

## RESEARCH ARTICLE

# Isolation, identification, and whole-genome sequencing of high-yield protease bacteria from Daqu of ZhangGong Laojiu

Yanbo Liu<sup>1,2,3,4</sup>, Junying Fu<sup>1,3,4</sup>, Linlin Wang<sup>1,3,4</sup>, Zhijun Zhao<sup>1,3,4</sup>, Huihui Wang<sup>1,3,4</sup>, Suna Han<sup>2</sup>, Xiyu Sun<sup>1,3,5</sup>, Chunmei Pan<sup>1,3,4\*</sup>

**1** College of Food and Biological Engineering (Liquor College), Henan University of Animal Husbandry and Economy, Zhengzhou, Henan Province, China, **2** Postdoctoral Programme, Henan Yangshao Distillery Co., Ltd., Mianchi, Henan Province, China, **3** Henan Liquor Style Engineering Technology Research Center, Henan University of Animal Husbandry and Economy, Zhengzhou, Henan Province, China, **4** Zhengzhou Key Laboratory of Liquor Brewing Microbial Technology, Henan University of Animal Husbandry and Economy, Zhengzhou, Henan Province, China, **5** ZhangGongLaoJiu Wine Co. Ltd., Ningling, Henan Province, China

\* [sige518888@163.com](mailto:sige518888@163.com)



## OPEN ACCESS

**Citation:** Liu Y, Fu J, Wang L, Zhao Z, Wang H, Han S, et al. (2022) Isolation, identification, and whole-genome sequencing of high-yield protease bacteria from Daqu of ZhangGong Laojiu. PLoS ONE 17(4): e0264677. <https://doi.org/10.1371/journal.pone.0264677>

**Editor:** Tofazzal Islam, Bangabandhu Sheikh Mujibur Rahman Agricultural University, BANGLADESH

**Received:** October 15, 2021

**Accepted:** February 15, 2022

**Published:** April 26, 2022

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0264677>

**Copyright:** © 2022 Liu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

## Abstract

A total of 296 strains of protease-producing bacteria were isolated and purified from medium-temperature Daqu produced by ZhangGong LaoJiu Wine Co. Ltd. After calculating the ratio of transparent ring diameter to colony diameter and measuring the protease activities, a strain of high-yield protease bacteria, called DW-7, was screened out with a protease activity of 99.54 U/mL. Through morphological observation, 16S rDNA sequence analysis, and physiological and biochemical tests, the isolated bacteria DW-7 was determined to be *Bacillus velezensis*. In addition, whole-genome sequencing (WGS), using PacBio and the Illumina platform, was performed. Gene annotation was then conducted using the Clusters of Orthologous Groups (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG), Non-Redundant Protein Sequence Database (NR), and Gene Ontology (GO) databases. The results showed that the genome of DW-7 was 3,942,829 bp long with a GC content of 46.45%. A total of 3,662 protein-encoding genes were predicted, with a total length of 3,402,822 bp. Additionally, 2,283; 2,796; and 2,127 genes were annotated in the COG, KEGG, and GO databases, respectively. A total of 196 high-yield protease genes were mainly enriched in the metabolism of alanine, aspartic acid, glutamate, glycine, serine, and threonine, as well as ABC transporter and transporter pathways.

## Introduction

Chinese Luzhou-flavored liquor is typically created via an oriental solid-state fermentation process. Liquor fermentation is a complex process that takes place in an open environment and is coordinated by various microorganisms and enzymes. The microorganisms in liquor mainly come from the Daqu, the surrounding environment, pit mud, and yellow water [1]. Daqu is a unique saccharification starter for solid liquor brewing in China. It not only contains

**Funding:** This work supported by the Major Science and Technology Projects of Henan Province of China(181100211400), Key Technologies Research and Development Program of Henan Province of China(202102110130), Scientific Research Foundation for Docotors of Henan University of Animal Husbandry and Economy(2018HNUAHEDF011) and Key Subject Projects of Henan University of Animal Husbandry and Economy(C3060020), all awarded to YL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

important crude enzymes and complex microflora, but it is also a source of important biologically active substances, such as aroma precursors that determine flavor characteristics [2, 3]. The quality of Daqu directly affects the quality, yield, and style of the liquor product.

The microorganisms in Daqu mainly include bacteria, yeasts, molds, and a small number of actinomycetes. Bacteria are the vital microorganism in Daqu, the amount and quantity of which directly determines its quality. *Bacillus sp.* is the dominant flora in Daqu, which produces various enzymes that are closely linked to the formation of flavor substances during the fermentation process [4]. It has been reported that many flavor substances in liquor are derived from proteins, and the functions of those proteins and their enzymes are significant for developing the flavor of liquor. For example, protease is an enzyme that hydrolyzes peptide bonds into amino acids or short peptides [5] and functions by maintaining enzyme activities under acidic, neutral, and alkaline conditions. Additionally, it is a metabolite of various microorganisms involved in the production of Daqu [6]. In the process of solid-state liquor fermentation, proteases in Daqu can promote the dissolution of raw materials, decompose proteins, produce amino acids, provide nitrogen sources, and flavor components, promote the growth of microorganisms, degrade yeast cells, improve the utilization of raw materials, cooperate in co-fermentation with other microbes, inhibit the production of fusel oil, and enhance the yield, quality, and aroma of the original liquor [7, 8].

Whole-genome sequencing (WGS) is a fast, low-cost, and highly efficient third-generation sequencing technology that can provide a complete bacterial genome sequence. The differences between species of the same genus can be better identified using WGS and by following gene annotations using online databases, such as GO, KEGG, COG, and NR. WGS has become a popular detection method and is widely used to identify microbial communities within intestinal flora, soil, and fungi [9, 10]. It is necessary for mining the core genome, identifying specific genes, and analyzing functional genomics, which ultimately contributes to the exploration of the diversity and biological characteristics of unknown microbial groups [11].

At present, domestic and foreign research mainly focuses on the physical and chemical properties and the purification technologies of proteases, as well as the methods to obtain high-yield protease strains and protease products through strain mutagenesis, genetic engineering, and other methods. For example, Lu et al. [12] screened out a high-yield, neutral protease bacterium from high-temperature yeast for making hard liquor, which was identified as *B. amyloliquefaciens*. Nie et al. [13] initially screened for microbes by calculating the transparent ring diameter, then re-screened by measuring protease activity, and finally identified two strains of high-yield proteases from Daqu in Jiang-flavor Chinese spirits, which were characterized as *B. subtilis*. However, the exact mechanism underlying the screening of high-yield proteases from Daqu of Chinese Luzhou-flavored liquor is not well understood. In the present study, we screened a strain of high-yield protease bacteria from the Daqu of ZhangGong Laojiu. We obtained the whole-genome sequence by performing WGS using the PacBio and Illumina platforms. Following gene annotations using online databases of GO, KEGG, COG, NR, and Swiss-Prot databases, we further revealed the target enriched genes and metabolic pathways. Our findings provide theoretical references for the mining of high-yield proteases and their biological functions.

## Materials and methods

### Materials and reagents

A medium-temperature starter for our study was provided by Henan ZhangGong Laojiu Wine Co., Ltd. Sodium carbonate anhydrous was obtained from Tianjin Dingshengxin Chemical Co., Ltd; the reagents L-tyrosine, agar powder, casein, trichloroacetic acid, and sodium

chloride were purchased from Tianjin Kemiou Chemical Reagent Co., Ltd; beef extract peptone was obtained from Beijing Aoboxing Biotechnology Co., Ltd; the Folin-Ciocalteu reagent was from Solarbio, Biotechnology Co., Ltd; the Ezup Column Bacteria Genomic DNA Purification Kit was from Sangon Biotech Co., Ltd.; DNA 1000 assay kit was from Lithuania Agilent Technologies; and the ABI StepOnePlus Real-Time PCR System was purchased from Life Technologies.

## Instruments

A constant-temperature incubator was purchased from Shanghai Shcimo Medical Device Manufacturing Co., Ltd; a high-speed refrigerated centrifuge was provided by Shanghai Anting Device Manufacturing Co., Ltd; a Dk-8D digital thermostatic laboratory water bath was from Guangzhou Hezhong Biotechnology Co., Ltd; a three-dimensional autoclave was obtained from Shanghai Shenan Medical Device Manufacturing Co., Ltd; 752 UV-VIS spectrophotometer was from Shanghai Jinghua Instruments; the biological purification table was from Suzhou Purification Equipment Co., Ltd; PacBio Single Molecule, real-time (SMRT) sequencing system was acquired from Pacific Biosciences of California; the Novaseq 6000 system was from Illumina; a 2100 Bioanalyzer Instrument was from Agilent Technologies; the Eppendorf 5427R centrifuge was from Eppendorf; the NanoDrop 2000 spectrophotometer was obtained from Thermal Fisher Scientific; the DYY-6C agarose gel electrophoresis kit was from Beijing Liuyi Biotechnology Co., Ltd; and finally, the PCR instrument was from Dongsheng Xingye Scientific Instrument Co., Ltd.

## Culture media

The Casein medium [14] contained: 3 g beef extract, 5 g sodium chloride, 10 g peptone, 20 g agar powder, and 4 g casein at the constant volume of 1000 mL of distilled water with a pH level between 7.0 and 7.2. The medium was sterilized at 121°C for 20 minutes.

The beef extract peptone agar medium contained: 3 g beef extract, 5 g sodium chloride, 10 g peptone, 20 g agar powder, and 4 g casein in a constant volume of 1000 mL of distilled water at pH 7.0–7.2. The medium was sterilized at 121°C for 20 min.

The fluid medium [15] contained: 3 g beef extract, 5 g sodium chloride, and 10 g peptone in a constant volume of 1000 mL of distilled water at pH 7.0–7.2, and sterilized at 121°C for 20 min.

Solid medium [16]: 20 g bran and 20 mL distilled water, and sterilized at 121°C for 20 min.

## Experimental methods

### Isolation of protease-producing bacteria.

- I. *Isolation*: Daqu was ground into powder, 10 g of which was added to an Erlenmeyer flask containing 90 mL of sterile distilled water. After gently shaking the solution in a shaking incubator at 150 rpm for 30 min [17], 1 mL of the suspension was transferred to a tube containing 9 mL of sterile water and diluted 10 times. 0.1 mL of  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$ , and  $10^{-6}$ -times diluted bacteria solution was plated and cultured at 37°C for 24 h in inverted position [18].
- II. *Initial screening*: Under aseptic conditions, a sterile pipette tip was used to place a small amount of purified bacteria on the casein medium plate, created three parallels, and cultured for 24 h in inverted position to observe the formation of a transparent ring. The ratio (D/d) of transparent ring diameter (D) to cell colony diameter (d) was calculated, and those with a larger D/d ratio were selected for re-screening [19].

## Screening of protease-producing bacteria

- I. *Preparation of the standard curve*: The standard curve was plotted using the absorbance value and the concentration of L-tyrosine as the ordinate and abscissa, respectively.
- II. *Preparation of crude enzyme solution*: To prepare the crude enzyme solution, bacteria were inoculated into solid medium of beef extract peptone, and cultured at 37°C for 24 h in inverted position. Subsequently, a single colony with obvious morphology was picked for inoculation into the liquid medium and then cultured on a shaker at 37°C for 24 hours. The bacterial suspension was inoculated into the solid culture medium at 10% of the inoculation volume, and incubated at 37°C for 3 days. Solid fermentation product (10 g) was added to a 250 mL Erlenmeyer flask containing 90 mL of distilled water and incubated in a 40°C water bath for 1 hour. Then, the mixture was stirred every 15 minutes. After centrifugation, the crude enzyme solution was obtained [20].
- III. *Measurement of protease activity using the Folin-Ciocalteu method*: Under the conditions of 1 mL of protease solution, at pH 7.5 and 40°C, the amount of protease required to hydrolyze casein to produce 1 µg L-tyrosine per minute was defined as the protease activity, calculated using the 0

$$\text{Protease activity (U/mL)} = \frac{A \times 4 \times N \times OD}{10};$$

The constant A was obtained from the standard curve and it is equal to the number of micrograms of L-tyrosine equivalent to  $OD_{680} = 1$ ; 4 represents 1 mL of solution collected from the total volume of 4 mL for measurement; N is the dilution multiple of protease solution to be measured; 10 represents the reaction time of 10 min, and OD represents the mean optical density of the parallel sample.

## Identification of protease-producing bacteria

- I. *Morphology identification*: The high-yield, protease-producing bacteria were activated on the plate, inoculated on the casein medium by the spot inoculation method, and subjected to Gram staining. The morphology was observed under a microscope [21].
- II. *Molecular characterization*: Using the DNA extracted from the high-yield, protease-producing bacteria as a template, the 16S rDNA gene in the bacterial genomic DNA was amplified using the bacterial universal primers 27F (5' -AGA GTT TCA TCT GGC TCA G-3') and 1492R (5' -GGT TAC CTT GTT ACG ACT T-3'). PCR was performed at 94°C for 5 minutes, followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 45 seconds, and extension at 72°C for 1.5 minutes. Finally, repair and extension at 72°C were performed for 10 minutes, and the sample was stored at 4°C. A 50 µL PCR amplification solution was prepared, and the amplification results were detected using 1% agarose electrophoresis to observe the specific products and their relative molecular masses. The amplified products were sent to Shanghai Sangon Biotech Co., Ltd. for sequencing [22]. Spliced sequences were compared with data in the NCBI database, and the standard bacteria with similar genetic relationships were analyzed using the MEGA 6.0, software to construct a phylogenetic tree used for visual observation of the genera and species of the bacteria [23].

*Physiological and biochemical tests*: Physiological and biochemical tests of the high-yield protease-producing bacteria were performed based on the 8<sup>th</sup> edition of *Bergey's Manual of Determinative Bacteriology* and the *Common Identification Methods of General Bacteria* [24].

## WGS analysis

- I. *DNA extraction*: Genomic DNA was extracted according to the instructions of the bacterial DNA extraction kit. To ensure the quality of WGS as well as the purity and concentration of the DNA, the samples were sent to Guangzhou Gene Denovo Biotechnology Co., Ltd. for sequencing.
- II. *PacBio sequencing*: g-TUBE for genomic DNA shearing was performed to process genomic DNA into 8–10 kbp fragments, followed by DNA end repair. The obtained SMRT Bell DNA template was sent off for quality detection, and an Agilent 2100 bioanalyzer was used to evaluate the insert size. Sequencing was performed using PacBio platform.
- III. *Illumina sequencing*: Genomic DNA fragments were prepared using dsDNA fragmentase, followed by end repair, phosphorylation, addition of polyA tails, and sequencing adapters. DNA fragments were purified using AMPure XP magnetic beads, and target fragments of 300–400 bp were subjected to PCR to establish the sequence library. PCR products were further purified using AMPure XP magnetic beads and examined using an Agilent 2100 bioanalyzer. Qualified products were sequenced using the Illumina Novaseq 6000 platform.
- IV. *Bioinformatic analyses*: Low-quality data in the raw data (reads) of PacBio and Illumina sequencing were filtered out of the dataset to ensure the accuracy and reliability of the bioinformatic analyses. Clean data were assembled to analyze genomes containing target strains, including genomic components of strains, tandem repeat prediction, tRNA prediction, rRNA prediction, sRNA prediction, genomic island prediction, transposons prediction, and prephage prediction using TRF (version 4.09) [25], tRNAscan (version 1.3.1) [26], rRNAmmer (version 1.2) [27], cmscan (version 1.1.2) [28], IslandPath-DIMOB (version 1.0.0) [29], TransposonPSI (version: 20100822) [30], and PHAST (version 2.0) [31], respectively. Gene function annotation was performed using the following online databases: COG database (<https://www.ncbi.nlm.nih.gov/COG/>), KEGG database (<https://www.genome.jp/kegg/>), NR database (<ftp://ftp.ncbi.nih.gov:21/blast/db/FASTA/>), SwissProt database (<https://www.uniprot.org/>), and GO database (<http://geneontology.org/>).

## Results

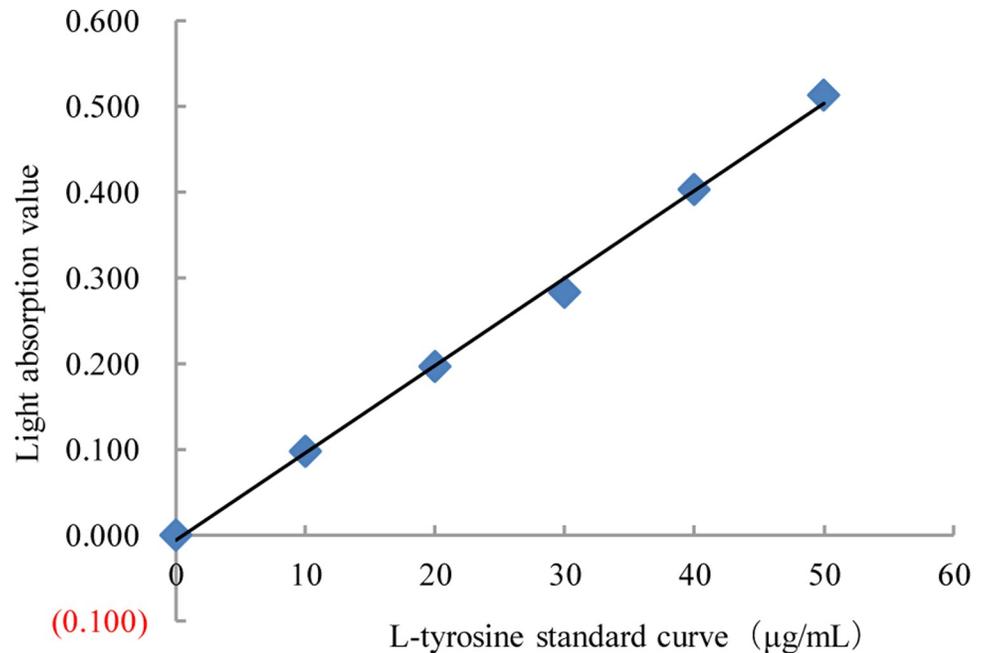
### Isolation of protease-producing bacteria

- I. *Isolation*: Through gradient dilution and plate coating, 296 strains of protease-producing bacteria were screened out from Daqu of ZhangGong Laojiu.
- II. *Initial screening*: A total of 296 strains of protease-producing bacteria were inoculated on casein medium and cultured at 37°C for 24 h in inverted position. The calculated D/d values are listed in [Table 1](#).

**Table 1. Diameter of transparent ring and strain diameter ratio.**

No.	d/cm	D/cm	D/d
DW-2	0.60±0.06	2.30±0.10	3.84±0.25
DW-7	0.57±0.15	2.50±0.06	4.39±0.98
DW-8	0.50±0.06	1.90±0.06	3.80±0.34
DW-22	0.60±0.06	2.20±0.12	3.69±0.62
DW-32	0.60±0.06	2.40±0.10	4.00±0.25
DW-40	0.50±0.00	1.80±0.06	3.60±0.12

<https://doi.org/10.1371/journal.pone.0264677.t001>



**Fig 1. L-tyrosine standard curve.**

<https://doi.org/10.1371/journal.pone.0264677.g001>

The D/d ratio of DW-7 was 4.39 and the highest among all of the isolated strains, which is likely attributed to the strong protease activity and degradation ability of casein. Protease activity in the subsequent highest six strains was measured to detect high-yield protease-producing bacteria.

### Screening of high-yield protease-producing bacteria

As shown in Fig 1, the regression equation of the L-tyrosine standard curve was  $y = 0.0102x - 0.0057$ , and the correlation coefficient (R) was 0.9977, indicating a significant linear relationship between protease activity and L-tyrosine production.

#### *I. Measurement of protease activity using the Folin-Ciocalteu method*

As shown in Table 2, the protease activity of DW-7 was the highest among the six strains tested at 99.54 U/mL. We subsequently examined the morphology, molecular biology, and physiological and biochemical properties of DW-7.

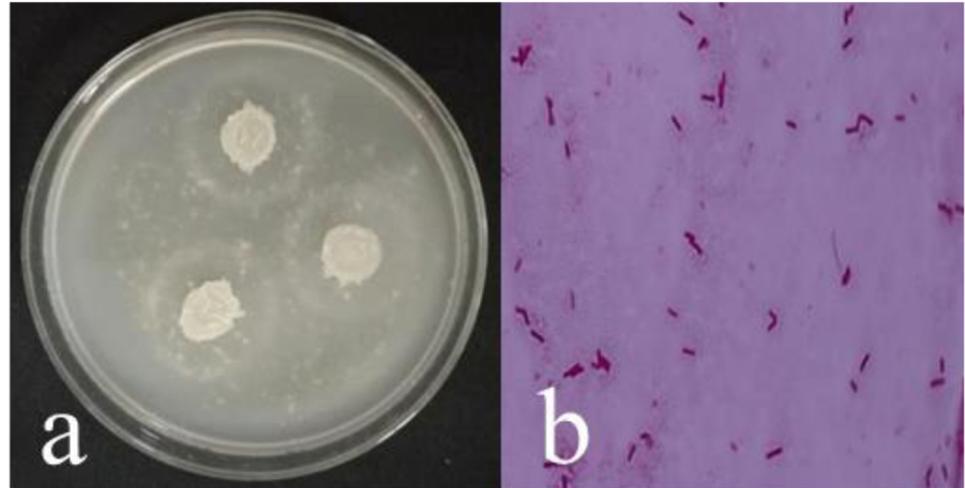
#### *II. Identification of high-yield protease-producing bacteria*

Cell colonies were milky white with a moist and viscous surface and wrinkles. They were round with smooth edges but lacked flagella (Fig 2A). Gram staining was performed to

**Table 2. Determination of protease activity.**

No.	OD-value	protease activity (U/mL)
DW-2	0.555±0.01	87.56±1.79
DW-7	0.631±0.01	99.54±2.23
DW-8	0.540±0.00	85.19±0.11
DW-22	0.525±0.00	82.82±0.00
DW-32	0.574±0.01	90.55±1.90
DW-40	0.497±0.03	78.41±4.24

<https://doi.org/10.1371/journal.pone.0264677.t002>



**Fig 2.** (a) Colony morphology of DW-7; (b) Morphology of DW-7 cells (magnification 1000X).

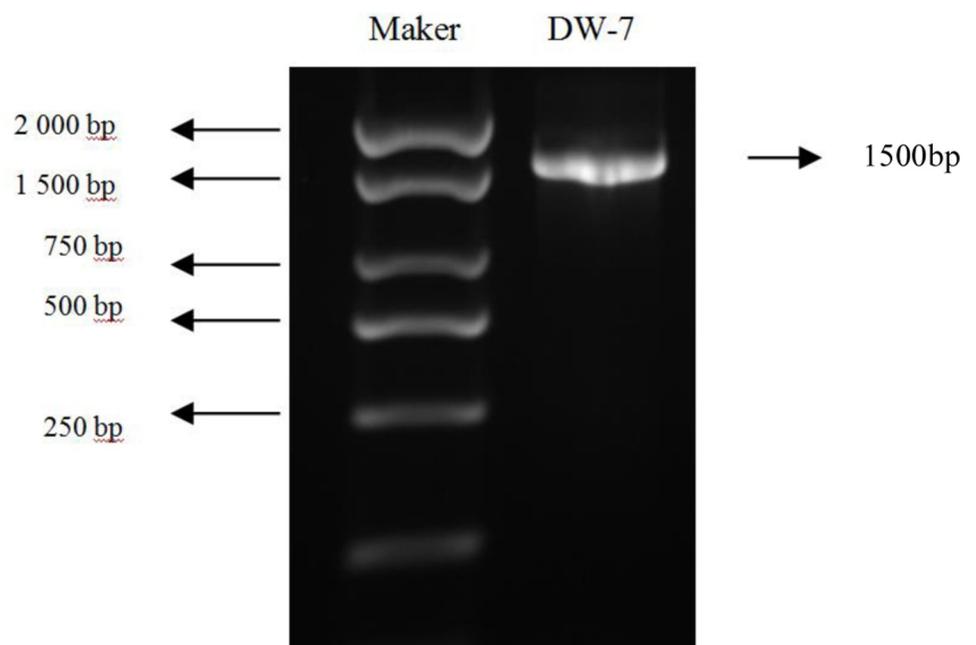
<https://doi.org/10.1371/journal.pone.0264677.g002>

observe the cell morphology under a microscope. The bacteria were stained red, suggesting the presence of Gram-negative bacteria. The individual bacterial cells were rod-shaped (Fig 2B).

### Molecular characterization

The DNA fragment of DW-7 was amplified by PCR and visualized by agarose gel electrophoresis. A phylogenetic tree was constructed using the MEGA 6.0. As shown in Fig 3, the size of the amplified product was approximately 1,500 bp, which is consistent with that of the target product.

Comparative analysis using the NCBI database showed that DW-7 shared the closest relationship with *Bacillus velezensis*, at 97.81% similarity (Fig 4).



**Fig 3.** Electrophoresis results of DW-7.

<https://doi.org/10.1371/journal.pone.0264677.g003>

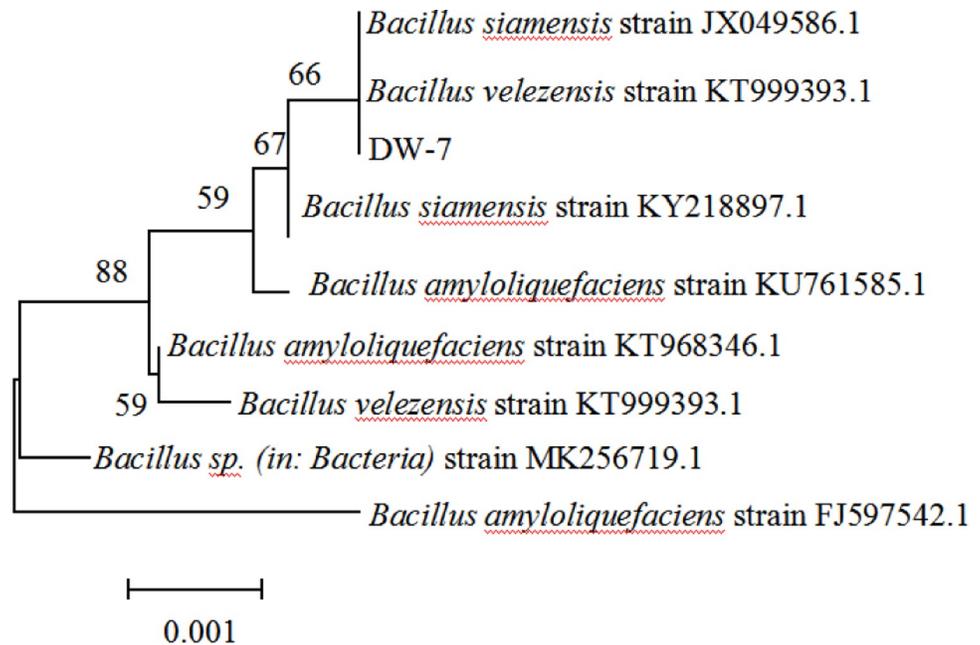


Fig 4. Phylogenetic tree of the DW-7 bacteria.

<https://doi.org/10.1371/journal.pone.0264677.g004>

### Physiological and biochemical test results

DW-7 presented a positive catalase test, meaning it was able to use D-mannitol to produce acid and survive on a medium containing 10% NaCl or at pH 5.0. Citrate was not used, but propionate was available (Table 3). Based on the physiological and biochemical test results, as well as 16S rDNA sequence analysis, DW-7 was confirmed to be *Bacillus velezensis*.

### Genome assembly of DW-7

WGS of DW-7 was performed using the PacBio sequencing platform, and the results showed that there were 193,753 effective reads after filtration, with a mean sequence length of 8380.3 bp (Table 4).

The DW-7 was comprised of a circular genome of 3,942,829 bp length. Protein-encoding genes were predicted using the NCBI database. A total of 3,662 genes were identified, with a combined length of 3,402,822 bp. The longest gene length was 16,299 bp, and the shortest was 75 bp. The GC content accounted for 46.46%. There were 64 transposons in the full genome of DW-7, with a total length of 4,515 bp. Additionally, there were 24 long interspersed repeats

Table 3. Physiological and biochemical analysis of the DW-7 bacterial strain.

Experiment	Result
Tolerance to 10% NaCl	+
Tolerance to pH 5.0	+
D-mannitol acid production	+
Catalase test	+
Propionate	+
Citrate	-

+ indicates a positive result, and – indicates a negative result.

<https://doi.org/10.1371/journal.pone.0264677.t003>

**Table 4. Genome statistics of the DW-7 strain.**

Type	Length	Amount
Sequence	3942829 bp	1
Gene	3402822 bp	3662
GC content	46.46%	
Total transposon length	4515 bp	64
Long scattered sequence	1869 bp	24
Short scattered sequence	1474 bp	22
Long terminal repeat sequence	185 bp	3
tandem repeat sequences	10086 bp	147
tRNA	6639 bp	86
rRNA	41367 bp	37
sRNA	2227 bp	17
Gene Islands	28970 bp	1
prophage	90882 bp	1

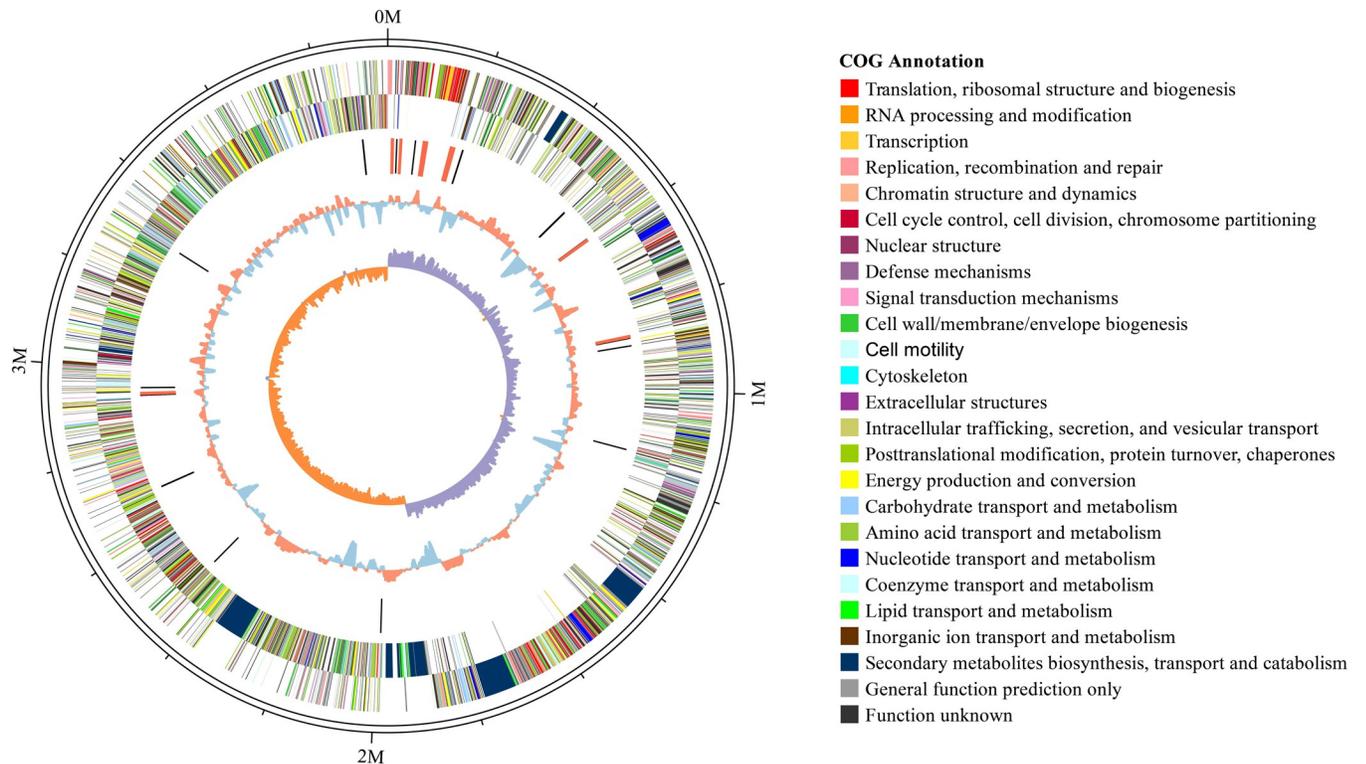
<https://doi.org/10.1371/journal.pone.0264677.t004>

(LINEs) with a total length of 1,869 bp accounting for 0.05% of the genome, and 22 short interspersed repeats (SINES) with a total length of 1,474 bp accounting for 0.04% of the genome. There were also three long terminal repeats (LTRs) with a combined length of 185 bp and 147 tandem repeats with a combined length of 10,068 bp, that together accounted for 0.26% of the genome. There were 86 tRNAs detected in the genome of DW-7, with a combined length of 6,639 bp and a mean length of 77 bp, as well as nine 5S rRNAs with a total length of 1,044 bp and a mean length of 116 bp. Nine 16S rRNAs with a combined length of 13,948 bp and a mean length of 1,549 bp were detected and nine 23S rRNAs with a combined length of 26,375 bp and a mean length of 2,930 bp were also detected. A total of 17 sRNAs with a combined length of 2,227 bp and a mean length of 131 bp were also detected. Finally, one genomic island was detected in the DW-7 with a total length of 28,970 bp as well as one prophage with a length of 90,882 bp.

### Bacterial genome components

Circular genome maps were depicted based on assembled genome sequences and prediction results of protein-encoding genes, which provided a visual of the genomic characteristics (Fig 5).

The outermost circle of the circular genome map of DW-7 is 3,402,822 bp in length. The second and third circles are positive and negative-strand coding regions, respectively. Different colors represent different functions annotated by the COG database. A total of 3,662 protein-encoding genes were identified, of which 2,796 (76.35%) genes were annotated by COG. Metabolism, information storage and processing, cell function and signal transmission, and atypical features were the four main functional categories. A total of 25 types were annotated, including general function prediction only, amino acid transport and metabolism, unknown function, transcription, carbohydrate transport, metabolism, and several other functions. The fourth circle is comprised of ncRNAs, in which tRNAs and rRNAs are colored in black and red, respectively. The fifth circle is the GC content, in which orange and blue indicate that the GC content was higher or lower than the mean value of the full genome, respectively. The sixth circle represents the GC-Skew value. Collectively, the circular genome maps of DW-7 directly showed the distribution of the DW-7 genome.



**Fig 5. Complete map of the genome of DW-7 bacteria.**

<https://doi.org/10.1371/journal.pone.0264677.g005>

### Gene annotation using COG

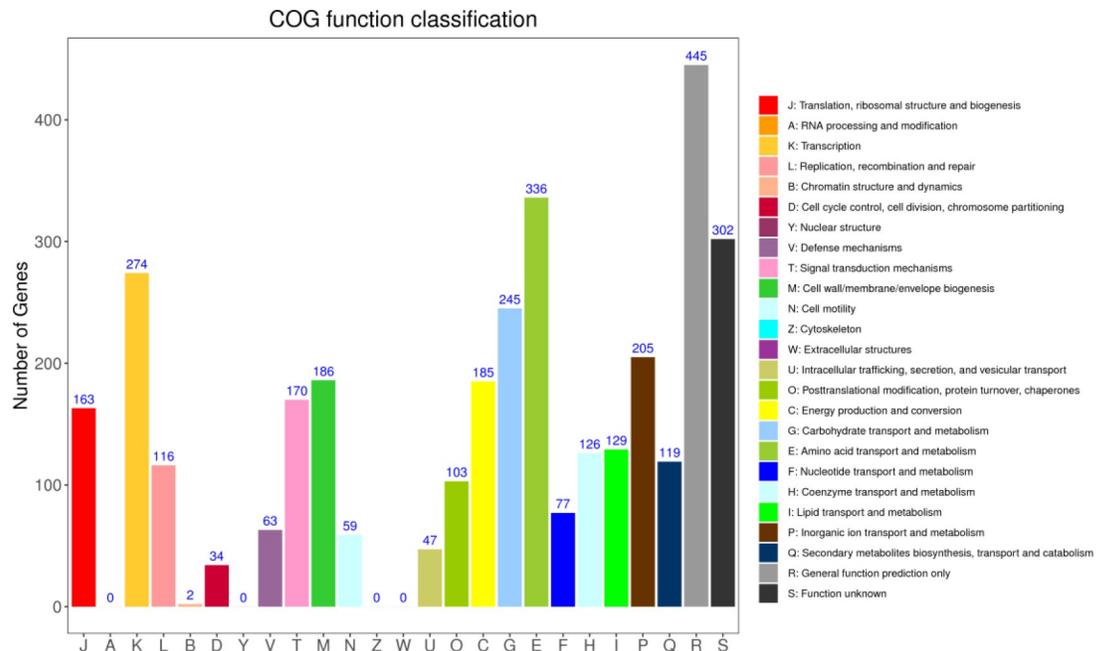
After comparison of the function annotations between the DIAMOND and COG databases, annotations were classified and depicted in Fig 6.

Function annotation of identified protein-encoding genes in DW-7 was performed using COG, and a total of 2,796 protein-coding genes involving 25 categories were annotated. Among them, 445 (15.92%) protein-encoding genes were enriched in the general function prediction only, which was the most enriched category, followed by 336 (12.02%) protein-encoding genes enriched in amino acid transport and metabolism category. Furthermore, 274 (9.80%), 245 (8.76%), and 205 (7.33%) protein-encoding genes were enriched in the transcription, carbohydrate transport and metabolism, and inorganic ion transport and metabolism categories, respectively. A combined total of 302 (10.80%) protein-encoding genes were enriched in all functional positions, which necessitates further exploration.

### Gene annotation using KEGG

The amino acid sequence of DW-7 was compared with data obtained from the KEGG database, and annotation results were obtained by matching the genes of target species and their corresponding function annotation information. A total of 2,249 genes in DW-7 were enriched in 23 pathways and five major functions of metabolism, including genetic information processing, environmental information processing, cell process, and biological system (Fig 7).

In the category of metabolic processes, 12 pathways were annotated, and 242 and 206 genes were enriched in the carbohydrate metabolism and amino acid metabolism pathways, respectively. Four pathways in the of genetic information processing category were annotated, with



**Fig 6. Functional classification of COG.**

<https://doi.org/10.1371/journal.pone.0264677.g006>

the translation pathway having the most annotations, in one of which 85 genes were enriched. There were 147 and 124 genes enriched in the membrane transport and signal transduction pathways, respectively, both of which are subcategories of the environmental information processing category. In addition, three pathways were annotated in the cell process category, in which the prokaryotic cell community was the most frequently annotated, in one of which 85 genes were enriched. In the biological systems category, four and three protein-encoding genes were enriched in the environmental adaptation and immune system subcategories, respectively.

According to the KEGG annotation information, 196 protein-encoding genes in DW-7 were enriched in the following pathways, all of which were correlated with presence of proteases involved in the metabolism of alanine, aspartic acid, glutamate, glycine, serine, threonine, ABC transporter, and transporter pathways. These protein-encoding genes included the KSG66-11125 (serine dehydratase), KSG66-13075 (aspartate kinase), KSG66-14990 (serine kinase), KSG66-10425 (aspartate aminotransferase), and several others.

### Gene annotation using nr and Swiss-Prot

The gene sequence of DW-7 was translated into the corresponding amino acid sequence, then compared with data from the nr database within the NCBI database (Fig 8).

A total of 3,662 genes in DW-7 were annotated using the nr database. Among them, 2,709; 389; 198; 107; 70; 53; and 41 genes were enriched in the primary seven identified species *B. velezensis*, *B. amyloliquefaciens*, *B. subtilis*, *B. abscessus*, and *Bacillus sp.5B6*, and *Streptococcus pneumoniae*, respectively.

A total of 3,399 genes in DW-7 were functionally annotated in the Swiss Prot database, which is a selected protein sequence database that describes protein function, structure, post-translational modification, mutations, and other characteristics.

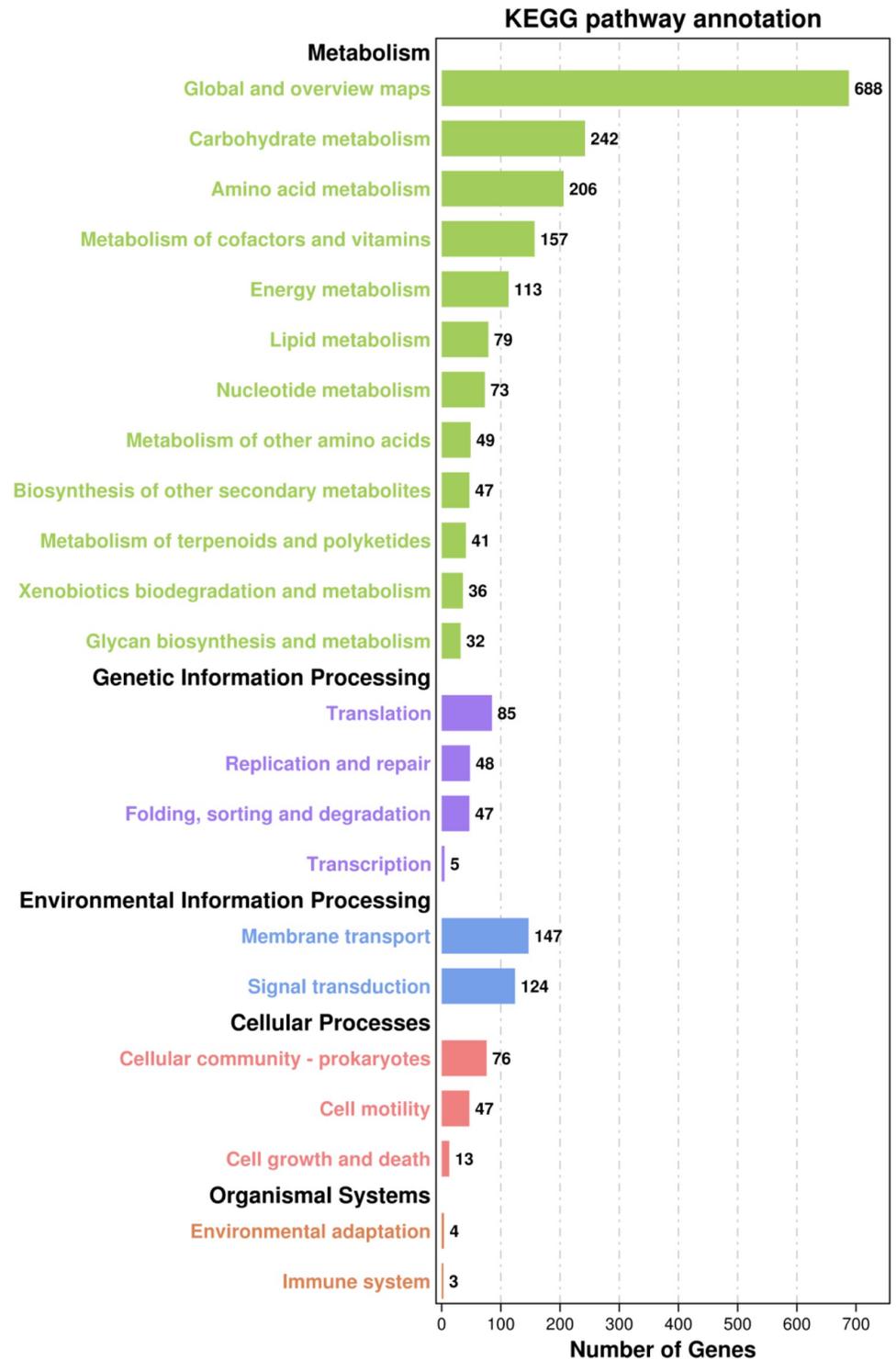


Fig 7. Bacterial gene functional annotation KEGG metabolic pathway.

<https://doi.org/10.1371/journal.pone.0264677.g007>

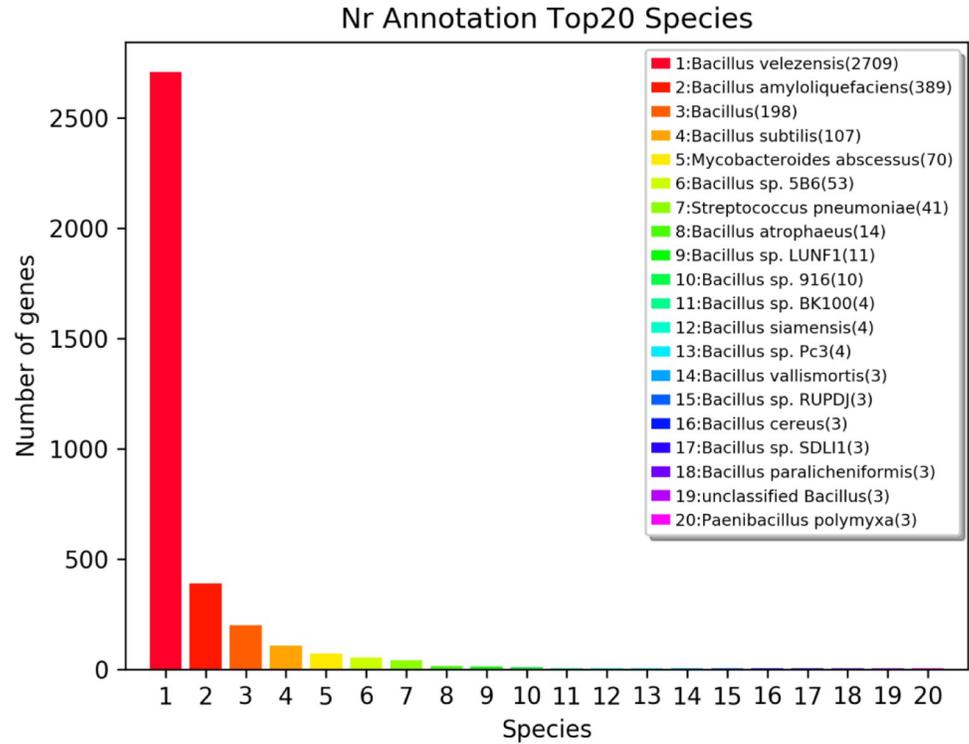


Fig 8. Annotated species statistics of Nr database (top 20 species).

<https://doi.org/10.1371/journal.pone.0264677.g008>

### Gene annotation using GO

According to the annotation information in the nr database, GO function annotations were identified to comprehensively describe the attributes of genes and gene products in DW-7 (Fig 9).

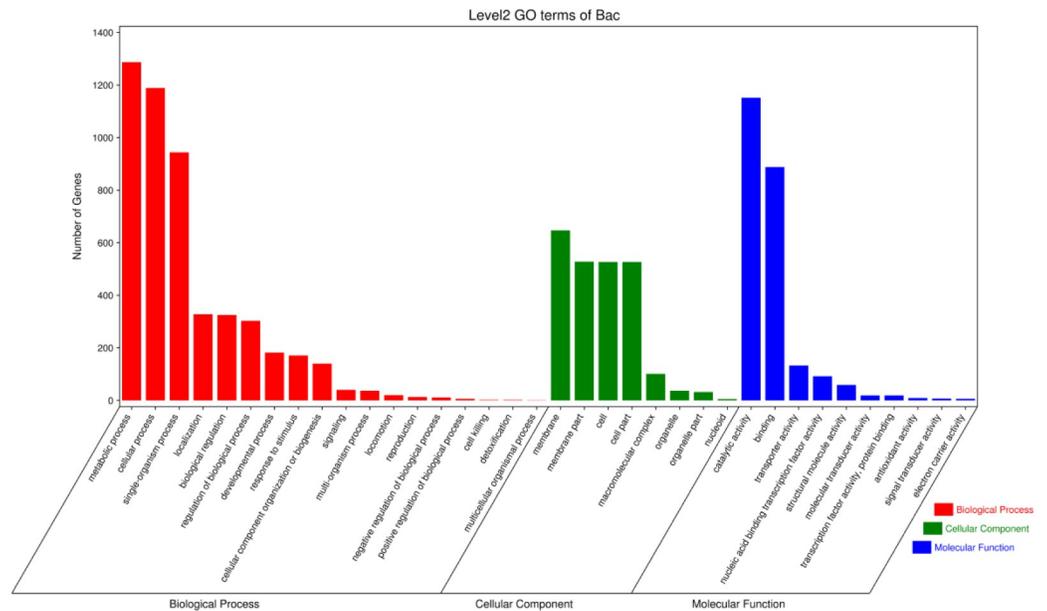


Fig 9. GO classification of bacterial gene function annotation.

<https://doi.org/10.1371/journal.pone.0264677.g009>

A total of 2,127 genes from DW-7 were annotated in the GO database. GO terms including biological process (BP), cellular component (CC), and molecular function (MF), comprising 18, 10, and 10 elements, respectively. The top three annotated BPs were the metabolic processes, biological processes, and single-cell tissue processes, in which 1,287, 1,189, and 944 genes were enriched, respectively. There were 647, 528, 527, and 527 genes enriched in the top four CCs, including the cell membranes, cell membrane compositions, cells, and genes related to cell components, respectively. A total of 1,152, 888, and 133 genes were enriched in the most annotated MFs of catalytic activity term: gene and binding, and transport activity, respectively.

## Discussion

In this study, 296 strains of protease-producing bacteria were isolated from Daqu from Zhang-Gong Laojiu Wine Co., Ltd. Additionally, one strain of high protease-producing bacteria, DW-7, with an enzyme activity of 99.54 U/mL was isolated by combining the transparent circle screening method and the forinol screening method. Morphological observations, 16S rDNA molecular biology identification, as well as physiological and biochemical tests, were performed on the isolated DW-7 which was identified as *Bacillus velezensis*. Few studies have examined high protease-producing strains using baijiu Daqu as samples. Yuan Xianling et al. [14] isolated and purified five strains of protease-producing bacteria from strong baijiu Daqu, including *Bacillus thuringiensis*, *Bacillus subtilis*, *Bacillus cereus*, *Bacillus multilocularis* and *Bacillus subtilis* with enzyme activities of 41.75 U/mL, 31.90 U/mL, 53.40 U/mL, 46.10 U/mL and 44.55 U/mL, respectively; Zhao Qunli [32] et al. used Moutai-flavored baijiu Daqu as the raw material, from which they screened a strain 3 J-1 with high enzyme activity and a protease activity of 70.15 U/mL; Huang et al. [33] also screened a strain from Moutai-flavored baijiu Daqu, identified as *A. hennebergii*, with a maximum enzyme activity of 71.13 U/mL after optimization of pH, carbon source, nitrogen source and metal ions; and a high enzyme activity level of 99.54 U/mL. Protease-producing bacteria are widely found in marine [34], animal intestine [35], soil [36], and other environments, and screening for these high protease-producing strains can be useful for industrial production and food processing.

Proteases are a class of enzymes that catalyze the hydrolysis of proteins into peptides and amino acids and degrade proteins into raw materials to form small peptides or amino acids used in yeast and lactic acid bacteria fermentation as a nitrogen source to promote the growth of brewing microorganisms [37]. Some of the amino acids in the protein degradation products are flavoring substances themselves, while others are precursors of aroma components that ultimately affect the formation and quality of the flavor of white wine [38]; and thus are indispensable enzymes for white wine production. Studies have shown that the hydrolytic enzyme system in soybean currants is mainly protease [39]. When the protease content is normal, it can significantly inhibit the production of hetero-ethanol oil [40], and the protease activity of superior grade barley is higher than that of normal grade barley [41]. Deng et al. compared different colors of high-temperature barley and found that QW protease activity was higher than QR, QY, and QB, while the windiness of bacterial enzyme-encoding genes was higher in QW [42]. Numerous studies have shown that bacteria of the genus *Bacillus* have a strong protease production capacity [2, 43].

To further investigate the protease metabolism mechanism and metabolic pathway, we performed WGS of DW-7, and determined that the genome length of this strain was 3,942,829 bp, with a GC content of 46.45%. The combined length of DW-7's coding genes was 3,402,822 bp, and by gene function annotation, there were 2,796 genes annotated in the COG database, with the main focus on general function prediction, amino acid transport and metabolism,

transcription, carbohydrate transport and metabolism, and inorganic ion transport and metabolism. In the KEGG database, 2,283 genes were annotated, mainly in carbohydrate metabolism, amino acid metabolic pathway, translation pathway, membrane transport, and signal transduction pathway. 2,127 genes were annotated in the GO database, mainly in metabolic function, cellular process, cell membrane, cell membrane components, catalytic activity, gene, and binding. Protease genes do not have biological functions after transcription and translation to form proteins, rather only after modification by a series of chemical reactions, such as ubiquitination, phosphorylation, acetylation, glycosylation, methylation, or lipidylation, can there be more protein types with complex structures, precise regulation, specific action, and perfect function [44]. In this study, COG function prediction, KEGG metabolic pathway annotation, and GO annotation of the genes encoded by this strain showed normal gene transcription and translation and a relatively high number of annotated genes, which indicated that the enzyme metabolism of this bacterium is active. Protease gene expression metabolic pathways were mainly focused on alanine, aspartate, glutamate, glycine, serine, and threonine, as well as ABC transporter proteins; and transporter protein. The main proteases involved in these pathways were serine dehydratase, aspartate kinase, serine kinase, and aspartate aminotransferase. Prajapati et al. [45] resolved the protease gene of *Bacillus amyloliquefaciens* strain KCP2 by WGS, while Li et al. [46] sequenced the genome of *Laceyella sacchari* FBKL4 and found that the genome contains genes encoding key enzymes, such as proteases and peptidases, along the tetramethylpyrazine metabolic pathway. Proteases are classified into four categories based on their catalytic role in a variety of physiological pathways, including serine proteases, aspartate proteases, cysteine proteases, and metalloproteases [47]. Among the proteases, serine proteases are the most common, accounting for nearly one-third of known proteases [48, 49]. The 196 protease-related gene annotations in this study were significantly enriched for serine dehydratase, aspartate kinase, serine kinase, and aspartate aminotransferase, which laterally reflected that the bacterium was active in serine protease and aspartate protease synthesis. This result also indicated that this high protease-producing *Bacillus* sp. metabolized mainly these two types of catalytic proteases.

## Conclusion

In this experiment, a strain of high-yield protease bacteria DW-7 with 99.54 U/mL protease activity was isolated from medium-temperature Daqu provided by ZhangGong LaoJiu Wine Co., Ltd. Through morphological observation, 16S rDNA sequence analysis, and physiological and biochemical tests, DW-7 was determined to be *Bacillus velezensis*. WGS analysis of DW-7 revealed 196 genes that were highly related to proteases predominantly enriched in the metabolism of alanine, aspartic acid, glutamate, glycine, serine, threonine, ABC transporter proteins, and transporter pathways.

Overall, our study preliminarily explored the biological characteristics and metabolic functions of the high-yield protease-producing strain DW-7 by WGS analysis. Abundant genomic information was obtained that provided useful references for identifying the functional characteristics of the strain, as well as the foundation for elucidating the background of the high-yield protease-producing feature. Our results are of great significance for the development of novel high-yield proteases from Daqu.

## Supporting information

**S1 File.**  
(PDF)

**S1 Raw data.**

(ZIP)

**S1 Raw images.**

(PDF)

**Acknowledgments**

We thank ZhangGong LaoJiu Wine Co. Ltd. for providing the samples and the distillery staff for their help in sampling. We also thank the members of the subject group for supporting this study and contributing to the experimental work. We would like to thank Editage ([www.editage.cn](http://www.editage.cn)) for English language editing.

**Author Contributions**

**Conceptualization:** Zhijun Zhao.

**Writing – original draft:** Yanbo Liu, Junying Fu, Linlin Wang.

**Writing – review & editing:** Yanbo Liu, Zhijun Zhao, Huihui Wang, Suna Han, Xiyu Sun, Chunmei Pan.

**References**

1. Chen Y, Li K, Liu T, Li R, Fu G, Wan Y, et al. Analysis of Difference in Microbial Community and Physicochemical Indices Between Surface and Central Parts of Chinese Special-Flavor *Baijiu Daqu*. *Front Microbiol*. 2021 Jan 14; 11:592421. <https://doi.org/10.3389/fmicb.2020.592421> PMID: 33519730
2. Liu J, Chen J, Fan Y, Huang X, Han B. Biochemical characterisation and dominance of different hydrolases in different types of Daqu—a Chinese industrial fermentation starter. *Journal of the Science of Food and Agriculture*. 2018, Jan 25; 98(1):113–121. <https://doi.org/10.1002/jsfa.8445> PMID: 28542883
3. Fan G, Fu Z, Sun B, Zhang Y, Wang X, Xia Y, et al. Roles of aging in the production of light-flavored Daqu. *Journal of bioscience and bioengineering*. 2019 Mar 19; 127(3):309–317. <https://doi.org/10.1016/j.jbiosc.2018.08.005> PMID: 30241902
4. Yang JG, Dou X, Ma YY. Diversity and dynamic succession of microorganisms during Daqu preparation for Luzhou-flavour liquor using second-generation sequencing technology. *Journal of the Institute of Brewing*. 2018, Oct 3; 124(4): 498–507. <https://doi.org/10.1002/jib.528>
5. Yin X, Yoshizaki Y, Ikenaga M, Han XL, Okutsu K, Futagami T, et al. Manufacturing impact of the solid-state saccharification process in rice-flavor baijiu production. *Journal of bioscience and bioengineering*. 2020 Mar; 129(3):315–321. <https://doi.org/10.1016/j.jbiosc.2019.09.017> PMID: 31718882
6. Zhang Y, Huang Y, Qiu S. Solid-state fermentation condition and its optimization of protease production by *Geotrichum candidum* in Moutai-flavor Daqu. *China Brewing*. 2012, Mar 15; 31(03):35–38. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=ZNGZ201203011&DbName=CJFQ2012>
7. Fan G, Sun B, Fu Z, Xia Y, Huang M, Xu C, et al. Analysis of physicochemical indices, volatile flavor components, and microbial community of a light-flavor *Daqu*. *Journal of the American Society of Brewing Chemists*. 2018, Mar 29; 76(3): 209–218. <https://doi.org/10.1080/03610470.2018.1424402>
8. Wu Q, Chen L, Xu Y. Yeast community associated with the solid state fermentation of traditional Chinese Maotai-flavor liquor. *International journal of food microbiology*. 2013, Sep 2; 166(2): 323–330. <https://doi.org/10.1016/j.ijfoodmicro.2013.07.003> PMID: 23978339
9. Tyler A D, Mataseje L, Urfano C J, Schmidt L, Antonation K S, Mulvey M R, et al. Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Scientific reports*. 2018 Jul 19; 8(1):10931. <https://doi.org/10.1038/s41598-018-29334-5> PMID: 30026559
10. McDermott PF, Tyson GH, Kabera C, Chen Y, Li C, Folster JP, et al. Whole-Genome Sequencing for Detecting Antimicrobial Resistance in Nontyphoidal Salmonella. *Antimicrob Agents Chemother*. 2016 Aug 22; 60(9):5515–20. <https://doi.org/10.1128/AAC.01030-16> PMID: 27381390
11. Ronholm J, Nasheri N, Petronella N, Pagotto F. Navigating Microbiological Food Safety in the Era of Whole-Genome Sequencing. *Clin Microbiol Rev*. 2016 Oct; 29(4):837–57. <https://doi.org/10.1128/CMR.00056-16> PMID: 27559074

12. Lu Z, Yan J, Li EZ, Liu JY, Lu ZZ, Zhang JS, et al. The screening of high-yielding neutral protease bacteria and the conditions of their enzyme production. *Heilongjiang Agricultural Science*. 2016 May 10;(05): 110–114. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=HLJN201605028&DbName=CJFQ2016>
13. Nie HF, Jiang YL, Shen Y, Chen L, Xu Y, Gao H, et al. The separation and identification of high-yielding protease bacteria in the sauce-scented large-shaped large-scale song. *Liquor-Making Science & Technology*. 2015 Jul 10;(12): 41–44. <https://doi.org/10.13746/j.njkj.2015157>
14. Yuan XL, Huang D, Peng J. Isolation of Protease-producing Bacteria Strains from Luzhou-flavor Daqu and Study on Its Protease-producing Conditions. *Liquor-Making Science & Technology*. 2012, Sep 18;(09):51–53. <https://doi.org/10.13746/j.njkj.2012.09.021>
15. Zhong XM, Fu J, Lan J, Li L, Yang XQ. Comparison of protease conditions and protease properties in *Actinomyces elegans*, *Aspegillus oryzae* and *Rhizopus oligosporus*. *Food and Fermentation Industry*. 2009, May 30; 35(05): 40–44. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=SPFX200905018&DbName=CJFQ2009>
16. Wang J, Wang XD, Luo XY, Qiu SY, Xiao P, Zhang XL, et al. Screening and identification of functional bacteria with high-yield protease from Moutai-flavor Daqu. *China Brewing*. 2015 Oct 15; 34(10): 43–46. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=ZNGZ201510015&DbName=CJFQ2015>
17. Wei CH, Huang ZG, Huang D, Deng J, Yang XD, Qiao XY. Separation and identification of high protease yielding strains and study on enzyme production performance of high-temperature Daqu. *Food & Machinery*. 2014 Jul 18; 30(4): 24–29. <https://doi.org/10.13652/j.issn.1003-5788.2014.04.007>
18. Ming HM, Dong RL, Xu DF, Guo Z, Zeng HL, Shen CP, et al. Separation and Preliminary Identification of Dominant Microbes in Nong-flavor Daqu. *Liquor-Making Science & Technology*. 2013 Sep 22;(12): 57–60. <https://doi.org/10.13746/j.njkj.2013.12.022>
19. Wang PP. Study on Isolation of Protease-producing Microorganism and Production of High-quality Protein Materials by Solid-state Fermentation. M.Sc. Thesis, Henan Agricultural University. 2010 Jun 1. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=2010270535.nh&DbName=CMFD2011>
20. Xiao W, Wang XD, Ban SD, Kong WB, Qiu SY, Hu PG. Screening and Identification of Thermophilic Bacterial Strain in Jiangxiang Daqu. *Liquor-Making Science & Technology*. 2014 Dec 8;(02): 50–53. <https://doi.org/10.13746/j.njkj.2014216>
21. Wang YT, Zheng RF, Han MY, Li FW. Screening and Identification of Bacillus Producing Protease and Amylase. *Feed Review*, 2018 Mar 25;(03): 4–7. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=SLBL201803002&DbName=CJFQ2018>
22. Faja OM, Sharad AA, Younis KM, Alwan MG, Mohammed BJ, Ahmad A. Isolation, detection of virulence genes, antibiotic resistance genes, plasmid profile, and molecular typing among *Vibrio parahaemolyticus* isolated in Malaysian seawater from recreational beaches and fish. *Vet World*. 2019 Jul 28; 12(7):1140–1149. <https://doi.org/10.14202/vetworld.2019.1140-1149> PMID: 31528045
23. Stackebrand TE, Goebel BM. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology*. 1994 Oct 1; 44(4): 846–849. <https://doi.org/10.1099/00207713-44-4-846>
24. Dvorak HF, Orenstein NS, Dvorak AM, Hammond ME, Roblin RO, Feder J, et al. Isolation of the cytoplasmic granules of guinea pig basophilic leukocytes: identification of esterase and protease activities. *J Immunol*. 1977 Jul; 119(1):38–46. PMID: 874325
25. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999 Jan 15; 27(2):573–80. <https://doi.org/10.1093/nar/27.2.573> PMID: 9862982
26. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 2016 Jul 8; 44(W1):W54–7. <https://doi.org/10.1093/nar/gkw413> PMID: 27174935
27. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007 Apr 22; 35(9):3100–8. <https://doi.org/10.1093/nar/gkm160> PMID: 17452365
28. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013 Nov 15; 29(22):2933–5. <https://doi.org/10.1093/bioinformatics/btt509> PMID: 24008419
29. Bertelli C, Brinkman FSL. Improved genomic island predictions with IslandPath-DIMOB. *Bioinformatics*. 2018 Jul 1; 34(13):2161–2167. <https://doi.org/10.1093/bioinformatics/bty095> PMID: 29905770
30. Haas B. TransposonPSI: an application of PSI-Blast to mine (retro-) transposon ORF homologies. Broad Institute, Cambridge, MA, USA, 2007. Available from: <http://transposonpsi.sourceforge.net>

31. Fouts DE. Phage\_Finder: automated identification and classification of prophage regions in complete bacterial genome sequences. *Nucleic Acids Res.* 2006; 34(20):5839–51. <https://doi.org/10.1093/nar/gkl732> PMID: 17062630
32. Zhao QL, Wu TX, Liu W. Isolation of a Bacteria Strain from Jiangxiang Daqu and Identification of Its Enzyme Activity. *Liquor-Making Science & Technology.* 2015 Jul 22;(11): 41–44. <https://doi.org/10.13746/j.njkj.2015105>
33. Huang Y, Wu Q, Xu Y. Isolation and identification of a black *Aspergillus* strain and the effect of its novel protease on the aroma of *Moutai-flavoured* liquor. *Journal of the Institute of Brewing.* 2014 Apr 22; 120(3): 268–276. <https://doi.org/10.1002/jib.135>
34. Thevarajoo S, Selvaratnam C, Chan KG, Goh KM, Chong CS. Draft genome sequence of *Vitellibacter vladvostokensis* KMM 3516<sup>(T)</sup>: a protease-producing bacterium. *Mar Genomics.* 2015 Oct; 23:49–50. <https://doi.org/10.1016/j.margen.2015.04.009> PMID: 25957696
35. Li F, Gao F, Tan J, Fan C, Sun H, Yan J, et al. Characterization and identification of enzyme-producing microflora isolated from the gut of sea cucumber *Apostichopus japonicus*. *Chin. J. Ocean. Limnol.* 2015 Jun 17; 34:153–162. <https://doi.org/https%3A//doi.org/10.1007/s00343-015-4149-z>
36. Chu WH. Optimization of extracellular alkaline protease production from species of *Bacillus*. *J Ind Microbiol Biotechnol.* 2007 Mar; 34(3):241–5. <https://doi.org/10.1007/s10295-006-0192-2> PMID: 17171551
37. Wang H, Gu Y, Zhou W, Zhao D, Qiao Z, Zheng J, et al. Adaptability of a Caproate-Producing Bacterium Contributes to Its Dominance in an Anaerobic Fermentation System. *Appl Environ Microbiol.* 2021 Sep 28; 87(20):e0120321. <https://doi.org/10.1128/AEM.01203-21> PMID: 34378978
38. Feng LF. Analysis of protease effect and isolation identification and application of functional bacteria in Daqu. M.Sc. Thesis, Shanxi Normal University. 2018 Jun 13. Available from: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=1019021541.nh&DbName=CMFD2019>
39. Hu BD, Qiu SY, Zhou HX, Wang XD. Relationships among Physicochemical Indices and Hydrolyzing Enzyme Systems and Enzymes-produced-ability in Jiangxiang Daqu. *Modern Food Science and Technology.* 2016 Nov 11; 33(2): 99–106. <https://doi.org/10.13982/j.mfst.1673-9078.2017.2.016>
40. Sun JX. Methods for Isoamyl Alcohol Control in Wine. *Modern Food Science and Technology.* 2012 Nov 15; 28(11):1541–1544. <https://doi.org/10.13982/j.mfst.1673-9078.2012.11.033>
41. Fan WY. Metaproteomic Analysis of Enzymatic Composition in Nong-flavor Daqu. M.Sc. Thesis, Jiangnan University, 2021 Jun 1. <https://doi.org/10.27169/d.cnki.gwqgu.2021.000945>
42. Deng L, Mao X, Liu D, Ning X Q, Shen Y, Chen B, et al. Comparative Analysis of Physicochemical Properties and Microbial Composition in High-Temperature *Daqu* With Different Colors. *Front Microbiol.* 2020 Nov 27; 11:588117. <https://doi.org/10.3389/fmicb.2020.588117> PMID: 33329462
43. Wu X, Jiang Q, Wang Z, Xu Y, Chen W, Sun J, et al. Diversity, enzyme production and antibacterial activity of *Bacillus* strains isolated from sesame-flavored liquor Daqu. *Arch Microbiol.* 2021 Nov; 203(9):5831–5839. <https://doi.org/10.1007/s00203-021-02552-8> PMID: 34491390
44. Zhang Y, Fang C, Bao H, Yuan W, Lu H. Discover the Post-Translational Modification Proteome Using Mass Spectrometry. *Chinese Journal of Chemistry.* 2021 Feb 19; 39(3): 550–558. <https://doi.org/10.1002/cjoc.202000515>
45. Prajapati VS, Ray S, Narayan J, Joshi CC, Patel KC, Trivedi UB, et al. Draft genome sequence of a thermostable, alkaliphilic  $\alpha$ -amylase and protease producing *Bacillus amyloliquefaciens* strain KCP<sub>2</sub>. *3 Biotech.* 2017 Dec; 7(6):372. <https://doi.org/10.1007/s13205-017-1005-1> PMID: 29071169
46. Li D, Huang W, Wang C, Qiu S. The complete genome sequence of the thermophilic bacterium *Laceyella sacchari* FBKL4.010 reveals the basis for tetramethylpyrazine biosynthesis in Moutai-flavor Daqu. *Microbiologyopen.* 2019 Dec; 8(12):e922. <https://doi.org/10.1002/mbo3.922> PMID: 31482696
47. Mamo J, Assefa F. The role of microbial aspartic protease enzyme in food and beverage industries. *Journal of Food Quality.* 2018 Jul 3; 15. <https://doi.org/10.1155/2018/7957269>
48. Harish BS, Uppuluri KB. Microbial serine protease inhibitors and their therapeutic applications. *Int J Biol Macromol.* 2018 Feb; 107(Pt B):1373–1387. <https://doi.org/10.1016/j.ijbiomac.2017.09.115> PMID: 28970170
49. Ekici OD, Paetzel M, Dalbey RE. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* 2008 Dec; 17(12):2023–37. <https://doi.org/10.1110/ps.035436.108> PMID: 18824507