FEMS Microbiology Letters, 365, 2018, fnx283

doi: 10.1093/femsle/fnx283
Advance Access Publication Date: 27 December 2017
Research Letter

RESEARCH LETTER – Food Microbiology

Comparative genomic survey of *Bacillus cereus sensu stricto* isolates from the dairy production chain in Brazil

Gabriel Augusto Marques Rossi[1,*], Higor Oliveira Silva[1], Carlos Eduardo Gamero Aguilar[1], Arina Lázaro Rochetti[2], Ben Pascoe[3,†], Guillaume Méric[3], Evangelos Mourkas[3], Matthew D. Hitchings[4], Luis Antonio Mathias[1], Vera Letticie de Azevedo Ruiz[2,‡], Heidge Fukumasu[2], Samuel K. Sheppard[3] and Ana Maria Centola Vidal[2]

[1]Departamento de Medicina Veterinária Preventiva e Reprodução Animal, UNESP – Univ. Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias (FCAV), Via de acesso Paulo Castellane, s/n, CEP 14884-900 Jaboticabal, São Paulo, Brazil, [2]Departamento de Medicina Veterinária, Avenida Duque de Caxias Norte 225, Universidade de São Paulo, Faculdade de Zootecnia e Engenharia de Alimentos (FZEA), CEP 13635-900 Pirassununga, São Paulo, Brazil, [3]The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, 4 South, Lab 0.39, Claverton Down, BA2 7AY Bath, UK and [4]Swansea University Medical School, Swansea, Swansea SA2 8PP, UK

*Corresponding author: Departamento de Medicina Veterinária Preventiva e Reprodução Animal, UNESP – Univ. Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias (FCAV), Via de acesso Paulo Castellane, s/n, CEP 14884-900 Jaboticabal, São Paulo, Brazil. Tel: +55 16 32092646; E-mail: gabrielrossiveterinario@gmail.com

One sentence summary: The population genomic analysis demonstrated the diversity of strains and variability of putative function among *Bacillus cereus* group isolates in Brazilian dairy production chain.

Editor: Miguel Gueimonde

†Ben Pascoe, http://orcid.org/0000-0001-6376-5121
‡Vera Letticie de Azevedo Ruiz, http://orcid.org/0000-0002-6983-8923

ABSTRACT

The genomes of 262 *Bacillus cereus* isolates were analyzed including 69 isolates sampled from equipment, raw milk and dairy products from Brazil. The population structure of isolates showed strains belonging to known phylogenetic groups II, III, IV, V and VI. Almost all the isolates obtained from dairy products belonged to group III. Investigation of specific alleles revealed high numbers of isolates carrying toxin-associated genes including *cytK* (53.62%), *hblA* (59.42%), *hblC* (44.93%), *hblD* (53.62%), *nheA* (84.06%), *nheB* (89.86%) and *nheC* (84.06%) with isolates belonging to groups IV and V having significant higher prevalence of *hblACD* and group IV of *CytK* genes. Strains from dairy products had significantly lower prevalence of *CytK* and *hblACD* genes compared to isolates from equipment and raw milk/bulk tanks. Genes related to sucrose metabolism were detected at higher frequency in isolates obtained from raw milk compared to strains from equipment and utensils. The population genomic analysis demonstrated the diversity of strains and variability of putative function among *B. cereus* group isolates in Brazilian dairy production, with large numbers of strains potentially able to cause foodborne illness. This

Received: 26 October 2017; **Accepted:** 23 December 2017

detailed information will contribute to targeted interventions to reduce milk contamination and spoilage associated with *B. cereus* in Brazil.

Keywords: population genomics; foodborne diseases; milk; next-generation sequencing

INTRODUCTION

Bacteria belonging to the *Bacillus cereus* group are widely distributed in the environment and are commonly isolated from soil, air, water, plants and feces. The ubiquity of these bacteria means that they frequently enter food production chains where human consumption of viable cells and preformed toxins can cause serious diarrheal disease and emetic syndrome, respectively. Food-poisoning outbreaks caused by this opportunistic pathogen have been widely reported (Zhou *et al*. 2014; Schimd *et al*. 2016), and *B. cereus* was responsible for an estimated 19% of reported foodborne outbreaks in the USA from 1998 to 2008 (Bennet, Walsh and Gould 2013). The actual number of outbreaks attributable to *B. cereus* may be higher, as detailed diagnostics to identify pathogenic strains are not widely used in clinical microbiology laboratories (Newell *et al*. 2010).

The genetic diversity within the *B. cereus* group (Raymond *et al*. 2010) and the mobility of virulence determinants among phenotypically defined species—often encoded on plasmids (Gonzalez, Brown and Carlton 1982; Okinaka *et al*. 1999; Raymond and Bonsall 2013)—have contributed to the difficulty in defining specific virulence and toxin production mechanisms in disease-causing lineages (Jeßberger *et al*. 2015). For example, the production of cereulide and diarrheal enterotoxin are not exclusive to *B. cereus sensu stricto* (Thorsen *et al*. 2006; Kovac *et al*. 2016), and some strains can harbor *B. anthracis* genes and are considered able to cause anthrax disease (Marston *et al*. 2016) and bacteremia (Sasahara *et al*. 2011; Schaefer *et al*. 2016). Detailed characterization of natural populations using comparative genomics techniques has the potential to identify strains that proliferate in food production systems and the genes associated with outbreak strains.

Bacillus cereus group isolates have been identified in numerous foods but food poisoning is principally associated with heat-treated foods where the ability to produce metabolically inert resistant spores facilitates survival in adverse processing conditions (Daelman *et al*. 2013). The abundance of *B. cereus* in animals, including cows, makes it highly prevalent in the dairy farm environment, contaminating raw milk by simple transfer during milking (Magnusson, Christiansson and Svensson 2007; Vissers *et al*. 2007; Coorevits *et al*. 2008; Masiello *et al*. 2014; McAuley *et al*. 2014). Pasteurization may induce sporulation, and the spores can subsequently survive the pasteurization process and contaminate dairy products (Bartoszewicz, Hansen and Swiecicka 2008; Miller *et al*. 2015). Adoption of cleaning systems and good hygiene practices in dairy processing can reduce spore contamination, the proliferation of vegetative cells and toxin formation (Shaheen *et al*. 2010; Kumari and Sarkar 2014), and *B. cereus* foodborne outbreaks are rarely linked with milk or dairy products consumption (Bennet, Walsh and Gould 2013). However, the risk for public health is dependent on careful management.

Increasing demand for cheap milk and dairy products, and poor standards, such as the presence of feces in the dairy environment, improper bedding (Magnusson, Christiansson and Svensson 2007) and use of silage (Vissers *et al*. 2007) can contribute to milk contamination with *B. cereus* spores. This is exacerbated by the capacity to form biofilms that can lead to persistent recontamination during the dairy production process (Salustiano *et al*. 2009; Kumari and Sarkar 2016). As a result, *B. cereus* remains a threat to public health as well as reducing production through spoilage of products. Furthermore, in countries such as Brazil, where there is higher consumption of dairy products sold through unregulated markets (Vidal-Martins *et al*. 2013), there is higher incidence of human infection from toxigenic *B. cereus* in raw milk and dairy products (Vidal-Martins *et al*. 2006; Aragon-Alegro *et al*. 2008; Salustiano *et al*. 2009; Chaves, Pires and Vivoni 2011; Reis *et al*. 2013).

In this study, we aim to clarify the population structure of *B. cereus* group isolates in Brazilian dairy production chains. Using structured sampling of *B. cereus* in dairy (farms, industries and products) and comparative genomics techniques, we investigate if specific lineages and genes are significantly overrepresented in particular production stages and products. This study provides evidence of the genomic factors that contribute to survival of *B. cereus* from farm to fork in Brazil.

MATERIAL AND METHODS

Bacterial sampling and genomes

A total of 466 samples were obtained from the dairy production chain (dairy farms (331 samples), industries (58) and products on supermarkets (77)) in the municipality of Pirassununga, located in the central-eastern region of the state of São Paulo, Brazil, from July to November 2016. A subset of 69 isolates were obtained from these samples and confirmed phenotypically to be *B. cereus sensu stricto* through isolation on Mannitol Egg Yolk Polymyxin agar plates. The remaining 397 samples were not contaminated with the *B. cereus* group or had other microorganisms concomitantly and were not sequenced. A set of 38 isolates were obtained from 26 distinct dairy farm environments in Pirassununga, isolated from the surface of bulk tanks (7), a manual milking can (1), milk cans (5), milk pipelines (4), raw milk (6), teats cups (5) and other utensils/equipment (10) that had direct contact with milk. Furthermore, nine isolates were obtained in two dairy processing plants located in the same region from the dairy trucks dumping raw milk (2), equipment (3), milk pipeline (1) and utensils (3) used in dairy production. The samples were collected using swabs or sterile tubes only for raw milk. The other 22 isolates were obtained from dairy products available for consumption, such as Minas cheese (1), UHT dairy beverage (1), cream prepared with cheese (1), instant cappuccino (1) and Brazilian curd cheese 'requeijão' (18). These 69 isolates were also categorized into three groups according to isolation origin as following: (i) equipments and utensils (34); (ii) raw milk and bulk tanks (13); (iii) dairy products (22). Furthermore, isolates were clustered into previously defined phylogenetic groups (Guinebretière *et al*. 2008).

Sequence data of 69 isolates obtained in Brazilian dairy production are archived in the NCBI GenBank repositor and the Short Read Archive (SRA) associated with BioProject: PRJNA390851 (Table S1, Supporting Information).

Assembled genomes are also available on FigShare (10.6084/m9.figshare.5120020). Full details and individual accession numbers of previously published genomes used as a reference (n = 193 strains) can be found in Table S2 (Supporting Information).

Bacterial culture and biochemical typing

Pre-enrichment of the samples was performed using Tryptic-Soy Browth (TSB) (Oxoid, Hampshire, UK) with Polymyxin B addition (20 μg/mL) (Stadhouders 1992) and incubated during 24 h at 30°C. Afterwards, the samples were transferred to Mannitol Egg Yolk Polymyxin Agar (Oxoid) plates (Mossel, Koopman and Jongerius 1967) using the streaking technique and incubated during 24 to 48 h at 30°C. One presumptive colony was transferred to tubes containing inclined Trypticase Soy Agar (TSA) (Oxoid) and incubated during 24 h at 30°C. These colonies were characterized using Gram staining and biochemical tests. Strains classified as *B. cereus sensu stricto* were those with the following biochemical characteristics: catalase (+), glucose fermentation (+), Voges-Proskauer (+), rhizoid growth (–), sheep's blood hemolysis (+), crystals production (–), motility (+/–) and nitrate reduction to nitrite (+/–) (MacFadin 1976; Sharif and Alaeddinoglu 1988; APHA 2001).

DNA extraction and genome sequencing

Bacteria were transferred from tubes containing TSA (Oxoid) to tubes containing TSB (Oxoid) and were incubated during 24 h at 30°C. Aliquots of 1.5 mL of suspended sample were transferred to eppendorf tubes to perform DNA extraction using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, St. Louis, MO) according to the manufacturer's instructions for Gram-positive bacteria. DNA quality and concentration were evaluated by fluorimetric quantitation using a DeNovix DS-11 system. Sequencing libraries were prepared using the Illumina Nextera® XT Library Prep Kit (v3) according to the manufacturer's protocol. Briefly, this process involves fragmentation of 1 μg genomic DNA by enzymatic fragmentation, labeling with adapters' ligation and PCR using the Illumina 3 primer set. AMPure® XP paramagnetic beads (Beckman Coulter, Inc., USA) were used to perform DNA cleanup after each step to remove small fragments of DNA. Libraries were normalized and pooled for sequencing in an Illumina MiSeq.

Genome assembly and analysis

Contiguous sequences (contigs) were assembled using SPAdes v.3.8.0 (Bankevich *et al.* 2012). Individual genes were aligned to *B. cereus* reference strain (ATCC 14579) using default BLAST parameters in BIGSdb (Jolley and Maiden 2010). The BLAST algorithm was used to scan all genomes for gene orthologs at each locus in the reference genome. An ortholog was defined as a reciprocal best hit of the sequence with >70% nucleotide identity over 50% of the difference in alignment length. MAFFT software was used to align gene orthologs on a gene-by-gene basis, and these data concatenated into contiguous sequence for each isolate genome (Sheppard, Jolley and Maiden 2012). Gene discovery, sequence export and gene-by-gene alignments were performed using MUSCLE (Edgar 2004).

Neighbour joining (NJ) trees were constructed using rapidnj (Simonsen, Mailund and Pedersen 2008) from 69 whole genome sequences compared with 193 publicly available *B. cereus* group genomes. For comparison, a maximum-likelihood (ML) phylogeny was also constructed using FastTree version 2.1 (Price, Dehal and Arkin 2010) and an approximation of the maximum-likelihood algorithm was visualized using Phylo.io (Robinson, Dylus and Dessimoz 2016) (Fig. S3, Supporting Information). A set of 69 isolates was also classified according into seven distinct phylogenetic groups according to *panC* sequence (Guinebretière *et al.* 2008) in the online tool https://www.tools.symprevius.org/bcereus/english.php. The multilocus sequence type of each isolate was defined based on allelic variation of seven housekeeping genes (*glp, gmK, ilvD, pta, pur, pyc* and *tpi*) (Priest *et al.* 2004). Available nucleotide sequences from NCBI were used to identify toxigenic genes from *B. cereus* group as follows: *cesA* (NC_010924), *cesB* (NC_010924.1), *cytK* (NC_005957.1), *hblA* (NC_005957.1), *hblC* (NC_005957.1), *hblD* (NC_005957.1), *nheA* (NC_005957.1), *nheB* (NC_003909.8) and *nheC* (NC_005957.1).

Statistical analysis

Fisher's exact test using 5% of significance ($P < 0.05$) was used to compare the prevalence of genes among phylogenetic groups and according to isolates' origin.

RESULTS

Population structure and epidemiology of milk-associated *B. cereus* group isolates

The clustering of isolates is consistent between NJ and ML genealogies among the phylogenetic groups (Fig. S3, Supporting Information). The population structure of 262 strains belonging to *B. cereus* group generally agreed with traditional genotyping designations based on the sequence of *panC* gene (Guinebretière *et al.* 2008) (Fig. 1), except by one strain that was classified as group IV using *panC* sequence but it was clustered with the strains belonging to group V using whole genome sequencing. No strain belonging to phylogenetic groups I or VII was detected in this study in which *B. pseudomycoides* and *B. cytotoxicus* are included (Guinebretière *et al.* 2008, 2010). Only one strain belonged to group VI, which is composed of *B. weihenstephanensis, B. mycoides and B. thuringiensis* strains. The other isolates obtained from samples of raw milk and bulk tanks and those obtained from samples of utensils and equipment from dairy farms and industries were distributed in different clusters belonging to phylogenetic groups II, III, IV or V (Guinebretière *et al.* 2008). However, almost all the isolates obtained from dairy products (mainly curd cheeses 'requeijão' cheeses) belonged to group III, which is composed of *B. cereus sensu stricto* (including emetic strains), *B. thuringiensis* and *B. anthracis* species. Using MLST, we detected sequence types ST18, ST46, ST144, ST182, ST223, ST247, ST554 on milking cans, teat-cups and other equipment and utensils from dairy farms and industries. In bulk tanks, ST18, ST73 and ST144 were detected. In dairy products, ST5, ST26 and ST92 were detected.

Assessment of toxigenic potential by genotyping

The potential to produce four toxins (cytotoxin Cytk, hemolysin BL, non-hemolytic enterotoxin and cereulide) was evaluated in isolates. The presence of putative toxin-producing genes was investigated in the genome including *cesA, cesB, cytK, hblA, hblC, hblD, nheA, nheB* and *nheC* (Ehling-Schulz *et al.* 2005; Ehling-Schulz *et al.* 2006; Ngamwongsatit *et al.* 2008). A total of 66 (95.65%) isolates carried at least one of these genes. All three (4.35%) isolates that did not carry any of these genes were

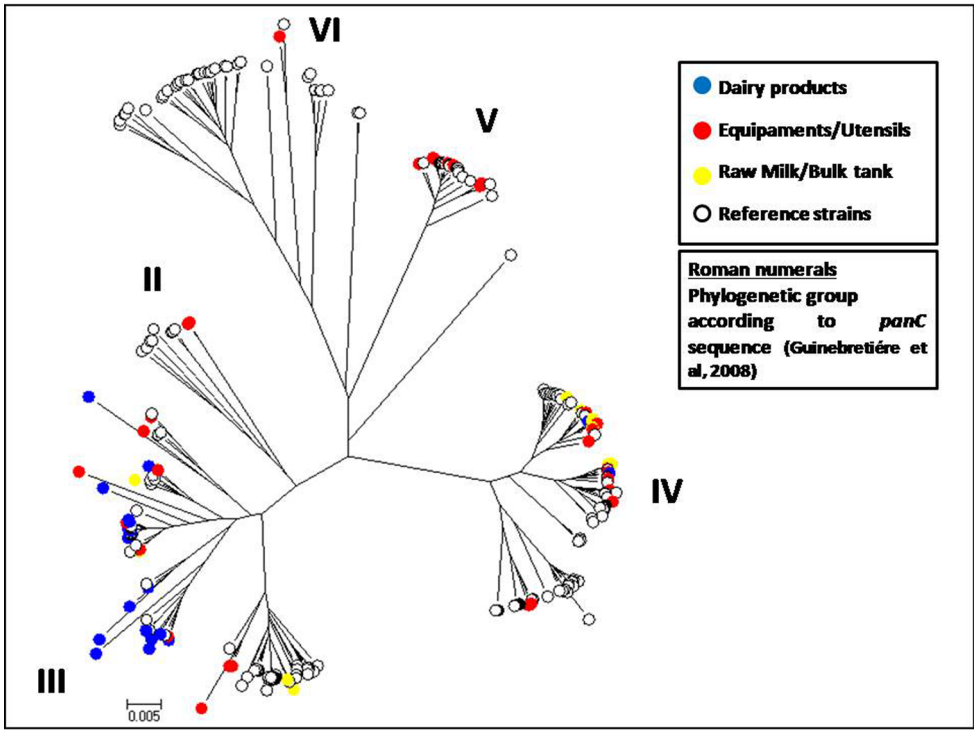

Figure 1. Population structure of 262 isolate genomes (69 from Brazilian dairy production chain) belonging to *B. cereus* group. The NJ trees was constructed using rapidnj (Simonsen, Mailund and Pedersen 2008) from whole genome sequencing (5210 genes). Isolate origin is presented in the leaves on the tree, classified in dairy products (blue), equipment and utensil (red), raw milk and bulk tank (yellow) and the background population comprised 193 strains from published collections (white). The phylogenetic groups are defined according to genotyping information of the *panC* gene (II to VI) (Guinebretière *et al.* 2008).

obtained from curd cheese requeijão. High numbers of isolates carried *cytK* (53.62%), *hblA* (59.42%), *hblC* (44.93%), *hblD* (53.62%), *nheA* (84.06%), *nheB* (89.86%) and *nheC* (84.06%) genes. The prevalence of putative toxin genes was evaluated based on isolate origin (Table 1) and phylogenetic group (Table 2).

The strains included in group IV and V had significantly higher prevalence of *hblA* ($P = 9.457 \times 10^{-11}$), *hblC* ($P = 8.846 \times 10^{-9}$) and *hblD* ($P = 2.432 \times 10^{-12}$) compared with those included in group III, using Fisher's exact test (5% of significance). Also, a significantly ($P = 1.082 \times 10^{-7}$) higher prevalence of *CytK* was observed in isolates included in group IV than in groups III and V. Fisher's exact test was also performed to evaluate statistical differences of gene prevalence among groups according to isolation origin. *CytK* was significantly ($P = 0.01$) more prevalent in isolates obtained from milk samples than those from processed dairy products. Also, *hblA* ($P = 8.312 \times 10^{-5}$), *hblC* ($P = 0.001$) and *hblD* ($P = 0.0002$) were significantly more prevalent in isolates obtained from equipment and utensils than in those obtained from dairy products. The presence of *cesAB* genes, related to emetic-toxin production, was observed in two isolates (2.89%): one from curd cheese requeijão and other from milk pipeline in a dairy farm.

Accessory genes associated with the later stage of dairy production chain

Some clues about the genes that may confer an advantage in survival through dairy processing can be obtained by quantifying genes that are enriched in strains isolated at the latter stages of the production chain. The difference in gene prevalence was determined by comparing *B. cereus* group isolates from three distinct categories: equipment and utensils; raw milk and silo tanks; dairy products. The greatest difference in gene prevalence between isolates obtained from raw milk and silo tanks and those from equipment and utensils was among genes putatively involved in sucrose metabolism. Specifically, the gene cassette composed of genes putatively encoding fructokinase (*FruC*), sucrose-6-phosphate hydrolase (*SacA*), sucrose-specific PTS (*SacP*) and sucrose open repressor (*sacR*) (Fouet *et al.* 1987)—which had the highest difference in prevalence among isolates obtained from milk compared to those from equipment (57%). The prevalence was 92% (11/12) in isolates from raw milk, while the value was 35% (11/34) from those obtained from equipment. Furthermore, the main difference (79%) comparing the prevalence of genes among the isolates from milk/tanks (92%—11/12) with those from dairy products (13%—3/22) was observed in the *amyS* gene, which is putatively responsible for alpha-amylase production and supports growth on starch/glycogen (Mols *et al.* 2007).

DISCUSSION

Bacillus cereus group isolates sampled from the Brazilian dairy production chain were distributed in several known phylogenetic groups (Guinebretière *et al.* 2008; Caamaño-Antelo *et al.* 2015). The population structure of strains belonging to *B. cereus* group generally agreed with traditional genotyping designations based on the sequence of *panC* gene (Guinebretière *et al.* 2008), except by one isolate. This difference could be attributed to the methods used in this study (NJ tree and traditional genotyping) as previously reported (Carroll *et al.* 2017) or detection of new putative *B. cereus* group species (Liu *et al.* 2017). The presence of only one strain belonging to group VI can be explained by the psychrotrophic characteristic of bacteria included in this

Table 1. Prevalence of genes related to toxins production (*cesAB*, *cytK*, *hblACD* and *nheABC*) in *B. cereus* group obtained from equipment and utensils, raw milk or bulk tank and dairy products in the state of São Paulo, Brazil, during 2016.

	Equipments/utensils				Raw milk/bulk tank				Dairy products				Total		
	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%	+	N	P (%)
cesA	1	34	2.94	0.52 to 14.92	0	13	0	0.00 to 22.81	1	22	4.55	0.81 to 21.80	2	69	2.9
cesB	1	34	2.94	0.52 to 14.92	0	13	0	0.00 to 22.81	1	22	4.55	0.81 to 21.80	2	69	2.9
cytK	19	34	55.88	39.19 to 72.57	11	13	84.62	57.76 to 95.67	7	22	31.82	16.36 to 52.68	37	69	53.62
hblA	27	34	79.41	65.82 to 93.00	9	13	69.23	42.37 to 87.32	5	22	22.73	10.12 to 43.44	41	69	59.42
hblC	21	34	61.76	45.43 to 78.10	7	13	53.85	29.14 to 76.79	3	22	13.64	4.75 to 33.34	31	69	44.93
hblD	24	34	70.59	55.27 to 85.90	9	13	69.23	42.37 to 87.32	4	22	18.18	7.31 to 38.52	37	69	53.62
nheA	30	34	88.24	73.38 to 95.33	12	13	92.31	66.69 to 98.63	16	22	72.73	51.85 to 86.85	58	69	84.06
nheB	32	34	94.12	80.91 to 98.37	13	13	100	77.19 to 100.00	17	22	77.27	56.56 to 89.88	62	69	89.86
nheC	29	34	85.29	69.87 to 93.55	13	13	100	77.19 to 100.00	16	22	72.73	51.85 to 86.85	58	69	84.06

+ = Number of positive isolates; N = Number of isolates; P = prevalence and C. I 95% = Confidence interval 95%.

Table 2. Prevalence of genes related to toxins production (*cesAB*, *cytK*, *hblACD* and *nheABC*) in *B. cereus* phylogenetic groups (Guinebretière *et al.* 2008) in the state of São Paulo, Brazil, during 2016.

	II				III				IV				V				VI			
	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%	+	N	P (%)	C. I 95%
cesA	0	2	0	0.0–65.76	1	34	2.94	0.52–14.92	1	22	4.55	0.81–21.80	0	8	0	0.0–32.44	0	1	0	0.0–79.35
cesB	0	2	0	0.0–65.76	1	34	2.94	0.52–14.92	1	22	4.55	0.81–21.80	0	8	0	0.0–32.44	0	1	0	0.0–79.35
cytK	0	2	0	0.0–65.76	12	34	35.29	21.49–52.09	22	22	100	85.13–100.0	3	8	37.5	13.68–69.43	0	1	0	0.0–79.35
hblA	2	2	100	34.24–100.0	8	34	23.53	12.44–40.0	22	22	100	85.13–100.0	8	8	100	67.56–100.0	1	1	100	20.65–100.0
hblC	2	2	100	34.24–100.0	4	34	11.76	4.67–26.62	16	22	72.73	51.85–86.85	8	8	100	67.56–100.0	1	1	100	20.65–100.0
hblD	2	2	100	34.24–100.0	5	34	14.71	6.45–30.13	22	22	100	85.13–100.0	7	8	87.5	52.91–97.76	1	1	100	20.65–100.0
nheA	2	2	100	34.24–100.0	27	34	79.41	63.20–89.65	20	22	90.91	72.18–97.47	8	8	100	67.56–100.0	1	1	100	20.65–100.0
nheB	2	2	100	34.24–100.0	30	34	88.24	73.38–95.33	22	22	100	85.13–100.0	7	8	87.5	52.91–97.76	1	1	100	20.65–100.0
nheC	2	2	100	34.24–100.0	29	34	85.29	69.87–93.55	19	22	86.36	66.67–95.25	7	8	87.5	52.91–97.76	1	1	100	20.65–100.0

+ = Number of positive isolates; N = Number of isolates; P = prevalence and C. I 95% = Confidence interval 95%.

group, which are rare in tropical countries (Von Stetten, Mayr and Scherer 1999). Furthermore, the exclusion of colonies with rhizoid morphology contributes to explain the non-detection of phylogenetic group I. Within these branching phylogenetic groups, specific strains of STs have been shown to be involved in foodborne outbreaks (Cardazzo *et al.* 2008; Yang *et al.* 2017). These include ST26, ST92 and ST144, isolated from curd cheese in this study, that have previously been involved in outbreaks (Hoffmaster *et al.* 2008). The genes *cytK*, *entFM*, *nhe* and *hbl* are reported in foodborne *B. cereus* STs (Cardazzo *et al.* 2008). ST26 and ST144 have also been associated with emesis (Priest *et al.* 2004; Hoffmaster *et al.* 2008; Didelot *et al.* 2009) pneumonia (Hoffmaster *et al.* 2008) and contaminated infant formula in China (Yang *et al.* 2017). Other STs detected in this study (ST173 and ST18) have been linked to food-poisoning outbreaks (Hoffmaster *et al.* 2008). With such a diversity of STs potentially causing disease, further research is required to understand the infection potential of the isolates described in this study.

Different phylogenetic groups have been shown to have different potential to cause disease. For example, group IV strains obtained from dairy origin pose a higher potential to produce hemolysin BL (Kovac *et al.* 2016). In our study, groups IV and V strains had higher prevalence of *hblA hblC* and *hblD* compared to those from group III. However, toxin production is dependent on regulatory mechanisms (Kovac *et al.* 2016) and intestinal conditions (Jeßberger *et al.* 2017). In our study, a high proportion of isolates belonged to phylogenetic group III (mainly curd cheeses 'requeijão' cheeses). This group is considered to be of high risk for foodborne disease (Guinebretière *et al.* 2008), containing *B. cereus sensu lato* lineages, including those considered as emetic, cytotoxic and 'highly cytotoxic' (Guinebretière *et al.* 2008, 2010). The presence of strains belonging to group III in heat-treated dairy products, such as observed in this study, could be related with a higher capacity of spores in resisting heat treatment processes (Luu-Thi, Khadka and Michiels 2014).

Alignment of sequences of whole genomes allows the identification of genes putatively involved in foodborne illness. For example, isolates carrying genes related to toxin production such as the *nhe* and *hbl* genes that are associated with diarrheal syndrome (Guinebretière *et al.* 2010) have been reported in isolates from milk (Granum, Brynestad and Kramer 1993). In our study, there was a high prevalence of the *cytK, hblA, hblC* and *hblD* genes in isolates from milk and equipment and utensils compared with those from dairy products. This potentially indicates different potential for toxin production in Brazilian dairy production, with consumption of raw milk being a potential risk for public health. The *ces* gene, associated with emetic-toxin production, was observed in two isolates in our study: one from curd cheese and the other from a milk pipeline in a dairy farm. The presence of emetic strains in dairy production chain is considered rare (Svensson *et al.* 2006; Cui *et al.* 2016), and the detection of a potential emetic strain in refrigerated 'requeijão' is a concern for consumers as production of cytotoxic isocereulide A is higher at low temperatures (Kranzler *et al.* 2016).

In addition to detecting putative toxin genes, the comparative genomic analysis allowed the identification of genes that may confer an advantage for survival or growth in dairy products. We detected a high prevalence of genes related to the sucrose pathway in isolates obtained from raw milk and bulk tank samples compared to those from equipment and utensils and dairy products. Sucrose utilization has been demonstrated among *B. cereus* strains isolated from milk (Wong, Chang and Fan 1988) but it is difficult to explain functionally as sucrose is not a natural component of milk (Jenkins and McGuire 2006). Sucrose is, however, a common plant carbohydrate, and the main *B. cereus* contamination sources for raw milk are spores in grazing systems from cow udders contaminated with soil, feces and bedding material (Christiansson, Bertilsson and Svensson 1999; Magnusson, Christiansson and Svensson 2007). This potentially explains the presence of strains able to metabolize sucrose in raw milk and bulk tanks.

The metabolism of carbohydrate is also related to toxin production (Ouhib, Clavel and Schmitt 2006). The induction of macrophage death at infection sites, and access to carbon sources and nutrients for bacterial growth, can occur through mechanisms requiring glucose and iron availability (Sineva *et al.* 2009; Tran *et al.* 2011, 2013; Guillemet *et al.* 2013). This potentially provides alternative explanations for the prevalence of putative sucrose pathway genes among *B. cereus* group isolates in raw milk. First, because macrophage killing potentially provides a competitive advantage within the microbial community (Ceuppens, Boon and Uyttendaele 2013) of raw milk where there are high numbers of viable macrophages (Langoni *et al.* 2011; Li *et al.* 2015). Second, because milk cows infected with subclinical mastitis, and hence elevated macrophage numbers are a major source of milk contamination (Bhatt *et al.* 2012). However, these explanations are not tested in this study.

In conclusion, we have demonstrated the utility of whole genome sequencing for characterizing the population structure of potentially toxigenic *B. cereus* group isolates in the dairy production in Brazil. Detailed information about the differential abundance of strains and specific genes across the production chain provides a basis for attributing the source of contamination and the detection of genetic elements that may underlie important phenotypes associated with persistence and pathogenicity. Clearly, the presence of potentially pathogenic strains in dairy production in Brazil requires the adoption of hygienic practices and avoidance of raw milk to improve food safety and reduce spoilage of dairy products.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSLE* online.

FUNDING

This work was supported by São Paulo Research Foundation (FAPESP) (Grant numbers 2014/13104–1, 2015/20874–0 and 2016/19214–9).

Conflict of Interest. None declared.

REFERENCES

American Public Health Association, Committee on Microbiological for Foods. *Compendium of Methods for the Microbiological Examination of Foods*, 4th edn. Washington, DC: American Public Health Association, 2001.

Aragon-Alegro LC, Palcich G, Lopes GV *et al.* Enterotoxigenic and genetic profiles of *Bacillus cereus* strains of food origin in Brazil. *J Food Protect* 2008, DOI: 10.4315/0362-028X-71.10.2115.

Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comp Biol* 2012, DOI: 10.1089/cmb.2012.0021.

Bartoszewicz M, Hansen BM, Swiecicka I. The members of the *Bacillus cereus* group are commonly present contaminants of fresh and heat-treated milk. *Food Microbiol* 2008, DOI: 10.1016/j.fm.2008.02.001.

Bennet SD, Walsh KA, Gould LH. Foodborne disease outbreaks caused by *Bacillus cereus, Clostridium perfringens*, and *Staphylococcus aureus* — United States, 1998–2008. *Clin Infect Dis* 2013, DOI: 10.1093/cid/cit244.

Bhatt VD, Ahir VB, Koringa PG *et al.* Milk microbiome signatures of subclinical mastitis-affected cattle analysed by shotgun sequencing. *J Appl Microbiol* 2012, DOI: 10.1111/j.1365-2672.2012.05244.x.

Caamaño-Antelo S, Fernández-No IC, Böhme K *et al.* Genetic discrimination of foodborne pathogenic and spoilage *Bacillus* spp. based on three housekeeping genes. *Food Microbiol* 2015, DOI: 10.1016/j.fm.2014.08.013.

Cardazzo B, Negrisolo E, Carraro L *et al.* Multiple-locus sequence typing and analysis of toxin genes in *Bacillus cereus* food-borne isolates. *Appl Environ Microb* 2008, DOI: 10.1128/AEM.01495-07.

Carroll LM, Kovac J, Miller RA *et al.* Rapid, High-Throughput identification of anthrax—causing and emetic *Bacillus cereus* group genome assemblies via BTyper, a computional tool for virulence-based classification of *Bacillus cereus* group isolates by using nucleotide sequecing data. *Appl Environ Microb* 2017, DOI: 10.1128/AEM.01096-17.

Ceuppens S, Boon N, Uyttendaele M. Diversity of *Bacillus cereus* group strains is reflected in their broad range of pathogenicity and diverse ecological lifestyles. *FEMS Microbiol Ecol* 2013, DOI: 10.1111/1574-6941.12110.

Chaves JQ, Pires ES, Vivoni AM. Genetic diversity, antimicrobial resistance and toxigenic profiles of *Bacillus cereus* isolated from food in Brazil over three decades. *Int J Food Microbiol* 2011, DOI: 10.1016/j.ijfoodmicro.2011.02.029.

Christiansson A, Bertilsson J, Svensson B. *Bacillus cereus* spores in raw milk: factors affecting the contamination of milk during the grazing period. *J Dairy Sci* 1999, DOI: 10.3168/jds.S0022-0302(99)75237-9.

Coorevits A, De Jonghe V, Vandroemme J *et al.* Comparative analysis of the diversity of aerobic spore-forming bacteria in raw milk from organic and conventional dairy farms. *Syst Appl Microbiol* 2008, DOI: 10.1016/j.syapm.2008.03.002.

Cui Y, Liu Y, Liu X *et al.* Evaluation of the toxicity and toxicokinetics of cereulide from an emetic *Bacillus cereus* strain of milk origin. *Toxins* 2016, DOI: 10.3390/toxins8060156.

Daelman J, Vermeulen A, Willemyns T *et al.* Growth/no growth models for heat-treated psychrotrophic *Bacillus cereus* spores under cold storage. *Int J Food Microbiol* 2013, DOI: 10.1016/j.ijfoodmicro.2012.11.017.

Didelot X, Barker M, Falush D *et al.* Evolution of pathogenicity in the *Bacillus cereus* group. *Syst Appl Microbiol* 2009, DOI: 10.1016/j.syapm.2009.01.001.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004, DOI: 10.1093/nar/gkh340.

Ehling-Schulz M, Fricker M, Grallert H *et al.* Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1. *BMC Microbiol* 2006, DOI: 10.1186/1471-2180-6-20.

Ehling-Schulz M, Vukov N, Schulz A *et al.* Identification and partial characterization of the nonribosomal peptide synthase gene responsible for cereulide production in emetic *Bacillus cereus*. *Appl Environ Microb* 2005, DOI: 10.1128/AEM.71.1.105-113.2005.

Fouet A, Arnaud M, Klier A *et al. Bacillus subtilis* sucrose-specific enzyme II of the phosphotransferase system: expression in *Escherichia coli* and homology to enzymes II from enteric bacteria. *P Natl Acad Sci USA* 1987;**84**:8773–7.

Gonzalez JM, Brown BJ, Carlton BC. Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringienisis* and *B. cereus*. *P Natl Acad Sci USA* 1982;**79**:6951–5.

Granum PE, Brynestad S, Kramer JM. Analysis of enterotoxin production by *Bacillus cereus* from dairy products, food poisoning incidents and non-gastrointestinal infections. *Int J Food Microbiol* 1993;**17**:269–79.

Guillemet E, Tran SL, Cadot C *et al.* Glucose 6P binds and activates HlyIIR to repress *Bacillus cereus* haemolysin hlyII gene expression. *PLoS One* 2013, DOI: 10.1371/journal.pone.0055085.

Guinebretière MH, Thompson FL, Sorokin A *et al.* Ecological diversification in the *Bacillus cereus* group. *Environ Microbiol* 2008, DOI: 10.1111/j.1462-2920.2007.01495.x.

Guinebretière MH, Velge P, Couvert O *et al.* Ability of *Bacillus cereus* group strains to cause food poisoning varies according to phylogenetic affiliation (groups I to VII) rather than species affiliation. *J Clin Microbiol* 2010, DOI: 10.1128/JCM.00921-10.

Hoffmaster AR, Novak RT, Marston CK *et al.* Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing. *BMC Microbiol* 2008, DOI: 10.1186/1471-2180-8-191.

Jenkins TC, McGuire MA. Major advances in nutrition: impact on milk composition. *J Dairy Sci* 2006, DOI: 10.3168/jds.S0022-0302(06)72198-1.

Jeßberger N, Krey VM, Rademacher C *et al.* From genome to toxicity: a combinatory approach highlights the complexity of enterotoxin production in *Bacillus cereus*. *Front Microbiol* 2015, DOI: 10.3389/fmicb.2015.00560.

Jeβberger N, Redemacher C, Krey KM *et al.* Simulating intestinal growth conditions enhances toxin production of enteropathogenic *Bacillus cereus*. *Front Microbiol* 2017, DOI: 10.3389/fmicb.2017.00627.

Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010, DOI: 10.1186/1471-2105-11-595.

Kovac J, Miller RA, Carroll LM *et al.* Production of hemolysin BL by *Bacillus cereus* group isolates of dairy origin is associated with whole-genome phylogenetic clade. *BMC Genomics* 2016, DOI: 10.1186/s12864-016-2883-z.

Kranzler M, Stollewerk K, Rouzeau-Szynalski K *et al.* Temperature exerts control of *Bacillus cereus* emetic toxin production on post-transcriptional levels. *Front Microbiol* 2016, DOI: 10.3389/fmicb.2016.01640.

Kumari S, Sarkar PK. In vitro model study for biofilm formation by *Bacillus cereus* in dairy chilling tanks and optimization of clean-in-place (CIP) regimes using response surface methodology. *Food Control* 2014, DOI: 10.1016/j.foodcont.2013.08.014.

Kumari S, Sarkar PK. *Bacillus cereus* hazard and control in industrial dairy processing environment. *Food Control* 2016, DOI: 10.1016/j.foodcont.2016.04.012.

Langoni H, Penachio DDS, Citadella JC *et al.* Quality and microbiological aspects of bovine milk. *Pesq Vet Bras* 2011, DOI: 10.1590/S0100-736X2011001200004.

Li N, Richoux R, Perruchot MH *et al.* Flow cytometry approach to quantify the viability of milk somatic cell counts after various physico-chemical treatments. *PLoS One* 2015, DOI: 0.1371/journal.pone.0146071.

Liu Y, Du J, Lai Q *et al.* Proposal of nine novel species of the *Bacillus cereus* group. *Int J Syst Evol Microbiol* 2017, DOI:10.1099/ijsem.0.001821.

Luu-Thi H, Khadka DB, Michiels CW. Thermal inactivation parameters of spores from different phylogenetic

groups of *Bacillus cereus*. *Int J Food Microbiol* 2014, DOI: 10.1016/j.ijfoodmicro.2014.07.027.

McAuley CM, McMillan K, Moore SC *et al*. Prevalence and characterization of foodborne pathogens from Australian dairy farm environments. *J Dairy Sci* 2014, DOI: 10.3168/jds.2014-8735.

MacFadin JF. *Biochemical Tests for Identification of Medical Bacteria*. Baltimore: The Williams & Wilkins, 1976, 312.

Magnusson M, Christiansson A, Svensson B. *Bacillus cereus* spores during housing of dairy cows: factors affecting contamination of raw milk. *J Dairy Sci* 2007, DOI: 10.3168/jds.2006-754.

Marston CK, Ibrahim H, Lee P *et al*. Anthrax toxin-expressing *Bacillus cereus* isolated from an anthrax-like eschar. *PLoS One* 2016, DOI: 10.1371/journal.pone.0156987.

Masiello SN, Martin NH, Watters RD *et al*. Identification of dairy farm management practices associated with the presence of psychrotolerant sporeformers in bulk tank milk. *J Dairy Sci* 2014, DOI: 10.3168/jds.2014-7938.

Miller RA, Kent DJ, Watterson MJ *et al*. Spore populations among bulk tank raw milk and dairy powders are significantly different. *J Dairy Sci* 2015, DOI: 10.3168/jds.2015-9943.

Mols M, de Been M, Zwietering MH *et al*. Metabolic capacity of *Bacillus cereus* strains ATCC 14579 and ATCC 10987 interlinked with comparative genomics. *Environ Microbiol* 2007, DOI: 10.1111/j.1462-2920.2007.01404.x.

Mossel DAA, Koopman MJ, Jongerius E. Enumeration of *Bacillus cereus* in foods. *Appl Microbiol* 1967;**15**:650–3.

Newell DG, Koopmans M, Verhoef L *et al*. Food-borne diseases — The challenges of 20 years ago still persist while new ones continue to emerge. *Int J Food Microbiol* 2010, DOI: 10.1016/j.ijfoodmicro.2010.01.021.

Ngamwongsatit P, Buasri W, Pianariyanon P *et al*. Broad distribution of enterotoxin genes (hblCDA, nheABC, cytK, and entFM) among *Bacillus thuringiensis* and *Bacillus cereus* as shown by novel primers. *Int J Food Microbiol* 2008, DOI: 10.1016/j.ijfoodmicro.2007.11.013.

Okinaka RT, Cloud K, Hampton O *et al*. Sequence and organization of pXO1, the large *Bacillus anthracis* plasmid harboring the anthrax toxin genes. *J Bacteriol* 1999;**181**:6509–15.

Ouhib O, Clavel T, Schmitt P. The production of *Bacillus cereus* enterotoxins is influenced by carbohydrate and growth rate. *Curr Microbiol* 2006, DOI: 10.1007/s00284-006-0094-6.

Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 2010, DOI: 10.1371/journal.pone.0009490.

Priest FG, Barker M, Baillie LW *et al*. Population structure and evolution of the *Bacillus cereus* group. *J Bacteriol* 2004, DOI: 10.1128/JB.186.23.7959-7970.2004.

Raymond B, Bonsall MB. Cooperation and the evolutionary ecology of bacterial virulence: the *Bacillus cereus* group as a novel study system. *Bioessays* 2013, DOI: 10.1002/bies.201300028.

Raymond B, Wyres KL, Sheppard SK *et al*. Environmental factors determining the epidemiology and population genetic structure of the *Bacillus cereus* group in the field. *PLoS Pathog* 2010, DOI: 10.1371/journal.ppat.1000905.

Reis ALS, Montanhini MTM, Bittencourt JVM *et al*. Gene detection and toxin production evaluation of hemolysin BL of *Bacillus cereus* isolated from milk and dairy products marketed in Brazil. *Braz J Microbiol* 2013;**44**:1195–8.

Robinson O, Dylus D, Dessimoz C. Phylo.io: Interactive viewing and comparison of large phylogenetic trees on the web. *Mol Biol Evol* 2016, DOI: 10.1093/molbev/msw080.

Salustiano VC, Andrade NJ, Soares NFF *et al*. Contamination of milk with *Bacillus cereus* by post-pasteurization surface exposure as evaluated by automated ribotyping. *Food Control* 2009, DOI: 10.1016/j.foodcont.2008.07.004.

Sasahara T, Hayashi S, Morisawa Y *et al*. *Bacillus cereus* bacteremia outbreak due to contaminated hospital linens. *Eur J Clin Microbiol* 2011, DOI: 10.1007/s10096-010-1072-2.

Schaefer G, Campbell W, Jenks J *et al*. Persistent *Bacillus cereus* bacteremia in 3 persons who inject drugs, San Diego, California, USA. *Emerg Infect Dis* 2016, DOI: 10.3201/eid2209.150647.

Schmid D, Rademacher C, Kanitz EE *et al*. Elucidation of enterotoxigenic *Bacillus cereus* outbreaks in Austria by complementary epidemiological and microbiological investigations, 2013. *Int J Food Microbiol* 2016, DOI: 10.1016/j.ijfoodmicro.2016.05.011.

Shaheen R, Svensson B, Andersson MA *et al*. Persistence strategies of *Bacillus cereus* spores isolated from dairy silo tanks. *Food Microbiol* 2010, DOI: 10.1016/j.fm.2009.11.004.

Sharif FA, Alaeddinoglu NG. A rapid and simple method for staining of the crystal protein of *Bacillus thuringiensis*. *J Ind Microbiol* 1988;**3**:227–9.

Sheppard SK, Jolley KA, Maiden MCJ. A gene-by-gene approach to bacterial population genomics: whole genome MLST of *Campylobacter*. *Genes* 2012, DOI: 10.3390/genes3020261.

Simonsen M, Mailund T, Pedersen CNS. Rapid neighbour-joining. In *Proceedings of the 8th Workshop in Algorithms in Bioinformatics (WABI)*, LNBI 5251, 113–22, Springer, 2008, DOI:10.1007/978-3-540-87361-7_10.

Sineva EV, Andreeva ZI, Shadrin AM *et al*. Expression of *Bacillus cereus* hemolysin II in *Bacillus subtilis* renders the bacteria pathogenic for the crustacean Daphnia magna. *FEMS Microbiol Lett* 2009, DOI: 10.1111/j.1574-6968.2009.01742.x.

Stadhouders J. The enumeration of spores and vegetative cells of *Bacillus cereus*. *Bull Int Dairy Fed* 1992;**275**:15–8.

Svensson B, Monthan A, Shaheen R *et al*. Occurrence of emetic toxin producing *Bacillus cereus* in the dairy production chain. *Int Dairy J* 2006, DOI: 10.1016/j.idairyj.2005.07.002.

Thorsen L, Hansen BM, Nielsen KF *et al*. Characterization of emetic *Bacillus weihenstephanensis*, a new cereulide-producing bacterium. *Appl Environ Microb* 2006, DOI: 10.1128/AEM.00170-06.

Tran SL, Guillemet E, Lereclus D *et al*. Iron regulates *Bacillus thuringiensis* haemolysin hlyII gene expression during insect infection. *J Invertebr Pathol* 2013, DOI: 10.1016/j.jip.2013.04.001.

Tran SL, Guillemet E, Ngo-Camus M *et al*. Hemolysin II is a *Bacillus cereus* virulence factor that induces apoptosis of macrophages. *Cell Microbiol* 2011, DOI: 10.1111/j.1462-5822.2010.01522.x.

Vidal-Martins AMC, Bürger KP, Gonçalves ACS *et al*. Evaluation of consumption of milk and its informal derivates and the knowledge of population about the public health problems in a municipal district in São Paulo State, Brazil. *Bol Ind Anim* 2013, DOI: 10.17523/bia.v70n3p221.

Vidal-Martins AMC, Rossi OD, Jr, Bürger KP *et al*. Enterotoxigenic *Bacillus cereus* in different phases of the UHT milk process. *Rev Bras Ciênc Vet* 2006;**13**:32–8.

Vissers MMM, Te Giffel MC, Driehuis F *et al*. Predictive modeling of *Bacillus cereus* spores in farm tank milk during grazing and housing periods. *J Dairy Sci* 2007, DOI: 10.3168/jds.S0022-0302(07)72629-2.

Von Stetten F, Mayr R, Scherer S. Climatic influence on mesophilic *Bacillus cereus* and psychtolerant *Bacillus weihen-*

stephanensis populations in tropical, temperate and alpine soil. *Environ Microbiol* 1999;**1**:503–15.

Wong HC, Chang MH, Fan JY. Incidence and characterization of *Bacillus cereus* isolates contaminating dairy products. *Appl Environ Microb* 1988;**54**:699–702.

Yang Y, Yu X, Zhan L *et al*. Multilocus sequence type profiles of *Bacillus cereus* isolates from infant formula in China. *Food Microbiol* 2017, DOI: 10.1016/j.fm.2016.09.007.

Zhou G, Bester K, Liao B *et al*. Characterization of three *Bacillus cereus* strains involved in a major outbreak of food poisoning after consumption of fermented black beans (douchi) in Yunan, China. *Foodborne Pathog Dis* 2014, DOI: 10.1089/fpd.2014.1768.