



# The Species Identification and Genomic Analysis of *Haemobacillus shengwangii*: A Novel Pathogenic Bacterium Isolated From a Critically Ill Patient With Bloodstream Infection

## OPEN ACCESS

### Edited by:

Zheng Zhang,  
Shandong University, China

### Reviewed by:

Quan Shen,  
Jiangsu University, China  
Fen Qu,  
Aviation General Hospital of China  
Medical University, China  
Vibhor Tak,  
AIIMS Jodhpur, India

### \*Correspondence:

Han Xia  
xiahan@hugobiotech.com  
Yinghua Yuan  
yuan\_yinghua@126.com  
Sheng Wang  
wangsheng@tongji.edu.cn

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 13 April 2022

Accepted: 17 May 2022

Published: 14 June 2022

### Citation:

Du Y, Li X, Liu Y, Mu S, Shen D,  
Fan S, Lou Z, Zhang S, Xia H, Yuan Y  
and Wang S (2022) The Species  
Identification and Genomic Analysis  
of *Haemobacillus shengwangii*:  
A Novel Pathogenic Bacterium  
Isolated From a Critically Ill Patient  
With Bloodstream Infection.  
*Front. Microbiol.* 13:919169.  
doi: 10.3389/fmicb.2022.919169

Yingying Du<sup>1†</sup>, Xuming Li<sup>2†</sup>, Yuhao Liu<sup>1†</sup>, Shikui Mu<sup>1</sup>, Dandan Shen<sup>3</sup>, Shu Fan<sup>2</sup>,  
Zheng Lou<sup>2</sup>, Shouqin Zhang<sup>1</sup>, Han Xia<sup>2\*</sup>, Yinghua Yuan<sup>3\*</sup> and Sheng Wang<sup>1\*</sup>

<sup>1</sup> Department of Critical Care Medicine, School of Medicine, Shanghai Tenth People's Hospital, Tongji University, Shanghai, China, <sup>2</sup> Department of Scientific Affairs, Hugo Biotech Co., Ltd., Beijing, China, <sup>3</sup> Department of Clinical Microbiology, School of Medicine, Shanghai Tenth People's Hospital, Tongji University, Shanghai, China

Since the first strain related to Thermicanaceae was reported in 1999, almost no literature on Thermicanaceae is available, particularly its genomics. We recently isolated a novel pathogenic bacterium, the  $\Delta$  strain DYY3, from the blood sample of a critically ill patient. The morphological, physiological, and biochemical characteristics of  $\Delta$  strain DYY3 were presented in this study, and the virulence factor genes and antibiotic resistance of DYY3 were also determined. Interestingly, the average nucleotide identity (ANI) and core-genes average amino acid identity (cAAI) analysis indicated that  $\Delta$  strain DYY3 was genus novel and species novel. Moreover, phylogenetic analysis based on both 16S rRNA gene and whole genomic core gene sequences suggested that  $\Delta$  strain DYY3 belonged to the family Thermicanaceae, and this novel taxon was thus named *Haemobacillus shengwangii* gen. nov., sp. nov. Besides, both the whole genome-based phylogenetic tree and amino acid identity analysis indicated that *Thermicanus aegyptius*, *Hydrogenibacillus schlegelii*, *Brockia lithotrophica*, and the newly discovered species *H. shengwangii* should belong to Thermicanaceae at the family level, and *T. aegyptius* was the closest species to *H. shengwangii*. We also constructed the first high-quality genome in the family Thermicanaceae using the next-generation sequencing (NGS) and single-molecule real-time (SMRT) sequencing technologies, which certainly contributed to further genomics studies and metagenomic-based pathogenic detection in the future.

**Keywords:** catheter-associated bloodstream infection, pathogenic bacterium, novel species identification, Thermicanaceae, genome *de novo* assembly, single-molecule real-time sequencing, comparative genomics

## INTRODUCTION

Catheter-related bloodstream infection (CRBSI) is a frequent and life-threatening condition in the intensive care unit (ICU), which is associated with increased morbidity, mortality, and healthcare costs (Schwab et al., 2018; Gerver et al., 2020; Zeng et al., 2021). For example, according to a prospective multi-center study in China, the average incidence of CRBSI and the mortality due to CRBSI in ICU were 1.5/1,000 catheter days and 18.09%, respectively (Zeng et al., 2021). A similar

incidence of CRBSI was reported in studies from European countries (Schwab et al., 2018; Gerver et al., 2020), and even higher rates were found in developing countries, up to 5.3/1,000 catheter days with 28–30% of mortality (Rosenthal et al., 2021).

To improve the clinical outcomes of patients with CRBSI, a rapid and accurate diagnosis of the causative pathogen is a critical step (Zhong et al., 2021). The current guideline recommends diagnosing CRBSI by hemoculture in suspected patients (Mermel et al., 2009), but this approach takes 48–96 h to isolate, identify, and perform antibiotic susceptibility tests (Tabak et al., 2018). Furthermore, it is often challenging to culture many fastidious or uncultivable pathogens in standard automated systems (Murray and Masur, 2012). Therefore, emerging technologies, especially high-throughput sequencing, were attempted to replace conventional culture-based methods and initiate timely targeted anti-infection therapy (Li et al., 2021). However, few clinical studies thus far identified and classified unknown species with extraordinary genetic distances from known species.

The only reported bacteria that might belong to the family Thermicaceae were *Thermicanus aegyptius*, *Hydrogenibacillus schlegelii*, *Brockia lithotrophica*, and *Carbobacillus altaicus*. *T. aegyptius* was first identified from soil and described as a fermentative microaerophile in 1999 (Gossner et al., 1999), and the reference genome was *T. aegyptius* DSM 12793, available online in 2013. *H. schlegelii* was originally named *Bacillus schlegelii* in 1979 (Schenk and Aragno, 1979) and was transferred to be a novel genus due to its massive divergence from other species in the genus *Bacillus* in 2013 (Kampfer et al., 2013). *H. schlegelii* was known for its ability of hydrogen-oxidizing (Barbosa et al., 2020) and was classified into order Bacillales and family Bacillaceae with NCBI taxonomy ID of 1484. *B. lithotrophica* was isolated from a hot spring in Russia and reported as a new taxon in 2013 (Perevalova et al., 2013). Finally, *C. altaicus* was still a candidate taxon classified into order Bacillales and family Bacillales *incertae sedis* with NCBI taxonomy ID 2163959. However, whether the classification method of the above bacteria is correct needs to be verified by genomic analysis, as this approach is increasingly being accepted as reliable data for bacterial taxonomy and species identification (Hayashi Sant'Anna et al., 2019).

In this study, the  $\Delta$  strain DYY3, isolated from the blood sample of a critically ill patient diagnosed with CRBSI, cannot be identified by VITEK-MS automatic microbiological analyzer and the 16S rRNA sequence analysis. Thus, the high-quality genome of  $\Delta$  strain DYY3 was constructed by the next-generation sequencing (NGS) and single-molecule real-time (SMRT) sequencing technologies, and multiple comparative genomics analyses were applied to identify this new strain.

## MATERIALS AND METHODS

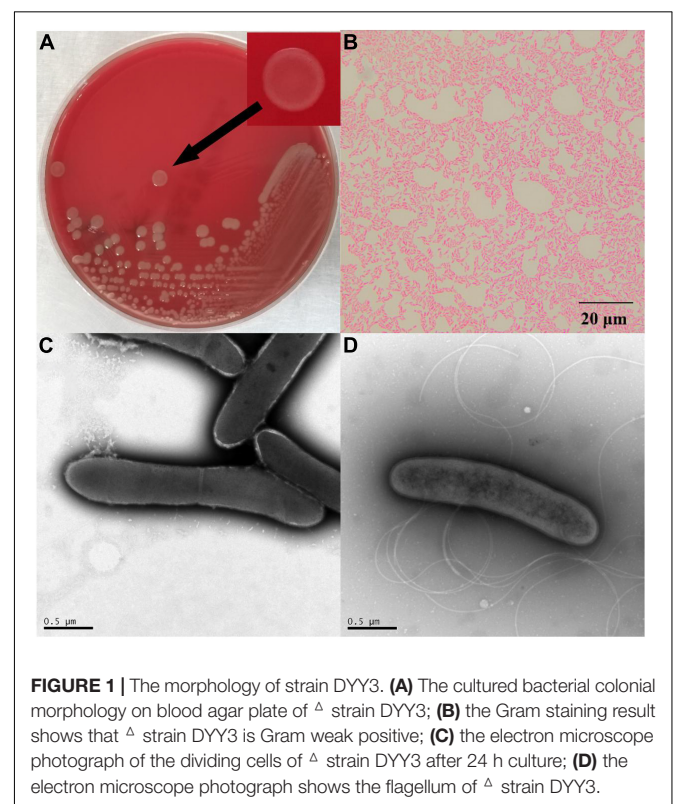
### Case Report

In January 2021, a 68-year-old female patient was admitted to the Shanghai Tenth People's Hospital ICU due to acute respiratory failure, aspiration pneumonia, and cerebral infarction. Invasive

mechanical ventilation, femoral vein catheterization, and urinary catheterization were performed during the treatment, and pulmonary infection was verified by fever, cough, and chest CT scanning. The patient's infection was effectively controlled initially by the empirical use of ceftazidime. However, the patient's body temperature, leukocyte count, and C-reactive protein were raised again after 9 days of anti-infection treatment. Considering the possibility of CRBSI, the femoral vein catheter was removed immediately, blood samples and the terminal of the central venous catheter were collected for bacterial culture, and vancomycin hydrochloride was added empirically to strengthen the anti-infection treatment. Three days later, both blood culture and catheter culture suggested unrecognized Gram-positive bacterial infection, and the drug sensitivity test showed that vancomycin was sensitive. Thus, vancomycin continued to be used, and the patient's bloodstream infection was cured in 2 weeks. The strain designated DYY3 was isolated from the blood sample and preserved in a  $-80^{\circ}\text{C}$  refrigerator to identify the unknown pathogenic bacterium further.

### Strain Isolation

The  $\Delta$  strain DYY3 was isolated from a critically ill patient's blood sample with a catheter-associated bloodstream infection. Briefly, the blood specimen was inoculated in a blood culture bottle (BD BACTEC Plus aerobic/F Culture Vials, Becton, Dickinson and Company, United States) at  $35^{\circ}\text{C}$  until it showed a positive result. For the isolation of  $\Delta$  strain DYY3, blood agar plates (bioMérieux, Marcy l'Etoile, France) were used, and the plates were incubated



**FIGURE 1 |** The morphology of strain DYY3. (A) The cultured bacterial colonial morphology on blood agar plate of  $\Delta$  strain DYY3; (B) the Gram staining result shows that  $\Delta$  strain DYY3 is Gram weak positive; (C) the electron microscope photograph of the dividing cells of  $\Delta$  strain DYY3 after 24 h culture; (D) the electron microscope photograph shows the flagellum of  $\Delta$  strain DYY3.

**TABLE 1** | Characteristics of *Haemobacillus shengwangii* and close species.

Characteristics	<i>Haemobacillus shengwangii</i> DYY3	<i>Thermicanus aegyptius</i> DSMZ 12793 <sup>T</sup>	<i>Brockia lithotrophica</i> Kam1851 <sup>T</sup>	<i>Hydrogenibacillus schlegelii</i> DSM 2000 <sup>T</sup>
Cell shape	Rod	Rod	Rod	Rod
Gram reaction	Weak positive	Weak positive	Positive	Positive
Motility	+	+	+	+
Spores	+	-	+	+
Flagellum	+	+	+	+
Colony color	Grayish white	Geige	NA*	Cream
Temperature for growth				
Range	20–45°C	37–65°C	46–78°C	37–80°C
Optimum	30–37°C	55–60°C	60–65°C	70–75°C
pH for growth				
Range	6–8	5.5–7.7	5.5–8.5	4.2–7.5
Optimum	7	6.5–7	6.5	6–7
NaCl concentration for growth (% w/v)				
Range	0–2	NA	NA	3–5
Optimum	0	NA	NA	3
Oxygen requirement	Facultative anaerobic	Facultative anaerobic	Strictly aerobic	Strictly aerobic
DNA G + C content (mol%)	40.62	50.3	63	67–68
Major fatty acids	C15:0 iso, C15:0 anteiso, C17:0 iso, C16:0 iso	NA	C16:0, C16:iso, C18:0, C17:0	C16:0 iso
Polar lipids <sup>#</sup>	DPG, PG, PE	NA	NA	DPG, PG
Quinone	MK7	NA	NA	MK7

<sup>#</sup>DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PE, phosphatidylethanolamine.  
\*NA, not available.

in a CO<sub>2</sub> incubator for 48 h. Later, dozens of single colonies were picked up from the blood agar plates. VITEK-MS automatic microbiological analyzer (bioMérieux, Marcy l'Etoile, France) was used to identify the taxonomic classification of the  $\Delta$  strain DYY3 according to the standard operation process using the VITEK MS IVD KB V3.2 database as the reference. The total length of the 16S rRNA sequence was amplified by PCR using the primers of 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'), and the amplified fragments were sequenced using a 3730XL sequencer (Applied Biosystems, United States).

## Morphology, Physiology, and Chemotaxonomy Analysis

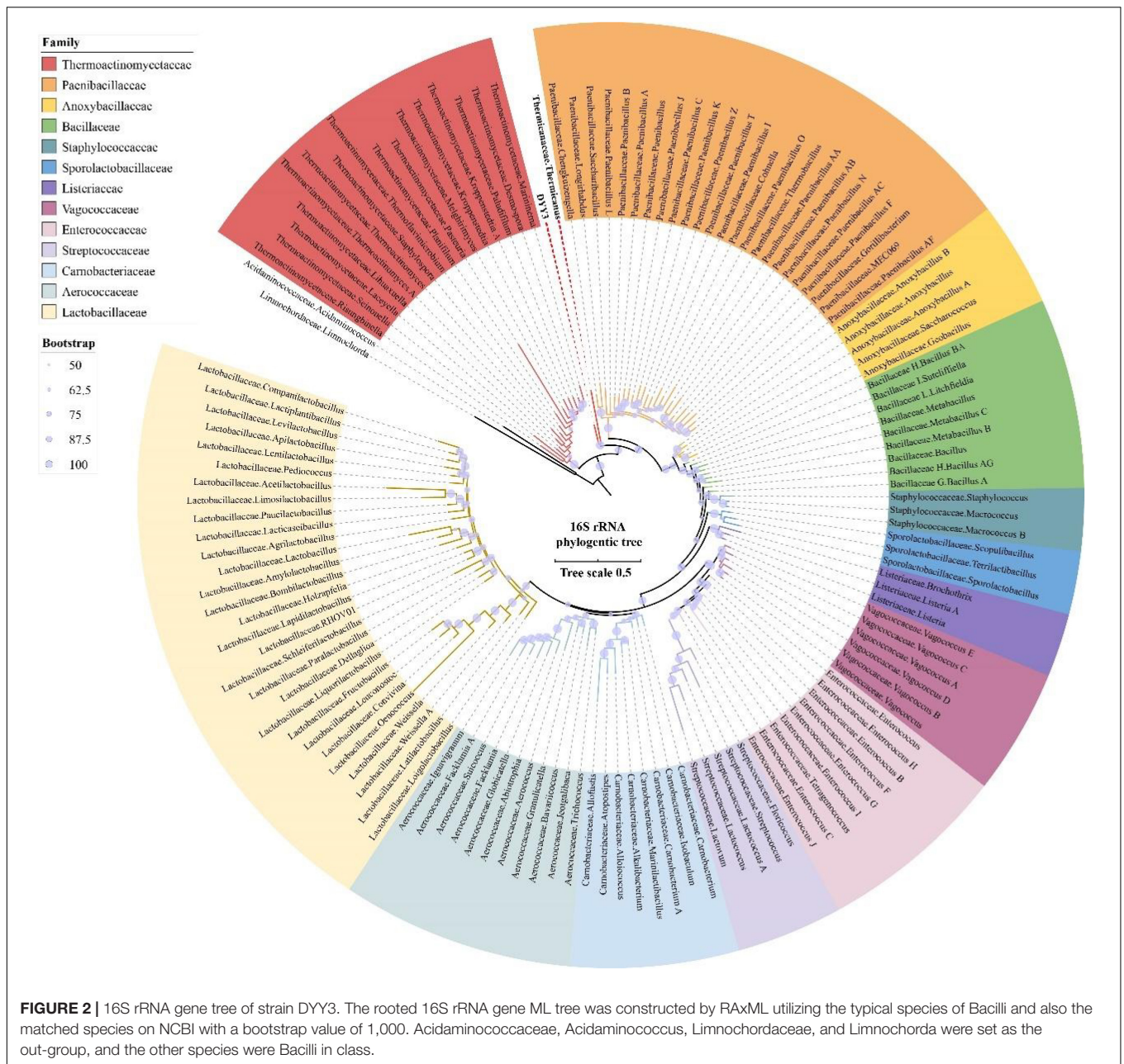
Gram staining of  $\Delta$  strain DYY3 was performed, referring to the procedures described by Wagner et al. (2018). The fresh biomass of DYY3 was stained with 1% (w/v) uranyl acetate, and the electron micrograph was taken by a transmission electron microscopy system JEM-1010 (JEOL, Japan). The growth tests were performed at various temperatures, NaCl concentrations, and pH levels using R2A agar plates (Difco, United States) as the culture medium. The tested temperatures were 4, 10, 15, 20, 30, 37, 40, 45, 50, and 55°C. The tested pH ranged from 5.0 to 11.0 with a gradient value of 1, K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub> buffer was used for pH 5–8, and NaHCO<sub>3</sub>/NaOH buffer was used for pH 9–11. The tested NaCl concentration ranged from 0 to 10% w/v using the interval of 1%. Acid production

tests, enzyme activity tests, and additional phenotypic tests were performed using API 50CHB, API ZYM, and API 20NE galleries (bioMérieux, Marcy l'Etoile, France), respectively. The utilization of carbon sources was tested using Biolog GPIII Microplates (Biolog, United States), and quinones were extracted and identified using the HPLC-20AT system (Shimadzu, Japan). The fresh biomass of  $\Delta$  strain DYY3 was hydrolyzed at 120°C for about 12 h to determine the composition of saccharides on its cell wall using ribose, arabinose, glucose, rhamnose, xylose, mannose, and galactose as references. The Cell Fatty Acid-Fatty Acid Methyl Ester (CFA-FAME) components were assayed by Agilent 6890 gas chromatograph (Agilent, United States), and the data were collected by the Sherlock Microbial Identification System (version 6.0, MIDI). The polar lipid analysis of  $\Delta$  strain DYY3 (1 g of freeze-dried cells) was performed and examined by thin-layer chromatography (TLC) on cellulose sheets. The spots for polar lipids were identified by spraying with 10% phosphomolybdic acid in ethanol, a-naphthol, and ninhydrin, respectively.

## Antibiotic Susceptibility Test

The minimum inhibitory concentrations (MICs) of the  $\Delta$  strain DYY3 to penicillin, ampicillin, vancomycin, gentamicin, erythromycin, ciprofloxacin, levofloxacin, clindamycin, trimethoprim/sulfisoxazole, rifampicin, and imipenem were determined by MicroScan Pos Combo Panel Type 33 (MicroScan, United States), with the interpretation of drug sensitivity results





**FIGURE 2 |** 16S rRNA gene tree of strain DYY3. The rooted 16S rRNA gene ML tree was constructed by RAxML utilizing the typical species of Bacilli and also the matched species on NCBI with a bootstrap value of 1,000. Acidaminococcaceae, Acidaminococcus, Limnochordaceae, and Limnochorda were set as the out-group, and the other species were Bacilli in class.

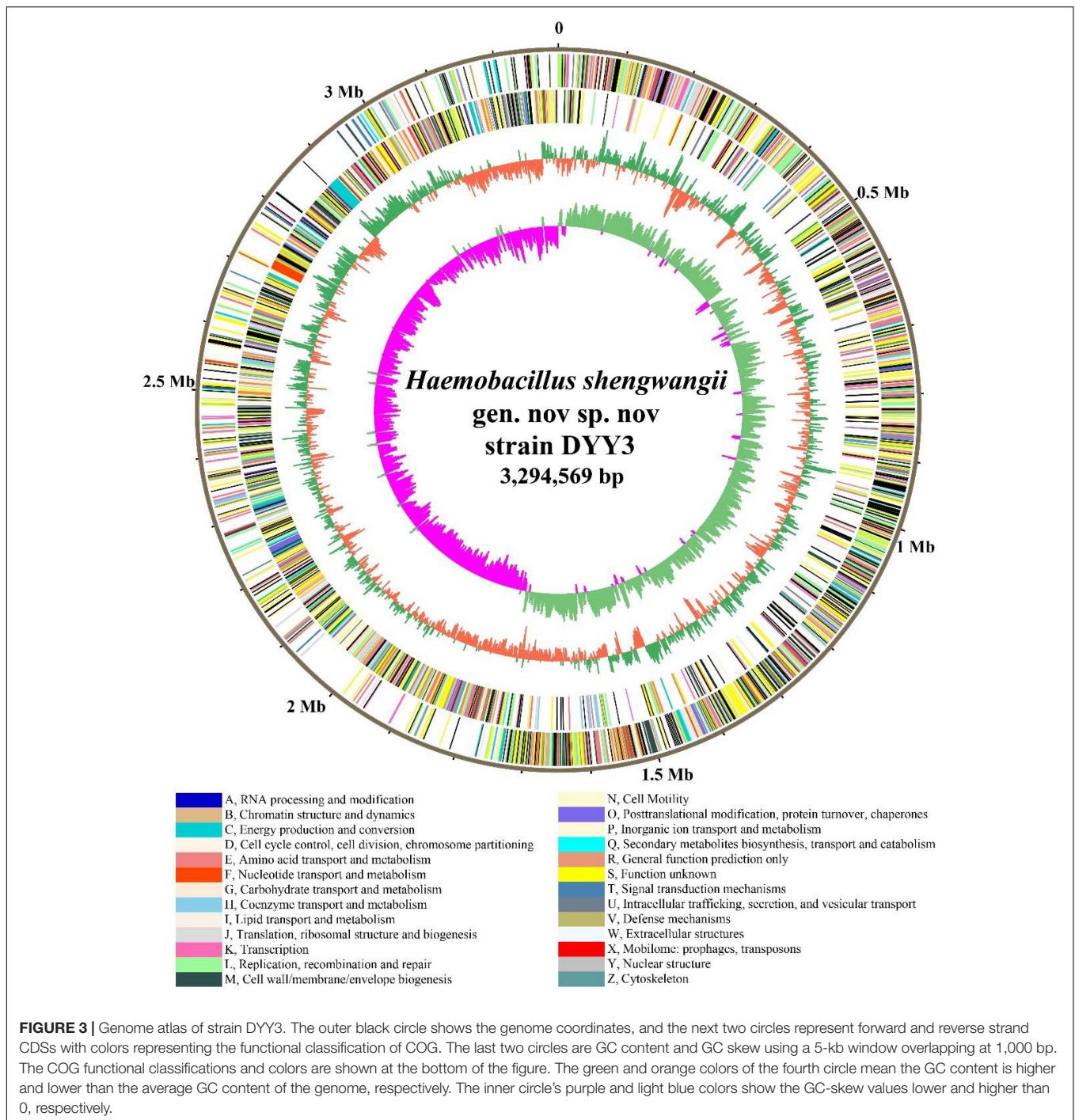
referred to the Clinical and Laboratory Standards Institute (CLSI) M45 guidelines.

### Genome Sequencing, Assembly, and Annotations

A 350-bp paired-end library was constructed and sequenced using the Illumina NovaSeq 6,000 sequencing platform (Illumina Inc., San Diego, CA, United States) with a PE150 layout. A 10-kb SMRT library was constructed and sequenced by the PacBio Sequel system (Pacific Biosciences, United States). The data were assembled by using Unicycler (version 0.4.7), and the genome was annotated by using Prokka (version 1.14.6) with default parameters. Prophages were annotated using phiSpy

(version 4.2.12), and genomic islands were identified by using Islanpath-DIMOB (version 1.0.6). The BUSCO database (version 5.2.2) was used to evaluate the completeness of the genome sequence (Manni et al., 2021). BLASTp (version 2.10.1) was used to query the non-redundant (nr) protein sequence database and hit with the highest score, and the identity higher than 60 was recognized as a match for each gene. Eggnog-mapper (version 2.0.1) with the parameter of “seed\_ortholog\_evalue 1e-5-m diamond” was used to query the eggNOG database. HMMER (version 3.3.2) with the parameter of “-E 1e-5” was used for Pfam (version 33.1) database annotation, and Diamond (version 0.9.24.125) with the parameter of “-e 1e-5” was used for the Swiss-Port annotation. The genome atlas was plotted





using CIRCOS.<sup>1</sup> The virulence factor genes were predicted by querying VFDB using a web-based VFAnalyzer. Antibiotic resistance genes were annotated by the CARD (version 3.1.4) database with BLASTp parameter of “-qcov\_hsp\_perc 80,” and hits with an identity less than 80 were filtered. The pathogen–host interaction associated genes were identified by the PHI database (version 4.12) with Diamond (version 0.9.24.125)

parameter of “-e 1e-5” and hits with an identity less than 50 were filtered.

### Phylogenetic Relationship Analysis

The 16S rRNA maximum likelihood (ML) phylogenetic tree was constructed by RAXML<sup>2</sup> using the GTR substitution matrix model. The whole genomic tree was constructed

<sup>1</sup><http://circos.ca/>

<sup>2</sup><https://raxml-ng.vital-it.ch/#/>

by IQ-TREE (version 1.6.12) with a bootstrap value of 1,000, referring to the core genes of the Genome Taxonomy Database (GTDB) (release202) (Rinke et al., 2021). One genome was selected as a representative for each genus of the principal families in class Bacilli. Two species from the family Acidaminococcaceae (phylum Firmicutes and class Negativicutes) and family Limnochordaceae (phylum Firmicutes and class Limnochordia) were set as out-group. The phylogenetic trees were plotted using iTOL.<sup>3</sup> The average nucleotide identity (ANI) values and the values of the pairwise core-genes average amino acid identity (cAAI) were calculated using fastANI (Jain et al., 2018) and CompareM (version 0.1.2)<sup>4</sup> with default parameters, respectively. Genes presented over 90% of genomes were used for cAAI calculation. The digital DNA-DNA hybridization (dddH) values were calculated using a genome-to-genome distance calculator (GGDC)<sup>5</sup>.

## RESULTS

### Morphology, Physiology, and Chemotaxonomy

After 48 h of cultivation at 30°C, the colonies of  $\Delta$  strain DYY3 on blood agar were 1–2 mm in size, grayish-white, round, smooth, and moist. The cells were weakly Gram-positive, rod-shaped, about 1.4–2.2  $\mu\text{m}$  in length, and 0.4–0.5  $\mu\text{m}$  in width, with flagellum and spore, motile, and facultative anaerobic (Figure 1). The  $\Delta$  strain DYY3 grew well on R2A blood agar plates with a temperature range of 20–45°C (preferred 30–37°C). The strain can grow in NaCl concentration ranging from 0 to 2% w/v but not in NaCl concentration over 2%. In addition, the strain can also grow at a pH of 6–8 with an optimum of 7.

API galleries tests revealed inactive biochemical reactions of  $\Delta$  strain DYY3 (Supplementary Table 1). The API ZYM assay indicated that alkaline phosphatase, esterase (C4), lipid esterase (C8), leucine aromatase, pancreatic rennet, acid phosphatase, and naphthol-AS-BI-phosphate hydrolase tests were positive. The API 20NE assay suggested a positive assimilation test for glucose, arabinose, mannitol, mannose, N-acetyl glucosamine, maltose, gluconate, capric acid, adipic acid, malic acid, citric acid, and phenylacetic acid. However, the  $\Delta$  strain DYY3 was only positive with 5-keto-gluconate in the API 50CHB assay.

Diphosphatidylglycerol (DPG), phosphatidylethanolamine (PE), and phosphatidylglycerol (PG) were identified to be the main polar lipids of  $\Delta$  strain DYY3, and the two-dimensional TCL of the polar lipids photographs was shown in Supplementary Figure 1. The  $\Delta$  strain DYY3 had no typical saccharides in the whole-cell hydrolysate experiment (Supplementary Figure 2). MK7 was the main methylanthraquinone in  $\Delta$  strain DYY3. The fatty acid analysis revealed that DYY3 synthesized mainly iso- and anteiso-branched saturated fatty acids, mainly including C15:0 iso (61.75%), C15:0 anteiso (13.29%), C17:0 iso (3.56%), and C16:0

iso (3.31%), and a spot of unsaturated fatty acids C17:1 iso w10c (3.54%). According to the RTSBA6 (version 6.21) database, *Bacillus* was the closest genus, but the similarities were not high (26.70%; refer to Supplementary Tables 2A,B). Carbon source utilization assays showed that  $\Delta$  strain DYY3 only used D-serine and glucuronamide as carbon sources (Supplementary Table 3). The VITEK-MS typing results cannot assign  $\Delta$  strain DYY3 to any species in the database, but later genetic analysis (see below) suggested that DYY3 belonged to the family Thermicanaceae. Hence, the morphological, physiological, and biochemical characteristics of the four species in the family Thermicanaceae are summarized in Table 1.

### 16S rRNA Phylogenetic Tree

To identify the phylogenetic relationship of  $\Delta$  strain DYY3, we first amplified and sequenced its full-length region of 16S rRNA genes. Then, we queried the 16S rRNA sequence at NCBI using online BLASTn. The result showed that the taxon of  $\Delta$  strain DYY3 was close to *Bacillus* at the genus level, and belonged to *Bacillus* at class level, and the best matching degree was 94.49%, and all the hit taxon belonged to Bacilli in class. Later, we extracted the entire length of the 16S rRNA gene sequence to construct a phylogenetic tree, utilizing the typical Bacilli species and the matched species on NCBI (Figure 2). The sequence information, blast results, and sequence similarity are shown in Supplementary Table 4. The phylogenetic tree showed that  $\Delta$  strain DYY3 was close to *Thermicanus* at the genus level and might belong to Thermicanaceae at the family level. However, this new taxon had six copies of the 16S rRNA gene, which were

TABLE 2 | Genome features of strain DYY3.

Genome features	Values
Chromosome	1 (circular)
Genome size	3,294,569
Genome coverage (NGS)	577
Genome coverage (TGS)	256
G + C content (%)	40.62%
BUSCO	100%
rRNAs (5, 16, 23S)	18 (6, 6, 6)
tRNAs	72 (35 families)
CDS genes	3,347
Genomic Island	13

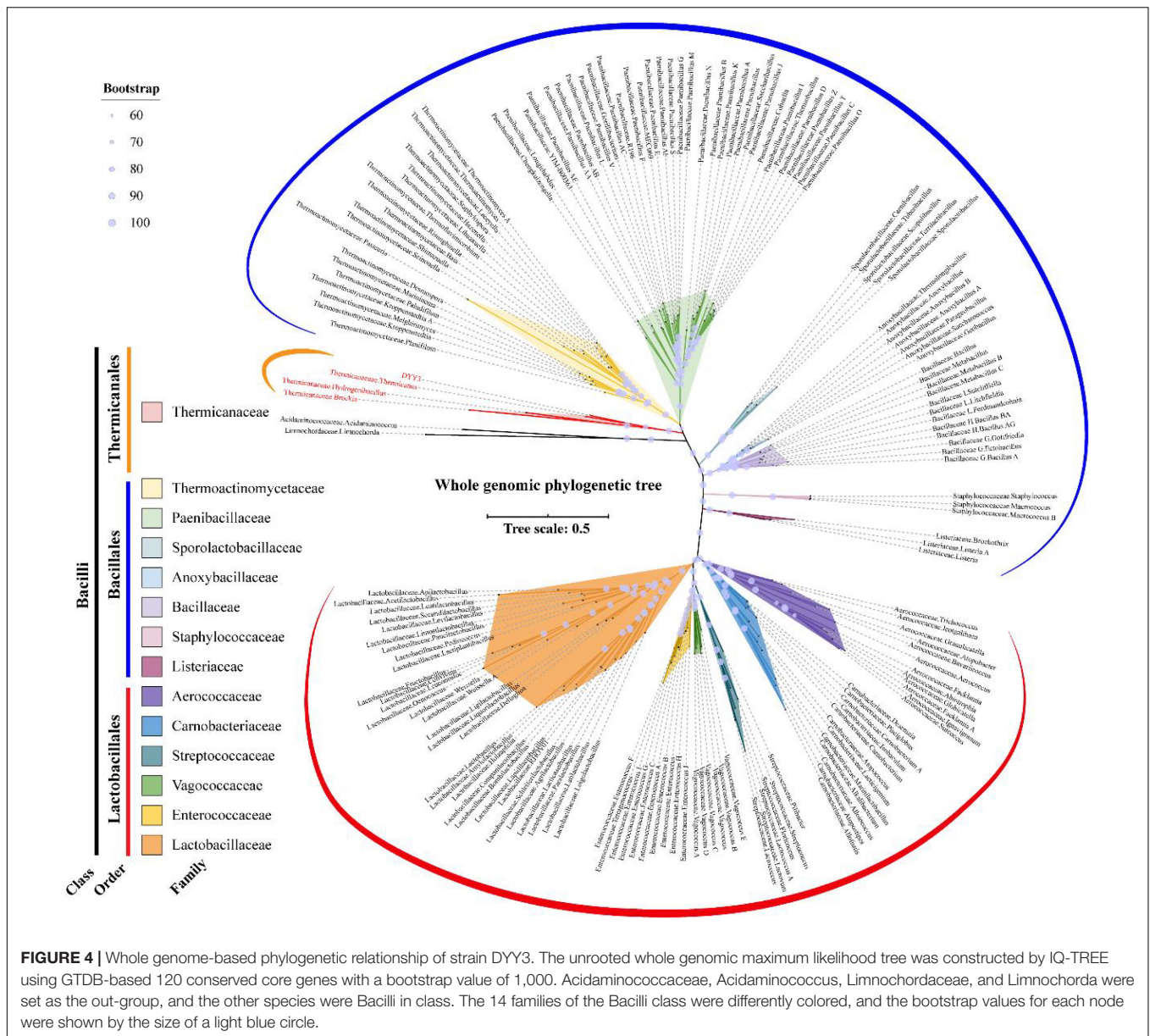
TABLE 3 | Gene functional annotation of strain DYY3.

Database	Annotated gene number	Percentage
Nr	1,834	54.80%
Pfam	2,685	80.22%
Swiss-Prot	2,034	60.77%
EggNOG	2,935	87.69%
GO	697	20.82%
KEGG	1,990	59.46%
COG	2,741	81.89%
Total	3,003	89.72%

<sup>3</sup><https://itol.embl.de/>

<sup>4</sup><http://github.com/dparks1134/CompareM>

<sup>5</sup><http://ggdc.dsmz.de/>



**FIGURE 4 |** Whole genome-based phylogenetic relationship of strain DYY3. The unrooted whole genomic maximum likelihood tree was constructed by IQ-TREE using GTDB-based 120 conserved core genes with a bootstrap value of 1,000. Acidaminococcaceae, Acidaminococcus, Limnochordaceae, and Limnochorda were set as the out-group, and the other species were Bacilli in class. The 14 families of the Bacilli class were differently colored, and the bootstrap values for each node were shown by the size of a light blue circle.

100% identical. Therefore, contradictory conclusions were drawn by the online NCBI blast results and sequence information of the 16S rRNA phylogenetic tree; thus,  $\Delta$  strain DYY3 cannot be assigned to any existing species or genus.

### Genome Features

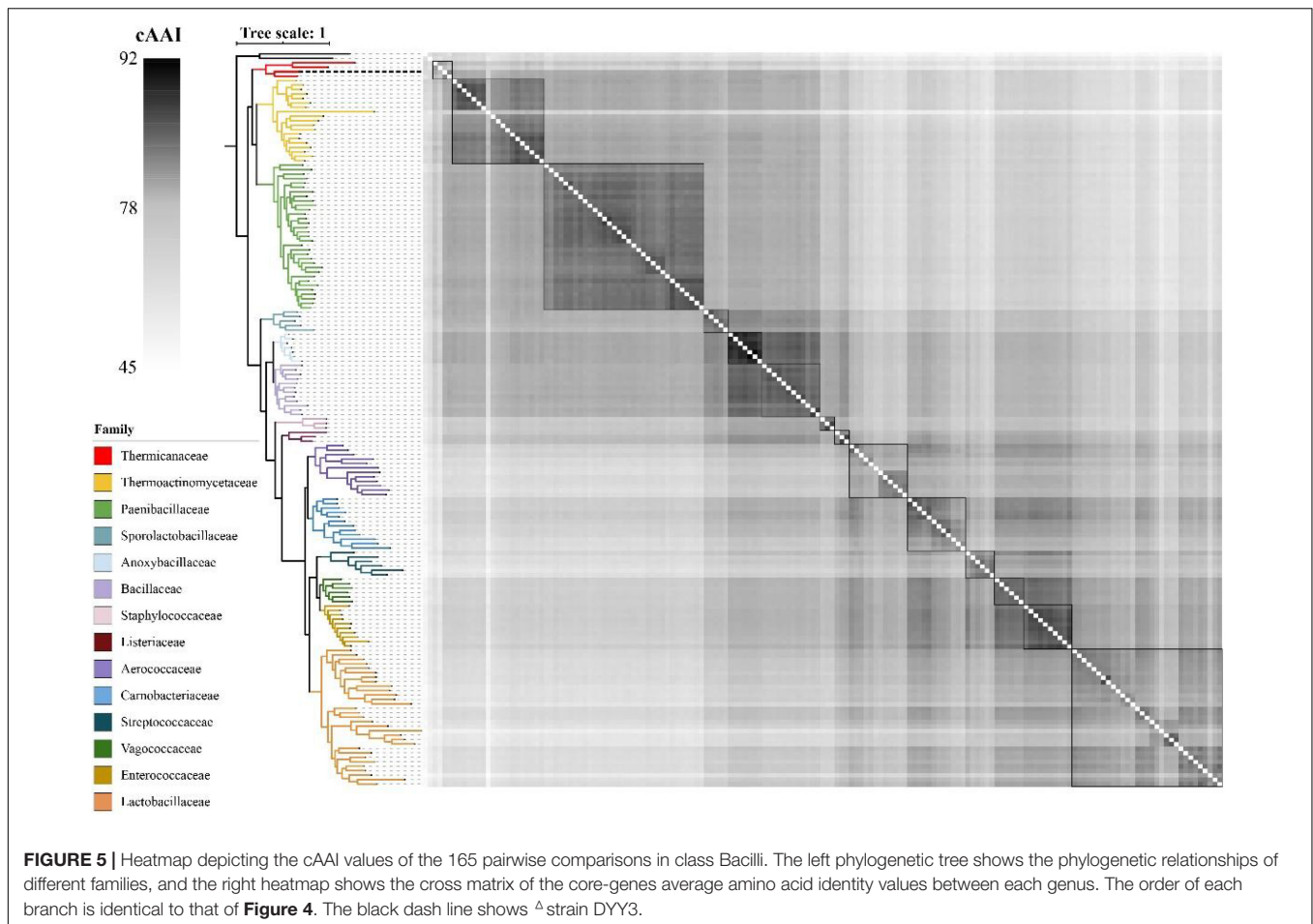
We combined NGS and SMRT technologies to construct the genome of  $\Delta$  strain DYY3 (Supplementary Tables 5A,B), and a circular genome with a total length of 3,294,569 bp and Guanine and Cytosine content of 40.62% was obtained (Figure 3 and Table 2). All conserved BUSCO genes (100%) were identified within the genome (Supplementary Table 6), indicating the high quality of this constructed genome. A total of 3,264 CDSs with an average length of 861 bp, 18 rRNAs, and 72 tRNAs were predicted. Besides, thirteen genomic islands

(Supplementary Table 7) and two partial prophage regions (Supplementary Table 8) were identified in the genome. A total of 3,003 (89.72%) coding sequence could be assigned with annotations, and 366 (10.94%) CDSs, including 22 database-based annotations, were annotated as hypothetical proteins (Table 3 and Supplementary Tables 9A–D).

### Phylogenetic Relationship

As shown in Figure 4, the whole genome-based phylogenetic tree was constructed, employing the protein sequence of the 120 conserved genes from the GTDB database, and 165 strains, including all the possible and available genomes of the Thermicanaceae family, were selected. The genomes' source, information, and dDDH are listed in Supplementary Table 10. This whole genome-based phylogenetic tree showed high





accordance with the 16S rRNA phylogenetic tree, demonstrating that DYY3 was close to the genus *Thermicanus* and belonged to the same taxon at the family level. Moreover, the sequences of majority groups were significantly distinct from calculating reliable values as the collinear regions were mainly less than 1% of the whole genomes. We then calculated cAAI to show the divergence between each genus, considering that the amino acid sequence is more conserved than the nucleotide sequence of gene coding regions. The plotted heatmap of the cAAI values agreed well with the core genes-based phylogenetic tree (**Figure 5** and **Supplementary Table 11**), suggesting that  $\Delta$  strain DYY3 belonged to Thermicanaceae at the family level. The average cAAI value between DYY3 and species from other families was 55.68, and the average cAAI value of each genus within the same existing family was 70.22 (**Table 4**). The cAAI value between *T. aegyptius* and DYY3 was 67.21, close to the average cAAI value for a single family, further demonstrating that  $\Delta$  strain DYY3 belonged to Thermicanaceae at the family level.

Finally, we compared the whole genomic similarity within the possible and available genomes of the Thermicanaceae family using BLAST-based ANI (ANIb) (**Table 5**). The ANIb between DYY3 and *T. aegyptius* was calculated to be 69.51%, indicating that they did not belong to the same species, as generally 95% ANI was found to recapitulate the majority

species (Jain et al., 2018; Olm et al., 2020; Parks et al., 2020). More importantly, *T. aegyptius*, *B. lithotrophica*, and *H. schlegelii* represented three different genera, and the ANIbs within the four species (e.g., DYY3) ranged from 68.86 to 75.40%, suggesting that  $\Delta$  strain DYY3 did not belong to any existing genus. Therefore,  $\Delta$  strain DYY3, the novel taxon, was named *Haemobacillus shengwangii* gen. nov., sp. nov. *Haemobacillus* referred to the strain isolated from the blood and was spore-forming and rod-shaped, and *Shengwangii* was named for appreciating the outstanding effort of Doctor Wang to save patients' lives.

## Virulence Factors

A total of 30 genes of 16 virulence factor classes were predicted to encode putative virulence factors in  $\Delta$  strain DYY3's genome by the virulence factor database (VFDB). DYY3 owned eight unique predicted virulence factor genes associated with adherence, anti-phagocytosis, immune evasion, intracellular survival, iron uptake, and motility (**Supplementary Table 12**). In addition, twenty genes were predicted associated with hypervirulence in DYY3's genome using the Pathogen Host Interactions database (PHI-base), with the highest number of the four species in the Thermicanaceae family (**Supplementary Tables 13A,B**).

**TABLE 4** | Statistical result of cAAI values between different families in Bacilli class.

Family name	Average cAAI values between DYY3 and each family	Average cAAI values between each genus in a single family	Genomes in each family
Thermoactinomycetaceae	59.59	68.25	19
Paenibacillaceae	60.47	70.38	33
Sporolactobacillaceae	58.73	70.23	5
Anoxybacillaceae	60.67	83.38	7
Bacillaceae	59.76	72.48	4
Staphylococcaceae	54.64	73.17	3
Listeriaceae	56.11	70.51	3
Aerococcaceae	51.74	61.95	12
Carnobacteriaceae	53.04	66.4	12
Streptococcaceae	50.91	66.75	6
Vagococcaceae	53.64	72.11	6
Enterococcaceae	53.65	75.34	10
Lactobacillaceae	50.91	61.86	31
<b>Average</b>	<b>55.68</b>	<b>70.22</b>	<b>12</b>
<b>Min</b>	<b>50.91</b>	<b>61.86</b>	<b>3</b>
<b>Max</b>	<b>60.67</b>	<b>83.38</b>	<b>33</b>
<b><i>T. aegyptius</i> vs. strain DYY3</b>		<b>67.21</b>	

## Antibiotic Resistance

Two genes, shared by  $\Delta$  strain DYY3 and *T. aegyptius* DSMZ 12793T, were predicted to be associated with antibiotic resistance (**Supplementary Table 14**). Of the two genes, one gene was related to ARO:3003438, annotated as AMR Gene Family associated with elfamycin. Another gene was related to ARO:3002838, annotated as LNU lincosamide nucleotidyltransferase associated with lincosamide. Besides, the MIC test revealed that  $\Delta$  strain DYY3 was sensitive to the most commonly used antibiotics, including penicillin, ampicillin, vancomycin, gentamicin, erythromycin, ciprofloxacin, levofloxacin, clindamycin, rifampicin, imipenem, and trimethoprim/sulfamethoxazole (**Supplementary Table 15**).

## DISCUSSION

Traditional morphological, physiological, and biochemical studies and comparative genomic analysis demonstrated that the  $\Delta$  strain DYY3 was a novel bacterial pathogen belonging to class Bacilli, order Thermicanales, family Thermicanaceae, and genus *Haemobacillus*. We thus proposed *H. shengwangii* gen.

nov., sp. nov. to be the name of the novel taxon. In addition, our data supported that *T. aegyptius*, *H. schlegelii*, *B. lithotrophica*, and *H. shengwangii* should belong to Thermicanaceae at a family level. Among the genome published in the family Thermicanaceae, *H. shengwangii* was the first genome constructed with high quality, which undoubtedly contributed to the future research of this family.

Based on 16S rRNA gene sequencing, querying, and phylogenetic tree analysis,  $\Delta$  strain DYY3 was identified as a species belonging to class Bacilli, and *T. aegyptius* was the most adjacent taxon. However, when we investigated the taxonomy background of the genus *Thermicanus*, the taxonomic status of this genus was inconsistent at the family level in different authoritative taxonomic databases, including the NCBI taxonomy database, GTDB, SILVA, and GBIF database. We then constructed the core genes-based phylogenetic tree using 16S representative species, including the typical genus of the representative families in class Bacilli and the genus associated with the family Thermicanaceae. The phylogenetic tree indicated that *T. aegyptius*, *H. schlegelii*, *B. lithotrophica*, and  $\Delta$  strain DYY3 were in the same tree clade, and the cAAI heatmap further verified their close genetic relationships. Hence, these four species should be classified into the same family. *T. aegyptius* was the closest species to  $\Delta$  strain DYY3, supported by the 16S rRNA gene tree and the whole genomic phylogenetic tree. The calculated ANIb (69.51%) between  $\Delta$  strain DYY3 and *T. aegyptius* str. DSM 12793 was far from the standard (95%) belonging to the same species, which indicated that  $\Delta$  strain DYY3 was a novel species. Since three of the four species in the family Thermicanaceae were known as different genera, the narrow range of the ANIb values (68.86–75.40%) suggested that DYY3 also did not belong to any existing genus. Therefore,  $\Delta$  strain DYY3 was verified belonging to a genus novel and species novel taxon in Thermicanaceae. This new taxon was proposed to be named *Haemobacillus shengwangii* gen. nov., sp. nov.

The MIC test showed that  $\Delta$  strain DYY3 was sensitive to all commonly used antibiotics. Ceftazidime was initially used in the patient's antibiotic treatment, which turned ineffective. Subsequently, vancomycin was empirically added to the anti-infection therapy, and the symptoms of infection were improved rapidly. Since *H. shengwangii* is a Gram-positive, penicillins, glycopeptides, and cephalosporins should be recommended. In addition, the MIC test also indicated that  $\Delta$  strain DYY3 did not have antibiotic resistance, although the CARD database predicted two antibiotic resistance genes associated with elfamycin and

**TABLE 5** | The ANIb values and alignment length of strain DYY3 and species in family Thermicanaceae.

Species	Alignment length			
	<i>H. shengwangii</i> DYY3	<i>T. aegyptius</i> DSM 12793	<i>B. lithotrophica</i> a DSM 22653	<i>H. schlegelii</i> MA48
<b>ANIb</b>				
<b>DYY3</b>		304,571	36,373	65,834
<b><i>T. aegyptius</i></b>	69.51%		77,647	176,981
<b><i>B. lithotrophica</i></b>	69.27%	68.86%		301,196
<b><i>H. schlegelii</i></b>	68.82%	69.75%	75.40%	

lincosamide, respectively. To further clarify the virulence of  $\Delta$  strain DYY3, its virulence factor genes were predicted based on sequence similarity comparison against the VFDB. The  $\Delta$  strain DYY3 owned eight unique predicted virulence factor genes associated with adherence, anti-phagocytosis, immune evasion, intracellular survival, iron uptake, and motility. Besides, twenty genes were predicted to be associated with hypervirulence in the  $\Delta$  strain DYY3 genome using PHI-base, with the highest number of the four species in the Thermicanaceae family. Given that *Bacillus* was the closest genus that can be queried in the database, the phylogenetic relationships of  $\Delta$  strain DYY3 and species in genus *Bacillus* were quite distant, and the sequence similarities of  $\Delta$  strain DYY3 and species in genus *Bacillus* were relatively low; we cannot exclude the possibility of  $\Delta$  strain DYY3 owning more genes associated with virulence factors.

Although blood culture is the golden standard for pathogenic diagnosis, it cannot fully support that  $\Delta$  strain DYY3 was the only pathogenic bacterium responsible for this patient's infection, as only a tiny proportion of pathogens are identifiable by the culture-based methods (Li et al., 2021). mNGS was more suitable for pathogenic detection when the pathogen was unculturable, novel, or variant species, such as DYY3, as all the nucleic acids can be sequenced and analyzed indiscriminately (Wilson et al., 2019; Price et al., 2021), and NGS has been widely used to perform comprehensive and precise diagnosis of pathogens with various sample types (Miller et al., 2019; Chen et al., 2021; Gu et al., 2021). Nonetheless, without a reference genome, even if the genome of a pathogen was sequenced, the species information cannot be disclosed by mNGS. Therefore, the availability of the *H. shengwangii* genome sequence could provide a valuable source for further comparative genomics analysis in the family Thermicanaceae and facilitate the family's detection rate when conducting an mNGS-based pathogenic detection or study.

According to the conventional taxonomic features, *T. aegyptius*, *H. schlegelii*, *B. lithotrophica*, and *H. shengwangii* shared the identical phenotypes of rod-shaped, Gram-positive, spore-forming, and motor ability, indicating that these were the standard features of family Thermicanaceae. However, a clear description of the Thermicanaceae needs further studies, as few reference species and literature on Thermicanaceae are currently available. In addition, we did not trace the source of  $\Delta$  strain DYY3, such as the patient's skin, living environment, and so on, and only one strain of this novel species was isolated so far. Finally, the pathogenic mechanism was not thoroughly investigated, although virulence factor genes and antibiotic resistance genes were predicted in this study.

## CONCLUSION

This study identified a novel pathogenic bacterium, *H. shengwangii* gen. nov., sp. nov., isolated from a critically ill patient with CRBSI. In addition to the traditional methods of species identification, we used multiple comparative genomics analyses to prove that  $\Delta$  strain DYY3 was genus novel and species novel in the family of Thermicanaceae. Moreover, the constructed high-quality *H. shengwangii* genome will contribute

to further genomics research and NGS-based pathogenic detection or study in the family Thermicanaceae.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

SW and YY conceived and designed the project. YD, YL, DS, and SM participated in the strains collection. YD, YL, and SZ performed the morphology physiology and chemotaxonomy identification test. XL, SF, and HX conducted the bioinformatic analysis. YD and XL wrote the manuscript. SW and ZL revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the Clinical Research Project of SHDC (No. SHDC2020CR6030-003) and the Three-Year Action Plan for Constructing the Shanghai Public Health System (GWV-3.1).

## ACKNOWLEDGMENTS

We thank Yu Kong and Xue Xu (Electron Microscopy Facilities of Center for Excellence in Brain Science and Technology, Chinese Academy of Sciences) for their assistance with EM sample preparation and image analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.919169/full#supplementary-material>

**Supplementary Figure 1** | Two-dimensional TLC of polar lipids photographs from strain DYY3. Figure (A–E) shows the dyeing results using ninhydrin, molybdenum blue 1-methylnaphthol, d reagent, and molybdophosphate. DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PE, phosphatidylethanolamine; PL1–5, unidentified phospholipids; GL, unidentified glycolipid; L1–2, unidentified lipids; AGPL, unidentified amino-glycerophospholipid.

**Supplementary Figure 2** | Whole-cell saccharide hydrolysate experiment. The results indicated that  $\Delta$  strain DYY3 did not have any characteristic sugars.



## REFERENCES

- Barbosa, R. G., van Veelen, H. P. J., Pinheiro, V., Sleutels, T., Verstraete, W., and Boon, N. (2020). Enrichment of hydrogen oxidizing bacteria from high temperature and salinity environments. *Appl. Environ. Microbiol.* 87:e02439-20. doi: 10.1128/AEM.02439-20
- Chen, Y., Feng, W., Ye, K., Guo, L., Xia, H., Guan, Y., et al. (2021). Application of metagenomic next-generation sequencing in the diagnosis of pulmonary infectious pathogens from bronchoalveolar lavage samples. *Front. Cell Infect. Microbiol.* 11:541092. doi: 10.3389/fcimb.2021.541092
- Gerwer, S. M., Mihalkova, M., Bion, J. F., Wilson, A. P. R., Chudasama, D., Johnson, A. P., et al. (2020). Surveillance of bloodstream infections in intensive care units in England, May 2016–April 2017: epidemiology and ecology. *J. Hosp. Infect.* 106, 1–9. doi: 10.1016/j.jhin.2020.05.010
- Gossner, A. S., Devereux, R., Ohnemuller, N., Acker, G., Stackebrandt, E., and Drake, H. L. (1999). *Thermicanus aegyptius* gen. nov., sp. nov., isolated from oxic soil, a fermentative microaerophile that grows commensally with the thermophilic acetogen *Moorella thermoacetica*. *Appl. Environ. Microbiol.* 65, 5124–5133. doi: 10.1128/AEM.65.11.5124-5133.1999
- Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z
- Hayashi Sant'Anna, F., Bach, E., Porto, R. Z., Guella, F., Hayashi Sant'Anna, E., and Passaglia, L. M. P. (2019). Genomic metrics made easy: what to do and where to go in the new era of bacterial taxonomy. *Crit. Rev. Microbiol.* 45, 182–200. doi: 10.1080/1040841X.2019.1569587
- Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114.
- Kampfer, P., Glaeser, S. P., and Busse, H. J. (2013). Transfer of *Bacillus schlegelii* to a novel genus and proposal of *Hydrogenibacillus schlegelii* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 63, 1723–1727. doi: 10.1099/ijs.0.045146-0
- Li, N., Cai, Q., Miao, Q., Song, Z., Fang, Y., and Hu, B. (2021). High-Throughput metagenomics for identification of pathogens in the clinical settings. *Small Methods* 5:2000792. doi: 10.1002/smt.202000792
- Manni, M., Berkeley, M. R., Seppely, M., Simao, F. A., and Zdobnov, E. M. (2021). BUSCO Update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* 38, 4647–4654. doi: 10.1093/molbev/msab199
- Mermel, L. A., Allon, M., Bouza, E., Craven, D. E., Flynn, P., O'Grady, N. P., et al. (2009). Clinical practice guidelines for the diagnosis and management of intravascular catheter-related infection: 2009 Update by the Infectious Diseases Society of America. *Clin. Infect. Dis.* 49, 1–45. doi: 10.1086/599376
- Miller, S., Naccache, S. N., Samayoa, E., Messacar, K., Arevalo, S., Federman, S., et al. (2019). Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* 29, 831–842. doi: 10.1101/gr.238170.118
- Murray, P. R., and Masur, H. (2012). Current approaches to the diagnosis of bacterial and fungal bloodstream infections in the intensive care unit. *Crit. Care Med.* 40, 3277–3282. doi: 10.1097/CCM.0b013e318270e771
- Olm, M. R., Crits-Christoph, A., Diamond, S., Lavy, A., Matheus Carnevali, P. B., and Banfield, J. F. (2020). Consistent metagenome-derived metrics verify and delineate bacterial species boundaries. *mSystems* [Preprint]. doi: 10.1128/mSystems.00731-19
- Parks, D. H., Chuvochina, M., Chaumeil, P. A., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* 38, 1079–1086.
- Perevalova, A. A., Kublanov, I. V., Baslerov, R. V., Zhang, G., and Bonch-Osmolovskaya, E. A. (2013). *Brockia lithotrophica* gen. nov., sp. nov., an anaerobic thermophilic bacterium from a terrestrial hot spring. *Int. J. Syst. Evol. Microbiol.* 63, 479–483. doi: 10.1099/ijs.0.041285-0
- Price, T. K., Realegeno, S., Mirasol, R., Tsan, A., Chandrasekaran, S., Garner, O. B., et al. (2021). Validation, Implementation, and Clinical Utility of Whole Genome Sequence-Based Bacterial Identification in the Clinical Microbiology Laboratory. *J. Mol. Diagn.* 23, 1468–1477. doi: 10.1016/j.jmoldx.2021.07.020
- Rinke, C., Chuvochina, M., Mussig, A. J., Chaumeil, P. A., Davin, A. A., Waite, D. W., et al. (2021). A standardized archaeal taxonomy for the genome taxonomy database. *Nat. Microbiol.* 6, 946–959. doi: 10.1038/s41564-021-00918-8
- Rosenthal, V. D., Duszynska, W., Ider, B. E., Gurskis, V., Al-Ruzzieh, M. A., Myatra, S. N., et al. (2021). International Nosocomial Infection Control Consortium (INICC) report, data summary of 45 countries for 2013–2018, Adult and Pediatric Units, Device-associated Module. *Am. J. Infect. Control* 49, 1267–1274. doi: 10.1016/j.ajic.2021.04.077
- Schenk, A., and Aragno, M. (1979). *Bacillus schlegelii*, a New Species of Thermophilic, Facultatively Chemolithoautotrophic Bacterium Oxidizing Molecular Hydrogen. *J. Gen. Microbiol.* 115, 333–341.
- Schwab, F., Geffers, C., Behnke, M., and Gastmeier, P. (2018). ICU mortality following ICU-acquired primary bloodstream infections according to the type of pathogen: A prospective cohort study in 937 Germany ICUs (2006–2015). *PLoS One* 13:e0194210. doi: 10.1371/journal.pone.0194210
- Tabak, Y. P., Vankeepuram, L., Ye, G., Jeffers, K., Gupta, V., and Murray, P. R. (2018). Blood Culture Turnaround Time in U.S. Acute Care Hospitals and Implications for Laboratory Process Optimization. *J. Clin. Microbiol.* 56:e00500-18. doi: 10.1128/JCM.00500-18
- Wagner, K., Springer, B., Pires, V. P., and Keller, P. M. (2018). Pathogen identification by multiplex lightmix real-time PCR assay in patients with meningitis and culture-negative cerebrospinal fluid specimens. *J. Clin. Microbiol.* 56:e01492-17. doi: 10.1128/JCM.01492-17
- Wilson, M. R., Sample, H. A., Zorn, K. C., Arevalo, S., Yu, G., Neuhaus, J., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* 380, 2327–2340. doi: 10.1056/NEJMoa1803396
- Zeng, C., Wu, A., Li, L., and Jia, H. (2021). Multi-center prospective study on central line-associated bloodstream infections in 79 ICUs of China. *BMC Infect. Dis.* 21:1208. doi: 10.1186/s12879-021-06871-5
- Zhong, Y., Zhou, L., Liu, X., Deng, L., Wu, R., Xia, Z., et al. (2021). Incidence, Risk factors, and attributable mortality of catheter-related bloodstream infections in the intensive care unit after suspected catheters infection: a retrospective 10-year cohort study. *Infect. Dis. Ther.* 10, 985–999. doi: 10.1007/s40121-021-00429-3

**Conflict of Interest:** XL, SF, HX, and ZL were employed by Hugo Biotechnologies Co., Ltd.

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

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

Copyright © 2022 Du, Li, Liu, Mu, Shen, Fan, Lou, Zhang, Xia, Yuan and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.