NTNU – Trondheim
Norwegian University of
Science and Technology

Multilocus Sequence Typing of close Neighbours to *Bacillus anthracis* isolated from Soil Samples

Kristin Hovland

Chemical Engineering and Biotechnology
Submission date: Januar 2014
Supervisor: Olav Vadstein, IBT
Co-supervisor: Jaran Strand Olsen, FFI

Norwegian University of Science and Technology
Department of Biotechnology

Acknowledgements

This thesis was performed as the final assignment for the Master of Science study in Chemical Engineering and Biotechnology at the Norwegian University of Science and Technology (NTNU). I want to thank Cand. scient. Jaran Strand Olsen and Professor Olav Vadstein for enabling me to write my final assignment at the Norwegian Defence and Research Establishment (FFI).

As my supervisor at FFI, I am sincerely grateful to Cand. scient Jaran Strand Olsen. His scientific knowledge and experience have been indispensable. Further would I like to thank Tone Aarskaug and Ingjerd Thrane for help and support at the laboratory. To the persons working at the Division for Protection: thank you for welcoming me to the FFI environment.

And at last, to my family and friends: Thank you for all support, patience and kindness through the course of my degree at NTNU.

Abstract

The *Bacillus cereus* group comprises of *B. anthracis*, *B. cereus*, *B. thuringiensis*, *B. mycoides*, *B. pseudomycoides*, *B. weihenstephanensis* and *B. cytotoxicus*. *B. anthracis* is the causative agent of anthrax, whilst *B. cereus* and *B. thuringiensis* (*B. anthracis* close neighbours) are mostly associated as food and insect pathogens respectively. During the last years, strains of *B. cereus* and *B. thuringiensis* have been the cause of serious infections in both humans and animals, and both phylogenetic and phenotypic characteristics associated with *B. anthracis* have been identified in some of these close neighbours. This has led to an increased interest for the different species and strains of *B. anthracis* close neighbours. The aim of this study was to evaluate the genetic variation among strains of the *Bacillus cereus* group, with a particular focus on isolates closely related to *B. anthracis*.

Soil samples collected in Etosha National Park, Namibia, near a carcass of a *B. anthracis* infected Zebra were used to isolate 169 *B. cereus* group members using a selective growth medium. Isolation was followed by DNA extraction and the DNA was used in real time qPCR using six *B. anthracis* specific markers. One of the PCR primer pairs amplified a VNTR (variable number of tandem repeat) region and 52 *B. cereus* group isolates were selected for MLST (multi locus sequence typing) due to having a VNTR region identical to *B. anthracis*. In addition, 8 isolates were selected for MLST due to the generation of PCR products when using *B. anthracis* plasmid markers.

Phylogenetic analyses were performed using an 11 loci MLST scheme (*adk, ccpA, ftsA, glpF, glpT, panC, pta, pycA, pyre, recF* and *sucC*) on 125 *B. cereus* group members (46 isolates from this study, 67 from Helgason *et al.* (2004), 5 in-house isolates and 7 from MLST Oslo), which resulted in 96 STs (sequences types). 39 *B. cereus* group members analysed during this study clustered in clade I. The two isolates (FFIBCgr36 and FFIBCgr46) that clustered closest to *B. anthracis* revealed only 10 and 12 point mutations that differentiating them from *B. anthracis*.

Sammendrag

Bacillus cereus gruppen inneholder den anthrax frembringende bakterien *B. anthracis*, de to bakteriene *B. cereus* and *B. thuringiensis* som ofte assosieres med matforgiftning og innsektsmidler, i tillegg til *B. mycoides*, *B. pseudomycoides*, *B. weihenstephanensis* and *B. cytotoxicus*. De senere årene har varianter av *B. cereus* og *B. thuringiensis* forårsaket svært alvorlige infeksjoner i både mennesker og dyr, og symptomene er blitt sammenlignet med anthrax. På grunn av de anthrax-lignende infeksjonene er disse isolatene blitt studert nærmere, og både fenotypiske og fylogenetiske karaktertrekk som minner om *B. anthracis* er blitt observert. Disse observasjonene har ført til en økt interesse for de forskjellige artene og variantene av *B. anthracis* sine "nære naboer". Hensikten med dette studiet var å se på den genetiske variasjonen i artene tilhørende *B. cereus*-gruppen, med hovedvekt på de "nære naboene" til *B. anthracis*.

Isolater fra *B. cereus*-gruppen ble isolert fra jordprøver hentet fra et område rundt et *B. anthracis* infisert sebrakadaver. Jordprøvene ble hentet i Etosha Nasjonalpark, Namibia, og dyrket på agar som er selektivt for *B. cereus*-gruppen. DNAet ble ekstrahert og "real time" qPCR med seks markører spesifikke for *B. anthracis* ble gjennomført for å velge ut "nære naboer" til *B. anthracis*. En av markørene amplifiserte en VNTR-region (variable number of tandem repeat), og basert på en VNTR-region lik *B. anthracis* ble 52 av isolatene valgt ut til MLST-analyse (multi locus sequence typing). I tillegg ble 8 ekstra isolater, med DNA som amplifiserte ved bruk av *B. anthracis* plasmidmarkører, valgt til MLST-analyse.

De fylogenetiske analysene ble gjennomført ved bruk av et MLST skjema med 11 loci (*adk, ccpA, ftsA, glpF, glpT, panC, pta, pycA, pyre, recF* and *sucC*) på totalt 125 medlemmer av *B. cereus*-gruppen (46 fra dette studiet, 67 fra Helgason *et al.* (2004), 5 "in-house" isolater og 7 fra MLST Oslo). Analysene resulterte i 96 STs (sekvenstyper). 39 av isolatene analysert i dette studiet grupperte seg i gruppe I sammen med *B. anthracis*, og to av disse isolatene (FFIBCgr36 og FFIBCgr46) grupperte seg nærmest *B. anthracis* og skilte seg fra *B. anthracis* med bare 10 og 12 punktmutasjoner hver.

Table of contents

1 Introduction

Bacillus anthracis is a member of the *Bacillus cereus* group and the etiological species of anthrax. New attention was shown this species in the fall of 2001 as spores from the Ames strain of *B. anthracis* had been sent via the US Postal Service to at least two news media offices and two democratic U.S Senators (Rasko *et al.* 2011). Eleven cases of pulmonary anthrax were reported due to exposure to these letters, and five people died as a result of the infection (Imperiale and Casadevall 2011). This was a serious reminder of the potential threat biological agents may cause.

1.1 The *Bacillus* genus and the *Bacillus cereus* group

Members of the genus *Bacillus* are endospore-forming, aerobic or facultatively anaerobic gram positive bacteria. They are diverse and adapted to an array of different environments such as fresh water, marine water, soil, plants, animals and air (Turnbull 1996; Madigan *et al.* 2009). Several *Bacillus* species are used in industrial production of riboflavin, streptadivin, β-lactamase, as well as insect and nematode toxins (Maughan and Van der Auwea 2011). However, difficulties to eliminate the spores of *Bacillus* species due to their resistance to radiation, heat, disinfectants and desiccation make them a frequent cause of contamination in medical and pharmaceutical material, and they are well known as spoilage organisms in the food industry (Turnbull 1996). Species of *Bacillus* are divided into five or six groups based on 16S rRNA phylogeny or phenotypic features respectively (Økstad and Kolstø 2011). The *Bacillus cereus* group (also known as *Bacillus cereus sensu lato (Maughan and Van der Auwea 2011)*) comprise of seven species; *B. cereus*, *B. anthracis*, *B. thuringiensis*, *B. mycoides*, *B. pseudomycoides*, *B. weihenstephanensis* and *B. cytotoxicus* (Soufiane and Côté 2013).

B. cereus, *B. anthracis* and *B. thuringiensis* are the three most frequently studied members of the *B. cereus* group. This is due to their ability to produce various toxins, and the high degree of genetic similarities between these species (Økstad and Kolstø 2011).It is also an on-going debate whether the three species should be considered three variants of the same species (Helgason *et al.* 2000).

1.2 *B. cereus, B. anthracis* and *B. thuringiensis*

Bacillus cereus

B. cereus is an environmental bacterium isolated worldwide. It is often associated as a food-borne pathogen in potatoes, vegetables, milk, herbs and spices (Ceuppens *et al.* 2013) and is frequently associated with gastrointestinal infections in humans, causing an emetic and a diarrheal form of infection. Plant based foods are the main source of infections, but *B. cereus* has also been isolated from meat, eggs and dairy products. *B. cereus* is an opportunist observed in patients with impaired immune system (Granum and Lund 1997; Bottone 2010). Nausea, vomiting and minor abdominal cramps characterize the emetic form of the infection and can be observed 1-5 hours after contamination. The diarrheal form of the infection is primarily abdominal cramps and diarrhoea, with onset 8-16 hours after intake of contaminated food (Turnbull 1996; Todar 2006). Intestinal infections by *B. cereus* are often caused by ingestion of heat treated food. This is due to the survival of *B. cereus* spores after heating and cooling. Most competing bacterial flora will be eliminated after such treatment and *B. cereus* will be the dominant bacteria when optimal growth conditions are restored (Granum and Lund 1997). Precaution, like good hygiene, is used to avoid food poisoning by *B. cereus* (Turnbull 1996).

Bacillus anthracis

B. anthracis is a pathogenic bacterium, a natural inhabitant in soil and able to infect nearly all mammals and cause anthrax (Klee *et al.* 2006). Anthrax outbreaks are most common in central and south America, sub-Saharan Africa, central and south-western Asia, and southern and eastern Europe (Shadomy 2012). Herbivores are the animals most susceptible to anthrax, but humans can also be infected through

contact with/ingestion of infected animals or animal products. Spread of *B. anthracis* spores by aerosols is a less common route of infections, but considered to be the preferred way to spread the bacteria if used as a biological threat agent (Inglesby *et al.* 2002).

Three types of anthrax are observed in humans, and the route of infection dictate the type of illness. Cutaneous anthrax is caused by infection of damaged skin and accounts for 95 % of all *B. anthracis* infections. Pulmonary and intestinal anthrax are caused by inhalation and ingestion of *B. anthracis* spores respectively (Turnbull 1996). Incubation time for cutaneous anthrax is 2-3 days. If observed early, treatment with antibiotics, are in most cases sufficient to cure cutaneous anthrax. However, if the infection is not treated, approximately 20 % of the cases may progress to septicaemia, shock and, in worst case, death. Cutaneous anthrax is less dangerous than both pulmonary and intestinal anthrax because the infected area is easily noticeable. Pulmonary and intestinal anthrax infections are often observed too late for treatment to be effective (Turnbull 1996).

The vegetative form of *B. anthracis* is mostly found in environments with low oxygen contents, i.e. the *B. anthracis* is found within the infected host. However, studies indicate vegetative cells of *B. anthracis* to be capable of survival for longer periods under certain conditions (Lindeque and Turnbull 1994; Saile and Koehler 2006). If *B. anthracis* is exposed to free oxygen, sporulation will often start immediately. Spores will be located in the soil until they are taken up by another host. This may take from hours to decades. Inside its new host the spores will go back to vegetative state (Epizootics 2008).

Bacillus thuringiensis

B. thuringiensis is an insect pathogen used as an insecticide to control some garden and crop pests (Todar 2006). The production of different Cry proteins are responsible for the pathogenicity of the bacteria (Schnepf *et al.* 1998). The diverse Cry proteins responsible for the different protoxins are located on plasmids and these plasmids

are the only known difference between *B. thuringiensis* and *B. cereus* (Schnepf *et al.* 1998). A potentially loss of plasmids will make it very difficult to distinguish *B. thuringiensis* from *B. cereus* (Helgason *et al.* 2000). To avoid loss of plasmids, *B. thuringiensis* isolates are often cultivated at low temperatures (Tourasse *et al.* 2006).

1.3 Anthrax like diseases and the identification of the different *B. cereus* group members

The potentially deadly outcomes of *B. anthracis* infections make scientists look for markers that can be used to quickly recognize and distinguish this species from its close relatives. *B. anthracis* harbour the pXO1 (181 kilo bases (kb)) and pXO2 (93.5 kb) plasmids, both essential for the pathogenicity (Okinaka *et al.* 1999). Plasmid pXO1 encodes the protective antigen (*pag*), lethal factor (*lef*) and edema factor (*cya*). Plasmid pXO2 encodes three capsule genes (*capA, capB and capC*) that together encode a poly-γ-D-glutamic acid capsule that protects the cell from the phagocytosis of the host's immune system (Dai *et al.* 1995; Turnbull 1996; Koehler 1999). Polymerase chain reaction (PCR) (section 1.4) is a method considered important to distinguish close neighbours on a genetic level (Klee *et al.* 2006). Primer pairs targeting different regions on the *B. anthracis* plasmids are frequently used to identify *B. anthracis* (Ezzel and Welkos 1999; Priest *et al.* 2004; Unknown 2009; Matero *et al.* 2011). A techoic acid ABC transporter that is assumed to be a highly conserved region in *B. anthracis* is also used as a genetic marker to identify *B. anthracis*. The primer pair is called BA5510 and the transporter is located on the *B. anthracis* chromosome (Olsen *et al.* 2007).

From 2001 to 2004 nine chimpanzees and one gorilla from two national parks (Côte d'Ivoire and Cameroon) in Africa died from anthrax-like diseases. The causes of infections were concluded not to be classical *B. anthracis* isolates. Classical phonotypical *B. cereus* traits were observed, such as motility, resistance to gamma phage, and resistance to penicillin G. In contrast, all the isolates contained plasmids with similar characteristics to the pXO1 and pXO2 plasmids found in *B. anthracis*

(Klee *et al.* 2006). The genome to one of the chimpanzees, isolates CI, has been sequenced and shows chromosomal similarities to *B. cereus* and similarities to the two *B. anthracis* plasmids (Klee *et al.* 2010). *B. cereus* isolate G9241 is another isolate that has been sequenced due to an anthrax-like disease. This isolate, which led to severe pneumonia in a welder, has a plasmid homolog to pXO1 (Okinaka *et al.* 2006) and a pXO2 analogue (pBC218) that encodes a novel polysaccharide capsule (Hoffmaster *et al.* 2004; Okinaka *et al.* 2006; Wilson *et al.* 2011). The pXO1 homolog in *B. cereus* G9241 is called pBCXO1 and shows >99 % amino acid identity to the pXO1 plasmid found in *B. anthracis (Hoffmaster et al. 2004)*.

B. cereus group member ATCC10987 was isolated from a chees spoilage study in Canada in 1930. The isolate did not lead to an anthrax-like disease, but has been used to study the genetic relationship to *B. anthracis.* This isolate have shown to contain possible virulent factors as well as a plasmid similar to the pXO1 plasmid found in *B. anthracis.* Chromosomal studies also reveal this isolate to have a higher degree of conserved genes to *B. anthracis* Ames than to other dairy strains (Rasko *et al.* 2004). *B. cereus* group member 200031002 is another non-*B. anthracis* isolate showing a high degree of chromosomal similarities to *B. anthracis*. This isolate were collected in 1960 from the Belgian Congo, and is thus far the isolate most similar to *B. anthracis* based on chromosomal markers *(Marston et al. 2006)*.

B. cereus strains holding plasmids analogue to the pXO1 and pXO2 plasmids makes the evolutionary history of the *B. cereus* group fascinating. For instance a nonsense mutation found in the *plcR* gene, a pleiotropic regulator, present in all *B. anthracis* strains is not observed in any other members of the *B. cereus* group (Mignot *et al.* 2001; Okinaka *et al.* 2006). This regulator affects virulence factors and genes encoding degradative enzymes, cell surface proteins and haemolytic and non-haemolytic enterotoxins (DeBell 2002). Genes regulated by the *plcR* gene are still present in *B. anthracis*, but they have been silenced due to the nonsense mutation (Mignot *et al.* 2001). This may explain the lack of haemolytic activity and some degradative enzymes in *B. anthracis* (DeBell 2002; Easterday *et al.* 2005). Studies have shown *B. anthracis* to harbour different DNA segments of variable numbers of

tandem repeats (VNTRs). Some of these VNTRs are present in other members of the *B. cereus* group and can be used as markers to identify strains that are closely related to *B. anthracis* (Valjevac *et al.* 2005). One marker tested by Valjevac *et al.* (2005), the Bcms17 primer pair, is alone assumed to be adequate to identify strains closely related to *B. anthracis*. The Bcms17 primer pair identifies a repeat of sixteen base pairs that is repeated ten times in *B. anthracis* (Valjevac *et al.* 2005).

Because of the identification of several *B. cereus* group members closely related to *B. anthracis*, it has become important to get more information on the different species and species combination within the *B. cereus* group.

1.4 Genotyping

Genotyping is widely used to evaluate associations between genetic variations and different phenotypes. There are many different techniques available and the different assays vary from analysing single genomic regions to larger genome and structure variations (Huang *et al.* 2011).

A variety of different genotyping methods have been used on members of the *B. cereus* group; multi locus enzyme electrophoresis (MLEE) (Helgason *et al.* 1998), pulsed-field gel electrophoresis (PFGE) (Harrell *et al.* 1995), restriction fragment length polymorphism (RFLP) (Jensen *et al.* 2005), amplified fragment length polymorphism (AFLP) (Hill *et al.* 2004), intergenic spacer regions (ISR) (Harrell *et al.* 1995) and multi locus sequence typing (MLST).

Multi Locus Sequence Typing

MLST is a method characterizing isolates of bacterial species based on several genes. This technique was first introduced in 1998 by Martin C. J. Maiden (Maiden *et al.* 1998) and is compared to MLEE. The work by MCJ Maiden lists two advantages of MLST over MLEE:

> First, far more variation can be detected, resulting in many more alleles per locus than are obtained with MLEE. Second, sequence data can be compared readily between laboratories, such that a typing method based on the sequences of gene fragments from a number of different housekeeping loci [multilocus sequence typing (MLST)] is fully portable and data stored in a single expanding central multilocus sequence database can be interrogated electronically via the Internet to produce a powerful resource for global epidemiology.
>
> (Maiden *et al.* 1998)

MLST uses the nucleotide sequence of a number of housekeeping genes (usually six or seven) localized on the bacterial chromosome. Genes or gene segments of interest (plural: loci) are amplified and the PCR products are sequenced (Maiden *et al.* 1998). Different nucleotide sequences obtained for the same locus correspond to distinct alleles and are labelled with a unique allele number. An allelic profile and a sequence type (ST) are assigned to every isolate and the allelic profile constitutes of the allelic numbers of each locus. Allele numbers in an allelic profile are listed in the order specified by the MLST scheme used. An allelic profile consists of one particular set of allelic numbers and every allelic profile is assign a unique ST number (Aanonsen and Spratt 2005). Several MLST schemes have been constructed in order to genotype the *B. cereus* group (Helgason *et al.* 2004; Ko *et al.* 2004; Priest *et al.* 2004).

Quantitative Polymerase Chain Reaction

Quantitative PCR (qPCR) is a method used to detect the presence of specific genes or DNA segments within an isolate. The method is based on DNA polymerase and its ability to synthesize a complementary strand of DNA from a template strand. In a three step temperature cycle (denaturation, annealing and elongation), the interaction of the primer pair and the polymerase enzyme results in synthesis of new DNA, complementary to the template DNA (Maier *et al.* 2009).

When using endpoint qPCR the presence of PCR products have to be verified by gel electrophoresis or an equivalent technique. By the use of RT-qPCR the amount of PCR product are measured digital after each temperature cycle. If a high number of temperature cycles are necessary to obtain amplification, it may be beneficial to use probes to eliminate false-positive results when using RT-qPCR. Probes are also recommended to differentiate between PCR products in a multiplex PCR reaction.

Dual-labelled probes are short sequences of single stranded oligonucleotides with a fluorescent reporter at one end, and a quencher at the other. The probe binds to its complementary strand during the annealing step (Figure 1.1). The proximity between the fluorescent reporter and the quencher prevents the reporter from fluorescing. During elongation polymerase will reach the probe and release the reporter. The fluorescent can then be measured and the intensity will be proportional to the amount of specific PCR products (QIAGEN 2002).

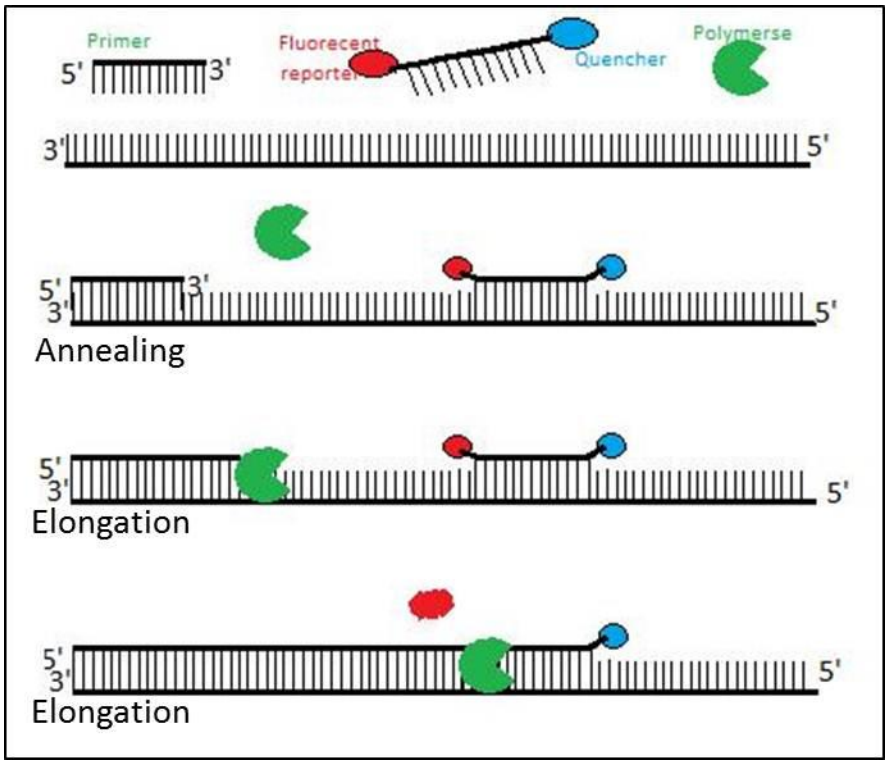


Figure 1.1: The mechanism of dual-labelled probes. Free reporters will fluorescent, and the intensity is proportional to the amount of specific PCR product.

1.5 Bioinformatics and phylogenetic analyses

Bioinformatics can be defined as the combination of biological and computer science, where the objective is to gather, store, analyse and integrate biological and genetic information (BioPlanet 2013). Many different bioinformatics tools are available, both for free online (e.g. Staden Package and MEGA) and to purchase (e.g. BioNumerics). Some software are specialised on one task like sequence editing (Gap4) or sequence assembly (ClustalW), while other software packages include tools for several operations from sequence editing to phylogenetic tree construction (e.g. BioNumerics, MEGA4 and Staden Package).

A phylogenetic tree, or dendrogram, is a visual representation of the genetic variation among different bacterial strains within one or several species. Generally there are two techniques to calculate a phylogenetic tree; 1) the distance- matrix method and 2) the discrete data method. The distance-matrix method is the simplest and fastest method. For all pairwise alignments a distance is calculated. The distance can be looked upon as the percentage sequence differences, and is used to assembly the phylogenetic tree. This technique is also called a clustering method or algorithmic method, and includes unweighted pair group method with arithmetic mean (UPGMA), neighbour-joining (NJ) and Fitch-Margoliash (F-M) (Baldauf 2003).

UPGMA assume the evolutionary rate of change to be constant along the branches and results in a rooted tree. The distance between sequences is approximately ultra-metric (the distance from root-to-tip is equal for all lineages) (Mount 2004).

The NJ method is based on a variable evolutionary rate of change along the branches and produces an unrooted tree (the tree can be rooted if desired). Trees predicted by the NJ method are additive, meaning that the distance between species on the tree is approximately the same distance as between the same species in the original matrix. The F-M method is similar to the NJ method except from the algorithm used to compare sequences (Mount 2004).

The discrete data method includes parsimony, maximum likelihood and Bayesian methods. This technique is more time consuming than the distance-matrix, but in addition it provides more information on every sequence in the alignment. A hypothesis for every column in the alignment is created making it possible to trace the evolution of the sequence (Baldauf 2003).

If the dendrogram is the information of interest, the clustering method is the recommended technique. In most cases the two techniques will provide the same grouping (Baldauf 2003). Bootstrapping is a common test to check the strength of a dendrogram topology. Bootstrapping indicates whether the topology of the tree is randomly selected among many nearly equal alternatives. Several subsamples of the dataset are used to create smaller dendrograms and for each subsample the frequency of different clades are measured. Bootstrap values above 70 % indicate trustworthy groupings (Hillis and Bull 1993).

Phylogenetic analyses using MLST data

Sequence information collected for MLST can be analysed by two basic approaches. The phylogenetic relationship can be determined based on allelic profiles, or the relationship can be established using the nucleotide sequences directly. The UPGMA method using the allelic profiles is recommended to analyse sequences obtained from populations where variations are mainly developed by recombination. For clonal organisms, where recombination is rare or absent, the use of nucleotide sequences directly are recommended (Maiden 2006).

Phylogenetic analyses of the B. cereus group

The *B. cereus* group is characterized mostly as a clonal population. This mean genetic variation is mainly based on vertical genetic transfers; they are in linkage disequilibrium (Helgason *et al.* 2004; Priest *et al.* 2004; Maiden 2006). Nonclonal populations on the other hand, comprise of horizontal gene transfers (HGT) that lead to randomized genetic diversity where evolutionary history of each locus may be different from one another. Within the same bacterial genus, groups of clonal

population have been observed alongside nonclonal populations (Maiden 2006). *B. anthracis* has been found to be the most clonal of the *B. cereus* group members (Keim *et al.* 2000). Studies indicate the *B. cereus* group to be mostly clonal, however HGT (recombination) have been suggested (Helgason *et al.* 2004; Priest *et al.* 2004; Barker *et al.* 2005).

For the different MLST schemes available for the *B. cereus* group, the results indicate the group members to be clustered into three main clades. Priest *et al.* (2004) analysed strains of *B. anthracis*, *B. cereus*, *B. thuringiensis*, *B. mycoides* and *B. weihenstephanensis* with a seven loci MLST scheme. A dendrogram based on maximum likelihood on the concatenated sequences revealed three main clades. Clade I included *B. anthracis*, mostly *B. cereus* and some *B. thuringiensis* strains. Clade II comprised mostly strains of *B. thuringiensis* and some *B. cereus* strains. Clade III included the remaining species (Priest *et al.* 2004). Barker *et al.* (2005) studied eight strains of *B. cereus* gathered from bacteraemia and soft tissue infections and performed a study using the same MLST scheme as Priest *et al.* (2004). A dendrogram constructed using the NJ method placed the new strains in clade I and clade II (Barker *et al.* 2005). Sorokin *et al.* (2006) characterized 134 *B. cereus* group members based on a six loci MLST scheme. Dendrograms based on sequences of individual locus and concatenated sequences using the NJ method, indicated that the isolates could be divided into three main clusters with a grouping similar to the clades presented by Priest *et al.* (2004) (Sorokin *et al.* 2006).

A general pattern based on studies on population structure indicates most clinical isolates to be clustered in clade I together with *B. anthracis* while clade II and III mostly comprise soil and dairy isolates (Tourasse *et al.* 2006). With no pathogenic isolates observed in clade III, pathogenic isolates seem to be evenly distributed among clade I and II, and recombination events seem to be less frequent for clade I and II compared to clade III (Didelot *et al.* 2009).

1.6 Aim of project

The aim of this master project was to evaluate the genetic variation among strains of the *Bacillus cereus* group, with a particular focus on isolates closely related to *B. anthracis*. The isolates were isolated from soil samples collected in Namibia.

The time frame of the project was 21 weeks, and the aim was addressed through the following approach:

- Isolates were screened for six markers known to be present either in the chromosome or on the plasmids in *B. anthracis*. The isolates were screened using RT-qPCR. One of the markers was a VNTR analysis.
- Two MLST schemes including a total of eleven loci were performed on selected isolates showing a VNTR similar to *B. anthracis* and/or generating a PCR product by using any of the three primer pairs BA5510, BA-*pag* or BA-*capB*
- BioNumerics, a bioinformatics tool, was used on the selected isolates to genotype and study the genetic relatedness of these isolates to *B. anthracis* and *B. anthracis* close neighbours.

2 Materials and Methods

2.1 Sampling

169 isolates of the *B. cereus* group were isolated from soil samples gathered from Etosha National Park in Namibia. The soil samples were collected in 2012 near a carcass of a *B. anthracis* infected zebra. Of the 169 isolates, 98 isolates were screened and analysed during this study, while the remaining 71 isolates were screened during a project conducted in spring 2013 (Hovland 2013).

The process to genotype and obtain dendrograms comprising *B. cereus* group members included several steps and different methods/techniques (Figure 2.1).

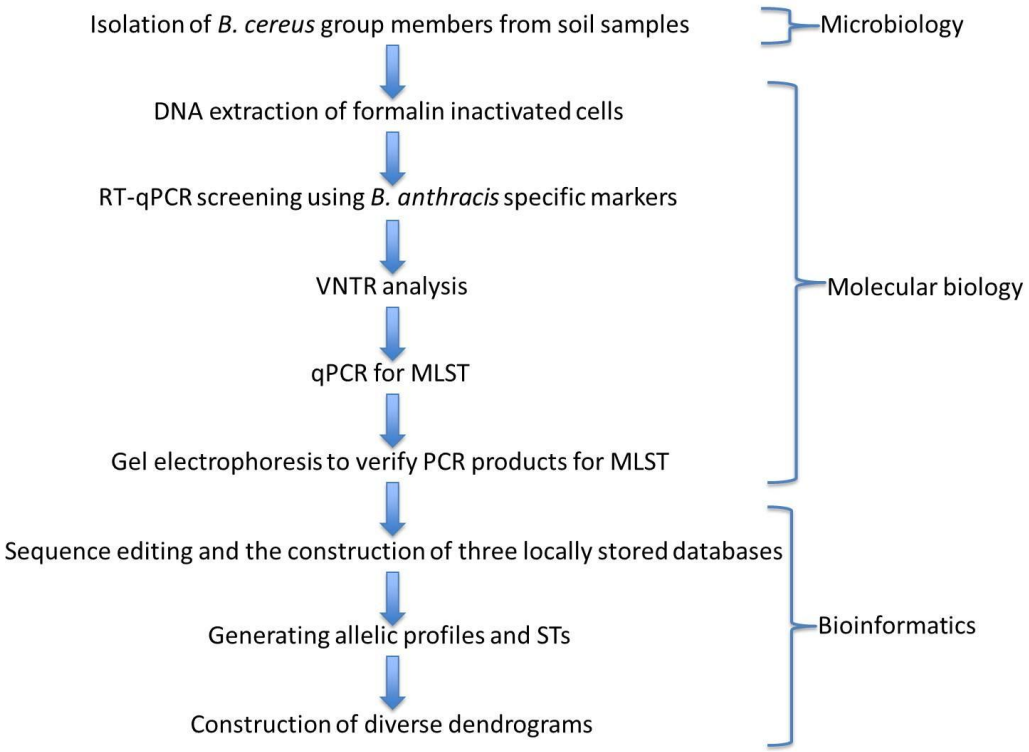


Figure 2.1: Flow chart covering the different steps performed to obtain dendrograms visualizing the genetic relationship of *B. cereus* group members.

2.2 Isolation, cultivation, DNA extraction and storage of bacteria

One gram of soil was shaken in 10 mL distilled water (dH_2O) for one minute to transfer the bacteria into the liquid. After five minutes of sedimentation, 100 µL of liquid was spread onto BACARA agar plates (AES Chemunex, France), a selective medium for *B. cereus* group members (Chemunex 2012). The agar plates were incubated for 18-24 hours at 32°C. Resulting colonies had a pink/orange core surrounded by a larger (2-4 mm) opaque/white zone of precipitate (Figure 2.2). Colonies of *B. anthracis* lacked the precipition zone due to absent lecithinase activity, i.e. the ability to break down lecithin to phosphorylcholine and diglycerides, both being insoluble (Chu 1949).

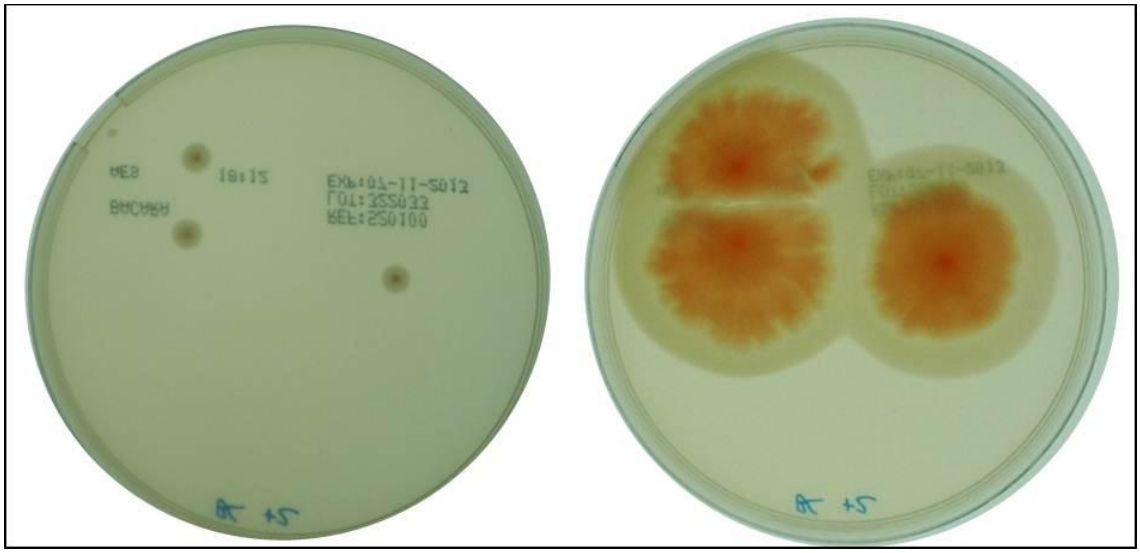


Figure 2.2: Colonies of a *B. cereus* group member on BACARA agar plates (AES Chemunex, France) incubated for 18-24 hours (left) and four days (right) at 30 °C.

Colonies from the BACARA agar plates were streaked on sheep blood agar (Oxoid, Germany) and incubated at room temperature for 18-24 hours. Single colonies were evaluated for haemolytic activity to rule out strains of *B. anthracis*. Haemolytic activity was observed as colonies having a clearing zone indicating the ability to lyse the red blood cells in the agar. Colonies used for molecular analysis were inactivated in

formalin (10 %) for 48 hours at 4 °C. Personnel working at the departments BSL3 (Biological safety level 3) laboratory performed these procedures.

Inactivated isolates were transported from the BSL3 laboratory to the BSL2 laboratory where DNA was extracted using QIAamp DNA Mini Kit (QIAGEN, Netherlands). The samples were centrifuged at 7000 rpm for 3 minutes before removal of the formalin by pipetting. The pellet was carefully suspended to avoid damage to the cell in Phosphate-Buffered Saline (PBS) (1 mL) followed by centrifugation at 7000 rpm for 3 minutes. Washing with PBS was repeated twice. Between the two washing steps and after the last, the PBS was removed by pipetting. The pellet was then suspended in 180 µL TE buffer (20 mM Tris-HCL, 2 mM EDTA with 20 % mg/mL lysosome and 1.2 % TritonX) and the steps listed in the protocol (Appendix A) were carried out. For the last step 150 µL elution solution was used instead of the original 200 µL to increase the DNA concentration. Extracted DNA was stored at -20 °C until used as qPCR template.

After screening the bacterial DNA using RT-qPCR (section 2.3) viable strains not containing any of the two plasmids pXO1 and pXO2, or the ABC transporter located on the *B. anthracis* chromosome were transported out of BSL3 laboratory. The viable strains were grown on Tryptic soy agar (Merck KGaA, Germany) at 30 °C for 18-24 hours and the resulting colonies were then transferred to Tryptic soy broth (Merck KGaA, Germany) containing 20 % glycerol for long-term storage at -72 °C.

2.3 Screening of bacterial DNA for DNA segments present in *B. anthracis*

DNA isolated from the 98 *B. cereus* group members was screened with RT-qPCR for six gene segments in three parallels. RT-qPCR analyses were performed using the LightCycler®480 (Roche, Switzerland) and the LightCycler®480 SYBR Green I master kit (SYBR Green) (Roche, Switzerland). The *plcR* primer pair was used

together with the LightCycler®480 Probe master kit (Probe master) (Roche, Switzerland).

The RT-qPCR analyses were conducted in a 20 µL reaction mixture (Table 2.1) that was similar for all, except for *plcR*, of the primer pairs. The reaction mixture used together with the *plcR* primer pair included probes (Table 2.2). Data related to the primer pairs used and RT-qPCR conditions used to screen the unknown insolates in this study are listed in Table 2.3 and 2.4 respectively.

Isolates showing late amplification (Ct-value>30) when using the BA-*pag*, BA-*capB* or BA5510 primer pairs (Table 2.4) were analysed further using two additional primer pairs to confirm initial PCR amplification. BA-*lef* targets one of the other toxins present on pXO1. *BA-capB(2)* targets another region displayed inside the same gene already screened for by the BA-*capB* primer pair.

Table 2.1: Reaction mixture used for RT-qPCR analyses using the primer pair BA-*pag*, BA-*capB*, BA5510, Bcms17, *adk*, BA-*lef* and BA-*capB(2)*.

Reagent	**Quantity [µL]**
Enzyme mix (SYBER Green)[1]	10
Primer pair mix (5 µM)	4
dH_2O	4
DNA template	2

[1] SYBR Green includes nucleotides, magnesium, enzyme and a dye.

Table 2.2: Reaction mixture used for RT-qPCR analyses using the *plcR* primer pair.

Reagent	Quantity [µL}
Enzyme mix (Probe Master)[2]	10
Primer pair mix (5 µM)	4
BA-probe (10 µM)[3]	0.5
BC-probe (10 µM)[4]	0.5
dH_2O	3
DNA template	2

Two dual-labelled probes with dye labelling VIC and FAM were added to the RT-qPCR mixture using the *plcR* primer pair. The FAM labelled probe hybridizes with the functional *plcR* gene of *B. cereus*, and the VIC labelled probe hybridizes the point mutated and dysfunctional sequence of *B. anthracis*. The Probe Master mix was used with the *plcR* primer pair rather than the SYBR Green mix. The Probe master mix does not include a DNA dye such as SYBR Green, and the increase in PCR products is detected by the dye labelled probes. Isolates hybridized with different probes were separated based on differences in fluorescent intensity.

[2] Probe Master includes nucleotides, magnesium and enzyme
[3] *B. anthracis* probe(VIC labeled)
[4] *B. cereus* probe (FAM labeled)

Table 2.3: Primers and RT-qPCR data used in screening of the unknown *B. cereus* group member isolates.

Primer pair	Primer sequences	Primer conc.[µM]	Annealing temp.[°C]	Cycle times[5] [sec]	Reference
BA-pag	*CGGATAGCGGCGGTTAATC* *CAAATGCTATTTTAAGGGCTTCTTTT*	*1.0*	*58*	*10/10/25*	*(Matero et al. 2011)*
BA-capB	*TTGGGAACGTGTGGATGATTT* *TCAGGGCGGCAATTCATAAT*	*1.0*	*58*	*10/10/25*	*(Matero et al. 2011)*
BA5510	*CTGCATTGATAGCAATTTCTTCA* *AGAAGGCAGGTTGATACATAAACTTTCCA*	*1.0*	*58*	*10/10/25*	*(Olsen et al. 2007)*
plcR	*CCAATCAATGTCATACTATTAATTTGACAC* *ATGCAAAAGCATTATACTTGGACAAT* *BA-probe: CAAAGCGCTTATTCGTATT* *BC-probe: AAAGCGCTTCTTCGTATT*	*1.0*	*58*	*10/30/25*	*(Easterday et al. 2005)*
Bcms17	*ATTGGACAAGAAAAACAAGGTACTG* *CGCTGATCTTCCATTTGCAT*	*1.0*	*58*	*10/10/25*	*(Valjevac et al. 2005)*
adk	*CAGCTATGAAGGCTGAAACTG* *CTAAGCCTCCGATGAGAACA*	*1.0*	*57*	*10/10/30*	*(Helgason et al. 2004)*
BA-lef[6]	*GCAGATTCCTATTGAGCCAAA* *GAATCACGAATATCAATTTGTAGC*	*1.0*	*58*	*10/10/30*	*(Olsen et al. 2007)*
BA-capB(2)[6]	*GTTGCCGCAAATTTTCTACG* *ACTCGTTTTTAATCAGCCCG*	*1.0*	*58*	*10/10/30*	*(Olsen et al. 2007)*

[5] Cycle times correspond to denaturation/annealing/elongation

[6] Primer pairs used if ambiguous results were obtained when using the primer pairs BA-*pag* or BA-*capB*

Table 2.4: RT-qPCR conditions used to screen the unknown isolates.

Analysis	Temperature [°C]	Cycles	Times [sec]
Denaturation	95	1	300
Denaturation	95		10
Annealing	57-58[8]	35-45[7]	10-30[8]
Elongation	72		25-30[8]
Melting curve	55→95	1	0.2 °C/sec

2.4 Fragment analyses

Fragment lengths from the VNTR analyses were examined using the Agilent 2100 Bioanalyzer (Agilent Technologies, US) with the Agilent DNA 1000 Kit (Agilent Technologies, US). The VNTR region of the unknown isolates was amplified using the Bcms17 primer pair and the PCR products were compared to the VNTR region of *B. anthracis*.

Traditional gel electrophoresis was used to verify the presence of PCR products when endpoint qPCR was used to amplify DNA segments for MLST. The gel (15 cm X 15 cm) was made of a 150 µL mixture (Table 2.5) consisting of dH2O, 5X TBE buffer (Tris-borate-EDTA electrophoresis buffer) (Table 2.6), LE Agarose (Lonza, USA) and Ethidium bromide solution (Sigma, USA).

To each well cast in the gel a mix of 7 µL 1X loading buffer (Table 2.6) and 3 µL PCR product were added. Gel electrophoresis was run for approximately 45 minutes at 100 volts.

[7] 35 RT-qPCR cycles were used with primer pairs BA-*pag,* BA-*capB*, BA5510, Bcms17, BA-*lef* and BA-*capB(2).* For the primer pairs *adk* and *plcR,* 40 and 45 RT-qPCR cycles were conducted respectively.
[8] See Table 2.3 for details on the different primer pairs

Table 2.5: Gel mixture for electrophoresis.

Reagent	Quantity
dH_2O	120 µL
5X TBE buffer[9]	30 µL
LE Agarose	1.5 g
Ethidium bromide solution	3 drops

Table 2.6: Solutions used for gel electrophoresis.

Solution	Amount	Manufacture
5X TBE buffer[9]	**1L**	
Trizma base (99.9 % titration)	54 g	Sigma-Aldrich, USA
Boric acid	27,5 g	E. Merck, Germany
EDTA (0.5 M)	20 mL	Calbiochem, USA
10X Loading buffer	**5 mL**	
Xylene cyanol FF (1 %)	0.25 mL	Sigma-Aldrich, USA
Bromophenol blue (1 %)	1.25 mL	Bio-Rad, USA
Grycerol (87 %)	3.45 mL	Merck KGaA, Germany
dH2O	0.05 mL	Millipore, USA

[9] Tris-borate-EDTA electrophoresis buffer

2.5 Amplifying housekeeping genes for multi locus sequence typing

Two MLST assays published by Helgason *et al.* (2004) and Tourasse *et al.* (2006) were used to analyse the genetic variation of *B. cereus* group members considered closely related to *B. anthracis*. The first MLST assay uses seven housekeeping genes; *adk* (adenylate kinase), *ccpA* (catabolite control protein A), *ftsA* (cell division protein), *glpT* (glycerol-3-phosphate permease), *pyrE* (orotate phosphoribosyltransferase), *recF* (DNA replication and repair protein) and *sucC* (succinyl coenzyme A synthetase, beta subunit) (Helgason *et al.* 2004). The second MLST assay includes another combination of seven housekeeping genes; *adk*, *glpT*, *glpF* (glycerol facilitator), *panC* (pantoate-beta-alanine ligase), *pycA* (pyruvate carboxylase), *ccpA* and *pta* (phosphate acetyltransferase) (Tourasse *et al.* 2006).

MLST was performed on 60 isolates belonging to the *B. cereus* group. Selection of the isolates was made based on VNTR regions of similar length as *B. anthracis* (a fragment length of 207-225 bp). MLST was also performed on isolates with DNA that amplified using any of the primer pairs BA-*capB*, BA-*pag* or BA5510. Endpoint qPCR was conducted in a 50 µL reaction mixture (Table 2.7) using the Taq DNA Polymerase 1000 Units kit (QIAGEN, Netherlands). Conditions for qPCR are listed in Table 2.8 and data related to the primer pairs used to amplify DNA for MLST are listed in Table 2.9. MWG Eurofins, Germany, performed the sequencing of the PCR products.

.

Table 2.7: Reaction mixture used for qPCR on *B. cereus* group isolates for multi locus sequence typing.

Reagent	Quantity [µL]
dNTP (2.5 mM)[10]	4
Polymerase enzyme	0.2
PCR buffer	5
Primer pair mix	0.5-1.5 µM[11]
$MgCl_2$	2.5-3.5 mM[11]
dH_2O	16-25[12]
DNA template	3.5

Table 2.8: qPCR conditions used to amplify the different loci for multi locus sequence typing.

Analysis	Temperature [°C]	Cycles	Times [sec]
Denaturation	95	1	300
Denaturation	95	45	10
Annealing	52-59[13]		15-20[13]
Elongation	72		30-45[13]
Extension	72	1	720

[10] dNTP is the four bases (A, C, G, T)
[11] See Table 2.9 for details on the different primer pairs
[12] dH_2O was filled up to a total of 50 µL
[13] See Table 2.9 for details on the different primer pairs

Table 2.9: Primers and conditions for qPCR used for multi locus sequence typing of the selected *B. cereus* group members.

Primer pair	Primer sequences	Primer conc. [µM]	Mg conc. [mM]	Annealing temp. [°C]	Cycle times[14] [sec]	Reference
sucC	GGCGGAACAGAAATTGAAGA TCACACTTCATAATGCCACCA	0.5	3.5	58	20/20/45	(Helgason *et al.* 2004)
recF	GCGATGGCGAAATCTCATAG CAAATCCATTGATTCTGATACATC	1.5	3.0	56	20/20/45	(Helgason *et al.* 2004)
ftsA	TCTTGACATCGGTACATCCA GCCTGTAATAAGTGTACCTTCCA	1.5	2.0	54	20/20/45	(Helgason *et al.* 2004)
pyrE	TCGCATCGCATTTATTAGAA CCTGCTTCAAGCTCGTATG	0.75	2.5	56	20/20/45	(Helgason *et al.* 2004)
adk	CAGCTATGAAGGCTGAAACTG CTAAGCCTCCGATGAGAACA	0.8	2.65	57	20/20/45	(Tourasse *et al.* 2006)
ccpA	GTTTAGGATACCGCCCAAATG TGTAACTTCTTCGCGCTTCC	0.8	2.65	59	20/20/45	(Tourasse *et al.* 2006)
glpT	TGCGGCTGGATGAGTGA AAGTAAGAGCAAGGAAGA	0.8	2.65	59	20/20/45	(Tourasse *et al.* 2006)
glpF	GCGTTTGTGCTGGTGTAAG CTGCAATCGGAAGGAAGAAG	0.8	2.65	52	20/15/30	(Tourasse *et al.* 2006)
panC	CGATATCCTCGTGATATTGATAGAG TCCGCATAATCTACAGTGCCTTTC	0.8	2.65	57	20/15/30	(Tourasse *et al.* 2006)
pycA	GCGTTAGGTGGAAACGAAAG CGCGTCCAAGTTTATGGAAT	0.8	2.65	57	20/15/30	(Tourasse *et al.* 2006)
pta	GCAGAGCGTTTAGCAAAAGAA TGCAATGCGAGTTGCTTCTA	0.8	2.65	58	20/15/30	(Tourasse *et al.* 2006)

[14] Cycle times correspond to to denaturation/annealing/elongation

2.6 Phylogenetic analyses

BioNumerics v.6.6 (Applied Maths, Belgium) was used for all bioinformatics analyses. A local database was established, consisting of 60 isolates from Etosha National Park, 67 from Helgason *et al.* (2004), 5 in-house isolates and 8 isolates from MLST Oslo (http://mlstoslo.uio.no). Isolates from MLST Oslo (Table 2.10) were chosen due to their resemblance to *B. anthracis*. NVH0391/98 was included to root the analyses.

Table 2.10: Isolates from MLST Oslo (http://mlstoslo.uio.no).

Key	Description	Source	Location	Referance
G9241	*B. cereus* G9241 (2000031673, B06_002)	Human, sputum and blood	USA, Lousiana	(Hoffmaster *et al.* 2004)
CI	*B. cereus biovar anthracis* CI	Great apes	Ivory Coast	(Klee *et al.* 2006)
97-27	*B. thuringiensis konkukian* 97-27 (97/027, CEB97/27, CEB97/027)	Human, tissue	Yugoslavia, Bosnia-Herzegovina, Sarajevo	(Han *et al.* 2006)
R3098/03	*B. cereus* R309803 (R_3098/03)	Human, blood	UK, England	(Raedts *et al.* 2011)
E33L	*B. cereus* E33L (Zebra Killer)	Zebra, swab of carcass	Namibia, Etosha National Park	(Han *et al.* 2006)
B4264	*B. cereus* B4264 (2002734361)			(Tourasse and Kolstø 2008)
NVH0391/98	*B. cereus*	Food poisoning outbreak		(Lund *et al.* 2000)
2000031002			Belgian Congo	(Marston *et al.* 2006)

The sequences obtained from MWG Eurofins of the 60 isolates from this study were supplied to the local database as AB1 files (Appendix B). Two AB1 files, representing the forward and the reverse strand of the PCR product, for all eleven loci investigated, were added to the database for every one of the 60 isolates. The sequences of the remaining 80 isolates were imported to the database as FASTA files (Appendix B).

An assembly and editing operation was performed on the 60 *B. cereus* group isolates from this study to align the two AB1 files corresponding to one locus to obtain a consensus sequence. The consensus sequences were thereafter edited. This included insertion, deletion and trimming of the sequence to comply the sequence length and start- and stop patterns of the locus as published (Helgason *et al.* 2004; Tourasse *et al.* 2006). For the loci published twice (*adk, ccpA* and *glpT*), the trimming patterns published by Tourasse *et al.* (2006) were used. For the 80 *B. cereus* group members from various sources, the consensus sequences were added, and only editing of the sequences lengths were required. When all sequences had achieved correct editing, the whole database was triplicated leading to three independent locally stored databases comprising all the same isolates and corresponding loci consensus sequences.

For each of the databases a MLST experiment was constructed. The different MLST experiments included the loci presented by the two different MLST schemes and the combined MLST scheme respectively. The first database covered the MLST scheme presented by Helgason *et al.* (2004). *B. cereus* group isolates in the database fulfilling the MLST criteria (i.g. having a satisfying consensus sequence for all loci) were selected and an allele profile and a ST were generated for each of the isolates. A concatenated sequence was also generated for each isolate. These procedures were then repeated in the two other databases for the two remaining MLST schemes.

Clustering analyses were performed using the concatenated sequences and the NJ method on the two MLST schemes published individually (database one and two),

and on a combined MLST scheme (database three) including all eleven loci (*adk, ccpA, ftsA, glpF, glpT, panC, pta, pycA, pyre, recF and sucC*). A dendrogram using the allelic profiles and the NJ method were also constructed for MLST scheme 3 to check the similarities in clustering compared to the dendrograms constructed using the NJ method and the concatenated sequences. Bootstrap analyses were performed on all cluster analyses. Mutation analyses were performed for all MLST loci, on *B. cereus* group members clustered in clade I in the dendrogram constructed using the concatenated sequences and MLST scheme 3 (the combined MLST scheme).

B. cereus group member 2000031002 (Table 2.10) did not fulfil the MLST criteria by not having a published sequence for one of the MLST locus (*panC),* and therefore not included in the MLST analyses and the construction of dendrograms. This non-*B. anthracis* isolate is thus far identified with the highest degree of chromosomal similarities to *B. anthracis*. A point mutation analysis was performed to compare this isolate with the two isolates that clustered closest to *B. anthracis* when using concatenated sequences.

3 Results

3.1 Isolation of *B. cereus* group members from soil samples

In total 169 *B. cereus* group members were isolated from soil samples using BACARA (AES Chemunex, France) and sheep blood agar plates (Oxoid, Germany). The isolates were selected due to the presence of an opaque/white zone of precipitate surrounding the core using BACARA agar plates (Figure 2.2).

3.2 RT-qPCR screening and fragment analyses

In this study, 98 *B. cereus* group member isolates were screened using RT-qPCR to obtain a selection of isolates closely related to *B. anthracis*. The remaining 71 *B. cereus* group members isolated from soil samples were screened during earlier work (Hovland 2013). Six different primer pairs (Table 2.3) known to amplify DNA from strains of *B. anthracis* were used. All 98 *B. cereus* group isolates contained DNA that amplified using the *adk* primer pair (Ct-values < 30). None of the isolates, except for FFIBCgr114, contained DNA that amplified with a Ct-value below 30 for any of the *B. anthracis* positive markers (BA-*pag*, BA-*capB* and BA5510). *B. cereus* group member FFIBCgr114 amplified when use of all six RT-qPCR markers, strongly indicating this isolate to be *B. anthracis*.

The *B. cereus* group member isolate FFIBCgr121 amplified late (Ct-value>30) when the BA-*capB*, BA-*pag,* BA-*lef* and BA-*capB(2)* primer pairs (Table 2.3) were used. No amplification was seen when the BA5510 primer pair was used, even though this primer pair is assumed to amplify DNA isolated from isolates of *B. anthracis*. On the other hand, DNA from the *B. cereus* group member isolate FFIBCgr119 amplified using the BA5510 primer pair but no amplification occurred using the pXO1 and pXO2 plasmid markers.

38 out of 98 *B. cereus* group members contained a VNTR region similar to *B. anthracis* (Table 3.1). In addition, DNA from 8 *B. cereus* group isolates (not having a VNTR region similar to *B. anthracis*) amplified in at least one parallel using any of the primer pairs BA-*pag*, BA-*capB* or BA5510 (Table 3.1). The 46 isolates of *B. cereus* group members were selected for MLST analyses. VNTR fragments accepted as similar to *B. anthracis* ranged from 207-225 base pairs (bp). The VNTR region of *B. anthracis* positive control was measured to be 216 bp.

In addition to the 46 isolates (Table 3.1), MLST was performed on an additional 14 isolates. RT-qPCR screening of the isolates, including MLST results obtained using MLST scheme 1 (isolates FFIBCgr10, FFIBCgr29, FFIBCgr36, FFIBCgr42-FFIBCgr51) have been performed earlier (Hovland 2013). Results from RT-qPCR and VNTR analyses for all 112 *B. cereus* group member isolates (98 isolates from this study and 14 isolates from earlier work) are shown in Appendix C.

Table 3.1: RT-qPCR from amplification and VNTR analyses of 46 *B. cereus* group isolates selected for MLST analyses; + amplification in all parallels, - no amplification, +/- amplification in two out of three parallels, -/+ amplification in one parallel. Using the Bcms17 primer pair, the VNTR regions lengths are listed. "na" means not analysed.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17 (bp)	*plcR*-BA[15]	BA-*capB(2)*	BA-*lef*
FFIBCgr57	+	-	-/+	-	-	/	-	-
FFIBCgr58	+	-/+	-/+	-	-	-	-	-
FFIBCgr60	+	-/+	-/+	-	195	-	-	-
FFIBCgr61	+	-	-/+	-	208	-	-	-
FFIBCgr62	+	-	-	-	222	-		
FFIBCgr63	+	-/+	-	-	-	-	-	-
FFIBCgr65	+	-/+	-/+	-	195	-	-	-
FFIBCgr66	+	-/+	+/-	-	223	-	-	-
FFIBCgr85	+	-	-/+	-	221	-	-	-
FFIBCgr87	+	-	-	-	220	/	na	na
FFIBCgr90	+	-	-	-	219	-	na	na
FFIBCgr93	+	-	-	-	221	-	na	na
FFIBCgr94	+	-	-	-	220	-	na	na
FFIBCgr95	+	-	-	-	221	-	na	na
FFIBCgr96	+	-	-	-	207	-	na	na
FFIBCgr99	+	-	-	-	218	-	na	na
FFIBCgr101	+	-	-	-	215	-	na	na
FFIBCgr103	+	-	-	-	214	-	na	na
FFIBCgr104	+	-	-	-	215	-	na	na
FFIBCgr105	+	-	-	-	215	-	na	na
FFIBCgr107	+	-	-	-	216	-	na	na
FFIBCgr108	+	-	-	-	216	-	na	na

[15] + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in only one detection channel, / no amplification

Table 3.1: Continued from previous page.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17 (bp)	*plcR*-BA[16]	BA-*capB(2)*	BA-*lef*
FFIBCgr109	+	-	-	-	215	-	na	na
FFIBCgr110	+	-	-	-	216	-	na	na
FFIBCgr111	+	-	+/-	-	111	-	-	-
FFIBCgr112	+	-	-	-	216	-	na	na
FFIBCgr113	+	+/-	+/-	-	216	-	-	-
FFIBCgr114	+	+	+	+	216	+	na	na
FFIBCgr115	+	+/-	+/-	-	188	-	-/+	-
FFIBCgr116	+	+	+/-	-	217	-	-	-
FFIBCgr117	+	+/-	+/-	-	215	-	-	-
FFIBCgr118	+	-/+	-	-	215	-	-	-
FFIBCgr119	+	-	-	+	207	-	na	na
FFIBCgr120	+	-	-	-	215	-	na	na
FFIBCgr121	+	+	+	-	215	-	+	+
FFIBCgr129	+	-	-	-	216	-	na	na
FFIBCgr135	+	-	-	-	213	*	na	na
FFIBCgr136	+	-	-	-	215	-	na	na
FFIBCgr139	+	-	-	-	215	-	na	na
FFIBCgr140	+	-	-	-	215	-	na	na
FFIBCgr141	+	-	-	-	215	-	na	na
FFIBCgr142	+	-	-	-	219	-	na	na
FFIBCgr144	+	-	-	-	214	*	na	na
FFIBCgr145	+	-	-	-	215	-	na	na
FFIBCgr148	+	+/-	+/-	-	-	-	-	-
FFIBCgr149	+	-/+	-/+	-	221	-	-	-

[16] + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in only one detection channel, / no amplification

3.3 MLST and phylogenetic results

Two MLST schemes (Helgason *et al.* 2004; Tourasse *et al.* 2006) (hereafter referred to as MLST-1 and MLST-2) targeting in total eleven different loci were used to analyse the 60 *B. cereus* group members.

Dendrograms deduced from the two different MLST schemes, individually and combined (MLST-3), were constructed using concatenated sequences and the NJ method for 125 *B. cereus* group members (46 isolates from this study, 67 from Helgason *et al.* (2004), 7 from MLST Oslo and 5 in-house isolates). The cluster analyses resulted in three dendrograms representing MLST scheme 1(A), 2(B) and 3(C) respectively (Figure 3.1).

The three dendrograms deduced from data from MLST scheme 1, 2 and 3 revealed three distinct clades that overall include the same *B. cereus* group member isolates. For the three dendrograms (A, B and C) 90, 95 and 96 STs were obtained leading to discriminatory abilities of 1.39, 1.32 and 1.30 respectively. The discriminatory ability indicates the degree of resolution for the MLST scheme. If every isolate is assign a unique ST (different from the other STs), the MLST scheme has a discriminatory ability of 1. Based on the discriminatory abilities (1.39, 1.32 and 1.30), results obtained using MLST-3 were examined further to investigate details in the clustering.

For the three dendrograms (Figure 3.1), clade I comprised of the clonal population of *B. anthracis* isolates and *B. anthracis* close neighbours. On average, dendrogram A, B and C all clustered 39 *B. cereus* group member analysed during this study in clade I (Figure 3.1). Two *B. cereus* group members (AH1247 and R3098/03) differed in clustering between the three dendrograms. AH1247 clustered one time in clade II (MLST-1) and two times in clade III (MLST-2 and 3), and R3098/03 clustered two times in clade II (MLST-1 and MLST-3) and one time in clade III (MLST-2).

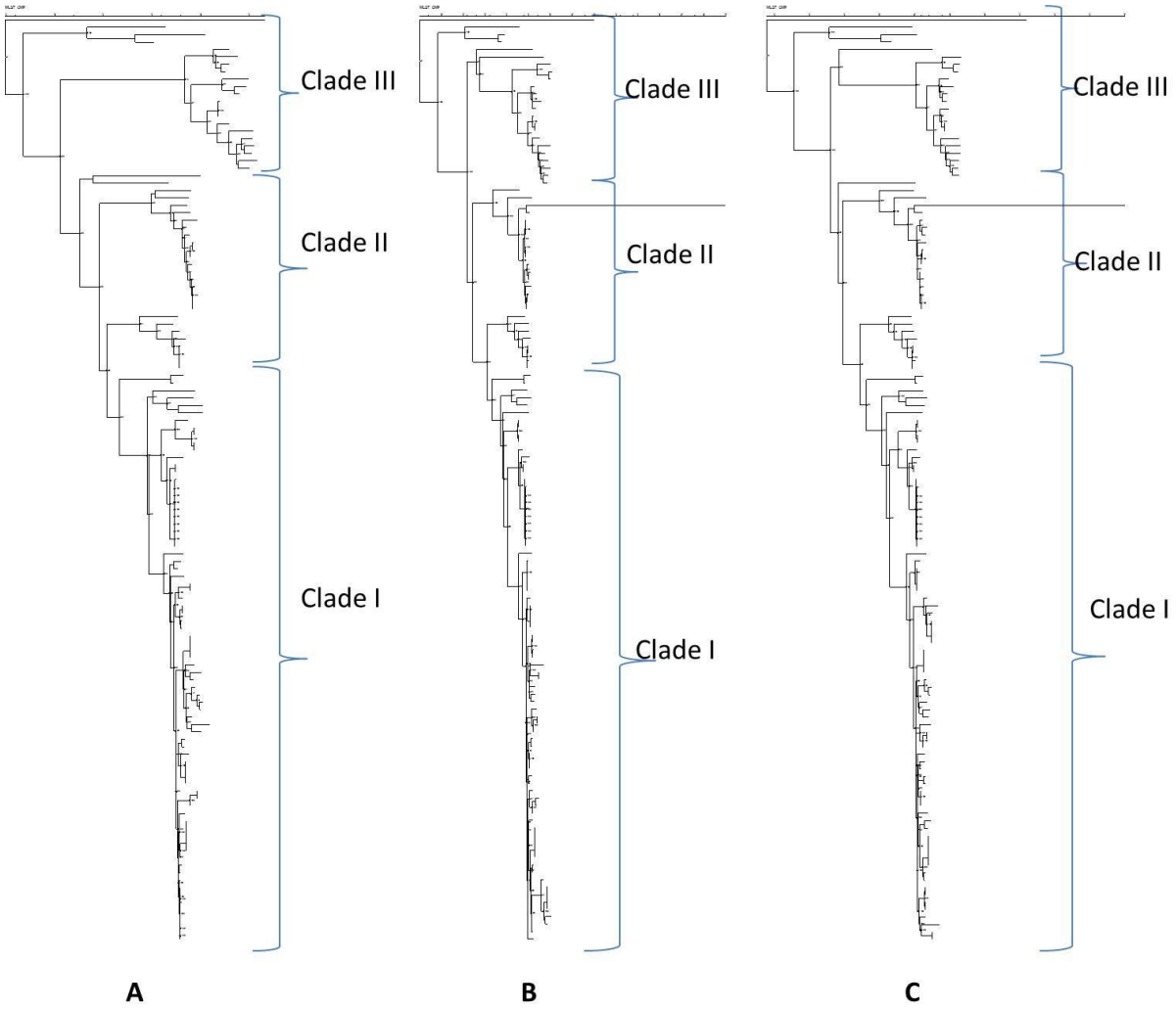


Figure 3.1: Dendrograms deduced from MLST-1(A)-, MLST-2(B)- and MLST-3(C) data of 125 *B. cereus* group members (46 isolates from this study, 67 from Helgason *et al.* (2004), 7 from MLST Oslo and 5 in-house isolates). Phylogenetic analyses were performed by the neighbour-joining method using concatenated sequences.

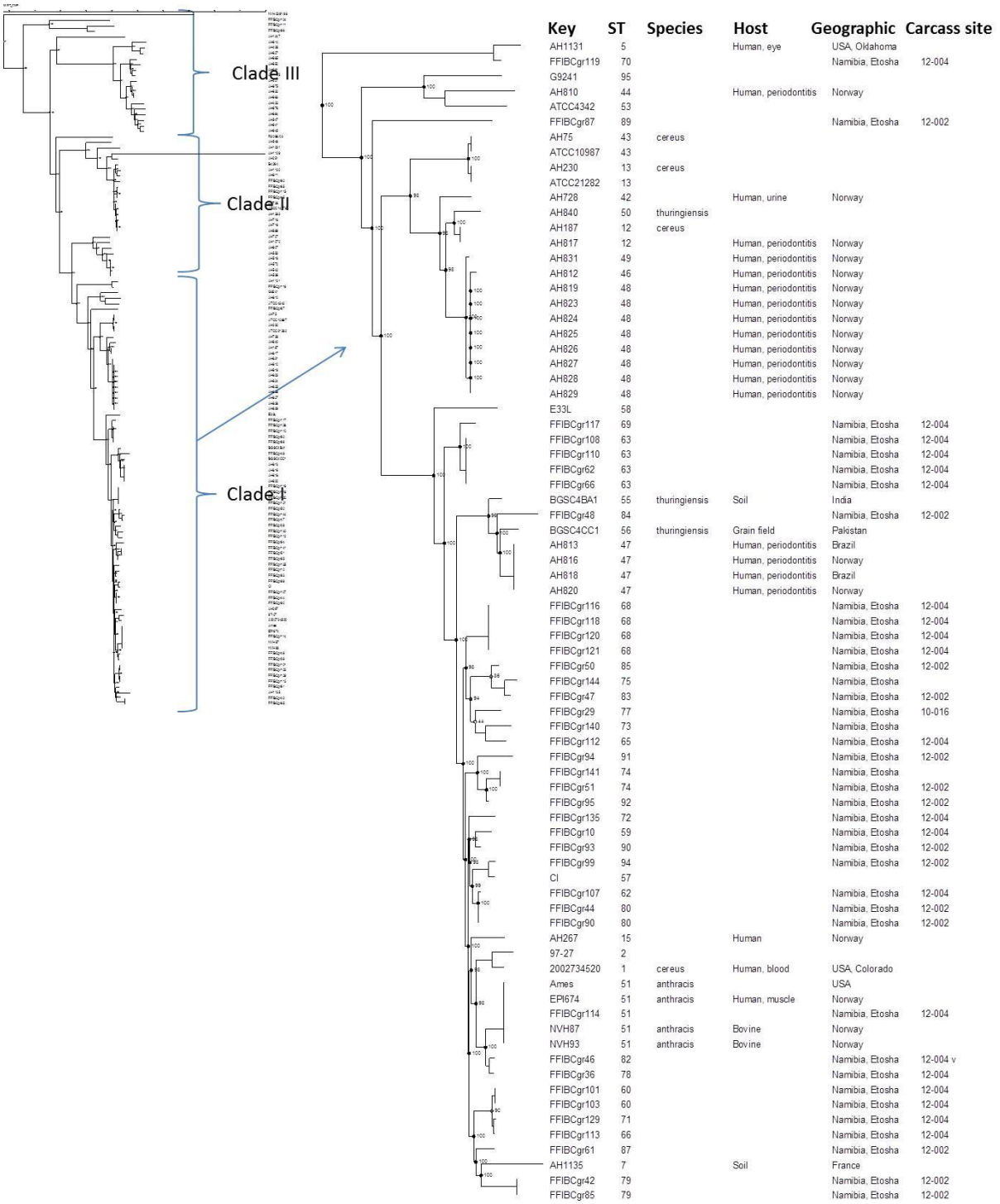
The lowest bootstrap value obtained for the three dendrograms (MLST-1, MLST-2 and MLST-3) were 75, 77 and 80 respectively. The exception was MLST-1 and MLST-3 that had one node each with bootstrap value of 62 and 51 respectively (both in clade I). The nodes both included the *B. cereus* group isolates FFIBCgr29, FFIBCgr112 and FFIBCgr140.

The three dendrograms did not include all 60 *B. cereus* group member isolates analysed during this study, due to insufficient PCR amplification for some loci. For this reason 14 isolates (FFIBCgr37, FFIBCgr43, FFIBCgr49, FFIBCgr57, FFIBCgr58, FFIBCgr63, FFIBCgr105, FFIBCgr109, FFIBCgr136, FFIBCgr139, FFIBCgr142, FFIBCgr145, FFIBCgr148 and FFIBCgr149) were omitted.

Phylogenetic results using MLST-3

The concatenated sequences (2658 base pairs) revealed a dendrogram (Figure 3.2) clustering 39 *B. cereus* group member from this study in clade I, 4 isolates (FFIBCgr45, FFIBCgr60, FFIBCgr65 and FFIBCgr115) in clade II and 3 isolates (FFIBCgr96, FFIBCgr104 and FFIBCgr111) clade III.

In total 96 STs were identified and 14 of the STs were represented more than once; (ST-12(2), ST-13(2), ST-19(2), ST-38(2), ST-43(2), ST-47(4), ST-48(8), ST-51(5), ST-60(2) ST-63(4), ST-68(4), ST-74(2), ST-79(2) and ST-80(2)). *B. cereus* group members isolated from Etosha National Park that revealed identical STs to each other (ST-47, ST-60, ST-63, ST-68, ST-74, ST-79 and ST-80), were collected from the same sampling spot (Appendix D). The *B. cereus* group member isolate FFIBCgr114 revealed a ST (ST-51) identical to various strains of *B. anthracis*.



Key	ST	Species	Host	Geographic	Carcass site
AH1131	5		Human, eye	USA, Oklahoma	
FFIBCgr119	70			Namibia, Etosha	12-004
G9241	95				
AH810	44		Human, periodontitis	Norway	
ATCC4342	53				
FFIBCgr87	89			Namibia, Etosha	12-002
AH75	43	cereus			
ATCC10987	43				
AH230	13	cereus			
ATCC21282	13				
AH728	42		Human, urine	Norway	
AH840	50	thuringiensis			
AH187	12	cereus			
AH817	12		Human, periodontitis	Norway	
AH831	49		Human, periodontitis	Norway	
AH812	46		Human, periodontitis	Norway	
AH819	48		Human, periodontitis	Norway	
AH823	48		Human, periodontitis	Norway	
AH824	48		Human, periodontitis	Norway	
AH825	48		Human, periodontitis	Norway	
AH826	48		Human, periodontitis	Norway	
AH827	48		Human, periodontitis	Norway	
AH828	48		Human, periodontitis	Norway	
AH829	48		Human, periodontitis	Norway	
E33L	58				
FFIBCgr117	69			Namibia, Etosha	12-004
FFIBCgr108	63			Namibia, Etosha	12-004
FFIBCgr110	63			Namibia, Etosha	12-004
FFIBCgr62	63			Namibia, Etosha	12-004
FFIBCgr66	63			Namibia, Etosha	12-004
BGSC4BA1	55	thuringiensis	Soil	India	
FFIBCgr48	84			Namibia, Etosha	12-002
BGSC4CC1	56	thuringiensis	Grain field	Pakistan	
AH813	47		Human, periodontitis	Brazil	
AH816	47		Human, periodontitis	Norway	
AH818	47		Human, periodontitis	Brazil	
AH820	47		Human, periodontitis	Norway	
FFIBCgr116	68			Namibia, Etosha	12-004
FFIBCgr118	68			Namibia, Etosha	12-004
FFIBCgr120	68			Namibia, Etosha	12-004
FFIBCgr121	68			Namibia, Etosha	12-004
FFIBCgr50	85			Namibia, Etosha	12-002
FFIBCgr144	75			Namibia, Etosha	
FFIBCgr47	83			Namibia, Etosha	12-002
FFIBCgr29	77			Namibia, Etosha	10-016
FFIBCgr140	73			Namibia, Etosha	
FFIBCgr112	65			Namibia, Etosha	12-004
FFIBCgr94	91			Namibia, Etosha	12-002
FFIBCgr141	74			Namibia, Etosha	
FFIBCgr51	74			Namibia, Etosha	12-002
FFIBCgr95	92			Namibia, Etosha	12-002
FFIBCgr135	72			Namibia, Etosha	12-004
FFIBCgr10	59			Namibia, Etosha	12-004
FFIBCgr93	90			Namibia, Etosha	12-002
FFIBCgr99	94			Namibia, Etosha	12-002
CI	57				
FFIBCgr107	62			Namibia, Etosha	12-004
FFIBCgr44	80			Namibia, Etosha	12-002
FFIBCgr90	80			Namibia, Etosha	12-002
AH267	15		Human	Norway	
97-27	2				
2002734520	1	cereus	Human, blood	USA, Colorado	
Ames	51	anthracis		USA	
EPI674	51	anthracis	Human, muscle	Norway	
FFIBCgr114	51			Namibia, Etosha	12-004
NVH87	51	anthracis	Bovine	Norway	
NVH93	51	anthracis	Bovine	Norway	
FFIBCgr46	82			Namibia, Etosha	12-004 v
FFIBCgr36	78			Namibia, Etosha	12-004
FFIBCgr101	60			Namibia, Etosha	12-004
FFIBCgr103	60			Namibia, Etosha	12-004
FFIBCgr129	71			Namibia, Etosha	12-004
FFIBCgr113	66			Namibia, Etosha	12-004
FFIBCgr61	87			Namibia, Etosha	12-002
AH1135	7		Soil	France	
FFIBCgr42	79			Namibia, Etosha	12-002
FFIBCgr85	79			Namibia, Etosha	12-002

Figure 3.2: Dendrogram deduced from MLST-3 data of 125 members of the *B. cereus* group (46 isolates from this study, 67 from Helgason *et al.* (2004), 5 in-house isolates and 7 isolates from MLST Oslo). Phylogenetic analysis was performed by the neighbour-joining method using concatenated sequences.

The number of alleles representing each MLST locus varied from 32 to 57 (Table 3.2). Several *B. cereus* group isolates (27, 9 and 6) had one, two or three alleles (the *ftsA, glpT, pta* and *recF* alleles) identical to *B. anthracis* respectively (Appendix E). The 6 *B. cereus* group isolates (AH820, AH818, AH816, AH813, FFIBCgr36 and FFIBCgr46) with three alleles identical to *B. anthracis* had all the *ftsA* and *recF* alleles identical to *B. anthracis*. In addition, the four *B. cereus* group members from Helgason *et al.* (2004) (AH820, AH818, AH816 and AH813) had the *pta* allele and the two *B. cereus* group members from this study (FFIBCgr36 and FFIBCgr46) had the *glpT* allele identical to *B. anthracis*.

Table 3.2: Number of alleles representing each of the MLST-3 loci.

Locus	Number of alleles	Locus	Number of alleles
adk	37	***pta***	49
ccpA	50	***pycA***	50
ftsA	32	***pyre***	52
glpF	57	***recF***	46
glpT	51	***sucC***	45
panC	52		

The mutation analyses of *B. cereus* group isolates clustered in clade I (Figure 3.2) revealed two isolates; FFIBCgr36 and FFIBgr46, to have 10 and 12 point mutations respectively that differentiate these isolates from *B. anthracis* (Appendix F). Other *B. cereus* group members in clade I clustering close to *B. anthracis*, 6 and 8 isolates had 16-20 and 21-25 point mutations respectively. Of these 14 isolates, 12 were *B. cereus* group members analysed during this study. The additional two isolates, an in-house isolate (82002734520), and the CI isolate (Klee *et al.* 2006), had 19 and 24 point mutations respectively. Locations of the point mutations for 16 *B. cereus* group members (isolates with up to 25 point mutations) revealed the mutations to be

situated with few exceptions at the same place in the DNA sequence for all the 16 *B. cereus* group member isolates (data not shown).

For the remaining isolates in clade I, the number of point mutations that differentiated these isolates from *B. anthracis* ranged from 26-151. *B. cereus* group member isolate AH1131 (Helgason *et al.* 2004) had 151 and were the isolate in clade I most distant to *B. anthracis*. Deletions and/or insertions of nucleotides were not observed in any of the MLST loci sequences and all mutations in clade I were categorised as silent mutations (data not shown).

The two *B. cereus* group isolates (FFIBCgr36 and FFIBCgr46) that clustered closest to *B. anthracis* in this study were compared to the *B. cereus* group isolate (2000031002) identified thus far with the highest chromosomal similarity to *B. anthracis.* A mutation analysis on ten loci revealed the three isolates to have 8, 11 and 4 point mutations that differentiated them from *B. anthracis* respectively (Table 3.3). The point mutations were located at the same base pair for the different isolates (data not shown).

Table 3.3: Point mutations differentiating three isolates (the isolate thus far with the highest chromosomal similarity to *B. anthracis* and the two isolates clustered closest to *B. anthracis* (Figure 3.2)) from *B. anthracis.*

Locus	***adk***	***ccpA***	***ftsA***	***glpF***	***glpT***	***pta***	**pycA**	***pyrE***	***recF***	***sucC***	***Total***
2000031002	1	0	0	1	0	0	1	0	0	1	**4**
FFIBCgr36	1	1	0	1	0	2	1	1	0	1	**8**
FFIBCgr46	1	1	0	1	0	2	3	2	0	1	**11**

Comparing the dendrogram deduced from concatenated sequences to the dendrogram deduced using allelic profiles (Figure 3.3) it revealed 10 *B. cereus* group member isolates to be relocated between different clades. *B. cereus* group isolates marked in red (AH1247, AH614, AH408 and AH627) were relocated from clade III to clade II, *B. cereus* group isolates marked in black (AH1131 and FFIBCgr119) were relocated from clade I to clade II and *B. cereus* group isolates marked in blue (FFIBCgr96, FFIBCgr104 and FFIBCgr111) were relocated from clade III to clade I. The *B. cereus* group member isolate AH1247 (marked in red) also changed clade when using concatenated sequences and the NJ method for all three MLST schemes (Figure 3.1).

Bootstrap values of 79 or higher were obtained for the dendrogram using allelic profiles, except for two nodes with bootstrap values of 46 (isolate B4264, AH1132 and AH811) and 48 (isolate AH681, AH1145 and AH685).

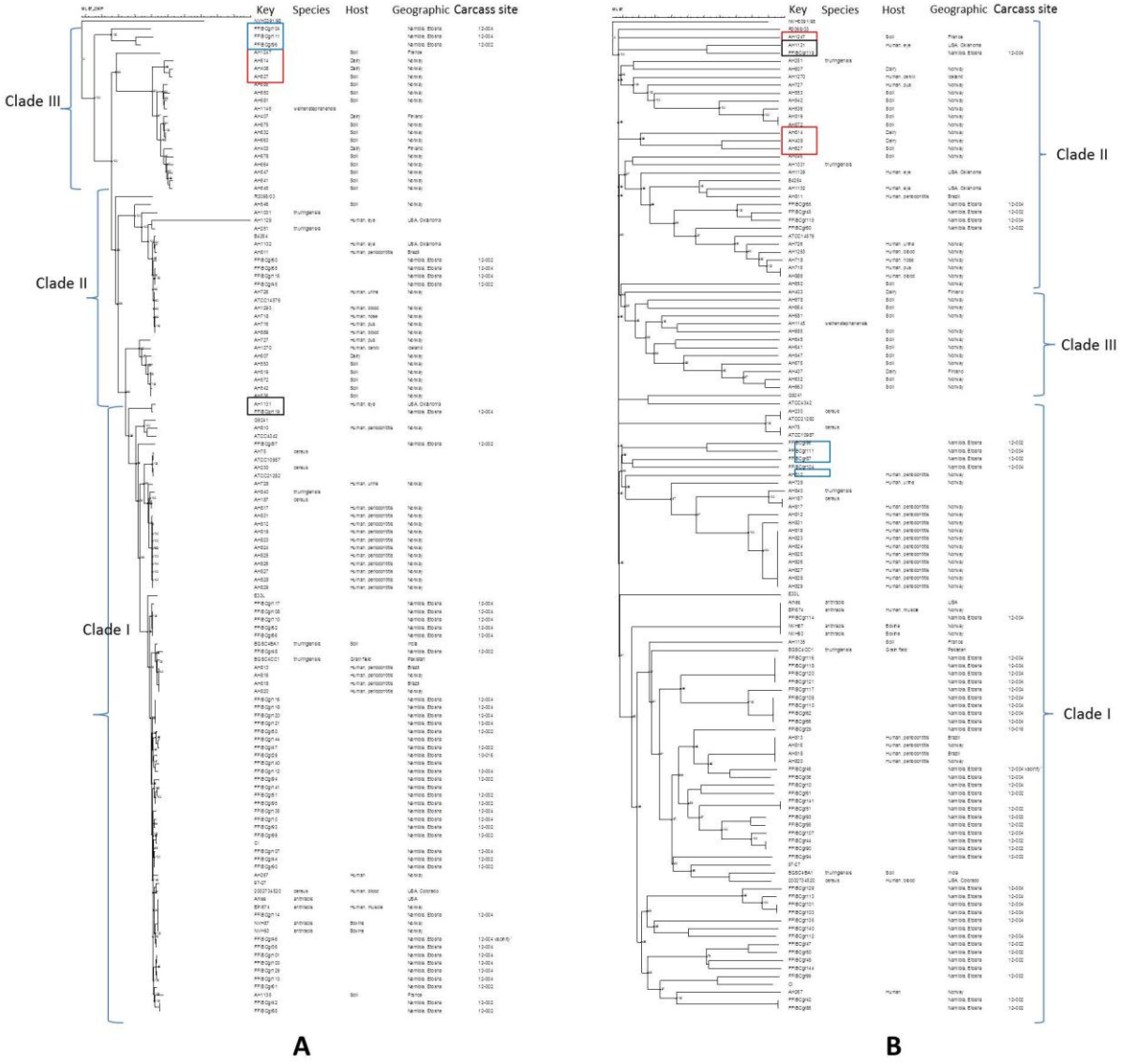


Figure 3.3: Dendrogram deduced from MLST-3 data of 125 members of the *B. cereus* group (46 ENP isolates, 67 isolates from Helgason *et al.* (2004), 5 in-house isolates and 7 isolates from MLST Oslo). Phylogenetic analysis was performed by the neighbour-joining method using concatenated sequences (A) and allelic profiles (B) respectively.

4 Discussion

4.1 qPCR screening and VNTR analyses

B. cereus group members isolated from soil samples collected in Etosha National Park (ENP), Namibia, were screened by RT-qPCR using specific markers (Table 2.3). The objective of this study was to study the genetic variation of *B. cereus* group members, with a particular focus on isolates closely related to *B. anthracis*.

It is an on-going debate whether *B. cereus*, *B. anthracis* and *B. thuringiensis* should be considered three variants of the same species or as three individual species (Helgason *et al.* 2000). This is due to the high degree of genetic similarities and that some of the phenotypically features differentiating these species are coded on mobile elements. Horizontal transfers of plasmids have been demonstrated for *B. cereus* and *B. thuringiensis* (Helgason *et al.* 2000; Rasko *et al.* 2007). However, such events have yet not been documented among *B. anthracis* subtypes (Pearson *et al.* 2004). The development of *B. anthracis* as a descendant of *B. cereus* and *B. thuringiensis* has been proposed (Keim *et al.* 1997; Valjevac *et al.* 2005), and later as a descendant only from *B. cereus (Okinaka et al. 2006)*. *B. anthracis* strains, with limited differences in their genomic sequences, are found globally. The high degree of genetic similarity makes it difficult to subtype the different strains, but it also indicates *B. anthracis* to be a recently emerged species (Keim and Wagner 2009). The vegetative life cycle of *B. anthracis* inside the host is short before the host dies and *B. anthracis* undergoes sporulation in the soil. Consequently the time available for accumulations of DNA mutations is limited (Harrell *et al.* 1995; Keim *et al.* 1997; Keim *et al.* 2000). The higher extend of mutations and recombination in *B. cereus* isolates may have led to a variety of genetic variants, more or less pathogenic. However, the introduction of the two plasmids (pXO1 and pXO2) appears to be of evolutionary advantage leading to a new variant adapted to a more restricted environmental niche (Keim and Wagner 2009).

Using a PCR marker amplifying a VNTR region in *B. cereus* group members (Valjevac *et al.* 2005), 39 % of the *B. cereus* group members isolated in this work (98 members isolated) were identified as close neighbours to *B. anthracis*, due to the amplification of an identical RT-qPCR fragment as *B. anthracis*. To select which isolates to undergo RT-qPCR screening, a pre-evaluation was performed using BACARA (AES Chemunex, France) agar plates, which are selective for *B. cereus* group members, to cultivate bacteria from the soil samples. Regardless of the pre-selection of isolates, 39 % identified as close neighbours, is a considerable amount compared to the 13 % of *B. cereus* group members isolated from various other sources (clinical, soil and dairy isolates) (Valjevac *et al.* 2005). The high amount of *B. anthracis* close neighbours isolated from soil samples collected in ENP may be a result of the soil samples being collected in an area known to be contaminated by *B. anthracis* (Valjevac *et al.* 2005). Some of the *B. cereus* group members isolated herein revealed identical STs to each other, hence increasing the portion of *B. anthracis* close neighbours. Previous studies reveal a general pattern of soil isolates mostly to be clustered in clade II and III (Helgason *et al.* 2004) compared to clinical *B. cereus* group members that mostly cluster in clade I (Tourasse *et al.* 2006). However, a study performed on environmental isolates from a variety of locations revealed soil isolates to have a greater tendency to harbour pXO1- and pXO2- like plasmids than isolates from other environmental environments (air, water and plants) (Hu *et al.* 2009). A general pattern indicate pXO1-like plasmids to be abundant in the environment (Bahl and Rosenberg 2010).

Several hypotheses may be addressed to answer the assumption that the portion of *B. anthracis* close neighbours may increase if the isolates are collected in an area contaminated by *B. anthracis*; 1) *B. anthracis* strains may have been exposed to point mutations during proliferation cycles leading to new strains not being complete copies; 2) *B. anthracis* living close to other members of the *B. cereus* group may have been involved in HGT leading to *B. anthracis* close neighbours (Battisti *et al.* 1985; Zwick *et al.* 2012); and 3) the environmental conditions in the area may have demanded the *B. cereus* group members to evolve to be able to survive, and in this case leading to isolates more similar to *B. anthracis* (Tourasse *et al.* 2006).

Only one *B. cereus* group member isolated during this study (FFIBCgr114) did indicate the presence of the two *B. anthracis* plasmids and the additional chromosomal marker tested (Table 3.1). Since *B. anthracis* harbour the pXO1 and pXO2 plasmids, and nearly none of the isolates from this study indicated the presence of these plasmids, they were assumed to be non-*B. anthracis* isolates. Even though the isolates clustered close to *B. anthracis* (Figure 3.2) only revealed few point mutations differentiating them from *B. anthracis* (Appendix F), did they not indicate the presence of the two plasmids. This largely eliminates hypothesis one; *B. anthracis* isolates to have undergone point mutations during cell proliferation.

The second hypothesis however, indicating the possibility of HGT may be a possible explanation. Even though MLST results used to study population structure of *B. cereus* group members give somewhat contradictory results as to whether the group is clonal or a result of recombination events, most papers indicate the *B. cereus* group to be weakly to largely clonal (Helgason *et al.* 2004; Priest *et al.* 2004). There are, however, also scientific evidences for HGT among *B. cereus* group members (Ko *et al.* 2004; Cardazzo *et al.* 2008), which supports the second hypothesis. For *B.* anthracis, these possible events would have taken place either inside the infected host or outside the infected host before sporulation of *B. anthracis*, since *B. anthracis* starts sporulation when exposed to high concentrations of free oxygen. On the other hand, the *B. cereus* group members analysed during this study did, as mention, not display the presence of any of the two pathogenic *B. anthracis* plasmids (except FFIBCgr114); hence the plasmids were not of the gene material transferred if HGT occurred. However, the isolates may harbour pXO1- and/or pXO2-like plasmids (Morten Søndenå personal communication, unpublished), such as plasmids found in the two *B. cereus* group members G9241 (Okinaka *et al.* 2006) and ATCC10987 (Rasko *et al.* 2004). A question to consider is whether the pXO1- and pXO2- like plasmids are plasmids from *B. anthracis* with some alterations due to additional mutations, or whether the pXO1- and pXO2- like plasmids are results of recombination from strains of *B. cereus* and/or *B. thuringiensis*.

Hypothesis three indicating the evolution of *B. cereus* group members to undertake point mutations and possibly HGT to adapt to a new environment is highly possible. The hypothesis leans towards *B. anthracis* being a descendant of *B. cereus (Okinaka et al. 2006)*. It is also possible to assume a combination of hypothesis 2 and hypothesis 3. The environmental conditions may be the reason for genetic alterations, but for HGT to occur the bacterial isolates have to be close to one another. Within-clade HGT events of the three *B. cereus* group clades have been shown to be performed at a higher rate than genetic import from external sources. However, clade I is shown to be the clade with most external genetic import compared to clade II and III (Didelot *et al.* 2009).

The lack of documented recombination events in diverse subtypes of *B. anthracis* do not support the theory of an increase of *B. anthracis* close neighbours due to contamination of *B. anthracis* (Pearson *et al.* 2004; Okinaka *et al.* 2006). However, as presented by the results in this study and indicated by Valjevac *et al.* (2005) the number of *B. anthracis* close neighbours in the samples is higher than would be expected by chance. The presence of pXO1- and pXO2- like plasmids in extensive *B. cereus* group members (Bahl and Rosenberg 2010) may indicate the possibility of HGT from *B. anthracis*, but this is still to be documented.

B. cereus group member FFIBCgr114 is shown to be a *B. anthracis* based on amplification of both pXO1 and pXO2 plasmid markers and the specific chromosomal *B. anthracis* marker (BA5510) (Table 3.1) .In addition, the isolate resulted in identical ST as *B. anthracis* isolates reported in previous work (Helgason *et al.* 2004). 19 of the *B. cereus* group members isolated during this study contained DNA that amplified at least one parallel using any of the primer pairs BA-*capB*, BA-*pag* or BA5510. Most of these amplifications had a high Ct-value (>30) and amplification was only obtained in one out of three parallels. Thus, these are assumed to be nonspecific amplification. This assumption is strengthened by the melting curve analysis on the PCR products. Only *B. cereus* group isolate FFIBCgr116, FFIBCgr119 and FFIBCgr121 had melting curves identical to *B. anthracis* for the PCR products (data not shown). 5 of the 19 *B. cereus* group members contained a VNTR region shorter

than *B. anthracis* (Table 3.1), which indicated the isolates not to be *B. anthracis* close neighbours.

Two of the *B. cereus* group isolates contained DNA that amplified in all three parallels using the BA5510 marker (isolate FFIBCGr119) and the four plasmids markers (BA-*pag,* BA-*capB,* BA-*lef* and BA-*capB(2)*) (isolate FFIBCgr121) (Table 3.1). Amplification of DNA from *B. cereus* group member FFIBCgr119 using the BA5510 primer pair obtainable a Ct-value lower than the *B. anthracis* positive control (data not shown). This isolate did, however, not contain DNA that amplified using the pXO1, pXO2 or *plcR* markers. Non-*B. anthracis* isolates indicating the presence of the chromosomal marker amplifying when using the BA5510 marker have earlier been observed twice (Olsen *et al.* 2007). The two isolates were both clustered close to *B. anthracis* using the MLST scheme published by Helgason *et al.* (2004). Isolate FFIBCgr119 clustered in both clade I and clade II (Figure 3.3) and revealed 141 point mutations in difference to *B. anthracis* (Appendix F). Even with DNA amplifying using the BA5510 primer pair, the isolate did not indicate the presence of the two *B. anthracis* plasmids (pXO1 and pXO2), and was hence assumed to be a non-*B. anthracis* isolate. B*. cereus* group member FFIBCgr121, on the other hand, obtained high Ct-values (>30) in all parallels using the *B. anthracis* plasmid markers, however low DNA concentration should not be the reason for late amplification based on the Ct-values obtained using the *adk* primer pair. The obtained amplifications could be disturbance and assembly of primer dimers, but this is unlikely due to the melting curve analysis on the PCR products. FFIBCgr121 may harbour pXO1- and pXO2- like plasmids (Morten Søndenå personal communication, unpublished) with alterations in the primer binding sequences causing only some of the primers to bind, hence high Ct-values. The isolate clustered close to *B. anthracis* (Figure 3.2) and revealed only 29 point mutations in difference to *B. anthracis* (Appendix F). Compared to *B. cereus* group member G9241, E33L and ATCC10987 (Hoffmaster *et al.* 2004; Rasko *et al.* 2004; Han *et al.* 2006) associated with anthrax like diseases (G9241 and E33L) and a pXO1-like plasmid (ATCC10987), FFIBCgr121 is assumed to be even closer related to *B. anthracis* based on chromosomal markers.

4.2 MLST and phylogenetic analyses

The NJ method was the chosen method to perform the phylogenetic analyses as most of the isolates were considered to be clonal (Helgason *et al.* 2004). The 60 *B*. *cereus* group members analysed during this study, and selected for MLST, were also assumed to be close relatives of *B. anthracis* due to the amplification of a VNTR region with identical length as *B. anthracis*. Both allelic profiles and concatenated sequences have in earlier studies been used to perform clustering of *B. cereus* group members (Helgason *et al.* 2004; Priest *et al.* 2004; Sorokin *et al.* 2006; Tourasse *et al.* 2006). The concatenated sequences were chosen to look for details in the clustering based on earlier work that indicate the *B. cereus* group to be mostly clonal (Helgason *et al.* 2004), and the use of nucleotide sequences are recommended when analysing clonal organisms (Maiden 2006).

Three different MLST schemes were used to study the genetic variation of *B. cereus* group members (Helgason *et al.* (2004), Tourasse *et al.* (2006) and a combination of the two). MLST analyses on 125 *B. cereus* group members (46 isolates form this study, 67 from Helgasin *et al.* (2004), 5 in-house isolates and 7 isolates from MLST Oslo) and dendrograms using the concatenated sequences and the NJ method were constructed for all three MLST schemes (MLST-1, MLST-2 and MLST-3) to look for differences in clustering. For MLST-1, MLST-2 and MLST-3 the clustering was mostly the same and the discriminatory ability calculated for the three different MLST schemes were highly congruent; 1.39, 1.32 and 1.30 respectively. MLST-3, with the discriminatory ability of 1.30, was the MLST scheme chosen to study details in the clustering. This was due to this MLST schemes assumingly better ability to split the different *B. cereus* group isolates compared to the two other MLST schemes. MLST-3 included eleven loci, whereas MLST-1 and MLST-2 comprised only of seven loci each. Due to the four additional loci sequenced when using MLST-3, the discriminatory ability was assumed to be closer to 1 than for the other two MLST schemes. However, only one ST differed when using MLST-2 and MLST-3 (95 and 96 respectively), which indicated that the four additional loci in MLST-3 do not have a large influence on the discriminatory ability. Discriminatory abilities obtained for all three MLST schemes were closer to 1 than the discriminatory abilities published by Helgason *et al.* (2004) and Priest *et al.* (2004) previously; 1.45 and 1.78 respectively.

B. cereus group members analysed during this study gave a surprisingly discriminatory ability of 1.28 (46 *B. cereus* group isolates from this study led to 36 STs). This is close to 1 considering all the isolates to be from soil samples collected in the same area. Helgason *et al.* (2004) and Priest *et al.* (2004) both analysed *B. cereus* group members from diverse sources and locations. With the *B. cereus* group members isolated herein (all from soil samples collected in the same area), it was assumed that several isolates would result in identical STs to each other. This result highlights how diverse the *B. cereus* group members are (Helgason *et al.* 1998).

Two of the *B. cereus* group members (FFIBCgr36 and FFIBCgr46) with three alleles identical to *B. anthracis* (Appendix E) had lowest number of point mutations, 10 and 12 respectively, which differentiated them from *B. anthracis* (Appendix F). Even with just a few point mutations, the two isolates are assumed to be non- *B. anthracis* isolates due to no amplifications of the plasmids- or chromosomal markers (Appendix C). FFIBCgr36 and FFIBCgr46 harboured more point mutations than *B. cereus* group member 2000031002 (Table 3.3), which thus far is the non-*B. anthracis* isolate most similar to *B. anthracis* based on chromosomal markers (Marston *et al.* 2006). Other *B. cereus* group isolates (AH820, AH818, AH816 and AH813), also having three alleles identical to *B. anthracis,* clustered quite a distance away from *B. anthracis* (Figure 3.2). The 4 isolates had 28 point mutations compared to *B. anthracis* (Appendix F). As the dendrogram (Figure 3.2) was based on concatenated sequences the high number of point mutations moved the isolates away from *B. anthracis* even with three identical alleles. Based on the MLST scheme published by Helgason *et al.* (2004), the 4 isolates just discussed clustered closest to *B. anthracis* and were separated from *B. anthracis* by only 6 point mutations. The additional 22 point mutations were exposed due to the four additional loci investigated when combining the two MLST schemes (Helgason *et al.* 2004; Tourasse *et al.* 2006). *B. cereus* group isolates that clustered closer to *B. anthracis* than the four isolates just discussed (isolates with up to 27 point mutations), all had on or two alleles identical to *B. anthracis*.

Dendrograms constructed based on concatenated sequences and allelic profiles using MLST-3 clustered relatively equally (Figure 3.3). However, some of the *B. cereus* group isolates relocated between clades depending on the different methods used. Concatenated sequences clustered four (AH1247, AH614, AH408 and AH627) and three (FFIBCgr96, FFIBCgr104 and FFIBCgr111) isolates in clade III, while allelic profiles clustered the same isolates in clade II and I respectively. Two additional isolates (AH1131 and FFIBCgr119) were relocated from clade I to clade II (Figure 3.3). *B. cereus* group member AH1247 also altered between clade II and clade III when concatenated sequences were used for all three MLST schemes (Figure 3.1). The nine *B. cereus* group members marked in red and blue (Figure 3.3) relocated from clade III to clade II and I indicate that the point mutations may be distributed on a few loci instead of evenly scattered among all eleven loci. This is confirmed by the mutation analyses (Appendix F), and may be a result of recombination events (Olsen *et al.* 2007). The three isolates (FFIBCgr96, FFIBCgr104 and FFIBCgr111) that relocated from clade III to clade I clearly indicate how concatenated sequences and allelic profiles can cluster the same isolates differently. The three isolates embrace many point mutations (270, 302 and 319) that differentiate them from *B. anthracis* and this causes the method using concatenated sequences to reject the isolates from being clustered in clade I. For comparison; the *B. cereus* group member in clade I that contains the most point mutations, harbour 151 point mutations. The three isolates however, have two or three loci nearly identical to *B. anthracis,* thus few point mutations. The allelic profiles of the three isolates do not differ exceptionally from the allelic profile of *B. anthracis,* hence, when using allelic profiles they are clustered in clade I.

Two *B. cereus* group members (AH1131 and FFIBCgr119) had point mutations evenly distributed on all eleven loci (Appendix F), but clustered in clade I when using concatenated sequences, and clustered in clade II when allelic profiles were used. The isolates clustered on the edge of clade I using the concatenated sequences, and thus speculations arise to whether they should have been clustered in clade II. Since the mutations are evenly distributed on all loci there is no indication of recombination and it is likely that these isolates are in the distinction between clade I and clade II, and different methods will cluster the isolates differently.

4.3 Further work

Based on the results obtained in this study, further study should investigate the phenotypically characteristics of the isolates from Etosha National Park that clustered in clade I. *B. cereus* group members FFIBCgr36 and FFIBCgr46 should be of special interest due to their chromosomal close resemblance to *B.* anthracis.

It can also be of interest to look more closely into where the soil samples were collected to see if there is a correlation between the isolates' resemblances to *B. anthracis* and the distance between the carcass and the sample point.

The isolates should also be analysed for the presence of pXO1- and pXO2- like plasmids and compared to earlier published isolates, *e.g* 2000031002 and G9241(Marston *et al.* 2006; Okinaka *et al.* 2006), known to harbour *B. anthracis* like plasmids. Regarding the pXO1- and pXO2- like plasmids it also is of interest to investigate whether they are transferred from strains of *B. anthracis* or *B. cereus/B. thuringiensis*.

5 Conclusion

In this study 169 *B. cereus* group members were isolated from soil samples collected in Etosha National Park, Namibia. Based on RT-qPCR screening using six primer pairs known to give amplification for *B. anthracis*, 60 isolates were assumed to be close neighbours of *B. anthracis* and were genotyped using three different MLST schemes. Two of the MLST schemes (Helgason *et al.* 2004; Tourasse *et al.* 2006) comprise of 7 loci each and the third MLST scheme (a combination of the first two MLST schemes) comprise of 11 loci and indicated a better discriminatory ability and were used to study the details in the clustering.

Phylogenetic analyses were performed using the third MLST scheme on 125 *B. cereus* group members (46 isolates from this study, 67 from Helgason *et al.* (2004), 5 in-house isolates and 7 isolates from MLST Oslo) and revealed 96 STs. The high number of STs supports the great diversity of the *B. cereus* group members earlier discovered.

Concatenated sequences and the NJ method clustered 39 isolates analysed during this study in clade I together with *B. anthracis*. This displays a high amount of *B. cereus* group members isolated from soil to be closely related to *B. anthracis* based on chromosomal markers. One isolate (FFIBCgr114) obtained a ST identical to other *B. anthracis* isolates. The two isolates (FFIBCgr36 and FFIBCrg46) that clustered closest to *B. anthracis* revealed a total of 10 and 12 point mutations that differentiated them from *B. anthracis* when 11 loci (*adk, ccpA, ftsA, glpF, glpT, panC, pta, pycA, pyre, recF* and *sucC*) were analysed. FFIBCgr36 and FFIBCgr46 were, however, more distant to *B. anthracis* than the *B. cereus* group member, isolate 2000031002, most similar to *B. anthracis* based on chromosomal markers.

References

Aanonsen, D. and B. G. Spratt (2005). "The multilocus sequence typing network: mlst.net." Nucleic Acids Research **33**: W728-W733.

Bahl, M. I. and K. Rosenberg (2010). "High abundance and diversity of Bacillus anthracis plasmid pXO1-like replicons in municipal wastewater." FEMS microbiology ecology **74**(1): 241-247.

Baldauf, S. L. (2003). "Phylogeny for the faint of heart: a tutorial." TRENDS in Genetics **19**(6): 345-351.

Barker, M., B. Thakker, et al. (2005). "Multilocus sequence typing reveals that Bacillus cereus strains isolated from clinical infections have distinct phylogenetic origins." FEMS microbiology letters **245**(1): 179-184.

Battisti, L., B. D. Green, et al. (1985). "Mating system for transfer of plasmids among Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis." Journal of Bacteriology **162**(2): 543-550.

BioPlanet (2013). "What is bioinformatics." Retrieved 13/11/2013, 2013, from http://www.bioplanet.com/what-is-bioinformatics/.

Bottone, E. J. (2010). "Bacillus cereus, a volatile human pathogen." Clinical microbiology reviews **23**(2): 382-398.

Cardazzo, B., E. Negrisolo, et al. (2008). "Multiple-locus sequence typing and analysis of toxin genes in Bacillus cereus food-borne isolates." Applied and Environmental Microbiology **74**(3): 850-860.

Ceuppens, S., N. Boon, et al. (2013). "Diversity of Bacillus cereus group strains is reflected in their broad range of phatogenicity and diverse ecological lifestyles." FEMS Microbiology Ecology **84**(3): 433-450.

Chemunex, A. (2012). "New B. cereus Selective Agar BACARA in FDA BAM." Retrieved 13/11/2013, 2013, from http://www.rapidmicrobiology.com/news/8h32.php.

Chu, H. P. (1949). "The Lecithinase of Bacillus cereus and its Comparison with Clostridium welchii α-toxin." Microbiology and Molecular Biology Reviews **3**(2): 255-273.

Dai, Z., J. C. Sirard, et al. (1995). "The atxA gene product activates transcription of the antrax toxine and is essential for virulence." Molecular Microbiology **16**(6): 1171-1181.

DeBell, R. M. (2002). "Progress in Research with the plcR Gene in Bacillus anthracis." The ASA Newsletter(102).

Didelot, X., M. Barker, et al. (2009). "Evolution of pathogenicity in the *Bacillus cereus* group." Systematic and Applied Microbiology **32**(2): 81-90.

Easterday, W. R., M. N. Van Ert, et al. (2005). "Use of single nucleotide polymorphisms in the plcR gene for specific identification of Bacillus anthracis." Journal of Clinical Microbiology **43**(4): 1995-1997.

Epizootics, I. O. o. (2008). Anthrax in humans and animals, World Health Organization.

Ezzel, J. W. and S. L. Welkos (1999). "The capsule of Bacillus anthracis, a review." Journal of Applied Microbiology(87): 250.

Granum, P. E. and T. Lund (1997). "Bacillus cereus and its food poisoning toxins." FEMS microbiology letters **157**(2): 223-228.

Han, C. S., G. Xie, et al. (2006). "Pathogenomic sequence analysis of Bacillus cereus and Bacillus thuringiensis isolates closely related to Bacillus anthracis." Journal of Bacteriology **188**(9): 3382-3390.

Harrell, L. J., G. L. Andersen, et al. (1995). "Genetic variability of Bacillus anthracis and related species." Journal of Clinical Microbiology **33**(7): 1847-1850.

Helgason, E., D. A. Caugant, et al. (1998). "Genetic diversity of Bacillus cereus/B. thuringiensis isolates from natural sources." Current microbiology **37**(2): 80-87.

Helgason, E., O. A. Okstad, et al. (2000). "Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis--one species on the basis of genetic evidence." Applied and Environmental Microbiology **66**(6): 2627-2630.

Helgason, E., N. J. Tourasse, et al. (2004). "Multilocus sequence typing scheme for bacteria of the Bacillus cereus group." Applied and Environmental Microbiology **70**(1): 191-201.

Hill, K. K., L. O. Ticknor, et al. (2004). "Fluorescent amplified fragment length polymorphism analysis of Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis isolates." Applied and Environmental Microbiology **70**(2): 1068-1080.

Hillis, D. M. and J. J. Bull (1993). "An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis." Systematic biology **42**(2): 182-192.

Hoffmaster, A. R., J. Ravel, et al. (2004). "Identification of anthrax toxin genes in a Bacillus cereus associated with an illness resembling inhalation anthrax." Proc Natl Acad Sci U S A **101**(22): 8449-8454.

Hovland, K. (2013). Close relatives to Bacillus anthracis.

Hu, X., G. Van der Auwera, et al. (2009). "Distribution, diversity, and potential mobility of extrachromosomal elements related to the Bacillus anthracis pXO1 and pXO2 virulence plasmids." Applied and Environmental Microbiology **75**(10): 3016-3028.

Huang, J., D. Mirel, et al. (2011). "Minimum information about a genotyping experiment (MIGEN)." Standards in genomic sciences **5**(2): 224.

Imperiale, M. J. and A. Casadevall (2011). "Bioterrorism: Lessons learned since the anthrax mailings." Americal Society for Microbiology; MBio **2**(6): e00232-00211.

Inglesby, T. V., T. O'Toole, et al. (2002). "Anthrax as a biological weapon, 2002." JAMA: the journal of the American Medical Association **287**(17): 2236-2252.

Jensen, G. B., N. Fisker, et al. (2005). "The possibility of discriminating within the *Bacillus cereus* group using *gyrB* sequencing and PCR-RFLP." International journal of food microbiology **104**(1): 113-120.

Keim, P., A. Kalif, et al. (1997). "Molecular evolution and diversity in Bacillus anthracis as detected by amplified fragment length polymorphism markers." Journal of Bacteriology **179**(3): 818-824.

Keim, P., L. Price, et al. (2000). "Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within Bacillus anthracis." Journal of Bacteriology **182**(10): 2928-2936.

Keim, P. S. and D. M. Wagner (2009). "Humans and evolutionary and ecological forces shaped the phylogeography of recently emerged diseases." Nature Reviews Microbiology **7**(11): 813-821.

Klee, S. R., E. B. Brzuszkiewicz, et al. (2010). "The genome of a Bacillus isolate causing anthrax in chimpanzees combines chromosomal properties of B. cereus with B. anthracis virulence plasmids." PLoS One **5**(7): e10986.

Klee, S. R., M. Ozel, et al. (2006). "Characterization of Bacillus anthracis-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon." Journal of Bacteriology **188**(15): 5333-5344.

Ko, K. S., J. W. Kim, et al. (2004). "Population structure of the Bacillus cereus group as determined by sequence analysis of six housekeeping genes and the plcR Gene." Infection and Immunity **72**(9): 5253-5261.

Koehler, T. M. (1999). "Bacillus anthracis Physiology and Genetics." Molecular Aspects of Medicine **30**(6): 386-396.

Lindeque, P. and P. Turnbull (1994). "Ecology and epidemiology of anthrax in the Etosha National Park, Namibia." The Onderstepoort journal of veterinary research **61**(1): 71-83.

Lund, T., M. L. De Buyser, et al. (2000). "A new cytotoxin from Bacillus cereus that may cause necrotic enteritis." Molecular microbiology **38**(2): 254-261.

Madigan, M. T., J. M. Martinko, et al. (2009). Brock Biology of Microorganisms, Pearson Education.

Maiden, M. C. (2006). "Multilocus sequence typing of bacteria." Annu. Rev. Microbiology **60**: 561-588.

Maiden, M. C., J. A. Bygraves, et al. (1998). "Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms." Proc Natl Acad Sci U S A **95**(6): 3140-3145.

Maier, R. M., I. L. Pepper, et al. (2009). Environmental microbiology, Access Online via Elsevier.

Marston, C. K., J. E. Gee, et al. (2006). "Molecular approaches to identify and differentiate Bacillus anthracis from phenotypically similar Bacillus species isolates." BMC microbiology **6**(1): 22.

Matero, P., H. Hemmilä, et al. (2011). "Rapid field detection assays for Bacillus anthracis, Brucella spp., Francisella tularensis and Yersinia pestis." Clinical Microbiology and Infection **17**(1): 34-43.

Maughan, H. and G. Van der Auwea (2011). "Bacillus taxonomy in the genomic era finds phenotypes to be essential though often misleading." Infections, Genetics and Evolution **11**(5): 789-797.

Mignot, T., M. Mock, et al. (2001). "The incompatibility between the PlcR-and AtxA-controlled regulons may have selected a nonsense mutation in Bacillus anthracis." Molecular microbiology **42**(5): 1189-1198.

Mount, D. W. (2004). "Sequence and genome analysis." Bioinformatics: Cold Spring Harbour Laboratory Press **2**.

Okinaka, R., K. Cloud, et al. (1999). "Sequence, assembly and analysis of pX01 and pX02." Journal of Applied Microbiology **87**(2): 261-262.

Okinaka, R., T. Pearson, et al. (2006). "Antrax, but not Bacillus anthracis?" PLOS Pathogens **2**(11): e122.

Olsen, J. S., G. Skogan, et al. (2007). "Genetic distribution of 295 Bacillus cereus group members based on adk-screening in combination with MLST (Multilocus Sequence Typing) used for validating a primer targeting a chromosomal locus in B. anthracis." Journal of Microbiological Methods **71**(3): 265-274.

Pearson, T., J. D. Busch, et al. (2004). "Phylogenetic discovery bias in Bacillus anthracis using single-nucleotide polymorphisms from whole-genome sequencing." Proc Natl Acad Sci U S A **101**(37): 13536-13541.

Priest, F. G., M. Barker, et al. (2004). "Population structure and evolution of the Bacillus cereus group." Journal of Bacteriology **186**(23): 7959-7970.

QIAGEN (2002). QuantiTect Probe RT-PCR Handbook: For quentitative, real time, one step RT-PCR using sequence specific probes.

Raedts, J., S. W. Kengen, et al. (2011). "Occurrence of L-iduronic acid and putative D-glucuronyl C5-epimerases in prokaryotes." Glycoconjugate Journal **28**(2): 57-66.

Rasko, D. A., J. Ravel, et al. (2004). "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXO1." Nucleic acids research **32**(3): 977-988.

Rasko, D. A., M. Rosovitz, et al. (2007). "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." Journal of Bacteriology **189**(1): 52-64.

Rasko, D. A., P. L. Worsham, et al. (2011). "Bacillus anthracis comparative genome analysis in support of the Amerithrax investigation." Proc Natl Acad Sci U S A **108**(12): 5027-5032.

Saile, E. and T. M. Koehler (2006). "Bacillus anthracis multiplication, persistence, and genetic exchange in the rhizosphere of grass plants." Applied and Environmental Microbiology **72**(5): 3168-3174.

Schnepf, E., N. Crickmore, et al. (1998). "Bacillus thuringiensis and its Pesticidal Crystal Proteins." Microbiology and Molecular Biology Reviews **62**(3): 775-806.

Shadomy, S. V. (2012). Infectious Diseases Related To Travel. New York, Oxford University Press,.

Sorokin, A., B. Candelon, et al. (2006). "Multiple-locus sequence typing analysis of Bacillus cereus and Bacillus thuringiensis reveals separate clustering and a distinct population structure of psychrotrophic strains." Applied and Environmental Microbiology **72**(2): 1569-1578.

Soufiane, B. and J. C. Côté (2013). "Bacillus weihenstephanensis characteristics are present in Bacillus cereus and Bacillus mycoides strains." FEMS Microbiology Letters: 1-11.

Todar, K. (2006). Todar's online textbook of bacteriology, University of Wisconsin-Madison Department of Bacteriology.

Tourasse, N. J., E. Helgason, et al. (2006). "The Bacillus cereus group: novel aspects of population structure and genome dynamics." Journal of Applied Microbiology **101**(3): 579-593.

Tourasse, N. J. and A.-B. Kolstø (2008). "Survey of group I and group II introns in 29 sequenced genomes of the Bacillus cereus group: insights into their spread and evolution." Nucleic acids research **36**(14): 4529-4548.

Turnbull, P. C. (1996). Bacillus. Medical microbiology. S. Baron. Galveston (TX), University of Texas Medical Branch at Galveston.

Unknown (2009). "Anthrax Strain (34F2) of Bacillus anthracis." National center of Emerging and Zoonotiv infectious Diseases.

Valjevac, S., V. Hilaire, et al. (2005). "Comparison of minisatellite polymorphisms in the Bacillus cereus complex: a simple assay for large-scale screening and identification of strains most closely related to Bacillus anthracis." Applied and Environmental Microbiology **71**(11): 6613-6623.

Wilson, M. K., J. M. Vergis, et al. (2011). "Bacillus cereus G9241 makes anthrax toxin and capsule like highly virulent B. anthracis ames but behaves like attenuated toxigenic nonencapsulated B. anthracis sterne in rabbits and mice." Infection and immunity **79**(9): 3012-3019.

Zwick, M. E., S. J. Joseph, et al. (2012). "Genomic characterization of the Bacillus cereus sensu lato species: backdrop to the evolution of Bacillus anthracis." Genome Research **22**(8): 1512-1524.

Økstad, O. A. and A. B. Kolstø (2011). Genomics of Bacillus Species. Genomics of foodborne bacterial pathogens**:** 29-53.

Appendix A:

Protocol for DNA extraction: QIAgamp DNA MINI KIT (QIAGEN, Netherlands)

Isolation of genomic DNA from Gram-positive bacteria
D2. Suspend bacterial pellet in 180 µl of the appropriate enzyme solution (20 mg/ml lysozyme or 200 µg/ml lysostaphin; 20 mM Tris·HCl, pH 8.0; 2 mM EDTA;1.2 % Triton).

D3. Incubate for at least 30 min at 37°C.

D4. Add 20 µl proteinase K and 200 µl Buffer AL. Mix by vortexing.

D5. Incubate at 56°C for 30 min and then for a further 15 min at 95°C. Note: Extended incubation at 95°C can lead to some DNA degradation.

D6. Centrifuge for a few seconds.

D7. Follow the "Protocol: DNA Purification from Tissues" from step 6

Protocol: DNA Purification from Tissues
6. Add 200 µl ethanol (96–100 %) to the sample and mix by pulse-vortexing for 15 s. After mixing, briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.
It is essential that the sample, Buffer AL, and the ethanol are mixed thoroughly to yield a homogeneous solution.
A white precipitate may form on addition of ethanol. It is essential to apply all of the precipitate to the QIAamp Mini spin column. This precipitate does not interfere with the QIAamp procedure or with any subsequent application. Do not use alcohols other than ethanol since this may result in reduced yields.

7. Carefully apply the mixture from step 6 (including the precipitate) to the QIAamp Mini spin column (in a 2 ml collection tube) without wetting the rim. Close the cap, and centrifuge at 6000 x g (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate.[17]
Close each spin column to avoid aerosol formation during centrifugation. It is essential to apply all of the precipitate to the QIAamp Mini spin column.

[17] Flow-through contains Buffer AL or Buffer AW1 and is therefore not compatible with bleach. See page 8 for safety information.
Tissues 36 QIAamp DNA Mini and Blood Mini Handbook 04/2010

Centrifugation is performed at 6000 x g (8000 rpm) in order to reduce noise. Centrifugation at full speed will not affect the yield or purity of the DNA. If the solution has not completely passed through the membrane, centrifuge again at a higher speed until all the solution has passed through.

8. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW1 without wetting the rim. Close the cap, and centrifuge at 6000 x g (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the collection tube containing the filtrate.[18]

9. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW2 without wetting the rim. Close the cap and centrifuge at full speed (20,000 x g; 14,000 rpm) for 3 min.

10. Recommended: Place the QIAamp Mini spin column in a new 2 ml collection tube (not provided) and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min.
This step helps to eliminate the chance of possible Buffer AW2 carryover.

11. Place the QIAamp Mini spin column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the filtrate. Carefully open the QIAamp Mini spin column and add 200 µl Buffer AE or distilled water. Incubate at room temperature for 1 min, and then centrifuge at 6000 x g (8000 rpm) for 1 min.

[18] Flow-through contains Buffer AL or Buffer AW1 and is therefore not compatible with bleach. See page 8 for safety information.

Appendix B:

File formats

An AB1 file is a chromatogram of the DNA fragments sequence. The different colours represent the different bases (red = T, green = A, blue = C and black = G).

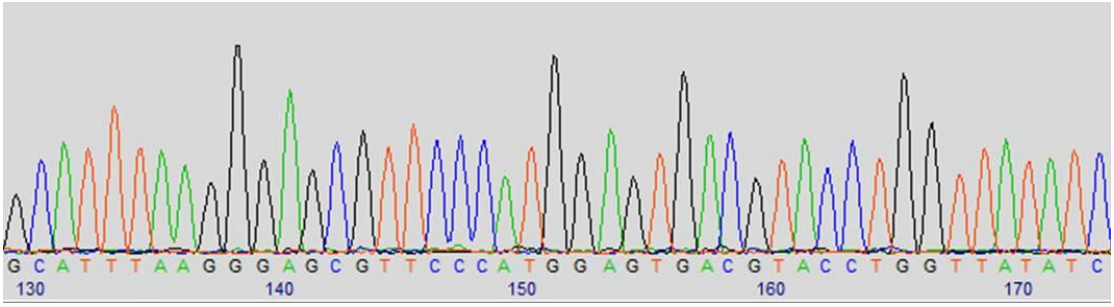


Figure B 1: AB1 file

A FASTA file is a text based format representing the nucleotide or amino acid sequence. It starts with a single line description followed by lines containing the sequence data. In this study nucleotide sequences were used.

>E33L|glpF
GCAGCGTATGCAGTTGGATCAATTAGTGGGGCACATTTGAATCCAGCTTTAACG
ATAGGATTAGCATTTAAGGGAGCGTTCCCATGGAGTGACGTACCTGGCTATATC
GCAGCACAAATGATTGGGGCAATTATCGGGGCAGTTATCGTATATTTACATTACT
TACCACACTGGAAAGAAACAGAAGATCCAGGAACAAAGTTAGGCGTATTTGCAA
CAGGTCCAGCAATTCCGAACACATTTGCAAACCTTTTAAGTGAAATGATTGGGA
CATTCGTTTTAGTATTTGGTATATTAGCAATTGGTGCAAATAAATTTGCAGATGGA
TTAAATCCATTTATCGTAGGTTTCTTAATTGTAAGTATTGGTTTA

Appendix C:

RT-qPCR and VNTR results

In total 112 isolates used in this study were screened with six *B. anthracis* specific markers using RT-qPCR. Two additional markers (BA-*capB(2)* and BA-*lef*) were used if ambiguous results. The first 14 isolates were screened during earlier work (Hovland 2013), and the remaining 98 isolates were screened during this study.

Table C 1: Results from RT-qPCR and VNTR analyses of the 112 *B. cereus* group members.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17	*plcR*-BA[1]	BA-*capB(2)*	BA-*lef*
FFIBCgr10	+	-	-	-	213	-		
FFIBCgr29	+	-	-	-	212	-		
FFIBCgr36	+	-	-	-	221	-		
FFIBCgr37	+	-	-	-	207	-		
FFIBCgr42	+	-	-	-	211	-		
FFIBCgr43	+	-	-	-	209	*		
FFIBCgr44	+	-	-	+	209	-		
FFIBCgr45	+	-	-	-	213	-		
FFIBCgr46	+	-	-	-	213	-		
FFIBCgr47	+	-	-	-	209	*		
FFIBCgr48	+	-	-	-	209	*		
FFIBCgr49	+	-	-	-	208	-		
FFIBCgr50	+	-	-	-	209	*		
FFIBCgr51	+	-	-	-	209	-		
FFIBCgr52	+	-	-	-	-	/		
FFIBCgr53	+	-	-	-	196	/		
FFIBCgr54	+	-	-	-	-	/		
FFIBCgr55	+	-	-	-	-	/		
FFIBCgr56	+	-	-	-	-	/		
FFIBCgr57	+	-	-/+	-	-	/	-	-
FFIBCgr58	+	-/+	-/+	-	-	-	-	-
FFIBCgr59	+	-	-	-	-	-		
FFIBCgr60	+	-/+	-/+	-	195	-	-	-
FFIBCgr61	+	-	-/+	-	208	-	-	-
FFIBCgr62	+	-	-	-	222	-		
FFIBCgr63	+	-/+	-	-	-	-	-	-
FFIBCgr64	+	-	-	-	-	-		

1) + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in one detection channel, / no amplification

Table C 1: Continued from previous page.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17	*plcR*-BA[1]	BA-*capB(2)*	BA-*lef*
FFIBCgr65	+	-/+	-/+	-	195	-	-	-
FFIBCgr66	+	-/+	+/-	-	223	-	-	-
FFIBCgr67	+	-	-	-	-	-		
FFIBCgr68	+	-	-/+	-	-	-	-	-
FFIBCgr69	+	-	-	-	-	-		
FFIBCgr70	+	-	-	-	-	-		
FFIBCgr71	+	-	-	-	-	-		
FFIBCgr72	+	-	-	-	-	/		
FFIBCgr73	+	-	-	-	-	-		
FFIBCgr74	+	-	-	-	-	-		
FFIBCgr75	+	-	-	-	-	-		
FFIBCgr76	+	-	-	-	194	-		
FFIBCgr77	+	-	-	-	192	-		
FFIBCgr78	+	-	-	-	191	-		
FFIBCgr79	+	-	-	-	191	/		
FFIBCgr80	+	-	-	-	192	-		
FFIBCgr81	+	-	-	-	193	-		
FFIBCgr82	+	-	-	-	193	/		
FFIBCgr83	+	-	-	-	194	-		
FFIBCgr84	+	-	-	-	194	/		
FFIBCgr85	+	-	-/+	-	221	-	-	-
FFIBCgr86	+	-	-	-	187	/		
FFIBCgr87	+	-	-	-	220	/		
FFIBCgr88	+	-	-	-	-	/		
FFIBCgr89	+	-	-	-	193	/		
FFIBCgr90	+	-	-	-	219	-		
FFIBCgr91	+	-	-	-	192	/		
FFIBCgr92	+	-	-	-	193	/		
FFIBCgr93	+	-	-	-	221	-		
FFIBCgr94	+	-	-	-	220	-		
FFIBCgr95	+	-	-	-	221	-		
FFIBCgr96	+	-	-	-	207	-		
FFIBCgr97	+	-	-	-	193	-		
FFIBCgr98	+	-	-	-	206	-		
FFIBCgr99	+	-	-	-	218	-		
FFIBCgr100	+	-	-	-	189	-		
FFIBCgr101	+	-	-	-	215	-		
FFIBCgr102	+	-	-	-	-	-		

1) + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in one detection channel, / no amplification

Table C 1: Continued from previous page.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17	*plcR*-BA[1]	BA-*capB(2)*	BA-*lef*
FFIBCgr103	+	-	-	-	214	-		
FFIBCgr104	+	-	-	-	215	-		
FFIBCgr105	+	-	-	-	215	-		
FFIBCgr106	+	-	-	-	-	-		
FFIBCgr107	+	-	-	-	216	-		
FFIBCgr108	+	-	-	-	216	-		
FFIBCgr109	+	-	-	-	215	-		
FFIBCgr110	+	-	-	-	216	-		
FFIBCgr111	+	-	+/-	-	111	-	-	-
FFIBCgr112	+	-	-	-	216	-		
FFIBCgr113	+	+/-	+/-	-	216	-	-	-
FFIBCgr114	+	+	+	+	216	+		
FFIBCgr115	+	+/-	+/-	-	188	-	-/+	-
FFIBCgr116	+	+	+/-	-	217	-	-	-
FFIBCgr117	+	+/-	+/-	-	215	-	-	-
FFIBCgr118	+	-/+	-	-	215	-	-	-
FFIBCgr119	+	-	-	+	207	-		
FFIBCgr120	+	-	-	-	215	-		
FFIBCgr121	+	+	+	-	215	-	+	+
FFIBCgr122	+	-	-	-	189			
FFIBCgr123	+	-	-	-	204	-		
FFIBCgr124	+	-	-	-	-	-		
FFIBCgr125	+	-	-	-	-	-		
FFIBCgr126	+	-	-	-	203	-		
FFIBCgr127	+	-	-	-	203	-		
FFIBCgr128	+	-	-	-	189	-		
FFIBCgr129	+	-	-	-	216	-		
FFIBCgr130	+	-	-	-	189	-		
FFIBCgr131	+	-	-	-	-			
FFIBCgr132	+	-	-	-	187			
FFIBCgr133	+	-	-	-	-	-		
FFIBCgr134	+	-	-	-	-	-		
FFIBCgr135	+	-	-	-	213	*		
FFIBCgr136	+	-	-	-	215	-		
FFIBCgr137	+	-	-	-	-	-		
FFIBCgr138	+	-	-	-	-			
FFIBCgr139	+	-	-	-	215	-		

1) + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in one detection channel, / no amplification

Table C 1 Continued from previous page.

Key	*adk*	BA-*pag*	BA-*capB*	BA5510	Bcms17	*plcR*-BA[1]	BA-*capB(2)*	BA-*lef*
FFIBCgr140	+	-	-	-	215	-		
FFIBCgr141	+	-	-	-	215	-		
FFIBCgr142	+	-	-	-	219	-		
FFIBCgr143	+	-	-	-	188	-		
FFIBCgr144	+	-	-	-	214	*		
FFIBCgr145	+	-	-	-	215	-		
FFIBCgr146	+	-	-	-	206	-		
FFIBCgr147	+	-	-	-	206	-		
FFIBCgr148	+	+/-	+/-	-	-	-	-	-
FFIBCgr149	+	-/+	-/+	-	221	-	-	-

1) + positive results for the *B. anthracis plcR* gene, - negative results for the *B. anthracis plcR* gene, * amplification in one detection channel, / no amplification

Appendix D:

Sampling spots for soil samples collected in Etosha National Park, Namibia

The different *B. cereus* group members isolated during this study connected with corresponding sample ID, carcass site, sample data and sample type information received from Wendy C. Turner (Project leader in Namibia).

Table D 1: Sampling information for the different soil samples and *B. cereus* group members isolated during this study

Key	Sample ID	Carcass Site	Sample date	Sample type
FFIBCgr1	24	10-013	02.04.12	soil
FFIBCgr2	24	10-013	02.04.12	soil
FFIBCgr3	24	10-013	02.04.12	soil
FFIBCgr4	24	10-013	02.04.12	soil
FFIBCgr5	23	10-013	02.04.12	soil
FFIBCgr6	22	10-022	31.03.12	soil
FFIBCgr7	22	10-022	31.03.12	soil
FFIBCgr8	22	10-022	31.03.12	soil
FFIBCgr9	22	10-022	31.03.12	soil
FFIBCgr10	12	12-004	03.04.12	nutrient soil
FFIBCgr11	12	12-004	03.04.12	nutrient soil
FFIBCgr12	12	12-004	03.04.12	nutrient soil
FFIBCgr13	12	12-004	03.04.12	nutrient soil
FFIBCgr14	12	12-004	03.04.12	nutrient soil
FFIBCgr15	10	12-004	06.04.12	nutrient soil
FFIBCgr16	10	12-004	06.04.12	nutrient soil
FFIBCgr17	10	12-004	06.04.12	nutrient soil
FFIBCgr18	10	12-004	06.04.12	nutrient soil
FFIBCgr19	10	12-004	06.04.12	nutrient soil
FFIBCgr20	10	12-004	06.04.12	nutrient soil
FFIBCgr21	10	12-004	06.04.12	nutrient soil
FFIBCgr22	10	12-004	06.04.12	nutrient soil
FFIBCgr23	10	12-004	06.04.12	nutrient soil
FFIBCgr24	10	12-004	06.04.12	nutrient soil
FFIBCgr25	10	12-004	06.04.12	nutrient soil
FFIBCgr26	10	12-004	06.04.12	nutrient soil
FFIBCgr27	10	12-004	06.04.12	nutrient soil
FFIBCgr28	21	10-016	31.03.12	soil
FFIBCgr29	21	10-016	31.03.12	soil
FFIBCgr30	21	10-016	31.03.12	soil

Table D 1: Continued from previous page.

Key	Sample ID	Carcass Site	Sample date	Sample type
FFIBCgr31	21	10-016	31.03.12	soil
FFIBCgr32	21	10-016	31.03.12	soil
FFIBCgr33	13	12-004	03.04.12	soil
FFIBCgr34	14	12-004	06.04.12	soil
FFIBCgr35	14	12-004	06.04.12	soil
FFIBCgr36	13	12-004	06.04.12	soil
FFIBCgr37	13	12-004	03.04.12	soil
FFIBCgr38	11	10-100	10.04.12	soil
FFIBCgr39	11	10-100	10.04.12	soil
FFIBCgr40	11	10-100	10.04.12	soil
FFIBCgr41	14	12-004	06.04.12	soil
FFIBCgr42	3	12-002	07.04.12	soil core
FFIBCgr43	3	12-002	07.04.12	soil core
FFIBCgr44	3	12-002	07.04.12	soil core
FFIBCgr45	3	12-002	07.04.12	soil core
FFIBCgr46	19C	12-004 vacinity`		soil profile
FFIBCgr47	4	12-002	07.04.12	soil core
FFIBCgr48	4	12-002	07.04.12	soil core
FFIBCgr49	4	12-002	07.04.12	soil core
FFIBCgr50	4	12-002	07.04.12	soil core
FFIBCgr51	4	12-002	07.04.12	soil core
FFIBCgr52	3	12-002	07.04.12	soil core
FFIBCgr53	3	12-002	07.04.12	soil core
FFIBCgr54	3	12-002	07.04.12	soil core
FFIBCgr55	3	12-002	07.04.12	soil core
FFIBCgr56	3	12-002	07.04.12	soil core
FFIBCgr57	3	12-002	07.04.12	soil core
FFIBCgr58	3	12-002	07.04.12	soil core
FFIBCgr59	3	12-002	07.04.12	soil core
FFIBCgr60	3	12-002	07.04.12	soil core
FFIBCgr61	3	12-002	07.04.12	soil core
FFIBCgr62	5	12-004	06.04.12	soil core
FFIBCgr63	5	12-004	06.04.12	soil core
FFIBCgr64	5	12-004	06.04.12	soil core
FFIBCgr65	5	12-004	06.04.12	soil core
FFIBCgr66	5	12-004	06.04.12	soil core
FFIBCgr67	19C	12-004 vacinity`		soil profile
FFIBCgr68	19C	12-004 vacinity`		soil profile
FFIBCgr69	19C	12-004 vacinity`		soil profile
FFIBCgr70	19C	12-004 vacinity`		soil profile
FFIBCgr71	19C	12-004 vacinity`		soil profile

Table D 1: Continued from previous page.

Key	Sample ID	Carcass Site	Sample date	Sample type
FFIBCgr72	19C	12-004 vacinity`		soil profile
FFIBCgr73	19C	12-004 vacinity`		soil profile
FFIBCgr74	19C	12-004 vacinity`		soil profile
FFIBCgr75	19C	12-004 vacinity`		soil profile
FFIBCgr76	4	12-002	07.04.12	soil core
FFIBCgr77	4	12-002	07.04.12	soil core
FFIBCgr78	4	12-002	07.04.12	soil core
FFIBCgr79	4	12-002	07.04.12	soil core
FFIBCgr80	4	12-002	07.04.12	soil core
FFIBCgr81	4	12-002	07.04.12	soil core
FFIBCgr82	3	12-002	07.04.12	soil core
FFIBCgr83	3	12-002	07.04.12	soil core
FFIBCgr84	3	12-002	07.04.12	soil core
FFIBCgr85	3	12-002	07.04.12	soil core
FFIBCgr86	4	12-002	07.04.12	soil core
FFIBCgr87	4	12-002	07.04.12	soil core
FFIBCgr88	4	12-002	07.04.12	soil core
FFIBCgr89	4	12-002	07.04.12	soil core
FFIBCgr90	4	12-002	07.04.12	soil core
FFIBCgr91	4	12-002	07.04.12	soil core
FFIBCgr92	4	12-002	07.04.12	soil core
FFIBCgr93	4	12-002	07.04.12	soil core
FFIBCgr94	4	12-002	07.04.12	soil core
FFIBCgr95	4	12-002	07.04.12	soil core
FFIBCgr96	4	12-002	07.04.12	soil core
FFIBCgr97	4	12-002	07.04.12	soil core
FFIBCgr98	4	12-002	07.04.12	soil core
FFIBCgr99	4	12-002	07.04.12	soil core
FFIBCgr100	4	12-002	07.04.12	soil core
FFIBCgr101	5	12-004	06.04.12	soil core
FFIBCgr102	5	12-004	06.04.12	soil core
FFIBCgr103	5	12-004	06.04.12	soil core
FFIBCgr104	5	12-004	06.04.12	soil core
FFIBCgr105	5	12-004	06.04.12	soil core
FFIBCgr106	5	12-004	06.04.12	soil core
FFIBCgr107	5	12-004	06.04.12	soil core
FFIBCgr108	5	12-004	06.04.12	soil core
FFIBCgr109	5	12-004	06.04.12	soil core
FFIBCgr110	5	12-004	06.04.12	soil core
FFIBCgr111	5	12-004	06.04.12	soil core
FFIBCgr112	10	12-004	06.04.12	nutrient soil

Table D 1: Continued from previous page.

Key	Sample ID	Carcass Site	Sample date	Sample type
FFIBCgr113	10	12-004	06.04.12	nutrient soil
FFIBCgr114	10	12-004	06.04.12	nutrient soil
FFIBCgr115	10	12-004	06.04.12	nutrient soil
FFIBCgr116	10	12-004	06.04.12	nutrient soil
FFIBCgr117	10	12-004	06.04.12	nutrient soil
FFIBCgr118	10	12-004	06.04.12	nutrient soil
FFIBCgr119	10	12-004	06.04.12	nutrient soil
FFIBCgr120	10	12-004	06.04.12	nutrient soil
FFIBCgr121	10	12-004	06.04.12	nutrient soil
FFIBCgr122	11	10-100	10.04.12	soil
FFIBCgr123	11	10-100	10.04.12	soil
FFIBCgr124	11	10-100	10.04.12	soil
FFIBCgr125	11	10-100	10.04.12	soil
FFIBCgr126	11	10-100	10.04.12	soil
FFIBCgr127	11	10-100	10.04.12	soil
FFIBCgr128	11	10-100	10.04.12	soil
FFIBCgr129	13	12-004	03.04.12	soil
FFIBCgr130	13	12-004	03.04.12	soil
FFIBCgr131	13	12-004	03.04.12	soil
FFIBCgr132	13	12-004	03.04.12	soil
FFIBCgr133	13	12-004	03.04.12	soil
FFIBCgr134	13	12-004	03.04.12	soil
FFIBCgr135	13	12-004	03.04.12	soil
FFIBCgr136	13	12-004	03.04.12	soil
FFIBCgr137	13	12-004	03.04.12	soil
FFIBCgr138	3			
FFIBCgr139	3			
FFIBCgr140	3			
FFIBCgr141	3			
FFIBCgr142	3			
FFIBCgr143	3			
FFIBCgr144	3			
FFIBCgr145	5			
FFIBCgr146	19B			
FFIBCgr147	3			
FFIBCgr148	17			
FFIBCgr149	19B			

Appendix E:

STs and allelic profiles

All 125 *B. cereus* group members analysed with MLST-3 were assign a unique ST and an allelic profile.

Table E 1: STs and allelic profiles generated for the 125 *B. cereus* group members analysed with MLST-3.

Key	ST	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*
2002734520	1	2	1	1	1	1	1	1	1	4	2	1
97-27	2	1	1	3	2	1	1	2	1	6	2	1
AH1031	3	3	4	4	3	3	2	3	2	7	4	3
AH1129	4	4	7	4	6	6	4	5	3	10	6	5
AH1131	5	5	8	5	7	7	5	6	4	11	7	6
AH1132	6	4	9	4	8	8	6	7	5	12	8	7
AH1135	7	1	10	1	4	9	3	4	6	13	5	1
AH1145	8	6	11	6	9	10	7	8	7	14	9	8
AH1247	9	7	12	7	10	11	8	9	8	15	10	9
AH1270	10	8	13	8	11	12	9	10	9	16	11	10
AH1293	11	4	16	11	12	15	10	11	10	19	14	7
AH187	12	11	17	12	13	16	11	13	11	20	15	13
AH817	12	11	17	12	13	16	11	13	11	20	15	13
AH230	13	13	19	14	15	18	13	15	13	22	16	15
ATCC21282	13	13	19	14	15	18	13	15	13	22	16	15
AH251	14	14	20	15	16	19	14	16	14	23	17	16
AH267	15	15	21	3	17	2	15	17	1	22	2	17
AH403	16	16	22	16	18	20	16	18	15	24	18	18
AH407	17	6	23	6	19	21	17	18	16	25	19	19
AH408	18	17	24	17	20	22	18	19	17	26	20	20
AH519	19	8	25	15	21	23	19	20	18	27	21	21
AH572	19	8	25	15	21	23	19	20	18	27	21	21
AH536	20	8	25	15	21	23	20	20	18	27	21	21
AH542	21	8	25	15	22	23	21	20	19	28	21	21
AH546	22	18	26	18	3	24	22	21	20	29	22	22
AH547	23	6	23	6	23	25	22	18	21	30	19	23
AH553	24	8	27	19	24	26	23	22	18	31	23	21
AH607	25	19	28	15	25	27	24	23	22	27	24	24
AH614	26	17	29	20	26	28	25	24	23	32	25	25
AH627	27	17	24	17	19	29	18	19	23	33	20	20
AH632	28	6	23	6	19	30	26	18	24	25	19	19
AH641	29	20	23	6	27	20	22	25	25	34	26	26
AH645	30	21	30	6	28	31	22	26	25	35	19	26
AH650	31	22	31	21	29	32	27	8	26	36	26	27

Table E 1: Continued from previous page.

Key	ST	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*
AH663	32	6	23	6	19	30	26	18	16	25	19	19
AH664	33	23	32	6	30	20	28	27	27	37	19	28
AH675	34	24	23	6	31	30	26	18	28	25	19	19
AH678	35	23	33	22	32	33	22	28	27	38	27	29
AH681	36	22	34	6	27	10	22	29	7	34	28	8
AH685	37	23	23	6	33	10	29	8	29	39	9	30
AH716	38	4	16	11	12	15	10	30	10	19	14	7
AH889	38	4	16	11	12	15	10	30	10	19	14	7
AH718	39	4	35	11	12	15	10	30	10	19	14	7
AH726	40	4	16	4	12	15	10	31	10	19	14	7
AH727	41	8	37	23	24	35	30	20	30	41	30	31
AH728	42	11	1	24	34	36	31	13	11	42	15	32
AH75	43	13	19	14	15	37	13	15	13	22	16	15
ATCC10987	43	13	19	14	15	37	13	15	13	22	16	15
AH810	44	25	38	25	35	38	32	32	31	42	31	33
AH811	45	4	39	4	8	39	6	7	32	19	32	7
AH812	46	11	1	26	13	16	11	33	33	43	15	34
AH813	47	1	1	1	1	40	33	1	34	44	3	1
AH816	47	1	1	1	1	40	33	1	34	44	3	1
AH818	47	1	1	1	1	40	33	1	34	44	3	1
AH820	47	1	1	1	1	40	33	1	34	44	3	1
AH819	48	11	1	26	13	16	11	33	11	20	15	34
AH823	48	11	1	26	13	16	11	33	11	20	15	34
AH824	48	11	1	26	13	16	11	33	11	20	15	34
AH825	48	11	1	26	13	16	11	33	11	20	15	34
AH826	48	11	1	26	13	16	11	33	11	20	15	34
AH827	48	11	1	26	13	16	11	33	11	20	15	34
AH828	48	11	1	26	13	16	11	33	11	20	15	34
AH829	48	11	1	26	13	16	11	33	11	20	15	34
AH831	49	11	1	26	13	16	11	33	35	20	15	34
AH840	50	11	17	12	13	37	11	13	11	20	15	13
Ames	51	27	2	1	36	2	34	1	36	3	3	35
EPI674	51	27	2	1	36	2	34	1	36	3	3	35
FFIBCgr114	51	27	2	1	36	2	34	1	36	3	3	35
NVH87	51	27	2	1	36	2	34	1	36	3	3	35
NVH93	51	27	2	1	36	2	34	1	36	3	3	35
ATCC14579	52	4	16	11	12	15	10	31	10	45	33	7
ATCC4342	53	12	18	13	14	17	12	14	12	21	34	14
B4264	54	4	9	4	37	39	10	7	5	47	35	7
BGSC4BA1	55	2	1	1	38	1	1	1	37	44	2	1
BGSC4CC1	56	1	3	30	39	48	35	1	34	1	2	1
CI	57	1	1	3	5	46	36	1	38	53	2	38
E33L	58	29	42	2	40	49	37	34	39	54	38	40

Table E 1: Continued from previous page.

Key	ST	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*
FFIBCgr10	59	1	1	1	17	1	38	35	40	48	2	1
FFIBCgr101	60	1	43	1	41	50	39	36	1	2	2	41
FFIBCgr103	60	1	43	1	41	50	39	36	1	2	2	41
FFIBCgr104	61	31	1	31	42	51	40	37	41	55	2	42
FFIBCgr107	62	1	1	1	43	40	36	1	1	44	2	1
FFIBCgr108	63	1	1	2	44	1	41	1	39	2	2	2
FFIBCgr110	63	1	1	2	44	1	41	1	39	2	2	2
FFIBCgr62	63	1	1	2	44	1	41	1	39	2	2	2
FFIBCgr66	63	1	1	2	44	1	41	1	39	2	2	2
FFIBCgr111	64	32	1	32	45	52	42	38	42	56	39	43
FFIBCgr112	65	1	1	1	46	53	43	39	1	1	40	38
FFIBCgr113	66	1	44	1	41	50	39	36	1	2	2	41
FFIBCgr115	67	33	16	4	12	15	10	40	43	19	41	44
FFIBCgr116	68	1	45	1	44	1	44	41	1	2	42	1
FFIBCgr118	68	1	45	1	44	1	44	41	1	2	42	1
FFIBCgr120	68	1	45	1	44	1	44	41	1	2	42	1
FFIBCgr121	68	1	45	1	44	1	44	41	1	2	42	1
FFIBCgr117	69	1	1	2	44	1	41	41	39	2	34	2
FFIBCgr119	70	34	46	33	47	7	5	6	4	11	7	45
FFIBCgr129	71	35	44	1	41	50	39	36	1	44	2	41
FFIBCgr135	72	36	1	34	48	1	45	36	1	52	43	38
FFIBCgr140	73	1	47	3	49	5	45	39	1	1	40	1
FFIBCgr141	74	1	48	1	50	55	46	35	1	44	44	1
FFIBCgr51	74	1	48	1	50	55	46	35	1	44	44	1
FFIBCgr144	75	38	49	3	51	56	1	1	1	44	34	46
FFIBCgr29	77	1	51	1	1	2	47	44	1	44	34	1
FFIBCgr36	78	1	1	1	1	2	1	45	1	1	3	1
FFIBCgr42	79	41	47	3	13	2	48	41	1	2	2	38
FFIBCgr85	79	41	47	3	13	2	48	41	1	2	2	38
FFIBCgr44	80	1	1	1	50	40	36	1	1	44	2	1
FFIBCgr90	80	1	1	1	50	40	36	1	1	44	2	1
FFIBCgr45	81	33	16	4	52	39	10	40	43	19	41	44
FFIBCgr46	82	1	1	1	1	2	35	45	45	44	3	1
FFIBCgr47	83	38	42	3	51	1	49	1	1	1	34	46
FFIBCgr48	84	15	42	3	51	1	50	1	34	44	34	46
FFIBCgr50	85	38	42	3	17	1	49	1	1	1	34	38
FFIBCgr60	86	4	16	11	56	15	10	47	46	19	48	47
FFIBCgr61	87	1	51	1	50	1	52	1	40	44	2	1
FFIBCgr65	88	33	16	4	52	15	10	40	47	19	49	44
FFIBCgr87	89	43	52	37	57	57	40	49	48	59	50	48
FFIBCgr93	90	1	1	1	17	58	46	1	1	44	2	1
FFIBCgr94	91	1	1	3	58	55	47	1	1	9	34	1
FFIBCgr95	92	1	1	1	17	55	46	1	1	44	2	1

Table E 1: Continued from previous page.

Key	ST	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*
FFIBCgr96	93	32	53	32	45	59	36	42	42	56	1	43
FFIBCgr99	94	1	54	3	5	2	36	50	38	1	2	38
G9241	95	44	55	13	14	60	53	51	49	60	51	49
NVH0391/98	96	45	56	38	59	61	54	52	50	61	52	50
R3098/03	97	46	57	39	60	62	55	53	51	62	53	51

Appendix F:

Mutation analyses

Mutation analyses were performed on all *B. cereus* group members clustered in clade I using concatenated sequences and the NJ method (Figure 3.2) (Table F1). Mutation analyses were also performed on isolates that relocated between clades when using concatenated sequences and allelic profiles (Figure 3.3) (Table F2 and F3).

Table F 1: Distribution of mutations compared to *B. anthracis* for the eleven MLST-3 loci for *B. cereus* group members clustered in clade I using concatenated sequences.

Key	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*	Total
Ames	0	0	0	0	0	0	0	0	0	0	0	**0**
NVH93	0	0	0	0	0	0	0	0	0	0	0	**0**
NVH87	0	0	0	0	0	0	0	0	0	0	0	**0**
FFIBCgr114	0	0	0	0	0	0	0	0	0	0	0	**0**
EPI674	0	0	0	0	0	0	0	0	0	0	0	**0**
FFIBCgr36	1	1	0	1	0	2	2	1	1	0	1	**10**
FFIBCgr46	1	1	0	1	0	1	2	3	2	0	1	**12**
FFIBCgr44	1	1	2	3	1	5	0	1	2	1	1	**18**
FFIBCgr90	1	1	2	3	1	5	0	1	2	1	1	**18**
FFIBCgr93	1	1	2	4	2	4	0	1	2	1	1	**19**
FFIBCgr107	1	1	2	4	1	5	0	1	2	1	1	**19**
2002734520	2	1	0	1	1	2	0	1	10	1	0	**19**
FFIBCgr10	1	1	0	4	1	5	1	3	2	1	1	**20**
FFIBCgr95	1	1	0	4	6	4	0	1	2	1	1	**21**
FFIBCgr101	1	3	0	2	1	5	2	1	3	1	2	**21**
FFIBCgr103	1	3	0	2	1	5	2	1	3	1	2	**21**
FFIBCgr113	1	3	0	2	1	5	2	1	3	1	2	**21**
FFIBCgr129	2	3	0	2	1	5	2	1	2	1	2	**21**
FFIBCgr61	1	6	0	3	1	6	0	3	2	1	1	**24**
CI	1	1	2	6	1	5	0	2	2	1	3	**24**
FFIBCgr99	1	3	2	6	0	5	1	2	1	1	3	**25**
FFIBCgr51	1	5	0	3	6	4	1	1	2	2	1	**26**
FFIBCgr141	1	5	0	3	6	4	1	1	2	2	1	**26**
AH267	2	3	2	4	0	5	1	1	7	1	0	**26**
BGSC4BA1	2	1	0	2	1	2	0	14	2	1	1	**26**
AH813	1	1	0	1	1	3	0	18	2	0	1	**28**
AH816	1	1	0	1	1	3	0	18	2	0	1	**28**
AH818	1	1	0	1	1	3	0	18	2	0	1	**28**
AH820	1	1	0	1	1	3	0	18	2	0	1	**28**

FFIBCgr135	4	1	2	5	1	3	2	1	4	2	3	**28**

Table F 1: Continued from previous page.

Key	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*	Total
97-27	1	1	2	6	1	2	1	1	12	1	0	**28**
FFIBCgr29	1	6	0	1	0	7	2	1	2	8	1	**29**
FFIBCgr116	1	2	0	4	1	4	3	1	3	9	1	**29**
FFIBCgr118	1	2	0	4	1	4	3	1	3	9	1	**29**
FFIBCgr120	1	2	0	4	1	4	3	1	3	9	1	**29**
FFIBCgr121	1	2	0	4	1	4	3	1	3	9	1	**29**
FFIBCgr50	5	2	2	4	1	4	0	1	1	8	3	**31**
FFIBCgr112	1	1	0	4	1	4	4	1	1	11	3	**31**
BGSC4CC1	1	2	0	5	1	1	0	18	1	1	1	**31**
FFIBCgr62	1	1	3	4	1	5	0	9	3	1	4	**32**
FFIBCgr66	1	1	3	4	1	5	0	9	3	1	4	**32**
FFIBCgr108	1	1	3	4	1	5	0	9	3	1	4	**32**
FFIBCgr110	1	1	3	4	1	5	0	9	3	1	4	**32**
FFIBCgr42	3	6	2	4	0	7	3	1	3	1	3	**33**
FFIBCgr85	3	6	2	4	0	7	3	1	3	1	3	**33**
FFIBCgr47	5	2	2	7	1	4	0	1	1	8	3	**34**
FFIBCgr140	1	6	2	4	1	3	4	1	1	11	1	**35**
FFIBCgr144	5	3	2	7	2	2	0	1	2	8	3	**35**
FFIBCgr94	1	1	2	6	6	7	0	1	4	8	1	**37**
FFIBCgr117	1	1	3	4	1	5	3	9	3	8	4	**42**
AH1135	1	5	0	8	6	6	4	3	9	2	0	**44**
FFIBCgr48	2	2	2	7	1	3	0	18	2	8	3	**48**
E33L	2	2	3	7	4	4	7	9	5	10	4	**57**
AH187	7	3	3	4	2	19	8	9	7	6	0	**68**
AH831	7	1	4	4	2	19	12	8	7	6	0	**70**
AH728	7	1	4	5	3	17	8	9	11	6	0	**71**
AH819	7	1	4	4	2	19	12	9	7	6	0	**71**
AH823	7	1	4	4	2	19	12	9	7	6	0	**71**
AH824	7	1	4	4	2	19	12	9	7	6	0	**71**
AH825	7	1	4	4	2	19	12	9	7	6	0	**71**
AH826	7	1	4	4	2	19	12	9	7	6	0	**71**
AH827	7	1	4	4	2	19	12	9	7	6	0	**71**
AH828	7	1	4	4	2	19	12	9	7	6	0	**71**
AH829	7	1	4	4	2	19	12	9	7	6	0	**71**
AH75	4	6	4	7	10	14	7	7	7	7	0	**73**
AH230	4	6	4	7	10	14	7	7	7	7	0	**73**
AH812	7	1	4	4	2	19	12	10	8	6	0	**73**
AH817	7	3	3	4	2	19	8	9	7	6	13	**81**
ATCC10987	4	6	4	7	10	14	7	7	7	7	13	**86**
ATCC21282	4	6	4	7	10	14	7	7	7	7	13	**86**

Table F 1: Continued from previous page.

Key	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*	Total
AH840	7	3	3	4	10	19	8	9	7	6	13	**89**
FFIBCgr87	6	7	3	6	12	20	8	11	8	12	8	**101**
AH810	3	7	7	9	12	28	8	12	11	5	0	**102**
G9241	5	8	6	8	5	26	10	11	12	6	12	**109**
ATCC4342	4	6	6	8	10	27	6	11	13	8	16	**115**
FFIBCgr119	9	6	7	7	9	30	10	18	12	16	17	**141**
AH1131	9	11	7	11	9	30	10	18	12	16	18	**151**

Table F 2: Distribution of mutations compared to *B. anthracis* for the MLST-3 loci for *B. cereus* group members relocated from clade III using concatenated sequences to clade I using allelic profiles (Figure 3.3).

Key	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*	Total
FFIBCgr96	21	2	27	33	46	5	22	52	20	8	34	**270**
FFIBCgr104	22	1	27	38	50	20	26	54	27	1	36	**302**
FFIBCgr111	21	1	27	33	49	6	23	52	20	53	34	**319**

Table F 3: Distribution of mutations compared to *B. anthracis* for the MLST-3 loci for B. cereus group members relocated from clade III using concatenated sequences to clade II using allelic profiles (Figure 3.3).

Key	*adk*	*ccpA*	*ftsA*	*glpF*	*glpT*	*panC*	*pta*	*pycA*	*pyrE*	*recF*	*sucC*	Total
AH1247	13	11	9	13	32	46	14	43	15	29	29	**254**
AH408	16	18	16	19	42	33	14	35	13	53	27	**286**
AH614	16	19	17	18	41	34	14	35	12	53	29	**288**
AH627	16	18	16	21	42	33	14	35	9	53	27	**284**