	YusJ
3308	YusK protein [<i>Bacillus subtilis</i>]	UniRef100_O32177	<i>Bacillus subtilis</i>	YusK
3309	YusL protein [<i>Bacillus subtilis</i>]	UniRef100_O32178	<i>Bacillus subtilis</i>	YusL
3310				
3311				
3312	YusN protein [<i>Bacillus subtilis</i>]	UniRef100_O32180	<i>Bacillus subtilis</i>	YusN
3313	Hypothetical protein yusU [<i>Bacillus subtilis</i>]	UniRef100_O32187	<i>Bacillus subtilis</i>	
3314	BH1040 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KE18	<i>Bacillus halodurans</i>	
3315	YusV protein [<i>Bacillus subtilis</i>]	UniRef100_O32188	<i>Bacillus subtilis</i>	YusV
3316				YfhA
3317	YfiZ protein [<i>Bacillus subtilis</i>]	UniRef100_O31568	<i>Bacillus subtilis</i>	YfiZ
3318	YfiY protein [<i>Bacillus subtilis</i>]	UniRef100_O31567	<i>Bacillus subtilis</i>	YfiY
3319	Hypothetical protein yusW precursor [<i>Bacillus subtilis</i>]	UniRef100_O32189	<i>Bacillus subtilis</i>	YusW
3320	YusX protein [<i>Bacillus subtilis</i>]	UniRef100_O32190	<i>Bacillus subtilis</i>	YusX
3321	D-alanyl-D-alanine carboxypeptidase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ERG0	<i>Oceanobacillus iheyensis</i>	DacB
3322	Hypothetical oxidoreductase yusZ [<i>Bacillus subtilis</i>]	UniRef100_P37959	<i>Bacillus subtilis</i>	YusZ
3323	Metalloregulation DNA-binding stress protein [<i>Bacillus subtilis</i>]	UniRef100_P37960	<i>Bacillus subtilis</i>	MrgA
3324	Probable serine protease yvtA [<i>Bacillus subtilis</i>]	UniRef100_Q9R9I1	<i>Bacillus subtilis</i>	YvtA
3325	Transcriptional regulatory protein cssR [<i>Bacillus subtilis</i>]	UniRef100_O32192	<i>Bacillus subtilis</i>	CssR
3326	Sensor protein cssS [<i>Bacillus subtilis</i>]	UniRef100_O32193	<i>Bacillus subtilis</i>	CssS
3327	YirB [<i>Bacillus subtilis</i>]	UniRef100_O32302	<i>Bacillus subtilis</i>	
3328	Putative HTH-type transcriptional regulator yuxN [<i>Bacillus subtilis</i>]	UniRef100_P40950	<i>Bacillus subtilis</i>	YuxN
3329	Fumarate hydratase class II [<i>Bacillus subtilis</i>]	UniRef100_P07343	<i>Bacillus subtilis</i>	CitG
3330				
3331	Spore germination protein A1 [<i>Bacillus subtilis</i>]	UniRef100_P07868	<i>Bacillus subtilis</i>	GerAA
3332	Spore germination protein A2 [<i>Bacillus subtilis</i>]	UniRef100_P07869	<i>Bacillus subtilis</i>	GerAB
3333	Spore germination protein A3 precursor [<i>Bacillus subtilis</i>]	UniRef100_P07870	<i>Bacillus subtilis</i>	GerAC
3334				
3335				YvqC
3336	YvqE protein [<i>Bacillus subtilis</i>]	UniRef100_O32198	<i>Bacillus subtilis</i>	YvqE
3337	YvqF protein [<i>Bacillus subtilis</i>]	UniRef100_O32199	<i>Bacillus subtilis</i>	YvqF
3338	YvqG protein [<i>Bacillus subtilis</i>]	UniRef100_O32200	<i>Bacillus subtilis</i>	YvqG
3339	Hypothetical protein yvqH [<i>Bacillus subtilis</i>]	UniRef100_O32201	<i>Bacillus subtilis</i>	YvqH
3340	Hypothetical protein yvqI [<i>Bacillus subtilis</i>]	UniRef100_O32202	<i>Bacillus subtilis</i>	YvqI
3341	Pectate lyase P358 [<i>Bacillus</i> sp. P-358]	UniRef100_Q8RR73	<i>Bacillus</i> sp. P-358	
3342	YvqK protein [<i>Bacillus subtilis</i>]	UniRef100_O34899	<i>Bacillus subtilis</i>	YvqK
3343	UPI00002E3648 UniRef100 entry	UniRef100_UPI00002E3648		FabG
3344				
3345				GbsB
3346				DapA
3347	Putative metal binding protein, YvrA [<i>Bacillus subtilis</i>]	UniRef100_O34631	<i>Bacillus subtilis</i>	YvrA
3348	Putative hemin permease, YvrB [<i>Bacillus subtilis</i>]	UniRef100_O34451	<i>Bacillus subtilis</i>	YvrB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3349	Putative metal binding protein, YvrC [<i>Bacillus subtilis</i>]	UniRef100_O34805	<i>Bacillus subtilis</i>	YvrC
3350	Transcriptional regulator, GntR family [<i>Bacillus anthracis</i>]	UniRef100_Q81SA7	<i>Bacillus anthracis</i>	YdhC
3351	Putative ketoreductase, YvrD [<i>Bacillus subtilis</i>]	UniRef100_O34782	<i>Bacillus subtilis</i>	YvrD
3352	UPI00003CC410 UniRef100 entry	UniRef100_UPI00003CC410		YfkK
3353	Transcriptional regulators, LysR family [<i>Bacillus cereus</i>]	UniRef100_Q81DJ6	<i>Bacillus cereus</i>	AlsR
3354	Exo-poly-alpha-D-galacturonidase, putative [<i>Thermotoga maritima</i>]	UniRef100_Q9WYR8	<i>Thermotoga maritima</i>	
3355	Altronate hydrolase [<i>Bacillus subtilis</i>]	UniRef100_O34673	<i>Bacillus subtilis</i>	UxaA
3356	Altronate oxidoreductase [<i>Bacillus subtilis</i>]	UniRef100_O34354	<i>Bacillus subtilis</i>	UxaB
3357	LacI repressor-like protein [<i>Bacillus subtilis</i>]	UniRef100_Q9JMQ1	<i>Bacillus subtilis</i>	ExuR
3358	Hypothetical symporter yjmB [<i>Bacillus subtilis</i>]	UniRef100_O34961	<i>Bacillus subtilis</i>	YjmB
3359	Uronate isomerase [<i>Bacillus subtilis</i>]	UniRef100_O34808	<i>Bacillus subtilis</i>	UxaC
3360	Putative sensory protein kinase, YvrG [<i>Bacillus subtilis</i>]	UniRef100_O34989	<i>Bacillus subtilis</i>	YvrG
3361	Putative DNA binding response regulator, YvrH [<i>Bacillus subtilis</i>]	UniRef100_P94504	<i>Bacillus subtilis</i>	YvrH
3362	Ferrichrome transport ATP-binding protein fhuC [<i>Bacillus subtilis</i>]	UniRef100_P49938	<i>Bacillus subtilis</i>	FhuC
3363	Ferrichrome transport system permease protein fhuG [<i>Bacillus subtilis</i>]	UniRef100_P49937	<i>Bacillus subtilis</i>	FhuG
3364	Ferrichrome transport system permease protein fhuB [<i>Bacillus subtilis</i>]	UniRef100_P49936	<i>Bacillus subtilis</i>	FhuB
3365	Putative arginine ornithine antiporter, YvsH [<i>Bacillus subtilis</i>]	UniRef100_O32204	<i>Bacillus subtilis</i>	YvsH
3366	Hypothetical protein yvsG precursor [<i>Bacillus subtilis</i>]	UniRef100_O32205	<i>Bacillus subtilis</i>	YvsG
3367	Putative molybdate binding protein, YvgJ [<i>Bacillus subtilis</i>]	UniRef100_O32206	<i>Bacillus subtilis</i>	YvgJ
3368				YcnB
3369	Putative reductase protein, YvgN [<i>Bacillus subtilis</i>]	UniRef100_O32210	<i>Bacillus subtilis</i>	YvgN
3370	Fructokinase [<i>Listeria monocytogenes</i>]	UniRef100_Q722A5	<i>Listeria monocytogenes</i>	YdhR
3371	Hypothetical protein ycbU precursor [<i>Bacillus subtilis</i>]	UniRef100_P42253	<i>Bacillus subtilis</i>	YcbU
3372	Hypothetical protein CPE0889 [<i>Clostridium perfringens</i>]	UniRef100_Q8XM01	<i>Clostridium perfringens</i>	
3373	YvgS protein [<i>Bacillus subtilis</i>]	UniRef100_O32215	<i>Bacillus subtilis</i>	YvgS
3374	Hypothetical UPF0126 protein yvgT [<i>Bacillus subtilis</i>]	UniRef100_O32216	<i>Bacillus subtilis</i>	YvgT
3375	Glutamate-rich protein grpB [<i>Bacillus cereus</i>]	UniRef100_Q81CT5	<i>Bacillus cereus</i>	YqkA
3376	Acetyltransferase, GNAT family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HJN8	<i>Bacillus thuringiensis</i>	YuaI
3377	Disulfide bond formation protein C [<i>Bacillus subtilis</i>]	UniRef100_O32217	<i>Bacillus subtilis</i>	BdbC
3378	Disulfide bond formation protein D precursor [<i>Bacillus subtilis</i>]	UniRef100_O32218	<i>Bacillus subtilis</i>	BdbD
3379	YvgW protein [<i>Bacillus subtilis</i>]	UniRef100_O32219	<i>Bacillus subtilis</i>	YvgW
3380	ABC transporter, substrate binding protein [<i>Agrobacterium tumefaciens</i>]	UniRef100_Q8U7T8	<i>Agrobacterium tumefaciens</i>	YvfK
3381	Putative sugar transport integral membrane protein [<i>Streptomyces coelicolor</i>]	UniRef100_Q9K442	<i>Streptomyces coelicolor</i>	YurN
3382	Putative sugar transport integral membrane protein [<i>Streptomyces coelicolor</i>]	UniRef100_Q9K441	<i>Streptomyces coelicolor</i>	YurM
3383	Alpha-glucosidase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97K36	<i>Clostridium acetobutylicum</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3384	Glucan-glucohydrolase [<i>Microbispora bispora</i>]	UniRef100_Q59506	<i>Microbispora bispora</i>	YbbD
3385	ABC transporter, ATP-binding protein [<i>Bacillus anthracis</i>]	UniRef100_Q81NK1	<i>Bacillus anthracis</i>	MntB
3386	ABC transporter, permease protein [<i>Bacillus anthracis</i>]	UniRef100_Q81NK2	<i>Bacillus anthracis</i>	MntD
3387	ABC transporter, manganese-binding protein [<i>Listeria monocytogenes</i>]	UniRef100_Q71YG9	<i>Listeria monocytogenes</i>	YcdH
3388	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q74NQ1	<i>Bacillus cereus</i>	YhjQ
3389	Phosphosugar-binding transcriptional regulator, RpiR family, putative [<i>Enterococcus faecalis</i>]	UniRef100_Q837Y3	<i>Enterococcus faecalis</i>	YfiA
3390	1-phosphofructokinase [<i>Enterococcus faecalis</i>]	UniRef100_Q837Y2	<i>Enterococcus faecalis</i>	FruK
3391	PTS system, fructose-specific family, IIBC components [<i>Enterococcus faecalis</i>]	UniRef100_Q837Y1	<i>Enterococcus faecalis</i>	FruA
3392	PTS system, IIA component [<i>Enterococcus faecalis</i>]	UniRef100_Q837Y0	<i>Enterococcus faecalis</i>	FruA
3393	Fructose/tagatose bisphosphate aldolase [<i>Vibrio vulnificus</i>]	UniRef100_Q7MC78	<i>Vibrio vulnificus</i>	FbaA
3394				
3395	Heavy metal-transporting ATPase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HF81	<i>Bacillus thuringiensis</i>	YvgX
3396	Copper chaperone copZ [<i>Bacillus subtilis</i>]	UniRef100_O32221	<i>Bacillus subtilis</i>	
3397	YvgZ protein [<i>Bacillus subtilis</i>]	UniRef100_O32222	<i>Bacillus subtilis</i>	YvgZ
3398	Intracellular proteinase [<i>Thermoplasma volcanium</i>]	UniRef100_Q97BN0	<i>Thermoplasma volcanium</i>	YraA
3399	Putative HTH-type transcriptional regulator yulB [<i>Bacillus subtilis</i>]	UniRef100_O05261	<i>Bacillus subtilis</i>	YulB
3400	Rhamnulose kinase/L-ficulose kinase [<i>Bacteroides thetaiotaomicron</i>]	UniRef100_Q8A1A3	<i>Bacteroides thetaiotaomicron</i>	YulC
3401	Putative sugar isomerase [<i>Streptomyces coelicolor</i>]	UniRef100_Q9XAB3	<i>Streptomyces coelicolor</i>	
3402	Hypothetical protein [<i>Polyangium cellulosum</i>]	UniRef100_Q9L8B7	<i>Polyangium cellulosum</i>	
3403	Hypothetical oxidoreductase yuxG [<i>Bacillus subtilis</i>]	UniRef100_P40747	<i>Bacillus subtilis</i>	YuxG
3404	YfhI protein [<i>Bacillus subtilis</i>]	UniRef100_O31577	<i>Bacillus subtilis</i>	YfhI
3405	YvmA [<i>Bacillus subtilis</i>]	UniRef100_O34307	<i>Bacillus subtilis</i>	YvmA
3406	YvmC [<i>Bacillus subtilis</i>]	UniRef100_O34351	<i>Bacillus subtilis</i>	YvmC
3407	Putative cytochrome P450 CYPX [<i>Bacillus subtilis</i>]	UniRef100_O34926	<i>Bacillus subtilis</i>	CypX
3408	Hypothetical protein [<i>Chromobacterium violaceum</i>]	UniRef100_Q7NRR6	<i>Chromobacterium violaceum</i>	
3409				YvnA
3410	Hypothetical protein BRA0099 [<i>Brucella suis</i>]	UniRef100_Q8FXH8	<i>Brucella suis</i>	
3411	Chloroplast Toc64-2 [<i>Physcomitrella patens</i>]	UniRef100_Q6RJN6	<i>Physcomitrella patens</i>	GatA
3412	Hypothetical oxidoreductase yvaA [<i>Bacillus subtilis</i>]	UniRef100_O32223	<i>Bacillus subtilis</i>	YvaA
3413	Putative acyl carrier protein phosphodiesterase 2 [<i>Bacillus subtilis</i>]	UniRef100_O32224	<i>Bacillus subtilis</i>	YvaB
3414	Hypothetical UPF0176 protein ybfQ [<i>Bacillus subtilis</i>]	UniRef100_O31457	<i>Bacillus subtilis</i>	YbfQ
3415	Putative DNA binding response regulator, YvrH [<i>Bacillus subtilis</i>]	UniRef100_P94504	<i>Bacillus subtilis</i>	
3416	YvrI protein [<i>Bacillus subtilis</i>]	UniRef100_O34843	<i>Bacillus subtilis</i>	YvrI
3417				
3418	Oxalate decarboxylase oxdC [<i>Bacillus subtilis</i>]	UniRef100_O34714	<i>Bacillus subtilis</i>	OxdC
3419	Hypothetical protein yvrL [<i>Bacillus subtilis</i>]	UniRef100_O34686	<i>Bacillus subtilis</i>	YvrL
3420	BH3120 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K886	<i>Bacillus halodurans</i>	
3421				

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3422	Group-specific protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q633V1	<i>Bacillus cereus</i> ZK	
3423				
3424				
3425				
3426	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
3427	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
3428	LexA repressor [<i>Listeria monocytogenes</i>]	UniRef100_Q720B9	<i>Listeria monocytogenes</i>	
3429	Hypothetical protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q630E5	<i>Bacillus cereus</i> ZK	
3430	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81XX1	<i>Bacillus anthracis</i>	
3431				
3432	Hypothetical protein RSc1705 [<i>Ralstonia solanacearum</i>]	UniRef100_Q8XYQ3	<i>Ralstonia solanacearum</i>	
3433				YxiD
3434	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HH72	<i>Bacillus thuringiensis</i>	
3435				BlyA
3436				
3437				
3438				
3439	YomQ protein [Bacteriophage SPBc2]	UniRef100_O64053	Bacteriophage SPBc2	YomQ
3440				XkdV
3441				
3442				
3443				YubE
3444	ORF38 [Bacteriophage phi-105]	UniRef100_Q9ZXE5	Bacteriophage phi-105	
3445				YqbO
3446	Hypothetical protein MW1895 [<i>Staphylococcus aureus</i>]	UniRef100_Q8NVQ7	<i>Staphylococcus aureus</i>	
3447				
3448				
3449	Prophage LambdaBa02, major tail protein, putative [<i>Bacillus anthracis</i>]	UniRef100_Q81W97	<i>Bacillus anthracis</i>	
3450				
3451	ORF32 [Bacteriophage phi-105]	UniRef100_Q9ZXF1	Bacteriophage phi-105	
3452				
3453	Gp7 protein [Bacteriophage phi3626]	UniRef100_Q8SBP7	Bacteriophage phi3626	
3454	ORF29 [Bacteriophage phi-105]	UniRef100_Q9ZXF4	Bacteriophage phi-105	
3455	ORF27 [Bacteriophage phi-105]	UniRef100_Q9ZXF6	Bacteriophage phi-105	
3456	ClpP family serine protease, possible phage related [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HW4	<i>Clostridium acetobutylicum</i>	ClpP
3457	ORF25 [Bacteriophage phi-105]	UniRef100_Q9ZXF8	Bacteriophage phi-105	
3458				
3459	Putative terminase large subunit [Bacteriophage phi3626]	UniRef100_Q8SBQ2	Bacteriophage phi3626	
3460	Phage terminase-like protein, small subunit [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97HW1	<i>Clostridium acetobutylicum</i>	
3461	Prophage LambdaBa02, HNH endonuclease family protein [<i>Bacillus anthracis</i>]	UniRef100_Q81W86	<i>Bacillus anthracis</i>	
3462	Spore coat protein D [<i>Bacillus subtilis</i>]	UniRef100_P07791	<i>Bacillus subtilis</i>	
3463	ORF16 [Bacteriophage phi-105]	UniRef100_Q9ZXB9	Bacteriophage phi-105	
3464				
3465				
3466	ORF14 [Bacteriophage phi-105]	UniRef100_Q9ZXC1	Bacteriophage phi-105	
3467	ORF13 [Bacteriophage phi-105]	UniRef100_Q9ZXC2	Bacteriophage phi-105	
3468	ORF12 [Bacteriophage phi-105]	UniRef100_Q9ZXC3	Bacteriophage phi-105	
3469				
3470	ORF11 [Bacteriophage phi-105]	UniRef100_Q9ZXC4	Bacteriophage phi-105	
3471	ORF10 [Bacteriophage phi-105]	UniRef100_Q9ZXC5	Bacteriophage phi-105	
3472	DNA, complete sequence [Bacteriophage phi-105]	UniRef100_Q9ZXC7	Bacteriophage phi-105	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3473	DNA, complete sequence [Bacteriophage phi-105]	UniRef100_Q9ZXC8	Bacteriophage phi-105	
3474				
3475				
3476				
3477	Orf 36 [<i>Staphylococcus aureus</i> bacteriophage PVL]	UniRef100_Q9T1Z5	<i>Staphylococcus aureus</i> bacteriophage PVL	
3478				
3479	Orf16 [Bacteriophage bIL312]	UniRef100_Q9AZE4	Bacteriophage bIL312	
3480	Gp26 protein [Bacteriophage phi3626]	UniRef100_Q8SBM8	Bacteriophage phi3626	
3481	CI like protein [Lactobacillus casei bacteriophage A2]	UniRef100_O64370	<i>Lactobacillus casei</i> bacteriophage A2	
3482	Int gene product [<i>Staphylococcus aureus</i>]	UniRef100_Q932E5	<i>Staphylococcus aureus</i>	YdcL
3483	SsrA-binding protein [<i>Bacillus subtilis</i>]	UniRef100_O32230	<i>Bacillus subtilis</i>	SmpB
3484	Ribonuclease R [<i>Bacillus subtilis</i>]	UniRef100_O32231	<i>Bacillus subtilis</i>	Rnr
3485	YvaK protein [<i>Bacillus subtilis</i>]	UniRef100_O32232	<i>Bacillus subtilis</i>	YvaK
3486				
3487				
3488				
3489				YvaO
3490				YvaO
3491	Glycine betaine/carnitine/choline transport system permease protein opuCD [<i>Bacillus subtilis</i>]	UniRef100_O34742	<i>Bacillus subtilis</i>	OpuCD
3492	Glycine betaine/carnitine/choline- binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_O32243	<i>Bacillus subtilis</i>	OpuCC
3493	Glycine betaine/carnitine/choline transport system permease protein opuCB [<i>Bacillus subtilis</i>]	UniRef100_O34878	<i>Bacillus subtilis</i>	OpuCB
3494	Glycine betaine/carnitine/choline transport ATP-binding protein opuCA [<i>Bacillus subtilis</i>]	UniRef100_O34992	<i>Bacillus subtilis</i>	OpuCA
3495	ORF-1 [<i>Bacillus subtilis</i>]	UniRef100_O34709	<i>Bacillus subtilis</i>	YvaV
3496				YvbG
3497				YvbI
3498	YvbJ protein [<i>Bacillus subtilis</i>]	UniRef100_O32247	<i>Bacillus subtilis</i>	YvbJ
3499	UPI00003CBDCC UniRef100 entry	UniRef100_UPI00003CBDCC		YfiY
3500				
3501	YoaZ [<i>Bacillus subtilis</i>]	UniRef100_O34947	<i>Bacillus subtilis</i>	YoaZ
3502	Enolase [<i>Bacillus subtilis</i>]	UniRef100_P37869	<i>Bacillus subtilis</i>	Eno
3503	2,3-bisphosphoglycerate-independent phosphoglycerate mutase [<i>Bacillus subtilis</i>]	UniRef100_P39773	<i>Bacillus subtilis</i>	Pgm
3504	Triosephosphate isomerase [<i>Bacillus subtilis</i>]	UniRef100_P27876	<i>Bacillus subtilis</i>	TpiA
3505	Phosphoglycerate kinase [<i>Bacillus subtilis</i>]	UniRef100_P40924	<i>Bacillus subtilis</i>	Pgk
3506	Glyceraldehyde-3-phosphate dehydrogenase 1 [<i>Bacillus subtilis</i>]	UniRef100_P09124	<i>Bacillus subtilis</i>	GapA
3507	Central glycolytic genes regulator [<i>Bacillus subtilis</i>]	UniRef100_O32253	<i>Bacillus subtilis</i>	CggR
3508	Hypothetical protein yvbT [<i>Bacillus subtilis</i>]	UniRef100_O32254	<i>Bacillus subtilis</i>	YvbT
3509	YvbW protein [<i>Bacillus subtilis</i>]	UniRef100_O32257	<i>Bacillus subtilis</i>	YvbW
3510	ABC-type multidrug transport system, ATPase component [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8R7K3	<i>Thermoanaerobacter tengcongensis</i>	YfiL
3511				
3512				
3513	YvbY protein [<i>Bacillus subtilis</i>]	UniRef100_O32259	<i>Bacillus subtilis</i>	YvbY
3514	YvfW protein [<i>Bacillus subtilis</i>]	UniRef100_O07021	<i>Bacillus subtilis</i>	YvfW
3515				YvfV
3516	Hypothetical protein yvfI [<i>Bacillus subtilis</i>]	UniRef100_O07007	<i>Bacillus subtilis</i>	YvfI
3517	Putative L-lactate permease yvfH [<i>Bacillus subtilis</i>]	UniRef100_P71067	<i>Bacillus subtilis</i>	YvfH

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3518	RNA polymerase sigma-54 factor [<i>Bacillus subtilis</i>]	UniRef100_P24219	<i>Bacillus subtilis</i>	SigL
3519	Hypothetical protein yvfG [<i>Bacillus subtilis</i>]	UniRef100_P71066	<i>Bacillus subtilis</i>	
3520	Hypothetical protein yvfF [<i>Bacillus subtilis</i>]	UniRef100_P71065	<i>Bacillus subtilis</i>	YvfF
3521	YvfE protein [<i>Bacillus subtilis</i>]	UniRef100_Q795J3	<i>Bacillus subtilis</i>	YvfE
3522	Hypothetical protein yvfD [<i>Bacillus subtilis</i>]	UniRef100_P71063	<i>Bacillus subtilis</i>	YvfD
3523	Hypothetical protein yvfC [<i>Bacillus subtilis</i>]	UniRef100_P71062	<i>Bacillus subtilis</i>	YvfC
3524				YvfB
3525				YvfA
3526	Hypothetical protein yveT [<i>Bacillus subtilis</i>]	UniRef100_P71059	<i>Bacillus subtilis</i>	YveT
3527	Hypothetical protein yveS [<i>Bacillus subtilis</i>]	UniRef100_P71058	<i>Bacillus subtilis</i>	YveS
3528	Hypothetical protein yveR [<i>Bacillus subtilis</i>]	UniRef100_P71057	<i>Bacillus subtilis</i>	YveR
3529	Hypothetical protein yveQ [<i>Bacillus subtilis</i>]	UniRef100_P71056	<i>Bacillus subtilis</i>	YveQ
3530	Hypothetical protein yveP [<i>Bacillus subtilis</i>]	UniRef100_P71055	<i>Bacillus subtilis</i>	YveP
3531	Hypothetical protein yveO [<i>Bacillus subtilis</i>]	UniRef100_P71054	<i>Bacillus subtilis</i>	YveO
3532	Hypothetical protein yveN [<i>Bacillus subtilis</i>]	UniRef100_P71053	<i>Bacillus subtilis</i>	YveN
3533	Hypothetical protein yveM [<i>Bacillus subtilis</i>]	UniRef100_P71052	<i>Bacillus subtilis</i>	YveM
3534				YveL
3535	Chain length determinant protein [<i>Bacillus cereus</i>]	UniRef100_Q72XH7	<i>Bacillus cereus</i>	YveK
3536	Hypothetical protein yveJ [<i>Bacillus subtilis</i>]	UniRef100_P71049	<i>Bacillus subtilis</i>	Slr
3537	General stress protein 14 [<i>Bacillus subtilis</i>]	UniRef100_P80871	<i>Bacillus subtilis</i>	YwrO
3538				YcbE
3539	UPI000027D43B UniRef100 entry	UniRef100_UPI000027D43B		YvcT
3540	4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262) (4-(phosphohydroxy)-L-threonine dehydrogenase) [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CUU4	<i>Oceanobacillus iheyensis</i>	
3541	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CUU5	<i>Oceanobacillus iheyensis</i>	
3542				GbsB
3543				DapA
3544	Transcriptional regulator [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8CUU7	<i>Oceanobacillus iheyensis</i>	RocR
3545	Hypothetical protein [Bdellovibrio bacteriovorus]	UniRef100_Q6MH73	<i>Bdellovibrio bacteriovorus</i>	
3546	NADH-dependent flavin oxidoreductase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ELB6	<i>Oceanobacillus iheyensis</i>	YqiG
3547	Levansucrase precursor [<i>Bacillus subtilis</i>]	UniRef100_P05655	<i>Bacillus subtilis</i>	SacB
3548	Hypothetical protein yveB [<i>Bacillus subtilis</i>]	UniRef100_O07003	<i>Bacillus subtilis</i>	YveB
3549	Hypothetical protein [<i>Pseudomonas aeruginosa</i>]	UniRef100_Q9I5B4	<i>Pseudomonas aeruginosa</i>	
3550	ATP-dependent Clp protease proteolytic subunit [<i>Bacillus subtilis</i>]	UniRef100_P80244	<i>Bacillus subtilis</i>	ClpP
3551	Hypothetical protein yvdD [<i>Bacillus subtilis</i>]	UniRef100_O06986	<i>Bacillus subtilis</i>	YvdD
3552				
3553	Possible metal-dependent phosphohydrolase; possible oxetanocin A resistance protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HM73	<i>Bacillus thuringiensis</i>	YdhJ

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3554	Hypothetical protein yvdB [<i>Bacillus subtilis</i>]	UniRef100_O06984	<i>Bacillus subtilis</i>	YvdB
3555	Hypothetical protein yvdA [<i>Bacillus subtilis</i>]	UniRef100_O06983	<i>Bacillus subtilis</i>	YvdA
3556	Probable 2-ketogluconate reductase [<i>Bacillus subtilis</i>]	UniRef100_O32264	<i>Bacillus subtilis</i>	YvcT
3557	BH0392 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFT5	<i>Bacillus halodurans</i>	YdfE
3558	BH0391 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFT6	<i>Bacillus halodurans</i>	YdfF
3559	Ribonuclease [<i>Paenibacillus polymyxa</i>]	UniRef100_Q7M0X3	<i>Paenibacillus polymyxa</i>	
3560	Barstar [<i>Bacillus amyloliquefaciens</i>]	UniRef100_P11540	<i>Bacillus amyloliquefaciens</i>	
3561	Hypothetical protein yvcN [<i>Bacillus subtilis</i>]	UniRef100_O06977	<i>Bacillus subtilis</i>	YvcN
3562	HPr-like protein crh [<i>Bacillus subtilis</i>]	UniRef100_O06976	<i>Bacillus subtilis</i>	
3563	Hypothetical protein yvcL [<i>Bacillus subtilis</i>]	UniRef100_O06975	<i>Bacillus subtilis</i>	YvcL
3564	Hypothetical UPF0052 protein yvcK [<i>Bacillus subtilis</i>]	UniRef100_O06974	<i>Bacillus subtilis</i>	YvcK
3565	Hypothetical UPF0042 protein yvcJ [<i>Bacillus subtilis</i>]	UniRef100_O06973	<i>Bacillus subtilis</i>	YvcJ
3566	Hypothetical protein yvcI [<i>Bacillus subtilis</i>]	UniRef100_O06972	<i>Bacillus subtilis</i>	YvcI
3567	Thioredoxin reductase [<i>Bacillus subtilis</i>]	UniRef100_P80880	<i>Bacillus subtilis</i>	TrxB
3568	Hypothetical protein yvcD [<i>Bacillus subtilis</i>]	UniRef100_O06968	<i>Bacillus subtilis</i>	YvcD
3569	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ETL1	<i>Oceanobacillus iheyensis</i>	YbbJ
3570	Histidine biosynthesis bifunctional protein hisIE [Includes: Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) (PRA-CH); Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31) (PRA-PH)] [<i>Bacillus subtilis</i>]	UniRef100_O34912	Includes: Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) (PRA-CH); Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31) (PRA-PH)	HisI
3571	Imidazole glycerol phosphate synthase subunit hisF [<i>Bacillus subtilis</i>]	UniRef100_O34727	<i>Bacillus subtilis</i>	HisF
3572	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase [<i>Bacillus subtilis</i>]	UniRef100_O35006	(5-phosphoribosylamino)methylideneamino	HisA
3573	Imidazole glycerol phosphate synthase subunit hisH [<i>Bacillus subtilis</i>]	UniRef100_O34565	<i>Bacillus subtilis</i>	HisH
3574	Imidazoleglycerol-phosphate dehydratase [<i>Bacillus subtilis</i>]	UniRef100_O34683	<i>Bacillus subtilis</i>	HisB
3575	Histidinol dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O34651	<i>Bacillus subtilis</i>	HisD
3576	ATP phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34520	<i>Bacillus subtilis</i>	HisG
3577				HisZ
3578	YvpB [<i>Bacillus subtilis</i>]	UniRef100_O34735	<i>Bacillus subtilis</i>	YvpB
3579	Pectate lyase Pel-28K [<i>Bacillus</i> sp. P-2850]	UniRef100_Q8L0R5	<i>Bacillus</i> sp. P-2850	YvpA
3580	Putative acetyltransferase [<i>Bacillus subtilis</i>]	UniRef100_O34993	<i>Bacillus subtilis</i>	YvoF
3581	Pyrophosphatase ppaX [<i>Bacillus subtilis</i>]	UniRef100_Q9JMQ2	<i>Bacillus subtilis</i>	HprP
3582	YvoD [<i>Bacillus subtilis</i>]	UniRef100_O34382	<i>Bacillus subtilis</i>	YvoD
3583	Prolipoprotein diacylglycerol transferase [<i>Bacillus subtilis</i>]	UniRef100_O34752	<i>Bacillus subtilis</i>	Lgt
3584	HPr kinase/phosphorylase (EC 2.7.1.—) (EC 2.7.4.—) (HPrK/P) (HPr(Ser) kinase/phosphorylase) [<i>Bacillus subtilis</i>]	UniRef100_O34483	<i>Bacillus subtilis</i>	HprK
3585	YfiV protein [<i>Bacillus subtilis</i>]	UniRef100_O31564	<i>Bacillus subtilis</i>	YfiV

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3586	YfiU protein [<i>Bacillus subtilis</i>]	UniRef100_O31563	<i>Bacillus subtilis</i>	YfiU
3587	YvnB [<i>Bacillus subtilis</i>]	UniRef100_O34986	<i>Bacillus subtilis</i>	YvnB
3588	YvlD [<i>Bacillus subtilis</i>]	UniRef100_O34648	<i>Bacillus subtilis</i>	YvlD
3589	YvlC [<i>Bacillus subtilis</i>]	UniRef100_O34719	<i>Bacillus subtilis</i>	
3590	YvlB [<i>Bacillus subtilis</i>]	UniRef100_O34628	<i>Bacillus subtilis</i>	YvlB
3591	Hypothetical protein yvlA [<i>Bacillus subtilis</i>]	UniRef100_O34322	<i>Bacillus subtilis</i>	YvlA
3592	Hypothetical protein yvkN [<i>Bacillus subtilis</i>]	UniRef100_O34604	<i>Bacillus subtilis</i>	
3593				
3594	UvrABC system protein A [<i>Bacillus subtilis</i>]	UniRef100_O34863	<i>Bacillus subtilis</i>	UvrA
3595	UvrABC system protein B [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q659H3	<i>Bacillus amyloliquefaciens</i>	UvrB
3596	CsbA protein [<i>Bacillus subtilis</i>]	UniRef100_P37953	<i>Bacillus subtilis</i>	
3597				
3598	YvjD [<i>Bacillus subtilis</i>]	UniRef100_O34375	<i>Bacillus subtilis</i>	YvjD
3599				YvzD
3600	Putative protease [<i>Bacillus subtilis</i>]	UniRef100_O35002	<i>Bacillus subtilis</i>	YvjB
3601	Tartrate dehydrogenase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESH5	<i>Oceanobacillus iheyensis</i>	YcsA
3602	Cell wall-binding protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K6X4	<i>Bacillus halodurans</i>	YvcE
3603	Cell division protein ftsX homolog [<i>Bacillus subtilis</i>]	UniRef100_O34876	<i>Bacillus subtilis</i>	FtsX
3604	Cell division ATP-binding protein [<i>Bacillus subtilis</i>]	UniRef100_O34814	<i>Bacillus subtilis</i>	FtsE
3605	Cytochrome c-551 [<i>Bacillus subtilis</i>]	UniRef100_O34594	<i>Bacillus subtilis</i>	CccB
3606	YvjA [<i>Bacillus subtilis</i>]	UniRef100_O34792	<i>Bacillus subtilis</i>	YvjA
3607	Peptide chain release factor 2 [<i>Bacillus subtilis</i>]	UniRef100_P28367	<i>Bacillus subtilis</i>	PrfB
3608	Preprotein translocase secA subunit [<i>Bacillus subtilis</i>]	UniRef100_P28366	<i>Bacillus subtilis</i>	SecA
3609	Hypothetical protein yvyD [<i>Bacillus subtilis</i>]	UniRef100_P28368	<i>Bacillus subtilis</i>	YvyD
3610				
3611	Flagellar protein fliT [<i>Bacillus subtilis</i>]	UniRef100_P39740	<i>Bacillus subtilis</i>	FliT
3612	Flagellar protein flIS [<i>Bacillus subtilis</i>]	UniRef100_P39739	<i>Bacillus subtilis</i>	FlIS
3613	Flagellar hook-associated protein 2 [<i>Bacillus subtilis</i>]	UniRef100_P39738	<i>Bacillus subtilis</i>	FlID
3614	Hypothetical protein yvyC [<i>Bacillus subtilis</i>]	UniRef100_P39737	<i>Bacillus subtilis</i>	YvyC
3615	Flagellin [<i>Bacillus subtilis</i>]	UniRef100_P02968	<i>Bacillus subtilis</i>	Hag
3616	Carbon storage regulator homolog [<i>Bacillus subtilis</i>]	UniRef100_P33911	<i>Bacillus subtilis</i>	
3617	Transmembrane protein [<i>Bacillus subtilis</i>]	UniRef100_P96503	<i>Bacillus subtilis</i>	YviF
3618	YviE protein [<i>Bacillus subtilis</i>]	UniRef100_P96502	<i>Bacillus subtilis</i>	YviE
3619	Flagellar hook-associated protein 3 [<i>Bacillus subtilis</i>]	UniRef100_P96501	<i>Bacillus subtilis</i>	FlgL
3620	Flagellar hook-associated protein 1 [<i>Bacillus subtilis</i>]	UniRef100_P39810	<i>Bacillus subtilis</i>	FlgK
3621	Hypothetical protein yvyG [<i>Bacillus subtilis</i>]	UniRef100_P39808	<i>Bacillus subtilis</i>	YvyG
3622	Negative regulator of flagellin synthesis [<i>Bacillus subtilis</i>]	UniRef100_P39809	<i>Bacillus subtilis</i>	
3623	Hypothetical protein yvyF [<i>Bacillus subtilis</i>]	UniRef100_P39807	<i>Bacillus subtilis</i>	YvyF
3624				ComFC
3625				
3626				ComFA
3627	DegV protein [<i>Bacillus subtilis</i>]	UniRef100_P32436	<i>Bacillus subtilis</i>	YviA
3628	Transcriptional regulatory protein degU [<i>Bacillus subtilis</i>]	UniRef100_P13800	<i>Bacillus subtilis</i>	DegU
3629				DegS
3630	Hypothetical UPF0029 protein yvyE [<i>Bacillus subtilis</i>]	UniRef100_P32437	<i>Bacillus subtilis</i>	YvyE

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3631	Putative transcriptional regulator [<i>Bacillus subtilis</i>]	UniRef100_P96499	<i>Bacillus subtilis</i>	YvhJ
3632	Hypothetical protein yunG [<i>Bacillus cereus</i> ZK]	UniRef100_Q63CU2	<i>Bacillus cereus</i> ZK	YunG
3633	Probable undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase [<i>Bacillus subtilis</i>]	UniRef100_O34753	<i>Bacillus subtilis</i>	TagO
3634	Putative teichuronic acid biosynthesis glycosyl transferase tuaH [<i>Bacillus subtilis</i>]	UniRef100_O32267	<i>Bacillus subtilis</i>	TuaH
3635	Putative teichuronic acid biosynthesis glycosyl transferase tuaG [<i>Bacillus subtilis</i>]	UniRef100_O32268	<i>Bacillus subtilis</i>	TuaG
3636	Teichuronic acid biosynthesis protein tuaF [<i>Bacillus subtilis</i>]	UniRef100_O32269	<i>Bacillus subtilis</i>	TuaF
3637	Teichuronic acid biosynthesis protein tuaE [<i>Bacillus subtilis</i>]	UniRef100_O32270	<i>Bacillus subtilis</i>	TuaE
3638	Putative UDP-glucose 4-epimerase [<i>Methanococcus jannaschii</i>]	UniRef100_Q57664	<i>Methanococcus jannaschii</i>	GalE
3639	UDP-glucose 6-dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_O32271	<i>Bacillus subtilis</i>	TuaD
3640	Putative teichuronic acid biosynthesis glycosyl transferase tuaC [<i>Bacillus subtilis</i>]	UniRef100_O32272	<i>Bacillus subtilis</i>	TuaC
3641	Teichuronic acid biosynthesis protein tuaB [<i>Bacillus subtilis</i>]	UniRef100_O32273	<i>Bacillus subtilis</i>	TuaB
3642	UPI00003CB85E UniRef100 entry	UniRef100_UPI00003CB85E		TuaA
3643				LytC
3644	Amidase enhancer precursor [<i>Bacillus subtilis</i>]	UniRef100_Q02113	<i>Bacillus subtilis</i>	LytB
3645				LytA
3646	Membrane-bound protein lytR [<i>Bacillus subtilis</i>]	UniRef100_Q02115	<i>Bacillus subtilis</i>	LytR
3647	UDP-N-acetylglucosamine 2-epimerase [<i>Bacillus subtilis</i>]	UniRef100_P39131	<i>Bacillus subtilis</i>	YvyH
3648	UTP--glucose-1-phosphate uridylyltransferase [<i>Bacillus subtilis</i>]	UniRef100_Q05852	<i>Bacillus subtilis</i>	GtaB
3649				TagH
3650	Teichoic acid translocation permease protein tagG [<i>Bacillus subtilis</i>]	UniRef100_P42953	<i>Bacillus subtilis</i>	TagG
3651	Teichoic acid biosynthesis protein [<i>Lactococcus lactis</i>]	UniRef100_Q9CH20	<i>Lactococcus lactis</i>	TagF
3652	CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase [<i>Bacillus subtilis</i>]	UniRef100_P13485	<i>Bacillus subtilis</i>	TagF
3653				TagD
3654				TagA
3655	Putative CDP-glycerol:glycerophosphate glycerophosphotransferase [<i>Bacillus subtilis</i>]	UniRef100_P27621	<i>Bacillus subtilis</i>	TagB
3656	Beta-N-acetylglucosaminidase precursor [<i>Bacillus subtilis</i>]	UniRef100_P39848	<i>Bacillus subtilis</i>	LytD
3657	Mannose-6-phosphate isomerase [<i>Bacillus subtilis</i>]	UniRef100_P39841	<i>Bacillus subtilis</i>	Pmi
3658				
3659	Anaerobic ribonucleoside-triphosphate reductase, putative [<i>Bacillus cereus</i>]	UniRef100_Q733N6	<i>Bacillus cereus</i>	
3660	2-keto-3-deoxygluconate permease [<i>Bacillus subtilis</i>]	UniRef100_P50847	<i>Bacillus subtilis</i>	KdgT
3661	KHG/KDPG aldolase [Includes: 4-hydroxy-2-oxoglutarate aldolase (EC 4.1.3.16) (2-keto-4-hydroxyglutarate aldolase) (KHG-aldolase); 2-dehydro-3-deoxy-phosphogluconate aldolase (EC 4.1.2.14) (Phospho-2-dehydro-3-deoxygluconate aldolase) (Phospho-2-keto-3-d	UniRef100_P50846	Includes: 4-hydroxy-2-oxoglutarate aldolase (EC 4.1.3.16) (2-keto-4-hydroxyglutarate aldolase) (KHG-	KdgA

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
			aldolase); 2-dehydro-3-deoxy-phosphogluconate aldolase (EC 4.1.2.14) (Phospho-2-dehydro-3-deoxygluconate aldolase) (Phospho-2-keto-3-deoxygluconate aldol	
3662	2-dehydro-3-deoxygluconokinase [<i>Bacillus subtilis</i>]	UniRef100_P50845	<i>Bacillus subtilis</i>	KdgK
3663	HTH-type transcriptional regulator kdgR [<i>Bacillus subtilis</i>]	UniRef100_P50844	<i>Bacillus subtilis</i>	KdgR
3664				KduI
3665	2-deoxy-D-gluconate 3-dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P50842	<i>Bacillus subtilis</i>	KduD
3666	Alcohol dehydrogenase, glutathione-dependent formaldehyde dehydrogenase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HGX1	<i>Bacillus thuringiensis</i>	AdhB
3667	YwtG protein [<i>Bacillus subtilis</i>]	UniRef100_P96742	<i>Bacillus subtilis</i>	YwtG
3668	Hypothetical transcriptional regulator ywtF [<i>Bacillus subtilis</i>]	UniRef100_Q7WY78	<i>Bacillus subtilis</i>	YwtF
3669	YwtE protein [<i>Bacillus subtilis</i>]	UniRef100_P96741	<i>Bacillus subtilis</i>	YwtE
3670	Gamma-DL-glutamyl hydrolase precursor [<i>Bacillus subtilis</i>]	UniRef100_P96740	<i>Bacillus subtilis</i>	YwtD
3671				
3672	YwtB protein [<i>Bacillus subtilis</i>]	UniRef100_P96738	<i>Bacillus subtilis</i>	YwtB
3673	YwtA protein [<i>Bacillus subtilis</i>]	UniRef100_P96737	<i>Bacillus subtilis</i>	YwtA
3674	YwsC protein [<i>Bacillus subtilis</i>]	UniRef100_P96736	<i>Bacillus subtilis</i>	YwsC
3675				RbsR
3676	Ribokinase [<i>Bacillus subtilis</i>]	UniRef100_P36945	<i>Bacillus subtilis</i>	RbsK
3677	High affinity ribose transport protein rbsD [<i>Bacillus subtilis</i>]	UniRef100_P36946	<i>Bacillus subtilis</i>	RbsD
3678	Ribose transport ATP-binding protein rbsA [<i>Bacillus subtilis</i>]	UniRef100_P36947	<i>Bacillus subtilis</i>	RbsA
3679	Ribose transport system permease protein rbsC [<i>Bacillus subtilis</i>]	UniRef100_P36948	<i>Bacillus subtilis</i>	RbsC
3680	D-ribose-binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P36949	<i>Bacillus subtilis</i>	RbsB
3681	Alpha-acetolactate decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_Q04777	<i>Bacillus subtilis</i>	AlsD
3682	Acetolactate synthase [<i>Bacillus subtilis</i>]	UniRef100_Q04789	<i>Bacillus subtilis</i>	AlsS
3683	HTH-type transcriptional regulator alsR [<i>Bacillus subtilis</i>]	UniRef100_Q04778	<i>Bacillus subtilis</i>	AlsR
3684	YwrD protein [<i>Bacillus subtilis</i>]	UniRef100_O05218	<i>Bacillus subtilis</i>	YwrD
3685	YwrC protein [<i>Bacillus subtilis</i>]	UniRef100_O05217	<i>Bacillus subtilis</i>	YwrC
3686	YwrB protein [<i>Bacillus subtilis</i>]	UniRef100_O05216	<i>Bacillus subtilis</i>	YwrB
3687	YwrA protein [<i>Bacillus subtilis</i>]	UniRef100_O05215	<i>Bacillus subtilis</i>	YwrA
3688	Putative protein-tyrosine phosphatase ywqE [<i>Bacillus subtilis</i>]	UniRef100_P96717	<i>Bacillus subtilis</i>	YwqE
3689	Hypothetical protein ywqC [<i>Bacillus subtilis</i>]	UniRef100_P96715	<i>Bacillus subtilis</i>	YwqC
3690	Hypothetical protein ywqB [<i>Bacillus subtilis</i>]	UniRef100_P96714	<i>Bacillus subtilis</i>	YwqB
3691	YwqA protein [<i>Bacillus subtilis</i>]	UniRef100_P94593	<i>Bacillus subtilis</i>	YwqA
3692	YwpJ protein [<i>Bacillus subtilis</i>]	UniRef100_P94592	<i>Bacillus subtilis</i>	YwpJ
3693	YwpI protein [<i>Bacillus subtilis</i>]	UniRef100_P94591	<i>Bacillus subtilis</i>	GlcR
3694	Fructose-6-phosphate aldolase 2 [<i>Escherichia coli</i> O157:H7]	UniRef100_P58424	<i>Escherichia coli</i> O157:H7	YwjH
3695	Hypothetical protein PM1968 [<i>Pasteurella multocida</i>]	UniRef100_Q9CJM9	<i>Pasteurella multocida</i>	FabG
3696	PTS system, sorbitol-specific IIA component [<i>Enterococcus faecalis</i>]	UniRef100_Q82YX6	<i>Enterococcus faecalis</i>	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3697	PTS system, sorbitol-specific IIBC components [<i>Enterococcus faecalis</i>]	UniRef100_Q82YX5	<i>Enterococcus faecalis</i>	
3698	UPI00002F4A3A UniRef100 entry	UniRef100_UPI00002F4A3A		
3699	Sorbitol operon activator [<i>Streptococcus mutans</i>]	UniRef100_Q9X673	<i>Streptococcus mutans</i>	
3700	Putative transcriptional regulator of sorbose uptake and utilization genes [<i>Shigella flexneri</i>]	UniRef100_Q83PA8	<i>Shigella flexneri</i>	DeoR
3701	YwpH protein [<i>Bacillus subtilis</i>]	UniRef100_P94590	<i>Bacillus subtilis</i>	YwpH
3702	Hypothetical protein ywpG [<i>Bacillus subtilis</i>]	UniRef100_P94589	<i>Bacillus subtilis</i>	YwpG
3703	YwpF protein [<i>Bacillus subtilis</i>]	UniRef100_P94588	<i>Bacillus subtilis</i>	YwpF
3704	Large-conductance mechanosensitive channel [<i>Bacillus subtilis</i>]	UniRef100_P94585	<i>Bacillus subtilis</i>	MscL
3705	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.—) ((3R)-hydroxymyristoyl ACP dehydrase) [<i>Bacillus anthracis</i>]	UniRef100_Q81JE0	acyl carrier protein	YwpB
3706	Flagellar hook-basal body complex protein flhP [<i>Bacillus subtilis</i>]	UniRef100_P39753	<i>Bacillus subtilis</i>	FlhP
3707	Flagellar hook-basal body complex protein flhO [<i>Bacillus subtilis</i>]	UniRef100_P39752	<i>Bacillus subtilis</i>	FlhO
3708	MreB-like protein [<i>Bacillus subtilis</i>]	UniRef100_P39751	<i>Bacillus subtilis</i>	Mbl
3709	Stage III sporulation protein D [<i>Bacillus subtilis</i>]	UniRef100_P15281	<i>Bacillus subtilis</i>	
3710				
3711	Hypothetical protein ywoH [<i>Bacillus subtilis</i>]	UniRef100_P94578	<i>Bacillus subtilis</i>	YwoH
3712	Hypothetical protein ywoG [<i>Bacillus subtilis</i>]	UniRef100_P94577	<i>Bacillus subtilis</i>	YwoG
3713	Hypothetical protein ywoA [<i>Bacillus subtilis</i>]	UniRef100_P94571	<i>Bacillus subtilis</i>	YwoA
3714	Manganese-dependent inorganic pyrophosphatase [<i>Bacillus cereus</i> ZK]	UniRef100_Q63AC7	<i>Bacillus cereus</i> ZK	PpaC
3715	Probable manganese catalase [<i>Bacillus subtilis</i>]	UniRef100_P80878	<i>Bacillus subtilis</i>	YdbD
3716	Hypothetical protein yqjF [<i>Bacillus subtilis</i>]	UniRef100_P54543	<i>Bacillus subtilis</i>	YqjF
3717	BH2605 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K9P0	<i>Bacillus halodurans</i>	YhcV
3718	Hypothetical protein yqxL [<i>Bacillus subtilis</i>]	UniRef100_P40948	<i>Bacillus subtilis</i>	YqxL
3719	Hypothetical UPF0053 protein yqhB [<i>Bacillus subtilis</i>]	UniRef100_P54505	<i>Bacillus subtilis</i>	YqhB
3720	Ammonium transporter nrgA [<i>Bacillus subtilis</i>]	UniRef100_Q07429	<i>Bacillus subtilis</i>	NrgA
3721	Nitrogen regulatory PII-like protein [<i>Bacillus subtilis</i>]	UniRef100_Q07428	<i>Bacillus subtilis</i>	NrgB
3722	ATP-dependent Clp protease proteolytic subunit [<i>Bacillus cereus</i>]	UniRef100_Q81CH1	<i>Bacillus cereus</i>	ClpP
3723	RNA polymerase ECF-type sigma factor [<i>Bacillus cereus</i> ZK]	UniRef100_Q63AF9	<i>Bacillus cereus</i> ZK	SigM
3724	Hypothetical protein ywnI [<i>Bacillus subtilis</i>]	UniRef100_P71044	<i>Bacillus subtilis</i>	SpoIIQ
3725	Hypothetical protein ywnH [<i>Bacillus subtilis</i>]	UniRef100_P71043	<i>Bacillus subtilis</i>	YwnH
3726	Hypothetical protein ywnF [<i>Bacillus subtilis</i>]	UniRef100_P71041	<i>Bacillus subtilis</i>	YwnF
3727	Probable cardiolipin synthetase 2 [<i>Bacillus subtilis</i>]	UniRef100_P71040	<i>Bacillus subtilis</i>	YwnE
3728	UPI00003CB764 UniRef100 entry	UniRef100_UPI00003CB764		YhcA
3729	Antibiotic/siderophore biosynthesis protein [<i>Bacillus cereus</i>]	UniRef100_Q81DP7	<i>Bacillus cereus</i>	
3730	Probable serine activating enzyme [<i>Bacillus subtilis</i>]	UniRef100_P45745	<i>Bacillus subtilis</i>	DhbF
3731	Isochorismatase [<i>Bacillus subtilis</i>]	UniRef100_P45743	<i>Bacillus subtilis</i>	DhbB
3732	2,3-dihydroxybenzoate-AMP ligase [<i>Bacillus subtilis</i>]	UniRef100_P40871	<i>Bacillus subtilis</i>	DhbE

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3733	Isochorismate synthase dhbC [<i>Bacillus subtilis</i>]	UniRef100_P45744	<i>Bacillus subtilis</i>	DhbC
3734	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P39071	<i>Bacillus subtilis</i>	DhbA
3735	YuiI protein [<i>Bacillus subtilis</i>]	UniRef100_O32102	<i>Bacillus subtilis</i>	YuiI
3736	Iron-uptake system permease protein feuC [<i>Bacillus subtilis</i>]	UniRef100_P40411	<i>Bacillus subtilis</i>	FeuC
3737	Iron-uptake system permease protein feuB [<i>Bacillus subtilis</i>]	UniRef100_P40410	<i>Bacillus subtilis</i>	FeuB
3738	Iron-uptake system binding protein precursor [<i>Bacillus subtilis</i>]	UniRef100_P40409	<i>Bacillus subtilis</i>	FeuA
3739	Putative HTH-type transcriptional regulator ybbB [<i>Bacillus subtilis</i>]	UniRef100_P40408	<i>Bacillus subtilis</i>	YbbB
3740	Hypothetical protein ywnG [<i>Bacillus subtilis</i>]	UniRef100_P71042	<i>Bacillus subtilis</i>	YwnC
3741	UPI00003CC069 UniRef100 entry	UniRef100_UPI00003CC069		
3742	IS1627s1-related, transposase [<i>Bacillus anthracis</i> str. A2012]	UniRef100_Q7CMD0	<i>Bacillus anthracis</i> str. A2012	
3743	Hypothetical protein ywnB [<i>Bacillus subtilis</i>]	UniRef100_P71037	<i>Bacillus subtilis</i>	YwnB
3744	Hypothetical protein ywnA [<i>Bacillus subtilis</i>]	UniRef100_P71036	<i>Bacillus subtilis</i>	YwnA
3745				
3746	Hypothetical protein ywmF [<i>Bacillus subtilis</i>]	UniRef100_P70963	<i>Bacillus subtilis</i>	YwmF
3747				MoaA
3748	FdhD protein homolog [<i>Bacillus subtilis</i>]	UniRef100_P39756	<i>Bacillus subtilis</i>	FdhD
3749	BH2528 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K9W7	<i>Bacillus halodurans</i>	YbfB
3750	Formate dehydrogenase alpha chain [<i>Bacillus halodurans</i>]	UniRef100_Q9K9W5	<i>Bacillus halodurans</i>	YjgC
3751				YwmD
3752	Hypothetical protein ywmC precursor [<i>Bacillus subtilis</i>]	UniRef100_P70960	<i>Bacillus subtilis</i>	YwmC
3753	Stage II sporulation protein D [<i>Bacillus subtilis</i>]	UniRef100_P07372	<i>Bacillus subtilis</i>	SpoIID
3754	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1 [<i>Bacillus subtilis</i>]	UniRef100_P70965	<i>Bacillus subtilis</i>	MurAA
3755	YwmB protein [<i>Bacillus subtilis</i>]	UniRef100_O32277	<i>Bacillus subtilis</i>	YwmB
3756	YwzB protein [<i>Bacillus subtilis</i>]	UniRef100_O32278	<i>Bacillus subtilis</i>	
3757	ATP synthase epsilon chain [<i>Bacillus subtilis</i>]	UniRef100_P37812	<i>Bacillus subtilis</i>	AtpC
3758	ATP synthase beta chain [<i>Bacillus subtilis</i>]	UniRef100_P37809	<i>Bacillus subtilis</i>	AtpD
3759	ATP synthase gamma chain [<i>Bacillus subtilis</i>]	UniRef100_P37810	<i>Bacillus subtilis</i>	AtpG
3760	ATP synthase alpha chain [<i>Bacillus amyloliquefaciens</i>]	UniRef100_Q659H2	<i>Bacillus amyloliquefaciens</i>	AtpA
3761	ATP synthase delta chain [<i>Bacillus subtilis</i>]	UniRef100_P37811	<i>Bacillus subtilis</i>	AtpH
3762	ATP synthase B chain [<i>Bacillus subtilis</i>]	UniRef100_P37814	<i>Bacillus subtilis</i>	AtpF
3763	ATP synthase a chain [<i>Bacillus subtilis</i>]	UniRef100_P37813	<i>Bacillus subtilis</i>	AtpB
3764				AtpI
3765	Uracil phosphoribosyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P39149	<i>Bacillus subtilis</i>	Upp
3766	Serine hydroxymethyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P39148	<i>Bacillus subtilis</i>	GlyA
3767	Hypothetical protein ywlG [<i>Bacillus subtilis</i>]	UniRef100_P39157	<i>Bacillus subtilis</i>	YwlG
3768	Hypothetical lacA/rpiB family protein ywlF [<i>Bacillus subtilis</i>]	UniRef100_P39156	<i>Bacillus subtilis</i>	YwlF
3769	Putative low molecular weight protein-tyrosine-phosphatase ywlE [<i>Bacillus subtilis</i>]	UniRef100_P39155	<i>Bacillus subtilis</i>	YwlE
3770	Hypothetical UPF0059 protein ywlD [<i>Bacillus subtilis</i>]	UniRef100_P39154	<i>Bacillus subtilis</i>	YwlD

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3771	Hypothetical protein ywIC [<i>Bacillus subtilis</i>]	UniRef100_P39153	<i>Bacillus subtilis</i>	YwIC
3772	Hypothetical protein ywIB [<i>Bacillus subtilis</i>]	UniRef100_P39152	<i>Bacillus subtilis</i>	YwIB
3773	Stage II sporulation protein R [<i>Bacillus subtilis</i>]	UniRef100_P39151	<i>Bacillus subtilis</i>	SpoIIR
3774	Hypothetical protein yhjI [<i>Bacillus subtilis</i>]	UniRef100_O07563	<i>Bacillus subtilis</i>	GlcP
3775	Hypothetical protein yhjJ [<i>Bacillus subtilis</i>]	UniRef100_O07564	<i>Bacillus subtilis</i>	YhjJ
3776	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81NC8	<i>Bacillus anthracis</i>	YhjK
3777	Hypothetical protein yhjL [<i>Bacillus subtilis</i>]	UniRef100_O07566	<i>Bacillus subtilis</i>	YhjL
3778	Hypothetical protein yhjM [<i>Bacillus subtilis</i>]	UniRef100_O07567	<i>Bacillus subtilis</i>	YhjM
3779	HemK protein homolog [<i>Bacillus subtilis</i>]	UniRef100_P45873	<i>Bacillus subtilis</i>	YwkE
3780	Peptide chain release factor 1 [<i>Bacillus subtilis</i>]	UniRef100_P45872	<i>Bacillus subtilis</i>	PrfA
3781	Hypothetical protein ywkD [<i>Bacillus subtilis</i>]	UniRef100_P45871	<i>Bacillus subtilis</i>	YwkD
3782	Hypothetical protein ywkC [<i>Bacillus subtilis</i>]	UniRef100_P45870	<i>Bacillus subtilis</i>	YwkC
3783	Hypothetical protein ywkB [<i>Bacillus subtilis</i>]	UniRef100_P45869	<i>Bacillus subtilis</i>	YwkB
3784	Probable NAD-dependent malic enzyme 2 [<i>Bacillus subtilis</i>]	UniRef100_P45868	<i>Bacillus subtilis</i>	YwkA
3785	Thymidine kinase [<i>Bacillus subtilis</i>]	UniRef100_Q03221	<i>Bacillus subtilis</i>	Tdk
3786	50S ribosomal protein L31 [<i>Bacillus subtilis</i>]	UniRef100_Q03223	<i>Bacillus subtilis</i>	
3787	Transcription termination factor rho [<i>Bacillus subtilis</i>]	UniRef100_Q03222	<i>Bacillus subtilis</i>	Rho
3788	Hypothetical protein ywjI [<i>Bacillus subtilis</i>]	UniRef100_Q03224	<i>Bacillus subtilis</i>	YwjI
3789	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 2 [<i>Bacillus subtilis</i>]	UniRef100_P19670	<i>Bacillus subtilis</i>	MurAB
3790	Transaldolase [<i>Bacillus subtilis</i>]	UniRef100_P19669	<i>Bacillus subtilis</i>	YwjH
3791	Probable fructose-bisphosphate aldolase 1 [<i>Bacillus subtilis</i>]	UniRef100_P13243	<i>Bacillus subtilis</i>	FbaA
3792	Sporulation initiation phosphotransferase F [<i>Bacillus subtilis</i>]	UniRef100_P06628	<i>Bacillus subtilis</i>	SpoOF
3793	Hypothetical protein ywjG [<i>Bacillus subtilis</i>]	UniRef100_P06629	<i>Bacillus subtilis</i>	YwjG
3794				PyrG
3795	DNA-directed RNA polymerase delta subunit [<i>Bacillus subtilis</i>]	UniRef100_P12464	<i>Bacillus subtilis</i>	RpoE
3796	Acyl-CoA dehydrogenase [<i>Bacillus cereus</i>]	UniRef100_Q814S9	<i>Bacillus cereus</i>	AcdA
3797	Acyl-CoA dehydrogenase [<i>Bacillus anthracis</i>]	UniRef100_Q81JV7	<i>Bacillus anthracis</i>	MmgC
3798	UPI00003CC303 UniRef100 entry	UniRef100_UPI00003CC303		MmgB
3799	Acetyl-CoA acetyltransferase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EM47	<i>Oceanobacillus iheyensis</i>	MmgA
3800	Hypothetical protein ywjF [<i>Bacillus subtilis</i>]	UniRef100_P45866	<i>Bacillus subtilis</i>	YwjF
3801				YwjE
3802	Hypothetical protein ywjC [<i>Bacillus subtilis</i>]	UniRef100_P45863	<i>Bacillus subtilis</i>	
3803	Arginyl-tRNA synthetase [<i>Bacillus subtilis</i>]	UniRef100_P46906	<i>Bacillus subtilis</i>	ArgS
3804				YwiB
3805	Agmatinase [<i>Bacillus subtilis</i>]	UniRef100_P70999	<i>Bacillus subtilis</i>	SpeB
3806	Spermidine synthase [<i>Bacillus subtilis</i>]	UniRef100_P70998	<i>Bacillus subtilis</i>	SpeE
3807	UPI00002A1DD1 UniRef100 entry	UniRef100_UPI00002A1DD1		
3808	Hypothetical protein ywhE [<i>Bacillus subtilis</i>]	UniRef100_P70997	<i>Bacillus subtilis</i>	YwhE

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3809	Hypothetical protein ywhD [<i>Bacillus subtilis</i>]	UniRef100_P70996	<i>Bacillus subtilis</i>	YwhD
3810	Hypothetical protein ywhC [<i>Bacillus subtilis</i>]	UniRef100_P70995	<i>Bacillus subtilis</i>	YwhC
3811	Probable tautomerase ywhB [<i>Bacillus subtilis</i>]	UniRef100_P70994	<i>Bacillus subtilis</i>	
3812	Conserved protein [<i>Methanosarcina mazei</i>]	UniRef100_Q8PUC6	<i>Methanosarcina mazei</i>	
3813	Hypothetical protein ywgA [<i>Bacillus subtilis</i>]	UniRef100_P71046	<i>Bacillus subtilis</i>	YwgA
3814	Hypothetical protein ywfo [<i>Bacillus subtilis</i>]	UniRef100_P39651	<i>Bacillus subtilis</i>	Ywfo
3815	YwzC protein [<i>Bacillus subtilis</i>]	UniRef100_O32280	<i>Bacillus subtilis</i>	
3816	Prespore specific transcriptional regulator rsfA [<i>Bacillus subtilis</i>]	UniRef100_P39650	<i>Bacillus subtilis</i>	RsfA
3817	Hypothetical protein yjdJ precursor [<i>Bacillus subtilis</i>]	UniRef100_O31651	<i>Bacillus subtilis</i>	
3818	Hypothetical protein ywfl [<i>Bacillus subtilis</i>]	UniRef100_P39648	<i>Bacillus subtilis</i>	Ywfl
3819	4-hydroxybenzoate 3-monoxygenase [<i>Pseudomonas</i> sp.]	UniRef100_Q59724	<i>Pseudomonas</i> sp.	YhjG
3820	Benzoate transporter protein [<i>Pseudomonas</i> sp. ND6]	UniRef100_Q6XUN5	<i>Pseudomonas</i> sp. ND6	YceI
3821	Putative 3,4-dihydroxyphenylacetate 2,3-dioxygenase [<i>Rhodopseudomonas palustris</i>]	UniRef100_Q6N986	<i>Rhodopseudomonas palustris</i>	
3822	4-oxalocrotonate tautomerase [<i>Pseudomonas putida</i>]	UniRef100_Q9Z431	<i>Pseudomonas putida</i>	
3823	4-oxalocrotonate decarboxylase [<i>Pseudomonas</i> sp. S-47]	UniRef100_Q84I96	<i>Pseudomonas</i> sp. S-47	
3824				DhaS
3825	Putative transcriptional regulator [<i>Agrobacterium tumefaciens</i>]	UniRef100_Q9F7C7	<i>Agrobacterium tumefaciens</i>	KipR
3826	Hypothetical protein SCO6305 [<i>Streptomyces coelicolor</i>]	UniRef100_Q93RT6	<i>Streptomyces coelicolor</i>	
3827	Phosphate acetyltransferase [<i>Bacillus subtilis</i>]	UniRef100_P39646	<i>Bacillus subtilis</i>	Pta
3828	Hypothetical protein ywfI [<i>Bacillus subtilis</i>]	UniRef100_P39645	<i>Bacillus subtilis</i>	YwfI
3829				YuaB
3830	Na ⁺ -transporting ATP synthase [<i>Bacillus halodurans</i>]	UniRef100_Q9KF87	<i>Bacillus halodurans</i>	YubG
3831	Hypothetical protein ywdL [<i>Bacillus subtilis</i>]	UniRef100_P39620	<i>Bacillus subtilis</i>	YwdL
3832	Multidrug ABC transporter, permease [<i>Bacillus cereus</i> ZK]	UniRef100_Q63D13	<i>Bacillus cereus</i> ZK	YfiN
3833	ABC transporter, permease protein [<i>Bacillus cereus</i>]	UniRef100_Q73AB4	<i>Bacillus cereus</i>	YfiM
3834	Multidrug ABC transporter, ATP-binding protein [<i>Bacillus cereus</i> ZK]	UniRef100_Q63D15	<i>Bacillus cereus</i> ZK	YfiL
3835	Sensor histidine kinase, putative [<i>Bacillus cereus</i>]	UniRef100_Q73AB6	<i>Bacillus cereus</i>	YfiJ
3836	DNA-binding response regulator, LuxR family [<i>Bacillus cereus</i>]	UniRef100_Q73AB7	<i>Bacillus cereus</i>	YfiK
3837				YwdK
3838	Putative purine permease ywdJ [<i>Bacillus subtilis</i>]	UniRef100_P39618	<i>Bacillus subtilis</i>	YwdJ
3839	Hypothetical protein ywdI [<i>Bacillus subtilis</i>]	UniRef100_P39617	<i>Bacillus subtilis</i>	YwdI
3840	Probable aldehyde dehydrogenase ywdH [<i>Bacillus subtilis</i>]	UniRef100_P39616	<i>Bacillus subtilis</i>	YwdH
3841	Uracil-DNA glycosylase [<i>Bacillus subtilis</i>]	UniRef100_P39615	<i>Bacillus subtilis</i>	Ung
3842				YwdF
3843	Phosphomethylpyrimidine kinase [<i>Bacillus subtilis</i>]	UniRef100_P39610	<i>Bacillus subtilis</i>	ThiD
3844	Hypothetical protein ywda [<i>Bacillus subtilis</i>]	UniRef100_P39609	<i>Bacillus subtilis</i>	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3845	Sucrose-6-phosphate hydrolase [<i>Bacillus subtilis</i>]	UniRef100_P07819	<i>Bacillus subtilis</i>	SacA
3846	PTS system, sucrose-specific IIBC component [<i>Bacillus subtilis</i>]	UniRef100_P05306	<i>Bacillus subtilis</i>	SacP
3847				SacT
3848				
3849	Minor extracellular protease vpr precursor [<i>Bacillus subtilis</i>]	UniRef100_P29141	<i>Bacillus subtilis</i>	Vpr
3850				
3851				YwcH
3852	Nitro/flavin reductase [<i>Bacillus subtilis</i>]	UniRef100_P39605	<i>Bacillus subtilis</i>	NfrA
3853	N-acyl-L-amino acid amidohydrolase [<i>Bacillus halodurans</i>]	UniRef100_Q9KCF8	<i>Bacillus halodurans</i>	YhaA
3854	Hypothetical protein ywcF [<i>Bacillus subtilis</i>]	UniRef100_P39604	<i>Bacillus subtilis</i>	RodA
3855				RodA
3856	Membrane protein [<i>Clostridium tetani</i>]	UniRef100_Q897A1	<i>Clostridium tetani</i>	
3857	Hypothetical protein ywcE precursor [<i>Bacillus subtilis</i>]	UniRef100_P39603	<i>Bacillus subtilis</i>	
3858	Clostripain [<i>Methanosarcina acetivorans</i>]	UniRef100_Q8TIY6	<i>Methanosarcina acetivorans</i>	
3859	YfmK protein [<i>Bacillus subtilis</i>]	UniRef100_O34536	<i>Bacillus subtilis</i>	YfmK
3860	ABC transporter, ATP-binding/permease protein [<i>Pseudomonas syringae</i>]	UniRef100_Q882Y2	<i>Pseudomonas syringae</i>	YgaD
3861	HlyB/MsbA family ABC transporter [<i>Gloeobacter violaceus</i>]	UniRef100_Q7NIB9	<i>Gloeobacter violaceus</i>	YwjA
3862				
3863				
3864				
3865				
3866	Response regulator [<i>Bacillus cereus</i> ZK]	UniRef100_Q63FE4	<i>Bacillus cereus</i> ZK	ComA
3867				
3868	Quinol oxidase polypeptide IV [<i>Bacillus subtilis</i>]	UniRef100_P34959	<i>Bacillus subtilis</i>	QoxD
3869	Quinol oxidase polypeptide III (EC 1.10.3.—) (Quinol oxidase aa3-600, subunit qoxC) (Oxidase aa(3)-600 subunit 3) [<i>Bacillus subtilis</i>]	UniRef100_P34958	<i>Bacillus subtilis</i>	QoxC
3870	Quinol oxidase polypeptide I (EC 1.10.3.—) (Quinol oxidase aa3-600, subunit qoxB) (Oxidase aa(3)-600 subunit 1) [<i>Bacillus subtilis</i>]	UniRef100_P34956	<i>Bacillus subtilis</i>	QoxB
3871				QoxA
3872	UDP-glucose 4-epimerase [<i>Thermoanaerobacter tengcongensis</i>]	UniRef100_Q8R8R8	<i>Thermoanaerobacter tengcongensis</i>	GalE
3873				
3874				
3875				
3876	Hypothetical protein ywqN [<i>Bacillus subtilis</i>]	UniRef100_P96726	<i>Bacillus subtilis</i>	YwqN
3877	Hypothetical protein ywcB [<i>Bacillus subtilis</i>]	UniRef100_P39600	<i>Bacillus subtilis</i>	
3878	Putative symporter ywcA [<i>Bacillus subtilis</i>]	UniRef100_P39599	<i>Bacillus subtilis</i>	YwcA
3879				XylB
3880	Xylose isomerase [<i>Bacillus</i> sp.]	UniRef100_P54272	<i>Bacillus</i> sp.	XylA
3881				XylR
3882	Thiamine-phosphate pyrophosphorylase [<i>Bacillus subtilis</i>]	UniRef100_P39594	<i>Bacillus subtilis</i>	ThiE
3883	Hydroxyethylthiazole kinase [<i>Bacillus subtilis</i>]	UniRef100_P39593	<i>Bacillus subtilis</i>	ThiM
3884	Putative HTH-type transcriptional regulator ywbI [<i>Bacillus subtilis</i>]	UniRef100_P39592	<i>Bacillus subtilis</i>	YwbI
3885	Putative HTH-type transcriptional regulator ywbI [<i>Bacillus subtilis</i>]	UniRef100_P39592	<i>Bacillus subtilis</i>	YwbI
3886				

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3887				YwqH
3888	Holin-like protein cidA [<i>Bacillus subtilis</i>]	UniRef100_P39591	<i>Bacillus subtilis</i>	YwbH
3889	Hypothetical protein ywbG [<i>Bacillus subtilis</i>]	UniRef100_P39590	<i>Bacillus subtilis</i>	YwbG
3890	HlyC domain protein [<i>Bacillus cereus</i>]	UniRef100_Q731T5	<i>Bacillus cereus</i>	
3891				OpuE
3892	Hypothetical UPF0064 protein ywbD [<i>Bacillus subtilis</i>]	UniRef100_P39587	<i>Bacillus subtilis</i>	YwbD
3893	Hypothetical protein ywbC [<i>Bacillus subtilis</i>]	UniRef100_P39586	<i>Bacillus subtilis</i>	YwbC
3894	Phosphoglycerate dehydrogenase [<i>Bacillus coagulans</i>]	UniRef100_Q8GCC8	<i>Bacillus coagulans</i>	YvcT
3895	Glycerate dehydrogenase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EP33	<i>Oceanobacillus iheyensis</i>	
3896	General stress protein A [<i>Bacillus subtilis</i>]	UniRef100_P25148	<i>Bacillus subtilis</i>	GspA
3897	YdfH protein [<i>Bacillus subtilis</i>]	UniRef100_P96685	<i>Bacillus subtilis</i>	YdfH
3898	YdfI protein [<i>Bacillus subtilis</i>]	UniRef100_P96686	<i>Bacillus subtilis</i>	YdfI
3899	YdfJ protein [<i>Bacillus subtilis</i>]	UniRef100_P96687	<i>Bacillus subtilis</i>	YdfJ
3900	Hypothetical protein ywaF [<i>Bacillus subtilis</i>]	UniRef100_P25149	<i>Bacillus subtilis</i>	YwaF
3901	MFS transporter, phthalate permease family [<i>Pseudomonas syringae</i>]	UniRef100_Q884B5	<i>Pseudomonas syringae</i>	YcbE
3902	Hypothetical protein OB2215 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EP98	<i>Oceanobacillus iheyensis</i>	YitF
3903	2-dehydro-3-deoxyphosphogluconate aldolase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EP99	<i>Oceanobacillus iheyensis</i>	KdgA
3904	2-keto-3-deoxygluconate kinase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPA0	<i>Oceanobacillus iheyensis</i>	KdgK
3905	Transcriptional regulator [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EPA1	<i>Oceanobacillus iheyensis</i>	KipR
3906				YwaD
3907	Hypothetical protein ywoD [<i>Bacillus subtilis</i>]	UniRef100_P94574	<i>Bacillus subtilis</i>	YwoD
3908				
3909	Hypothetical protein ywaC [<i>Bacillus subtilis</i>]	UniRef100_P39583	<i>Bacillus subtilis</i>	YwaC
3910				
3911				DltA
3912	Protein dltB [<i>Bacillus subtilis</i>]	UniRef100_P39580	<i>Bacillus subtilis</i>	DltB
3913	D-alanine-poly(phosphoribitol) ligase subunit 2 [<i>Bacillus subtilis</i>]	UniRef100_P39579	<i>Bacillus subtilis</i>	
3914	Protein dltD precursor [<i>Bacillus subtilis</i>]	UniRef100_P39578	<i>Bacillus subtilis</i>	DltD
3915	Putative branched-chain-amino-acid aminotransferase [<i>Bacillus subtilis</i>]	UniRef100_P39576	<i>Bacillus subtilis</i>	YwaA
3916	Putative macrolide-efflux determinant, YvgJ [<i>Bacillus subtilis</i>]	UniRef100_O32203	<i>Bacillus subtilis</i>	YvgJ
3917	Probable 6-phospho-beta-glucosidase [<i>Bacillus subtilis</i>]	UniRef100_P46320	<i>Bacillus subtilis</i>	LicH
3918	PTS system, lichenan-specific IIA component [<i>Bacillus subtilis</i>]	UniRef100_P46319	<i>Bacillus subtilis</i>	LicA
3919	PTS system, lichenan-specific IIC component [<i>Bacillus subtilis</i>]	UniRef100_P46317	<i>Bacillus subtilis</i>	LicC
3920	PTS system, lichenan-specific IIB component [<i>Bacillus subtilis</i>]	UniRef100_P46318	<i>Bacillus subtilis</i>	
3921	Probable licABCH operon regulator [<i>Bacillus subtilis</i>]	UniRef100_P46321	<i>Bacillus subtilis</i>	LicR
3922				
3923	Putative 3-methyladenine DNA glycosylase [<i>Bacillus subtilis</i>]	UniRef100_P94378	<i>Bacillus subtilis</i>	YxIJ
3924	Transcriptional regulator [<i>Rhizobium loti</i>]	UniRef100_Q988D9	<i>Rhizobium loti</i>	YdhC
3925	Citrate synthase III [<i>Bacillus subtilis</i>]	UniRef100_P45858	<i>Bacillus subtilis</i>	MmgD
3926	2-methylisocitrate dehydratase [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HJ23	<i>Bacillus thuringiensis</i>	MmgE
3927	Putative carboxyvinyl-carboxyphosphonate phosphorylmutase [<i>Bacillus subtilis</i>]	UniRef100_P54528	<i>Bacillus subtilis</i>	YqiQ

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
3928	Chromate transporter [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HIU6	<i>Bacillus thuringiensis</i>	YwrA
3929	BH0575 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KFB0	<i>Bacillus halodurans</i>	
3930				YwjA
3931	YesT protein [<i>Bacillus subtilis</i>]	UniRef100_O31523	<i>Bacillus subtilis</i>	YesT
3932				YweA
3933				YweA
3934	Drug resistance transporter Bcr/CflA subfamily; possible bicyclomycin resistance protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HPJ5	<i>Bacillus thuringiensis</i>	YdgK
3935	YxlG protein [<i>Bacillus subtilis</i>]	UniRef100_P94375	<i>Bacillus subtilis</i>	YxlG
3936	YxlF protein [<i>Bacillus subtilis</i>]	UniRef100_P94374	<i>Bacillus subtilis</i>	YxlF
3937	YxlE protein [<i>Bacillus subtilis</i>]	UniRef100_P94373	<i>Bacillus subtilis</i>	
3938				
3939	Hypothetical protein yxlC [<i>Bacillus subtilis</i>]	UniRef100_P94371	<i>Bacillus subtilis</i>	
3940	RNA polymerase sigma factor sigY [<i>Bacillus subtilis</i>]	UniRef100_P94370	<i>Bacillus subtilis</i>	SigY
3941	YxlH protein [<i>Bacillus subtilis</i>]	UniRef100_P94376	<i>Bacillus subtilis</i>	YxlH
3942	Catalase [<i>Bacillus halodurans</i>]	UniRef100_Q9KDA8	<i>Bacillus halodurans</i>	KatX
3943	Catalase [<i>Bacillus subtilis</i>]	UniRef100_Q9AQQ9	<i>Bacillus subtilis</i>	KatA
3944	Peroxide operon transcriptional regulator [<i>Listeria monocytogenes</i>]	UniRef100_Q71YY4	<i>Listeria monocytogenes</i>	PerR
3945	Ferrochelatase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ERX9	<i>Oceanobacillus iheyensis</i>	HemH
3946	Hypothetical protein yhcI [<i>Bacillus subtilis</i>]	UniRef100_P54593	<i>Bacillus subtilis</i>	YhcI
3947	Bacitracin transport ATP-binding protein bcrA [<i>Bacillus cereus</i>]	UniRef100_Q81AB3	<i>Bacillus cereus</i>	YhcH
3948	MrsE protein [<i>Bacillus</i> sp.]	UniRef100_Q9RC25	<i>Bacillus</i> sp.	
3949				
3950	ABC-2 type transport system ATP-binding protein [<i>Pyrobaculum aerophilum</i>]	UniRef100_Q8ZZG9	<i>Pyrobaculum aerophilum</i>	YdbJ
3951				
3952	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KFL9	<i>Bacillus halodurans</i>	
3953				
3954	ClyA protein [Plasmid pAD1]	UniRef100_Q52055	Plasmid pAD1	IspA
3955	Lantibiotic mersacidin transporter system [<i>Bacillus halodurans</i>]	UniRef100_Q9KFM8	<i>Bacillus halodurans</i>	YgaD
3956				
3957	Lantibiotic mersacidin [<i>Bacillus halodurans</i>]	UniRef100_Q9KFM5	<i>Bacillus halodurans</i>	
3958	Lantibiotic mersacidin modifying enzyme [<i>Bacillus halodurans</i>]	UniRef100_Q9KFM4	<i>Bacillus halodurans</i>	
3959	Pectate lyase Pel-34K [<i>Bacillus</i> sp. P-2850]	UniRef100_Q8L0S6	<i>Bacillus</i> sp. P-2850	PeI
3960	Hypothetical UPF0031 protein yxkO [<i>Bacillus subtilis</i>]	UniRef100_P94368	<i>Bacillus subtilis</i>	YxkO
3961	Transport ATP-binding protein cydD [<i>Bacillus subtilis</i>]	UniRef100_P94367	<i>Bacillus subtilis</i>	CydD
3962	Transport ATP-binding protein cydC [<i>Bacillus subtilis</i>]	UniRef100_P94366	<i>Bacillus subtilis</i>	CydC
3963	Cytochrome d ubiquinol oxidase subunit II [<i>Bacillus subtilis</i>]	UniRef100_P94365	<i>Bacillus subtilis</i>	CydB
3964	Cytochrome d ubiquinol oxidase subunit I [<i>Bacillus subtilis</i>]	UniRef100_P94364	<i>Bacillus subtilis</i>	CydA
3965	Hypothetical transport protein ywcJ [<i>Bacillus subtilis</i>]	UniRef100_P39608	<i>Bacillus subtilis</i>	YwcJ
3966	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q735J7	<i>Bacillus cereus</i>	
3967	Probable NAD-dependent malic enzyme 1 [<i>Bacillus subtilis</i>]	UniRef100_P54572	<i>Bacillus subtilis</i>	MleA
3968	Malate-2H(+)/Na(+)-lactate antiporter [<i>Bacillus subtilis</i>]	UniRef100_P54571	<i>Bacillus subtilis</i>	MleN
3969				AnsB

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
3970	L-asparaginase [<i>Bacillus subtilis</i>]	UniRef100_P26900	<i>Bacillus subtilis</i>	AnsA
3971				AnsR
3972	Hypothetical protein ykgB [<i>Bacillus subtilis</i>]	UniRef100_O34499	<i>Bacillus subtilis</i>	YkgB
3973	Hypothetical sensory transduction protein yxdJ [<i>Bacillus subtilis</i>]	UniRef100_P42421	<i>Bacillus subtilis</i>	YxdJ
3974	Hypothetical sensor-like histidine kinase yxdK [<i>Bacillus subtilis</i>]	UniRef100_P42422	<i>Bacillus subtilis</i>	YxdK
3975	Hypothetical ABC transporter ATP-binding protein yxdL [<i>Bacillus subtilis</i>]	UniRef100_P42423	<i>Bacillus subtilis</i>	YxdL
3976	Hypothetical protein yxdM [<i>Bacillus subtilis</i>]	UniRef100_P42424	<i>Bacillus subtilis</i>	YxdM
3977	Hypothetical protein yxeA precursor [<i>Bacillus subtilis</i>]	UniRef100_P54940	<i>Bacillus subtilis</i>	YxeA
3978	YxkH protein [<i>Bacillus subtilis</i>]	UniRef100_P94361	<i>Bacillus subtilis</i>	YxkH
3979	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KFA7	<i>Bacillus halodurans</i>	YdeL
3980	Hypothetical protein orfD [<i>Lactococcus lactis</i>]	UniRef100_O86288	<i>Lactococcus lactis</i>	YdaM
3981	Citrate/malate transporter [<i>Bacillus subtilis</i>]	UniRef100_P94363	<i>Bacillus subtilis</i>	YxkJ
3982	Glycerophosphodiester phosphodiesterase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMK9	<i>Oceanobacillus iheyensis</i>	YhdW
3983	Hypothetical protein yhjA precursor [<i>Bacillus subtilis</i>]	UniRef100_O07555	<i>Bacillus subtilis</i>	
3984				
3985	YfnI [<i>Bacillus subtilis</i>]	UniRef100_O06487	<i>Bacillus subtilis</i>	YfnI
3986	Hypothetical lipoprotein yfjD precursor [<i>Bacillus subtilis</i>]	UniRef100_O31555	<i>Bacillus subtilis</i>	YfjD
3987	YxkD protein [<i>Bacillus subtilis</i>]	UniRef100_P94357	<i>Bacillus subtilis</i>	YxkD
3988	Arginine repressor, argR [<i>Bacillus cereus</i>]	UniRef100_Q81II2	<i>Bacillus cereus</i>	AhrC
3989	Arginine deiminase [<i>Bacillus cereus</i>]	UniRef100_Q73E87	<i>Bacillus cereus</i>	YkgA
3990	Omithine carbamoyltransferase [<i>Bacillus cereus</i>]	UniRef100_Q73E86	<i>Bacillus cereus</i>	ArgF
3991	Arginine/ornithine antiporter [<i>Bacillus cereus</i>]	UniRef100_Q73E85	<i>Bacillus cereus</i>	YvsH
3992	Carbamate kinase [<i>Bacillus cereus</i>]	UniRef100_Q81IH8	<i>Bacillus cereus</i>	
3993	Transcriptional regulator, Crp family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HP25	<i>Bacillus thuringiensis</i>	Fnr
3994	Beta-glucosidase [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ES64	<i>Oceanobacillus iheyensis</i>	YdhP
3995	PTS system, lichenan-specific Iic component [<i>Erwinia carotovora</i>]	UniRef100_Q6D101	<i>Erwinia carotovora</i>	LicC
3996	Hypothetical protein [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HC13	<i>Bacillus thuringiensis</i>	
3997				SigV
3998	Anaerobic ribonucleoside-triphosphate reductase activating protein [<i>Methanobacterium thermoautotrophicum</i>]	UniRef100_O26387	<i>Methanobacterium thermoautotrophicum</i>	
3999	Hypothetical protein yxeC [<i>Bacillus subtilis</i>]	UniRef100_P54942	<i>Bacillus subtilis</i>	
4000				
4001				
4002	Hypothetical protein [<i>Anabaena</i> sp.]	UniRef100_Q9AIM4	<i>Anabaena</i> sp.	FtsH
4003	Peptidase T [<i>Bacillus subtilis</i>]	UniRef100_P55179	<i>Bacillus subtilis</i>	PepT
4004	Beta-fructosidase FruA [<i>Bacillus megaterium</i>]	UniRef100_Q8GM36	<i>Bacillus megaterium</i>	SacA
4005	Sugar transporter FruP [<i>Bacillus megaterium</i>]	UniRef100_Q8GM37	<i>Bacillus megaterium</i>	YwbF
4006	ABC transporter membrane-spanning permease-sugar transport [<i>Streptococcus pneumoniae</i>]	UniRef100_Q8DNN9	<i>Streptococcus pneumoniae</i>	YesQ
4007	ABC transporter membrane-spanning permease-sugar transporter [<i>Streptococcus pneumoniae</i>]	UniRef100_Q8DNN8	<i>Streptococcus pneumoniae</i>	YurN

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
4008				YesO
4009	Repressor FruR [<i>Bacillus megaterium</i>]	UniRef100_Q8GM38	<i>Bacillus megaterium</i>	RbsR
4010	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q73818	<i>Bacillus cereus</i>	
4011				YjeA
4012	Hypothetical protein yxjG [<i>Bacillus subtilis</i>]	UniRef100_P42318	<i>Bacillus subtilis</i>	YxjG
4013	YkcB protein [<i>Bacillus subtilis</i>]	UniRef100_O34575	<i>Bacillus subtilis</i>	YkcB
4014	Putative glycosyl transferase ykcC [<i>Bacillus subtilis</i>]	UniRef100_O34319	<i>Bacillus subtilis</i>	YkcC
4015				GalE
4016				
4017				
4018				
4019	Hypothetical 10.1 kDa protein [<i>Bacillus pseudofirmus</i>]	UniRef100_O50571	<i>Bacillus pseudofirmus</i>	
4020				
4021	Hypothetical transport protein yxjA [<i>Bacillus subtilis</i>]	UniRef100_P42312	<i>Bacillus subtilis</i>	YxjA
4022	Catalase 2 [<i>Bacillus subtilis</i>]	UniRef100_P42234	<i>Bacillus subtilis</i>	KatE
4023	Catalase 2 [<i>Bacillus subtilis</i>]	UniRef100_P42234	<i>Bacillus subtilis</i>	KatE
4024	DNA gyrase inhibitory protein, GyrI [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97D17	<i>Clostridium acetobutylicum</i>	YosT
4025	Probable glycosyl hydrolase [<i>Erwinia carotovora</i>]	UniRef100_Q6D774	<i>Erwinia carotovora</i>	YckE
4026	PTS system, beta-glucoside-specific enzyme II, ABC component [<i>Bacillus halodurans</i>]	UniRef100_Q9KG19	<i>Bacillus halodurans</i>	BglP
4027	Transcription antiterminator licT [<i>Bacillus subtilis</i>]	UniRef100_P39805	<i>Bacillus subtilis</i>	LicT
4028	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ES40	<i>Oceanobacillus iheyensis</i>	YkoC
4029	Cation ABC transporter ATP-binding protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ES39	<i>Oceanobacillus iheyensis</i>	YkoD
4030	Hypothetical conserved protein [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ES38	<i>Oceanobacillus iheyensis</i>	YkoE
4031	Transcriptional activator of extracellular enzyme genes [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8ESZ0	<i>Oceanobacillus iheyensis</i>	TenA
4032	Isocitrate lyase [<i>Bacillus cereus</i>]	UniRef100_Q73C38	<i>Bacillus cereus</i>	YqiQ
4033	Malate synthase A [<i>Bacillus cereus</i>]	UniRef100_Q73C39	<i>Bacillus cereus</i>	
4034				
4035	Hypothetical protein yycA [<i>Bacillus subtilis</i>]	UniRef100_P37483	<i>Bacillus subtilis</i>	YycA
4036	Bacitracin export permease protein bceB [<i>Bacillus subtilis</i>]	UniRef100_O34741	<i>Bacillus subtilis</i>	YtsD
4037	Bacitracin export ATP-binding protein bceA [<i>Bacillus subtilis</i>]	UniRef100_O34697	<i>Bacillus subtilis</i>	YtsC
4038	Sensor protein bceS [<i>Bacillus subtilis</i>]	UniRef100_O35044	<i>Bacillus subtilis</i>	YtsB
4039	Sensory transduction protein bceR [<i>Bacillus subtilis</i>]	UniRef100_O34951	<i>Bacillus subtilis</i>	YtsA
4040	Beta-glucosidase [<i>Bacillus subtilis</i>]	UniRef100_P40740	<i>Bacillus subtilis</i>	BglIH
4041	PTS system, beta-glucoside-specific I/ABC component [<i>Bacillus subtilis</i>]	UniRef100_P40739	<i>Bacillus subtilis</i>	BglP
4042	Hypothetical protein yxeG [<i>Bacillus subtilis</i>]	UniRef100_P54946	<i>Bacillus subtilis</i>	YxeG
4043	Hypothetical protein yxel [<i>Bacillus subtilis</i>]	UniRef100_P54948	<i>Bacillus subtilis</i>	YxeI
4044				
4045				
4046	Hypothetical protein yxiA precursor [<i>Bacillus subtilis</i>]	UniRef100_P42293	<i>Bacillus subtilis</i>	YxiA
4047	ATP-dependent RNA helicase dbpA [<i>Bacillus subtilis</i>]	UniRef100_P42305	<i>Bacillus subtilis</i>	DeaD
4048	Hypothetical protein OB2811 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMN1	<i>Oceanobacillus iheyensis</i>	
4049	Hypothetical protein OB2810 [<i>Oceanobacillus iheyensis</i>]	UniRef100_Q8EMN2	<i>Oceanobacillus iheyensis</i>	

TABLE 1-continued

		Predicted functions		
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
4050				YxiD
4051				
4052	BH4015 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5S3	<i>Bacillus halodurans</i>	YxiB
4053	Pyrimidine-nucleoside phosphorylase [<i>Bacillus subtilis</i>]	UniRef100_P39142	<i>Bacillus subtilis</i>	Pdp
4054	Pyrimidine nucleoside transport protein [<i>Bacillus subtilis</i>]	UniRef100_P39141	<i>Bacillus subtilis</i>	NupC
4055	Deoxyribose-phosphate aldolase [<i>Bacillus subtilis</i>]	UniRef100_P39121	<i>Bacillus subtilis</i>	Dra
4056	Deoxyribonucleoside regulator [<i>Bacillus subtilis</i>]	UniRef100_P39140	<i>Bacillus subtilis</i>	DeoR
4057	Hypothetical protein [<i>Bacillus anthracis</i>]	UniRef100_Q81X49	<i>Bacillus anthracis</i>	
4058	YdaJ protein [<i>Bacillus subtilis</i>]	UniRef100_P96584	<i>Bacillus subtilis</i>	YdaJ
4059	YdaK protein [<i>Bacillus subtilis</i>]	UniRef100_P96585	<i>Bacillus subtilis</i>	YdaK
4060	YdaL protein [<i>Bacillus subtilis</i>]	UniRef100_O31487	<i>Bacillus subtilis</i>	YdaL
4061	YdaM protein [<i>Bacillus subtilis</i>]	UniRef100_P96587	<i>Bacillus subtilis</i>	YdaM
4062	YdaN protein [<i>Bacillus subtilis</i>]	UniRef100_O31488	<i>Bacillus subtilis</i>	YdaN
4063	Valine-pyruvate aminotransferase [<i>Anabaena</i> sp.]	UniRef100_Q8YTB2	<i>Anabaena</i> sp.	AlaT
4064	Hypothetical protein [<i>Leifsonia xyli</i>]	UniRef100_Q6AEW0	<i>Leifsonia xyli</i>	
4065	Hypothetical protein [<i>Leifsonia xyli</i>]	UniRef100_Q6AEW1	<i>Leifsonia xyli</i>	YwfB
4066	Similarities with putative carboxylase [<i>Photobacterium luminescens</i>]	UniRef100_Q7NA29	<i>Photobacterium luminescens</i>	YwfE
4067	Transporter, Drug/Metabolite Exporter family [<i>Bacillus cereus</i> ZK]	UniRef100_Q63AK7	<i>Bacillus cereus</i> ZK	YoaV
4068	Probable fructose-bisphosphate aldolase 2 [<i>Bacillus subtilis</i>]	UniRef100_P42420	<i>Bacillus subtilis</i>	FbaB
4069	IolI protein [<i>Bacillus subtilis</i>]	UniRef100_P42419	<i>Bacillus subtilis</i>	IolI
4070	IolH protein [<i>Bacillus subtilis</i>]	UniRef100_P42418	<i>Bacillus subtilis</i>	IolH
4071	Myo-inositol dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_Q6B6R7	<i>Bacillus subtilis</i>	Idh
4072	Myo-inositol transport protein [<i>Bacillus subtilis</i>]	UniRef100_P42417	<i>Bacillus subtilis</i>	IolF
4073	IolE protein [<i>Bacillus subtilis</i>]	UniRef100_P42416	<i>Bacillus subtilis</i>	IolE
4074	Probable malonic semialdehyde oxidative decarboxylase [<i>Bacillus subtilis</i>]	UniRef100_P42415	<i>Bacillus subtilis</i>	IolD
4075	Protein iolC [<i>Bacillus subtilis</i>]	UniRef100_P42414	<i>Bacillus subtilis</i>	IolC
4076	IolB protein [<i>Bacillus subtilis</i>]	UniRef100_P42413	<i>Bacillus subtilis</i>	IolB
4077	Probable methylmalonate-semialdehyde dehydrogenase [acylating] [<i>Bacillus subtilis</i>]	UniRef100_P42412	<i>Bacillus subtilis</i>	MmsA
4078	DNA-binding protein iolR [<i>Bacillus subtilis</i>]	UniRef100_P46337	<i>Bacillus subtilis</i>	IolR
4079	IolS protein [<i>Bacillus subtilis</i>]	UniRef100_P46336	<i>Bacillus subtilis</i>	IolS
4080	UPI00003CB380 UniRef100 entry	UniRef100_UPI00003CB380		GlpQ
4081	Putative glycerol-3-phosphate transporter [<i>Staphylococcus aureus</i>]	UniRef100_Q6GJY0	<i>Staphylococcus aureus</i>	GlpT
4082	Chaperone protein htpG [<i>Bacillus subtilis</i>]	UniRef100_P46208	<i>Bacillus subtilis</i>	HtpG
4083				YfmM
4084	Hypothetical protein yxeH [<i>Bacillus subtilis</i>]	UniRef100_P54947	<i>Bacillus subtilis</i>	YxeH
4085	Glycerol dehydrogenase [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97IL4	<i>Clostridium acetobutylicum</i>	AraM
4086	Phosphoserine phosphatase family protein [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97IL5	<i>Clostridium acetobutylicum</i>	YkrX
4087	BH0833 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9KEL9	<i>Bacillus halodurans</i>	
4088				YxeB
4089				
4090	BH3956 protein [<i>Bacillus halodurans</i>]	UniRef100_Q9K5Y0	<i>Bacillus halodurans</i>	YraH
4091				
4092	Ferrous iron transport protein B [<i>Bacillus cereus</i> ZK]	UniRef100_Q632N2	<i>Bacillus cereus</i> ZK	
4093	Lmo2104 protein [<i>Listeria monocytogenes</i>]	UniRef100_Q929R6	<i>Listeria monocytogenes</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
4094	YfiQ protein [<i>Bacillus subtilis</i>]	UniRef100_Q31559	<i>Bacillus subtilis</i>	YfiQ
4095	ABC transporter permease protein [<i>Bacillus cereus</i>]	UniRef100_Q81E11	<i>Bacillus cereus</i>	YxdM
4096	ABC transporter ATP-binding protein [<i>Bacillus cereus</i>]	UniRef100_Q81D39	<i>Bacillus cereus</i>	YxdL
4097	Sensor histidine kinase [<i>Bacillus cereus</i> ZK]	UniRef100_Q63AY4	<i>Bacillus cereus</i> ZK	YxdK
4098	UPI00003CBD9F UniRef100 entry	UniRef100_UPI00003CBD9F		YxdJ
4099	Hypothetical protein [<i>Bacillus cereus</i>]	UniRef100_Q72X82	<i>Bacillus cereus</i>	
4100	YoaH [<i>Bacillus subtilis</i>]	UniRef100_Q34576	<i>Bacillus subtilis</i>	YoaH
4101	Amino acid transporter [<i>Bacillus halodurans</i>]	UniRef100_Q9K609	<i>Bacillus halodurans</i>	YfiA
4102	Alanine dehydrogenase [<i>Bacillus halodurans</i>]	UniRef100_Q9KAF8	<i>Bacillus halodurans</i>	Ald
4103	Hypothetical family 53 glycosyl hydrolase yvfO precursor [<i>Bacillus subtilis</i>]	UniRef100_O07013	<i>Bacillus subtilis</i>	YvfO
4104	Hypothetical protein yvfN [<i>Bacillus subtilis</i>]	UniRef100_O07012	<i>Bacillus subtilis</i>	LacA
4105	Hypothetical protein yvfM [<i>Bacillus subtilis</i>]	UniRef100_O07011	<i>Bacillus subtilis</i>	YvfM
4106	Maltose/maltodextrin transport system [<i>Bacillus halodurans</i>]	UniRef100_Q9KBA8	<i>Bacillus halodurans</i>	YvfL
4107	Hypothetical protein yvfK [<i>Bacillus subtilis</i>]	UniRef100_O07009	<i>Bacillus subtilis</i>	YvfK
4108				LacR
4109	Galactokinase [<i>Streptococcus gordonii</i>]	UniRef100_Q840N8	<i>Streptococcus gordonii</i>	GalK
4110	UDP-glucose 4-epimerase [<i>Listeria innocua</i>]	UniRef100_Q928B6	<i>Listeria innocua</i>	GalE
4111	UPI00003CB9B6 UniRef100 entry	UniRef100_UPI00003CB9B6		GalT
4112	Transcriptional repressor of the xylose operon [<i>Bacillus halodurans</i>]	UniRef100_Q9KDW7	<i>Bacillus halodurans</i>	XylR
4113				GntR
4114				GntK
4115				GntP
4116				GntZ
4117	Alcohol dehydrogenase [<i>Bacillus cereus</i>]	UniRef100_Q818A4	<i>Bacillus cereus</i>	GbsB
4118	Alkyl hydroperoxide reductase subunit C [<i>Bacillus subtilis</i>]	UniRef100_P80239	<i>Bacillus subtilis</i>	AhpC
4119	NADH dehydrogenase [<i>Bacillus subtilis</i>]	UniRef100_P42974	<i>Bacillus subtilis</i>	AhpF
4120				
4121				
4122				
4123				YvfR
4124	Hypothetical protein yvfS [<i>Bacillus subtilis</i>]	UniRef100_O07017	<i>Bacillus subtilis</i>	YvfS
4125				YvfT
4126	Hypothetical protein yvfU [<i>Bacillus subtilis</i>]	UniRef100_O07019	<i>Bacillus subtilis</i>	YvfU
4127				
4128				
4129				YtrE
4130				
4131				
4132				
4133				
4134	Hypothetical protein [<i>Desulfotalea psychrophila</i>]	UniRef100_Q6ARQ4	<i>Desulfotalea psychrophila</i>	
4135	Hypothetical protein [<i>Desulfotalea psychrophila</i>]	UniRef100_Q6ARQ5	<i>Desulfotalea psychrophila</i>	
4136				
4137				
4138	Putative zinc metallopeptidase [<i>Clostridium tetani</i>]	UniRef100_Q894N2	<i>Clostridium tetani</i>	

TABLE 1-continued

Predicted functions				
SEQ ID NO.	Description	UniRef Accession No.	Organism	<i>Bacillus subtilis</i> homolog (Gene Name)
4139	Type I restriction-modification system restriction subunit [<i>Bdellovibrio bacteriovorus</i>]	UniRef100_Q6MH63	<i>Bdellovibrio bacteriovorus</i>	
4140	Lin0523 protein [<i>Listeria innocua</i>]	UniRef100_Q92ED7	<i>Listeria innocua</i>	
4141				
4142	Hypothetical protein CAC1662 [<i>Clostridium acetobutylicum</i>]	UniRef100_Q97IH9	<i>Clostridium acetobutylicum</i>	
4143	Type I restriction-modification system, M subunit [<i>Desulfovibrio vulgaris</i>]	UniRef100_Q72BC7	<i>Desulfovibrio vulgaris</i>	
4144				
4145				
4146				
4147	UPI00002C7192 UniRef100 entry	UniRef100_UPI00002C7192		
4148	Cassette chromosome recombinase B [<i>Staphylococcus aureus</i>]	UniRef100_Q8RPD2	<i>Staphylococcus aureus</i>	SpoIVCA
4149	Hypothetical UPF0247 protein yyda [<i>Bacillus subtilis</i>]	UniRef100_Q45601	<i>Bacillus subtilis</i>	YydA
4150				
4151				
4152	YycN protein [<i>Bacillus subtilis</i>]	UniRef100_O32293	<i>Bacillus subtilis</i>	YycN
4153				
4154	UPI00003CA59D UniRef100 entry	UniRef100_UPI00003CA59D		
4155	Hypothetical serine protease yyxA [<i>Bacillus subtilis</i>]	UniRef100_P39668	<i>Bacillus subtilis</i>	YyxA
4156	YycJ protein [<i>Bacillus subtilis</i>]	UniRef100_Q45611	<i>Bacillus subtilis</i>	YycJ
4157	YycI protein [<i>Bacillus subtilis</i>]	UniRef100_Q45612	<i>Bacillus subtilis</i>	YycI
4158	YycH protein [<i>Bacillus subtilis</i>]	UniRef100_Q45613	<i>Bacillus subtilis</i>	YycH
4159	Sensor protein yycG [<i>Bacillus subtilis</i>]	UniRef100_Q45614	<i>Bacillus subtilis</i>	YycG
4160	Transcriptional regulatory protein yycF [<i>Bacillus subtilis</i>]	UniRef100_P37478	<i>Bacillus subtilis</i>	YycF
4161	Phosphohydrolase [<i>Bacillus cereus</i>]	UniRef100_Q81103	<i>Bacillus cereus</i>	
4162	Adenylosuccinate synthetase [<i>Bacillus subtilis</i>]	UniRef100_P29726	<i>Bacillus subtilis</i>	PurA
4163				
4164	Replicative DNA helicase [<i>Bacillus subtilis</i>]	UniRef100_P37469	<i>Bacillus subtilis</i>	DnaC
4165	Hypothetical protein yycD [<i>Bacillus subtilis</i>]	UniRef100_P37480	<i>Bacillus subtilis</i>	
4166	Hypothetical protein yyzB [<i>Bacillus subtilis</i>]	UniRef100_O32296	<i>Bacillus subtilis</i>	
4167	Hypothetical transport protein yycB [<i>Bacillus subtilis</i>]	UniRef100_P37482	<i>Bacillus subtilis</i>	YycB
4168	N-acetylglucosamine-6-phosphate deacetylase [<i>Bacillus halodurans</i>]	UniRef100_Q9KFQ7	<i>Bacillus halodurans</i>	NagA
4169				
4170	Transcriptional regulator [<i>Bacillus halodurans</i>]	UniRef100_Q9KFQ9	<i>Bacillus halodurans</i>	NagB YvoA
4171	Hypothetical protein VP0543 [<i>Vibrio parahaemolyticus</i>]	UniRef100_Q87S81	<i>Vibrio parahaemolyticus</i>	YbbI
4172	PTS system, n-acetylglucosamine-specific enzyme II, ABC component [<i>Bacillus halodurans</i>]	UniRef100_Q9KF24	<i>Bacillus halodurans</i>	NagP
4173	50S ribosomal protein L9 [<i>Bacillus subtilis</i>]	UniRef100_P37437	<i>Bacillus subtilis</i>	RplI
4174	Hypothetical protein yybT [<i>Bacillus subtilis</i>]	UniRef100_P37484	<i>Bacillus subtilis</i>	YybT
4175	Hypothetical protein yybS [<i>Bacillus subtilis</i>]	UniRef100_P37485	<i>Bacillus subtilis</i>	YybS
4176	Spore coat protein F precursor [<i>Bacillus subtilis</i>]	UniRef100_P23261	<i>Bacillus subtilis</i>	CotF
4177	Universal stress protein, Usp family [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HIV0	<i>Bacillus thuringiensis</i>	YxiE
4178	Sulfate permease [<i>Bacillus thuringiensis</i>]	UniRef100_Q6HIU9	<i>Bacillus thuringiensis</i>	YbaR
4179	Hypothetical transport protein yybO [<i>Bacillus subtilis</i>]	UniRef100_P37489	<i>Bacillus subtilis</i>	YybO
4180	30S ribosomal protein S18 [<i>Bacillus subtilis</i>]	UniRef100_P21475	<i>Bacillus subtilis</i>	

TABLE 1-continued

Predicted functions				<i>Bacillus subtilis</i> homolog (Gene Name)
SEQ ID NO.	Description	UniRef Accession No.	Organism	
4181	Single-strand binding protein [<i>Bacillus subtilis</i>]	UniRef100_P37455	<i>Bacillus subtilis</i>	Ssb
4182	30S ribosomal protein S6 [<i>Bacillus subtilis</i>]	UniRef100_P21468	<i>Bacillus subtilis</i>	
4183	GTP-dependent nucleic acid-binding protein engD [<i>Bacillus subtilis</i>]	UniRef100_P37518	<i>Bacillus subtilis</i>	YyaF
4184	Lin2921 protein [<i>Listeria innocua</i>]	UniRef100_Q926W9	<i>Listeria innocua</i>	
4185	Hypothetical protein yyaD [<i>Bacillus subtilis</i>]	UniRef100_P37520	<i>Bacillus subtilis</i>	YyaD
4186	Hypothetical protein yyaC [<i>Bacillus subtilis</i>]	UniRef100_P37521	<i>Bacillus subtilis</i>	YyaC
4187	Stage 0 sporulation protein J [<i>Bacillus subtilis</i>]	UniRef100_P26497	<i>Bacillus subtilis</i>	Spo0J
4188				Soj
4189	Hypothetical protein [<i>Burkholderia pseudomallei</i> K96243]	UniRef100_Q63L10	<i>Burkholderia pseudomallei</i> K96243	YdfG
4190	RNA polymerase sigma factor, ECF subfamily [<i>Bacillus cereus</i> ZK]	UniRef100_Q63BC0	<i>Bacillus cereus</i> ZK	SigM
4191	YyaA protein [<i>Bacillus subtilis</i>]	UniRef100_P37524	<i>Bacillus subtilis</i>	YyaA
4192	Methyltransferase gidB [<i>Bacillus subtilis</i>]	UniRef100_P25813	<i>Bacillus subtilis</i>	GidB
4193	Glucose inhibited division protein A [<i>Bacillus subtilis</i>]	UniRef100_P25812	<i>Bacillus subtilis</i>	GidA
4194	Probable tRNA modification GTPase trmE [<i>Bacillus subtilis</i>]	UniRef100_P25811	<i>Bacillus subtilis</i>	ThdF
4195	Jag protein [<i>Bacillus subtilis</i>]	UniRef100_Q01620	<i>Bacillus subtilis</i>	Jag
4196				SpoIIIJ
4197	Ribonuclease P protein component [<i>Bacillus subtilis</i>]	UniRef100_P25814	<i>Bacillus subtilis</i>	RnpA

TABLE 2

Features of the <i>Bacillus licheniformis</i> genome and comparison with genomes of other <i>Bacillus</i> species.						
Feature	<i>B. licheniformis</i>	<i>B. subtilis</i> ^a	<i>B. halodurans</i> ^b	<i>Oceanobacillus iheyensis</i> ^c	<i>B. anthracis</i> ^d	<i>B. cereus</i> ^e
Chromosome Size (bp)	4,222,336	4,214,630	4,202,352	3,630,528	5,227,293	5,426,909
G + C content (mol %)	46.2	43.5	43.7	35.7	35.4	35.4
Protein coding sequences	4197	4106	4066	3496	5508	5366
Average length (bp)	873	896	879	883	800	835
Percent of coding region	87	87	85	85	84	84
Ribosomal RNA operons	7	10	8	7	11	13
Number of tRNAs	81	86	78	69	95	108
Phage-associated genes	71	268	42	27	62	124
Transposase genes of IS-elements	10	0	93	14	18	10

^aKunst, F., Ogasawara, N., Mozser, I., Albertini, A. M., Alloni, G., Azebedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., and Borchert, S. et al. (1997) *Nature* 390, 249-256.

^bTakami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hiram, C., Nakamura, Y., Ogasawara, N. et al. (2000) *Nucleic Acids Res.* 28, 4317-4331.

^cTakami, H., Takaki, Y., and Uchiyama, I. (2002) *Nucleic Acids Res.* 30, 3927-3935.

^dRead, T. D., Peterson, S. N., Tourasse, N., Baillie, L. W., Paulsen, I. T., Nelson, K. E., Tettelin, H., Fouts, D. E., Eisen, J. A., and Gill, S. R. et al. (2003) *Nature* 423, 81-86.

^eIvanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelson, B., Kapatral, V., Bhattacharyya, A., Reznik, G., Mikhailova, N., and Lapidus, A. et al. (2003) *Nature* 423, 87-91.

TABLE 3

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome					
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function	<i>B. subtilis</i> Gene
11	dacA	[1-32]	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 5)	“Molecular Function: serine carboxypeptidase activity (GO:0004185), Biological Process: proteolysis and peptidolysis (GO:006508)”	DacA
45	yabE	[1-32]	conserved hypothetical containing domain DUF348	YabE	YabE
67	divC	[1-66]	cell-division initiation protein	“required for both vegetative and sporulation septum formation, Biological Process: cell cycle (GO:0007049)”	DivC
157	BL01016	[1-28]	N-acetylmuramoyl-L-alanine amidase	“cell wall hydrolase, Molecular Function: N-acetylmuramoyl-L-alanine amidase activity (GO:0008745), Biological Process: peptidoglycan catabolism (GO:0009253)”	CwID
159	gerD	[1-27]	GerD	“germination response to L-alanine and to the combination of glucose, fructose, L-asparagine, and KCl (early stage)”	GerD
161	ybaN	[1-42]	“Polysaccharide deacetylase, Carbohydrate Esterase Family 4”	“Biological Process: carbohydrate metabolism (GO:0005975), Molecular Function: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)”	YbaN
162	pbpX	[1-43]	penicillin-binding protein		PbpX
167	ybbC	[1-23]	conserved hypothetical protein YbbC		YbbC
168	ybbD	[1-27]	“Glycoside hydrolase, family 3”	“Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)”	YbbD
169	ybbE	[1-22]	putative beta-lactamase YbbE		YbbE
227		[1-34]	putative ribose ABC transporter (ribose-binding protein)	“ribose transport,”	RbsB
228	yomI	[1-32]	putative lytic transglycosylase YomI		YomI
247	penP	[1-35]	beta-lactamase precursor	“Molecular Function: aspartic-type endopeptidase activity (GO:0004190), Biological Process: proteolysis and peptidolysis (GO:0006508)”	PenP
249	tatAD	[1-35]	component of the twin-arginine pre-protein translocation pathway	“Biological Process: protein transport (GO:0015031), Cellular Component: integral to membrane (GO:0016021)”	
267		[1-26]	“putative Proteinase inhibitor I4, serpin”	Molecular Function: serine-type endopeptidase inhibitor activity (GO:0004867)	
268	BL01663	[1-22]	Pectin lyase-like		YbdN
287	BL01793	[1-31]	putative lipoprotein		YcdA
303	yvbX	[1-33]	“conserved protein, Glycoside Hydrolase Family 18, YvbX”	“Molecular Function: catalytic activity (GO:0003824), Biological Process: carbohydrate metabolism (GO:0005975)”	YvbX
304		[1-27]	“Chitinase precursor, Glycoside Hydrolase Family 18”	“Molecular Function: catalytic activity (GO:0003824), Biological Process: carbohydrate metabolism (GO:0005975)”	
312	yvcE	[1-31]	“putative peptidoglycan hydrolase, DL-endopeptidase II family”		YvcE
348	yckB	[1-34]	putative extracellular solute-binding protein YckB	“Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810), Cellular Component: periplasmic space (sensu Gram-negative Bacteria) (GO:0030288)”	YckB
353	nucA	[1-37]	nuclease NucA		NucA
359	BL01722	[1-34]	putative extracellular solute-binding protein	“Molecular Function: transporter activity (GO:0005215), Biological	YvfK

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
373		[1-25]	hypothetical protein	Process: transport (GO:0006810)"		
382	BL01829	[1-25]	"putative extracellular solute-binding protein, family 3"	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810), Cellular Component: periplasmic space (sensu Gram-negative Bacteria) (GO:0030288)"		YckK
403	BL01746	[1-28]	ribose ABC transporter (ribose-binding protein)	"ribose transport,"		RbsB
409	phy	[1-30]	phytase	"hydrolysis of phytate into inorganic phosphate and myo-inositol,"		Phy
426	yclQ	[1-25]	Periplasmic binding protein	"Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		YclQ
438		[1-26]	hypothetical protein			
472	BL02821	[1-34]	Cell envelope-related transcriptional attenuator			YvhJ
483		[1-42]	hypothetical protein			
513	ydcC	[1-28]	conserved membrane protein YdcC			YdcC
552	yrhM	[1-68]	YrhM			YrhM
558	ywpE	[1-26]	conserved protein YwpE	"Biological Process: biosynthesis (GO:0009058), Molecular Function: transferase activity (GO:0016740)"		
576		[1-34]	putative transporter	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810), Cellular Component: membrane (GO:0016020)"		
605	yjeAA	[1-37]	conserved hypothetical protein			YjeA
607	amyL	[1-30]	"alpha amylase, Glycoside Hydrolase Family 13"	"Molecular Function: alpha-amylase activity (GO:0004556), Biological Process: carbohydrate metabolism (GO:0005975)"		TreA
610	yvdG	[1-28]	putative maltose/maltodextrin transport system substrate-binding protein	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		YvdG
619		[1-37]	hypothetical protein			
620		[1-28]	hypothetical protein			
621	ydjM	[1-41]	YdjM			YdjM
622	ydjN	[1-21]	YdjN			YdjN
660	yerB	[1-28]	conserved protein YerB			YerB
665	yerH	[1-26]	YerH			YerH
683	BL05063	[1-35]	mannan endo-1,4-beta-mannosidase	Degradation of mannan polysaccharides		
733		[1-29]	putative carboxypeptidase	"Molecular Function: carboxypeptidase A activity (GO:0004182), Biological Process: proteolysis and peptidolysis (GO:0006508)"		
773	yfkD	[1-26]	conserved protein YfkD			YfkD
776	BL03088	[1-24]	"Polysaccharide deacetylase, Carbohydrate Esterase Family 4"	"Biological Process: carbohydrate metabolism (GO:0005975), Molecular Function: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)"		YfjS
784	yfjL	[1-33]	putative transporter YfjL	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810), Cellular Component: membrane (GO:0016020)"		YfjL
787	BL03096	[1-26]	hypothetical protein			

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function	<i>B. subtilis</i> Gene	
813	BL03024	[1-26]	"putative extracellular solute-binding protein, family 1"	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"	YtcQ	
841	appAC	[1-30]	oligopeptide ABC transporter (binding protein)	"required for initiation of sporulation, competence development, and oligopeptide transport, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"	OppA	
866	ssuA	[1-24]	aliphatic sulfonate ABC transporter	(binding lipoprotein)	SsuA	
876		[1-23]	hypothetical protein			
886		[1-23]	hypothetical protein			
892		[1-28]	hypothetical protein		YhcC	
900	BL03178	[1-27]	putative lipoprotein		YhcJ	
902	BL03161	[1-31]	hypothetical protein		YhcM	
904	BL03176	[1-29]	hypothetical protein		YhcN	
905	BL03175	[1-30]	hypothetical protein		YhcP	
934	lytE	[1-26]	cell wall hydrolase phosphatase-associated protein	"cell wall lytic activity, Biological Process: cell wall catabolism (GO:0016998), Biological Process: cell wall catabolism (GO:0016998)"	LytE	
937		[1-26]	putative glucose dehydrogenase			
965	BL02841	[1-41]	"Polysaccharide deacetylase, Carbohydrate Esterase Family 4"	"Biological Process: carbohydrate metabolism (GO:0005975), Molecular Function: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)"	YheN	
996	prsA	[1-30]	molecular chaperone PrsA	"essential for the stability of secreted proteins at stages following translocation across the membrane, Molecular Function: isomerase activity (GO:0016853)"	PrsA	
1001	BL02907	[1-31]	hypothetical protein		YhaH	
1032	yhfQ	[1-32]	putative transferase	"Biological Process: metabolism (GO:0008152), Molecular Function: transferase activity (GO:0016740)"	YhfQ	
1044	epr	[1-27]	extracellular serine protease	"Molecular Function: subtilase activity (GO:0004289), Biological Process: proteolysis and peptidolysis (GO:0006508), Molecular Function: subtilase activity (GO:0004289), Biological Process: proteolysis and peptidolysis (GO:0006508)"	Epr	
1061		[1-24]	hypothetical protein			
1063	msmE	[1-29]	multiple sugar-binding protein MsmE	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"	MsmE	
1073	BL01323	[1-28]	conserved hypothetical protein		LytB	
1077	yvgL	[1-29]	molybdate transport system substrate-binding protein	"Molecular Function: molybdate-transporting ATPase activity (GO:0015412), Biological Process: molybdate ion transport (GO:0015689)"	YvgL	
1113	BL01309	[1-24]	Short-chain dehydrogenase/reductase SDR	"Biological Process: metabolism (GO:0008152), Molecular Function: oxidoreductase activity (GO:0016491)"	YcdF	
1150	appA	[1-32]	oligopeptide ABC transporter (oligopeptide-binding protein)	"oligopeptide transport, Molecular Function: transporter activity	AppA	

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
1156	oppA	[1-29]	oligopeptide ABC transporter (binding protein)	(GO:0005215), Biological Process: transport (GO:0006810)" "required for initiation of sporulation, competence development, and oligopeptide transport, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		OppA
1165	BL03332	[1-31]	conserved hypothetical protein			YflP
1216	abnAA	[1-26]	Glycoside Hydrolase Family 43	"degradation of plant cell wall polysaccharide, Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"		AbnA
1219		[1-30]	hypothetical protein			
1226	BL01957	[1-23]	conserved hypothetical protein			YoeB
1227	BL01953	[1-26]	Peptidoglycan-binding protein	"Biological Process: cell wall catabolism (GO:0016998), Biological Process: cell wall catabolism (GO:0016998)"		YocH
1283	yesO	[1-28]	putative transport system substrate-binding protein YesO	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		YesO
1291	yesW	[1-33]	putative polysaccharide lyase family 11 protein			YesW
1299	lplA	[1-32]	lipoprotein	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		LplA
1311	dppE	[1-27]	dipeptide ABC transporter (dipeptide-binding protein)	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		DppE
1319	pelI	[1-25]	"pectate lyase family 1, PelI"			Pel
1332		[1-30]	hypothetical protein			
1341		[1-27]	hypothetical protein			
1345		[1-27]	hypothetical protein			YonS
1368		[1-28]	hypothetical protein			
1415	pbpC	[1-30]	penicillin-binding protein 3	Molecular Function: penicillin binding (GO:0008658)		PbpC
1448		[1-28]	hypothetical protein			
1475	BL05139	[1-27]	glycerophosphoryl diester phosphodiesterase	"hydrolysis of deacylated phospholipids, Biological Process: glycerol metabolism (GO:0006071), Molecular Function: glycerophosphodiester phosphodiesterase activity (GO:0008889)"		GlpQ
1479		[1-27]	hypothetical protein			
1493		[1-27]	hypothetical protein			
1496	BL03556	[1-25]	putative Cell wall hydrolase			YkvT
1498	ykvV	[1-28]	hypothetical protein	Contains thioredoxin domain 2		YkvV
1501		[1-25]	hypothetical protein			
1512	ykwC	[1-24]	Hypothetical oxidoreductase			YkwC
1513	BL03646	[1-27]	Allergen V5/Tpx-1 related	Cellular Component: extracellular (GO:0005576)		YkwD
1527		[1-28]	hypothetical protein			
1553	BL01601	[1-32]	"putative secretion protein, protein transporter"	"Molecular Function: protein transporter activity (GO:0008565), Biological Process: protein secretion (GO:0009306), Cellular Component: membrane (GO:0016020)"		YknX

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome					
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function	<i>B. subtilis</i> Gene
1565		[1-40]	hypothetical protein		
1576	ykyA	[1-29]	Hypothetical protein ykyA		YkyA
1601	BL02997	[1-27]	hypothetical protein		YlaJ
1617	yIbC	[1-34]	conserved hypothetical protein YIbC	Cellular Component: extracellular (GO:0005576)	YIbC
1625	yIbL	[1-29]	YIbL	Molecular Function: protein binding (GO:0005515)	YIbL
1635	ftsL	[1-59]	cell-division protein	"septum formation (early stage),"	FtsL
1636	pbpB	[1-42]	penicillin-binding protein 2B	"formation of the cell-division septum (late stage),"	PbpB
1644	divIb	[1-57]	cell-division initiation protein	"probably involved in stabilizing or promoting the assembly of the division complex (septum formation),"	DivIb
1645	BL02246	[1-29]	conserved hypothetical protein		YIxW
1646	BL02248	[1-30]	hypothetical protein		YIxX
1650	bprA	[1-28]	bacillopeptidase F	"Molecular Function: subtilase activity (GO:0004289), Biological Process: proteolysis and peptidolysis (GO:0006508)"	Bpr
1651	bprB	[1-31]	bacillopeptidase F	"Molecular Function: subtilase activity (GO:0004289), Biological Process: proteolysis and peptidolysis (GO:0006508)"	Bpr
1752	fliL	[1-32]	flagellar protein	"required for flagellar formation, Biological Process: ciliary/flagellar motility (GO:0001539), Biological Process: chemotaxis (GO:0006935), Cellular Component: flagellar basal body (sensu Bacteria) (GO:0009425)"	FliL
1781	celA	[1-34]	Glycoside Hydrolase Family 9	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"	
1783	celC	[1-34]	Glycoside Hydrolase Family 5	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"	
1784	celD	[1-32]	Glycoside Hydrolase family 5	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"	
1820	BL03655	[1-31]	conserved hypothetical protein		YmdA
1839		[1-29]	hypothetical protein		
1904	yoaO	[1-38]	YoaO		YoaO
1914	yoaW	[1-26]	YoaW		YoaW
1926		[1-38]	phage-like protein		
1929	nucB	[1-34]	nuclease		NucB
1931	BL05188	[1-29]	Peptidoglycan-binding protein	Biological Process: cell wall catabolism (GO:0016998)	YneA
1946	yneN	[1-29]	putative thiol: disulfide interchange protein YneN	"Molecular Function: electron transporter activity (GO:0005489), Biological Process: electron transport (GO:0006118)"	YneN
1983	bglC	[1-50]	"endo-1,4-beta-glucanase, Glycoside hydrolase Family 5"	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975),"	BglC
1990	ywoF	[1-22]	"Polysaccharide Lyase Family 9, YwoF"		YwoF
1995	BL00297	[1-27]	conserved hypothetical protein		YddR
2011	yjmF	[1-31]	Short-chain dehydrogenase/reductase SDR	"Biological Process: metabolism (GO:0008152), Molecular Function: oxidoreductase activity (GO:0016491)"	YjmF
2016	detB	[1-29]	possible C4-dicarboxylate binding protein	"", Biological Process: transport (GO:0006810), Cellular	DctB

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
2028	dacC	[1-29]	penicillin-binding protein (D-alanyl-D-alanine carboxypeptidase)	Component: periplasmic space (sensu Gram-negative Bacteria) (GO:0030288)" "Molecular Function: serine carboxypeptidase activity (GO:0004185), Biological Process: proteolysis and peptidolysis (GO:0006508)"		DacC
2060		[1-37]	hypothetical protein			
2072	BL00899	[1-28]	"Spore germination B3 GerAC like, C-terminal"			YndF
2105		[1-33]	conserved hypothetical protein			
2108	BL01303	[1-26]	Peptidoglycan-binding protein	"Biological Process: cell wall catabolism (GO:0016998), Biological Process: cell wall catabolism (GO:0016998)"		YocH
2123		[1-36]	hypothetical protein			
2132	BL01404	[1-24]	hypothetical protein			YoaH
2152	yvgO	[1-25]	conserved protein YvgO			YvgO
2160	lytF	[1-27]	gamma-D-glutamate-meso-diaminopimelate muropeptidase (major autolysin) (CWBP49')	"cell wall lytic activity, Biological Process: cell wall catabolism (GO:0016998)"		LytF
2167	yoaJ	[1-26]	YoaJ			YoaJ
2177	ctpA	[1-37]	carboxy-terminal processing protease	“, Biological Process: proteolysis and peptidolysis (GO:0006508), Molecular Function: serine-type peptidase activity (GO:0008236)"		CtpA
2180	yodJ	[1-31]	putative carboxypeptidase			YodJ
2207	ypmS	[1-34]	conserved protein YpmS			YpmS
2208	ypmR	[1-24]	conserved protein YpmR	Molecular Function: catalytic activity (GO:0003824)		YpmR
2209	ypmQ	[1-25]	conserved hypothetical YpmQ	Biological Process: electron transport (GO:0006118)		YpmQ
2218	ypjP	[1-30]	conserved hypothetical protein YpjP			YpjP
2266	ponA	[1-63]	"penicillin-binding proteins, Glycosyl Transferase Family 51"	"involved in division septum formation,"		PonA
2271	aspB	[1-23]	aspartate aminotransferase	"Molecular Function: transaminase activity (GO:0008483), Biological Process: biosynthesis (GO:0009058)"		AspB
2272	ypmB	[1-25]	conserved protein YpmB			YpmB
2315	BL02789	[1-29]	hypothetical protein			YphF
2328	sleB	[1-34]	spore cortex-lytic enzyme			SleB
2330	ansZA	[1-18]	putative Asparaginase/glutaminase	"Biological Process: amino acid metabolism (GO:0006520), Biological Process: amino acid metabolism (GO:0006520)"		YccC
2334	BL02228	[1-27]	putative hydrolase	Molecular Function: hydrolase activity (GO:0016787)		YpbG
2344	yocH	[1-31]	putative peptidoglycan hydrolase YocH	"Biological Process: cell wall catabolism (GO:0016998), Biological Process: cell wall catabolism (GO:0016998)"		
2347	BL00652	[1-32]	hypothetical protein			LytR
2348	yheN	[1-53]	"Carbohydrate Esterase Family 4, YheN"	"Biological Process: carbohydrate metabolism (GO:0005975), Molecular Function: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)"		YheN
2349		[1-29]	hypothetical protein			
2360	dacB	[1-28]	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 5*)	"required for spore cortex synthesis, Molecular Function: serine carboxypeptidase activity (GO:0004185), Biological		DacB

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
2371	BL01892	[1-44]	conserved hypothetical protein	Process: proteolysis and peptidolysis (GO:0006508)*		YpuD
2376	BL03279	[1-24]	conserved hypothetical protein			YpuA
2392	dacF	[1-33]	penicillin binding protein (putative D-alanyl-D-alanine carboxypeptidase)	“required for spore cortex synthesis, Molecular Function: serine carboxypeptidase activity (GO:0004185), Biological Process: proteolysis and peptidolysis (GO:0006508)”		DacF
2400	BL01161	[1-28]	“Phosphotransferase system, lactose/cellobiose-specific IIB subunit”	“Molecular Function: sugar porter activity (GO:0005351), Biological Process: phosphoenolpyruvate-dependent sugar phosphotransferase system (GO:0009401)”		
2418	lip	[1-31]	lipase	Molecular Function: catalytic activity (GO:0003824)		Lip
2420	appAA	[1-32]	oligopeptide ABC transporter (oligopeptide-binding protein)	“oligopeptide transport, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)”		AppA
2435		[1-24]	hypothetical protein			
2436		[1-29]	hypothetical protein			
2438	BL00808	[1-38]	hypothetical protein			YmaC
2449		[1-35]	hypothetical protein			
2450		[1-31]	conserved hypothetical protein			SsuA
2456	BL01380	[1-28]	“Phosphotransferase system, lactose/cellobiose-specific IIB subunit”	“Molecular Function: sugar porter activity (GO:0005351), Biological Process: phosphoenolpyruvate-dependent sugar phosphotransferase system (GO:0009401)”		
2481	BL01514	[1-36]	Cell wall hydrolase/autolysin	“Molecular Function: N-acetylmuramoyl-L-alanine amidase activity (GO:0008745), Biological Process: peptidoglycan catabolism (GO:0009253)”		YqiI
2482	yqiH	[1-33]	YqiH			YqiH
2484		[1-26]	hypothetical protein			
2499	spoIIIAH	[1-32]	SpoIIIAH	“mutants block sporulation after engulfment,”		SpoIIIAH
2505	spoIIIAB	[1-23]	SpoIIIAB	“mutants block sporulation after engulfment,”		SpoIIIAB
2517	yqhL	[1-24]	conserved protein YqhL			YqhL
2518	opuAC	[1-31]	glycine betaine ABC transporter (glycine betaine-binding protein)	“glycine betaine transport, Molecular Function: transporter activity (GO:0005215), Molecular Function: binding (GO:0005488), Biological Process: transport (GO:0006810)”		OpuAC
2528	tasA	[1-28]	translocation-dependent antimicrobial spore component			TasA
2530	yqxM	[1-44]	YqxM			YqxM
2531	yqzG	[1-24]	YqzG			YqzG
2533	comGG	[1-38]	probably part of the DNA transport machinery ComGG	“required for exogenous DNA-binding,”		ComGG
2546	yqgU	[1-28]	conserved protein YqgU			YqgU
2548	fhuD	[1-32]	ferrichrome ABC transporter (ferrichrome-binding protein)	“Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)”		FhuD
2560	yqzC	[1-33]	conserved protein YqzC			YqzC
2565	pstS	[1-33]	phosphate ABC transporter (binding protein)	“involved in high-affinity phosphate uptake, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)”		PstS
2566	pbpA	[1-44]	penicillin-binding protein 2A	“involved in the synthesis of peptidoglycan associated with		PbpA

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
				cell wall elongation in spore outgrowth, Molecular Function: penicillin binding (GO:0008658)"		
2571	yqfZ	[1-40]	conserved protein YqfZ			
2616	BL01411	[1-37]	conserved hypothetical protein			YqfA
2632	yqxA	[1-29]	YqxA			YqxA
2633	spoIIP	[1-55]	SpoIIP	"required for dissolution of the septal cell wall,"		SpoIIP
2652	BL02075	[1-36]	putative lipoprotein	Molecular Function: catalytic activity (GO:0003824)		YqeF
2697	appAB	[1-36]	oligopeptide ABC transporter (oligopeptide-binding protein)	"oligopeptide transport, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		AppA
2706		[1-32]	putative phosphoesterase			
2707	BL02428	[1-27]	carboxylesterase family	Molecular Function: catalytic activity (GO:0003824)		PnbA
2710	BL03118	[1-28]	hypothetical protein			YndA
2712	sacC	[1-24]	Glycoside Hydrolase Family 32	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"		SacC
2729	BL00468	[1-33]	Periplasmic iron-binding protein	"Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		YxeB
2745	yrrS	[1-50]	conserved protein YrrS			YrrS
2752	yrrL	[1-44]	conserved protein YrrL			YrrL
2771	yrvJ	[1-29]	N-acetylmuramoyl-L-alanine amidase YrvJ			YrvJ
2793	ymaC	[1-24]	phage-related protein YmaC			YmaC
2797	coxA	[1-27]	spore cortex protein			CoxA
2801	nadB	[1-37]	L-aspartate oxidase	"required for NAD biosynthesis,"		NadB
2818	mreC	[1-33]	cell-shape determining protein	"", Biological Process: regulation of cell shape (GO:0008360)"		MreC
2866		[1-28]	hypothetical protein			
2872	gerM	[1-34]	spore germination protein GerM	"germination (cortex hydrolysis) and sporulation (stage II, multiple polar septa),"		GerM
2898		[1-33]	hypothetical protein			
2911	araN	[1-32]	sugar-binding protein	"L-arabinose transport, Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		AraN
2916	abnAB	[1-28]	Glycoside Hydrolase Family 43	"degradation of plant cell wall polysaccharide, Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)"		AbnA
2939	pelB	[1-31]	"pectate lyase, Polysaccharide Lyase Family 1"			PelB
3005	ytxE	[1-51]	YtxE	Cellular Component: outer membrane (sensu Gram-negative Bacteria) (GO:0009279)		YtxE
3049	ytcQ	[1-30]	putative multiple sugar transport system substrate-binding protein YtcQ	"Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)"		YtcQ
3059	BL00013	[1-30]	putative lipoprotein			YusA
3081	ytlA	[1-33]	putative sulfonate transport system substrate-binding protein YtlA			YtlA
3088	ytkA	[1-29]	conserved protein YtkA			YtkA
3093	BL05310	[1-29]	Periplasmic solute binding protein	"Molecular Function: binding (GO:0005488), Cellular Component: periplasmic space		YcdH

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
3217		[1-22]	hypothetical protein	(sensu Gram-negative Bacteria) (GO:0030288) Molecular Function: catalytic activity (GO:0003824)		
3235	yuiC	[1-34]	conserved protein YuiC			YuiC
3264	yutC	[1-24]	conserved protein YutC			YutC
3266	BL02121	[1-22]	putative metalloproteinase	"Molecular Function: metalloendopeptidase activity (GO:0004222), Biological Process: proteolysis and peptidolysis (GO:0006508)"		YunA
3267		[1-32]	hypothetical protein			
3269	yunB	[1-41]	conserved protein YunB			YunB
3284	yurYA	[1-40]	extracellular ribonuclease	Molecular Function: nuclease activity (GO:0004518)		YurI
3298	yusA	[1-28]	putative ABC transport system substrate-binding protein YusA			YusA
3318	yfiY	[1-29]	ABC transport system substrate-binding protein	"Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		YfiY
3319	yusW	[1-28]	YusW			YusW
3333	gerAC	[1-25]	spore germination protein A3 precursor	"germination response to L-alanine and related amino acids (earliest stage),"		GerAC
3338	yvqG	[1-39]	conserved protein YvqG	"Biological Process: metabolism (GO:0008152), Molecular Function: oxidoreductase activity (GO:0016491)"		YvqG
3349	yvrC	[1-37]	putative iron transport system substrate-binding protein YvrC	"Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		YvrC
3380	BL02684	[1-27]	putative Extracellular solute-binding protein	Molecular Function: transporter activity (GO:0005215)		YvfK
3384	BL02680	[1-31]	putative Glycoside Hydrolase Family 3			YbbD
3387	mntA	[1-28]	manganese transport system substrate-binding protein	"Molecular Function: binding (GO:0005488), Cellular Component: periplasmic space (sensu Gram-negative Bacteria) (GO:0030288)"		YcdH
3429		[1-33]	hypothetical protein			
3465		[1-32]	hypothetical protein			
3492	opuCC	[1-32]	glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)	"high affinity transport of glycine betaine, carnitine, and choline, Molecular Function: transporter activity (GO:0005215), Molecular Function: binding (GO:0005488), Biological Process: transport (GO:0006810)"		OpuCC
3499	BL03470	[1-29]	"Periplasmic binding protein, putative iron transporter"	"Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		YfiY
3510		[1-26]	hypothetical protein			
3547	sacB	[1-30]	"levansucrase, Glycoside Hydrolase Family 68"	"Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: sugar utilization (GO:0007587)"		SacB
3548	levB	[1-33]	"endolevanase, Glycoside Hydrolase Family 32"			YveB
3559		[1-29]	putative ribonuclease	"Molecular Function: RNA binding (GO:0003723), Molecular Function: endoribonuclease activity (GO:0004521)"		
3578	yvpB	[1-39]	putative cysteine protease YvpB	"Molecular Function: cysteine-type endopeptidase activity (GO:0004197), Biological		YvpB

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
3579	yvpA	[1-29]	Polysaccharide Lyase Family 3	Process: proteolysis and peptidolysis (GO:0006508)*		YvpA
3587	yvnB	[1-29]	YvnB			YvnB
3600	ctpB	[1-37]	"Peptidase S41A, C-terminal protease"	"Biological Process: proteolysis and peptidolysis (GO:0006508), Molecular Function: serine-type peptidase activity (GO:0008236)"		YvjB
3602	BL03388	[1-29]	"Spectrin repeat, Rudiment single hybrid motif"	"		YvcE
3643	lytC	[1-14]	N-acetylmuramoyl-L-alanine amidase (major autolysin) (CWB49)	"involved in cell separation, cell wall turnover, antibiotic-induced lysis, motility and general cell lysis induced by sodium azide, Molecular Function: N-acetylmuramoyl-L-alanine amidase activity (GO:0008745), Biological Process: peptidoglycan catabolism"		LytC
3644	lytB	[1-26]	modifier protein of major autolysin	LytC (CWB49)		LytB
3645	BL03297	[1-25]	membrane bound lipoprotein	involved in the secretion of major autolysin		LytA
3646	lytR	[1-32]	membrane-bound protein	"attenuator role for lytABC and lytR expression."		LytR
3656	lytD	[1-29]	"N-acetylglucosaminidase (major autolysin), Glycoside Hydrolase Family 73"	"involved in cell separation, cell wall turnover, antibiotic-induced lysis and motility, Molecular Function: amidase activity (GO:0004040), Biological Process: peptidoglycan catabolism (GO:0009253)"		LytD
3670	pgdS	[1-34]	gamma-DL-glutamyl hydrolase	"gamma-DL-glutamyl hydrolase cleaving the gamma-glutamyl bond only between D- and L-glutamic acids of PGA, DL-endopeptidase II family"		YwtD
3671	ywtC	[1-25]	YwtC			YwtB
3672	pgsAA	[1-47]	poly-gamma-glutamate synthesis protein			RbsB
3680	rbsB	[1-28]	ribose ABC transporter (ribose-binding protein)	"ribose transport,"		
3738	feuA	[1-28]	iron-binding protein	"component of iron-uptake system, Molecular Function: iron ion transporter activity (GO:0005381), Biological Process: high affinity iron ion transport (GO:0006827)"		FeuA
3751	ywmD	[1-25]	YwmD			YwmD
3752	ywmC	[1-24]	YwmC			YwmC
3753	spoIID	[1-33]	SpoIID	"required for complete dissolution of the asymmetric septum,"		SpoIID
3755	ywmB	[1-30]	conserved protein YwmB			YwmB
3773	spoIIR	[1-29]	SpoIIR	required for processing of pro-sigma-E		SpoIIR
3801	ywjE	[1-28]	putative Phospholipase	"Molecular Function: catalytic activity (GO:0003824), Biological Process: metabolism (GO:0008152)"		YwjE
3808	BL03962	[1-41]	"Penicillin-binding protein, Glycosyl Transferase Family 51"	"Molecular Function: penicillin binding (GO:0008658), Biological Process: cell wall biosynthesis (sensu Bacteria) (GO:0009273)"		YwhE
3812		[1-25]	hypothetical protein			
3829	BL03904	[1-32]	hypothetical protein			YuaB
3849	vpr	[1-29]	extracellular serine protease	"Molecular Function: subtilase activity (GO:0004289), Biological Process: proteolysis and peptidolysis (GO:0006508)"		Vpr

TABLE 3-continued

Extracellular proteins predicted in the <i>Bacillus licheniformis</i> genome						
SEQ ID NO.	Gene Name	Signal Peptide Location	Putative Product	Putative Function		<i>B. subtilis</i> Gene
3906	ywaD	[1-31]	Putative aminopeptidase	“Biological Process: proteolysis and peptidolysis (GO:0006508), Molecular Function: peptidase activity (GO:0008233)”		YwaD
3914	dltD	[1-27]	precursor DltD	D-alanine esterification of lipoteichoic acid and wall teichoic acid (D-alanine transfer from undecaprenol-P to the poly(glycerophosphate) chain of LTA)		DltD
3920	licB	[1-24]	phosphotransferase system (PTS) lichenan-specific enzyme IIB component	“Molecular Function: sugar porter activity (GO:0005351), Biological Process: phosphoenolpyruvate-dependent sugar phosphotransferase system (GO:0009401)”		
3933	BL05372	[1-31]	hypothetical protein			YweA
3954	lanP	[1-25]	serine protease	“subtilase activity (GO:0004289), proteolysis and peptidolysis (GO:0006508), lantibiotic leader peptide processing”		IspA
3970	ansA	[1-29]	L-asparaginase	“Molecular Function: asparaginase activity (GO:0004067), Biological Process: amino acid metabolism (GO:0006520)”		AnsA
3977	BL02958	[1-25]	conserved hypothetical protein			YxeA
3978	yxkH	[1-28]	“Polysaccharide deacetylase, Carbohydrate Esterase Family 4”	“Biological Process: carbohydrate metabolism (GO:0005975), Molecular Function: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)”		YxkH
3982	BL02963	[1-30]	Glycerophosphoryl diester phosphodiesterase	“Biological Process: glycerol metabolism (GO:0006071), Molecular Function: glycerophosphodiester phosphodiesterase activity (GO:0008889)”		YhdW
3983	BL05379	[1-28]	hypothetical protein			
3996		[1-73]	hypothetical protein			
4011	BL05383	[1-34]	hypothetical protein			YjeA
4016		[1-31]	hypothetical protein	Biological Process: electron transport (GO:0006118)		
4046	yxiA	[1-27]	“Glycoside hydrolase, family 43”	“Molecular Function: hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553), Biological Process: carbohydrate metabolism (GO:0005975)”		YxiA
4058	ydaJ	[1-29]	putative Glycoside transferase			YdaJ
4080	glpQ	[1-29]	glycerophosphoryl diester phosphodiesterase	“hydrolysis of deacylated phospholipids, Biological Process: glycerol metabolism (GO:0006071), Molecular Function: glycerophosphodiester phosphodiesterase activity (GO:0008889)”		GlpQ
4099	yxeA	[1-34]	conserved hypothetical protein			
4103	yvfO	[1-28]	Glycosyl Hydrolase Family 53			YvfO
4107	cycB	[1-32]	“putative extracellular solute-binding protein, family 1 CycB”	“Molecular Function: transporter activity (GO:0005215), Biological Process: transport (GO:0006810)”		YvfK
4121		[1-28]	hypothetical protein			
4128		[1-34]	hypothetical protein			
4130		[1-42]	hypothetical protein			
4158	yycH	[1-36]	conserved hypothetical YycH			YycH

TABLE 4

Codon Usage Table for Chromosome				
# Codon	Amino acid	Fract	/1000	Number
GCA	A	0.230	18.394	22495
GCC	A	0.272	21.724	26567
GCG	A	0.301	24.082	29451
GCT	A	0.197	15.766	19281
TGC	C	0.662	5.113	6253
TGT	C	0.338	2.613	3195
GAC	D	0.432	22.129	27063
GAT	D	0.568	29.114	35605
GAA	E	0.704	51.721	63253
GAG	E	0.296	21.797	26657
TTC	F	0.354	16.034	19609
TTT	F	0.646	29.223	35738
GGA	G	0.316	22.644	27693
GGC	G	0.374	26.825	32806
GGG	G	0.171	12.287	15026
GGT	G	0.139	9.957	12177
CAC	H	0.386	8.412	10288
CAT	H	0.614	13.369	16350
ATA	I	0.101	7.217	8826
ATC	I	0.487	34.741	42487
ATT	I	0.412	29.376	35926
AAA	K	0.693	49.125	60078
AAG	K	0.307	21.739	26586
CTA	L	0.032	3.106	3798
CTC	L	0.152	14.848	18158
CTG	L	0.263	25.592	31298
CTT	L	0.229	22.330	27309
TTA	L	0.143	13.920	17023
TTG	L	0.181	17.639	21572
ATG	M	1.000	26.541	32458
AAC	N	0.505	19.159	23431
AAT	N	0.495	18.780	22967
CCA	P	0.112	4.189	
CCC	P	0.120	4.477	5475
CCG	P	0.520	19.405	23731
CCT	P	0.248	9.236	11295
CAA	Q	0.465	16.656	20369
CAG	Q	0.535	19.179	23455
AGA	R	0.218	9.611	11754
AGG	R	0.146	6.448	7886
CGA	R	0.074	3.275	4005
CGC	R	0.253	11.173	13664
CGG	R	0.194	8.538	10441
CGT	R	0.114	5.044	6168
AGC	S	0.263	15.893	19436
AGT	S	0.063	3.820	4672
TCA	S	0.202	12.173	14887
TCC	S	0.154	9.287	11357
TCG	S	0.171	10.289	12583
TCT	S	0.147	8.865	10841
ACA	T	0.327	16.969	20752
ACC	T	0.188	9.765	11942
ACG	T	0.373	19.404	23730
ACT	T	0.112	5.833	7134
GTA	V	0.144	9.710	11875
GTC	V	0.361	24.342	29769
GTG	V	0.238	16.042	19619
GTT	V	0.257	17.354	21223
TGG	W	1.000	10.303	12600
TAC	Y	0.398	13.468	16471
TAT	Y	0.602	20.348	24885
TAA	*	0.551	1.962	2400
TAG	*	0.157	0.558	682
TGA	*	0.292	1.041	1273

TABLE 5

Signal Peptide Codon Usage Table				
Codon	Amino acid	Fract	/1000	Number
GCA	A	0.294	42.315	106
GCC	A	0.219	31.537	79
GCG	A	0.242	34.731	87
GCT	A	0.244	35.13	88
TGC	C	0.543	9.98	25
TGT	C	0.457	8.383	21
GAC	D	0.231	2.395	6
GAT	D	0.769	7.984	20
GAA	E	0.784	15.968	40
GAG	E	0.216	4.391	11
TTC	F	0.335	21.956	55
TTT	F	0.665	43.513	109
GGA	G	0.341	22.355	56
GGC	G	0.262	17.166	43
GGG	G	0.183	11.976	30
GGT	G	0.213	13.972	35
CAC	H	0.316	2.395	6
CAT	H	0.684	5.19	13
ATA	I	0.142	9.98	25
ATC	I	0.466	32.735	82
ATT	I	0.392	27.545	69
AAA	K	0.767	61.876	155
AAG	K	0.233	18.762	47
CTA	L	0.05	7.186	18
CTC	L	0.098	13.972	35
CTG	L	0.235	33.533	84
CTT	L	0.16	22.754	57
TTA	L	0.221	31.537	79
TTG	L	0.235	33.533	84
ATG	M	1	51.497	129
AAC	N	0.482	10.778	27
AAT	N	0.518	11.577	29
CCA	P	0.164	4.79	12
CCC	P	0.192	5.589	14
CCG	P	0.452	13.174	33
CCT	P	0.192	5.589	14
CAA	Q	0.475	11.577	29
CAG	Q	0.525	12.774	32
AGA	R	0.333	9.98	25
AGG	R	0.08	2.395	6
CGA	R	0.107	3.194	8
CGC	R	0.16	4.79	12
CGG	R	0.173	5.19	13
CGT	R	0.147	4.391	11
AGC	S	0.171	14.77	37
AGT	S	0.093	7.984	20
TCA	S	0.241	20.758	52
TCC	S	0.148	12.774	32
TCG	S	0.134	11.577	29
TCT	S	0.213	18.363	46
ACA	T	0.262	11.178	28
ACC	T	0.234	9.98	25
ACG	T	0.327	13.972	35
ACT	T	0.178	7.585	19
GTA	V	0.216	14.371	36
GTC	V	0.186	12.375	31
GTG	V	0.335	22.355	56
GTT	V	0.263	17.565	44
TGG	W	1	7.585	19
TAC	Y	0.405	5.988	15
TAT	Y	0.595	8.782	22
TAA	*	0	0	0
TAG	*	0	0	0
TGA	*	0	0	0

Example 6

Identification of Genes Induced by Glucose Limitation Using DNA Microarrays

[0270] *Bacillus licheniformis* ATCC 14580 was grown in duplicate shake flasks containing 100 ml of Spizizen I medium (Anagnostopoulos and Spizizen, 1961, *Journal of Bacteriology* 81: 741-748). The inocula for the shake flasks were obtained from 10 ml of mid-log cells (approximately 80-90 Klett units) and a 0 hour total RNA sample was extracted from 30 ml of the mid-log cells, as described below. Total cellular RNA was also purified from 10 ml of the Spizizen I shake flasks that were sampled at 2, 4, 6 and 8 hours post inoculum. All of the cell samples were mixed with two volumes of RNAProtect RNA stabilizing reagent (QIAGEN, Inc., Valencia, Calif.) for 5 minutes. The cells were subsequently pelleted at 2800×g, 4° C. for 10 minutes, the supernatant was discarded, the cell pellets were frozen at -80° C., and cellular RNA was purified using a FastRNA Pro Blue kit (QBiogene, Carlsbad, Calif.) using the protocol supplied by the manufacturer. The frozen cell pellets from each sample were resuspended in 2 ml RNAPro solution provided in the FastRNA Pro Blue kit and RNA was obtained from two lysis matrix B vials.

[0271] Sixteen replicate cDNA targets for the 0 time point and two technical replicate cDNA targets for each of the rest of the time points were prepared and hybridized to *Bacillus licheniformis* DNA microarrays prepared as described by Berka et al., 2002, *Molecular Microbiology* 43: 1331-1345. Dye swap hybridizations using the zero time point as reference were prepared as described in Hu et al. [In G. Parmigiani, E. S. Garrett, R. A. Irizarry, and S. L. Zeger (eds.) *The Analysis of Gene Expression Data*, pp. 318-319, Springer-Verlag, New York (2003)].

[0272] The arrays were scanned with an Axon 4000B scanner and formatted for analysis with GenePix Pro version 5.0 (Axon Instruments, Inc, Redwood City, Calif.). The fluorescence intensity data from GenePix was imported directly into S-Plus ArrayAnalyzer version 2 software (Insightful Corporation, Seattle, Wash.) for statistical analysis. The raw intensity data were normalized using the lowess (locally weighted regression) function provided in S-Plus ArrayAnalyzer software, and the genes that were differentially expressed at each

time point relative to the time zero reference were identified by employing a multiple comparison t-test with Bonferroni correction as outlined in the user's guide to S-Plus ArrayAnalyzer software (Insightful Corp., Seattle, Wash.). The family-wise error rate (FWER) was set at 0.1. The list of genes at each time point that passed these statistical criteria were used to query a pathway-genome database established using Pathway Tools Suite software (Karp et al., 2002, *Bioinformatics* 18 suppl. 1: S225-S232), and the gene expression profiles were painted onto a schematic of metabolic pathways. In doing so, those pathways that were significantly altered at each time point were identified.

[0273] As a result of this analysis the following observations were made: (1) Early in the cultures, genes encoding enzymes of the glycolytic pathway were slightly induced reflecting the change from partially depleted glucose levels in the inoculum culture to ample glucose in the shake flask medium; (2) As glucose was progressively depleted in the shake flask culture (times 6 and 8 hours), genes encoding glycolytic pathway enzymes were down-regulated; (3) genes that encode enzymes for phenylalanine, tryptophan, histidine, and arginine biosynthetic pathways were up-regulated during stationary phase suggesting that cells had exhausted critical amino acids, and they were turning on the genes for pathways to synthesize more of these essential nutrients; (4) as cells entered stationary phase, cell division, DNA replication, and ribosome production all decreased to a low level. This was reflected in the shut-down of purine and pyrimidine biosynthetic pathways and down-regulation of genes encoding many ribosomal proteins.

[0274] The invention described and claimed herein is not to be limited in scope by the specific aspects herein disclosed, since these aspects are intended as illustrations of several aspects of the invention. Any equivalent aspects are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. In the case of conflict, the present disclosure including definitions will control.

[0275] Various references are cited herein, the disclosures of which are incorporated by reference in their entireties.

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=US20120070881A1>). 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).

1. An isolated gene which encodes a polypeptide having aspartate 1-decarboxylase activity, selected from the group consisting of:

- (a) a gene comprising a nucleotide sequence having at least 90% sequence identity with the nucleotide sequence of SEQ ID NO: 2275; and
- (b) a gene comprising a nucleotide sequence which hybridizes under high stringency conditions with the nucle-

otide sequence of SEQ ID NO: 2275, or the complementary strand thereof, wherein high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and 50% formamide, and washing three times each for 15 minutes using 2×SSC, 0.2% SDS at 65° C.

2. The isolated gene of claim 1, which comprises a nucleotide sequence having at least 90% sequence identity with SEQ ID NO: 2275.

3. The isolated gene of claim 1, which comprises a nucleotide sequence having at least 95% sequence identity with SEQ ID NO: 2275.

4. The isolated gene of claim 1, which comprises a nucleotide sequence having at least 97% sequence identity with SEQ ID NO: 2275.

5. The isolated gene of claim 1, which comprises the nucleotide sequence of SEQ ID NO: 2275.

6. The isolated gene of claim 1, which hybridizes under high stringency conditions with the nucleotide sequence of SEQ ID NO: 2275, or the complementary strand thereof, wherein high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and 50% formamide, and washing three times each for 15 minutes using 2×SSC, 0.2% SDS at 65° C.

7. The isolated gene of claim 1, which hybridizes under very high stringency conditions with the nucleotide sequence of SEQ ID NO: 2275, or the complementary strand thereof, wherein very high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and 50% formamide, and washing three times each for 15 minutes using 2×SSC, 0.2% SDS at 70° C.

8. The isolated gene of claim 1, which is contained in *Bacillus licheniformis* ATCC 14580.

9. A nucleic acid construct comprising the isolated gene of claim 1 operably linked to one or more control sequences that direct the production of the polypeptide in a host cell.

10. A recombinant expression vector comprising the nucleic acid construct of claim 9.

11. A recombinant host cell comprising the nucleic acid construct of claim 9.

12. An isolated gene which encodes a polypeptide having aspartate 1-decarboxylase activity, wherein the polypeptide has at least 90% sequence identity with the amino acid sequence of SEQ ID NO: 6472.

13. The isolated gene of claim 12, wherein the polypeptide has at least 95% sequence identity with the amino acid sequence of SEQ ID NO: 6472.

14. The isolated gene of claim 12, wherein the polypeptide has at least 97% sequence identity with the amino acid sequence of SEQ ID NO: 6472.

15. The isolated gene of claim 12, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 6472.

16. A nucleic acid construct comprising the isolated gene of claim 12 operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.

17. A recombinant expression vector comprising the nucleic acid construct of claim 16.

18. A recombinant host cell comprising the nucleic acid construct of claim 16.

19. A recombinant host cell comprising a polynucleotide which encodes a polypeptide having aspartate 1-decarboxylase activity and at least 90% sequence identity with the amino acid sequence of SEQ ID NO: 6472, wherein the polynucleotide is operably linked to a promoter heterologous to the host cell.

20. The recombinant host cell claim 19, wherein the polypeptide has at least 95% sequence identity with the amino acid sequence of SEQ ID NO: 6472.

* * * * *