



# Draft Genome Sequence of *Bacillus amyloliquefaciens* Strain CB, a Biological Control Agent and Plant Growth-Promoting Bacterium Isolated From Cotton (*Gossypium* L.) Rhizosphere in Coimbatore, Tamil Nadu, India

## OPEN ACCESS

### Edited by:

Dhaval K. Acharya,  
B N Patel Institute of  
Paramedical, India

### Reviewed by:

Nicolas Plaza,  
Autonomous University of Chile, Chile  
Maria Carolina Quecine,  
University of São Paulo, Brazil  
Ramesh K. Kothari,  
Saurashtra University, India

### \*Correspondence:

Nakkeeran Sevugapperumal  
nakkeeranayya@tnau.ac.in  
Vimalkumar S. Prajapati  
vimalprajapati100@gmail.com  
orcid.org/0000-0003-4257-1728

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

Received: 01 May 2021

Accepted: 08 July 2021

Published: 06 August 2021

### Citation:

Sevugapperumal N, Prajapati VS,  
Murugavel V and Perumal R (2021)  
Draft Genome Sequence of  
*Bacillus amyloliquefaciens* Strain CB,  
a Biological Control Agent and Plant  
Growth-Promoting Bacterium Isolated  
From Cotton (*Gossypium* L.)  
Rhizosphere in Coimbatore,  
Tamil Nadu, India.  
Front. Genet. 12:704165.  
doi: 10.3389/fgene.2021.704165

Nakkeeran Sevugapperumal<sup>1\*†</sup>, Vimalkumar S. Prajapati<sup>2\*†</sup>, Vanthana Murugavel<sup>1</sup> and Renukadevi Perumal<sup>1</sup>

<sup>1</sup> Department of Plant Pathology, Center for Plant Protection Studies, Tamil Nadu Agricultural University, Coimbatore, India,

<sup>2</sup> Division of Microbiology and Environmental Biotechnology, Aspee Shakilam Biotechnology Institute, Navsari Agricultural University, Surat, India

**Keywords:** *Bacillus amyloliquefaciens*, illumina hi seq, whole genome shotgun sequencing, biocontrol, PGPR, NGS

## INTRODUCTION

Although rhizobacteria have been widely explored for their plant growth-promoting capabilities and to manage various fungal and bacterial diseases in plants, viral diseases are an ongoing challenge in the agricultural sector (Vinodkumar et al., 2017). As most plant viral diseases are transmitted through vectors, researchers around the globe are utilizing biotechnological approaches to generate resistant lines. Various antagonistic bio-agents contribute to host defense, and various *Bacillus* species have been shown to produce these agents to protect against a wide range of pathogens. Several reports have demonstrated the antiviral efficacy of various *Bacillus* species against the cotton leaf curl virus (Ramzan et al., 2016), the cucumber mosaic virus in tomato (Zehnder et al., 2000), the tomato mottle virus in tomato (Murphy et al., 2000), and the tobacco mosaic virus in tobacco (Wang et al., 2009).

This bacterium is well known for the production of antibacterial, antiviral, and antifungal substances like *Bacillomycin* D, *Surfactin*, and *Bacillaene*, which protect the plant from pathogenic organisms (Chen et al., 2009). Additionally, the proteases and amylases produced by *Bacillus amyloliquefaciens* are used in industrial applications (Prajapati et al., 2015, 2017). *Bacillus* species belonging to this group are reported to have 24 diverse antimicrobial peptide (AMP) genes, which lead to the production of numerous compounds such as iturin, bacilysin, bacillomycin, fengycin, surfactin, mersacidin, ericin, subtilin, subtilosin, and mycosubtilin (Chung et al., 2008; Mora et al., 2011). Moreover, *Bacillus* species synthesize various volatile and non-volatile compounds that synergistically restrict plant diseases (Fernando et al., 2005; Mora et al., 2011). *B. amyloliquefaciens* CB has been used to prevent stem rot of carnations, and it was observed that minimum percentage disease incidence and maximum plant growth promotion occurred in plant treated with isolate CB. Further detailed experimentation will be carried out to evaluate the in-depth potential of the *B. amyloliquefaciens* CB.

Here, we report a draft genome sequence of *B. amyloliquefaciens* strain CB, which was isolated from rhizospheric soil of the cotton plant, collected from a cotton farm on the Tamil Nadu Agricultural University (TNAU) campus in Coimbatore, Tamil Nadu, India. This bacterium is gram positive with long rod-shaped, aerobic motile rods arranged singly or in chains. *B. amyloliquefaciens* belongs to the group of free-living soil bacteria, which aid to suppress plant pathogens and assist in promoting plant growth.

## VALUE OF THE DATA

The *B. amyloliquefaciens* CB draft genome can be used as a base/reference sequence to explore and map specific genes related to AMPs and other important enzymes. It could be a valuable resource to conduct comparative analyses among different species related to *B. amyloliquefaciens*, which may have similar biocontrol properties.

## METHODS AND DATA ANALYSIS

Bacterial DNA from the CB strain was extracted using phenol-chloroform methodology, and purification was performed using a Genomic DNA Clean and Concentrator (Zymo Research, Irvine, CA, USA). One nanogram of highly purified and good-quality DNA was used for the DNA fragment libraries prepared using a Nextera XT DNA sample preparation kit. Sequencing was performed on (2 × 150 paired-end reads with the Illumina v2 reagent kit) (Illumina, San Diego, CA, USA) an Illumina HiSeq system using the standard protocols described by the manufacturer. In total, 4,623,289 reads were obtained, and quality-based read trimming was done using the Trimmomatic software (version 0.30) (Bolger et al., 2014) followed by quality checking with FastQC (version 3.0) (Andrews, 2010). The genome was assembled using GS De Novo Assembler v. 2.6, ABySS v. 1.5.1, Celera Assembler v. 8.3rc2, Edena v. 3.131028, Megahit v. 1.1.2, SOAPdenovo v. 2.04, Velvet v. 1.2.10, SPAdes v. 3.1.1, and SPAdes v. 3.11.0; and the final assembly was merged using CISA v. 1.3 and submitted to the National Center for Biotechnology Information (NCBI) GenBank having accession no. WODE00000000. The total sequence length has been counted up to be 4,113,229 bp consisting of 11 scaffolds/contigs, a contig N50 of 675,513 with L50 of 3, with the largest contig size of the submitted assembly being 1,050,139. The genome size was estimated to be 4.11 MB with a guanine–cytosine (GC) content of 46.30%. Gene annotation was performed using NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016), which identified 3,847 protein-coding sequences (Table 1).

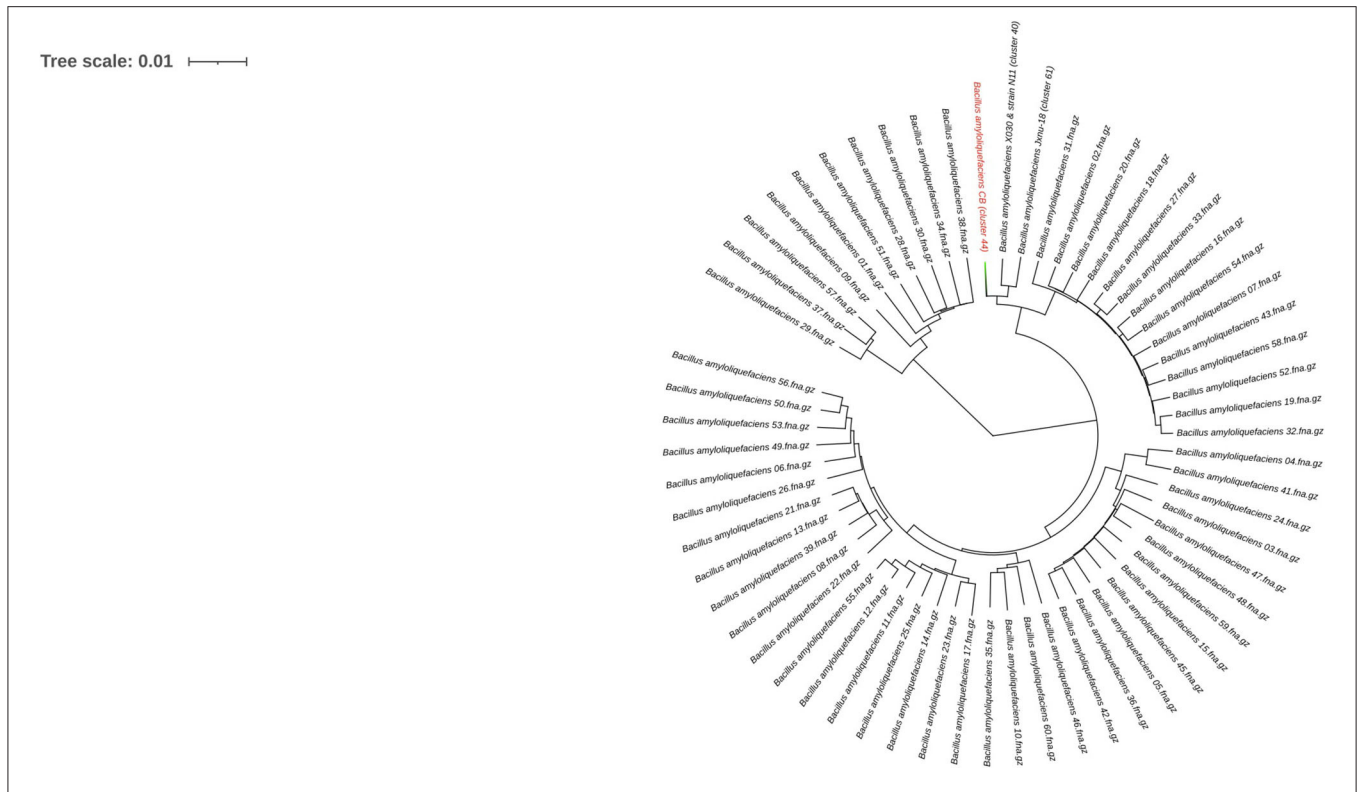
To infer the phylogenetic relationship, all the 119 assemblies (Supplementary File 1) of *B. amyloliquefaciens* accessible in the NCBI database were considered for the Bacsort analysis (<https://github.com/rrwick/Bacsort>) including the *B. amyloliquefaciens* CB. A total of 61 clusters of the considered assemblies were generated (cluster accession provided in Supplementary File 2), and FastANI was employed to generate the matrix of all pairwise distance between the clusters. FastANI algorithm

**TABLE 1** | Genomic features of *Bacillus amyloliquefaciens* strain CB annotated using National Center for Biotechnology Information—Prokaryotic Genome Annotation Pipeline (NCBI-PGAP) v. 4.10.

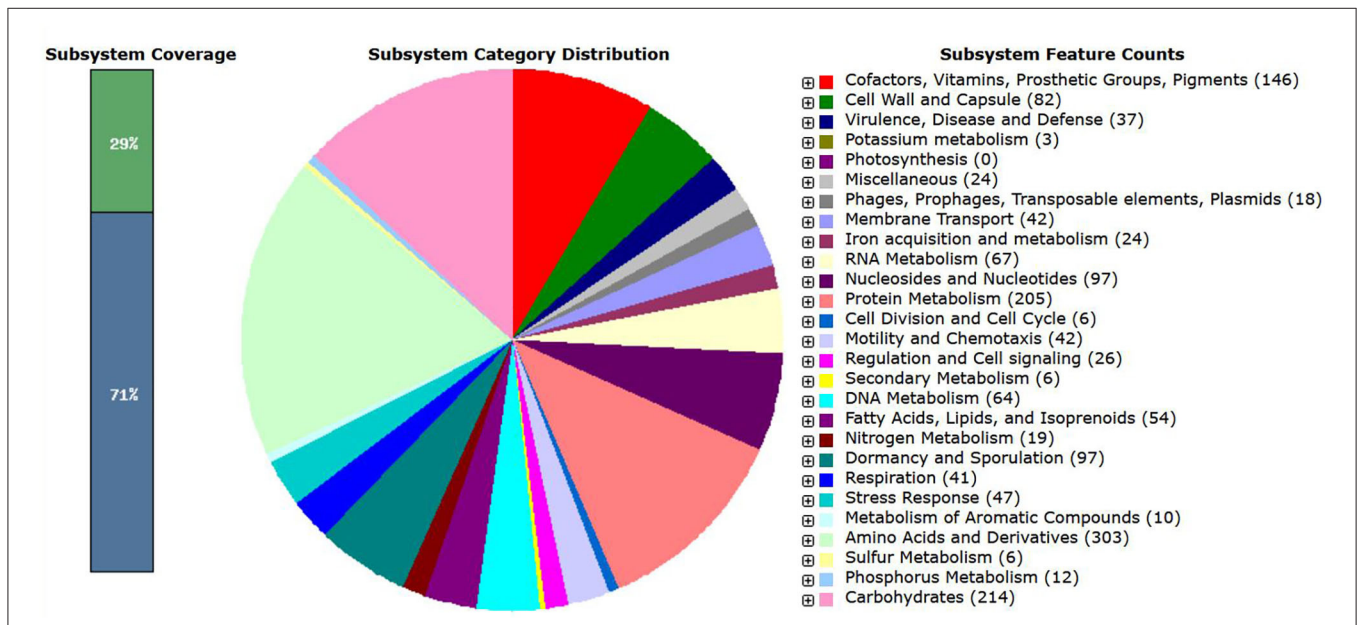
Items	Counts
Total genes	4,012
Total CDS	3,911
Coding genes	3,847
Coding CDS	3,847
Genes (RNA)	101
rRNAs	8, 8, 7 (5S, 16S, 23S)
tRNAs	73
ncRNAs	5
Total pseudo genes	64

(<https://github.com/ParBLiSS/FastANI>) generates pairwise Average Nucleotide Identity (ANI) measurements using the only sequence shared by two assemblies (Supplementary File 3), which makes it less swayed due to the accessory genome and produce more accurate trees. The phylogeny tree was created by BIONJ algorithm with bootstrap value of 1,000 to form the generated data and was drawn precisely using Interactive Tree Of Life (iTOL) v5, which is an online tool for the display, annotation, and management of phylogenetic trees (Letunic and Bork, 2021) (Figure 1). Out of 61 clusters, two distinct nodes were generated, in which 51 leaves and 10 leaves form a separate group. Fifty-one leaves split into another group having 17 leaves and 34 leaves, which generate other groups consequently as shown in Figure 1. The *B. amyloliquefaciens* strain CB forms a separate cluster (44) having branch length 0.00582, while its nearby cluster (40) includes two strains, *B. amyloliquefaciens* X030 and *B. amyloliquefaciens* N11 (branch length 0.00456), while the cluster (61) comprises the *B. amyloliquefaciens* strain Jxnu-18 (branch length 0.00520). Clusters 44, 40 and 61 originated from a common node having branch length 0.00141 (Figure 1).

The genome of *B. amyloliquefaciens* CB was also mapped to the seed subsystem to obtain the high-quality genome annotation through Rapid Annotation using the Subsystem Technology (RAST; version 2.0) (<http://rast.nmpdr.org>) (Overbeek et al., 2014). The total 325 subsystem with 29% subsystem coverage resulted for *B. amyloliquefaciens* strain CB through RAST server (Figure 2). The present investigation revealed that highest number of the genes was allocated to the subsystem category of amino acids and derivatives (303 genes) followed by carbohydrates (214 genes); protein metabolism (205 genes); cofactors, vitamins, prosthetic groups, and pigments (146 genes); nucleosides and nucleotides (97 genes); dormancy and sporulation (97 genes); cell wall and capsule (82 genes); RNA metabolism (67 genes); DNA metabolism (64 genes); fatty acids, lipids, and isoprenoids (54 genes); stress response (47 genes); motility and chemotaxis (42 genes); membrane transport (42 genes); respiration (41 genes); and virulence, disease, and defense (37 genes). A total of 24 genes were found to be associated with iron acquisition and metabolism as well 24 genes for some other miscellaneous applications. More precisely in the



**FIGURE 1 |** Genome lineage tree of *Bacillus amyloliquefaciens* strain CB. BIONJ algorithm was used to construct the Newick format tree from the distance matrix of all the clustered assemblies (119 different strains of *B. amyloliquefaciens*). The generated Newick tree file was analyzed using the Interactive Tree Of Life (iTOL) v. 5 (strain CB highlighted in red color text with light green color branch).



**FIGURE 2 |** Schematic overview of subsystem coverage, distribution of subsystem category, and subsystem feature counts anticipated in *Bacillus amyloliquefaciens* strain CB using RAST server.

category of miscellaneous application, 10 genes were specifically associated to iron–sulfur cluster assembly, five genes for niacin-choline transport and metabolism, and one gene for single-rhodanese-domain proteins. A total of 26 genes were found to be associated with regulation and cell signaling, while 18 genes were collectively specified for phages, prophages, transposable elements, plasmids, and 12 genes for phosphorus metabolism.

Most of the *Bacillus* spp. belonging to this group of genera have been reported to have antifungal potential and have been utilized for the management of the various fungal diseases; however, their efficacy against viral diseases is still not known (Vinodkumar et al., 2018). The PGAP annotation confirms that the *B. amyloliquefaciens* strain CB genome has gene locus srfAA, srfAD, srfAB, and srfAC, which produce various peptides like surfactin non-ribosomal peptide synthetase and surfactin biosynthesis thioesterase. It has been well documented that lipopeptides like surfactin have acquired more attention due to their high surface activity and antibiotic potential. Moreover, surfactin also possesses antiviral, antitumor, and hemolytic activities (Wang et al., 2010), which required further intensive experimentation for characterization to understand its exact mechanism for such action.

The whole-genome shotgun sequence of *B. amyloliquefaciens* strain CB and its annotation report presented here provide a resource for comparative analysis with other genera of *Bacillus* and can be used for engineering purposes where characteristics of the strain CB are desired. The genome representation of *B. amyloliquefaciens* strain CB showed antagonistic potential due to various AMPs imparting various properties like antifungal, antibacterial, and antiviral as well plant growth promotion, leading to strong future prospects for uplifting the sustainable agriculture.

## DATA AVAILABILITY STATEMENT

*Bacillus amyloliquefaciens* strain CB, whole genome shotgun sequencing project data have been deposited at DDBJ/ENA/GenBank under the accession number WODE00000000. The version described in this data report

## REFERENCES

- Andrews, S. (2010). *Fastqc - a Quality Control Tool for High Throughput Sequence Data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Bolger, Anthony M., Marc, Lohse., and Bjoern, Usadel. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–212. doi: 10.1093/bioinformatics/btu170
- Chen, X.H., Koumoutsis, A., Scholz, R., Schneider, K., Vater, J., Süssmuth, R., et al. (2009). Genome analysis of *Bacillus amyloliquefaciens* FZB42 reveals its potential for biocontrol of plant pathogens. *J. Biotechnol.* 140, 27–37. doi: 10.1016/j.jbiotec.2008.10.011
- Chung, S., Kong, H., Buyer, J., Lakshman, D.K., Lydon, J., Kim, S.D. et al. (2008). Isolation and partial characterization of *Bacillus subtilis* ME488 for suppression of soil borne pathogens of cucumber and pepper. *Appl. Microbiol. Biotechnol.* 80, 115–123. doi: 10.1007/s00253-008-1520-4

is the first version having accession number WODE00000000.1. The assembled contigs and its annotation files (CDS, gff, and proteins) are available in [https://www.ncbi.nlm.nih.gov/assembly/GCA\\_011754125.1/#/st](https://www.ncbi.nlm.nih.gov/assembly/GCA_011754125.1/#/st) repository with all the annotations details in Readme file.

## AUTHOR CONTRIBUTIONS

NS: funding and modeling the study. VP: genome assembly, annotations, and analysis. VP and NS: manuscript preparation. VM and RP: sampling and sequencing and other miscellaneous stuff. All authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

We all authors are acknowledging Dr. Anthony E. Zamora, Assistant Professor [Medicine (Hematology and Oncology) and Microbiology; Immunology], Medical College of Wisconsin, USA for proof reading and correcting the manuscript for its language competence.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.704165/full#supplementary-material>

**Supplementary File 1** | List of the all the genomes/assemblies of *B. amyloliquefaciens* considered in the presented study (strain, biosample and bioproject information, assembly level, GC%, Scaffold, CDS).

**Supplementary File 2** | Cluster assemblies: Cluster of assemblies were generated with removing the redundancy of assemblies. Similar assemblies are forming one cluster and have only a single representative chosen based on assembly N50 (Cluster\_accessions file lists the cluster name, followed by a tab, followed by a comma-delimited list of the assemblies in that cluster, with the representative assembly marked with a \*).

**Supplementary File 3** | Distance matrix: FastANI produces pairwise ANI measurements of all the generated assemblies's cluster using only the sequence shared by two assemblies.

- Fernando, W. G. D., Ramarathnam, R., and Krishnamoorthy, A. S. (2005). Identification and use of potential bacterial organic antifungal volatiles in biocontrol. *Soil Biol. Biochem.* 37, 955–964. doi: 10.1016/j.soilbio.2004.10.021
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, 293–296. doi: 10.1093/nar/gkab301
- Mora, I., Cabrefiga, J., and Montesinos, E. (2011). Antimicrobial peptide genes in *Bacillus* strains from plant environments. *Int. Microbiol.* 14, 213–223. doi: 10.2436/20.1501.01.151
- Murphy, J.F., Zehnder, G.W., Schuster, D.J., Sikora, E.J., Polston, J.E., and Kloepper, J.W. (2000). Plant growth-promoting rhizobacterial mediated protection in tomato against tomato mottle virus. *Plant Dis.* 84, 779–784. doi: 10.1094/PDIS.2000.84.7.779
- Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, 206–214. doi: 10.1093/nar/gkt1226
- Prajapati, V.S., Ray, S., Narayan, J., Joshi, C.G., Patel, K.C., Trivedi, U.B., et al. (2017). Draft genome sequence of a thermostable, alkaliphilic  $\alpha$ -amylase and



- protease producing *Bacillus amyloliquefaciens* strain KCP2. 3. *Biotech* 7:372. doi: 10.1007/s13205-017-1005-1
- Prajapati, V.S., Trivedi, U.B., and Patel, K.C. (2015). A statistical approach for the production of thermostable and alklophilic alpha-amylase from *Bacillus amyloliquefaciens* KCP2 under solid-state fermentation. *Biotech* 5, 211–220. doi: 10.1007/s13205-014-0213-1
- Ramzan, M., Tabassum, B., Nasir, I.A., Khan, A., Tariq, M., Awan, M.F., et al. (2016). Identification and application of biocontrol agents against 147 cotton leaf curl virus disease in *Gossypium hirsutum* under greenhouse conditions. *Biotechnol. Equip.* 30, 469–478. doi: 10.1080/13102818.2016.1148634
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569
- Vinodkumar, S., Nakkeeran, S., Renukadevi, P., and Malathi, V.G. (2017). Biocontrol potentials of antimicrobial peptide producing *Bacillus* species: multifaceted antagonists for the management of stem rot of carnation caused by *Sclerotinia sclerotiorum*. *Front. Microbiol.* 8:446. doi: 10.3389/fmicb.2017.00446
- Vinodkumar, S., Nakkeeran, S., Renukadevi, P., and Mohankumar, S. (2018). Diversity and antiviral potential of rhizospheric and endophytic *Bacillus* species and phyto-antiviral principles against tobacco streak virus in cotton. *Agric. Ecosyst. Environ.* 267, 42–51. doi: 10.1016/j.agee.2018.08.008
- Wang, S., Huijun, W., Junqing, Q., Lingli, M., Jun, L., Yanfei, X., et al. (2009). Molecular mechanism of plant growth promotion and induced systemic resistance to tobacco mosaic virus by *Bacillus* spp. *J. Microbiol. Biotechnol.* 19, 1250–1258. doi: 10.4014/jmb.0901.008
- Wang, Yu., Zhaoxin, Lu., Xiaomei, Bie., and Fengxia, L.V. (2010). Separation and extraction of antimicrobial lipopeptides produced by *Bacillus amyloliquefaciens* ES-2 with macroporous resin. *Eur. Food Res. Technol.* 231, 189–196. doi: 10.1007/s00217-010-1271-1
- Zehnder, G.W., Yao, C., Murphy, J.F., Sikora, E.R., and Kloepper, J.W. (2000). Induction of resistance in tomato against cucumber mosaic cucumovirus by plant growth-promoting rhizobacteria. *Biocontrol* 45 (1), 127–137. doi: 10.1023/A:1009923702103

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Sevugapperumal, Prajapati, Murugavel and Perumal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.