Comparative genome analysis and mining of secondary metabolites of Paenibacillus polymyxa

Buqing Wang¹²,³, Huicai Cheng²,³, Wenjiang Qian¹², Wenya Zhao²,³, Cong Liang²,³, Chao Liu¹,², Guanhui Cui¹²,³, Hongwei Liu²,³* and Liping Zhang¹²,³*

¹Hebei University of Technology, Tianjing 300000, China
²Institute of Biology, Hebei Academy of Sciences, Shijiazhuang 050081, China
³Main Crops Disease of Microbial Control Engineering Technology Research Center in Hebei Province, Shijiazhuang 050081, China

(Received 9 October 2019, accepted 4 March 2020; J-STAGE Advance published date: 2 July 2020)

Paenibacillus polymyxa is a well-known Gram-positive biocontrol bacterium. It has been reported that many P. polymyxa strains can inhibit bacteria, fungi and other plant pathogens. Paenibacillus polymyxa employs a variety of mechanisms to promote plant growth, so it is necessary to understand the biocontrol ability of bacteria at the genome level. In the present study, thanks to the widespread availability of Paenibacillus genome data and the development of bioinformatics tools, we were able to analyze and mine the genomes of 43 P. polymyxa strains. The strain NCTC4744 was determined not to be P. polymyxa according to digital DNA–DNA hybridization and average nucleotide identity. By analysis of the pan-genome and the core genome, we found that the pan-genome of P. polymyxa was open and that there were 3,192 core genes. In a gene cluster analysis of secondary metabolites, 797 secondary metabolite gene clusters were found, of which 343 are not similar to known clusters and are expected to reveal a large number of new secondary metabolites. We also analyzed the plant growth-promoting genes that were mined and found, surprisingly, that these genes are highly conserved. The results of the present study not only reveal a large number of unknown potential secondary metabolite gene clusters in P. polymyxa, but also suggest that plant growth promotion characteristics are evolutionary adaptations of P. polymyxa to plant-related habitats.

Key words: comparative genome, core genome, Paenibacillus polymyxa, secondary metabolite

INTRODUCTION

Paenibacillus is a genus of facultative anaerobic Gram-positive bacteria that can produce stress-resistant spores and have unique physiological characteristics (McSpadden Gardener, 2004; Lal and Tabacchioni, 2009). Its members exist in various habitats and can secrete a variety of bioactive substances. They have a wide range of application prospects in medical, food, agriculture and industrial fields (Shishido et al., 1996; Guemouri-Athmani et al., 2000). In addition, Paenibacillus species, especially Paenibacillus polymyxa, are common in soil, and are used as model bacteria (Ash et al., 1993). Paenibacillus polymyxa is a plant growth-promoting rhizobacterium (PGPR) with a wide range of hosts. Plant-associated P. polymyxa strains colonize the plant rhizosphere, promote plant growth and suppress competing phytopathogens, and can ameliorate plant diseases caused by fungi, bacteria and nematodes (Beatty and Jensen, 2002; Niu et al., 2013; Castaneda-Alvarez et al., 2016). Because its phenotypic characteristics are similar to those of Bacillus, the species used to be known as Bacillus polymyxa, before it was classified into the Paenibacillus genus.

As a common biocontrol bacterium, P. polymyxa is useful for the implementation of biological control strategies for plant pathogens (Gurr and You, 2015). It colonizes the rhizosphere of various crops and effectively inhibits Phytophthora palmivora and Pythium aphanidermatum (Khan et al., 2008). In addition, strains colonizing the root tip form a biofilm, enhancing the stress resistance of plants (Timmusk et al., 2005). Paenibacillus polymyxa is widely used in biological organic fertilizers and biological pesticides (Son et al., 2010); it prevents patho-
 gens from invading plants by direct and indirect mechanisms. *Paenibacillus polymyxa* also employs a variety of mechanisms to promote plant growth, including the synthesis of auxin, mitogen and extracellular polysaccharide, as well as the enhancement of soil permeability, phosphorus dissolution and nitrogen fixation (Bloemberg and Lugtenberg, 2001; Anand and Chanway, 2013).

The term “pan-genome” refers to the collection of all genetic information of different strains of a certain species, which can be divided into three parts: the core genome (shared by all strains), the non-essential genome (shared by several but not all strains) and the unique genome (found in only one strain) (Medini et al., 2005). Although a large number of *P. polymyxa* genomes have been studied, to the best of our knowledge no analysis has been carried out on the complete *P. polymyxa* pan-genome (Xie et al., 2016). In the current genomic era, pan-genome analysis is an important tool for identifying functional genes of microbial genomes (Medini et al., 2008; Vernikos et al., 2015; Choi, 2019). Different from comparative genomics (interspecies) analysis, pan-genomic (intrasppecies) analysis can be used to predict the openness of the genome and to measure the conservation of functional genes. Moreover, when a new *P. polymyxa* genome sequence becomes available, the new *P. polymyxa* pan-genome size can be predicted (Yang et al., 2016). Traditional isolation and purification methods sometimes limit the identification of secondary metabolites of different strains. Some orphan gene clusters can be found by bioinformatics analysis of genome data, e.g., biosynthesis gene clusters with unknown secondary metabolites. Through the characterization of these gene clusters, new active substances can be identified (Singh and Sareen, 2014; Shi et al., 2019).

In this study, the genomic data of 43 *P. polymyxa* strains were downloaded from NCBI, and the complete genomes of 13 strains representing the current pan-genome were analyzed to identify the species boundary, characterize the database (Altschul, 1990), so as to mine and identify known active substances and potential unknown secondary metabolite synthesis genes clusters. Gene clusters for non-ribosomal peptides and polyketide compounds were analyzed using PRISM (Skinnider et al., 2017). Potential bacteriocins were mined and analyzed using the online software BAGEL 3 (van Heel et al., 2013).

**Mining and identification of secondary metabolite gene clusters** The *P. polymyxa* genome data (Supplementary Table S1) were downloaded, and gene clusters for secondary metabolites were excavated and analyzed using the online software antiSMASH (Blin et al., 2019). BLAST was used locally to compare the sequences of the excavated gene clusters with those in the database (Altschul, 1990), so as to mine and identify known active substances and potential unknown secondary metabolite synthesis genes clusters. Gene clusters for non-ribosomal peptides and polyketide compounds were analyzed using PRISM (Skinnider et al., 2017). Potential bacteriocins were mined and analyzed using the online software BAGEL 3 (van Heel et al., 2013).

**Pan-genomic analysis of *P. polymyxa*** The GenBank format data of 13 completely sequenced *P. polymyxa* strains (see Fig. 4 for their identification) were downloaded from NCBI and used as data input files. The genomes were analyzed using the Gene Family (GF) algorithm in the PGAP (Zhao et al., 2012) module of PGAweb. The output files were downloaded, the Orthologs_Cluster1.txt file was selected, and PanGP (Zhao et al., 2014) was used for data fitting. The algorithm was set to Traverse All; the other parameters were set to their default values.

**Digging for plant growth-promoting genes** Using the nitrogenase-encoding genes *nifS*, *nifHDK* and *budAB* in PGPR as bait, similar nitrogen fixation sequences were found in *Klebsiella*, *Bacillus*, *Rhizobium* and *Serratia* (Bruto et al., 2014). The nitrogenase biosynthesis pathway was successfully identified by integrating the genes *nifB*, *hesA*, *groRIS*, *groEL* and *nifE* from *P. polymyxa*. The ribonuclease-encoding gene BL03719, which is induced by phosphate starvation in *Pseudomonas* sp., of *Pasteurella multocida* was used to mine the studied genomes (Nguyen et al., 2016). The genes *yyvD*, *citA*, *mmgD*, *mmgE* and *yqQ*, which are involved in mediating the glucose and nitrogen starvation reaction.
of Bacillus licheniformis, were identified in the genome of P. polymyxa (Voigt et al., 2010). A fusaricidin expression vector, which contains the fusGFEDCBAA operon, was constructed using a promoter, the amyL terminator and the potential regulator AbrB, and transformed into Bacillus polymyxa. Fusaricidin, which has antifungal activity, was successfully expressed (Li et al., 2013). A new cellu-lase expression system was constructed in P. polymyxa using the promoter of the cellulase model gene celB from Clostridium thermocellum (Heinze et al., 2018). Different genomes were searched for exogenous enzymes, and SignalP 4.0 was used to detect secreted polypeptides (Petersen et al., 2011). The enzymes targeted were cel- lulase, xylanase, lipase and ferulic acid esterase. The B. licheniformis ATCC 14580 spore germination operon genes gerA, gerK and ynd were used to search for homolo-gous genes in the P. polymyxa genome (Borch-Pedersen et al., 2016).

**Enzymes affecting plant growth** A series of enzymes produced by Bacillus spp. that promote the growth and development of plants can be found in different genomes of P. polymyxa. The synthesis of (R,R)-2,3-butanediol by P. polymyxa was promoted by homologous recombination knockout of the diacetylreductase gene dudA (Zhang et al., 2018). A novel polysaccharide lyase gene was found in the KP-1 genome, and the gene combination of P. polymyxa was found by homology (Zhao et al., 2018). The genome A18, which encodes the gene involved in the response to X2-CBM3, was selected as bait for mining the genome of P. polymyxa (Pasari et al., 2017). Bacillus can secrete CIPC hydrolase, 3-CA dioxygenase and chlorocatechol 1, 2-dioxygenase, which can degrade the herbicide phenyl carbamate (CIPC) and reduce herbicide damage to animals and plants (Pujar et al., 2018). Paenibacillus polymyxa natto kinase (NK), trehalose-6-phosphoric acid hydrolase (TreA), glycine oxidase (BiLGO), gamma glutamyl transeptidase (BiSGGT), methyl pectin enzyme (BiLPM), meso-2,3-butanediol dehydrogenase (BiBBDH), glucose transferase (YjIC) and tagatose sugar-1-phosphokinase (TagK) were included in our genome mining work. Our research also included polymyxin E, which is synthesized by the P. polymyxa enzymes encoded by pmxA, pmxE and ectB and plays an important role in microbial growth and metabolism (Yu et al., 2015). The genes oxyR, aphCF, katEG and NRF2, which protect plants from oxidative stress, were also found.

**Antibacterial compounds** Genes similar to the chlor-amphenicol resistance gene cat, the erythromycin resis-tance gene ermD and the streptomycin resistance genes aph and adaK were found in the genome of P. polymyxa by BLAST searches (Kim and Timmusk, 2013). The genes bmyB, bacA, ituD, srfAA and fenDZWP, involved in the synthesis of antibiotics, have been used as bait for genome mining of P. polymyxa (Zhang et al., 2015a). The synthesis of bacterial peptides was enhanced by knock-out of the amino acid permease enzyme gene yhdG (Li et al., 2018). The wool thiopptide genes lanM, lanA and lanP from B. subtilis were used to mine the genome of P. polymyxa for homologous genes (Park et al., 2017). Based on homology mining, the iron-chelating genes of P. polymyxa abrB, yvnA and yvmB were synthesized and the genome was analyzed (Wang et al., 2018).

**Plant disease resistance** The bcrABC-homologous gene yuoA, which contributes to bacterial peptide resis-tance, was found in the Bacillus genome (Ohki et al., 2003). The P. polymyxa diaminobutyric acid synthase gene ectB was introduced, and the B. subtilis gene abrB was knocked out to enhance polymyxin synthesis (Park et al., 2012). The B. subtilis β-lactam antibiotic gene blaI was used as bait for a genomic BLAST search (Filée et al., 2002). Antioxidant enzyme genes and homologues have been found in different P. polymyxa genomes (Raza et al., 2011). The outer protein biosynthesis genes ectABC from B. subtilis were used as bait for mining homologs in the genome of P. polymyxa (Kuhlmann and Bremer, 2002).

Sources of plant growth- and protection-related genes in other bacteria that were used to search for homologs in P. polymyxa are listed in Supplementary Table S2.

**RESULTS**

**Present situation of P. polymyxa strains** In this study, 43 P. polymyxa strains were selected for genome mining (Supplementary Table S1); the complete genomes of 13 strains were available (see Fig. 4). It was found that the P. polymyxa genome size is between 4.06 and 6.42 Mb, the GC content is between 44.70% and 46.60%, and the number of predicted genes lies between 4,234 and 5,840. Most strains were found in food and soil in China, UK, South Korea, India, Germany and Canada. In previous studies, 70% similarity of DDH between two genomes was determined as the appropriate threshold of species division (Stackebrandt and Goebel, 1994), with the DDH value obtained by pairwise genome sequence comparisons via the GGDC application (Meier-Kolthoff et al., 2014); ANI analysis has also often been used in species divi-sion, with a 95–96% similarity cut-off used for species division (Goris et al., 2007; Richter and Rosselló-Móra, 2009). Based on the above species classification criteria, DDH analysis (Fig. 1A) and ANI analysis (Fig. 1B) showed that the DDH value between strain NCTC4744 and any other strain was 0, and the ANI value between strain NCTC4744 and other strains was invariably less than 83%. Because of the consistency between the two species classification techniques, strain NCTC4744 is evidently not a “true” P. polymyxa strain. These results were confirmed by our genome-wide phylogeny analysis, in which
the 43 strains were classified into two main evolutionary branches. One branch consists of strain NCTC4744, and the other consists of the remaining 42 strains (Fig. 1C). We conclude that this strain NCTC4744 does not belong to \textit{P. polymyxa}.

**Mining of gene clusters for secondary metabolites of \textit{P. polymyxa}** Using bioinformatics software, we identified the secondary metabolites and metabolic pathways of \textit{P. polymyxa}. We found that there are 11 groups of secondary metabolites in \textit{P. polymyxa} (Fig. 2A), among which non-ribosomal polypeptides (Nrps), polyketones (Pks) and lanolin thiopptides (lanthipeptides) are widespread in \textit{P. polymyxa}. A total of 797 secondary metabolite gene clusters were identified (Fig. 2B). The number of gene clusters encoding Nrps-processing enzymes was the largest, accounting for 50.94% of the total. Among all secondary metabolite gene clusters excavated in \textit{P. polymyxa}, 454 were similar to the gene clusters for 25 known substances such as fusaricidin, paeninodin, paenilan, polymyxin, and tridecaptin (Fig. 2C). The remaining 343 secondary metabolite gene clusters are not similar to known clus-
Fig. 2. Mining of secondary metabolites of *P. polymyxa*. (A) Heat map reflecting the number of different types of secondary metabolites in different strains. (B) Cumulative presentation of the number of secondary metabolite gene clusters in the different strains. (C) Heat map reflecting the number of secondary metabolites in known categories in different strains. In the heat map, different colors represent different numbers of secondary metabolite gene clusters excavated in *P. polymyxa*.
B. WANG et al.

In the 13 strains of *P. polymyxa* whose genomes were fully sequenced, 8,623 protein-coding genes were present in the pan-genome. Generally, core genome genes in the pan-genome determine the basic biological properties of the species. In the 13 genomes (Fig. 3A), 3,192 genes were identified as constituting the core genome, accounting for 37.01% of the pan-genome of *P. polymyxa*. Based on pan-genome analysis, the relationship between the number of new genes per genome and the number of included genomes was calculated using PanGP software (Fig. 3B). The results show that at the current genome size, when a newly sequenced genome is added, the number of genes can be estimated to increase by 197. The gene accumulation curve (Fig. 3C) shows that the pan-genome of *P. polymyxa* increases as the number of sequencing genomes increases, while the core genome gradually
decreases. Therefore, we conclude that the pan-genome of *P. polymyxa* is open.

**Growth-promoting genes in *P. polymyxa* based on homology mining** Through bioinformatics-based prediction and homology mining of *P. polymyxa*, many putative growth-promoting genes have been identified (Belbahri et al., 2017). We evaluated the growth-promoting potential of the identified plant growth-promoting genes by bioinformatics analysis (Fig. 4), and found that almost all of the identified genes are present in all (completely and incompletely) sequenced *P. polymyxa* strains.

**DISCUSSION**

In the past, *P. polymyxa* strains have been identified as *B. polymyxa* due to their similarity to the *Bacillus* genus (Ash et al., 1993). Considering the high phenotypic similarity between *P. polymyxa* and its closely related *Paenibacillus* spp., it is impossible to distinguish these species from one another by conventional analytical methods (Ash et al., 1993; Dunlap et al., 2015). Although 16S rRNA sequencing is widely used to classify bacterial species, it is difficult and controversial to classify bacterial species based on a limited set of characteristics, as this has led to misclassification errors in a large number of studies (Hahnke et al., 2018). In this study, two species classification techniques, DDH and ANI, supplemented by genome comparison and phylogenetic relationship reconstruction, were used to classify *P. polymyxa*. Our findings showed that 43 strains of *P. polymyxa* were identified as two species (Fig. 1), among which 42 strains represented “true” *P. polymyxa*. The strain NCTC4744 does not match *P. polymyxa*, and should be classified as a separate species. These results are in contrast to the traditional species division based on 16S rRNA sequencing. The genome-based classification methods DDH and ANI can resolve confusion and ambiguity in the classification of highly similar bacteria (Stackebrandt and Goebel, 1994; Goris et al., 2007; Richter and Rosselló-Mora, 2009). Therefore, we suggest that *P. polymyxa* should be classified on the basis of genomic system development.

We have identified multiple secondary metabolite
synthesis gene clusters in *P. polymyxa* (Fig. 2), including Nrps, terpenes, Pks, bacteriocins and lanthipeptides, confirming that *P. polymyxa* is a rich source of secondary metabolites (Cimermancic et al., 2014). Pks and lanthipeptides are common pathogen inhibitors, and strains producing these metabolites have been widely used in food and agriculture (Ansari et al., 2004; Plat et al., 2013). Using bioinformatics tools to predict and mine secondary metabolites, we found that as the number of *P. polymyxa* genomes increased, so did the number of secondary metabolites mined by antiSMASH. Letzel et al. (2014) excavated 211 published genomes of anaerobic bacteria and found that more than 25% of the strains had synthetic gene clusters for post-translationally modified peptides. Zhang et al. (2015b) excavated 830 published actinomycete genomes and found 1,163 synthetic gene clusters for lanthipeptides. In comparison with these types of bacteria, we excavated 11 classes that included 797 secondary metabolite synthesis gene clusters from 43 strains of *P. polymyxa*. In addition, when we analyzed the secondary metabolite gene clusters by bioinformatics tools, we found that 343 metabolite gene clusters were not similar to known substances; these may be considered as orphan gene clusters. By identifying such orphan gene clusters, it is possible to find novel secondary metabolites (Singh and Sareen, 2014; Shi et al., 2019).

Our study showed that the core genome of *P. polymyxa* accounted for about 37% of the pan-genome size, and identified 3,192 genes as the core genome of *P. polymyxa* in 13 fully sequenced strains. These results were similar to those of other microbes. For example, analysis of the genomes of 42 *Brucella* strains showed that the core genome accounted for 58% of the total genome size (Zhao et al., 2014), while the core genome of 18 *Akkermansia muciniphila* strains contains 1,035 genes (Xing et al., 2019). In previous studies, the number of newly identified genes for each newly sequenced *Escherichia coli* strain was about 300 (Xing et al., 2019). We infer from the pan-genome size curve of *P. polymyxa* that as the genomes of more strains are sequenced, many new genes will continue to be identified.

PGPRs can produce metabolites beneficial to plant growth and also inhibit plant pathogens and harmful microorganisms, thus promoting plant growth and development (Bhattacharyya and Jha, 2012). *Paenibacillus polymyxa* is a well-known rhizosphere bacterium that employs a variety of mechanisms to promote plant growth (Jeong et al., 2019). We mined and analyzed genomes of *P. polymyxa* that had been collected from different environments and geographical locations (Supplementary Table S1 and Fig. 4), and obtained much information about the interaction between *P. polymyxa* and host plants. In previous work, comparative genome analysis revealed that 35 *Paenibacillus* spp. had highly conserved plant growth-promoting traits (Xie et al., 2016). We found that characteristic factors influencing processes such as nitrogen fixation, root colonization, production of IAA and antimicrobial peptides, biomass degradation, and plant disease resistance exist in almost all genomes of *P. polymyxa*, regardless of whether the genome of the strain was sequenced completely. These results indicate that *P. polymyxa* has highly conserved plant growth-promoting traits. Therefore, based on previous findings (Belbahri et al., 2017), we speculate that plant growth-promoting traits are evolutionary characteristics of a strain’s adaptation to plant-related habitats.

In this study, through gene cluster analysis of the secondary metabolites of 43 *P. polymyxa* strains, of which 13 strains were completely sequenced, we identified one strain that clearly lies outside the species boundary, characterized the diversity of secondary metabolites, and highlighted the conservation of genes that promote plant growth. Based on DDH and ANI, supplemented by phylogenetic tree analysis of the genome, we determined that the strain NCTC4744 was not *P. polymyxa*. By pan-genomic analysis, we found that the pan-genome of *P. polymyxa* was open, and there were 3,192 core genes in *P. polymyxa*. Analysis of growth-promoting genes showed that *P. polymyxa* strains have highly conserved plant growth-promoting traits. Moreover, we have found 343 potential unknown secondary metabolite gene clusters in the genome of *P. polymyxa*. These results will help us to further understand *P. polymyxa*, and to excavate more novel secondary metabolites in *P. polymyxa*.

This study was supported by grants from the High-level Talents Funding Projects of Hebei Province (B2018003019), Science and Technology Planning Project of Hebei Academy of Sciences (19310, 20309), and Key R&D Program Projects in Hebei Province (19226513D). We thank LetPub (www.letpub.com) for linguistic assistance during the preparation of this manuscript.

**REFERENCES**

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Anand, R., and Chanway, C. P. (2013) *nif* gene sequence and arrangement in the endophytic diazotroph; *Paenibacillus polymyxa* strain P2b-2R. Biol. Fertil. Soils 49, 965–970.

Ansari, M. Z., Yadav, G., Gokhale, R. S., and Mohanty, D. (2004) Nucleic Acids Res. 32, W405–W413.

Ash, C., Priest, F. G., and Collins, M. D. (1993) Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Antonie Van Leeuwenhoek 64, 253–260.

Beatty, P. H., and Jensen, S. E. (2002) *Paenibacillus polymyxa* produces fusaricidin-type antifungal antibiotics active against *Leptosphaeria maculans*, the causative agent of blackleg disease of canola. Can. J. Microbiol. 48, 159–169.

Belbahri, L., Chenari Bouket, A., Rekik, I., Alenezi, F. N., Vallat, A., Luptakova, L., Petrovova, E., Oszako, T., Cherrad, S., Vacher, S., et al. (2017) Comparative genomics of *Bacillus amyloliquefaciens* strains reveals a core genome with traits...
for habitat adaptation and a secondary metabolites rich accessory genome. Front. Microbiol. 8, 1438.

Bhattacharyya, P. N., and Jha, D. K. (2012) Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. World J. Microbiol. Biotechnol. 28, 1237–1350.

Blin, K., Kim, H. U., Medema, M. H., and Weber, T. (2019) Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. Brief Bioinform. 1103–1113.

Bloomberg, G. V., and Lugtenberg, B. J. J. (2001) Molecular basis of plant growth promotion and biocontrol by rhizobacteria. Curr. Opin. Plant Biol. 4, 343–350.

Borch-Pedersen, K., Lindbäck, T., Madslien, E. H., Kidd, S. W., O'Sullivan, K., Granum, P. E., and Aspholm, M. (2016) The cooperative and interdependent roles of GerA, GerK, and Ynd in germination of Bacillus licheniformis spores. Appl. Environ. Microbiol. 82, 4279–4287.

Bruto, M., Prigent-Combaret, C., Muller, D., and Moënne-Loccoz, Y. (2014) Analysis of genes contributing to plant-beneficial functions in plant growth-promoting rhizobacteria and related Proteobacteria. Sci. Rep. 4, 6261.

Castaneda-Alvarez, C., Prodan, S., Rosales, I. M., and Aballay, E. (2016) Exoenzymes and metabolites related to the nematicidal effect of rhizobacteria on Xiphinema index Thorne & Allen. J. Appl. Microbiol. 120, 413–424.

Choi, H.-K. (2019) Translational genomics and multi-omics integrated approaches as a useful strategy for crop breeding. Genes Genomics 41, 133–146.

Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Wieland Brown, L. C., Mavrommatis, K., Pati, A., Godfrey, P. A., Koehrsen, M., Clardy, J., et al. (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 58, 412–421.

Duplan, C. A., Kim, S.-J., Kwon, S.-W., and Rooney, A. P. (2015) Phylogenomic analysis shows that Bacillus amyloliquefaciens subsp. plantarum is a later heterotypic synonym of Bacillus methylotrophicus. Int. J. Syst. Evol. Microbiol. 65, 2104–2109.

Fiëre, P., Benlafya, K., Delmarchelle, M., Moutzourelis, G., Frère, J. M., Brans, A., and Joris, B. (2002) The fate of the BlaP repressor during the induction of the Bacillus licheniformis BlaP β-lactamase. Mol. Microbiol. 44, 685–694.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81–91.

Guemouri-Athmani, S., Berge, O., Bourrain, M., Mavingui, P., Thiéry, J. M., Bhatnagar, T., and Heulin, T. (2000) Diversity of Paenibacillus polymyxa populations in the rhizosphere of wheat (Triticum durum) in Algerian soils. Eur. J. Soil Biol. 36, 149–159.

Gurr, G. M., and You, M. (2015) Conservation biological control of pests in the molecular era: new opportunities to address old constraints. Front. Plant Sci. 6, 1255.

Hahnke, R. L., Meier-Kolthoff, J. P., García-López, M., Mukherjee, S., Huntemann, M., Ivanova, N. N., Woyke, T., Kyrpides, N. C., Klenk, H.-P., and Göker, M. (2018) Corrigendum: genome-based taxonomic classification of Bacteroidetes. Front. Microbiol. 9, 304.

Heine, S., Zimmermann, K., Ludwig, C., Heinzlimreir, S., Schwarz, W. H., Zverlov, V. V., Liebl, W., and Kornberger, P. (2018) Evaluation of promoter sequences for the secretion production of a Clostridium thermocellum cellulase in Paenibacillus polymyxa. Appl. Microbiol. Biotechnol. 102, 10147–10159.

Jeong, H., Choi, S.-K., Ryu, C.-M., and Park, S.-H. (2019) Chronicle of a soil bacterium: Paenibacillus polymyxa E681 as a tiny guardian of plant and human health. Front. Microbiol. 10, 467.

Khan, Z., Kim, S. G., Jeon, Y. H., Khan, H. U., Son, S. H., and Kim, Y. H. (2008) A plant growth promoting rhizobacterium, Paenibacillus polymyxa strain GBR-1, suppresses root-knot nematode. Bioresour. Technol. 99, 3016–3023.

Kim, S.-B., and Timmus, S. (2013) A simplified method for gene knockout and direct screening of recombinant clones for application in Paenibacillus polymyxa. PLoS One 8, e68092.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kuhlmann, A. U., and Bremer, E. (2002) Osmotically regulated synthesis of the compatible solute ectoine in Bacillus pasteurii and related Bacillus spp. Appl. Environ. Microbiol. 68, 772–783.

Lai, S., and Tabacchioni, S. (2009) Ecology and biotechnological potential of Paenibacillus polymyxa: a minireview. Indian J. Microbiol. 49, 2–10.

Letzel, A.-C., Pirot, S. J., and Hertweck, C. (2014) Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. BMC Genomics 15, 983.

Li, S., Zhang, R., Wang, Y., Zhang, N., Shao, J., Qiu, M., Shen, B., Yin, X., and Shen, Q. (2013) Promoter analysis and transcription regulation of fas gene cluster responsible for fusaricidin synthesis of Paenibacillus polymyxa SQR-21. Appl. Microbiol. Biotechnol. 97, 9479–9489.

Li, Y., Wu, F., Cai, D., Zhan, Y., Li, J., Chen, X., Chen, H., Chen, S., and Ma, X. (2018) Enhanced production of bacitracin by knocking out of amino acid permease gene yhdG in Bacillus licheniformis DW2. Sheng Wu Gong Cheng Xue Bao 34, 916–927.

McSpadden Gardener, B. B. (2004) Ecology of Bacillus and Paenibacillus spp. in agricultural systems. Phytopathology 94, 1252–1258.

Medini, D., Donati, C., Tettelin, H., Masiignani, V., and Rappuoli, R. (2005) The microbial pan-genome. Curr. Opin. Genet. Dev. 15, 589–594.

Medini, D., Serruto, D., Parkhill, J., Relman, D. A., Donati, C., Moxon, R., Falkow, S., and Rappuoli, R. (2008) Microbiology in the post-genomic era. Nat. Rev. Microbiol. 6, 419–430.

Meier-Kolthoff, J. P., Klenk, H. P., and Göker, M. (2014) Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. Int. J. Syst. Evol. Microbiol. 64, 352–356.

Nguyen, T. T., Nguyen, M. H., Nguyen, H. T., Nguyen, H. A., Le, T. H., Schweder, T., and Jürgen, B. (2016) A phosphate starvation-inducible ribonuclease of Bacillus licheniformis. J. Microbiol. Biotechnol. 26, 1464–1472.

Niu, B., Vater, J., Ruckert, C., Blom, J., Lehmann, M., Ru, J.-J., Chen, X.-H., Wang, Q., and Borris, R. (2013) Polymyxin P is the active principle in suppressing phytopathogenic Erwinia spp. by the biocontrol rhizobacterium Paenibacillus polymyxa M-1. BMC Microbiol. 13, 137.

Ohki, R., Tateno, K., Okada, Y., Okajima, H., Asai, K., Sadaie, Y., Murata, M., and Aiso, T. (2003) A bacitracin-resistant Bacillus subtilis gene encodes a homologue of the membrane-spanning subunit of the Bacillus licheniformis ABC transporter. J. Bacteriol. 185, 51–59.

Park, J.-E., Kim, H.-R., Park, S.-Y., Choi, S.-K., and Park, S.-H.
