JOBIMB, 2024, Vol 12, No 2, 61-68
https://doi.org/10.54987/jobimb.v12i2.1067

JOURNAL OF BIOCHEMISTRY, MICROBIOLOGY AND BIOTECHNOLOGY

Website: http://journal.hibiscuspublisher.com/index.php/JOBIMB/index

JOBIMB VOL 12 NO 2 2024
Bacterial acrylamidase

Genomic Analysis of Antibiotics and Secondary Metabolites Biosynthetic Genes Clusters in *Bacillus cereus* Group

Tasiu Mahmud[1]*, Ibrahim Alhaji Sabo[2], Yahaya Ubah Ya'u[3] and Farida Isah El-hassan[1]

[1]Department of Microbiology, Faculty of Science, Aliko Dangote University of Science and Technology, P.M.B 3244, Wudil, Kano State, Nigeria.
[2]Department of Microbiology, Faculty of Pure and Applied Sciences, Federal University Wukari, P.M.B 1020, Wukari, Taraba State, Nigeria.
[3]Department of Science Laboratory Technology, Faculty of Science, Aliko Dangote University of Science and Technology, P.M.B 3244, Wudil, Kano State, Nigeria.

*Corresponding author:
Tasiu Mahmud
Department of Microbiology,
Faculty of Science,
Aliko Dangote University of Science and Technology,
P.M.B 3244, Wudil,
Kano State,
Nigeria.
Email: tasiu.microlabs@gmail.com

HISTORY

Received: 19th Sep 2024
Received in revised form: 21st Nov 2024
Accepted: 25th Dec 2024

KEYWORDS

Clusters
Bioinformatics
Genome
Antibiotics
Secondary metabolites

ABSTRACT

Genome mining of the *Bacillus cereus* group using bioinformatic web resource was conducted to explore the antibiotic potentials of this highly similar group of bacteria. *Bacillus cereus* group consists of eight closely related *Bacillus* species, they are gram-positive, spore-forming, aerobic, facultative anaerobic rod-shaped bacteria, and they have low G+C-content when compared with *B. subtilis*. They include *B. cereus, B. anthracis, B. thuringiensis, B. mycoides, B. pseudomycoides B. weihenstephanensis, B. cytotoxicus, B. toyonensis.* The genera *Bacillus* are known to produce antibiotic compounds through biosynthetic gene clusters in their genome which are responsible for the formation of single or multiple natural specialized products. In the present study, whole genome sequences of Eight strains representing each of the *B. cereus* group bacteria: *B. cereus* NC7401, *B. anthracis* strain STI1, *B. thuringiensis* strain TG-5, *B. pseudomycoides* strain 219298, *B. mycoides* strain Gnyt1, *B. cytotoxicus* strain E8.1, *B. toyonensis* strain UTDF19-29B and *B. weihenstephanensis* strain WSBC 10204 were retrieved from GeneBank database of the National Center For Biotechnology Information (NCBI). The genomic data was analyzed for antibiotic and secondary metabolites biosynthetic gene clusters (BGCs) using antibiotic and secondary metabolites analysis shell version 4.0 (antiSMASH 4.0) pipeline with default parameters. The average size of the genome was approximately 5.035 mb, however, *Bacillus weihenstephanensis* strain WSBC 10204 has the highest genome size of 5,608,349 bp. The findings of this study identified and reported up to 76 biosynthetic gene clusters in the 8 genomes of the *B. cereus* group. These clusters were classified into 14 different BGCs, which include 14 RiPP-like, 16 NRPS, 8 NPR-metallophore, 8 betalactone. Comparatively, RRE-containing, Type I polyketides synthase (T1PKS), and Ras-RiPP clusters were scarce across the 8 genomes, and they are only found in the genome of *B. cytotoxicus*. Similarly, there are 33 (49.25%) unknown clusters for which no known homologous or similar BGCs could be identified, this revealed the potential novelty associated with these clusters.

INTRODUCTION

Recent advancements in molecular techniques such as whole genome sequencing coupled with bioinformatics tools have become a powerful tool for the identification of biosynthetic gene clusters (BGCs), thereby expanding natural product discovery [1-2]. The genome mining approach has revealed the enormous potential of *Bacillus cereus* group for producing specialized natural products that can be transformed into useful drugs [3-4]. Biosynthetic gene clusters are organized groups of two or more genes on the bacterial chromosomes that together encode for enzymes responsible for the biosynthesis of specialized metabolites [5]. A bacterial genome usually contains 20-40 biosynthetic gene clusters; one biosynthetic gene cluster encodes for an enzyme that can catalyze the formation of one or several similar natural compounds. Genomic studies in bacteria have

shown that nearly 30-40% of all genes are organized in an operonic structure [6]. *Bacillus cereus* group consists of eight closely related *Bacillus* species; they are gram-positive, spore-forming, aerobic, facultative anaerobic, rod-shaped bacteria, and they have low G+C-content when compared with *B. subtilis* [7].

The *Bacillus cereus* group members are also known as *B. cereus* sensu lato and include *B. cereus, B. anthracis, B. thuringiensis, B. mycoides, B. pseudomycoides, B. weihenstephanensis, B. cytotoxicus, B. toyonensis* [8]. However, there are three extensively studied species among the group due to their pathogenic potential, and these include *B. anthracis,* which causes anthrax in humans and animals, *B. cereus*, which causes food-borne infection and *B. thuringiensis* which is an insect (Lepidopteron) pathogen [9].

Antibiotics are chemical compounds produced by microorganisms (bacteria and fungi) that inhibit or destroy other microorganisms; however, microbial secondary metabolites are low molecular mass products that are used for several purposes, including antibiotics, antitumor agents, enzyme inhibitors, etc. Out of 22,000 known microbial secondary metabolites, nearly 1540 (7%) were produced by *Bacillus* species [10]. The objective of the study is to identify the distribution of biosynthetic gene clusters that encode for antibiotics and secondary metabolites in the genomes of *Bacillus cereus* group using a genome mining approach, this can explore the potential of *B. cereus* group for specialized metabolite and offer new avenue for antibiotic discovery and development [11].

METHODOLOGY

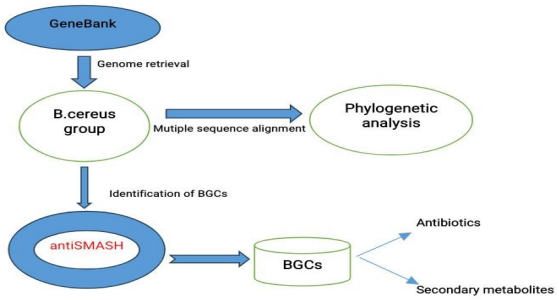

Fig.1. Overview of genome mining approaches to identify BGCs in the *B. cereus* group.

Data collection

Whole genome sequences of eight bacterial strains representing the *B. cereus* group were retrieved from the NCBI GeneBank database (https://www.ncbi.nlm.nih.gov/). The bacterial strains include *Bacillus anthracis* strain *STI1*, *Bacillus pseudomycoides* strain *219298, Bacillus cereus NC7401, Bacillus mycoides* strain *Gnyt1*, *Bacillus thuringiensis* strain *TG-5*, *Bacillus cytotoxicus* strain *E8.1*, *Bacillus toyonensis* strain *UTDF19-29B* and *Bacillus weihenstephanensis* strain *WSBC 10204* [2-5]**.**

Phylogenetic analysis

To determine the genetic relationship between *B. cereus* and other closely related members of the group, the sequences of the 16S rRNA gene were filtered from the genome and used for the multiple sequence alignment using the BLASTn algorithm. The phylogenetic tree was generated from the multiple sequence alignment using the neighbor-joining method with a maximum sequence difference of 0.75 [3].

Identification of antibiotics and Secondary metabolites biosynthetic genes clusters

In order to identify the genes involved in the production of antibiotics and secondary metabolites, the genome of each bacterial strain was mined using an antibiotic and secondary metabolite analysis shell (antiSMASH v4.0) pipeline [2-3]. The detection strictness was set to 'relaxed' and all other important parameters such as ClusterBlast, KnownClusterBlast, SubClusterBlast, MIBiG Cluster comparisons, Activesite Finder, RREFinder Cluster Pfam analysis, Pfam-based GO term annotation, TFBS analysis, TIGRFam analysis were enabled with default settings. The antiSMASH analysis was determined, and the summary of the annotated BGCs and region of each identified cluster was generated [12].

RESULTS AND DISCUSSION

Phylogenetic results based on 16S rRNA gene sequences of the *B. cereus* group showed that *B. anthracis* species is more closely related to the *B. cereus* species with 99.93% similarity; they are located on the same branch (**Fig. 2**). However, the other members of the *Bacillus cereus* group which also showed similarities with *B. cereus* species include *Bacillus thuringiensis* strain TG-5 (99.66%), *B. weihenstephanensis* strain WSBC 10204 (99.62%), *B. toyonensis* strain UTDF19-29B (99.59), *B. pseudomycoides* strain 219298 (99.59%), *Bacillus mycoides* strain Gnyt1 (99.38%) and *Bacillus cytotoxicus* strain E8.1 (98.00%). These findings were consistent with previous reports of Ehling-Schulz et al. [7] which showed that *B. anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* were the most closely related species of the *B. cereus* group.

The average size of the genome was approximately 5.035mb, with *Bacillus weihenstephanensis* strain *WSBC 10204* having the highest genome size of 5,608,349 bp, however, the genome size of the other members of the *Bacillus cereus* group include; *Bacillus mycoides* strain Gnyt1 *(*5,597,907 bp), *Bacillus thuringiensis* strain TG-5 (5,432,867 bp), *Bacillus toyonensis* strain UTDF19-29B (5,240,743), *Bacillus anthracis* strain STI1*(*5,229,731 bp), *Bacillus pseudomycoides* strain 219298 (5,228,024 bp), *Bacillus anthracis* strain STI1 *(*5,221,581 bp) and *Bacillus cytotoxicus* strain E8.1 having the smallest genome size of 3.99 mb (4,132,005 bp) (**Figs. 3** to **10**).

The results of the antSMASH analysis were presented in **Tables 1-8**, the clusters were classified into 14 different BGCs, which include 16 NRP (21.05%), 14 RiPP-like(18.42%), 8 NPR-metallophore (10.52%), 8 betalactone (10.52%), 8 terpene (10.52%), 5 NI-siderophore (6.57%), 4 NRP-like (5.26%), 3 lassopeptide (3.94%), 2 ranthipeptide (2.63 %).While, RRE-containing, Type I polyketides synthase (T1PKS), and Ras-PiPP clusters were scarce across the 8 genomes and they are only found in the genome of *B. cytotoxicus*. The finding of this study showed that there was a relatively homologous distribution of BGCs among the *B. cereus* group; for instance, the cluster of NRPS has high abundance in each of the *Bacillus* genome analyzed; however, there is a uniform distribution of NRP-metallaphore, betalactone and terpene across the 8 genomes. This finding was largely similar to another recent study by Grubbs et al. [3].

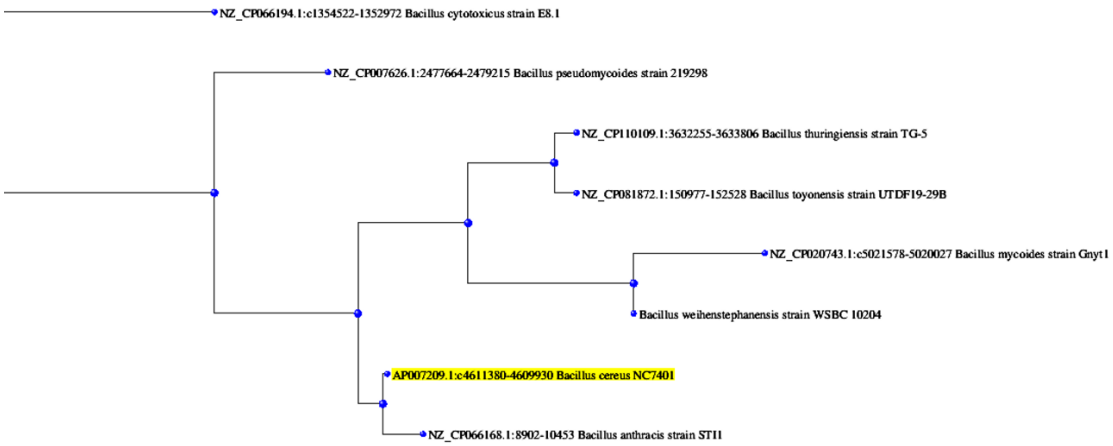

Fig. 2. Phylogenetic tree of multiple sequence alignment of eight *B. cereus group*. The tree was constructed using the neighbor-joining method, which had the maximum sequence difference of 0.75.

Table 1. Identified secondary metabolites region using strictness 'relaxed from *Bacillus cereus* NC7401.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1. (368,212-411,400) nt	NRPS	Thailastatin A , NRP+PK	10
2. (1,273,573-1,297,080)nt	LAP	-	-
3.(2,240,152-2,291,878)nt	NRP-metallophore,NRPS	Bacillibactin, NRP	71
4.(2,382,301-2,407,539)nt	Betalactone	Fengycin, NRP	40
5.(2,445,906-2,456,187)nt	RiPP-like	-	-
6. (2,506,802-2,517,059)nt	PiPP-like	-	-
7. (3,275,699-3,297,552)nt	Terpene	Molybdenum cofactor, Other	-

Key: NRPS-Non-ribosomal peptide, NRP-Non-ribosomal peptide, PK-Polyketide, RiPP-Ribosomally synthesize post-translationally modified peptide, LAP- linea azoline containing peptides, nt-Nucleotide.

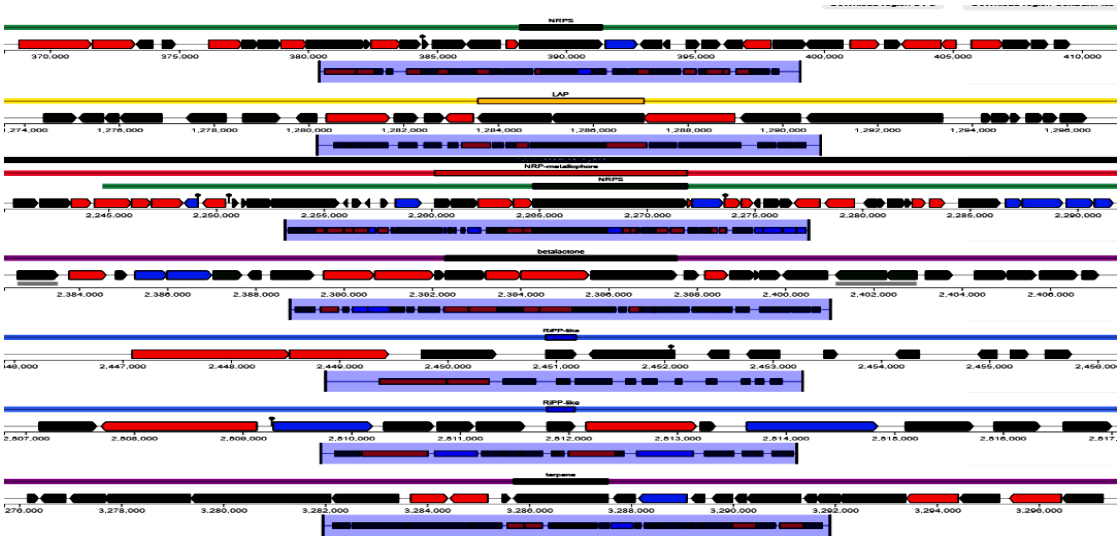

Fig. 3. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus cereus NC7401*.

Table 2. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus anthracis* strain STI1.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1. (1,208,117-1,231,624)n	LAP	-	-
2. (1,851,197-1,882,902)nt	NI-siderophore	Petrobactin, Other	100
3. (2,182,467-2,234,206)nt	NRP-metallophone,NRPS	Bacillibactin, NRP	85
4.(2,363,819-2,389,057)nt	Betalactone	Fengycin NRP	40
5.(2,490,513-2,500,779)nt	RiPP-like	-	-
6. (2,513,360-2,560,370)nt	NRPS	-	-
7. (3,311,990-3,333,843)nt	Terpene	Molybdenum cofactor, Other	17

Key: NRPS-Non-ribosomal peptide, NRP-Non-ribosomal peptide, PK-Polyketide, RiPP-Ribosomally synthesize post-translationally modified peptide, LAP- linea azoline containing peptides, nt-Nucleotide.

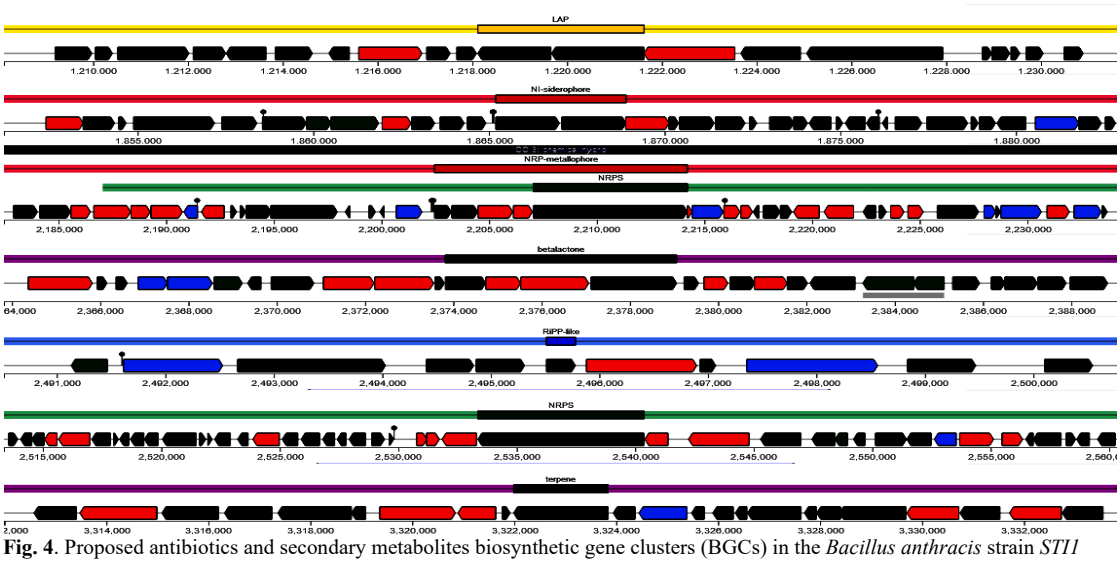

Fig. 4. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus anthracis* strain *STI1*

Table 3. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus thuringiensis* strain TG-5.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1. (224,439-276,187)nt	NRP-metallophore, NRPS	Bacillibactin, NRP	85
2. (462,121-528,029)nt	NRPS	-	-
3. (601,459-626,697)nt	Betalactone	Fengycin, NRP	40
4.(679,217-689,582)nt	RiPP-like	-	-
5.(796,030-806,290)nt	RiPP-like	-	-
6. (822,885-896,898)nt	NRPS	-	-
7. (1,615,320-1,637,173)nt	Terpene	Molybdenum cofactor, other	17
8. (3,886,052-3,929,633)nt	NRPS-like	-	-
9. (4,700,704-4,724,210)nt	LAP	-	-
10. (5,321,134-5,352,841)	NI-siderophore	Petrobactin, other	100

Key: NRPS-Non-ribosomal peptide synthethase, NRP-Non-ribosomal peptide, PK-Polyketide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides.

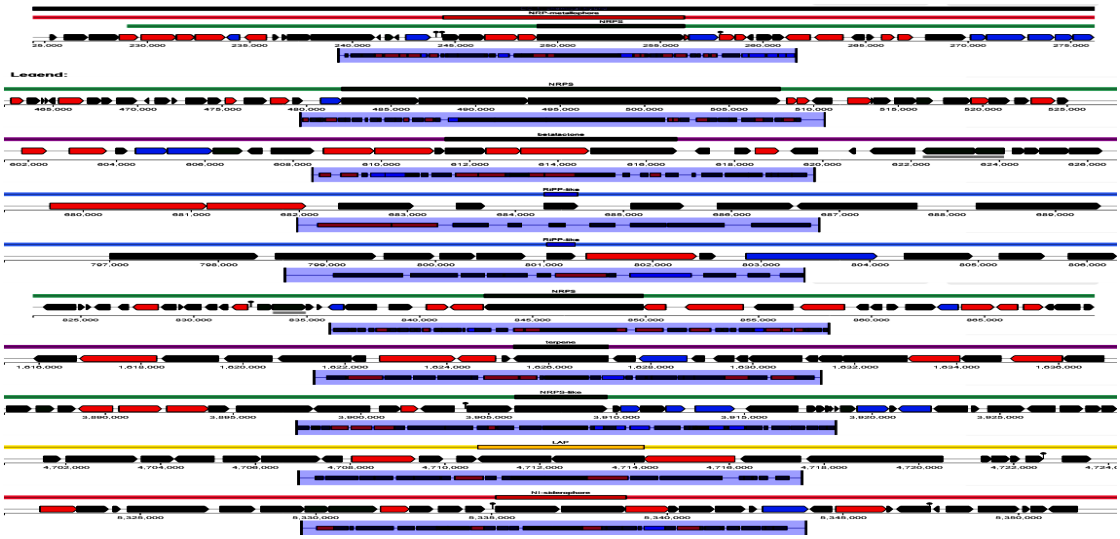

Fig. 5. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus thuringiensis* strain TG-5.

Table 4. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus mycoides* strain Gnyt1.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1 (441,549-485,130)nts	NRPS	-	-
2 (1,303,400-1,326,934)nt	LAP	-	-
3 (1,993,475-2,025,192)nt	NI-siderophore	Petrobactin, other	100
4(2,384,193-2,435,945)nt	NRP-metallophore, NRPS	Bacillibactin, NRP	85
5(2,533,024-2,598,932)nt	NRPS	Saccharothrixin D,E,F,G,H,I,J,K,L,M,PK	6
6 (2,641,249-2,666,486)nt	Betalactone	Fengycin, NRP	40
7 (2,690,177-2,737,034)nt	NRPS	-	-
8(2,759,340-2,769,621)nt	RiPP-like	-	-
9 (3,626,426-3,648,279)nt	Terpene	-	-
10 (3,740,868-3,764,776)nt	Lassopeptides	Paeninodin, RiPP	100

Key: NRPS-Non-ribosomal peptide synthethase, NRP-Non-ribosomal peptide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides.

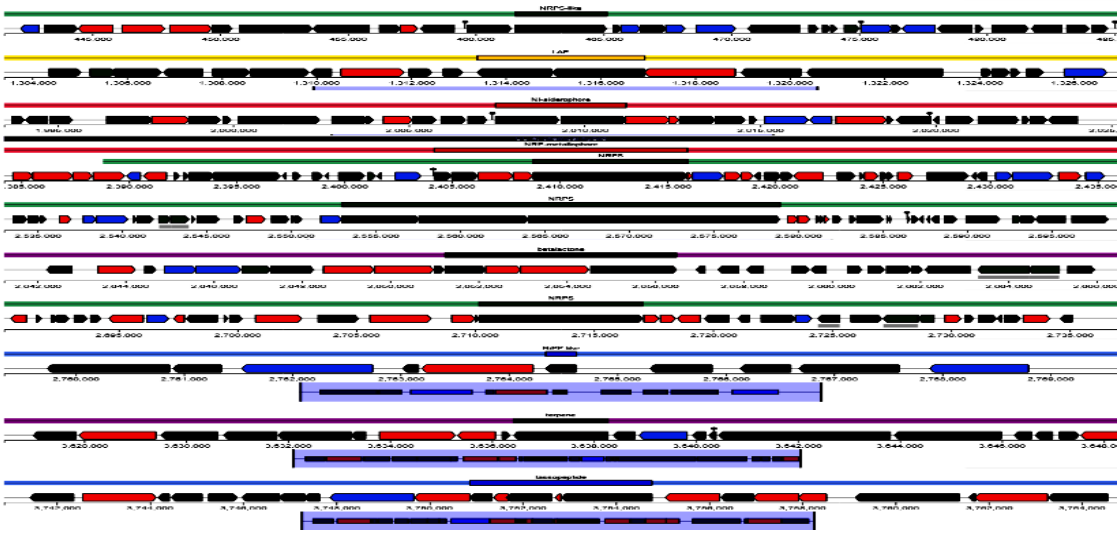

Fig. 6. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus mycoides* strain Gnyt1.

Table 5. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus pseudomycoides* strain 219298.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1 (502,032- 523,891)nts	Terpene	-	-
2 (1,851,197-1,882,902)nt	Lassopeptides	Paeninodin, RiPP	100
3 (3,769,602-3,821,846)nt	Ranthipeptides,,LAP+Thiopeptides	-	-
4(2,363,819-2,389,057)nt	NRPS	Desmamide A, B and C, NRP+PK	18
5(2,490,513-2,500,779)nt	NRP-metallophore, NRPS	Bacillibactin, NRP	85
6 (4,959,067-4,984,304)nt	Betalactone	Fengycin, NRP	40
7 (5,019,001-5,030,867)nt	RiPP-like	-	-

Key: NRPS-Non-ribosomal peptide synthethase, NRP-Non-ribosomal peptide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides, PK-Polyketides.

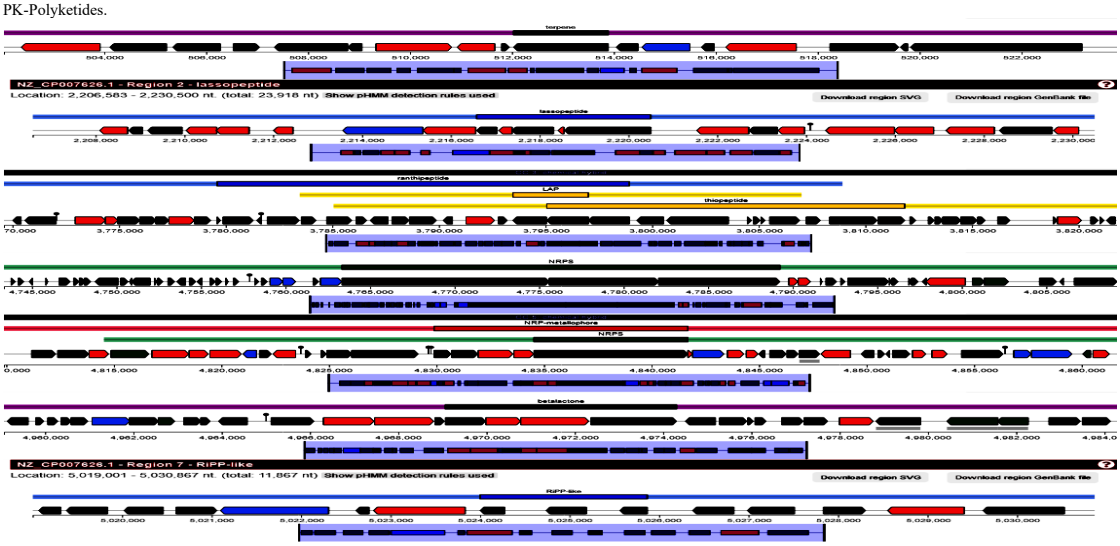

Fig. 7. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus pseudomycoides* strain 219298.

Table 6. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus cytotoxicus* strain E8.1.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1.(524,559- 548,116)nt	LAP	-	-
2.(1,201,621-1,245,226)nt	NRPS-like	Cittilin A/cittilin B RiPP	7
3 (1,468,131-1,490,634)nt	Ras-RiPP	Locillomycin B and C, NRP+PK	14
4.(2,103,626-2,114,066)nt	RiPP-like	-	-
5.(2,443,844-2,500,516)nt	NRPS, NRPS-like, T1PKS	Heme D1, Other	17
6.(3,248,014-3,273,253)nt	Betalactone	Fengycin, NRP	40
7.(3,289,641-3,382,449)nt	NRP-metallophore,NRPS	Bacillibactin, NRP	85
8.(3,469,128-3,490,243)nt	RRE-containing	-	-
9.(3,557,469-3,578,891)nt	Ranthipeptides	-	-
10.(3,621,039-3,642,898)nt	Terpene	-	-

Key: NRPS-Non-ribosomal peptide synthethase, T3PKS- Type I polyketide synthase, NRP-Non-ribosomal peptide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides, RRE-Ribosomally-synthesized recognition element.

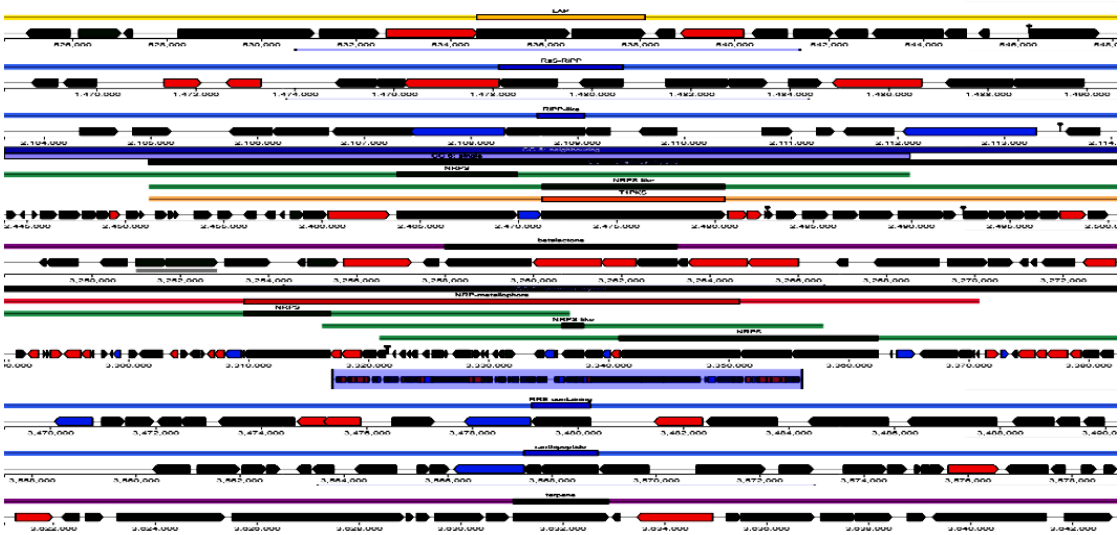

Fig. 8. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus cytotoxicus* strain E8.1.

Table 7. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus toyonensis* strain UTDF19-29B.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1.(405,036-448,617)	NRPS-like	-	-
2.(1,197,300-1,220,806)	LAP	-	-
3.(1,906,193-1,937,906)	NI-siderophore	Petrobactin, Other	100
4.(2,217,489-2,269,238)	NRP-metallophore, NRPS	Bacillibactin, NRP	40
5.(2,406,148-2,431,386)	Betalactone	Fengycin, NRP	40
6.(2,490,915-2,501,274)	RiPP-like	-	-
7.(2,561,488-2,571733)	RiPP-like	-	-
8.(3,360,403-3,382,256)	Terpene	Molybdenum cofactor, other	17
9.(3,492,023-3,515,930)	Lasso peptide	Paeninodin, RiPP	80

Key: NRPS-Non-ribosomal peptide synthethase, NRP-Non-ribosomal peptide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides.

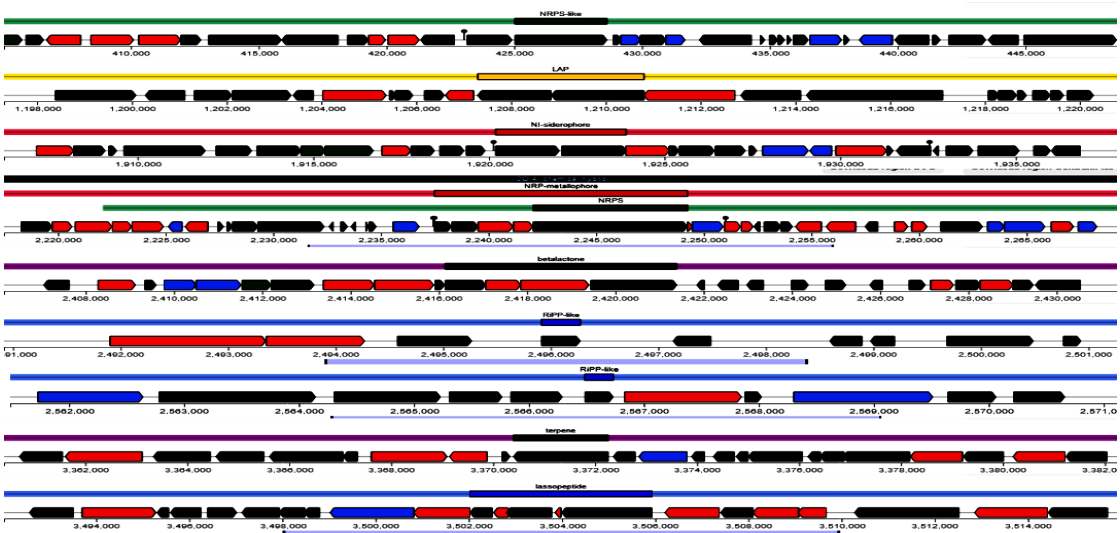

Fig. 9. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus toyonensis* strain UTDF19-29B.

Table 8. Identified antibiotics and secondary metabolites region using strictness 'relaxed' from *Bacillus weihenstephanensis* strain WSBC 10204.

Cluster location	Cluster type	Most similar known cluster	Similarity (%)
1.(1,265,786-1,287,639)nt	Terpene	-	-
2 (2,188,753-2,199,028) nt	RiPP-like	-	-
3 (2,222,147-2,269,157) nt	NRPS	-	-
4 (2,292,145-2,317,383)nt	Betalactone	Fengycin, NRP	40
5(2,432,093-2,442,401)nt	RiPP-like	-	-
6 (2,475,746-2,527,498)nt	NRP-metallophore, NRPS	Bacillibactin, NRP	85
7 (2,833,056-2,864,773)nt	NI-siderophore	Petrobactin, other	100
8 (2,897,496-2,967,400)nt	NRPS, TIPKS	Paenilamicin A2, B1 and B2, NRP+PK	35
9.(3,537,957-3,561,491) nt	LAP	-	-
10.(5,033,998-5,099,930)nt	NRPS	-	-
11.(5,491,939-5,504,158)nt	RiPP-like	-	-
12.(5,567,791-5,579,974)nt	RiPP-like	-	-

Key: NRPS-Non-ribosomal peptide synthetase, T3PKS- Type I polyketide synthase, NRP-Non-ribosomal peptide, RiPP-Ribosomally synthesize post-translationally modified peptide, nt- Nucleotide, LAP- linea azoline containing peptides.

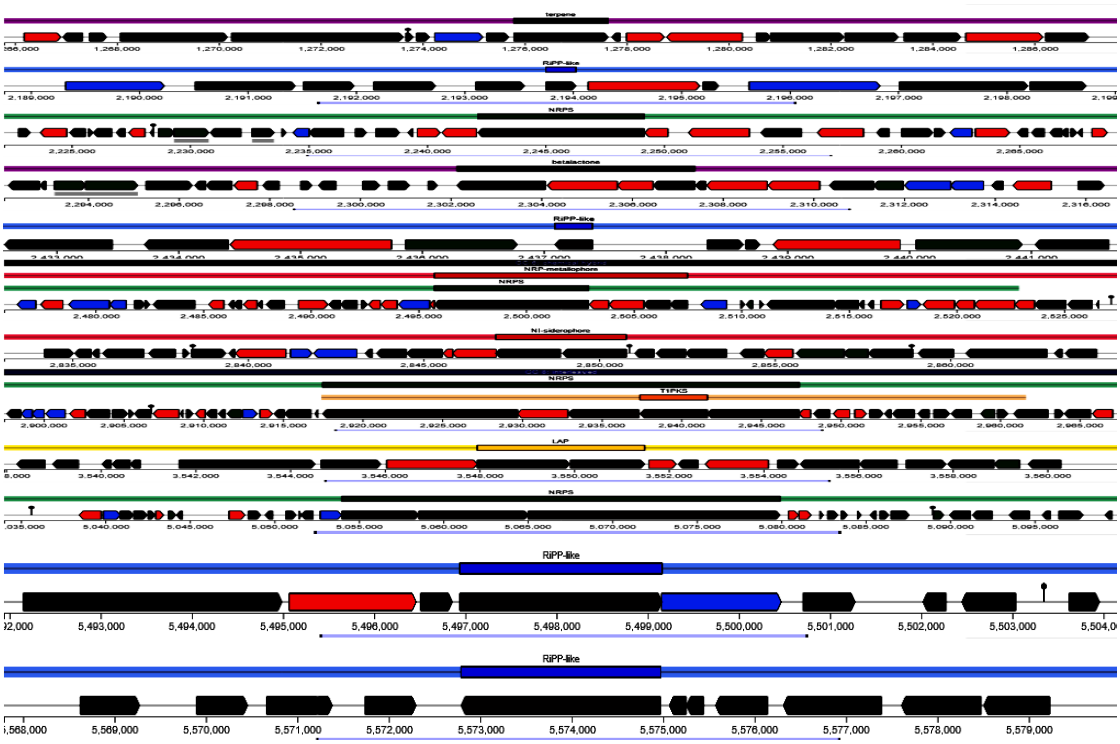

Fig. 10. Proposed antibiotics and secondary metabolites biosynthetic gene clusters (BGCs) in the *Bacillus weihenstephanensis* strain WSBC 10204.

Similarly, the cumulative output from antiSMASH revealed a total of 76 biosynthetic gene clusters across the 8 genomes of the *B. cereus* group; this is equivalent to an average of 9.50 BGCs per genome; this finding is slightly different from the average of BGCs obtained from another study by Liming et al.[4] which reported a higher average of BGCs in *Bacillus* species.

It's important to note that there were 44 (57.99%) known clusters and 32 (42.10%) unknown clusters for which no known homologous or similar BGCs could be identified; this indicated that a substantial number of the BGCs are orphans and may be considered as potential gene clusters for novel antibiotics and secondary metabolites, these clusters do not share any similarity with any known clusters according to antiSMASH genome mining resource, because they do not fit into any other category. Hence, they can be potential genes that may code for a novel class of natural products; the clusters include Non-ribosomal peptides synthetase, linea azoline containing peptides (LAP), non-ribosomal peptides-metallophone, ribosomally-synthesized and unspecified post-translationally modified peptides products (RiPP), terpene, ranthipeptides, non-ribosomal peptides synthetase fragments (NRP-like) and RRE-containing cluster.

The genome with the highest diversity and abundance of biosynthetic gene clusters was *Bacillus weihenstephanensis* strain WSBC 10204; this finding showed that the strain with the largest genome size also had the highest number of clusters. The predicted antibiotics across the 8 genomes include bacillibactin, fengycin, paeninodicin, locillomycin B,C and ranthipeptides.The bioactive natural products predicted by the antiSMASH are thailastatin A, molybdenume cofactor, saccharothrixin D-M, thiopeptides, desmamide A, B, C,cittilin A, B, and heme D1.

CONCLUSION

This study demonstrates that genome mining using bioinformatics tools such as the antiSMASH web resource is reliable for expanding research and exploring bacteria's antibiotic and secondary metabolite potential. The phylogenetic analysis of the *Bacillus cereus* group, which is based on 16S rRNA gene sequences, revealed that *B. anthracis* is closely related to *B. cereus*, with a similarity of 99.93%. The genomic analyses of the *B. cereus* group have highlighted a relatively homologous distribution of the biosynthetic gene clusters (BGCs), having a total of 76 BGCs that are identified across the eight genomes, averaging 9.5 BGCs per genome. Among these, 57.99% are known clusters, and 42.10% are found to be orphan clusters, representing significant potential for the future discovery of novel antibiotics and novel secondary metabolites. Notably, *Bacillus weihenstephanensis* strain WSBC 10204 exhibited the greatest variety and a wealth of BGCs, with 12 biosynthetic gene clusters, implying that the number of biosynthetic gene clusters may depend on the genome size. The study associated clusters for non-ribosomal peptides synthetases (NRPS), ribosomally-synthesized and post-translationally modified peptides (RiPPs), metallophores, and terpenes. While NRPS clusters were abundant across the genomes, unique clusters such as ranthipeptides, RRE-containing clusters, and linea azoline-containing peptides (LAP) were found exclusively in specific genomes, particularly *B. cytotoxicus*. The predicted antibiotics, which include bacillibactin, fengycin, paeninodicin, locillomycin B and C, and ranthipeptides, along with several bioactive products such as thailastatin A, molybdenum cofactor, saccharothrixin D-M, thiopeptides, desmamide A-C, cittilin A-B, and heme D1, has highlighted the immense biosynthetic potential of bacterium from the *B. cereus* group. This diversity underscores the role of *in silico* genome mining in revealing orphan clusters that may code for novel natural products. These findings line up with the aforementioned studies and grant a roadmap for further exploration using advanced molecular techniques, such as next-generation sequencing, CRISPR-Cas9 gene editing, and genome mining. The discovery of novel BGCs and their products implies avenues for advancing innovative antibiotics and focusing on the growing threat of antimicrobial resistance.

DATA AVAILABILITY STATEMENT

The genome sequences that supported the findings of this study are available in the NCBI GeneBank database under the accession number >NZ_CP066168.1 (*Bacillus anthracis* strain STI1),>NZ_CP007626.1 (*Bacillus pseudomycoides* strain 219298), >AP007209.1 (*Bacillus cereus* NC7401),>NZ_CP020743.1(*Bacillus mycoides* strain Gnyt1), >NZ_CP110109.1 (*Bacillus thuringiensis* strain TG-5), NZ_CP066194.1 (*Bacillus cytotoxicus* strain E8.1), >NZ_CP081872.1 (*Bacillus toyonensis* strain UTDF19-29B) and >CP009746.1 (*Bacillus weihenstephanensis* strain WSBC 10204). All data retrieved or generated are not modified or copy-edited.

DECLARATION OF COMPETING INTEREST

The authors declared that they have no competing or personal interest that would influence this article. All the authors concur with the work, and there are no conflicts of interest between them.

REFERENCES

1. Put H, Gerstmans H, Vande CH, Faarten F, Michiels J, Masschelein J. *Bacillus subtilis* as a host for natural product discovery and engineering of biosynthetic gene clusters. Nat Prod Rep. 2024; 3–7. DOI :10.1039/d3np00065f.
2. Meesil W, Paramaporn M, Sutthirat S, Triwit R, Narisara C, Ricardo AR, et al. Genome mining reveals novel biosynthetic gene clusters in entomopathogenic bacteria. Sci Rep. 2023; 13:20764–9. DOI:10.1038/s41598-023-47121-9.
3. Grubbs KJ, Bleich RM, Santa Maria KC, Allen SE, Farag S, Ag-Biome T, et al. Large-scale bioinformatics analysis of Bacillus genomes uncovers conserved roles of natural products in bacterial physiology. Mol Biol Physiol. 2017; 2(6):1–16. DOI:10.1128/mSystems.00040-17.
4. Liming X, Youzhi M, Ali C, Yan L, Zihao L, Xinli S, et al. Biosynthetic gene cluster profiling predicts the positive association between antagonism and phylogeny in Bacillus. Nat Commun. 2022;13:1023. DOI:10.1038/s41467-022-28668-z.
5. Hu D, Sun C, Jin T, Fan G, Mok KM, Li K, et al. Exploring the potential of antibiotic production from rare actinobacteria by whole-genome sequencing and guided MS/MS analysis. Front Microbiol. 2020; 11:1540. DOI:10.3389/fmicb.2020.01540.
6. Martina A, Marius S, Evi S, Nadine, Peter S. Mining bacterial genomes for secondary metabolite gene clusters. Methods Mol Biol. 2017; 1520:9–21. DOI:10.1007/978-1-4939-6634-9.
7. Ehling-Schulz M, Theresa MK, Didier L. The *Bacillus cereus* group: Bacillus species with pathogenic potential. Microbiol Spectr. 2019;7(3):2–6. DOI: 10.1128/microbiolspec.GPP3-0032.
8. Carroll M, Rachel AC, Martin W A, Jasna K. Keeping up with the *Bacillus cereus* group: taxonomy through the genomics era and beyond. Crit Rev Food Sci Nutr. 2022;62(28):7677–702. DOI:10.1080/10408398.2021.1916735.
9. Anderson I, Alexei S, Vinayak K, Gary R, Anamitra B, Natalia M, et al. Comparative genome analysis of *Bacillus cereus* group genomes with *Bacillus subtilis*. FEMS Microbiol Lett. 2005; 250:175–84.
10. Yahya G, Ebada A, Khalaf EM, Mansour B, Nouh NA, Mosbah RA, et al. Soil-associated Bacillus species: a reservoir of bioactive compounds with potential therapeutic activity against human pathogens. Microorganisms. 2021;9(1131):2–18. DOI: 10.3390/microorganisms9061131
11. Lee N, Hwang S, Kim J, Cho S, Palsson B, Cho BK. Mini-review: genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in Streptomyces. Comput Struct Biotechnol J. 2020;18:1548–56. DOI:10.1016/j.csbj.2020.06.024.
12. Kai B, Simon S, Marnix HM, Tilmann W. The antiSMASH database version 4: additional genomes and BGCs, new sequence-based searches and more. Nucleic Acids Res. 2023. DOI:10.1093/nar/gkad984.