

# Genome Mining of Antimicrobial Gene Clusters from Phyllospheric Bacteria Isolated from *Solanum Lycopersicum* and *Lactuca Sativa* to Identify Biocontrol Properties

Claudia Y. Muñoz

University of Groningen

Lu Zhou

University of Groningen

Yunhai Yi

University of Groningen

Oscar P. Kuipers (✉ [o.p.kuipers@rug.nl](mailto:o.p.kuipers@rug.nl))

University of Groningen

---

## Research Article

**Keywords:** Biocontrol, Phyllosphere, Bacillus, Paenibacillus, Antimicrobials, Biosynthetic gene clusters, NRPs, Bacteriocins, RiPPs

**Posted Date:** June 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-633887/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Biocontrol agents are sustainable eco-friendly alternatives for chemical pesticides that cause adverse effects in the environment and toxicity in animals including humans. An improved understanding of the phyllosphere microbiology is of vital importance for biocontrol development. Most studies have been directed towards beneficial plant-microbe interactions and ignore the pathogens that might affect humans when consuming vegetables. In this study we extended this perspective and investigated potential biocontrol strains isolated from the tomato and lettuce phyllosphere that can promote plant growth and antagonize mammalian pathogens as well as plant pathogens. Subsequently, we mined into their genomes for discovery of antimicrobial biosynthetic gene clusters (BGCs), several of which are good candidates to produce protectants against microbial plant and mammalian pathogens.

## Results

The antimicrobial activity of 69 newly isolated strains from a healthy tomato and lettuce phyllosphere against several plant and mammalian pathogens was determined with plates assays. Three strains with the highest antimicrobial activity against the relevant pathogens were selected and characterized (*Bacillus subtilis* STRP31, *Bacillus velezensis* SPL51, and *Paenibacillus sp.* PL91). All three strains showed a plant growth promotion effect by the production of volatile compounds (VOCs) on tomato and lettuce. In addition, genome mining of these isolates showed the presence of a large variety of biosynthetic gene clusters. A total of 39 BGCs were identified, of which several are already known, such as bacilysin, bacillibactin, surfactin, subtilomycin, etc., but also several novel ones. Further analysis revealed that among the novel BGCs, one NRPS and two bacteriocins are encoded which were analyzed in more depth.

## Conclusions

Several antimicrobial BGCs were found in the selected strains, including the rediscovery of known ones, but also the discovery of novel ones. Our study serves as support for subsequent examination and characterization of novel antimicrobial metabolites, and the possibility of developing biocontrol agents.

## Background

There is an increasing demand by consumers for nutritious foods that improve physical performance and reduce risks of diseases. Vegetables represent a widely consumed food worldwide. *Solanum lycopersicum* better known as tomato is one of the most important vegetable plants in the world (1). Tomato is a healthy food that supplies a wide range of vitamins needed for the organism, since it contains high levels of zinc, potassium, anthocyanins and lycopene, which provide a high antioxidant power. Along with other compounds they reduce risk of contracting cancer, among other benefits according to several epidemiological-food- and health studies (1, 2). On the other hand, *Lactuca sativa*

better known as lettuce is another important crop with a growing interest from people due to its healthy and beneficial properties and richness in antioxidants (e.g., vitamins C, E and carotenoids) (3).

Agriculture in the last century has faced multiple challenges, including the need to produce more food to feed a growing population, adapting to climate change, controlling crop diseases that cause significant losses, and adopting more efficient and sustainable production methods. An increased number of fungi, bacteria and viruses are causing plant diseases, several of which are the reason of major economic losses (4). For this reason, food security has become one of the main points of attention in human-driven development, and therefore any plant pathogen causing substantial crop yield losses needs to be minimized (5). Currently, tomato and lettuce production losses due to biotic agents (insect pests, microbe-or- virus caused diseases and weeds) have been controlled mainly by spraying crops with a vast amount of synthetic chemical pesticides. However, the extensive use of them caused adverse side effects which represent a serious threat to living organisms and the environment. In addition, for many plant pathogens, fungicide-resistant populations have made many fungicides ineffective (6).

Some mammalian pathogens, such as *Bacillus cereus*, attach and form biofilm on lettuce leaf surfaces posing a risk of causing disease in humans upon consumption (7). *Klebsiella pneumoniae* and *Escherichia coli* are other known pathogens that have been reported as causative agents of foodborne diseases, being found in different sources including raw vegetables (8). Contamination by pathogenic bacteria of vegetables can occur during the harvesting period, post harvesting, handling, storage, transportation, and processing by customers. To ensure safety level of vegetables and others, research of biocontrol agents against mammalian pathogen is needed as well.

It is now well established that plant-associated microorganisms play essential roles in plant health and development and contribute to the environmental equilibrium (9). The use of beneficial microorganisms is a promising method to fight against crop diseases and increase yields to ensure sufficient crop production (10). The structural and functional analysis of microbial genomes and the proteins encoded by genes of important plant-associated microbes, which can be possible biocontrol agents will provide insights into several aspects of these molecular interactions and will be crucial for the development of more efficient control measures (11).

Historically, research investigating the factors associated with plant microbe interactions has focused on the rhizosphere, which comprises the area in the soil around plant roots, but much less is known about the phyllosphere. The phyllosphere represents the aboveground or aerial parts of the plants dominated by the leaves in contrast to the rhizosphere that represents the below-ground part of plants. Environmental factors, including UV radiation, changes in relative humidity, temperature, leaf wetness, pollution, nitrogen fertilization as well as biotic factors, such as leaf age and the presence of other microorganisms are factors that microbe endure in such environment (12).

Gram-positive, aerobic spore forming bacteria, like *Bacillus* and *Paenibacillus spp.*, have been widely reported to be effective in stimulating plant growth and are well known as producers of a broad array of

antimicrobials having between 5 to 8% of the total genome devoted to the biosynthesis of secondary metabolites (13).

*Bacillus* and closely related species antimicrobials production is highly diverse, depending on their biosynthesis pathways, and chemical nature. Antimicrobial compounds can be classified into bacteriocins (both linear and ribosomally synthesized and postranslationally modified peptides (RiPPs)), non-ribosomally synthesized peptides (NRPs), and Polyketides (14). NRPs and PKs natural products are synthesized via multi domain mega enzymes known as non-ribosomally synthesized peptides synthases (NRPSs) and polyketide synthases (PKSs) which are arranged into units called 'modules' that work in an assembly-line system to build polymeric peptide chains with a determined function (15). PKSs gather small acetic acid-type acyl construction blocks into polyketides through C–C bonds, and NRPSs gather amino acids into peptides through amide bonds. NRPSs and PKSs employ a similar strategy for the biosynthesis of different classes of natural products (16). On the other hand, bacteriocins are ribosomally synthesized. Bacteriocin BGCs are smaller than the previous mentioned compounds, and carry one or more precursor peptide gene(s), which allows a finer prediction of the ending product structure based on the properties of enzymes involved in their biosynthesis and on the chemical structure of the initial peptide substrate. Bacteriocins can be classified in peptides that undergo post-translational modifications (class I) also known as Ribosomally synthesized and post-translationally modified peptides (RiPPs), or largely unmodified peptides, that sometimes contain disulfide bonds (class II) (17).

Reported and well-known NRPs antimicrobials produced by *B. subtilis* and *B. velezensis* are cyclic lipopeptides exhibiting strong surfactant and antimicrobial activities, such as surfactins, bacillibactins and fengycins (18, 19). *B. velezensis*, the biocontrol model also employs polyketides such as macrolactin, bacillaene, and difficidin, which play significant roles in both pathogen suppression and plant growth promotion (20). NRPs fusaricidin and polymyxin produced by *Paenibacillus* strains, contributes to antagonism against phytopathogens like *Erwinia spp.* (21), while well studied *Bacillus*-originated bacteriocins include subtilocin and subtilomycin (22).

In the last decade volatile compounds (VOCs) produced by some plant-associated bacteria and their biological function, have attracted increased attention. Among other characteristics, they have been proved to have the potential as antimicrobials and have plant growth promoting properties. The VOCs 2,3-butanediol and acetoin, could trigger growth promotion in *Arabidopsis thaliana* rhizosphere. *Bacillus megaterium* XTBG34 produces 2-pentylfuran which promotes the growth of *Arabidopsis thaliana* plants after 15 days of treatment (23, 24). In our study we aimed to isolate and screen novel phyllospheric bacteria with antimicrobial properties, and further mined into their genomes to identify known or novel biosynthetic gene clusters (BGCs) that are potentially involved in phytopathogen, and plant-originated mammalian pathogen antagonism.

## Results

### Isolation and antimicrobial activity

A total of 69 strains were isolated from internal and external leaf tissue from tomato and lettuce plants in early growth stage coming from a small farm in Roden, The Netherlands (Table S1). The antimicrobial activity of the 69 strains was first evaluated against five major tomato and lettuce pathogens, i.e., *Rhizoctonia solani* (25), *Botrytis cinerea* (26), and *Phytophthora ultimum* (27) and two gram-negative bacterial pathogens, i.e. *Pseudomonas syringae* pv tomato DC300 (28), and *Erwinia carotovora* (29, 30), covering the two major classes of pathogens, i.e. fungi and Gram-negative bacteria. The three most outstanding strains with the highest antimicrobial activity were selected and their antimicrobial spectrum was extended, to evaluate activity against a wider range of pathogens and to further evaluate properties of plant growth promotion.

In addition, raw and not-processed vegetables, such as tomato and lettuce in salads, are considered as a source of food-borne diseases as well. Some mammalian pathogens can inhabit plants as secondary hosts. In the present study some mammalian pathogens (*Klebsiella pneumoniae* (31), *Bacillus cereus* (32) *Escherichia coli* (33) and *Candida albicans* (34)) were also included in the antimicrobial activity tests as indicators. Three bacteria with the highest antimicrobial activity were selected: STRP31 isolated from tomato, and SPL51, PL91 both isolated from lettuce. These strains can antagonize Gram-positive pathogenic bacteria, Gram-negative pathogenic bacteria, a yeast, and fungal pathogens (Table 1 and Fig. 1). The other 66 strains showed a significant lower activity towards bacterial and fungal pathogens tested and had a more limited spectrum of activity.

Table 1

Antimicrobial activity of bacterial strains *Bacillus velezensis* SPL51, *Bacillus subtilis* STRP31 and *Paenibacillus* sp. PL91

| Pathogen types         | Species or strains   | SPL51                  | STRP31 | PL91 |
|------------------------|--|------------------------|--------|------|
| Fungi                  | <i>Rhizoctonia solani</i>                                      | ●                      | ●      | ●    |
|                        | <i>Botrytis cinerea</i>  | ●                      | ●      | ●    |
|                        | <i>Verticillium dahliae</i>                                    | ●                      | ●      | ●    |
|                        | <i>Fusarium culmorum</i>                                       | ●                      | ●      | ●    |
|                        | <i>Fusarium oxysporum</i>                                      | ●                      | ●      | ●    |
|                        | <i>Alternaria solani</i>                                       | ●                      | ●      | ●    |
|                        | <i>Septoria lycopersici</i>                                    | ●                      | ●      | ●    |
|                        | <i>Aspergillus niger</i>                                       | ●                      | ●      | ●    |
|                        | Oomycete   | <i>Phytium ultimum</i> | ●      | ●    |
| Yeast                  | <i>Candida albicans</i>  | ++                     | +      | -    |
| Gram-negative bacteria | <i>Escherichia coli</i> WA321 *                                | +                      | +      | -    |
|                        | <i>Klebsiella pneumonia</i> *                                  | ++                     | +      | -    |
|                        | <i>Pseudomonas syringae</i> pv. <i>tomato</i> DC300            | ++                     | +      | ++   |
|                        | <i>Erwinia corotovor</i> a subsp. <i>brasiliensis</i> LMG21371 | ++                     | -      | -    |
| Gram-positive bacteria | <i>Bacillus cereus</i> *                                       | ++                     | +      | +    |

In the antibacterial assay, no inhibition (-), inhibitory zone < 5mm (+), inhibitory zone ≥ 5mm (++) . In the antifungal assay, no inhibition (●), clear inhibition (●). \*Mammalian pathogens

## Genome Sequencing Of The Selected Strains And Phylogenetic Analyses

The genomes of the selected strains were sequenced, and their DNA sequences were reported previously (35). Phylogenetic analysis using these whole-genome sequences was conducted with GTDB-Tk v1.5.0 (36) and a phylogenetic dendrogram was constructed with the obtained sequences by IQ-TREE and visualized by iTOL web-based tool (37) (Fig. 2). The species names of the three strains were designated as their most closely related strains namely *B. velezensis* SPL51, *B. subtilis* STRP31, and unidentified *Paenibacillus* sp. PL91.

# Plant Growth Promotion Effect

In vitro experiments were conducted to investigate the effect of volatile organic compounds (VOCs) produced by the three strains (SPL51, STRP31 and PL91) on plant growth of tomato and lettuce plants. Comparisons were made using *B. subtilis 168* as negative control. We observed that the bacterial isolates exerted a significant influence on the increase of plant biomass of tomato and lettuce plants. However, we did not observe a significant difference in the length of the plants when comparing with the controls. No significant differences were noted when the results of plants exposed to the VOCs of *B. subtilis 168* were compared to the un-inoculated control. The plant growth promoting potential of VOCs produced by SPL51, STRP31 and PL91 is depicted in Fig. 3.

## Genome Mining For Bgcs

The genome mining of biosynthetic gene clusters of antimicrobial compounds was conducted with antiSMASH 5.0 (59) and BAGEL4 (60). Before applying to the pipelines, the draft genome of each strain was assembled into a pseudomolecule using closely related strains as a reference using Medusa web server (61) (<http://combo.dbe.unifi.it/medusa>). The genes predicted from both pipelines were further confirmed with protein BLAST. BGCs that showed equal or less than 80% protein identity to the reported ones were regarded as novel.

## Potential Novel Bgcs

After genome mining, the three more interesting, potentially novel BGCs were chosen based on sequence similarity to other compounds and comparing the precursor peptide. Selected BGCs are listed in Fig. 5. In the genome of *Paenibacillus sp.* PL91 a modular NRPS gene cluster with all essential modules (starting module, elongation module, termination module) was found (Fig. 5a). The BGC consists of nine genes and has a total size of 88 kb, it shows 25% of sequence similarity to the cyclic antimicrobial decapeptide tyrocidine, produced by *Brevibacillus brevis* NBRC, which exhibits antifungal and antibacterial activity against Gram-positive bacteria. Tyrocidine A, B and C produced by *Brevibacillus* spp. share structural similarities, summarized as Cyclo (Phe - Pro<sub>2</sub> - Phe/Trp<sub>3</sub> - Phe/Trp<sub>4</sub> - Asn<sub>5</sub> - Gln<sub>6</sub> - Tyr<sub>7</sub> - Val<sub>8</sub> - Orn<sub>9</sub> - Leu<sub>10</sub>). Residues 1,2, and 5 to 10 are conserved among the three tyrocidine members produced by *Brevibacillus* (50). However, in the NRPs BGC found in PL91 neither the order of genes nor the predicted amino acid composition shows similarity to tyrocidine, indicating the likely novelty of the final product.

Another BGC found in *Paenibacillus sp.* PL91 was a linear azol(in)e-containing peptide (LAP), an important subgroup of RiPPs with a distinguishing heterocyclic ring of oxazoles and thiazoles derived from serine/ threonine and cysteine by enzymatic cyclodehydration and dehydrogenation (51). This BGC consists of five genes and has a total size of 41 kb. The precursor peptide encoded by the core biosynthetic gene has 83 amino acids and shows no similarity with sequences of other bacteriocins.

Since it is a bacteriocin it might be possible that it is partially responsible for the antimicrobial activity against Gram-positive bacteria, which needs to be further investigated.

An interesting lasso peptide BGC, discovered in *Paenibacillus* sp. PL91, has 80% sequence similarity to paeninodin. Paeninodin was firstly discovered in *Paenibacillus dendritiformis* C454. It was firstly reported in the year 2016 for being a novel lasso peptide tailored by a new class of kinases (52). Lasso peptide biosynthesis requires at least three genes, referred to as the A, B, and C proteins. The A gene encodes the precursor peptide, which is modified by the B and C proteins into the mature natural product (53). The Lasso peptide BGC found in strain PL91 consists of six genes and has a total size of 23 kb. When we compared lasso peptide BGC in PL91 with the paeninodin lasso peptide, we found they have a highly similar gene organization, and the same number genes too. However, the suggested functionality of the genes is not completely the same. After further analysis with Blastp, the results reveal that the first gene might be an isopeptide-bond forming cyclase protein with cyclization function. The second gene corresponds to the precursor gene. The corresponding precursor peptide compared to the paeninodin produced by *Paenibacillus dendritiformis* C454 is not the same, thus, this finding suggests a probable novelty of the final product (Fig.5c). The third gene is suggested to correspond to an aldolase, which is also present in other *Paenibacillus* spp., while the third gene of *Paenibacillus dendritiformis* is Pade-K, a gene encoding a kinase, which is not present in the lasso peptide BGC of *Paenibacillus* sp. PL91. The fourth gene corresponds to be a PqqD family chaperone with a possible function as a maturation enzyme. The fifth gene appears to be a B2 protein, with a proposed function as a maturation enzyme as well, B2 part of 'split B', protease. Finally, the sixth gene corresponds to be an ABC transporter, with a proposed export function (Table 2).

**Table 2.** Proposed functions of ORFs in the putative lasso peptide biosynthetic gene cluster

identified by genome mining in *Paenibacillus* sp. PL91 and reported *Bacillus Dendritiformis* C454

| Gene          | Paeninodin<br><i>Paenibacillus dendritiformis</i>                                   | Gene | Lasso peptide BGC<br><i>Panibacillus</i> PL91 |
|---------------|---|------|---|
| <i>padeC</i>  | Maturation enzyme, cyclization, adenylation   | 1    | Isopeptide bond-forming cyclase               |
| <i>padeA</i>  | Precursor peptide   | 2    | Precursor peptide                             |
| <i>padeK</i>  | Putative kinase   | 3    | Aldolase                                      |
| <i>padeB1</i> | Maturation enzyme PqqD superfamily. B1 part of 'split B', protease                  | 4    | Maturation enzyme PqqD chaperone              |
| <i>padeB2</i> | Maturation enzyme, transglutaminase-like superfamily B2 part of 'split B', protease | 5    | B2 protein maturation enzyme                  |
| <i>padeD</i>  | Export ABC type transport system  | 6    | ABC transporter ATP-binding protein           |

## Discussion

In this study we have identified three promising new biocontrol strains, isolated from tomato and lettuce leaves. We show the presence of a large variety and number of biosynthetic gene clusters among the three selected strains. However, the majority of BGCs identified in *Bacillus velezensis* SPL51 and *Bacillus subtilis* STRP31 were already known, and both strains represent two highly studied species with biocontrol potential, although these could have additional advantages. *Paenibacillus sp.* PL91 harbored most of the unknown BGCs, including a potentially novel NRP, and two bacteriocin (one LAP and one lasso peptide) BGCs. We suspect that the novel BGCs found in this study have antibacterial and (or) antifungal activity from the antimicrobial assay realized, but this needs to be confirmed by future experiments.

Bacteriocins have many properties including their potency (as determined in vitro and in vivo), their low toxicity, the availability of both broad- and narrow-spectrum peptides, the possibility of in situ production by probiotics and the fact that they can be bioengineered. One of the novel BGCs found is a lasso peptide BGC. Lasso peptides are characterized by having peptide bonded proteinogenic amino acids with an extra linkage between the N-terminus of the peptide and an aspartate or glutamate sidechain, which gives them the resembling characteristic topology of a slipknot. Most lasso peptides carry Gly, Cys, Ser or Ala at the N-terminus, and their structure confers them an excellent stability against thermal, chemical, and proteolytic degradation. These features make them attractive for the generation of more potent compounds by engineering. Some lasso peptides are known for their antimicrobial activity, enzyme inhibition, anticancer properties, and antiviral activity. Microcin J25, specialicin, achromosin, sphaericin, actinokineosin, lassomycin are some examples of lasso peptides with antimicrobial activity (54). This

makes them attractive not only for biocontrol purposes but also for pharmaceutical research. The lasso peptide present in strain PL91 has significant differences when compared to other known lasso peptides, suggesting that it is a novel variant of paeninodin, thus making it highly interesting for antimicrobial activity and drug discovery.

On the other hand, the LAP BGC found in PL91 has no similarity with any other known compound. LAPs main components include a precursor peptide (A), and a heterotrimeric enzyme complex composed of a dehydrogenase (B) and cyclodehydratase (C and D). The ribosomal synthesis of the precursor peptide is the first step in the biosynthesis. After the generation of the enzyme-substrate complex, ATP-dependent cyclodehydration takes place, developing into an azoline heterocycle. The most characteristic feature in a LAP is the genetically and biochemically conserved installation of azol(in)e rings on non-macrocyclized natural products), however, other post-translational modifications can be found, such as acetylation, methylation, and dehydration (50). Known LAP include streptolysin S (*Streptococcus pyogenes*), microcin B17 (*Escherichia coli*), plantazolicin (*Bacillus velezensis*), goadsporin (*Streptomyces* sp. TP- A0584) and clostridiolysin S (*Clostridium botulinum*), owing to the difficulty inherent to the structure elucidation of heterocycles many LAPs remain unidentified (55).

It remains difficult to predict the exact compound products from genome sequence data only. The functionality of the novel BGCs found in our *Paenibacillus* sp. PL91 strain remains elusive, and further studies will be necessary, including purification of the compound, High Performance Liquid Chromatography (HPLC) mass spectrometry (MS/MS) and Nuclear Magnetic Resonance (NMR) spectroscopy. These studies have already been initiated.

## Conclusions

Biological control represents an alternative safe method to counter the extensive use of chemical pesticides which has caused severe environmental and health problems.

In this study, we identified three novel potential biocontrol strains, out of 69 strains screened, that can antagonize phytopathogens and mammalian pathogens, showing potential to be used as biocontrol agents and promote plant growth as well. Further mining into the genomes of the potential biocontrol strains the *in silico* prediction revealed a great number of BGCs, including known and potential novel ones. We found that BGCs in *Bacillus* species frequently encode conserved known compounds. The *Paenibacillus* sp. PL91 strain showed the largest number of possible novel BCGs; in our study we found one intact NRPs and two bacteriocins. Continued efforts will be directed to identify these interesting secondary metabolites as well as their contribution to biocontrol.

## Methods

### Plant material, bacterial isolation and strains

Plant material used in this study came from a small farm in the town of Roden, the Netherlands. The isolation of bacteria strain from tomato and lettuce leaves have been reported in our previous research. The genomic sequences of the most outstanding strains (SPL51, STRP31, PL81, PL91) were deposited into GenBank under accession no. of JABBYI000000000, JABBYF000000000, JABBYH000000000, and JABBYG000000000 respectively (35).

## Genome Sequence And Phylogenetic Analysis

Strains were grown overnight in Luria Bertani (LB) broth at 30°C, 220 rpm. Bacterial cells were harvested at the exponential growth phase with an  $OD_{600} = 1.0$ , DNA was extracted with a GenElute Bacterial Genomic DNA kit (Sigma) according to the manufacturer's protocol. The purity of the extracted DNA was checked using the Nanodrop (Thermo Fisher Scientific) and 1% agarose gel. Prediction of protein-encoding regions and automatic functional annotation was performed using the Rapid Annotations using Subsystem Technology (RAST) server (56). Genome-scale comparison of the four bacterial strains and other relevant strains were conducted GTDB-Tk v1.5.0 (36). A multiple sequence alignment of the 118 identified markers from 22 genomes was generated. It was employed to build a maximum likelihood phylogenetic tree by using IQ-TREE multicore version 1.6.12 (57). Substitution models LG + F + R4 was selected by ModelFinder (58). iTOL web-based tool (37) was used for tree visualization, using the previous alignment.

### In vitro **antagonistic assays**

Each fungus was inoculated in PDA agar and incubated at 28°C for specific time according to the fungus growth rate. Briefly, an agar plug (5mm diameter) with fungal hyphae was inoculated into a PDA plate at the center. 5µl overnight cultures ( $OD_{600} 1.0$ ) of the collection isolates, were subsequently spotted 2cm away from the plug in 4 corners symmetrically. Plates were sealed with parafilm and incubated at 28°C for each fungus specific time, and inhibition was documented.

### Verticillium dahliae and Septoria lycopersici **antifungal test**

Plugs were cutted from whole plates with each different fungus, and then vortexed in 10 ml of sterile water in order to collect the spores, mixed with 15 ml PDA agar, and the solution poured into plates. After drying the *SPL51*, *STRP31*, *PL91* overnight culture with  $OD_{600} 1.0$ , 5 µl was spotted at the center, followed by two to eight days incubation at 28°C, and inhibition was registered.

To test antibacterial activity, overnight cultures of bacterial pathogens were grown until  $OD_{600} = 1.0$ ; then mixed with LB agar media, at a final concentration of  $1 \times 10^6$  cells/ml. The mixed media was poured into petri dishes to get pathogen- fusion agar plates. Subsequently 7 µl of ( $1 \times 10^8$  cells /mL) culture of each isolated strain was inoculated at the center of the plate, followed by incubation between 24 and 48 hr. The presence or absence of a halo was monitored.

## Plant growth promoting assay by volatile compounds (VOCs) in vivo

Tomato seeds were sterilized by submerging them in 2% sodium hypochlorite for 15 min and then washed with sterile water 5 times to remove completely sodium hypochlorite. Seeds were germinated on large Petri dishes with Murashigie Skoog medium (Duchefa Biochemie) and incubated in a climate chamber with the next conditions: Light-time (16 hours) at 24+/- 2 Celsius degrees, 4000–6000 lux (or photon flux density of 200  $\mu\text{mol m}^{-2} \text{S}^{-1}$ ); Dark time (8 hours) at 21+/-2 Celsius degrees; humidity 75+/-5%. Once seeds germinated after 3 or 6 days (lettuce and tomato respectively), seedlings were transferred to a new MS Petri dish and 10  $\mu\text{l}$  of overnight cultures with  $\text{OD}_{600}$  1.0 were inoculated onto a separate small petridish inside the MS plate creating a separate compartment in order to evaluate plant growth promotion by volatile compounds. Plates were sealed with parafilm and after co-culture of 7 and 10 days (lettuce and tomato respectively) in the same conditions mentioned before, fresh weight and dry weight of the plants was measured as well as total length. All methods were performed in accordance with the relevant guidelines/regulations/legislation

## Abbreviations

### BCAs

Biocontrol agents; **BGCs**:Biosynthetic gene clusters; **NRPs**:Nonribosomal peptides; **NRPSs**:Nonribosomal peptide synthetases; **PKs**:Polyketides; **PKSs**:Polyketide synthetases; **RiPPs**:Ribosomally produced and posttranslationally modified peptides; **VOCs**:Volatile organic compounds; **A**:Adenylation; **ACP**:Acyl-carrier protein; **AT**:Acyltransferase; **Atd**:Trans-acyltransferase docking; **C**:Condensation; **CAL**:Co-enzyme A ligase domain; **DH**:Dehydratase; **E**:Epimerization; **ISR**:Induced systemic resistance; **KR**:Keto-reductase; **KS**:Keto-synthase; **PCP**:Peptidyl carrier protein; **PGPR**:Plant growth-promoting rhizobacteria; **TE**:Thioesterase.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Data availability

The whole genome data are available at DDBJ/EMBL/GenBank under the bio project accession PRJNA623850 ( <https://www.ncbi.nlm.nih.gov/bioproject/623850> ).

### Competing interests

The authors declare that they have no conflict interests.

## Funding

Claudia Y. Muñoz was financially supported by the Mexican National Council for Science and Technology (CONACYT) scholarship (471987).

Lu Zhou was financially supported by the China Scholarship Council (201606910037).

Yunhai Yi was financially supported by the China Scholarship Council (201904910477).

## Authors' contributions

Conceived and designed the project: CM, LZ and OK. Conceived and designed the experiments: CM, LZ and OK. Execution of experimental work and interpretation of the data: CM, LZ, YY. Manuscript writing: CM. Corrected the manuscript: CM, LZ, OK. All authors read and approved the final manuscript.

## Authors' details

Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, the Netherlands

## References

1. Kimura S, Sinha N. Tomato (*Solanum lycopersicum*): A Model Fruit-Bearing Crop. *CSH Protoc.* 2008;2008:pdb.emo105.
2. Campbell JK, Canene-Adams K, Lindshield BL, Boileau TW, Clinton SK, Erdman JW Jr. Tomato phytochemicals and prostate cancer risk. *J Nutr.* 2004;134(12).
3. Nicolle C., Cardinault N., Gueux E, Jaffrelo L, Rock E, Mazur A, Amouroux P, Rémésy C. Health effect of vegetable-based diet: lettuce consumption improves cholesterol metabolism and antioxidant status in the rat. *Clin Nutr.* 2004; 23(4):605–14.
4. Nazarov, P. A., Baleev, D. N., Ivanova, M. I., Sokolova, L. M., & Karakozova, M. V. Infectious Plant Diseases: Etiology, Current Status, Problems and Prospects in Plant Protection. *Acta naturae*, 2020;12(3), 46–59.
5. Smith, P., & Gregory, P. Climate change and sustainable food production. *Proceedings of the Nutrition Societ.* 2013;72(1), 21–28.
6. Chen X., Pizzatti C., Bonaldi M., Saracchi M., Erlacher A., Kunova A., Berg G., & Cortesi, P. Biological Control of Lettuce Drop and Host Plant Colonization by Rhizospheric and Endophytic Streptomycetes. *Frontiers in microbiology.* 2016;7, 714.
7. Elhariry H. M. Attachment strength and biofilm forming ability of *Bacillus cereus* on green-leafy vegetables: cabbage and lettuce. *Food microbiology.* 2011; 28(7), 1266–1274.
8. Ongena M, Jacques P. *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol.* 2008;16(3):115–25.

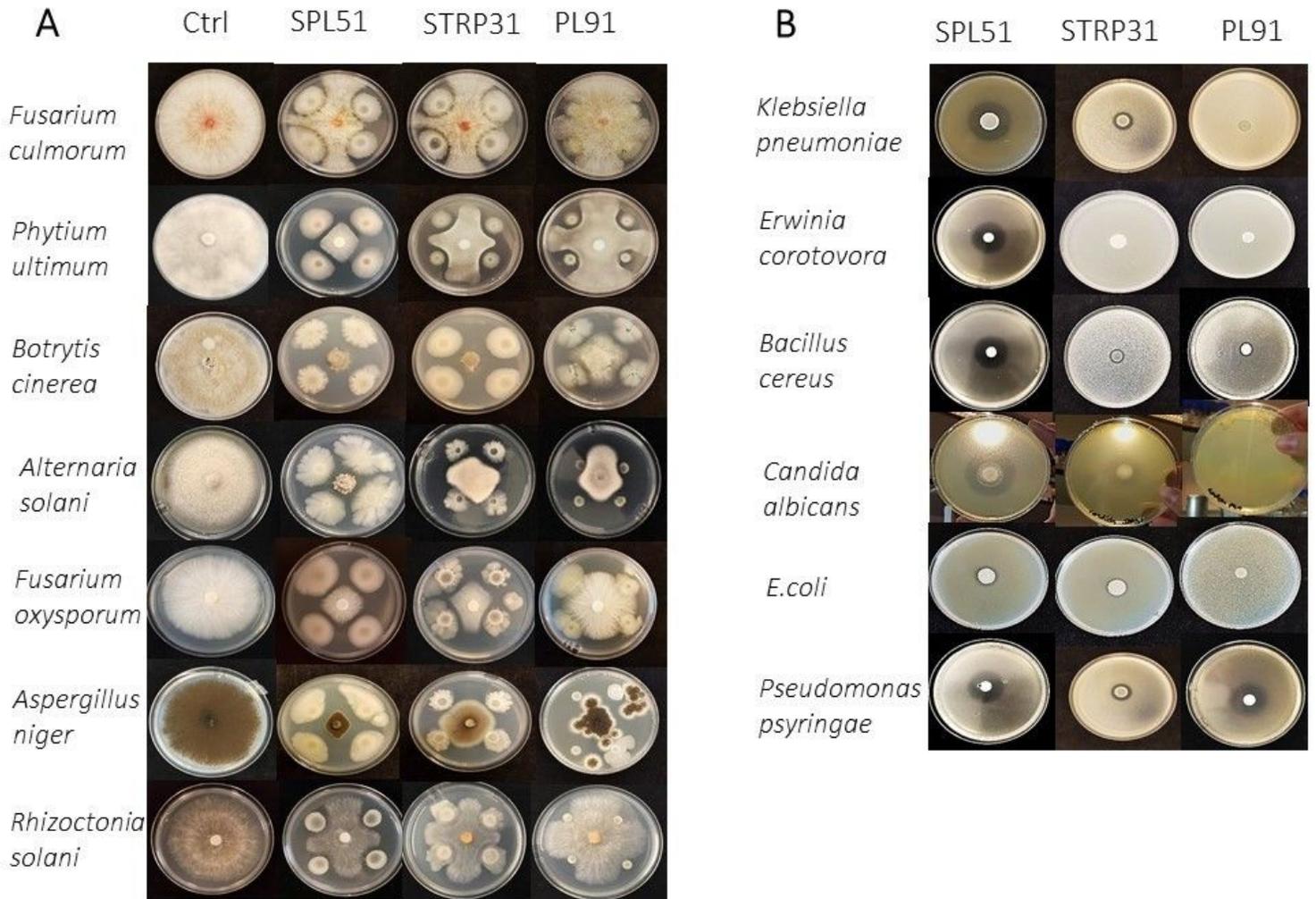
9. Heydari A, Pessarakli M., A Review on Biological Control of Fungal Plant Pathogens Using Microbial Antagonists. *Journal of Biological Sciences* Journal of Biological Sciences Volume. 2010;10 (4): 273–290.
10. Montesinos E. Plant-associated microorganisms: a view from the scope of microbiology. *International Microbiology*. 2003; 6(4), 221–223.
11. Vorholt, J. Microbial life in the phyllosphere. *Nat Rev Microbiol*. 2012;**10**, 828–840
12. Djordje Firilvica Dimkić Tanja Berić Jelena Lozo Slaviša Stanković Biological control of plant pathogens by *Bacillus* species. *Journal of Biotechnology*. 2018;285, 44.
13. Zhao, X., & Kuipers, O. P. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. *BMC genomics*. 2016;**17**(1): 882.
14. Weissman, K. The structural biology of biosynthetic megaenzymes. *Nat Chem Biol*. 2015; **11**, 660–670.
15. Miyanaga A, Kudo F, Eguchi T. Protein-protein interactions in polyketide synthase-nonribosomal peptide synthetase hybrid assembly lines. *Nat. Prod. Rep.*, 2018;**35**, 1185.
16. M. Montalbán-López et. al., New developments in RiPP discovery, enzymology and engineering *Nat. Prod. Rep.*, 2021, **38**, 130–239
17. Kaspar F, Neubauer P, Gimpel M. Bioactive Secondary Metabolites from *Bacillus subtilis*: A Comprehensive Review. *J Nat Prod*. 2019;**82**(7):2038–2053.
18. Zhao X, Kuipers OP. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. *BMC Genomics*. 2016;**17**(1):882.
19. Rabbee, M. F., Ali, M. S., Choi, J., Hwang, B. S., Jeong, S. C., & Baek, K. H. *Bacillus velezensis*: A Valuable Member of Bioactive Molecules within Plant Microbiomes. *Molecules* (Basel, Switzerland). 2019;**24**(6), 1046.
20. Niu, B., Vater, J., Rueckert, C. *et al.* Polymyxin P is the active principle in suppressing phytopathogenic *Erwinia spp.* by the biocontrol rhizobacterium *Paenibacillus polymyxa* M-1. *BMC Microbiol*. 2013; **13**, 137.
21. Dunlap, C. A., Schisler, D. A., Bowman, M. J., & Rooney, A. P. Genomic analysis of *Bacillus subtilis* OH 131.1 and co-culturing with *Cryptococcus flavescens* for control of *Fusarium* head blight. *Plant Gene*. 2015; **2**, 1–9.
22. Avalos, M., van Wezel, G. P., Raaijmakers, J. M., & Garbeva, P. Healthy scents: microbial volatiles as new frontier in antibiotic research? *Current Opinion in Microbiology*. 2018; **45**, 84–91.
23. Zou C, Li Z, Yu D. *Bacillus megaterium* strain XTBG34 promotes plant growth by producing 2-pentylfuran. *J Microbiol*. 2010;**48**(4):460–6.
24. Ajayi-Oyetunde, O.O. and Bradley, C.A. *Rhizoctonia solani*: taxonomy, population biology and management of rhizoctonia seedling disease of soybean. *Plant Pathol*. 2018;**67**: 3–17.
25. Rupp, S., Weber, R. W., Rieger, D., Detzel, P., & Hahn, M. Spread of *Botrytis cinerea* Strains with Multiple Fungicide Resistance in German Horticulture. *Frontiers in microbiology*. 2017;**7**, 2075.

26. Salman, M., Abuamsha, R. Potential for integrated biological and chemical control of damping-off disease caused by *Pythium ultimum* in tomato. *BioControl*. 2012; **57**, 711–718.
27. Xin, X. F., & He, S. Y. *Pseudomonas syringae* pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. *Annual review of phytopathology*. 2013; *51*, 473–498.
28. Alippi, A. M., Bó, E. D., Ronco, L. B., Casanova, P. E., & Aguilar, O. M. Tomato as a New Host of *Erwinia carotovora* subsp. *carotovora* in Argentina. *Plant disease*. 1997; *81*(2), 230.
29. Canik Orel, D. *Erwinia persicina* as the new causal agent of lettuce soft rot. *Eur J Plant Pathol*. 2020; **158**, 223–235.
30. SRIHARMINDA PAHM HARTANTYO, MAN LING CHAU, TSE HSIEN KOH, MIN YAP, TSENG YI, DELPHINE YAN HONG CAO, RAMONA ALIKIITEAGA GUTIÉRREZ, LEE CHING NG; Foodborne *Klebsiella pneumoniae*: Virulence Potential, Antibiotic Resistance, and Risks to Food Safety. *J Food Prot*. 2020; *83* (7): 1096–1103.
31. Tewari, A., & Abdullah, S. *Bacillus cereus* food poisoning: international and Indian perspective. *Journal of food science and technology*. 2015; *52*(5), 2500–2511.
32. Dublan MdIA, Ortiz-Marquez JCF, Lett L, Curatti L. Plant-Adapted *Escherichia coli* Show Increased Lettuce Colonizing Ability, Resistance to Oxidative Stress and Chemotactic Response. *PLoS ONE*. 2014; *9*(10): e110416.
33. Rajkowska, K., & Kunicka-Styczyńska, A. Typing and virulence factors of food-borne *Candida* spp. isolates. *International Journal of Food Microbiology*. 2018; *279*, 57–63.
34. Muñoz CY, de Jong A, Kuipers OP. Draft genome sequences of a *Bacillus subtilis* strain, a *Bacillus velezensis* strain, a *Paenibacillus* strain, and an *Acinetobacter baumannii* strain, all isolated from the phyllosphere of *Lactuca sativa* or *Solanum lycopersicum*. *Microbiol Resour Announc*. 2021; *10*:e01092-20.
35. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. 2019;*36*(6):1925–7.
36. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 2016;*44*(W1):W242-5.
37. Entfellner, E., Frei, M., Christiansen, G., Deng, L., Blom, J., & Kurmayer, R. Evolution of Anabaenopeptin Peptide Structural Variability in the Cyanobacterium *Planktothrix*. *Frontiers in microbiology*. 2017; *8*, 219.
38. Özcengiz G, Ögülür İ. Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *N Biotechnol*. 2015;*32*(6):612–9.
39. May, Jürgen J.; Wendrich, Thomas M.; Marahiel, Mohamed A. "The *dhb* Operon of *Bacillus subtilis* Encodes the Biosynthetic Template for the Catecholic Siderophore 2,3-Dihydroxybenzoate-Glycine-Threonine Trimeric Ester Bacillibactin". *Journal of Biological Chemistry*. 2001;*276* (10): 7209–7217.

40. Wu, L., Wu, H., Chen, L., Yu, X., Borriss, R., & Gao, X. Difficidin and bacilysin from *Bacillus amyloliquefaciens* FZB42 have antibacterial activity against *Xanthomonas oryzae* rice pathogens. *Scientific reports*. 2015;5, 12975.
41. Sur, S., Romo, T. D., & Grossfield, A. Selectivity and Mechanism of Fengycin, an Antimicrobial Lipopeptide, from Molecular Dynamics. *The journal of physical chemistry*. 2018; B, 122(8), 2219–2226.
42. Butcher, R. A., Schroeder, F. C., Fischbach, M. A., Straight, P. D., Kolter, R., Walsh, C. T., & Clardy, J. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(5), 1506–1509.
43. Yuan, J., Zhao, M., Li, R., Huang, Q., Rensing, C., Raza, W., & Shen, Q. Antibacterial Compounds-Macrolactin Alters the Soil Bacterial Community and Abundance of the Gene Encoding PKS. *Frontiers in microbiology*. 2016; 7, 1904.
44. Scholz, R., Molohon, K. J., Nachtigall, J., Vater, J., Markley, A. L., Süssmuth, R. D., Mitchell, D. A., & Borriss, R. Plantazolicin, a novel microcin B17/streptolysin S-like natural product from *Bacillus amyloliquefaciens* FZB42. *Journal of bacteriology*. 2011; 193(1), 215–224.
45. Zhi, Y., Wu, Q. & Xu, Y. Genome and transcriptome analysis of surfactin biosynthesis in *Bacillus amyloliquefaciens* MT45. *Sci Rep*. 2017; 7, 40976.
46. Phelan, R. W., Barret, M., Cotter, P. D., O'Connor, P. M., Chen, R., Morrissey, J. P., Dobson, A. D., O'Gara, F., & Barbosa, T. M. Subtilomycin: a new lantibiotic from *Bacillus subtilis* strain MMA7 isolated from the marine sponge *Haliclona simulans*. *Marine drugs*. 2013; 11(6), 1878–1898.
47. Babasaki, K., Takao, T., Shimonishi, Y., & Kurahashi, K. Subtilosin A, a new antibiotic peptide produced by *Bacillus subtilis* 168: isolation, structural analysis, and biogenesis. *Journal of biochemistry*. 1985;98(3), 585–603.
48. Zhu, S., Hegemann, J. D., Fage, C. D., Zimmermann, M., Xie, X., Linne, U., & Marahiel, M. A. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. *The Journal of biological chemistry*. 2016; 291(26), 13662–13678.
49. Yang, X., & Yousef, A. E. Antimicrobial peptides produced by *Brevibacillus* spp.: structure, classification and bioactivity: a mini review. *World Journal of Microbiology and Biotechnology*. 2018;34(4).
50. Arnison, Paul G et al. "Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature." *Natural product reports*. 2013; 30,1.
51. Zhu S, Hegemann JD, Fage CD, Zimmermann M, Xie X, Linne U, Marahiel MA. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. *J Biol Chem*. 2016; 24;291(26):13662-78.
52. Maksimov MO, Link AJ. "Prospecting genomes for lasso peptides". *Journal of Industrial Microbiology & Biotechnology*. 2014; 41 (2): 333–44.
53. Cheng C, Hua ZC. Lasso Peptides: Heterologous Production and Potential Medical Application. *Front Bioeng Biotechnol*. 2020 Sep 28;8:571165.

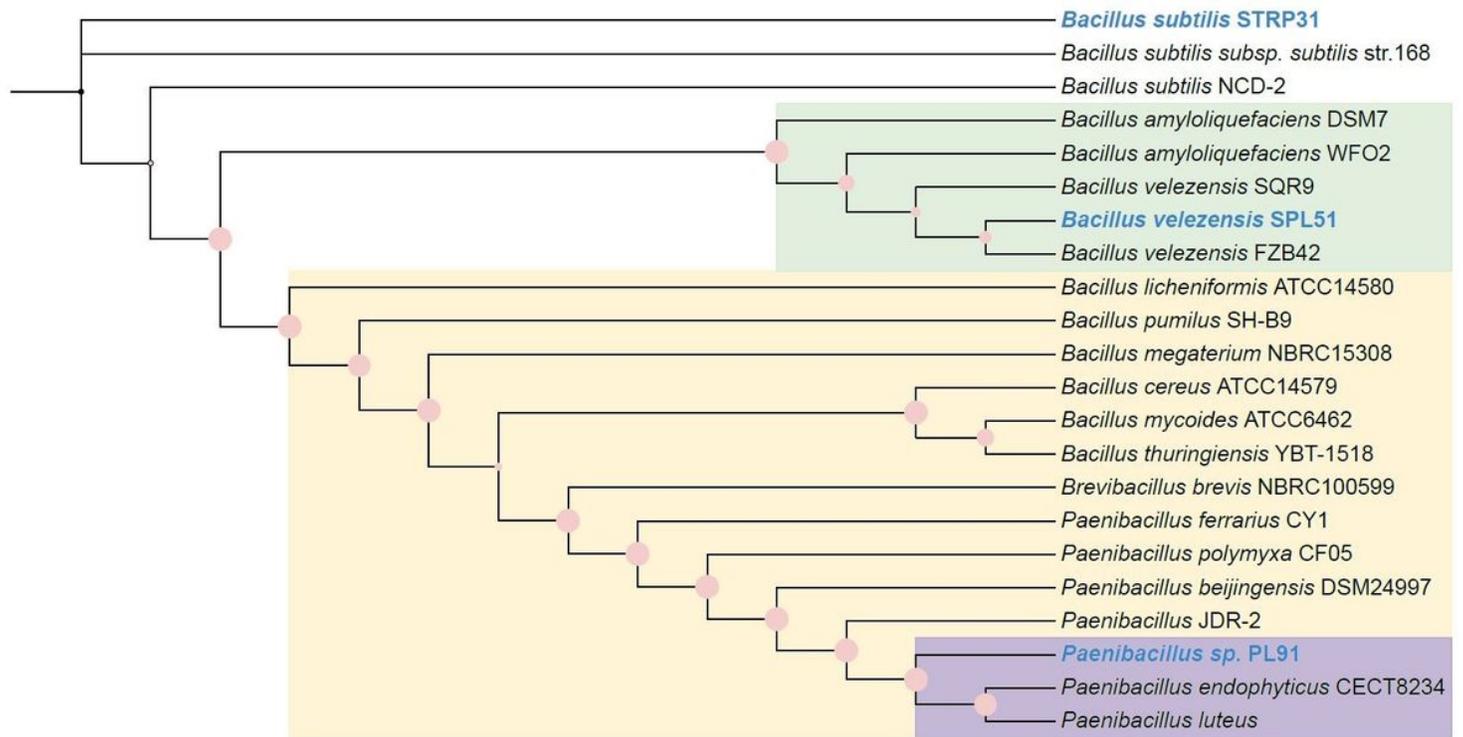
54. Letzel, A. C., Pidot, S. J., & Hertweck, C. Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. *BMC genomics*. 2014; 15(1), 983.
55. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics*. 2008;9:75.
56. L.-T. Nguyen, H.A. Schmidt, A. von Haeseler, B.Q. Minh IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 2015;32:268–274.
57. S. Kalyaanamoorthy, B.Q. Minh, T.K.F. Wong, A. von Haeseler, and L.S. Jermin ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates, *Nature Methods*. 2017; 14:587–589.
58. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 2019;47(W1):W81–7
59. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* 2018;46(W1): W278–81.
60. Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, Crescenzi P, Fani R, Fondi MJB: MeDuSa: a multi-draft based scaffold. 2015;31:2443–2451
61. 61. Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, Crescenzi P, Fani R, Fondi MJB: MeDuSa: a multi-draft based scaffold. 2015;31:2443–2451

## Figures



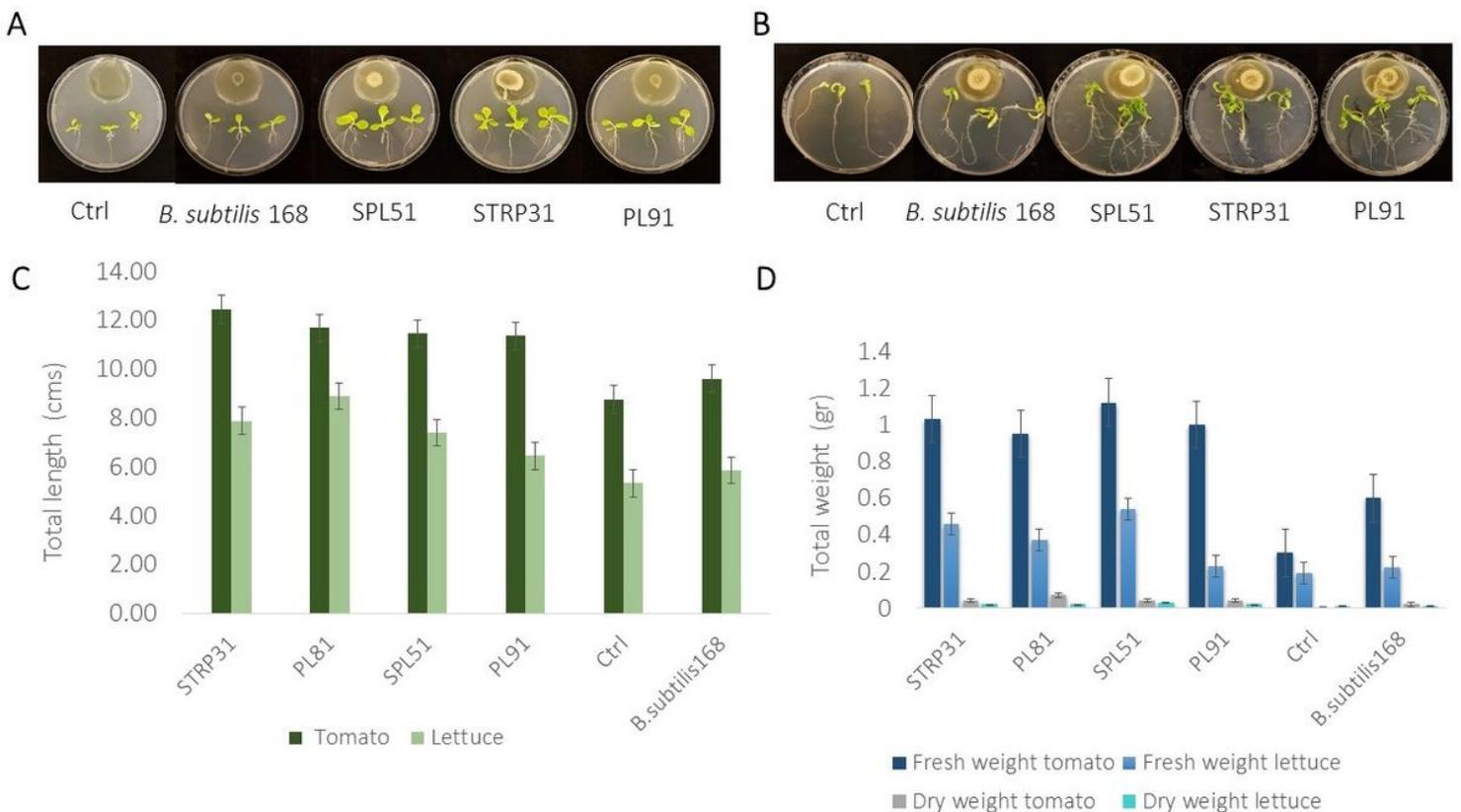
**Figure 1**

A) Fungal inhibition assay with selected strains *B. velezensis* SPL51, *B. subtilis* STRP31 and *Paenibacillus* sp. PL91. B) Bacterial inhibition assay with selected strains *B. velezensis* SPL51, *B. subtilis* STRP31 and *Paenibacillus* sp. PL91.



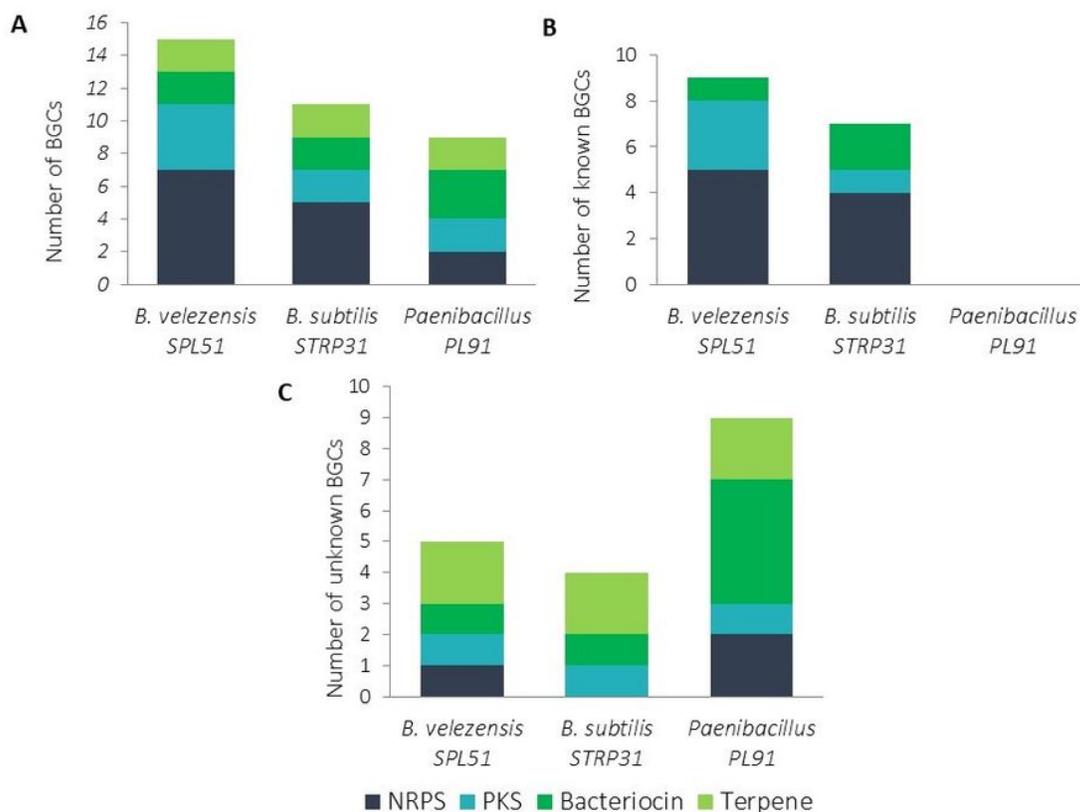
**Figure 2**

Phylogenetic analysis of the selected isolated strains with antimicrobial activity in this study. The sequences of reference strains were retrieved from the NCBI database (Table S3 for accession number). The reference strains are in black, and our sequenced strains are highlighted with blue. The pink dots on branches represent bootstrapping values, ranging from 84 to 100.



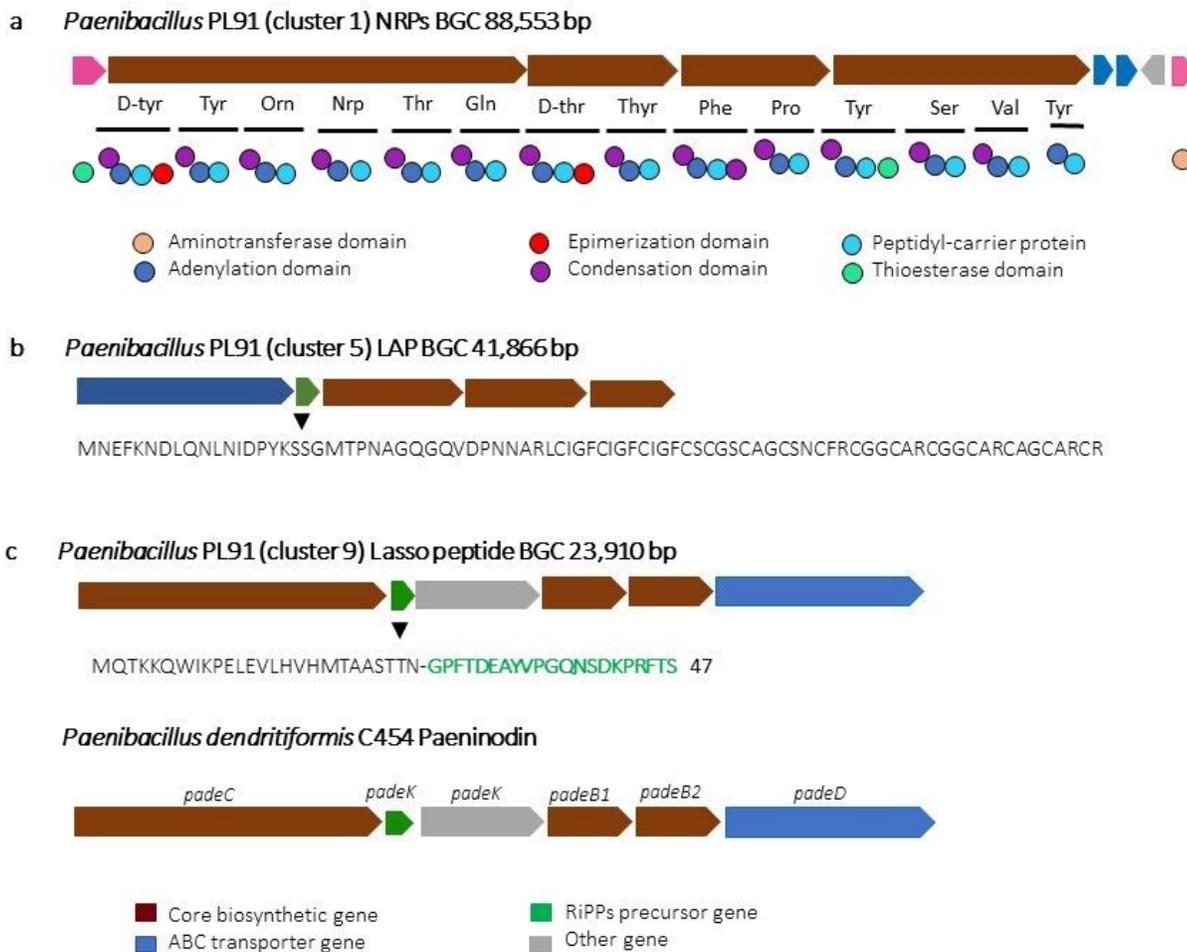
### Figure 3

Plant growth promotion by volatiles. A) Phenotypic changes in lettuce seedlings exposed to volatiles emitted by SPL51, STP31, and PL91 strains inoculated on agar LB medium (10  $\mu$ l OD=1.0), the control LB agar only, and *B. subtilis* 168 used as second control. Pictures were taken 7 days after exposure. B) Phenotypic changes in tomato seedlings exposed to volatiles emitted by SPL51, STP31, PL81, and PL91 strains inoculated on agar LB medium (10  $\mu$ l OD=1.0) the control LB agar only, and *B. subtilis* 168 used as second control. Pictures were taken 10 days after exposure. C) Total length tomato and lettuce after treatment (mean  $\pm$  standard error [SE], n = 4). D) Biomass, fresh weight and dry weight (mean  $\pm$  standard error [SE], n = 4).



### Figure 4

Numbers of BGCs harbored by the strains. (a) total number of BGCs in the strains. (b) number of reported BGCs in the genomes of strains. (c) number of unknown BGCs found in the strains. BGCs that have different numbers of genes or show equal or less than 80% protein identity to the reported ones were regarded as novel.



**Figure 5**

Potential novel BCGs found in the genomes of the selected strains. a) an NRPS present in *Paenibacillus* PL91. b) a LAP present in *Paenibacillus* PL91, c) A lasso peptide harbored by *Paenibacillus* PL91 compared to a Paeninodin lasso peptide produced by *Paenibacillus dendritiformis* C454. The arrows represent the coding regions of the genes identified in the genomes.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryfile.pdf](#)