

RESEARCH

Open Access



Organization, evolution and function of fengycin biosynthesis gene clusters in the *Bacillus amyloliquefaciens* group

Qingchao Zeng^{1,2}, Jianbo Xie², Yan Li¹, Xinyi Chen¹, Xiaofei Gu¹, Panlei Yang¹, Guangcan Hu³ and Qi Wang^{1*}

Abstract

The *Bacillus velezensis* strain PG12, belonging to the *Bacillus amyloliquefaciens* group, is an endophytic bacterium known for its antimicrobial activities against crop pathogens. However, our knowledge of the molecular basis underlying its biocontrol activity and the relatedness of different strains in the *Bacillus amyloliquefaciens* group is limited. Here, we sequenced and analyzed the genome of PG12 to test its taxonomic affiliation and identified genes involved in the biocontrol activity. The phylogenomic analysis results indicate that PG12 belongs to *B. velezensis*, a subgroup of the *B. amyloliquefaciens* group. By comparing the genomes of 22 strains in this group, we confirmed that it comprises three different phylogenetic lineages: *B. amyloliquefaciens*, *B. velezensis* and *B. siamensis*. Three secondary metabolism gene clusters related to the production of lipopeptides, namely fengycin, iturin and surfactin, were identified in the genomes of the *B. amyloliquefaciens* group. The core genome of *B. velezensis* is enriched in secondary metabolism genes compared with *B. siamensis* and *B. amyloliquefaciens*. Three of the five genes pertaining to the gene cluster responsible for fengycin biosynthesis (*fenBCD*) were found in *B. velezensis* and *B. siamensis*, but not in *B. amyloliquefaciens*. Phenotypic analysis showed that the $\Delta fenA$ mutant of PG12 displayed significantly decreased biofilm formation and swarming motility, which indicates that fengycin contributes to the colonization and pathogen control abilities of PG12. Our results also suggest that *B. siamensis* and *B. velezensis* have acquired the *fenBCD* genes from *Paenibacillus* spp. by horizontal gene transfer (HGT). Taken together, the results provide insights into the evolutionary pattern of the *B. amyloliquefaciens* group strains and will promote further researches on their taxonomy and functional genomics.

Keywords: *Bacillus amyloliquefaciens*, Biocontrol, Comparative genomic, Evolution, Genome sequencing

Background

Pathogenic microorganisms affecting plant health are a major and chronic threat to food production worldwide (Compant et al. 2005). The annual yield loss of crop plants due to microbial diseases is approximately 25%. To reduce crop loss, farmers rely heavily on agrochemicals (Schäfer and Adams 2015; Wu et al. 2015). However, overuse of chemicals to enhance crop yield and control

plant diseases has resulted in an irreversible loss of soil quality along with serious health and environmental problems (Prashar et al. 2013). As a result, many countries are now restricting the use of a wide range of fungicides and pesticides (Rahman 2013). In this changing context, disease control in plants by beneficial bacteria is steadily increasing as an alternative to chemical pesticides (Wu et al. 2015). Several bacterial species, such as *Pseudomonas* spp., *Bacillus* spp. and *Streptomyces* spp. have been commercialized as biological control agents (Paterson et al. 2017). Among them, *Bacillus* spp. have become increasingly important in agriculture

*Correspondence: wangqi@cau.edu.cn

¹ Department of Plant Pathology, MOA Key Lab of Pest Monitoring and Green Management, College of Plant Protection, China Agricultural University, Beijing 100193, China

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

and agro-food industry where spore-forming strains are favored due to their long-term viability (Qiao et al. 2014).

The genus *Bacillus* represents a large group of Gram-positive bacteria belonging to the Firmicutes phylum. They are capable of forming stable dormant structures called endospores in nutrient-poor and stressful environmental conditions (Hamdache et al. 2013). The bacteria from this group can inhabit a large variety of ecological niches, including soil, water, plant surfaces and rhizosphere (Siefert et al. 2000; Feng et al. 2014; Kim et al. 2015; Gao et al. 2019). *B. cereus*, *B. subtilis* and *B. amyloliquefaciens* are three *Bacillus* species known for supporting plant growth and protecting plants from diseases (Chen et al. 2016; Fan et al. 2017b; Zeng et al. 2018). The biocontrol activity mediated by *Bacillus* can be divided into direct and indirect forms. Effective biological control of pathogens by *Bacillus* is achieved by a combination of mechanisms, including production of antibiotic, induction of host resistance and promotion of plant growth (Paterson et al. 2017; Jayapala et al. 2019). Among them, the production of antibiotics is one of the primary mechanisms for achieving biocontrol effect (Li et al. 2021). The antibiotic activity of *Bacillus* species is partly driven by the secretion of active lipopeptides. These cyclic lipopeptides belong to three main families: surfactin (surfactin, lichenisin, pumilacidin and halobacilin), iturin (iturin, bacillomycin and mycosubtilin) and fengycin (fengycin, plipastatin and maltacin) (Zhao et al. 2017; Fira et al. 2018). They are produced by nonribosomal peptide synthetase and exhibit broad-spectrum antimicrobial activity (Cochrane and Vederas 2016). Different lipopeptides can act synergistically to inhibit the growth of phytopathogens (Koumoutsi et al. 2014). In addition, studies have revealed that lipopeptides play an important role in maintaining colonization behaviors such as biofilm formation and swarming motility (Luo et al. 2014; Fan et al. 2017b).

According to phylogenomic analyses, the *B. amyloliquefaciens* group comprises three species: *B. siamensis*, mainly occurring in Asian food, some of which are capable of controlling phytopathogens (Fan et al. 2017a, 2019); *B. velezensis*, the main source of bioformulations used in organic agriculture; *B. amyloliquefaciens*, known for its ability to produce industrial enzymes (Wu et al. 2015; Fan et al. 2017a). Currently, many basic characteristics of these species remain unclear. With the advent of comparative genomics and the availability of an increasing number of whole genome sequences, it is possible to understand the *B. amyloliquefaciens* group strains at genome level. In our previous study, we isolated a *B. velezensis* strain PG12 from apple fruit, which exhibits distinctive inhibition to apple ring rot pathogen and broad-spectrum antagonistic activities against other

pathogens. PG12 demonstrates an outstanding biocontrol performance, especially against *Botryosphaeria dothidea*. It produces several antagonistic compounds, including iturin and fengycin (Chen et al. 2016). Recently, the released data on genome sequences of numerous *Bacillus* strains have to some extent revealed the underlying molecular basis of their biocontrol performance (Kim et al. 2015; Li et al. 2021).

Based on the above background, the aims of the present study were to investigate: (1) the genomic features of PG12 and its phylogenetic relationship; (2) the significant genomic differences between the *B. amyloliquefaciens* group strains; (3) the difference in the specific genes involved in biocontrol activity between the *B. amyloliquefaciens* group strains. Consequently, we sequenced and analyzed the genome of PG12, and meanwhile, we collected previously published genomes of the *B. amyloliquefaciens* group strains to perform a comparative genomic analysis. The results will expand our knowledge about the *B. amyloliquefaciens* group strains, with emphasis on their taxonomical status. Understanding the genomic features enables us to gain novel insights into the ecology and evolution of the *B. amyloliquefaciens* group.

Results

Genome features of *B. velezensis* PG12

To identify genetic factors involved in the biocontrol activity of *B. velezensis* PG12, the genome of this strain was sequenced and analyzed. The genome assembly is 3,990,845 bp with 22 scaffolds ($N_{50} = 2,085,242$ bp, G + C content of 46.45%). The number of predicted protein-coding genes is 3,884, including 3,001 genes with a putative function and 883 hypothetical protein-encoding genes (Additional file 1: Figure S1). These protein-coding genes have an average length of 917 bp and account for 89.25% of the PG12 genome. Furthermore, a total of 56 tRNA-coding genes and 8 rRNA genes were predicted in the genome sequence (Table 1).

Phylogenomic analysis of *B. velezensis* PG12

Preliminary molecular identification based on 16S rDNA and *gyrA* sequences assigned PG12 strains to *B. amyloliquefaciens* (Chen et al. 2016). Here, we performed a phylogenomic reconstruction based on 624 single-copy conserved genes using the maximum-likelihood (ML) method. PG12 was clearly clustered within the *B. velezensis* clade, with *B. velezensis* CAU B946 as its closest relative (Fig. 1). They have an average nucleotide identity (ANI) of 99.96% (Additional file 2: Table S1).

Genome comparison of several *B. velezensis* strains, including PG12, revealed a high sequence conservation throughout the genome of these strains (Additional file 1:

Table 1 Genomic features of *B. velezensis* PG12

Attribute	PG12	
	Value	% of total
Genome size (bp)	3,995,119	100
DNA coding region (bp)	3,562,014	89.16
Total genes	3,956	100
Protein-coding genes	3,884	98.17
Protein-coding genes with function prediction	3,002	75.88
Protein-coding genes assigned to COGs	2,920	73.81
rRNA genes	11	0.28
tRNA genes	60	1.52

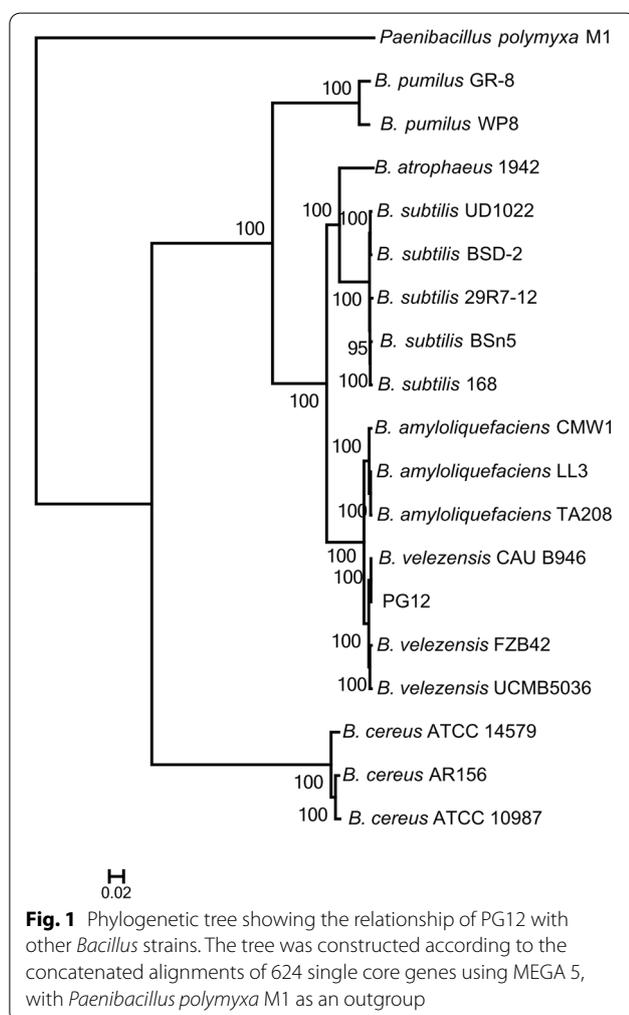


Figure S2). This was confirmed by the pan-genome analysis carried out among four *B. velezensis* strains. Genome-wide comparisons of orthologous clusters between PG12 and closely related *B. velezensis* strains indicated

that these *B. velezensis* strains contained 3,226 genes in their core genome, covering from 83.1% to 88.7% of their genomes (Additional file 1: Figure S3).

Distinct phylogenetic and evolutionary history of the *B. amyloliquefaciens* group strains

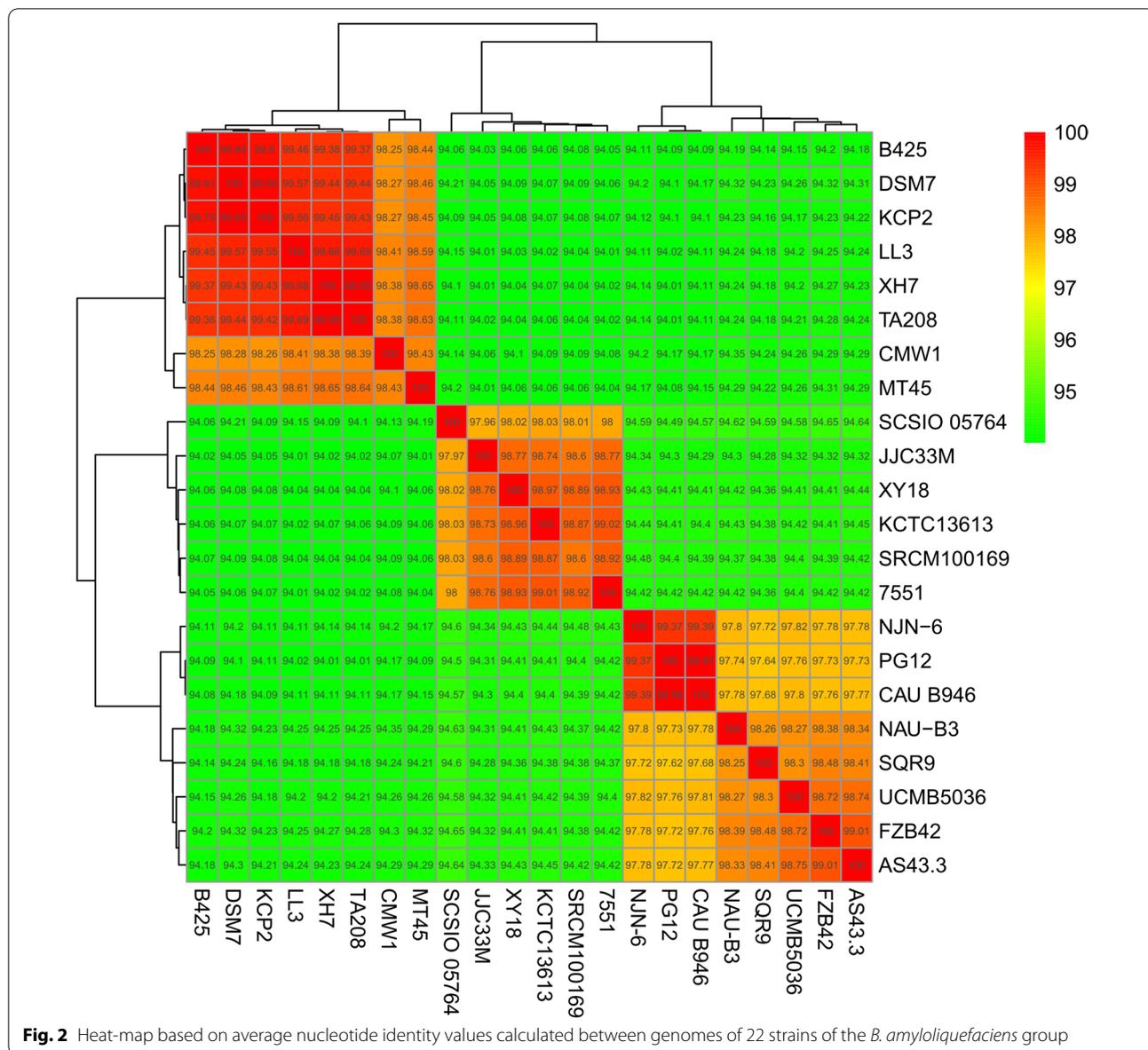
A phylogenetic tree of the *B. amyloliquefaciens* group strains based on 860 conserved single-copy genes was constructed, and the result showed that this group strains were clearly separated into three different phylogenetic lineages, i.e. *B. amyloliquefaciens*, *B. velezensis* and *B. siamensis* (Additional file 1: Figure S4), consistent with previous reports (Fan et al. 2017a; Chun et al. 2019). Our ANI analysis further supports the distribution of *B. amyloliquefaciens* group strains in three clusters (Fig. 2).

Genes involved in secondary metabolism are enriched in *B. velezensis* strains

The pan-genome analysis of 22 *B. amyloliquefaciens* group strains was conducted and yielded a pan-genome with 7,753 genes. The core, accessory and unique genomes comprise 2,806, 2,287 and 2,660 genes, respectively. The strain-specific genes for these selected *B. amyloliquefaciens* group strains ranged from 37 to 324 (Fig. 3). Meanwhile, the sizes of pan-genome and core-genome of *B. velezensis*, *B. siamensis* and *B. amyloliquefaciens* were estimated using the above genome data. *B. velezensis* strains have a core genome of 3,138 orthologous genes (76.8%–86.3% of the protein repertoire); while *B. amyloliquefaciens* and *B. siamensis* strains have a core genome of 3,116 (73.9%–81.6% of the protein repertoire) and 3,144 orthologous genes (73.1%–85.8% of the protein repertoire), respectively. The pan-genome sizes of *B. velezensis*, *B. siamensis* and *B. amyloliquefaciens* appeared to reach infinity when the number of added genomes increased infinitely, suggesting that these species have open pan-genomes (Additional file 1: Figure S5). Furthermore, we compared the core genome among the *B. amyloliquefaciens* group strains by using the Cluster of Orthologous Group of proteins (COG) assignments to determine whether there are differences in the proportion of the core genome attributable to particular cellular processes. The core genomes of *B. velezensis* strains were found to be highly enriched in secondary metabolism genes compared with those of *B. amyloliquefaciens* and *B. siamensis* strains (Table 2) (Fisher's exact test; $P < 0.05$).

The distribution of lipopeptide gene clusters in the *B. amyloliquefaciens* group strains

Three gene clusters coding for the cyclic lipopeptides: surfactin, bacillomycin and fengycin were identified in the genomes of the *B. amyloliquefaciens* group strains (Fig. 4). The most obvious difference between the



mentioned three subgroup strains was found in the gene cluster coding for fengycin. The fengycin gene cluster (*fenA*, *fenB*, *fenC*, *fenD* and *fenE*) is conserved in all the *B. velezensis* and *B. siamensis* strains that we tested. In contrast, only two genes (*fenA* and *fenE*) were found in *B. amyloliquefaciens* strains (Fig. 4 and Additional file 1: Figure S6). By contrast, the surfactin and iturin gene clusters are highly conserved in the *B. amyloliquefaciens* group, and only slight differences were noticed between different subgroup strains (Fig. 4).

In all the examined *B. velezensis* and *B. siamensis* strains, we observed striking difference ($P < 0.01$) in the G + C content between the average genome (46.36%)

and *fenBCD* gene sequences (49.77%, 51.70% and 50.82% for *fenB*, *fenC* and *fenD*, respectively) (Additional file 1: Figure S7). The observed bias in G + C content of *fenBCD* suggests a potential HGT might occur in these genes (Xie et al. 2014). To gain insights into the origin of *fenBCD* genes in *B. velezensis* strains, we constructed phylogenetic trees based on the protein sequence of *fenB*, *fenC* or *fenD* (Additional file 1: Figure S8–S10). The *fenB*, *fenC* and *fenD* sequences of *B. velezensis* clustered with their orthologs from *Paenibacillus* spp., suggesting a transfer of the fengycin gene cluster from *Paenibacillus* spp. to *B. velezensis*.

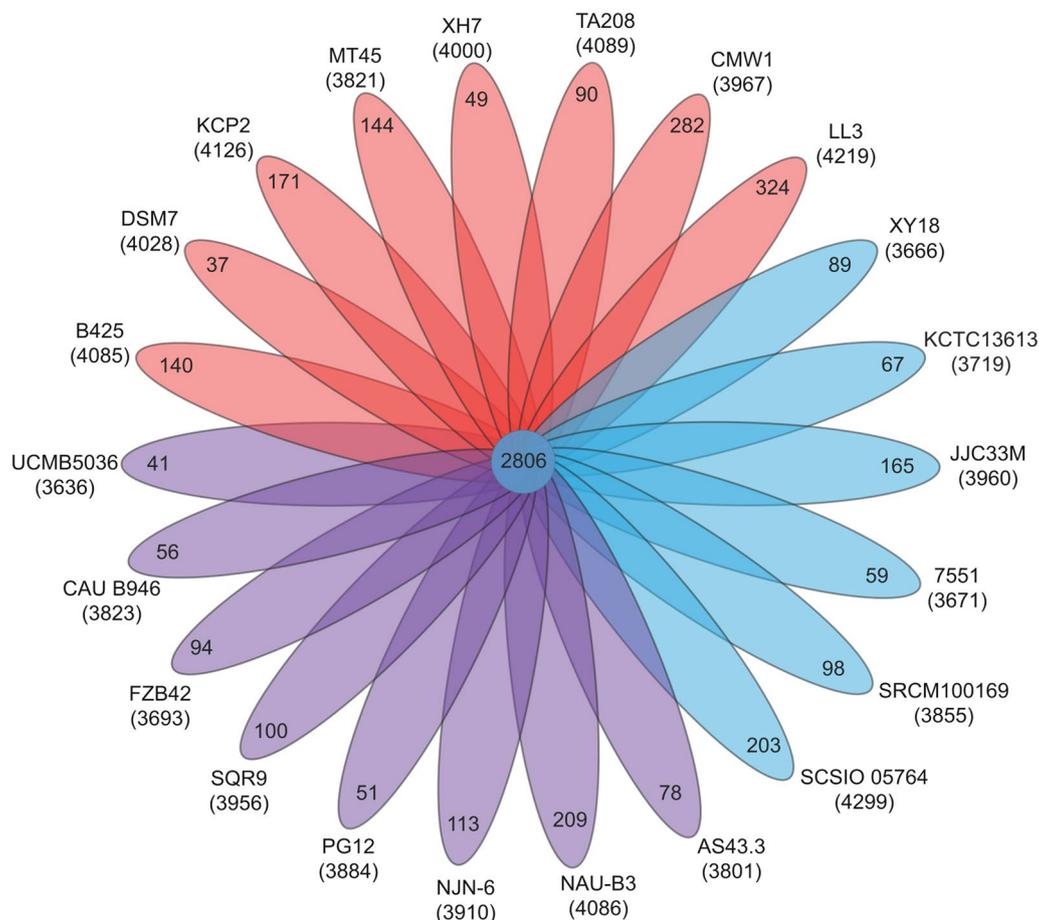


Fig. 3 Genomic diversity of the *B. amyloliquefaciens* group strains. Each strain is represented by a colored oval: *B. velezensis* (purple), *B. amyloliquefaciens* (red) and *B. siamensis* (blue). The number of orthologous coding sequences (CDS) shared by all strains (the core genome) is in the center. Overlapping regions show the number of CDSs conserved only within the specified genomes. Numbers in non-overlapping portions of each oval indicate the number of CDSs unique to each strain. The total number of protein-coding genes within each genome is listed below the strain name

The *fenA* mutant exhibited decreased swarming motility, biofilm formation and antimicrobial activities

In *B. subtilis*, swarming motility and biofilm formation are both considered as important phenotypic features favoring lipopeptide secretion (Allard-Massicotte et al. 2016; Fan et al. 2017b). To investigate if the biosynthesis of fengycin also affects these features, we produced a $\Delta fenA$ mutant from the parental strain PG12. Compared with the wild-type PG12 strain, the $\Delta fenA$ mutant exhibited a substantial reduction in its swarming motility (Fig. 5a, d). PG12 formed colonies with dense wrinkles and compact structure on MSgg plate, and a thick and wrinkled floating pellicle in MSgg broth. In contrast, the $\Delta fenA$ mutant formed a flat colony on MSgg plate and a thin and less wrinkled floating layer in MSgg broth (Fig. 5b, c, e). Thus, deletion of the *fenA* gene significantly influenced the biofilm formation and motility of PG12. In

addition, the $\Delta fenA$ mutant showed a decreased ability to inhibit the growth of plant pathogens, such as *Colletotrichum* sp. (strawberry) and *Fusarium moniforme* (Fig. 5f and Additional file 1: Figure S11).

Discussion

In our previous study, *B. velezensis* PG12 was isolated from apple fruit and showed high antifungal activity, which indicates that this bacterial strain has the potential to be used to control fungal pathogens infecting crop plants (Chen et al. 2016). In the present study, we displayed that both *B. velezensis* strains including PG12 and *B. siamensis* strains contain three secondary metabolism gene clusters related to the production of lipopeptides, such as surfactin, iturin and fengycin. Identification of this large spectrum of antibiotic-biosynthetic genes in the genome of PG12 to some extent explains its

Table 2 Comparison of COG assignments between the *B. amyloliquefaciens* group strains

Functional categories	Core genomes			p-value		
	Bv	Ba	Bs	Bv vs Ba	Bv vs Bs	Ba vs Bs
C: Energy production and conversion	163	157	163	0.9541	0.8195	0.7734
D: Cell cycle control, cell division, chromosome partitioning	31	30	29	1	0.8974	0.8967
E: Amino acid transport and metabolism	242	240	232	0.8488	0.8855	0.7361
F: Nucleotide transport and metabolism	78	73	78	0.8689	0.8713	0.7413
G: Carbohydrate transport and metabolism	168	157	148	0.7749	0.417	0.6372
H: Coenzyme transport and metabolism	166	164	163	0.9096	0.9547	0.9545
I: Lipid transport and metabolism	71	72	72	0.7999	0.8654	1
J: Translation, ribosomal structure and biogenesis	159	148	164	0.7248	0.6463	0.4141
K: Transcription	184	177	179	0.9566	1	0.9563
L: Replication, recombination and repair	100	98	100	1	0.8855	0.9423
M: Cell wall/membrane/envelope biogenesis	160	161	158	0.863	0.9081	0.8626
N: Cell motility	21	20	21	1	1	1
O: Posttranslational modification, protein turnover, chaperones	89	86	88	1	0.9392	0.9386
P: Inorganic ion transport and metabolism	145	148	143	0.7185	0.9517	0.7634
Q: Secondary metabolite biosynthesis, transport and catabolism	59	36	37	0.02966	0.03925	1
R: General function prediction only	300	301	297	0.7301	0.8627	0.8627
S: Function unknown	294	290	289	0.8614	0.9303	0.965
T: Signal transduction mechanism	80	79	81	0.9359	0.8108	0.936
U: Intracellular trafficking, secretion and vesicular transport	26	26	26	1	1	1
V: Defense mechanism	46	44	46	1	0.9166	0.9154

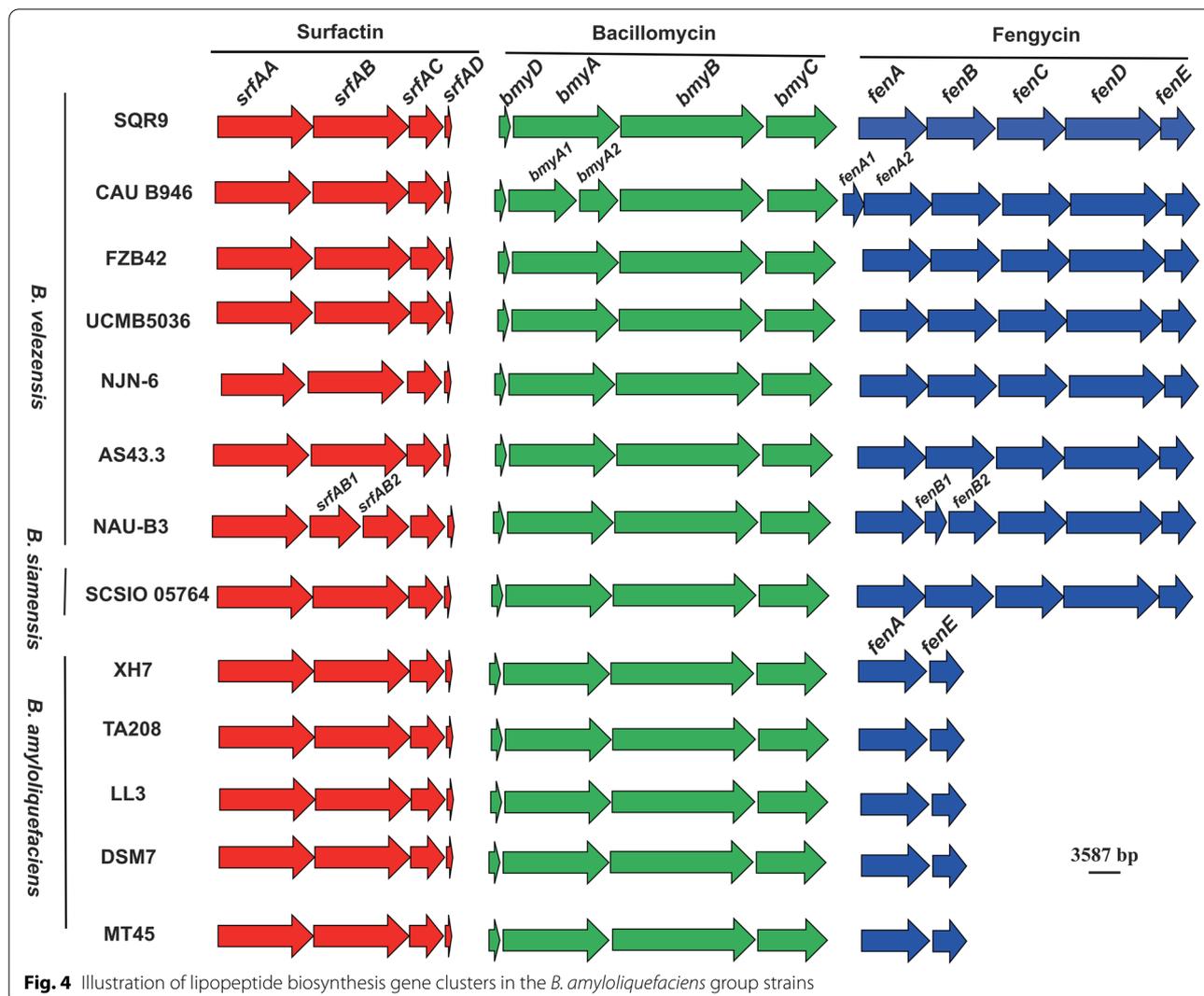
Bv, Ba and Bs represent *B. velezensis*, *B. amyloliquefaciens* and *B. siamensis*, respectively

high competitive ability against other microorganisms. However, it is still not known whether this biocontrol capacity is expressed in the field under different environmental conditions, as biocontrol effectiveness can be critically influenced by the resident microbial communities (Goudjal et al. 2014). Therefore, further field tests on the biocontrol effect of PG12 need to be performed in the future.

The species in the genus of *Bacillus* are both taxonomically and metabolically diverse, and the taxonomic status of some *Bacillus* strains constantly brings confusion to researchers (Hamdache et al. 2013). *B. subtilis*, *B. amyloliquefaciens* and *B. pumilus*, used in the agro-food industry, belong to a group of phylogenetically homogeneous species. It is therefore difficult to distinguish these species according to microbiological phenotypic characteristics. Moreover, phylogenetic analysis based on the 16S rRNA gene also fails to differentiate these species due to the highly conserved nature of the gene (Rooney et al. 2009). The *B. cereus* group comprises eleven closely related species, but the phylogenetic and taxonomic relationships within this group are still debated (Liu et al. 2015). With the development of sequencing technology, more and more genomic data of *Bacillus* strains are released. Construction of phylogenetic trees based on genome sequence will help us to correct errors and

improve the taxonomy quality of *Bacillus* strains. For instance, the biocontrol strains that were ever classified as *B. subtilis* based on 16S rRNA and metabolic profiles were reclassified as *B. amyloliquefaciens* based on their genome sequence (Magno-Perez-Bryan et al. 2015). On this count, we used 860 single-copy core genes to obtain a robust phylogenetic analysis of the *B. amyloliquefaciens* group strains in this study. The result showed that the tree can be divided into three clusters (Additional file 1: Figure S4). A heat-map analysis based on ANI values also support to divide this group into three subgroups (Fig. 2). With genome sequence data, we can realize accurate identification and taxonomy of *Bacillus* isolates, and further reveal their distinct features at the molecular level.

The *B. amyloliquefaciens* group harbors a wide range of ecotypes, including soil borne *B. amyloliquefaciens* and plant-associated *B. siamensis* and *B. velezensis* (Reva et al. 2004; Fan et al. 2017a). The most important biocontrol and plant-growth-promoting strains were found in *B. velezensis*, and have been successfully used in agriculture (Borriss 2011). There are significant differences in genome sequences between plant-associated and non-plant-associated strains. Rueckert et al. (2011) performed a genome comparison between FZB42 (a type strain of plant-associated *B. velezensis*) and DSM7 (a type strain of non-plant-associated *B. amyloliquefaciens*)



and found significant differences in their genomic sequences (Rueckert et al. 2011). COG function classification and manual curation of the core genomes of the *B. amyloliquefaciens* group strains showed that the genes involved in the biosynthesis of secondary metabolism were more abundant in *B. velezensis* (Table 2). Fan et al. (2017a) also found that the gene cluster involved in macrolactin synthesis is present in *B. velezensis*, but not in *B. siamensis* nor in *B. amyloliquefaciens*. In addition, the number of metabolites produced by *B. velezensis* is higher than those by *B. siamensis* and *B. amyloliquefaciens*. As a consequence, the strains harboring more antibiotic-synthesizing genes have an increased ability to limit the colonization of crop plants by fungal pathogens through the biosynthesis of antibiotics.

Horizontal gene transfer is recognized as an important factor in driving evolution and differentiation of different organisms. Three of the five genes pertaining to fengycin

biosynthesis gene cluster (*fenBCD*) were not found in *B. amyloliquefaciens*, whereas were identified in *B. siamensis* and *B. velezensis*, with the possibility to acquire them from *Paenibacillus* spp. by HGT. The $\Delta fenA$ mutant of PG12 displayed a significantly decreased antifungal activity, swarming motility and biofilm formation (Fig. 5 and Additional file 1: Figure S11), confirming the key role of this type of lipopeptide in maintaining bacterial antagonistic ability. It also demonstrated that fengycin as an antagonistic compound plays an important role in bio-control activities (Liu et al. 2021; Xiao et al. 2021). Additionally, fengycin was reported as an elicitor produced by *B. velezensis* SQR9 and *B. subtilis* GLB191, and it is able to induce plant immunity by triggering induced systemic resistance in *Arabidopsis* and grape plants in response to pathogen infection (Wu et al. 2018; Li et al. 2019). Therefore, fengycin-producing strain PG12 is a very promising candidate for the biological control of fungal diseases

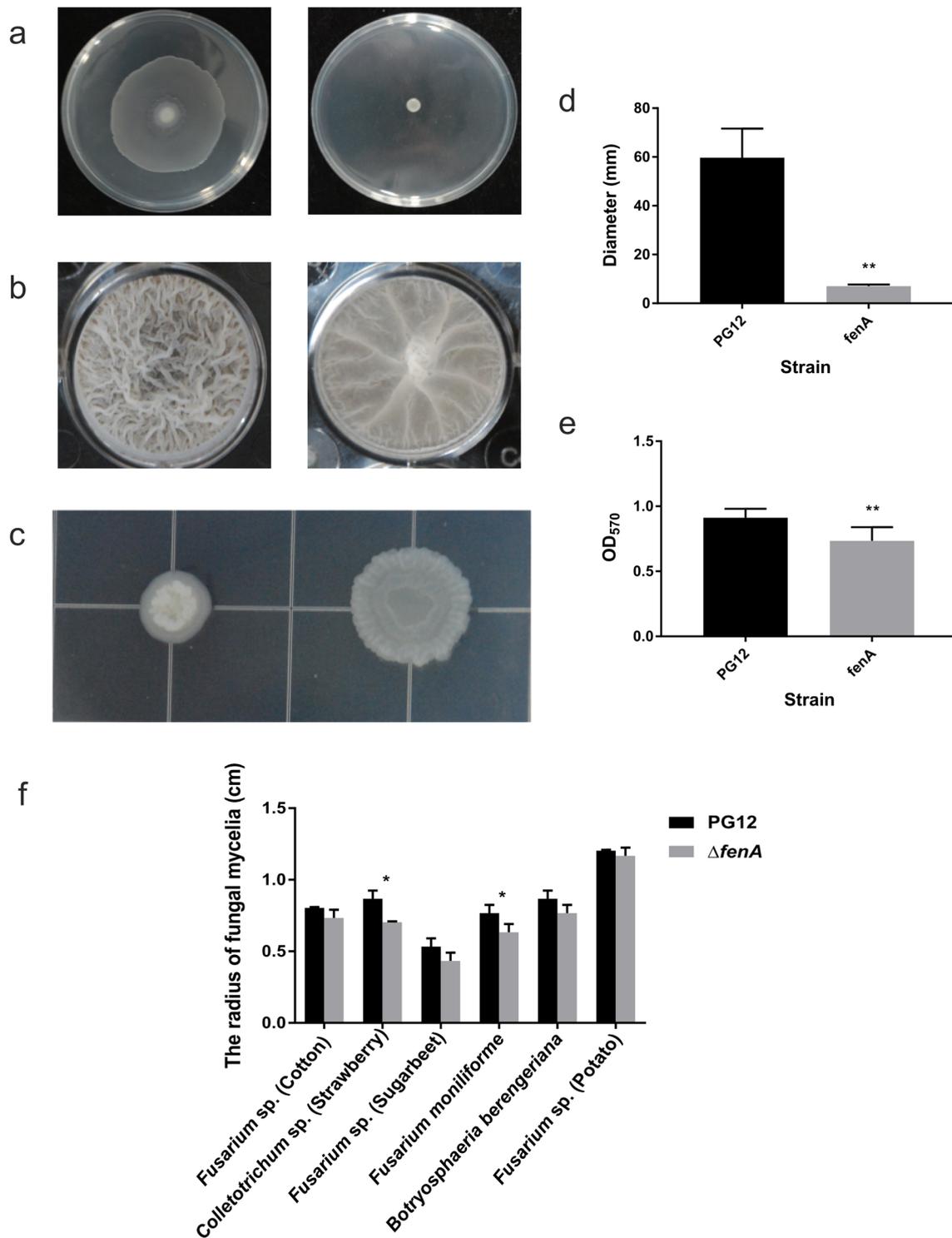


Fig. 5 Comparative analysis of the biocontrol properties of *B. velezensis* PG12 and $\Delta fenA$. **a** The swarming ability of PG12 and $\Delta fenA$. **b** Pellicle formation on MSgg liquid media. **c** Colony morphology on MSgg media. **d** Quantitative analysis of swarming ability. **e** Quantitative analysis of biofilm formation. **f** Measurement of the inhibition zones of PG12 and $\Delta fenA$. Data are an average of three replicates \pm the standard deviation of three independent experiments. Bars represent standard deviations. ** and * represent $P < 0.01$ and $P < 0.05$ (t-test), respectively

in crop plants. Further studies will also need to characterize the role of the *fenBCD* genes and the regulation mechanism of fengycin biosynthesis gene clusters. Altogether, our results promote comparative genomic studies of *Bacillus* strains, and meanwhile allow to gain insights into the underlying mechanistic basis of their biocontrol activity.

Conclusion

The biocontrol strain PG12 was identified as *B. velezensis* based on its genomic sequence. Three lipopeptides gene clusters were identified in the genome of PG12. Furthermore, fengycin was found to contribute to the biocontrol activity of PG12 through maintaining its colonization properties and meanwhile inhibiting pathogen growth. Importantly, genes related to secondary metabolism are more abundant in the core genome of *B. velezensis*. Moreover, *B. siamensis* and *B. velezensis* probably acquired the fengycin biosynthesis cluster gene through HGT events during their evolutionary process. Our study offers novel insights into the evolution of the *B. amyloliquefaciens* group strains and also theoretical basis for further application of beneficial strains of this group in agriculture.

Methods

Strains used in this study

B. velezensis PG12 was isolated from apple fruit (Chen et al. 2016). Genome data of other *Bacillus* strains used in this study were downloaded from the National Center for Biotechnology Information (NCBI) database. *Fusarium* sp., *Colletotrichum* sp., *Botryosphaeria berengeriana* and *Fusarium moniliforme* were provided by the Department of Plant Pathology of the China Agricultural University. All these fungal strains were stored in paraffin at 4°C before use.

Genome sequencing, assembly, and annotation

The draft sequence of the *B. velezensis* PG12 strain was produced using Illumina HiSeq 2500 paired-end (350 bp) sequencing platform by Berry Genomics (Berry Genomics, Beijing). The assembly was performed using SOAPdenovo (v.2.04), resulting in 22 scaffolds for PG12 (Luo et al. 2012; Zeng et al. 2019). Predictions of protein-coding genes were implemented in Prokka (v. 1.11) (Seemann 2014). Functional annotation was carried out using the Basic Local Alignment Search Tool (BLAST) against the Cluster of Orthologous Groups of proteins (COG), NCBI nr protein database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database and InterPro database (Zeng et al., 2018). Ordering of contigs of PG12 was achieved using the program Mauve (Rissman et al. 2009). The genome sequence of *B. velezensis* CAU B946 was

used as a reference sequence. The final annotated chromosome was plotted using CIRCOS software (Krzywinski et al. 2009). The secondary metabolite gene clusters were examined using the antiSMASH v4.0.0rc1 program (Blin et al. 2017).

Phylogenomic analyses

A ML phylogenetic tree of *Bacillus* species was constructed based on 624 or 860 single-copy core proteins shared by *Bacillus* and *Paenibacillus polymyxa* M1 according to the following methods: (1) multiple alignment of amino acid sequences by mafft (v. 7.310) (Katoh and Standley 2013); (2) conserved blocks from multiple alignment using Gblocks (Castresana 2000); (3) construction of ML tree using RAxML (v. 8.2.10) software and PROTGAMMALGX model with 100 bootstrap replicates (Zeng et al. 2018). The tree was displayed using molecular evolutionary genetic analysis (MEGA). Then, ANI value was calculated using Jspecies software with MUMmer (NUCmer) alignment (Richter and Rossello-Mora 2009). Single gene alignments were performed with MEGA 5. The neighbor-joining trees were constructed using the same software with 1,000 bootstrap replicates (Kumar et al. 2008).

For genome similarity assessment, genome comparisons were performed by using BRIG program (v. 0.95). The analysis was done by using the blastn option. The genome sequence of *B. velezensis* FZB42 was used as a reference sequence, whereas the genome sequences of strains (CAU B946, PG12 and UCMB5036) were used as query sequences.

Core and pan-genome analyses

Identification of the core and pan-genomes of the 22 *B. amyloliquefaciens* group strains were performed with the PGAP pipeline (Zhao et al. 2012). The protein similarity method from PGAP was used to detect a set of core orthologs which were clustered using at least 50% protein sequence identity to each other and 50% overlap with the longest sequence with an e-value threshold of 1e-5. The core and strain-specific genomes was extracted from the orthologous table by using a custom Perl script (Zeng et al. 2018). Functional annotation of core genes of *B. velezensis*, *B. amyloliquefaciens* and *B. siamensis* was performed using the COG database. The core and pan-genomes, as well as estimated respective sizes and trajectories, were assessed using models and regression algorithms proposed by Tettelin et al. (2005, 2008). Both the core and pan-genome were visualized through PanGP (v1.0.1) using default parameters to generate distribution plots of (i) total genes, and (ii) conserved genes found upon progressive sampling of “n” genomes (Zhao et al. 2014).

Phenotypic characterization of *B. velezensis* PG12

Cultures and conditions

B. velezensis PG12 was grown at 37°C in Luria–Bertani (LB) broth or on solid LB medium (Chen et al. 2016). For assays of biofilm formation, MSgg medium was used. The recipe for MSgg is as follows: 5 mM potassium phosphate (pH 7.0), 100 mM MOPS (morpholine propane sulfonic acid) (pH 7.0), 2 mM MgCl₂, 700 mM CaCl₂, 50 mM MnCl₂, 50 mM FeCl₃, 1 mM ZnCl₂, 2 mM thiamine, 0.5% glycerol, 0.5% glutamic acid, 50 mg/mL tryptophan, 50 mg/mL threonine, and 50 mg/mL phenylalanine (Branda et al. 2001; Fan et al. 2017b). When necessary, the antibiotic kanamycin was added to a final concentration of 50 µg/mL.

Construction of *fenA* deletion mutant

The *fenA* deletion mutant was generated from PG12 by homologous recombination. Primers UF/UR and DF/DR were used to amplify the upstream region (1,585 bp) and downstream region (1,500 bp) of *fenA*, respectively. The kanamycin resistance gene was amplified from plasmid pUB110 using primer pair Kan-F/Kan-R. To construct the knockout vector, the upstream, downstream and resistance gene region were ligated using gene splicing by overlap extension (4,040 bp). Then, the target product was ligated into pMD19-T to generate the knockout plasmid pTFENA, which was subsequently transferred into *B. velezensis* PG12 by electroporation. Kanamycin-sensitive clones were isolated, and the mutants were identified by PCR with primer pair *fenA*-out-F/*fenA*-out-R and confirmed by Sanger sequencing (Additional file 2: Table S2, S3).

Analysis of biofilm formation

Biofilm formation was monitored in MSgg medium. PG12 was grown in LB medium at 37°C overnight. Then, 4 µL of cell culture was added to each well of a 12-well microplate, with each well containing 4 mL of MSgg medium, and incubated statically at 28°C for up to 60 h. To quantify biofilm formation, liquid cultures beneath the biofilm were carefully drawn off and the biofilm in each well was washed three times with 1 mL of sterile saline and fixed with 2 mL of 99% (v/v) methanol for 15 min, followed by air-drying. The dried biofilms were stained with 2 mL 0.1% crystal violet (CV) for 10 min. After excess CV was removed, the dried biofilms were dissolved with 5 mL of 33% (v/v) glacial acetic acid and then were diluted to 10 times for OD₅₇₀ detection (Weng et al. 2013; Fan et al. 2017b). For colony architecture, 2 µL of cell culture was spotted onto the surface of a Petri plate containing MSgg medium and 1.5% agar, and incubated at 28°C for 60 h.

Assays of swarming motility

The swarming motility of PG12 was tested using standard protocol with minor modification. PG12 cultures in LB liquid medium were prepared with shaking (200 rpm) at 37°C to an OD₆₀₀ of 0.9, one mL of cells were collected by centrifugation at 6,000 g for 5 min, washed with phosphate-buffered saline (PBS), and resuspended in 100 µL PBS. Three µL of cell suspension was spotted on the center of a Petri plate containing LB medium and 0.7% agar, and the plates were incubated at room temperature for 9 h to allow cell growth in order to clearly visualize the swarming zone (Chen et al. 2012; Fan et al. 2017b). The plates were air-dried for 2 h in a laminar flow hood, and then the diameter of the swarming zone was measured.

In vitro antagonism test

The antagonistic activities of the wild-type and mutant strains against fungal pathogens were assessed as follows. In the plate confrontation assay, the fungi were cultured on potato-dextrose-agar plates (PDA) at 28°C for 3–5 days. A 5-mm-diameter mycelial block was then cut from the margin of an actively growing fungal culture and transferred into the center of a fresh PDA plate. After 1 day of incubation, 2 µL of overnight bacterial culture grown in LB medium was spotted on the PDA plate 2.5 cm away from the center where the mycelial block was placed. The antifungal activities of the examined bacterial strains were evaluated by measuring the radial extension of mycelium after 5–7 days of incubation at 28°C.

Abbreviations

ANI: Average nucleotide identity; Blast: Basic local alignment search tool; COG: Cluster of Orthologous Group of proteins; HGT: Horizontal gene transfer; KEGG: Kyoto Encyclopedia of Genes and Genomes; LB: Luria–Bertani; ML: Maximum likelihood; NCBI: National Center for Biotechnology Information; PBS: Phosphate-buffered saline; PDA: Potato dextrose agar.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42483-021-00103-z>.

Additional file 1: Figure S1. Genome map of *B. velezensis* PG12. Each circle has a different genome information, and circles from outside to inside: (1) scale marks (unit: Mb), (2, 3) protein-coding genes on the forward and reverse strands, respectively (color-coded by the functional categories), (4, 5) rRNA (blue) and tRNA (red) on the forward and reverse strands, respectively, (6) GC content (positive: red; negative: blue), and (7) GC skew (above average: aquamarine; below average: orange). **Figure S2.** Genome comparisons of FZB42 with other three *B. velezensis* strains. The inner circle represents the complete genome sequence of FZB42, and the shade of each color shows the similarities between the strains (CAU B946, UCMB5036 or PG12) and FZB42. **Figure S3.** Venn diagram showing shared orthologous protein clusters in the genomes of PG12 and its closely related *B. velezensis* strains CAU B946, FZB42 and UCMB5036. Each strain is represented by an oval. The number of orthologous protein-coding genes shared by all strains is in the center. Overlapping regions show the number of coding sequences conserved only within the specified genomes. Numbers in the non-overlapping portions of each oval show

the number of CDS unique to each strain. The total number of protein-coding genes within each genome is listed below the strain name. **Figure S4.** Phylogenetic tree showing the relationship of the *B. amyloliquefaciens* group strains. The tree is based on the concatenated alignments of 860 single core genomes. *Paenibacillus polymyxa* M1 was used as an outgroup. **Figure S5.** Pan- and core-genome plots of the *B. amyloliquefaciens* group strains. **a** *B. velezensis*; **b** *B. siamensis*; **c** *B. amyloliquefaciens*. The lines in yellow and purple represent the pan- and core-genomes, respectively. **Figure S6.** Organization of lipopeptide gene clusters in draft genome sequence of the *B. amyloliquefaciens* group strains. Different genes are marked with different colors. **Figure S7.** Comparative analysis of G + C contents between genomes and the fengycin gene clusters (*fenB*, *fenC* and *fenD*) in the *B. amyloliquefaciens* group strains. **Figure S8.** Neighbor-joining phylogenetic tree constructed based on the protein sequences of *fenB*. **Figure S9.** Neighbor-joining phylogenetic tree constructed based on the protein sequence of *fenC*. **Figure S10.** Neighbor-joining phylogenetic tree constructed based on the protein sequence of *fenD*. **Figure S11.** The antifungal activity of PG12 and the $\Delta fenA$ mutant. A mycelium plug of fungal pathogen was inoculated onto the center of a PDA plate, and then 2 μ L of PG12 and $\Delta fenA$ cultures in LB medium were added at a distance of 2.5 cm from the rim of the fungal plug. Clear inhibition zones were observed after incubation for 5–7 days.

Additional file 2: Table S1 Average nucleotide identity pairwise comparisons among the *B. amyloliquefaciens* group strains. **Table S2.** Strains and plasmids used in this study. **Table S3.** Primers used in this study.

Acknowledgements

We are thankful to Prof. Francis Martin of Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University for help with revising the manuscript.

Authors' contributions

QZ performed the experiments and wrote the manuscript. QZ, XG and YL analyzed the data. XC, PY and GH prepared the figures and tables. QW and JX helped to design the experiments and revise the manuscript. All authors read and approved the final manuscript.

Funding

The research was financially supported by the National Natural Science Foundation of China (3162074), the National Key R&D Program of China (2017YFD0201106; 2019YFD1002003) and the Beijing Municipal Natural Science Foundation (6172018).

Availability of data and materials

Not applicable.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Plant Pathology, MOA Key Lab of Pest Monitoring and Green Management, College of Plant Protection, China Agricultural University, Beijing 100193, China. ²Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing 100083, China. ³Institute of Citrus Science, Yichang 443000, China.

Received: 24 June 2021 Accepted: 26 October 2021
Published online: 09 November 2021

References

- Allard-Massicotte R, Tessier L, Lecuyer F, Lakshmanan V, Lucier J, Garneau D, et al. *Bacillus subtilis* early colonization of *Arabidopsis thaliana* roots involves multiple chemotaxis receptors. *Mbio*. 2016;7(6):e01664-16. <https://doi.org/10.1128/mBio.01664-16>.
- Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res*. 2017;45:W36-41. <https://doi.org/10.1093/nar/gkx319>.
- Borriis R. Use of plant-associated *Bacillus* strains as biofertilizers and biocontrol agents. In: Maheshwari D, editor. *Bacteria in Agrobiology: Plant Growth Responses*. Heidelberg: Springer. https://doi.org/10.1007/978-3-642-20332-9_3.
- Branda SS, Gonzalez-Pastor JE, Ben-Yehuda S, Losick R, Kolter R. Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci USA*. 2001;98(20):11621-6. <https://doi.org/10.1073/pnas.191384198>.
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 2000;17(4):540-52. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>.
- Chen Y, Chai Y, Gu J, Losick R. Evidence for cyclic Di-GMP-mediated signaling in *Bacillus subtilis*. *J Bacteriol*. 2012;194(18):5080-90. <https://doi.org/10.1128/JB.01092-12>.
- Chen X, Zhang Y, Fu X, Li Y, Wang Q. Isolation and characterization of *Bacillus amyloliquefaciens* PG12 for the biological control of apple ring rot. *Post-harvest Biol Technol*. 2016;115:113-21. <https://doi.org/10.1016/j.postharvbio.2015.12.021>.
- Chun BH, Kim K, Jeong SE, Jeon CO. Genomic and metabolic features of the *Bacillus amyloliquefaciens* group-*B. amyloliquefaciens*, *B. velezensis*, and *B. siamensis*- revealed by pan-genome analysis. *Food Microbiol*. 2019;77:146-57. <https://doi.org/10.1016/j.fm.2018.09.001>.
- Cochrane SA, Vederas JC. Lipopeptides from *Bacillus* and *Paenibacillus* spp.: a gold mine of antibiotic candidates. *Med Res Rev*. 2019;36(1):4-31. <https://doi.org/10.1002/med.21321>.
- Compant S, Duffy B, Nowak J, Clement C, Barka EA. Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl Environ Microbiol*. 2005;71(9):4951-9. <https://doi.org/10.1128/AEM.71.9.4951-4959.2005>.
- Fan B, Blom J, Klenk H, Borriis R. *Bacillus amyloliquefaciens*, *Bacillus velezensis*, and *Bacillus siamensis* form an "operational group *B. amyloliquefaciens*" within the *B. subtilis* species complex. *Front Microbiol*. 2017a;8:22. <https://doi.org/10.3389/fmicb.2017.00022>.
- Fan H, Zhang Z, Li Y, Zhang X, Duan Y, Wang Q. Biocontrol of bacterial fruit blotch by *Bacillus subtilis* 9407 via surfactin-mediated antibacterial activity and colonization. *Front Microbiol*. 2017b;8:1973. <https://doi.org/10.3389/fmicb.2017.01973>.
- Fan B, Wang C, Song X, Ding X, Wu L, Wu H, et al. *Bacillus velezensis* FZB42 in 2018: The Gram-positive model strain for plant growth promotion and biocontrol. *Front Microbiol*. 2019;10:1279. <https://doi.org/10.3389/fmicb.2019.01279>.
- Feng L, Ma T, Zhang J, Xu F, Shi L. Draft genome sequence of *Bacillus subtilis* QH-1, a chromium-reducing bacterial strain isolated in Qinghai province. *China Genome Announce*. 2014;2(2):e00182-e214. <https://doi.org/10.1128/genomeA.00182-14>.
- Fira D, Dimkic I, Beric T, Lozo J, Stankovic S. Biological control of plant pathogens by *Bacillus* species. *J Biotechnol*. 2018;285:44-55. <https://doi.org/10.1016/j.jbiotec.2018.07.044>.
- Gao T, Ding M, Yang C, Fan H, Chai Y, Li Y. The phosphotransferase system gene *ptsH* plays an important role in MnSOD production, biofilm formation, swarming motility, and root colonization in *Bacillus cereus* 905. *Res Microbiol*. 2019;170(2):86-96. <https://doi.org/10.1016/j.resmic.2018.10.002>.
- Goudjal Y, Toumatia O, Yekkour A, Sabaou N, Mathieu F, Zitouni A. Biocontrol of *Rhizoctonia solani* damping-off and promotion of tomato plant growth by endophytic actinomycetes isolated from native plants of Algerian Sahara. *Microbiol Res*. 2014;169(1):59-65. <https://doi.org/10.1016/j.micres.2013.06.014>.
- Hamdache A, Azarken R, Lamarti A, Aleu J, Collado IG. Comparative genome analysis of *Bacillus* spp. and its relationship with bioactive nonribosomal peptide production. *Phytochem Rev*. 2013;12(4):685-716. <https://doi.org/10.1007/s11101-013-9278-4>.
- Jayapala N, Mallikarjunaiah NH, Puttaswamy H, Gavirangappa H, Ramachandrappa NS. Rhizobacteria *Bacillus* spp. induce resistance against

- anthracnose disease in chili (*Capsicum annum* L.) through activating host defense response. *Egypt J Biol Pest Control*. 2019;29:45. <https://doi.org/10.1186/s41938-019-0148-2>.
- Katoh K, Standley DM. MAFFT: multiple sequence alignment software Version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80. <https://doi.org/10.1093/molbev/mst010>.
- Kim B, Lee S, Ahn J, Song J, Kim W, Weon H. Complete genome sequence of *Bacillus amyloliquefaciens* subsp. *plantarum* CC178, a phyllosphere bacterium antagonistic to plant pathogenic fungi. *Genome Announce*. 2015;3(1):e1314-68. <https://doi.org/10.1128/genomeA.01368-14>.
- Koumoutsis A, Chen XH, Henne A, Liesegang H, Hitzeroth G, Franke P, et al. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. *J Bacteriol*. 2014;186(4):1084–96. <https://doi.org/10.1128/JB.186.4.1084-1096.2004>.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639–45. <https://doi.org/10.1101/gr.092759.109>.
- Kumar S, Nei M, Dudley J, Tamura K. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings Bioinform*. 2008;9(4):299–306. <https://doi.org/10.1093/bib/bbn017>.
- Li Y, Heloir M, Zhang X, Geissler M, Trouvelot S, Jacquens L, et al. Surfactin and fengycin contribute to the protection of a *Bacillus subtilis* strain against grape downy mildew by both direct effect and defence stimulation. *Mol Plant Pathol*. 2019;20(8):1037–50. <https://doi.org/10.1111/mpp.12809>.
- Li C, Cheng P, Zheng L, Li Y, Chen Y, Wen S, et al. Comparative genomics analysis of two banana Fusarium wilt biocontrol endophytes *Bacillus subtilis* R31 and TR21 provides insights into their differences on phyto-beneficial trait. *Genomics*. 2021;113(3):900–9. <https://doi.org/10.1016/j.jygeno.2021.02.006>.
- Liu Y, Lai Q, Goeker M, Meier-Kolthoff JP, Wang M, Sun Y, et al. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep*. 2015;5:14082. <https://doi.org/10.1038/srep14082>.
- Liu H, Zeng Q, Yalimaimaiti N, Wang W, Zhang R, Yao J. Comprehensive genomic analysis of *Bacillus velezensis* AL7 reveals its biocontrol potential against *Verticillium* wilt of cotton. *Mol Genet Genomics*. 2021. <https://doi.org/10.1007/s00438-021-01816-8>.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*. 2012. <https://doi.org/10.1186/s13742-015-0069-2>.
- Luo C, Zhou H, Zou J, Wang X, Zhang R, Xiang Y, et al. Bacillomycin L and surfactin contribute synergistically to the phenotypic features of *Bacillus subtilis* 916 and the biocontrol of rice sheath blight induced by *Rhizoctonia solani*. *Appl Microbiol Biot*. 2014;99(4):1897–910. <https://doi.org/10.1007/s00253-014-6195-4>.
- Magno-Perez-Bryan MC, Martinez-Garcia PM, Hierrezuelo J, Rodriguez-Palenzuela P, Arrebola E, Ramos C, et al. Comparative genomics within the *Bacillus* genus reveal the singularities of two robust *Bacillus amyloliquefaciens* biocontrol strains. *Mol Plant-Microbe Interact*. 2015;28(10):1102–16. <https://doi.org/10.1094/MPMI-02-15-0023-R>.
- Paterson J, Jahanshah G, Li Y, Wang Q, Mehnaz S, Gross H. The contribution of genome mining strategies to the understanding of active principles of PGPR strains. *FEMS Microbiol Ecol*. 2017;93:fw249. <https://doi.org/10.1093/femsec/fw249>.
- Prashar P, Kapoor N, Sachdeva S. Biocontrol of plant pathogens using plant growth promoting bacteria. In: Lichtfouse editor. *Sustainable agriculture Reviews*. Sustainable Agricultural Reviews. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-5961-9_10.
- Qiao J, Wu H, Huo R, Gao X, Borriss R. Stimulation of plant growth and biocontrol by *Bacillus amyloliquefaciens* subsp. *plantarum* FZB42 engineered for improved action. *Chem Biol Technol Agric*. 2014;1:12. <https://doi.org/10.1186/s40538-014-0012-2>.
- Rahman MM. Insecticide substitutes for DDT to control mosquitoes may be causes of several diseases. *Environ Sci Pollut Res*. 2013;20(4):2064–9. <https://doi.org/10.1007/s11356-012-1145-0>.
- Reva ON, Dixelius C, Meijer J, Priest FG. Taxonomic characterization and plant colonizing abilities of some bacteria related to *Bacillus amyloliquefaciens* and *Bacillus subtilis*. *FEMS Microbiol Ecol*. 2004;48(2):249–59. <https://doi.org/10.1016/j.femsec.2004.02.003>.
- Richter M, Rossello-Mora R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA*. 2009;106(45):19126–31. <https://doi.org/10.1073/pnas.0906412106>.
- Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics*. 2009;25(16):2071–3. <https://doi.org/10.1093/bioinformatics/btp356>.
- Rooney AP, Price NPJ, Ehrhardt C, Swezey JL, Bannan JD. Phylogeny and molecular taxonomy of the *Bacillus subtilis* species complex and description of *Bacillus subtilis* subsp. *inaquosorum* subsp. nov. *Int J Syst Evol Microbiol*. 2009;59(10):2429–36. <https://doi.org/10.1099/ijs.0.009126-0>.
- Rueckert C, Blom J, Chen X, Reva O, Borriss R. Genome sequence of *B. amyloliquefaciens* type strain DSM7T reveals differences to plant-associated *B. amyloliquefaciens* FZB42. *J Biotechnol*. 2011;155(1):78–85. <https://doi.org/10.1016/j.jbiotec.2011.01.006>.
- Schäfer T, Adams T. The importance of microbiology in sustainable agriculture. In: Lugtenberg B. editor. *Principles of plant-microbe interactions*. Switzerland: Springer. https://doi.org/10.1007/978-3-319-08575-3_2.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9. <https://doi.org/10.1093/bioinformatics/btu153>.
- Siefert JL, Larios-Sanz M, Nakamura LK, Slepceky RA, Paul JH, Moore E, et al. Phylogeny of marine *Bacillus* isolates from the Gulf of Mexico. *Curr Microbiol*. 2000;41(2):84–8. <https://doi.org/10.1007/s002840010098>.
- Tettelin H, Massignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome." *Proc Natl Acad Sci USA*. 2005;102(39):13950–5. <https://doi.org/10.1073/pnas.0506758102>.
- Tettelin H, Riley D, Cattuto C, Medini D. Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol*. 2008;11(5):472–7. <https://doi.org/10.1016/j.mib.2008.09.006>.
- Weng J, Wang Y, Li J, Shen Q, Zhang R. Enhanced root colonization and biocontrol activity of *Bacillus amyloliquefaciens* SQR9 by *abrB* gene disruption. *Appl Microbiol Biot*. 2013;97(19):8823–30. <https://doi.org/10.1007/s00253-012-4572-4>.
- Wu L, Wu H, Qiao X, Gao X, Borriss R. Novel routes for improving biocontrol activity of *Bacillus* based bioinoculants. *Front Microbiol*. 2015;6:1395. <https://doi.org/10.3389/fmicb.2015.01395>.
- Wu G, Liu Y, Xu Y, Zhang G, Shen Q, Zhang R. Exploring elicitors of the beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 to induce plant systemic resistance and their interactions with plant signaling pathways. *Mol Plant-Microbe Interact*. 2018;31(5):560–7. <https://doi.org/10.1094/MPMI-11-17-0273-R>.
- Xiao J, Guo X, Qiao X, Zhang X, Chen X, Zhang D. Activity of fengycin and iturin A isolated from *Bacillus subtilis* Z-14 on *Gaeumannomyces graminis* var. *tritici* and soil microbial diversity. *Front Microbiol*. 2021;12:682. <https://doi.org/10.3389/fmicb.2021.682437>.
- Xie J, Du Z, Bai L, Tian C, Zhang Y, Xie J, et al. Comparative genomic analysis of N-2-fixing and non-N-2-fixing *Paenibacillus* spp.: organization, evolution and expression of the nitrogen fixation genes. *PLoS Genet*. 2014;10(3):e1004231. <https://doi.org/10.1371/journal.pgen.1004231>.
- Zeng Q, Xie J, Li Y, Gao T, Xu C, Wang Q. Comparative genomic and functional analyses of four sequenced *Bacillus cereus* genomes reveal conservation of genes relevant to plant-growth-promoting traits. *Sci Rep*. 2018;8:17009. <https://doi.org/10.1038/s41598-018-35300-y>.
- Zeng Q, Xie J, Li Y, Chen X, Wang Q. Draft genome sequence of an endophytic biocontrol bacterium, *Bacillus velezensis* PG12, isolated from apple fruit. *Microbiol Resour Announce*. 2019;8(41):e00468-e519. <https://doi.org/10.1128/MRA.00468-19>.
- Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. PGAP: pan-genomes analysis pipeline. *Bioinformatics*. 2012;28(3):416–8. <https://doi.org/10.1093/bioinformatics/btr655>.
- Zhao Y, Jia X, Yang J, Ling Y, Zhang Z, Yu J, et al. PanGP: a tool for quickly analyzing bacterial pan-genome profile. *Bioinformatics*. 2014;30(9):1297–9. <https://doi.org/10.1093/bioinformatics/btu017>.
- Zhao H, Shao D, Jiang C, Shi J, Li Q, Huang Q, et al. Biological activity of lipopeptides from *Bacillus*. *Appl Microbiol Biot*. 2017;101:5951–60. <https://doi.org/10.1007/s00253-017-8396-0>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.