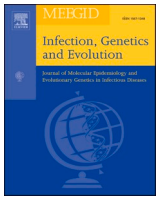
Infection, Genetics and Evolution 123 (2024) 105649

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid



# A genome-based investigation of the *Priestia* species isolated from anthrax endemic regions in Kruger National Park

Thuto Gomolemo Magome [a,*], Sunday Ochonu Ochai [b,c,d], Ayesha Hassim [b], Cornelius Carlos Bezuidenhout [a], Henriette van Heerden [b], Kgaugelo Edward Lekota [a]

[a] Unit for Environmental Sciences and Management, Microbiology, North-West University, Potchefstroom, South Africa
[b] Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa
[c] International Centre for Antimicrobial Resistance Solutions, Copenhagen S, 2300, Denmark
[d] Antimicrobial Research Unit, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

ARTICLE INFO

*Keywords:*
*Priestia*
*Bacillus anthracis*
pan-genomics
Whole genome sequencing
anthrax

ABSTRACT

*Priestia* is a genus that was renamed from the genus *Bacillus* based on the conserved signature indels (CSIs) in protein sequences that separate *Priestia* species from *Bacillus*, with the latter only including species closely related to *B. subtilis* and *B. cereus*. Diagnosis of anthrax, a zoonotic disease, is implicated by tripartite anthrax virulence genes (*lef, pagA*, and *cya*) and poly-γ-D-glutamic acid capsular genes *cap-ABCDE* of *Bacillus anthracis.* Due to the amplification of anthrax virulence genes in *Priestia* isolates, the search for homologous anthrax virulence genes within the *Priestia* genomes (n = 9) isolated from animal blood smears was embarked upon through whole genome sequencing. In silico taxonomic identification of the isolates was conducted using genome taxonomy database (GTDB), average nucleotide identity (ANI), and multi-locus sequence typing (MLST), which identified the genomes as *P. aryabhattai* (n = 5), *P. endophytica* (n = 2) and *P. megaterium* (n = 2). A pan-genome analysis was further conducted on the *Priestia* genomes, including the screening of virulence, antibiotic resistance genes and mobile genetic elements on the sequenced genomes. The oligoribonuclease NrnB protein sequences showed that *Priestia* spp. possess a unique CSI that is absent in other *Bacillus* species. Furthermore, the CSI in *P. endophytica* is unique from other *Priestia* spp. Pan-genomic analysis indicates that *P. endophytica* clusters separately from *P. aryabhattai* and *P. megaterium*. In silico BLASTn genome analysis using the SYBR primers, Taqman probes and primers that target the chromosomal marker (Ba-1), protective antigen (*pagA*), and lethal factor (*lef*) on *B. anthracis*, showed partial binding to *Priestia* regions encoding for hypothetical proteins, pyridoxine biosynthesis, hydrolase, and inhibitory proteins. The antibiotic resistance genes (ARG) profile of *Priestia* spp. showed that the genomes contained no more than two ARGs. This included genes conferring resistance to rifamycin and fosfomycin on *P. endophytica*, as well as clindamycin on *P. aryabhattai* and *P. megaterium*. *Priestia* genomes lacked *B. anthracis* plasmids and consisted of plasmid replicon types with unknown functions. Furthermore, the amplification of *Priestia* strains may result in false positives when qPCR is used to detect the virulence genes of *B. anthracis* in soil, blood smears, and/or environmental samples.

## 1. Introduction

The members of the genus *Priestia* are Gram-positive, aerobic, rod-shaped, endospore forming and non-haemolytic bacteria (Gupta et al., 2020). Species belonging to this genus are found in diverse environments, including soil, inner tissues of cotton plants, sea sediment and the rhizosphere of willow roots (Reva et al., 2002). The genus *Priestia* was reclassified from the genus *Bacillus* following the identification of unique conserved signature indels (CSIs) found in the oligoribonuclease NrnB from the DHH superfamily of proteins that phylogenetically separates the members of *Bacillus* and *Priestia*. Thus, *Bacillus* species not closely related to *B. subtilis* and *B. cereus* were reclassified into the *Priestia* genus (Gupta et al., 2020). The representative taxa of this genus include *P. abyssalis*, formerly *B. abyssalis* (You et al., 2013), *P. aryabhattai*, formerly *B. aryabhattai* (Shivaji et al., 2009), *P. endophytica*, formerly *B. endophyticus* (Reva et al., 2002), *P. filamentosa*, formerly

* Corresponding author.
*E-mail address:* magomethuto22@gmail.com (T.G. Magome).

https://doi.org/10.1016/j.meegid.2024.105649
Received 2 June 2024; Received in revised form 21 July 2024; Accepted 23 July 2024
Available online 25 July 2024

*B. filamentosus* (Sonalkar et al., 2015), *P. flexa*, formerly *B. flexus* (Priest et al., 1988), *P. koreensis*, formerly *B. koreensis* (Lim et al., 2006), *P. megaterium*, formerly *B. megaterium*; (Bary, 1884), *P. paraflexa*, formely *B. paraflexus* (Chandna et al., 2013), *P. qingshengii*, formerly *B. qingshengii* (Xi et al., 2014), *P. taiwanensis*, formerly *B. taiwanensis* (Liu et al., 2015).

Although phylogenetically and taxonomically distinctive, the morphology of *Bacillus* and *Priestia* is similar in that they are rod-shaped bacteria that are Gram-positive (Gupta et al., 2020). Both genera are aerobic, some species such as *B. anthracis and Priestia* species are non-haemolytic and can produce endospores (Stear, 2005). *Bacillus* is a relatively well studied genus with a wide range of species such as *B. cereus*, *B. thuringiensis* and *B. anthracis* from the *B. cereus* group (Alcaraz et al., 2010). *Bacillus anthracis* is a bacterium that causes the zoonotic disease anthrax, which is mainly associated with herbivores and domestic animals. However, there are reported human cases that often involve populations close to infected livestock (Hamutyinei Dhliwayo et al., 2022; World Health Organization and International Office of Epizootics, 2008). *Bacillus anthracis* can be differentiated from *Priestia* species based on non-motility, penicillin and γ-phage sensitive phenotypic characteristics (Stear, 2005). The virulence of *B. anthracis* is attributed to the tripartite toxin genes (protective antigen-*pagA*, lethal factor-*lef*, and edema factor-*cya*) encoded on the pXO1 (182 kb) plasmid (Mock and Mignot, 2003; Turk, 2007) and the poly-γ-D-glutamic acid (PGA), which encodes a five-gene operon *cap-ABCDE* (involved in capsule biosynthesis and anchoring of the PGA) on the pXO2 (95 kb) plasmid (Collier and Young, 2003; Lekota et al., 2018).

Species closely related *to B. anthracis*, termed atypical *B. cereus* strains (G9241, BC-AK, FL2013 and 03BB102), were implicated in human cases presenting with anthrax-like disease in the United States. These strains contained anthrax virulence factors with *B. anthracis* like plasmids pBCXO1 and pBXO2 (Hoffmaster et al., 2006). Moreover, historical bone samples from Western African countries revealed the presence of anthrax-like disease associated with *B. cereus* infection in wildlife animals and livestock and the strains were classified as *B. cereus* biovar *anthracis* strains (CA, CAR and DRC) that were found to carry *B. anthracis*-like plasmids virulence plasmids pBCXO1 and pBCXO2 (Antonation et al., 2016; Baldwin, 2020). This discovery suggested that there may be other *Bacillus* species closely related to *B. anthracis* that may carry anthrax-like virulence plasmids, especially in endemic regions which may be under-reported in routine diagnostic research studies in anthrax endemic regions. Research studies that continue to monitor anthrax in the endemic regions of South Africa include a combination of classical microbiological tests and anthrax molecular genetic markers targeting the pXO1 and pXO2 plasmids and virulence factors from blood, soil and tissue samples. This is in effort to monitor *B. anthracis* and other *Bacillus* species that may potentially harbour anthrax-virulence factors and plasmids (Lekota et al., 2016; Lekota et al., 2018; Ochai et al., 2024).

To date there has been no evidence of any other *Bacillus* species other than *B. anthracis* causing anthrax disease in endemic regions of South Africa. However, isolates belonging to *Bacillus* and *Priestia* genera have indicated the presence of PGA capsular genes similar to those found in *B. anthracis* (Lekota et al., 2016; Lekota et al., 2018). The frequent detection of other *Bacillus* species amplifying anthrax virulence markers during routine diagnosis has been of growing concern, especially when considering the reliability of some diagnostic markers for anthrax virulence genes recommended by World Health Organization (WHO) when testing against other *Bacillus* species from the same isolation source. Although, alternative markers have been suggested for detecting *B. anthracis* in environmental samples (Bassy et al., 2018; Zincke et al., 2020) that offer potential improvements in accuracy and reliability over existing methods. Identification and/or diagnosis that solely depends on molecular screening tests can still be subjective (Hoffmaster et al., 2006; Lekota et al., 2016).

The use of 16S rRNA gene sequencing for the identification of closely related *Bacillus* species is not reliable, especially in whole environmental DNA (Helgason et al., 2000). The housekeeping gene gyrase subunit *B* (*gyrB*) has therefore been used for the identification of certain *Bacillus* species, offering enhanced resolution for distinguishing closely related bacteria (Wang et al., 2007; Larsen et al., 2014; Ochai et al., 2024). Whole genome sequencing (WGS) provides an in-depth analysis by effectively discriminating highly related lineages among bacterial species and subspecies groups (Köser et al., 2012; Wang and Ash, 2015; Quainoo et al., 2017). Moreover, WGS provides higher phylogenetic resolution to differentiate between *Priestia* and *Bacillus* genera, while providing insight into resistance, virulence factors and genetic variation among different species (Khromykh and Solomon, 2015).

In our previous study (Ochai et al., 2024), isolates were obtained from blood smear samples collected from animals during an anthrax outbreak in the anthrax-endemic region of KNP from 2014 to 2016, and the microbiological characteristics were recorded. These isolates were screened for anthrax virulence genes using qPCR with SYBR green and Taqman probe assays. Sequence analysis using the housekeeping gene *gyrB* identified *Priestia* spp. and *B. anthracis* isolates during the same outbreak. In the current study, we employed whole genome sequencing on the *Priestia* spp. and *B. anthracis* isolates obtained by Ochai et al. (2024) using qPCR with anthrax virulence markers for comparative analysis. Additionally, we investigated the genetic diversity of the *Priestia* group to contribute to the research work associated with the genus *Priestia*.

## 2. Materials and methods

### 2.1. Sample collection and screening

This study is based on a culture collection of blood smears (denoted with 'AX') obtained from animal carcasses from an anthrax outbreak between 2014 and 2016 in KNP (Table 1, Supplementary Table S1). The work of Ochai et al. (2024) describes the sample processing and screening in detail. Briefly, the isolates were primarily subjected to classical phenotypic methods that included microscopy, colony morphology and haemolysis, γ-phage and penicillin sensitivity tests as described by the WHO (World Health Organization and International Office of Epizootics, 2008; Ochai et al., 2024). The isolates (denoted with 'AX') were screened for the presence of anthrax-toxin markers (*pagA*, *lef*, and *capB*) as prescribed by World Health Organization and International Office of Epizootics, 2008 and the *B. anthracis* chromosomal marker Ba-1 using SYBR green and Taqman probe qPCR-based method described by Zincke et al. (2020). The results are summarized

**Table 1**
*Bacillus anthracis* and *Priestia* isolates from South Africa isolated from blood smears obtained from animals during an anthrax outbreak from 2014–2016.*

| Isolate/strain number | Date of smear | Host species | Location | Genus species |
|---|---|---|---|---|
| AX2014–1007 | 2014/11/22 | Impala | KNP | *Priestia* |
| AX2014-729BE | 2014/03/15 | Impala | KNP | *Priestia* |
| AX2014–950A | 2014/09/18 | Impala | KNP | *Priestia* |
| AX2015–1136 | 2015//01/26 | Impala | KNP | *B. anthracis* |
| AX2015–1152 | 2015/01/29 | Nyala | KNP | *B. anthracis* |
| AX2015–1270 | 2015/02/13 | Zebra | KNP | *B. anthracis* |
| AX2015-1277A | 2015/02/15 | Buffalo | KNP | *B. anthracis* |
| AX2015-1469A | 2015/07/07 | Buffalo | KNP | *Priestia* |
| AX2015–1511 | 2015/09/06 | Kudu | KNP | *Priestia* |
| AX2015-1523BE | 2015/09/09 | Giraffe | KNP | *Priestia* |
| AX2016-1708NH | 2016/04/05 | Buffalo | KNP | *Priestia* |
| AX2016-1708NH2 | 2016/04/05 | Buffalo | KNP | *Priestia* |
| RL201528 | 2015 | Impala | KNP | *Priestia* |

KNP: Kruger National Park.

* See metadata information for the nine sequenced *Priestia* genomes in Supplementary Table S1

in the Supplementary Table S2 extracted from the study of Ochai et al. (2024).

### 2.2. Genomic extractions and sequencing

Genomic DNA of the *Priestia* isolates (n = 9) and *B. anthracis* (n = 4) were extracted from overnight pure cultures using the Pure link Genomic DNA kit (ThermoFisher Scientific, USA) following the manufacturer's protocol. The concentration and quality of the DNA were determined using the Qubit 2.0 fluorometer (Thermofisher-Scientific, USA). The DNA was sent to the Agricultural Research Council-Biotechnology Platform (ARC-BTP) for whole genome sequencing. Sequence libraries of the isolates were constructed using the MGIEasy FS DNA Prep Kit (BGI, China) according to the manufacturer's protocol. The prepared libraries were sequenced using the BGI MGISEQ-2000 platform (BGI Shenzhen,China) with paired-end 150 nt strategy.

### 2.3. Genome assembly, identification and annotation

The quality of the sequenced reads was assessed using FastQC software 0:10.1 (Andrews et al., 2010). The paired-end trimmed reads were de novo assembled using the MEGAHIT pipeline (Li et al., 2015). The minimum contig length was set to 500 bp, and kmer sizes 21, 33, 55, 77, 99, and 127 were used for the assembly. CheckM (Parks et al., 2015) was additionally used to assess the quality and contamination percentage in individual assembled genomes. Quast v 2.3 (Gurevich et al., 2013) was used to evaluate the draft genome assemblies. In silico identification of the isolates was conducted using the Pub-MLST species-ID search tool (https://pubmlst.org/species-id) (Jolley et al., 2018). The GTDB-Tk v1.7.0 (Chaumeil et al., 2020) which incorporates the Fast Average Nucleotide Identity (Fast ANI) on KBase app (Arkin et al., 2018), was also employed to identify the genomes. Assembled genomes were annotated using Prokka v1.11 (Seemann, 2014) and RASTtk (Aziz et al., 2008).

### 2.4. Conserved signature indel phylogeny

The oligoribonuclease NrnB protein sequences were extracted from *Priestia* (n = 9) and *B. anthracis* (n = 4) annotated genomes sequenced in this study. Multiple sequence alignment was conducted using MAFFT V.7.0 (Katoh et al., 2009). The alignment included *Priestia* reference protein sequences of the oligoribonuclease NrnB as described by Gupta et al. (2020). The alignment also included *B. subtilis* and *B. cereus* group (*B. anthracis, B. cereus, B. thuringiensis* and *B. wiedmannii*) oligoribonuclease NrnB protein sequences for comparison and phylogenetic placement. A phylogenetic tree was constructed using the maximum-likelihood method in MAFFT. The phylogenetic tree was visualized using ITOL v.5 (Letunic and Bork, 2021).

### 2.5. In silico amplification of anthrax-like toxin markers on Priestia genomes

The search for homologous anthrax virulence genes among the sequenced *Priestia* genomes was conducted through a series of BLASTn search in RAST-tk (rast.nmpdr.org) (Aziz et al., 2008; Altschul et al., 1997) using the SYBR green primers and Taqman probe sequences listed in supplementary Table S4. The alignment of the SYBR green primers and Taqman probe sequences on the *Priestia* genomes was obtained to identify the annotated regions on the *Priestia* genomes, which was compared to that of the sequenced *B. anthracis* genomes, including the *B. anthracis* Ames Ancestor reference strain (NCBI Accession: GCF_000007845.1).

### 2.6. Identification of mobile elements, virulence factors, and ARGs

The Mob-Recon tool was used to detect the plasmid replicons from draft genomes in this study (Robertson and Nash, 2018). Proksee-tk was used for characterization of the plasmid replicons detected (Grant et al., 2023). ABRicate pipeline was used to screen for antibiotic resistance and virulence genes using the following databases. Antibiotic resistance determinants were identified in each assembled genome using the ResFinder [−db ResFinder] (Feldgarden et al., 2019; Zankari et al., 2012) with the minimum identity and coverage thresholds of 75 (− minid 75) and 50% (−mincov 50), respectively. Virulence factors in the sequenced genomes were mined using the virulence factor database [−db vfdb] (Chen et al., 2016; Liu et al., 2019), using minimum identity and coverage thresholds of 75% (−minid 75) and 50% (−mincov 50), respectively. The Phage Search Tool with Enhanced Sequence Translation (PHASTEST) server (Zhou et al., 2011; Arndt et al., 2016; Wishart et al., 2023) was used for the rapid identification of prophage sequences in the bacterial genomes.

### 2.7. Pan-genome construction and phylogeny

A pan-genome analysis of *Priestia* spp. was constructed using 234 publicly available genomes (Supplementary Table S3) obtained from GenBank as well as the *Priestia* isolates (n = 9) sequenced in this study. All the *Priestia* (n = 243) were annotated using Prokka v1.11 (Seemann, 2014), and pan-genomic analysis was conducted using Roary v3.6.8 (Page et al., 2015). Similarity searches between the coding domain sequences (CDS) of assembled genomes were conducted using pair-wise BLASTp (Altschul et al., 1997) and Markov Cluster Algorithm (MCL). Clusters were created, paralogs identified, and the isolates were ordered by presence/absence of orthologs for selected strains of *Priestia* including our sequenced genomes (Page et al., 2015). Pan-genome clusters were defined as follows: Core-genes present in all isolates; soft core-genes present in at least 95% of isolates; shell-genes present between 15 and 95% of isolates; cloud-genes in <15% of isolates (Tettelin et al., 2015). Average nucleotide identity analysis with selected *Priestia* spp. (n = 50) was conducted using the Intergrated Prokaryotes Genome and Pan-genome Analysis v1.09 (Liu et al., 2022). Further pan-genome analysis was carried using Anvio-7.1 (Eren et al., 2021). The phylogenetic trees were visualized using ITOL v.5 (Letunic and Bork, 2021).

## 3. Results

### 3.1. Genome identification and classification

The Pub-MLST species identification substantiated with the Genome Taxonomy Database (GTDB) (Fast ANI threshold ≥95%) was used for genome identification. The genomes were identified as follows; *B. anthracis* (AX2015–1136, AX2015–1152, AX2015–1270 and AX2015-1277A), *P. endophytica* (AX2014-729BE and RL201528), *P. megaterium* (AX2014–1007 and AX2015–1511) and *P. aryabhattai* (AX2014–950A, AX2015-1469A, AX2015-1523BE, AX2016-1708NH and AX2016-1708NH2) (Supplementary Table S1). The multiple alignment of the oligoribonuclease NrnB protein sequences showed that all the *Priestia* spp. isolates (n = 9) from this study and available GenBank *Priestia* spp. sequences consist of a single amino acid insertion which Gupta et al. (2020) identified as one of the CSI which that absent from the compared *Bacillus* spp., including the *B. anthracis* strains (AX2015–1136, AX2015–1152, AX2015–1270, and AX2015-1277A) sequenced in this study (Fig. 1A). The CSI on the *P. endophytica* is denoted by 'T' and for *P. aryabhattai*, while the *P. megaterium* strains are denoted by 'K' in the same position (Fig. 1A). Isolates sequenced in this study grouped in respective clades, i.e., AX2014-729BE and RL201528 strains as *P. endophytica*, AX2014–1007, AX2015–1511 strains as *P. megaterium*, AX2014–950A, AX2015-1469A, AX2015-1523BE, AX2016-1708NH, and AX2016-1708NH2 strains as *P. aryabhattai*. The *B. anthracis* strains (AX2015–1136; AX2015–1152; AX2015–1270, and AX2015-1277A) form their own cluster (Fig. 1B).

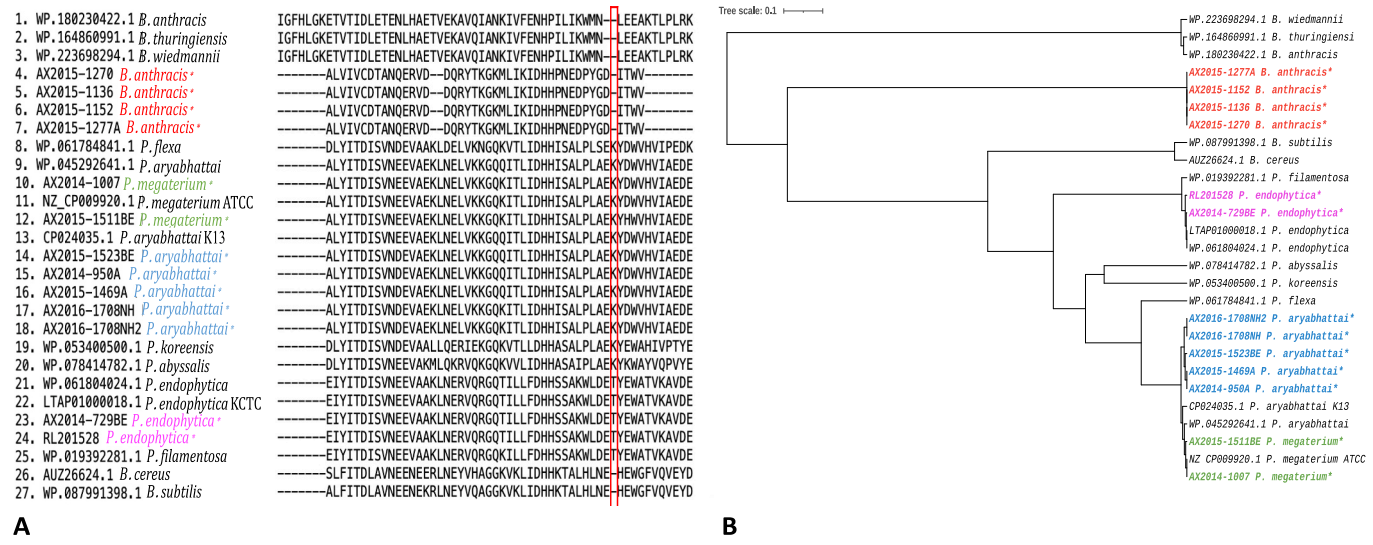



**Fig. 1.** (A) Multiple sequence alignment depicting the CSI portion of the oligoribonuclease NrnB protein sequences from the genomes sequenced in this study (indicated by *) against verified reference protein sequence of *Priestia* spp. and *Bacillus cereus* group (*B. anthracis*, *B. cereus, B. thuringiensis*, and *B. wiedmannii*). The red block shows the singular CSI and differences between the genomes of *Priestia* other *Bacillus* species. (B) Represents the phylogenetic placement using the oligoribonuclease NrnB DHHA-1 domain-contain protein sequences constructed using the maximum-likelihood tree on MAFFT. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### *3.2. Microbiological characteristics and molecular detection of anthrax virulence factors*

The use of classical microbiological methods characterized all the investigated isolates in this study as non-haemolytic. *Bacillus anthracis* isolates exhibited penicillin and gamma-phage sensitivity. *Priestia* isolates were gamma-phage resistant and penicillin sensitive with only *P. megaterium* AX2015–1511 exhibiting penicillin resistance. Based on the qPCR results (Supplementary Table S2), anthrax virulence genes *pagA, lef* and the chromosomal marker Ba-1 were detected in *B. anthracis* strains (AX2015–1136, AX2015–1152, AX2015–1270, and AX2015-1277A). The capsular *capB* gene was found in three *B. anthracis* strains except in the *B. anthracis* AX2015–1136 strain and *Priestia* spp. isolates. The *P. megaterium* strains AX2014-1007, AX2015-1511, as well as the *P. aryabhattai* AX2015-1469A strain were detected using the qPCR method targeting chromosomal marker Ba-1 and anthrax virulence genes (*pagA* and *lef*). *Priestia* megaterium AX2015-1523BE showed positive results for the *lef* and *pagA* genes, whereas *P. aryabhattai* AX2014–950A strain and *P. endophytica* AX2014-729BE strain showed positive results for the *lef* and chromosomal marker Ba-1 genes. The amplification of anthrax virulence *pagA* and *lef* genes in the *Priestia* genomes was further explored using whole genome sequencing. BLASTn sequence alignment of the primers targeting *pagA, lef* and Ba-1 to the *Priestia* genomes was explored using in silico analysis and the amplified genes were further compared on the annotated regions of *B. anthracis* and *Priestia* genomes (Supplementary Fig. S1; Supplementary Table S6-S9).

#### *3.2.1. Anthrax virulence sequence alignment and gene annotations on B. anthracis and Priestia genomes*

The BLASTn results demonstrated that on the sequenced *B. anthracis* genomes, including the reference strain *B. anthracis* Ames Ancestor, all of the searched primer sequences (Supplementary Table S4) targeting *pagA*, *lef, capB*, and chromosomal marker Ba-1 completely aligned to the target regions of the genomes sequences (Supplementary Fig. S1A and Supplementary Table S6-S9). The gene annotation for the target regions in *B. anthracis* isolates were identified as anthrax toxin moiety, protective antigen (*pagA*) (Supplementary Fig. S1B), calmodulin sensitive adenylate cyclase lethal factor (*lef*) on *B. anthracis* pXO1 plasmid (Supplementary Table S8). The chromosomal marker Ba-1 encodes for a hypothetical protein, which suggests that the function of the Ba-1 is unknown (Supplementary Table S9).

The BLASTn results targeting *pagA, lef* and chromosomal marker Ba-1 on *Priestia* genomes exhibited partial alignment of the forward and reverse SYBR and Tagqman probe primers sequences (Supplementary Fig. S1A and Supplementary Table S6-S9). The *pagA* sequence primers were identified as pyridoxine biosynthesis, spore germination, Interlin-like protein, CoA ligase for biotin synthase and hypothetical proteins on *P. aryabhattai* AX2015-1523BE, AX2016-1708NH, AX2015-1469A and *P. megaterium* AX2014–1007 (Supplementary Fig. S1). The *lef* Taqman probe sequence showed a complete alignment to the region identified histone acetyl transferase found in *P. endophytica* strain AX2014-729BE (Supplementary Table S8). In summary, the primer sequences targeting the anthrax virulence factors and the chromosomal marker presented with partial alignments to the *Priestia* genomes. The annotated regions on *Priestia* spp. were different from *B. anthracis* annotated regions suggesting the absence of anthrax-virulence markers *pagA, lef* and the chromosomal marker Ba-1 on *Priestia* genomes.

### *3.3. Identification of mobile genetic elements*

#### *3.3.1. Plasmids replicons detection and features*

Partial plasmids sequences on the sequenced *Priestia* spp. genomes were determined to compare the genetic features of the plasmid replicons against the *B. anthracis* plasmids, pXO1 and pXO2. Genomes identified as *P. aryabhattai* (n = 5) presented a total of seven partial plasmid replicons (Table 2). One plasmid replicon was detected in each isolate of *P. aryabhattai* AX2014–950A (AC355: non-mobilizable) and AX2016-1708NH (NH_8e51f: mobilizable). Typing analysis of plasmid ACC355 showed that it was closely related to the *P. megaterium* strain JX285 plasmid 7 (CP028089) and contained the gene cluster *iol*-ABC-DEGIJOST which is potentially involved in the *myo*-inositol catabolic pathway. *Priestia aryabhattai* strain AX2015-1469A contained two partial plasmids with one mobilizable (B_5d1f9) and one non-mobilizable (A_ce92), which are closely related to the plasmids of *P. megaterium* strain QM B1551 (CP001985) and *P. megaterium* strain YC4-R4 (CP026737). *Priestia aryabhattai* strain AX2015-1523BE presented with three partial plasmid regions, i.e., one mobilizable (A_36e68) and two

**Table 2**
Characterization of the incomplete plasmid sequences detected on the *Bacillus anthracis* and *Priestia* spp. isolates.

| Species | Strain: Plasmid ID | Size (bp) | G + C | Predicted mobility | Mash nearest neighbour | Mash neighbour identification |
|---|---|---|---|---|---|---|
| *B. anthracis* | Ames Ancestor:AA800 (pXO1) | 181,677 | 32.52 | Conjugative | NC_007322 | *B. anthracis* str. Ames Ancestor |
| | Ames Ancestor:AA955 (pXO2) | 94,830 | 33.04 | Conjugative | CP001972 | *B. anthracis* str. A16 |
| *B. anthracis* | AX2015–1136:AA800 | 179,083 | 32.54 | Conjugative | CP009330 | *B. anthracis* |
| *B. anthracis* | AX2015–1152:AA800 | 179,711 | 32.52 | Conjugative | CP009330 | *B. anthracis* |
| | AX2015–1152:AA955 | 94,847 | 33.05 | Conjugative | CP009329 | *B. anthracis* |
| *B. anthracis* | AX2015–1270:AA800 | 177,927 | 32.51 | Conjugative | CP023002 | *B. anthracis* |
| | AX2015–1270c:AA955 | 96,291 | 33.06 | Conjugative | CP009339 | *B. anthracis* |
| *B. anthracis* | AX2015-1277A:AA800 | 179,907 | 32.51 | Conjugative | CP009330 | *B. anthracis* |
| | AX2015-1277A:AA955 | 94,885 | 33.06 | Conjugative | CP009339 | *B. anthracis* |
| *P. aryabhattai* | AX2014–950A:AC355 | 37,346 | 35.03 | Non-mobilizable | CP028089 | *P. megaterium* |
| *P. aryabhattai* | AX2015-1469A:A_ce92 | 26,467 | 35.64 | Non-mobilizable | CP026737 | *P. megaterium* |
| | AX2015-1469A:B_5d1f9 | 7287 | 38.52 | Mobilizable | CP001985 | *P. megaterium* QM B1551 |
| *P. aryabhattai* | AX2015-1523BE:A_36e68 | 18,405 | 35.27 | Mobilizable | NC_010010 | *P. megaterium* QM B1551 |
| | AX2015-1523BE:B_79b79 | 35,943 | 36.66 | Mon-mobilizable | CP028090 | *P. megaterium* |
| | AX2015-1523BE:C_d49e58 | 25,619 | 34.00 | Non-mobilizable | CP018878 | *P. megaterium* |
| *P. aryabhattai* | AX2016-1708NH:NH_8e51f | 8815 | 35.88 | Mobilizable | CP018880 | *P. megaterium* |
| | AX2016-1708NH2 | – | – | – | – | – |
| *P. endophytica* | AX2014-729BE:AC211 | 54,133 | 35.49 | Non-mobilizable | CP015325 | *P. filamentosa* |
| | AX2014-729BE:AB701 | 9160 | 34.90 | Non-mobilizable | CP015329 | *P. filamentosa* |
| *P. endophytica* | RL201528 | – | – | – | – | – |
| *P. megaterium* | AX2014–1007:AD244 | 41,146 | 34.58 | Non-mobilizable | CP035096 | *P. megaterium* NBRC 15308 = ATCC 14582 |
| | AX2014–1007:AF111 | 5990 | 35.07 | Non-mobilizable | CP035099 | *P. megaterium* NBRC 15308 = ATCC 14582 |
| | AX2014–1007:A_ 1db94 | 4977 | 38.49 | Mobilizable | CP035098 | *P. megaterium* NBRC 15308 = ATCC 14581 |
| | AX2014–1007:B_ 25cd8b | 21,699 | 35.90 | Non-mobilizable | CP026737 | *P. megaterium* |
| | AX2014–1007:C_ ec1cb | 3598 | 37.43 | Mobilizable | CP032536 | *P. megaterium* NCT-2 |
| *P. megaterium* | AX2015–1511:AF112 | 9009 | 35.49 | Mobilizable | CP032537 | *P. megaterium* NCT-2 |
| | AX2015–1511:AF497 | 118,143 | 32.54 | Non-mobilizable | CP032531 | *P. megaterium* NCT-2 |
| | AX2015–1511:A_3519c | 37,573 | 36.65 | Mobilizable | CP026737 | *P. megaterium* |
| | AX2015–1511:B_712f3 | 55,060 | 34.70 | Non-mobilizable | CP018876 | *P. megaterium* |

non-mobilizable (B_79b79 and C_d49e58) that are closely related to *P. megaterium* strain QM B1551 plasmid pBM300 (NC_010010), *P. megaterium* strain SGAir0080 plasmid unnamed_7 (CP028090) and *P. megaterium* strain JX285 plasmid 5 (CP018878).

In the *P. megaterium* strain AX2014–1007, two mobilizable (A_1db94 and C_ec1cb) and three non-mobilizable (AD244, AF111 and B_ 25cd8b) plasmid replicons were detected. The small, acid-soluble spore protein (SASP), that was previously used as a *B. anthracis* marker was detected in plasmid AD244 of *P. megaterium* strain AX2014–1007. *Priestia megaterium* strain AX2015–1511 contained four plasmid replicons, i.e., two mobilizable (AF112 and A_3519c) and two non-mobilizable (AF497 and B_712f3). Two non-mobilizable plasmid replicons (AB701 and AC2112) were detected in the *P. endophytica* strain AX2014-729BE. The plasmid replicons are closely related to the plasmids pBEH7 (CP015329) and pBEH3 (CP015325) found in *P. filamentosa* strain Hbe603.

The *B. anthracis* strains (AX2015–1152; AX2015–1270, and AX2015-1277A) each presented with two conjugative plasmid replicons AA800 and AA955 which were identified as *B. anthracis* related plasmids pXO1 (Accession CP009330) and pXO2 (Accession CP009329), respectively (Table 2). Annotated features of plasmid AA800 contained the virulence genes i.e. lethal factor (*lef*) and the protective antigen (*pagA*). The plasmid AA955 that is related to pXO2 contained the gene cluster *capABCD* for capsule biosynthesis. The transition state regulatory protein (*AbrB*) was present in the *B. anthracis* plasmid AA800 and *P. megaterium* strain AX2015–1511 plasmid B_712f3. Tyrosine recombinase gene *XerC* was present in plasmid AA800 of *B. anthracis* strain AX2015–1136 and in plasmid C_d49e58 of *P. aryabhattai* strain AX2015-1523BE, *P. endophytica* AX2014-729BE plasmid AC211 and *P. megaterium* AX2014–1007 plasmid AD244. *Bacillus anthracis* strain AX2015–1136 only contained a conjugative plasmid AA800 that was related to the *B. anthracis* plasmid pXO1 (CP009330). No plasmid replicon types were detected in the draft genome of *P. aryabhattai* AX2016-170NH2 and *P. endophytica* RL201528 strains (Table 2).

### 3.3.2. Determination of prophages in Priestia species and B. anthracis genomes

One fully intact prophage region, namely PHAGE_Bacill_BM5_NC_029069, was detected in two isolates: *P. endophytica* AX2014-729BE and *P. megaterium* AX2015–1511. The prophage regions had sizes of 38.3 kb and 41.8 kb, encoding 49 and 36 proteins, respectively. In *P. aryabhattai* strain AX2014-1469 A, one questionable prophage (PHAGE_Paenib Tripp_NC_028930) region with a length of 26.6 kb encoding 32 proteins was detected.

The prophage regions PHAGE_Staphy_vB_SepS_SEP9_NC_023582, PHAGE_Bacill_phi4B1_NC_028886, PHAGE_Bacill_1_NC_009737 were detected in the *B. anthracis* strains (AX2015–1136, AX2015–1152, AX2015–1270; and AX2015-1277A). Prophage PHAGE_Bacill_1_NC_009737 presented twice in *B. anthracis* strains AX2015–1136, AX2015–1152 and AX2015-1277 A. Prophage PHAGE_Bacill_PfE FR_5_NC_031055 was detected in *B. anthracis* strains AX2015–1270 and AX2015-1277 A. PHAGE_Bacill_vB_BceS_MY192_NC_048633 was the common prophage found intact in the *B. anthracis* AX2015–1136 and AX2015–1152 genomes. PHAGE_Paenib_PG1_NC_021558 was detected and flagged as incomplete in *B. anthracis* strain AX2015–1152. The *B. anthracis* strain AX2015–1270 contained an intact PHAGE_Bacill_phBC6A52_NC_004821. None of the prophage regions detected in *Priestia* genomes were identical to those detected in the *B. anthracis* strains from this study.

### 3.3.3. Determination of antibiotic resistance genes and virulence factors in Priestia species and B. anthracis strains

*Priestia endophytica* strains (AX2014-729BE and RL201528) contained the rifamycin resistance gene *rphB* and fosfomycin resistance *FosB* gene. *FosB* gene was also present in the *B. anthracis* genomes (Fig. 2). The clindamycin resistance gene *lsaB* was present in the *P. aryabhattai* AX2014–950A and *P. megaterium* strains (AX2014–1007 and AX2015–1511). *Priestia megaterium* strain AX2015–1511 contained the extended-spectrum β-lactamase gene *blaTEM-116_7*. All nine sequenced *Priestia* spp. strains contained the *capC* gene associated with the polyglutamate (PGA) synthesis. Three *B. anthracis* strains (AX2015–1152,


**Fig. 2.** Heat map showing the presence and absence of antibiotic resistance genes (ARGs) and virulence genes (VGs) detected in the draft genomes of the sequenced *Bacillus anthracis* and *Priestia* spp. isolates including the reference strain of *B. anthracis* Ames Ancestor (GCF_000007845.1).

AX2015–1270, and AX2015-1277A) presented with all the PGA capsule genes *cap-ABCDE* found in the pXO2 plasmid except for *B. anthracis* strain AX2015–1136 (Fig. 2). This corresponds with the absence of the pXO2 plasmid (AA955) as indicated in Table 2 and *capB* gene using qPCR. Eight *Priestia* genomes excluding *P. endophytica* strain AX2014-729BE and *B. anthracis* genomes contained the ATP-dependent *clpP* subunit. Endopeptidase *clpC* was detected in three *P.* aryabhatai genomes (AX2015-1523BE, AX2016-1708NH and AX2016-1708NH2) and *P. endophytica* strain RL201528 (Fig. 2). *Bacillus anthracis* isolates contained the protective antigen (*pagA*), lethal factor (*lef*) and the oedema factor (*cya*) virulence genes. The non-haemolytic enterotoxins *nhe-ABC*, including the immune inhibitor A metalloprotease *inhA* and the thiol-activated cytolysin *BAS3109* were detected only on the *B. anthracis* genomes.

### 3.4. Phylogenomic and pan-genomic relationship of Priestia species

The *P. aryabhattai* ($n = 47$), *P. endophytica* ($n = 10$) and *P. megaterium* ($n = 177$) retrieved from GenBank and the nine sequenced genomes (Supplementary Table S1) in this study were used to determine the pan-genomic structure of the *Priestia* spp. (Supplementary Fig. S2). The pan-genome was defined by a total of 94,821 genes, of which shell and cloud genes were 6822 and 87,976 respectively. The phylogeny showed that *P. endophytica* isolates grouped distantly from *P. megaterium* and *P. aryabhattai*, except for four strains classified as *P. megaterium* (CH447-14T and S2) and *P. aryabhattai* (PHB10 and GW320) that was in the same cluster as the *P. endophytica* strains (Supplementary Fig. S2). A hierarchical phylogenetic structure comparing ANI showed an ANI score of <75% when comparing these strains to *P. endophytica* strains. Moreover, the strains *P. megaterium* (CH447-14T and S2) and *P. aryabhattai* (PHB10 and GW320) presented with an ANI <75% when compared to other *Priestia* genomes, which indicates an uncommon pattern within *Priesta* species.

*Priestia megaterium* (AX2014–1007 and AX2015–1511) and *P. aryabhattai* (AX2014–950A, AX2015-1469A, AX2015-1523BE, AX2016-1708NH and AX2016-1708NH2) isolates from this study were correctly assigned in their respective clusters and shared an ANI value >90%, with the exception of the four unique strains of *P. megaterium* (CH447-14T and S2) and *P. aryabhattai* (PHB10 and GW320), whereby the ANI value was <76% (Fig. 3; Supplementary Fig. S2). However, gene cluster analysis which includedgenomes classified as *P. aryabhattai* and *P. megaterium* indicated a total of 86,337 genes present between *P. aryabhattai* and *P. megaterium* strains including the sequenced genomes from this study. The gene cluster in Fig. 3B was reconstructed using 47 *Priestia* genomes, defined by a total of 22,059 genes across the group of strains. Triphosphoribosyl-dephospho-CoA synthase, AAA family ATPase, Hsp20/alpha crystal, and the aldehyde dehydrogenase family protein where some of the unique proteins that separate *P. endophytica* from *P. aryabhattai* and *P. megaterium.*

## 4. Discussion

In our previous study, *Priestia* species collected from blood smears obtained during anthrax outbreaks were identified using the *gyrB* marker in Ochai et al. (2024). These species were previously classified under the genus *Bacillus* until the discovery of two CSIs in two protein sequences from the DHH superfamily of proteins that were used to reclassify the species under the genus *Priestia* (Gupta et al., 2020). The isolates were identified as *P. endophytica*; AX2014-729BE and RL201528, *P. megaterium*; AX2014–1007; AX2015–1511; *P. aryabhattai* AX2014–950A, AX2015-1469A, AX2015-1523BE, AX2016-1708NH and AX2016-1708NH2 using GTDB at ANI value ≥95% for genus species. *Bacillus anthracis* isolated from blood smears during the 2014–2016 outbreaks, confirmed with microscopy and PCR were also sequenced

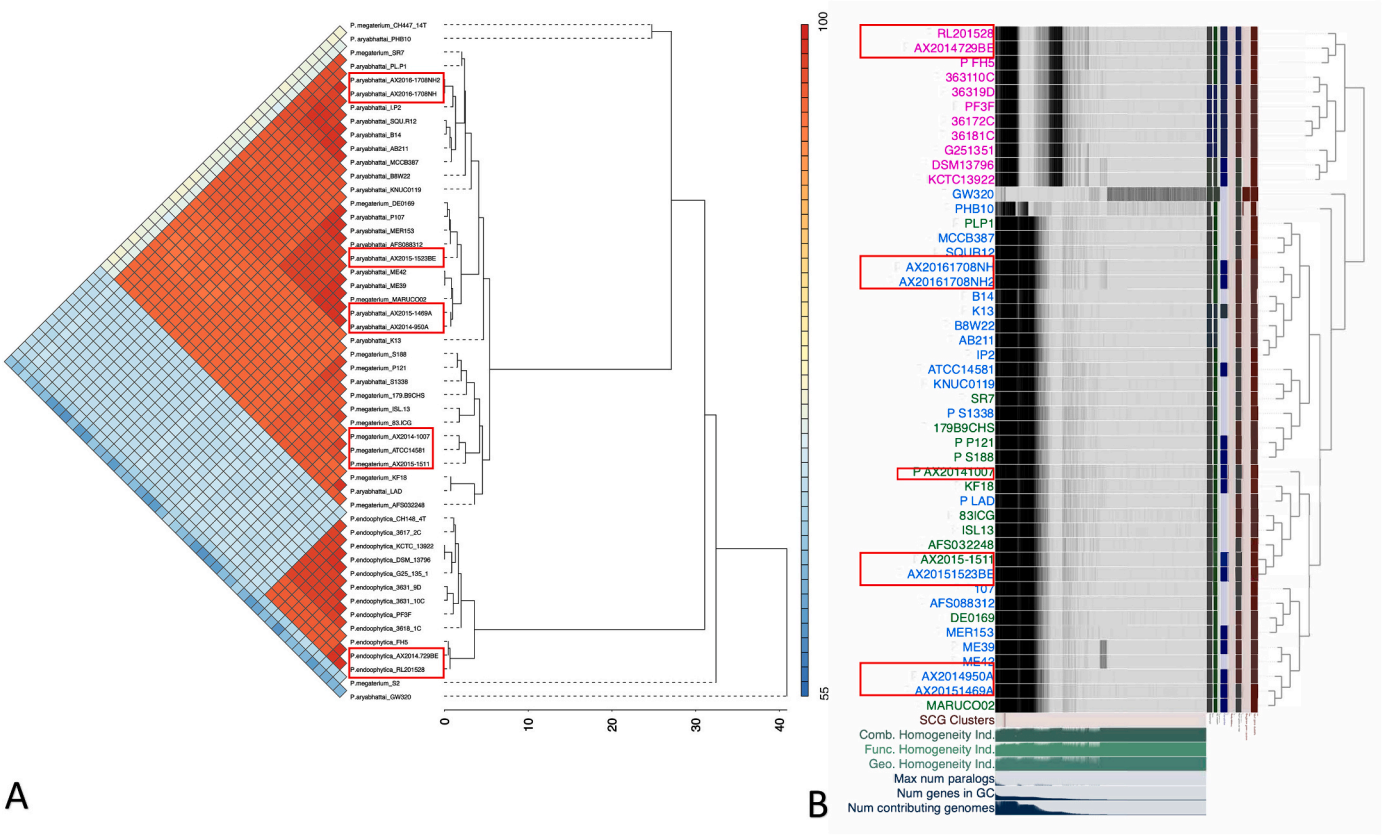



**Fig. 3.** (A) Phylogenetic reconstruction showing average nucleotide identity (ANI) between selected *P. megaterium, P. aryabhattai* and *P. endophytica* that closely grouped together including the sequenced *Priestia* genomes (red box). (B) Phylogeny *Priestia* species indicating the core and accessory genes between selected *Priestia* genomes including the sequenced *Priestia* genomes (red box). Contruction and visualization of the pan-genome analysis was conducted using ANVI'O. Central dendrogram clustering of samples is ordered by gene cluster presence/absence. Items order: presence absence (D, Euclidean; L, Ward). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and compared to the *Priestia* sequences (Ochai et al., 2024). The oligoribonuclease NrnB sequence alignment and phylogeny revealed that *Priestia* species were distinct from the *B. anthracis* and other *B. cereus* group isolates (*B. cereus, B. thuringiensis* and *B. wiedmannii*). This correlates with the study of Gupta et al. (2020), reporting that *Priestia* species contain a single CSI in the oligoribonuclease NrnB protein, which is not present in other *Bacilllus* species (Fig. 1A).

The pan-genome phylogeny of the *Priestia* group revealed a close relation between some species of *P. megaterium* and *P. aryabhattai* (Fig. 2) which included the *P. megaterium* and *P. aryabhattai* isolated from the blood smears by Ochai et al. (2024) and characterized in this study. For example, *P. aryabhattai* AX2014–950A and AX2015-1469A strains grouped with the GenBank genome *P. megaterium* MARUCO02 (GCF_025837055.1) that was isolated from marine mangrove-inhabited sediments. This pattern has been observed in previous studies that investigated *Priestia* genomes using 16S rRNA and *gyrB* for identification (Li et al., 2022). The *P. megaterium* and *P. aryabhattai* genomes were defined by 86,337 core genes. These two species groups are suggestively a single species with variable strains when based on an ANI threshold of > 95% (Maghembe et al., 2023). However, Sam (2023) contends they are a separate species based on whole-genome genetic distance and the difference in orthologous genes that showed divergence in protein sequences between *P. aryabhattai* and *P. megaterium*. Four strains classified as *P. megaterium* (CH447-14T and S2) and *P. aryabhattai* (PHB10 and GW320) presented anomalies when compared to other *Priestia* genomes, particularly strain GW320 (GCF_015832345.1), which was isolated from the rhizospheric soil of cucumbers in the Republic of Korea. It showed a different gene cluster fingerprint from the other genomes with an ANI score <70% when compared to other *Priestia* species. Based on the commonly accepted value of an ANI score of 95%, these strains may need to be revisited. *Priestia endophytica* grouped separately from *P. megaterium* and *P. aryabhattai*, this can be attributed to a difference in the shell and cloud genes (Fig. 3B), including the notable difference in the CSI on the *P. endophytica* which is denoted by 'T' and for *P. aryabhattai* and *P. megaterium* isolates, denoted by 'K' in the same position (Fig. 1A). Although categorized as endophytic bacteria, *P. endophytica* species are considered to be soil plant-endophytic bacteria known for their specific metabolic properties in association with other microorganisms (Jabborova et al., 2022). On the other hand, *P. megaterium* and *P. aryabhattai* reportedly exhibit more plant-growth-promoting abilities (Xu et al., 2022), likely attributed to their large size, which has been reported to contain more coding genes (Chandra et al., 2021; Khalifa and Alsowayeh, 2023).

The common prophage region PHAGE_Bacill_BM5_NC_029069 detected in *P. endophytica* AX2014-729BE and *P. megaterium* AX2015–1511 genomes has been reported in *P. filamentosa* strain Hbe603 (previously identified as *B. endophytica*) and *P. endophytica* 3617_2C (Lekota et al., 2018). Comparative analysis of the prophage regions showed that the *B. anthracis* prophage regions were not similar to those detected in the *Priestia* isolates. The rifampin phosphotransferase protein (*rphB*) was exclusively detected in *P. endophytica* isolates (AX2014-729BE and RL201528). This enzyme is known to confer resistance to rifamycin by deactivating rifampin (Pawlowski et al., 2016). Such resistance occurs due to the mutations in the β-subunit of the RNA polymerase encoded on the *rpoB* gene that was investigated in *B. anthracis, B. cereus* (Vogler et al., 2002) and *Paenibacillus* (Pawlowski et al., 2016). The clindamycin resistance gene *lsaB* was present in the *B. anthracis* and *P. aryabhattai* AX2014–950A and

*P. aryabhattai* (AX2014–1007 and AX2015–1511). Resistance to clindamycin was reported in *P. megaterium* strains (Agersø et al., 2018). The fosfomycin resistance gene is commonly detected in *B. anthracis* and was present, including in the two *P. endophytica* genomes (AX2014-729BE and RL201528) from this study. The presence of ARGs in *Priestia* genomes warrants careful consideration, especially in agricultural settings, as it could serve as a potential pathway for the spread of ARG among bacterial populations (Scaccia et al., 2021). Conversely, certain strains such as the *P. megaterium* strain INA01082 have been investigated for their ability to produce antimicrobial substances (Malanicheva et al., 2012). Exploring antimicrobial activity and the potential risks and benefits in other *Priestia* species may assist in the fight against antibiotic resistance. The *clpP* gene was present in *Priestia* isolates, except in the *P. endophytica* strain AX2015-729BE. This gene encodes for caseinolytic protease proteolytic subunit serine proteases and is advantageous in sustaining the bacteria to survive in stress conditions (Liu et al., 2021). ATP-dependent proteases (*clpC* and *clpP*) are involved in the cellular quality control systems and protein degradation in *B. subtilis* (Krüger et al., 2001).

The presence of *B. anthracis* related toxin genes *capB* and *pagA* in *Priestia* spp. has been previously highlighted in the study of Lekota et al. (2016). Most of the primers prescribed for the capsule region in the pXO2 plasmid (i.e *capB* and *capC*) and those targeting the chromosome (*sasp*, *Ba813* and *rpoB*) of *B. anthracis* have been shown to produce false-positive results with indigenous soil flora and other non *B. anthracis* isolates (Beyer et al., 1999; Ramisse et al., 1999; Ellerbrok et al., 2002; Lekota et al., 2016; Lekota et al., 2018). The protein sequence of the gene *capC* of *P. endophytica* is reportedly 82% similar to that of *B. anthracis* (Lekota et al., 2018). However, it is not unique to *B. anthracis* as it has been detected in other *Bacillus* species, such as *B. subtilis* (Ashiuchi et al., 1999), *B. licheneformis* (McLean et al., 1990), the *B. cereus* group (Klee et al., 2010) as well as in the *Priestia* isolates from the present study. The study of Ochai et al. (2024) reported four typical *B. anthracis* isolates containing the *pagA* and *lef* genes of the pXO1 and only three of the four isolates contained the *capB* gene found on the pXO2. Regions encoding for the anthrax virulence factors, namely the *pagA* and *lef* have been detected on the pXO1-like plasmids in *B. cereus* sensu stricto strains (G9241, FL2013 and 03BB102). These strains reportedly caused anthrax-like disease in humans and/or animals (Klee et al., 2010; Hoffmaster et al., 2006; Baldwin, 2020). However, these atypical strains were more closely related to *B. cereus* than *B. anthracis*, suggesting that anthrax toxin genes may be carried by other *Bacillus* species. None of the *B. anthracis* conjugative plasmids were found in the *Priestia* genomes. The *Priestia* genomes sequenced in this study contained partial plasmids replicons that are mobilizable and non-mobilizable.

*B. anthracis* can be distinguished from other closely related species by standard microbiological methods (non-haemolytic, penicillin and gamma phage sensitivity) and presence of anthrax-toxin genes by PCR (Marston et al., 2006). *Priestia* spp. (previously *Bacillus)* from this study were non-haemolytic, susceptible to penicillin except for *P. megaterium* AX2015–1511. At least, one *B. anthracis* gene marker (*pagA, lef* and Ba-1) was detected through qPCR screening (Ochai et al., 2024). Discrepancies in the diagnosis/identification of *B. anthracis* have been reported by Marston et al. (2006). Relying solely on one standard microbiology method can lead to inaccurate identification, classification and potential virulence of organisms (Carroll et al., 2022; Ochai et al., 2024).

The *pagA* is a crucial component of the edema and lethal toxins that are key virulence factors involved in the severity of anthrax infections (Hoffmaster and Koehler, 1999). Sequencing alignment comparison of the amplified *pagA* regions identified in *Priestia* and other *Bacillus* isolates in the study of Ochai et al. (2024), showed that these regions were conserved. The Taqman probe used to target the *lef* in *B. anthracis* exhibited complete alignment to some of the *Priestia* genomes. However, whole genome comparison of the alignment regions on *Priestia* and *B. anthracis* genomes indicate that these were non-specific bindings on the *Priestia* genome as the annotated regions were identified as a histone acetyltransferase protein and/or hypothetical proteins. The presence of the *B. anthracis* chromosomal Ba-1 marker in *Priestia* isolates may be of concern as the chromosomal region of Ba-1 encodes for a hypothetical protein in *B. anthracis.* In some of the *Priestia* genomes the partial regions were annotated as hypothetical proteins (Zincke et al., 2020). Further investigation on the specificity of the gene marker and knowledge of the function of the Ba-1 region is essential to determine the uniqueness and role of Ba-1 in *B. anthracis*. The length of the primer sequence also affects the results during PCR amplification as they may not provide sufficient coverage, and thus non-specific binding on the isolates is more likely to occur (Ghyselinck et al., 2013). The presence of false positives may provide incorrect identification of *B. anthracis*, especially during active outbreaks, more especially in pooled communityDNA samples. For diagnostic purposes, it is recommended that combining microscopy and molecular markers (i.e Ba-1/*pagA*/*lef*/*capB*) may provide more accurate results for the detection of *B. anthracis* in community DNA when bacterial cultures are not present (Ochai et al., 2024).

In conclusion, the results from this study show the importance of using whole genome sequencing in investigating the genetic variations present in the *Priestia* and *B. anthracis* genomes. False positives when *B. anthracis* is diagnosed without phenotypic tests and only with PCR and/or qPCR indicates the need to determine the essential diagnostic markers for the identification of *B. anthracis*. Through WGS analysis, we were able to show that anthrax virulence genes (*pagA*, and *lef*) and *B. anthracis* anthrax-like plasmids (pXO1 and pXO2) are not present in the genus *Priestia.* We identified various partial plasmid replicons, that have not been identified in the *Priestia* genomes and that require further investigation. The ARGs profile of *Priestia* spp. showed that the isolates contained no more than two antibiotic-resistance genes. Pan-genome analysis, considering the presence and absence of genes and protein structure, revealed differences among *P. endophytica, P. megaterium* and *P. aryabhattai* from this study, despite their closer relationship.

## Funding

This work is based on research supported by the National Research Foundation (NRF) of South Africa, Grant UID 129525. Views expressed are those of the authors and not of the funders.

## CRediT authorship contribution statement

**Thuto Gomolemo Magome:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Sunday Ochonu Ochai:** Writing – review & editing, Methodology, Investigation, Data curation. **Ayesha Hassim:** Writing – review & editing, Supervision, Methodology, Data curation. **Cornelius Carlos Bezuidenhout:** Writing – review & editing, Supervision. **Henriette van Heerden:** Writing – review & editing, Validation, Supervision, Project administration, Conceptualization. **Kgaugelo Edward Lekota:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization, Validation.

## Declaration of competing interest

None.

## Data availability

*Priestia* species SRA sequences were deposited in NCBI GenBank under bioproject PRJNA1111811 and *B. anthracis* were deposited under bioproject PRJNA1112894 with submission numbers SAMN41438185: AX2015-1277A,SAMN41438186: AX2015–1136, SAMN41438187: AX2015–1152, SAMN41438188: AX2015–1270.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2024.105649.

## References


Agersø, Y., Stuer-Lauridsen, B., Bjerre, K., Jensen, M.G., Johansen, E., Bennedsen, M., Brockmann, E., Nielsen, B., 2018. Antimicrobial susceptibility testing and tentative epidemiological cutoff values for five *Bacillus* species relevant for use as animal feed additives or for plant protection. Appl. Environ. Microbiol. 84 (19), e01108–e01118.

Alcaraz, L.D., Moreno-Hagelsieb, G., Eguiarte, L.E., Souza, V., Herrera-Estrella, L., Olmedo, G., 2010. Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. BMC Genomics 11, 1–17.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25 (17), 3389–3402.

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., Wingett, S., 2010. FastQC. A Quality Control Tool for High Throughput Sequence Data, 370.

Antonation, K.S., Grützmacher, K., Dupke, S., Mabon, P., Zimmermann, F., Lankester, F., Peller, T., Feistner, A., Todd, A., Herbinger, I., de Nys, H.M., 2016. *Bacillus cereus* biovar *anthracis* causing anthrax in sub-Saharan Africa—chromosomal monophyly and broad geographic distribution. PLoS Negl. Trop. Dis. 10 (9), e0004923.

Arkin, A.P., Cottingham, R.W., Henry, C.S., Harris, N.L., Stevens, R.L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M.W., 2018. KBase: the United States department of energy systems biology knowledgebase. Nat. Biotechnol. 36 (7), 566–569.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44 (W1), W16–W21.

Ashiuchi, M., Soda, K., Misono, H., 1999. A poly-γ-glutamate synthetic system of *Bacillus subtilis* IFO 3336: gene cloning and biochemical analysis of poly-γ-glutamate produced by *Escherichia coli* clone cells. Biochem. Biophys. Res. Commun. 263 (1), 6–12.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9, 1–15.

Baldwin, V.M., 2020. You can't *B. Cereus*–a review of *Bacillus cereus* strains that cause anthrax-like disease. Front. Microbiol. 11, 530339.

Bary, A., 1884. Vergleichende morphologie und biologie der pilze, mycetozoen und bacterien. Engelmann.

Bassy, O., Jiménez-Mateo, O., Ortega, M.V., Granja, C., Cabria, J.C., 2018. Rapid identification of *Bacillus anthracis* by real-time PCR with dual hybridization probes in environmental swabs. Mol. Cell. Probes 37, 22–27.

Beyer, W., Pocivalsek, S., Öhm, R.B., 1999. Polymerase chain reaction-ELISA to detect Bacillus anthracis from soil samples—limitations of present published primers. J. Appl. Microbiol. 87 (2), 229–236.

Carroll, L.M., Matle, I., Kovac, J., Cheng, R.A., Wiedmann, M., 2022. Laboratory misidentifications resulting from taxonomic changes to *Bacillus cereus* group species, 2018–2022. Emerg. Infect. Dis. 28 (9), 1877.

Chandna, P., Mayilraj, S., Kuhad, R.C., 2013. *Bacillus paraflexus* sp. nov., isolated from compost. Int. J. Syst. Evol. Microbiol. 63 (Pt_12), 4735–4743.

Chandra, A., Chandra, P., Tripathi, P., 2021. Whole genome sequence insight of two plant growth-promoting bacteria (*B. Subtilis* BS87 and *B. Megaterium* BM89) isolated and characterized from sugarcane rhizosphere depicting better crop yield potentiality. Microbiol. Res. 247, 126733.

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2020. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database.

Chen, L., Zheng, D., Liu, B., Yang, J., Jin, Q., 2016. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. Nucleic Acids Res. 44 (D1), D694–D697.

Collier, R.J., Young, J.A., 2003. Anthrax toxin. Annu. Rev. Cell Dev. Biol. 19, 45.

Ellerbrok, H., Nattermann, H., Özel, M., Beutin, L., Appel, B., Pauli, G., 2002. Rapid and sensitive identification of pathogenic and apathogenic *Bacillus anthracis* by real-time PCR. FEMS Microbiol. Lett. 214 (1), 51–59.

Eren, A., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., Fink, I., Pan, J.N., Yousef, M., Fogarty, E.C., Trigodet, F., 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 6 (1), 3–6.

Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Slotta, D.J., Tolstoy, I., Tyson, G.H., Zhao, S., Hsu, C.H., McDermott, P.F., Tadesse, D.A., 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob. Agents Chemother. 63 (11), 10–1128.

Ghyselinck, J., Pfeiffer, S., Heylen, K., Sessitsch, A., De Vos, P., 2013. The effect of primer choice and short read sequences on the outcome of 16S rRNA gene based diversity studies. PLoS One 8 (8), e71360.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.Y., Graham, M., Van Domselaar, G., Stothard, P., 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 51 (W1), W484–W492.

Gupta, R.S., Patel, S., Saini, N., Chen, S., 2020. Robust demarcation of 17 distinct *Bacillus* species clades, proposed as novel *Bacillaceae* genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus *Bacillus* limiting it only to the members of the subtilis and Cereus clades of species. Int. J. Syst. Evol. Microbiol. 70 (11), 5753–5798.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29 (8), 1072–1075.

Hamutyinei Dhliwayo, T., Chonzi, P., Madembo, C., Juru, T.P., Chadambuka, A., Gombe, N.T., Chikerema, S., Shambira, G., Umeokonkwo, C.D., Tshimanga, M., 2022. Anthrax outbreak investigation in Tengwe, Mashonaland West Province, Zimbabwe, 2022. PLoS One 17 (12), e0278537.

Helgason, E., Økstad, O.A., Caugant, D.A., Johansen, H.A., Fouet, A., Mock, M., Hegna, I., Kolstø, A.B., 2000. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*—one species on the basis of genetic evidence. Appl. Environ. Microbiol. 66 (6), 2627–2630.

Hoffmaster, A.R., Koehler, T.M., 1999. Autogenous regulation of the *Bacillus anthracis* Pag operon. J. Bacteriol. 181 (15), 4485–4492.

Hoffmaster, A.R., Hill, K.K., Gee, J.E., Marston, C.K., De, B.K., Popovic, T., Sue, D., Wilkins, P.P., Avashia, S.B., Drumgoole, R., Helma, C.H., 2006. Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely related to *Bacillus anthracis* and harbor *B. Anthracis* virulence genes. J. Clin. Microbiol. 44 (9), 3352–3360.

Jabborova, D., Davranov, K., Jabbarov, Z., Bhowmik, S.N., Ercisli, S., Danish, S., Singh, S., Desouky, S.E., Elazzazy, A.M., Nasif, O., Datta, R., 2022. Dual inoculation of plant growth-promoting *Bacillus endophyticus* and *Funneliformis mosseae* improves plant growth and soil properties in ginger. ACS Omega 7 (39), 34779–34788.

Jolley, K.A., Bray, J.E., Maiden, M.C., 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST. Org website and their applications. Wellcome Open Res. 3.

Katoh, K., Asimenos, G., Toh, H., 2009. Multiple alignment of DNA sequences with MAFFT. Bioinformat. DNA Sequen. Analy. 39–64.

Khalifa, A., Alsowayeh, N., 2023. Whole-genome sequence insight into the plant-growth-promoting bacterium *Priestia filamentosa* strain AZC66 obtained from Zygophyllum coccineum rhizosphere. Plants 12 (10), 1944.

Khromykh, A., Solomon, B.D., 2015. The benefits of whole-genome sequencing now and in the future. Mol. Syndromol. 6 (3), 108–109.

Klee, S.R., Brzuszkiewicz, E.B., Nattermann, H., Brüggemann, H., Dupke, S., Wollherr, A., Franz, T., Pauli, G., Appel, B., Liebl, W., Couacy-Hymann, E., 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.

Köser, C.U., Ellington, M.J., Cartwright, E.J., Gillespie, S.H., Brown, N.M., Farrington, M., Holden, M.T., Dougan, G., Bentley, S.D., Parkhill, J., Peacock, S.J., 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog. 8 (8), e1002824.

Krüger, E., Zühlke, D., Witt, E., Ludwig, H., Hecker, M., 2001. Clp-mediated proteolysis in gram-positive bacteria is autoregulated by the stability of a repressor. EMBO J. 20 (4), 852–863.

Larsen, N., Thorsen, L., Kpikpi, E.N., Stuer-Lauridsen, B., Cantor, M.D., Nielsen, B., Brockmann, E., Derkx, P.M., Jespersen, L., 2014. Characterization of *Bacillus* spp. strains for use as probiotic additives in pig feed. Appl. Microbiol. Biotechnol. 98, 1105–1118.

Lekota, K.E., Hassim, A., Mafofo, J., Rees, J., Muchadeyi, F.C., Van Heerden, H., Madoroba, E., 2016. Polyphasic characterization of *Bacillus* species from anthrax outbreaks in animals from South Africa and Lesotho. J. Infect. Develop. Countries 10 (08), 814–823.

Lekota, K.E., Bezuidt, O.K.I., Mafofo, J., Rees, J., Muchadeyi, F.C., Madoroba, E., Van Heerden, H., 2018. Whole genome sequencing and identification of *Bacillus endophyticus* and *B. Anthracis* isolated from anthrax outbreaks in South Africa. BMC Microbiol. 18 (1), 1–15.

Letunic, I., Bork, P., 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49 (W1), W293–W296.

Li, D., Liu, C.M., Luo, R., Sadakane, K., Lam, T.W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31 (10), 1674–1676.

Li, Q., Hou, Z., Zhou, D., Jia, M., Lu, S., Yu, J., 2022. A plant growth-promoting bacteria *Priestia megaterium* JR48 induces plant resistance to the crucifer black rot via a salicylic acid-dependent signaling pathway. Front. Plant Sci. 13, 1046181.

Lim, J.M., Jeon, C.O., Lee, J.C., Ju, Y.J., Park, D.J., Kim, C.J., 2006. *Bacillus koreensis* sp. nov., a spore-forming bacterium, isolated from the rhizosphere of willow roots in Korea. Int. J. Syst. Evol. Microbiol. 56 (1), 59–63.

Liu, B., Liu, G.H., Sengonca, C., Schumann, P., Wang, M.K., Xiao, R.F., Zheng, X.F., Chen, Z., 2015. *Bacillus taiwanensis* sp. nov., isolated from a soil sample from Taiwan. Int. J. Syst. Evol. Microbiol. 65 (Pt_7), 2078–2084.

Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47 (D1), D687–D692.

Liu, H., Prajapati, V., Prajapati, S., Bais, H., Lu, J., 2021. Comparative genome analysis of *Bacillus amyloliquefaciens* focusing on phylogenomics, functional traits, and prevalence of antimicrobial and virulence genes. Front. Genet. 12, 724217.

Liu, D., Zhang, Y., Fan, G., Sun, D., Zhang, X., Yu, Z., Wang, J., Wu, L., Shi, W., Ma, J., 2022. IPGA: a handy integrated prokaryotes genome and pan-genome analysis web service. iMeta 1 (4), e55.

Maghembe, R.S., Mdoe, F.P., Makaranga, A., Mpemba, J.A., Mark, D., Mlay, C., Moto, E. A., Mtewa, A.G., 2023. Complete genome sequence data of *Priestia megaterium* strain MARUCO02 isolated from marine mangrove-inhabited sediments of the Indian Ocean in the Bagamoyo coast. Data Brief 48, 109119.

Malanicheva, I.A., Kozlov, D.G., Sumarukova, I.G., Efremenkova, O.V., Zenkova, V.A., Katrukha, G.S., Reznikova, M.I., Tarasova, O.D., Sineokii, S.P., El'-Registan, G.I., 2012. Antimicrobial activity of *Bacillus megaterium* strains. Microbiology 81, 178–185.

Marston, C.K., Gee, J.E., Popovic, T., Hoffmaster, A.R., 2006. Molecular approaches to identify and differentiate *Bacillus anthracis* from phenotypically similar *Bacillus* species isolates. BMC Microbiol. 6, 1–7.
McLean, R.J., Beauchemin, D., Clapham, L., Beveridge, T.J., 1990. Metal-binding characteristics of the gamma-glutamyl capsular polymer of *Bacillus licheniformis* ATCC 9945. Appl. Environ. Microbiol. 56 (12), 3671–3677.
Mock, M., Mignot, T., 2003. Anthrax toxins and the host: a story of intimacy. Cell. Microbiol. 5 (1), 15–23.
Ochai, S.O., Hassim, A., Dekker, E.H., Magome, T., Lekota, E.K., Makgabo, S.M., de Klerk-Loris, L.M., van Schalkwyk, L.O., Kamath, P.L., Turner, W.C., van Heerden, H., 2024. Comparing microbiological and molecular diagnostic tools for the surveillance of anthrax. medRxiv pp.2024-04.[pre-print].
Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015. Roary: rapid large-scale prokaryote pan-genome analysis. Bioinformatics 31 (22), 3691–3693.
Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25 (7), 1043–1055.
Pawlowski, A.C., Wang, W., Koteva, K., Barton, H.A., McArthur, A.G., Wright, G.D., 2016. A diverse intrinsic antibiotic resistome from a cave bacterium. Nat. Commun. 7 (1), 13803.
Priest, F.G., Goodfellow, M., Todd, C., 1988. A numerical classification of the genus *Bacillus*. Microbiology 134 (7), 1847–1882.
Quainoo, S., Coolen, J.P., van Hijum, S.A., Huynen, M.A., Melchers, W.J., van Schaik, W., Wertheim, H.F., 2017. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. Clin. Microbiol. Rev. 30 (4), 1015–1063.
Ramisse, V., Patra, G., Vaissaire, J., Mock, M., 1999. The Ba813 chromosomal DNA sequence effectively traces the whole *Bacillus anthracis* community. J. Appl. Microbiol. 87 (2), 224–228.
Reva, O.N., Smirnov, V.V., Pettersson, B., Priest, F.G., 2002. *Bacillus endophyticus* sp. nov., isolated from the inner tissues of cotton plants (*Gossypium* sp.). Int. J. Syst. Evol. Microbiol. 52 (1), 101–107.
Robertson, J., Nash, J.H., 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb. Genom. 4 (8), e000206.
Sam, S., 2023. Genomic and Metagenomic Analysis of the Ongoing Speciation between *Priestia Megaterium* and *Priestia Aryabhattai*. Doctoral thesis,. Nanyang Technology University, Singapore.
Scaccia, N., Vaz-Moreira, I., Manaia, C.M., 2021. The risk of transmitting antibiotic resistance through endophytic bacteria. Trends Plant Sci. 26 (12), 1213–1226.
Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (14), 2068–2069.
Shivaji, S., Chaturvedi, P., Begum, Z., Pindi, P.K., Manorama, R., Padmanaban, D.A., Shouche, Y.S., Pawar, S., Vaishampayan, P., Dutt, C.B.S., Datta, G.N., 2009. *Janibacter hoylei* sp. nov., *Bacillus isronensis* sp. nov. and *Bacillus aryabhattai* sp. nov., isolated from cryotubes used for collecting air from the upper atmosphere. Int. J. Syst. Evol. Microbiol. 59 (12), 2977–2986.
Sonalkar, V.V., Mawlankar, R., Venkata Ramana, V., Joseph, N., Shouche, Y.S., Dastager, S.G., 2015. *Bacillus filamentosus* sp. nov., isolated from sediment sample. Antonie Van Leeuwenhoek 107, 433–441.
Stear, M.J., 2005. OIE manual of diagnostic tests and vaccines for terrestrial animals (mammals, birds and bees) 5th Edn. Volumes 1 & 2. World Organization for Animal Health 2004. ISBN 92 9044 622 6.€ 140. Parasitology 130 (6), 727.
Tettelin, H., Chancey, S., Mitchell, T., Denapaite, D., Schähle, Y., Rieger, M., Hakenbeck, R., 2015. Genomics, genetic variation, and regions of differences. In: Streptococcus pneumoniae. Academic Press, pp. 81–107.
Turk, B.E., 2007. Manipulation of host signalling pathways by anthrax toxins. Biochem. J. 402 (3), 405–417.
Vogler, A.J., Busch, J.D., Percy-Fine, S., Tipton-Hunton, C., Smith, K.L., Keim, P., 2002. Molecular analysis of rifampin resistance in *Bacillus anthracis* and *Bacillus cereus*. Antimicrob. Agents Chemother. 46 (2), 511–513.
Wang, A., Ash, G.J., 2015. Whole genome phylogeny of *Bacillus* by feature frequency profiles (FFP). Sci. Rep. 5 (1), 13644.
Wang, L.T., Lee, F.L., Tai, C.J., Kasai, H., 2007. Comparison of *gyr*B gene sequences, 16S rRNA gene sequences and DNA–DNA hybridization in the *Bacillus subtilis* group. Int. J. Syst. Evol. Microbiol. 57 (8), 1846–1850.
Wishart, D.S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J.R., Stothard, P., Gautam, V., 2023. PHASTEST: faster than PHASTER, better than PHAST. Nucleic Acids Res. 51 (W1), W443–W450.
World Health Organization and International Office of Epizootics, 2008. Anthrax in Humans and Animals. World Health Organization.
Xi, J., He, L.Y., Huang, Z., Sheng, X.F., 2014. *Bacillus qingshengii* sp. nov., a rock-weathering bacterium isolated from weathered rock surface. Int. J. Syst. Evol. Microbiol. 64 (Pt_7), 2473–2479.
Xu, H., Gao, J., Portieles, R., Du, L., Gao, X., Borras-Hidalgo, O., 2022. Endophytic bacterium *Bacillus aryabhattai* induces novel transcriptomic changes to stimulate plant growth. PLoS One 17 (8), e0272500.
You, Z.Q., Li, J., Qin, S., Tian, X.P., Wang, F.Z., Zhang, S., Li, W.J., 2013. *Bacillus abyssalis* sp. nov., isolated from a sediment of the South China Sea. Antonie Van Leeuwenhoek 103, 963–969.
Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67 (11), 2640–2644.
Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S., 2011. PHAST: a fast phage search tool. Nucleic Acids Res. 39 (suppl_2), W347–W352.
Zincke, D., Norris, M.H., Cruz, O., Kurmanov, B., McGraw, W.S., Daegling, D.J., Krigbaum, J., Hoang, T.T.H., Khanipov, K., Golovko, G., Hadfield, T., 2020. TaqMan assays for simultaneous detection of *Bacillus anthracis* and *Bacillus cereus* biovar *anthracis*. Pathogens 9 (12), 1074.