

## OPEN ACCESS

EDITED BY  
Ben Pascoe,  
University of Oxford, United Kingdom

REVIEWED BY  
Marcos Godoy,  
San Sebastián University, Chile  
Nattinee Kittiwat,  
Department of Livestock  
Development, Thailand

\*CORRESPONDENCE  
Yin Wang  
✉ 10334@sicau.edu.cn

†These authors share first authorship

RECEIVED 28 December 2023  
ACCEPTED 05 February 2024  
PUBLISHED 21 February 2024

CITATION  
Zhou Y, Tu T, Yao X, Luo Y, Yang Z, Ren M,  
Zhang G, Yu Y, Lu A and Wang Y (2024)  
Pan-genome analysis of *Streptococcus suis*  
serotype 2 highlights genes associated with  
virulence and antibiotic resistance.  
*Front. Microbiol.* 15:1362316.  
doi: 10.3389/fmicb.2024.1362316

COPYRIGHT  
© 2024 Zhou, Tu, Yao, Luo, Yang, Ren, Zhang,  
Yu, Lu and Wang. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC  
BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Pan-genome analysis of *Streptococcus suis* serotype 2 highlights genes associated with virulence and antibiotic resistance

You Zhou<sup>1†</sup>, Teng Tu<sup>1†</sup>, Xueping Yao<sup>1</sup>, Yan Luo<sup>1</sup>, Zexiao Yang<sup>1</sup>,  
Meishen Ren<sup>1,2,3,4</sup>, Ge Zhang<sup>2,3,4</sup>, Yuanyuan Yu<sup>2,3,4</sup>, Aiping Lu<sup>2,3,4</sup>  
and Yin Wang<sup>1\*</sup>

<sup>1</sup>Key Laboratory of Animal Diseases and Human Health of Sichuan Province, College of Veterinary Medicine, Sichuan Agricultural University, Chengdu, China, <sup>2</sup>Law Sau Fai Institute for Advancing Translational Medicine in Bone and Joint Diseases (TMBJ), School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China, <sup>3</sup>Guangdong-Hong Kong-Macao Greater Bay Area International Research Platform for Aptamer-based Translational Medicine and Drug Discovery (HKAP), Kowloon Tong, Hong Kong SAR, China, <sup>4</sup>Institute of Integrated Bioinformedicine and Translational Science (IBTS), School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

*Streptococcus suis* serotype 2 (SS2) is a Gram-positive bacterium. It is a common and significant pathogen in pigs and a common cause of zoonotic meningitis in humans. It can lead to sepsis, endocarditis, arthritis, and pneumonia. If not diagnosed and treated promptly, it has a high mortality rate. The pan-genome of SS2 is open, and with an increasing number of genes, the core genome and accessory genome may exhibit more pronounced differences. Due to the diversity of SS2, the genes related to its virulence and resistance are still unclear. In this study, a strain of SS2 was isolated from a pig farm in Sichuan Province, China, and subjected to whole-genome sequencing and characterization. Subsequently, we conducted a Pan-Genome-Wide Association Study (Pan-GWAS) on 230 strains of SS2. Our analysis indicates that the core genome is composed of 1,458 genes related to the basic life processes of the bacterium. The accessory genome, consisting of 4,337 genes, is highly variable and a major contributor to the genetic diversity of SS2. Furthermore, we identified important virulence and resistance genes in SS2 through pan-GWAS. The virulence genes of SS2 are mainly associated with bacterial adhesion. In addition, resistance genes in the core genome may confer natural resistance of SS2 to fluoroquinolone and glycopeptide antibiotics. This study lays the foundation for further research on the virulence and resistance of SS2, providing potential new drug and vaccine targets against SS2.

## KEYWORDS

SS2, pan-genome, GWAS, the core genome, the accessory genome, virulence, antibiotic resistance

## 1 Introduction

*Streptococcus suis* is a gram-positive coccus with a capsule. In previous studies, *Streptococcus suis* is divided into 35 serotypes (type 1–34, type 1/2), according to the different capsular polysaccharide (CPS) antigens (Fittipaldi et al., 2012). However, recent studies indicated that certain serotypes (type 20, 22, 26, and 32–34) are not associated with

*Streptococcus suis* (Hill et al., 2005; Nishibori et al., 2013). Among them, *Streptococcus suis* serotype 2 (SS2) is the most common and virulent. SS2 primarily infects through wounds, causing acute septicemia, meningitis, arthritis, endocarditis, pneumonia, and other diseases in pigs (Zhu et al., 2006). It can also infect humans, leading to bacterial meningitis or toxic shock-like syndrome (Tang et al., 2006; Jiang et al., 2020).

With the advancement of sequencing technologies, an increasing number of species' genome sequences are being uploaded (Pareek et al., 2011). Particularly, the hybrid assembly of third-generation sequencing data and second-generation sequencing data is currently the mainstream approach for completing bacterial genome drafts (Loman and Pallen, 2015). Whole-genome sequences contribute to deepening our understanding of individual organisms (Orsini et al., 2016). Unfortunately, a single genome cannot reflect how genetic variations drive the pathogenic mechanisms within bacterial species (Tettelin et al., 2005; Maturana and Cárdenas, 2021). With the introduction of the pan-genome concept, the genetic variation trends and scope within bacterial populations can be described (Medini et al., 2005). The pan-genome consists of the core genome present in all strains and the accessory genome present in one or more strains (Vernikos, 2020). Moreover, with the increase in the number of genomes, core genomes and accessory genomes may exhibit differences in functionality, metabolic pathways, and resistance (Rasko et al., 2008; Mira et al., 2010). Currently, there have been reports of an open pan-genome composed of 19 SS2 strains, showing notable differences in pili and prophage regions (Guo et al., 2021). However, due to the limited number of samples, the specific differences in the core genome and accessory genome remain unknown.

The development of SS2 disease typically begins with the colonization of bacteria in the upper respiratory tract (Xia et al., 2019). During this process, various virulence factors of SS2 contribute to initial adhesion, immune evasion, and host invasion (Ji et al., 2016; Li et al., 2017). For example, CPS is the most representative virulence factor of *Streptococcus suis* (Zheng et al., 2013). It not only forms the basis for serotyping but also protects the bacteria from phagocytosis (Xia et al., 2019). Although several virulence factors have been proven to play a role in the early stages of infection, there are still quite a few virulence factors that remain unclear.

Currently, the most effective approach for treating *Streptococcus suis* involves antibiotics, as there are no commercially available vaccines (Palmieri et al., 2011). However, with the overuse of antibiotics in farming, the increasing prevalence of multidrug-resistant strains is limiting the options for effective antibiotics. Studies have reported a slow increase in resistance of *Streptococcus suis* to tetracyclines and macrolide/lincosamide antibiotics, which are widely used in the global animal sector (Dechêne-Tempier et al., 2023). Hence, it is necessary to explore relevant resistance genes to investigate the types of resistance within the SS2 population.

In this study, a strain of *Streptococcus suis* serotype 2 was isolated from a pig farm in Sichuan Province, China, and the complete genome sequence of the strain was obtained. Subsequently, we reported a whole-genome association study on the genomic sequences of 230 naturally isolated strains of SS2.

This research unveiled significant distinctions between the core genome and accessory genome of SS2, such as in the transport and metabolism of essential substances, resistance to external conditions, and genetic reproduction. Additionally, our findings revealed virulence genes associated with the SS2 infection process and suggested its potential intrinsic antibiotic resistance.

## 2 Materials and methods

### 2.1 Diseased pig and bacteria isolation

The samples of diseased pig were collected from a pig farm located in Sichuan, China. All diseased pigs exhibited symptoms of anorexia, tremors, and joint swelling. The nasal secretions, joint fluid, and blood from diseased pigs were individually streaked onto trypticase soy agar (TSA) containing 5% fetal bovine serum and cultured at a constant temperature of 37°C for 18 h. Subsequently, grayish-white, semi-transparent, and smoothly surfaced circular colonies were selected for Gram staining and identification of isolated strains through 16S rRNA gene sequencing (Lane, 1991).

### 2.2 DNA extraction

In this study, the Bacterial DNA Kit (OMEGA, Norcross, USA) was used to extract bacterial DNA. Due to the direct correlation between Nanopore sequencing quality and input DNA quality, unnecessary centrifugation and shaking should be avoided during the DNA extraction process to minimize DNA fragmentation. The DNA solutions were quantified using a Nanodrop 2000 and a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, CA, USA).

### 2.3 Library preparation and sequencing

The Ligation Sequencing Kit (SQK-LSK109) was used to construct Nanopore sequencing libraries according to the manufacturer's instructions, and MinION (Oxford Nanopore, Cambridge, UK) was used for sequencing. Subsequently, Illumina libraries were constructed using a Nextera XT Kit (Illumina, San Diego, CA, USA) followed by 150 bp paired-end sequencing on either the NextSeq 550 platform (Illumina, San Diego, CA, USA).

### 2.4 Base-calling and data processing

Firstly, guppy (V4.0.11, <https://community.nanoporetech.com>, accessed on 5 June 2023) was used for base-calling with a high accuracy model. Secondly, NanoFilt (V2.8.0, <https://anaconda.org/bioconda/nanofilt>, accessed on 5 June 2023) was used to filter the raw data, with a filtering threshold set at a Q-value >10 and a minimum read length of 1,000 bp, aiming to select a dataset with higher quality. Finally, the Illumina sequencing data were subjected to quality control via fastp (V0.23.3, <https://github.com/OpenGene/fastp>, accessed on 5 June 2023), followed by filtering to remove adapters and low-quality reads.

## 2.5 Genome assembly and integrity assessment

Firstly, Flye (V2.9.1, <https://github.com/fenderglass/Flye>, accessed on 10 September 2023) were employed for the initial assembly using filtered Nanopore sequencing data. Secondly, Pilon (V1.23, <https://github.com/broadinstitute/pilon>, accessed on 10 September 2023) was used for error correction of the bacterial genome via Illumina sequencing data. Thirdly, Bandage (V0.9.0, <https://github.com/rrwick/Bandage>, accessed on 10 September 2023) was used to check whether the contig formed a circular structure. Subsequently, Quast (V5.2.0, <https://github.com/ablab/quast>, accessed on 10 September 2023) was evaluated the corrected and uncorrected genome sequences. Finally, Busco (V5.4.7, <https://busco.ezlab.org/>, accessed on 12 September 2023) was used to assess the assembled bacterial genome.

## 2.6 Minimal inhibitory concentration (MIC)

It was carried out according to the standard microdilution method recommended by the Clinical and Laboratory Standards Institute (CLSI; Shryock, 2002; CLSI, 2018; Feßler et al., 2023). Seven types of commonly used SS2 drugs were selected, namely macrolides, lincomycins,  $\beta$ -lactams, cephalosporins, tetracyclines, glycopeptides and fluoroquinolones. *Streptococcus pneumoniae* ATCC 49619 was used as a quality control strain for drug susceptibility.

## 2.7 Information of bacterial strains

In this study, the complete genomes of 230 SS2 strains were retrieved to construct pangenome. These strains were available in March 2023 from NCBI (<ftp://ftp.ncbi.nih.gov/genomes/all/>). In addition, all strains were subjected to Busco assessment, with completeness exceeding 95%, to exclude genomes of low quality. Information about the 230 strains is summarized in [Supplementary Table 1](#).

## 2.8 Pan-genome construction

To sustain the consistency and reliability of gene prediction and annotation, the Prokaryotic Genome Annotation System (Prokka) pipeline (V1.14.5 <https://github.com/tseemann/prokka>, accessed on 5 September 2023) was uniformly applied to all the 230 SS2 genomes. Based on the GFF3 files produced by Prokka, the Roary program (V3.13.0, <https://github.com/sanger-pathogens/Roary>, accessed on 6 October 2023) was used to construct the pangenome with a minimum percentage identity of 95% between each predicted protein homolog. The decision not to choose 100% of the strains was to avoid misclassification caused by low-quality genomes or genome defects in individual strains, ensuring that true core genes were not overlooked during the annotation and classification process.

## 2.9 Gene annotation tool

In order to annotate the core genome and accessory genome, a variety of annotation tools and databases were utilized, as indicated in the [Table 1](#).

# 3 Result

## 3.1 Genome assembly and quality assessment

The attempt to directly assemble the bacterial genome using the filtered data with Flye resulted in a single contig of length 2,093,244 bp, closely matching the average size of the SS2 genome. This outcome suggested that the DNA extraction quality was good. After filtering, it was possible to assemble the bacterial genome directly. However, due to single-base errors in Nanopore sequencing, Illumina sequencing data will be used for correction in subsequent analyses. Illumina sequencing generated a total of 4.2 GB raw data, including 5,863,541 reads with a length of 151 bp and a GC content of 42%. The corrected and uncorrected genome sequences were evaluated using Quast, and the results are shown in [Table 2](#). Pilon mainly corrected small indels and GC content since a contig forming a circular structure was already assembled in Flye. The final total length of the contig was 2,096,025 bp, with a GC content of 41.96%. Busco was employed for the integrity assessment of the genome sequences before and after correction, as shown in [Figure 1](#). The genome sequence directly assembled using Flye exhibited 13 fragments and five missing genes. After correction with Illumina sequencing data, the integrity of the obtained genome was further improved. The number of genome fragments was reduced, and the missing genes were corrected. Therefore, the genome sequence corrected by Pilon is considered the final complete genome sequence of SS2 obtained in this study.

## 3.2 Genome composition analysis

Using Prodigal (V2.6.3, <https://github.com/hyattpd/Prodigal>, accessed on 10 September 2023) for gene prediction resulted in 2019 genes of varying lengths with a GC content of 41.96%, as shown in [Table 3](#) and [Figure 2](#). Genome composition analysis revealed 57 tRNA, 12 rRNA (including 4 5S\_rRNA, 4 16S\_rRNA, and 4 23S\_rRNA), and nine sRNA. Additionally, 25 scattered repetitive sequences (16 short scattered repeats, seven long scattered repeats, and 2 DNA elements) and 46 tandem repeat sequences were identified. The genome also contained two CRISPR sequences, 4 GIs (Genomic Islands), and one prophage sequence. The relevant data has been uploaded to NCBI with the Bioproject ID PRJNA1041968, named cnzyss2-311.

## 3.3 Antimicrobial susceptibility profiles

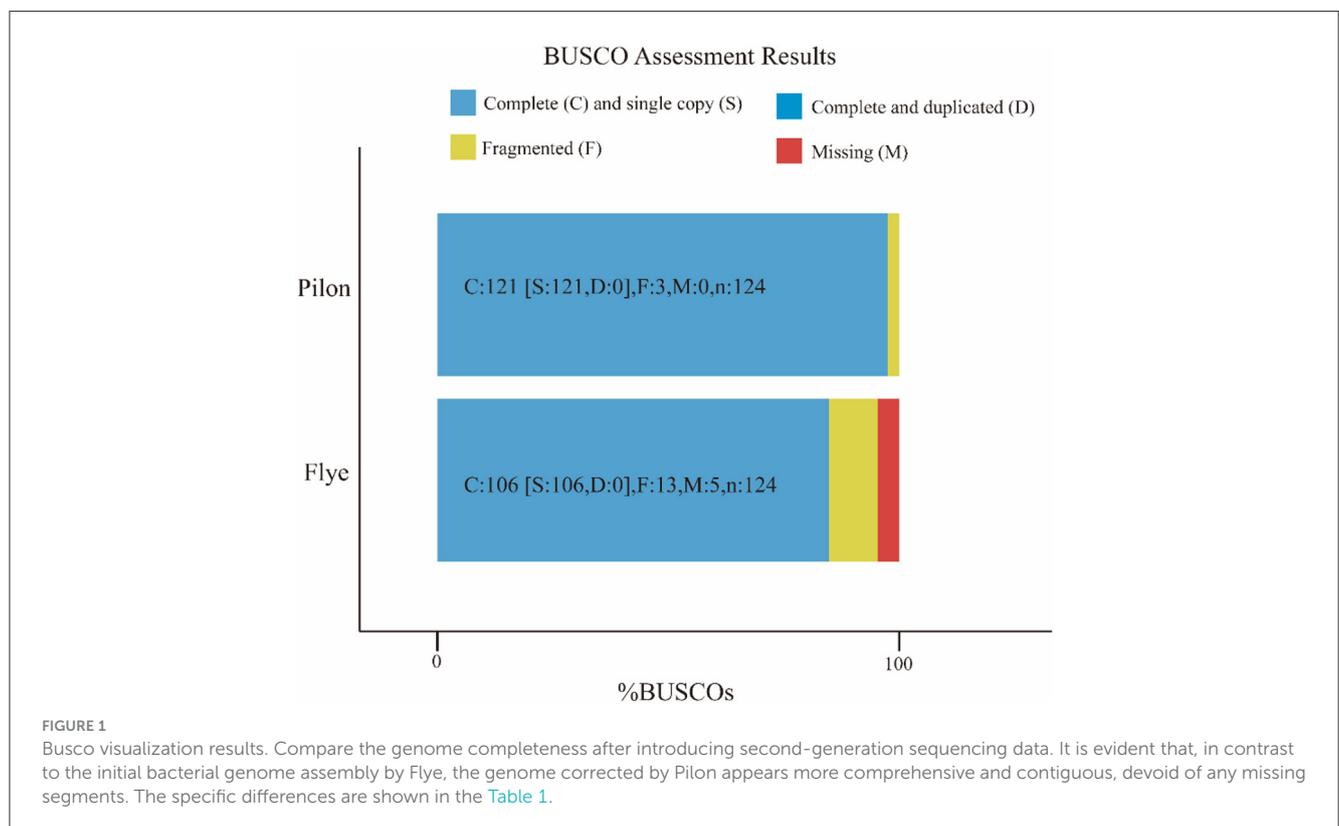
Using the broth microdilution method, the minimum inhibitory concentration (MIC) of isolated strains of *Streptococcus suis* type 2 was determined. The results were

TABLE 1 Gene annotation tools and databases.

Name	URL	Purpose
eggNOG-mapper	<a href="https://github.com/eggnogdb/eggno-mapper">https://github.com/eggnogdb/eggno-mapper</a> (Cantalapiedra et al., 2021)	COG annotation
Interproscan	<a href="https://ftp.ebi.ac.uk/pub/software/unix/iprscan/">https://ftp.ebi.ac.uk/pub/software/unix/iprscan/</a> (Jones et al., 2014)	GO annotation
KEGG	<a href="https://www.genome.jp/kegg/">https://www.genome.jp/kegg/</a> (Kanehisa and Goto, 2000)	KEGG annotation
VFDB	<a href="http://www.mgc.ac.cn/VFs/">http://www.mgc.ac.cn/VFs/</a> (Liu et al., 2022)	Virulence factor annotation
CAZy	<a href="http://www.cazy.org/">http://www.cazy.org/</a> (Drula et al., 2022)	Carbohydrate enzyme annotation
PHI	<a href="http://www.phi-base.org/">http://www.phi-base.org/</a> (Urban et al., 2020)	PHI annotation
CARD	<a href="https://card.mcmaster.ca">https://card.mcmaster.ca</a> (Alcock et al., 2023)	Drug resistance gene annotation

TABLE 2 Quast assessment results.

Class	Software	Length (bp)	GC (%)	Contig	N50
Preliminary results	Flye	2,094,582	41.23	1	2,094,582
Results after correction	Pilon	2,096,025	41.96	1	2,096,025



interpreted according to the CLSI standard (Table 4). The research results indicated that the isolated SS2 strain in this experiment exhibits high resistance to various antibiotics of different classes, including glycopeptides, tetracyclines,  $\beta$ -lactams, cephalosporins, and macrolides, reaching multidrug resistance. Additionally, it showed varying sensitivity to fluoroquinolones, including susceptibility to ofloxacin (S), intermediate resistance to levofloxacin (I), and resistance to trovafloxacin (R).

### 3.4 Construction and phylogenetic analysis of the SS2 pan-genome

The pan-genome of 230 SS2 strains was constructed after uniform annotation of their genomic sequences. The results indicated that the SS2 pan-genome comprises a total of 5,792 genes. Among them, there were 1,353 core genes shared by several strains ranging from 99 to 100%. In addition, there are 105 soft-core genes carried by 95–99% of strains; 541 shell genes carried by 15–95% of

strains; and 3,796 cloud genes shared by 15% of strains (Figure 3A). However, considering that the functions of the four gene sets were not fully understood, only the core genome and accessory genome were used in the following analysis.

According to the composition of the SS2 pan-genome, it was found that the core genome of SS2 consists of 1,458 genes, accounting for 25.12% of the pan-genome, while the accessory genome was composed of 4,337 genes, indicating a high degree of genome variability in SS2. As shown in Figure 3B, the pan-genome was open, allowing continuous acquisition of foreign genes to adapt to different environments.

### 3.5 Gene function annotation

The eggNOG-mapper (V5.0, <https://github.com/eggnogdb/eggnog-mapper>, accessed on 10 September 2023) was employed for functional annotation by aligning with the Cluster of Orthologous Groups of proteins (COG) database to analyze and infer gene functions. The genes obtained in Section 3.2 were annotated through COG analysis and the specific results are shown in Figure 4A. The functions of the genes in the SS2 strain obtained in this study are mainly concentrated in categories such as E (Amino Acid Transport and Metabolism), G (Carbohydrate Transport and Metabolism), J (Translation, Ribosomal Structure,

and Biogenesis), K (Transcription), L (Replication, Recombination, and Repair), R (General Function Prediction Only), and other functions directly related to bacterial survival. Moreover, ~140 genes still have unknown functions. In addition, for the assembled genome sequence, combined with the predicted results of coding genes, a circular genome map was drawn to comprehensively display the features of the genome, and the result is shown in Figure 4C.

The COG annotation results for the pan-genome are shown in Figure 4B and Table 5. The results indicated that the core genome participated in various aspects of bacterial life processes. Moreover, almost all functions included a certain number of core genomes. The main functions of the core genome were concentrated in COG categories such as C (energy production and conversion), E (amino acid transport and metabolism), F (nucleotide transport and metabolism), J (translation, ribosomal structure, and biogenesis), P (inorganic ion transport and metabolism), which were biologically significant for maintaining bacterial metabolism.

The accessory genome was annotated to all functional categories. It was annotated to categories not involved in the core genome (B, chromatin structure and dynamics). Additionally, COG annotation of the accessory genome was mainly concentrated in D (cell cycle control, cell division, and chromosome partitioning), L (replication, recombination, and repair), G (carbohydrate transport and metabolism), H (coenzyme transport and metabolism), indicating that the accessory genome was not only involved in bacterial division and reproduction but also involved the transport metabolism of necessary substances, showing a certain degree of essentiality. On the one hand, the accessory genome in K (transcription), O (post-translational modification), and U (intracellular trafficking, secretion, and vesicular transport)

TABLE 3 Genetic statistics.

Counts	Total length	Maximum length	Minimum length	GC content
2,019	1,837,305	6,309	90	41.96%

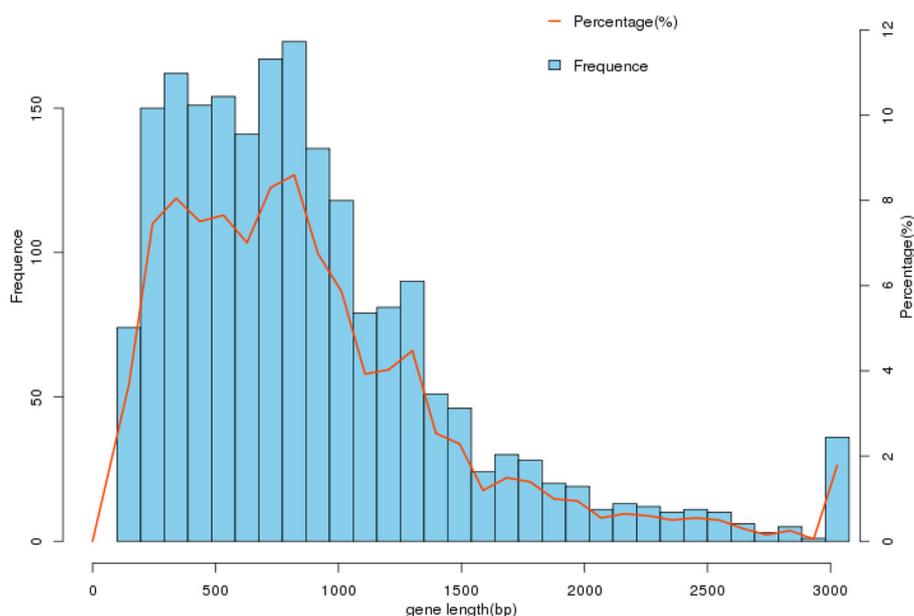
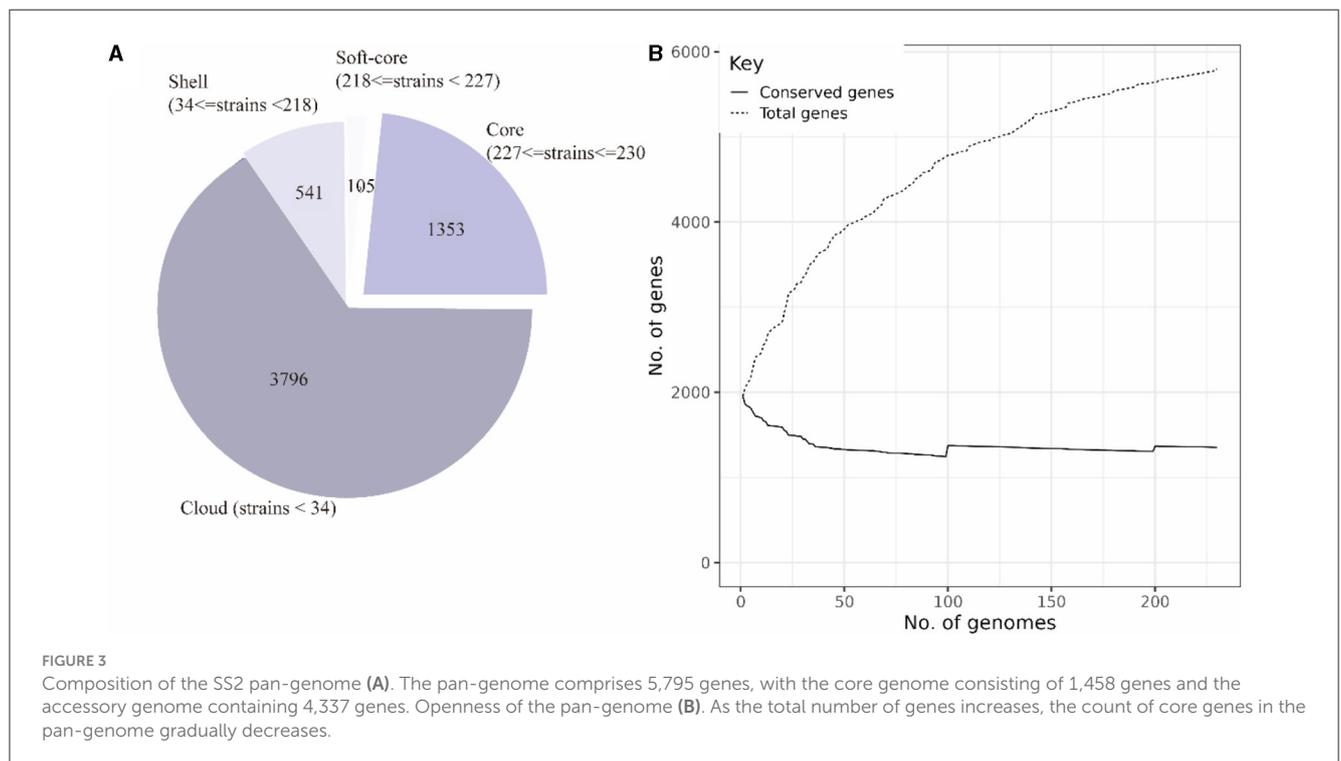


FIGURE 2

Gene length distribution. In the gene set cznyss2-311, the gene lengths are predominantly distributed within the range of 0–3,000 bp, with the highest number of genes falling between 500 and 1,000 bp. The specific differences are shown in the Table 2.

TABLE 4 Determination of minimum inhibitory concentration (MIC) results for cnzyss2-311 strain using microdilution broth method.

Antibiotic category		MIC ( $\mu\text{g/mL}$ )	Susceptible (S)	Intermediate (I)	Resistant (R)
Macrolides	Erythromycin	8	$\leq 0.25$	0.5	$\geq 1$
Cephalosporins	Ceftazidime	2	$\leq 0.5$	-	-
	Cefepime	32	$\leq 0.5$	-	-
$\beta$ -lactams	Penicillin	32	$\leq 0.12$	-	-
	Penicillin G sodium salt	16	$\leq 0.06$	0.12–1	$\geq 2$
	Amoxicillin	64	$\leq 0.5$	1	$\geq 2$
Tetracyclines	Doxycycline	4	$\leq 0.5$	1	$\geq 2$
	Tetracycline	16	$\leq 2$	4	$\geq 8$
Lincomycins	Lincomycin	64	-	-	$\geq 2$
Glycopeptides	Vancomycin	2	$\leq 1$	-	-
Fluoroquinolones	Ofloxacin	0.25	$\leq 2$	4	$\geq 8$
	Levofloxacin	4	$\leq 2$	4	$\geq 8$
	Trovafloxacin	4	$\leq 1$	2	$\geq 4$



annotated quantities were significantly higher than the core genome, which might reveal the crucial role of the accessory genome in bacterial protein expression and transport processes. On the other hand, the accessory genome was concentrated in M (cell wall/membrane/envelope biogenesis), T (signal transduction mechanisms), and V (defense mechanisms). These three categories were related to the perception, response, defense, and adaptation of bacteria to changes in the external environment, maintaining survival and reproduction. Moreover, it might provide selective advantages and enrich population diversity. However, whether in the core genome or accessory genome, a large number of genes were

still annotated to categories of unknown function (S), requiring further research.

### 3.6 The GO annotation of the core genome and the accessory genome

The two gene sets were aligned and annotated with the Interproscan (<https://ftp.ebi.ac.uk/pub/software/unix/iprscan/>, accessed on 25 September 2023), and the results were shown

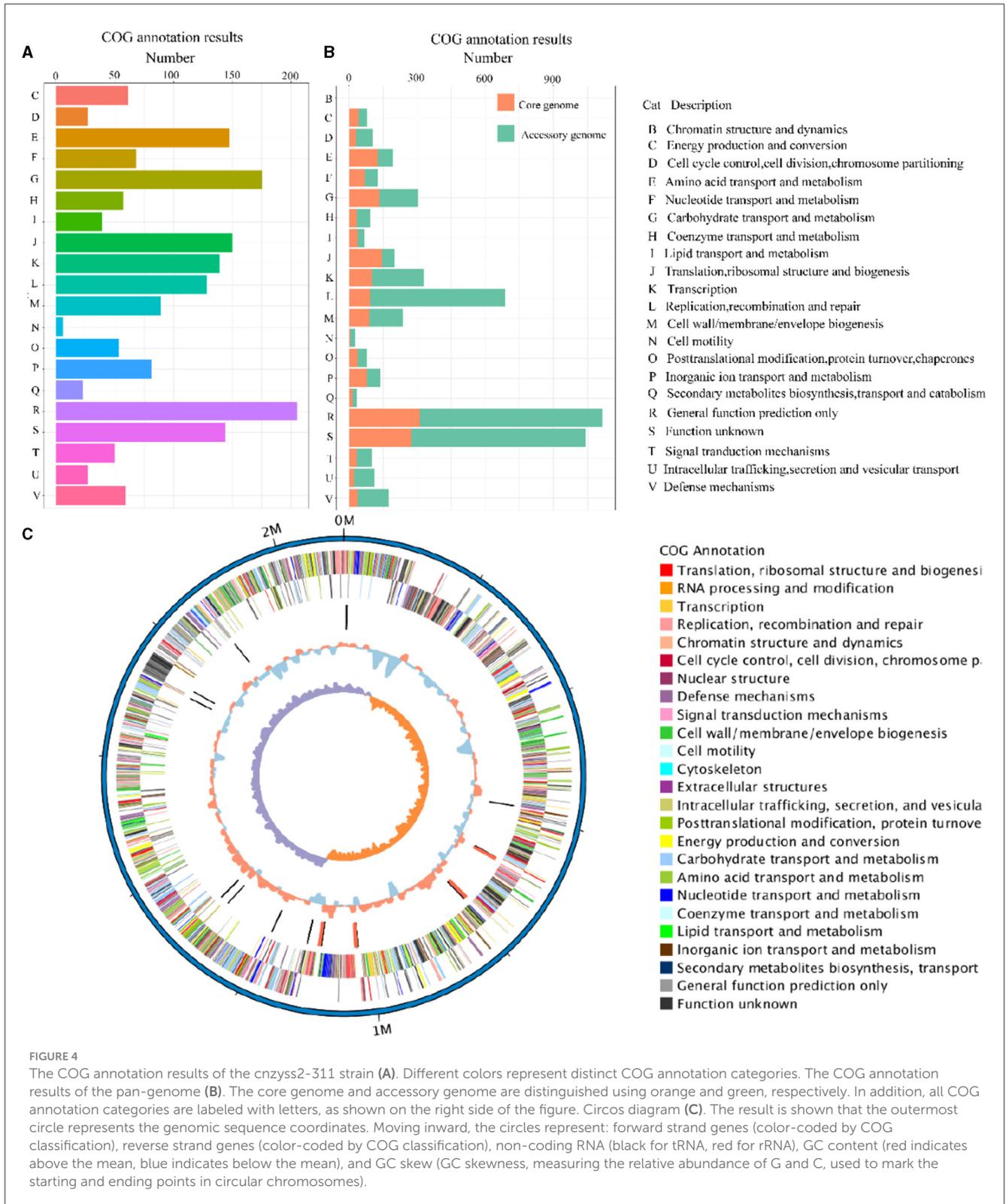


FIGURE 4

The COG annotation results of the cnzys2-311 strain (A). Different colors represent distinct COG annotation categories. The COG annotation results of the pan-genome (B). The core genome and accessory genome are distinguished using orange and green, respectively. In addition, all COG annotation categories are labeled with letters, as shown on the right side of the figure. Circos diagram (C). The result is shown that the outermost circle represents the genomic sequence coordinates. Moving inward, the circles represent: forward strand genes (color-coded by COG classification), reverse strand genes (color-coded by COG classification), non-coding RNA (black for tRNA, red for rRNA), GC content (red indicates above the mean, blue indicates below the mean), and GC skew (GC skewness, measuring the relative abundance of G and C, used to mark the starting and ending points in circular chromosomes).

in the Figure 5. Compared to the accessory genome, although the core genome was widely involved in biological processes, molecular functions, and cellular components, the core genome had more annotations in the cellular component category. This reflected that the cellular structural positions where

gene products execute functions were mainly encoded by the core genome. In addition, the accessory genome was also extensively involved in various processes that maintained bacterial survival, significantly enriching molecular function and biological process categories, showing certain essentiality, and possibly

TABLE 5 Cluster of Orthologous Groups of Proteins (COG) annotation of core genome and accessory genome.

Class	Description	Core genome	Accessory genome
B	Chromatin structure and dynamics	NA	1
C	Energy production and conversion	42	38
D	Cell cycle control, cell division, and chromosome partitioning	29	75
E	Amino acid transport and metabolism	124	68
F	Nucleotide transport and metabolism	69	59
G	Carbohydrate transport and metabolism	133	169
H	Coenzyme transport and metabolism	33	60
I	Lipid transport and metabolism	37	29
J	Translation, ribosomal structure, and biogenesis	145	53
K	Transcription	98	230
L	Replication, recombination and repair	93	593
M	Cell wall/membrane/envelope biogenesis	88	148
N	Cell motility	6	19
O	Posttranslational modification, protein turnover, and chaperones	37	41
P	Inorganic ion transport and metabolism	77	61
Q	Secondary metabolites biosynthesis, transport, and catabolism	15	18
R	General function prediction only	312	805
S	Function unknown	275	767
T	Signal transduction mechanisms	33	66
U	Intracellular trafficking, secretion, and vesicular transport	21	89
V	Defense mechanisms	37	137

providing beneficial supplementation when the core genome was damaged.

Furthermore, the core genome and accessory genome had unique annotations in GO. Protein folding chaperone and translation regulator activities were unique GO annotations for the core genome, indicating the crucial role of the core genome in protein formation and regulation. The unique GO annotations for the accessory genome included biological processes involved in interspecies interactions, multicellular organism processes, viral processes, and cell skeleton movement activities. It revealed that the accessory genome might play a key role in resisting phage invasion, interacting with hosts, and providing a foundation for dynamic changes in cells.

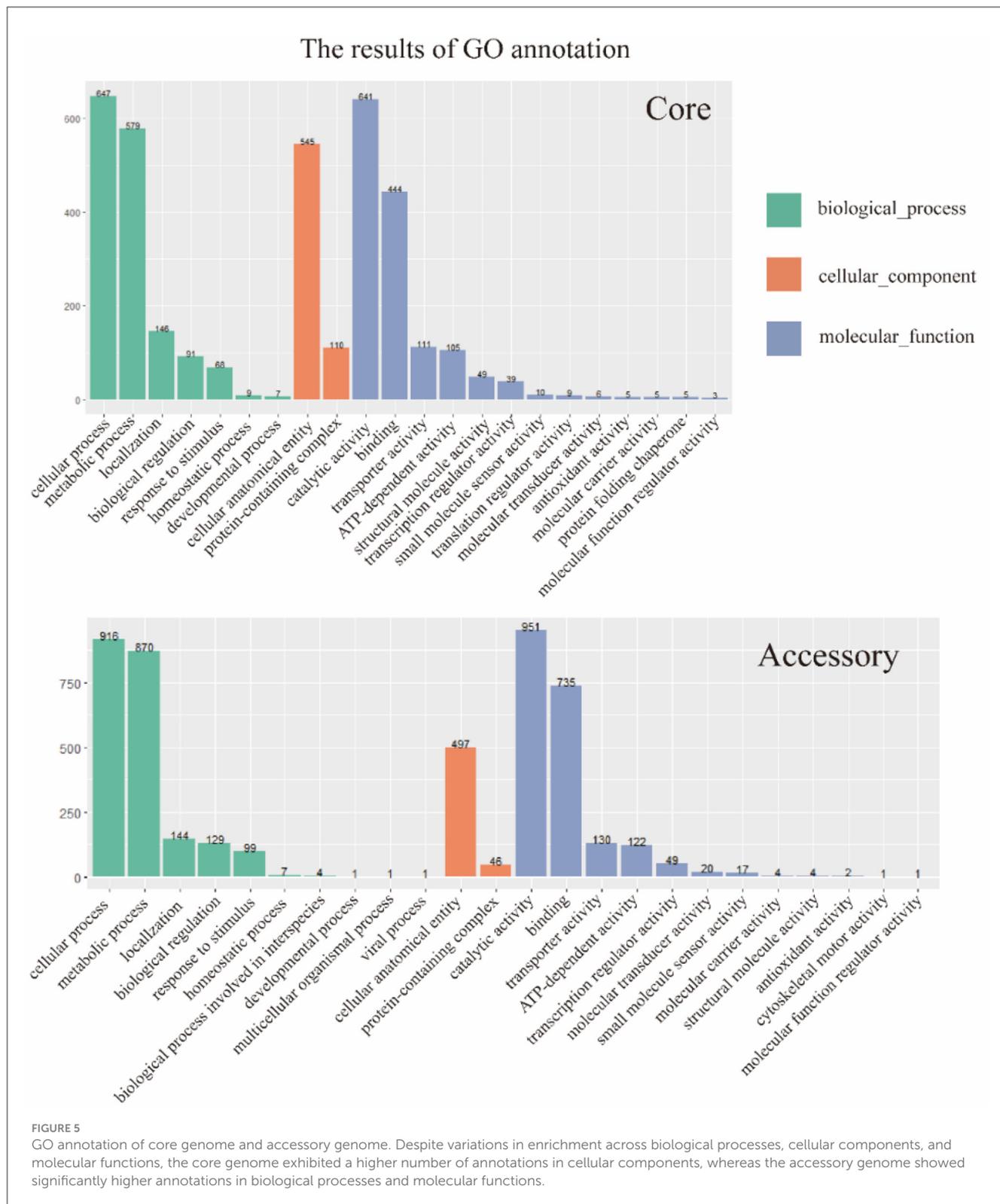
### 3.7 The KEGG annotation results of the core genome and accessory genome

The results of the alignment of the two gene sets with the Kyoto Encyclopedia of Genes and Genomes (KEGG, <https://www.genome.jp/kegg/>, accessed on 25 September 2023) database (Kanehisa and Goto, 2000) were shown in the Figure 6. The results indicated that the core genome was widely involved in the metabolism, genetic information processing, environmental information processing, and cellular processes of the bacterium. The annotation count was generally higher than that of the accessory genome. Moreover, it was annotated to pathways not covered by the accessory genome, such as transcription and cell motility, demonstrating the biological essentiality of the core genome. Similarly, the accessory genome also exhibited varying numbers of annotations in these four aspects, indicating a certain degree of biological essentiality, serving as a complement and modification to the core genome.

### 3.8 The VFDB annotation results of the core genome and accessory genome

The results of the annotation of the two gene sets with the virulence factor database (VFDB, <http://www.mgc.ac.cn/VFs/>, accessed on 20 September 2023) database (Liu et al., 2022) to identify potential virulence factors. The results were shown below (Figure 7 and Table 6). The core genome and accessory genome obtained annotations for the following categories of virulence factors: adhesion, enzymes, immune invasion, proteases, anti-phagocytosis, phagosome capture, and toxins. Compared to the core genome, the accessory genome was completely lacking in the categories of anti-phagocytosis and phagosome capture, focusing primarily on adhesion and protease categories, and annotating virulence factors (toxins) not involved in the core genome.

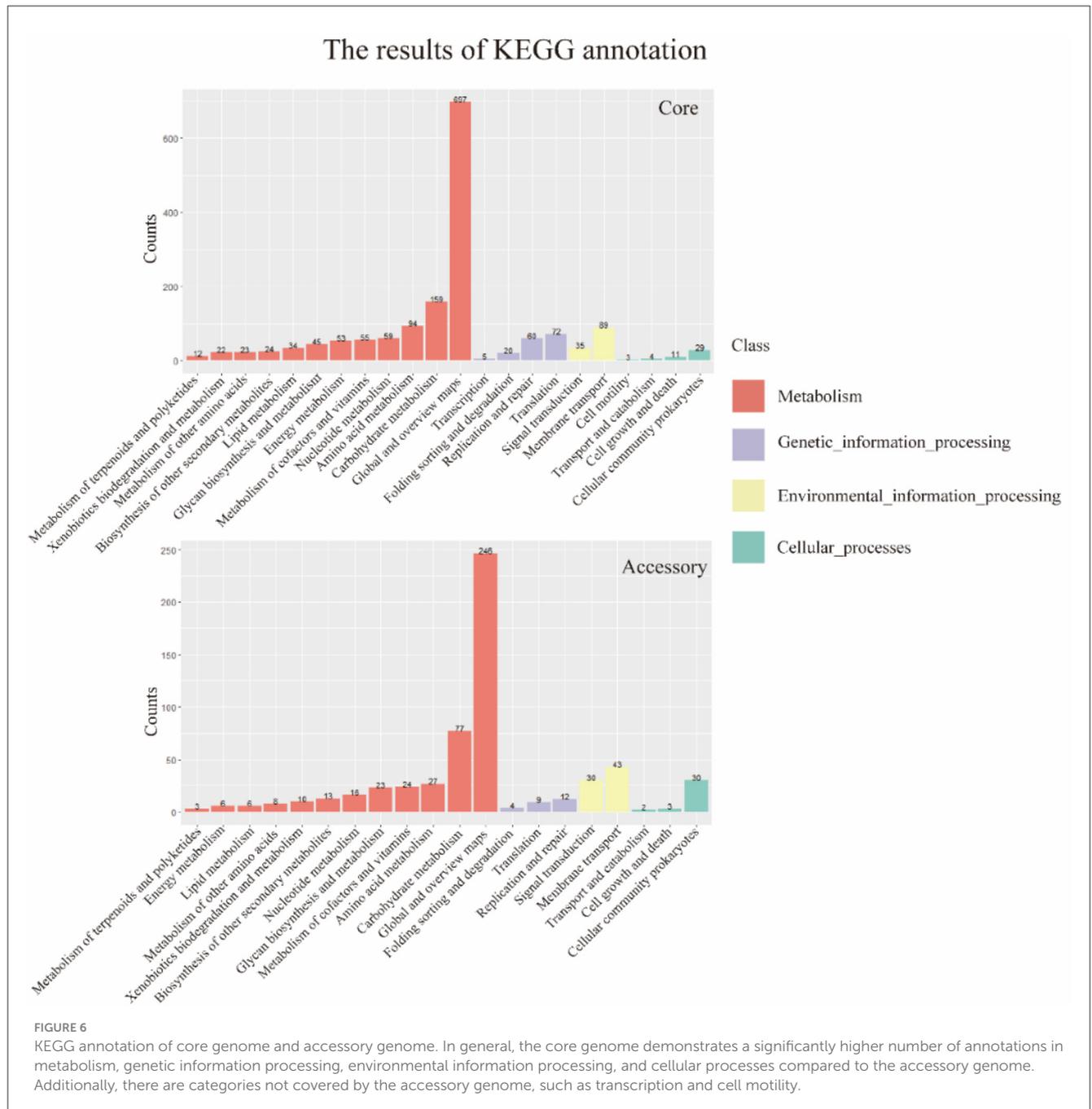
Our findings indicated the presence of a diverse array of virulence factors in the pan-genome, potentially providing pathways for various pathogenic mechanisms of SS2. This diversity enhances the bacterium's survival opportunities and contributes to more severe pathological responses.



### 3.9 The CAZy annotation of results of the core genome and accessory genome

The alignment and annotation of the two gene sets with the Carbohydrate-Active enZymes (CAZy, <http://www.cazy.org/>,

accessed on 20 September 2023) database (Drula et al., 2022) were shown in the Figure 8 and Table 7. Both the core genome and accessory genome were involved in all categories, with the core genome primarily annotated as carbohydrate esterases (CE) and glycosyltransferases (GT). In contrast, the accessory genome



was annotated with much more glycoside hydrolases (GH) and carbohydrate-binding modules (CBM) than the core genome.

### 3.10 The PHI annotation of the core genome and the accessory genome

The alignment of the two gene sets with the Pathogen-Host Interaction (PHI, <http://www.phi-base.org/>, accessed on 20 September 2023) database (Urban et al., 2020) and subsequent annotation revealed that pathogenic genes were commonly present in both the core and accessory genomes (Figure 9 and Table 8). Furthermore, genes associated with different PHI phenotypes

exhibited much higher abundance in the accessory genome compared to their abundance in the core genome. This observation was closely related to bacterial serum resistance, adhesion capabilities, and invasiveness.

### 3.11 The CARD annotation of the cnzys2-331 and pan-genome

The results of the annotation using RGI (V6.0.3, <https://github.com/arpard/rgi>, accessed on 20 September 2023) against the comprehensive antibiotic resistance database (CARD, <https://card.mcmaster.ca>, accessed on 20 September 2023; Alcock et al.,

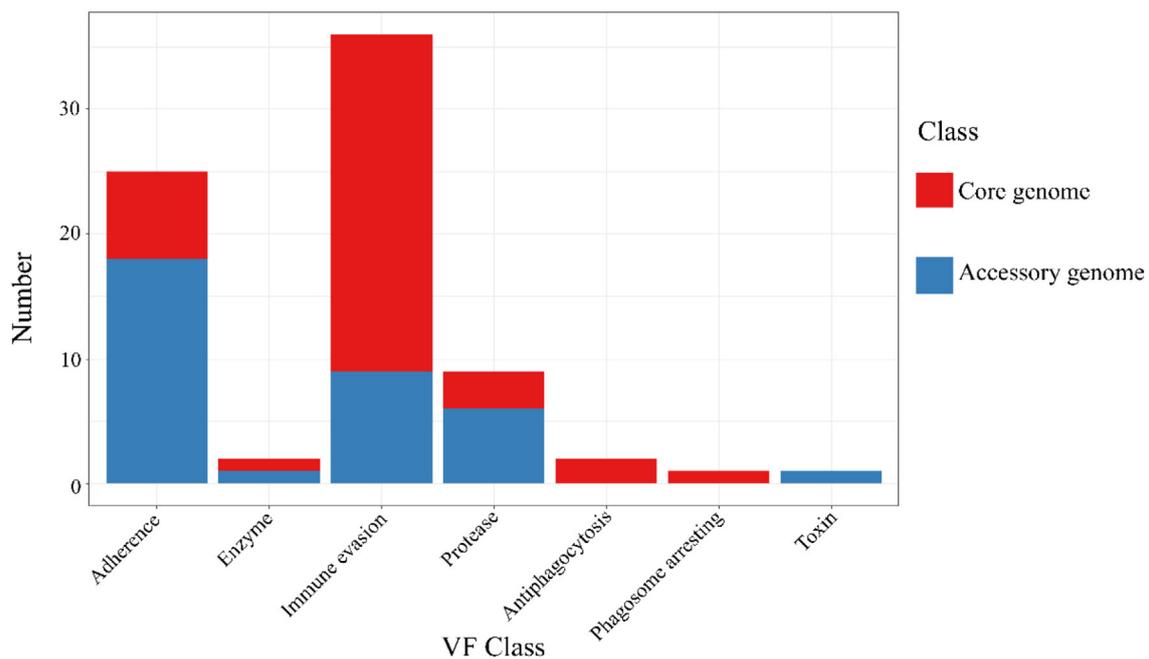


FIGURE 7

VFDB annotation of core genome and accessory genome. While both the core genome and accessory genome show varying degrees of annotations in adhesion, enzymes, immune invasion, and proteases, antiphagocytosis and phagosome arresting are unique annotation categories found in the core genome, whereas toxin is a unique annotation category in the accessory genome.

2023) are shown in Figure 10. In the perfect or strict hits mode, four types of resistance genes were annotated in *cnzyss2-331*, namely *ErmB*, *tet(O)*, *patB*, and *vanY* gene in *vanB* cluster, exhibiting resistance to streptogramins, macrolides, lincosamides, tetracyclines, fluoroquinolones, and glycopeptide antibiotics. These findings aligned well with the MIC test results in Table 4. Our results indicated that CARD annotation correlates with the antibiotic resistance phenotype.

The core genome and accessory genome were compared and annotated against the CARD database. The resistance mechanisms were categorized into six main types: antibiotic target alteration, antibiotic efflux, antibiotic inactivation, antibiotic target protection, antibiotic target replacement, and reduced antibiotic permeability (Figure 10 and Table 9). In the perfect or strict hits mode, the core genome was annotated with only two resistance genes, specifically, glycopeptide resistance genes and fluoroquinolone resistance genes. In contrast, the accessory genome contributed the majority of resistance genes (aminoglycosides, phenicol, streptogramin, macrolides, lincosamides, nucleosides, diaminopyrimidines, and tetracyclines), demonstrating a diverse array of resistance mechanisms, including antibiotic target alteration, antibiotic efflux, antibiotic inactivation, antibiotic target protection and antibiotic target replacement.

## 4 Discussion

Pan-genome analysis provides a comprehensive understanding of the overall lineage of SS2. It can be used to screen and identify various virulence and antibiotic resistance genes. In this study, the

strain *cnzyss2-311* was initially assembled into a genomic sketch using third-generation sequencing data, and second-generation sequencing data were used for error correction, resulting in the completion of the full genomic sequence. Subsequently, pan-genome analysis of SS2 revealed that only 1353 genes were shared among different individuals, constituting the core genome (Figure 3). The core genome serves as the essential framework supporting the rest of the genome, rather than the minimal set of genes required for bacterial survival (Medini et al., 2005; Tettelin et al., 2008). If we expand the definition of the core (including genes that are only partially missing in a small fraction of the genomes), the core genome consists of 1,458 genes, including both core and soft genes. These genes are present in at least 95% of the sampled genomes. The results suggested that SS2 has an open pan-genome, and the number of core genome genes does not significantly change with an increasing number of strains and consistent with previous research (Guo et al., 2021). In contrast, genes present in certain strains and those unique to individual strains constitute the accessory genome, comprising 4,337 genes. The accessory genome reflects the genetic diversity and unique genomic features present in specific strains (Kim et al., 2020). Pan-genome analysis provides insights into the genomic variability of SS2, highlighting both the conserved core genes and the flexible accessory genes that contribute to its adaptability and diversity across strains.

The core genome and accessory genome are extensively involved in various aspects of bacterial life activities, including metabolism and genetic variation (Figures 4–6). The core genome primarily focuses on the transport, metabolism, and translation processes of essential substances for life. Similarly, the accessory

TABLE 6 The virulence factor database (VFDB) annotation of the core genome and the accessory genome.

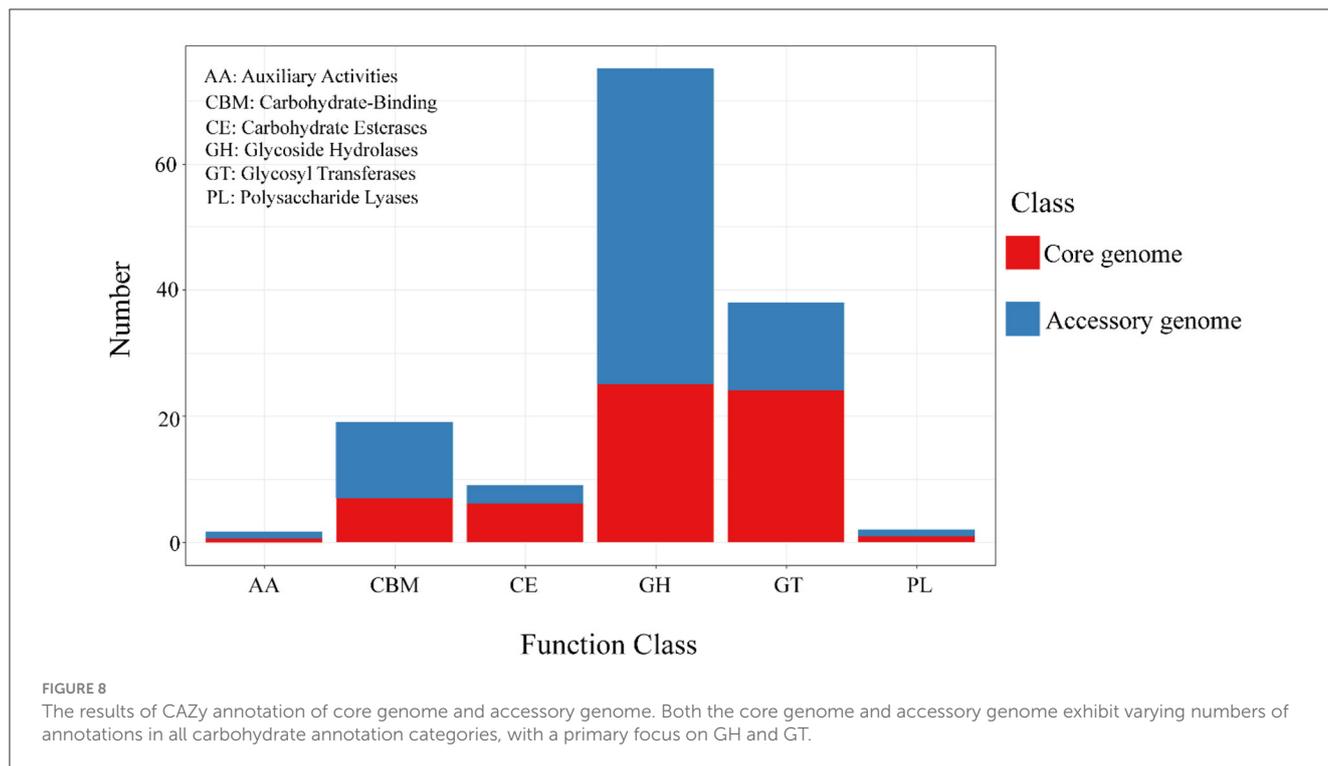
VF class	Virulence factors	Related genes	Core genome	Accessory genome	
Adherence	Streptococcal lipoprotein rotamase A	<i>slrA</i>	1	NA	
	Streptococcal plasmin receptor/GAPDH	<i>plr/gapA</i>	1	NA	
	LrA islet	<i>srtC</i>	2	NA	
	Fibronectin-binding proteins	<i>pavA</i>	1	NA	
	Laminin-binding protein	<i>Lmb</i>	1	NA	
	Sortase A	<i>srtA</i>	1	1	
	Agglutinin receptor	Undetermined	NA	8	
	Choline-binding proteins		<i>mrp</i>	NA	3
			<i>srtC4</i>	NA	2
		<i>srtB</i>	NA	1	
		<i>cbpD</i>	NA	3	
Enzyme	Streptococcal enolase	<i>eno</i>	1	NA	
	Mitogenic factor 3	<i>mf3</i>	NA	1	
Immune evasion	Polysaccharide capsule (Bacillus)	<i>manA</i>	1	NA	
	Capsule	Undetermined	26	9	
Protease	Serine protease	<i>htrA/degP</i>	1	NA	
	Trigger factor	<i>tig/ropA</i>	1	NA	
	Zinc metalloproteinase	<i>zmpC</i>	1	2	
	C5a peptidase	<i>scpA/scpB</i>	NA	3	
	Extracellular factor	<i>epf</i>	NA	1	
Antiphagocytosis	Capsule (Enterococcus)	<i>cdsA</i>	1	NA	
		<i>cps2j</i>	1	NA	
Phagosome arresting	Nucleoside diphosphate kinase	<i>ndk</i>	1	NA	
Toxin	Suilysin	<i>sly</i>	NA	1	

genome exhibits certain biological characteristics, especially in functions such as carbohydrate transport and metabolism, transcription, etc. These examples indicate that the primary function of the core genome is to control the normal morphology, reproduction, and execution of basic biological functions in bacteria. In addition, genes participating in fundamental biological processes have been discovered in the accessory genome. It is speculated that when the core genome is damaged, the accessory genome may have alternative functions to maintain certain biological processes in bacteria.

Similarly, the core genome and accessory genome exhibit differences in the expression of carbohydrate enzymes (Figure 8). For example, the annotation of the core genome revealed a high abundance of CE and GT. These are associated with the hydrolysis of carbohydrates and the formation of glycosidic bonds (Venegas et al., 2022). It plays a crucial role in bacteria's acquisition of external carbon sources, adhesion to host cells, and biofilm formation (Middleton et al., 2021; Na et al., 2021). In addition, the accessory genome exhibits significantly higher levels of GH compared to the core genome. GH is mainly involved in the hydrolysis of polysaccharide glycosidic bonds, enabling the degradation of

complex host polysaccharides, facilitating bacterial acquisition of carbohydrate nutrients, and promoting the colonization of *Streptococcus suis* (Chen et al., 2020; Hamre and Sørli, 2020; Redman et al., 2020).

In general, bacterial adhesion to the surface of objects to form a biofilm is crucial for further infecting the host (Fittipaldi et al., 2012). SS2 typically colonizes the upper respiratory tract, and crossing the mucosal barrier is a prerequisite for SS2 to cause invasive infections (Xia et al., 2019). The annotation results of virulence factors also suggest that, aside from undefined capsule-related genes, the most abundantly annotated virulence genes in the pan-genome belong to the adhesion category. Streptococcal lipoteichoic acid rotamase A (SlrA) has been shown to indirectly promote host cell adhesion and invasion, as well as prevent phagocytosis of *Streptococcus pneumoniae*, making it a potential therapeutic target for preventing bacterial colonization (Cron et al., 2009). Additionally, SS2 possesses a homologous lipoteichoic acid protein gene (*slrA*), but lacks choline-binding proteins (CBPs; Hermans et al., 2006). Interestingly, the virulence factor annotation results also show that CBPs are completely absent in the core genome of SS2. CBPs have been shown to play a crucial role in



the growth, autolysis, and biofilm formation of *S. pneumoniae* and are essential for its pathogenicity (Galán-Bartual et al., 2015). Moreover, in Gram-positive bacteria, various surface proteins decorated with proteins are accomplished by sortase (Spirig et al., 2011). Although not essential for bacterial viability, sortase may be an important virulence factor, as it is involved in mediating bacterial adhesion to host tissues, host cell entry, evasion and inhibition of the immune response, and acquisition of essential nutrients by surface proteins (Marraffini et al., 2006). Our results show that streptococcal sortase A (*srtA*) and pilin-sorting enzyme C (*srtC*) are fixed encoded sortases in SS2. SrtA recognizes specific LPXTG motifs, anchoring a diverse array of functionally distinct proteins to the cell wall, playing a crucial role in adhesion to extracellular matrix (ECM) components (Spirig et al., 2011; Li et al., 2013; Alharthi et al., 2021). Considering the essential role of SrtA in the pathogenic mechanisms of Gram-positive bacteria, it is also considered an ideal target for potential drugs (Spirig et al., 2011). SrtC sorting enzyme not only anchors smaller substrates but also participates in pilus assembly, suggesting that the formation of pili is an inherent ability of SS2 to promote microbial adhesion and biofilm formation (LeMieux et al., 2008; Qian et al., 2018; Faulds-Pain et al., 2019). In addition, during the process of adhering to endothelial cells and epithelial cells in host tissues, SS2 mediates host cell invasion through its own ECM-binding protein, binding to host ECM components fibronectin (FN) and laminin (LN; Schwarz-Linek et al., 2003; Wahid et al., 2005; Raganathan et al., 2009). Interestingly, highly conserved hexapeptide sequences (LPXTGE) have been identified in the C-terminus of known surface proteins in Gram-positive cocci and play an important role in cell adhesion, such as the aforementioned sortases, fibronectin-binding proteins (FnBPs), and laminin-binding proteins (LNBPs; Fischetti et al., 1990; Tenenbaum et al., 2007; Yamaguchi et al., 2013).

**TABLE 7** The Carbohydrate-Active enZymes (CAZy) annotation of the core genome and the accessory genome.

CAZy classification	Core genome	Accessory genome
Auxiliary Activities (AA)	1	1
Carbohydrate-Binding Modules (CBM)	7	12
Carbohydrate Esterases (CE)	6	3
Glycoside Hydrolases (GH)	25	50
Glycosyl Transferases (GT)	24	14
Polysaccharide Lyases (PL)	1	1

However, there is also research indicating that surface proteins encoded by the *pavA* gene, which lacks a typical LPXTG anchoring motif, still play an important role in the adhesion and virulence of *Streptococcus pneumoniae* (Pracht et al., 2005). In conclusion, the core genome of SS2 contains diverse adhesion factors and adhesion proteins, and surface proteins with LPXTG anchoring motifs may not be the sole criterion driving the interaction between the bacterium and the host.

Interestingly, in the core genome, enolase (*eno*) was identified as an essential virulence factor for *Streptococcus suis* serotype 2 (Xia et al., 2019). It imparts the bacterium with the ability to directly produce toxicity to porcine brain microvascular endothelial cells (PBMECs), promoting cell apoptosis, inhibiting the expression of tight junctions, or activating the ERK and p38MAPK signaling pathways in porcine brain microvascular endothelial cells to secrete the pro-inflammatory factor IL-8, and increasing the permeability

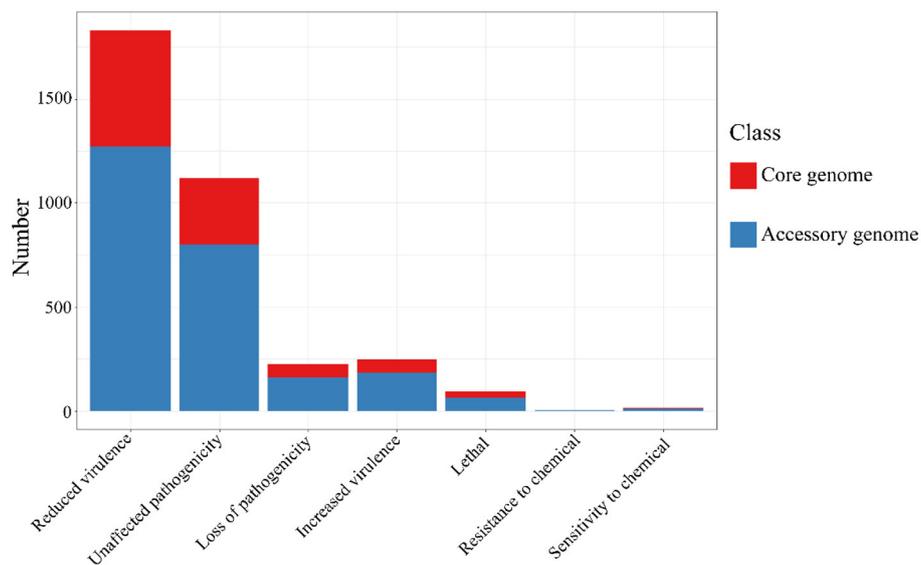


FIGURE 9

The results of PHI annotation of core genome and accessory genome. The accessory genome shows varying numbers of annotations in all PHI phenotype categories, including categories not covered by the core genome, such as resistance to chemical.

TABLE 8 The Pathogen-Host Interactions (PHI) annotation of the core genome and the accessory genome.

PHI phenotype classification	Core genome	Accessory genome
Reduced virulence	560	1,271
Unaffected pathogenicity	321	797
Loss of pathogenicity	66	157
Increased virulence	66	183
Lethal	29	61
Resistance to chemical	1	2
Sensitivity to chemical	3	9

of the blood-brain barrier (Liu H. et al., 2021). Therefore, even in strains considered non-virulent or low-virulent, there is still a possibility of penetrating the blood-brain barrier, which may be one of the reasons for SS2 causing host meningitis. Moreover, classical virulence factors of *Streptococcus suis*, such as sly, muramidase-released protein (mrp), and extracellular factor (ef), are found exclusively in the accessory genome (Li et al., 2017). For example, suilysin (sly), as one of the cholesterol-dependent bacterial cytolysins, exhibits direct cytotoxicity (Vötsch et al., 2019). Sly induces cell membrane rupture, decreased cytoplasm density, and even exudation through perforation, leading to lesions and damage to epithelial cells and fibroblasts from different host sources (Liu M. et al., 2021). It is a recognized important virulence factor of SS2. This phenomenon suggests that these virulence factors could serve as criteria for assessing the virulence of SS2, as their presence varies among different strains, as previously reported.

In addition, only in the core genome, virulence genes related to anti-phagocytosis and phagosome capture, such as *cps2J*, *cdsA*,

and *ndk*, were found. *Cps* is a recognized virulence factor, not only an important basis for the serotype classification of *Streptococcus suis*, but also an important factor hindering phagocytosis. Its presence requires the combined action of immunoglobulins and complement to promote phagocytosis (Zhao et al., 2015). In addition, researchers have found that the capsule switch in *Streptococcus suis* can be achieved through gradual evolution with a combination of minor mutation, deletion, and recombination in the *cps* locus (Zhu et al., 2020). Therefore, this may also be the reason why *cps* genes were found in both the core genome and accessory genome. Phosphatidylglycerol transferase, encoded by *cdsA* gene, is a major enzyme in the synthesis of cell membrane phospholipids (Adams et al., 2017; Sawasato et al., 2019). Although studies have found that it mediates resistance to daptomycin in streptococci and enterococci and may resist innate and exogenous antimicrobial peptide attacks, the specific anti-phagocytic mechanism remains unclear (Mishra et al., 2017; Tran et al., 2019). In addition, nucleoside diphosphate kinase (*ndk*) is a nucleotide metabolism enzyme that not only maintains the ribonucleotide and deoxyribonucleotide pools in cells but also participates in the regulation of virulence-related traits related to quorum sensing systems (QS), Type III Secretion Systems (T3SS), and virulence factor production systems in *Pseudomonas aeruginosa* (Neeld et al., 2014; Yu et al., 2016, 2017). Moreover, the *ndk* in *Mycobacterium tuberculosis* has macrophage toxicity and plays a crucial role in evading the host immune system's elimination (Chopra et al., 2003). In summary, the core genome serves as the fundamental assurance for the bacterium to invade the host, while the virulence factors present in the accessory genome complement the core genome advantageously, providing various pathways for bacterial adhesion to the host and evasion of the immune system. Moreover, various virulence factors existing in the core genome indicated that strains previously considered non-virulent may still



contribute to infection. The diverse virulence factors present in the accessory genome can be viewed as an effective complement to the core genome, influencing the pathogenicity of the bacterium. The redundancy of various virulence factors in *Streptococcus suis* allows them to compensate for the loss of another factor. Finally, non-virulent strains also have the possibility of acquiring virulence genes and transforming into virulent strains (Griffith, 1928; Tribble et al., 2012).

It is generally believed that antimicrobial resistance can be categorized into intrinsic resistance and acquired resistance based on its underlying causes (Olaitan et al., 2014; Zhang and Feng, 2016). While acquired resistance is commonly considered the major factor leading to the development of resistant bacteria,

intrinsic resistance may still pose challenges to the treatment of bacterial infections (Sirichoat et al., 2020). According to the annotation results from the CARD database, the core genome of SS2 contains two types of antibiotic resistance genes, namely *patB* and *vanY* (Figure 10 and Table 9). *PatB* is a natural resistance gene in *Streptococcus pneumoniae*, conferring resistance to fluoroquinolone antibiotics through the ABC family (El Garch et al., 2010). This result supported previous research, indicating that SS2 achieves efflux of fluoroquinolones through SatAB (an ABC transporter homologous to PatA and PatB; Escudero et al., 2011). Gram-positive bacteria like lactobacilli have natural resistance to glycopeptide antibiotics (such as vancomycin) controlled by the *vanY* gene (Finch et al., 2010; Alby and Miller, 2018). The results

TABLE 9 The comprehensive antibiotic research database (CARD) annotation (perfect or strict hits).

Class	Antibiotic resistance ontology (ARO) term	Drug class	Resistance mechanism	AMR gene family
cnzys2-311	<i>ErmB</i>	Streptogramin, macrolide, lincosamide, streptogramin B, and streptogramin A	Antibiotic target alteration	Erm 23S ribosomal RNA methyltransferase
	<i>tet(O)</i>	Tetracycline	Antibiotic target protection	Tetracycline-resistant ribosomal protection protein
	<i>patB</i>	Fluoroquinolone	Antibiotic efflux	ATP-binding cassette (ABC) antibiotic efflux pump
	<i>vanY gene in vanB cluster</i>	Glycopeptide	Antibiotic target alteration	Sugar-peptide resistance gene cluster
Core genome	<i>VanY gene in vanB cluster</i>	Glycopeptide	Antibiotic target alteration	Sugar-peptide resistance gene cluster
	<i>PatB</i>	Fluoroquinolone	Antibiotic efflux	ATP-binding cassette (ABC) antibiotic efflux pump
Accessory genome	<i>AAC6_30_AAC6_Ib</i>	Aminoglycoside	Antibiotic inactivation	Aminoglycoside bifunctional resistance protein
	<i>ANT(6)-Ia</i>	Aminoglycoside	Antibiotic inactivation	ANT(6)
	<i>aad(6)</i>	Aminoglycoside	Antibiotic inactivation	ANT(6)
	<i>APH(3′)-IIIa</i>	Aminoglycoside	Antibiotic inactivation	APH(3′)
	<i>catA8</i>	Phenicol	Antibiotic inactivation	Chloramphenicol acetyltransferase (CAT)
	<i>ErmB</i>	Streptogramin, macrolide, lincosamide, streptogramin B, and streptogramin A	Antibiotic target alteration	Erm 23S ribosomal RNA methyltransferase
	<i>LnuC</i>	Lincosamide	Antibiotic inactivation	Lincosamide nucleotidyltransferase (LNU)
	<i>LnuB</i>	Lincosamide	Antibiotic inactivation	Lincosamide nucleotidyltransferase (LNU)
	<i>LsaE</i>	Lincosamide, pleuromutilin, and streptogramin	Antibiotic target protection	Lsa-type ABC-F protein
	<i>SAT-4</i>	Nucleoside	Antibiotic inactivation	Streptothricin acetyltransferase (SAT)
	<i>DfrF</i>	Diaminopyrimidine	Antibiotic target replacement	Trimethoprim resistant dihydrofolate reductase dfr
	<i>Tet45</i>	Tetracycline	Antibiotic efflux	MFS antibiotic efflux pump
	<i>Tet40</i>	Tetracycline	Antibiotic efflux	MFS antibiotic efflux pump
	<i>TetM</i>	Tetracycline	Antibiotic target protection	Tetracycline-resistant ribosomal protection protein
	<i>TetW</i>	Tetracycline	Antibiotic target protection	Tetracycline-resistant ribosomal protection protein
	<i>TetO</i>	Tetracycline	Antibiotic target protection	Tetracycline-resistant ribosomal protection protein
	<i>tet(O/W/32/O)</i>	Tetracycline	Antibiotic target protection	Tetracycline-resistant ribosomal protection protein

suggest that SS2 may possess natural resistance to fluoroquinolone and glycopeptide antibiotics. Remarkably, despite the single strain showing sensitivity to ofloxacin (S, fluoroquinolones) based on MIC results, it demonstrated resistance to levofloxacin (I, fluoroquinolones), trovafloxacin (S, fluoroquinolones), and vancomycin (S, Glycopeptides; Table 4). The distinct phenotypes of the strain toward these three fluoroquinolones suggest that the resistance mechanism mediated by *patB* may not uniformly affect antibiotics of the same class.

Additionally, bacteria can acquire resistance to a particular class of antibiotics through various mechanisms such as mutation, transformation, and integration of exogenous DNA (Magiorakos et al., 2012; Impey et al., 2020). The CARD annotation for the single strain identified only four classes of resistance genes, and the represented resistant drugs align well with the MIC results. When compared to the CARD annotation results for the individual strain cnzys2-311, the pan-genome shows a more diverse overall resistance profile with a greater variety of resistance mechanisms

(Figure 10). Notably, the tetracycline resistance gene *tet* was carried at high frequencies, consistent with previous reports (Haenni et al., 2018). Therefore, the accessory genome is a major contributor to bacterial acquisition of antibiotic resistance. Our results indicate a close correlation between antibiotic resistance phenotypes and the composition of CARD-related genes, highlighting the accessory genome as a primary source of antibiotic resistance. The acquisition and loss of resistance genes contribute to the challenges in clinical treatment. Antibiotic therapy is currently a major approach for managing *Streptococcus suis* infections in pigs, and this may lead to the widespread dissemination of antimicrobial resistance (AMR) genes, adding pressure to clinical treatment. Thus, it is essential to regulate antibiotic use in clinical practice. These annotation results indicate that SS2's main reservoir of resistance genes comes from the accessory genome. Bacteria can acquire exogenous resistance genes through various means, and the continuous acquisition or loss of resistance genes plays a crucial role in adapting to environmental pressures. This phenomenon is likely a direct cause of the current challenges in clinical treatment.

In this study, various databases and annotation tools were employed, as shown in Table 1. Specifically, eggno-mapper was used for alignment with the COG database, providing annotation results including COG annotations, GO annotations, and KEGG annotations. Considering that the GO descriptions in the COG database might not be the latest version, we conducted a reannotation using the Interproscan database. Additionally, we excluded some descriptions unrelated to bacteria in the KEGG annotations, such as human cancers, to ensure the accuracy of KEGG annotations.

Our analysis has some limitations. Firstly, we are confined to publicly available genomes retrieved from NCBI, including a total of 229 isolates, and the genomes assembled in this study. Lastly, as our analysis is limited to assumed proteins in the database, further experiments are required to validate these presumed genes associated with virulence and drug resistance.

In summary, we have demonstrated the characteristic differences between the core genome and accessory genome of SS2. Additionally, we have identified potential virulence genes associated with SS2, laying the foundation for the exploration of new virulence factors. During the early stages of infection, multiple genes influence the adhesion of the bacterium and may contribute to the pathogenicity of SS2. The presence of diverse virulence factors in SS2 suggests redundancy, possibly serving as advantageous supplements to other virulence factors. Finally, our findings indicate that SS2 may exhibit natural resistance to fluoroquinolone and glycopeptide antibiotics.

## 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.

## Author contributions

YZ: Writing – original draft, Writing – review & editing. TT: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. XY: Data curation, Project administration, Resources, Supervision, Writing – review & editing. YL: Data curation, Investigation, Project administration, Writing – review & editing. ZY: Conceptualization, Data curation, Investigation, Visualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – review & editing. MR: Data curation, Formal analysis, Supervision, Software, Validation, Visualization, Writing – review & editing. GZ: Conceptualization, Software, Visualization, Methodology, Supervision, Validation, Writing – review & editing. YY: Project administration, Funding acquisition, Visualization, Writing – review & editing. AL: Formal analysis, Resources, Supervision, Validation, Visualization, Writing – review & editing. YW: Conceptualization, Funding acquisition, Validation, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was supported by the Sichuan Province Science and Technology Planning Program (2021ZDZX0010 and 2021YJ0270), Provincial Natural Science Foundation of Sichuan (2023NSFSC1216), China Postdoctoral Science Foundation (2022M722300), and Hong Kong Scholars Program 2022, China Postdoctoral Science Foundation (XJ2022047).

## Conflict of interest

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

## Publisher's note

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

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1362316/full#supplementary-material>

## References

- Adams, H. M., Joyce, L. R., Guan, Z., Akins, R. L., and Palmer, K. L. (2017). *Streptococcus mitis* and *S. oralis* lack a requirement for CdsA, the enzyme required for synthesis of major membrane phospholipids in bacteria. *Antimicrob. Agents Chemother.* 61:16. doi: 10.1128/AAC.02552-16
- Alby, K., and Miller, M. B. (2018). *Principles and Practice of Pediatric Infectious Diseases* (Amsterdam: Elsevier), 1467–1478.e1464.
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., et al. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucl. Acids Res.* 51,690–699. doi: 10.1093/nar/gkac920
- Alharthi, S., Alavi, S. E., Moyle, P. M., and Ziora, Z. M. (2021). Sortase A (SrtA) inhibitors as an alternative treatment for superbug infections. *Drug Disc. Tod.* 26, 2164–2172. doi: 10.1016/j.drudis.2021.03.019
- Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* 38, 5825–5829. doi: 10.1093/molbev/msab293
- Chen, P., Liu, R., Huang, M., Zhu, J., Wei, D., Castellino, F. J., et al. (2020). A unique combination of glycoside hydrolases in *Streptococcus suis* specifically and sequentially acts on host-derived  $\alpha$ -Gal-epitope glycans. *J. Biol. Chem.* 295, 10638–10652. doi: 10.1074/jbc.RA119.011977
- Chopra, P., Singh, A., Koul, A., Ramachandran, S., Drlica, K., Tyagi, A. K., et al. (2003). Cytotoxic activity of nucleoside diphosphate kinase secreted from *Mycobacterium tuberculosis*. *Eur. J. Biochem.* 270, 625–634. doi: 10.1046/j.1432-1033.2003.03402.x
- CLSI, I. (2018). *Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals*. Wayne, PA: Clinical and Laboratory Standards Institute.
- Cron, L. E., Bootsma, H. J., Noske, N., Burghout, P., Hammerschmidt, S., and Hermans, P. W. (2009). Surface-associated lipoprotein PpmA of *Streptococcus pneumoniae* is involved in colonization in a strain-specific manner. *Microbiology* 155, 2401–2410. doi: 10.1099/mic.0.026765-0
- Dechêne-Tempier, M., Jouy, E., Bayon-Auboyer, M.-H., Bougeard, S., Chauvin, C., Libante, V., et al. (2023). Antimicrobial resistance profiles of *Streptococcus suis* isolated from pigs, wild boars, and humans in France between 1994 and 2020. *J. Clin. Microbiol.* 61, e00164–e00123. doi: 10.1128/jcm.00164-23
- Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henriessat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucl. Acids Res.* 50,571–577. doi: 10.1093/nar/gkab1045
- El Garch, F., Lismond, A., Piddock, L. J., Courvalin, P., Tulkens, P. M., and Van Bambeke, F. (2010). Fluoroquinolones induce the expression of *patA* and *patB*, which encode ABC efflux pumps in *Streptococcus pneumoniae*. *J. Antimicrob. Chemother.* 65, 2076–2082. doi: 10.1093/jac/dkq287
- Escudero, J. A., San Millan, A., Gutierrez, B., Hidalgo, L., La Ragione, R. M., Abuoun, M., et al. (2011). Fluoroquinolone efflux in *Streptococcus suis* is mediated by SatAB and not by SmrA. *Antimicrob. Agents Chemother.* 55, 5850–5860. doi: 10.1128/AAC.00498-11
- Faulds-Pain, A., Shaw, H. A., Terra, V. S., Kellner, S., Brockmeier, S. L., and Wren, B. W. (2019). The *Streptococcus suis* sortases SrtB and SrtF are essential for disease in pigs. *Microbiology* 165, 163–173. doi: 10.1099/mic.0.000752
- Fefler, A. T., Wang, Y., Burbick, C. R., Diaz-Campos, D., Fajt, V. R., Lawhon, S. D., et al. (2023). Antimicrobial susceptibility testing in veterinary medicine: performance, interpretation of results, best practices and pitfalls. *One Health Adv.* 1:26. doi: 10.1186/s44280-023-00024-w
- Finch, R. G., Greenwood, D., Whitley, R. J., and Norrby, S. R. (2010). *Antibiotic and Chemotherapy e-book*. Amsterdam: Elsevier Health Sciences.
- Fischetti, V., Pancholi, V., and Schneewind, O. (1990). Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. *Mol. Microbiol.* 4, 1603–1605. doi: 10.1111/j.1365-2958.1990.tb02072.x
- Fittipaldi, N., Segura, M., Grenier, D., and Gottschalk, M. (2012). Virulence factors involved in the pathogenesis of the infection caused by the swine pathogen and zoonotic agent *Streptococcus suis*. *Fut. Microbiol.* 7, 259–279. doi: 10.2217/fmb.11.149
- Galán-Bartual, S., Pérez-Dorado, I., García, P., and Hermoso, J. A. (2015). Structure and function of choline-binding proteins. *Streptococcus pneumoniae* 9, 207–230. doi: 10.1016/B978-0-12-410530-0.00011-9
- Griffith, F. (1928). The significance of pneumococcal types. *Epidemiol. Infect.* 27, 113–159. doi: 10.1017/S0022172400031879
- Guo, G., Du, D., Yu, Y., Zhang, Y., Qian, Y., and Zhang, W. (2021). Pan-genome analysis of *Streptococcus suis* serotype 2 revealed genomic diversity among strains of different virulence. *Transbound. Emerg. Dis.* 68, 637–647. doi: 10.1111/tbed.13725
- Haenni, M., Lupo, A., and Madec, J.-Y. (2018). Antimicrobial resistance in *Streptococcus* spp. *Microbiol. Spectr.* 6:2017. doi: 10.1128/microbiolspec.ARBA-0008-2017
- Hamre, A. G., and Sørli, M. (2020). Kinetic relationships with processivity in *Serratia marcescens* family 18 glycoside hydrolases. *Biochem. Biophys. Res. Commun.* 521, 120–124. doi: 10.1016/j.bbrc.2019.10.089
- Hermans, P. W., Adrian, P. V., Albert, C., Estevao, S., Hoogenboezem, T., Luijckendijk, I. H., et al. (2006). The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. *J. Biol. Chem.* 281, 968–976. doi: 10.1074/jbc.M510014200
- Hill, J. E., Gottschalk, M., Brousseau, R., Harel, J., Hemmingsen, S. M., and Goh, S. H. (2005). Biochemical analysis, *cpn60* and 16S rDNA sequence data indicate that *Streptococcus suis* serotypes 32 and 34, isolated from pigs, are *Streptococcus orisratti*. *Vet. Microbiol.* 107, 63–69. doi: 10.1016/j.vetmic.2005.01.003
- Impey, R. E., Hawkins, D. A., Sutton, J. M., and Soares Da Costa, T. P. (2020). Overcoming intrinsic and acquired resistance mechanisms associated with the cell wall of gram-negative bacteria. *Antibiotics* 9:623. doi: 10.3390/antibiotics9090623
- Ji, X., Sun, Y., Liu, J., Zhu, L., Guo, X., Lang, X., et al. (2016). A novel virulence-associated protein, *vapE*, in *Streptococcus suis* serotype 2. *Mol. Med. Rep.* 13, 2871–2877. doi: 10.3892/mmr.2016.4818
- Jiang, F., Guo, J., Cheng, C., and Gu, B. (2020). Human infection caused by *Streptococcus suis* serotype 2 in China: report of two cases and epidemic distribution based on sequence type. *BMC Infect. Dis.* 20, 1–6. doi: 10.1186/s12879-020-4943-x
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., Mcanulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kim, Y., Gu, C., Kim, H. U., and Lee, S. Y. (2020). Current status of pan-genome analysis for pathogenic bacteria. *Curr. Opin. Biotechnol.* 63, 54–62. doi: 10.1016/j.copbio.2019.12.001
- Lane, D. (1991). *16S/23S rRNA Sequencing. Nucleic Acid Techniques in Bacterial Systematics*. John Wiley and sons.
- LeMieux, J., Woody, S., and Camilli, A. (2008). Roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J. Bacteriol.* 190, 6002–6013. doi: 10.1128/JB.00379-08
- Li, Q., Fu, Y., Ma, C., He, Y., Yu, Y., Du, D., et al. (2017). The non-conserved region of MRP is involved in the virulence of *Streptococcus suis* serotype 2. *Virulence* 8, 1274–1289. doi: 10.1080/21505594.2017.1313373
- Li, W., Wan, Y., Tao, Z., Chen, H., and Zhou, R. (2013). A novel fibronectin-binding protein of *Streptococcus suis* serotype 2 contributes to epithelial cell invasion and *in vivo* dissemination. *Vet. Microbiol.* 162, 186–194. doi: 10.1016/j.vetmic.2012.09.004
- Liu, B., Zheng, D., Zhou, S., Chen, L., and Yang, J. (2022). VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucl. Acids Res.* 50,912–917. doi: 10.1093/nar/gkab1107
- Liu, H., Lei, S., Jia, L., Xia, X., Sun, Y., Jiang, H., et al. (2021). *Streptococcus suis* serotype 2 enolase interaction with host brain microvascular endothelial cells and RPSA-induced apoptosis lead to loss of BBB integrity. *Vet. Res.* 52, 1–15. doi: 10.1186/s13567-020-00887-6
- Liu, M., Xia, X., Liu, X., and Kasianenko, O. (2021). Research progress on the pathogenic mechanism of *Streptococcus suis* 2. Scientific messenger of LNU of veterinary medicine and biotechnologies. *Ser. Vet. Sci.* 23, 30–35. doi: 10.32718/nlvvet10405
- Loman, N. J., and Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* 13, 787–794. doi: 10.1038/nrmicro3565
- Magiorakos, A. P., Srinivasan, A., Carey, R. B., Carmeli, Y., Falagas, M., Giske, C., et al. (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* 18, 268–281. doi: 10.1111/j.1469-0691.2011.03570.x
- Marruffini, L. A., Dedent, A. C., and Schneewind, O. (2006). Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 70, 192–221. doi: 10.1128/MMBR.70.1.192-221.2006
- Maturana, J. L., and Cárdenas, J. P. (2021). Insights on the evolutionary genomics of the *Blautia* genus: potential new species and genetic content among lineages. *Front. Microbiol.* 12:660920. doi: 10.3389/fmicb.2021.660920
- Medini, D., Donati, C., Tettelin, H., Massignani, V., and Rappuoli, R. (2005). The microbial pan-genome. *Curr. Opin. Genet. Dev.* 15, 589–594. doi: 10.1016/j.gde.2005.09.006

- Middleton, D. R., Aceil, J., Mustafa, S., Paschall, A. V., and Avci, F. Y. (2021). Glycosyltransferases within the *psrP* locus facilitate pneumococcal virulence. *J. Bacteriol.* 203:20. doi: 10.1128/JB.00389-20
- Mira, A., Martín-Cuadrado, A. B., D'auria, G., and Rodríguez-Valera, F. (2010). The bacterial pan-genome: a new paradigm in microbiology. *Int. Microbiol.* 13, 45–57. doi: 10.2436/20.1501.01.110
- Mishra, N. N., Tran, T. T., Seepersaud, R., García-De-La-Mària, C., Faull, K., Yoon, A., et al. (2017). Perturbations of phosphatidate cytidylyltransferase (CdsA) mediate daptomycin resistance in *Streptococcus mitis/oralis* by a novel mechanism. *Antimicrob. Agents Chemother.* 61:16. doi: 10.1128/AAC.02435-16
- Na, L., Li, R., and Chen, X. (2021). Recent progress in synthesis of carbohydrates with sugar nucleotide-dependent glycosyltransferases. *Curr. Opin. Chem. Biol.* 61, 81–95. doi: 10.1016/j.cbpa.2020.10.007
- Neeld, D., Jin, Y., Bichsel, C., Jia, J., Guo, J., Bai, F., et al. (2014). *Pseudomonas aeruginosa* injects NDK into host cells through a type III secretion system. *Microbiology* 160:1417. doi: 10.1099/mic.0.078139-0
- Nishibori, T., Nishitani, Y., Nomoto, R., and Osawa, R. (2013). Reappraisal of the taxonomy of *Streptococcus suis* serotypes 20, 22, 26, and 33 based on DNA-DNA homology and *sodA* and *recN* phylogenies. *Vet. Microbiol.* 162, 842–849. doi: 10.1016/j.vetmic.2012.11.001
- Olaitan, A. O., Morand, S., and Rolain, J.-M. (2014). Mechanisms of polymyxin resistance: acquired and intrinsic resistance in bacteria. *Front. Microbiol.* 5:643. doi: 10.3389/fmicb.2014.00643
- Orsini, M., Cuccuru, G., Uva, P., and Fotia, G. (2016). Bacterial genomic data analysis in the next-generation sequencing era. *Data Min. Techniq. Life Sci.* 21, 407–422. doi: 10.1007/978-1-4939-3572-7\_21
- Palmieri, C., Varaldo, P. E., and Facinelli, B. (2011). *Streptococcus suis*, an emerging drug-resistant animal and human pathogen. *Front. Microbiol.* 2:235. doi: 10.3389/fmicb.2011.00235
- Pareek, C. S., Smoczynski, R., and Tretyn, A. (2011). Sequencing technologies and genome sequencing. *J. Appl. Genet.* 52, 413–435. doi: 10.1007/s13353-011-0057-x
- Pracht, D., Elm, C., Gerber, J., Bergmann, S., Rohde, M., Seiler, M., et al. (2005). Pava of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation. *Infect. Immun.* 73, 2680–2689. doi: 10.1128/IAI.73.5.2680-2689.2005
- Qian, Y., Zhang, Y., Yu, Y., Li, Q., Guo, G., Fu, Y., et al. (2018). SBP1 is an adhesion-associated factor without the involvement of virulence in *Streptococcus suis* serotype 2. *Microb. Pathog.* 122, 90–97. doi: 10.1016/j.micpath.2018.06.008
- Ragunathan, P., Spellerberg, B., and Ponnuraj, K. (2009). Structure of laminin-binding adhesin (Lmb) from *Streptococcus agalactiae*. *Acta Crystallogr. Sect. D* 65, 1262–1269. doi: 10.1107/S0907444909038359
- Rasko, D. A., Rosovitz, M., Myers, G. S., Mongodin, E. F., Fricke, W. F., Gajer, P., et al. (2008). The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190, 6881–6893. doi: 10.1128/JB.00619-08
- Redman, W. K., Welch, G. S., and Rumbaugh, K. P. (2020). Differential efficacy of glycoside hydrolases to disperse biofilms. *Front. Cell. Infect. Microbiol.* 10:379. doi: 10.3389/fcimb.2020.00379
- Sawasato, K., Sato, R., Nishikawa, H., Iimura, N., Kamemoto, Y., Fujikawa, K., et al. (2019). CdsA is involved in biosynthesis of glycolipid MPLase essential for membrane protein integration *in vivo*. *Sci. Rep.* 9:1372. doi: 10.1038/s41598-018-37809-8
- Schwarz-Linek, U., Werner, J. M., Pickford, A. R., Gurusiddappa, S., Kim, J. H., Pilka, E. S., et al. (2003). Pathogenic bacteria attach to human fibronectin through a tandem  $\beta$ -zipper. *Nature* 423, 177–181. doi: 10.1038/nature01589
- Shryock, T. R. (2002). *Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals: Approved Standard*. Clinical and Laboratory Standards Institute.
- Sirichoat, A., Flórez, A. B., Vázquez, L., Buppasiri, P., Panya, M., Lulitanond, V., et al. (2020). Antibiotic resistance-susceptibility profiles of *Enterococcus faecalis* and *Streptococcus* spp. from the human vagina, and genome analysis of the genetic basis of intrinsic and acquired resistances. *Front. Microbiol.* 11:1438. doi: 10.3389/fmicb.2020.01438
- Spirig, T., Weiner, E. M., and Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* 82, 1044–1059. doi: 10.1111/j.1365-2958.2011.07887.x
- Tang, J., Wang, C., Feng, Y., Yang, W., Song, H., Chen, Z., et al. (2006). Streptococcal toxic shock syndrome caused by *Streptococcus suis* serotype 2. *PLoS Med.* 3:e151. doi: 10.1371/journal.pmed.0030151
- Tenenbaum, T., Spellerberg, B., Adam, R., Vogel, M., Kim, K. S., and Schrotten, H. (2007). *Streptococcus agalactiae* invasion of human brain microvascular endothelial cells is promoted by the laminin-binding protein Lmb. *Microb. Infect.* 9, 714–720. doi: 10.1016/j.micinf.2007.02.015
- Tettelin, H., Massignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc. Natl. Acad. Sci. U. S. A.* 102, 13950–13955. doi: 10.1073/pnas.0506758102
- Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. doi: 10.1016/j.mib.2008.09.006
- Tran, T. T., Mishra, N. N., Seepersaud, R., Diaz, L., Rios, R., Dinh, A. Q., et al. (2019). Mutations in *cdsA* and *pgsA* correlate with daptomycin resistance in *Streptococcus mitis* and *S. oralis*. *Antimicrob. Agents Chemother.* 63:18. doi: 10.1128/AAC.01531-18
- Tribble, G. D., Rigney, T. W., Dao, D. H. V., Wong, C. T., Kerr, J. E., Taylor, B. E., et al. (2012). Natural competence is a major mechanism for horizontal DNA transfer in the oral pathogen *Porphyromonas gingivalis*. *MBio* 3:11. doi: 10.1128/mBio.00231-11
- Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S. Y., et al. (2020). PHI-base: the pathogen-host interactions database. *Nucl. Acids Res.* 48, 613–620. doi: 10.1093/nar/gkz904
- Venegas, F. A., Koutaniemi, S., Langeveld, S. M., Bellemare, A., Chong, S.-L., Dilokpimol, A., et al. (2022). Carbohydrate esterase family 16 contains fungal hemicellulose acetyl esterases (HAEs) with varying specificity. *N. Biotechnol.* 70, 28–38. doi: 10.1016/j.nbt.2022.04.003
- Vernikos, G. (2020). A review of pangenome tools and recent studies. *The Pangenome* 4, 89–112. doi: 10.1007/978-3-030-38281-0\_4
- Vötsch, D., Willenborg, M., Oelemann, W. M., Brogden, G., and Valentin-Weigand, P. (2019). Membrane binding, cellular cholesterol content and resealing capacity contribute to epithelial cell damage induced by suliyisin of *Streptococcus suis*. *Pathogens* 9:33. doi: 10.3390/pathogens9010033
- Wahid, R. M., Yoshinaga, M., Nishi, J., Maeno, N., Sarantuya, J., Ohkawa, T., et al. (2005). Immune response to a laminin-binding protein (Lmb) in group A streptococcal infection. *Pediatr. Int.* 47, 196–202. doi: 10.1111/j.1442-200x.2005.02038.x
- Xia, X., Qin, W., Zhu, H., Wang, X., Jiang, J., and Hu, J. (2019). How *Streptococcus suis* serotype 2 attempts to avoid attack by host immune defenses. *J. Microbiol. Immunol. Infect.* 52, 516–525. doi: 10.1016/j.jmii.2019.03.003
- Yamaguchi, M., Terao, Y., and Kawabata, S. (2013). Pleiotropic virulence factor-*Streptococcus pyogenes* fibronectin-binding proteins. *Cell. Microbiol.* 15, 503–511. doi: 10.1111/cmi.12083
- Yu, H., Rao, X., and Zhang, K. (2017). Nucleoside diphosphate kinase (Ndk): a pleiotropic effector manipulating bacterial virulence and adaptive responses. *Microbiol. Res.* 205, 125–134. doi: 10.1016/j.micres.2017.09.001
- Yu, H., Xiong, J., Zhang, R., Hu, X., Qiu, J., Zhang, D., et al. (2016). Ndk, a novel host-responsive regulator, negatively regulates bacterial virulence through quorum sensing in *Pseudomonas aeruginosa*. *Sci. Rep.* 6:28684. doi: 10.1038/srep28684
- Zhang, G., and Feng, J. (2016). The intrinsic resistance of bacteria. *Yi chuan Hereditas.* 38, 872–880. doi: 10.16288/j.ycz.16-159
- Zhao, J., Pan, S., Lin, L., Fu, L., Yang, C., Xu, Z., et al. (2015). *Streptococcus suis* serotype 2 strains can induce the formation of neutrophil extracellular traps and evade trapping. *FEMS Microbiol. Lett.* 362:fnv022. doi: 10.1093/femsle/fnv022
- Zheng, J. X., Li, Y., Zhang, H., Fan, H. J., and Lu, C. P. (2013). Identification and characterization of a novel hemolysis-related gene in *Streptococcus suis* serotype 2. *PLoS ONE* 8:e74674. doi: 10.1371/journal.pone.0074674
- Zhu, H., He, J., Jing, H., Wang, Z., and Duan, Q. (2006). Isolation and identification of *Streptococcus suis* serotype 2 from sick-pig samples of Sichuan province. *Wei Sheng wu xue bao Acta Microbiologica Sinica* 46, 635–638.
- Zhu, Y., Dong, W., Ma, J., Zhang, Y., Zhong, X., Pan, Z., et al. (2020). Comparative genomic analyses provide clues to capsule switch in *Streptococcus suis*. *bioRxiv* 2020:377622. doi: 10.1101/2020.11.11.377622