Research Article

Received: 2 May 2024 Revised: 20 August 2024 Published online in Wiley Online Library: 26 September 2024

(wileyonlinelibrary.com) DOI 10.1002/ps.8433

Genome mining of *Bacillus thuringiensis* strain SY49.1 reveals novel candidate pesticidal and bioactive compounds

Semih Yılmaz,[a*] Abeer Babiker Idris,[b] Abdurrahman Ayvaz,[c] Rıdvan Temizgül,[c] Aysun Çetin[d] and Mohammed A Hassan[e,f]

Abstract

BACKGROUND: *Bacillus thuringiensis* SY49.1 (*Bt* SY49.1) strain has promising insecticidal and fungicidal activity against phytopathogens and pests. Therefore, we selected this strain for whole-genome sequencing (WGS), annotation and analysis, with the aim of identifying genes responsible for producing putative pesticidal toxins, antimicrobial metabolites and plant growth-promoting features.

RESULTS: Our results showed that the SY49.1 genome is 6. 32 Mbp long with a GC content of 34.68%. Genome mining revealed the presence of multiple gene inventories for the biosynthesis of bioactive compounds such as insecticidal delta endotoxins, secondary metabolites, and several plant growth-promoting proteins. Multiple sequence alignment revealed residue variations in the toxic core of Cry1Ab when compared with known Cry1Ab sequences from *Bt* nomenclature databases. This suggests that the *cry1Ab* of SY49.1 is a new kind of its group. Among the predicted secondary metabolites, we found a kurstakin with a predicted peptide that differs from the known kurstakin peptide available in the NORINE database. In addition, lipopeptides extracted from SY49.1 suppressed the growth of *Verticillium dahliae* and *Fusarium oxysporum*.

CONCLUSION: We anticipate that the complete genome of *Bt* SY49.1 may provide a model for properly understanding and studying antimicrobial compound mining, genetic diversity among the *B. cereus* group, and pathogenicity against insects. This is the first report on the WGS and mining of the *Bt* strain isolated from Turkey.

Supporting information may be found in the online version of this article.

Keywords: *Bacillus thuringiensis*; whole-genome sequencing; genome mining; pesticide; fungicide; PGPR

1 INTRODUCTION

Bacillus thuringiensis (*Bt*) is a versatile bacterium found in diverse environments and produces parasporal crystals that contain powerful toxins specifically targeting insect larvae. These toxins, such as *δ*-endotoxins, have been successfully integrated into genetically modified crops owing to their specificity and safety.[1] *Bt*, particularly in the *Bacillus* genus, also is recognized for producing biologically active compounds that inhibit phytopathogens, addressing economic losses in agriculture.[2] With increasing concerns about toxic and nonbiodegradable antimicrobial agents, secondary metabolites such as polyketides (PKs) and nonribosomal peptides (NRPs) from *Bacillus* are explored for their environmentally friendly antimicrobial properties.[3] *Bt* strains have been found to exhibit antibacterial, antibiofilm, antifungal and emulsifying activities.[1,2,4] They are known to be significant sources of antimicrobial compounds, such as bacteriocins and lipopeptides, and antifungal compounds, such as zwittermycin, lipopeptides and chitinase.[5]

Several strains of *Bt* have raised global interest for various biopest applications because of their specific pesticidal activities. However, the development of insect resistance to *Bt* toxin proteins has been widely reported.[6] This has led to an urgent search for novel *Bt* with new insecticide toxins as an effective

* *Correspondence to: S Yilmaz, Department of Agricultural Biotechnology, Faculty of Agriculture, Erciyes University, Talas, Kayseri, 38652, Turkey. E-mail: ylmazsemh@yahoo.com*

a *Department of Agricultural Biotechnology, Faculty of Agriculture, Erciyes University, Kayseri, Turkey*

b *Department of Agricultural Sciences and Technologies, Graduate School of Natural and Applied Sciences, Erciyes University, Kayseri, Turkey*

c *Department of Biology, Faculty of Science, Erciyes University, Kayseri, Turkey*

d *Department of Medical Biochemistry, Faculty of Medicine, Erciyes University, Kayseri, Turkey*

e *Department of Bioinformatics, Africa City of Technology, Khartoum, Sudan*

f *Sanimed international lab and management l.l.C, Abu Dhabi, UAE*

management strategy. Indeed, genome mining on *Bt* genomes has proven to be effective in guiding the discovery and identification of new putative pesticides, bioactive compounds and secondary metabolites.[7] The availability of whole-genome sequencing (WGS) for bacteria along with bioinformatics tools has sped up and facilitated the ability to identify potential metabolic processes that could be utilized in biotechnological applications.[8] To the best of our knowledge, there have been no reports in the literature regarding the WGS and mining of *Bt* isolated from Turkish soils.

In a previous study, we characterized a highly entomopathogenic *Bt* SY49.1 strain in terms of both its *16S*-ITS *rDNA* region and *cry* genes contents.[9] This strain was isolated from Adana, a city in Turkey's southern region distinguished by its distinct climate and geographical location on the fertile and watery delta of the Seyhan and Ceyhan rivers. Interestingly, PCR results indicated that SY49.1 harbors lepidopter-dipter-specific *cry2A* gene, as well as lepidopter-specific *cry1Ab*, *cry1Aa/Ad*, *cry1C*, *cry2Aa, cry9A*, *cry9C*, lepidopter-dipter-coleopter-specific *cry1B*, and nematode-specific *cry5* gene.[10] Additionally, it showed *in vitro* activity against lepidopteran pests (*Ephestia kuehniella*, *Plodia interpunctella* and *Thaumetopoea wilkinsoni* Tams) and a dipteran pest (*Culex pipiens*).[9,11–15] It is worth mentioning that the *Bacillus thuringiensis* delta-endotoxin nomenclature committee designed the Cry2Aa sequence of *Bt* SY49.1 as Cry2Aa18 being a new member of *Bt* toxins (accession no. KX243304; protein id. ANF99565).[16] Additionally, *Bt* SY49.1 produced a chitinase enzyme that inhibits the growth of fungi by breaking down the chitin-containing cell wall; therefore, it showed *in vitro* antagonism against plant pathogenic fungi.[17]

Considering the promising characteristics of *Bt* SY49.1, we selected this strain for WGS, annotation and analysis, with the aim of identifying the genes responsible for the production of putative pesticidal toxins, antimicrobial metabolites and plant growth-promoting features. Additionally, we extracted lipopeptides from *Bt* SY.49.1 to assess their antifungal activities against *Fusarium oxysporum* and *Verticillium dahliae*, both of which are responsible for significant economic losses on a global scale. Studying the effectiveness of extracted lipopeptides offers a more targeted, efficient and customizable approach for various applications compared to examining the whole bacterium.

2 MATERIALS AND METHODS

2.1 Bacterial isolation and identification

Bt strain SY49.1 was obtained from the bacterial stock culture at the Department of Enzyme and Microbial Biotechnology, Faculty of Agriculture, Erciyes University. The methods for isolation and identification, as well as the characterization of Cry proteins using scanning electron microscopy, are described in our previous works.[9,10]

Strain SY49.1 has been deposited in the Adana *Bacillus* Strains Culture Collection (ABSCC) which is World Data Centre for Microorganisms (WDCM) collection number 1288. The strain's number in the collection is 49.

2.2 Extraction and antifungal activities of lipopeptides

Bt SY 49.1 was initially subcultured in Medium Optimal for Lipopeptide Production (MOLP) at 35 °C for 72 h to enhance lipopeptide production. Lipopeptides were extracted using the acid precipitation method, as described previously.[18] The precipitated lipopeptides were suspended in 80% (v/v) methanol, and 100 µL of the extract was applied as spots on paper discs placed on potato dextrose agar (PDA) plates. The fungal plug was positioned in the center of each PDA plate. Control was established using a fungus grown in the absence of any inhibitory substances. The plates were incubated at 28 ± 2 °C for 7 days. Fungal growth inhibition was evaluated by comparing the mycelium growth in both the presence and absence of lipopeptides.

2.3 Whole-genome sequencing, assembly and annotation

High-quality bacterial DNA was extracted using an EasyPure bacterial genomic DNA kit, according to the protocol provided by the manufacturer. The quality and the concentration of the DNA were checked by agarose gel electrophoresis and NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA), respectively. The WGS was performed commercially using Illumina HiSeq technology at BGI-Hong Kong Co., Ltd.

A 760.2-Mbp paired-end read library with a 500-bp insert size was generated from the DNA sample. However, 10–20% of the data were eliminated after filtering for low quality, Ns' bases, adapter contamination and duplication reads. *De novo* assembly was performed using SOAPdenovo (v1.05). The assembly error was corrected by aligning the reads to the assembly result using the SOAPaligner (v2.21) tool and counting the mapping information of reads. A GC content and depth correlative analysis, and K-mer analysis were performed to evaluate whether heterogeneous sequences exist in the sample or not. The coverage depth and coverage ratio of *Bt* SY49.1 strain to a reference sequence (NC_017208, *Bt* Chinensis serovar) were 48.79x and 80.29%, respectively.

The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), the Joint Genome Institute– Integrated Microbial Genomes and Microbiomes (JGI-IMG/M) pipeline and Rapid Prokaryotic Genome Annotation (Prokka), and gene functions were predicted using Rapid Annotation using Subsystems Technology (RAST) and the Pathosystems Resource Integration Center (PATRIC).

2.4 Phylogenomic and comparative analyses

Phylogenomic analysis of strain SY49.1, along with the closest reference and representative genomes, was conducted using the PATRIC server. Additionally, sequences of *16S rRNA* genes from 13 *Bacillus* strains, including SY49.1, were selected for molecular phylogenetic analysis. These species represent members of the *B. cereus sensu lato* group, with *B. subtilis* IAM 12118 selected as an outgroup. The Kimura 2-parameter (K2 + G) model was used with 1000 bootstrap replicates to construct a maximum-likelihood (ML) phylogenetic tree using MEGA7 software.

For species assignment, we measured pairwise genome-wide average nucleotide identity (gANI) metric and alignment fractions (AFs) within the JGI-IMG/M server using the Microbial Species Identifier (MiSI) calculator. Strain assignment was established, based on digital DNA: DNA hybridization (dDDH) and DNA G + C content, using the Genome-to-Genome Distance Calculator (GGDC 3.0) server.

Functional comparison (sequence-based) of the *Bt* SY49.1 genome was performed with closely related *Bacillus* species, *Bt* str. 97–27 (NC_005957), *Bt* str. Al Hakam (NC_008600), *B. cereus* ATCC 10987 (NC_003909), *B. cereus* ATCC 14579 (NC_004722), *B. cereus* E33L (NC_006274) and *B. anthracis* str. Ames Ancestor

(NC_007530); *B. subtilis* subsp. *subtilis* str.186 (NC_000964) was used as an outgroup in the map. The comparison was done by using bidirectional and unidirectional best hit (protein BLAST) implemented in the RAST server.

2.5 Genome mining

The putative insecticidal protein genes were predicted using the Cry processor and BT toxin scanner tool, available in the *Bt* nomenclature website.[16] The predicted coding sequences were compared to the nonredundant database using BLASTX from NCBI, and the coding sequences with a hit length > 100 aa and bit score > 90 were selected. Also, they were BLAST-ed against the BPPRC database based on the *Bt* toxin available in the *Bt* nomenclature list. Furthermore, multiple sequence alignment was performed using BioEdit software, and the Project HOPE server was used to analyze the structural effects of a point mutation in a protein sequence. Additionally, other toxins and virulence factors of SY49.1 were identified using the VFanalyzer tool.

Secondary metabolites, which are often encoded in contiguous biosynthetic gene clusters (BGCs), were predicted, annotated and analyzed using antiSMASH v4.0, NRPS-PKS domain search program and the PKS/NRPS Analysis website. Using the NORINE database (http://norine.univ-lille.fr/norine/), the structural patterns found in the peptides of secondary metabolites were compared to all known NRPs.[19] Additionally, CRISPR repeats were predicted by using the CRISPRfinder webserver. The PHASTER webserver was used for the identification and annotation of prophage sequences within the *Bt* SY49.1 genome. Also, the IslandViewer 4 webserver was applied to predict and interactively visualize genomic islands (GI).

3 RESULTS AND DISCUSSION

3.1 Genome properties

Bt SY49.1 was originally isolated from Turkish soil and characterized by producing crystals containing protein toxins (Cry proteins).[11] The general features of *Bt* SY49.1 and its genome sequence information are summarized in Table 1. The complete genome sequence of the *Bt* SY49.1 strain is 6 322 230 bp long with a 34.68% GC content, which is similar in size and content to the genomes of other *Bt* strains.[1,2,20] It contains 270 scaffolds with an N50 value of 79 855 bp. In total, 6562 protein-encoding genes (PEGs), 18 RNAs and 429 pseudogenes were annotated by NCBI-PGAP. Conversely, the CRT 1.8.2-JGI IMG/M pipeline predicted 7105 PEGs, 24 RNAs and 187 pseudogenes, as illustrated in Table 2. The 6982 protein-coding genes predicted in the IMG annotation pipeline were placed in 25 general clusters of orthologous (COG) functional gene catalogs. The distribution of these protein-coding genes based on COG function is presented in Table 3. The RAST annotation has sorted the 6562 PEGs into 342 functional subsystem pathways using SEED subsystems. The most abundant of them are genes that are associated with amino acids and derivatives metabolism 392 (18.8%), followed by carbohydrate metabolisms 264 (12.67%), cofactors, pigment and prosthetic groups 158 (7.59%), protein metabolism 156 (7.5%), and dormancy and sporulation 109 (9. 23%). Moreover, two sequences were predicted with CRISPR repeats and one sequence with Cas cluster, whereas the IMG annotation pipeline predicted four CRISPR repeats. The *Bt* SY49.1 WGS project has been deposited at NCBI GenBank under accession no. NZ_JAHKEZ000000000, assembly no. GCF_018791885, BioProject accession no. PRJNA734785 and BioSample accession no. SAMN19533886.

Table 1. General features of *Bt* strain SY49.1 and its genome

Property	Term
Bacterial taxonomy	Domain: Bacteria
	Phylum: Firmicutes
	Class: Bacilli
	Order: Bacillales
	Family: Bacillaceae
	Genus: *Bacillus*
	Species: *Bacillus thuringiensis*
	Strain: SY49.1
Gram stain	Positive
Cell shape	Rod
Motility	Motile
Sporulation	Spore (subterminal)
Habitat	Soil
Oxygen requirement	Aerobic
Temperature; pH	35 °C; 7,0
Biotic relationship	Free-living
Pathogenicity	Insecticidal and fungicidal
Geographic location	Adana, Turkey
Sample collection	2008
latitude, longitude	37.00 N, 35.321335 E
Genome sequence information	
Finishing quality	Complete genome
Libraries used	Illumina paired-end
Sequencing technology	Illumina HiSeq
Fold coverage	48.79x
Assemblers	SOAPDENOVO v1.05
Gene-calling method	Best-placed reference protein set; GeneMarkS; Prodigal; CRT
Locus tag	KN035
GenBank ID	JAHKEZ010000000
IMG Submission ID	264 041
WDCM CCINFO/ WDCM CCINFO No.	*Bt* ABSCC 49 /1288

The 6.32-Mbp genome map of *Bt* SY49.1 is presented in Supporting Information, Fig. S1.

3.2 Phylogenomic analyses

The phylogenetic tree of the complete genome and *16S rRNA* of SY49.1 supports the placement of strain *Bt* SY49.1 within the *B. cereus* group (see Fig. S2). The *Bt* SY49.1genome is highly similar to that of *Bt* serovar kurstaki based on average nucleotide identity and digital DNA: DNA hybridization (> 99%; see Table S1).

Comparison of the *Bt* SY49.1 genome with the six closely related strains of *Bt*, *B. cereus* and *B. anthracis*, revealed strain-specific genes that encode hypothetical proteins, phage-like proteins, prophages, mobile genetic elements, and transposases. These findings were cross-validated and visualized in the IslandViewer 4 server which showed that most of these genes grouped into genomic islands (see Fig. S3 and Table S2). The PHASTER server identified 10 prophage regions, of which one region is intact (score >90), seven regions are incomplete (score <70) and two regions are questionable (score 70–90).

Table 2. Genome statistics

Attribute	Value	% of total
Genome size (bp)	6 322 230	100.00%
DNA coding (bp)	5 300 308	83.84%
DNA G + C (bp)	2 192 777	34.68%
DNA contigs	708	100.00%
DNA scaffolds	270	100.00%
Total genes	7105	100.00%
Protein coding genes	6982	98.27%
Regulatory and miscellaneous features	99	1.39%*
RNA genes	24	0.34%
Pseudogenes	187	2.63%
Genes with function prediction	4919	69.23%
Genes with enzymes	1436	20.21%
Genes in internal clusters	2314	32.57%
Genes assigned to COGs	4719	66.42%
Genes with Pfam domains	5070	71.36%
Genes with signal peptides	266	3.74%
Genes with transmembrane helices	1931	27.18%
CRISPR repeats	4	0.056%

*Regulatory or miscellaneous genes are genes that are not classified as CDS, a type of RNA, or a pseudogene, but as 'unknown' or 'other' by the source provider.

3.3 Insights from the genome sequence of *Bt* SY49.1

3.3.1 *Identification of insecticidal toxins and virulence factors*

Bt SY49.1 was found to be flagellated, sporulating with a subterminal endospore and producing insecticidal parasporal inclusions (Fig. 1). These phenotypes are supported by gene inventories observed in the genome of *Bt* SY49.1. *Bt* insecticidal toxins are a useful and eco-friendly option for controlling pests. However, insect resistance to *Bt* toxins is a significant limitation. Therefore, identifying novel *Bt* strains or toxins is of scientific and economic interest, and it is critical for effective insect pest control and insect resistance management. Genome mining of the *Bt* SY49.1 using Cry processor, BT toxin scanner and BLAST analysis tool retrieved eight genes that encoded *Bt* putative insecticidal proteins. Seven of these proteins had high homology (100% similarity) with Cry family proteins: {Cry1Aa, MBU0451991 and MBU0451992}; {Cry1Ac, MBU0453390 and MBU0453356}; {Cry1Ia, MBU0451993}; {Cry2Aa, MBU0451997}; and {Cry2Ab, MBU0452330}. In previous studies, these families (Cry1 and Cry2) have been detected in SY49.1 and characterized using conventional PCR and SDS-PAGE methods.[10,15] Bioassay investigations of SY49.1 spores/crystals mixture and Cry proteins showed insecticidal activity against *Ephestia kuehniella Zeller* (Lepidoptera: Pyralidae), *Plodia interpunctella* (Hübner) (Lepidoptera: Pyralidae), *Thaumetopoea wilkinsoni Tams* (Lepidoptera: Thaumotepoeidae), *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae) and *Culex pipiens* (Diptera: Culicidae) larvae.[11–15]

Table 3. Number of protein-coding genes associated with general COG functional categories

Code	Description	Value	Percent
E	Amino acid transport and metabolism	495	8.84%
G	Carbohydrate transport and metabolism	302	5.39%
D	Cell cycle control, cell division, chromosome partitioning	94	1.68%
N	Cell motility	86	1.54%
M	Cell wall/membrane/envelope biogenesis	295	5.27%
B	Chromatin structure and dynamics	2	0.04%
H	Coenzyme transport and metabolism	247	4.41%
Z	Cytoskeleton	3	0.05%
V	Defense mechanisms	176	3.14%
C	Energy production and conversion	252	4.5%
W	Extracellular structures	15	0.27%
S	Function unknown	478	8.54
R	General function prediction only	611	10.91%
P	Inorganic ion transport and metabolism	342	6.11%
U	Intracellular trafficking, secretion and vesicular transport	58	1.04%
I	Lipid transport and metabolism	184	3.29%
X	Mobilome: prophages, transposons	155	2.77%
F	Nucleotide transport and metabolism	137	2.45%
O	Post-translational modification, protein turnover, chaperones	237	4.23%
L	Replication, recombination and repair	191	3.41%
Q	Secondary metabolites biosynthesis, transport and catabolism	142	2.54%
T	Signal transduction mechanisms	261	4.66%
K	Transcription	501	8.95%
J	Translation, ribosomal structure and biogenesis	334	5.97%
-	Not in COG	2386	33.58%

Pest Management Science published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

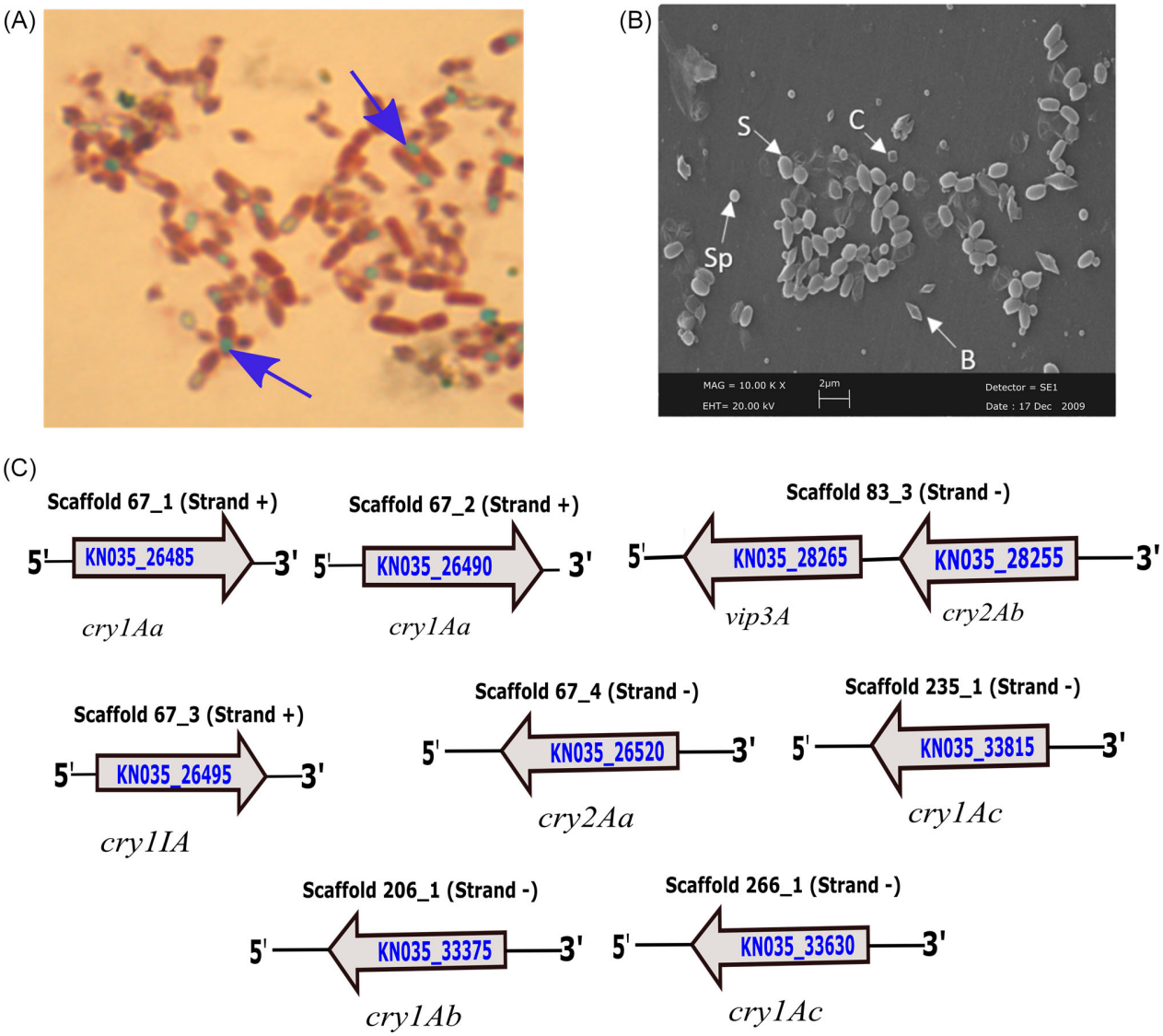

Figure 1. (A) Microscopic view of *Bt* SY49.1 after spore-staining. Arrows indicate the endospores. (B) Scanning electron micrograph of the different types of crystal proteins and spores (S) produced by *Bt* SY49.1. B, bipyramidal; C, cubic; Sp, spherical. (C) Gene synteny of the candidate insecticidal genes identified in the *Bt* SY49.1 genome.

Regarding the structure of Cry proteins, three domains (D-I, D-II and D-III) required for toxicity (toxic core) are present in the N-terminal half, whereas the C-terminal half is made up of protoxin domains (D-IV to D-VII) and is not observed in smaller proteins. Domain I in the toxic core is responsible for membrane insertion and pore formation, whereas Domains II and III contribute to the structural stability of the toxin and its interaction with insect receptors.[21] Early investigations into the role of the protoxin, as well as the weight of evidence accumulated to date, suggest that the protoxin is unnecessary for insecticidal activity; but its function is likely to be related to crystal formation, stability, and selective solubilization in the insect gut.[22] As a result, 3′-truncated *cry* genes of larger proteins, such as Cry1, could encode for proteins with insecticidal activity.[23] However, minor variations in deduced amino acid sequences of the toxic core of the Cry toxins can result in significant changes in the protein's toxicity and insecticidal activities.[24] In this study, multiple sequences alignment (MAS) of {KN035_33375, MBU0453309} with known Cry1Ab proteins revealed residue variations (F148L and H206Y) in domain I (α4 and α6) of the toxic core (Fig. 2). Based on Project HOPE analysis, there are differences in the size, charge and hydrophobicity value between wild-type residues and mutant residues in both locations 148 and 206, and the mutated residues can affect the function but not damage the Cry protein. This suggests that the *cry1Ab* of *Bt* SY49.1 is a new kind of its group. Mutagenesis studies have been conducted to investigate the mechanism of action of Cry toxins and to identify specific sites that can be modified to enhance the protein's efficacy against insect pests. Previous studies have demonstrated that mutations in Domain I of Cry1Ab can potentially enhance toxin activity by accelerating Cry1Ab unfolding, aiding membrane integration and increasing the toxin's pore-forming capability, although contrasting findings have been reported.[25–29] However, the Cry1Ab of *Bt* SY49.1 had demonstrated its effectiveness in a previous study. In that study, we detected, cloned and expressed the *cry1Ab* gene of SY49.1, and it exhibited significant insecticidal activity against *lepidopteran*

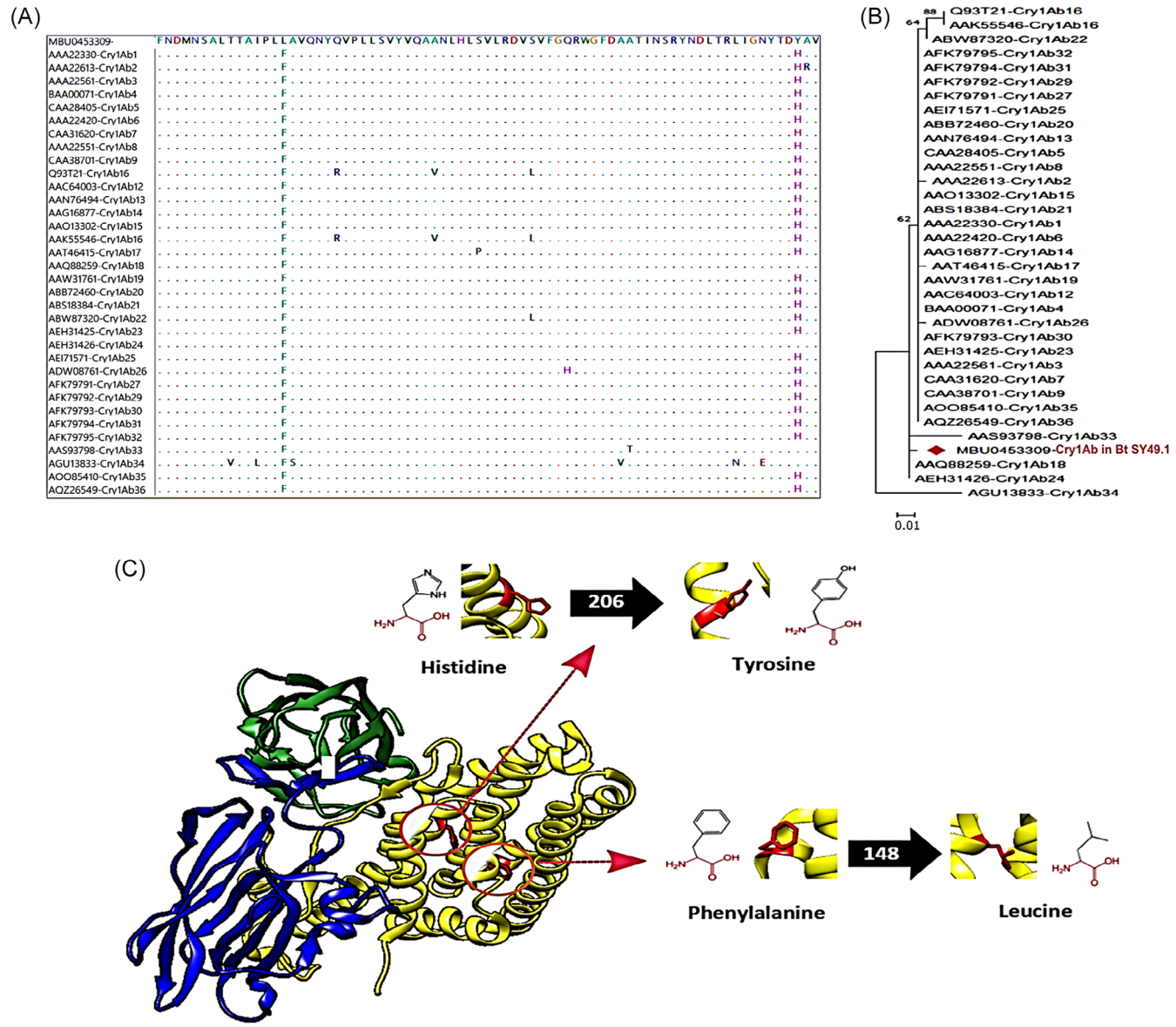

Figure 2. (A) Multiple sequences alignment of SY49.1 Cry1Ab toxic core {KN035_33375, MBU0453309} with known Cry1Ab proteins obtained from the *Bt* nomenclature database (Panneerselvam *et al.*, 2022). (B) Neighbor-joining phylogenetic tree of SY49.1 Cry1Ab and other known Cry1Ab proteins based on amino acid sequences. The evolutionary distances were computed using the JTT matrix-based method and the bootstrap was 1000 replicates. Evolutionary analyses were conducted in MEGA11. (C) 3D structure of the active Cry1Ab toxin. The three Domains are colored as follows: Domain I (Yellow), Domain II (blue) and Domain III (green). Mutated residues (F148L and H206Y) observed in Domain I are shown in a detailed view.

pests, *Ephestia kuehniella* and *Plodia interpunctella*.[14] Furthermore, {KN035_28265, MBU0452331} encoded proteins with 100% homology to the Vip3A protein.

However, in addition to the *Bt* pesticidal protein genes, the SY49.1 strain has been found to contain 46 coding sequences from various classes of virulence factors and some *Bt* strains with similar virulence factors have been used as pesticides[30] (see Table S3). However, these factors are similar to those found in other strains of the *B. cereus* group, and some of them include toxins that can cause food-borne diseases.[31] These factors include hemolysins (*HlyIII*, *cytK2* and *hbl A-D*), Non-hemolytic enterotoxins (*nhe A-C*) and anthrolysin O/cereolysin O family cholesterol-dependent cytolysin (*alo*). Although *Bt* has been proven to be safe for biocontrol use, it has been hypothesized that *Bt* could be responsible for part of the *B. cereus*-associated illnesses.[32] Therefore, it is necessary to conduct further *in vitro* and *in vivo* experiments to assess the pathogenicity of SY49.1 in mammalian cells before using it as a biopesticide.

It is worth mentioning that strain SY49.1 has a protein belonging to the hemolysin XhlA family which is essential for the virulence of the insect pathogen *Xenorhabdus nematophila*.[33] Also, SY49.1 encodes several virulence factor genes involved in adherence, stress, host colonization, immune evasion and iron acquisition (Table S3), which could aid in the avoidance of the host immune system and the effective degradation and utilization of insect tissues.[30] In addition, SY49.1 has genes encoding ABD transporters, which in the last decade, have emerged as significant contributors to combat pest resistance against insecticidal pore-forming proteins from *Bt*.[34] These findings suggest that SY49.1 is a highly toxic strain to insects because its genome

encodes a large number of insecticidal toxicity-related genes (ITRGs).[30]

3.3.2 *Plant growth-promoting rhizobacteria (PGPR) traits*

We found that the *Bt* SY49.1 strain has genes/gene clusters responsible for nitrogen fixation as well as plant-promoting growth traits; thus, it could be a good candidate for the development of a bio-fertilizing inoculant that provides plants with essential compounds and protects them from phytopathogens.

Soil contains two forms of phosphate: organic phosphates, primarily in the form of inositol phosphates (or phytates) and phosphoesters, and mineral phosphates which are present as hydroxyapatites, calcium phosphates and rock phosphates. Both mineral and organic phosphates must be solubilized before plants can use them, and phosphate-solubilizing bacteria play a vital role in this process.[35] Although the genetic basis of mineral phosphate solubilization is not yet well understood, the best-characterized mechanism for inorganic phosphate solubilization involves the production of organic acids that chelate divalent cations, thus making the phosphate available to microbial cells and plants.[36] The SY49.1 genome contained several phosphatases predicted to be involved in the mineralization of organic phosphorus, including acid phosphatase (EC 3.1.3.2), metallophosphatase and two alkaline phosphatases [PhoX (EC 3.1.3.1) and the rhodanese/Cdc25 phosphatase superfamily]. Regarding the capability of SY49.1 to solubilize inorganic phosphate, dehydrogenase enzymes involved in D-gluconate and ketogluconates metabolism were found.

It is worth mentioning that SY49.1 has genes encoding chitinolytic enzymes, including chitinase (EC 3.2.1.14), chitin-binding protein, glucosamine-6-phosphate deaminase (EC 3.5.99.6), *N*-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) and predicted transcriptional regulator of *N*-Acetylglucosamine utilization (GntR family). Additionally, various gene clusters associated with bacterial siderophore or siderophore precursors [i.e. petrobactin, bacillibactin and 2,3-dihydroxybenzoic acid (2,3-DHB)] were discovered in the genome. For *Bt*, the utilization of siderophores for iron acquisition is important not only for the biocontrol of phytopathogens through iron competition, but also for providing iron to the plant. Various plants have heterologous iron uptake mechanisms that utilize iron-bacterial siderophores for iron acquisition (for a review, see Batista et al.[37]). In addition, the potential ability of the SY49.1 strain to produce auxins is one of its most interesting properties. This strain possesses the *ipdC* gene, which codes for indole pyruvate decarboxylase, a key enzyme involved in the biosynthesis of IAA from tryptophan via the indole pyruvate pathway. Auxins are known to play important roles in plant growth, development and stress response, and for microbial cells, they can control various aspects of cell division, differentiation and elongation.[38]

Also, 2,3-butanediol dehydrogenase (which catalyzes the oxidation of 2,3-butanediol to acetoin), and enzymes for nitrogen fixation are encoded in the genome of *Bt* SY49.1. An overview of the PGPR traits identified in the *Bt* SY49.1 genome is presented in Table S4.

3.3.3 *Biosynthesis of secondary metabolites (SMs)*

The antiSMASH 4.0 server predicted seven gene clusters in the *Bt* SY49.1 genome that are responsible for the biosynthesis of secondary metabolites which include lantipeptide, antimicrobial peptides and siderophores (Table 4).

Microorganisms produce natural compounds that serve as the basis for many drugs, including antibiotics and antivirals, but drug resistance and genetic alteration are great threats to controlling microbial infection. As a result, researchers are exploring strategies such as using naturally occurring antimicrobial peptides (AMPs) and modifying existing drugs.[2] In this study, the SY49.1 genome was found to carry gene clusters including lantibiotics and antimicrobial peptides (fengycin and zwittermycin A) (Table 4). Fengycin increases the plasma membrane permeability of the target cell and exhibits strong fungi toxic activity,[39] The aminopolyol antifungal compound (zwittermycin A) was previously shown to suppress oomycete diseases in plants.[40] Additionally, the SY49.1 genome was found to carry gene clusters with complete homology to the biosynthetic gene cluster of the

Table 4. Secondary metabolites (SM) clusters identified in the *Bt* SY49.1 genome using antiSMASH 4.0 server

SM Cluster (scaffold)	Predicted cluster	Length (Kb)	Type	Key protein from biosynthetic core	Locus tag
SM-1 (7_9)	Bacillibactin	11.88	NRPS[†]	Siderophore 2,3-dihydroxybenzoate (DhbF)	KN035_06960–06985
SM-2 (16_3)	-*	3.58	NRPS-like	-	KN035_12900
SM-3 (28)	Zwittermicin A	57.42	Type I PKS[‡] / NRPS	ZmaA, ZmaB, ZmaK, ZmaO and ZmaQ.	KN035_18280–18 405
SM-4 (28)	Thuricin	7.43	Lanthipeptide-class-II	Thuricin modification protein (thuM)	KN035_18415–18 440
SM-5 (43_1)	Kurstakin*	27.26	NRPS	KrsA, KrsB, and KrsC	KN035_22580–22 600
SM-6 (46_2)	Fengycin/ Plipastin	11.55	NRP	YngG and YngI	KN035_23155–23 210
SM-7 (57_1)	Petrobactin	6.09	Siderophore	Petrobactin biosynthesis proteins AsbA and AsbB	KN035_25240–25 260

*Predicted secondary metabolites which are not identified by antiSMASH software.
[†] NRPS, nonribosomal peptide synthetase.
[‡] PKS, polyketide synthase.

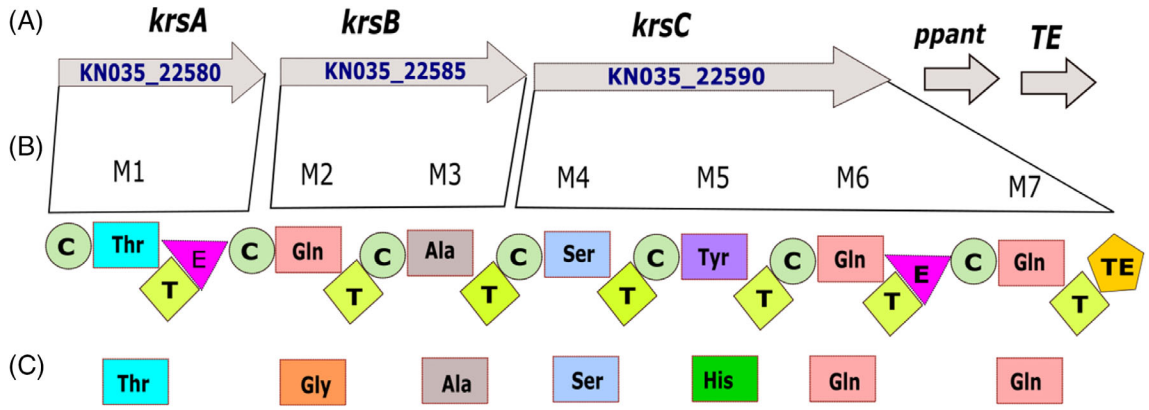

Figure 3. Kurstakin operon of SY49.1. (A) The kurstakin biosynthesis genes, *krsA*, *krsB* and *krsC*, followed by a gene coding a phosphopantetheinyl transferase (*ppant*) and second thioesterase domain (TE). (B) Modules (M) and domains identified in the deduced protein sequence of *krsA*, *krsB* and *krsC*. Condensation (C), adenylation (A), thiolation (T) and epimerization (E) domains. The predicted amino acids are presented in red boxes. (C) Amino acid sequence of the peptide moiety of kurstakin as found in the NORINE database (Flissi *et al.*, 2020). [Correction added after first online publication on 16 October 2024; Figure 3 has been updated.]

thuricin (class II lantibiotic). Lantibiotics are gene-encoded peptide antibiotics that imply great potential for medical, agricultural and industrial applications as a consequence of their special mode-of-action and low potential for resistance. Thus, the discovery or detection of lantibiotics is urgent and appealing.[41] The antimicrobial assay of thuricin exhibited its activity against *Micrococcus flavus*, and the conserved disulfide bridge was significant to the function of this bovicin HJ50-like lantibiotic.[42]

Interestingly, the antiSMASH 4.0 server predicted the secondary metabolite SM-5 as belonging to NRPS without annotation. Further bioinformatics analyses, which were previously mentioned in the Section 2, were carried out to characterize the SM-5 operon and predict the produced peptide. SM-5 showed high similarity with the kurstakin operon of *Bt* and *B. cereus*. Kurstakin was discovered in 2000 and is considered a biomarker for *Bt*. This lipoheptapeptide family displays antifungal activities, especially against *Stachybotrys charatum*.[43] However, the predicted heptapeptide of SM-5 is different from the known kurstakin peptide available in the NORINE database, and therefore could be a new variant in the kurstakin family (Fig. 3). However, it is worth noting that natural variants are not unique to kurstakins. For example, natural variants have been observed in surfactins as a result of alterations in the adenylation domains of modules two, four and seven of surfactin synthetases.[44] Generally, the diversity of lipopeptide chemical structures and properties can be influenced by various factors, including the producer strain and ecological differences among microbial producers.[45] Additionally, the content of each variant of a lipopeptide is crucial in determining the characteristics of metabolites as well as the biological and biochemical traits of microbes.[46]

3.4 Fungal growth inhibition by extracted lipopeptides

Lipopeptides, extensively researched over the last decade, offer a range of applications, from antimicrobial and antifungal to immunosuppressant and antitumor properties. These bacterial-derived compounds are advantageous for their low toxicity, eco-friendliness, surface activity, adaptability to extreme conditions, and recyclability. They find wide-ranging use in cosmetics, food, agriculture and pharmaceuticals, promoting bacterial survival and multicellular growth, enhancing colonization, motility and swarming abilities.[47] In this study, lipopeptides extracted from cell-free culture broth of strain SY49.1 significantly suppressed the growth of *V. dahlia* but showed weak activity against *F. oxysporum*. *F. oxysporum* is the causative agent of Fusarium wilt and inflicts severe damage on a wide range of important crops, including potatoes, cloves, tomatoes, beans, cucurbits (cucumber, watermelon, pumpkin and melon), peanuts, eggplant, strawberries, sugar beet and alfalfa, and various ornamental plants.[48] Although *V. dahliae* is responsible for Verticillium wilt, this disease predominantly affects crops such as sunflowers, sesame, tomatoes, okra, peppers, beans, eggplants, cowpeas, potatoes, parsley, apricots, peanuts, melons, peaches, cloves and olives.[49]

4 CONCLUSIONS

Several gene inventories encoded in the complete genome of strain SY49.1 could be used to infer the antibacterial, antifungal and insecticidal properties, as well as PGPR traits. This indicates that the strain could have several potential utilities as a source of antibiotic compounds, as well as being a biocontrol agent for fungal phytopathogens and insects. We anticipate that the complete genome of the *Bt* SY49.1 strain may provide a model for properly understanding and studying antimicrobial compound mining, genetic diversity among the *B. cereus* group, and pathogenicity against insect pests and plant diseases.

ACKNOWLEDGEMENTS

This work is funded by Erciyes University Scientific Project Unit under the code FBD-08-540, and part of the study was supported by the TÜBİTAK project under the code 22AG012. Also, we gratefully acknowledge the collaboration of the staff at the GenKök research center.

DATA AVAILABILITY STATEMENT

The strain SY49.1 has been deposited in the Adana Bacillus Strains Culture Collection (ABSCC)/WDCM CCINFO No. 1288. The strain's number in the collection is 49. Also, Bt SY49.1 whole-genome sequencing project has been deposited at NCBI GenBank under the accession number NZ_JAHKEZ000000000, BioProject accession number PRJNA734785, and BioSample accession number SAMN19533886. Other available data can be found in the supplementary material.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

AUTHOR CONTRIBUTIONS

Semih Yılmaz: Conceptualization (equal), Funding acquisition (lead), Methodology (equal), Investigation (equal), Resources (equal), Supervision (lead), and Writing- review and editing (equal). **Abeer Babiker Idris:** Conceptualization (equal), Formal analysis (lead), Investigation (equal), Methodology (equal), Software (lead), Writing- original draft (lead), and Writing- review and editing (equal). **Abdurrahman AYVAZ:** Supervision (lead), Resources (support), and Writing- review and editing (equal). **Rıdvan Temizgül:** Methodology (equal) and Investigation (equal). **Aysun Çetin:** Resources (equal), and Methodology (equal). **Mohammed A Hassan:** Conceptualization (equal), Investigation (support), Supervision (support), and Writing- review and editing (equal).

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Reyaz AL, Balakrishnan N and Udayasuriyan V, Genome sequencing of Bacillus thuringiensis isolate T414 toxic to pink bollworm (Pectinophora gossypiella Saunders) and its insecticidal genes. *Microb Pathog* **134**:103553 (2019).
2 Daas MS, Rosana ARR, Acedo JZ, Douzane M, Nateche F, Kebbouche-Gana S *et al.*, Insights into the draft genome sequence of bioactives-producing Bacillus thuringiensis DNG9 isolated from Algerian soil-oil slough. *Stand Genomic Sci* **13**:25 (2018).
3 Meena KR and Kanwar SS, Lipopeptides as the antifungal and antibacterial agents: applications in food safety and therapeutics. *Biomed Res Int* **2015**:473050 (2015).
4 Djenane Z, Nateche F, Amziane M, Gomis-Cebolla J, el-Aichar F, Khorf H *et al.*, Assessment of the antimicrobial activity and the entomocidal potential of bacillus thuringiensis isolates from Algeria. *Toxins (Basel)* **9:**139 (2017).
5 Cochrane SA and Vederas JC, Lipopeptides from Bacillus and Paenibacillus spp.: a gold mine of antibiotic candidates. *Med Res Rev* **36**:4–31 (2016).
6 Kaze M, Brooks L and Sistrom M, Antimicrobial resistance in Bacillus-based biopesticide products. *Microbiology* **167**:001074 (2021).
7 Su X, Li L, Pan J, Fan X, Ma S, Guo Y *et al.*, Identification and partial purification of thuricin 4AJ1 produced by Bacillus thuringiensis. *Arch Microbiol* **202**:755–763 (2020).
8 Trivella DBB and de Felicio R, The tripod for bacterial natural product discovery: genome mining, silent pathway induction, and mass spectrometry-based molecular networking. *mSystems* **3**:e00160-17 (2018).
9 Yılmaz S, Ayvaz A, Akbulut M, Azizoglu U and Karabörklü S, A novel Bacillus thuringiensis strain and its pathogenicity against three important pest insects. *J Stored Prod Res* **51**:33–40 (2012).
10 Yılmaz S, Ayvaz A and Azizoglu U, Diversity and distribution of cry genes in native Bacillus thuringiensis strains isolated from wild ecological areas of East-Mediterranean region of Turkey. *Trop Ecol* **58**: 605–610 (2017).
11 Yılmaz S *et al.*, Toxicity of native bacillus thuringiensis isolates on the larval stages of pine processionary moth Thaumetopoea wilkinsoni at different temperatures. *Turkish J Agric For* **37**:163–172 (2013).
12 Yilmaz S, Azizoglu U, Ayvaz A, Temizgul R, Atciyurt ZB and Karabörklü S, Cloning and expression of cry2Aa from native Bacillus thuringiensis strain SY49-1 and its insecticidal activity against Culex pipiens (Diptera: Culicidae). *Microb Pathog* **105**:81–85 (2017).
13 Azizoglu U, Ayvaz A, Yılmaz S and Temizgul R, The synergic and antagonistic activity of Cry1Ab and Cry2Aa proteins against lepidopteran pests. *J Appl Entomol* **140**:223–227 (2016).
14 Azizoglu U, Ayvaz A, Yılmaz S, Karabörklü S and Temizgul R, Expression of cry1Ab gene from a novel Bacillus thuringiensis strain SY49-1 active on pest insects. *Braz J Microbiol* **47**:597–602 (2016).
15 Azizoglu U, Yilmaz S, Ayvaz A, Karabörklü S and Atciyurt ZB, Mosquitocidal potential of native Bacillus thuringiensis strain SY49-1 against disease vector, Culex pipiens (Diptera: Culicidae). *Trop Biomed* **34**: 256–262 (2017).
16 Panneerselvam S, Mishra R, Berry C, Crickmore N and Bonning BC, BPPRC database: a web-based tool to access and analyse bacterial pesticidal proteins. *Database (Oxford)* **2022**:baac022 (2022).
17 Azizoglu ZB, Yilmaz S, Azizoglu U, Karabörklü S, Temizgul R and Ayvaz A, Molecular characterization of the chitinase genes of native Bacillus thuringiensis isolates and their antagonistic activity against three important phytopathogenic fungi. *Biologia (Bratisl)* **76**:2745–2755 (2021).
18 Alvarez F, Castro M, Príncipe A, Borioli G, Fischer S, Mori G *et al.*, The plant-associated Bacillus amyloliquefaciens strains MEP2 18 and ARP2 3 capable of producing the cyclic lipopeptides iturin or surfactin and fengycin are effective in biocontrol of sclerotinia stem rot disease. *J Appl Microbiol* **112**:159–174 (2012).
19 Flissi A, Ricart E, Campart C, Chevalier M, Dufresne Y, Michalik J *et al.*, Norine: update of the nonribosomal peptide resource. *Nucleic Acids Res* **48**:D465–D469 (2020).
20 Wang P, Zhang C, Guo M, Guo S, Zhu Y, Zheng J *et al.*, Complete genome sequence of Bacillus thuringiensis YBT-1518, a typical strain with high toxicity to nematodes. *J Biotechnol* **171**:1–2 (2014).
21 Lucena WA *et al.*, Molecular approaches to improve the insecticidal activity of Bacillus thuringiensis cry toxins. *Toxins (Basel)* **6**:2393–2423 (2014).
22 Bravo A, Pacheco S, Gómez I and Soberón M, Chapter two - mode of action of Bacillus thuringiensis cry pesticidal proteins, in *Insecticidal Proteins and RNAi in the Control of Insects*, Vol. **65**, ed. by Jurat-Fuentes JLBT-A. Academic Press, USA, pp. 55–92 (2023).
23 Darsi S, Divya Prakash G and Udayasuriyan V, Cloning and characterization of truncated cry1Ab gene from a new indigenous isolate of Bacillus thuringiensis. *Biotechnol Lett* **32**:1311–1315 (2010).
24 Sasaki J, Asano S, Hashimoto N, Lay BW, Hastowo S, Bando H *et al.*, Characterization of a cry2A gene cloned from an isolate of Bacillus thuringiensis serovar sotto. *Curr Microbiol* **35**:1–8 (1997).
25 Alzate O, Osorio C, Florez AM and Dean DH, Participation of valine 171 in alpha-Helix 5 of Bacillus thuringiensis Cry1Ab delta-endotoxin in translocation of toxin into Lymantria dispar midgut membranes. *Appl Environ Microbiol* **76**:7878–7880 (2010).
26 Girard F *et al.*, Cysteine scanning mutagenesis of alpha4, a putative pore-lining helix of the Bacillus thuringiensis insecticidal toxin Cry1Aa. *Appl Environ Microbiol* **74**:2565–2572 (2008).
27 Rodríguez-Almazán C, Zavala LE, Muñoz-Garay C, Jiménez-Juárez N, Pacheco S, Masson L *et al.*, Dominant negative mutants of Bacillus thuringiensis Cry1Ab toxin function as anti-toxins: demonstration of the role of oligomerization in toxicity. *PLoS One* **4**:e5545 (2009).
28 Vachon V, Préfontaine G, Rang Ć, Coux F, Juteau M, Schwartz JL *et al.*, Helix 4 mutants of the Bacillus thuringiensis insecticidal toxin Cry1Aa display altered pore-forming abilities. *Appl Environ Microbiol* **70**:6123–6130 (2004).
29 Jiménez-Juárez N, Muñoz-Garay C, Gómez I, Saab-Rincon G, Damian-Almazo JY, Gill SS *et al.*, Bacillus thuringiensis Cry1Ab mutants affecting oligomer formation are non-toxic to Manduca sexta larvae. *J Biol Chem* **282**:21222–21229 (2007).
30 Zhu L, Peng D, Wang Y, Ye W, Zheng J, Zhao C *et al.*, Genomic and transcriptomic insights into the efficient entomopathogenicity of Bacillus thuringiensis. *Sci Rep* **5**:14129 (2015).
31 Bonis M, Felten A, Pairaud S, Dijoux A, Maladen V, Mallet L *et al.*, Comparative phenotypic, genotypic and genomic analyses of Bacillus thuringiensis associated with foodborne outbreaks in France. *PLoS One* **16**:e0246885 (2021).
32 Rosenquist H, Smidt L, Andersen SR, Jensen GB and Wilcks A, Occurrence and significance of Bacillus cereus and Bacillus thuringiensis in ready-to-eat food. *FEMS Microbiol Lett* **250**:129–136 (2005).
33 Cowles KN and Goodrich-Blair H, Expression and activity of a Xenorhabdus nematophila haemolysin required for full virulence towards Manduca sexta insects. *Cell Microbiol* **7**:209–219 (2005).

34 Heckel DG, The essential and enigmatic role of ABC transporters in Bt resistance of Noctuids and other insect pests of agriculture. *Insects* **12**:389 (2021).
35 Rodríguez H and Fraga R, Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol Adv* **17**:319–339 (1999).
36 Alvarez F, Simonetti E, Draghi WO, Vinacour M, Palumbo MC, do Porto DF *et al*., Genome mining of Burkholderia ambifaria strain T16, a rhizobacterium able to produce antimicrobial compounds and degrade the mycotoxin fusaric acid. *World J Microbiol Biotechnol* **38**:114 (2022).
37 Masalha J, Kosegarten H, Elmaci Ö and Mengel K, The central role of microbial activity for iron acquisition in maize and sunflower. *Biol Fertil Soils* **30**:433–439 (2000).
38 Batista BD, Dourado MN, Figueredo EF, Hortencio RO, Marques JPR, Piotto FA *et al*., The auxin-producing Bacillus thuringiensis RZ2MS9 promotes the growth and modifies the root architecture of tomato (Solanum lycopersicum cv. Micro-Tom). *Arch Microbiol* **203**:3869–3882 (2021).
39 Deleu M, Paquot M and Nylander T, Effect of fengycin, a lipopeptide produced by Bacillus subtilis, on model biomembranes. *Biophys J* **94**:2667–2679 (2008).
40 Emmert EAB, Klimowicz AK, Thomas MG and Handelsman J, Genetics of zwittermicin a production by Bacillus cereus. *Appl Environ Microbiol* **70**:104–113 (2004).
41 Willey JM and van der Donk WA, Lantibiotics: peptides of diverse structure and function. *Annu Rev Microbiol* **61**:477–501 (2007).
42 Wang J, Ma H, Ge X, Zhang J, Teng K, Sun Z *et al*., Bovicin HJ50-like lantibiotics, a novel subgroup of lantibiotics featured by an indispensable disulfide bridge. *PLoS One* **9**:e97121 (2014).
43 Béchet M, Caradec T, Hussein W, Abderrahmani A, Chollet M, Leclère V *et al*., Structure, biosynthesis, and properties of kurstakins, nonribosomal lipopeptides from Bacillus spp. *Appl Microbiol Biotechnol* **95**:593–600 (2012).
44 Théatre A *et al*., The Surfactin-like Lipopeptides from Bacillus spp.: natural biodiversity and synthetic biology for a broader application range. *Front Bioeng Biotechnol* **9**:623701 (2021).
45 Price NPJ, Rooney AP, Swezey JL, Perry E and Cohan FM, Mass spectrometric analysis of lipopeptides from Bacillus strains isolated from diverse geographical locations. *FEMS Microbiol Lett* **271**:83–89 (2007).
46 Zhao Y, Yang SZ and Mu BZ, Quantitative analyses of the isoforms of surfactin produced by Bacillus subtilis HSO 121 using GC-MS. *Anal Sci Int J Japan Soc Anal Chem* **28**:789–793 (2012).
47 Sarwar A, Brader G, Corretto E, Aleti G, Abaidullah M, Sessitsch A *et al*., Qualitative analysis of biosurfactants from bacillus species exhibiting antifungal activity. *PLoS One* **13**:e0198107 (2018).
48 Çolak A and Biçici M, Determination differentiating of Fusarium oxysporum formae speciales and determination incidence, severity, and prevalence of Fusarium wilt and crown-root rot in protected tomato growing areas of East Mediterranean region of Turkey. *Plant Prot Bull* **51**:331–345 (2011).
49 Pegg GF and Brady BL, *Verticillium Wilts*. CABI, UK (2002).

Pest Management Science published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.