Reclassification of the biocontrol agents *Bacillus subtilis* BY-2 and Tu-100 as *Bacillus velezensis* and insights into the genomic and specialized metabolite diversity of the species

Alex J. Mullins[2]†, Yinshui Li[1]†, Lu Qin[1], Xiaojia Hu[1], Lihua Xie[1], Chiming Gu[1], Eshwar Mahenthiralingam[2], Xing Liao[1],* and Gordon Webster[2],*

Abstract

The genomes of two historical *Bacillus* species strains isolated from the roots of oilseed rape and used routinely in PR China as biocontrol agents to suppress *Sclerotinia* disease were sequenced. Average nucleotide identity (ANI) and digital DNA–DNA hybridization analyses demonstrated that they were originally misclassified as *Bacillus subtilis* and now belong to the bacterial species *Bacillus velezensis*. A broader ANI analysis of available *Bacillus* genomes identified 292 *B. velezensis* genomes that were then subjected to core gene analysis and phylogenomics. Prediction and dereplication of specialized metabolite biosynthetic gene clusters (BGCs) defined the prevalence of multiple antimicrobial-associated BGCs and highlighted the natural product potential of *B. velezensis*. By defining the core and accessory antimicrobial biosynthetic capacity of the species, we offer an in-depth understanding of *B. velezensis* natural product capacity to facilitate the selection and testing of *B. velezensis* strains for use as biological control agents.

INTRODUCTION

Phytopathogens are a major constraint on global food production, causing plant disease in the field and during post-harvest storage. The over-reliance on synthetic pesticides in modern agriculture and the damage they cause to soil fertility, the accumulation of toxic residues and the emergence of pathogen resistance have led to a need for more sustainable disease management practices [1]. Biological control of plant pathogens using naturally occurring antagonistic micro-organisms (biopesticides) is one such practice, and it has recently attracted renewed research interest due to the requirement for environmentally friendly options [2]. Several micro-organisms have been developed as biopesticides and used routinely in the field, and among the most successful are members of the *Bacillus subtilis* species complex, known to suppress disease caused by bacterial, fungal, viral and nematode plant pathogens [2–5].

The important *B. subtilis* species complex [6] includes approximately 20 closely related species [7] and belongs to the phylum *Firmicutes*. However, the taxonomy is notoriously difficult to discriminate based on traditional phenotypic and 16S rRNA gene sequencing methods [7, 8]. For example, comparison of the complete 16S rRNA gene sequences of *Bacillus amyloliquefaciens* DSM 7^T and *B. subtilis* 168^T revealed >99% sequence identity [7]. The taxonomy of the *B. subtilis* species complex was recently updated and clarified through genomic analysis [7, 9–11]. The complex now includes four monophyletic groups: the clade I *B. subtilis* group; the clade II *B. amyloliquefaciens* group; the clade III *Bacillus licheniformis* group; and the clade IV *Bacillus pumilus* group [7]. The clade II *B. amyloliquefaciens* group contains strains that are more proficient at rhizosphere colonization and biocontrol than other members of the *B. subtilis* species complex [8, 12], and average nucleotide identity (ANI), digital DNA–DNA hybridization (dDDH) and core gene phylogeny demonstrate that clade II comprises three discrete but closely

related species: *B. amyloliquefaciens*, *Bacillus velezensis* and *Bacillus siamensis* [7, 13].

From the biotechnological point of view, a key feature of the *B. amyloliquefaciens* group is their diverse metabolism and the ability to produce a wide range of different antagonistic compounds resulting in broad antimicrobial activity [2, 13, 14]. The model biocontrol bacterium *B. velezensis* FZB42 (formerly *B. amyloliquefaciens* subspecies *plantarum*) has a large part of its genome dedicated to the synthesis of specialized metabolites, with the ability to produce numerous structurally diverse antimicrobial compounds, including cyclic lipopeptides (e.g. surfactin, bacillomycin-D, fengycin, bacillibactin), polyketides (e.g. macrolactin, bacillaene, difficidin), siderophores and bacteriocins [5, 15–17]. In comparison with *B. amyloliquefaciens* and *B. siamensis*, the species *B. velezensis* has a higher number of biosynthetic genes involved in specialized metabolite production [7, 13].

In this study we aimed to sequence the genomes of two important *Bacillus* species strains (BY-2 and Tu-100) used routinely in PR China as biocontrol agents for the pathogen of oilseed rape, *Sclerotinia sclerotiorum* [18], which causes a disease of major importance in PR China [19], with annual yield losses of 10–80% [18]. Previously both strains had been identified as *B. subtilis* [20, 21], but via genomic analysis we have revised their taxonomic status within the *B. subtilis* species complex as *B. velezensis* positioned in the clade II *B. amyloliquefaciens* group. In addition, the two biocontrol strains were analysed for their specialized metabolite biosynthetic potential via genome mining alongside 290 additional *B. velezensis* genomes. Understanding the intricacies of the antimicrobial repertoire within *B. velezensis* will allow improved biocontrol formulations and applications to be developed.

METHODS

Bacterial isolates

Bacillus strains BY-2 and Tu-100 were originally isolated from the internal root tissues [22] and the rhizosphere [23], respectively, of oilseed rape (*Brassica napus*) plants grown in field plots located in Wuhan, Hubei Province, PR China. Both strains were shown to control *Sclerotinia* disease on oilseed rape [4, 20, 21]. Strains were routinely maintained in lysogeny broth (LB) [24] supplemented with streptomycin (50 µg ml⁻¹) and rifampicin (50 µg ml⁻¹) for BY-2 and kanamycin (30 µg ml⁻¹) and rifampicin (50 µg ml⁻¹) for Tu-100 to prevent contamination; BY-2 and Tu-100 are naturally resistant to these antibiotics at these concentrations. Strains BY-2 and Tu-100 are held in the Agricultural Culture Collection of China (Beijing, PR China) as ACCC 06137 and ACCC 05846, respectively.

Genome sequencing, assembly and annotation

Strains BY-2 and Tu-100 were grown in LB at 28 °C in 250 ml volume Erlenmeyer flasks and shaken at 150 r.p.m. for 48 h. Genomic DNA was then extracted using a modified cetyltrimethylammonium bromide (CTAB) extraction method as

Impact Statement

This study revisits two historically successful biocontrol *Bacillus* strains and applies genomics and bioinformatics to reclassify the strains as *Bacillus velezensis*. *B. velezensis* is a promising plant growth-promoting and biocontrol species and, as such, a deeper understanding of the specialized metabolite potential would be of benefit to the agricultural research community. We assess the specialized metabolite potential of 292 *B. velezensis* genomes via a biosynthetic gene cluster prediction and dereplication strategy. This builds on existing knowledge by providing a holistic view of the prevalence of multiple antimicrobial-associated gene clusters and defines the core and accessory gene clusters of the species. Our study offers a comprehensive analysis of *B. velezensis* natural product biosynthesis capacity and potential.

described previously [25], and DNA concentration, quality and integrity were determined with a Qubit Fluorometer (Invitrogen) and a NanoDrop Spectrophotometer (Thermo Scientific). Sequencing libraries were generated using the TruSeq DNA Sample Preparation kit (Illumina) and the Template Prep kit (Pacific Biosciences). Whole-genome data were then generated using an Illumina NovaSeq (paired-end, 2×150 bp) sequencing platform and a PacBio sequel system at Wuhan Yanxing Biotechnology Co., Ltd . PacBio reads were assembled into contigs by HGAP v4 implementing CANU v1.7.1. Contigs were then polished with Pilon v1.18 using the Illumina reads. Genes were predicted by GeneMarkS v4.32 [26] and classified using eggNOG v5.0 [27] according to the clusters of orthologous groups (COGs) functional classification. A genome map was drawn from coding sequences (CDSs) and non-coding RNA predictions by Wuhan Yanxing Biotechnology Co., Ltd using the CGView server [28].

Genomic data source and analysis

All genomes submitted as *Bacillus* (Taxon ID 1386) were downloaded from the European Nucleotide Archive (ENA) using the ENA enaBrowserTools interface scripts (https://github.com/enasequence/enaBrowserTools). Analyses were performed using the Cloud Infrastructure for Microbial Bioinformatics (CLIMB) [29].

Phylogenomics of *B. velezensis*

B. velezensis genomes were identified by comparing downloaded *Bacillus* genomes to the *B. velezensis* type strain NRRL B-41580ᵀ using the average nucleotide identity analysis tool FastANI v1.2 [30]. Genomes with an ANI value ≥95% were further verified using the alignment based (ANIm) ANI tool PyANI v0.2.9 [31] implementing a 95% ANI species threshold [30, 32]. Genomes were annotated using Prokka v1.14.5, and the annotation files used to generate a core-gene alignment with Roary v3.13.0 [33] implementing MAFFT v7.455. A maximum-likelihood phylogeny was constructed
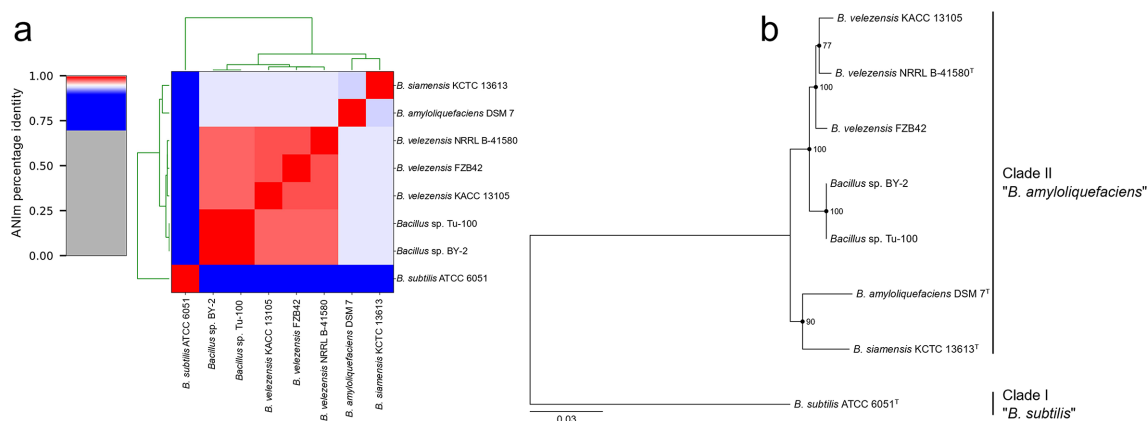
Fig. 1. Pairwise average nucleotide identity analysis and core gene phylogeny. (a) *Bacillus* strains BY-2 and Tu-100 were compared to representatives of the *B. subtilis* species complex clade II *B. amyloliquefaciens* operational group and the *B. subtilis* type strain ATCC 6051[T] (*B. subtilis* species complex clade I, operational group *B. subtilis*) using the ANI tool PyANI (ANIm). (b) A phylogeny was constructed using 313 core genes identified from the genomes used in the ANI analysis. Nodes are labelled with bootstrap values. Error bar represents substitutions per site.

using RAxML v8.2.12 with a GTR substitution matrix and GAMMA model of rate heterogeneity supported by 100 bootstraps. In addition, the Type (strain) Genome Server (TYGS) bioinformatics platform (https://tygs.dsmz.de) was used for whole genome-based taxonomic analysis of BY-2, Tu-100, FZB42 and KACC 13105 against type strains [34]. This platform provides both species assignment and digital DNA–DNA hybridization (dDDH) values to the closest type strain genomes available. The dDDH values (Table S1, available in the online version of this article) used in this study were calculated using formula d_4 [sum of all identities found in high-scoring segment pairs (HSPs) divided by overall HSP length] [35]. The formula d_4 is preferred as it is independent of genome length and robust against the use of incomplete draft genomes. Digital DNA-DNA hybridisation values of <70% were used to indicate different species [36].

Biosynthetic gene cluster prediction and dereplication

Specialized metabolite biosynthetic gene clusters (BGCs) were predicted for *B. velezensis* genomes with antiSMASH 5.0.0. A k-mer based dereplication of BGCs was performed as described previously [37]. In brief, a pairwise comparison of BGC k-mers was achieved using the distance estimation tool Mash v2.2.2 with a variable distance threshold. Predicted BGCs with a sequence length below 6 kbp were excluded as these likely represented BGCs fragmented across multiple contigs. Due to the adjacency of the iturin BGC and fengycin BGC in *B. velezensis* genomes, a local BLASTN v2.9.0+ search was performed to confirm their distribution.

### *In vitro* fungal inhibition assays

Bacillus strains BY-2 and Tu-100 were grown in LB at 28 °C in 50 ml volume Erlenmeyer flasks and shaken at 150 r.p.m. After 48 h, growth bacteria were removed by centrifugation (5000 r.p.m. for 10 min; Eppendorf centrifuge 5427 R)

and the spent media containing bacterial metabolites were sterilized using a 0.2 μm sterile filter. *In vitro* fungal inhibition assays were then conducted on 20 ml potato dextrose agar (PDA) contained in 9.0 cm plastic Petri dishes. Four holes were cut out of the PDA using a 5.0 mm diameter corer at equal distance around the outer part of the plate. To each assay plate (three replicate plates per fungi tested), the holes were filled with (i) 100 μl of iturin A (1.0 mg ml⁻¹), (ii) 100 μl of BY-2 media, (iii) 100 μl of Tu-100 media and (iv) 100 μl of water as a control. A 5.0 mm diameter plug of leading-edge growth of *S. sclerotiorum* strain Ss-1 and *Rhizoctonia solani* strain Rs-2 was placed in the centre of the four treatments. Assay plates were incubated for a further 48 h at 22 °C (*S. sclerotiorum*) or 28 °C (*R. solani*), and the distance between the edge of the fungal colony and treatment was measured and averaged per plate.

RESULTS AND DISCUSSION

Reclassifying *B. subtilis* strains BY-2 and Tu-100 as *B. velezensis*

The species classification of *Bacillus* strains BY-2 and Tu-100 was determined by comparing their genomes to more than 4800 *Bacillus* genomes using a k-mer ANI analysis. Both BY-2 and Tu-100 were similar (≥95% ANI) to *Bacillus* genomes classified as *B. velezensis*. To confirm this classification, BY-2 and Tu-100 genomes were further compared by alignment-based ANI, digital DDH and core gene phylogenomics to type strains of the *B. amyloliquefaciens* clade [7] and *B. velezensis* strains formerly classified as different species (Fig. 1, Table S1). All three genome comparison methods gave congruent results and confirmed the placement of BY-2 and Tu-100 with *B. velezensis*, separate from other members of the *B. amyloliquefaciens* clade and the outgroup *B. subtilis* ATCC 6051[T].
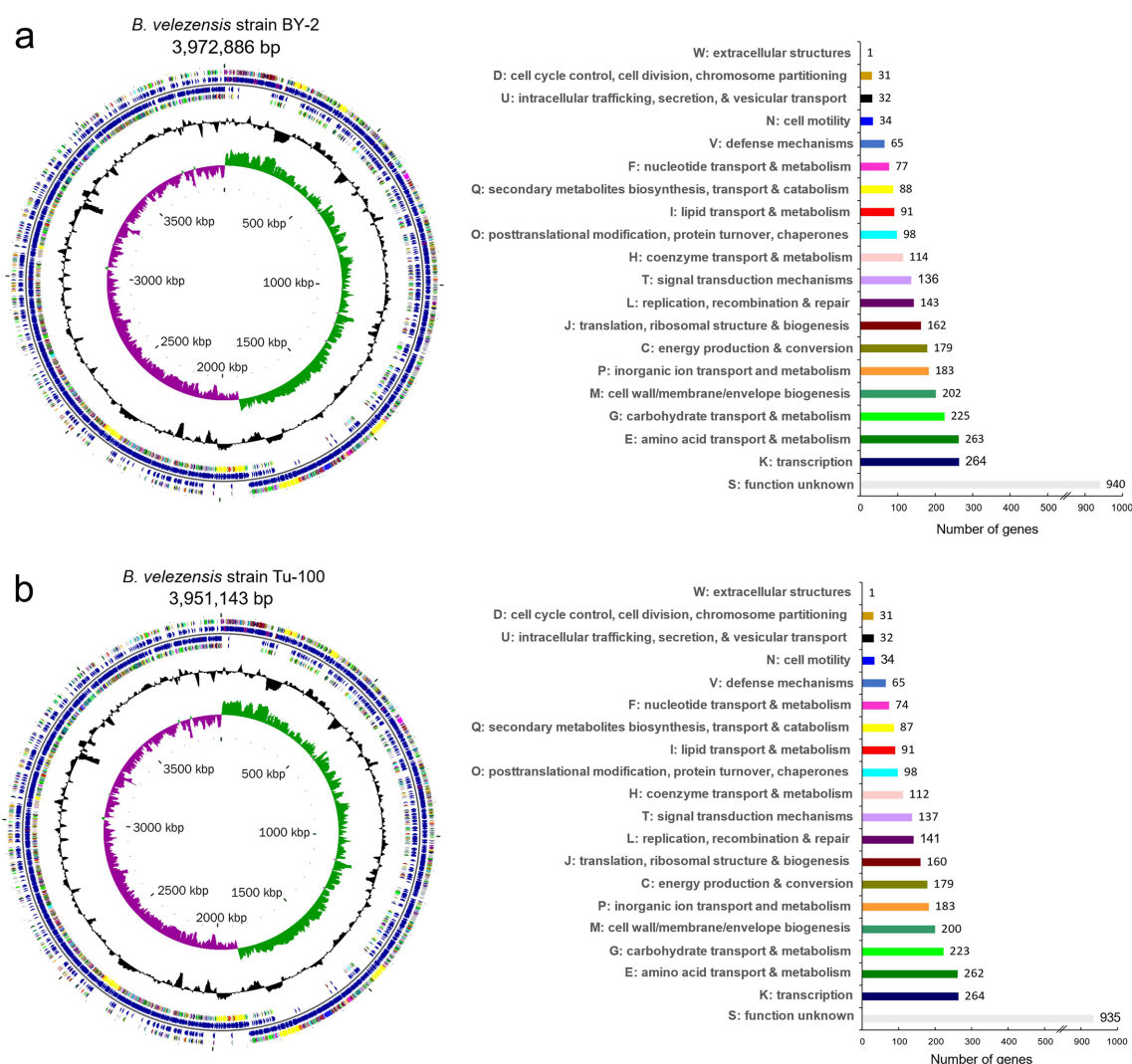
Fig. 2. Visual representation and COG functional classification of the *B. velezensis* BY-2 and Tu-100 genomes. (a) *B. velezensis* BY-2 chromosome map and COG functional annotation of 3328 predicted proteins. (b) *B. velezensis* Tu-100 chromosome map and COG functional annotation of 3309 predicted proteins. From outer circle to the centre: CDSs on forward strand (coloured according to COG categories listed in the bar chart); all CDSs (blue) and RNA (brown/purple) genes on forward strand; all CDSs (blue) and RNA (brown/purple) genes on reverse strand; CDSs on reverse strand (coloured according to COG categories); GC content (black); GC skew (positive GC skew values are plotted in purple, and negative values are in green); scale bar. The map was generated using Circular Genome Viewer (CGView).

The *B. velezensis* BY-2 and Tu-100 genomes

The genomes of *B. velezensis* strains BY-2 and Tu-100 consisted of a single circular chromosomal replicon that was 3.97 and 3.95 Mbp in size, respectively, with a G+C content of 46.5% (Fig. 2, Table S2), and each also contained one 0.2 Mbp plasmid (Fig. S1). The genome metrics of both strains fitted within the size range of 3.81–4.24 Mbp and 45.9–46.8% GC reported for other completed genomes of *B. velezensis* [38]. The BY-2 and Tu-100 genomes were predicted to carry 3915 and 3893 genes, respectively, of which approximately 85% (3328 and 3309 CDSs, respectively) were classified as protein coding sequences with COG function (Table S3). Both BY-2 and Tu-100 carried 114 (2.9%) and 108 (2.8%)

rRNA/tRNA coding genes, respectively, representing 9 and 8 copies each of the rRNA operon genes (5S, 16S and 23S rRNA genes; Fig. 2). The most abundant COG functions in the genomes of BY-2 and Tu-100 were associated with metabolism, and the proportion of genes related to metabolic functions was 36.7% (1220/3328) and 36.6% (1211/3309) of genes classified, respectively (Table S3). Specific functions of transcription (K), amino acid transport and metabolism (E), carbohydrate transport and metabolism (G), and cell wall/membrane/envelope biogenesis (M) were the most abundant COG categories in the genomes (Fig. 2, Table S3). Interestingly only 2.6% of the genes (Fig. 2) in both genomes by COG were characterized as secondary metabolites biosynthesis,
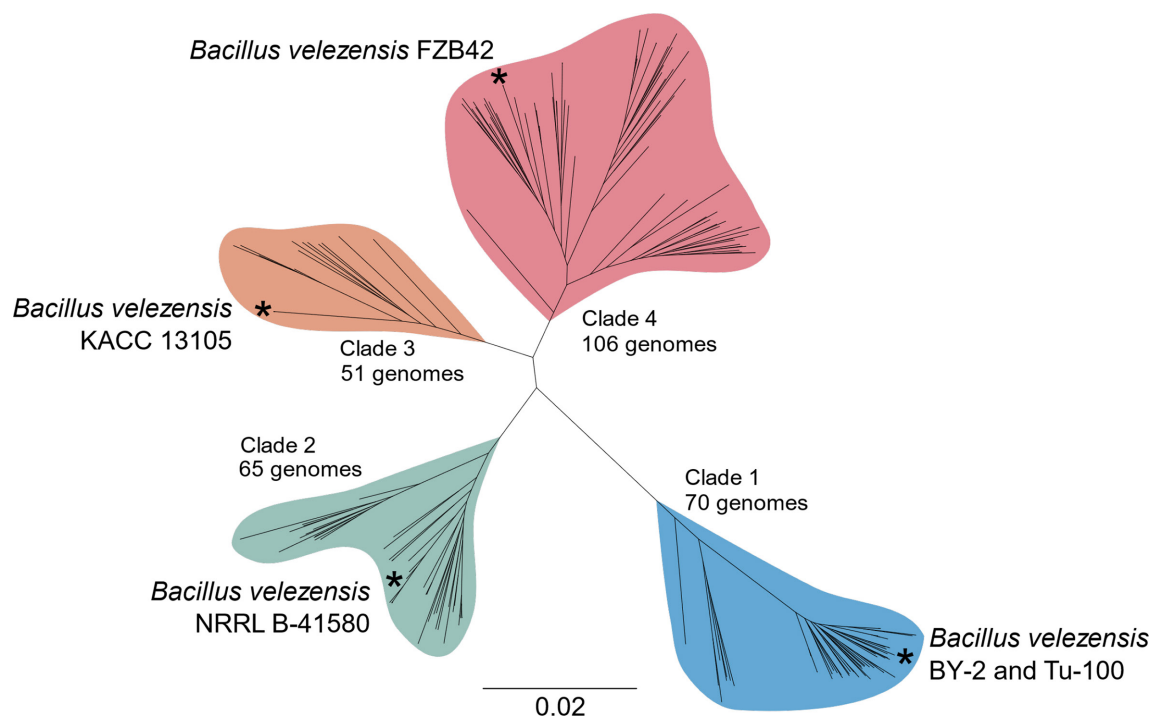
Fig. 3. Unrooted core gene phylogeny of *B. velezensis*. The phylogenetic tree was constructed based on an alignment of 1301 core genes identified in 292 *B. velezensis* genomes. The phylogeny was divided into four broad clades and labelled: clade 4, *B. velezensis* FZB42 (formerly *B. amyloliquefaciens* subspecies *plantarum*); clade 3, *B. velezensis* KACC 13105 (formerly *B. methylotrophicus*); clade 2, *B. velezensis* type strain NRRL B-41580^T; clade 1, *B. velezensis* strains BY-2 and Tu-100. The number of *B. velezensis* genomes in each clade is listed.

transport and catabolism (Q), whereas antiSMASH analysis predicted that 20.7% of the genomes for both strains were involved in specialized metabolite production (Table S4). A similar discrepancy (2.8 and 20.2%, respectively) in the two methods was also reported for the plant growth-promoting *B. velezensis* strain WRN014 [39]. This may be due to genes predicted within a BGC by antiSMASH not being successfully categorized or directly linked with secondary metabolism using the COG database. In addition, antiSMASH analysis of the model biocontrol strain *B. velezensis* FZB42 also suggests that it has 20.4% of its genome dedicated to BGCs (data not shown) and this high percentage (approximately 1/5 of the genome) linked to specialized metabolism is a feature of *B. velezensis* strains [39].

Genomic diversity of *B. velezensis* reveals subspecies structure

A core gene analysis and phylogeny were constructed to better understand the genomic diversity of the biological control species *B. velezensis*. An analysis of 292 genomes defined a total of 26965 genes, and a core genome composed of 1301 genes. The average *B. velezensis* genome possessed 3871 genes, indicating that, on average, 66.4% of the genes in each strain represent accessory functions. The core gene phylogeny of the 292 genomes divided broadly into four *B. velezensis* clades, with the *B. velezensis* type strain NRRL

B-41580^T and heterotypic synonym type strains, *B. velezensis* (formerly *B. methylotrophicus*) KACC 13105 and *B. velezensis* (formerly *B. amyloliquefaciens* subspecies *plantarum*) FZB42, representing distinct clades (Figs 3 and S2). The fourth group (clade 1) was occupied by the closely related *B. velezensis* strains BY-2 and Tu-100 (Figs 3 and S2).

Specialized metabolite potential of *B. velezensis* based on biosynthetic gene cluster diversity

Biosynthetic gene clusters were predicted for all 292 *B. velezensis* genomes to capture the potential of the species to synthesize specialized metabolites. Following the removal of fragmented BGCs, 4563 BGCs were defined and subsequently dereplicated into 41 distinct BGCs (Fig. 4; see File S1 for more details). Characterized specialized metabolite BGCs represented 15 of the 41 BGCs predicted across *B. velezensis*. A well-defined threshold was observed from this analysis that demarcated core BGCs, occurring in ≥90% of *B. velezensis* strains, and accessory BGCs, occurring in <90% of strains (Fig. 4). Only 12 BGCs (29%) were defined as core, while the remaining 29 BGCs (71%) constituted the accessory specialized metabolite capacity of the species. The majority of core BGCs (8 out of 12) were associated with known characterized antimicrobial metabolites, including iturin, fengycin, difficidin and macrolactin H. Other known antimicrobial BGCs were accessory and occurred less frequently
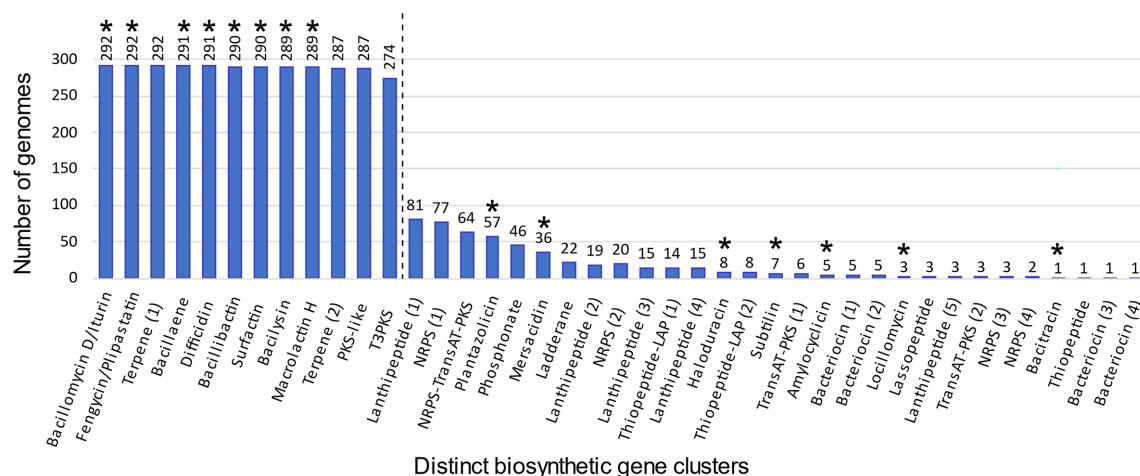
Fig. 4. Occurrence of distinct biosynthetic gene clusters in *B. velezensis*. Following biosynthetic gene cluster (BGC) prediction and dereplication, a total of 41 distinct BGCs were identified across 292 *B. velezensis* genomes. NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; LAP, linear azol(in)e-containing peptides. The presence of an asterisk indicates BGCs that have been characterized previously. The natural demarcation of core specialized metabolite BGCs occurring in >90% strains is indicated by a dashed line.

in *B. velezensis* (Fig. 4). The plantazolicin and mersacidin loci were carried by 20 and 12% of genomes, respectively; and the remaining five characterized BGCs (haloduracin, subtilin, amylocyclicin, locillomycin and bacitracin) each occurred in less than 3% of genomes examined (Fig. 4). The most populous metabolite classes of uncharacterized BGCs were lanthipeptides, five BGCs; non-ribosomal peptide synthetases (NRPS), four BGCs; and bacteriocins, four BGCs. Interestingly, there was also evidence of two uncharacterized transAT-polyketide synthase (transAT-PKS)-encoding BGCs, and a hybrid NRPS-transAT-PKS-encoding BGC. A limited number of the accessory BGCs appeared to be *B. velezensis* clade-specific, with the majority exhibiting a broad distribution across the phylogeny (Fig. 3). The phosphonate BGC was restricted to *B. velezensis* clade 2 comprising the *B. velezensis* type strain NRRL B-41580[T]. However, each of the remaining clade-restricted metabolites, including the subtilin and locillomycin BGCs, only occurred in eight or fewer genomes and, as such, were sub-clade specific and not defining features of the broader phylogenetic clades. The two highly similar *B. velezensis* biocontrol strains BY-2 and Tu-100 possessed the same suite of specialized metabolite biosynthetic gene clusters as each other (Table S4). In addition to the 12 core BGCs defined above (Fig. 4), BY-2 and Tu-100 also carried 2 accessory BGCs, lanthipeptide (4) and the NRPS-transAT-PKS (Table S4). A similar study was conducted on the genus *Bacillus* analysing 1566 genomes, predicting 19962 BGCs, and clustering similar BGCs [40]. The prevalence of characterized specialized metabolite BGCs, relative abundance of metabolite classes, and existence of uncharacterized BGCs across multiple species were highlighted to understand the natural product capacity of the genus [40]. This study focused mainly on the *Bacillus cereus* group species, which represented 90% of the genomes analysed, and the remaining genomes outside the *B. cereus* group lacked representatives of *B. velezensis*.

B. velezensis strains BY-2 and Tu-100 produce antifungal compounds

To confirm whether *B. velezensis* produces antifungal compounds *in vitro* an antagonism assay against two fungal plant pathogens was conducted. Clear bioactivity was observed for spent growth media of BY-2 and Tu-100 against the pathogens *S. sclerotiorum* and *R. solani* (Fig. S3). Interestingly, activity against the two pathogens was also observed for the pure lipopeptide iturin A (Fig. S3), a compound known to be produced by *B. velezensis* [14] and encoded for in both the BY-2 and Tu-100 genomes. Previously, it was shown by PCR amplification of the *ituC* and *ituD* that Tu-100 contained the iturin gene cluster, and it was possible to detect iturin in culture filtrates by using thin layer chromatography [20]. However, it is likely that the bioactivity of BY-2 and Tu-100 metabolites within the spent growth media was due to a mixture of compounds produced by several of the 14 BGCs detected in their genomes, which include known antifungal, antibacterial and nematocidal metabolites: iturin, fengycin, surfactin, bacillaene, difficidin and bacilysin (Fig. 4, Table S4) [41, 5].

Summary

B. velezensis has emerged as a very promising biological control bacterium, with multiple studies highlighting its ability to control crop pathogens and produce a multitude of antimicrobial metabolites. This impressive specialized metabolite repertoire has been defined by the presence of the metabolites themselves or their corresponding biosynthetic gene clusters in specific strains [7, 5]. Other studies have investigated the broader distribution of metabolite classes, such as the iturinic lipopeptides, across the entire *B. subtilis* species complex [14], providing insights into the broader antimicrobial properties of *B. velezensis* and related species.

In contrast, this study has explored the prevalence of known antimicrobial metabolite BGCs deposited in MiBIG across 292 *B. velezensis* genomes and highlighted the unexplored specialized metabolite potential through the prediction of uncharacterized BGCs. The detection of many known metabolites in *B. velezensis*, including iturin, plantazolicin and amylocyclicin, was congruent with the literature [5]. However, we also found evidence of other BGCs previously uncharacterized in *B. velezensis*, such as haloduracin and subtilin. Knowledge of the distribution of antimicrobial BGCs, and core versus accessory BGCs, will contribute to identifying optimal strains for use as biological control agents and allow improved biocontrol formulations and applications to be developed.

References
1. Compant S, Duffy B, Nowak J, Clément 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;71:4951–4959.

2. Fira D, Dimkić I, Berić T, Lozo J, Stanković S. Biological control of plant pathogens by *Bacillus* species. *J Biotechnol* 2018;285:44–55.

3. Borriss R. Use of plant-associated *Bacillus* strains as biofertilizers and biocontrol agents in agriculture. In: Maheshwari D (editor). *Bacteria in Agrobiology: Plant Growth Responses*. Berlin, Heidelberg: Springer; 2011.

4. Hu X, Roberts DP, Jiang M, Zhang Y. Decreased incidence of disease caused by *Sclerotinia sclerotiorum* and improved plant vigor of oilseed rape with *Bacillus subtilis* Tu-100. *Appl Microbiol Biotechnol* 2005;68:802–807.

5. Rabbee MF, Ali MS, Choi J, Hwang BS, Jeong SC *et al. Bacillus velezensis:* A valuable member of bioactive molecules within plant microbiomes. *Molecules* 2019;24:1046.

6. Fritze D. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. *Phytopathology* 2004;94:1245–1248.

7. Fan B, Blom J, Klenk HP, Borriss R. *Bacillus amyloliquefaciens, Bacillus velezensis* and *Bacillus siamensis* form an "Operational Group B. amyloliquefaciens" within the *B. subtilis* species complex.. *Front Microbiol* 2017;8:22.

8. 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;48:249–259.

9. Dunlap CA, Kim S-J, Kwon SW, Rooney AP. Phylogenomic analysis shows that *Bacillus amyloliquefaciens* subsp. *plantarum* is a later heterotypic synonym of *Bacillus methylotrophicus*. *Int J Syst Evol Microbiol* 2015;65:2104–2109.

10. Dunlap CA, Kim SJ, Kwon SW, Rooney AP. *Bacillus velezensis* is not a later heterotypic synonym of *Bacillus amyloliquefaciens*; *Bacillus methylotrophicus*, *Bacillus amyloliquefaciens* subsp. *plantarum* and '*Bacillus oryzicola*' are later heterotypic synonyms of *Bacillus velezensis* based on phylogenomics. *Int J Syst Evol Microbiol* 2016;66:1212–1217.

11. 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;59:2429–2436.

12. Hossain MJ, Ran C, Liu K, Ryu C-M, Rasmussen-Ivey CR *et al.* Deciphering the conserved genetic loci implicated in plant disease control through comparative genomics of *Bacillus amyloliquefaciens* subsp. *plantarum*. *Front Plant Sci* 2015;6:631.

13. Chun BH, Kim KH, 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;77:146–157.

14. Dunlap CA, Bowman MJ, Rooney AP. Iturinic lipopeptide diversity in the *Bacillus subtilis* species group - important antifungals for plant disease biocontrol applications. *Front Microbiol* 2019;10:10.

15. Chen XH, Koumoutsi A, Scholz R, Schneider K, Vater J *et al.* Genome analysis of *Bacillus amyloliquefaciens* FZB42 reveals its potential for biocontrol of plant pathogens. *J Biotechnol* 2009;140:27–37.

16. Fan B, Wang C, Song X, Ding X, Wu L *et al. Bacillus velezensis* FZB42 in 2018: the Gram-positive model strain for plant growth promotion and biocontrol. *Front Microbiol* 2018;9:2491.

17. Rückert C, Blom J, Chen X, Reva O, Borriss R. Genome sequence of *B. amyloliquefaciens* type strain DSM7(T) reveals differences to plant-associated *B. amyloliquefaciens* FZB42. *J Biotechnol* 2011;155:78–85.

18. Hu X, Roberts DP, Xie L, Qin L, Li Y *et al.* Seed treatment containing *Bacillus subtilis* BY-2 in combination with other *Bacillus* isolates for control of *Sclerotinia sclerotiorum* on oilseed rape. *Biological Control* 2019;133:50–57.

19. Derbyshire MC, Denton-Giles M. The control of sclerotinia stem rot on oilseed rape (*Brassica napus*): current practices and future opportunities. *Plant Pathol* 2016;65:859–877.

20. Hu X, Roberts DP, Maul JE, Emche SE, Liao X *et al.* Formulations of the endophytic bacterium *Bacillus subtilis* Tu-100 suppress *Sclerotinia sclerotiorum* on oilseed rape and improve plant vigor in field trials conducted at separate locations. *Can J Microbiol* 2011;57:539–546.

21. Hu X, Roberts DP, Xie L, Maul JE, Yu C *et al.* Formulations of *Bacillus subtilis* BY-2 suppress *Sclerotinia sclerotiorum* on oilseed rape in the field 2014;70:54–64.

22. Jiang M, Zhao R, Hu X, Zhang Y, Wang G. Colonization of antifungal endobacterium BY-2 in oilcrop rape and its control effect on disease caused by *Sclerotinia sclerotiorum*. *Acta Phytopathologica Sinica* 2007;37:192–196.

23. Hu X, Huan Q, Zhang X. Isolating and inoculating rhizosphere bacteria to promote oilseed rape growth. *Oil Crops China* 1992;4:67–69.

24. Miller JH. *Experiments in Molecular Genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory; 1972.

25. Wang F, Hong K. CTAB method for genomic DNA extraction from *Nonomuraea*. *Microbiology China* 2010;37:1211–1215.

26. Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 2001;29:2607–2618.

27. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource

based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;47:D309–D314.

28. Grant JR, Stothard P. The CGView server: a comparative genomics tool for circular genomes. *Nucleic Acids Res* 2008;36:W181–W184.

29. Connor TR, Loman NJ, Thompson S, Smith A, Southgate J *et al*. CLIMB (the cloud infrastructure for microbial bioinformatics): an online resource for the medical microbiology community. *Microb Genom* 2016;2:e000086.

30. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114.

31. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods* 2016;8:12–24.

32. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;106:19126–19131.

33. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S *et al*. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.

34. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.

35. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.

36. Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2:117–134.

37. Mullins AJ, Murray JAH, Bull MJ, Jenner M, Jones C *et al*. Genome mining identifies cepacin as a plant-protective metabolite of the biopesticidal bacterium *Burkholderia ambifaria*. *Nat Microbiol* 2019;4:996–1005.

38. Grady EN, MacDonald J, Ho MT, Weselowski B, McDowell T *et al*. Characterization and complete genome analysis of the surfactin-producing, plant-protecting bacterium *Bacillus velezensis* 9D-6. *BMC Microbiol* 2019;19:5.

39. Wang J, Xing J, Lu J, Sun Y, Zhao J *et al*. Complete genome sequencing of *Bacillus velezensis* WRN014, and comparison with genome sequences of other *Bacillus velezensis* strains. *J Microbiol Biotechnol* 2019;29:794–808.

40. Grubbs KJ, Bleich RM, Santa Maria KC, Allen SE, Farag S *et al*. Large-scale bioinformatics analysis of *Bacillus* genomes uncovers conserved roles of natural products in bacterial physiology. *mSystems* 2017;2:00040–17.

41. Caulier S, Nannan C, Gillis A, Licciardi F, Bragard C *et al*. Overview of the antimicrobial compounds produced by members of the *Bacillus subtilis* group. *Front Microbiol* 2019;10:302.