

## COMPLETE GENOME SEQUENCING AND GENE FUNCTION PREDICTION OF *BACILLUS VELEZENSIS* ANY11

LONG, W.<sup>1†#</sup> – CHEN, Y.<sup>1,2#</sup> – XIN, W.<sup>1,2#</sup> – TANG, J.<sup>1,2</sup> – XIAO, X.<sup>1</sup> – MOU, B.<sup>1</sup> – LI, J.<sup>1</sup> – LU, F.<sup>1</sup> –  
ZHANG, T.<sup>1</sup> – LIAO, H.<sup>1</sup> – HAN, X.<sup>3</sup> – YANG, L.<sup>1</sup> – MA, M.<sup>2</sup> – YANG, Y.<sup>1\*</sup> – WANG, H.<sup>1\*</sup> – LI, J.<sup>1\*</sup>

<sup>1</sup>*Bamboo Pests Control and Resources Development Key Laboratory of Sichuan Province,  
College of Life Science, Leshan Normal University, Leshan, Sichuan 614000, China*

<sup>2</sup>*College of Resources, Sichuan Agricultural University, Wenjiang, Chengdu, Sichuan 611130,  
China*

<sup>3</sup>*College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, Hunan  
410128, China  
(phone: +86-181-8834-0209)*

<sup>#</sup>*These authors contributed equally*

<sup>\*</sup>*Corresponding authors*

*e-mail: rsyyj@126.com (Y. Yang); lijiafengleshan@163.com (J. Li);  
hanyu.wang6143@outlook.com (H. Wang)*

(Received 3<sup>rd</sup> Jan 2025; accepted 19<sup>th</sup> Mar 2025)

**Abstract.** *Bacillus velezensis*, a member of phylum Firmicutes, is a Gram-positive rod-shaped and endospore-forming bacterium inhabiting diverse environments. The preliminary analysis of the whole genome of *Bacillus velezensis* ANY11, isolated from bamboo fiber humus in the intestines of *Cyrtotrachelus buqueti*, revealed that the genome is approximately 3,949,880 bp in length and contains 4265 coding genes. Among these, 12815, 2473, 3193, and 171 genes were annotated in the GO, KEGG, COG, and CAZy databases, respectively. Additionally, the virulence, pathogenicity, and antibiotic resistance of *Bacillus velezensis* ANY11 were analyzed using PHI, VFDB, and CARD databases. Based on genomic sequencing and gene function analysis, *Bacillus velezensis* is believed to possess certain disease resistance and the capability to hydrolyze lignocellulose, suggesting its potential role in plant disease prevention and control, as well as in the hydrolysis and reuse of lignocellulose. The bioinformatics from this study may lay a foundation for the production of biofertilizers and biopesticides, as well as for biomass utilization in clean ethanol production.

**Keywords:** *Bacillus velezensis*, gene annotation, gene function prediction, whole genome, whole genome sequencing

### Introduction

*Bacillus velezensis* was first isolated from the mouth of the Vélez River in Málaga, southern Spain, in 2005 (Wang et al., 2019). This microorganism is widely distributed across various environments, including plant rhizospheres, soil, rivers, human food, animal intestines, and seawater, making it easy to isolate and cultivate (Ye et al., 2018). As a Gram-positive, rod-shaped multifunctional bacterium belonging to the genus *Bacillus*, *B. velezensis* can produce spores under unfavorable conditions. These spores can be converted into powder without killing the bacteria, providing a significant advantage over other biocontrol bacteria (Laloo et al., 2010).

Due to its harmlessness to humans and animals and its environmental friendliness, *B. velezensis* has been commercially utilized as a biofertilizer and biopesticide. Extensive research has demonstrated its broad-spectrum antifungal activity against various plant

pathogenic fungi, making it effective for the prevention and control of plant diseases (Park et al., 2019). Significant research is currently focused on the antifungal activity of *B. velezensis*. Genomic sequencing has been conducted to predict the gene clusters responsible for synthesizing secondary metabolites and exploring antibacterial substances and their mechanisms. By investigating the antibacterial genes of these strains, our study provides a bioinformatics foundation for the efficient development and application of *Bacillus velezensis* in agricultural production.

Lignocellulose, a widely distributed, renewable, and abundantly available biomass resource, is an essential raw material for bioethanol production. However, the hydrolysis of lignocellulose, particularly the degradation of cellulose, poses a significant challenge in the production of lignocellulosic bioethanol (Mansour et al., 2016). In nature, several examples demonstrate lignocellulose degradation, with herbivorous insects being among the most dominant organisms. Within these insects, gut symbiotic microorganisms play a crucial role in breaking down lignocellulose (Zhang et al., 2008). Therefore, the intestines of scavenger insects serve as ideal sites for isolating lignocellulose-degrading microorganisms.

Studies have shown that *Bacillus velezensis* can produce highly active neutral enzymes, such as proteases and cellulases, which break down cellulose into oligosaccharides or monosaccharides (Peixoto et al., 2011). Furthermore, *Bacillus velezensis* contains genes encoding enzymes involved in the degradation of cellulose, xylan, lignin, starch, mannan, galactosides, and arabinose (Chen et al., 2018). Previous research isolated *Bacillus velezensis* LC1 from bamboo fiber humus, and whole-genome sequencing, along with cellulase activity measurements, indicated that this symbiotic bacterium possesses the ability to degrade lignocellulose and exhibits cellulase activity, highlighting its potential to further convert bamboo lignocellulosic components into ethanol (Li et al., 2020).

In this experiment, *Bacillus velezensis* ANY11, isolated from bamboo fiber humus in the intestines of *Cyrtotrachelus buqueti*, was used. The whole-genome sequencing, genomic analysis, and functional annotation conducted in this study provide a foundation for future research into its potential applications in biofertilizers, plant disease and pest control, as well as biocatalytic production of renewable ethanol.

## Experimental materials and procedures

### *Test materials*

*Bacillus velezensis* ANY11 was isolated from bamboo fiber humus in the intestines of *Cyrtotrachelus buqueti*.

LB Medium: 0.5% yeast extract, 1% peptone, 1% sodium chloride, pH adjusted to 7.0.

### *Test methods*

The *Bacillus velezensis* strain was inoculated into LB liquid medium and cultured with constant shaking at 220 r/min at 30°C for 12-24 h to activate the bacteria. After activation, the bacterial cells were collected by centrifugation.

Genomic DNA from the sample was extracted using the STE method. The purity and integrity of the DNA were subsequently assessed by agarose gel electrophoresis, and quantitation was performed using Qubit. Large fragments were then recovered using the BluePippin automated nucleic acid fragment recovery system, followed by end repair

and A-tailing. The samples were then mixed in equimolar amounts, and adapter ligation was performed using the SQK-LSK109 ligation kit (Oxford Nanopore Technologies) to construct a 1D library. Finally, sequencing was carried out using the Nanopore platform. The DNA samples that passed the quality check of electrophoresis were randomly fragmented into approximately 350 bp-long segments using an ultrasonic disruptor. The processed DNA fragments were then used for library preparation. After passing library quality control, the different libraries were sequenced on both the Nanopore PromethION and Illumina NovaSeq PE150 platforms based on their effective concentrations and targeted sequencing outputs. With the data of each sample passing the quality control, a second-generation and third-generation sequencing was performed using Unicycler software (Wick et al., 2017). Chromosome and plasmid sequences were screened, and the chromosome sequences were assembled into a circular genome.

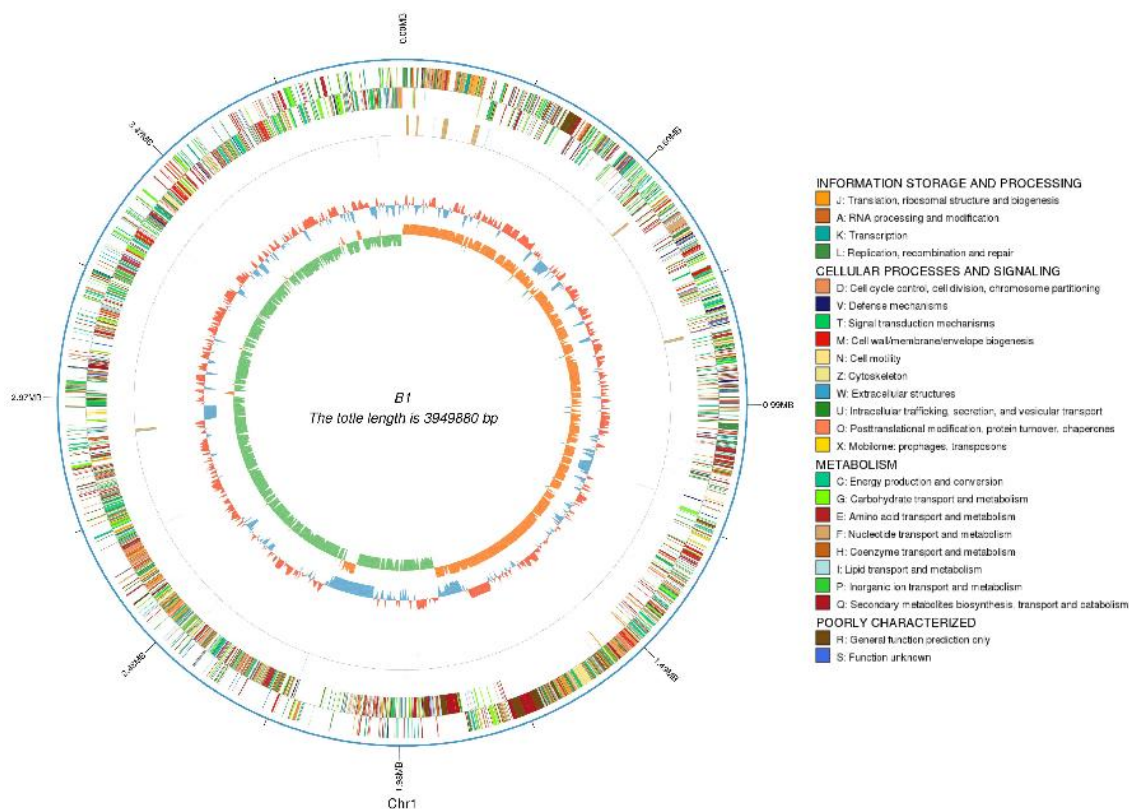
Functional annotation of the predicted proteins was performed through comparative analysis against several established databases, including the Non-Redundant (NR) database for species classification (Liu et al., 2017), Gene Ontology (GO) for cellular component, molecular function, and biological process assignments (Hsiao et al., 2003), the Kyoto Encyclopedia of Genes and Genomes (KEGG) for metabolic pathway analysis (Kanehisa et al., 2006), and the Clusters of Orthologous Groups (COG) database for phylogenetic classification (Galperin et al., 2015). The Carbohydrate-Active Enzyme (CAZy) database was utilized to identify enzyme families involved in carbohydrate metabolism. Repetitive sequences, including interspersed elements (SINEs and LINEs) and tandem repeats (minisatellite and microsatellite DNA), were identified and classified. Genomic islands (GIs) were predicted based on sequence composition and the presence of mobility-associated genes. Secondary metabolite gene clusters were predicted and analyzed to identify genes potentially involved in secondary metabolite biosynthesis. Finally, potential pathogenicity factors, virulence factors, and antibiotic resistance genes were identified through screening against the Pathogen Host Interactions (PHI) database (Urban et al., 2015), the Virulence Factors Database (VFDB) (Chen et al., 2012), and the Comprehensive Antibiotic Resistance Database (CARD), respectively.

## Result and analysis

### *Sequencing and assembly of Bacillus velezensis ANY11*

The complete genome sequence of *Bacillus velezensis* ANY11 was assembled into a circular chromosome measuring 3,949,880 bp in length, with a GC content of 45.85% (Fig. 1). Gene prediction revealed the presence of 4265 coding genes in *Bacillus velezensis* ANY11, with a total length of 3,530,166 bp, constituting 89.37% of the entire genome. The distribution of gene lengths is illustrated in Figure 2.

Repetitive sequences (DNA repeats) are identical or complementary fragments present in multiple copies within the genome and play a role in gene regulatory networks. These repeated sequences, based on their distribution in the linear genome, can be classified as interspersed or tandem repeats. Interspersed repeats include Short Interspersed Nuclear Elements (SINEs) and Long Interspersed Nuclear Elements (LINEs), the latter of which are often transpositionally active. In contrast, tandem repeats can be further divided into Minisatellite DNA and Microsatellite DNA. Repetitive sequences evolve at a faster rate, with certain sequences being species-specific. These species-specific repetitive elements can serve as genetic markers, aiding in the study of evolutionary relationships among different species.



**Figure 1.** Whole genome map of *Bacillus velezensis* ANY11. The outermost circle represents the position coordinates of the genome sequence. Moving from the outside to the inside, the following elements are included: gene function annotation results (including COG/KOG annotation information based on the actual project), ncRNA, and GC content of the genome. The GC content is calculated using a window size of (chromosome length/1000) bp and a step size of (chromosome length/1000) bp. The inward blue sections indicate regions where the GC content is lower than the average GC content of the entire genome, while the outward red sections indicate the opposite. The higher the peak, the greater the deviation from the average GC content. Additionally, the GC skew value is calculated using the same window and step sizes, with the specific algorithm being  $(G - C)/(G + C)$ . The inward green sections represent regions where the content of G is lower than that of C, while the outward orange sections indicate the opposite

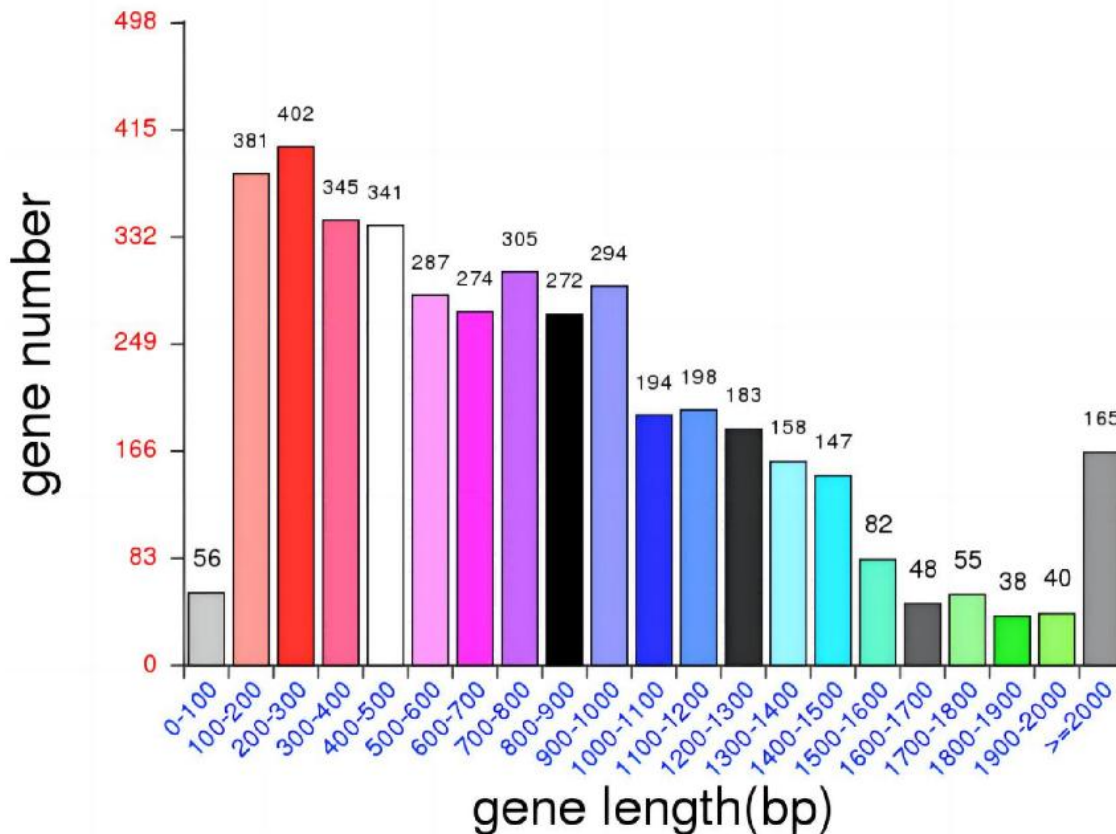
Through the prediction of repetitive sequences, 13 SINEs were identified, each measuring 1084 bp in length, along with 34 LINEs averaging 2196 bp, 148 Minisatellite DNAs totaling 9512 bp, and one Microsatellite DNA measuring 39 bp in the genome of *Bacillus velezensis* ANY11 (Table 1).

Genomic Islands (GIs) are segments integrated into the genomes of bacteria, phages, or plasmids through horizontal gene transfer, playing a critical role in bacterial evolutionary adaptation and pathogenicity (Liu et al., 2017). Through detection, seven genomic islands (GIs) are predicted in *Bacillus velezensis* ANY11 with a total length of 285,922 bp. The distribution of genes within these GIs is shown in Figure 3.

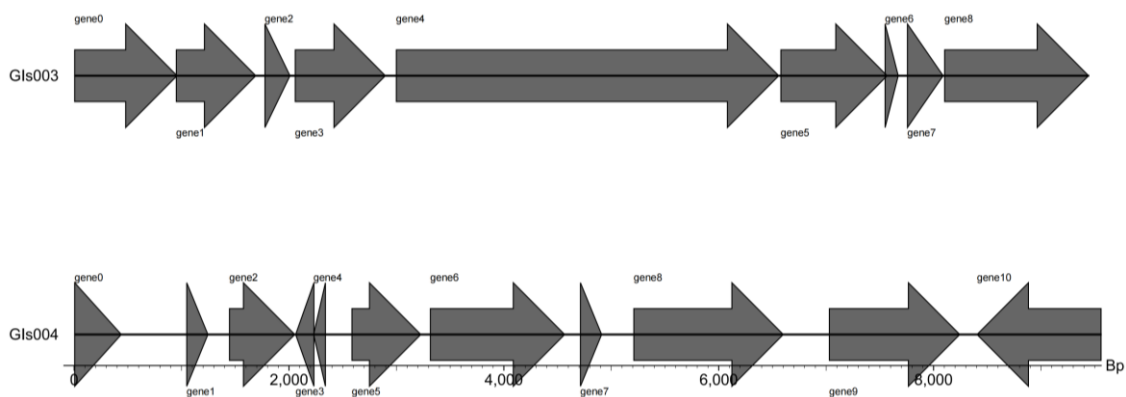
An analysis of the secondary metabolite gene clusters in *Bacillus velezensis* ANY11 was conducted and the prediction revealed a total of 10 gene clusters, with the number of genes contained in each cluster shown in Figure 4.

**Functional annotation of the *Bacillus velezensis* ANY11 genome**

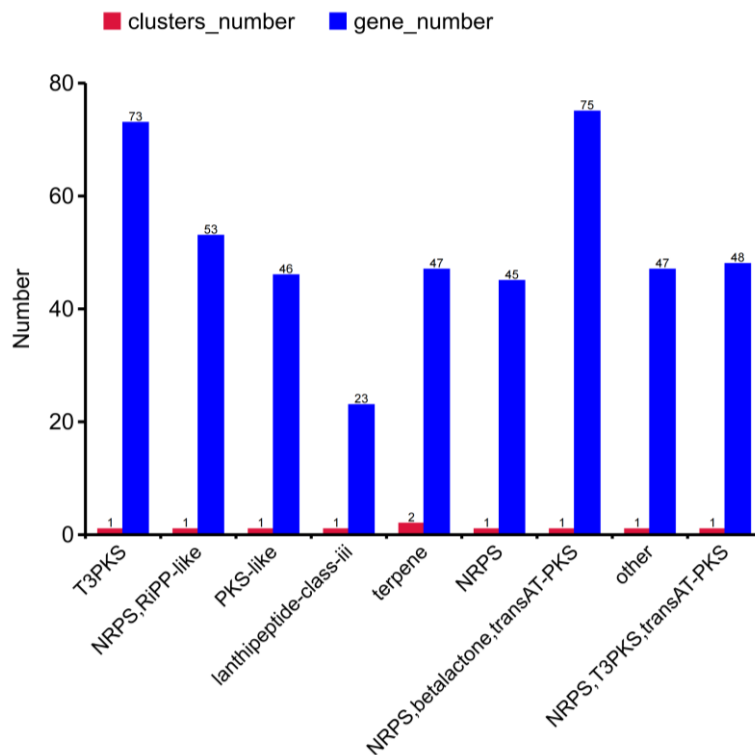
Figure 5 illustrates the species annotation of *Bacillus velezensis* ANY11 in the NR database, revealing that the predicted protein sequences are predominantly enriched in *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus velezensis*.



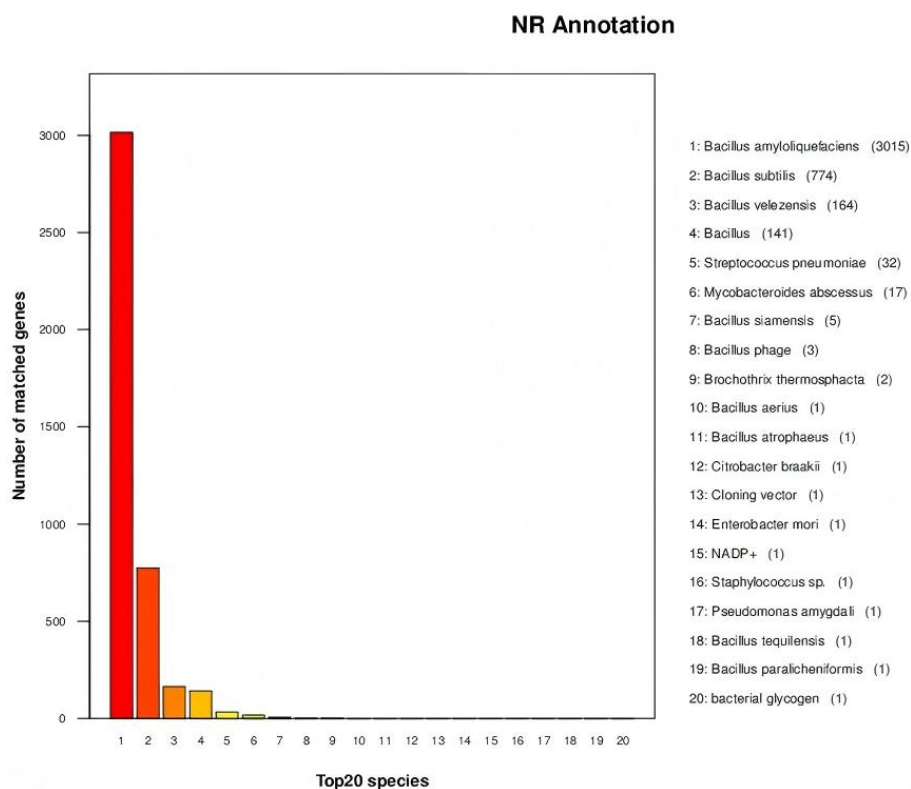
**Figure 2.** Distribution of gene lengths in the *Bacillus velezensis* ANY11 Genome. The horizontal axis represents the gene length, and the vertical axis represents the number of corresponding genes



**Figure 3.** Statistical chart of gene distribution in the genomic island. The horizontal axis represents the length scale (only genomic islands with lengths less than 15kb are displayed in the figure)



**Figure 4.** Statistical chart of gene clusters and their corresponding gene count



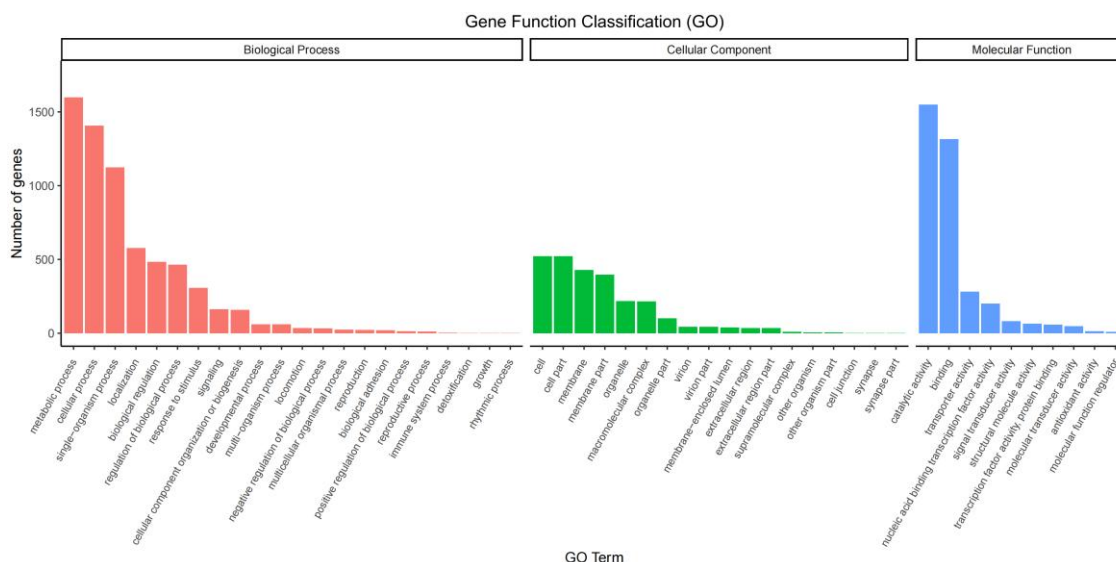
**Figure 5.** Species annotation of *Bacillus velezensis* ANY11 protein sequences in the non-redundant (NR) database. The horizontal axis represents species ID, and the vertical axis represents the number of genes annotated

**Table 1.** Statistics of the results of repetitive sequences of *Bacillus velezensis* ANY11

Type	Number	Total length (bp)	In genome (%)
LTR	110	9155	0.2318
DNA	31	1759	0.0445
LINE	34	2196	0.0556
SINE	13	1084	0.0274
RC	2	88	0.0022
Minisatellite DNA	148	9512	0.2408
Microsatellite DNA	1	39	0.001

Type: Type of interspersed repeat sequence; Number: Number of repeat sequences; Total Length (bp): Total length of repeat sequences; In genome (%): Percentage of repeat sequences in the genome; LTR: Long terminal repeat; DNA: DNA transposon; LINE: Long interspersed nuclear element; SINE: Short interspersed nuclear element; RC: Rolling circle; Minisatellite DNA: Minisatellite DNA; Microsatellite DNA: Microsatellite DNA

Figure 6 presents the statistical structure of the three major categories in the GO database for *Bacillus velezensis* ANY11 genes. The data indicates that 6570, 2622, and 3623 functional genes are annotated to biological process, cellular component, and molecular function, respectively. Among these, the gene-encoded products are primarily involved in two major biological processes: catalytic activity (1559 genes, 12.17%) and binding (1316 genes, 10.27%).



**Figure 6.** Gene function annotation GO functional classification. The horizontal axis represents the GO terms of the next level under the three major categories of GO, while the vertical axis represents the number of genes annotated to that term (including its sub-terms). The three different classifications represent the three basic categories of GO terms: biological process, cellular component, and molecular function (from left to right)

As illustrated in Figure 7, a total of 2473 genes annotated in the database were analyzed, revealing their enrichment in six major functional categories within the KEGG Pathway: Cellular Processes, Environmental Information Processing, Genetic

Information Processing, Human Diseases, Metabolism, and Organismal Systems. Among them, the Metabolism function had the highest number of annotations, with a total of 1699 genes. Within Metabolism, the Global and overview maps pathway (635 genes) was the most frequently annotated subcategory followed by the Carbohydrate metabolism pathway (230 genes).

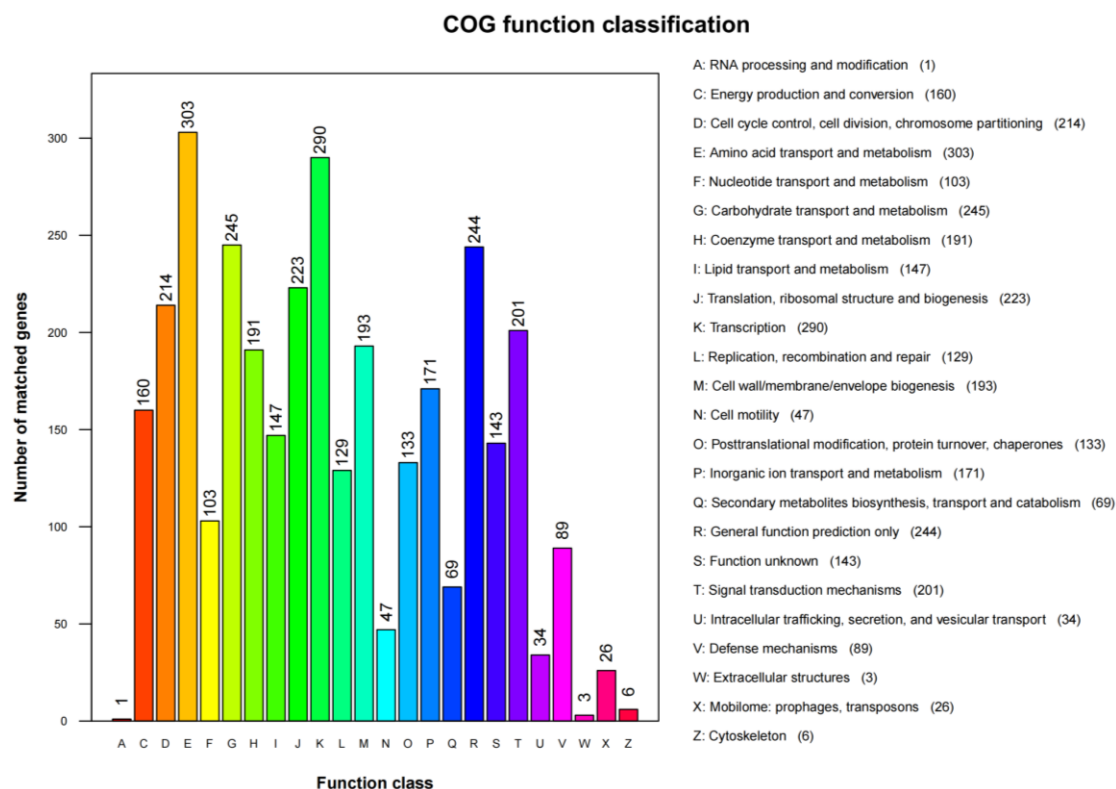


**Figure 7.** Gene function annotation KEGG metabolic pathway classification. The numbers on the bar chart represent the number of genes annotated. The other axis represents the level 1 functional category codes from the database, with the explanations for the codes provided in the corresponding legend

The COG functional annotation classification diagram for *Bacillus velezensis* ANY11 (Fig. 8) reveals 24 categories of annotations. Among these, the most common genes are related to amino acid transport and metabolism (303 genes, 9.00%), followed by transcription-related genes (290 genes, 8.62%). Additionally, genes involved in carbohydrate transport and metabolism (245 genes, 7.28%) rank third, closely followed by those dedicated to general function prediction only.

The CAZy database annotation results revealed enzyme families associated with carbohydrate degradation, modification, and biosynthesis within the complete genome of *Bacillus velezensis* ANY11 (Fig. 9). The five major categories within this database

include Glycoside Hydrolases (GHs), Glycosyl Transferases (GTs), Polysaccharide Lyases (PLs), Carbohydrate Esterases (CEs), and Auxiliary Activities (AAs). The largest number of genes encoded in the genome of *Bacillus velezensis* ANY11 pertains to Glycoside Hydrolases, with a total of 73 genes, accounting for 42.69%. Following this, genes encoding Carbohydrate Esterases (44 genes, 25.73%) and Glycosyl Transferases (33 genes, 21.05%) rank second and third, respectively.



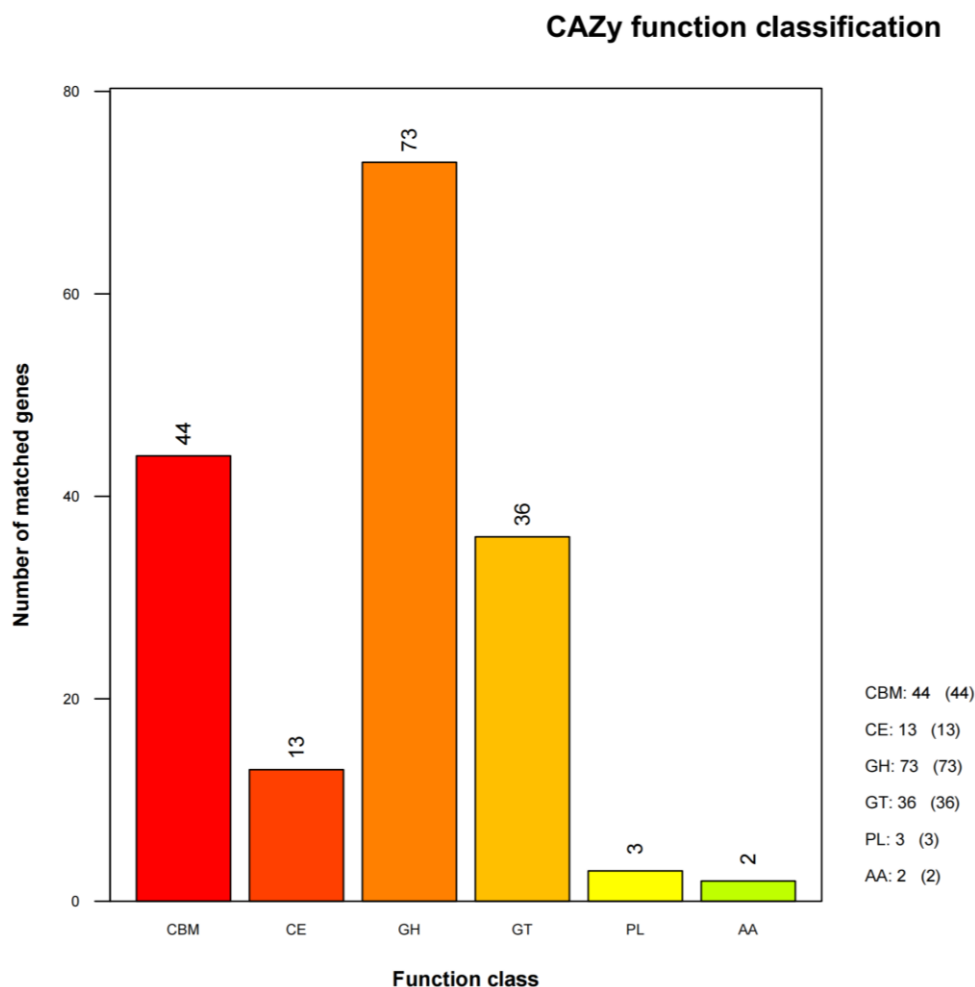
**Figure 8.** Gene function annotation COG functional classification chart. The horizontal axis represents the COG functional categories, and the vertical axis represents the number of genes annotated

Table 2 compares the number of genes encoding glycoside hydrolases (GH), glycosyltransferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), and auxiliary activities (AA) across several *Bacillus velezensis* strains with lignocellulose-degrading capabilities, including *B. velezensis* ANY11. The results indicate that *B. velezensis* ANY11 strain is enriched in GH family genes associated with cellulose and hemicellulose degradation than most other *Bacillus velezensis* strains. Moreover, it possesses a large number of genes essential for carbohydrate metabolism, highlighting its significant potential for degrading carbohydrate polysaccharides.

### **Analysis of the virulence or pathogenicity of *Bacillus velezensis* ANY11**

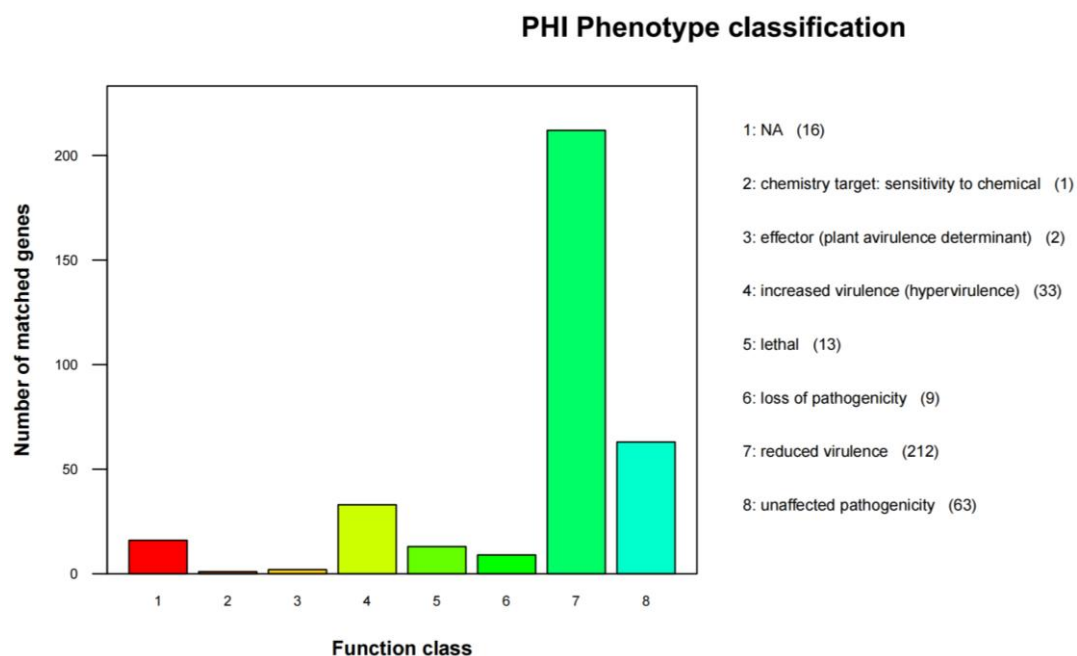
The PHI database, which specializes in pathogen-host interactions, plays a crucial role in identifying target genes for drug intervention studies. It also encompasses antifungal compounds and their corresponding target genes. Each gene entry in the database comprises nucleic acid and amino acid sequences, along with detailed

descriptions of the predicted protein functions during host infection (Urban et al., 2015). As shown in *Figure 10*, out of the 349 genes annotated in the PHI database, 212 genes were annotated as reduced pathogenicity; 63 genes were annotated as having unaffected pathogenicity; 33 genes were annotated as virulence enhanced; 13 genes were annotated as lethal factors; 9 genes were annotated as pathogenic loss; 2 genes were annotated as effectors related to plant avirulence determinant; 1 gene was annotated as a chemically sensitive target; and 13 genes were annotated as unknown.



**Figure 9.** Statistical chart of CAZy functional classification and corresponding gene counts. The horizontal axis represents the classification types from the CAZy database, while the vertical axis represents the number of genes annotated

The VFDB provides species information, basic characteristics of virulence genes, and detailed descriptions of their functions and pathogenic mechanisms (Chen et al., 2012). A comparison of the genomic data of *Bacillus velezensis* ANY11 with the VFDB database identified four virulence genes with known functions (Identity threshold of  $\geq 70\%$ ) (*Table 3*). Among these, the virulence genes ClpC (Xu et al., 2018) and BslA (Kern et al., 2010) are associated with bacterial adhesion, ClpP is related to bacterial growth and metabolism (Roy et al., 2019), and Capsule is involved in bacterial protection (Taylor et al., 2005).



**Figure 10.** Distribution chart of phenotypic mutation types of pathogen PHI. The horizontal axis represents the phenotypic mutation types, while the vertical axis represents the number of genes annotated

**Table 2.** Comparison of CAZy database classification annotation results for other *Bacillus velezensis*

Strains	CBM	CE	GH	GT	PL	AA
<i>B. velezensis</i> ANY11	44	13	73	36	3	2
<i>B. velezensis</i> SSF6 (Zhang et al., 2023)	39	14	67	34	3	1
<i>B. velezensis</i> LC1 (Li et al., 2020)	15	30	43	38	3	6
<i>B. velezensis</i> 157 (Chen et al., 2018)	23	18	43	33	3	5
<i>B. velezensis</i> SQR9 (Zhang et al., 2015)	18	20	42	33	3	5
<i>B. velezensis</i> UCM, B5113 (Niazi et al., 2014)	22	18	41	32	3	5
<i>B. velezensis</i> SYBC_H47 (Li et al., 2016)	18	21	37	32	3	5
<i>B. velezensis</i> FZB42 (Li et al., 2020)	21	10	37	36	3	1

The CARD annotation results summarize the names of resistance-related genes and the types of antibiotics to which they confer resistance. By aligning the amino acid sequences of *Bacillus velezensis* ANY11 with the CARD database and combining the target species' genes with their corresponding resistance function annotations, the results were obtained and are summarized in Table 4. According to Table 4, the genome of *Bacillus velezensis* ANY11 contains eight antibiotic resistance genes. Specifically, the

clbA, BcI, and qacJ genes exhibit multi-drug resistance, suggesting that *Bacillus velezensis* ANY11 possesses a degree of antibiotic resistance.

**Table 3.** The VFDB annotation results for *Bacillus velezensis* ANY11 reveal four virulence genes with known functions

Gene ID	VF name	Related genes	Identity (%)	Functions
GM003742	ClpC	clpC endopeptidase Clp ATP-binding chain C	78.6	An ATPase promoting early escape from the phagosome of macrophages; ClpC is also required for adhesion and invasion, possibly by modulating the expression of InlA, InlB and ActA
GM003613	ClpP	clpP ATP-dependent Clp protease proteolytic subunit	77.9	Serine protease involved in proteolysis and is required for growth under stress conditions
GM003119	BslA	bslA/yuaB hydrophobin BslA	73.9	Encodes a major S-layer protein that is located within the pXO1 pathogenicity island that also codes for toxin genes. BslA mediates adherence and entry to epithelial cells by binding integrin $\alpha 2 \beta 1$ and complement component C1q; forms a highly hydrophobic coat around <i>B. subtilis</i> biofilms
GM002521	Capsule	gnd 6-phosphogluconate dehydrogenase	71.1	Assisting in evading the host immune system by protecting bacteria from opsonophagocytosis and serum killing

**Table 4.** The CARD annotation results for *Bacillus velezensis* ANY11

Gene ID	ARO name	Drug class	Resistance mechanism
GM000467	vanT gene in vanG cluster	Glycopeptide antibiotic	Antibiotic target alteration
GM000514	clbA	Lincosamide antibiotic; streptogramin antibiotic; streptogramin A antibiotic; oxazolidinone antibiotic	Antibiotic target alteration
GM000547	<i>Nocardia farcinica</i> rox	Rifamycin antibiotic	Antibiotic inactivation
GM001271	BcI	Cephalosporin; penem	Antibiotic inactivation
GM001819	qacJ	Disinfecting agents and antiseptics	Antibiotic efflux
GM002063	vanY gene in vanB cluster	Glycopeptide antibiotic	Antibiotic target alteration
GM002563	vanT gene in vanG cluster	Glycopeptide antibiotic	Antibiotic target alteration
GM002728	tet (45)	Tetracycline antibiotic	Antibiotic efflux

## Discussion

As an effective detection technique, whole-genome sequencing has been extensively utilized to predict and investigate microbial functions. To acquire the genomic sequence information of *B. velezensis* ANY11, isolated from bamboo fiber

humus, whole-genome sequencing was performed. Following the initial construction, correction, and optimization, the genome was assembled into a circular chromosome measuring 3,949,880 bp in length, with predicted coding genes spanning 3,530,166 bp, accounting for 89.37% of the entire genome. Furthermore, the analysis of repetitive sequences revealed 190 scattered repeats and 176 tandem repeats. The non-coding RNAs identified include 86 tRNAs, 9 16sRNAs, 9 23sRNAs, 9 5sRNAs, and 9 sRNAs.

Genomic islands (GIs) can harbor genes involved in pathogenicity, metabolism, antibiotic resistance, and symbiosis. In this context, the prediction of genomic islands in *B. velezensis* ANY11 was conducted. Additionally, prophages were predicted that could contribute to the bacterium's acquisition of antibiotic resistance, enhanced environmental adaptability, increased adhesiveness, or pathogenicity. The results indicated the presence of seven GIs, eleven prophages, and ten gene clusters in the whole genome of *Bacillus velezensis* ANY11. Annotation results from the PHI and VFDB databases revealed that strain B1 contains a total of 349 virulence factor genes, suggesting that *B. velezensis* ANY11 may release virulence factors to inhibit the growth of pathogenic bacteria. Previous studies have shown that *Bacillus velezensis* ZF2 exhibits significant antagonistic activity against 14 plant pathogenic bacteria. Furthermore, genomic sequencing predicted 14 gene clusters related to the synthesis of secondary metabolites, potentially encoding various known antimicrobial compounds (Xu et al., 2020). Research on *Bacillus velezensis* Bvel1 has shown that it secretes metabolites that inhibit *Botrytis cinerea* and identified 13 antibacterial biosynthetic gene clusters. Further analysis confirmed that these gene clusters encode several known antibacterial compounds, including iturin A, fengycin, surfactin, bacilysin, difficidin, bacillaene, and bacillibactin (Nifakos et al., 2021). Additionally, analysis of the genome of *Bacillus velezensis* HNA3 revealed 77 genes encoding various antibacterial enzymes, such as chitinase, endoglucanase, lysozyme, and peptidoglycan (Zaid et al., 2022). *Bacillus velezensis* is recognized as a potential biocontrol agent due to its significant inhibitory effects against various pathogenic fungi. The *B. velezensis* ANY11 strain, isolated from the intestine of *Cyrtotrachelus buqueti*, has demonstrated considerable antibacterial capability based on genomic analysis. However, further gene mining is necessary to identify specific genes encoding antibacterial enzymes and gene clusters related to disease resistance. A thorough exploration of the antibacterial genes of *B. velezensis* ANY11 will provide crucial insights into its antibacterial mechanisms and potential applications, serving as a molecular basis for its development and utilization in agricultural production and biopesticide development.

Significant genetic capabilities of *B. velezensis* ANY11 in carbohydrate degradation were identified through gene annotation using functional databases, including NR, GO, KEGG, COG, and CAZy. Annotations from the NR database revealed a strong correlation between the genes of *B. velezensis* ANY11 and those of *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus velezensis*. Previous studies have established that both *Bacillus amyloliquefaciens* (Mei et al., 2020) and *Bacillus subtilis* are effective in decomposing lignocellulose. Extensive research on various cellulases produced by *Bacillus subtilis* has demonstrated their impact on the digestibility and structural modification of lignocellulosic biomass (Malik et al., 2021). Additionally, the *Bacillus velezensis* SSF6 strain, isolated from humus soil, has been shown to possess lignocellulose-decomposing abilities. Physiological and biochemical analyses performed using the Biolog Gen III MicroStation automatic microbial identification

system revealed 24 positive reactions, indicating its capability to utilize carbon sources such as cellobiose, sucrose, and fructose (Zhang et al., 2023). Similarly, *Bacillus velezensis* LC1, also isolated from bamboo fiber humus in the intestine of *Cyrtotrachelus buqueti*, demonstrated lignocellulose degradation potential. When cultured on media and stained with Congo red, it produced clear zones around the colonies, indicating cellulolytic activity (Li et al., 2020). Through KEGG database analysis, 230 genes associated with carbohydrate metabolism pathways were identified. The CAZy database helped isolate five main types of carbohydrate enzymes present in the genome of *B. velezensis* ANY11, including various genes from glycoside hydrolase (GH) families, which are crucial for the degradation of cellulose and hemicellulose (Zhang et al., 2023).

When compared to other strains known for their lignocellulose degradation and cellulase activity, it is speculated that *B. velezensis* ANY11 also possesses the ability to degrade lignocellulose. Moreover, this strain exhibits a greater number of GH family genes associated with cellulose and hemicellulose degradation than most other *Bacillus velezensis* strains, suggesting it has a superior capacity for lignocellulose degradation (Table 2). Therefore, *B. velezensis* ANY11, isolated from bamboo fiber humus in the intestines of *Cyrtotrachelus buqueti*, demonstrates the potential to decompose lignocellulose and facilitate its conversion to ethanol, making it a promising candidate for the production of bioethanol through biomass decomposition. However, further research is needed to identify the specific cellulase genes within the genome of this strain. Such exploration will be crucial for enhancing lignocellulose degradation rates and improving ethanol production efficiency.

## Conclusion

In this study, the genome of *B. velezensis* ANY11, isolated from bamboo fiber humus in the intestines of *Cyrtotrachelus buqueti*, was analyzed, leading to the acquisition of the complete genome sequence of the strain. The genome was annotated using functional databases to infer the functions associated with this strain. Initial tests were also conducted to assess the virulence and pathogenicity of *B. velezensis* ANY11, with virulence factors being annotated through the Pathogen-Host Interactions (PHI) database and the VFDB. Additionally, drug resistance genes were annotated using relevant databases. The analysis of these data reveals that *B. velezensis* ANY11 possesses genes associated with antifungal activity, suggesting that it may exhibit inhibitory effects on pathogenic fungi, positioning it as a potential biocontrol agent for enhancing agricultural production. Furthermore, the strain demonstrates the potential to degrade lignocellulose. This finding emerged from a comparative analysis of annotations from the CAZy database. Further research into the specific genes responsible for lignocellulose-degrading enzymes within this strain could provide important bioinformatics insights into the genetic basis of lignocellulose degradation and the production of clean fuel ethanol from lignocellulose. The genomic sequencing and analysis conducted in this study contribute new microbial resources for lignocellulose degradation, promising to open up broad prospects for industrial production application.

**Acknowledgements.** This work was financially supported by the Scientific Research and Cultivation Plan of Leshan Normal University: No.2022SSDJS016 and No.2022SSDJS011.

## REFERENCES

- [1] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. (2000): Gene ontology: tool for the unification of biology. – *Nat Genet* 25(1): 25-29.
- [2] Chen, L., Xiong, Z., Sun, L., Yang, J., Jin, Q. (2012): VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. – *Nucleic Acids Res* 40(Database issue): D641-D645.
- [3] Chen, L., Gu, W., Xu, H.-Y., Yang, G.-L., Shan, X.-F., Chen, G., Kang, Y.-H., Wang, C.-F., Qian, A. (2018): Comparative genome analysis of *Bacillus velezensis* reveals a potential for degrading lignocellulosic biomass. – *3 Biotech* 8: 1-5.
- [4] Galperin, M. Y., Makarova, K. S., Wolf, Y. I., Koonin, E. V. (2015): Expanded microbial genome coverage and improved protein family annotation in the COG database. – *Nucleic Acids Res* 43(Database issue): D261-D269.
- [5] Hsiao, W., Wan, I., Jones, S. J., Brinkman, F. S. L. (2003): IslandPath: aiding detection of genomic islands in prokaryotes. – *Bioinformatics* 19(3): 418-420.
- [6] Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., Hirakawa, M. (2006): From genomics to chemical genomics: new developments in KEGG. – *Nucleic Acids Res* 34(Database issue): D354-D357.
- [7] Kern, J., Schneewind, O. J. M. (2010): BslA, the S-layer adhesin of *B. anthracis*, is a virulence factor for anthrax pathogenesis. – *Mol Microbiol* 75(2): 324-332.
- [8] Laloo, R., Maharajh, D., Görgens, J., Gardiner, N. (2010): A downstream process for production of a viable and stable *Bacillus cereus* aquaculture biological agent. – *Appl Microbiol Biotechnol* 86(2): 499-508.
- [9] Li, X., Zhang, Y., Wei, Z., Guan, Z., Cai, Y., Liao, X. (2016): Antifungal activity of isolated *Bacillus amyloliquefaciens* SYBC H47 for the biocontrol of peach gummosis. – *PloS ONE* 11(9): e0162125.
- [10] Li, Y., Lei, L., Zheng, L., Xiao, X., Tang, H., Luo, C. (2020): Genome sequencing of gut symbiotic *Bacillus velezensis* LC1 for bioethanol production from bamboo shoots. – *Biotechnol Biofuels* 13: 34.
- [11] Liu, G., Kong, Y., Fan, Y., Geng, C., Peng, D., Sun, M. (2017): Data on genome analysis of *Bacillus velezensis* LS69. – *Data Brief* 13: 1-5.
- [12] Malik, W. A., Javed, S. (2021): Biochemical characterization of cellulase from *Bacillus subtilis* strain and its effect on digestibility and structural modifications of lignocellulose rich biomass. – *Front Bioeng Biotechnol* 9: 800265.
- [13] Mansour, A. A., Da Costa, A., Arnaud, T., Lu-Chau, T. A., Fdz-Polanco, M., Moreira, M. T., Cacho Rivero, J. A. (2016): Review of lignocellulolytic enzyme activity analyses and scale-down to microplate-based assays. – *Talanta* 150: 629-637.
- [14] Mei, J., Shen, X., Gang, L., Xu, H., Wu, F., Sheng, L. (2020): A novel lignin degradation bacteria-*Bacillus amyloliquefaciens* SL-7 used to degrade straw lignin efficiently. – *Bioresour Technol* 310: 123445.
- [15] Niazi, A., Manzoor, S., Asari, S., Bejai, S., Meijer, J., Bongcam-Rudloff, E. (2014): Genome analysis of *Bacillus amyloliquefaciens* subsp. *plantarum* UCMB5113: a rhizobacterium that improves plant growth and stress management. – *PLoS ONE* 9(8): e104651.
- [16] Nifakos, K., Tsalgatidou, P. C., Thomloudi, E.-E., Skagia, A., Kotopoulis, D., Baira, E., Delis, C., Papadimitriou, K., Markellou, E., Venieraki, A., Katinakis, P. (2021): Genomic analysis and secondary metabolites production of the endophytic *Bacillus velezensis* Bvel1: a biocontrol agent against *Botrytis cinerea* causing bunch rot in post-harvest table grapes. – *Plants (Basel)* 10(8).
- [17] Park, G., Nam, J., Kim, J., Song, J., Kim, P. I., Min, H. J., Lee, C. W. J., Bot, K. C. S. (2019): Structure and mechanism of surfactin peptide from *Bacillus velezensis* antagonistic to fungi plant pathogens. – *B Korean Chem Soc* 40(7): 704-709.

- [18] Peixoto, S. B., Cladera-Olivera, F., Daroit, D. J., Brandelli, A. J. A. R. (2011): Cellulase-producing *Bacillus* strains isolated from the intestine of Amazon basin fish. – *Aquac Res* 42(6): 887-891.
- [19] Roy, S., Zhu, Y., Ma, J., Roy, A. C., Zhang, Y., Zhong, X., Pan, Z., Yao, H. (2019): Role of ClpX and ClpP in *Streptococcus suis* serotype 2 stress tolerance and virulence. – *Microbiol Res* 223-225.
- [20] Taylor, C. M., Roberts, I. S. (2005): Capsular polysaccharides and their role in virulence. – *Nephron* 12: 55-66.
- [21] Urban, M., Pant, R., Raghunath, A., Irvine, A. G., Pedro, H., Hammond-Kosack, K. E. (2015): The Pathogen-Host Interactions database (PHI-base): additions and future developments. – *Nucleic Acids Res* 43(Database issue): D645-D655.
- [22] Wang, J., Xing, J., Lu, J., Sun, Y., Zhao, J., Miao, S., Xiong, Q., Zhang, Y., Zhang, G. (2019): Complete genome sequencing of *Bacillus velezensis* WRN014, and comparison with genome sequences of other *Bacillus velezensis* strains. – *J Microbiol Biotechnol* 29(5): 794-808.
- [23] Wick, R. R., Judd, L. M., Gorrie, C. L., Holt, K. E. (2017): Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. – *PLoS Comput Biol* 13(6): e1005595.
- [24] Xu, S., Xie, X., Zhao, Y., Shi, Y., Chai, A., Li, L., Li, B. (2020): Whole-genome analysis of *Bacillus velezensis* ZF2, a biocontrol agent that protects *Cucumis sativus* against *Corynespora* leaf spot diseases. – *3 Biotech* 10(4): 186.
- [25] Xu, X., Huang, L., Su, Y., Yan, Q. (2018): The complete genome sequence of *Vibrio aestuarianus* W-40 reveals virulence factor genes. – *Microbiologyopen* 7(3): e00568.
- [26] Ye, M., Tang, X., Yang, R., Zhang, H., Li, F., Tao, F., Li, F., Wang, Z. (2018): Characteristics and application of a novel species of *Bacillus*: *Bacillus velezensis*. – *ACS Chem Biol* 13(3): 500-505.
- [27] Zaid, D. S., Cai, S., Hu, C., Li, Z., Li, Y. (2022): Comparative genome analysis reveals phylogenetic identity of *Bacillus velezensis* HNA3 and genomic insights into its plant growth promotion and biocontrol effects. – *Microbiol Spectr* 10(1): e0216921.
- [28] Zhang, H., Jackson, T. A. (2008): Autochthonous bacterial flora indicated by PCR-DGGE of 16S rRNA gene fragments from the alimentary tract of *Costelytra zealandica* (Coleoptera: Scarabaeidae). – *J Appl Microbiol* 105(5): 1277-1285.
- [29] Zhang, N., Yang, D., Wang, D., Miao, Y., Shao, J., Zhou, X., Xu, Z., Li, Q., Feng, H., Li, S., Xhen, Q., Zhang, R. (2015): Whole transcriptomic analysis of the plant-beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm formation regulated by maize root exudates. – *BMC genomics* 16: 1–20.
- [30] Zhang, T., Wei, S., Liu, Y., Cheng, C., Ma, J., Yue, L., Gao, Y., Cheng, Y., Ren, Y., Su, S., Zhao, X., Lu, Z. (2023): Screening and genome-wide analysis of lignocellulose-degrading bacteria from humic soil. – *Front Microbiol* 14: 1167293.