Complete genome sequence of *Bacillus cereus* FORC_005, a food-borne pathogen from the soy sauce braised fish-cake with quail-egg

Dong-Hoon Lee¹, Hye Rim Kim²†, Han Young Chung², Jong Gyu Lim², Suyeon Kim², Se Keun Kim², Hye-Jin Ku¹, Heebal Kim⁴, Sangryeol Ryu³, Sang Ho Choi²* and Ju-Hoon Lee¹*

**Abstract**

Due to abundant contamination in various foods, the pathogenesis of *Bacillus cereus* has been widely studied in physiological and molecular level. *B. cereus* FORC_005 was isolated from a Korean side dish, soy sauce braised fish-cake with quail-egg in South Korea. While 21 complete genome sequences of *B. cereus* has been announced to date, this strain was completely sequenced, analyzed, and compared with other complete genome sequences of *B. cereus* to elucidate the distinct pathogenic features of a strain isolated in South Korea. The genomic DNA containing a circular chromosome consists of 5,349,617-bp with a GC content of 35.29 %. It was predicted to have 5170 open reading frames, 106 tRNA genes, and 42 rRNA genes. Among the predicted ORFs, 3892 ORFs were annotated to encode functional proteins (75.28 %) and 1278 ORFs were predicted to encode hypothetical proteins (748 conserved and 530 non-conserved hypothetical proteins). This genome information of *B. cereus* FORC_005 would extend our understanding of its pathogenesis in genomic level for efficient control of its contamination in foods and further food poisoning.

**Keywords:** Complete genome sequence, *Bacillus cereus*, Food-borne pathogen, Food poisoning, Diarrhea

**Introduction**

*Bacillus cereus* is one of the major food-borne pathogens, even though it is usually underreported due to its relatively mild symptoms and short duration [1, 2]. It has been known to contaminate diverse types of foods including meat, milk, eggs, and especially various vegetables. *B. cereus* also has been known to produce several pathogenic compounds and virulence factors including spores, dodecadepsipeptide cereulide, and enterotoxins. *B. cereus* has spore-forming ability so it may survive to cause problems in pasteurization and even sterilization in food processing, and this spore is highly hydrophobic to allow them to adhere to food transfer pipelines [1, 3, 4]. An extracellular protein, dodecadepsipeptide cereulide, is also known to be associated with emesis after food ingestion. In addition, *B. cereus* produces three different enterotoxins including hemolysin BL, nonhemolytic enterotoxin, and cytotoxin K, causing diarrhea after *B. cereus* infection [4, 5].

While the pathogenesis of *B. cereus* has been studied in physiological and molecular levels, characterization and pathogenesis studies of the genomes of *B. cereus* have been recently conducted to extend our understanding about its pathogenicity and virulence factors. To date, 21 different genomes of *B. cereus*, isolated from many other countries, have been completely sequenced and analyzed. However, the complete genome sequence of *B. cereus*, isolated from Korean foods, has never been announced previously. To elucidate the genome sequence and its genomic features of Korean *B. cereus*, its...
genome was completely sequenced, analyzed and compared with previously reported *B. cereus* complete genome sequences. Here, we present the complete genome sequence, annotation data, and genomic features of *B. cereus* FORC_005, which was isolated from a contaminated Korean side dish that caused food-borne illness in South Korea, and its evolutionary relationships with other previously reported complete genome sequences using comparative genomics.

**Organism information**

**Classification and features**

*B. cereus* is a Gram-positive, rod-shaped, motile, and spore-forming bacterium. It is often found in various habitats including soil, water, and even food materials (fresh vegetables and food animals). In particular, this bacterium has been well-known food-borne pathogen causing diarrhea, vomiting, and nausea by enterotoxin production. It is a facultative anaerobe that can survive in the temperature range of 10–50 °C, pH range of 4.9–9.3, and salinity of up to 7.5 % NaCl. *B. cereus* belongs to the family *Bacillaceae*, the order Bacillales, the class Bacilli, and the phylum Firmicutes. In this study, *B. cereus* FORC_005 was isolated from a contaminated Korean side dish, soy sauce braised fish-cake with quail-egg, which was suspected to be an original pathogen of food-borne outbreak in March 2014, by Incheon Health and Environment Institute, South Korea. Morphology observation using Transmission Electron Microscopy (TEM; JEM–2100, JEOL, Tokyo, Japan) showed that the strain FORC_005 is rod-shaped with about 4 μm long and 0.5 μm wide in diameter, and can perform motility with peritrichous flagella, suggesting that the strain FORC_005 has typical morphology of *B. cereus* (Fig. 1 and Table 1). In addition, the 16S rRNA sequence analysis and phylogenetic tree analysis of the strain FORC_005 and other *Bacillus* species revealed that this strain was identified to *B. cereus* (data not shown) and positioned at the group of *B. cereus* in the phylogenetic tree, indicating that this strain indeed belongs to *B. cereus* (Fig. 2).

**Genome project history**

**Genome sequencing information**

The complete genome sequence and annotation data of *B. cereus* FORC_005 have been deposited in the GenBank database under the accession number CP009686. Genome sequencing of the strain FORC_005 is a part of the Food-borne Pathogen Omics Research Center project supported by Ministry of Food and Drug Safety South Korea, which aims for the collection and database construction of complete genome sequences of various food-borne pathogens in South Korea. A summary of the project information and its association with MIGS version 2.0 compliance [6] was presented in Table 2.

**Growth conditions and genomic DNA preparation**

*B. cereus* FORC_005 was aerobically cultivated at 30 °C for 12 h with Brain Heart Infusion (BHI; Difco, Detroit, MI, USA) media, and the cells were harvested by centrifugation at 16,000 x g for 1 min. Total genomic DNA was extracted and purified using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions for Gram-positive bacteria. Bacterial cells (about 2 X 10⁹ CFU) were harvested by centrifuging for 10 min at 5000 x g and pellet was resuspended with 180 μl of enzymatic lysis buffer (20 mM Tris · Cl, pH 8.0, 2 mM sodium EDTA, 1.2 % Triton X-100, and 20 mg/ml lysozyme) and this mixture was incubated at least 30 min at 37 °C. In addition to the mixture, 25 μl of protease K and 200 μl of Buffer AL, which are included in the kit, were mixed by vortexing before incubation at 56 °C for 30 min. And then 200 μl of absolute ethanol was added and mixed thoroughly by vortexing. The mixture was transformed to DNeasy Mini spin column in a new 2 ml collection tube and centrifuged at 6000 x g for 1 min to remove the flow-through. Then, 500 μl of Buffer AW2 was added and centrifuged for 3 min at 20,000 x g to wash the genomic DNA in the
The column was placed into a clean 1.5 ml tube and 200 μl of Buffer AE was directly added onto the DNeasy membrane. After incubation at room temperature for 1 min, the column was centrifuged for 1 min at 6000 x g to elute the purified genomic DNA. The concentration and purity of the purified DNA was determined using a NanoVue spectrophotometer (GE Healthcare, Little Chalfont, UK).

**Genome sequencing and assembly**

The genome sequence was determined using a hybrid-genome sequencing approach with PacBio RS II (Pacific Biosciences, Menlo Park, CA, USA) and Illumina MiSeq (Illumina, San Diego, CA, USA). Library construction for PacBio RS II was initialized by ligating universal hairpin adaptors to both ends of DNA fragments using SMRTbell Template Prep kit 1.0 (Pacific Biosciences), followed by purification using AMPure PB bead system (Pacific Biosciences) for the removal of small fragments sized <1.5 kb. Subsequent DNA polymerase binding with template DNAs was conducted using DNA/Polymerase Binding kit P6 v2 with C2 chemistry (Pacific Biosciences), followed by loading of SMRTbells using MagBeads kit (Pacific Biosciences) for greater number of reads at longer lengths per SMRT Cell. Sequencing was conducted using DNA sequencing Bundle 2.0 (Pacific Biosciences), on the PacBio RS II platform. 300 bp paired-end library for Illumina MiSeq (Illumina) was initialized by simultaneously fragmenting template DNAs and tagging them with sequencing adapters using Nextera DNA Sample Preparation kit and Index kit (Illumina), followed by purification of prepared template DNA fragments using MinElute reaction clean up kit (Illumina) and AMPure XP bead (Beckman Coulter, Menlo Park, CA, USA).

### Table 1: Classification and general features of *Bacillus cereus* strain FORC_005 [6]

| MIGS ID | Property                | Term                                | Evidence code*     |
|---------|-------------------------|-------------------------------------|--------------------|
|         | Classification          | Domain *Bacteria*                   | TAS [7]            |
|         |                         | Phylum *Firmicutes*                 | TAS [8–10]         |
|         |                         | Class *Firmibacteria*              | TAS [8, 9, 11]     |
|         |                         | Order *Bacilloales*                | TAS [8, 9, 12, 13] |
|         |                         | Family *Bacillaceae*               | TAS [8, 9, 13, 14] |
|         |                         | Genus *Bacillus*                   | TAS [8, 9, 13, 15] |
|         |                         | Species *Bacillus cereus*          | TAS [8, 9, 13, 16] |
|         | Strain FORC_005         |                                     |                    |
|         | Gram stain              | Positive                           | TAS [17, 18]       |
|         | Cell shape              | Rod                                | TAS [17, 18]       |
|         | Motility                | Motile with peritrichous flagella  | TAS [17, 18]       |
|         | Sporulation             | Endospore-forming                  | TAS [8]            |
|         | Temperature range       | 10 °C–50 °C                        | TAS [8, 17]        |
|         | Optimum temperature     | 28–35 °C                           | TAS [8]            |
|         | pH range; Optimum       | 4.9–9.3; 6.0–7.0                    | TAS [17]           |
|         | Carbon source           | Glucose, Aesculin                   | TAS [19]           |
|         | Habitat                 | Ubiquitous; especially in soil     | TAS [17, 18]       |
|         | Salinity                | 0–7.5 % NaCl (w/v)                 | TAS [17]           |
|         | Oxygen requirement      | Facultative anaerobes              | TAS [17, 18]       |
|         | Biotic relationship     | free-living                        | NAS                |
|         | Pathogenicity           | Diarrhea, emesis in human          | NAS                |
|         | Geographic location     | Incheon, South Korea               | IDA                |
|         | Isolation               | Korean soy sauce braised fish-cake with quail-egg | IDA |
|         | Sample collection       | March 2014                         | IDA                |
|         | Latitude                | 37.27 N                            | IDA                |
|         | Longitude               | 126.42 E                           | IDA                |

* Evidence codes - IDA inferred from direct assay, TAS traceable author statement (i.e., a direct report exists in the literature), NAS non-traceable author statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [20].
Brea, CA, USA). Sequencing was conducted using MiSeq Reagent kit (600 cycle; Illumina). All kits were used according to the manufacturer’s instructions. Sequencing reads from Illumina MiSeq system were assembled using the CLC Genomics Workbench v7.0.4 (CLC bio, Aarhus, Denmark), and the reads from PacBio system were assembled using the PacBio SMRT Analysis v2.0 (Pacific Biosciences). Finally, the initially assembled scaffolds were gathered and re-assembled to obtain one contig using CLC Genomics Workbench program.

Genome annotation

Initial prediction and annotation of all open reading frames, and tRNA/rRNA gene prediction was carried out using Glimmer3 by the Rapid Annotation using Subsystem Technology server [23], and was confirmed using the GeneMarkS ORF prediction program [24]. Predicted ribosome binding sites by RBSfinder (J. Craig Venter Institute, Rockville, MD, USA) were used to confirm the predicted ORFs. The Global Annotation of Multiplexed On-site Blasted DNA-Sequences program and InterProScan5 program with conserved protein domain databases were used for the annotation of confirmed ORFs [25, 26]. Artemis16 was used for handling of genome sequence and annotated data [27]. The functional categorization and classification of all predicted ORFs were conducted using the RAST server-based SEED viewer and Clusters of Orthologous Groups -based WebMGA programs [28, 29]. Circular genome map, showed in Fig. 3, was generated using GenVision (DNASTAR, Madison, WI, USA) based on all predicted ORFs with COG information, tRNAs and rRNAs, GC-content, and gene cluster information. Detection and identification of virulence factors were carried out using BLAST search with protein sequences of VFs in the database [30]. Signal peptides, transmembrane helices, and Clustered Regularly Interspaced Short Palindromic Repeats were identified by using SignalP server v.4.1 [31], TMHMM server v.2.0 [32], and CRISPRfinder [33], respectively.

Genome properties

Table 3 contains the main FORC_005 statistics. It contains one chromosomal double-stranded DNA and no plasmid. The chromosome consists of 5,349,617 bp in DNA length with the GC content of 35.29 %, containing 5170 ORFs, 106 tRNA genes, and 42 rRNA genes consisting of 14 complete rRNA operons. Among the predicted 5170 ORFs, 3892 ORFs (75.28 %) were annotated to encode functional proteins and 1278 ORFs were hypothetical proteins (748 conserved and 530 non-

---

**Table 2** Project information

| MGIS ID | Property | Term |
|---------|----------|------|
| MGIS 31 | Finishing quality | Finished |
| MGIS-28 | Libraries used | MiSeq 300 bp paired end, PacBio P4-C2 chemistry, size selected 10 kb library 2cell |
| MGIS 29 | Sequencing platforms | Illumina MiSeq, PacBio RS II |
| MGIS 31.2 | Fold coverage | 494.66 |
| MGIS 30 | Assemblers | CLC genomics Workbench v7.0.4, PacBio SMRT Analysis v2.0 |
| MGIS 32 | Gene calling method | GeneMarkS, Glimmer3 |
| Locus Tag | FORCS |
| Genbank ID | CP009686 |
| GenBank Date of Release | April 22, 2015 |
| GOLD ID | Gp0107273 |
| BIOPROJECT | PRJNA261557 |
| MGIS 13 | Source Material Identifier | FORC_005 |
| Project relevance | Agricultural |
In addition, 4202 ORFs (81.28%) were assigned to the related COG functional categories, and are listed in Table 4.

Insights from the genome sequence
Pathogenesis and virulence factors
Frequently, *B. cereus* causes diarrhea and emesis after ingestion of the contaminated food. These food-borne illnesses are reported to be associated with specific toxin genes. The genome analysis of *B. cereus* FORC_005 revealed that there are three major toxins including cytotoxin K, hemolysin BL, and non-hemolytic enterotoxin [5]. These toxins are involved in severe diarrhea after infection of *B. cereus*. Cytotoxin K is encoded by a single gene, *cytK* (FORC5_0979). However, other two toxins are encoded by two different gene clusters, a hemolysin BL gene cluster (*hblABDC*; FORC5_2954 to FORC5_2957) and a non-hemolytic enterotoxin gene cluster (*nheABC*; FORC5_1734 to FORC5_1736)). In addition, hemolysin III (*hlyIII*; FORC5_2063) was detected in the genome for additional hemolysis activity.

Anthrolysin O, one of the cholesterol-dependent cytolysins, was detected in the genome (FORC5_1940), which has been suggested to be a pore-forming protein...
often found in many Gram-positive bacteria. This hemolytic and cytolytic protein was reported to be associated with cholesterol binding in the host cell plasma membrane, pore formation via its oligomerization, and transfer of virulence factors to the host cell cytoplasm [34, 35]. In addition, an internalin (FORC5_1206) was detected in the genome, suggesting that it may play an important role in host cell invasion. The predicted functions of these two host invasive proteins in the genome revealed that regulation and control of this initial step in the occurrence of food-borne pathogenesis and illness may be important for the host protection.

Bacillus has been known to have two kinds of bacterial protection systems from the host cell immune system, including polysaccharide capsule (PSC) and polyglutamic acid capsule [36, 37]. While PSC biosynthesis is common in B. cereus, only PSC biosynthesis gene cluster (FORC5_4952 to FORC5_4971) was detected in the genome of B. cereus FORC_005, suggesting that the strain can protect itself from the host cell immune defense system for its further pathogenesis in the host cells. Therefore, this bacterial protection system is also considered as one of the virulence factors in B. cereus.

Table 3 Genome statistics

| Attribute                        | Value     | % of Total |
|----------------------------------|-----------|------------|
| Genome size (bp)                 | 5,349,617 | 100.00     |
| DNA coding (bp)                  | 4,476,678 | 83.68      |
| DNA G+C (bp)                     | 1,887,736 | 35.29      |
| DNA scaffolds                     | 1         | -          |
| Total genes                      | 5318      | 100.00     |
| Protein coding genes             | 5170      | 97.22      |
| RNA genes                        | 148       | 2.78       |
| Pseudo genes                     | 134       | 2.52       |
| Genes in internal clusters       | 0         | 0.00       |
| Genes with function prediction   | 3892      | 73.19      |
| Genes assigned to COGs           | 4202      | 81.28      |
| Genes with Pfam domains          | 4718      | 88.72      |
| Genes with signal peptides       | 383       | 7.20       |
| Genes with transmembrane helices | 1544      | 29.03      |
| CRISPR repeats                   | 4         | -          |

The total is based on either the size of the genome in base pairs or the total number of total gene in the annotated genome.

Table 4 Number of genes associated with general COG functional categories

| Code | Value | % age | Description                                                                 |
|------|-------|-------|-----------------------------------------------------------------------------|
| J    | 215   | 4.16  | Translation, ribosomal structure and biogenesis                            |
| A    | 1     | 0.02  | RNA processing and modification                                             |
| K    | 378   | 7.31  | Transcription                                                               |
| L    | 165   | 3.19  | Replication, recombination and repair                                       |
| B    | 1     | 0.02  | Chromatin structure and dynamics                                            |
| D    | 46    | 0.89  | Cell cycle control, Cell division, chromosome partitioning                 |
| V    | 105   | 2.03  | Defense mechanisms                                                          |
| T    | 230   | 4.45  | Signal transduction mechanisms                                              |
| M    | 249   | 4.82  | Cell wall/membrane biogenesis                                               |
| N    | 59    | 1.14  | Cell motility                                                               |
| U    | 52    | 1.01  | Intracellular trafficking and secretion                                     |
| O    | 109   | 2.11  | Posttranslational modification, protein turnover, chaperones               |
| C    | 202   | 3.91  | Energy production and conversion                                            |
| G    | 264   | 5.11  | Carbohydrate transport and metabolism                                       |
| E    | 392   | 7.58  | Amino acid transport and metabolism                                         |
| F    | 121   | 2.34  | Nucleotide transport and metabolism                                         |
| H    | 171   | 3.31  | Coenzyme transport and metabolism                                           |
| I    | 127   | 2.46  | Lipid transport and metabolism                                              |
| P    | 229   | 4.43  | Inorganic ion transport and metabolism                                      |
| Q    | 92    | 1.78  | Secondary metabolites biosynthesis, transport and catabolism                |
| R    | 558   | 10.79 | General function prediction only                                            |
| S    | 436   | 8.43  | Function unknown                                                            |
| -    | 968   | 18.72 | Not in COGs                                                                 |

The total is based on the total number of protein coding genes in the genome.
Comparative genome analysis

To elucidate the evolutionary relationship of \textit{B. cereus} FORC\_005 with other complete genomes of \textit{B. cereus}, 16S rRNA sequence-based phylogenetic tree analysis and whole genome-based average nucleotide identity (ANI) analysis were conducted. Comparative phylogenetic tree analysis revealed that \textit{B. cereus} and \textit{B. anthracis} strains formed a group including \textit{B. cereus} FORC\_005, suggesting that they may have evolved from a common ancestor (Fig. 2). Subsequent ANI analysis using the complete genome sequences of strain FORC\_005 and other 21 \textit{B. cereus} strains revealed the closest evolutionary relationship between the strains FORC\_005 and B4264 with the ANI value of 98.68 (Fig. 4). The strain B4264 was initially isolated from a male patient with a fatal pneumonia in 1969 [38], indicating that this strain is a clinical isolate. However, the strain FORC\_005 was isolated from a \textit{B. cereus}-contaminated Korean food, suggesting that it may have pathogenesis for potential food-borne outbreak. Therefore, the genome information of the strain FORC\_005 may be important to extend our knowledge about the study of food-borne outbreak via ingestion of the contaminated foods in genomic level, even though it is a food isolate \textit{B. cereus} strain.

Conclusions

While 21 complete genome sequences have been announced to date, all strains were originated from other countries. In this study, newly isolated \textit{B. cereus} FORC\_005 from a contaminated food in South Korea was selected for genomic study to elucidate the genomic pathogenesis of a \textit{B. cereus} strain from South Korea and taxonomical location with complete genome sequences of other \textit{B. cereus} strains.

The strain FORC\_005 showed that anthrolysin O and internalin in the strain FORC\_005 may occur or help the initial step of host cell invasion. In addition, the polysaccharide capsule biosynthesis gene cluster may protect the bacterial pathogen from the host cell immune defense system after host cell invasion. Subsequently, the pathogenesis-associated enterotoxin genes may cause severe diarrhea. The enterotoxin-associated genes in the genome encode three different enterotoxins including cytotoxin K, hemolysin BL, and non-hemolytic enterotoxin, suggesting that this strain may cause human diarrhea. Therefore, the genome of this strain has a complete set of genes or gene clusters for host cell invasion, bacterial protection from the host cell immune system, and enterotoxin production for diarrhea, suggesting that this strain is a food-borne pathogenic bacterium indeed. Comparative phylogenetic tree analysis and ANI analysis of the strain FORC\_005 and other \textit{B. cereus} strains revealed that a food isolate strain FORC\_005 is the closest to a clinical isolate strain B4264, supporting this.

In conclusion, the genomic study of \textit{B. cereus} FORC\_005 provides important information about the genomic features and pathogenesis mechanism of a food isolate \textit{B. cereus}, which is highly similar to a clinical isolate B4264. Furthermore, this genome information would be useful for development of novel biocontrol approach to regulate the pathogenesis of food isolate \textit{B. cereus} strains.

Abbreviations

FORC: Food-borne pathogen omics research Center; TEM: Transmission electron microscopy; PSC: Polysaccharide capsule; ANI: Average nucleotide identity.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

SHC, and JHL initiated and supervised the study. DHL and JHL drafted the manuscript. HRK, HYC, SK, SKK, HJK conducted wetlab work and performed electron microscopy. DHL and HRK worked on the genome sequencing and annotated the genome. DHL, HRK, HK, SK, SC, and JHL discussed, analyzed the data and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by a grant (14162MFD972) from the Ministry of Food and Drug Safety, Korea in 2015 (to JHL and SHC) and the R&D Convergence Center Support Program of the Ministry of Agriculture, Food and Rural Affairs, Republic of Korea (to SHC).
Author details

1Department of Food Science and Biotechnology, Institute of Life Science and Resources, Kyung Hee University, Yongin, Republic of Korea. National Research Laboratory of Molecular Microbiology and Toxicology, Department of Agricultural Biotechnology, and Center for Food Safety and Toxicology, Seoul National University, Seoul, Republic of Korea. 2Department of Food and Animal Biotechnology, Department of Agricultural Biotechnology, Research Institute of Agriculture and Life Sciences, and Center for Food and Bioconvergence, Seoul National University, Seoul, Republic of Korea. 3Department of Agricultural Biotechnology, Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea.

Received: 23 April 2015 Accepted: 31 October 2015

Published online: 11 November 2015

References

1. Granum PE, Lund T. Bacillus cereus and its food poisoning toxins. FEMS Microbiol Lett. 1997;157(2):223–8.
2. Portnoy BL, Goeppert JM, Harmon SM. An outbreak of Bacillus cereus food poisoning resulting from contaminated vegetable sprouts. Am J Epidemiol. 1976;103(6):589–94.
3. Andersons A, Ronner U, Granum PE. What problems does the food industry have with the spore-forming pathogens Bacillus cereus and Clostridium perfringens? Int J Food Microbiol. 1995;28(2):145–55.
4. Botrell EJ. Bacillus cereus, a volatile human pathogen. Clin Microbiol Rev. 2010;23(3):382–98.
5. Jefferies N, Dietrich R, Bock S, Didier A, Märtlbauer E. Bacillus cereus enterotoxins act as major virulence factors and exhibit distinct cytotoxicity to different human cell lines. Toxicon. 2014;77:49–57.
6. Field D, Garty G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol. 2008;26(5):541–7.
7. Woeser CR, Kandler O, Wöhrle ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87(12):4576–9.
8. Ludwig W, Schleifer K, Whitman WB. Class I. Bacilli class. nov. In: Parte AC, Whitman WB, De Vos P, Garty GM, Jones D, Krieg NR, et al., editors. Bergey’s Manual of Systematic Bacteriology. Volume 3 (The Firmicutes). 2nd ed. New York: Springer; 2000. p. 21–128.
9. Garty GM, Lilburn TG, Cole JR, Harrison SN, Euzéby J, Tindal B. The Bacteria: Phylum Firmicutes, Class “Bacill”, In: Garty GM, editor. Taxonomy Outline of Bacteria and Archaea, release 7.7. East Lansing: Michigan State University Board of Trustees; 2007. p. 333–98.
10. Gribbons N, Murray R. Proposals concerning the higher taxa of bacteria. Int J Syst Bacteriol. 1980;30(1):225–30.
11. Euzéby J. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol. 2010;60(Pt 5):1009–10.
12. Prevot A: Dictionnaire des bactéries pathogènes. Hauduroy P, Ehringer G, Prevot A, editors. Paris: Masson; 1932. p. 21.
13. Skerman VBD, McGowan V, Sneath PHA. Approved lists of bacterial names. Int J Syst Bacteriol. 1980;30(1):225–30.
14. Fischer A. Untersuchungen über Bakterien. In Jahrbücher für Wissenschaftliche Botanik. 1895;27:1–163.
15. Cohn F. Untersuchungen über Bakterien. In Berl Biol Pflanz. 1872;1:127–224.
16. Frankland GC, Frankland PF. Studies on some new micro-organisms obtained from air. Philos. Trans. R. Soc. Lond. B 178(1887):257–87.
17. Rajkowski KT, Bennett RW. Bacillus cereus. In: Millodot MD, Bier JW, editors. International Handbook of Foodborne Pathogens. New York: Marcel Dekker, Inc; 2003. p. 40–52.
18. Kramer JM, Gilbert RJ. Bacillus cereus and Other Bacillus Species. In: Doyle MP, editor. Foodborne Bacterial Pathogens. New York: Marcel Dekker, Inc; 1989. p. 22–70.
19. Logan NA, Berkeley RC. Identification of Bacillus strains using the API system. Seyl Gen Microbiol. 1984;30(7):1871–8.
20. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–9.
21. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80.
22. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.
23. Aizt R, Bartels D, Best AA, DeJongh M, Dizg T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75-2164-9-75.
24. Reimerer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 2001;29(12):2607–18.
25. Altermann E, Kleinheinner TR. GAMOLa: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes. Omics A Journal of Integrative Bio. 2002;7(2):161–9.
26. Zdobnov EM, Apweiler R. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics. 2001;17(9):847–8.
27. Rutherford K, Parkhill J, Cook J, Hornsell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16(10):944–5.
28. Dizg T, Ahlert S, Cuestas D, Olson R, Overbeek R, von Krönke V, et al. Accessing the SEED genome databases via Web services API tools for programmers. BMC Bioinformatics. 2010;11:319-2105-11-319.
29. Wu S, Zhu Z, Fu L, Niu B, Li W. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics. 2011;12:444-2164-12-444.
30. Chen L, Xiong Z, Sun L, Yang J, Jin Q, VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 2012;40(Database issue):D641–5.
31. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol. 2004;340(4):783–95.
32. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model application to complete genomes. J Mol Biol. 2001;315(5):567–80.
33. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007;35(Web Server issue):W52–7.
34. Palmer M. The family of thiol-activated, cholesterol-binding cytolsins. Toxicon. 2001;39(11):1681–90.
35. Madden JC, Ruiz N, Caparon M. Cytolysin-mediated translocation (CMT): A functional equivalent of type III secretion in gram-positive bacteria. Cell. 2001;104(1):143–52.
36. Sue D, Hoffmaster AR, Popovic T, Wilkins PP. Capsule production in Bacillus cereus strains associated with severe pneumonia. J Clin Microbiol. 2006;44(9):3426–8.
37. Han CS, Xie G, Challacombe JF, Ahlert S, Bhotika SS, Brown N, et al. Pathogenicity sequence analysis of Bacillus cereus and Bacillus thuringiensis isolates closely related to Bacillus anthracis. J Bacteriol. 2006;188(9):3382–90.
38. J Craig Venter Institute. [http://gcid.jcvi.org/projects/msc/bacillus/]
39. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Proc Natl Acad Sci U S A. 2007;104(1):200–5.
40. Richter M, Rossello-Mora R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A. 2009;106(45):19126–31.
41. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Proc Natl Acad Sci U S A. 2007;104(1):200–5.
42. Richter M, Rossello-Mora R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A. 2009;106(45):19126–31.
43. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997;25(13):3389–402.