Antimicrobial Activity Screening of Rhizosphere Soil Bacteria from Tomato and Genome-Based Analysis of their Antimicrobial Biosynthetic Potential

CURRENT STATUS: Under Review

Lu Zhou, Chunxu Song, Zhibo Li, Oscar P. Kuipers

Lu Zhou
Rijksuniversiteit Groningen

Chunxu Song
China Agricultural University

Zhibo Li
Rijksuniversiteit Groningen

Oscar P. Kuipers
Rijksuniversiteit Groningen

ORCiD: https://orcid.org/0000-0001-5596-7735

Prescreen

10.21203/rs.3.rs-26812/v1

Subject Areas

Epigenetics & Genomics
Keywords

*Bacillus, Plant Growth Promoting Rhizobacteria (PGPR), Genome mining, Tomato, Rhizosphere, biosynthetic gene clusters, non-ribosomal synthesized peptides, polyketides, bacteriocins*
Abstract

Background: Tomato plant growth is frequently hampered by a high susceptibility to pests and diseases. Traditional chemical control causes a serious impact on both the environment and human health. Therefore, seeking environment-friendly and cost-effective green methods in agricultural production becomes crucial nowadays. Plant Growth Promoting Rhizobacteria (PGPR) can promote plant growth through biological activity. Their use is considered to be a promising sustainable approach for crop growth. Moreover, a vast number of biosynthetic gene clusters (BGCs) for secondary metabolite production are being revealed in PGPR, which helps to find potential anti-microbial activities for tomato disease control.

Results: We isolated 351 bacterial strains (181 of which are Bacillus sp.) from healthy tomato, rhizosphere soil, and tomato tissues. In vitro antagonistic assays revealed that 34 Bacillus strains have antimicrobial activity against Erwinia carotovora, Pseudomonas syringae; Rhizoctonia solani; Botrytis cinerea; Verticillium dahliae and Phytophthora infestans. The genomes of 10 Bacillus and Paenibacillus strains with good antagonistic activity were sequenced. Via genome mining approaches, we identified 120 BGCs encoding NRPs, PKs-NRPs, PKs, terpenes and bacteriocins, including known compounds such as fengycin, surfactin, bacillibactin, subtilin, etc. In addition, several novel BGCs were identified. We discovered that the NRPs and PKs-NRPs BGCs in Bacillus species are encoding highly conserved known compounds as well as various novel variants.

Conclusions: This study highlights the great number of varieties of BGCs in Bacillus strains. These findings pave the road for future usage of Bacillus strains as biocontrol agents for tomato disease control and are a resource arsenal for novel antimicrobial discovery.

Background

Tomato (Solanum lycopersicum) is the second most important vegetable crop worldwide after potato, based on the sizes of their growth areas [1]. However, tomato crops face serious threats of disease, partially due to the use of cultivars susceptible to diseases that are causing substantial production losses [2]. The overuse of chemical pesticides has contaminated soils and has caused harmful effects on human beings [3]. Accordingly, putting biocontrol agents isolated from nature into the soil is environmentally friendly and useful for tomato crop disease control. One way to improve plant growth is by using plant growth-promoting rhizobacteria (PGPR), since PGPR have the ability to colonize the roots and express their plant growth promotion activities in the rhizosphere [4].

The rhizosphere, a narrow zone of soil that surrounds and is influenced by plant roots, gives home to an overwhelming variety of organisms, in particular microorganisms such as bacteria, fungi, oomycetes, archaea, protozoa and algae [5, 6]. This complex microbial community has profound effects on plant growth since it facilitates nutrient absorption and provides health protection to plants [7]. Among all the microorganisms, PGPR has been largely described for their biocontrol capabilities. They can promote plant growth either indirectly by suppression of diseases with secreted antimicrobials or directly by the improvement of physiological metabolic processes such as N2 fixation, phosphate solubilization and IAA production [8].

Among PGPR, the group of Gram-positive Bacillus strains has been studied less intensively, compared to widely used Gram-negative bacteria, like Pseudomonas strains [9]. One of the most efficient Gram-positive bacteria that promote plant growth belongs to the genus Bacillus. Bacillus subtilis is used in agriculture to protect plants from several plant pathogens since it can either indirectly protect plants by inducing systemic resistance (ISR) against a broad range of pathogens or directly excrete antimicrobials [10–13]. Besides, Bacillus species can produce hard, resistant endospores to allow them to resist adverse environmental conditions and permit easy formulation and storage of the commercial products [14].

The Bacillus species offer a plethora of antagonistic compounds displaying a broad range of biological functions, which have good potential to be used as biocontrol agents for tomato disease control [15]. All the bioactive
secondary metabolites are encoded by biosynthetic gene clusters (BGCs). Based on their products, BGCs are classified as ribosomally synthesized peptides (linear bacteriocin BGCs and ribosomally-produced an postranslationally modified peptides [RiPPs]), non-ribosomally synthesized peptides synthetases (NRPSs) BGCs and polyketide synthases (PKSs) BGCs [16]. Here, we set out to find novel BGCs in Bacillus strains which encode potentially active compounds to inhibit plant-pathogens. Based on genome mining, 10 selected (out of 351) promising Bacillus strains newly isolated from rhizosphere soils of healthy tomato plants and tissues were characterized with respect to anti-pathogen activities. Subsequently, several novel BGCs were discovered, which have potential functions in tomato pathogens antagonism.

Results

Isolation of Bacteria and In Vitro Antagonistic Assay

A total of 351 bacterial strains were isolated from healthy tomato rhizosphere soil and tomato plant tissue collected in either the Netherlands or Spain, 181 strains of which were considered as Bacillus-like strains based on the morphology of colonies and 80 °C treatment given before (spores surviving temperature). In order to identify potential PGPR strains, all the Bacillus-like strains were preliminarily screened in vitro antagonistic activity against six major tomato plant pathogens, i.e. Erwinia carotovora, Pseudomonas syringae, Rhizoctonia solani, Botrytis cinerea, Verticillium dahliae, and Phytophthora infestans. The results revealed that 34 Bacillus-like strains could inhibit different bacterial, fungal and oomycetal plant pathogens growth on plates (Fig. 1). A phylogenetic analysis based on 16S rRNA revealed that these strains belong to the species of Bacillus subtilis (18 strains), Bacillus velezensis (3 strains), Bacillus endophyticus (3 strains), Bacillus megaterium (4 strains), Bacillus aryabhattai (2 strains), Bacillus cereus (2 strains), Bacillus firmus (1 strain) and Paenibacillus xylanexedens (1 strain). In total, 14 strains showed inhibition on bacterial, fungal, and oomycetal pathogens. Among them, 12 strains belong to the Bacillus subtilis group, while the remaining 2 strains belong to the Bacillus velezensis group. The others only showed antagonistic activity against fungi and oomycetes. A total of 10 strains (BH5, BH6, DH12, EH2, EH5, EH11, FH5, FH17, TH16 and edo6), showing high antagonistic activity, were genome sequenced for further research.

Genome Sequencing and Biosynthesis Gene Cluster Mining

The genomes of 10 isolated strains were sequenced, assembled and annotated as described in a previous study [17]. Based on whole genome phylogenetic analyses, the 10 Bacillus strains were clustered into four clades as presented in Fig. 2. All of them were tightly clustered together with reported PGPR strains from the Bacillus class, such as B. subtilis Bsn5, B. velezensis FZB42 and P. polymyxa E681. This suggests that they probably can promote plant growth as well, which needs to be further investigated. Strains BH5, BH6, DH12, EH2, EH5, and EH11 fall into the B. subtilis group, FH17 and TH16 were identified as B. velezensis, while FH5 and edo6 belong to the B. endophyticus and P. xylanexedens, respectively. The ten strains were selected from the rhizosphere soil and plant tissues due to their activities against tomato phytopathogens, which indicated the presence of some important antimicrobial gene clusters. By using antiSMASH 5.0 [18] and BAGEL4 [19], a total of 120 BGCs were found, averaging 12 clusters per genome. All the BGCs were designated as those encoding NRPSs, PKSs, terpenes, hybrid NRPS/PKSs, bacteriocins, RiPPs and others (Table 1A). The BGCs encoding surfactin [20], fengycin [20], bacillibactin [21], subtilosin A [22], bacilaene [23], macrolactin [24], difficidin [25], and subtilin [26] were discovered in the genomes. Besides, some BGCs encoding unknown compounds, were also identified (Table 1B). Most of the unknown BGCs (76.47%) are PKSs BGCs, which cannot be assigned to any known compounds. 73.07% bacteriocins BGCs encodes potential novel peptides. 27.78% and 27.27% of NRPSs and Hybrids BGCs are still unknown. These findings provide a great opportunity of new bioactive compounds discovery.

A.
| Strains                          | Predicted BGCs | NRPS | PKS | Hybrid NRPS/PKS | Terpene | Bacteriocin | Other |
|---------------------------------|----------------|------|-----|----------------|---------|-------------|-------|
| Bacillus subtilis BH5           | 12             | 4    | 1   | 1              | 2       | 3           | 1     |
| Bacillus subtilis BH6           | 12             | 4    | 1   | 1              | 2       | 3           | 1     |
| Bacillus subtilis DH12          | 12             | 4    | 1   | 1              | 2       | 3           | 1     |
| Bacillus subtilis EH2           | 10             | 3    | 1   | 1              | 2       | 2           | 1     |
| Bacillus subtilis EH5           | 11             | 3    | 1   | 1              | 2       | 3           | 1     |
| Bacillus subtilis EH11          | 12             | 4    | 1   | 1              | 2       | 3           | 1     |
| Bacillus endophyticus FH5       | 10             | 2    | 1   | 1              | 2       | 3           | 1     |
| Bacillus velezensis FH17        | 15             | 5    | 4   | 1              | 2       | 1           | 2     |
| Bacillus velezensis TH16        | 12             | 4    | 4   | 1              | 1       | 1           | 1     |
| Paenibacillus xylanexedens edo6 | 14             | 3    | 2   | 2              | 1       | 4           | 1     |

B.
Table 1. Distribution of BGC totals in 10 isolated strains (A) and percentages of BGCs encoding unknown compounds identified in genome sequence (B).

| BGC Types | Total BGCs | % Unknown | Known compounds |
|-----------|------------|-----------|-----------------|
| NRPSs     | 36         | 27.78     | surfactin (8 BGCs), fengycin (8 BGCs), bacillibactin (10 BGCs) |
| PKSs      | 17         | 76.47     | macrolactin (2 BGCs), difficidin (2 BGCs) |
| Hybrids   | 11         | 27.27     | bacillaene (8 BGCs) |
| Bacteriocin | 26         | 73.07     | subtilin (2 BGCs), subtilosin A (6 BGCs) |

Novel NRPs and PKs BGCs identified from the 10 strains

The majority of BGCs could be assigned to known compounds, whereas 5 clusters represented probably novel NRPs and NRPs/PKs hybrid BGCs for which no or low similarity BGCs could be identified in the MIBiG database (Fig. 3).

Two novel gene clusters were identified from *B. endophyticus* FH5. One NRPs (Fig. 3a) BGC consists of three genes and has a total size of 25 kb. Three genes are encoding 24 domains, which includes 7 condensation (C) domains, 7 adenylation (A) domain, 7 thiolation (T) domain, 2 epimerization (E) domain and 1 thioesterase (TE) domain. All the domains are essential components in this BGC and catalyze primary formation of a lipopeptide product. This BGC is showing no similarity to any known BGCs reported. The other one (Fig. 3b) is a Type I PKs-NRPs hybrid BGC with a size of approximately 30 kb. The PKs module consists of a ketosynthase (KS) domain, a acyltransferase (AT) domain, an acyl carrier protein (ACP) domain and a terminal reductase (TD) domain. It likely incorporates the polyketide moiety of malonyl-CoA, while the NRPs modules incorporate six amino acid residues. Based on antiSMASH analysis, only 28% genes show similarity to the known paenilamicin BGC. Paelinamicin [28], synthesized by *pam* BGC from *Paenibacillus larvae* DSM25430, has antibacterial and antifungal activity. The *pam* cluster consists of five NRPs genes, two Type I PKs genes, and two Type I PKs-NRPs hybrid genes, and has a size of ~60 kb. In contrast, the Type I PKs-NRPs hybrid BGC identified in *B. endophyticus* FH5 consists of only three NRPS genes and one Type I PKS gene. All of them differ from the *pam* cluster of *Paenibacillus larvae* DSM25430.

In the genome of *P. xylanexedens* edo6, two novel trans-AT PKs-NRPs hybrid gene clusters (cluster 13 and cluster 12) were discovered, which have the sizes of almost 35 kb and 28 kb, respectively (Fig. 3c and 3d). The order and domain of the genes of both hybrid clusters differ from each other. Specifically, Cluster 13 has an additional dehydratase domain variant (DHT) playing an important role during polyketide biosynthesis through the dehydration of the nascent polyketide intermediate to provide olefins [29], which cannot be found in cluster 12. In addition to the differences observed at the domain level of core biosynthetic genes, regulator and transporter genes are also different. Moreover, only 33% and 21% of the genes of cluster 13 and cluster 12 exhibit similarity to known pellasoren and xenocoumacin BGCs respectively. Pellasoren [30] was isolated from myxobacterium, which has shown to possess potential anti-cancer activity. The known pellasoren BGC, is a Type I PKs-NRPs hybrid cluster identified from *Sorangium cellulosum* So ce38 and consists of six genes of Type I PKs and one single gene of NRPs as compared to the trans-AT PKs-NRPs hybrid gene (cluster 13) of *P. xylanexedens* edo6, which in turn consists of four trans-AT PKs genes and one trans-AT PKs-NRPs hybrid gene. Xenocoumacin [31] is the main antibacterial and anti-fungal compound produced by *Xenorhabdus nematophila*. The known xenocoumacin BGC, also being a Type I PKs-NRPs hybrid cluster, which was identified from *Xenorhabdus nematophila* ATCC 19061,
consists of four genes of Type I PKs and two genes of NRPs whereas cluster 12 from P. xylanexedens edo6 consists of one single trans-AT domain gene, one gene of trans-AT PKs and one gene of trans-AT PKs-NRPs hybrid.

One novel NRPs BGC was discovered both in B. velezensis FH17 and TH16 (Fig. 3e). This BGC contains seven genes with a size of approximately 33 kb. Whereas seven modules are only encoded by two core biosynthetic genes, seven amino acids are incorporated into the final product. This BGC shows no similarity to any known clusters. Furthermore, a single heterocyclization (Cy) domain in the first module is found. This domain first catalyzes amide bond formation, and then the intramolecular cyclodehydration between the side chain of the first amino acid (Cys) and the backbone carbonyl carbon takes place to form a thiazoline ring [32]. This ring is important for the structure and function of this lipopeptide product. So far, many well-known drugs for anti-microbial and anti-cancer activity exhibit thiazoline rings [33], such as Sulfathiazole (anti-microbial drug), Ritonavir (anti-retroviral drug), Tiazofurin (anti-neoplastic drug) and Abafungin (anti-fungal drug) [34]. These findings suggest the potential anti-microbial activity of the compounds produced by this BGC in B. velezensis FH17 and TH16.

**Novel Ribosomally synthesized and Post-translationally modified Peptides (RiPPs) identified in the 10 strains**

A total of nine novel bacteriocin BGCs were identified from the 10 strains (Fig. 4). All of them are belong to RiPPs (less than 10 kDa). These peptides are ribosomally synthesized, and undergo posttranslational modifications (PTMs), resulting in different structures and properties, mainly showing anti-bacterial activity against closely related producer strains [35].

Two novel gene clusters were identified as class I lanthipeptide BGCs. One lanthipeptide BGC was identified from both B. subtilis DH12 and EH11 with a size of ~6 kb (Fig. 4a). This BGC consists of four genes. The precursor peptide contains 59 amino acids, which shows no similarity to any known bacteriocins. Another one lanthipeptide BGC (Fig. 4b) was identified from P. xylanexedens edo6 with a size of ~9 kb. This BGC contains seven genes. The precursor peptide encoded by the core biosynthetic gene contains 59 amino acids, which also shows no similarity to any known bacteriocins.

Three novel BGCs were identified as class II lanthipeptide BGCs. All of them belong to two-component lanthipeptides consisting of two peptides. The individual peptides of two-component lanthipeptides only have little or no antimicrobial activity, but the two peptides act in synergy to exhibit significantly higher activity in equimolar concentrations [36]. Both B. subtilis BH5 and BH6 harbor the same two-component lanthipeptide BGC (Fig. 4c). It consists of six genes with a size of ~9 kb. This BGC has 70% of genes showing similarity to staphylococcin C55 α/β BGC [37]. The presursors of two core biosynthetic genes (α and β) of this BGC identified contain 65 and 67 amino acids respectively. The C terminus (from C36 to K65) of the α precursor is belonging to the plantaricin C family of lantibiotics with a identity of 83.33% to the known peptide staphylococcin C55 α. Whereas the C terminus (from I38 to C67) of the β precursor shows 62.07% identity to lacticin 3147 A2 [38]. The second novel class II lanthipeptide BGC was discovered from B. subtilis EH5 (Fig. 4d). This BGC has six genes with a length of ~9 kb. The presursors of two core peptide genes (α and β) contain 65 and 67 amino acids respectively. It is also showing 70% gene sequence similarity to staphylococcin C55 α/β BGC. The C terminus (from C36 to C64) of the α precursor has a similarity of 79.31% to the known peptide staphylococcin C55 α and the C terminus (from W38 to C63) of the β precursor is showing 72% identity to lacticin 3147 A2. The third BGC was identified from B. endophyticus FH5 (Fig. 4e). It is comprised of nine genes with a size of ~10 kb. Its precursors of two peptides (α and β) contain 58 and 54 amino acids respectively. There is no similarity found to any known BGCs. The C terminal region (from A28 to C58) of the α precursor has a similarity of 53.33% to the known peptide plantaricin W α [39] and the C terminus (from A23 to D54) of the β precursor is showing 56.25% identity to haloduracin β [40]. Furthermore, the precursor β in this potential novel BGC found in B. endophyticus FH5 has four replicates, indicating potential high amount production of β peptide.

Two novel gene clusters were identified as class III lanthipeptide BGCs. This Class contains RiPPs that are modified by the multifunctional enzymes LanKC. LanKC firstly phosphorylates the Ser/Thr residues in the substrate peptide and then similarly catalyzes modification of the substrate to form the final product, as the
class II lanthipeptide LanM enzyme [41]. The one identified from \textit{B. subtilis} EH2 contains ten genes with a size of \(~8\) kb (Fig. 4f). No similarity was found to any known BGCs. The full precursor contains 58 amino acids. The predicted cleavage site by antiSMASH is between T27 and G28. The C terminus (from G28 to N58) of the precursor has no identity to any known RiPPs. The other class III lanthipeptide BGC is harbored by \textit{B. velezensis} TH16 (Fig. 4g). This one contains five genes with a length of \(~5\) kb. The core biosynthetic gene encodes a 45-amino acid precursor peptide. 35\% genes of this BGC show similarity to locillomycin [42], which is a cyclic lipopeptide (NRPs) discovered from \textit{B. subtilis} 916. The predicted cleavage site is between V21 and D22 by antiSMASH and the C terminus (from D22 to C45) of the precursor has no identity to any known RiPPs.

Two novel lasso peptide BGCs were identified from the genomes of \textit{P. xylanexedens} edo6 and \textit{B. endophyticus} FH5. The one from \textit{P. xylanexedens} edo6 contains eight genes with a size of \(~8\) kb (Fig. 4h). It shows that gene sequences are 60\% similar to that of the paeninodin BGC [43]. The precursor peptide contains 45 amino acids. The predicted cleavage site is between M22 and A23. The core peptide (from A23 to S45) shows 33.3\% identity to the paeninodin [43] from \textit{P. dendritiformis} C454. Another novel lasso peptide BGC was mined from \textit{B. endophyticus} FH5 (Fig. 4i). This BGC comprised of six genes. It is showing 80\% genes similarity to paeninodin. Its precursor peptide contains 45 amino acids. The cleavage site is between M20 and A21. The core peptide (from A21 to S45) has 76\% identity to the paeninodin.

Large-scale genome-based analysis of the bioactive potential of \textit{Bacillus}

Lipopetides produced by the \textit{Bacillus} genus are involved in the biocontrol mechanisms of plant pathogens [44]. To gain a general overview of BGCs distributed in the genomes of \textit{Bacillus} genus, the diversity of BGCs in the genomes of \textit{Bacillus} isolated was investigated. The complete genomes of 555 \textit{Bacillus} strains from 60 species of \textit{Bacillales} were downloaded from Genbank and analyzed by antiSMASH 5.0 [18]. A total of 9459 BGCs were predicted and identified, which included NRPs (2377 BGCs), RiPPs (1564 BGCs), Type I PKs (517 BGCs), PKs-NRPs hybrids (309 BGCs), PKs (including Trans AT-PKs and Type III PKs) (1369 BGCs), Terpene (970 BGCs), Saccharide (62 BGCs) and Others (2291 BGCs). The BGCs were then analyzed using BiG-SCAPE [45], a program that constructs sequence similarity networks of BGCs and groups them into Gene Cluster Families (GCFs). For visualization, the distance matrix between BGCs generated by BiG-SCAPE was used in Cytoscape [46]. The similarity network of predicted BGCs revealed that a large number of BGCs are present in \textit{Bacillus} strains, and are distributed throughout different kinds of secondary metabolites (Fig. 5). Based on our investigation, some of the NRPs BGCs were conserved among the BGCs identified in the \textit{Bacillus} species. 259 out of 2377 (10.85\%) NRPs BGCs were encoding surfactin, 330 (13.88\%) BGCs were encoding bacillibactin, 110 (4.63\%) NRPs BGCs were encoding fengycin, 158 (6.65\%) NRPs BGCs were encoding petroactin [47]. And 38 (1.60\%) NRPs BGCs were encoding lichenysin [48]. Thus, a total of \(~38\)% of the NRPs BGCs are correlated to already reported compounds. Additionally, most of PKs-NRPs hybrid BGCs (67.64\%) were encoding bacillicene. Unlike the well-described NRPs and PKs-NRPs hybrid BGCs, the PKs BGCs were mostly attributed to unknown products with the exception of macrolactin [49] and difficidin [25]. Notably, 1357 out of 1564 (87.76\%) RiPPs BGCs were also unknown. Overall, the distribution of known and unknown BGCs vary dramatically across the different kinds of metabolites in \textit{Bacillus} species, in which the NRPs BGCs are the most abundant ones, comprising 2377 BGCs. Many of them are conserved and already characterized, but still a large number of unknown NRPs BGCs are identified for further study.

Discussion

\textit{Bacillus} strains attract more and more attention due to their ability to produce hard, resistant endospores and antibiotics which have the potential to be used as biocontrol agents. In this study, We found 34 \textit{Bacillus} strains (out of 351 strains) have antagonistic activity against six major tomato plant pathogens (\textit{E. carotovora}, \textit{P. syringae}, \textit{R. solani}, \textit{B. cinerea}, \textit{V. dahliae}, and \textit{P. infestans}). \textit{E. carotovora} and \textit{P. syringae} are Gram-negative phytopathogens that cause bacterial soft rot disease and bacterial speck disease in tomato, respectively [50, 51]. \textit{R. solani}, \textit{B. cinerea} and \textit{V. dahliae} are fungal pathogens that cause damping-off, gray mold and verticillium wilt diseases of tomato [3, 52, 53]. \textit{P. infestans}, is an oomycetal pathogen that resembles fungi in lifestyle and morphology but without an evolutionary relationship to fungi, and can cause late blight disease of tomato [54,
Based on the genome mining, a total of 14 novel BGCs were revealed from the genomes of 10 sequenced strains. Five novel clusters are identified as NRPs and PKs-NRPs hybrid BGCs. These categories of BGCs encode non-ribosomally synthesized peptides synthetases (NRPSs) and hybrid polyketides synthases and non-ribosomally synthesized peptides synthetases (PKS-NRPSs), which are modular multienzymes. NRPSs construct peptides from amino acids and PKS-NRPSs construct hybrid molecules from acyl-CoA moieties together with amino acids [56]. So far, some known bioactive compounds identified from Bacillus strains are belonging to these modular biosynthetic compounds, such as surfactin [20], iturin [57], bacillomycin [20], fengcin [20] and difficidin [25]. All of them are antimicrobials which can be used for biocontrol in agriculture. It is meaningful to investigate the compounds produced by the five novel modular BGCs identified. Due to the antagonistic assays in vitro, we speculate some of them have antibacterial and (or) antifungal activities. This needs to be characterized by experiments in the future. Meanwhile, nine novel RiPPs BGCs were discovered. They are categorized into lanthipeptide I/II/III and lasso peptide BGCs. Lanthipeptides (also called lantibiotics for those with antibacterial activities) are ribosomally synthesized post-translationally modified peptides having thioether cross-linked amino acids, lanthionines, as a structural element [58]. They have potentials to be used as therapeutics. Subtilin, from lanthipeptide I, is one of the most studied bacteriocins from the Bacillus strains [15]. It is synthesized by spa BGC which is possessing strong antibiotic activities [59]. Mersacidin, is produced by mrs BGC in Bacillus sp. HIL Y-85/54728 which is belonging to lanthipeptide II. It has activity against Gram-positive bacteria including Staphylococcus aureus, Streptococcus pneumoniae and Enterococcus faecium [59]. However, To date, only Lanthipeptide I and II have been reported from Bacillus strains. With an increasing number of genomes, more and more lanthipeptides BGCs are certainly discovered, but only a few of them have been characterized by experimental researches. Lasso peptides contain a macrocyclic linkage between an Asp or Glu side chain to the N-terminus of the core peptide. The C-terminal tail is threaded through the macrocycle, giving a lariat topology for which the lasso peptides are named [60]. Until now, some lasso peptides are reported with antimicrobial activity. i.e. lariatin [61], lassomycin [62] and microcin J25 [63]. However, we can not predict the compounds produced by these nine novel lanthipeptides and lasso peptides BGCs are antimicrobials or not, which need to be investigated by further experiments.

The result of large-scale genome mining of bioactive potential of Bacillus shows the distribution of known and unknown BGCs varying dramatically across the different kinds of metabolites in Bacillus species. NRPs and PKs-NRPs BGCs in Bacillus species are encoding highly conserved known compounds as well as various novel variants. These findings are consistent with KirK J. Grubbs [64] reported. They found that many of the highly conserved BGCs in Bacillus genomes set encode NRPSs; a few additional PKS-NRPSs, such as zwittermicin and bacillaene, are also well conserved. These conserved compounds have important roles influencing the physiology and development of Bacillus species. In addition, a lot of novel BGCs including variants are also discovered from Bacillus species, which are deserved to be characterized by experimental investigation. The genus Bacillus is well known for the natural products with antibacterial and antifungal activities, which has a strong potential to be applied to agriculture for plant diseases control [65]. Therefore, novel antimicrobials discovery is in need of identification and characterization of novel BGCs in Bacillus strains.

Conclusions

This work showed that 10 Bacillus and Paenibacillus strains, selected from 351 that were isolated from the rhizosphere soil of healthy tomato plants and their tissues, have strong in vitro antagonistic activity against tomato bacterial, fungal and oomycetal pathogens. Based on genome mining, we identified a large number of BGCs from their genome sequences encoding known and unknown compounds, which form a great source for pharmaceutical compound discovery. Furthermore, a total of 14 novel BGCs were characterized in detail, including 2 NRPs, 3 PKs-NRPs hybrid and 9 RiPPs BGCs. In addition, from the large-scale bioinformatics analysis of the genomes from Bacillus genus, we found that NRPS and PKS-NRPS BGCs resources hidden in Bacillus species are frequently encoding highly conserved known compounds including surfactin, fengycin, bacillibactin, petrobactin, lichenysin and bacillaene.
Sample Collection, Bacteria Isolation, and culture conditions

Healthy tomato plants (cultivar: Boludo) and their rhizosphere soil were carefully collected during spring (February 2017) from tomatoes grown in a garden in the village of Roden in the Netherlands and Almería in Spain, which were given to us by company Koppert and with their consent. The bacterial isolation was performed as described previously [17]. Briefly, 1 g rhizosphere soil was suspended in 9 ml of 10 mM sterilized MgSO₄ buffer. Then the suspension was diluted 10³-10⁶ times with 10 mM sterilized MgSO₄ buffer. After dilution, all the samples were heat-treated at 80 °C for 15 min. and subsequently spread on Luria-Bertani (LB) agar plates. The plates were incubated at 28 °C for 24-48 h to obtain single colonies. For the isolation of endophytes, 1 g tomato leaves were surface-sterilized for 1 min. in 70% ethanol and 3 min. in 0.5% NaClO solution supplemented with one droplet Tween 80 per 100 ml solution and then rinsed 5 times with sterilized deionized water. After surface sterilization, the leaves were macerated in 9 ml of 10 mM sterilized MgSO₄ buffer with a sterilized mortar to obtain the suspension. The following steps were the same as isolation from rhizosphere soil. The surface-sterilization process was checked by spreading aliquots of the last rinsing deionized water on LB agar plates (if no organism growth was observed after 7 days, surface sterilization was considered to be successful). All the isolated strains were stored in 25% glycerol solution at -80 °C untill further investigation.

Screening of Antimicrobial Activity

In vitro antagonistic activity assays were performed on dual culture plates as described before with slight modification [66]. All the isolated strains were screened against different bacterial, fungal and oomycetal plant pathogens, such as E. carotovora, P. syringae, R. solani, B. cinerea, V. dahliae, and P. infestans.

To test antibacterial activity, bacterial pathogens were mixed with pre-cooled LB agar media (around 55 °C) at a final concentration of 1 × 10⁶ cells/ml. Then the mixed media was poured into Petri dishes to obtain pathogen-fusion agar plates. 5 ul of 1 × 10⁸ cells/ml overnight culture of each isolated strain was inoculated at the center of plates. All the plates were incubated at 28 °C for 2 days before the clear halo surrounding the strain isolated was measured.

Antagonistic activity of isolated strains against R. solani was tested as follows. A 0.5-cm mycelium plug of 3-day-old R. solani was placed at the center of the 1/5th PDA plate, and 5 ul of 1 × 10⁸ cells/ml overnight culture of each isolated strain was inoculated at a distance of 2 cm from the fungus. Plates were incubated at 28 °C for 3 days and inhibition of fungal growth was recorded as the diameter of the inhibition zone (mm). As a control, LB media was used in place of the bacterial suspension.

Antagonistic activity determination of isolated strains against B. cinerea, V. dahliae, and P. infestans was performed similarly as the antibacterial activity assay. Spores of B. cinerea and V. dahliae were collected respectively from 7-day-old and 20-day-old PDA plates with sterilized Mili-Q water by washing the mycelium. All the spores were counted using a Thoma chamber and were then mixed into 1/5th PDA media and adjusted to 1 × 10⁷ spores/ml. In addition, sporangia of P. infestans were harvested by washing the 30-day-old RSA plates with sterilized mili-Q water and then counted and mixed into 1/5th PDA media at the final concentration of 4000 sporangia/ml. Before mixing, the sporangia suspension was stimulated to release zoospores by chilling for 1-3 hours at 4 °C. Subsequently, the mixed 1/5th PDA media was poured into Petri dishes. 5 ul of 1 × 10⁸ cells /ml overnight culture of each isolated strain was inoculated at the center of plates. All the plates were incubated at 28 °C. The clear halo surrounding each isolated strain was monitored and measured depending on the pathogens’ growth rate.

Genome Sequencing, Phylogenetic Analysis, Antimicrobial Compounds Mining, and Metabolite BGC Network Analysis

Genomic DNA of isolates was extracted with a GenElute Bacterial Genomic DNA kit (Sigma) according to the
The draft genomes were assembled and deposited in GeneBank. The whole genomes were aligned using the GEGENEES tool [67] with other reference genomes of reported PGPR strains. A phylogenetic tree was generated with iTOL version 4.4.2 and Splitstree [68]. For identification of biosynthesis gene clusters (BGCs), antiSMASH 5.0 [18, 69] and BAGEL4 [19] were used. Each draft genome was assembled into a pseudomolecule based on multiple closely related strains as a reference with Medusa web server (http://combo.dbe.unifi.it/medusa) [70]. The similarity network between BGCs was calculated with BiG-SCAPE (https://git.wageningenur.nl/medema-group/BiG-SCAPE) [45]. A cut-off of 0.3 was used for the analysis. To visualize, the network file generated by BiG-SCAPE was exported and annotated in Cytoscape v3.7.0 (http://www.cytoscape.org/) [46]. Default parameters were used for all software unless noted.

**Abbreviations**

PGPR: plant growth promoting rhizobacteria; BGCs: biosynthetic gene clusters; ISR: inducing systemic resistance; NRPSs: non-ribosomally synthesized peptides synthetases; PKSs: polyketides synthases; NRPs: non-ribosomally synthesized peptides; PKs: polyketides; PKS-NRPSs: hybrid polyketides synthases and non-ribosomally synthesized peptides synthetases; RiPPs: ribosomally synthesized and post-translationally modified peptides; PTMs: posttranslational modifications; C: condensation; A: adenylation; T: thiolation; E: epimerization; TE: thioesterase; KS: ketosynthase; AT: acyltransferase; ACP: acyl carrier protein; TD: reductase; DHT: dehydratase domain variant; Cy: heterocyclization

**Declarations**

**Acknowledgements**

We thank the Koppert company in Netherlands for providing tomato plants and their rhizosphere soil and Dr. Sjoerd van der Meulen for garden samples.

**Funding**

L. Zhou was financially supported by the China Scholarship Council (201606910037). C. Song was supported by a grant of NWO-STW for the Back to the Roots project. Z. Li was financially supported by the China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of manuscript.

**Author’s contributions**

LZ and OPK conceived the study and designed experiments. LZ and CS conducted the experiments. LZ performed the bioinformatic analysis and wrote the draft manuscript. LZ, CS, ZL and OPK corrected the manuscript. All authors have read and approved the manuscript.

**Corresponding author**

Correspondence to Oscar P. Kuipers o.p.kuipers@rug.nl

**Availability of data and materials**

The whole genome data are available at DDBJ/EMBL/GenBank under the bioproject accession PRJNA503984 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA503984/)

**Ethics declarations**

**Ethics approval and consent to participate**

Not applicable
Consent for publication

Not applicable

Competing interests

The authors declare that they have no conflict interests

References

1. Olaniyi J, Akanbi W, Adejumo T, Ak OJAJoFS: Growth, fruit yield and nutritional quality of tomato varieties. 2010, 4:398-402.
2. Dias MP, Bastos MS, Xavier VB, Cassel E, Astarita LV, Santarem ER: Plant growth and resistance promoted by Streptomyces spp. in tomato. Plant Physiol Biochem 2017, 118:479-493.
3. Asaka O, Shoda MJAEM: Biocontrol of Rhizoctonia solani damping-off of tomato with Bacillus subtilis RB14. Appl Environ Microbiol 1996, 62:4081-4085.
4. Barea JM, Pozo MJ, Azcon R, Azcon-Aguilar C: Microbial co-operation in the rhizosphere. J Exp Bot 2005, 56:1761-1778.
5. Raaijmakers JM, Paulitz TC, Steinberg C, Alabouvette C, Moënne-Loccoz Y: The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. Plant and Soil 2009, 321:341-361.
6. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH: Going back to the roots: the microbial ecology of the rhizosphere. Nat Rev Microbiol 2013, 11:789-799.
7. Lynch J, Whipps JJP, soil: Substrate flow in the rhizosphere. 1990, 129:1-10.
8. Whipps JMJJoeB: Microbial interactions and biocontrol in the rhizosphere. 2001, 52:487-511.
9. Silo-Suh LA, Lethbridge BJ, Raffel SJ, He H, Clardy J, Handelsman JJAEM: Biological activities of two fungistatic antibiotics produced by Bacillus cereus UW85. 1994, 60:2023-2030.
10. Wang S, Wu H, Qiao J, Ma L, Liu J, Xia Y, Gao XJJMB: Molecular mechanism of plant growth promotion and induced systemic resistance to tobacco mosaic virus by Bacillus spp. 2009, 19:1250-1258.
11. Ryu C-M, Farag MA, Hu C-H, Reddy MS, Kloepper JW, Paré PWJPP: Bacterial volatiles induce systemic resistance in Arabidopsis. 2004, 134:1017-1026.
12. Choudhary DK, Johri BN: Interactions of Bacillus spp. and plants--with special reference to induced systemic resistance (ISR). Microbiol Res 2009, 164:493-513.
13. Ongen M, Duby F, Jourdan E, Beaudry T, Jadin V, Dommes J, Thonart P: Bacillus subtilis M4 decreases plant susceptibility towards fungal pathogens by increasing host resistance associated with differential gene expression. Appl Microbiol Biotechnol 2005, 67:692-698.
14. Perez-Garcia A, Romero D, de Vicente A: Plant protection and growth stimulation by microorganisms: biotechnological applications of Bacilli in agriculture. Curr Opin Biotechnol 2011, 22:187-193.
15. Caulier S, Nannan C, Gillis A, Licciardi F, Bragard C, Mahillon J: Overview of the Antimicrobial Compounds Produced by Members of the Bacillus subtilis Group. Frontiers in Microbiology 2019, 10.
16. Crits-Christoph A, Diamond S, Butterfield CN, Thomas BC, Banfield JF: Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. Nature 2018, 558:440-444.
17. Zhou L, Song C, de Jong A, Kuipers OP: Draft Genome Sequences of 10 Paenibacillus and Bacillus sp. Strains Isolated from Healthy Tomato Plants and Rhizosphere Soil. Microbiol Resour Announc 2019, 8.
18. 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.
19. 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:W278-W281.
20. Koumoutsi A, Chen XH, Henne A, Liesegang H, Hitzeroth G, Franke P, Vater J, Borris R: Structural and
Functional Characterization of Gene Clusters Directing Nonribosomal Synthesis of Bioactive Cyclic Lipopeptides in Bacillus amyloliquefaciens Strain FZB42. Journal of Bacteriology 2004, 186:1084-1096.

21. Chen XH, Koumoutsi A, Scholz R, Borriess R: More than anticipated - production of antibiotics and other secondary metabolites by Bacillus amyloliquefaciens FZB42. J Mol Microbiol Biotechnol 2009, 16:14-24.

22. Kawulka KE, Sprules T, Diaper CM, Whittal RM, McKay RT, Mercier P, Zuber P, Vederas JCJB: Structure of subtilosin A, a cyclic antimicrobial peptide from Bacillus subtilis with unusual sulfur to α-carbon cross-links: formation and reduction of α-thio-α-amino acid derivatives. Biochemistry 2004, 43:3385-3395.

23. Moldenhauer J, Chen XH, Borriess R, Piel J: Biosynthesis of the antibiotic bacillaene, the product of a giant polypeptide synthase complex of the trans-AT family. Angew Chem Int Ed Engl 2007, 46:8195-8197.

24. <2007-Macrolactin is the Polyketide Biosynthesis Product of the pks2 Cluster of Bacillus amyloliquefaciens FZB42.pdf>.

25. Chen XH, Vater J, Piel J, Franke P, Scholz R, Schneider K, Koumoutsi A, Hitzeroth G, Grammel N, Strittmatter AW, et al: Structural and functional characterization of three polyketide synthase gene clusters in Bacillus amyloliquefaciens FZB 42. J Bacteriol 2006, 188:4024-4036.

26. Klein C, Kaletta C, Entian K: Biosynthesis of the lantibiotic subtilin is regulated by a histidine kinase/response regulator system. Appl Environ Microbiol 1993, 59:296-303.

27. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft J, Van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V: MIBig 2.0: a repository for biosynthetic gene clusters of known function. Nucleic acids research 2020, 48:D454-D458.

28. Garcia-Gonzalez E, Muller S, Hertlein G, Heid N, Sussmuth RD, Genersch E: Biological effects of paenilamicin, a secondary metabolite antibiotic produced by the honey bee pathogenic bacterium Paenibacillus larvae. Microbiologyopen 2014, 3:642-656.

29. Fiers WD, Dodge GJ, Sherman DH, Smith JL, Aldrich CC: Vinlylogous Dehydration by a Polyketide Dehydratase Domain in Curacin Biosynthesis. J Am Chem Soc 2016, 138:16024-16036.

30. Jahns C, Hoffmann T, Muller S, Gerth K, Washausen P, Hofle G, Reichenbach H, Kalesse M, Muller R: Pellasoren: structure elucidation, biosynthesis, and total synthesis of a cytotoxic secondary metabolite from Sorangium cellulosum. Angew Chem Int Ed Engl 2012, 51:5239-5243.

31. Reimer D, Luxenburger E, Brachmann AO, Bode HB: A new type of pyrrolidine biosynthesis is involved in the late steps of xenocoumacin production in Xenorhabdus nematophila. Chembiochem 2009, 10:1977-1981.

32. Bloudoff K, Fage CD, Marahiel MA, Schmeing TM: Structural and mutational analysis of the nonribosomal peptide synthetase heterocyclization domain provides insight into catalysis. Proc Natl Acad Sci U S A 2017, 114:95-100.

33. Kumawat MK: Thiazole containing heterocycles with antimalarial activity. Current drug discovery technologies 2018, 15:196-200.

34. Rouf A, Tanyeli C: Bioactive thiazole and benzothiazole derivatives. European journal of medicinal chemistry 2015, 97:911-927.

35. Zhao X, Kuipers OP: Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. BMC Genomics 2016, 17:882.

36. Xin B, Zheng J, Liu H, Li J, Ruan L, Peng D, Sajid M, Sun M: Thusin, a Novel Two-Component Lantibiotic with Potent Antimicrobial Activity against Several Gram-Positive Pathogens. Front Microbiol 2016, 7:1115.

37. Navaratna MA, Sahl H-G, Tagg JR: Identification of genes encoding two-component lantibiotic production in Staphylococcus aureus C55 and other phage group II S. aureus strains and demonstration of an association with the exfoliative toxin B gene. Infection and immunity 1999, 67:4268-4271.

38. Ryan MP, Jack RW, Josten M, Sahl H-G, Jung G, Ross RP, Hill C: Extensive post-translational modification, including serine to D-alanine conversion, in the two-component lantibiotic, lacticin 3147. Journal of Biological Chemistry 1999, 274:37544-37550.

39. Holo H, Jeknic Z, Daeschel M, Stevanovic S, Nes IF: Plantaricin W from Lactobacillus plantarum
belongs to a new family of two-peptide lantibiotics. *Microbiology* 2001, **147**:643-651.

40. Lawton EM, Cotter PD, Hill C, Ross RP: Identification of a novel two-peptide lantibiotic, haloduracin, produced by the alkaliphile Bacillus halodurans C-125. *FEMS Microbiol Lett* 2007, **267**:64-71.

41. Zhang Q, Yu Y, Velasquez JE, van der Donk WA: Evolution of lanthipeptide synthetases. *Proc Natl Acad Sci U S A* 2012, **109**:18361-18366.

42. Luo C, Liu X, Zhou H, Wang X, Chen Z: Nonribosomal peptide synthase gene clusters for lipopeptide biosynthesis in Bacillus subtilis 916 and their phenotypic functions. *Appl Environ Microbiol* 2015, **81**:422-431.

43. Zhu S, Hegemann JD, Fage CD, Zimmermann M, Xie X, Linne U, Marahiel MA: Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. *J Biol Chem* 2016, **291**:13662-13678.

44. Ongena M, Jacques P: Bacillus lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol* 2008, **16**:115-125.

45. Yeong M: BiG-SCAPE: exploring biosynthetic diversity through gene cluster similarity networks. 2016.

46. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, Pico AR, Bader GD, Ideker T: A travel guide to Cytoscape plugins. *Nature methods* 2012, **9**:1069.

47. Cendrowski S, MacArthur W, Hanna P: *Bacillus anthracis* requires siderophore biosynthesis for growth in macrophages and mouse virulence. *Molecular microbiology* 2004, **51**:407-417.

48. Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, Bäumer S, Henne A, Liesegang H, Merkl R: The complete genome sequence of Bacillus licheniformis DSM13, an organism with great industrial potential. *Journal of molecular microbiology and biotechnology* 2004, **7**:204-211.

49. Schneider K, Chen X-H, Vater J, Franke P, Nicholson G, Borris R, Süßmuth RD: Macrolactin is the polyketide biosynthesis product of the pks2 cluster of Bacillus amyloliquefaciens FZB42. *Journal of natural products* 2007, **70**:1417-1423.

50. Chandrasekaran M, Chun SC: Expression of PR-protein genes and induction of defense-related enzymes by Bacillus subtilis CBR05 in tomato (Solanum lycopersicum) plants challenged with *Erwinia carotovora* subsp. *carotovora*. *Biosci Biotechnol Biochem* 2016, **80**:2277-2283.

51. Cai R, Lewis J, Yan S, Liu H, Clarke CR, Campanile F, Almeida NF, Studholme DJ, Lindeberg M, Schneider D, et al: The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. *PLoS Pathog* 2011, **7**:e1002130.

52. Sanchez-Bel P, Troncho P, Gamir J, Pozo MJ, Camanes G, Cerezo M, Flors V: The Nitrogen Availability Interferes with Mycorrhiza-Induced Resistance against Botrytis cinerea in Tomato. *Front Microbiol* 2016, **7**:1598.

53. Han Q, Wu F, Wang X, Qi H, Shi L, Ren A, Liu Q, Zhao M, Tang C: The bacterial lipopeptide iturins induce *Verticillium dahliae* cell death by affecting fungal signalling pathways and mediate plant defence responses involved in pathogen-associated molecular pattern-triggered immunity. *Environ Microbiol* 2015, **17**:1166-1188.

54. Kots K, Meijer HJ, Bouwmeester K, Govers F, Ketelaar T: Filamentous actin accumulates during plant cell penetration and cell wall plug formation in *Phytophthora infestans*. *Cell Mol Life Sci* 2017, **74**:909-920.

55. Zuluaga AP, Vega-Arreguin JC, Fei Z, Ponnala L, Lee SJ, Matas AJ, Patev S, Fry WE, Rose JK: Transcriptional dynamics of *Phytophthora infestans* during sequential stages of hemibiotrophic infection of tomato. *Mol Plant Pathol* 2016, **17**:29-41.

56. Weissman KJ: The structural biology of biosynthetic megaenzymes. *Nat Chem Biol* 2015, **11**:660-670.

57. Tsuge K, Akiyama T, Shoda M: Cloning, sequencing, and characterization of the iturin A operon. *J Bacteriol* 2001, **183**:6265-6273.

58. Ongey EL, Neubauer P: Lanthipeptides: chemical synthesis versus in vivo biosynthesis as tools for pharmaceutical production. *Microb Cell Fact* 2016, **15**:97.

59. Barbosa J, Caetano T, Mendo S: Class I and Class II Lanthipeptides Produced by Bacillus spp. *J Nat Prod* 2015, **78**:2850-2866.

60. Tietz JL, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA: A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol* 2017, **13**:470-478.

61. Inokoshi J, Matsuhama M, Miyake M, Ikeda H, Tomoda H: Molecular cloning of the gene cluster for...
lariatin biosynthesis of Rhodococcus jostii K01-B0171. Applied microbiology and biotechnology 2012, 95:451-460.

62. Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissell A, et al: Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. Chem Biol 2014, 21:509-518.

63. Wilson KA, Kalkum M, Ottesen J, Yuzenkova J, Chait BT, Landick R, Muir T, Severinov K, Darst SA: Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. Journal of the American Chemical Society 2003, 125:12475-12483.

64. Grubbs KJ, Bleich RM, Santa Maria KC, Allen SE, Farag S, AgBiome T, Shank EA, Bowers AA: Large-Scale Bioinformatics Analysis of Bacillus Genomes Uncovers Conserved Roles of Natural Products in Bacterial Physiology. mSystems 2017, 2.

65. Mohamad OAA, Li L, Ma JB, Hatab S, Xu L, Guo JW, Rasulov BA, Liu YH, Hedlund BP, Li WJ: Evaluation of the Antimicrobial Activity of Endophytic Bacterial Populations From Chinese Traditional Medicinal Plant Licorice and Characterization of the Bioactive Secondary Metabolites Produced by Bacillus atrophaeus Against Verticillium dahliae. Front Microbiol 2018, 9:924.

66. Chen Y, Yan F, Chai Y, Liu H, Kolter R, Losick R, Guo JH: Biocontrol of tomato wilt disease by Bacillus subtilis isolates from natural environments depends on conserved genes mediating biofilm formation. Environ Microbiol 2013, 15:848-864.

67. Agren J, Sundstrom A, Hafstrom T, Segerman B: Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One 2012, 7:e39107.

68. Huson DH, Bryant D: Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 2006, 23:254-267.

69. Letunic I, Bork P: Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 2019, 47:W256-W259.

70. 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.
Figure 1

Neighbor-joining tree based on 16S rRNA genes of 34 isolated strains showing antagonistic activity against tomato plant bacterial, fungal and oomycetal pathogens. Red represents the inhibition of bacterial pathogens and blue indicates the inhibition of fungal and oomycetal pathogens. The red and blue scale bar represent the radius of inhibition halo observed (mm).
Tree scale: 0.1

Bacillus endophyticus KCTC 13922
Bacillus endophyticus FH5
Bacillus endophyticus DSM 13796
Bacillus endophyticus Hbe603
Bacillus endophyticus 2102

Bacillus subtilis subsp. subtilis str. BAB-1
Bacillus subtilis NCD-2
Figure 2

Phylogenetic position of 10 isolated Bacillus and Paenibacillus strains with high significant antagonistic activity against tomato pathogens. A maximum likelihood (ML) tree was constructed based whole genome sequences analysis using Gegenee.
B. endophyticus FH5 (cluster 5) Type I PKS-NRPS hybrid BGC 30553bp

P. xylanexedens EDO6 (cluster 13) trans-AT PKS-NRPS hybrid BGC 35857bp

P. xylanexedens EDO6 (cluster 12) trans-AT PKS-NRPS hybrid BGC 28523bp

B. velezensis FH17 (cluster 7) TH16 (cluster 12) NRPS BGC 33403bp
Figure 3

Novel NRPS and PKS BGCs identified from the isolated Bacillus and Paenibacillus strains
Figure 4

Novel bacteriocin BGCs identified from the isolated Bacillus and Paenibacillus strains
Figure 5

The predicted Biosynthetic Gene Clusters (BGCs) Similarity network of 555 Bacillus strains genomes showing their diversity, distribution. The lines between the nodes represent genes shared between BGCs.