



# Comparative Genomic Analysis of *Streptococcus dysgalactiae* subspecies *dysgalactiae* Isolated From Bovine Mastitis in China

Siyu Xu<sup>1</sup>, Yang Liu<sup>1</sup>, Jian Gao<sup>1</sup>, Man Zhou<sup>1</sup>, Jingyue Yang<sup>1</sup>, Fumeng He<sup>1</sup>, John P. Kastelic<sup>2</sup>, Zhaoju Deng<sup>1\*</sup> and Bo Han<sup>1\*</sup>

<sup>1</sup> Department of Clinical Veterinary Medicine, College of Veterinary Medicine, China Agricultural University, Beijing, China, <sup>2</sup> Department of Production Animal Health, Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

## OPEN ACCESS

### Edited by:

Apichai Tuanyok,  
University of Florida, United States

### Reviewed by:

Ben Pascoe,  
University of Bath, United Kingdom  
Sophie Payot,  
INRAE Grand Est-Nancy, France

### \*Correspondence:

Zhaoju Deng  
zhaaju2020@hotmail.com  
Bo Han  
hanbo@cau.edu.cn

### Specialty section:

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

**Received:** 02 August 2021

**Accepted:** 24 September 2021

**Published:** 22 October 2021

### Citation:

Xu S, Liu Y, Gao J, Zhou M, Yang J, He F, Kastelic JP, Deng Z and Han B (2021) Comparative Genomic Analysis of *Streptococcus dysgalactiae* subspecies *dysgalactiae* Isolated From Bovine Mastitis in China. *Front. Microbiol.* 12:751863. doi: 10.3389/fmicb.2021.751863

*Streptococcus dysgalactiae* subsp. *dysgalactiae* (SDSD) is one of the most prevalent pathogens causing bovine mastitis worldwide. However, there is a lack of comprehensive information regarding genetic diversity, complete profiles of virulence factors (VFs), and antimicrobial resistance (AMR) genes for SDSD associated with bovine mastitis in China. In this study, a total of 674 milk samples, including samples from 509 clinical and 165 subclinical mastitis cases, were collected from 17 herds in 7 provinces in China from November 2016 to June 2019. All SDSD isolates were included in phylogenetic analysis based on 16S rRNA and multi-locus sequence typing (MLST). In addition, whole genome sequencing was performed on 12 representative SDSD isolates to screen for VFs and AMR genes and to define pan-, core and accessory genomes. The prevalence of SDSD from mastitis milk samples was 7.57% (51/674). According to phylogenetic analysis based on 16S rRNA, 51 SDSD isolates were divided into 4 clusters, whereas based on MLST, 51 SDSD isolates were identified as 11 sequence types, including 6 registered STs and 5 novel STs (ST521, ST523, ST526, ST527, ST529) that belonged to 2 distinct clonal complexes (CCs) and 4 singletons. Based on WGS information, 108 VFs genes in 12 isolates were determined in 11 categories. In addition, 23 AMR genes were identified in 11 categories. Pan-, core and accessory genomes were composed of 2,663, 1,633 and 699 genes, respectively. These results provided a comprehensive profiles of SDSD virulence and resistance genes as well as phylogenetic relationships among mastitis associated SDSD in North China.

**Keywords:** bovine mastitis, *Streptococcus dysgalactiae* subsp. *dysgalactiae*, whole genome sequencing, virulence factor genes, antimicrobial resistance genes

## INTRODUCTION

*Streptococcus dysgalactiae* is an important bovine mastitis causing pathogen worldwide (Cameron et al., 2016; Gao et al., 2017; de Campos et al., 2021), which could result in economic loss and deteriorated animal welfare (Heikkilä et al., 2018). *Streptococcus dysgalactiae* consists of 2 subspecies: *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) and *Streptococcus dysgalactiae*

subsp. *dysgalactiae* (SDSD). They can be distinguished based on their hemolytic properties [SDSD (a-hemolytic) and SDSE (b-hemolytic)] on blood agar (Alves-Barroco et al., 2021). SDSD is classically described as an animal pathogen, causing animal diseases, such as bovine mastitis (Zhang et al., 2018), and mostly results in persistent (sub)clinical mastitis (Botrel et al., 2010). Meanwhile, studies indicate that SDSD isolates can also be isolated from human with breast cancer and infect human primary keratinocyte cells *in vitro* (Alves-Barroco et al., 2019b; Koh et al., 2020), indicating the potential to infect human. Therefore, SDSD in cow milk should be considered a threat to public health.

The pathogenicity of *Streptococcus dysgalactiae* attributes to various virulence factors (VFs), which involved in the following categories: adherence, enzyme (including hyaluronidase, mitogenic factor, streptococcal enolase, and streptodornase et al.), immune evasion, immunoreactive antigen, iron uptake, manganese uptake, protease, superantigen, and toxin formation process, based on *Streptococcus* virulence factors database (VFDB). These virulence genes facilitate the process of adherence to epithelial cells and internalization as well as the subsequent dissemination into host cells (O'Halloran et al., 2016; Alves-Barroco et al., 2019a). Abuse of antimicrobials in mastitis treatment contributes the most to the spread of antimicrobial resistance (AMR) in mastitis associated SDSD isolates (McDougall et al., 2014; Cameron et al., 2016; Zhang et al., 2018). Specifically, the emergence of multiple-drug resistant strains can be a major challenge in the treatment of bovine mastitis and is a growing concern for public health (McDougall et al., 2014; El Garch et al., 2020).

Population structure of SDSD remains unknown, although various genotypes of *S. dysgalactiae* have been described (Rato et al., 2013; Santos-Sanches et al., 2015). Genetic characteristics of SDSD associated with bovine mastitis, including virulence factors and AMR genes, have been profiled using PCR-based methods (Tian et al., 2019). However, these methods are insufficient to capture the subtle genomic differences among strains. Whole genome sequencing (WGS) with high resolution could be the preferred method to determine strain diversity and to infer phylogenetic relationships as well as to identify virulence and resistance genes in SDSD isolates with subtle differences. Therefore, the objectives of this study were to (1) estimate the apparent prevalence of SDSD in samples from (sub)clinical mastitis; (2) determine genetic diversity and evolution of SDSD isolates by 16s RNA sequencing and whole genome sequencing analysis; (3) identify VFs and AMR genes in those SDSD isolates.

## MATERIALS AND METHODS

### Statement of Ethics

This study was conducted in accordance with ethical guidelines and standard biosecurity and institutional safety procedures of China Agricultural University (CAU; Beijing, China). Prior to the start of the study, ethical approval was granted by the Departmental Committee of the College of Veterinary Medicine, CAU.

## Milk Sample Collection and *Streptococcus dysgalactiae* subsp. *dysgalactiae* Identification

A total of 674 milk samples, including 509 clinical mastitis (CM) and 165 subclinical mastitis (SCM) milk samples, were obtained from 17 large dairy farms (each had > 1,000 lactating cows) in 7 provinces (Tianjin = 19 samples; Shanxi = 15 samples; Hebei = 189 samples; Heilongjiang = 87 samples; Inner Mongolia = 95 samples; Shandong = 259 samples; and Shaanxi = 10 samples) in China from November 2016 to June 2019. The CM samples, including changes in the milk (e.g., clots, discoloration, flakes, and wateriness) and either with or without visible abnormalities of the udder and/or systemic symptoms (e.g., red, swollen, firm or painful udder, or fever), were collected from a single quarter with abnormalities. The SCM samples were defined as intramammary infection without clinical symptoms and could be detected by California mastitis test and somatic cell counts (Godden et al., 2017). All milk samples were collected aseptically, stored in an ice box and subsequently transported to Mastitis Reference Laboratory at the College of Veterinary Medicine, CAU, Beijing, China.

Pathogens were identified according to bacteriological culture, colony morphology, gram staining, biochemical tests and 16S rRNA sequencing (NMC, 2017). In short, an aliquot (50  $\mu$ l) of milk sample was spread on a tryptone soy agar (TSA; Oxoid, Basingstoke, United Kingdom) with 5% defibrinated sheep blood (Land Bridge Technology, Beijing, China) and incubated aerobically at 37°C in a humidified condition for 24 h. Bacterial colony morphology was recorded and samples that yielded  $\geq 3$  morphologically distinct colonies were considered as contaminated and were excluded from the subsequent analysis. Gram staining and catalase testing were conducted to discriminate *Staphylococci* (gram positive, catalase-positive) and *Streptococci-enterococci* group (gram-positive, catalase-negative). Esculin test was used to differentiate esculin-positive cocci and *Streptococci* (esculin-negative). *Streptococci* were identified at subspecies level using 16S rRNA gene sequencing (Beijing Sunbiotech Inc., Beijing, China) and BLAST with sequences deposited in National Center for Biotechnology Information (NCBI). A total of 51 isolates were identified as SDSD and stored in Brain Heart infusion broth (BHI; Aobox, China) containing 25% glycerol and stored at  $-80^{\circ}\text{C}$  for subsequent analysis.

### DNA Extraction

Genomic DNA was extracted with a bacterial genomic DNA extraction kit (CoWin Biosciences, Beijing, China) for each bacterial isolate according to manufacturer's instructions for Gram-positive bacteria. Additional RNase A (4  $\mu$ l, 100 mg/ml) was used to exclude RNA. The extracted DNA was quantified with a NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

### 16S rRNA Phylogenetic Analysis

The 16S rRNA gene sequences of a total of 51 SDSD isolates, along with sequence from the selected reference strain ATCC 43078 (Accession no. NR\_115275 REGION: 11.767) from

GenBank, were edited and aligned using ClustalW multiple sequence alignment algorithm and then phylogenetic tree was constructed via Maximum Likelihood (Tamura-Nei model) with MEGA-X v10.1.8 (Tamura and Nei, 1993). Confidence values for each branch of the phylogenetic tree was estimated using bootstrapping with 1,000 resamplings. Loci with < 95% site coverage, including fewer than 5% alignment gaps, missing data and ambiguous bases at any positions (partial deletion option) (Kumar et al., 2018), were eliminated. The phylogenetic tree was visualized by the iTOL web server<sup>1</sup> (Letunic and Bork, 2019).

## Multi-Locus Sequence Typing

MLST was applied to determine sequence types (STs) of the SDSL isolates based on 7 housekeeping genes, namely *gki*, *gt*, *murI*, *mutS*, *recP*, *xpt*, and *atoB*. Sequence at each locus was assigned with an allele number, and the corresponding combination of the 7 allele numbers for each isolate was submitted to the PubMLST database<sup>2</sup> to obtain the ST of the isolate. In this study, a clonal complex (CC) was defined as a group of STs in which every ST shared at least 5 of 7 identical allele profiles with at least 1 other ST in the group. The minimum spanning tree (MST) was constructed by the geoBURST algorithm and visualized by the PhyloViz web server<sup>3</sup> to infer phylogenetic relationships among STs of original isolates.

## Genome Assembly and Annotation

A total of 10 SDSL isolates (of 3 dominant STs) from clinical mastitis cases and 2 from subclinical mastitis cases were randomly selected for whole genome sequencing. The DNA samples with concentration  $\geq 40$   $\mu\text{g/mL}$  and purity indices  $A_{260}/_{280} \geq 1.8$ ,  $A_{260}/_{230} \geq 2.0$  were submitted for whole genome sequencing with an Illumina HiSeq 4000 system (Illumina, San Diego, CA, United States) at the Beijing Genomics Institute (Shenzhen, China). Genomic DNA was sheared randomly to construct 3 read libraries with lengths of 350 bp by a Bioruptor ultrasonicator (Diagenode, Denville, NJ, United States) and physicochemical methods. The paired-end fragment libraries were sequenced according to the Illumina HiSeq 4000 protocol. Raw reads of low quality from paired-end sequencing (those with consecutive bases covered by fewer than 5 reads) were discarded. Subsequently, the reads were assembled using SOAPdenovo Version 1.05.<sup>4</sup> The total number of reads, sequences, and contigs, genome size, N50 and total length were obtained. Gene prediction was performed on the SDSL genomes assembly by glimmer3<sup>5</sup> with Hidden Markov models. tRNA, rRNA and sRNAs were identified by tRNAscan-SE, RNAmmer, and the Rfam database, respectively. The tandem repeats annotation was obtained using the Tandem Repeat Finder.<sup>6</sup> Prophage regions were predicted using the

PHage Search Tool Enhanced Release (PHASTER) web server<sup>7</sup> (Arndt et al., 2016).

## Whole Genome Analysis Identification of Virulence-Associated Genes and Antimicrobial Resistance Genes

Virulence-associated genes were identified based on the core dataset in VFDB Version 2019-07 (Liu et al., 2019). Antimicrobial resistance genes were identified by comparing the SDSL genomes Comprehensive Antibiotic Research Database (CARD) Version V6. Amino acid sequences of predicted genes were aligned against the proteins in these databases using blastp. A gene was assigned to a virulence or antimicrobial resistance protein by the highest score hit containing a minimum identity of 40%. The *E*-value  $5.3\text{E}-08$  was the highest among those *E*-values. Pseudo genes were excluded from the virulence and AMR genes. Location of the AMR genes were retrieved from MobileElementFinder<sup>8</sup> and ISFinder<sup>9</sup> and PHASTER (see text footnote 7).

## Pan-Genome Analysis

The pan-genome of 12 SDSL isolates was computed with Bacterial Pan Genome Analysis tool (BPGA) Version 1.3, using a USEARCH algorithm to cluster orthologous gene families (Chaudhari et al., 2016). Core genes were defined

<sup>7</sup><http://phaster.ca/>

<sup>8</sup><https://cge.cbs.dtu.dk/services/MobileElementFinder/>

<sup>9</sup><https://www-is.biotoul.fr/blast.php>

**TABLE 1** | *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolates (*n* = 51) recovered from 509 bovine clinical mastitis (CM) and 165 subclinical mastitis (SCM) milk samples collected from 17 large dairy farms in 7 provinces of China.

Province	Farm	Date	Mastitis type	No. samples	No. isolates	Rate (%)
Tianjin	A	2017/09	CM	19	2	10.5
Shanxi	B	2018/01	CM	15	1	6.7
Hebei	C	2018/03	CM	19	8	42.1
Heilongjiang	D	2018/04	CM	26	1	3.8
Hebei	E	2018/04	CM	9	1	11.1
Heilongjiang	F	2018/05	CM	25	1	4.0
Inner Mongolia	G	2018/06	CM	15	2	13.3
Shandong	H	2017/12	CM	9	1	11.1
Hebei	I	2018/06	CM	16	1	6.3
Shaanxi	J	2018/06	CM	10	1	10.0
Heilongjiang	K	2018/09	CM	36	3	8.3
Inner Mongolia	L	2017/06	CM	19	3	15.8
Hebei	M	2017/05	CM	40	2	5.0
Hebei	N	2017/04	CM	25	3	12.0
Inner Mongolia	O	2016/11	CM	61	5	8.2
Shandong	P	2019/03	CM	125	12	9.6
Hebei	Q	2019/06	CM	40	2	5.0
Shandong	P	2019/03	SCM	125	1	0.8
Hebei	Q	2019/06	SCM	40	1	2.5
Total	17	/	/	674	51	7.57

<sup>1</sup><https://itol.embl.de/>

<sup>2</sup><http://pubmlst.org/sdysgalactiae/>

<sup>3</sup><https://online.phyloviz.net/>

<sup>4</sup><https://github.com/aquaskyline/SOAPdenovo2>

<sup>5</sup><http://www.ccb.umd.edu/software/glimmer/>

<sup>6</sup><http://tandem.bu.edu/trf/trf.html>

**TABLE 2 |** *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolate (n = 51) recovered from 509 bovine clinical mastitis (CM) and 165 subclinical mastitis (SCM) milk samples collected from farms in China and their genotypes.

Isolate	Province	Farm	Date	Sample type	Allelic numbers							ST	CC
					<i>gki</i>	<i>gtr</i>	<i>murl</i>	<i>mutS</i>	<i>recP</i>	<i>xpt</i>	<i>atoB</i>		
SDSD_1	Tianjin	A	2017/09	CM	43	43	35	38	50	69	39	ST460	/
SDSD_2	Tianjin	A	2017/09	CM	43	43	35	38	50	69	39	ST460	/
SDSD_4	Shanxi	B	2018/01	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_5	Hebei	C	2018/03	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_6	Hebei	C	2018/03	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_7	Hebei	C	2018/03	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_9	Hebei	C	2018/03	CM	40	43	35	38	50	64	39	ST523	CC1
SDSD_10	Hebei	C	2018/03	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_11	Hebei	C	2018/03	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_12	Hebei	C	2018/03	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_13	Hebei	C	2018/03	CM	39	43	35	55	48	67	40	ST521	CC 2
SDSD_15	Heilongjiang	D	2018/04	CM	40	43	35	38	50	64	39	ST523	CC1
SDSD_16	Hebei	E	2018/04	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_17	Heilongjiang	F	2018/05	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_19	Inner Mongolia	G	2018/06	CM	41	43	35	55	50	64	39	ST529	CC1
SDSD_20	Inner Mongolia	G	2018/06	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_26	Shandong	H	2017/12	CM	41	43	35	34	48	67	8	ST527	/
SDSD_27	Hebei	I	2018/06	CM	40	43	35	34	48	67	40	ST461	CC2
SDSD_28	Shaanxi	J	2018/06	CM	41	43	35	36	50	64	39	ST305	CC1
SDSD_29	Heilongjiang	K	2018/09	CM	41	43	35	55	50	67	39	ST454	CC1
SDSD_30	Heilongjiang	K	2018/09	CM	41	43	35	55	50	67	39	ST454	CC1
SDSD_32	Heilongjiang	K	2018/09	CM	41	43	35	55	50	67	39	ST454	CC1
SDSD_34	Inner Mongolia	L	2017/06	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_35	Inner Mongolia	L	2017/06	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_36	Hebei	M	2017/05	CM	41	43	35	34	48	67	8	ST527	/
SDSD_37	Hebei	M	2017/05	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_38	Hebei	N	2017/04	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_39	Hebei	N	2017/04	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_40	Hebei	N	2017/04	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_41	Inner Mongolia	O	2016/11	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_42	Inner Mongolia	O	2016/11	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_43	Inner Mongolia	O	2016/11	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_44	Inner Mongolia	O	2016/11	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_45	Inner Mongolia	L	2017/06	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_46	Inner Mongolia	O	2016/11	CM	39	43	35	55	48	67	40	ST521	CC2
SDSD_47	Shandong	P	2019/03	CM	40	43	35	38	50	64	39	ST523	CC1
SDSD_48	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_49	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_50	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_51	Shandong	P	2019/03	CM	39	43	35	33	47	64	39	ST298	/
SDSD_52	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_53	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_54	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_55	Shandong	P	2019/03	CM	40	43	35	38	50	64	39	ST523	CC1
SDSD_56	Shandong	P	2019/03	CM	39	43	35	33	47	64	39	ST298	/
SDSD_57	Shandong	P	2019/03	SCM	40	43	35	55	50	69	40	ST526	/
SDSD_58	Shandong	P	2019/03	CM	40	43	35	38	50	64	39	ST523	CC1
SDSD_59	Shandong	P	2019/03	CM	41	43	35	34	50	64	39	ST453	CC1
SDSD_60	Hebei	Q	2019/06	SCM	39	43	35	55	48	67	40	ST521	CC2
SDSD_62	Hebei	Q	2019/06	CM	41	43	35	55	50	67	39	ST454	CC1
SDSD_63	Hebei	Q	2019/06	CM	41	43	35	55	50	67	39	ST454	CC1

as the genes that exist in all the genomes, accessory genes were defined as genes present in  $\geq 1$  genomes but not all the genomes, while unique genes were defined as the genes only found in a single genome, according to BPGA analysis. Functional annotations of core, accessory, and unique genes were obtained after comparing sequences to those present in the clusters of orthologous groups (COGs) of proteins and Kyoto encyclopedia of genes and genomes (KEGG) databases. The phylogenetic tree was constructed based on core genes of 12 SDS D genome sequences, combined with 12 *Streptococci* genome sequences obtained from the National Center for Biotechnology Information (NCBI) (details in **Supplementary Table 1**) using BPGA v1.3.

### Statistical Analyses

Test for proportions was applied to compare the proportion of the genes fall into the functional categories among core, accessory and unique genome using SPSS 23.0 software (SPSS Inc., Chicago, IL, United States) and significance was considered at  $p < 0.05$  in a two-tailed test.

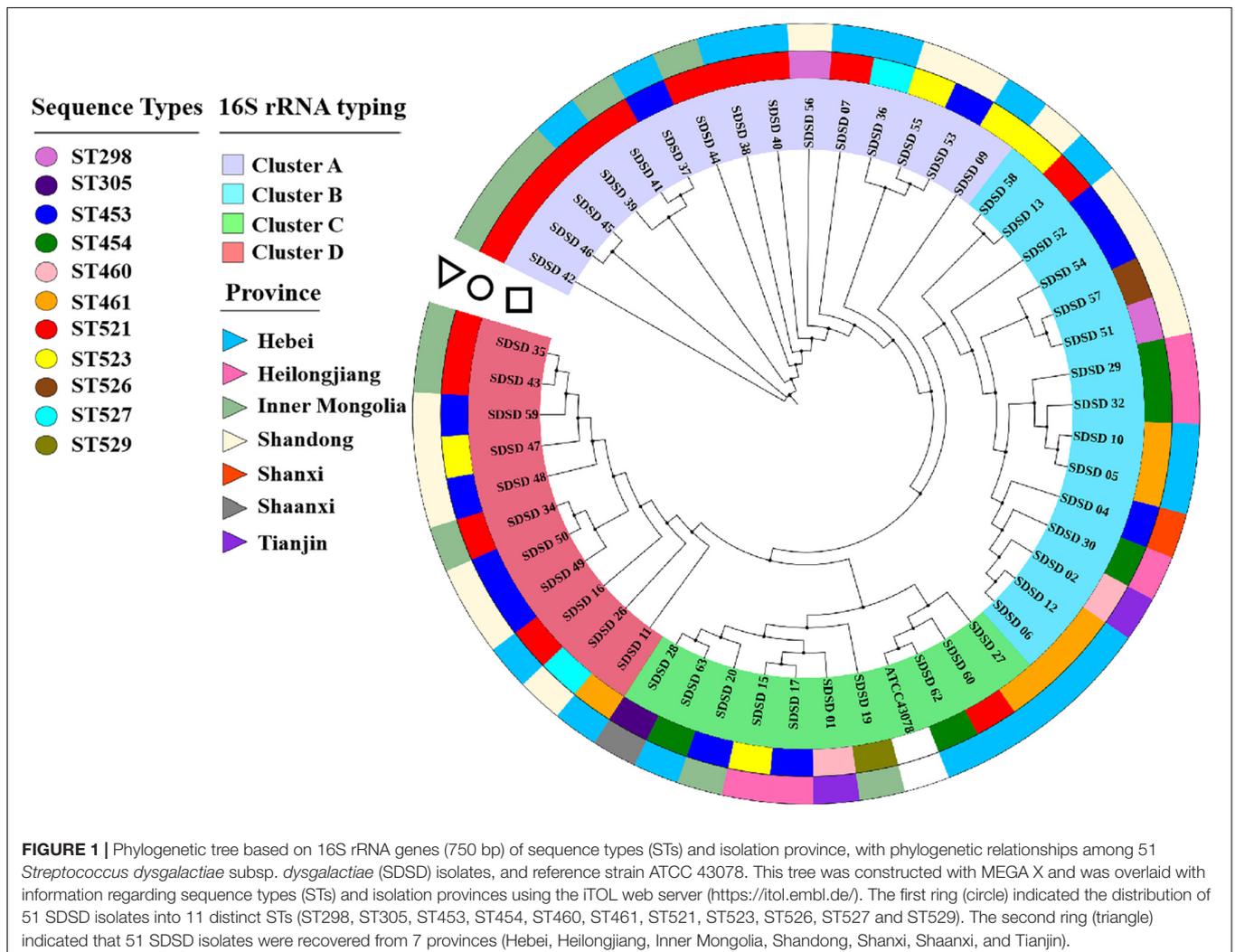
## RESULTS

### Isolates

Detailed information of SDS D isolate is shown in **Tables 1, 2**. A total of 51 (7.57%) isolates were obtained from 674 milk samples, 49 isolates from 509 CM samples and 2 isolates from 165 SCM samples. The apparent prevalence of SDS D in CM and SCM samples was 49/509 (9.6%) and 2/165 (1.2%), respectively.

### Phylogenetic Analysis Based on 16S rRNA

The phylogenetic analysis based on 16S rRNA was conducted to determine phylogenetic relationships among the isolates (**Figure 1**). Four clusters were found: Cluster A consisted of 15 isolates collected from 3 provinces (Hebei, Inner Mongolia and Shandong); Cluster B consisted of 15 isolates from 5 provinces (Hebei, Shandong, Heilongjiang, Shanxi and Tianjin); Cluster C consisted of 10 isolates from 5 provinces (Hebei, Inner Mongolia, Heilongjiang, Shaanxi and Tianjin) and ATCC 43078; and Cluster D consisted of 11 isolates isolated from 3 provinces (Shandong,



Inner Mongolia and Hebei). SDDS 62 was the isolate that most closely related to the reference genome ATCC 43078.

## Multi-Locus Sequence Typing and Minimum-Spanning Tree

A total of 11 distinct STs were identified in the 51 SDDS isolates from 7 provinces in China (Table 2 and Figure 1), 5 of which were novel, namely ST521, ST523, ST526, ST527, and ST529. Of these STs, ST521 (15 isolates from 7 herds in 2 provinces) was the most predominant one, followed by ST453 (11 isolates from 6 herds in 5 provinces) and ST523 (5 isolates from 3 herds in 3 provinces). The phylogenetic analysis and MST (Figure 2) involving 445 STs were performed. The 11 STs were grouped into 2 CCs and 4 singletons. In CC1, ST453 was the main evolutionary starter and 4 STs were involved, whereas ST461 was the main evolutionary starter in CC2.

## Whole Genome Sequencing, Assembly, and Annotation

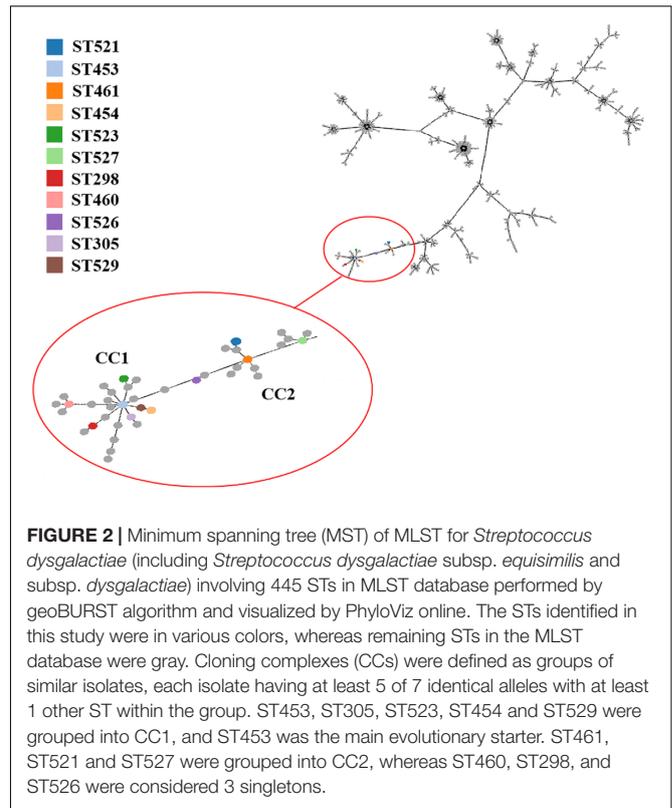
In this study, 12 SDDS genomes were sequenced with Illumina HiSeq 4000 system's protocol. Detailed description on whole genome sequencing, assembly and annotation are shown in Table 3. Seven million reads per isolate and an average genome depth of 59.2 were obtained from the Illumina sequencing. The genome sizes of 12 isolates varied from 2.15 to 2.36 Mb. And the average number of genes identified in all the isolates was 2,042. Furthermore, SDDS47 showed the highest diversity in genes, RNA and prophage regions. Whole genome sequences of the 12 SDDS isolates were deposited in GenBank with the following accession numbers: JAIEZU0000000000 (SDDS04), JAIEZX0000000000 (SDDS09), JAIFRQ0000000000 (SDDS15), JAIEZZ0000000000 (SDDS16), JAIEZY0000000000 (SDDS17), JAIFAA0000000000 (SDDS20), JAIFRR0000000000 (SDDS34), JAIFRS0000000000 (SDDS37), JAIFRT0000000000 (SDDS47), JAIFRU0000000000 (SDDS48); JAIFRV0000000000 (SDDS57); JAIFRW0000000000 (SDDS60).

## Identification of Virulence-Associated Genes

A total of 108 virulence-associated genes were annotated by VFDB in all 12 SDDS strains, SDDS16 had the highest number of VFs ( $n = 114$ ) and SDDS57 had the least ( $n = 107$ ). These VF genes belonged to 11 main virulence categories: adherence, enzyme, immune evasion, immune reactive antigen, iron, manganese uptake, protease, peptidase, superantigen, toxin-related genes and others. The occurrence and distribution of virulence genes are shown in Figures 3A, 4, respectively.

Among the 26 adherence-related genes, 18 genes (*abpB*, *cysB*, *flmH*, *fbpA*, *fnbA*, *groEL*, *gapA*, *gpbB*, *msrA/B*, *plr*, *shr*, *slrA*, *srtA*, *cbpD*, *lmb*, *lap*, *ssp-5*, and *tuf*) were present in all SDDS isolates. In addition, the number of some genes had ST-specific distributions. For example, isolates of ST521 harbored 3 *lmb* and 2 *ssp-5* gene copies, respectively. Whereas isolates of ST453, ST523, and ST526 contained 2 and 1, respectively.

Regarding enzyme-related genes, *dltA*, *eno*, *glnA1*, *hylB*, *luxS*, *lplA1*, *lgt*, *mf*, *relA*, *sak*, *stp*, *sodB*, and *sda* were present in all



**FIGURE 2 |** Minimum spanning tree (MST) of MLST for *Streptococcus dysgalactiae* (including *Streptococcus dysgalactiae* subsp. *equisimilis* and subsp. *dysgalactiae*) involving 445 STs in MLST database performed by geoBURST algorithm and visualized by PhyloViz online. The STs identified in this study were in various colors, whereas remaining STs in the MLST database were gray. Cloning complexes (CCs) were defined as groups of similar isolates, each isolate having at least 5 of 7 identical alleles with at least 1 other ST within the group. ST453, ST305, ST523, ST454 and ST529 were grouped into CC1, and ST453 was the main evolutionary starter. ST461, ST521 and ST527 were grouped into CC2, whereas ST460, ST298, and ST526 were considered 3 singletons.

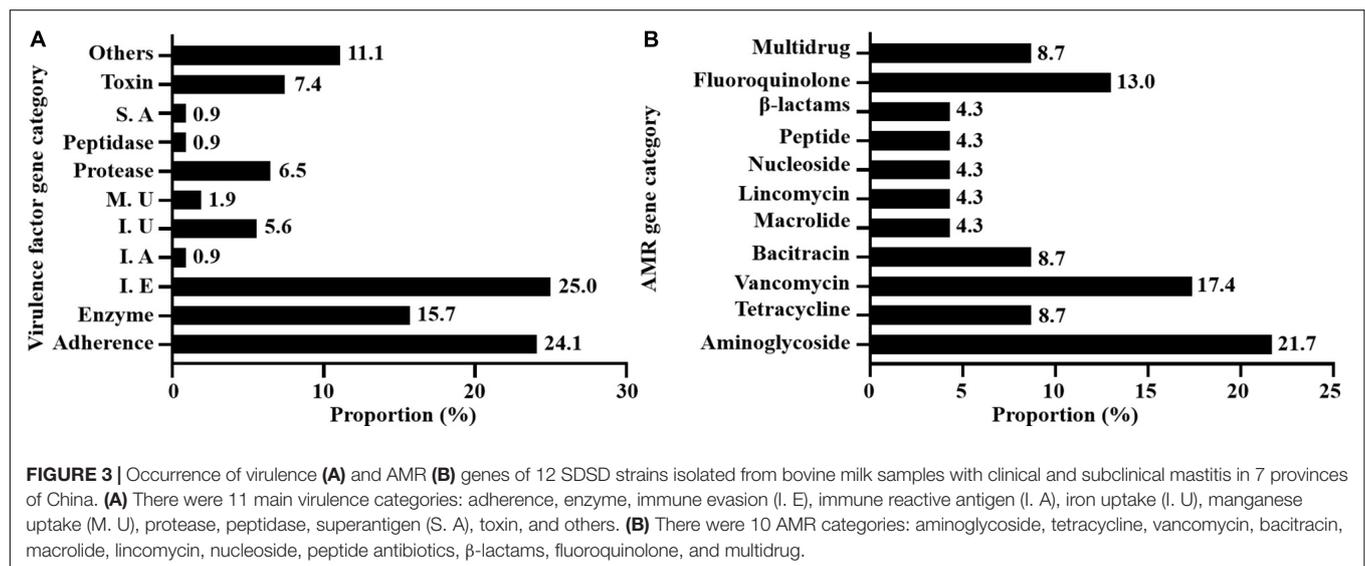
SDDS isolates. There were 27 immune evasion-related genes, including 5 capsular genes (*rgpA* to *G*, except *rgpC*) and another 21 immune evasion-related genes were detected in all isolates. Among the 7 protease-related genes present in all STs, *clpE* had uniform distribution: isolates of ST453 had 3 *clpE* copies, whereas isolates of other STs only had 2. Notably, all isolates carried 4 cytolysin regulator gene *cylR2* copies, which belong to the toxin-related genes category.

## Identification of Antimicrobial Resistance Genes

A total of 23 AMR genes in the genomes of 12 SDDS isolates were identified, the occurrence and distribution of AMR genes are presented in Figures 3B, 5, respectively. Based on gene number, 23 AMR genes varied from 0 to 3 copies in the 12 SDDS isolates. Two isolates of ST453, namely SDDS20 and SDDS37, both had the most abundant 21 AMR genes, belonging to 9 distinct classes of antimicrobials (aminoglycoside, tetracycline, vancomycin, bacitracin, fluoroquinolone, lincomycin, nucleoside, peptide, and  $\beta$ -lactamase). Moreover, all 12 isolates harbored 1 *lmrp* gene which encodes a multidrug resistance efflux pump. We found a total of 4 AMR genes (from 4 isolates in ST453 and ST521) were located at mobile genetic elements, most of which were located in insertion sequences [*ANT6-IA*, *SAT-4* and *APH(3')-IIIA*] and one in transposon [*APH(3')-Ia*]. Meanwhile, two integrative conjugative elements (ICEs) ICETn6009 were identified in SDDS20 and SDDS37 isolates. However, there was no AMR gene carried by ICETn6009 and prophage regions.

**TABLE 3** | Sequencing, assembly statistics and annotation of 12 *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolates.

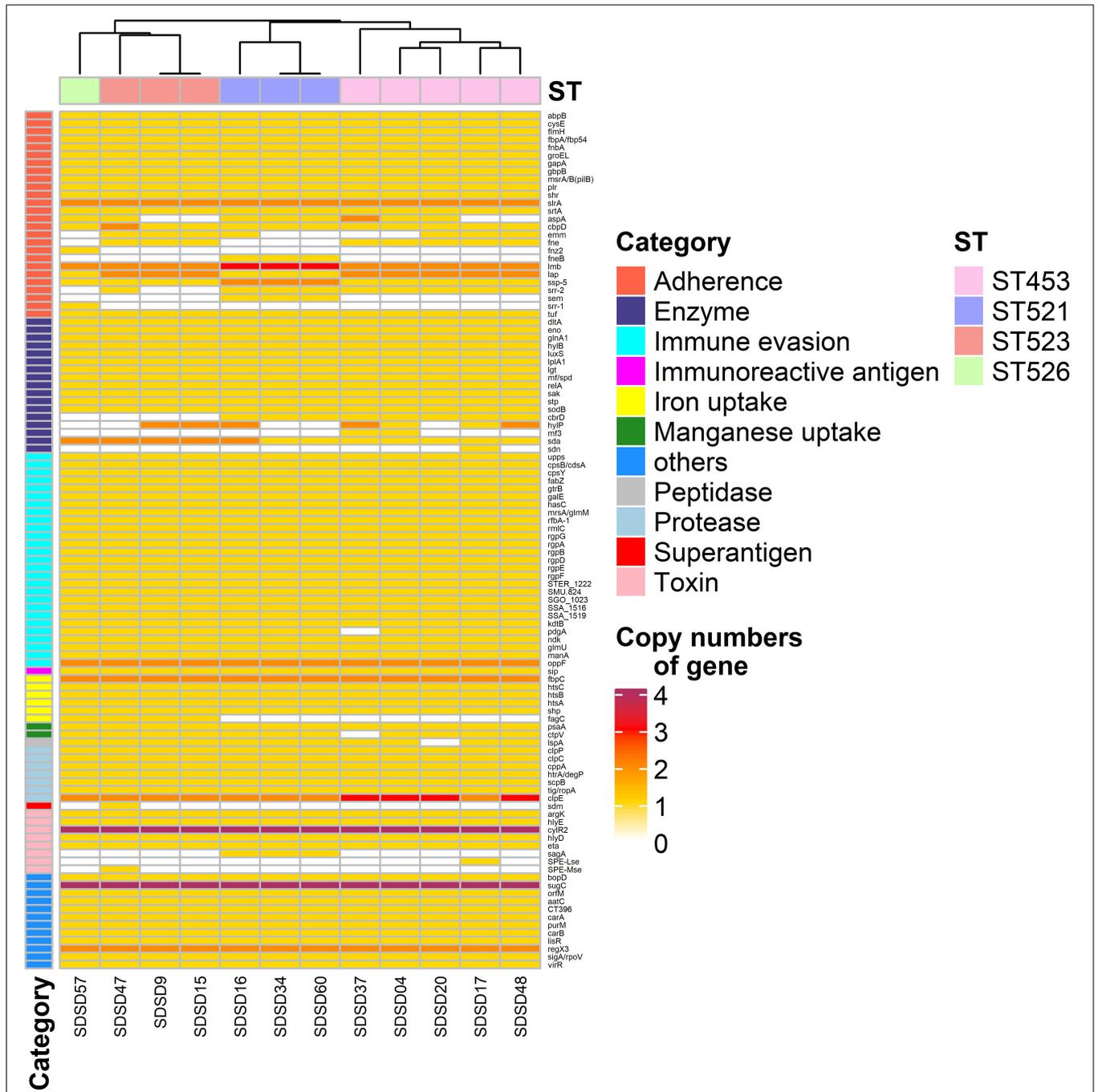
Isolate	Sequencing statistics			Assembly statistics				Annotation					
	Total No. reads ( $\times 1,000$ )	Total no. sequences (Mb)	Average read length (bp)	Genome depth	Genome size (Mb)	No. of contigs	N50 (bp)	Total length (Kb)	rRNA	tRNA	sRNA	Tandem repeats	Prophage regions
SDSD04	7,011	1,051	350	58.63	2.27	77	2,055,641	2,035	3	47	38	57	2
SDSD09	7,011	1,051	350	61.96	2.15	81	2,025,205	2,017	3	45	62	31	5
SDSD15	7,011	1,051	350	60.86	2.19	81	2,025,342	2,017	3	45	62	31	5
SDSD16	7,011	1,051	350	58.65	2.27	66	2,070,894	2,025	3	47	41	41	2
SDSD17	7,011	1,051	350	56.42	2.36	85	2,063,783	2,067	3	47	46	38	2
SDSD20	7,011	1,051	350	57.55	2.32	71	2,059,105	2,009	3	47	53	45	1
SDSD34	7,011	1,051	350	60.88	2.19	56	2,029,365	1,964	3	47	44	40	1
SDSD37	7,011	1,051	350	56.44	2.36	74	2,197,553	2,170	3	47	51	57	3
SDSD47	7,011	1,051	350	59.75	2.23	91	2,097,043	2,099	3	45	62	39	6
SDSD48	7,011	1,051	350	59.77	2.23	74	2,042,841	2,038	3	47	47	37	3
SDSD57	7,011	1,051	350	57.61	2.31	78	2,064,798	2,093	3	47	58	47	4
SDSD60	7,011	1,051	350	61.96	2.15	67	2,034,354	1,971	3	47	45	40	1



### Pan-Genome Analysis

The pan-genome of the 12 SDS D contained 2,663 genes. The core genome (shared by all *S. dysgalactiae* isolates) consisted of 1,633 genes. The accessory genome comprised 699 genes, and the unique genome included 293 genes. The number of core genes was fairly constant at ~1,600 genes, whereas the size of genes in the pan-genome continued to increase as the number of strains increased (Figure 6). Functional annotation of genes in the pan-genome revealed the distribution of functional categories among 3 pan-genome sets (Figure 7). Among which, metabolism was identified as the most abundant functional category in the core genes. The overall percentage of metabolic functions in the core genes was 38.6%, whereas that in the accessory and unique genes were 26.5 and 9.7%, respectively. Metabolism was almost 1.5 and 4 times more enhanced in the core genes compared to accessory and unique genes (Figure 7A). The functional category of information storage and processing had a higher

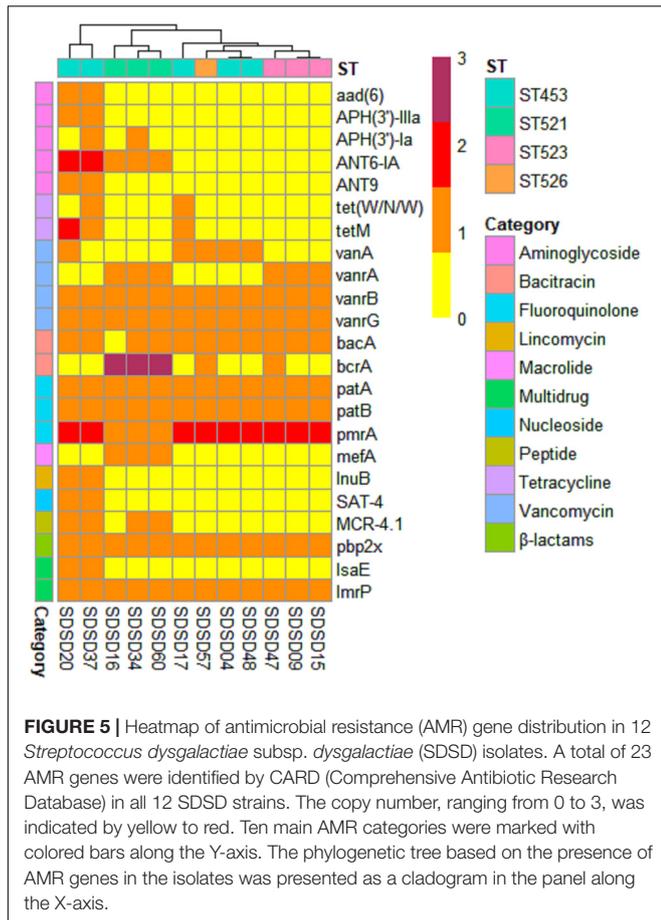
proportion in unique genes than in core and accessory genes. The functions of transcription and replication, recombination and repair were enhanced in unique genes, whereas the functions of translation, ribosomal structure and biogenesis were enhanced in core genes (Figure 7B). According to blasting with KEGG databases, this distinct distribution was observed (Figure 7C). The functional category of metabolism was the most abundant in three pan-genome sets components, with the overall proportion of 65.7, 60.0 and 53.1% in core, accessory and unique genes, respectively. Notably, the function of carbohydrate metabolism was enhanced in accessory genes rather than in core or unique genes. Meanwhile, the functional category of genetic information processing was significantly enhanced in unique genes than in core or accessory genes. Interestingly, the functions of replication and repair were enhanced in unique genes, whereas, the functions of folding, sorting and degradation, transcription, and translation were enhanced in core genes (Figure 7D).



**FIGURE 4 |** Heatmap of virulence factor gene number distribution in 12 *Streptococcus dysgalactiae* subsp. *dysgalactiae* (SDSD) isolates. A total of 108 virulence-associated genes were determined by Virulence Factors of Database (VFDB) in all 12 SDSD strains, and the copy number, ranging from 0 to 4 were indicated by yellow to red. Eleven main virulence categories were marked with colored bars along the Y-axis. The phylogenetic tree based on the presence of virulence genes of the isolates was presented as a cladogram in the panel along the X-axis.

A phylogenetic tree was constructed based on the core genes of 24 *Streptococci* genomes, including 12 SDSD genomes obtained from the Illumina sequencing and 12 obtained from NCBI (Figure 8). The phylogenetic tree revealed that the 24 *Streptococci* were divided into 3 phylogenetic groups. Group A only contained *S. pneumoniae* CGSP14, and group C included

*S. dysgalactiae* strains isolated from human, fish and swine. All SDSD isolates from cattle belonged to group B. Isolates of ST521 were more closely related to SDSD NCTC4670 collected from humans, whereas isolates with ST453, ST523 and ST526 were more closely related to SDSD ATCC27957 and NCTC13731 (both were isolated from dairy cows).



**FIGURE 5 |** Heatmap of antimicrobial resistance (AMR) gene distribution in 12 *Streptococcus dysgalactiae* subsp. *dysgalactiae* (SDSD) isolates. A total of 23 AMR genes were identified by CARD (Comprehensive Antibiotic Research Database) in all 12 SDSD strains. The copy number, ranging from 0 to 3, was indicated by yellow to red. Ten main AMR categories were marked with colored bars along the Y-axis. The phylogenetic tree based on the presence of AMR genes in the isolates was presented as a cladogram in the panel along the X-axis.

## DISCUSSION

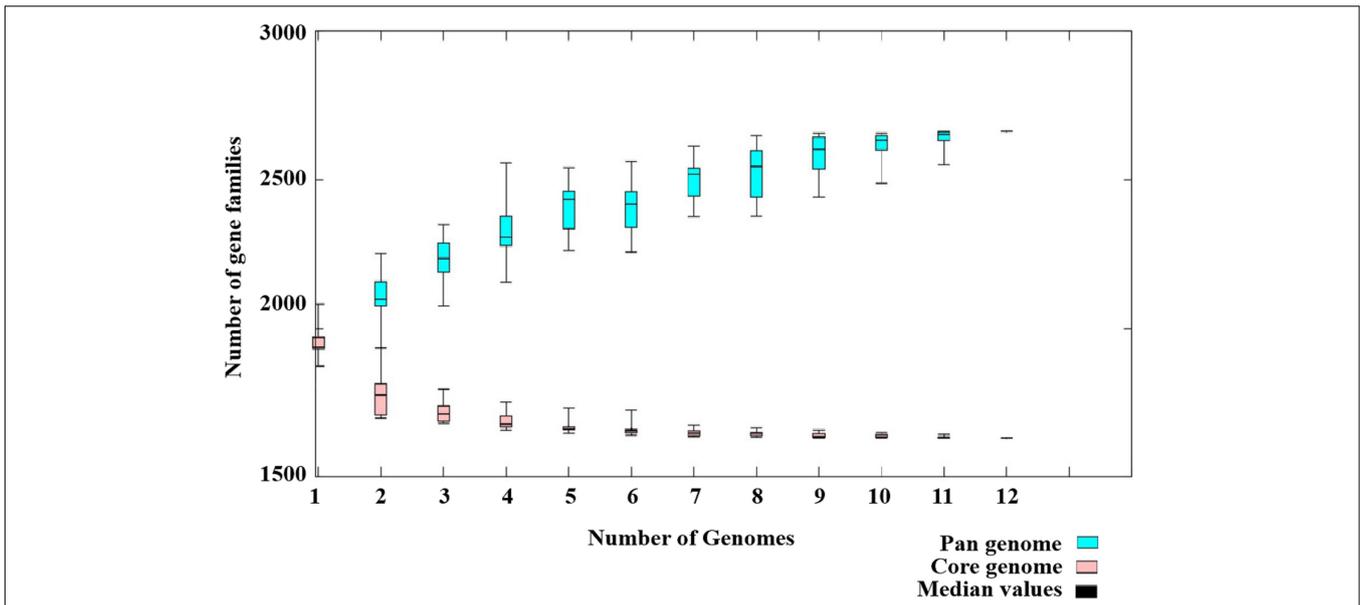
Several studies have described the epidemiology of *Streptococcus dysgalactiae* in dairy herds (Cameron et al., 2016; Horpiencharoen et al., 2019). In addition, studies using PCR-based techniques to identify virulence and AMR genes in bovine *S. dysgalactiae* isolates have been published (Fernandez et al., 2013; Horpiencharoen et al., 2019). However, no in-depth studies have investigated the population structure, profiles of virulence and AMR genes as well as the phylogenetic relationships of bovine mastitis derived SDSD isolates. Therefore, we conducted phylogenetic analysis based on 16S rRNA and MLST of 51 SDSD isolates obtained from 17 dairy herds in 7 provinces of China. In addition, we also conducted WGS of 12 representative SDSD isolates to determine the distribution of virulence and antimicrobial resistance genes. The core genomes were more related to metabolism, and the unique genomes were more relevant to genetic replication and repair according to functional pan-genomes analysis. Consequently, our study suggests that core genes of metabolism can be used as a tool to identify *Streptococcus* species in the microbiome. The phylogenetic tree based on the core genome of 24 *Streptococcus* revealed that human *S. dysgalactiae* isolates have closer genetic relationship with fish and swine *S. dysgalactiae* isolates than bovine SDSD isolates, which is consistent with Alves-Barroco et al. (2021).

The apparent prevalence of SDSD in CM was 9.6%, roughly consistent with the figures in previous studies in China, America, and Europe (Cameron et al., 2016; Soltau et al., 2017; Zhang et al., 2018), but lower than the 28.0% in Australia (Shum et al., 2009). The apparent prevalence of SDSD in SCM samples was 1.2%, lower than other studies (Botrel et al., 2010; Andrea Dieser et al., 2014; Leelahapongsathon et al., 2014).

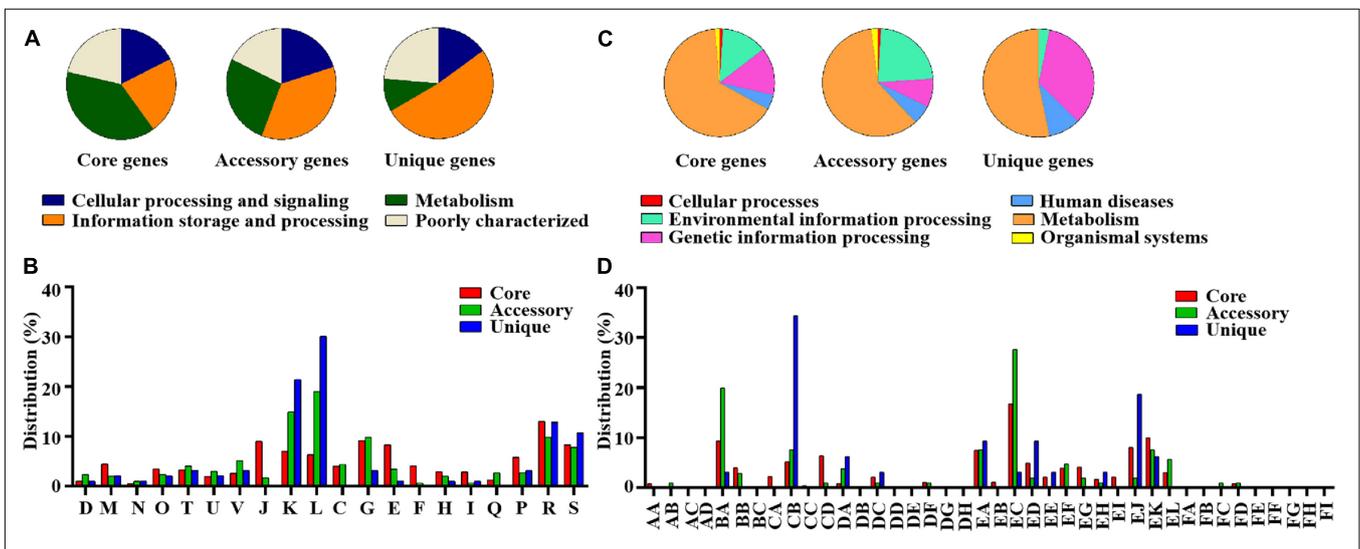
Understanding the phylogenetic relationships among strains is important to determine the transmission of pathogens. The phylogenetic tree based on 16S rRNA sequences revealed 4 phylogenetic clusters. Cluster A-D contained 15, 15, 10, and 11 isolates derived from 3, 5, 5, and 3 provinces, respectively. Some identical STs (ST521, ST453, ST523, ST454, and ST527) were identified in multiple herds in this study, whereas the other STs were only detected in single herds. That identified STs were in different herds, which is consistent with a previous study (Lundberg et al., 2014). Furthermore, movement of infected animals likely promotes spreading of pathogens between herds (Widgren and Frössling, 2010).

Using the whole genome sequences of 12 representative SDSD genomes, we observed differences in the number of predicted prophage regions in 12 SDSD genomes among different strains and STs. A total of 35 prophage regions were detected in 12 SDSD genomes. The average number of prophages per genome was 2.9—the highest number of prophages was presented in ST453 (5.3) and the lowest was 1.3 in ST521. These values suggest that the number of prophage regions in SDSD could possibly be associated with ST/CC, which is in line with Lichvariková et al. (2020).

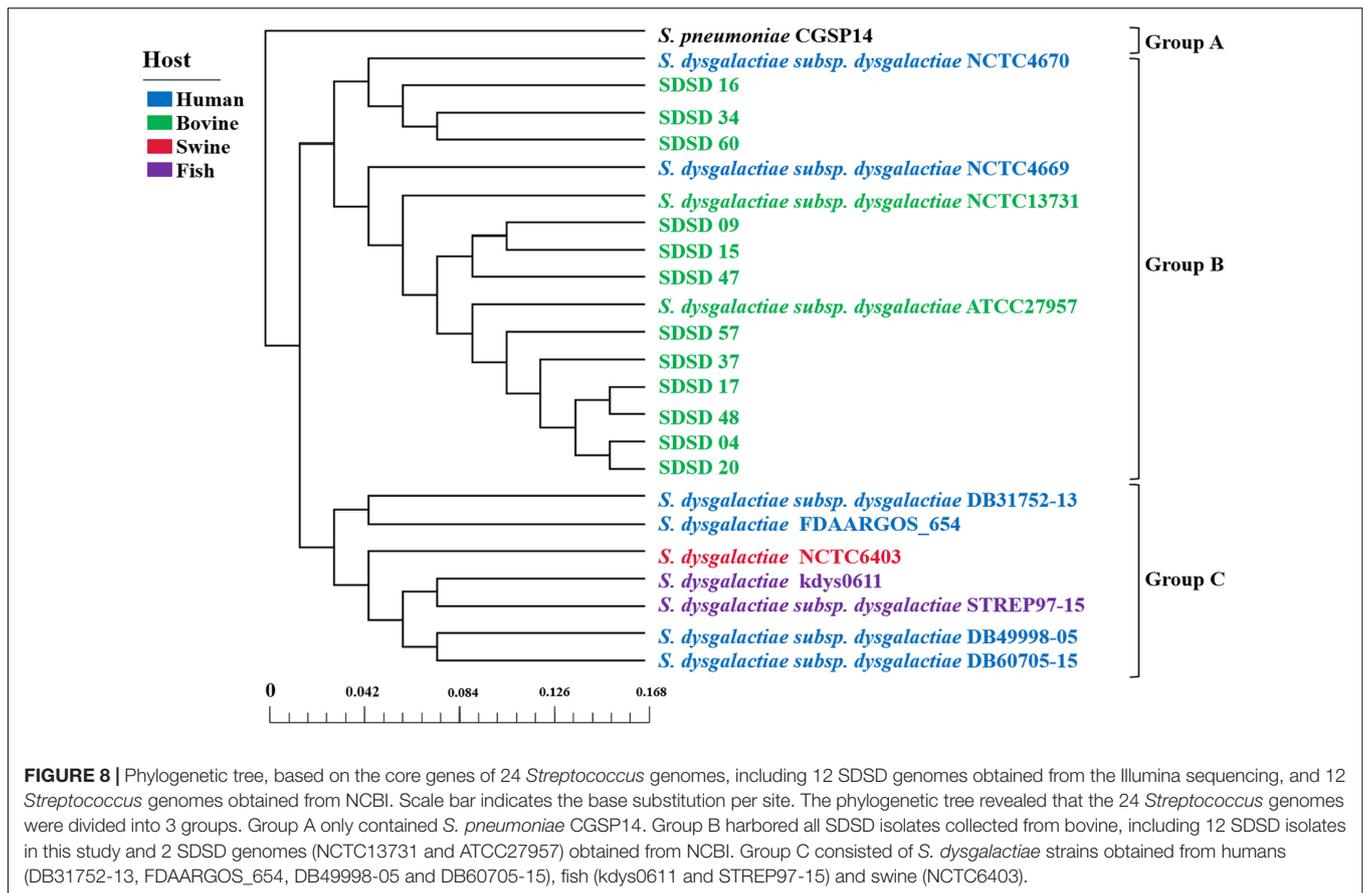
Adherence-related genes could facilitate adhesion and biofilm formation, which are important factors in *Streptococci* pathogenesis (Widgren and Frössling, 2010). Adhesion is the first step in biofilm formation or invasion into host cells, promoting survival of microorganisms in infected tissues and facilitating development of mastitis (Dahesh et al., 2012; Alves-Barroco et al., 2019a). Eleven adhesion-related genes [*abpB* (Huang et al., 2015), *fbpA* (Bao et al., 2014), *fnBA* (Yeswanth et al., 2017), *shr* (Dahesh et al., 2012), *slrA* (Bober et al., 2011), *cbpD* (Roig-Molina et al., 2020), *fnz2* (Yi et al., 2013), *fneB* (Lannergård et al., 2005), *lmb* (Zhang et al., 2014), *seM* (Timoney et al., 2010), *srr-1* (Sheen et al., 2011)] encode a number of surface proteins. These surface proteins are identified as important virulence factors that involve bacterial adhesion to the epithelium of the host cell mediated by microbial surface components recognizing adhesive matrix molecules, consequently contributing to host cell attachment and tissue colonization. *Lmb* was the most abundant adherence related gene in the 12 SDSD isolates. In previous studies, the *lmb* gene had a relatively high prevalence in *Streptococcus uberis* isolated from cattle (Fessia et al., 2019), but was less prevalent in bovine-associated *Streptococcus dysgalactiae* and *Streptococcus agalactiae* (Ding et al., 2016; Tian et al., 2019). Three biofilm-related genes (*srtA*, *aspA*, and *emm*) were identified in 12 SDSD genomes. Biofilms are communities of microorganisms attached to a surface and are involved in chronic and recurrent infections in animals and humans (Veerachamy et al., 2014). Compared to planktonic bacteria, those with biofilms are more resistant to antibiotics (Rabin et al., 2015). Therefore, these genes



**FIGURE 6 |** Pan-genome of 12 *Streptococcus dysgalactiae* subsp. *dysgalactiae* (SDSD) isolates. The pan-genome of 12 SDSD tested in this study had 2,663 genes; the size of the genome in the pan-genome continued to increase as the number of strains increased. However, the number of core genomes (shared by 100% of SDSD isolates) was fairly constant at ~1,600 genes.



**FIGURE 7 |** Differential distribution of COG and KEGG functional categories in core, accessory and unique genes: **(A)** Proportion of 4 classes of COG functional categories in core, accessory and unique genes. **(B)** COG functional sub-categories in core, accessory and unique genes. **(C)** Proportion of 6 classes of KEGG functional categories in core, accessory and unique genes. **(D)** KEGG functional sub-categories in core, accessory and unique genes. D, Cell cycle control, cell division, and chromosome partitioning; M, Cell wall/membrane/envelope biogenesis; N, Cell motility; O, Post-translational modification, protein turnover, and chaperones; T, Signal transduction mechanisms; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; C, Energy production and conversion; G, Carbohydrate transport and metabolism; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; Q, Secondary metabolites biosynthesis, transport, and catabolism; P, Inorganic ion transport and metabolism; R, General function prediction only; S, Function unknown. AA, Cell growth and death; AB, Cell motility; AC, Cellular community; AD, Transport and catabolism; BA, Membrane transport; BB, Signal transduction; BC, Signaling molecules and interaction; CA, Folding, sorting and degradation; CB, Replication and repair; CC, Transcription; CD, Translation; DA, Cancers; DB, Cardiovascular diseases; DC, Drug resistance; DD, Endocrine and metabolic diseases; DE, Immune diseases; DF, Infectious diseases; DG, Neurodegenerative diseases; DH, Substance dependence; EA, Amino acid metabolism; EB, Biosynthesis of other secondary metabolites; EC, Carbohydrate metabolism; ED, Energy metabolism; EE, Glycan biosynthesis and metabolism; EF, Lipid metabolism; EG, Metabolism of cofactors and vitamins; EH, Metabolism of other amino acids; EI, Metabolism of terpenoids and polyketides; EJ, Nucleotide metabolism; EK, overview; EL, Xenobiotics biodegradation and metabolism; FA, Circulatory system; FB, Development; FC, Digestive system; FD, Endocrine system; FE, Environmental adaptation; FF, Excretory system; FG, Immune system; FH, Nervous system; FI, Sensory system.



may contribute to the persistent infection induced by SDSL (Kaczorek et al., 2017).

Enzyme (*dltA*, *eno*, *sda*, *hylB*, and *mf/spd*), immune evasion (*hasC*, *rgpA* to *G*, except *rgpC*, *oppF* etc.), immune reactive antigen (*sip*), iron and manganese uptake (*shp*, *htsA* to *C*), protease (*htrA*, *clp*), and peptidase (*lspA*) genes were widely distributed among all SDSL. These genes enable *S. dysgalactiae* to colonize (Roy et al., 2019; Diaz-Dinamarca et al., 2020), spread (Florindo et al., 2018) and survive in host tissues (De et al., 2017; Roy et al., 2019) and cause infection (Calonzi et al., 2020). *HylB* was the most abundant enzyme-related gene, and this high prevalence was consistent with previous studies (Ding et al., 2016; Calonzi et al., 2020; Lin et al., 2021). Hyaluronidase encoded by *hylB* promoted intracellular survival of *Streptococci* and expression of pro-inflammatory cytokines, indicating it may have a key role in pathogenesis of *Streptococci* mastitis, including SDSL (Wang et al., 2014).

A total of 23 AMR genes were identified with various copy numbers in 12 SDSL isolates according to the annotation of CARD. According to previous surveys, 5 categories of antimicrobials, including  $\beta$ -lactams (penicillin G, cefalexin, and ceftriaxone), macrolide (erythromycin), aminoglycoside (kanamycin and streptomycin), and tetracyclines, are frequently used in treatment of bovine mastitis on Chinese dairy farms (Shi et al., 2010; Zhang et al., 2018). The emergence of AMR genes are mainly through spontaneous mutations and acquisition of AMR

genes from other bacteria (MacLean and Millan, 2019). The high prevalence and breadth diversity of AMR genes, which encode proteins conferring resistance to the 5 categories of antibiotics above, might be attributed to abuse of antimicrobials in Chinese dairy herds. In this study, only one integrative conjugative element *ICETn6009* was detected in SDSL20 and SDSL37, both of which belonged to ST453 and harbored the most abundant AMR genes. *ICETn6009*, a Tn916 based element with Gram-positive *mer* operon and directly linked to the *tet(M)* gene, was identified from two Gram-positive and three Gram-negative genera, including *Streptococcus* (Soge et al., 2008). Meanwhile, 4 AMR genes were located in insertion sequences from 4 isolates in ST453 and ST521 and one in transposon from SDSL 34 in ST521. This may indicate isolates of these two STs were more capable in acquisition of antimicrobial resistance genes from MGEs (Oppegaard et al., 2020).

In general, the average identity values were relatively high for the AMR genes relevant to specific antibiotics: Aminoglycoside [*ANT6-IA*, *APH(3')-IIIa*, *APH(3')-Ia*; 100.00%], Peptide (*MCR-4.1*; 100.00%), Lincomycin (*InuB*; 100.00%), and Nucleoside (*SAT-4*; 99.43%). These high identity values indicate that these genes were highly likely to be AMR genes (Velez et al., 2017). Four antimicrobial resistance mechanisms are involve in the AMR genes identified in 12 SDSL genomes, including antibiotic inactivation [*aad(6)*, *APH(3')-IIIa*, *APH(3')-Ia*, *ant6-IA*, *aad9-IB*, *InuB*, and *SAT4*], antibiotic target alteration

(*vanA*, *vanRA*, *vanRB*, *vanRG*, *bacA*, *MCR-4.1*, and *pbp2X*), antibiotic target protection [*tet(W/N/W)*, *tetM*, *mefA* and *IsaE*], and antibiotic efflux (*bcrA*, *patA*, *patB*, and *pmrA*) (Christaki et al., 2020). Aminoglycoside phosphotransferase encoded by *APH(3')-IIIa* in gram-positive bacteria could phosphorylate the aminoglycoside agents, such as amikacin, neomycin, and kanamycin (Woegerbauer et al., 2014). And *lnuB*, encoding a 3-lincosamide-O-nucleotidyltransferase which could inactivate lincosamides by adenylation in position 3, was reported in *S. agalactiae* (Haenni et al., 2018). Mutation in *pbp2x* could reduce the affinity of the bacteria to  $\beta$ -lactam antibiotics, which has been considered as the major mechanism of *Streptococcus* drug resistance (McDougall et al., 2020; Vannice et al., 2020). *TetM* directly interacts and alters the conformation of nucleotide within 16S rRNA that comprises part of the tetracycline binding site, leading to tetracycline dissociation from the ribosome and prevents rebinding (Dönhöfer et al., 2012). *PatA* and *patB* work as heterodimers to interact together to make a functional drug efflux transporter, conferring efflux-mediated fluoroquinolone resistance in *Streptococcus* (Boncoeur et al., 2012; Baylay et al., 2015). However, the expression of AMR gene is regulated by other components, such as membrane sensor protein (Depardieu et al., 2007). For example, component regulatory systems *bceRS* take part in the regulation of bacitracin resistance in *Streptococcus* (Ouyang et al., 2010). Therefore, AMR phenotypes should be determined to further disclose the relationship between the phenotype and the genotype of SDS in respect to AMR.

## CONCLUSION

To the best of our knowledge, this is the first report on characterization of *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolated from mastitis cows in China using whole-genome sequencing. The apparent prevalence of SDS was estimated at 7.6%. Eleven sequence types (ST298, ST305, ST453, ST454, ST460, ST461, ST521, ST523, ST526, ST527, and ST529) were determined according to MLST. A total of 108 VFs in 11 categories (adherence, enzyme, immune evasion, immune reactive antigen, iron and manganese uptake, protease, peptidase, superantigen, toxin, and others) and 23 AMR genes in 11 categories (aminoglycoside, tetracycline, vancomycin, bacitracin, fluoroquinolone, lincomycin, nucleoside, peptide, macrolide,  $\beta$ -lactams antimicrobials, and multidrug) were identified according to whole genome sequence analysis. These results can be used to study bovine-associated SDS virulence and resistance

## REFERENCES

- Alves-Barroco, C., Caço, J., Roma-Rodrigues, C., Fernandes, A. R., Bexiga, R., Oliveira, M., et al. (2021). New insights on *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolates. *Front. Microbiol.* 12:686413. doi: 10.3389/fmicb.2021.686413
- Alves-Barroco, C., Roma-Rodrigues, C., Raposo, L. R., Bras, C., Diniz, M., Caco, J., et al. (2019b). *Streptococcus dysgalactiae* subsp. *dysgalactiae* isolated from milk of the bovine udder as emerging pathogens: in vitro and in vivo infection of human cells and zebrafish as biological models. *MicrobiologyOpen* 8:e6231. doi: 10.1002/mbio.3.623

genotypes, as well as to better understand the phylogenetic relationships among mastitis-derived SDS in Northern China.

## 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

This study was conducted in accordance with ethical guidelines and standard biosecurity and institutional safety procedures of China Agricultural University (CAU; Beijing, China). Prior to the start of the study, ethical approval was granted by the Departmental Committee of the College of Veterinary Medicine, CAU.

## AUTHOR CONTRIBUTIONS

BH conceived and designed the experiment. SX, YL, MZ, JY, FH, JG, and ZD performed the experiments, analyzed the data, and contributed to drafting the manuscript. SX, JG, JK, ZD, and BH wrote, reviewed, and edited the manuscript. All authors approved the final version of the manuscript.

## FUNDING

This study was supported financially by the National Natural Science Foundation of China (Nos. 31772813 and 31850410474), the Ningxia Key R&D Project (No. 2019BBF02027), the Hebei Key R&D Project (19226607D), the National Key R&D Project (No. 2016YFD0501203), and the High-end Foreign Experts Recruitment Program (No. GDT20171100013).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.751863/full#supplementary-material>

- Alves-Barroco, C., Roma-Rodrigues, C., Balasubramanian, N., Guimaraes, M. A., Ferreira-Carvalho, B. T., Muthukumar, J., et al. (2019a). Biofilm development and computational screening for new putative inhibitors of a homolog of the regulatory protein BrpA in *Streptococcus dysgalactiae* subsp. *dysgalactiae*. *Int. J. Med. Microbiol.* 309, 169–181. doi: 10.1016/j.ijmm.2019.02.001
- Andrea Dieser, S., Vissio, C., Cristina Lasagno, M., Ines Bogno, C., Jose Larriestra, A., and Monica Odierno, L. (2014). Prevalence of pathogens causing subclinical mastitis in argentinean dairy herds. *Pak. Vet. J.* 34, 124–126.
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387

- Bao, Y., Liang, Z., Booyjzsen, C., Mayfield, J. A., Li, Y., Lee, S. W., et al. (2014). Unique genomic arrangements in an invasive serotype M23 strain of *Streptococcus pyogenes* identify genes that induce hypervirulence. *J. Bacteriol.* 196, 4089–4102. doi: 10.1128/JB.02131-14
- Baylay, A. J., Ivens, A., and Piddock, L. J. V. (2015). A novel gene amplification causes upregulation of the PatAB ABC transporter and fluoroquinolone resistance in *Streptococcus pneumoniae*. *Antimicrob. Agents Ch.* 59, 3098–3108. doi: 10.1128/AAC.04858-14
- Bober, M., Mörgelin, M., Olin, A. I., von Pawel-Rammingen, U., and Collin, M. (2011). The membrane bound LRR lipoprotein Slr, and the cell wall-anchored M1 protein from *Streptococcus pyogenes* both interact with type I collagen. *PLoS One* 6:e20345. doi: 10.1371/journal.pone.0020345
- Boncoeur, E., Durmort, C., Bernay, B., Ebel, C., Di Guilmi, A. M., Croizé, J., et al. (2012). PatA and PatB form a functional heterodimeric ABC multidrug efflux transporter responsible for the resistance of *Streptococcus pneumoniae* to fluoroquinolones. *Biochemistry-US* 51, 7755–7765. doi: 10.1021/bi300762p
- Botrel, M. A., Haenni, M., Morignat, E., Sulpice, P., Madec, J. Y., and Calavas, D. (2010). Distribution and antimicrobial resistance of clinical and subclinical mastitis pathogens in dairy cows in Rhone-Alpes. *France. Foodborne Pathog. Dis.* 7, 479–487. doi: 10.1089/fpd.2009.0425
- Calonzi, D., Romano, A., Monistero, V., Moroni, P., Luini, M. V., Biscarini, F., et al. (2020). Technical note: development of multiplex PCR assays for the molecular characterization of *Streptococcus uberis* strains isolated from bovine mastitis. *J. Dairy Sci.* 103, 915–921. doi: 10.3168/jds.2019-16823
- Cameron, M., Saab, M., Heider, L., McClure, J. T., Rodriguez-Lecompte, J. C., and Sanchez, J. (2016). Antimicrobial susceptibility patterns of environmental *Streptococci* recovered from bovine milk samples in the maritime provinces of Canada. *Front. Vet. Sci.* 3:79. doi: 10.3389/fvets.2016.00079
- Chaudhari, N. M., Gupta, V. K., and Dutta, C. (2016). BPGA- an ultra-fast pan-genome analysis pipeline. *Sci. Rep.* 6:24373. doi: 10.1038/srep24373
- Christaki, E., Marcou, M., and Tofarides, A. (2020). Antimicrobial resistance in bacteria: mechanisms, evolution, and persistence. *J. Mol. Evol.* 88, 26–40. doi: 10.1007/s00239-019-09914-9913
- Dahesh, S., Nizet, V., and Cole, J. N. (2012). Study of streptococcal hemoprotein receptor (Shr) in iron acquisition and virulence of M1T1 group a *Streptococcus*. *Virulence* 3, 566–575. doi: 10.4161/viru.21933
- De, A., Liao, S., Bitoun, J. P., Roth, R., Beatty, W. L., Wu, H., et al. (2017). Deficiency of RgpG causes major defects in cell division and biofilm formation, and deficiency of LytR-A-Psr family proteins leads to accumulation of cell wall antigens in culture medium by *Streptococcus mutans*. *Appl. Environ. Microb.* 83:e0092817. doi: 10.1128/AEM.00928-17
- de Campos, J. L., Kates, A., Steinberger, A., Sethi, A., Suen, G., Shutske, J., et al. (2021). Quantification of antimicrobial usage in adult cows and preweaned calves on 40 large Wisconsin dairy farms using dose-based and mass-based metrics. *J. Dairy Sci.* 104, 4727–4745. doi: 10.3168/jds.2020-19315
- Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., and Courvalin, P. (2007). Modes and modulations of antibiotic resistance gene expression. *Clin. Microbiol. Rev.* 20, 79–114. doi: 10.1128/CMR.00015-06
- Diaz-Dinamarca, D. A., Hernandez, C., Escobar, D. F., Soto, D. A., Muñoz, G. A., Badilla, J. F., et al. (2020). Mucosal vaccination with *Lactococcus lactis*-Secreting surface immunological protein induces humoral and cellular immune protection against group B *Streptococcus* in a murine model. *Vaccines (Basel)* 8:146. doi: 10.3390/vaccines8020146
- Ding, Y., Zhao, J., He, X., Li, M., Guan, H., Zhang, Z., et al. (2016). Antimicrobial resistance and virulence-related genes of *Streptococcus* obtained from dairy cows with mastitis in inner Mongolia. *China. Pharm. Biol.* 54, 162–167. doi: 10.3109/13880209.2015.1025290
- Dönhöfer, A., Franckenberg, S., Wickles, S., Berninghausen, O., Beckmann, R., and Wilson, D. N. (2012). Structural basis for TetM-mediated tetracycline resistance. *Proc. Natl. Acad. Sci. U S A.* 109, 16900–16905. doi: 10.1073/pnas.1208037109
- El Garch, F., Youala, M., Simjee, S., Moyaert, H., Klee, R., Truszkowska, B., et al. (2020). Antimicrobial susceptibility of nine udder pathogens recovered from bovine clinical mastitis milk in Europe 2015–2016: VetPath results. *Vet. Microbiol.* 245:1086440. doi: 10.1016/j.vetmic.2020.108644
- Fernandez, G., Barreal, M. L., Pombo, M. B., Ginzo-Villamayor, M. J., Gonzalez-Manteiga, W., Prieto, A., et al. (2013). Comparison of the epidemiological behavior of mastitis pathogens by applying time-series analysis in results of milk samples submitted for microbiological examination. *Vet. Res. Commun.* 37, 259–267. doi: 10.1007/s11259-013-9570-1
- Fessia, A. S., Dieser, S. A., Raspanti, C. G., and Odierno, L. M. (2019). Genotyping and study of adherence-related genes of *Streptococcus uberis* isolates from bovine mastitis. *Microb. Pathog.* 130, 295–301. doi: 10.1016/j.micpath.2019.03.027
- Florindo, C., Barroco, C. A., Silvestre, I., Damião, V., Gomes, J. P., Spellerberg, B., et al. (2018). Capsular type, sequence type and microbial resistance factors impact on DNase activity of *Streptococcus agalactiae* strains from human and bovine origin. *Eur. J. Microbiol. Immunol. (Bp)* 8, 149–154. doi: 10.1556/1886.2018.00026
- Gao, J., Barkema, H. W., Zhang, L., Liu, G., Deng, Z., Cai, L., et al. (2017). Incidence of clinical mastitis and distribution of pathogens on large Chinese dairy farms. *J. Dairy Sci.* 100, 4797–4806. doi: 10.3168/jds.2016-12334
- Godden, S. M., Royster, E., Timmerman, J., Rapnicki, P., and Green, H. (2017). Evaluation of an automated milk leukocyte differential test and the California Mastitis Test for detecting intramammary infection in early- and late-lactation quarters and cows. *J. Dairy Sci.* 100, 6527–6544. doi: 10.3168/jds.2017-12548
- Haenni, M., Lupo, A., Madec, J., Aarestrup, F. M., Schwarz, S., Shen, J., et al. (2018). Antimicrobial Resistance in *Streptococcus spp.* *Microbiol Spectr.* 6, 2–6. doi: 10.1128/microbiolspec.ARBA-0008-2017
- Heikkilä, A. M., Liski, E., Pyörälä, S., and Taponen, S. (2018). Pathogen-specific production losses in bovine mastitis. *J. Dairy Sci.* 101, 9493–9504. doi: 10.3168/jds.2018-14824
- Horpiencharoen, W., Thongratsakul, S., and Poolkhet, C. (2019). Risk factors of clinical mastitis and antimicrobial susceptibility test results of mastitis milk from dairy cattle in western Thailand: bayesian network analysis. *Prev. Vet. Med.* 164, 49–55. doi: 10.1016/j.prevetmed.2019.01.014
- Huang, K., Yuan, Z., Li, J., Zhang, Q., Xu, Z., Yan, S., et al. (2015). Identification and characterization a surface-associated arginine peptidase in *Streptococcus suis* serotype 2. *Microbiol. Res.* 170, 168–176. doi: 10.1016/j.micres.2014.08.001
- Kaczorek, E., Małaczewska, J., Wójcik, R., and Siwicki, A. K. (2017). Biofilm production and other virulence factors in *Streptococcus spp.* isolated from clinical cases of bovine mastitis in Poland. *BMC Vet. Res.* 13:398. doi: 10.1186/s12917-017-1322-y
- Koh, T. H., Binte, A. R. N., and Sessions, O. M. (2020). Comparative genomic analysis of *Streptococcus dysgalactiae* subspecies *dysgalactiae*, an occasional cause of zoonotic infection. *Pathology* 52, 262–266. doi: 10.1016/j.pathol.2019.09.016
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Lannergård, J., Flock, M., Johansson, S., Flock, J. I., and Guss, B. (2005). Studies of fibronectin-binding proteins of *Streptococcus equi*. *Infect. Immun.* 73, 7243–7251. doi: 10.1128/IAI.73.11.7243-7251.2005
- Leelahapongsathon, K., Schukken, Y. H., and Suriyasathaporn, W. (2014). Quarter, cow, and farm risk factors for intramammary infections with major pathogens relative to minor pathogens in Thai dairy cows. *Trop. Anim. Health Prod.* 46, 1067–1078. doi: 10.1007/s11250-014-0603-8
- Letunic, I., and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Lichvariková, A., Soltys, K., Szemes, T., Slobodnikova, L., Bukovska, G., Turna, J., et al. (2020). Characterization of clinical and carrier *Streptococcus agalactiae* and prophage contribution to the strain variability. *Viruses* 12:1323. doi: 10.3390/v12111323
- Lin, L., Huang, X., Yang, H., He, Y., He, X., Huang, J., et al. (2021). Molecular epidemiology, antimicrobial activity, and virulence gene clustering of *Streptococcus agalactiae* isolated from dairy cattle with mastitis in China. *J. Dairy Sci.* 104, 4893–4903. doi: 10.3168/jds.2020-19139
- Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47, D687–D692. doi: 10.1093/nar/gky1080
- Lundberg, A., Nyman, A., Unnerstad, H. E., and Waller, K. P. (2014). Prevalence of bacterial genotypes and outcome of bovine clinical mastitis due to *Streptococcus dysgalactiae* and *Streptococcus uberis*. *Acta Vet. Scand.* 56:80. doi: 10.1186/s13028-014-0080-0

- MacLean, R., and Millan, A. (2019). The evolution of antibiotic resistance clinically relevant evolution studies are needed to help fight the spread of antibiotic resistance. *Science* 365, 1082–1083. doi: 10.1126/science.aax3879
- McDougall, S., Clausen, L., Ha, H. J., Gibson, I., Bryan, M., Hadjirin, N., et al. (2020). Mechanisms of  $\beta$ -lactam resistance of *Streptococcus uberis* isolated from bovine mastitis cases. *Vet. Microbiol.* 242:108592. doi: 10.1016/j.vetmic.2020.108592
- McDougall, S., Hussein, H., and Petrovski, K. (2014). Antimicrobial resistance in *Staphylococcus aureus*, *Streptococcus uberis* and *Streptococcus dysgalactiae* from dairy cows with mastitis. *N. Z. Vet. J.* 62, 68–76. doi: 10.1080/00480169.2013.843135
- NMC (2017). *Laboratory Handbook on Bovine Mastitis*. Madison, WI: National Mastitis Council.
- O'Halloran, F., Beecher, C., Chaurin, V., Sweeney, T., and Giblin, L. (2016). Lactoferrin affects the adherence and invasion of *Streptococcus dysgalactiae* ssp. *dysgalactiae* in mammary epithelial cells. *J. Dairy Sci.* 99, 4619–4628. doi: 10.3168/jds.2015-10465
- Oppegaard, O., Skrede, S., Mylvaganam, H., and Kittang, B. R. (2020). Emerging threat of antimicrobial resistance in  $\beta$ -Hemolytic streptococci. *Front. Microbiol.* 11:797. doi: 10.3389/fmicb.2020.00797
- Ouyang, J., Tian, X., Versey, J., Wishart, A., and Li, Y. (2010). The BceABRS four-component system regulates the bacitracin-induced cell envelope stress response in *Streptococcus mutans*. *Antimicrob. Agents Chemother.* 54, 3895–3906. doi: 10.1128/AAC.01802-1809
- Rabin, N., Zheng, Y., Opoku-Temeng, C., Du, Y., Bonsu, E., and Sintim, H. O. (2015). Biofilm formation mechanisms and targets for developing antibiofilm agents. *Future Med. Chem.* 7, 493–512. doi: 10.4155/fmc.15.6
- Rato, M. G., Bexiga, R., Florindo, C., Cavaco, L. M., Vilela, C. L., and Santos-Sanches, I. (2013). Antimicrobial resistance and molecular epidemiology of *Streptococci* from bovine mastitis. *Vet. Microbiol.* 161, 286–294. doi: 10.1016/j.vetmic.2012.07.043
- Roig-Molina, E., Sánchez-Angulo, M., Seele, J., García-Asencio, F., Nau, R., Sanz, J. M., et al. (2020). Searching for antipneumococcal targets: choline-binding modules as phagocytosis enhancers. *ACS Infect. Dis.* 6, 954–974. doi: 10.1021/acinfecdis.9b00344
- Roy, S., Zhu, Y., Ma, J., Roy, A. C., Zhang, Y., Zhong, X., et al. (2019). Role of ClpX and ClpP in *Streptococcus suis* serotype 2 stress tolerance and virulence. *Microbiol. Res.* 223–225, 99–109. doi: 10.1016/j.micres.2019.04.003
- Santos-Sanches, I., Chambel, L., and Tenreiro, R. (2015). Pulsed-field gel electrophoresis (PFGE): application in population structure studies of bovine mastitis-causing *Streptococci*. *Methods Mol. Biol.* 1247, 323–334. doi: 10.1007/978-1-4939-2004-4\_23
- Sheen, T. R., Jimenez, A., Wang, N. Y., Banerjee, A., van Sorge, N. M., and Doran, K. S. (2011). Serine-rich repeat proteins and pili promote *Streptococcus agalactiae* colonization of the vaginal tract. *J. Bacteriol.* 193, 6834–6842. doi: 10.1128/JB.00094-11
- Shi, D., Hao, Y., Zhang, A., Wulan, B., and Fan, X. (2010). Antimicrobial resistance of *Staphylococcus aureus* isolated from bovine mastitis in China. *Transbound. Emerg. Dis.* 57, 221–224. doi: 10.1111/j.1865-1682.2010.01139.x
- Shum, L. W. C., McConnel, C. S., Gunn, A. A., and House, J. K. (2009). Environmental mastitis in intensive high-producing dairy herds in New South Wales. *Aust. Vet. J.* 87, 469–475. doi: 10.1111/j.1751-0813.2009.00523.x
- Soge, O. O., Beck, N. K., White, T. M., No, D. B., and Roberts, M. C. (2008). A novel transposon, Tn6009, composed of a Tn916 element linked with a *Staphylococcus aureus* mer operon. *J. Antimicrob. Chemother.* 62, 674–680. doi: 10.1093/jac/dkn255
- Soltan, J. B., Einax, E., Klengel, K., Katholm, J., Failing, K., Wehrend, A., et al. (2017). Within-herd prevalence thresholds for herd-level detection of mastitis pathogens using multiplex real-time PCR in bulk tank milk samples. *J. Dairy Sci.* 100, 8287–8295. doi: 10.3168/jds.2016-12385
- Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526. doi: 10.1093/oxfordjournals.molbev.a040023
- Tian, X. Y., Zheng, N., Han, R. W., Ho, H., Wang, J., Wang, Y. T., et al. (2019). Antimicrobial resistance and virulence genes of *Streptococcus* isolated from dairy cows with mastitis in China. *Microb. Pathogenesis.* 131, 33–39. doi: 10.1016/j.micpath.2019.03.035
- Timoney, J. F., DeNegri, R., Sheoran, A., and Forster, N. (2010). Affects of N-terminal variation in the SeM protein of *Streptococcus equi* on antibody and fibrinogen binding. *Vaccine.* 28, 1522–1527. doi: 10.1016/j.vaccine.2009.11.064
- Vannice, K. S., Ricaldi, J., Nanduri, S., Fang, F. C., Lynch, J. B., Bryson-Cahn, C., et al. (2020). *Streptococcus pyogenes* pbp2x mutation confers reduced susceptibility to  $\beta$ -Lactam antibiotics. *Clin. Infect. Dis.* 71, 201–204. doi: 10.1093/cid/ciz1000
- Veerachamy, S., Yarlagadda, T., Manivasagam, G., and Yarlagadda, P. K. (2014). Bacterial adherence and biofilm formation on medical implants: a review. *Proc. Inst. Mech. Eng. H* 228, 1083–1099. doi: 10.1177/0954411914556137
- Velez, J. R., Cameron, M., Rodriguez-Lecompte, J. C., Xia, F., Heider, L. C., Saab, M., et al. (2017). Whole-Genome sequence analysis of antimicrobial resistance genes in *Streptococcus uberis* and *Streptococcus dysgalactiae* isolates from canadian dairy herds. *Front. Vet. Sci.* 4:63. doi: 10.3389/fvets.2017.00063
- Wang, Z., Guo, C., Xu, Y., Liu, G., Lu, C., and Liu, Y. (2014). Two novel functions of hyaluronidase from *Streptococcus agalactiae* are enhanced intracellular survival and inhibition of proinflammatory cytokine expression. *Infect. Immun.* 82, 2615–2625. doi: 10.1128/IAI.00022-14
- Widgren, S., and Frössling, J. (2010). Spatio-temporal evaluation of cattle trade in Sweden: description of a grid network visualization technique. *Geospat. Health* 5, 119–130. doi: 10.4081/gh.2010.192
- Woegerbauer, M., Zeininger, J., Springer, B., Hufnagl, P., Indra, A., Korschneck, I., et al. (2014). Prevalence of the aminoglycoside phosphotransferase genes aph(3')-IIIa and aph(3')-IIa in *Escherichia coli*, *Enterococcus faecalis*, *Enterococcus faecium*, *Pseudomonas aeruginosa*, *Salmonella enterica* subsp. *enterica* and *Staphylococcus aureus* isolates in Austria. *J. Med. Microbiol.* 63(Pt 2), 210–217. doi: 10.1099/jmm.0.065789-0
- Yeswanth, S., Chaudhury, A., and Sarma, P. V. G. K. (2017). Quantitative expression analysis of SpA, FnbA and rsp genes in *Staphylococcus aureus*: actively associated in the formation of biofilms. *Curr. Microbiol.* 74, 1394–1403. doi: 10.1007/s00284-017-1331-x
- Yi, L., Wang, Y., Ma, Z., Zhang, H., Li, Y., Zheng, J. X., et al. (2013). Contribution of fibronectin-binding protein to pathogenesis of *Streptococcus equi* ssp. *Zoonotic Pathog. Dis.* 67, 174–183. doi: 10.1111/2049-632X.12029
- Zhang, S., Piepers, S., Shan, R., Cai, L., Mao, S., Zou, J., et al. (2018). Phenotypic and genotypic characterization of antimicrobial resistance profiles in *Streptococcus dysgalactiae* isolated from bovine clinical mastitis in 5 provinces of China. *J. Dairy Sci.* 101, 3344–3355. doi: 10.3168/jds.2017-14031
- Zhang, Y., Shao, Z., Wang, J., Wang, L., Li, X., Wang, C., et al. (2014). Prevalent distribution and conservation of *Streptococcus suis* Lmb protein and its protective capacity against the Chinese highly virulent strain infection. *Microbiol. Res.* 169, 395–401. doi: 10.1016/j.micres.2013.09.007

**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.

Copyright © 2021 Xu, Liu, Gao, Zhou, Yang, He, Kastelic, Deng and Han. 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.