Check for updates

Screening and genome-wide analysis of lignocellulose-degrading bacteria from humic soil

Tianjiao Zhang[1,2,3,4†], Shuli Wei[1,2,3,4†], Yajie Liu[1,2,3,4], Chao Cheng[5], Jie Ma[1,2,3,4], Linfang Yue[2], Yanrong Gao[1,2,3,4], Yuchen Cheng[2,3,4], Yongfeng Ren[2,3,4], Shaofeng Su[1,2,3,4*], Xiaoqing Zhao[1,2,3,4*] and Zhanyuan Lu[1,2,3,4*]

[1]School of Life Science, Inner Mongolia University, Hohhot, China, [2]Inner Mongolia Academy of Agriculture and Husbandry Science, Hohhot, China, [3]Key Laboratory of Black Soil Protection And Utilization (Hohhot), Ministry of Agriculture and Rural Affairs, Hohhot, China, [4]Inner Mongolia Key Laboratory of Degradation Farmland Ecological Restoration and Pollution Control, Hohhot, China, [5]School of Life Science, Jining Normal University, Ulanqab, China

Crop straw contains huge amounts of exploitable energy, and efficient biomass degradation measures have attracted worldwide attention. Mining strains with high yields of cellulose-degrading enzymes is of great significance for developing clean energy and industrial production of related enzymes. In this study, we reported a high-quality genome sequence of *Bacillus velezensis* SSF6 strain using high-throughput sequencing technology (Illumina PE150 and PacBio) and assessed its lignocellulose degradation potential. The results demonstrated that the genome of *B. velezensis* SSF6 was 3.89 Mb and contained 4,015 genes, of which 2,972, 3,831 and 158 genes were annotated in the COGs (Clusters of Orthologous Groups), KEGG (Kyoto Encyclopedia of Genes and Genomes) and CAZyme (Carbohydrate-Active enZymes) databases, respectively, and contained a large number of genes related to carbohydrate metabolism. Furthermore, *B. velezensis* SSF6 has a high cellulose degradation capacity, with a filter paper assay (FPA) and an exoglucanase activity of 64.48 ± 0.28 and 78.59 ± 0.42 U/mL, respectively. Comparative genomic analysis depicted that *B. velezensis* SSF6 was richer in carbohydrate hydrolase gene. In conclusion, the cellulose-degrading ability of *B. velezensis* SSF6 was revealed by genome sequencing and the determination of cellulase activity, which laid a foundation for further cellulose degradation and bioconversion.

KEYWORDS

cellulose-degrading bacteria, *Bacillus velezensis*, whole-genome sequencing, comparative genomic analysis, carbohydrate-active enzyme

1. Introduction

Crop straw is one of the most abundant biological sources on Earth (Menshawy et al., 2022; Yang et al., 2022). At present, owing to the problems of abundant straw varieties and large yields, complex straw composition and structure, and low conversion rate of the straw industry (Marriott et al., 2016), as well as the influence of rough treatment methods such as straw incineration and burial, a large amount of straw resources are wasted and accompanied by

serious environmental pollution. he return of straw to the field is a crucial step in the innocuous treatment of straw, which can significantly increase soil nitrogen (Ahmed et al., 2020), phosphorus (Bai et al., 2015), potassium (Ma et al., 2010), and other nutrients, which are advantageous for crop growth and development (Chen et al., 2022b), and improve soil nutrients. Therefore, developing quick and efficient techniques for straw treatment is essential.

Agricultural straw primarily comprises of lignocellulosic biomass (LCB) (Ragauskas et al., 2006). Microbial degradation of lignocellulose is a biological treatment method with ecological benefits, compared to physical and chemical methods. By utilizing a variety of members of the carbohydrate-active enzyme (CAZyme) family in concert, cellulose-degrading microorganisms, which act as an intrinsic driving force for the degradation and transformation of biomass, such as straw, degrade carbohydrates into reducing sugars (Sharma et al., 2016), and at the same time degraded lignocellulose is advantageous for microbial growth (Dar et al., 2018). Enzymatic hydrolysis of lignocellulose is a key strategy for the degradation of cellulose because of its specificity, high conversion rate, and ecological character (Pollegioni et al., 2015; Fatani et al., 2021). The ability to rapidly degrade biomass depends on the successful identification of novel strains that generate cellulases, for example, fungi, bacteria and other microorganisms (Abd Elhameed et al., 2020) that can produce highly active cellulases isolated from soil, decaying branches and leaves, animal intestines, and other stuff (Alonso-Pernas et al., 2017; Tian et al., 2017; Ayumi et al., 2018; Wang et al., 2022).

Bacillus velezensis, an endospore-forming gram-positive bacterium belonging to the phylum Firmicutes, is widespread in waterway dregs, soils, and plants, indicating its high ecological adaptability (Balderas-Ruíz et al., 2020; Xu et al., 2020). To date, *B. velezensis* has been widely studied for its ability to efficiently express hydrolases, antibacterial proteins, lipopeptides, and plant hormones, to promote plant growth, and to inhibit plant diseases caused by bacteria and fungi. It has also been reported that *B. velezensis* has a good ability to produce cellulase (Pereira et al., 2019; Lu et al., 2020; Shin et al., 2021; Song et al., 2022). High-throughput sequencing technology is an effective method to analyze the whole genome of *B. velezensis* and mine its related functional genes is an effective method for studying the characteristics of the strain, clarifying its enzyme activity characteristics for degrading lignocellulose, and further increasing the application value of the strain in biomass transformation.

2. Materials and methods

2.1. Sampling, screening, and detection of strains

Humic soil samples were collected from Qingshuihe County (E 111° 0.68′, N 39° 0.92′), Hohhot City, Inner Mongolia Autonomous Region, China. We weighed 5 g of soil into 45 mL of sterile water and diluted it to different concentration of 10^{-1} - 10^{-9} g/mL. Then, 200 μL of 10^{-7} to 10^{-9} dilutions was applied to carboxymethylcellulose sodium culture (CMC) agar medium (K_2HPO_4 2.5 g/L, Na_2HPO_4 2.5 g/L, peptone 2 g/L, yeast extract 0.5 g/L, carboxymethylcellulose sodium 20 g/L, agar 20 g/L), cultured at 37°C for 24 h. According to the method in Teather and Wood (1982), cellulose degrading bacteria

were screened by 0.2% (W/V) Congo red dye. Meanwhile, the selected strains were stained with 1% (W/V) iodine solution (Anand et al., 2010). The isolated strains were inoculated into microcrystalline cellulose (Avicel) agar medium (Avicel 10 g/L, $(NH_4)_2SO_4$ 1.4 g/L, K_2HPO_4 2.5 g/L, $CaCl•2H_2O$ 0.3 g/L, $MgSO_4•7H_2O$ 0.3 g/L, peptone 2 g/L, yeast extract 0.5 g/L, agar 20 g/L), CMC agar medium and starch agar medium (beef extract 5 g/L, peptone 10 g/L, NaCl 5 g/L, starch 2 g/L, agar 20 g/L). The hydrolytic capacity ratio (HCR) of each strain was determined and expressed as a transparent circle diameter ratio. The screened colonies were confirmed to be single colonies by multiple purification cultures and microscopic examinations, and they were stored in liquid Luria-Bertani (LB) medium (tryptone 10 g/L, yeast extract 5 g/L, NaCl 10 g/L) with 30% glycerol at −80°C.

The morphology of the strain was observed by scanning electron microscope (SEM). The cultures were placed on sterile glass covers in a petri dish. After gently rinsing with PBS, fixation with electron microscope solution (No: G1102, Servicebio, China) was fixed at room temperature for 2 h. The fixed samples were rinsed three times with 0.1 M phosphate-buffered saline (PBS) at pH 7.4, with each rinse lasting 15 min Dehydration with 50, 70, 80, 90, and 100% ethanol for 15 min, respectively, and each concentration was repeated three times. Finally, the sample is dried in a critical point dryer (K850, Quorum, England), coated with gold by an ion sputtering apparatus (No: MC1000, HITACHI, Japan), and observed under a scanning electron microscope (No: SU8100, HITACHI, Japan).

2.2. Molecular identification of bacteria

After the activation of the candidate strains, genomic DNA was extracted according to the instructions of the bacterial whole-genome extraction kit (No: DP302, Tiangen Biochemical Technology Co., Ltd., China). Subsequently, 16S ribosomal ribonucleic acid (rRNA) genes was amplified by polymerase chain reaction (PCR) using sequence-specific primers: 27F (5'-AGAGTTTGATCCTGGCTCA-3′) and 1492R (5'-GGTTACCTTGTTACGACTT-3′) (Dobrzyński et al., 2022) in a thermal cycler (Bole T100, United States). The reaction conditions were as follows: 94°C predenaturation (5 min), followed by 35 denaturation cycles (94°C for 30 s), annealing (55°C for 45 s), and extension (72°C for 2 min), and the final repair extension was set at 72°C for 10 min. PCR products were characterized by 1% agarose gel electrophoresis and quantified using a NanoDrop™ One ultra-micro spectrophotometer (Thermo Scientific, USA). The PCR products were sequenced using the Sanger method. Blastn was used to search 16S rRNA gene fragment sequences in the National Center for Biotechnology Information (NCBI) nucleotide database to determine their closest taxonomic relatives (Dashtban et al., 2010). The sequences were used to construct a phylogenetic tree along with other reference genes obtained from NCBI GenBank. A phylogenetic tree was constructed using the MEGA X software neighbor-joining method (1000 bootstrap replications; Kumar et al., 2018).

2.3. Cellulase activity assay

Before being placed in liquid CMC medium for culturing, the isolated strains were cultured in liquid LB medium for 18 h. Endocellulase, exocellulase, and glucanase tests were conducted using

sodium carboxymethyl cellulose, microcrystalline cellulose, and salicin solutions, and the total cellulase activity was determined using the FPA (Wang et al., 2020). The dinitrosalicylic acid method was used to estimate the reducing sugars released during hydrolysis (Mansour et al., 2016; Yadav and Dubey, 2018). One unit (U) of enzyme activity was defined as the amount of enzyme required to release 1 μmol of reducing sugars per milliliter per minute. The Michaelis–Menten equation was used to calculate the kinetic parameters of the enzymatic reaction of cellulose from strain SSF6. The cellulase activity was calculated using the following formula:

$$\text{Enzyme Activity}\left(\text{U / mL}\right) = \frac{\text{Glucose production}\left(\text{mg}\right) \times 1000}{\text{Enzyme dosage}\left(\text{mL}\right) \times \text{time}\left(\text{min}\right)}$$

2.4. Genome sequencing, species assignment and annotation

The extracted genomic DNA was entrusted to Beijing Novogene Bioinformatics Technology Co., Ltd. using an Illumina PE150 system and PacBio high-throughput sequencing technology. Genome assembly was performed using the SMRT Link (version 5.0.1).[1] The initially assembled data were subjected to low-quality read filtering (less than 500 bp), error correction (selection of long read sequences over 6,000 bp), correction (minimum mass value filtering result 20, minimum read depth 4, maximum read depth 1,000), and cyclization starting checkpoint correction to obtain the final completed map sequence (Ardui et al., 2018; Reiner et al., 2018). GeneMarkS software (version 4.17)[2] was used for coding gene prediction and filtering (Besemer et al., 2001). Repeat Masker software (version 4.0.5) (Saha et al., 2008) was used for scattered repetitive sequence prediction, and the TRF (Tandem Repeats Finder, version 4.07b) (Benson, 1999) was used to search for tandem repeats in DNA sequences. Transfer RNA (tRNA) genes were predicted using tRNAscan-SE (Lowe and Eddy, 1997). Ribosomal RNA (rRNA) genes were analyzed using the rRNAmmer software (version 1.2) (Lagesen et al., 2007), and small nuclear RNAs (sRNAs) were predicted using the Rfam database software (Nawrocki et al., 2015). The PhiSpy tool (version 2.3) (Zhou et al., 2011) was used to predict prophages.

The genome sequence of *Bacillus* was queried and downloaded from the GenBank genome database.[3] Typing of the SSF6 genome assembly was determined by calculating Average nucleotide identity (ANI) values through the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016). Meanwhile, orthologous ANI (OrthoANI) was calculated using the orthologous Average Nucleotide Identity Tool (version 1.40) (Lee et al., 2016). The digital DNA–DNA hybridization (dDDH) analysis was carried out using the Genome-to-Genome Distance Calculator (GGDC version 2.1) (Li et al., 2002). BLAST was used to compare the identified genes to the commonly used databases of the NR (NonRedundant Protein), KEGG (Kanehisa et al., 2004, 2006), and the COGs of Proteins (Galperin et al., 2015).

For gene function annotation, the CAZy was used (Cantarel et al., 2009).

2.5. Comparative genomics analysis

The complete genome sequence of the *B. velezensis* FZB42 (NC_009725) (Fan et al., 2017, 2018) strain was downloaded from the NCBI Genome for comparative genomics analysis with *B. velezensis* SSF6. MUMmer alignment software (version 3.23) was used to detect individual SNPs (Kurtz et al., 2004), and SNP functions were annotated according to positional relationships and interactions between SNPs and genes. Protein sequences of multiple samples to be analyzed were clustered using CD-hit software (version 4.6.1) and plotted using R software (version 3.2.4).

2.6. Statistical analysis

GraphPad Prism 9 was used to make glucose standard curves and characteristic enzyme curves. SPSS 20.0 software (IBM SPSS, Chicago, IL) was used for one-way ANOVA analysis of variance for statistical analysis. The data are expressed as the mean ± standard deviation, and a $p < 0.05$ was considered statistically significant.

3. Results

3.1. Isolation of cellulose-degrading bacteria

Four strains with good cellulose degradation function, SSF1, SSF4, SSF6, and SSF15, were screened from soil using a CMC selective medium. These four isolated strains, when stained with Congo red solution, produced clear hydrolysis circles around the colonies, the hydrolysis diameters are 4.82 ± 0.13, 7.92 ± 0.16, 26.20 ± 0.44, and 11.56 ± 0.45 mm, respectively, indicating that they have cellulose hydrolysis ability (Table 1). Cellulose-degrading bacteria were screened based on the HCR ratio (Kang et al., 2022), and the hydrolysis ratios of the four isolated strains in the two media were compared. The hydrolysis ratios of each strain were 1.50 ± 0.07, 1.36 ± 0.11, 3.61 ± 0.05, and 3.57 ± 0.32, respectively. Comprehensive analysis revealed that the strain SSF6 had a higher cellulose hydrolysis ratio (Table 2; Figures 1Aa,b) and exhibited excellent cellulose

TABLE 1 HCR determination for the diameter of the degradation circle of the isolated strain on CMC medium.

Strain	Clearing zone diameter (D, mm)	Colony diameter (d, mm)	HCR (D/d value)
SSF1	4.82 ± 0.13	3.71 ± 0.15	1.30 ± 0.07[c]
SSF4	7.92 ± 0.16	6.48 ± 0.05	1.22 ± 0.03[d]
SSF6	26.20 ± 0.44	6.06 ± 0.08	4.32 ± 0.16[a]
SSF15	11.56 ± 0.45	5.49 ± 0.12	2.11 ± 0.10[b]

Different lowercase letters mean the significance at $p < 0.05$. D, The diameter of the hydrolysis circle; d, The diameter of the colony; HCR (D/d value), The hydrolysis ratio of the strain.

---

1   https://www.pacb.com/support/software-downloads/

2   http://topaz.gatech.edu/

3   https://www.ncbi.nlm.nih.gov/genome/

TABLE 2 HCR determination for the diameter of the degradation circle of the isolated strain on Avicel medium.

Strain	Clearing zone diameter (D, mm)	Colony diameter (d, mm)	HCR (D/d value)
SSF1	7.48 ± 0.21	4.99 ± 0.13	1.50 ± 0.07[b]
SSF4	3.16 ± 0.15	2.33 ± 0.16	1.36 ± 0.11[b]
SSF6	22.82 ± 0.36	6.32 ± 0.15	3.61 ± 0.05[a]
SSF15	10.87 ± 0.81	3.05 ± 0.08	3.57 ± 0.32[a]

Different lowercase letters mean the significance at $p < 0.05$. D, The diameter of the hydrolysis circle; d, The diameter of the colony; HCR (D/d value), The hydrolysis ratio of the strain.

degradation ability. At the same time, we found that the strain SSF6 also had a strong starch degradation ability (3.20 ± 0.12; Table 3; Figures 1Ac), while the other three strains could not degrade starch (Table 3). After comprehensive consideration, strain SSF6 was selected for subsequent experiment and analysis.

3.2. Identification of strain SSF6

The results of morphological identification demonstrated that the SSF6 colony had a smooth and gray surface surrounded by wrinkles, aerobic growth, positive Gram staining, and microscopic rod-shaped bacteria with spores (Figures 1Ad,e, B). Physiological and biochemical analyzes of strain SSF6 were performed using the Biolog GEN III MicroStation automated microbial identification system. There were 24 positive reactions in the carbon source utilization test, including the ability to use cellobiose, sucrose, and fructose as substrates (Supplementary Table S1). Strain SSF6 displayed sensitivity to L-alanine, L-aspartate, L-glutamic acid, and D-aspartate substrates (Supplementary Table S2) and sodium butyrate, sodium sulfite, lithium chloride, and sodium lactate (Supplementary Table S3). Amplified 16S rRNA fragment from strain SSF6 genomic DNA, a 1,500 bp fragment was obtained and submitted to NCBI and performed sequence blast analysis to construct a phylogenetic tree. The results showed that strain SSF6 was closely related to *B. amyloliquefaciens* strain BV2007 (MT613661.1), *B. velezensis* strain 2630 (MT611652.1), and *B. velezensis* strain FZB42 (ON041103.1; Supplementary Figure S1). Therefore, strain SSF6 was identified as *Bacillus* sp.

3.3. Determination of cellulase activity in the isolated strains

To further confirm the cellulose degradation ability of bacterial SSF6, the cellulase activity of the strain in CMC medium was determined, including filter paper activity (FPA), endoglucanase, exoglucanase, and β-glucosidase activities. The results illustrated that the filter paper cellulase activity of *B. velezensis* SSF6 was 64.48 ± 0.28 U/mL, endoglucanase activity was 54.39 ± 0.46 U/mL, exoglucanase activity was 78.59 ± 0.42 U/mL, and β-glucosidase activity was 58.96 ± 0.05 U/mL (Figure 1C). The results of the enzymatic reaction demonstrated that the rate of enzymatic reaction was influenced by the substrate concentration. Exoglucanase was most

obviously impacted by the microcrystalline cellulose concentration, and it had the highest reaction rate when the substrate concentration was saturated, followed by endoglucanase and β-glucosidase (Supplementary Table S4).

3.4. Characterization of the whole genome of strain SSF6

Bacterial lignocellulolytic activity can be better understood using genomic information. In this study, the genome of strain SSF6 was sequenced and the functional genes involved in lignocellulosic degradation were analyzed. Strain SSF6 was assembled into a circular genome after steps of assembly, correction and optimization (Figure 2A). Strain SSF6 had a genome size of 3,891,780 bp, contained 46.67% GC content, three contigs, and N50 contig length 3,893,584 bp (Supplementary Figure S2).
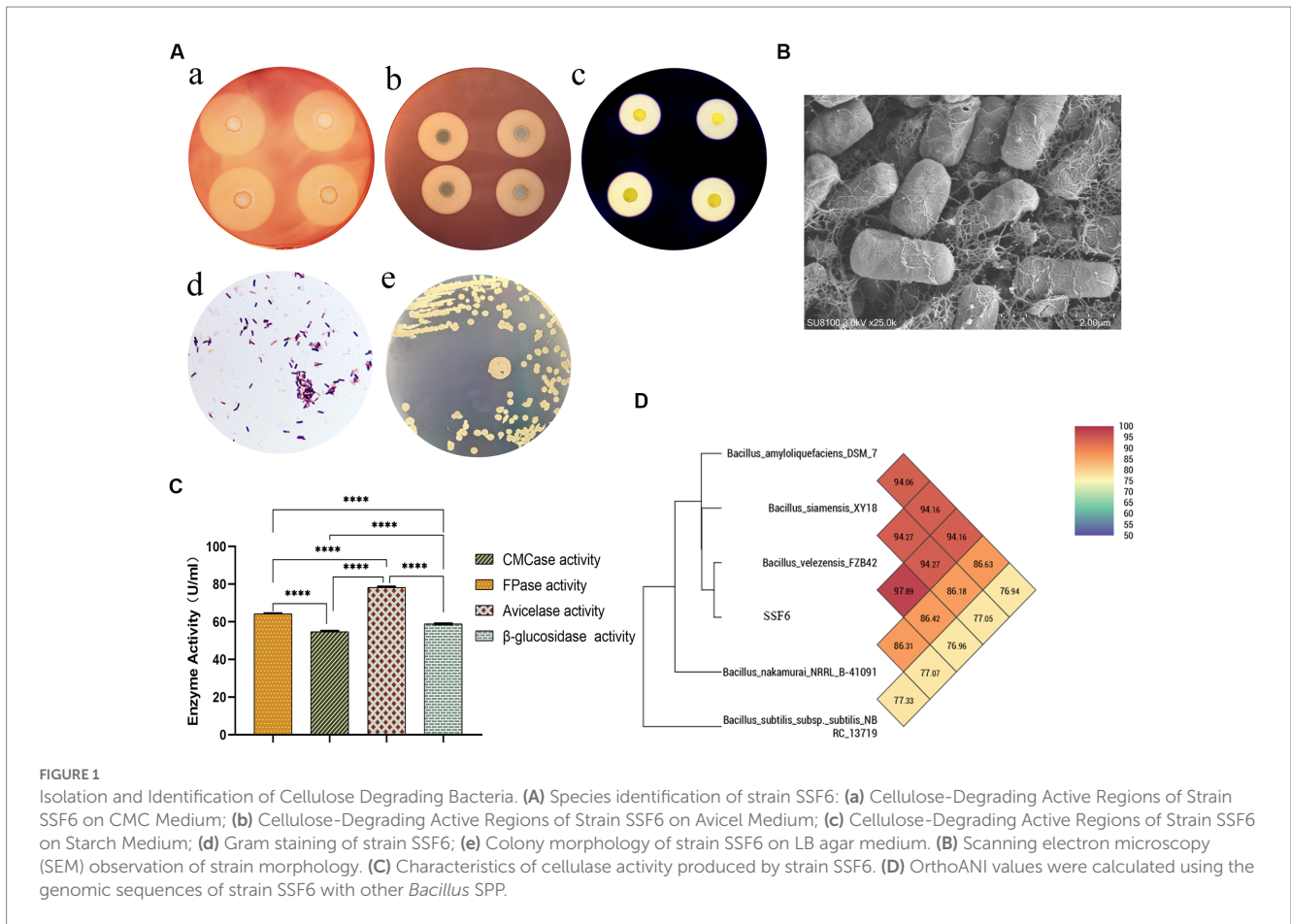
Through genomic analysis of the strain, 4,015 genes with a combined size of 3,499,518 bp were predicted in strain SSF6. The total number of repeats was 380,205 of which were scattered repeats (132 LTR, 19 DNA, 38 LINE, 13 SINE, 2 RC, and 1 unknown), and 175 were tandem repeats. The noncoding RNAs included 86 tRNAs, 27 rRNAs (9 16S rRNAs, 9 5S rRNAs, and 9 23S rRNAs), and 10 sRNAs. Eighteen prophages were predicted, with a total fragment length of 628,164 bp.

The ANI (95–96%) and DNA–DNA hybridization value (DDH,70%) calculated based on genomic nucleic acid sequences have become the gold standard for species classification (Choi et al., 2021). According to ANI calculation by PGAP, strain SSF6 was predicted to be *B. velezensis* with high confidence. The results showed that the top 8 were all *B. velezensis*, with ANI values greater than 97% (Supplementary Table S5). At the same time, OrthoANI values were calculated with five genomic sequences, including SSF6, and the results showed that strain SSF6 had the highest value compared with *B. velezensis* FZB42 (97.59%), followed by *B. siamensis* KCTC 13613 (94.34%) and *B. amyloliquefaciens* DSM 7 (94.02%; Figure 1D). Further calculation of OrthoANI values for 100 selected *B. velezensis* genomes showed that all *B. velezensis* genomes had OrthoANI values greater than 97% (Supplementary Table S6). The dDDH values of strain SSF6 and 100 strains *B. velezensis* genome sequences ranged from 96.32 to 99.1%, where the probability of DDH value ≥70% was greater than 90% (Supplementary Table S6). In summary, strain SSF6 was identified as *B. velezensis* and named *B. velezensis* SSF6.

3.5. Gene function annotation

According to the NR database, the number of *B. velezensis* genes annotated by *B. velezensis* SSF6 was the highest (1846), followed by *B. amyloliquefaciens* (1090). This further confirmed that the strain SSF6 was *B. velezensis* (Figure 2B). COG database annotation revealed that the most enriched genes were those involved in the transport and metabolism of amino acids (302 genes) and carbohydrates (249 genes; Supplementary Figure S3). A total of 112 COGs were annotated as being involved in carbohydrate metabolism, including COG2814 (predicted arabinose efflux permease AraJ, MFS family), COG0726 (Peptidoglycan/xylan/chitin deacetylase, PgdA/NodB/CDA1 family), COG1349 (DNA-binding transcriptional regulator of sugar

FIGURE 1
Isolation and Identification of Cellulose Degrading Bacteria. **(A)** Species identification of strain SSF6: **(a)** Cellulose-Degrading Active Regions of Strain SSF6 on CMC Medium; **(b)** Cellulose-Degrading Active Regions of Strain SSF6 on Avicel Medium; **(c)** Cellulose-Degrading Active Regions of Strain SSF6 on Starch Medium; **(d)** Gram staining of strain SSF6; **(e)** Colony morphology of strain SSF6 on LB agar medium. **(B)** Scanning electron microscopy (SEM) observation of strain morphology. **(C)** Characteristics of cellulase activity produced by strain SSF6. **(D)** OrthoANI values were calculated using the genomic sequences of strain SSF6 with other *Bacillus* SPP.

TABLE 3 HCR determination for the diameter of the degradation circle of the isolated strain on Avicel medium.

Strain	Clearing zone diameter (D, mm)	Colony diameter (d mm)	HCR (D/d value)
SSF1	-	-	-
SSF4	-	-	-
SSF6	21.37 ± 0.32	6.70 ± 0.16	3.20 ± 0.12
SSF15	-	-	-

Different lowercase letters mean the significance at $p < 0.05$. D, The diameter of the hydrolysis circle; d, The diameter of the colony; HCR (D/d value), The hydrolysis ratio of the strain.

metabolism, DeoR/GlpR family), and COG0697 (Permease of the drug/metabolite transporter (DMT) superfamily).
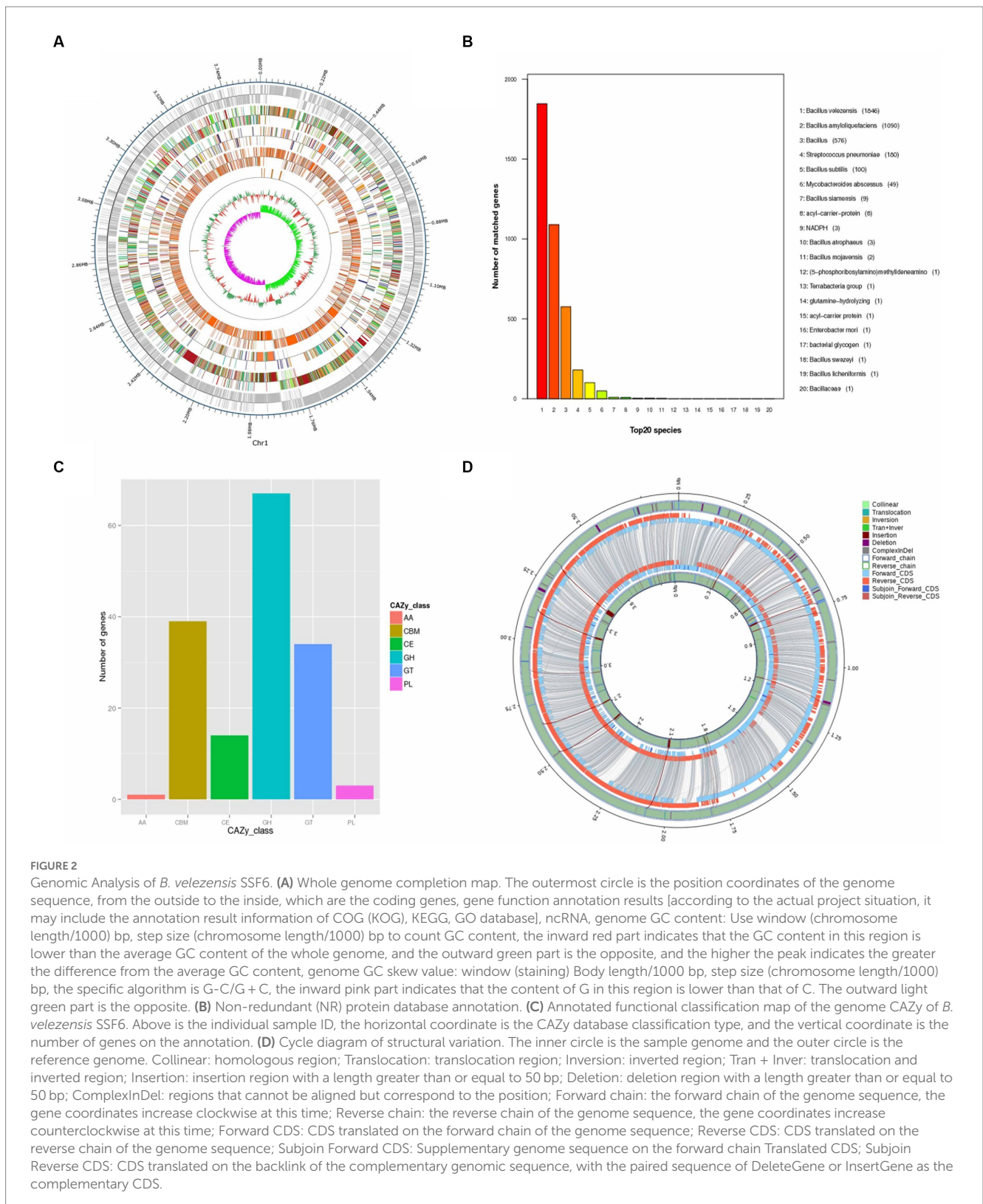
KEGG integrates genomic, chemical, and system function information and can graphically represent many metabolic pathways and the relationship between various pathways to comprehensively elucidate metabolic pathways. A total of 375 genes related to carbohydrate metabolism were annotated in KEGG (Supplementary Figure S4). Amino sugar and nucleotide sugar metabolism (ko00520, 41 genes), pyruvate metabolism (ko00620, 39 genes), glycolysis/gluconeogenesis (ko00010, 36 genes), starch and sucrose metabolism (ko00500, 34 genes), and the pentose phosphate pathway (ko00030, 25 genes) were the dominant energy metabolism pathways that play a key role in cellulose degradation

(Supplementary Table S7). These findings suggested that the *B. velezensis* SSF6 genome contained many genes necessary for metabolizing of carbohydrates and other nutrients, indicating that the strain SSF6 had a significant capacity for carbohydrate polysaccharide degradation.

The genome contained 158 CAZyme genes (3.9% of the total number of genes) with five major classifications: glycoside hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs) (Figure 2C). A total of 67 GH genes were annotated in the genome and assigned to 36 GH families, with the GH1, GH23, GH13-3, and GH4 having the most members. Additionally, carbohydrate-binding module (CBM) genes were identified and assigned to 6 CBM families, which enhanced the catalytic activity by targeting enzymes associated with specific cell wall components (Duan et al., 2017). Moreover, 34 glycosyl transferase (GT) genes were assigned to 9 GT families, along with 3 polysaccharide lyase (PL) genes, and 1 auxiliary activity (AA) gene.

3.6. Comparative genomic analysis

The results of genome comparison between *B. velezensis* SSF6 and *B. velezensis* FZB42 showed that the number of genes of the former was greater than that of the latter. In addition to providing molecular evidence for phenotypic differences and similarities, the study of the core genome is of great significance for determining functional

FIGURE 2
Genomic Analysis of *B. velezensis* SSF6. **(A)** Whole genome completion map. The outermost circle is the position coordinates of the genome sequence, from the outside to the inside, which are the coding genes, gene function annotation results [according to the actual project situation, it may include the annotation result information of COG (KOG), KEGG, GO database], ncRNA, genome GC content: Use window (chromosome length/1000) bp, step size (chromosome length/1000) bp to count GC content, the inward red part indicates that the GC content in this region is lower than the average GC content of the whole genome, and the outward green part is the opposite, and the higher the peak indicates the greater the difference from the average GC content, genome GC skew value: window (staining) Body length/1000 bp, step size (chromosome length/1000) bp, the specific algorithm is G-C/G + C, the inward pink part indicates that the content of G in this region is lower than that of C. The outward light green part is the opposite. **(B)** Non-redundant (NR) protein database annotation. **(C)** Annotated functional classification map of the genome CAZy of *B. velezensis* SSF6. Above is the individual sample ID, the horizontal coordinate is the CAZy database classification type, and the vertical coordinate is the number of genes on the annotation. **(D)** Cycle diagram of structural variation. The inner circle is the sample genome and the outer circle is the reference genome. Collinear: homologous region; Translocation: translocation region; Inversion: inverted region; Tran + Inver: translocation and inverted region; Insertion: insertion region with a length greater than or equal to 50 bp; Deletion: deletion region with a length greater than or equal to 50 bp; ComplexInDel: regions that cannot be aligned but correspond to the position; Forward chain: the forward chain of the genome sequence, the gene coordinates increase clockwise at this time; Reverse chain: the reverse chain of the genome sequence, the gene coordinates increase counterclockwise at this time; Forward CDS: CDS translated on the forward chain of the genome sequence; Reverse CDS: CDS translated on the reverse chain of the genome sequence; Subjoin Forward CDS: Supplementary genome sequence on the forward chain Translated CDS; Subjoin Reverse CDS: CDS translated on the backlink of the complementary genomic sequence, with the paired sequence of DeleteGene or InsertGene as the complementary CDS.

differences and similarities between the strains. Core genome analyzes of 2 *Bacillus* genomes were performed. The total number of core genes was 3,341. SSF6-specific genes (557) were more abundant than FZB42-specific genes (340). A total of 156 genes were annotated as having unknown functions (Figure 2D). The core genes related to carbohydrates were annotated to 27 GH (46 genes), 6 CE (13 genes), 8 GT (29 genes), 1 AA (1 gene), 6 CBM (32 genes), and 3 PL (3 genes) families. However, no CAZyme-encoding genes were detected in FZB42, such as the GH13-5 and GH43-8 families. There were 16,027 nonsynonymous SNPs in the SSF6 genome, distributed among 24

glycoside hydrolase families, including GH1, GH4, and GH23 (Supplementary Table S8).

3.7. Lignocellulose gene analysis

Genes involved in lignocellulose degradation have been detected in the genome of *B. velezensis* SSF6. A total of 26 cellulase genes were annotated in the *B. velezensis* SSF6 genome, including 2 endoglucanase genes, 10 exoglucanase genes, and 14 β-glucosidase genes (Table 4). The proteins encoded by the endoglucanase, exoglucanase, β-glucosidase genes belong to the GH51, GH1, and GH3 and GH4 families. The number of cellulase genes and family species was similar in SSF6 and FZB42, indicating that *B. velezensis* SSF6 has a strong cellulose degradation potential.

Bacillus velezensis SSF6 contained 28 genes encoding hemicellulases, including 12 β-xylosidase genes, 9 xylanase genes, and 7 additional enzyme genes (mannosidase, α-L-arabinofuranosidase, and polyarabinose exonuclease genes; Table 4). The β-Xylosidases primarily belong to the GH1, GH3, and GH51 families, the xylanases to the CE4 and CE7 families, and the mannosidases, α-L-arabinofuranosidases, and polyarabinose exonucleases to the GH1, GH3, and GH53 families (Table 4).

4. Discussion

Recently, a range of microbes capable of degrading lignin, hemicellulose, and cellulose have been identified in microecological settings, such as soil, compost, anaerobic sludge, and plants, including *Serratia marcescens* (Tang et al., 2022), *B. velezensis* (Li Y. et al., 2020), *Paenibacillus* (Yadav and Dubey, 2018), *Cellulomonas*, *Cytophaga* (Tan et al., 2022) and other microorganisms. Based on different cellulose degradation mechanisms, these microorganisms produce a variety of cellulase with industrial value, resulting in huge economic value. The examination of *B. velezensis* focused primarily on the improvement in enzyme production conditions. At present, the related research of *B. velezensis* mainly focuses on the screening of strains, the optimization of fermentation conditions (Nair et al., 2018; Li F. et al., 2020; Djelid et al., 2022), and the prediction of lignocellulosic degradation function of strains by high-throughput sequencing technology and the mining of corresponding genes (Chen et al., 2018; Tang et al., 2021). As a member of a class of microorganisms with cellulose degradation potential, the *B. velezensis* genome contains abundant carbohydrase genes. They can secrete various cellulases. In this study, genome-wide and comparative genomic analyzes were performed for *B. velezensis* SSF6 to uncover functional genes involved in cellulose degradation. This study will benefit the mining of cellulose-degrading enzyme resources and the development and utilization of *B. velezensis* strains.

Bacillus velezensis SSF6 had a total cellulose enzyme activity of 64.48 ± 0.28 U/mL, which was higher than that of *B. velezensis* M2 (33.03 U/mL) (Li F. et al., 2020) and *B. licheniformis* KY962963 (6.19 IU/mL) (Shah and Mishra, 2020). Endoglucanase activity was 54.39 ± 0.46 U/mL, exoglucanase activity was 78.59 ± 0.42 U/mL, and β-glucosidase activity was 58.96 ± 0.05 U/mL, which was higher than that of *B. stratosphericus* BHUJPV-H5 (0.35, 0.02, and 1.33 U/mL, respectively), *B. subtilis* BHUJPV-H12 (0.21, 0.03, and 1.24 U/mL,

TABLE 4 Annotated common genes encoding lignocellulose degrading enzymes of *B. velezensis* SSF6 strains.

Classifcation	CAZy	Count	Predicted function	EC numbers
Cellulose-related	GH1	8	beta-glucosidase	EC 3.2.1.21
	GH1	8	6-phospho-beta-glucosidase	EC 3.2.1.86
	GH1	8	6-phospho-beta-galactosidase	EC 3.2.1.85
	GH3	2	beta-glucosidase	EC 3.2.1.21
	GH4	4	6-phospho-beta-glucosidase	EC 3.2.1.86
	GH4	4	α-glucosidase	EC 3.2.1.20
	GH5	1	endo-1,4-β-glucanase	EC 3.2.1.4
	GH32	3	endo-levanase	EC 3.2.1.65
	GH51	1	endoglucanase	EC 3.2.1.4
Hemicellulose-related	GH1	8	beta-mannosidase	EC 3.2.1.25
	GH1	8	beta-glycosidase	EC 3.2.1.-
	GH1	8	beta-xylosidase	EC 3.2.1.37
	GH3	2	beta-xylosidase	EC 3.2.1.37
	GH3	2	alpha-L-arabinofuranosidase	EC 3.2.1.55
	GH26	1	beta-mannanase	EC 3.2.1.78
	GH51	2	beta-xylosidase	EC 3.2.1.37
	GH51	2	alpha-L-arabinofuranosidase	EC 3.2.1.55
	GH53	1	endo-beta-1,4-galactanase	EC 3.2.1.89
	CE4	8	acetyl xylan esterase	EC 3.2.1.72
	CE7	1	acetyl xylan esterase	EC 3.2.1.72

respectively), *B. subtilis* BHUJPV-H19 (0.23, 0.01, and 2.55 U/mL, respectively), and *B. subtilis* BHUJPV-H23 (0.26, 0.02, and 1.87 U/mL, respectively) (Singh et al., 2019). Normally, cellulose is exploited by the synergistic action of three enzymes, endoglucanase, exoglucanase, and β-glucosidase, which hydrolyze cellulose to glucose monomers (Horn et al., 2012). Comparative analysis revealed that *B. velezensis* SSF6 has a strong comprehensive cellulose degradation ability and may be widely used in agriculture in the future. In the future, enzyme production conditions can be further improved and large-scale cellulase production can be achieved through biotechnology, thereby creating significant economic value (Chen et al., 2018; Zhang et al., 2018).

Genomics can provide information about the functional potential of microorganisms. Similar to *B. velezensis* FZB42, the genome size of *B. velezensis* SSF6 was estimated to be 3,891,780 bp with a 46.67% GC content. Through various database annotations, a significant number of amino acid transport, carbohydrate, and metabolic activities have been anticipated. Additionally, many glycoside hydrolases (GHs) are involved in carbohydrate metabolism (An et al., 2021). It contains various hydrolases that act on glycosidic linkages and can hydrolyse polysaccharide substances, such as cellulose, starch, xylan, and mannose (Du et al., 2021). The GH13 family has 36 members, including α-amylase (EC 3.2.1.1) (Janeček et al., 2014), cyclodextrin glucosyltransferase (EC 2.4.1.19), and α-glucosidase (EC 3.2.1.20) (Zhou et al., 2015), and can hydrolyse starch. Both GH1 and GH4 have the potential for cellulose degradation and can effectively degrade lignocellulosic biomass. GH43 is an important component of xylan degradation and is related to hemicellulose degradation.

In conclusion, these carbohydrases catalyze carbohydrate degradation, modification, and biosynthesis and have many applications. The CBM family enhances the catalytic activity by targeting enzymes linked to specific cell wall components (Duan et al., 2017). The CBM5 module improves the affinity of enzymes, such as endoglucanase, chitinase, and lytic polysaccharide monooxygenase (LPMO), for crystalline cellulose and chitin, improving their efficiency in binding to substrates over a wider pH range (Manjeet et al., 2019). In Bcl PMO10A, CBM5 promotes substrate binding and protects the enzyme from deactivation (Mutahir et al., 2018). CE7 contains acetyl xylan esterase and cephalosporin C deacetylase, which are important catalytic enzymes for synthesizing cephalosporin antibiotics (Vincent et al., 2003). CE4 and CE7 promote xylan dissolution and are involved in hemicellulose degradation. The AA10 family's LPMOs catalyze the oxidative degradation of crystalline polysaccharides, such as cellulose and chitin. Additionally, they act on the xylan, mannan, and cellulose structures of lignocellulosic biomass to provide more binding checkpoints for glycoside hydrolases, thereby enhancing the accessibility of cellulases to substrates and promoting substrate degradation (Pierce et al., 2017).

Several researchers have highlighted the lignocellulose degradation capabilities of *B. velezensis* and predicted cellulase interactions by elucidating the related capabilities of carbohydrases in the genome (Chen et al., 2022a). Furthermore, abundant cellulases and xylanases increase the release of monosaccharides during straw saccharification through synergistic effects (Zeng et al., 2021), which promotes the conversion of cellulose and hemicellulose into soluble sugar. The genome of *B. velezensis* SSF6 used in this study was rich in cellulase and hemicellulase genes. Compared with *B. velezensis* FZB42, strain SSF6 had more GT4 and GH28 genes. GT4 had the most genes among the GT families. These enzymes utilize not only nucleotide sugar donors but also simple phosphosaccharide and lipid

phosphosaccharide donors and have potential therapeutic implications. The GH28 and CE8 families are classified as pectinases and play important roles in pectin degradation. These results indicated that *B. velezensis* SSF6 can as the potential to degrade lignocellulose. However, some genes still have not been annotated, and it is still unclear whether the anticipated gene-phenotype can obtain through expression of these genes, so further investigation and inquiry are still required in future scientific research efforts.

5. Conclusion

Owing to its potential to offer additional value, such as sustainable energy from biomass and leftover agricultural resources, cellulose conversion is a topic of interest in biotechnology. The results of the present study showed that *B. velezensis* strain SSF6 has an efficient capacity for cellulose degradation. Whole genome sequencing and comparative genomics analysis of *B. velezensis* SSF6 revealed that it contained a large number of genes from different glycosyl hydrolase (GH) families that are essential for cellulose and hemicellulose biodegradation, indicating that strain SSF6 has broad application prospects for industrial enzyme production in the future. In this study, the genetic basis of lignocellulosic degradation was revealed through genome sequencing and analysis, which provided a new microbial resource for lignocellulosic degradation.

Data availability statement

The datasets presented in this study can be found in NCBI public database. The accession number for genome sequence of *Bacillus velezensis* strain SSF6 is PRJNA934860. The 16S rRNA gene sequence of *Bacillus velezensis* SSF6 strain is OR229485.

Author contributions

YG and YL: sample collection and strain culture. JM, CC, and LY: experimental study. TZ and SW: research supervision, data analytics, and manuscript writing. YC and YR: critical revision of a manuscript. SS: design of research ideas and key modifications to manuscripts. ZL and XZ: research sponsors and key changes to manuscripts. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1167293/full#supplementary-material

References

Abd Elhameed, E., Sayed, A. R. M., Radwan, T. E. E., and Hassan, G. (2020). Biochemical and molecular characterization of five *Bacillus* isolates displaying remarkable carboxymethyl cellulase activities. *Curr. Microbiol.* 77, 3076–3084. doi: 10.1007/s00284-020-02135-8

Ahmed, W., Qaswar, M., Jing, H., Wenjun, D., Geng, S., Kailou, L., et al. (2020). Tillage practices improve rice yield and soil phosphorus fractions in two typical paddy soils. *J. Soils Sediments* 20, 850–861. doi: 10.1007/s11368-019-02468-3

Alonso-Pernas, P., Bartram, S., Arias-Cordero, E. M., Novoselov, A. L., Halty-deLeon, L., Shao, Y., et al. (2017). In vivo isotopic labeling of symbiotic Bacteria involved in cellulose degradation and nitrogen recycling within the gut of the Forest cockchafer (*Melolontha hippocastani*). *Front. Microbiol.* 8:1970. doi: 10.3389/fmicb.2017.01970

An, X., Chen, X., and Wang, Y. (2021). Cellulolytic bacterium characterization and genome functional analysis: an attempt to lay the foundation for waste management. *Bioresour. Technol.* 321:124462. doi: 10.1016/j.biortech.2020.124462

Anand, A. A. P., Vennison, S. J., Sankar, S. G., Prabhu, D. I. G., Vasan, P. T., Raghuraman, T., et al. (2010). Isolation and characterization of bacteria from the gut of *Bombyx mori* that degrade cellulose, xylan, pectin and starch and their impact on digestion. *J. Insect Sci.* 10:107. doi: 10.1673/031.010.10701

Ardui, S., Ameur, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res.* 46, 2159–2168. doi: 10.1093/nar/gky066

Ayumi, S., Junjarus, S., Phakhinee, T., and Sirilak, B. (2018). Characterization of an anaerobic, thermophilic, alkaliphilic, high lignocellulosic biomass-degrading bacterial community, ISHI-3, isolated from biocompost. *Enzym. Microb. Technol.* 118, 66–75. doi: 10.1016/j.enzmictec.2018.07.001

Bai, Y., Wang, L., Lu, Y., Yang, L., Zhou, L., Ni, L., et al. (2015). Effects of long-term full straw return on yield and potassium response in wheat-maize rotation. *J. Integr. Agric.* 14, 2467–2476. doi: 10.1016/S2095-3119(15)61216-3

Balderas-Ruíz, K. A., Bustos, P., Santamaria, R. I., González, V., Cristiano-Fajardo, S. A., Barrera-Ortíz, S., et al. (2020). *Bacillus velezensis* 83 a bacterial strain from mango phyllosphere, useful for biological control and plant growth promotion. *AMB Express* 10:163. doi: 10.1186/s13568-020-01101-8

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29, 2607–2618. doi: 10.1093/nar/29.12.2607

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238. doi: 10.1093/nar/gkn663

Chen, L., Chen, W., Zheng, B., Yu, W., Zheng, L., Qu, Z., et al. (2022a). Fermentation of NaHCO3-treated corn germ meal by *Bacillus velezensis* CL-4 promotes lignocellulose degradation and nutrient utilization. *Appl. Microbiol. Biotechnol.* 106, 6077–6094. doi: 10.1007/s00253-022-12130-7

Chen, L., Gu, W., Xu, H., Yang, G.-L., Shan, X.-F., Chen, G., et al. (2018). Complete genome sequence of *Bacillus velezensis* 157 isolated from *Eucommia ulmoides* with pathogenic bacteria inhibiting and lignocellulolytic enzymes production by SSF. *3 Biotech* 8, 114–110. doi: 10.1007/s13205-018-1125-2

Chen, L., Sun, S., Yao, B., Peng, Y., Gao, C., Qin, T., et al. (2022b). Effects of straw return and straw biochar on soil properties and crop growth: a review. *Front. Plant Sci.* 13:986763. doi: 10.3389/fpls.2022.986763

Choi, J., Nam, J., and Seo, M.-H. (2021). Complete genome sequence of *Bacillus velezensis* NST6 and comparison with the species belonging to operational group *B. amyloliquefaciens*. *Genomics* 113, 380–386. doi: 10.1016/j.ygeno.2020.12.011

Dar, M. A., Shaikh, A. A., Pawar, K. D., and Pandit, R. S. (2018). Exploring the gut of *Helicoverpa armigera* for cellulose degrading bacteria and evaluation of a potential strain for lignocellulosic biomass deconstruction. *Process Biochem.* 73, 142–153. doi: 10.1016/j.procbio.2018.08.001

Dashtban, M., Maki, M., Leung, K. T., Mao, C., and Qin, W. (2010). Cellulase activities in biomass conversion: measurement methods and comparison. *Crit. Rev. Biotechnol.* 30, 302–309. doi: 10.3109/07388551.2010.490938

Djelid, H., Flahaut, S., Vander Wauven, C., Oudjama, Y., Hiligsmann, S., Cornu, B., et al. (2022). Production of a halotolerant endo-1,4-β-glucanase by a newly isolated *Bacillus velezensis* H1 on olive mill wastes without pretreatment: purification and characterization of the enzyme. *Arch. Microbiol.* 204:681. doi: 10.1007/s00203-022-03300-2

Dobrzyński, J., Wróbel, B., and Górska, E. B. (2022). Cellulolytic properties of a potentially lignocellulose-degrading *Bacillus* sp. 8E1A strain isolated from bulk soil. *Agronomy* 12:665. doi: 10.3390/agronomy12030665

Du, C., Li, C., Cao, P., Li, T., Du, D., Wang, X., et al. (2021). *Massilia cellulosiltytica* sp. nov., a novel cellulose-degrading bacterium isolated from rhizosphere soil of rice (*Oryza sativa* L.) and its whole genome analysis. *Antonie Van Leeuwenhoek* 114, 1529–1540. doi: 10.1007/s10482-021-01618-3

Duan, C.-J., Huang, M.-Y., Pang, H., Zhao, J., Wu, C.-X., and Feng, J.-X. (2017). Characterization of a novel theme C glycoside hydrolase family 9 cellulase and its CBM-chimeric enzymes. *Appl. Microbiol. Biotechnol.* 101, 5723–5737. doi: 10.1007/s00253-017-8320-7

Fan, B., Blom, J., Klenk, H.-P., and Borriss, R. (2017). *Bacillus amyloliquefaciens*, *Bacillus velezensis*, and *Bacillus siamensis* form an "operational group *B. amyloliquefaciens*" within the *B. subtilis* species complex. *Front. Microbiol.* 8:22. doi: 10.3389/fmicb.2017.00022

Fan, B., Wang, C., Song, X., Ding, X., Wu, L., Wu, H., et al. (2018). *Bacillus velezensis* FZB42 in 2018: the gram-positive model strain for plant growth promotion and biocontrol. *Front. Microbiol.* 9:2491. doi: 10.3389/fmicb.2018.02491

Fatani, S., Saito, Y., Alarawi, M., Gojobori, T., and Mineta, K. (2021). Genome sequencing and identification of cellulase genes in *Bacillus paralicheniformis* strains from the Red Sea. *BMC Microbiol.* 21:254. doi: 10.1186/s12866-021-02316-w

Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261–D269. doi: 10.1093/nar/gku1223

Horn, S. J., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V. G. (2012). Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* 5:45. doi: 10.1186/1754-6834-5-45

Janeček, Š., Svensson, B., and MacGregor, E. A. (2014). α-Amylase: an enzyme specificity found in various families of glycoside hydrolases. *Cell. Mol. Life Sci.* 71, 1149–1170. doi: 10.1007/s00018-013-1388-z

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., et al. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* 34, D354–D357. doi: 10.1093/nar/gkj102

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, 277D–2280D. doi: 10.1093/nar/gkh063

Kang, M., Su, X., Yun, L., Shen, Y., Feng, J., Yang, G., et al. (2022). Evaluation of probiotic characteristics and whole genome analysis of *Bacillus velezensis* R-71003 isolated from the intestine of common carp (*Cyprinus carpio* L.) for its use as a probiotic in aquaculture. *Aquacult. Rep.* 25:101254. doi: 10.1016/j.aqrep.2022.101254

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12

Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H.-H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160

Lee, I., Ouk Kim, Y., Park, S.-C., and Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* 66, 1100–1103. doi: 10.1099/ijsem.0.000760

Li, W., Jaroszewski, L., and Godzik, A. (2002). Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics* 18, 77–82. doi: 10.1093/bioinformatics/18.1.77

Li, Y., Lei, L., Zheng, L., Xiao, X., Tang, H., and Luo, C. (2020). Genome sequencing of gut symbiotic *Bacillus velezensis* LC1 for bioethanol production from bamboo shoots. *Biotechnol. Biofuels* 13:34. doi: 10.1186/s13068-020-1671-9

Li, F., Xie, Y., Gao, X., Shan, M., Sun, C., Niu, Y. D., et al. (2020). Screening of cellulose degradation bacteria from min pigs and optimization of its cellulase production. *Electron. J. Biotechnol.* 48, 29–35. doi: 10.1016/j.ejbt.2020.09.001

Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.955

Lu, K., Jin, Q., Lin, Y., Lu, W., Li, S., Zhou, C., et al. (2020). Cell-free fermentation broth of *Bacillus velezensis* strain S3-1 improves Pak Choi nutritional quality and changes the bacterial community structure of the rhizosphere soil. *Front. Microbiol.* 11:2043. doi: 10.3389/fmicb.2020.02043

Ma, E., Zhang, G., Ma, J., Xu, H., Cai, Z., and Yagi, K. (2010). Effects of rice straw returning methods on N2O emission during wheat-growing season. *Nutr. Cycl. Agroecosyst.* 88, 463–469. doi: 10.1007/s10705-010-9369-1

Manjeet, K., Madhuprakash, J., Mormann, M., Moerschbacher, B. M., and Podile, A. R. (2019). A carbohydrate binding module-5 is essential for oxidative cleavage of chitin by a multi-modular lytic polysaccharide monooxygenase from *Bacillus thuringiensis* serovar kurstaki. *Int. J. Biol. Macromol.* 127, 649–656. doi: 10.1016/j.ijbiomac.2019.01.183

Mansour, A. A., Da Costa, A., Arnaud, T., and Lu-Chau, T. A. (2016). Review of lignocellulolytic enzyme activity analyses and scale-down to microplate-based assays. *Talanta* 150, 629–637. doi: 10.1016/j.talanta.2015.12.073

Marriott, P. E., Gómez, L. D., and McQueen-Mason, S. J. (2016). Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol.* 209, 1366–1381. doi: 10.1111/nph.13684

Menshawy, M. N., Abdel-Hamid, A. M., Mohamed, S. K., and El-Katatny, M. H. (2022). Isolation and molecular identification of cellulose/hemicellulose degrading bacteria from agricultural compost and determination of their hydrolytic potential. *S. Afr. J. Bot.* 149, 617–621. doi: 10.1016/j.sajb.2022.06.045

Mutahir, Z., Mekasha, S., Loose, J. S. M., Abbas, F., Vaaje-Kolstad, G., Eijsink, V. G. H., et al. (2018). Characterization and synergistic action of a tetra-modular lytic polysaccharide monooxygenase from *Bacillus cereus*. *FEBS Lett.* 592, 2562–2571. doi: 10.1002/1873-3468.13189

Nair, A. S., Al-Battashi, H., Al-Akzawi, A., Annamalai, N., Gujarathi, A., Al-Bahry, S., et al. (2018). Waste office paper: a potential feedstock for cellulase production by a novel strain *Bacillus velezensis* ASN1. *Waste Manag.* 79, 491–500. doi: 10.1016/j.wasman.2018.08.014

Nawrocki, E. P., Burge, S. W., Bateman, A., Daub, J., Eberhardt, R. Y., Eddy, S. R., et al. (2015). Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 43, D130–D137. doi: 10.1093/nar/gku1063

Pereira, J. Q., Ritter, A. C., Cibulski, S., and Brandelli, A. (2019). Functional genome annotation depicts probiotic properties of *Bacillus velezensis* FTC01. *Gene* 713:143971. doi: 10.1016/j.gene.2019.143971

Pierce, B. C., Agger, J. W., Zhang, Z., Wichmann, J., and Meyer, A. S. (2017). A comparative study on the activity of fungal lytic polysaccharide monooxygenases for the depolymerization of cellulose in soybean spent flakes. *Carbohydr. Res.* 449, 85–94. doi: 10.1016/j.carres.2017.07.004

Pollegioni, L., Tonin, F., and Rosini, E. (2015). Lignin-degrading enzymes. *FEBS J.* 282, 1190–1213. doi: 10.1111/febs.13224

Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A., et al. (2006). The path forward for biofuels and biomaterials. *Science* 311, 484–489. doi: 10.1126/science.1114736

Reiner, J., Pisani, L., Qiao, W., Singh, R., Yang, Y., Shi, L., et al. (2018). Cytogenomic identification and long-read single molecule real-time (SMRT) sequencing of a Bardet-Biedl syndrome 9 (BBS9) deletion. *NPJ Genom. Med.* 3:3. doi: 10.1038/s41525-017-0042-3

Saha, S., Bridges, S., Magbanua, Z. V., and Peterson, D. G. (2008). Empirical comparison of ab initio repeat finding programs. *Nucleic Acids Res.* 36, 2284–2294. doi: 10.1093/nar/gkn064

Shah, F., and Mishra, S. (2020). In vitro optimization for enhanced cellulose degrading enzyme from *Bacillus licheniformis* KY962963 associated with a microalgae

Chlorococcum sp. using OVAT and statistical modeling. *SN Appl. Sci.* 2:1923. doi: 10.1007/s42452-020-03697-9

Sharma, A., Tewari, R., Rana, S. S., Soni, R., and Soni, S. K. (2016). Cellulases: classification, methods of determination and industrial applications. *Appl. Biochem. Biotechnol.* 179, 1346–1380. doi: 10.1007/s12010-016-2070-3

Shin, J.-H., Park, B.-S., Kim, H.-Y., Lee, K.-H., and Kim, K. S. (2021). Antagonistic and plant growth-promoting effects of *Bacillus velezensis* BS1 isolated from rhizosphere soil in a pepper field. *Plant Pathol. J.* 37, 307–314. doi: 10.5423/PPJ.NT.03.2021.0053

Singh, S., Jaiswal, D. K., Sivakumar, N., and Verma, J. P. (2019). Developing efficient thermophilic cellulose degrading consortium for glucose production from different agro-residues. *Front. Energy Res.* 7:61. doi: 10.3389/fenrg.2019.00061

Song, S., Jeon, E. K., and Hwang, C.-W. (2022). Characteristic analysis of soil-isolated *Bacillus velezensis* HY-3479 and its antifungal activity against Phytopathogens. *Curr. Microbiol.* 79, 357–311. doi: 10.1007/s00284-022-03060-8

Tan, Y., Song, W., Gao, L., Zhang, W., and Lu, X. (2022). *Cytophaga hutchinsonii* chu_2177, encoding the O-antigen ligase, is essential for cellulose degradation. *J. Microbiol.* 60, 364–374. doi: 10.1007/s12275-022-1531-3

Tang, H., Li, Y.-Q., and Li, Z. (2022). Efficient saccharification of bamboo biomass by secretome protein of the cellulolytic bacterium *Serratia marcescens* LY1 based on whole-genome and secretome analysis. *Renew. Energy* 193, 32–40. doi: 10.1016/j.renene.2022.04.146

Tang, H., Zheng, L., Li, Y., Lei, L., Yang, X., and Luo, C. (2021). Comparative genomic and secretomic characterisation of endophytic *Bacillus velezensis* LC1 producing bioethanol from bamboo lignocellulose. *Arch. Microbiol.* 203, 3089–3099. doi: 10.1007/s00203-021-02306-6

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569

Teather, R. M., and Wood, P. J. (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Appl. Environ. Microbiol.* 43:777. doi: 10.1128/aem.43.4.777-780.1982

Tian, X., Yang, T., He, J., and Chu, Q. (2017). Fungal community and cellulose-degrading genes in the composting process of Chinese medicinal herbal residues. *Bioresour. Technol.* 241, 374–383. doi: 10.1016/j.biortech.2017.05.116

Vincent, F., Charnock, S. J., Verschueren, K. H. G., Turkenburg, J. P., Scott, D. J., Offen, W. A., et al. (2003). Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the *Bacillus subtilis* enzyme at 1.9A resolution. *J. Mol. Biol.* 330, 593–606. doi: 10.1016/S0022-2836(03)00632-6

Wang, X., Tian, L., Li, Y., and Zhong, C. (2022). Effects of exogenous cellulose-degrading bacteria on humus formation and bacterial community stability during composting. *Bioresour. Technol.* 359:127458. doi: 10.1016/j.biortech.2022.127458

Wang, Z. Y., Wang, R. X., Zhou, J. S., Cheng, J. F., and Li, Y. H. (2020). An assessment of the genomics, comparative genomics and cellulose degradation potential of *Mucilaginibacter polytrichastri* strain RG4-7. *Bioresour. Technol.* 297:122389. doi: 10.1016/j.biortech.2019.122389

Xu, W., Zhang, L., Goodwin, P. H., Xia, M., Zhang, J., Wang, Q., et al. (2020). Isolation, identification, and complete genome assembly of an endophytic *Bacillus velezensis* YB-130, potential biocontrol agent against *Fusarium graminearum*. *Front. Microbiol.* 11:598285. doi: 10.3389/fmicb.2020.598285

Yadav, S., and Dubey, S. K. (2018). Cellulose degradation potential of *Paenibacillus lautus* strain BHU3 and its whole genome sequence. *Bioresour. Technol.* 262, 124–131. doi: 10.1016/j.biortech.2018.04.067

Yang, W., Li, X., and Zhang, Y. (2022). Research Progress and the development trend of the utilization of crop straw biomass resources in China. *Front. Chem.* 10:904660. doi: 10.3389/fchem.2022.904660

Zeng, J., Huang, W., Tian, X., Hu, X., and Wu, Z. (2021). Brewer's spent grain fermentation improves its soluble sugar and protein as well as enzymatic activities using *Bacillus velezensis*. *Process Biochem.* 111, 12–20. doi: 10.1016/j.procbio.2021.10.016

Zhang, Z.-Y., Raza, M. F., Zheng, Z., Zhang, X., Dong, X., and Zhang, H. (2018). Complete genome sequence of *Bacillus velezensis* ZY-1-1 reveals the genetic basis for its hemicellulosic/cellulosic substrate-inducible xylanase and cellulase activities. *3 Biotech* 8:465. doi: 10.1007/s13205-018-1490-x

Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485

Zhou, C., Xue, Y., and Ma, Y. (2015). Evaluation and directed evolution for thermostability improvement of a GH 13 thermostable α-glucosidase from *Thermus thermophilus* TC11. *BMC Biotechnol.* 15:97. doi: 10.1186/s12896-015-0197-x