Research article

# Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing

Alex R Hoffmaster*, Ryan T Novak, Chung K Marston, Jay E Gee, Leta Helsel, James M Pruckler and Patricia P Wilkins

Address: National Center for Zoonotic, Vector-borne, and Enteric Diseases, Centers for Disease Control and Prevention, Atlanta, USA

Email: Alex R Hoffmaster* - amh9@cdc.gov; Ryan T Novak - bnk4@cdc.gov; Chung K Marston - cdk5@cdc.gov; Jay E Gee - xzg4@cdc.gov; Leta Helsel - loh1@cdc.gov; James M Pruckler - jmp3@cdc.gov; Patricia P Wilkins - pma1@cdc.gov

* Corresponding author

## Abstract

**Background:** *Bacillus cereus* is most commonly associated with foodborne illness (diarrheal and emetic) but is also an opportunistic pathogen that can cause severe and fatal infections. Several multilocus sequence typing (MLST) schemes have recently been developed to genotype *B. cereus* and analysis has suggested a clonal or weakly clonal population structure for *B. cereus* and its close relatives *B. anthracis* and *B. thuringiensis*. In this study we used MLST to determine if *B. cereus* isolates associated with illnesses of varying severity (e.g., severe, systemic vs. gastrointestinal (GI) illness) were clonal or formed clonal complexes.

**Results:** A retrospective analysis of 55 clinical *B. cereus* isolates submitted to the Centers for Disease Control and Prevention between 1954 and 2004 was conducted. Clinical isolates from severe infections (n = 27), gastrointestinal (GI) illness (n = 18), and associated isolates from food (n = 10) were selected for analysis using MLST. The 55 isolates were diverse and comprised 38 sequence types (ST) in two distinct clades. Of the 27 isolates associated with serious illness, 13 clustered in clade 1 while 14 were in clade 2. Isolates associated with GI illness were also found throughout clades 1 and 2, while no isolates in this study belonged to clade 3. All the isolates from this study belonging to the clade 1/cereus III lineage were associated with severe disease while isolates belonging to clade1/cereus II contained isolates primarily associated with severe disease and emetic illness. Only three STs were observed more than once for epidemiologically distinct isolates.

**Conclusion:** STs of clinical *B. cereus* isolates were phylogenetically diverse and distributed among two of three previously described clades. Greater numbers of strains will need to be analyzed to confirm if specific lineages or clonal complexes are more likely to contain clinical isolates or be associated with specific illness, similar to *B. anthracis* and emetic *B. cereus* isolates.

## Background

The phylogenetically related species of the *Bacillus cereus* group include: *B. cereus*, *B. anthracis*, *B. thuringiensis*, *B.* mycoides and two recently described species, *B. pseudomycoides* and *B. weihenstephanensis* [1-4]. *B. cereus*, *B. thuringiensis*, and *B. anthracis* have been the most characterized

due to their pathogenic nature. *B. anthracis* is the etiologic agent of anthrax, *B. thuringiensis* is an insect pathogen, and *B. cereus* can be associated with a variety of human infections. *B. cereus* is most frequently associated with food poisoning, characterized by strains producing emetic or diarrheal toxins. It is also an opportunistic pathogen resulting in localized (wound, ocular, oral) and systemic infections (bacteremia, septicemia, endocarditis, meningitis, pneumonia) especially in immunocompromised patients [5,6]. However, although *B. cereus* is ubiquitous in the environment, particularly soil, it does not generally pose a health threat.

Systemic infection due to *B. cereus* can be associated with significant morbidity and mortality. Typically the species of *Bacillus* most commonly associated with serious human illness is *B. anthracis*. However, pulmonary infections caused by *B. cereus* can be severe and fatal but have been described predominantly in patients with significant risk factors [7-14]. There are rare reports in the literature of fulminating pneumonia associated with *B. cereus* infections in patients not known to be immune compromised [15-18]. Interestingly, *B. cereus* strains have been shown to harbor diverse plasmids which can share conserved sequences with *B. anthracis* virulence plasmids pXO1 and pXO2 [19-21], and isolates associated with a few of these severe infections were shown to contain *B. anthracis* virulence genes. *B. cereus* G9241, an isolate associated with a severe pneumonia case, contained a plasmid, pBCXO1, which was 99.6% similar to the *B. anthracis* pXO1 virulence plasmid [17]. More recently several *B. cereus* isolates, associated with fatal pneumonias in Texas, were described that harbored *B. anthracis* toxin and/or capsule virulence genes (i.e., *pagA*, *lef*, *cya*, and *capBCA*) [17,22]. G9241 and the Texas isolates were recovered from patients who were not obviously immunocompromised; however, all three infections were in metal workers, an occupation that may have affected their susceptibility to respiratory infections [23-25].

The population structure of pathogenic bacteria varies over a wide range, from strictly clonal to effectively panmictic [26]. A study of localized populations of the *B. cereus* group in soil using multilocus sequence typing (MLST) suggested frequent recombination among *B. weihenstephanensis*, while *B. cereus* and *B. thuringiensis* appeared clonal [27]. Two additional MLST studies, each using different MLST gene sets, also suggested a weakly clonal to clonal *B. cereus* population structure with evolutionary changes occurring through point mutations [4,28].

While the overall population structure of the *B. cereus* group is not yet clear due to limited numbers of strains examined, data suggest that in addition to *B. anthracis*,

some pathogenic *B. cereus* isolates associated with periodontal disease and emetic gastrointestinal illness are limited to specific clonal groups [29,30]. A recent MLST study found 30 of 48 isolates of clinical origin were associated with clonal complexes and some groups contained strains isolated from similar human sources [31]. In contrast, two studies using MLST analysis on smaller numbers (n = 8 and n = 10) of *B. cereus* isolates from severe infections were found to be diverse and not restricted to single clonal groups or lineages [3,32]. In this study we used MLST to examine the phylogenetic diversity and relatedness of 55 clinical *B. cereus* isolates, from a historical collection of *B. cereus* recovered from 1954–2004.

## Methods
### Bacterial isolates
A total of 55 strains of *B. cereus* associated with severe (septicemia and pneumonia) or gastrointestinal illness were analyzed (Table 1 and 2). The strains were selected from over 400 *Bacillus* in the collection of the Special Bacteriology Reference Laboratory, Centers for Disease Control and Prevention. The isolates represent a diverse collection of strains received between 1954 and 2004 from 19 states in the U.S. All isolates found in the collection associated with severe systemic human infections of the blood or lungs were selected for this study. This resulted in the identification of 27 isolates from 26 epidemiologically distinct infections. In addition, a convenient sampling of 28 isolates associated with cases of GI illness or food remnants connected to eight foodborne outbreaks and five sporadic cases of food poisoning were included for comparison. Foodborne isolates were generally selected if associated with two or more clustered cases. All strains shared common biochemical/phenotypic characteristics that are consistent with the identification of *Bacillus cereus*. They were motile rods with peritrichous flagella, and hemolytic with β-hemolysis or lavender-green coloration under heavy growth indicating proteolysis. They produced catalase, lecithinase, and reduced nitrate. They fermented D-glucose and maltose but were variable in fermenting lactose and sucrose. There was no fermentation of D-xlyose and D-mannitol. Production of oxidase and urease was variable while there was no production of indole. Toxin crystals were not detected for these isolates. 16S DNA sequencing was determined as described previously [33]. Isolates were cultured on trypticase soy agar containing 5% sheep blood (Becton Dickinson Microbiology Systems, Sparks, MD) at 37°C overnight. Bacterial genomic DNA extractions for PCR were prepared as described previously [34]. Emetic isolates were identified by PCR as described by Ehling-Shulz et al. [35].

### MLST
The isolates were characterized by the MLST scheme described by Priest et al. [28] and performed with modifi-

**Table 1: Characteristics of isolates associated with severe disease**

| ID | ST[a] | Origin[b] | Source | Year isolated | Clade[c] | Additional information |
|---|---|---|---|---|---|---|
| 03BB102[d] | 11 | TX | Blood | 2003 | 1 | Pneumonia, fatal |
| F4976 | 24 | OR | Lung | 1983 | 2 | Pneumonia |
| E6345 | 26 | NH | Lung | 1979 | 1 | |
| D4214 | 62 | TX | Blood | 1975 | 1 | |
| B5780 | 76 | TX | Blood | 1970 | 1 | |
| G9241[e] | 78 | LA | Sputum | 1994 | 1 | Pneumonia |
| G9898[f] | 78 | LA | Sputum | 1996 | 1 | Pneumonia, fatal |
| 03BB87[d] | 78 | TX | Blood | 2003 | 1 | Pneumonia, fatal |
| 04S 00334 | 85 | CA | Blood | 2004 | 2 | |
| B4266 | 89 | VA | Lung | 1969 | 2 | Pneumonia, fatal |
| H1548 | 90 | NC | Blood | 2000 | 1 | Septicemia, fatal |
| B1357 | 91 | LA | Lung | 1968 | 1 | |
| F9930 | 93 | GA | Lung | 1987 | 2 | Biopsy |
| F3920 | 94 | CA | Lung | 1982 | 2 | Pneumonia |
| G1440 | 94 | LA | Lung | 1988 | 2 | Pneumonia, biopsy |
| G4200 | 95 | MS | Blood | 1990 | 2 | Pneumonia, fatal |
| D8096 | 100 | LA | Blood | 1971 | 2 | |
| I952 | 101 | GA | Blood | 1955 | 2 | Septicemia, fatal |
| E3911 | 102 | TX | Blood | 1978 | 1 | |
| F4794 | 111 | MI | Blood | 1982 | 2 | Pneumonia |
| G3317 | 121 | Israel | Blood | 1989 | 1 | |
| F5191[g] | 122 | CA | Lung | 1983 | 1 | Biopsy |
| C3276 | 136 | TX | Blood | 1972 | 2 | |
| E1149 | 138 | MO | Lung | 1978 | 2 | |
| F5190[g] | 140 | CA | Lung | 1983 | 1 | Biopsy |
| 2003-I-1253 | 142 | OR | Blood | 2003 | 2 | Fatal |
| SB460 | 146 | NY | Blood | 1987 | 2 | Septicemia |

[a] ST = sequence type
[b] Country or U.S.A. state abbreviation
[c] Clade as described by Priest et al. [28]
[d] Hoffmaster et al. [22]
[e] Hoffmaster et al. [17]
[f] Sue et al. [42]
[g] Distinct isolates from same patient biopsy

cations described previously [36]. Briefly, approximately 10 ng of DNA extracted from bacterial cultures were used as the template in the PCR reactions with the primers described on the *B. cereus* MLST website http://www.pubmlst.org/bcereus to generate amplicons from the seven MLST gene loci (*glpF, gmk, ilvD, pta, pur, pycA* and *tpi*). Primers described for "Option 1" were used in the cases of the *ilvD* and *gmk* gene primers. Amplification products of the correct size and similar concentrations, based on visual inspection of agarose gels, were obtained from all 55 isolates and were then purified by using a QiaAmp PCR purification kit (Qiagen Inc., Valencia, CA) according to the instructions of the manufacturer.

The nucleotide sequences of both DNA strands of the PCR fragments were determined using the amplification primers and the BigDye terminator cycle sequencing kit version 3.1 according to the manufacturer's instructions, except 3 μl of BigDye was used instead of 8 μl (Applied BioSystems, Foster City, CA). Sequencing products were separated from unincorporated dye terminators by Centri-Sep column purification (Princeton Separations, Adelphia, NJ) and resolved with an Applied Biosystems model 3100 automated DNA sequencing system.

### Data analysis
Sequence editing and alignment was completed using the Genetics Computer Group Wisconsin Package version 10.3 (Accelrys, San Diego, CA). The chromatograms from the ABI 3100 sequencer were exported, visually examined, and assembled in SEQMERGE. The sequences of each of the seven housekeeping genes were edited to the previously described allele lengths (between 348 and 504 bp). These were then assigned allele numbers based on those already described in the *B. cereus* MLST database http://www.pubmlst.org/bcereus. Isolates were assigned a sequence type (ST) on the basis of the combination of the seven alleles. Allelic sequences differing from known alleles were assigned new allele numbers and ST and added to the *B. cereus* MLST database.

**Table 2: Characteristics of GI-associated isolates**

| ID | STa | Originb | Source | Year isolated | Cladec | Addtional information |
|---|---|---|---|---|---|---|
| F6722 | 4 | AL | Food | 1985 | 2 | Food poisoning, diarrheal toxin |
| F663d | 26 | NC | Food | 1981 | I | Outbreak, emetic toxink |
| F665d | 26 | NC | Stool | 1981 | I | Outbreak, emetic toxink |
| F667d | 26 | NC | Stool | 1981 | I | Outbreak, emetic toxink |
| H3074.97e | 26 | TX | Stool | 1997 | I | Outbreak, emetic toxink |
| H2573.97f | 26 | England | Food | 1997 | I | Outbreak, emetic toxink |
| H2576.97f | 26 | England | Food | 1997 | I | Outbreak, emetic toxink |
| H9310 | 26 | MN | Vomit | 2000 | I | Outbreak, emetic toxink |
| H9311 | 26 | MN | Food | 2000 | I | Outbreak, emetic toxink |
| H9312 | 26 | MN | Vomit | 2000 | I | Outbreak, emetic toxink |
| G9844g | 56 | NE | Stool | 1996 | 2 | Outbreak |
| G9842g | 56 | NE | Stool | 1996 | 2 | Outbreak, diarrheal toxin |
| G9843g | 56 | NE | Stool | 1996 | 2 | Outbreak |
| G2008e | 73 | PA | Stool | 1988 | 2 | Food Poisoning |
| F666d | 92 | NC | Stool | 1981 | I | Outbreak |
| C1784i | 96 | FL | Food | 1972 | 2 | Diarrhea |
| C1617 | 97 | NC | Stool | 1972 | 2 | |
| F439j | 98 | MI | Stool | 1981 | 2 | Diarrhea, diarrheal toxin |
| F438j | 98 | MI | Stool | 1981 | 2 | Diarrhea, diarrheal toxin |
| G2055h | 99 | PA | Stool | 1988 | 2 | Food poisoning, diarrheal toxin |
| G2054h | 120 | PA | Food | 1988 | I | Food poisoning, diarrheal toxin |
| C1785i | 123 | FL | Food | 1972 | I | Diarrhea |
| C1783i | 123 | FL | Stool | 1972 | I | Diarrhea |
| 3297 | 124 | AZ | Stool | 1957 | 2 | Food poisoning |
| F5157 | 137 | CA | Food | 1983 | 2 | Food poisoning, diarrheal toxin |
| F7720 | 141 | NM | Stool | 1986 | I | Diarrheal toxin |
| H3076.97e | 144 | TX | Stool | 1997 | I | Outbreak, emetic toxink |
| H3081.97e | 144 | TX | Food | 1997 | I | Outbreak, emetic toxink |

a ST = sequence type
b Country or U.S.A. state abbreviation
c Clade as described by Priest et al. [28]
d North Carolina 1981 food outbreak, 18 people, linked to rice
e Texas 1997 food outbreak, linked to rice
f England 1997 food outbreak, linked to rice
g Nebraska food outbreak, three people, source unknown
h Food poisoning, one person, source unknown
i Food poisoning linked to one of two food sources tested
j Different patients, cases linked
k Presence of emetic toxin genetic determinants were detected by PCR [35].

The seven gene fragments from each of the 55 isolates were concatenated and downloaded from the MLST website. To assess the relationships of the isolates in this study with other strains, a phylogenetic tree was derived from the aligned concatenated sequences using the neighbor-joining method [37] using the Kimura 2-parameter model [38] and 1000 step bootstrap analysis of the data using Mega 3 [39], as described previously [36]. Simpson's index of diversity was calculated as described by Hunter and Gaston [40].

## Results

Forty-five clinical and 10 food-associated isolates of *B. cereus* were selected from over 400 *Bacillus* submitted to the CDC Special Bacteriology Reference Laboratory between 1954 and 2004 and were analyzed using MLST.

Twenty-seven strains were recovered from severe atypical *Bacillus* infections in 26 patients presenting with either pneumonia or septicemia (Table 1). Eight of these 26 cases were fatal. Twenty-eight isolates were recovered from sporadic cases of GI illness, foodborne outbreak investigations or associated foods. The 55 isolates represented 38 STs (Table 3), 27 of which were new to the MLST database http://www.MLST.net. The number of alleles per locus observed in the isolates in this study ranged from 18–23 and averaged 19.9. The Simpson's index of diversity was calculated using a subset of epidemiologically distinct isolates and found to be 0.989 (where 1.0 is equal to absolute discrimination).

A neighbor-joining tree was constructed using the seven concatenated allele sequences from the isolates in this

**Table 3: Sequence types and alleles grouped by lineage**

| ST[a] | N[b] | glpF | gmk | ilvD | pta | pur | pycA | tpi | Clade/lineage[c] |
|---|---|---|---|---|---|---|---|---|---|
| 90 | 1 | 6 | 4 | 41 | 5 | 43 | 46 | 3 | 1/Cereus I |
| 92 | 1 | 6 | 4 | 42 | 4 | 16 | 6 | 3 | 1/Cereus I |
| 120 | 1 | 6 | 4 | 3 | 4 | 46 | 6 | 36 | 1/Cereus I |
| 144 | 1 | 67 | 2 | 63 | 5 | 36 | 3 | 4 | 1 |
| 26 | 4 | 3 | 2 | 31 | 5 | 16 | 3 | 4 | 1/Cereus II |
| 140 | 1 | 52 | 2 | 21 | 5 | 19 | 3 | 38 | 1/Cereus II |
| 91 | 1 | 51 | 2 | 21 | 5 | 19 | 3 | 38 | 1/Cereus II |
| 141 | 1 | 54 | 2 | 21 | 5 | 19 | 3 | 4 | 1/Cereus II |
| 121 | 1 | 53 | 2 | 59 | 54 | 47 | 3 | 2 | 1/Cereus II |
| 122 | 1 | 53 | 2 | 59 | 17 | 47 | 3 | 2 | 1/Cereus II |
| 11 | 1 | 34 | 1 | 32 | 1 | 33 | 37 | 24 | 1/Cereus III |
| 62 | 1 | 38 | 1 | 32 | 1 | 18 | 33 | 24 | 1/Cereus III |
| 76 | 1 | 34 | 15 | 28 | 38 | 30 | 29 | 33 | 1/Cereus III |
| 78 | 3 | 24 | 22 | 33 | 37 | 34 | 38 | 5 | 1/Cereus IV |
| 123 | 1 | 48 | 30 | 33 | 37 | 44 | 48 | 45 | 1/Cereus IV |
| 101 | 1 | 33 | 33 | 44 | 19 | 2 | 17 | 17 | 2/Toworthi |
| 96 | 1 | 33 | 8 | 44 | 19 | 2 | 17 | 17 | 2/Toworthi |
| 89 | 1 | 14 | 8 | 40 | 19 | 2 | 17 | 17 | 2/Toworthi |
| 94 | 2 | 13 | 29 | 9 | 14 | 9 | 12 | 31 | 2/Toworthi |
| 73 | 1 | 13 | 8 | 9 | 14 | 9 | 12 | 31 | 2/Toworthi |
| 99 | 1 | 13 | 8 | 9 | 14 | 46 | 12 | 31 | 2/Toworthi |
| 24 | 1 | 12 | 8 | 9 | 14 | 11 | 12 | 10 | 2/Toworthi |
| 100 | 1 | 13 | 8 | 46 | 28 | 9 | 12 | 7 | 2/Toworthi |
| 142 | 1 | 13 | 8 | 8 | 11 | 9 | 12 | 7 | 2/Toworthi |
| 4 | 1 | 13 | 8 | 8 | 11 | 11 | 12 | 7 | 2/Toworthi |
| 95 | 1 | 13 | 29 | 9 | 11 | 2 | 12 | 8 | 2/Toworthi |
| 93 | 1 | 14 | 29 | 43 | 11 | 9 | 47 | 8 | 2/Toworthi |
| 98 | 1 | 40 | 28 | 14 | 12 | 2 | 50 | 7 | 2/Kurstaki |
| 85 | 1 | 47 | 28 | 14 | 12 | 2 | 36 | 7 | 2/Kurstaki |
| 137 | 1 | 37 | 28 | 14 | 12 | 37 | 36 | 7 | 2/Kurstaki |
| 136 | 1 | 31 | 8 | 14 | 44 | 2 | 36 | 7 | 2/Kurstaki |
| 138 | 1 | 37 | 9 | 14 | 12 | 12 | 14 | 7 | 2/Kurstaki |
| 124 | 1 | 14 | 32 | 48 | 45 | 58 | 51 | 7 | 2/Kurstaki |
| 146 | 1 | 68 | 20 | 23 | 8 | 4 | 8 | 39 | 2/Thuringiensis |
| 97 | 1 | 15 | 6 | 47 | 8 | 4 | 8 | 7 | 2/Thuringiensis |
| 56 | 1 | 15 | 7 | 7 | 2 | 7 | 26 | 13 | 2/Sotto |
| 111 | 1 | 43 | 26 | 35 | 42 | 39 | 41 | 30 | 2 |
| 102 | 1 | 49 | 31 | 45 | 43 | 45 | 49 | 37 | 1 |

[a] ST = sequence type
[b] Number of isolates identified from epidemiologically distinct (unrelated) events.
[c] Clade as described by Priest et al. [28]

study and selected isolates from the MLST database (Figure 1). The isolates in this study were not monophyletic by morbidity or mortality; instead, they were distributed among the *Bacillus* group clades 1 and 2 and in numerous lineages as previously defined [28].

### *Clade 1*
Thirteen of 27 isolates associated with severe illness grouped in clade 1, while nine of 18 GI-associated isolates and seven of 10 food isolates were also grouped to this clade. Isolates associated with severe illness were found in all three previously described Cereus lineages within clade 1. Isolates recovered from GI illness grouped in Cereus I and II lineages, but not Cereus III.

All of the emetic *B. cereus* isolates were represented by two (ST 26 and 144) of the four previously described STs of emetic strains [41]. ST 26 was the most commonly identified ST in this and other reports [28,41]. Over half (6/9) of the GI-associated clinical isolates and five of the seven food isolates in clade 1 were associated with emetic GI illness (Figure 1).

Three STs (ST 102, 78 and 123) found to cluster in clade 1 did not group into described lineages [28]. The ST 78 and 123 isolates were sufficiently divergent to possibly warrant a designation of a fourth Cereus lineage (Cereus IV) in clade 1 (Figure 1). All three isolates of ST 78 were associated with severe pneumonia cases, of which two
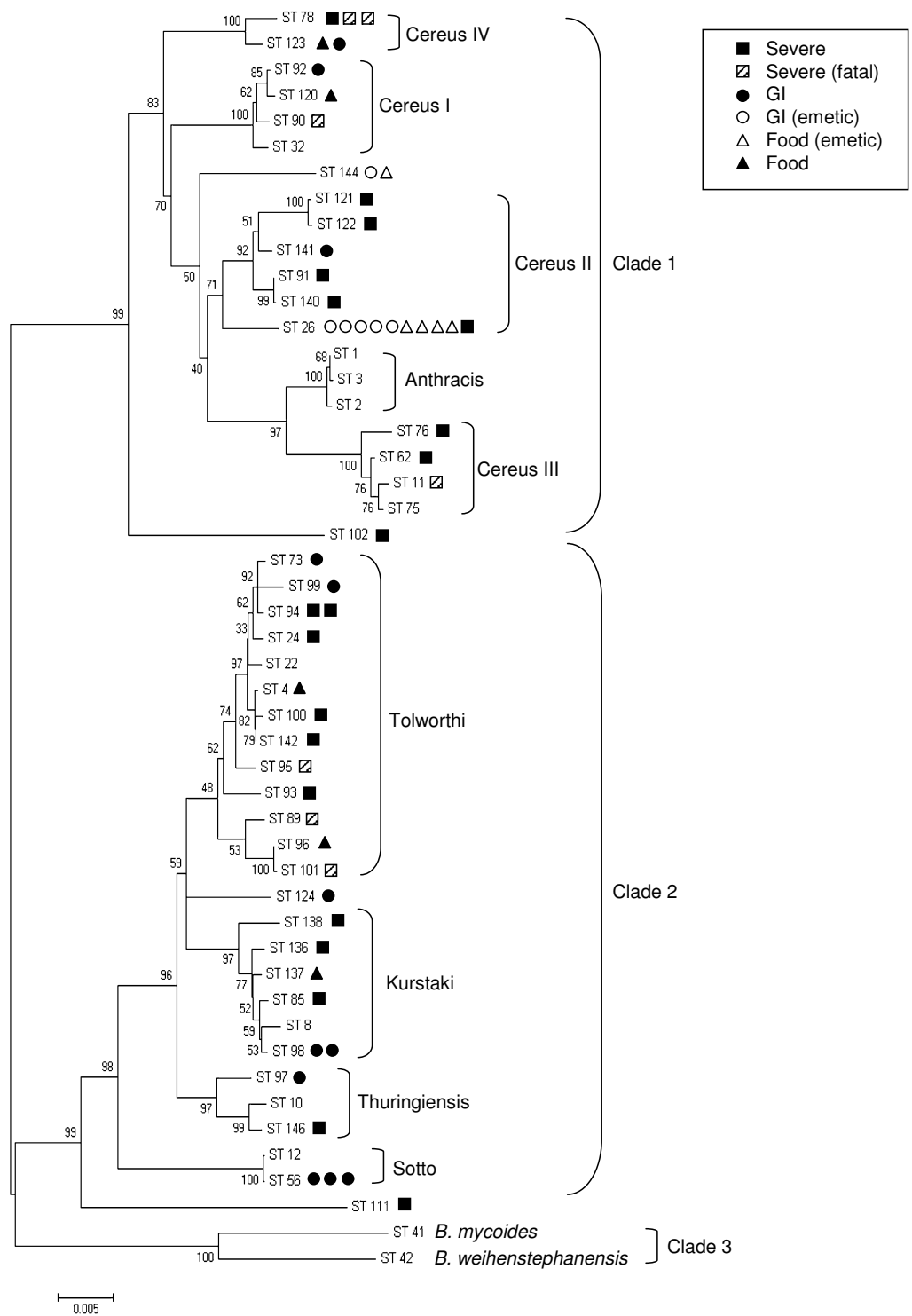
**Figure 1**
**Relationships between isolates of this study and other select reference isolates using concatenated sequences from seven housekeeping alleles.** Tree was constructed using the neighbor joining method and percent bootstrap confidence levels were calculated using 1000 resamplings of the original data. Clades and lineages are labeled as designated by Priest et al. [28] with the addition of Cereus IV.

were fatal. Both isolates of ST 123 were associated with a diarrheal disease and group with ST 78 isolates, based on alignment of the concatenated sequences, however, these two STs share only two of seven allele types (Table 3).

### Clade 2

Fourteen of 27 isolates associated with severe illness grouped in clade 2, while nine of 18 GI-associated isolates and three of 10 food isolates were also grouped into this clade. Isolates recovered from severe illness were found in three of four previously described Cereus lineages within clade 2, but not in the Sotto lineage [28]. Isolates associated with GI illness grouped in all four clade 2 lineages (Figure 1). A single isolate associated with a pneumonia case (ST 111) clustered to clade 2, but did not belong to any of the described lineages.

## Discussion

In this study we used MLST to examine the phylogenetic diversity and relatedness of *B. cereus* isolates that were associated with different clinical presentations. Allelic diversity among the isolates in the study was examined and the average number of alleles per locus was determined to be 19.9. Previously, Helgason et al. found an average of 33.6 alleles per locus in 67 *Bacillus* isolates [4] and Priest et al. found an average of 30.5 alleles per locus in 105 isolates [28]. Adjusted for the number of isolates included in the study, the allelic diversity in this study was less than that described by Helgason and greater than the diversity reported by Priest et al. (0.5 > this study-0.36 > Priest-0.29). We used the Simpson's index of diversity to measure the discriminatory power of this subtyping method (i.e., likelihood of two isolates from epidemiologically distinct events having the same ST). The overall discrimination index was 0.989 (where 1.0 is equal to absolute discrimination), making MLST a very useful tool for molecular epidemiology. However, this value is clearly biased by the types of isolates that are included. For instance, the discrimination would be quite poor for analysis of foodborne isolates associated with emetic disease (for which only four STs have been observed).

Only three STs were seen more than once from epidemiologically distinct cases (ST 26, 78 and 94). ST 26 was associated with four separate emetic GI outbreaks. This was not unexpected since ST 26 has been the most common ST associated with emetic isolates in previous studies [28,41]. ST 78 and ST 94 are not closely related (i.e., ST 78 is in clade 1 and ST 94 is in clade 2) however, all the isolates of these STs (n = 5) were associated with pneumonia cases. Interestingly, all three of the ST 78 isolates were the cause of pneumonia in metal workers (two cases were fatal). In addition, we have previously shown all three isolates (G9241, G9898, and 03BB87) produce a capsule and were PCR positive for pBC218 plasmid genes (putative

polysaccharide polymerase and translocase genes) hypothesized to be required for a polysaccharide capsule production in G9241 [17,22,42]. Thus far, we have only seen these plasmid genes in isolates of ST 78. Perhaps this is an under sampled clonal group harboring a specific virulence plasmid similar to *B. anthracis*.

The clade 1/Cereus III lineage is most closely related to the Anthracis lineage and was previously shown to include both *B. cereus* and *B. thuringiensis* [28]. All the Cereus III isolates analyzed for this study were associated with systemic illness. This lineage contained isolates from Texas, including 03BB102 (ST 11) which was isolated from a fatal case of pneumonia in a welder [22], and a ST 62 clinical isolate recovered from a septicemic patient (TX, 1975). The other isolate in this lineage, B5780, was recovered from blood and shares ST 76 with another invasive isolate in the MLST database.

The clade 1/Cereus II lineage, which is also closely related to the Anthracis/Cereus III lineages, contained isolates from both systemic and GI illnesses. Nine of the 10 GI-associated isolates, recovered from both clinical specimens and food samples, in cereus II are emetic isolates and were shown to produce a functional emetic toxin peptide or were PCR positive for the emetic genetic determinants. The single ST 26 isolate that was negative for these emetic loci, E6345, was recovered from lung. It is possible this isolate never had these determinants or, since they are plasmid-borne [43], the plasmid may have been cured prior to or during laboratory manipulations. We have limited information on the case from which E6345 originated. It is possible that the pneumonia may have resulted from aspiration due to infection with bacteria that originally produced emetic toxin. The clade 1/Cereus II lineage appears to contain mostly isolates associated with severe systemic infections and emesis.

Several studies have suggested that *B. cereus* isolates that cluster closely with *B. anthracis* are more likely to be associated with clinical cases as opposed to environmental sources [4,44,45]. This suggests that it may be possible to infer the virulence of a strain (i.e., pathogenic potential) based on ST. The MLST data from this study suggests strains associated with severe disease presentations (pneumonia, septicemia, and bacteremia) are phylogenetically diverse within clades 1 and 2 but continue to appear restricted to these clades and are not present in clade 3. Thus, while clinical isolates are diverse, there may be differences in pathogenic potential at the clade level of resolution. In addition, some lineages appear to have more clinical isolates associated with severe or emetic illness such as the Cereus II and III lineages. Whether these or additional lineages or clonal complexes contain isolates primarily associated with specific diseases may become

clearer as more clinical and environmental isolates are added to the database. Attempts to correlate ST with pathogenic potential of *B. cereus* are complicated by several factors. First, the pathogenic potential of environmental isolates is unknown. The source of an isolate (clinical versus environmental) does not necessarily correlate with pathogenic potential. *B. cereus* is not an obligate pathogen, therefore, isolates like those recovered from clinical specimens should also be found in the environment. Studies have shown that environmental isolates of the *B. cereus* group harbor a variety of *B. cereus* virulence genes, such as hemolysins, enterotoxins, cytotoxin K and phospholipase C [46,47]. Host susceptibility and whether or not *B. cereus* is associated with the primary infection or is a secondary infection may also complicate efforts to segregate strains or clonal complexes by virulence or pathogenic potential.

## Conclusion

Clinical isolates in general are diverse and well represented throughout clades 1 and 2, while no clinical isolates have been identified in clade 3. STs of isolates associated with severe disease were distributed throughout clades 1 and 2; however, the majority of GI isolates within clade 1 were limited to emetic strains of *B. cereus*. There was some evidence, based on limited numbers, of isolates within lineages or clonal complexes being associated with severe disease or a specific illness such as ST 78 with pneumonia. However, further work will be needed to clearly understand the relationship of isolates within the clades and lineages and how they may be related to the potential of isolates to cause a variety of illnesses. A more robust picture of the *B. cereus* group is not yet realized due to low number of sequenced isolates and biased sampling. The existence of multiple MLST schemes for the *B. cereus* group (n = 5) further hampers efforts. A recent report, however, describes the creation of a supertree database (SuperCAT) that combines and integrates data from all the published schemes [48] which should aid in the effort to better view the *B. cereus* group phylogeny. The true population structure of *B. cereus* will emerge as more diverse isolates, including clinical isolates with good epidemiological data, are analyzed and added to the public databases.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

ARH and PPW contributed to study design, analysis, and writing of the manuscript. RTN performed MLST and contributed to study design and writing of manuscript. CKM performed MLST and contributed to writing of manuscript. JEG performed MLST, data analysis and contributed to writing of manuscript. LH and JP contributed to study design and writing of manuscript. The manuscript has been reviewed and approved by all authors.

## References

1.  Lechner S, Mayr R, Francis KP, Pruss BM, Kaplan T, Wiessner-Gunkel E, Stewart GS, Scherer S: ***Bacillus weihenstephanensis* sp. nov. is a new psychrotolerant species of the *Bacillus cereus* group.** *Int J Syst Bacteriol* 1998, **48(Pt 4):**1373-1382.
2.  Jensen GB, Hansen BM, Eilenberg J, Mahillon J: **The hidden lifestyles of *Bacillus cereus* and relatives.** *Environ Microbiol* 2003, **5(8):**631-640.
3.  Barker M, Thakker B, Priest FG: **Multilocus sequence typing reveals that *Bacillus cereus* strains isolated from clinical infections have distinct phylogenetic origins.** *FEMS Microbiol Lett* 2005, **245(1):**179-184.
4.  Helgason E, Tourasse NJ, Meisal R, Caugant DA, Kolsto AB: **Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group.** *Appl Environ Microbiol* 2004, **70(1):**191-201.
5.  Drobniewski FA: ***Bacillus cereus* and related species.** *Clin Microbiol Rev* 1993, **6(4):**324-338.
6.  Kotiranta A, Lounatmaa K, Haapasalo M: **Epidemiology and pathogenesis of *Bacillus cereus* infections.** *Microbes Infect* 2000, **2(2):**189-198.
7.  Bekemeyer WB, Zimmerman GA: **Life-threatening complications associated with *Bacillus cereus* pneumonia.** *Am Rev Respir Dis* 1985, **131(3):**466-469.
8.  Coonrod JD, Leadley PJ, Eickhoff TC: ***Bacillus cereus* pneumonia and bacteremia. A case report.** *Am Rev Respir Dis* 1971, **103(5):**711-714.
9.  Frankard J, Li R, Taccone F, Struelens MJ, Jacobs F, Kentos A: ***Bacillus cereus* pneumonia in a patient with acute lymphoblastic leukemia.** *Eur J Clin Microbiol Infect Dis* 2004, **23(9):**725-728.
10. Funada H, Machi T, Matsuda T: ***Bacillus cereus* pneumonia with empyema complicating aplastic anemia–a case report.** *Kansenshogaku Zasshi* 1991, **65(4):**477-480.
11. Gascoigne AD, Richards J, Gould K, Gibson GJ: **Successful treatment of *Bacillus cereus* infection with ciprofloxacin.** *Thorax* 1991, **46(3):**220-221.
12. Leff A, Jacobs R, Gooding V, Hauch J, Conte J, Stulbarg M: ***Bacillus cereus* pneumonia. Survival in a patient with cavitary disease treated with gentamicin.** *Am Rev Respir Dis* 1977, **115(1):**151-154.
13. Sliman R, Rehm S, Shlaes DM: **Serious infections caused by *Bacillus* species.** *Medicine (Baltimore)* 1987, **66(3):**218-223.
14. Strauss R, Mueller A, Wehler M, Neureiter D, Fischer E, Gramatzki M, Hahn EG: **Pseudomembranous tracheobronchitis due to *Bacillus cereus*.** *Clin Infect Dis* 2001, **33(5):**E39-41.
15. Jonsson S, Clarridge J, Young EJ: **Necrotizing pneumonia and empyema caused by *Bacillus cereus* and *Clostridium bifermentans*.** *Am Rev Respir Dis* 1983, **127(3):**357-359.
16. Miller JM, Hair JG, Hebert M, Hebert L, Roberts FK Jr, Weyant RS: **Fulminating bacteremia and pneumonia due to *Bacillus cereus* .** *J Clin Microbiol* 1997, **35(2):**504-507.
17. Hoffmaster AR, Ravel J, Rasko DA, Chapman GD, Chute MD, Marston CK, De BK, Sacchi CT, Fitzgerald C, Mayer LW, Maiden M, Priest FG, Barker M, Jiang L, Cer RZ, Rilsonte J, Peterson SN, Weyant RS, Galloway DR, Read TD, Popovic T, Fraser CM: **Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax.** *Proc Natl Acad Sci USA* 2004, **101(22):**8449-8454.
18. Avashia SB, Riggins WS, Lindley C, Hoffmaster A, Drumgoole R, Nekomoto T, Jackson PJ, Hill KK, Williams K, Lehman L, Libal MC,

Wilkins PP, Alexander J, Tvaryanas A, Betz T: **Fatal pneumonia among metalworkers due to inhalation exposure to *Bacillus cereus* containing *Bacillus anthracis* toxin genes.** *Clin Infect Dis* 2007, **44(3)**:414-416.

19. Rasko DA, Rosovitz MJ, Okstad OA, Fouts DE, Jiang L, Cer RZ, Kolsto AB, Gill SR, Ravel J: **Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the *B. cereus*-group plasmids, including *Bacillus anthracis* pXO1.** *J Bacteriol* 2007, **189(1)**:52-64.

20. Pannucci J, Okinaka RT, Sabin R, Kuske CR: ***Bacillus anthracis* pXO1 plasmid sequence conservation among closely related bacterial species.** *J Bacteriol* 2002, **184(1)**:134-141.

21. Pannucci J, Okinaka RT, Williams E, Sabin R, Ticknor LO, Kuske CR: **DNA sequence conservation between the *Bacillus anthracis* pXO2 plasmid and genomic sequence from closely related bacteria.** *BMC Genomics* 2002, **3(34)**:.

22. Hoffmaster AR, Hill KK, Gee JE, Marston CK, De BK, Popovic T, Sue D, Wilkins PP, Avashia SB, Drumgoole R, Helma H, Ticknor LO, Okinaka RT, Jackson PJ: **Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely related to *Bacillus anthracis* and harbor *B. anthracis* virulence genes.** *J Clin Microbiol* 2006, **44(9)**:3352-3360.

23. Antonini JM: **Health effects of welding.** *Crit Rev Toxicol* 2003, **33(1)**:61-103.

24. Coggon D, Inskip H, Winter P, Pannett B: **Lobar pneumonia: an occupational disease in welders.** *Lancet* 1994, **344(8914)**:41-43.

25. Doig AT, Challen PJ: **Respiratory hazards in welding.** *Ann Occup Hyg* 1964, **111**:223-231.

26. Smith JM, Smith NH, O'Rourke M, Spratt BG: **How clonal are bacteria?** *Proc Natl Acad Sci USA* 1993, **90(10)**:4384-4388.

27. Sorokin A, Candelon B, Guilloux K, Galleron N, Wackerow-Kouzova N, Ehrlich SD, Bourguet D, Sanchis V: **Multiple-locus sequence typing analysis of *Bacillus cereus* and *Bacillus thuringiensis* reveals separate clustering and a distinct population structure of psychrotrophic strains.** *Appl Environ Microbiol* 2006, **72(2)**:1569-1578.

28. Priest FG, Barker M, Baillie LW, Holmes EC, Maiden MC: **Population structure and evolution of the *Bacillus cereus* group.** *J Bacteriol* 2004, **186(23)**:7959-7970.

29. Ehling-Schulz M, Svensson B, Guinebretiere MH, Lindback T, Andersson M, Schulz A, Fricker M, Christiansson A, Granum PE, Martlbauer E, Nguyen-The C, Salkinoja-Salonen M, Scherer S: **Emetic toxin formation of *Bacillus cereus* is restricted to a single evolutionary lineage of closely related strains.** *Microbiology* 2005, **151(Pt 1)**:183-197.

30. Helgason E, Caugant DA, Olsen I, Kolsto AB: **Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections.** *J Clin Microbiol* 2000, **38(4)**:1615-1622.

31. Tourasse NJ, Helgason E, Okstad OA, Hegna IK, Kolsto AB: **The *Bacillus cereus* group: novel aspects of population structure and genome dynamics.** *J Appl Microbiol* 2006, **101(3)**:579-593.

32. Vassileva M, Torii K, Oshimoto M, Okamoto A, Agata N, Yamada K, Hasegawa T, Ohta M: **Phylogenetic analysis of *Bacillus cereus* isolates from severe systemic infections using multilocus sequence typing scheme.** *Microbiol Immunol* 2006, **50(9)**:743-749.

33. Sacchi CT, Whitney AM, Mayer LW, Morey R, Steigerwalt A, Boras A, Weyant RS, Popovic T: **Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus anthracis* .** *Emerg Infect Dis* 2002, **8(10)**:1117-1123.

34. Hoffmaster AR, Fitzgerald CC, Ribot E, Mayer LW, Popovic T: **Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States.** *Emerg Infect Dis* 2002, **8(10)**:1111-1116.

35. Ehling-Schulz M, Fricker M, Scherer S: **Identification of emetic toxin producing *Bacillus cereus* strains by a novel molecular assay.** *FEMS Microbiol Lett* 2004, **232(2)**:189-195.

36. Marston CK, Gee JE, Popovic T, Hoffmaster AR: **Molecular approaches to identify and differentiate *Bacillus anthracis* from phenotypically similar *Bacillus* species isolates.** *BMC Microbiol* 2006, **6**:22.

37. Saitou N, Nei M: **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4(4)**:406-425.

38. Kimura M: **A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences.** *J Mol Evol* 1980, **16(2)**:111-120.

39. Kumar S, Tamura K, Nei M: **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.** *Brief Bioinform* 2004, **5(2)**:150-163.

40. Hunter PR, Gaston MA: **Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity.** *J Clin Microbiol* 1988, **26(11)**:2465-2466.

41. Vassileva M, Torii K, Oshimoto M, Okamoto A, Agata N, Yamada K, Hasegawa T, Ohta M: **A new phylogenetic cluster of cereulide-producing *Bacillus cereus* strains.** *J Clin Microbiol* 2007, **45(4)**:1274-1277.

42. Sue D, Hoffmaster AR, Popovic T, Wilkins PP: **Capsule production in *Bacillus cereus* strains associated with severe pneumonia.** *J Clin Microbiol* 2006, **44(9)**:3426-3428.

43. Hoton FM, Andrup L, Swiecicka I, Mahillon J: **The cereulide genetic determinants of emetic *Bacillus cereus* are plasmid-borne.** *Microbiology* 2005, **151(Pt 7)**:2121-2124.

44. Helgason E, Okstad OA, Caugant DA, Johansen HA, Fouet A, Mock M, Hegna I, Kolsto : ***Bacillus anthracis, Bacillus cereus,* and *Bacillus thuringiensis*–one species on the basis of genetic evidence.** *Appl Environ Microbiol* 2000, **66(6)**:2627-2630.

45. Hill KK, Ticknor LO, Okinaka RT, Asay M, Blair H, Bliss KA, Laker M, Pardington PE, Richardson AP, Tonks M, Beecher DJ, Kemp JD, Kolsto A, Lee Wong AC, Keim P, Jackson PJ: **Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis, Bacillus cereus,* and *Bacillus thuringiensis* isolates.** *Appl Environ Microbiol* 2004, **70(2)**:1068-1080.

46. Hendriksen NB, Hansen BM, Johansen JE: **Occurrence and pathogenic potential of *Bacillus cereus* group bacteria in a sandy loam.** *Antonie Van Leeuwenhoek* 2006, **89(2)**:239-249.

47. Moravek M, Dietrich R, Buerk C, Broussolle V, Guinebretiere MH, Granum PE, Nguyen-The C, Martlbauer E: **Determination of the toxic potential of *Bacillus cereus* isolates by quantitative enterotoxin analyses.** *FEMS Microbiol Lett* 2006, **257(2)**:293-298.

48. Tourasse NJ, Kolsto AB: **SuperCAT: a supertree database for combined and integrative multilocus sequence typing analysis of the *Bacillus cereus* group of bacteria (including *B. cereus, B. anthracis* and *B. thuringiensis*).** *Nucleic Acids Res* 2008, **36**:D461-8.