Article

Not peer-reviewed version

Plant-Associated Representatives of the *Bacillus cereus* Group Harbor a Rich Biosynthetic Potential of Antimicrobial Compounds and are Efficient in Suppressing Plant Pathogens

Joachim Vater , Thanh Tam Le Thi , Jennifer Jähne , Stefanie Herfort , Christian Blumenscheit , Andy Schneider , Thi Luong Pham , Phuong Thao Le Thi , Jochen Blom , Silke R. Klee , Thomas Schweder , Peter Lasch , Rainer Borriss *

Posted Date: 10 October 2023

doi: 10.20944/preprints202310.0635.v1

Keywords: Bacillus cereus; phylogenomics; DNA-DNA hybridization (ddH); biocontrol; plant growth promotion (PGP); biosynthesis gene cluster (BGC); kurstakin; thumolycin

Article

Plant-Associated Representatives of the *Bacillus cereus* Group Harbor a Rich Biosynthetic Potential of Antimicrobial Compounds and Are Efficient in Suppressing Plant Pathogens

Joachim Vater [1], Le Thi Thanh Tam [2], Jennifer Jähne [1], Stefanie Herfort [1], Christian Blumenscheit [1], Andy Schneider [1], Pham Thi Luong [2], Le Thi Phuong Thao [2], Jochen Blom [3], Silke Klee [4], Thomas Schweder [5,6], Peter Lasch [1] and Rainer Borriss [5,7,*]

[1] Proteomics and Spectroscopy Unit (ZBS6), Center for Biological Threats and Special Pathogens, Robert Koch Institute, Berlin, Germany
[2] Division of Pathology and Phyto-Immunology, Plant Protection Research Institute (PPRI), Ha Noi, Viet Nam
[3] Bioinformatics and Systems Biology, Justus-Liebig Universität Giessen, Giessen, Germany
[4] Center for Biological Threats and Special Pathogens, Robert Koch Institute, Berlin, Germany
[5] Institute of Marine Biotechnology e.V. (IMaB), Greifswald, Germany
[6] Pharmaceutical Biotechnology, University of Greifswald, Germany
[7] Institute of Biology, Humboldt University Berlin, Berlin, Germany
* Correspondence: Rainer Borriss; rainer.borriss@rz.hu-berlin.de

Abstract: Seventeen bacterial strains able to suppress plant pathogens have been isolated from healthy Vietnamese crop plants, and taxonomically assigned as members of the *Bacillus cereus* group. In order to prove their potential as biocontrol agents, we performed a comprehensive analysis which included whole genome sequencing of selected strains, and mining for genes and gene clusters involved in synthesis of endo- and exotoxins, and secondary metabolites, such as antimicrobial peptides (AMPs). Kurstakin, thumolycin and other AMPs were detected and characterized by different mass spectrometric methods, such as MALDI-TOF-MS, and LIFT-MALDI-TOF/TOF fragment analysis. Based on their whole genome sequences, the plant-associated isolates were assigned to the following species and subspecies: *B. cereus* subsp. *cereus* (6), *B. cereus* subsp. *bombysepticus* (5), *Bacillus tropicus* (2), and *Bacillus pacificus.* Three isolates represented novel genomospecies. Genes encoding entomopathogenic crystal and vegetative proteins were detected in *B. cereus* subsp. *bombysepticus* TK1. *In vitro* assays revealed that many plant-associated isolates enhanced plant growth, and suppressed plant pathogens. Our findings indicated that plant-associated representatives of the *B. cereus* group are a rich source of putative antimicrobial compounds with potential in sustainable agriculture. However, the presence of virulence genes might restrict their application as biologicals in agriculture.

Keywords: Bacillus cereus; phylogenomics; DNA-DNA hybridization (ddH); biocontrol; plant growth promotion (PGP); biosynthesis gene cluster (BGC); kurstakin; thumolycin

1. Introduction

Replacement of harmful chemical pesticides by environmentally friendly biological means is a pressing need in present agriculture worldwide. Microbes, such as bacteria and fungi were proven as being promising candidates for development of efficient agents useful in sustainable agriculture. At present, endospore-forming *Bacillus* spp., and gram-negative *Pseudomonas* spp. are the most used constituents of bioformulations applied in biological plant protection. The main advantage of bioformulations based on *Bacillus* endospores is their longevity, which makes their stability comparable with that of chemical fungicides [1].

During a survey for plant-beneficial bacteria as part of the microbiome of healthy Vietnamese crop plants, a number of Gram-positive, endospore-forming bacteria, able to suppress common plant pathogens, have been isolated from healthy crop plants grown in fields infested with plant pathogens. Based on their draft genome sequences, the isolates were taxonomically assigned as being members of the *Bacillaceae*-family, representing four main taxonomic groups: *Lysinibacillus* spp., *Brevibacillus* spp., the *Bacillus subtilis* species complex, and the *Bacillus cereus* group [2]. Our further studies revealed that by contrast to *Lysinibacillus* sp., the plant-associated Brevibacilli harbored a multitude of interesting antimicrobial peptides with a strong potential to suppress phytopathogenic bacteria, fungi, and nematodes [3]. The *Bacillus velezensis* isolates TL7 and S1, members of the *B. subtilis* species complex, were identified in large scale trials as being the most promising candidates for developing efficient biocontrol agents [4]. In this study we focus on the plant-associated isolates belonging to the *B. cereus* group in order to investigate their potential for biocontrol and plant growth promotion.

The *B. cereus* group, also known as *B. cereus sensu lato* (*s.l.*), comprises a steadily increasing number of species [5] including the human-pathogenic *B. anthracis*, the entomopathogenic *B. thuringiensis*, and the opportunistic pathogen *B. cereus sensu stricto* (*s.s.*). According to Bartlett et al. [6], the three species are established human pathogens. They are closely related, and harbor very similar genome sequences, which not necessarily justify their delineation in different species. Traditionally, they have been discriminated due to properties mainly encoded by extrachromosomal elements.

B. anthracis (risk level 3) has been identified in 1876 by the German physician Robert Koch as the causative agent of anthrax [7], the first disease which has been linked to a microbe. Its virulence based on the ability to form exotoxins and capsules, which is encoded by the plasmids pXO1 and pXO2 [8].

The plasmid encoded production of the highly toxic cereulide is restricted to rare emetic *B. cereus* strains occurring in some foods, whilst production of the diarrheal-inducing enterotoxins (haemolysin BL, HBL; non-hemolytic enterotoxin, NHE; and cytotoxin K, CytK) is common in *B. cereus s.l.* [9]. Due to the extreme stability of the cyclic dodecadepsipeptide celeuride, which is withstanding current food processing techniques, their emetic *B. cereus* producer strains are of particular concern for human health [10].

B. thuringiensis (Bt), isolated 1901 by Ishikawa as "B. sotto" from silkworm, and some years later as *B. thuringiensis* by Berliner from meal moth [11], is an insect pathogen that is successfully used in agriculture as a biopesticide based on the production of diverse crystal toxins, also known as δ-endotoxins [12].

However, the *B. cereus* taxonomy solely based on presence of virulence plasmids with specific function becomes increasingly questionable in light of the recent phylogenomic data. Occurrence of *B. cereus* strains containing pXO1 plasmids [13], and of crystal protein harboring *B. thuringiensis* strains, which are phylogenetically related to *B. anthracis* [14], make the identification of these species to a difficult task. Moreover, occurrence of virulence genes in *B. cereus s.l.* species can´t be excluded. Therefore, *B. cereus s.l.* strains with potential for use in sustainable agriculture can be a risk for public health, and need to be carefully checked for their genomic content, also in case that their taxonomic delineation suggests to be a "safe" species.

In this study, we aimed to elucidate the potential of plant-associated members of the *B. cereus* species complex for biological plant protection. Genome based phylogenetic analyses revealed that most of the isolates were clustered within two subspecies of the *B. cereus s.s.* species. Interestingly, the isolate *B. cereus* TK1 harbored genes encoding two different crystal proteins, and one vegetative insectopathogenic toxin (Vip3). Genome mining for biosynthetic gene clusters probably involved in synthesis of antimicrobial peptides (AMPs), and direct mass-spectrometric investigation of the synthesized AMPs, revealed that the isolates are promising candidates for use in sustainable agriculture. In this context, the lipopeptides kurstakin and thumolycin seem to be of special importance. A special highlight of our research was resolving the primary structure of the plasmid-encoded thumolycin pentapeptide by LIFT-MALDI-TOF/TOF fragment analysis. The inhibiting action of the *B. cereus* s.l. isolates against plant pathogens was corroborated in direct assays performed

 doi:10.20944/preprints202310.0635.v1

with pathogenic ooymycetes, fungi, and nematodes. Finally, plant growth promoting activity of some of the isolates was demonstrated. Regardless of these promising results, we have to consider the risk for public health due to the presence of virulence genes, when the *B. cereus s.l.* isolates are applied as biological means in agriculture.

2. Materials and Methods

2.1. Strain isolation and cultivation

Isolation from Vietnamese healthy crop plants, and purification of the strains was performed as described previously [4]. Cultivation of the bacterial strains and DNA isolation have been previously described [15].

2.2. Reconstruction of complete genomes

The genome sequences of *Bacillus cereus* A22, *B. cereus* A24, *B. cereus* HD1.4B, and *B. cereus* HD2.4 were reconstructed using a combined approach of two sequencing technologies which generated short paired-end reads and long reads. The resulted sequences were then used for hybrid assembly. Short-read sequencing has been previously described [15]. Long read sequencing was done in house with the Oxford Nanopore MinION with the flowcell (R9.4.1) as described previously [3]. The quality of assemblies was assessed by determining the ratio of falsely trimmed protein by using Ideel (https://github.com/phiweger/ideel). Genome coverage of the obtained contigs was 50 x in average. Genome annotation and visualization was performed as described previously [3].

2.3. Screening of the virulence genes

The screening of virulence genes in WGS and complete genomes was performed by using a combined analysis of the PATRIC annotation system [21] and tblastN in the 17 genomes. The most characteristic genes from *B. anthracis* Vollum, including four genes of the pXO1 plasmid (*cya*, *lef*, *pag*A and *rep*X) and six genes of the pXO2 plasmid (*cap*A, *cap*B, *cap*C, *cap*D, *cap*E and *rep*S), were used as the reference sequences. The tblastN threshold for both similarity and coverage was >30%, and all BLAST results were cross-checked against the PATRIC annotation, available at the Bacterial and Viral Bioinformatics Resource Center, BV-BRC, https://www.bv-brc.org/.

The criteria for the presence of virulence plasmids were established as described by Liu et al. [9]. Sequences representing the different types of δ-endotoxins (Suppl. Table S1) were extracted from the NCBI data bank. Searches for presence of genes encoding crystal proteins, and toxins in the 17 plant-associated *B. cereus* genome sequences were performed with tblastN using the respective protein sequences as query.

2.4. Taxonomical phylogeny assessment

Species and subspecies delineation were performed using the Type (Strain) Genome Server (TYGS) platform [4]. Information on nomenclature was provided by the List of Prokaryotic names with Standing in Nomenclature (LPSN, available at https://lpsn.dsmz.de [17]. The EDGAR3.0 pipeline [18] was used for elucidating taxonomic relationships as described previously [4].

2.5. Genome mining

In silico prediction of gene clusters involved in secondary metabolite synthesis was performed using the antiSMASH pipeline version 6 [19], the bioinformatic tool described by Bachmann and Ravel [20], and BAGEL4 [21].

2.6. Sample preparation and mass-spectrometric detection of the bioactive peptides

Bioactive compounds of the investigated *B. cereus s.l.* strains were detected and identified by MALDI-TOF MS, as outlined previously [22,23]. A Bruker Autoflex Speed TOF/TOF mass spectrometer (Bruker Daltonics; Bremen, Germany) was used with Smartbeam laser technology

 doi:10.20944/preprints202310.0635.v1

applying a 1 kHz frequency-triple Nd-YAG-laser (λ_{ex} = 355 nm). Samples (2 µl) of surface extracts and culture supernatants were mixed with 2 µl matrix solution (a saturated solution of α-hydroxy-cinnamic acid in 50% aqueous ACN containing 0.1 % TFA) spotted on the target, air dried and measured. Mass spectra were obtained by positive-ion detection in reflector mode. Monoisotopic masses were observed. Parent ions were detected with a resolution of 10.000. Sequence analysis of peptide products was performed by MALDI-LIFT-TOF/TOF mass spectrometry in laser induction decay (LID) mode [24]. The product ions in the LIFT-TOF/TOF fragment spectra were obtained with a resolution of 1000.

2.7. Antibacterial, antifungal, nematocidal, and plant growth-promoting activity assays

Assays for activity against plant pathogens (oomycetes, fungi, neatodes) were performed as previously described [4]. The root-knot nematode *Meloidogyne* sp. was isolated from roots of infested pepper plants according to Hooper et al. [25]. Tomato plantlets were grown in pots with natural soil under controlled conditions in greenhouse. Test bacteria and second stage juveniles (J2) nematodes were added to the pots two weeks after transplanting. Ten weeks after infesting with the nematodes the number of knots in tomato plants was estimated [26].

Plant growth promotion assays were performed with *Arabidopsis thaliana* seedlings as described previously [27].

2.8. Data analysis

The data obtained from biocontrol and plant growth promotion experiments were analyzed using one-factorial analysis of variance (ANOVA). Mean values were calculated from the results of the replicates (n≥3). The Fisher´s least significant difference (LSD) test was conducted as *post-hoc* test for estimating significant differences ($p \leq 0.05$) between the mean values as described previously [4].

2.9. Gene bank accession numbers of complete genome sequences

Bacillus cereus A22 chromosome: CP085498.1, *Bacillus cereus* A22 plasmid P1: CP085499.1, *Bacillus cereus* A22 plasmid P2: CP085500.1, *Bacillus cereus* A24 chromosome: CP085501.1, *Bacillus cereus* A24 plasmid P1: CP085502.1, *Bacillus cereus* A24 plasmid P2: CP085503.1, *Bacillus* cereus HD1.4B chromosome: CP0855510.1, *Bacillus cereus* HD1.4B plasmid P1: CP085511.1, *Bacillus cereus* HD1.4B plasmid P2: CP085512.1, *Bacillus cereus* HD1.4B plasmid P3: CP085513.1, *Bacillus cereus* HD2.4 chromosome: CP0855506.1, *Bacillus cereus* HD2.4 plasmid P1: CP085507.1, *Bacillus cereus* HD2.4 plasmid P2: CP085508.1, *Bacillus cereus* HD2.4 plasmid P3: CP085509.1.

3. Results and Discussion

3.1. Comparative genome analysis of the isolates from Vietnamese crop plants representing the Bacillus cereus s.l. complex

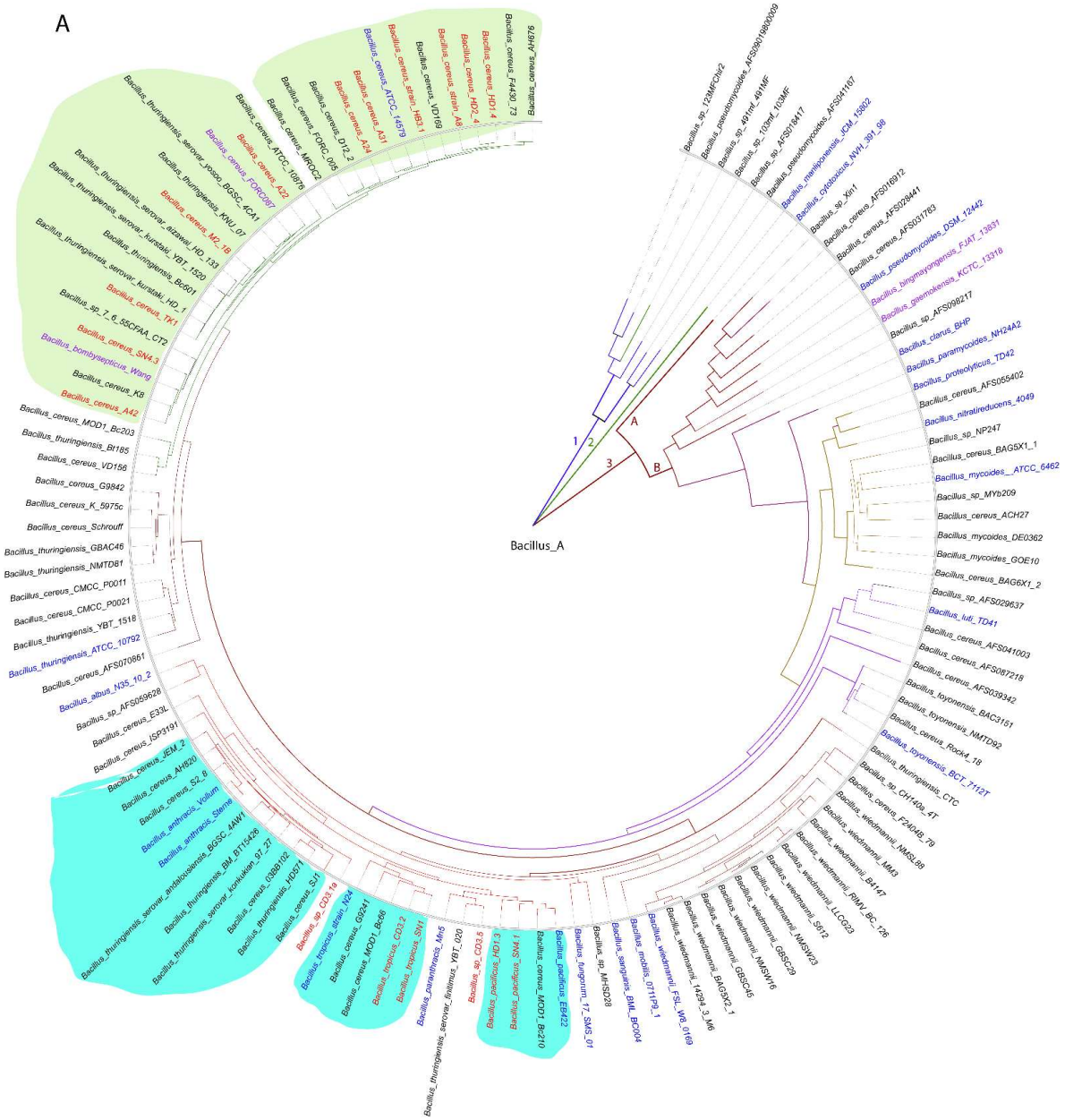
3.1.1. Genome-based species and subspecies delineation of the plant-associated isolates belonging to the B. cereus group

Seventeen of the endospore-forming bacterial strains, isolated from Vietnamese crop plants have been previously assigned to the *Bacillus cereus sensu lato* group [2]. All isolates displayed the typical features of *B. cereus*: they developed phospholipase C and hemolytic activity. The phylogenetic tree obtained from the 16S rRNA sequences supported their previous taxonomic assignment as members of the *B. cereus s.l.* group, but possessed an average branch support of only 28.5 % (Suppl. Figure S1), which is not sufficient for robust species delineation.

We used a whole-genome based approach for robust delineating of the taxonomic position of the plant-associated *B. cereus s.l.* isolates. The phylogenomic tree containing a total of 128 *B. cereus s.l.* genomes, mainly extracted from the NCBI data bank, yielded three main branches (1,2,3). Branch 3 was subdivided into the clusters 3A and 3B. All of our isolates were distributed within the cluster 3B,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

and found related to the clusters formed by the type strains of *B. cereus*, *B. anthracis*, *B. tropicus*, and *B. pacificus* (Figure 1A).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

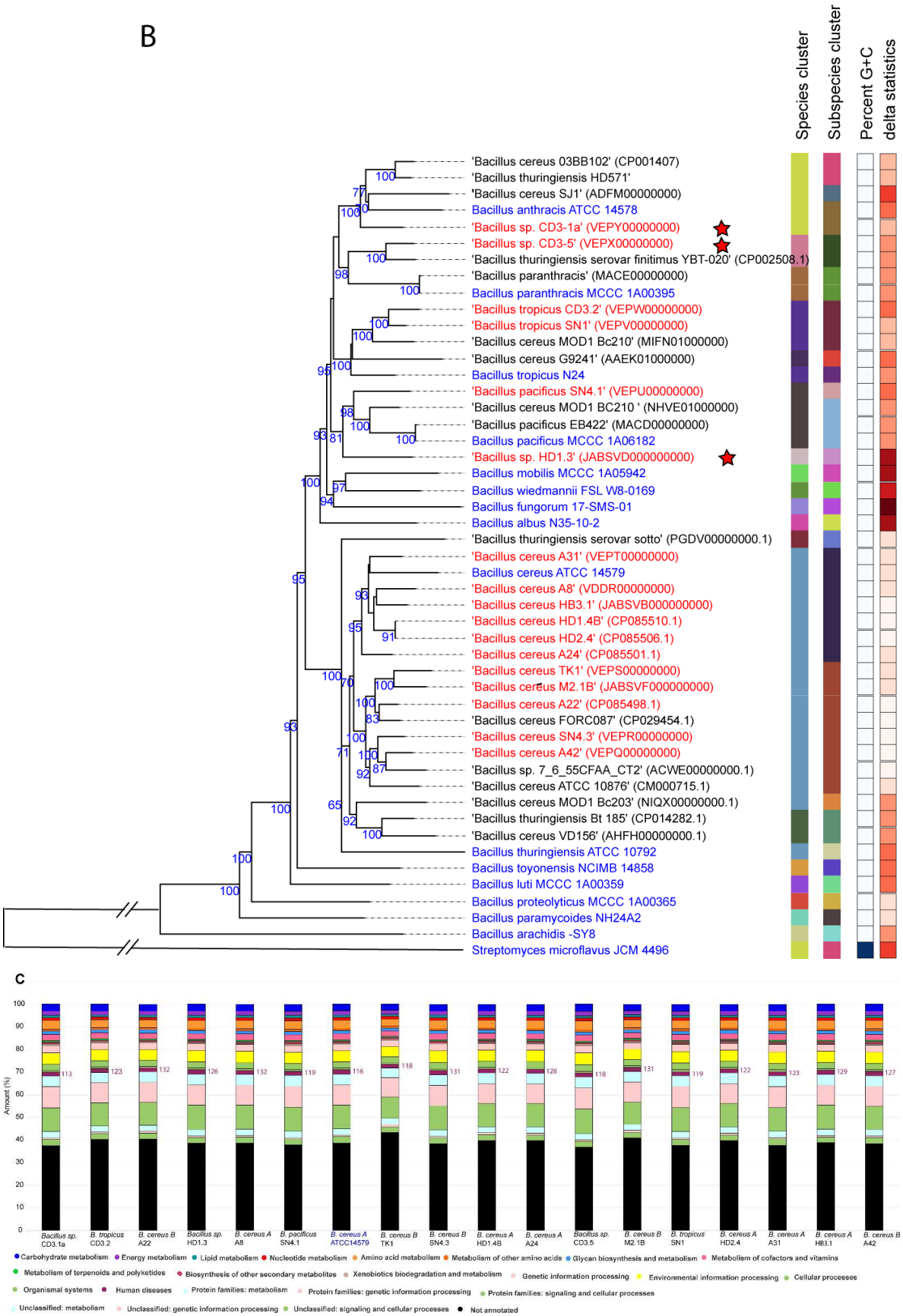


Figure 1. A: Approximately-maximum likelihood phylogenetic tree for 128 *Bacillus cereus* group genomes, calculated by EDGAR using the Fast tree software (http://www.microbesonline.org/fasttree/). The unrooted tree was built out of a core of 1054 genes per genome. The core has 280,349 AA residues/bp per genome, 35,884,672 in total. Bacillus_A is a term

 doi:10.20944/preprints202310.0635.v1

used from GTDB (https://gtdb.ecogenomic.org) for genomospecies belonging to the *B. cereus* group. The tree is divided into three main branches (1-3). Cluster 3 is further divided into the subclusters 3A and 3B. The 21 type strains are labelled in blue letters. The isolates, investigated in this study, are indicated by red letters and belong all to the subcluster 3B. Related subclusters containing the plant-associated Vietnamese isolates are marked by unregular colored fields. The ANI values are documented in Suppl. Figure S3. **B**: *Bacillus cereus* tree inferred with FastMe 2.1.6.1 [29] from GBDP distances calculated from whole-proteome data using the Type (Strain) Genome Server TYGS (https://tygs.dsmz.de). Analysis was performed using both, Maximum Likelihood and Maximum Parsimony, with 16 type strains (blue letters) and 17 genome sequences obtained from the *Bacillus cereus* strains isolated from Vietnamese crop plants (red letters). In addition, 16 strains with similar proteomes obtained from the NCBI data bank, were included yielding a total of 49 proteomes. The branch lengths are scaled in terms of GBDP distance formula d_5. Putative novel genomospecies were indicated by red stars. The numbers above branches are GBDP pseudo-bootstrap support values > 60% from replications, with an average branch support of 87.5 %. The first two colored columns to the right of each name refer to the genome-based species and subspecies clusters, specified by dDDH cut-off values of 70% and 79%, respectively. **C**: Functional KEGG category analysis of plant-associated *B. cereus* group isolates. Type strain *B. cereus* ATCC 14579 was included in the analysis. The number of genes associated with human diseases is indicated. 125216 KEGG functional categories (including not annotated sequences) were found in the selected 18 contigs.

Similar, but more detailed results were obtained when using the Type (Strain) Genome server TYGS [28]. Our survey resulted in assigning of six species and seven subspecies clusters (Figure 1B, Suppl. Figure S2). 15 of the isolates were assigned to four valid species, *B. cereus* (11), *Bacillus pacificus* (1), *Bacillus tropicus* (2), and *Bacillus anthracis* (1). In case of *B. cereus*, the dDDH values obtained after comparison of 11 isolates with the type strain ATCC14579 exceeded the species cut off (>70%, Suppl. Table S2). On the genomic level, two subclusters were distinguished: Six isolates, yielding dDDH values above the subspecies cutoff (>79%), represented the subspecies ´A´ (*B. cereus* subsp. *cereus*), whilst five isolates showed dDDH values ranging from 72% to 74%, when compared with ATCC14579. The latter cluster formed together with *B. cereus* FORC087.1 a second subcluster ´B´, which was clearly distinguished from subcluster ´A´ (Suppl. Table S2). When the genomes of the members of the *B. cereus* subcluster ´B´ were compared with the genome of *Bacillus bombysepticus* Wang [30], their dDDH values exceeded the subspecies cutoff (>79%, Suppl. Table S2). The direct comparison of FORC087 with *B. bombysepticus* Wang yielded a dDDH value of 87.8%, suggesting their close relationship. In the Genome Taxonomy Database, GTDB [31], the genomospecies *Bacillus_A bombysepticus* harbored 667 members, whilst the *B. cereus* genomospecies represented by *B. cereus* ATCC 14579 harbored only 310 genomes (GTDB release 08-RS214, 28th April 2023).

Although *B. bombysepticus* is still not listed as valid species in the List of Prokaryotic names with Standing in Nomenclature, LPSN [32], we propose to designate the *B. cereus* subcluster ´B´ as genomosubspecies *B. cereus* subsp. *bombysepticus*, taking into account that the members of the ´*bombysepticus* group` shared dDDH values above the species cutoff with *B. cereus* ATCC 14579.

The genomes of two other isolates, SN1 and CD3-2, were assigned, according to their dDDH and Fast ANI values, to *Bacillus tropicus*. However, when compared with the *B. tropicus* type strain N24, their dDDH and ANI values were below the subspecies cutoff indicating that these isolates form together with *B. cereus* MOD1 Bc210 the subcluster ´B´, distinct from the *B. tropicus* type strain (Figure 1B, Suppl. Table S2).

The isolate *Bacillus* sp. CD3-1a was found clustered together with the *B. anthracis* type strains. However, this species delineation appeared to be questionable, since we did not detect the *B. anthracis* virulence plasmids pXO1 and pXO2 in the draft genome of CD3-1a (see next section).

Two isolates, *Bacillus* sp. CD3-5, and *Bacillus sp.* HD1.3, although distantly related to *B. tropicus* and *B. pacificus*, could not been assigned to any species present in the TYGS database (17-04-2023), and might represent novel genomospecies (Figure 1B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

3.1.2. Occurrence of virulence genes might restrict application of B. cereus s.l. isolates

Functional KEGG analysis revealed presence of genes possibly involved in human disease in the genomes of all plant-associated *B. cereus s.l.* isolates (Figure 1C).

Within the *B. cereus* group, occurrence of virulence factors, which are closely linked to disease symptoms [33,34], and of entomopathogenic Cry toxins [35] have been reported. In past, these elements have been widely applied to assign *B. anthracis*, *B. cereus* and *B. thuringiensis*.

Since the genome sequence of the isolate CD3-1a formed a cluster together with the *B. anthracis* type strain ATCC 14578 (Figure 1), we checked the CD3-1a draft genome for the presence of sequences of the characteristic anthrax toxin plasmids pXO1 and pXO2. No sequences similar to the genes encoding the Rep proteins RepX (pXO1) and RepS (pXO2) were detected in CD3-1a. Moreover, no sequences exhibiting significant similarity with the anthrax genes of pXO1 (*cya*, *pagA*, *lef*) and pXO2 (*capABCDE*) were found in CD3-1a, and in the other plant-associated *B. cereus s.l.* isolates, excluding their taxonomic delineation as representative of the human-pathogenic *B. anthracis* species (Suppl. Table S3).

To distinguish "true" *B. anthracis* isolates from non-anthrax-causing representatives of the *B. cereus* group, detecting of the dhp chromosomal marker sequences, which are indicating the presence of *B. anthracis* specific prophages, by multiplex PCR assays [36] was performed. None of the *B. anthracis* specific dhp fragments could be amplified when using chromosomal DNA prepared from CD3-1a. In addition, amplification of the protective antigen *pagA* gene, and of the capsule *capB* gene using CD3-1a DNA was not achieved. By contrast to *B. anthracis*, but similar to the other isolates, CD3-1a was hemolytic when cultivated on blood agar, and the genes encoding the hemolysin BL (HBL) toxin were present onto the chromosome (Suppl. Table S3).

Interestingly, the isolate *B. pacificus* SN4.1 harbored a pXO1-like *repX* gene, and the isolate *B. tropicus* CD3.2 harbored a sequence resembling to the pXO2-like *repS* gene, in their draft genomes (Figure 2, Suppl. Table S3) suggesting that the *rep* genes characteristic for the pOX plasmids can occur in other members of the *B. cereus s.l.* species complex. This is in line with previous findings of Liu et al. [9].

Next, we proved the *B. cereus s.l.* isolates for the presence of other virulence genes involved in production of toxins responsible for foodborne diseases of human beings. Celeurin, the causative agent of the emetic syndrome [37], is known to be non-ribosomally synthesized by giant peptide synthetases encoded by the *ces* gene cluster. None of our isolates harbored this gene cluster (Figure 2), suggesting that the plant-associated isolates did not represent emetic *B. cereus* strains.

By contrast, the HBL/NHE enterotoxin operons encoding the non-hemolytic enterotoxin A (NHE), and the haemolysin component BL (HBL) [38], occurred in nearly all isolates with one exception. *B. pacificus* SN4.1 harbored the genes responsible for synthesis of the NHE enterotoxin, but not the genes for hemolysin synthesis (Suppl. Table S3). HBL and NHE are the causative agents of the diarrhoeal syndrome in human beings, which is caused by ingestion of vegetative cells and spores that produce enterotoxins in the small intestine [39].

Due to these findings, we can not exclude that the plant-associated *B. cereus s.l.* isolates can cause the diarrhoeal syndrome in human beings. Application of plant-associated *B. cereus s.l.* strains in crop protection agents represents a possible risk for public health and should be considered with care.

3.1.3. Genes encoding insecticidal proteins in B. cereus subsp. bombysepticus TK1

Furthermore, we proved the occurrence of *cry* genes encoding entomocidal proteins (δ-endotoxins). Sequences completely matching with the crystal proteins Cry1A1 and Cry2Ba1 were detected in *B. cereus* subsp. *bombysepticus* TK1. The gene encoding the cytolytic CytK protein was detected in in all *B. cereus* group isolates (Figure 2). Synthesis of δ-endotoxins is considered as a typical feature of *B. thuringiensis* [40] However, in line with our results, Liu et al. [9] found, that the ability to synthesize δ-endotoxins is widespreaded in different members of the *B. cereus* species complex. Thus, the presence or absence of *cry* genes cannot be used to discriminate between the *B. cereus* and *B. thuringiensis* species [9].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

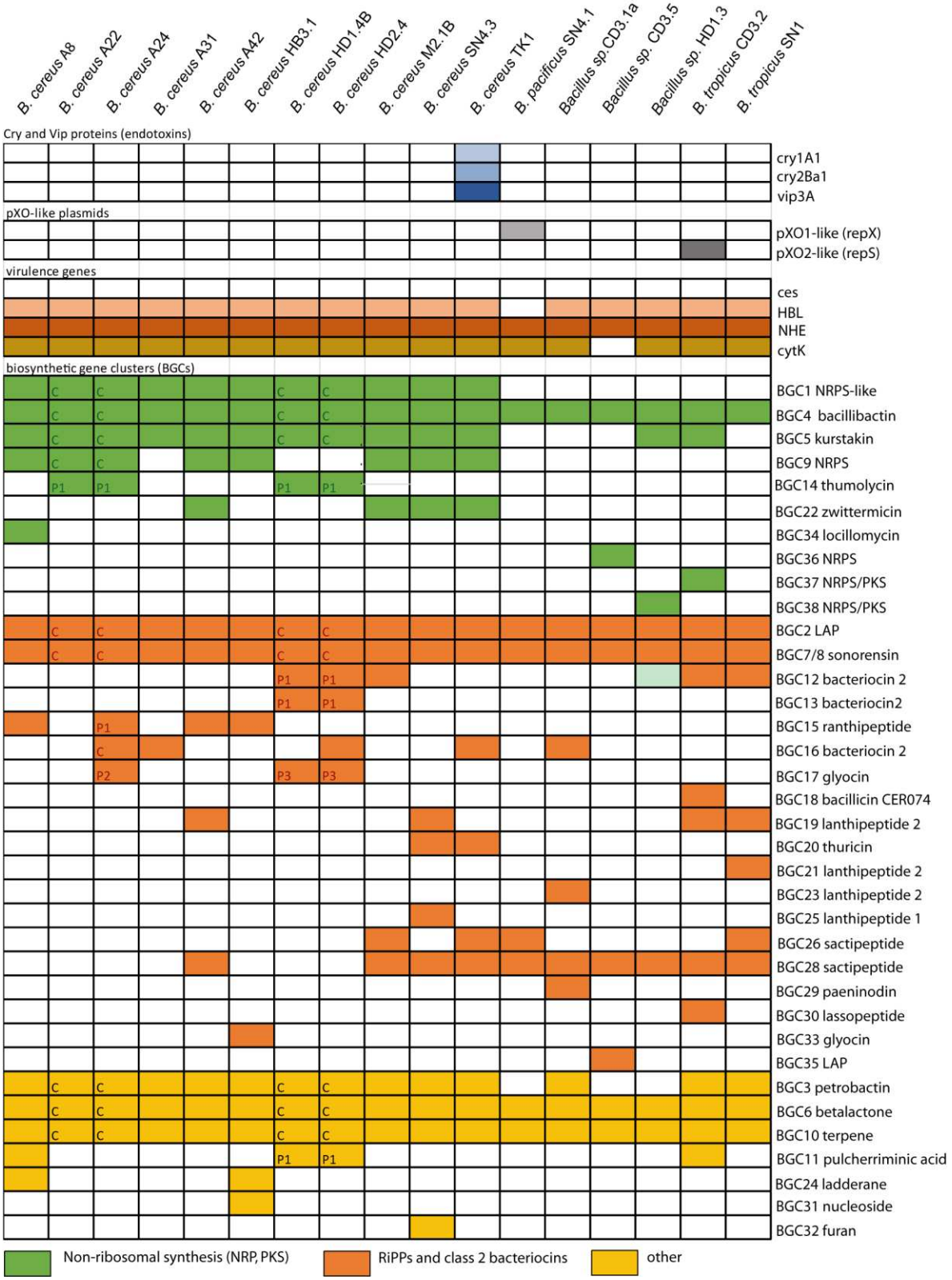


Figure 2. Occurrence of genes encoding entomopathogenic crystal and vegetative proteins, the replication proteins RepS and RepX of *B. anthracis* pXO plasmids, and virulence genes (HBL/NHE, *cyt*K). The gene cluster for synthesizing celeuride (ces) was not detected in any of the isolates.The 36 biosynthetic gene clusters (BGCs) encoding secondary metabolites in the 17 *B. cereus s.l.* isolates were identified by AntiSMASH6.0 and BAGEL4. The location of the BGC on either the chromosome [C], or the plasmids [P1, P3] is indicated, when available. Further information in Suppl. Tables S3-S8.

 doi:10.20944/preprints202310.0635.v1

In addition to Cry proteins, TK1 harbored a gene for synthesis of the vegetative insecticidal protein, Vip3. Vip proteins are referred as second-generation insecticidal proteins. Vip3 proteins have insecticidal activity against Lepidopteran pests [41], and can be used for the management of various detrimental pests.

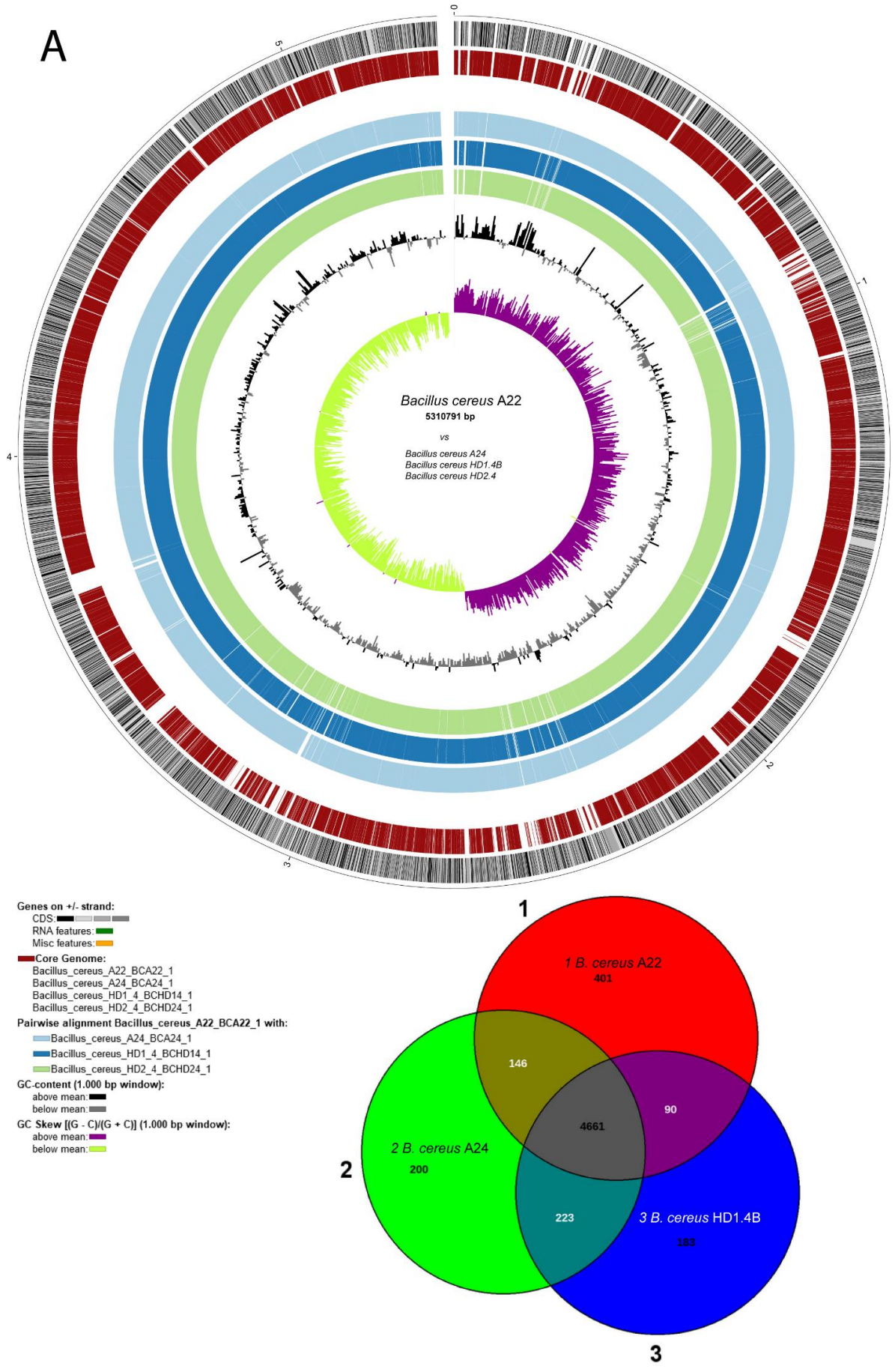
3.1.4. Plasmid-encoded virulence genes and biosynthetic gene clusters in B. cereus isolates A22, A24, HD1.4B and HD2.4

A first survey of the draft genome sequences for the presence of gene clusters encoding lipopeptides revealed that *B. cereus ssp. bombysepticus* A22, and the *B. cereus ssp. cereus* strains A24, HD1.4B, and HD2.4 harbored gene clusters similar to the thumolycin gene cluster, previously detected in *B. thuringiensis* BMB171 [42]. This finding prompted us to sequence completely the four strains using the nanopore sequencing technology (see Materials and Methods). The complete genomes consisted of one single chromosomal DNA molecule and extrachromosomal DNA elements, bearing the features of plasmid DNA (Figure 3). The chromosomes of all four isolates contained more than 5,000 kb. The large P1 plasmids of A22 and A24 contained 480,744 bps, and 471,669 bps, respectively**.** The smaller P2 plasmid of A22 contained 93,778 bps. Small plasmids, not exceeding 12 kb were detected in A24 (P2), HD1.4B (P3), and HD2.4 (P3). The DNA elements found in HD1.4B and HD2.4 were nearly identical suggesting that both isolates represented clones of the same strain. Both, harbored one chromosome and three plasmids of nearly identical size and gene content (Suppl. Table S4). The presence of plasmid-specific Rep proteins in all extrachromosomal elements was corroborated by using the SEED and the RAST annotation system [43] (Suppl. Figure S4).

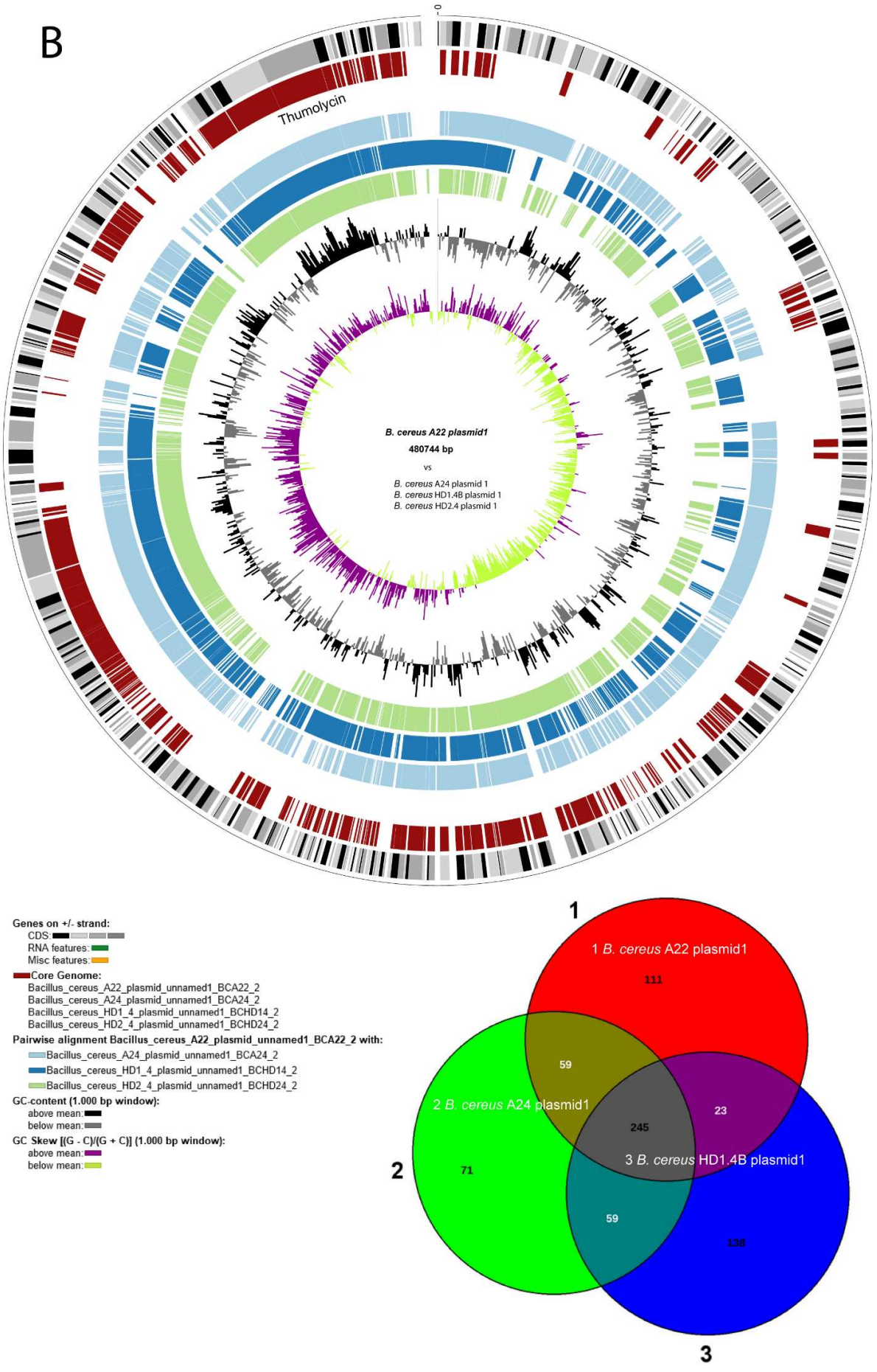
Comparing *B. cereus ssp. cereus* A22 with the representatives of the *B. cereus bombysepticus* clade (*B. bombysepticus* Wang, FORC087, ATCC10876, A42, M2.1B, SN4.3, and TK1) yielded 113 singletons including two catalases, 5-methylcytosine-specificic restriction enzyme A and other restriction enzymes. Comparing *B. cereus ssp. bombysepticus* A24 with the representatives of the *B. cereus cereus* clade (ATCC14579, A22, HD1.4B, HD2.4, A8, HB3.1, and A31) revealed 367 singletons including proteins involved in conjugative transfer, the AlwI family type II restriction endonuclease, and urease subunits and associated proteins.

The potential virulence factor, phosphatidylinositol-specific phospholipase C (PI-PLC), a characteristic marker of the *B. cereus* group [38], was encoded by the large P1 plasmids of A22, A24, HD1.4B and HD2.4. PI-PLCs catalyze the cleavage of the membrane lipid phosphatidylinositol (PI), or its phosphorylated derivatives, to produce diacylglycerol (DAG) and the water-soluble head group, phosphorylated *myo*-inositol [44].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1



 doi:10.20944/preprints202310.0635.v1

Figure 3. Circular plots of the *Bacillus cereus* A22 chromosome (**A**) and plasmid P1 (**B**) generated with Biocircos. The Venn diagrams below show the comparison of A22 with the chromosomes and plasmid P1 of *B. cereus* A24, and HD1.4B. From outer to inner circle: Genes (CDS) on + (1)/- strand (2); core genome, brown (3); GC-content (1,000 bp window) above mean: black, below mean: grey (4); GC Skew [(G-C)/(+C9] (1,000 bp window), above mean: purple, below mean: light green (5). The grey line within the inner circle shows deviations of the average GC-content The 30 kb thumolycin gene cluster is part of the core genome in all four plasmid P1 species.

Annotation of the chromosomal elements detected in A22, A24, HD1.4B and HD2.4 is summarized in Suppl. Table S4. Surprisingly, the plasmid P1 sequences from HD1.4B, and HD2.4 harbored three genes with similarity to the NHE/ HBL enterotoxin operons. These genes are known to be located on the chromosome. In fact, the chromosomes of the four isolates including HB1.4, and HD2.4B harbored the complete NHE/HBL gene set (Suppl. Figure S5).

Many metabolic features were found encoded by the large P1 plasmids harboring more than 500 coding genes. Besides the thumolycin gene cluster, present in the large plasmids of all four isolates, two other BGCs encoding pulcheriminic acid, and the bacteriocin cerein 7B precursor were found located in the large plasmids of HD1.4B and HD2.4.

Interestingly, in addition to the chromosomal encoded type 1 restriction modification systems (RM) [45], type 1 RM gene clusters encoding the subunits M, S, and R were present in the 481 kb plasmid P1 of A24, and in the 94 kb plasmid P2 of A22. A fragmentary type III RM system consisting of RMIII helicase and the methylation subunit flanked by UvrD helicase and a transposase was detected onto the P1 plasmid of HD2.4

A gene cluster detected in plasmid P1 of the *B. cereus* strains was similar to the anthrose BGC, previously described in *B. anthracis* Sterne [46]. The anthrose containing oligosaccharide attached at the surface of the exosporium, might contributes to enhanced survival rates under multiple stress conditions. Our results are in line with previous results of Dong et al. [47] demonstrating that the presence of anthrose-containing exosporia is not restricted to *B. anthracis*.

The complete operon for *myo*-inositol catabolism was detected in the large plasmids of all the four *B. cereus* isolates. The gene cluster was found harboring the genes encoding the same enzymes as the *myo*-inositol operon previously detected in the chromosome of *B. subtilis* [48]. Presence of repeats and mobile elements in the flanking regions suggested that the operon might be acquired by horizontal gene transfer (Suppl. Figure S6).

3.2. Genome mining for biosynthetic gene clusters (BGCs) encoding secondary metabolites

Antimicrobial compounds belong to structurally diverse groups of molecules, such as nonribosomal peptides (NRP) and polyketides (PK), ribosomal synthesized and posttranslationally modified (RiPPs) and unmodified (class 2 bacteriocins) peptides [49,50]. Genome mining using the software pipelines of antiSMASH6.0 [19], PKS/NRPS Analysis [20], and BAGEL4 [21] was performed with the genomes of all the Vietnamese isolates of the *B. cereus sensu lato* complex. The results were subsequently compared against the MIBiG database [51] in order to distinguish between characterized and uncharacterized BGCs. Our survey yielded a total of 209 BGCs representing 36 different gene clusters involved in biosynthesis of secondary metabolites. Only few, such as the siderophores petrobactin (BGC0000942) and bacillibactin (BGC0000309), zwittermicin (BGC0001059), locillomycin (BGC0001005) and pulcherrimic acid (BGC0002103) were found listed in the MIBiG data bank. Two BGCs, kurstakin, and thumolycin, although not listed in the MIBiG repository, could be identified due to their similarity to genes already deposited in the NCBI data bank. Most of the BGCs exhibited no or only low similarity to known BGCs present in the MIBiG data bank. Five BGCs encoding bacillibactin, RiPPs (2), betalactone (1), and terpene (1) were found conserved in all *B. cereus s.l.* isolates (Figure 2). An overview about the BGC species detected in the plant-associated *B. cereus* isolates is presented in Suppl. Table S5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

3.2.1. Non-ribosomally synthesized antimicrobial peptides (NRP) and polyketides (PK)

NRPs are secondary metabolites, which are synthesized through giant multi-modular peptide synthetases [52]. The complete *krs* gene cluster (BGC5) encoding the cyclic lipoheptapeptide kurstakin [53] was found widely distributed, and occurred in the genomes of most *B. cereus sensu lato* isolates, except *B. pacificus* SN4.1, *B. tropicus* SN1, and *Bacillus sp.* CD3-1a and CD3.5. Kurstakin is responsible for biofilm formation [54]. Although, that kurstakin was present in the majority of the investigated isolates (13/17 strains), the kurstakin gene cluster, containing the genes *krsE*, *krsA*, *krsB*, *krsC*, *sfp*, and *krsD* (Figure 4A), is not listed in the MIBiG data bank.

The lipopeptide thumolycin, recently detected in *B. thuringiensis* BMB171, enabled the bacterium to develop a broad spectrum of antimicrobial and nematocidal activities [42]. Unfortunately, the structure of the lipopeptide is still not resolved. We detected the thumolycin (*tho*) gene cluster (BGC14) in plasmids of the *B. cereus* strains A22, A24, HD1.4B, and HD2.4. The genes of the *tho* cluster were spanning around 30 kb. Two multimodular non-ribosomal peptide synthetases (ThoH and ThoI) synthesized a putative pentapeptide Orn-D-X-Leu/Ile-XS-Leu (Figure 4B). The *tho*C, *tho*D, and *tho*E encoded proteins are probably involved in synthesis of the fatty acid chain [42].

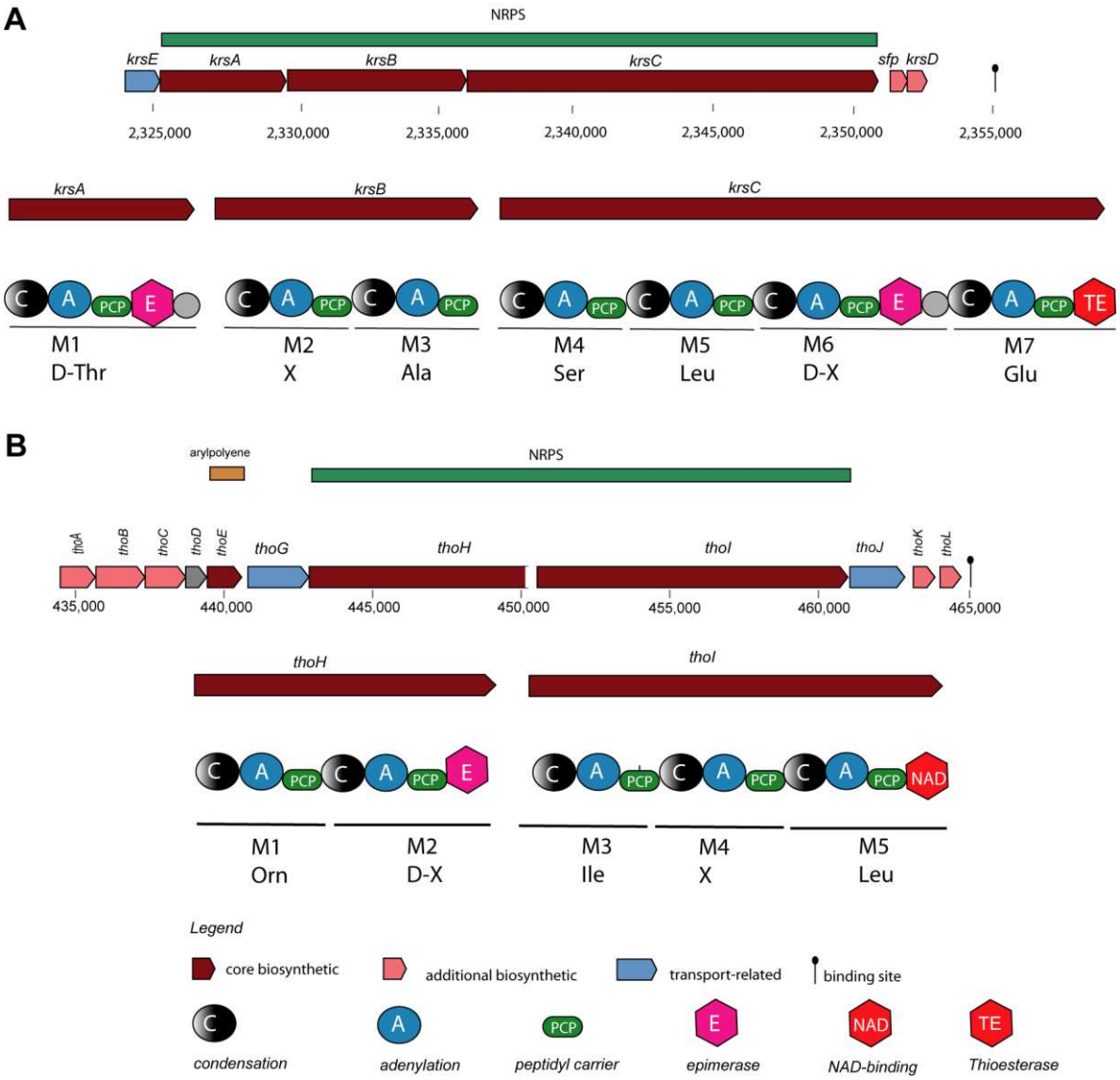


Figure 4. Gene cluster and domain structure involved in non-ribosomal synthesis of cyclic lipopeptides in *B. cereus* A22. **A**: The kurstakin (*krs*) gene cluster is located on the chromosome of A22 between 2,325 – 2,355 kb. The amino acid sequence deduced from the adenylation domains was experimentally corrected and completed by LIFT-MALDI-TOF/TOF MS (see Figure 7). **B**: The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

thumolycin (*tho*) gene cluster resides in the *B. cereus* A22 plasmid 1 between 435 kb to 465 kb. The domain structure of ThoH and ThoI including the amino acids deduced from their adenylation domains is shown. The complete amino acid sequence determined by LIFT-MALDI -TOF/TOF MS is shown in Figure 8. Further domains were detected in ThoC (A), ThoE (KS), ThoK (TE), and ThoL (ACPS).

Fragments of the locillomycin gene cluster [55] (BGC34) were detected in *B. cereus* A8 (Suppl. Table S5). To the best of our knowledge, until now the locillomycin gene cluster has been detected only in members of the *B.subtilis* species complex. The gene cluster for synthesis of the catecholic iron siderophore bacillibactin, 2,3-dihydroxybenzoyl-Gly-Thr trimeric ester, has been previously reported in the genomes of *B. subtilis* [56] (BGC0000309), and *B. velezensis* FZB42 [57] (BGC0001185). Its non-ribosomal synthesis was found dependent on Sfp (phosphopantetheinyl transferase, [58]. BGCs with similar structure as in BGC0000309, and BGC0001185 were detected in all 17 *B. cereus sensu lato* isolates investigated in this study. Whilst the core structure of the bacillibactin transcription unit was well conserved, a *sfp* gene in the flanking region was identified as unique feature for the *B. cereus* bacillibactin operon (Figure 5A). This is in contrast to the operon structure in the *B. subtilis* species complex, where the *sfp* gene is located fare remote, downstream from the surfactin operon [57].

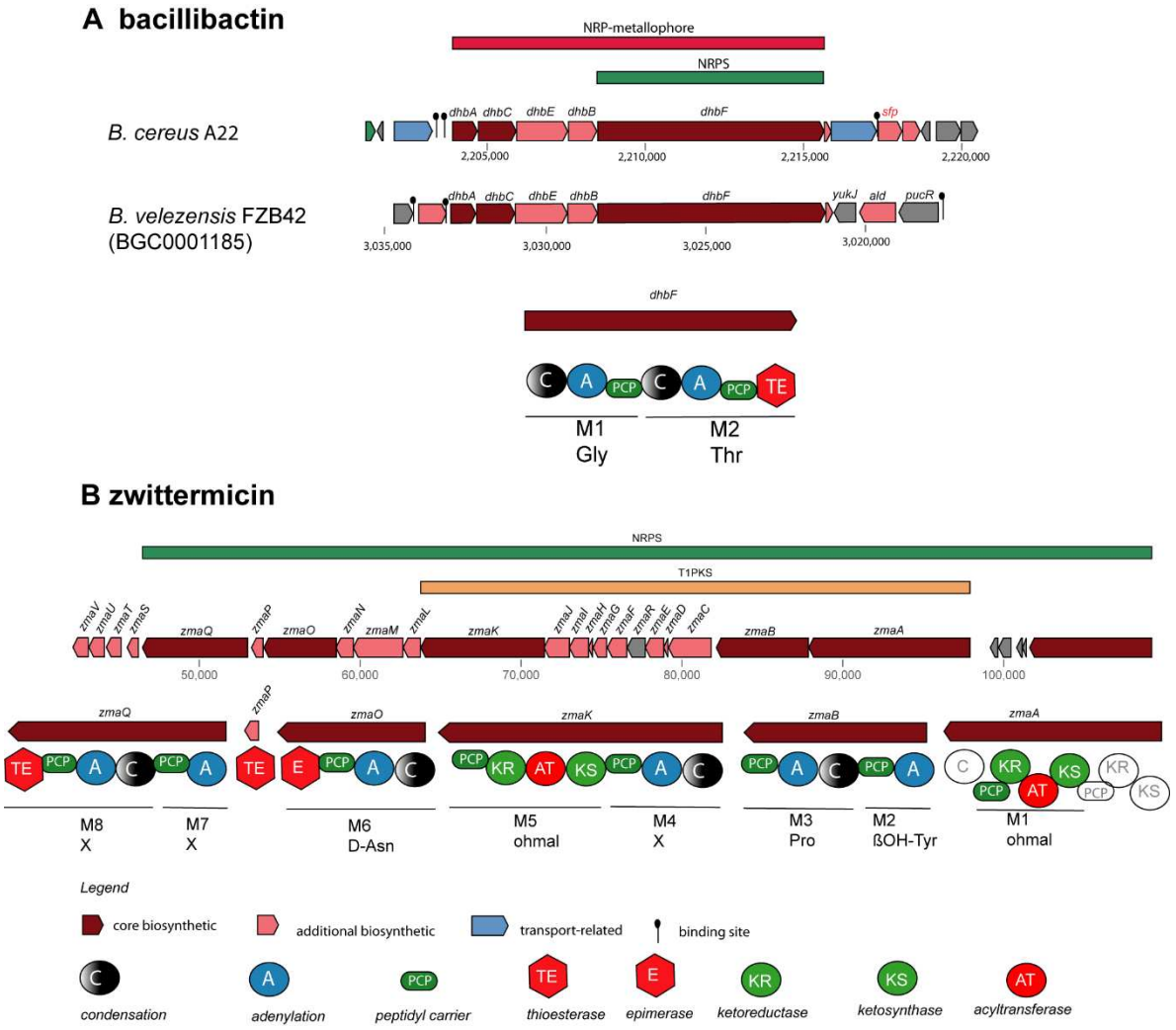


Figure 5. Gene clusters involved in non-ribosomal synthesis of bacillibactin and zwittermicin. The siderophore gene cluster in the *B. cereus* A22 chromosome, **A**: Comparison with the corresponding *B. velezensis* FZB42 gene cluster revealed the presence of the *sfp* gene downstream of the ballibactin transcription unit as unique feature. **B**: The gene cluster for synthesis of the PK/NRP hybrid zwittermicin in *B. cereus* SN4-3.

 doi:10.20944/preprints202310.0635.v1

The gene cluster for synthesis of the aminopolyol antibiotic zwittermicin was detected in four *B. cereus* genomes. The highly polar zwittermicin A (ZmA) possesses antiprotist and antibacterial activities, and consists of numerous ethanolamine and glycolyl moieties flanked by N-terminal D-serine, and an unusual amide generated from ß-ureidoalanine. The aminopolyol structure of the final product results from different processing events of the NRPS/PK hybrid precursor molecule (Figure 5B) in which a multitude of gene products of the *zma* gene cluster is involved [59].

In addition, numerous uncharacterized NRPs, PKS, and NRP/PKS hybrids were found (Suppl. Tables S5 and S6). A unique gene in *B. tropicus* CD3.2, located downstream of an uncharacterized NRP+PK cluster (BGC37, Suppl. Figure S7), encoded a putative necrose inducing protein (NPP1 family) [60].

3.2.2. Gene clusters representing RiPPs, and bacteriocins

By contrast to polyketides and peptides which are synthesized independent from ribosomes, numerous peptides with antimicrobial activity (bacteriocins) are synthesized by a ribosome-dependent mechanism. According to Zhao and Kuipers [49] several groups of ribosomally synthesized peptides (RPs) are distinguished:

- Class I: post-translationally modified peptides smaller than 10 kDa
- Class II: small (<10 Da), unmodified peptides with or without leader sequence.
- Class III: peptides larger than 10 kDa.

RiPPs, such as lanthipeptides (class1 and class2), linear azol(in)e-containing peptides (LAPs), lassopeptides, sactipeptides, thiopeptides, and representatives of the class II unmodified bacteriocins, such as UviB peptides (holin-like proteins), were detected in the *B. cereus* group isolates applying the antiSMASH and BAGEL4 toolkits (Suppl. Figures S8 and S9). Many RiPP biosynthetic proteins recognize and bind their cognate precursor peptide through a domain known as RiPP recognition element (RRE) [61]. Detection of RRE domains by using antiSMASH supported genome mining was helpful in identifying BGCs involved in synthesis of RiPPs, which did not contain known core peptide encoding sequences [62].

A total of 19 different BGCs encoding RiPPs and unmodified class 2 bacteriocins were detected. Only six encoded precursor peptides with apparent similarity to known RiPPs: BGC20 (Thuricin), BGC17 and BGC18 (sublancin/CER074), BGC29 (paeninodin), BGC26: thurincin H) (Suppl. Tables S5 and S7).

Antimicrobial lanthipeptides (lantibiotics) are post-translationally highly modified, and contain the thioether amino acid lanthionine as well as several other modified amino acids [63]. LanA precursor peptides consist of an N-terminal leader peptide and a C-terminal core region. The first step in posttranslational modification is the activation and elimination of water from the Ser and Thr residues forming dehydroalanine (DhA) and dehydrobutyrine (DhB), respectively. Then, ß-thioether cross-links are generated between DhA, DhB and the Cys residues. The modifying enzymes involved in formation of the thioether link in class AI lanthipeptides are the dehydratase LanB and the cyclase LanC. Modification of A2 lanthipeptides is accomplished by LanM, containing the dehydratase and the cyclase domain in one protein. Class A3 and A4 lanthipeptides are modified by LanKC and LanL, respectively [64]. We detected a BGC encoding a representative of the A1 lanthipeptides in *B. cereus* SN4.3 (BGC25). Four genes encoding precursor peptides similar to paenibacillin and subtilomycin were identified within BGC25. BGCs encoding A2 lantibiotics similar to plantaricin (BGC19), thuricin (BGC20), lichenicidin (BGC21), salivaricin (BGC23), and paenibacillin (BGC25) occurred in several *B. cereus* isolates (Suppl. Tables S5 and S7).

Gene clusters encoding LAPS were found widely distributed in *B. cereus* and related species. LAPS are characterized by posttranslational modification of the precursor peptide yielding thiazol(in)e and (methyl)oxazol(in)e heterocycles. Modifying enzymes are the FMN-dependent dehydrogenase (SagB), and cyclodehdratase (SagC, YcaO) [65]. BGC2 (Suppl. Table S7) was identified as being member of the TOMM class (thiazole/oxazole-modified microcins) characterized by a gene cluster consisting of a cyclodehydratase gene, and associated genes encoding dehydrogenase and a

 doi:10.20944/preprints202310.0635.v1

maturation protein. The core region of the TOMM precursor leader peptide contained a region enriched with Cys residues (BGC7/8), which is typically for the hetero-cycloanthracin/sonorensin family [66].

A glyocin encoding gene cluster (BGC17) was detected in plasmid P2 of *B. cereus* A24. Glyocins are defined as post-translationally glycosylated RiPPs with antimicrobial activity [67]. BGC17 resembled sublancin, which is an S-linked glycopeptide coding a SunS family peptide S-glcosyltransferase, and a bacillicin CER074 peptide (BGC0001863) containing a glucose attached to a cysteine residue [68]. A second gene cluster (BGC33) harboring genes encoding a glycosyltransferase and a putative 75 aa precursor peptide was detected in *B. cereus* HB3.1 (Suppl. Table S5).

Lassopeptides are characterized by an N-terminal macrolactam ring threaded by the C-terminal tail. A cysteine protease B, and a lactam synthetase C are necessary for posttranslational modification of the precursor peptide [69]. Two gene clusters, probably encoding lassopeptides, were identified. BGC29 harbored, besides a structural gene (*paeA*) for synthesis of paeninodin lasso peptide [70], the genes for synthesis of the essential components of posttranslational modification, *paeB1* (PQQD family protein), *paeB2* (cysteine protease), and *paeC* (lactam ring closing cyclase). Four copies of the lasso precursor gene, and split B1 and B2 genes were detected in BGC30 (Suppl. Table S7).

Two gene clusters (BGC26, BGC28), encoding the radical S-adenosylmethionine (rSAM) enzyme, necessary for posttranslational modification of sactipeptides, occurred in representatives of the *B. cereus* group. A well-known representative of sactipeptides is subtilosin A (SboA) synthesized by *Bacillus subtilis*. The rSAM enzyme (AlbA) catalyzes the linkage of a thiol with an α-carbon of a functional amino acid residue [71]. BGC26 harbored genes involved in synthesis and rSAM-dependent modification of a thurincin H-like precursor peptide. A gene encoding a protein containing an N-terminal radical SAM domain (pfam04055) and a C-terminal pfam08756 domain with a CxCxxxxC motif (BmbF) was detected in BGC28. By contrast to *B. subtilis* the YfkA and YfkB regions originally reported as separate ORFs in *B. subtilis* were found fused in the *B. cereus* gene cluster (Suppl. Table S7).

Like the structurally related sactipeptides, the thioether linkage in ranthipeptides is generated via a radical-initiated mechanism. However, ranthipeptides do not contain α-carbon links, and were recently designated as non-α thioether peptides [72]. The ranthipeptide gene cluster (BGC15) detected in four *B. cereus* isolates harbored a gene encoding the rSAM protein belonging to the MoaA/NifB/PqqE/SkfB superfamily (Suppl. Table S7).

A gene cluster (BGC7/8), involved in synthesis and modification of an 82 aa precursor thiopeptide belonging to the heterocycloanthracin/sonorensin family [73], occurred in all *B. cereus* group isolates. Its C-terminal region contained an extended repeat region with Cys at every third residue (Suppl. Table S7).

Two different subclasses of bacteriocin class II peptides were detected: The holin-like BhlA encoding genes (BGC12 and BGC16), and a cluster (BGC13) harboring a gene encoding a cerein-like prepeptide belonging to the Blp family. Similar as in lanthipeptides, the Blp family prepeptides are characterized by a conserved GlyGly processing site between the N-terminal leader and the C-terminal core peptide region [74]. The BhlA holin of *Bacillus pumilus* causes bacterial death by cell membrane disruption [75]. Besides the genes encoding leaderless BhlA peptides, the holin gene clusters harbored genes encoding muramidases (GH25 glycosyl hydrolases) that hydrolyze the peptidoglycan cell wall (Suppl. Table S7, Suppl. Figures S8 and S9)).

3.2.3. Other antimicrobial secondary metabolites

Seven BGCs encoding other antimicrobial secondary metabolites were identified in the plant-associated *B. cereus* isolates (Suppl. Figure S7, Suppl. Table S8). Three of them, BGC3 (=BGC0000942, petrobactin), BGC11 (=BGCBGC0002103, pulcherriminic acid), and BGC32 (=BGC0000914, furan) were similar to known BGCs deposited in the MiBIG data bank.

The *asbABCDEF* gene cluster is responsible for biosynthesis of petrobactin, a catecholate siderophore that functions in both, iron acquisition and virulence [76]. We detected the petrobactin gene cluster in the genomes of 14 isolates. Only *B. pacificus* SN4-1 and HD1-3, and *Bacillus sp.* CD3.5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

did not harbor the BGC0000942 cluster, which is common in most representatives of the *B. cereus* group [77].

Pulcherriminic acid is a cyclic dipeptide, able to chelate Fe^{3+} [78]. Due to its high affinity to Fe ions, *Bacillus* strains producing pulcherriminic acid compete successfully with other microorganisms in low iron environments. A gene cluster similar to the pulcherriminic acid synthesis cluster in *B. subtilis* (BGC0000914) was detected in the *B. cereus* HD1.4B and HD2.4 plasmid sequences (Suppl. Table S4), and in the draft genomes of *B. cereus* A8 and *B. tropicus* CD3-2.

Several genes of BGC32, possibly involved in synthesis of a furan-like metabolite, showed striking similarity to the methylenomycin A gene cluster in *Streptomyces coelicolor* [79].

Four BGCs did not show similarity to any characterized biosynthetic gene clusters. BGC6 possibly encoded a ß-lactone harbored genes with similarity to genes flanking the plipastatin BGC in *B. subtilis*. BGC24 contained several genes of the carbohydrate metabolism probably involved in synthesis of ladderane. BGC10 and BGC31 encoded enzymes for synthesis of terpene and nucleoside metabolites, respectively.

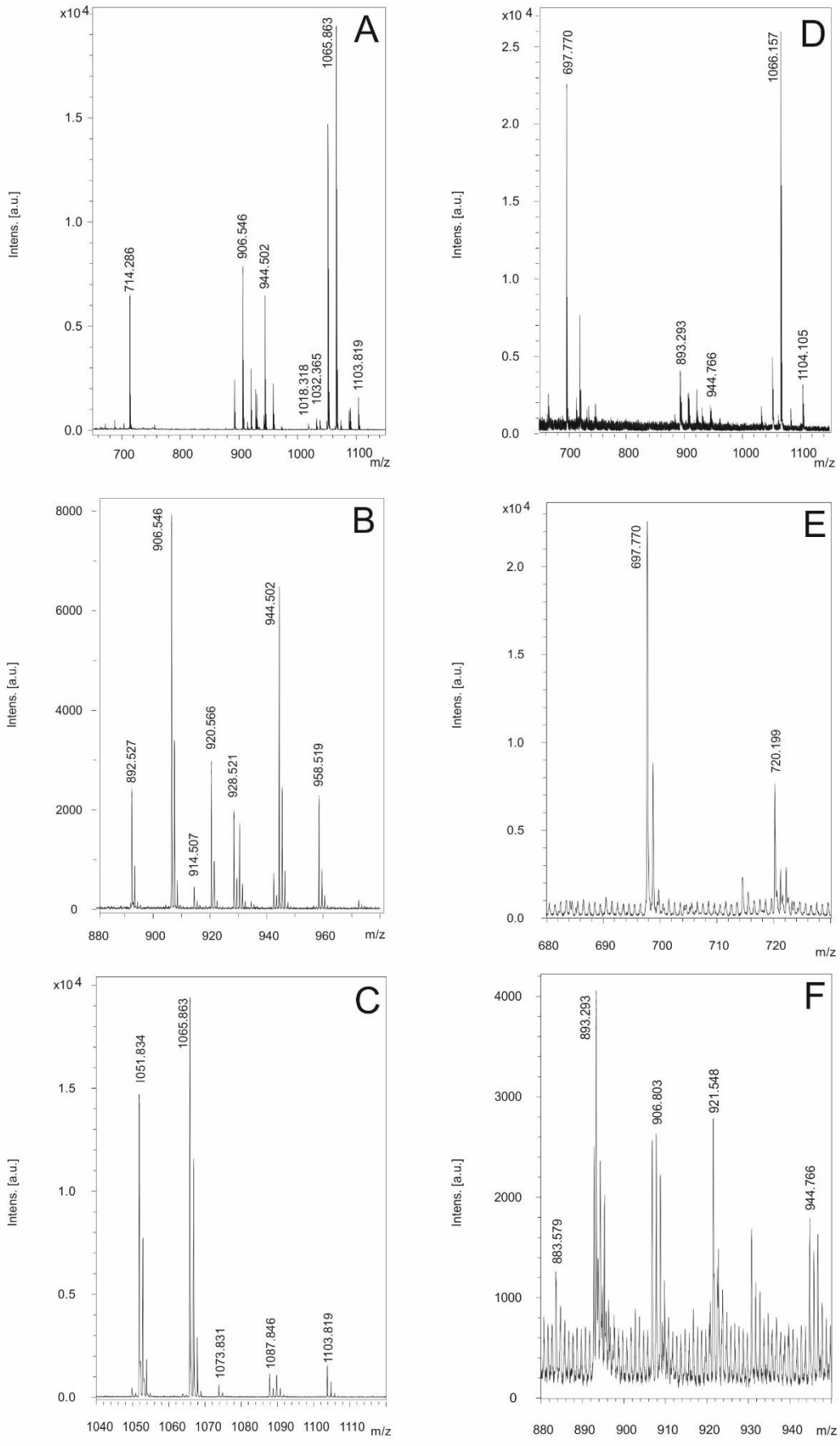
3.3. Detection of bioactive peptides by MALDI-TOF mass spectrometry

The genome mining data summarized in Figure 2 indicate the presence of BGCs on the genomic level. However, the real biosynthetic capacity of the investigated isolates can only be verified by isolation and structural analysis of the compounds actually produced. In Figure 6 we demonstrate the production of the non-ribosomally formed secondary metabolites of strain A22 as a representative for *B. cereus* detected by MALDI-TOF MS. As an overview, Figure 6 A-C shows mass spectra for the compounds found in a surface extract of A22 taken from cell materials grown on agar plates in the Landy medium for 48 h. Two prominent products were observed. Fig 5B shows the mass peaks for two kurstakins with chain lengths of their fatty acid component of 12 and 13 carbon atoms, respectively. The following mass data were found: C12-kurstakins: $[M + H,Na,K]^+$ = 892,5/914,5/930.5 Da; C13-kurstakins: $[M + H,Na,K]^+$ = 906.5/928.5/944.5 Da. In addition, yet unknown compounds with mass numbers of 1051.8 and 1065.9 were found which dominate the MALDI-TOF mass spectrum of the surface extract in Figure 6A. Presumably, there are two isomers which differ by a methylene group (Figure 6C). This compound cannot be correlated to any of the BGCs of the antiSMASH-profile of strain A22.

Figure 6 D-F show mass spectra of the products of A22 released into the culture broth on growth in liquid cultures for 48 in the Landy medium. Here the arylpolyene lipopeptide thumolycin and both siderophores bacillibactin and petrobactin were detected. Figure 6E exhibits the mass peaks of thumolycin (m/z = 697.8) and petrobactin (m/z = 720.2). A relatively small part of the kurstakins is released into the culture filtrate, while the main part retains attached at the outer surface of A22. In Figure 6F kurstakin mass peaks (*m/z* = 906.8/930.8 and 944.8) overlap with those of the siderophore bacillibactin [M + H,Na,K]+ = 883.6/905.6/921 Da. These results demonstrate that kurstakins were predominantly found attached to the outer surface of *B. cereus* cells, while thumolycin and both siderophores bacillibacin and petrobactin are released into the culture medium. Similar profiles have been obtained for strains A24, HD1.4B and HD2.4. All other investigated *B, cereus* isolates did not produce thumolycin.

The structure of both lipopeptide products of strain A22, kurstakin and thumolycin were investigated in detail by LIFT-MALDI-TOF/TOF fragment analysis [24]. Figure 7 and Table 1 show the sequence determination of a C13 kurstakin with a parent ion [M + H}+ = 906.504 Da derived from product ions obtained by LIFT-MALDI-TOF/TOF fragment ion spectra. In Table 1 the structure of this compound is modelled from nearest neighbour-relationships using di-, tri- and tetrapeptide fragments.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

Figure 6. Bioactive compounds produced by *B. cereus* A22. **A-C:** Compounds detected in surface extracts of this strain. **A:** MALDI-TOF mass spectrum of a surface extract of strain A22 grown on agar plates in the Landy medium in the mass range from *m/z* = 650-1150. **B:** Kurstakins observed between *m/z* = 880-980. **C:** Detection of a yet unknown prominent product with mass numbers of 1051.83 and 1065.86 Da. **D-F**: Compounds found in the culture filtrate of strain A22 grown in the Landy medium for 48 h. **D:** MALDI-TOF mass spectrum of a culture filtrate of strain A22 grown in the Landy medium for 48 h in the mass range from *m/z* = 690 – 1120. **E**: Detection of the arylpolyene-lipopeptide thumolycin and the siderophore petrobactin with mass numbers of 697.77 and 720.20. Da. **F:** Kurstakins and the siderophore bacillibactin ([M + H;K]$^+$= 883.58 and 921.55 Da) found in the mass range between *m/z*= 880-950.

$b_n - H_2O$(*found*)	-	280.123	337.149	408.107	495.159	632.284	760.445	
b_n (*found*)	197.030	298.133	355.112	-	-	-	778.446	906.504
b_n (*calc.*)	197.190	298.238	355.259	426.296	513.328	650.387	778.446	906.504
	C13-FA	**Thr**	**Gly**	**Ala**	**Ser**	**His**	**Gln**	**Gln**
y_n (*calc.*)	906.502	710.322	609.275	552.253	481.216	394.184	257.125	129.066
y_n (*found*)	906.504	710.266	609.163	552.133	481.106	394.084	257.050	129.016
$y_n - H_2O$(*found*)		692.185	-	-	463.103	-	-	-

Figure 7. Mass spectrometric sequence determination of a C13-kurstakin produced by *B. cereus* A22 with a parent ion [M + H]$^+$ = 906.504 derived from product ion patterns obtained by LIFT-MALDI-TOF/TOF fragment ion spectra. FA: fatty acid component.

Table 1. Modeling of the structure of a C13-Kurstakin (*m/z* = 906.5) by nearest neighbor relationships obtained by MALDI-TOF MS.

a) dipeptide fragments	*m/z* Calc.	*m/z* found
C13-FA-Thr	298.238	298.141/280.129
Thr-Gly	159.077	159.007/141.022
Gly-Ala	129.066	129.016
Ala-Ser	159.077	159.007/141.022
Ser-His	225.099	225.034/207.019
His-Gln	266.125	266.060
Gln-Gln	257.125	257.050
b) tripeptide fragments		
C13-FA-Thr-Gly	341.244	341.170/323.167
Thr-Gly-Ala	230.114	230.026/212.025
Gly-Ala-Ser	216.098	216.024/198.022
Ala-Ser-His	296.136	- /278.050
Ser-His-Gln	353.157	353.084/335.070
His-Gln-Gln	394.184	394.101/335.070
c) tetrapeptide fragments		
C13-FA-Thr-Gly-Ala	426.296	- /408.121
Thr-Gly-Ala-Ser	317.146	317.057/299.070
Gly-Ala-Ser-His	353.157	353.084/335.070
Ala-Ser-His-Gln	424.194	-
Ser-His-Gln-Gln	481.216	481.120/463.120

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

With the same technique we investigated thumolycin which is a combination of a pentapeptide attached to a yet unknown arylpolyene lipidic residue. By LIFT-MALDI-TOF/TOF fragment analysis we obtained the complete sequence of the pentapeptide part for the first time which is compatible with the initial results from Zheng et al. [42], and the module organization of the corresponding BGC derived from antiSMASH 6.0 genome mining. The structure of this pentapeptide is shown in Figure 8.

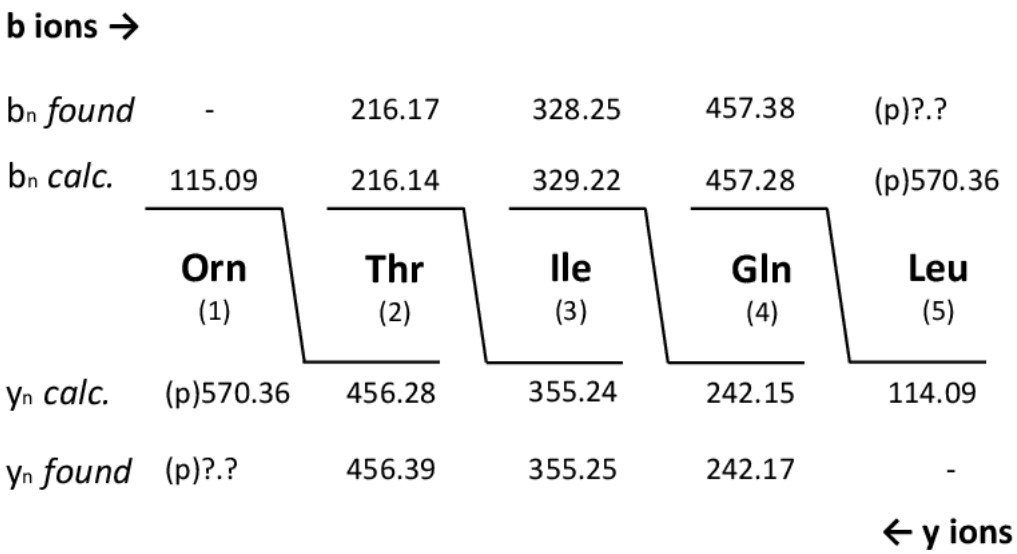


Figure 8. Sequence of the pentapeptide part of the lipopeptide thumolycin derived from product ion pattern obtained by LIFT-MALDI-TOF/TOF fragment ion spectra. The yet unknown arylpolyene lipid part of thumolycin of unknown length is linked to the Orn residue. The amino acids Orn, Ile, and Leu were also predicted by their adenylation domain sequences from genome mining using antiSMASH 6.0 (Figure 3B).

In summary, by MALDI-TOF mass spectrometry we have detected all compounds produced by *B. cereus* strains non-ribosomally. Investigation of the RiPPs, such as lanthipeptides, sactipeptides and bacteriocins are in preparation.

3.4. B. cereus s. l. strains suppressed plant pathogens and promoted plant growth

3.4.1. Antifungal and nematocidal activity

Our *in vitro* bioassays revealed, that only few of the *B. cereus* group isolates inhibited efficiently phytopathogenic fungi, and nematodes. Antifungal activity was examined *in vitro* using *Fusarium oxysporum* known for causing fusarium wilt disease [80], and *Phytophthora palmivora,* one of the most detrimental plant pathogens in Vietnam [81]. Whilst most of the isolates did not significantly suppress growth of the pathogenic fungi, strong antifungal activity was exerted by *B. cereus* HB3.1 and in *Bacillus sp.* HD.3 (Figure 9).

Root-knot nematodes, such as *Meloidogyne spp*., are one of the most important plant pathogens in tropical and temperate agriculture, and are responsible for significant harvest losses of main Vietnamese crops, such as coffee and black pepper [82]. In order to analyze antagonistic activity of the *B. cereus* isolates, we tested at first their suppressing effect against the model nematode *Caenorhabditis elegans.* Fast and slow death rates were estimated in a bioassay under laboratory conditions. Most of the investigated *B. cereus* isolates were able to kill considerable amounts of the nematodes as revealed in both test systems (Figure 10A).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

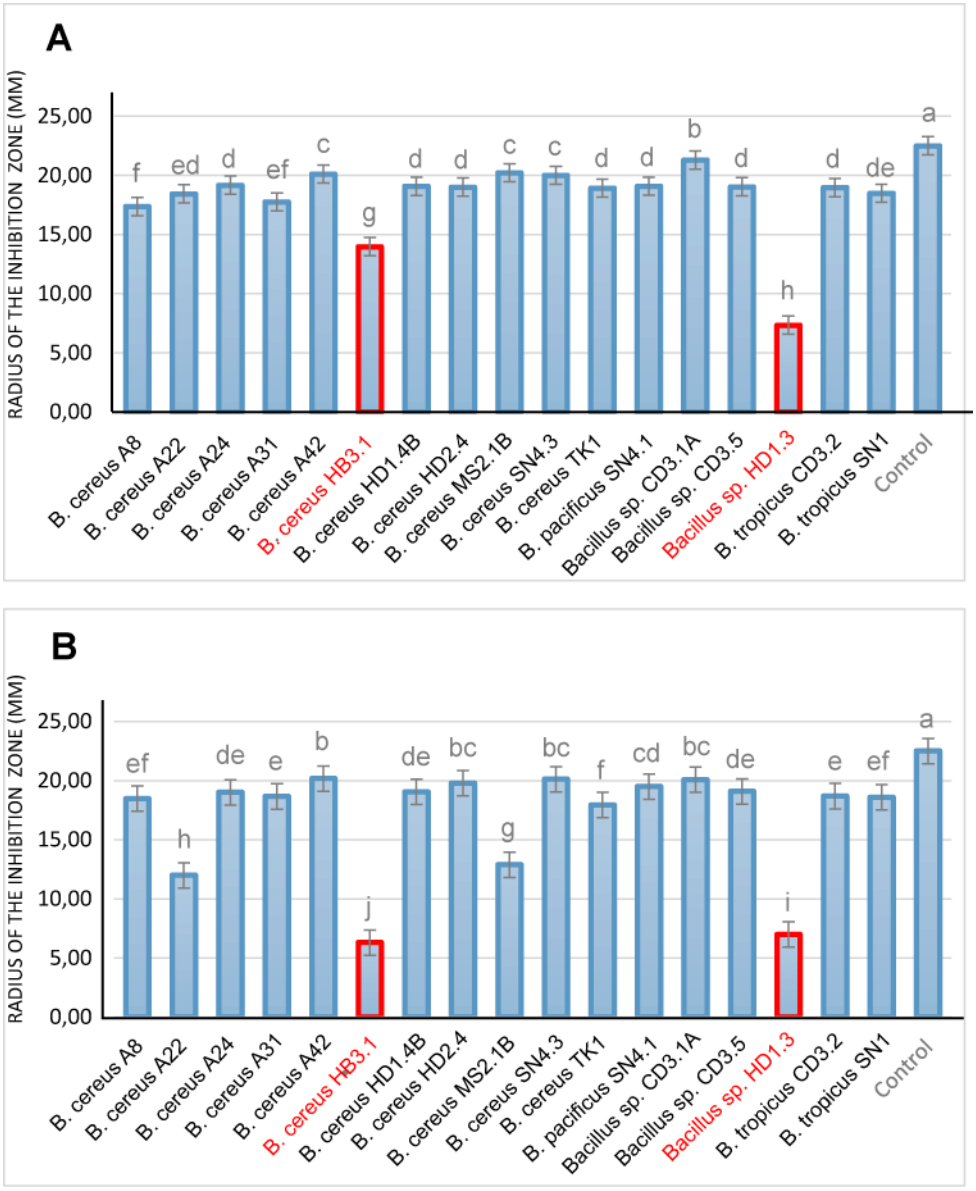


Figure 9. *In vitro* assay of antifungal activity of *B. cereus* group strains isolated from Vietnamese crop plants. A: Suppression of *Phytophthora palmivora*; B: Suppression of *Fusarium oxysporum*. Strains with enhanced antifungal action are indicated by red letters. All diagrams showed the means of at least three replicates (n≥ 3). Negative controls were performed without treatment with the bacteria. Columns with superscripts with the same letter are not significantly different according to Fisher´s Least Significance Difference (LSD) Test ($p \leq 0.05$).

In order to examine inhibiting effects against phytopathogenic nematodes more directly, we isolated a representative strain of *Meloidogyne sp*. from the galls of infested black pepper plant roots according to the hypochlorite procedure [83]. The inhibiting effect exerted by the test bacteria on disease development was examined in a greenhouse experiment. Ten weeks after transplanting of the tomato plantlets in soil, formation of root knots was visually registered and used as measure for calculating the disease index according to Bridge and Page [26]. All *B. cereus* isolates were found efficient against nematodes. In presence of HB3.1, HD1.4B, HD2.4, MS2.1B, and SN4.3 the killing rates (estimated as slow and fast killing rates) of *Caenorhabditis elegans* exceeded 60% (Figure 9A). The *Melodoine* sp. caused disease index of tomato plants was found reduced by more than 50% after application of *B. cereus* HB3.1, HD2.4, MS2.1B, SN4.3 (Figure 9B). Similar rates were previously detected in representatives of *B. velezensis* [4] and *Brevibacillus* spp. [3].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

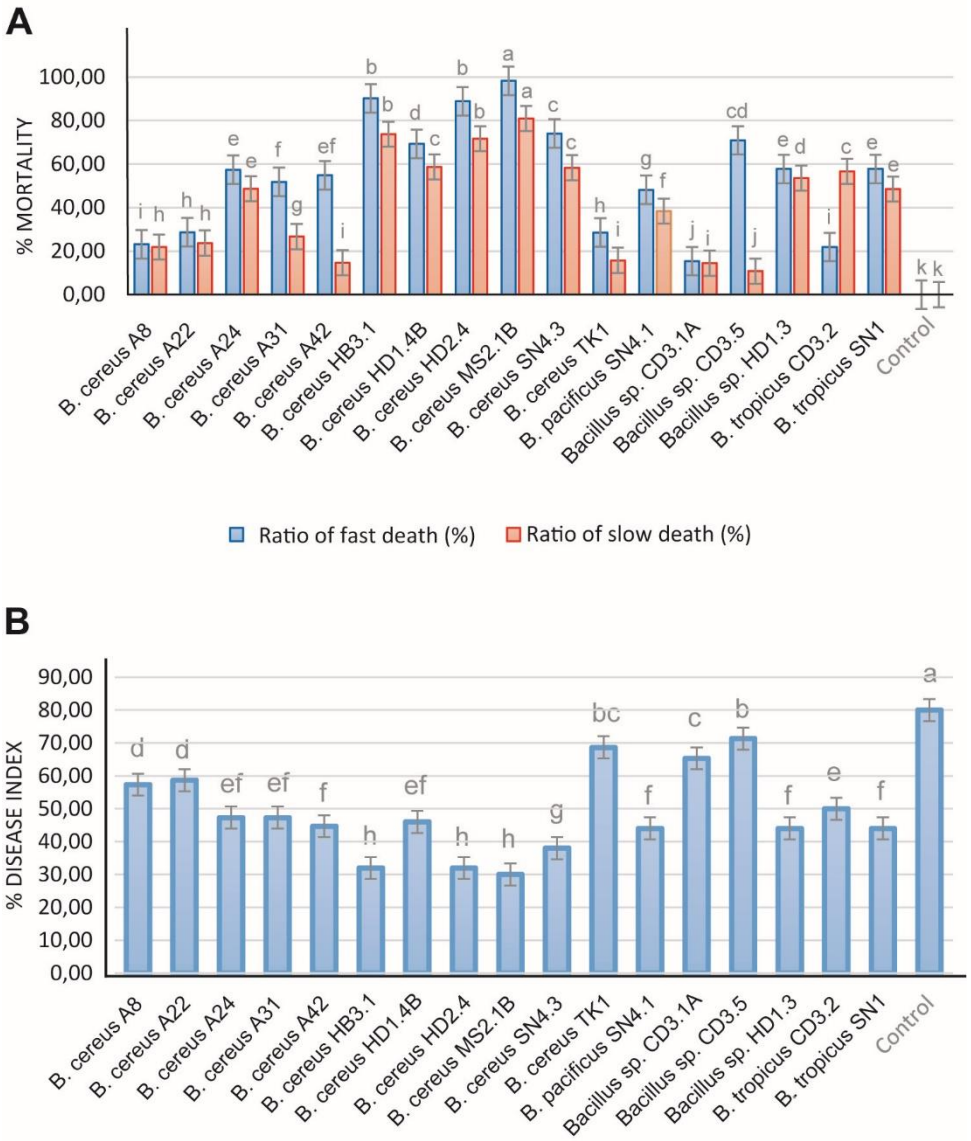


Figure 10. Nematocidal activity of *B. cereus* isolates **A**: Bioassay with *Caenorhabditis elegans*. Slow killing activity was determined onto NGM plates, and fast killing activity in liquid medium as described previously [3]. **B:** Determination of the biocontrol action of the *B. cereus* isolates on the root-knot nematode *Meloidogyne* sp. in greenhouse experiments. Tomato plants infested with *Meloidogyne* sp. were used for the test (counting of "knots" in tomato roots). The increase compared to the control without adding with the *Bacillus* isolates is shown. All experiments were carried out with three independent repetitions and randomized design. The bars above the columns indicate the standard error (SE). Different letters at each treatment indicate significance between inoculated and uninoculated conditions at $P \leq 0.05$ level after the t-test.

3.4.2. Plant growth promotion

We examined the effect of the Vietnamese *B. cereus* isolates in the *Arabidopsis thaliana* biotest system [27]. *B. cereus* HD1.4B and *B. cereus* HD2.4 enhanced growth of the *Arabidopsis* seedlings by more than 20% (Figure 11). However, using the same biotest system, the increase rates observed for some plant-associated *Brevibacillus* and *B. velezensis* strains isolated during the same survey [2] were higher and estimated to be in the range of 30 to 40 % [3,4].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2023 doi:10.20944/preprints202310.0635.v1

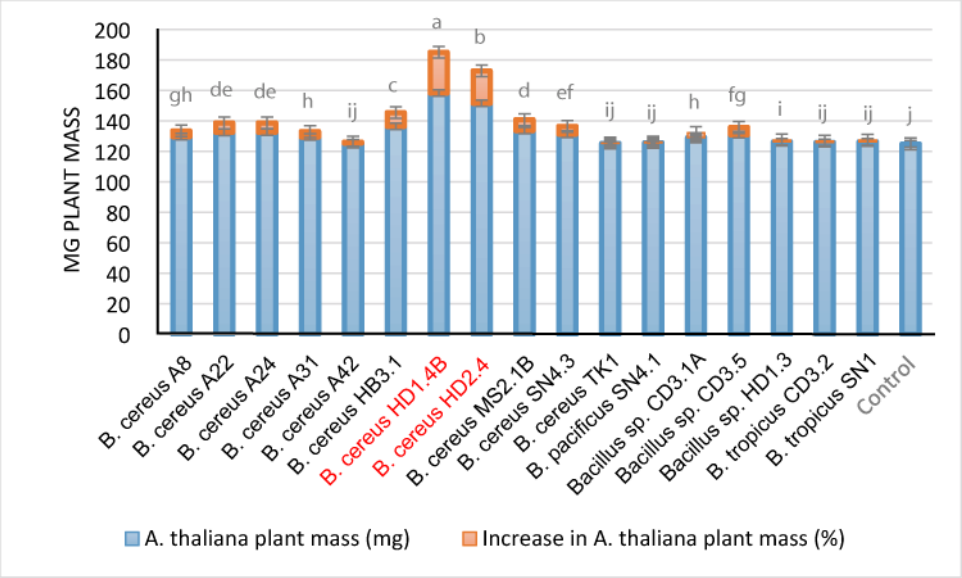


Figure 11. Growth promoting effects of *B. cereus* isolates on *Arabidopsis thaliana* seedlings. The blue columns in the diagram represent the fresh weight obtained after 21 days under controlled conditions in the growth chamber. The % increase compared to the untreated control (red columns) is indicated on top of the columns. Each treatment value is presented as means of three replications (n = 3) with standard error. Different letters at each treatment indicate significance between inoculated and uninoculated conditions at $P \leq 0.05$ level after the t-test.

4. Conclusions

In this study we have shown that plant-associated representatives of the *B. cereus* group were able to suppress important plant pathogens, such as fungi (*Fusarium oxysporum*), oomycetes (*Phytophthora palmivora*), and root-knot forming nematodes (*Meloidogyne sp*.) The plant-growth-promoting activity of some of the isolates could also be demonstrated.

Genome mining revealed that the members of the *B. cereus* group are rich in gene clusters probably involved in synthesis of antimicrobial peptides efficient in inhibiting plant pathogens, and triggering plant induced systemic resistence [84]. A total of 36 different biosynthetic gene clusters (BGCs), many of them not listed in the MiBIG data bank, were detected in the 17 isolates obtained from Vietnames crop plants. Mass-spectrometric analysis revealed that in addition to some hitherto unknown compounds, several species of the antimicrobial lipopeptides kurstakin and thumolycin, and the siderophores bacillibactin and petrobactin were expressed in many of the isolates. The arylpolyene lipopeptide thumolycin was reported to possess interesting antimicrobial and nematocidal activities, but its primary structure was not resolved [42]. Here, the primary structure of the pentapeptide part has been resolved, and the Orn residue was identified as being linked with the yet unknown arylpolyene lipid part.

Besides antimicrobial peptides, biocontrol of plant pathogens can be exerted by δ-endotoxins (parasporal inclusion proteins) traditionally known to be produced by *B. thuringiensis*, a close relative of *B. cereus*. *cry* genes encoding the entomocidal crystal proteins Cry1Aa1 and Cry2Ba1 were detected in the genome of *B. cereus* ssp. *bombysepticus* TK1. In the same strain another gene, encoding the vegetative insecticidal protein Vip3, was also detected suggesting that presence of insecticidal proteins is not restricted to *B. thuringiensis*. The encouraging results described above might cause to develop selected representatives of the *B. cereus* group as biocontrol agents. However, as known food poisoning organisms and members of risk group 2, the potential of the *B. cereus* group isolates to produce toxins needed to be carefully examined, before they can be applied in sustainable agriculture.

In this context we showed, that no gene clusters encoding *B. anthracis* pXO plasmid-related toxins were present in all the *B. cereus s.l.* genomes investigated here. Furthermore, we could rule out synthesis of the heat-stable celeuride toxin, the causative agent of the emetic syndrome. However,

the regular appearance of chromosomal localized virulence genes, encoding the heat-labile enterotoxins HBL and NHE, might restrict the direct application of *B. cereus* s.l. strains in biological plant protection.

However, biological use of interesting *B. cereus* BGCs, and genes encoding entomocidal proteins can be achieved by their heterologous expression in safe plant-beneficial host strains, which has been already demonstrated with *B. velezensis* FZB42 [85,86].

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Tree inferred with FastMe 2.1.6.1 from GBDP distances calculated from 16S rDNA gene sequences.; Figure S2: GBDP tree (whole genome sequence based) inferred with FastMe 2.1.6.1 from GBDP distances calculated from genome sequences; Figure S3: Heat map of the FastANI values of 128 *B. cereus* group genomes calculated and drawn by the EDGAR software package; Figure S4: Figure S5: Localization of virulence genes and gene clusters on chromosomes and plasmids of B. cereus isolates. The cytK gene and the NHE/HBL gene clusters were chromosomally localized. The complete set of NHE and HBL genes was chromosomally localized in all four completely sequenced strains (A22, A24, HD1.4B, HD2.4). The P1 plasmid sequences of HD1.4B and HD2.4 harbored genes with similarity to the NHE/HBL enterotoxin family; Figure S6: Plasmid encoded catabolic operons, biosynthetic gene clusters (BGCs) and restriction/modification systems; Figure S7: BGCs in the *B. cereus* group isolates encoding NRPS/PKS and other secondary metabolites; Figure S8: RiPP gene clusters detected by applying the BAGEL4 software (http://bagel4.molgenrug.nl/) in the genomes of the Vietnamese *Bacillus cereus* group genomes; Figure S9: RiPP gene clusters detected by applying the antiSMASH version 6 software in the Vietnamese *Bacillus cereus* group genomes. Table S1: Types of entomocidal Cry toxins; Table S2: Species and subspecies delineation of the 17 Vietnamese *B. cereus* sensu lato isolates according to their dDDH and Fast ANI values; Table S3: List of the 17 *B. cereus* group strains used inthis study with detailed annotation; Table S4: Genome annotation of the four completely sequenced B. cereus strains by RASTk; Table S5: Gene clusters (BGCs) of the Vietnamese Bacillus cereus group isolates; Table S6: Gene clusters (BGCs) of the Vietnamese Bacillus cereus group isolates involved in nonribosomal synthesis of antimicrobial pepttides (RiPPs and bacteriocins); Suppl. Table S7: Gene clusters (BGCs) of the Vietnamese Bacillus cereus group isolates involved in ribosomal synthesis of antimicrobial peptides (RiPPs and bacteriocins). Suppl. Table S8: Gene clusters (BGCs) of the Vietnamese Bacillus cereus group isolates involved in synthesis of other secondary metabolites.

Author Contributions: Conceptualization, R.B., L.T.T.T., T.S., P.L.; software, J.B.: investigation, J.V., J.J., S.H, L.T.T.T., A.S., P.T.L., L.T.T.P., S.K., R.B.; resources, T.S., P.L.; data curation, J.J. S.H., J.V., C.B.; writing, R.B., J.V., L.T.T.T.; review and editing, J.V., S.K., T.S., P.L., R.B.; project administration, R.B., T.S., P.L.; funding acquisition, R.B., T.S., P.L., L.T.T.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported through the project ENDOBICA by the Bundesministerium für Bildung und Forschung (BMBF) (grant no 031B0582A/031B0582B), the National Foundation for Science and Technology Development (NAFOSTED); code no. 106.03-2017.28), and the Ministry of Science and Technology (MOST) in Vietnam (code no. NDT.40.GER/18).

Data Availability Statement: Gene bank accession numbers of complete genome sequences available in the NCBI data bank are listed in section 2 Materials and Methods.

Acknowledgments: Mrs. Silke Becker and Petra Lochau is thanked for excellent technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Borriss R. (2011). "Use of plant-associated *Bacillus* strains as biofertilizers and biocontrol agents,", in Bacteria in Agrobiology: Plant Growth Responses. Ed. Maheshwari D. K. (Heidelberg: Springer;), 41–76. doi: 10.1007/978-3-642-20332-9_3.
2. Tam, L.T.T.; Jähne, J.; Luong, P.T.; Thao, L.T.P.; Chung, L.T.K.; Schneider, A.; Blumenscheit, C.; Lasch, P.; Schweder, T.; Borriss, R. Draft genome sequences of 59 endospore-forming Gram-positive bacteria associated with crop plants grown in Vietnam. *Microbiol. Resour. Announc.* **2020**, 9, e01154-20..
3. Jähne, J., Le Thi, T. T., Blumenscheit, C., Schneider, A., Pham, T. L., Le Thi, P. T., Blom, J., Vater, J., Schweder, T., Lasch, P., & Borriss, R. Novel Plant-Associated Brevibacillus and Lysinibacillus Genomospecies Harbor a Rich Biosynthetic Potential of Antimicrobial Compounds. *Microorganisms* **2023**, *11*(1), 168. https://doi.org/10.3390/microorganisms11010168

4. Thanh Tam LT, Jähne J, Luong PT, Phuong Thao LT, Nhat LM, Blumenscheit C, Schneider A, Blom J, Kim Chung LT, Anh Minh PL, Thanh HM, Hoat TX, Hoat PC, Son TC, Weinmann M, Herfort S, Vater J, Van Liem N, Schweder T, Lasch P, Borriss R. Two plant-associated *Bacillus velezensis* strains selected after genome analysis, metabolite profiling, and with proved biocontrol potential, were enhancing harvest yield of coffee and black pepper in large field trials. *Front Plant Sci.* **2023**, 23;14:1194887. doi: 10.3389/fpls.2023.1194887
5. Liu Y, Du J, Lai Q, Zeng R, Ye D, Xu J, Shao Z. Proposal of nine novel species of the Bacillus cereus group. *Int J Syst Evol Microbiol.* **2017**, 67(8):2499-2508. doi: 10.1099/ijsem.0.001821
6. Bartlett A, Padfield D, Lear L, Bendall R, Vos M. A comprehensive list of bacterial pathogens infecting humans. *Microbiology* (Reading). **2022**, 168(12). doi: 10.1099/mic.0.001269
7. Koch, R. (1876). Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des Bacillus Anthracis. Cohns beiträge zur Biologie der Pflanzen, 2 (2), 277
8. Young JA, Collier RJ: Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu Rev Biochem* **2007**, 76:243-65
9. Liu Y, Lai Q, Göker M, Meier-Kolthoff JP, Wang M, Sun Y, Wang L, Shao Z. Genomic insights into the taxonomic status of the Bacillus cereus group. *Sci Rep*. **2015**, 5:14082. doi: 10.1038/srep14082
10. Yang, S., Wang, Y., Liu, Y., Jia, K., Zhang, Z., & Dong, Q. Cereulide and Emetic Bacillus cereus: Characterizations, Impacts and Public Precautions. *Foods*, **2023**, *12*(4). doi.org/10.3390/foods12040833
11. Berliner, E. (1915). Über die Schlafsucht der Mehlmottenraupe (Ephestia kühniella Zell) und ihren Erreger *Bacillus thuringiensis* n. sp. *Zeitschrift für angewandte Entomologie*, Berlin
12. Palma L, Muñoz D, Berry C, Murillo J, Caballero P. Bacillus thuringiensis toxins: an overview of their biocidal activity. *Toxins* (Basel). **2014**, 6(12):3296-325. doi: 10.3390/toxins6123296.
13. Hoffmaster AR, Hill KK, Gee JE, Marston CK, De BK, Popovic T, Sue D, Wilkins PP, Avashia SB, Drumgoole R, Helma CH, Ticknor LO, Okinaka RT, Jackson PJ. Characterization of Bacillus cereus isolates associated with fatal pneumonias: strains are closely related to Bacillus anthracis and harbor B. anthracis virulence genes. *J Clin Microbiol.* **2006**, 44(9):3352-60. doi: 10.1128/JCM.00561-06.
14. Kolstø AB, Tourasse NJ, Økstad OA. What sets Bacillus anthracis apart from other Bacillus species? *Annu Rev Microbiol.* **2009**, 63:451-76. doi: 10.1146/annurev.micro.091208.073255
15. Blumenscheit C., Jähne J., Schneider A., Blom J., Schweder T., Lasch P., Borriss, R. Genome sequence data of *Bacillus velezensis* BP1.2A and BT2.4. *Data Brief.* **2022**, 41, 107978. doi: 10.1016/j.dib.2022.107978.
16. Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, Davis JJ, Dempsey DM, Dickerman A, Dietrich EM, Kenyon RW, Kuscuoglu M, Lefkowitz EJ, Lu J, Machi D, Macken C, Mao C, Niewiadomska A, Nguyen M, Olsen GJ, Overbeek JC, Parrello B, Parrello V, Porter JS, Pusch GD, Shukla M, Singh I, Stewart L, Tan G, Thomas C, VanOeffelen M, Vonstein V, Wallace ZS, Warren AS, Wattam AR, Xia F, Yoo H, Zhang Y, Zmasek CM, Scheuermann RH, Stevens RL. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res.* **2023,** 51(D1):D678-D689. doi: 10.1093/nar/gkac1003.
17. Meier-Kolthoff J. P., Sardà Carbasse J., Peinado-Olarte R. L., Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acid Res.* **2022**, 50, D801–D807. doi: 10.1093/nar/gkab902
18. Dieckmann, M.A.; Beyvers, S.; Nkouamedjo-Fankep, R.C.; Hanel, P.H.G.; Jelonek, L.; Blom, J.; Goesmann, A. EDGAR3.0: Comparative genomics and phylogenomics on a scalable infrastructure. *Nucleic Acids Res*. **2021**, 49, W185–W192..
19. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; van Wezel, G.P.; Medema, M.H.; Weber, T. antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, 49, W29–W35..
20. Bachmann BO, Ravel J.. Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **2009**, 458:181-217. doi: 10.1016/S0076-6879(09)04808-3.
21. van Heel, A.J.; de Jong, A.; Song, C.; Viel, J.H.; Kok, J.; Kuipers, O.P. BAGEL4: A user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* **2018**, 46, W278–W281..
22. Vater J, Herfort S, Doellinger J, Weydmann M, Borriss R, Lasch P. Genome Mining of the Lipopeptide Biosynthesis of Paenibacillus polymyxa E681 in Combination with Mass Spectrometry: Discovery of the Lipoheptapeptide Paenilipoheptin. *Chembiochem.* **2018**, 19(7):744-753. doi: 10.1002/cbic.201700615
23. Mülner P, Schwarz E, Dietel K, Junge H, Herfort S, Weydmann M, Lasch P, Cernava T, Berg G, Vater J. Profiling for Bioactive Peptides and Volatiles of Plant Growth Promoting Strains of the *Bacillus subtilis* Complex of Industrial Relevance. *Front Microbiol.* **2020**, 11:1432. doi: 10.3389/fmicb.2020.01432
24. Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J., and Holle, A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem*. **2003**, 376, 952-956. doi: 10.1007/s00216-003-2057.0

25. Hooper D.J., Hallmann J., Subbotin S.A. Methods for extraction, processing and detection of plant and soil nematodes. In: Luc M., Sikora R., Bridge J., editors. Plant Parasitic Nematodes in Subtropical and Tropical Agriculture. CAB International; Wallingford, UK**: 2005**. pp. 53–86..
26. Bridge J., Page S.L.J. Estimation of Root Knot Nematode Infestation Levels in Roots Using a Rating Chart. Trop. Pest Manag. **1980**, 26:296–298. doi: 10.1080/09670878009414416..
27. Budiharjo A, Chowdhury SP, Dietel K, Beator B, Dolgova O, Fan B, Bleiss W, Ziegler J, Schmid M, Hartmann A, Borriss R. Transposon mutagenesis of the plant-associated Bacillus amyloliquefaciens ssp. plantarum FZB42 revealed that the nfrA and RBAM17410 genes are involved in plant-microbe-interactions. *PLoS One*, **2014**, 9(5):e98267. doi: 10.1371/journal.pone.0098267
28. Meier-Kolthoff, J.P., Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* **10**, 2182 https://doi.org/10.1038/s41467-019-10210-3
29. Lefort V, Desper R, Gascuel O. (). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol*. **2015**, 32: 2798–2800. DOI: 10.1093/molbev/msv150
30. Cheng T, Lin P, Jin S, Wu Y, Fu B, Long R, Liu D, Guo Y, Peng L, Xia Q. Complete Genome Sequence of Bacillus bombysepticus, a Pathogen Leading to Bombyx mori Black Chest Septicemia. *Genome Announc*. **2014**, 2(3):e00312-14. doi: 10.1128/genomeA.00312-14.
31. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **2022,** 50(D1):D785-D794. doi: 10.1093/nar/gkab776
32. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol.* **2020**, 70(11):5607-5612. doi: 10.1099/ijsem.0.004332.
33. Agata N., Ohta M. & Mori M. Production of an emetic toxin, cereulide, is associated with a specific class of *Bacillus cereus*. *Curr Microbiol* **1996**, 33, 67–69.
34. Baillie L. & Read T. D. *Bacillus anthracis*, a bug with attitude! *Curr Opin Microbiol* **2001**, 4, 78–81.
35. Bravo A, Gómez I, Porta H, García-Gómez BI, Rodriguez-Almazan C, Pardo L, Soberón M. Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. *Microb Biotechnol* **2013**, 6, 17–26.
36. Radnedge L, Agron PG, Hill KK, Jackson PJ, Ticknor LO, Keim P, Andersen GL. Genome differences that distinguish Bacillus anthracis from Bacillus cereus and Bacillus thuringiensis. *Appl Environ Microbiol*. **2003**, 69(5):2755-64. doi: 10.1128/AEM.69.5.2755-2764.2003.
37. Yang S, Wang Y, Liu Y, Jia K, Zhang Z, Dong Q. Cereulide and Emetic *Bacillus cereus*: Characterizations, Impacts and Public Precautions. *Foods*. **2023**, 12(4):833. doi: 10.3390/foods12040833.
38. Beecher DJ, Wong ACL. Cooperative, synergistic and antagonistic haemolytic interactions between haemolysin BL, phosphatidylcholine phospholipase C and sphingomyelinase from Bacillus cereus. *Microbiology* (Reading). **2000**, 146 Pt 12:3033-3039. doi: 10.1099/00221287-146-12-3033
39. Dietrich R, Jessberger N, Ehling-Schulz M, Märtlbauer E, Granum PE. The Food Poisoning Toxins of *Bacillus cereus*. *Toxins* (Basel). **2021**, 13(2):98. doi: 10.3390/toxins13020098.
40. Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P. Bacillus thuringiensis: a century of research, development and commercial applications. *Plant Biotechnol J*. **2011**, 9(3):283-300. doi: 10.1111/j.1467-7652.2011.00595.x.
41. Gupta M, Kumar H, Kaur S. Vegetative Insecticidal Protein (Vip): A Potential Contender From *Bacillus thuringiensis* for Efficient Management of Various Detrimental Agricultural Pests. *Front Microbiol*. **2021**, 12:659736. doi: 10.3389/fmicb.2021.659736
42. Zheng D, Zeng Z, Xue B, Deng Y, Sun M, Tang YJ, Ruan L. Bacillus thuringiensis produces the lipopeptide thumolycin to antagonize microbes and nematodes. *Microbiol Res.* **2018**, 215:22-28. doi: 10.1016/j.micres.2018.06.004
43. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014 Jan;42(Database issue):D206-14. doi: 10.1093/nar/gkt1226
44. Griffith OH, Ryan M. Bacterial phosphatidylinositol-specific phospholipase C: structure, function, and interaction with lipids. *Biochim Biophys Acta*. **1999**, 1441(2-3):237-54. doi: 10.1016/s1388-1981(99)00153-5
45. Gao Y, Cao D, Zhu J, Feng H, Luo X, Liu S, Yan XX, Zhang X, Gao P. Structural insights into assembly, operation and inhibition of a type I restriction-modification system. *Nat Microbiol.* **2020**, 5(9):1107-1118. doi: 10.1038/s41564-020-0731-z
46. Dong S, McPherson SA, Wang Y, Li M, Wang P, Turnbough CL Jr, Pritchard DG. Characterization of the enzymes encoded by the anthrose biosynthetic operon of Bacillus anthracis. *J Bacteriol.* **2010**, 192(19):5053-62. doi: 10.1128/JB.00568-10.
47. Dong S, McPherson SA, Tan L, Chesnokova ON, Turnbough CL Jr, Pritchard DG. Anthrose biosynthetic operon of Bacillus anthracis. *J Bacteriol*. **2008**, 190(7):2350-9. doi: 10.1128/JB.01899-07

48. Yoshida K, Yamaguchi M, Morinaga T, Kinehara M, Ikeuchi M, Ashida H, Fujita Y. myo-Inositol catabolism in Bacillus subtilis. *J Biol Chem*. **2008**, 283(16):10415-24. doi: 10.1074/jbc.M708043200
49. Zhao X, Kuipers OP. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. *BMC Genomics*. **2016**, 17(1):882. doi: 10.1186/s12864-016-3224-y. PMID: 27821051
50. Tracanna, V.; de Jong, A.; Medema, M.H.; Kuipers, O.P. Mining prokaryotes for antimicrobial compounds: from diversity to function. *FEMS Microbiol Rev.* **2017,** 41(3),417-429. doi: 10.1093/femsre/fux014. PMID: 28402441
51. Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, Lee S, Meijer D, Recchia MJJ, Reitz ZL, van Santen JA, Selem-Mojica N, Tørring T, Zaroubi L, Alanjary M, Aleti G, Aguilar C, Al-Salihi SAA, Augustijn HE, Avelar-Rivas JA, Avitia-Domínguez LA, Barona-Gómez F, Bernaldo-Agüero J, Bielinski VA, Biermann F, Booth TJ, Carrion Bravo VJ, Castelo-Branco R, Chagas FO, Cruz-Morales P, Du C, Duncan KR, Gavriilidou A, Gayrard D, Gutiérrez-García K, Haslinger K, Helfrich EJN, van der Hooft JJJ, Jati AP, Kalkreuter E, Kalyvas N, Kang KB, Kautsar S, Kim W, Kunjapur AM, Li YX, Lin GM, Loureiro C, Louwen JJR, Louwen NLL, Lund G, Parra J, Philmus B, Pourmohsenin B, Pronk LJU, Rego A, Rex DAB, Robinson S, Rosas-Becerra LR, Roxborough ET, Schorn MA, Scobie DJ, Singh KS, Sokolova N, Tang X, Udwary D, Vigneshwari A, Vind K, Vromans SPJM, Waschulin V, Williams SE, Winter JM, Witte TE, Xie H, Yang D, Yu J, Zdouc M, Zhong Z, Collemare J, Linington RG, Weber T, Medema MH (2023). MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Res. 51(D1):D603-D610. doi: 10.1093/nar/gkac1049. PMID: 36399496; PMCID: PMC9825592.
52. Finking, R.; Marahiel, M.A. Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol.* **2004**, 58,453-88. doi: 10.1146/annurev.micro.58.030603.123615.
53. Béchet M, Caradec T, Hussein W, Abderrahmani A, Chollet M, Leclère V, Dubois T, Lereclus D, Pupin M, Jacques P. Structure, biosynthesis, and properties of kurstakins, nonribosomal lipopeptides from Bacillus spp. *Appl Microbiol Biotechnol.* **2012,** 95(3):593-600. doi: 10.1007/s00253-012-4181-2. PMID: 22678024
54. Yu YY, Zhang YY, Wang T, Huang TX, Tang SY, Jin Y, Mi DD, Zheng Y, Niu DD, Guo JH, Jiang CH (2023). Kurstakin Triggers Multicellular Behaviors in *Bacillus cereus* AR156 and Enhances Disease Control Efficacy Against Rice Sheath Blight. *Plant Dis.* **2023,** 107(5):1463-1470. doi: 10.1094/PDIS-01-22-0078-RE
55. Luo C, Liu X, Zhou X, Guo J, Truong J, Wang X, Zhou H, Li X, Chen Z. Unusual Biosynthesis and Structure of Locillomycins from Bacillus subtilis 916. *Appl Environ Microbiol*. **2015**, 81(19):6601-9. doi: 10.1128/AEM.01639-15
56. May JJ, Wendrich TM, Marahiel MA. The dhb operon of Bacillus subtilis encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin. *J Biol Chem*. **2001**, 276(10):7209-17. doi: 10.1074/jbc.M009140200
57. Chen XH, Koumoutsi A, Scholz R, Borriss R. More than anticipated - production of antibiotics and other secondary metabolites by Bacillus amyloliquefaciens FZB42. *J Mol Microbiol Biotechnol*. **2009**, 16(1-2):14-24. doi: 10.1159/000142891.
58. Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, Morgenstern B, Voss B, Hess WR, Reva O, Junge H, Voigt B, Jungblut PR, Vater J, Süssmuth R, Liesegang H, Strittmatter A, Gottschalk G, Borriss R. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium Bacillus amyloliquefaciens FZB42. *Nat Biotechnol*. **2007**, 25(9):1007-14. doi: 10.1038/nbt1325.
59. Kevany BM, Rasko DA, Thomas MG. Characterization of the complete zwittermicin A biosynthesis gene cluster from Bacillus cereus. *Appl Environ Microbiol*. **2009**, 75(4):1144-55. doi: 10.1128/AEM.02518-08
60. Fellbrich G, Romanski A, Varet A, Blume B, Brunner F, Engelhardt S, Felix G, Kemmerling B, Krzymowska M, Nürnberger T. NPP1, a Phytophthora-associated trigger of plant defense in parsley and Arabidopsis. *Plant J.* **2002**, 32(3):375-90. doi: 10.1046/j.1365-313x.2002.01454.x
61. Burkhart BJ, Hudson GA, Dunbar KL, Mitchell DA. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat Chem Biol*. **2015**, 11(8):564-70. doi: 10.1038/nchembio.1856 Genome-Mining Tool for Class-Independent RiPP Discovery. *mSystems*. **2020,** 5(5):e00267-
62. Kloosterman AM, Shelton KE, van Wezel GP, Medema MH, Mitchell DA. RRE-Finder: a 20. doi: 10.1128/mSystems.00267-20
63. Sahl HG, Jack RW, Bierbaum G. Biosynthesis and biological activities of lantibiotics with unique post-translational modifications. *Eur J Biochem*. **1995**, 230(3):827-53. doi: 10.1111/j.1432-1033.1995.tb20627.x
64. Wang J, Ma H, Ge X, Zhang J, Teng K, Sun Z, Zhong J. Bovicin HJ50-like lantibiotics, a novel subgroup of lantibiotics featured by an indispensable disulfide bridge. *PLoS One*. **2014**, 9(5):e97121. doi: 10.1371/journal.pone.0097121
65. Walker MC, Eslami SM, Hetrick KJ, Ackenhusen SE, Mitchell DA, van der Donk WA. Precursor peptide-targeted mining of more than one hundred thousand genomes expands the lanthipeptide natural product family. *BMC Genomics*. **2020**, 21(1):387. doi: 10.1186/s12864-020-06785-7

66. Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJ, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süssmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey JM, van der Donk WA. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep*. **2013**, 30(1):108-60. doi: 10.1039/c2np20085f
67. Ren H, Biswas S, Ho S, van der Donk WA, Zhao H. Rapid Discovery of Glycocins through Pathway Refactoring in Escherichia coli. *ACS Chem Biol*. **2018**, 13(10):2966-2972. doi: 10.1021/acschembio.8b00599
68. Oman TJ, Boettcher JM, Wang H, Okalibe XN, van der Donk WA. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nat Chem Biol*. **2011**, 7(2):78-80. doi: 10.1038/nchembio.509
69. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Lasso peptides: an intriguing class of bacterial natural products. *Acc Chem Res.* **2015**, 48(7):1909-19. doi: 10.1021/acs.accounts.5b00156
70. Zhu S, Hegemann JD, Fage CD, Zimmermann M, Xie X, Linne U, Marahiel MA. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. *J Biol Chem*. **2016**, 291(26):13662-78. doi: 10.1074/jbc.M116.722108
71. Flühe L, Knappe TA, Gattner MJ, Schäfer A, Burghaus O, Linne U, Marahiel MA. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. *Nat Chem Biol.* **2012**, 8(4):350-7. doi: 10.1038/nchembio.798. Erratum in: Nat Chem Biol. 2012 Aug;8(8):737
72. Hudson GA, Burkhart BJ, DiCaprio AJ, Schwalen CJ, Kille B, Pogorelov TV, Mitchell DA. Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J Am Chem Soc.* **2019**, 141(20):8228-8238. doi: 10.1021/jacs.9b01519
73. Chopra L, Singh G, Choudhary V, Sahoo DK. Sonorensin: an antimicrobial peptide, belonging to the heterocycloanthracin subfamily of bacteriocins, from a new marine isolate, Bacillus sonorensis MT93. *Appl Environ Microbiol*. **2014**, 80(10):2981-90. doi: 10.1128/AEM.04259-13
74. Franz CM, van Belkum MJ, Holzapfel WH, Abriouel H, Gálvez A. Diversity of enterococcal bacteriocins and their grouping in a new classification scheme. *FEMS Microbiol Rev*. **2007**, 31(3):293-310. doi: 10.1111/j.1574-6976.2007.00064.x
75. Aunpad R, Panbangred W. Evidence for two putative holin-like peptides encoding genes of Bacillus pumilus strain WAPB4. *Curr Microbiol*. **2012**, 64(4):343-8. doi: 10.1007/s00284-011-0074-3
76. Lee JY, Janes BK, Passalacqua KD, Pfleger BF, Bergman NH, Liu H, Håkansson K, Somu RV, Aldrich CC, Cendrowski S, Hanna PC, Sherman DH. Biosynthetic analysis of the petrobactin siderophore pathway from Bacillus anthracis. *J Bacteriol*. **2007**, 189(5):1698-710. doi: 10.1128/JB.01526-06
77. Koppisch AT, Dhungana S, Hill KK, Boukhalfa H, Heine HS, Colip LA, Romero RB, Shou Y, Ticknor LO, Marrone BL, Hersman LE, Iyer S, Ruggiero CE. Petrobactin is produced by both pathogenic and non-pathogenic isolates of the Bacillus cereus group of bacteria. *Biometals*. **2008**, 21(5):581-9. doi: 10.1007/s10534-008-9144-9
78. Yuan S, Yong X, Zhao T, Li Y, Liu J. Research Progress of the Biosynthesis of Natural Bio-Antibacterial Agent Pulcherriminic Acid in *Bacillus*. *Molecules*. **2020**, 25(23):5611. doi: 10.3390/molecules25235611
79. Corre C, Song L, O'Rourke S, Chater KF, Challis GL. 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by Streptomyces coelicolor genome mining. *Proc Natl Acad Sci U S A*. **2008**, 105(45):17510-5. doi: 10.1073/pnas.0805530105
80. Arie T. *Fusarium* diseases of cultivated plants, control, diagnosis, and molecular and genetic studies. *J Pestic Sci*. **2019**, 44(4):275-281. doi: 10.1584/jpestics.J19-03
81. Chi, N.M.; Thu, P.Q.; Nam, H.B.; Quang, D.Q.; Phong, L.V.; Van, N.D.; Trang, T.T.; Kien, T.T.; Tam, T.T.T; Dell, B. Management of *Phytophthora palmivora* disease in *Citrus reticulata* with chemical fungicides. *J Gen Plant Pathol.* **2020**, 86, 494–502. https://doi.org/10.1007/s10327-020-00953-z
82. Eisenback, J. D.; and Triantaphyllou, H.H. Root-knot nematodes: Meloidogyne species and races. Pp. 191-274, In, W. R. Nickle, ed., *Manual of Agricultural Nematology*, Marcell Dekker: **1991** New York
83. Hussey, P.S.; Barker, K.R. A comparison of methods of collecting inocula of Meloidogyne spp., including a new technique. *Plant Disease Reporter*. **1973**, 57, 1025-1028
84. Chowdhury SP, Hartmann A, Gao X, Borriss R. Biocontrol mechanism by root-associated Bacillus amyloliquefaciens FZB42 - a review. *Front Microbiol*. **2015**, 6:780. doi: 10.3389/fmicb.2015.00780
85. Qiao J.-Q., Wu H.-J., Huo R., Gao X.-W., Borriss R. Stimulation and biocontrol by *Bacillus amyloliquefaciens* subsp. plantarum FZB42 engineered for improved action. *Chem. Biol. Technol. Agric.* **2014**, 1:12 10.1186/s40538-014-0012-2

86. Wu L, Wu HJ, Qiao J, Gao X, Borriss R. Novel Routes for Improving Biocontrol Activity of Bacillus Based Bioinoculants. *Front Microbiol.* **2015**, 6:1395. doi: 10.3389/fmicb.2015.01395.