New Insights on Streptococcus dysgalactiae subsp. dysgalactiae Isolates

Cinthia Alves-Barroco, João Caço, Catarina Roma-Rodrigues, Alexandra R. Fernandes, Ricardo Bexiga, Manuela Oliveira, Lélia Chambel, Rogério Tenreiro, Rosario Mato and Ilda Santos-Sanches

Streptococcus dysgalactiae subsp. dysgalactiae (SDSD) has been considered a strict animal pathogen. Nevertheless, the recent reports of human infections suggest a niche expansion for this subspecies, which may be a consequence of the virulence gene acquisition that increases its pathogenicity. Previous studies reported the presence of virulence genes of Streptococcus pyogenes phages among bovine SDSD (collected in 2002–2003); however, the identity of these mobile genetic elements remains to be clarified. Thus, this study aimed to characterize the SDSD isolates collected in 2011–2013 and compare them with SDSD isolates collected in 2002–2003 and pyogenic streptococcus genomes available at the National Center for Biotechnology Information (NCBI) database, including human SDSD and S. dysgalactiae subsp. equisimilis (SDSE) strains to track temporal shifts on bovine SDSD genotypes. The very close genetic relationships between humans SDSD and SDSE were evident from the analysis of housekeeping genes, while bovine SDSD isolates seem more divergent. The results showed that all bovine SDSD harbor Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas IIA system. The widespread presence of this system among bovine SDSD isolates, high conservation of repeat sequences, and the polymorphism observed in spacer can be considered indicators of the system activity. Overall, comparative analysis shows that bovine SDSD isolates carry speK, speC, speL, speM, spd1, and sdn virulence genes of S. pyogenes prophages. Our data suggest that these genes are maintained over time and seem to be exclusively a property of bovine SDSD strains. Although the bovine SDSD genomes characterized in the present study were not sequenced, the data set, including the high homology of superantigens (SAgs) genes between bovine SDSD and S. pyogenes strains, may indicate that events of horizontal genetic transfer occurred before habitat separation. All bovine SDSD isolates were negative for genes of operon encoding streptolysin S, except for sagA gene, while the presence of this operon was detected in all SDSE and human SDSD strains. The data set of this study suggests that the separation between the subspecies...
“dysgalactiae” and “equisimilis” should be reconsidered. However, a study including the most comprehensive collection of strains from different environments would be required for definitive conclusions regarding the two taxa.

Keywords: Streptococcus dysgalactiae subsp. dysgalactiae, CRISPR typing, phylogenetic relationships, virulence genes, phage genes

INTRODUCTION

Streptococcus dysgalactiae (SD) includes two subspecies: S. dysgalactiae subsp. dysgalactiae (SDSD) and S. dysgalactiae subsp. equisimilis (SDSE) (Jensen and Kilian, 2012). While SDSE has been recognized as an increasingly important human pathogen, SDSD has been considered a strictly animal pathogen and is commonly associated with bovine mastitis (Zadoks et al., 2011; Rato et al., 2013). In the last years, although rare, human diseases by this subspecies have been reported (Koh et al., 2009; Park et al., 2012; Jordal et al., 2015; Chennapragada et al., 2018), and it is urgent to understand the role of this subspecies in human pathogenesis. Recently, we have reported bovine SDSD isolates’ ability to adhere to and internalize different human cell lines, including keratinocytes (Roma-Rodrigues et al., 2016; Alves-Barroco et al., 2018)

SDSD and SDSE, together with Streptococcus agalactiae, Streptococcus canis, and Streptococcus pyogenes, are the major pathogens of the pyogenic group. Previous studies reveal that these species are closely related. Phylogenetic analysis involving the Streptococcus core genome within the pyogenic clade displayed overall consensus. Though the issue is not entirely clear-cut, the most likely scenario is that the species tree has S. dysgalactiae and S. pyogenes as sister species and, together with S. canis, establishes a very closely related branch. These data suggest that they have recently descended from a common ancestor (Suzuki et al., 2011; Jensen and Kilian, 2012; Lefébure et al., 2012). The divergence among species has been mainly associated with mobile genetic elements (MGEs) (Lefébure et al., 2012). Indeed, about a quarter of streptococci’s genome consists of exogenous and mobile DNA, e.g., transposons, prophage, and plasmid. Thus, horizontal gene transfer (HGT) is considered the primary mechanism for generating diversity in this genus (Baiano and Barnes, 2009; Haenni et al., 2010; De la Casa-Esperón, 2012; Richards et al., 2012; Wong and Yuen, 2012; Rohde and Cleary, 2016). Overall, the MGEs are self-transmissible elements that are omnipresent in bacteria, harboring genes responsible for their maintenance, regulation, and mobility and genes encoding antimicrobial resistance, as well as other determinants, such as genes that encode virulence factors (Grohmann et al., 2003; McNeilly and McMillan, 2014).

Thus, many MGEs serve as shuttles for genes that are beneficial to bacteria during their proliferation in a host environment. Although several mechanisms can mediate genetic transfer in pyogenic streptococci, the transduction might be particularly relevant since bacteriophages have been detected in a considerable proportion (McShan and Nguyen, 2016). In S. pyogenes and SDSE, more than 20% of the accessory genomes consist of prophage DNA, highlighting the critical role of transduction in the evolution of these bacteria (Lier et al., 2015; Yamada et al., 2019).

The presence of phage-carrying S. pyogenes virulence genes among SDSD isolates from bovine mastitis was previously reported, namely, the streptococcal pyrogenic exotoxin genes (speK, speC, speL, and speM), the phage DNase (spd1), and determinants of antibiotic resistance located on MGEs (Rato et al., 2011; Abdelsalam et al., 2015). It was hypothesized that these virulence genes might contribute to the increased pathogenic potential; thus, SDSD should not be ignored as a potential human pathogen.

Despite the benefits of HGT, due to the vast diversity of the bacteriophages in the environment, the bacteria are under constant phage attack. To survive, the bacteria have devised various strategies that involve innate and adaptive immunity and interfere with the infection process (Barrangou and Marraffini, 2014). Among all of these strategies, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas system is the one that provides the bacteria with the ability to “memorize” a piece of the invaders’ genetic material (Bhaya et al., 2011; Bondy-Denomy et al., 2013; Rath et al., 2015; Makarova et al., 2015; Makarova et al., 2018) by acquiring foreign nucleic acid fragments called spacers. Together with host Cas proteins, the spacers behave as an adaptive immune system in bacteria and archaea. Besides, the sequential order of the spacers in the array provides chronological information about the bacteria’s exposure to foreign nucleic acids (Lier et al., 2015), providing insights into host–phage interactions within a specific niche. These data show indicators of the evolution of the bacteria that carry them (Lier et al., 2015).

To date, studies on the presence of S. pyogenes phages among SDSD isolates are scarce. For example, it is not yet clear whether these phages (and other MGEs) contribute to more pathogenic SDSE phenotypes. The CRISPR/Cas system in the SDSD and its impact on HGT are examples of how little is known, so far, about this microorganism.

In the present study, SDSD isolates collected in two different periods (2002–2003, partially studied, and 2011–2013) were characterized and compared to obtain new information about molecular virulence profiles, antibiotypes, and population structure. We also investigated the CRISPR/Cas systems in SDSD strains to gain insights into defense systems and their impact on HGT. Additionally, pyogenic streptococcus genomes available at National Center for Biotechnology Information (NCBI) database, including human SDSD and SDSE strains, were included in this study to characterize CRISPR/Cas system, prophage regions, giving further insights on horizontal acquisitions of genetic elements particularly associated with resistance traits or virulence to track temporal shifts of SDSD genotypes.
MATERIALS AND METHODS

Bacterial Isolates
Out of a total of 194 Gram-positive cocci, 37 SDSD were isolated and included in this study (collection II). The isolates were collected from 2011 to 2013, from 250 animals with mastitis of 12 dairy farms located in Portugal. The presumptive identification of the SDSD isolates was performed by traditional phenotypic tests such as colony morphology, type of hemolysis, Lancefield serologic groups using the SLIDEX Strepto Plus (Biomérieux, Marcy-l’Étoile, France), and absence of catalase.

For a comparative analysis, 18 strains of bovine SDSD (collection I), partially characterized (Rato et al., 2010, 2011, 2013) and collected between 2002 and 2003 from 248 animals with mastitis on eight farms located in Portugal, were included in the study.

Genomic DNA Extraction
Genomic DNA was extracted according to Alves-Barroco et al. (2019).

Identification by 16S rRNA Sequence
The identification of the SDSD isolates was confirmed by analysis of 16S rRNA gene sequencing. Amplification and sequencing were performed with the same primers (Takahashi et al., 1997). The sequencing of PCR products was carried out by STAB-Vida (Lisbon, Portugal). The sequences were analyzed and aligned using CLC Main Workbench 20.1 alignment program editor (QIAGEN, Venlo, Netherlands) and then compared with sequences from the NCBI database using the Basic Local Alignment Search Tool (BLAST)\(^1\). Subsequently, the nucleotide sequences were subjected to phylogenetic analysis using MEGAX (Kumar et al., 2018).

Pulsed-Field Gel Electrophoresis
The agarose plugs containing the genomic DNA from streptococcal strains were prepared as previously described (Rato et al., 2008). All pulsed-field gel electrophoresis (PFGE) patterns were analyzed visually and by computer-assisted cluster analysis using BioNumerics v. 6.6 (Applied Maths, Sint-Martens-Latem, Belgium). Levels of similarity were calculated using Dice coefficient, and unweighted pair group method with arithmetic mean (UPGMA) was used for clustering to produce band-based dendrograms (1.5% and no optimization). Groups of patterns with similarity of 100% were considered indistinguishable and were assigned to the same subtype of PFGE type.

Phylogenetic Analysis
For the phylogenetic analysis of bovine SDSD isolates, the Multilocus Sequence Analysis (MLSA) scheme was performed according to the method previously described for SDSE based on sequences of seven housekeeping genes (McMillan et al., 2010). Seven housekeeping genes, namely, glucose kinase \(gki\), glutamine transport protein \(gtr\), glutamate racemase \(mtr\), DNA mismatch repair protein \(mutS\), transketolase \(recP\), xanthine phosphoribosyl transferase \(xpt\), and acetate-acytetyl-coathyloioase \(alto\), were amplified; and all amplified DNA fragments were sequenced. The sequences were edited, aligned, and subjected to phylogenetic analysis using CLC Bio Main Workbench 20.1 alignment program editor (QIAGEN, Venlo, Netherlands). Housekeeping gene sequences were trimmed, concatenated, and subjected to phylogenetic analysis. For concatenated housekeeping alleles, phylogenetic relationships were inferred using the minimum evolution method, and the distances were computed using the maximum composite likelihood method.

For the phylogenetic analysis, bovine SDSD isolates representatives of collection I and II (a total of 15 strains) were selected based on their different resistance, virulence, and PFGE profiles. Additionally, the sequences of SDSD strains \((n = 12)\), SDSE strains \((n = 15)\), S. pyogenes strains \((n = 11)\), and S. canis strains \((n = 3)\) (Supplementary Table 1 in Supplementary File 1), available in the NCBI Nucleotide database, were included in this study for comparison.

Virulence Gene Screening
The following genetic determinants of \(S.\ pyogenes\) (i.e., \(speA\), \(speB\), \(speC\), \(speF\), \(speG\), \(speH\), \(speK\), \(speL\), \(speM\), \(spgg\), \(smZ\), \(spd1\), and \(sdn\)) were screened in bovine SDSD isolates by PCR. Additionally, we investigated the presence of operon encoding streptolysin S (SLS), namely, \(sagA\), \(sagB\), \(sagE\), and \(sagL\) genes. For each 25 \(\mu l\) PCR mixture were added 100 ng of bacterial DNA, 1 × reaction buffer for NZYTaq DNA polymerase, 2.5 mM of MgCl\(_2\), 0.4 mM of dNTPs NZYMix, 1 U of NZYTaq DNA polymerase (NZYTech, Lisbon, Portugal), and 1 \(\mu M\) of each primer (Thermo Fisher Scientific, Waltham, MA, United States). Gene expression assays by reverse-transcriptase PCR (RT-PCR) were performed according to Alves-Barroco et al. (2019). Negative results were confirmed in at least three independent assays. VSD13 and VSD7 were used as positive control. Primer sequences and amplicon expected size PCR conditions for amplifying all the genetic determinants are listed in Supplementary Table 2 (Supplementary File 2).

Screening of virulence genes and comparative analysis in \textit{silico} among SDSD (from bovine, human, and fish strains), SDSE, \(S.\ canis\), and \(S.\ pyogenes\) sequences (see Supplementary Table 1 in Supplementary File 1) were performed using the CLC Bio Main Workbench. Alignment of \(speC\), \(speK\), \(speL\), \(speM\), and \(spd1\) sequence of bovine SDSD against the GenBank database using BLAST was performed to trace homolog genes among streptococcal species.

Antimicrobial Resistance Patterns
All SDSD strains were tested for antimicrobial resistance by disk diffusion method (Oxoid Ltd., Basingstoke, United Kingdom), and the determination of minimum inhibitory concentration (MIC) follows the guidelines from the Clinical and Laboratory Standards Institute (CLSI) for antimicrobial susceptibility tests for bacteria isolated from animals (CLSI M31-A3, 2008). The following antimicrobials were selected for testing, based on several criteria: (a) licensing for mastitis treatment in cattle—amoxicillin–clavulanic acid 30 \(\mu g\) (AMC), gentamicin 10 \(\mu g\)

---

\(^1\)www.ncbi.nlm.nih.gov/BLAST
(CN), penicillin 10 units (P), pirlimycin 2 µg (PRL), and streptomycin 10 µg (S); and (b) use in human medicine—chloramphenicol 30 µg (CHL), erythromycin 15 µg (ERY), tetracycline 30 µg (TET), and vancomycin 30 µg (VA). The results were interpreted according to the CLSI guidelines (CLSI M31-A3, 2008). The antimicrobial manufacturers’ instructions were followed when CLSI guidelines were not available. The reference strain *Staphylococcus aureus* ATCC 29213 was used for quality control for both methods.

**Macrolide Resistance Phenotypes**

All the isolates were tested for resistance only to macrolides (M phenotype) or to macrolides, lincosamides, and streptogramin-B (MLSB phenotype). The MLSB resistance, inducible (iMLSB) or constitutive (cMLSB), was checked by a double-disc test with erythromycin and pirlimycin disks. Resistance only to lincosamides (L phenotype) and to lincosamides and streptogramin-A (LSA phenotype) was described previously (Malbruny et al., 2004; Rato et al., 2013).

**Antimicrobial Resistance Gene Detection**

All the isolates with M or MLSB phenotypes (Pires et al., 2005) were screened to detect *ermA, ermB*, and *mefA* macrolide resistance genes. Genes *tetM, tetT, tetW, tetL, tetQ, tetO, tetK*, and *tetS* were screened among all tetracycline-resistant strains (Pires et al., 2005). All strains were tested for the presence of *linB* gene (lincosamide resistance) (Bozdogan et al., 1999). Primer sequences and amplicon expected sizes are listed in Supplementary Table 2 (Supplementary File 2). Negative results were confirmed in at least three independent assays. VSD13 and VSD19 strains were used as positive control.

**CRISPR Array Sequencing and CRISPR-Associated (Cas) Gene Screening**

In this study, the genome sequence of an SDSE strain ATCC 12394 (GenBank accession number: NC_017567) was used to design specific primers to amplify the CRISPR array and *cas* genes (Supplementary Table 2 in Supplementary File 2). PCRs were performed in a thermocycler (Biometra, Göttingen, Germany) in a final volume of 25 µL containing 100 ng of bacterial DNA, 1 × PCR buffer, 2.5 mM of MgCl₂, 0.4 mM of dNTPs, 1 U of Taq polymerase (NZYTech, Lisbon, Portugal), and 1 µM of each primer (NZYTech, Lisbon, Portugal). PCR conditions consisted of an initial denaturation cycle (95°C for 5 min), followed by 35 cycles of denaturation (95°C for 30 s), annealing (52°C for 30 s), and extension (72°C for 80 s). A final extension at 72°C for 7 min was also performed. Milli-Q water was used as a negative control in each PCR. The PCR products of CRISPR array regions were purified using the Wizard PCR Preps DNA Purification System (Promega, Madison, WI, United States). Sequencing was performed by STAB-Vida (Lisbon, Portugal) using specific primers (CRISPR Sdys For and CRISPR Sdys Rev; Supplementary Table 2 in Supplementary File 2). The obtained DNA sequences were analyzed and edited using the software CLC Bio Main Workbench. For each CRISPR array sequence, the spacers, repeats, and flanking regions were determined using CRISPRfinder and CRISPRtionary². All experiments were carried out in triplicate (Grissa et al., 2007).

The similarity of the spacers, repeats, and flanking regions was analyzed using software currently available for CRISPR miner to infer a large-scale microbe–phage infection network¹. The spacers from each strain were selected and used to BLAST (blastn with parameters settings: e-value < 0.01, and up to three mismatches allowed) against both bacteria and archaea virus genomes of an integrated database from the NCBI viral Genome Resource. Matches with a query cover above 90% with 100% identity were considered.

**Prophage Identification**

Prophinder (Lima-Mendez et al., 2008) was used to screen for prophage-specifying regions among the genome of SDSD, SDSE, *S. canis*, and *S. pyogenes* strains (accession numbers of genomes are available in Supplementary Table 1 in Supplementary File 1). Prophage sequences were annotated using RAST (Aziz et al., 2008). Intact prophages were manually analyzed to confirm the presence of genes indispensable to produce functional phages: left attachment site (attL); lysogeny; DNA replication; transcriptional regulation; head; tail; lysis modules; and right attachment site (attR).

**Statistical Analysis**

GraphPad Prism version 7.0 was used for statistical analysis. Data analysis was performed using the chi-square method. Statistical significance was considered when *p* < 0.05.

**RESULTS**

**Bacterial Isolates and Identification**

The SDSD isolates’ identification was confirmed using 16S rRNA gene sequencing. Out of 194 Gram-positive cocci, 37 isolates (19%) were identified as *S. dysgalactiae* species. Sequence analysis of the rRNA 16S from bovine SDSD isolates showed between 99.2% and 100% nucleotide identity to SDSD ATCC 2795, SDSD NCTC4670, SDSD NCTC13731, and SDSD NCTC4669 strains deposited in the NCBI Nucleotide database. UPGMA dendrogram shows the multiple alignments of rRNA 16S sequences of bovine SDSD (from this study), SDSD, SDSE, and *S. pyogenes* strains deposited in the NCBI database. See Supplementary Figure 1 (Supplementary File 3).

**Pulsed-Field Gel Electrophoresis Profiles**

The 55 SDSD isolates from subclinical mastitis isolated in Portugal during 2002–2003 (collection I) and 2011–2013 (collection II) were resolved into 45 PFGE types. The 55 SDSD isolates from subclinical mastitis isolated in Portugal during 2002–2003 (collection I) and 2011–2013 (collection II) were resolved into 45 PFGE types. Epidemiologically groups or clusters (at 70% similarity) revealed two major clonal groups, including close to 90% of all bovine isolates. The macro-restriction profiles are listed in

¹[https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index](https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index)

²[http://www.microbiome-bigdata.com/CRISPRminer/index.php/Home/Index/index](http://www.microbiome-bigdata.com/CRISPRminer/index.php/Home/Index/index)
**TABLE 1 | Characterization of bovine SDSD isolates from collection I (VSD1 to VSD11 and VSD13 to VSD19) (2002–2003) and from collection II (VSD20 to VSD55) (2011–2013).**

| Strain | Farm | PFGE type | Virulence genotype | Resistance genotype | Resistance phenotype |
|--------|------|-----------|--------------------|---------------------|---------------------|
| VSD1   | F    | A-1       | sagA, speC, speK, spd1 | tet                 |                     |
| VSD2   | F    | J-1       | sagA                |                     |                     |
| VSD3   | J    | L-1       | sagA                | ermA/Tr, tetM       | cMLSB + tet         |
| VSD4   | E    | M-1       | sagA, speC, speK, spd1 | linB, tetM         | L + tet             |
| VSD5   | E    | F-1       | sagA, sdn          | tetM                |                     |
| VSD6   | I    | D-1       | sagA, speK, speL   |                     |                     |
| VSD7   | D    | G-1       | sagA, speM, sdn    | tetO                |                     |
| VSD8   | G    | C-1       | sagA, sdn          |                     |                     |
| VSD9   | B    | N-1       | sagA                | ermB, tetO          | cMLSB + tet         |
| VSD10  | G    | C-1       | sagA, sdn          |                     |                     |
| VSD11  | F    | B-1       | sagA, speC, speK, spd1 | linB, tetO     | L + tet             |
| VSD12  | G    | I-1       | sagA, speC, speK, speL, speM, spd1 | ermB, tetO | cMLSB + tet         |
| VSD13  | F    | C-2       | sagA                | ermB, tetO          | cMLSB + tet         |
| VSD14  | J    | D-1       | sagA, speK, speL   |                     |                     |
| VSD15  | B    | A-2       | sagA, speC, speK, spd1 |                     |                     |
| VSD16  | C    | E-1       | sagA                |                     |                     |
| VSD17  | I    | D-1       | sagA, speK, speL   |                     |                     |
| VSD18  | E    | M-1       | sagA, speC, speK, spd1 | linB, tetM     | L + tet             |
| VSD19  | S    | O-1       | sagA, speL          | ermB, tetO          | cMLSB + tet         |
| VSD20  | Q    | P-1       | sagA, speL, speM   |                     |                     |
| VSD21  | Q    | Q-1       | sagA, speM          |                     |                     |
| VSD22  | V    | R-1       | sagA, speM, sdn    |                     |                     |
| VSD23  | V    | S-1       | sagA, speC, speK, spd1 |                     |                     |
| VSD24  | N    | T-1       | sagA, sdn          |                     |                     |
| VSD25  | N    | V-1       | sagA, speK          |                     |                     |
| VSD26  | M    | X-1       | sagA                |                     |                     |
| VSD27  | X    | Z-1       | sagA, sdn          |                     |                     |
| VSD28  | M    | K-1       | sagA, speM          |                     |                     |
| VSD29  | T    | Y-1       | sagA, speL, speM   |                     |                     |
| VSD30  | M    | W-1       | sagA                |                     |                     |
| VSD31  | O    | AA-1      | sagA, speK, sdn    |                     |                     |
| VSD32  | O    | AA-1      | sagA, speK, sdn    |                     |                     |
| VSD33  | P    | AB-1      | sagA, speM          |                     |                     |
| VSD34  | N    | Z-1       | sagA, speL, speM, sdn |                     |                     |
| VSD35  | M    | W-1       | sagA, speL, speM   |                     |                     |
| VSD36  | O    | AG-1      | sagA, speK, sdn    |                     |                     |
| VSD37  | U    | W-1       | sagA                |                     |                     |
| VSD38  | U    | W-1       | sagA                |                     |                     |
| VSD39  | U    | AD-1      | sagA                |                     |                     |
| VSD40  | R    | AE-1      | sagA, sdn          |                     |                     |
| VSD41  | M    | W-1       | sagA                |                     |                     |
| VSD42  | M    | W-1       | sagA                |                     |                     |
| VSD43  | M    | AF-1      | sagA, speK, sdn    |                     |                     |
| VSD44  | S    | AG-1      | sagA, speC, speK, spd1 |                     |                     |
| VSD45  | U    | AH-1      | sagA                |                     |                     |
| VSD46  | U    | AJ-1      | sagA                |                     |                     |
| VSD47  | M    | AJ-1      | sagA, speC, speK, spd1 |                     |                     |
| VSD48  | M    | AJ-1      | sagA, speC, speK, spd1 |                     |                     |
| VSD49  | M    | AL-1      | sagA, speC, spd1, speK |                     |                     |
| VSD50  | R    | V-1       | sagA, speL, speK   |                     |                     |
| VSD51  | M    | AM-1      | sagA, speL, speM   |                     |                     |
| VSD52  | O    | AN-1      | sagA                |                     |                     |

(Continued)
Table 1. Two or more strains share six PFGE patterns, and 31 PFGE patterns were of single isolates. Results demonstrated that there are no distinct groups among isolates of collection I and collection II.

Phylogenetic Analysis
Alignment of concatenated sequences of seven housekeeping genes (gki, gtr, murI, mutS, recP, xpt, and atoB) was created using the standard parameters of ClustalW. Then a minimum evolution tree was obtained from sequences from bovine SDSD isolates (determined in this study) and sequences of a selection of SDSD, SDSE, S. pyogenes, and S. canis strains available in the NCBI database (Supplementary Table 1 in Supplementary File 1 and Figure 1).

Streptococcus canis and S. pyogenes constitute monophyletic clusters. The S. dysgalactiae cluster was divided into two groups: cluster 1—human and fish SDSD strains, and SDSE; cluster 2—bovine SDSD isolates. Data show that human SDSD strains are closer to SDSE than bovine SDSD isolates. Interestingly, this phylogenetic relationship between SDSD and SDSE obtained by MLSA does not reflect the relationship based on the 16S rRNA sequences (Figure 1 and Supplementary Figure 1 in Supplementary File 3). The accession numbers of gki, gtr, murI, mutS, recP, xpt, and atoB genes are present in Supplementary Table 3 (Supplementary File 5).

Virulence Genes
All the 37 bovine SDSD isolates (collection II), screened by PCR, carry at least one of the following superantigens (SAgs) genes (speC, speK, speL, and speM), DNase (spd1), streptodornase (sdn), or SLS (sagA). None of the strains carry speA, speB, speF, speG, speH, and speM SAgs genes; smeZ mitogenic exotoxin gene; and sagB, sagE, and sagH genes.

The results showed that the SAgs genes were unevenly distributed among all the 37 bovine SDSD isolates (see Table 1). The screening revealed the following distribution: speC, 20%; speK, 40%; speL, 20%; speM, 23.6%; spd1, 20%; and sdn, 27.2% positive isolates. Transcriptional analysis showed that all virulence genes were transcribed in all SDSD isolates that carried them.

Similar virulence gene patterns were found among bovine SDSD from collection I (Rato et al., 2010, 2011) and II. No relationships were found between virulence profiles and the farm or macro-restriction patterns (see Table 1).

The comparison of the distribution of virulence genes between SDSD isolates from collection I and II is shown in Figure 2. Chi-square statistical analysis revealed that each independent virulence gene proportion was not statistically different between collections.

The comparison of the homologous sequences of bovine SDSD SAgs genes with S. pyogenes pyrogenic exotoxin gene from the NCBI databases, showed high identity: >98% identity for speC; >97% for speK; >98% for speM; and >97% for speL. In addition, a high identity was also found for DNAses spd1 (>98%) (Supplementary Tables 4–7 in Supplementary File 4).

For a comparative analysis, the presence of virulence genes in the genomes of the streptococci listed in Supplementary Table 1 (Supplementary File 1) was analyzed. None of the speB, speC, speF, speG, speH, speK, speL, speM, and smeZ virulence genes were detected among human SDSD or SDSE strains. Streptodornase (sdn) was found among the human SDSE strains (3/15), S. pyogenes strains (2/15), bovine SDSD isolates (5/20), but it was not detected in human SDSD or S. canis strains. spegg gene was found in SDSE, S. pyogenes, and human SDSD, but not in bovine SDSD; and the smeZ was found only among S. pyogenes strains (Supplementary Figure 2 in Supplementary File 3). The nucleotide sequences of spegg gene in human and fish SDSD strains revealed >99% identity with its homolog in S. pyogenes from the NCBI database.

In addition, all the bovine SDSD isolates from collection I and the genomes listed in Supplementary Table 1 (Supplementary File 1) were tracked for the presence of the SLS operon. The comparative analysis showed that of the nine genes that are part of the SLS operon, the bovine SDSD isolates only had sagA gene (Table 1), while the SDSE, human and fish SDSD, S. pyogenes, and S. canis strains harbored the complete operon (Figures 3A,B and Supplementary Figure 2 in Supplementary File 3). Figure 3C shows the amino acid sequence of SagA protein of the bovine SDSD isolates deduced through in silico analysis, compared with SDSE, SDSD, S. canis, and S. pyogenes sequences. The alignment of SagA peptide demonstrated that the leader-peptide region is highly conserved, while the pro-peptide region is more variable. In this last region, the bovine SDSD isolates present a deletion of two cysteine residues.

Antimicrobial Resistance Patterns
Among the 37 bovine isolates from collection II, the highest antimicrobial resistance patterns were observed for tetracycline (90%), followed by streptomycin (69%) and gentamicin (60%). None of the streptococcal isolates tested showed resistance to amoxicillin/clavulanic acid, cefazolin, cefoperazone, chloramphenicol, penicillin, rifaximin, and vancomycin (Table 1 and Figure 4). Two isolates resistant to pirlimycin and erythromycin presented the cMLSB phenotype (Table 1),
while the phenotype M (erythromycin-resistant only) was not detected. The two strains showed the cMLSB phenotype (constitutive MLSB). Strains of the cMLSB phenotype harbored \textit{ermB} gene, while \textit{ermA} and \textit{mefA} (macrolide resistance genes) were not detected (Table 1). The resistance to tetracycline was associated with \textit{tetM}, \textit{tetK}, and \textit{tetO} (\textit{tetS}, \textit{tetT}, \textit{tetW}, \textit{tetL}, and \textit{tetQ} negative). Figure 4 compares the resistance phenotypes distribution and genotypes among the 55 bovine SDSD isolates from collection I (Rato et al., 2013) and collection II. Chi-square statistical analysis revealed that each independent resistance phenotype proportion was not statistically different between collections. However, statistically significant differences were observed between \textit{tetM}, \textit{tetK}, and \textit{tetO}.

**CRISPR System Analysis and Prophage Identification**

SDSD isolates from collection I of 2002–2003 (\(N = 18\)) and collection II of 2011–2013 (\(N = 37\)) were analyzed by sequencing the CRISPR loci. Overall, the analysis of CRISPR array discloses a high intrinsic diversity among bovine SDSD understudy, demonstrating the distinguish groups among isolates from both collections (Figure 5).

Among the 55 bovine SDSD isolates analyzed for the CRISPR array, 49 were PCR positive (but only 41 strains were CRISPR spacer positive; see Supplementary File 6), and four were PCR negative. For two strains, the amplified product showed high homology (>99%) with ISSdy transposase (ORF A, GenBank accession: VDZ39966.1 and ORF B, GenBank accession: VDZ39965.1).

A high degree of direct repeat (DR) sequence (5′GTTTTAGAGCTGTGCTGTTTCGAATGGTTCCAAAAC′3) conservation was observed among bovine SDSD isolates. The DR showed sequence similarity to typical repeat sequences of other streptococci species, such as \textit{S. agalactiae} with 100% of identity (Lier et al., 2015), \textit{Streptococcus thermophilus} with an average identity of about 85% (Hu et al., 2020), and \textit{S. pyogenes} with >95% of identity (Nozawa et al., 2011).

A total of 158 spacers were identified among the 41 CRISPR spacer positive strains, of which 80 (50%) were found to be unique. The number of spacers in each CRISPR array was eight on average, ranging from a maximum of 26 to a minimum of 2 (Supplementary Tables 8, 9 in Supplementary File 6). Most strains (39 of 41) shared at least one spacer; these strains harbored six and three unique spacers, respectively. The SDSD VSD5, VSD9, VSD11, VSD29, VSD33, VSD37, and VSD40 isolates were CRISPR spacer negative (Supplementary File 6).

Three major groups were observed based on the common spacer sequences or absence of spacers (Figure 5 and Supplementary Table 9 in Supplementary File 6). Some old spacers were detected in bovine SDSD that harbored the same combinations of SAgs. For example, the 41, 42, 62, 63, 64, 65, 66, 118, 119, 120, and 121 spacers were detected among isolates from collection I (VSD16 and VSD19) and collection II (VSD24, VSD47, VSD48, and VSD49) harboring \textit{speC},
speK, and spd1 genes (Figure 5 and Supplementary File 6). Some of these spacers were also detected in SDSD NCTC4669 (isolated in 1935). About 72% (114/158) spacers matched with sequences present in the genome of S. pyogenes, S. agalactiae, S. canis, and S. dysgalactiae with a similarity >95. A BLAST was carried...
Streptococcus phage 11b; str. ILRI005; S. agalactiae and phages belonging to other species, namely, regions. These regions correspond to phage isolated from the data was determined by a chi-square test ($\chi^2$), where a probability value ($p$) $\leq$ 0.05 was considered as statistically significant.

### DISCUSSION

In the present study, the MLSA was performed according to the method previously described (McMillan et al., 2010). In a previous study, the phylogeny of S. canis, S. dysgalactiae, Streptococcus equi, and S. pyogenes species based on MLSA (map, pfl, ppaC, pyk, rpoB, sodA, and tuf genes) was reconstructed according to this analysis; S. dysgalactiae includes two separate clusters corresponding to the “dysgalactiae” and “equisimilis” subspecies (Jensen and Kilian, 2012). Surprisingly, our results show the separation of bovine SDSD from human SDSD strains in different clusters, the latter being closer to the SDSE (Figure 1).

In this study, we searched for the presence of different S. pyogenes virulence genes (encoded by MGE or either
Some of the SAgS genes are phage- and transposon-associated among S. pyogenes strains (Proft and Fraser, 2016). Previous studies have related some of these genes to an increase in the pathogenic potential of SDSD isolates (Rato et al., 2010, 2011). Our results show that the SAgS genes (speC, speK, speL, and speM) and spd1 are unevenly distributed among bovine SDSD isolates (from collection I and collection II) but not among human and fish SDSD strains (Table 1 and Supplementary Figure 2 in Supplementary File 3).

The speC and spd1 (initially identified in the M1 phage) genes were always detected together, and it was also found to be associated with speK (initially identified in the M3 phage) among bovine SDSD isolates, indicating the probable poly-lysogeny, as described for S. pyogenes (Giovanetti et al., 2015). We also found the speC–spd1–speK genes in the Strept T12 NC_028700 prophage of the bovine SDSD NCTC4669 strain.

Alignment of speC, speK, speL, speM, and spd1 sequence of bovine SDSD against the GenBank database, using BLAST, revealed a nucleotide identity of 97% to 99% with the homologous S. pyogenes virulence (Supplementary File 5). Although the genomes of the bovine SDSD isolates from collection I and II had not been sequenced, the high homology of SAgS and the presence of these genes in prophage regions among S. pyogenes could indicate that similar phages (or even the same) are shared among the bovine SDSD and S. pyogenes. The high identity of the SAgS and the different habitats of SDSD (animals) and S. pyogenes (human pathogen) may suggest that the HGT events occurred before habitat separation.

We analyzed the presence of virulence genes among streptococcal genomes listed in Supplementary Table 1 (Supplementary File 1; Supplementary Figure 2 in Supplementary File 3). We found the streptococcal pyrogenic exotoxin G gene (spegg) in SDSE, S. pyogenes, and human SDSD, but not in bovine SDSD. The presence of spegg has been chromosomal) among bovine SDSD isolates, namely, genes encoding DNAses, SAgS, and SLS.

![Dendrogram based on spacer CRISPR IIA cluster of Streptococcus dysgalactiae subsp. dysgalactiae isolates of mastitis in cattle collected in 2002–2005 (Rato et al., 2010, 2011, 2013) (VSD1 to VSD19—collection I) and 2011–2013 (VSD20 to VSD56—collection II). Streptococcus pyogenes phages encode genes speC, speK, speL, speM, spd1, and spd1. The dendrogram was obtained by means of Dice coefficient and the agglomerative clustering of the unweighted pair group method with arithmetic mean (UPGMA).](#)
sequences from the human and fish SDSD strains revealed that this gene has not been reported so far among human SDSD strains. Nucleotide sequences from the human and fish SDSD strains (Abdelsalam et al., 2010); SDSD, Streptococcus dysgalactiae subsp. dysgalactiae.

SDSD (α-hemolytic) and SDSE (β-hemolytic) have been distinguished based on their hemolytic properties on blood agar in routine laboratory practice. The hemolytic activity is mainly attributed to SLS production among the pyogenic group streptococci. In this study, the results showed the loss of sagB-I genes among bovine SDSD isolates. At the same time, the presence of the SLS operon was observed in human and fish SDSD, SDSE, S. pyogenes, and S. canis (β-hemolytic strains). Our data are the first to suggest that hemolytic differences between human SDSD strains (β-hemolytic) and bovine SDSD isolates (α-hemolytic) may be related to the loss of SagB-I and mutation of sagA gene observed in bovine SDSD isolates (Figure 3). Likely, the exact incidence of SDSD human infections on a global basis can be underestimated, mainly because of the failure to distinguish SDSD from SDSE strains in clinical laboratories.

Even though SDSD remains susceptible to most prescribed antibiotics, a significant number of treatment failures have been reported (Trieu-Cuot et al., 1993; Arthur et al., 1995; González Pedraza Avalés and Ortiz Zaragoza, 2000; Schwarz et al., 2004; Woodford, 2005; Pletz et al., 2006; Kimura et al., 2013; Cattoir, 2016; Doumith et al., 2017; Lai et al., 2017). Although tetracycline is not indicated for bovine mastitis-associated streptococci treatment in Portugal nowadays, high resistance to this class of antibiotics has been observed among bovine SDSD isolates. We observed a significant increase in the frequency of tetM, tetK, and tetO tetracycline resistance genes among bovine SDSD of collection II compared with bovine strains of the collection I. This frequency is likely associated with the fact that tetracycline has been extensively used to treat infections in bovine for many decades; tetracycline resistance determinants are linked to other resistance genes. Therefore, tetracycline resistance genes can be acquired via MGEs (e.g., plasmids and/or transposons), which also harbored.

### TABLE 2 | Identification and nucleotide sequence of the bacteriophage spacers in a CRISPR array of SDSD.

| Spacer | Sequence | Bacteriophage | SDSD strain |
|--------|----------|--------------|-------------|
| sp4    | ATAAATTTTTGGTACGAGTCTACCGT       | Streptococcus phage phi3396 | NCTC4669 |
| sp6    | ATGAGAAATAGCGATATAAACAGGAA       | Temperate phage phiNH1.1    | NCTC4669, VSD16 |
| sp7    | TTAGGGCGGAGTATTAAAGGTGTGTGTCT    | Streptococcus phage 315.6   | NCTC4669 |
| sp11   | TTGTTTTTGAGGTCCGTGAATCAATAAAA    | Streptococcus phage 315.2   | NCTC4669, VSD16 |
| sp14   | AGGTATTCTCAGCAACACTCCCATACAGTA   | Streptococcus phage 315.6 or S. phage P9 | NCTC4669, VSD16 |
| sp15   | AGGTATTCTCAGCAACACTCCCATACAGTA   | Streptococcus phage P9      | NCTC4669 |
| sp17   | TTAGGGCGGAGTATTAAAGGTGTGTGTCT    | Streptococcus phage 315.6 or S. phage P9 | NCTC4669 |
| sp25   | TACGAACCCCTCAATAATGCACAGTGA      | Streptococcus phage phi3396 | VSD1, VSD2 |
| sp30   | AAGGAGCTCAATAGCTCCGATGATATG      | Streptococcus phage 315.1   | VSD1, VSD2 |
| sp35   | AGAATTCTTCTGGAAACTCTGCCATGATG    | Streptococcus phage 315.1   | VSD1, VSD2 |
| sp36   | AAAAGGCCTCTCAAATGCAACAGTGA       | Streptococcus phage 315.3   | VSD10, VSD8 |
| sp49   | AGGTATTCTCAGCAACACTCCCATACAGTA   | Streptococcus phage 315.1   | VSD10, VSD8 |
| sp50   | CTTTTTGGACCGGGCTCCCATAGCGATT    | Streptococcus phage 315.2   | VSD14, VSD21, VSD22, VSD45 |
| sp52   | TATACTATCCTCAGCTGTAACACTCCGAG    | Streptococcus phage 315.3   | VSD14, VSD21, VSD22, VSD3 |
| sp53   | AGAACGATATGACGGATATCCGTAAGAATAC | Bacillus virus IEBH or Bacillus virus 250 | VSD14, VSD22, VSD3 |
| sp55   | ACGATTTCTAGTCAACATACGATCAAGC    | Streptococcus phage 315.5   | VSD14, VSD3 |
| sp58   | TTAAGGCGCATGCTTGTATGAGGGC       | Listeria phage B054          | VSD14, VSD21, VSD22, VSD27, VSD31, VSD36, VSD38, VSD41 |
| sp81   | GGTGATTTGCTAAGACTGGTGAGAAAGA    | Streptococcus phage 315.5 or S. phage phi3396 | VSD19, VSD47, VSD48, VSD49 |
| sp88   | ACTACCAACAGGACAGTTTGCGACCTCTGT  | Temperate phage phiNH1.1 or S. phage 315.4 | VSD20, VSD30, VSD32, VSD39, VSD46, VSD63, VSD65 |
| sp103  | TTCCCTGATGAAACCTTAAAGG          | Streptococcus phage phi3396 | VSD25 |
| sp104  | AAAACTACAACAGCCTTGCTAAAAGCACC   | Streptococcus phage phi3396 | VSD25 |
| sp105  | TAAATAACGATTTGATATTGCTAAGTGG   | Flavobacterium phage 11b or S. phage SpS1 | VSD25, VSD35 |
| sp106  | CCATAGCAATACGAGAAGATT          | Synechococcus phage S-CAM7 | VSD25 |
| sp116  | AGATAGATAGAAAAGAAGATTTTTTQAT    | Temperate phage phiNH1.1 or S. s phage 315.1 | VSD32 |
| sp138  | TCAACAGCATCCTTATAAGACATCTCAAAGC | Streptococcus phage 315.3   | VSD45 |
| sp139  | AGCTCGGCAAGCGATAATGCTGCAATAGG  | Streptococcus phage 315.5   | VSD45 |
| sp140  | GCTGGGAGCTGGGTGCTGTGGGATTTTAT  | Streptococcus phage 315.4   | VSD45 |
| sp149  | AGATTGCCTTCTACAGGATGTTTTTGT    | Streptococcus phage 315.5   | VSD47, VSD48, VSD49, VSD16 |
| sp158  | AAATATTGATGCGGATCGGATGCGG       | Streptococcus phage 315.3   | VSD49, VSD16 |
TABLE 3 | Identification of CRISPR/Cas systems and prophage or prophage-like regions.

| Strain | CRISPR system | Prophage regions* |
|--------|---------------|-------------------|
| SD FDAARGOS 654 | Type IA | 12; Type IIC | 19 1 5 |
| SD NCTC13759 | Type IA | 3 2 8 |
| SD NCTC6403 | Type IIC | 19 1 2 |
| SD SCOR-SD1 | Type IA | 15; Type IIC | 20 0 3 |
| SDS D ATCC 27957 | Type IA | 3 2 6 |
| SDS D B31752-13 | Type IA | 33; Type IIC | 11 0 5 |
| SDS D B49998-05 | Type IIC | 28 0 4 |
| SDS D B53993-17 | Type IA | 5 0 2 |
| SDS D B60705-15 | Type IIC | 42 3 8 |
| SDS D Kdys0611 | Type IIC | 22 1 12 |
| SDS D NCTC13731 | Type IA | 3 2 8 |
| SDS D NCTC4669 | Type IA | 14; Type IIC | 22 3 13 |
| SDS D NCTC4670 | Type IICA | 4 21 11 |
| SDS D STREP97-15 | Type IIC | 7 3 20 |
| SDS D NCTC11554 | Type IA | 13; Type IIC | 34 1 5 |
| SDS D NCTC5370 | No detected 1 4 |
| SDS D NCTC5371 | Type IIC | 42 1 5 |
| SDS D NCTC6179 | Type IA | 3; Type IIC | 9 2 4 |
| SDS D NCTC6181 | Type IA | 5; Type IIC | 9 2 6 |
| SDS D NCTC7136 | Type IA | 25; Type IIC | 15 3 4 |
| SDS D NCTC9413 | Type IIC | 11 1 5 |
| SDS D NCTC9414 | Type IIC | 42 1 6 |
| SDS D 167 | No detected 1 4 |
| SDS D AC-2713 | Type IA | 15; Type IIC | 20 1 8 |
| SDS D ATCC 12394 | Type IA | 30; Type IIC | 26 0 4 |
| SDS D GGS 124 | Type IIC | 19 1 4 |
| SDS D RE378 | Type IA | 14; Type IIC | 8 0 5 |
| Streptococcus canis B700072 | Type IIC | 18 0 4 |
| S. canis NCTC12191 | Type IA | 8; Type IIC | 2 3 |
| S. canis HL, 100 | Type IA | 12; Type IIC | 21 1 4 |
| Streptococcus pyogenes emm64.3 | No detected 1 3 |
| S. pyogenes emm70 | No detected 1 3 |
| S. pyogenes emm87.1 | No detected 1 1 |
| S. pyogenes FDAARGOS, 668 | Type IA | 7; Type IIC | 3 2 4 |
| S. pyogenes JS12 | Type IIC | 3 1 2 |
| S. pyogenes NCTC10085 | Type IA | 4; Type IIC | 3 2 2 |
| S. pyogenes NCTC12052 | Type IIC | 3 1 7 |
| S. pyogenes NCTC12696 | Type IIC | 0 3 6 |
| S. pyogenes NCTC13736 | No detected 4 5 |
| S. pyogenes NCTC13738 | No detected 3 5 |
| S. pyogenes NCTC8322 | Type IA | 7; Type IIC | 10 3 4 |
| S. pyogenes SP1336 | Type IA | 6; Type IIC | 0 2 4 |

SDSE, Streptococcus dysgalactiae subsp. equisimilis; SDSD, S. dysgalactiae subsp. dysgalactiae.

*Complete prophages | total of putative prophage regions. Bovine SDSD isolates: ATCC 27957, NCTC13731, NCTC4669, and NCTC4670. Human SDSD strains: DB31752-13, DB49998-05, DB60705-15, and DB53993-17. Fish SDSD strain: SDS D Kdys0611 and SDS D STREP97-15.

erythromycin resistance genes (Benciani et al., 2007, 2010; Emaneini et al., 2014; Santoro et al., 2014; Cattoir, 2016), contributing to a multiresistant phenotype. In our SDSD isolates, we identified the presence of ermB/tetO/tetK and ermB/tetO macrolide/tetracycline resistance genes, suggesting a possible horizontal gene co-transference of these genes.

Since the sequential order of the spacers in the CRISPR array provides chronological information about the bacteria’s exposure to foreign nucleic acids (Lier et al., 2015), providing insights into host–phage interactions within a specific niche, herein, we investigated the occurrence of CRISPR/Cas systems and CRISPR spacer content in bovine SDSD isolates to gain insight into the population diversity and the impact of this system on HGT.

Our results showed that all bovine SDSD strains characterized in the present study carried the CRISPR/Cas IIA system. In addition, we also observed a high degree of polymorphism in the CRISPR spacers providing a high ability to discriminate between strains. About 18.4% of spacers match known prophages with over 96% identity indicating a cross-species exchange of genetic material. The most frequent bacteriophages are listed in Table 2. Some studies have suggested that some common phages act as spacer donators (Le Rhun et al., 2019; Hu et al., 2020); thus, different spacers in different strains but identified as the same phage may be correlated with phage's evolution (Sorek et al., 2013; Landsberger et al., 2018; Watson et al., 2018). About 28% of the sequence spacers of SDSD were not identified in the database, probably because of the rapid evolution of phage sequences, and sequences of new phages are not available.

The widespread presence of this CRISPR/Cas system among SDSD bovine strains, the high conservation of repeated sequences, and the polymorphism observed among the CRISPR IIA array spacers of these strains may lead us to consider them as indicators of the activity of the system (Figure 5 and Supplementary Table 9 in Supplementary File 6).

To elucidate the relationship between the prophages and CRISPR systems among SDSD strains, we analyzed CRISPR array, CRISPR-associated (Cas) proteins, and prophage structure using complete genome sequences of SDSD listed in Supplementary Table 1 (Supplementary File 1). For a comparative analysis, complete genome of SDSE, S. canis, and S. pyogenes were also analyzed.

We observed that the CRISPR II-A system had increased the number of spacers relative to CRISPR I-C. Indeed, CRISPR II-A locus is the most widespread type among streptococci strains and has the most significant number of spacers. Several studies concluded that type II-A CRISPR might be the primary phage prevention or elimination system in the streptococcus genome (Serbanescu et al., 2015; Yamada et al., 2019; Hu et al., 2020).

The genomes of SDSD and SDSE show a high number of spacers compared with those of S. pyogenes (Table 3). Some studies have speculated on the loss of the CRISPR array and/or the CRISPR/Cas locus in some clinical isolates. In this regard, it was proposed that the lack of CRISPR/Cas could indicate an adaptation to acquire new virulence genes. It has also been hypothesized that the low number of spacers acquired by S. pyogenes could be a strategy to acquire virulence genes and thus increase its pathogenic capacity (Le Rhun et al., 2019). Among our SDSD strains, there was no statistically significant difference in the total number of prophages in the strains with type IC and type IIA systems. A more significant
### TABLE 4 | Prophage regions identified by computational tools for the genome of *Streptococcus dysgalactiae* subsp. *dysgalactiae* strains.

| Strain          | Region length | Score | Total proteins | Phage most common     | GC%  |
|-----------------|---------------|-------|----------------|-----------------------|------|
| NCTC4669        | 54.7 kb       | 140   | 82             | Strept T12 (NC_028700)| 37.71|
|                 | 49.9 kb       | 150   | 65             | Strept 315.4 (NC_004587)| 38.96|
| STREP97-15      | 50 kb         | 130   | 56             | Strept 315.5 (NC_004588)| 37.96|
|                 | 43.4 kb       | 150   | 66             | Strept phi3936 (NC_000918)| 36.85|
|                 | 42.1 kb       | 150   | 51             | Strept phiNJ2 (NC_019418)| 38.29|
| NCTC4670        | 55.5 kb       | 150   | 84             | Strept T12 (NC_028700)| 38.44|
| ATCC 27957      | 62.5 kb       | 130   | 79             | Strept 315.3 (NC_004586)| 38.80|
|                 | 58.8 kb       | 150   | 76             | Strept 315.3 (NC_004586)| 38.68|
| NCTC13731       | 79.2 kb       | 150   | 104            | Strept T12 (NC_028700)| 38.32|
|                 | 50 kb         | 140   | 68             | Strept 315.4 (NC_004587)| 39.59|
| NCTC13759       | 79.2 kb       | 150   | 104            | Strept T12 (NC_028700)| 38.32|
|                 | 50 kb         | 140   | 67             | Strept 315.4 (NC_004587)| 39.59|
| DB60705-15      | 34.1 kb       | 140   | 49             | Strept A25 (NC_028697)| 41.13|
|                 | 36.6 kb       | 150   | 55             | Strept phiNJ2 (NC_019418)| 37.20|
|                 | 27.4 kb       | 140   | 31             | Escher RCS47 (NC_042128)| 34.78|
| Kdys0611        | 36 kb         | 100   | 24             | Geobac GBSV1 (NC_008376)| 40.00|

Bovine SDSD isolates: ATCC 27957, NCTC13731, NCTC4669, and NCTC4670. Human SDSD strains: DB49998-05 and DB60705-15. Fish SDSD strain: SDSD Kdys0611 and SDSD STREP97-15.

**GC**, guanine–cytosine; **SDSD**, *S. dysgalactiae* subsp. *dysgalactiae*.

number of prophage regions would be expected in the absence of the CRISPR/Cas system; however, we observe SDSD strains with an active CRISPR/Cas system harboring several prophage regions per genome.

Despite the different origins of SDSD strains, most prophages are related and were identified initially in *S. pyogenes* (Table 4). The phages’ specificity can explain the high frequency of *S. pyogenes* prophages among SDSD isolates; that is, the interaction between phage particles and bacteria generally involves specific receptors located in the bacterial cell membranes (Stone et al., 2019). The fact that SDSD and *S. pyogenes* do not share the same host makes it seem likely that the transfer of phages harboring streptococcal pyrogenic exotoxins has occurred before the divergence of ecological niches between the two species.

In conclusion, our results showed a significant increase in the frequencies of the *tetM*, *tetO*, and *tetK* genes over time, although tetracycline is not indicated for bovine mastitis-associated streptococcus treatment in Portugal nowadays. Our data suggest that carriage of *S. pyogenes* prophage virulence genes (*speC, speK, speL, speM, and spd1*) is maintained over time and seems to be a property of bovine SDSD, but not of human and fish SDSD and SDSE strains. The high homology of SAgs genes between bovine SDSD and *S. pyogenes* strains may suggest that the HGT events occurred before habitat separation.

This study documented for the first time the widespread presence of the CRISPR/Cas IIA system among bovine SDSD isolates, and the polymorphism observed in spacer content can be considered indicators of the system activity. However, no correlation was observed between the number of spacers CRISPR IIA and the number of prophages in the SDSD genomes. Further studies are necessary to understand better the CRISPR/Cas systems among SDSD strains, such as other CRISPR systems, and its impact on the evolution of SDSD virulence.

Our data are the first to suggest that hemolytic differences between human SDSD strains (β-hemolytic) and bovine SDSD isolates (α-hemolytic) may be related to the loss of SagB-I and mutation of *sagA* gene observed in bovine SDSD isolates. The SagA alignment reveals higher homology between human and fish SDSD and SDSE strains when compared with the bovine SDSD isolates. The close genetic relationship between non-bovine SDSD and SDSE was also clear from phylogenetic analysis based on MLSA, while bovine SDSD isolates seem more divergent.

The data set of this study suggests that the separation between the subspecies “*dysgalactiae*” and “*equisimilis*” should be reconsidered. However, a study including the most comprehensive collection of strains from different environments would be required for definitive conclusions regarding the two taxa.

**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**

IS-S, CA-B, AF, and RM contributed to the idea or design of the research. RB and MO collaborated with the collection of bovine mastitis samples. CA-B contributed to the experimental work, analysis of sequences, and performed complete genomes. JC and
CR-R contributed to the research of virulence genes. LC and RT contributed with the support in the BioNumerics software and assistance in the analysis. CA-B wrote down the first draft and the subsequent revisions of the manuscript. RM and AF performed the revision and editing of the final manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

This work was supported by the Unidade de Ciências Biomoleculares Aplicadas-UCIBIO, which is financed by national funds from FCT/MEC (UIDP/04378/2020 and UIDB/04378/2020) and also by projects PTDC/CVT-EPJ/4651/2012 and PTDC/CVT-EPJ/6685/2014. FCT-MEC is also acknowledged for grant SFHR/BD/118350/2016 to CA-B.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.686413/full#supplementary-material
of Strepotococcus agalactiae, S. dysgalactiae subsp. dysgalactiae, and S. uberis. Appl. Environ. Microbiol. 76, 7957–7965. doi: 10.1128/AEM.00805-10
Hu, T., Cui, Y., and Xu, Q. (2020). Characterization and comparison of CRISPR loci in Strepotococcus thermophilus. Arch. Microbiol. 202, 695–710. doi: 10.1007/s00203-019-01780-3
Jasir, A., Tanna, A., Efstratiou, A., and Schalén, C. (2001). Unusual occurrence of M type 77, antibiotic-resistant group B streptococci in southern Sweden. J. Clin. Microbiol. 39, 586–590. doi: 10.1128/CM.39.2.586-590.2001
Jensen, A., and Kilian, M. (2012). Delineation of *Strepotococcus dysgalactiae* subsp. its subspecies, and its clinical and phylogenetic relationship to *Strepotococcus pyogenes*. J. Clin. Microbiol. 50, 113–126. doi: 10.1128/CM.05900-11
Jordal, S., Glambek, M., Oppegaard, O., and Kittang, B. R. (2015). New tricks from *Strepotococcus* subsp. *Strepotococcus agalactiae*. Front. Vet. Microbiol. 6, 539–542. doi: 10.1093/jac/dk423
Koh, T. H., Sng, I. H., Yuen, S. M., Thomas, C. K., Tan, P. L., Tan, S. H., et al. (2009). *Strepotococcus* subserotype 9 strain GZ0565 contains a type VII secretion system putative substrate EsxA that contributes to bacterial virulence and a vanZ-like gene that confers resistance to teicoplanin and dalbavancin in *Strepotococcus agalactiae*. Vet. Microbiol. 205, 26–33. doi: 10.1016/j.vetmic.2017.04.030
Landsberg, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, M., et al. (2014). CRISPR-Cas superfamily: Where from here? Front. Microbiol. 5:676. doi: 10.3389/fmicb.2014.00676
McShan, W. M., and Nguyen, S. V. (2016). “The Bacteriophages of *Strepotococcus pyogenes*. 2016 Feb 10 [Updated 2016 Mar 25],” in Bacteriophages of *Strepotococcus pyogenes*: Basic Biology to Clinical Manifestations [Internet], eds J. J. Ferretti, D. L. Stevens, and V. A. Fischetti (Oklahoma (OK): University of Oklahoma Health Sciences Center).
Ng, L. K., Martin, I., Alfa, M., and Mulvey, M. (2001). Multiplex PCR for the detection of tetracycline resistant genes. Mol. Cell. Probes. 15, 209–215. doi: 10.1006/mcpp.2001.0363
Nozawa, T., Furukawa, N., Aikawa, C., Watanabe, T., Haobam, B., Kurokawa, K., et al. (2011). CRISPR inhibition of prophage acquisition in *Strepotococcus pyogenes*. PLoS One 6:e15943. doi: 10.1371/journal.pone.0015943
Park, M. J., Eun, I. S., Jung, C. Y., Ko, Y. C., Kim, Y. J., Kim, C. K., et al. (2012). *Strepotococcus dysgalactiae* subsp. *dysgalactiae* infection after total knee arthroplasty: a case report. Knee Surg. Relat. Res. 24, 120–123. doi: 10.5792/ksr.2012.24.2.120
Pires, R., Rolo, D., Gama-Norton, L., Morais, A., Lito, L., Salgado, M. J., et al. (2005). Group A *Strepotococci* from carriage and disease in Portugal: evolution of antimicrobial resistance and T antigenic types during 2000-2002. Microb. Drug Resist. 11, 360–370. doi: 10.1089/mdr.2005.11.360
Pletz, M. W., McGee, L., Van Beneden, C. A., Petit, S., Bardsley, M., Barlow, M., et al. (2006). Fluoroquinolone resistance in invasive *Strepotococcus pyogenes* isolates due to spontaneous mutation and horizontal gene transfer. Antimicrob. Agents Chemother. 50, 943–948. doi: 10.1128/AAC.50.3.943-948.2006
Profij, T., and Fraser, J. D. (2016). “Strepotococcus Superantigens: Biological properties and potential role in disease,” in *Strepotococcus pyogenes*: Basic Biology to Clinical Manifestations [Internet], eds J. J. Ferretti, D. L. Stevens, and V. A. Fischetti (Oklahoma (OK): University of Oklahoma Health Sciences Center).
Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie 117, 119–128. doi: 10.1016/j.bioch.2015.03.025
Rato, M. G., Beixiga, 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
Rato, M. G., Beixiga, R., Nunes, S. F., Cavaco, L. M., Vilela, C. L., and Santos-Sanches, I. (2008). Molecular epidemiology and population structure of bovine *Strepotococcus uberis*. J. Dairy Sci. 91, 4542–4551. doi: 10.3168/jds.2007-0907
Rato, M. G., Beixiga, R., Nunes, S. F., Vilela, C. L., and Santos-Sanches, I. (2010). Human group A *Strepotococci* virulence genes in bovine group C streptococci. Emerg. Infect. Dis. 16, 116–119. doi: 10.3201/eid1601.090632
Rato, M. G., Nerlich, A., Bergmann, R., Beixiga, R., Nunes, S. F., Vilela, C. L., et al. (2011). Virulence gene pool detected in bovine group C *Strepotococcus agalactiae* subsp. *dysgalactiae* isolates by use of a group A S. pyogenes virulence microarray. J. Clin. Microbiol. 49, 2470–2479. doi: 10.1128/JCM.00008-11
Richards, V. P., Zadoks, R. N., Vilema, P. D., Lefebure, T., Lang, P., Werner, B., et al. (2012). Genome characterization and population genetic structure of the zoonotic pathogen *Strepotococcus canis*. BMC Microbiol. 12:293. doi: 10.1186/1471-2180-12-293
Rothe, M., and Cleary, P. P. (2016). “Adhesion and invasion of *Strepotococcus pyogenes* into host cells and clinical relevance of intracellular streptococci. 10,” in Bacteriophages of *Strepotococcus pyogenes*: Basic Biology to Clinical Manifestations [Internet], eds J. J. Ferretti, D. L. Stevens, and V. A. Fischetti (Oklahoma City (OK): University of Oklahoma Health Sciences Center).
Roma-Rodrigues, C., Alves-Barroso, C., Raso, J. R., Costa, M. N., Fortunato, E., Baptista, P. V., et al. (2016). Infection of human keratinocytes by *Strepotococcus dysgalactiae* subsp. *dysgalactiae* isolated from milk of the bovine udder. Microbes Infect. 18, 290–293. doi: 10.1016/j.micinf.2015.11.005
Santoro, F., Vianna, M. E., and Roberts, A. P. (2014). Variation on a theme; an overview of the Ts916/Tn1545 family of mobile genetic elements in the oral and nasopharyngeal streptococci. Front. Microbiol. 5:335. doi: 10.3389/fmicb.2014.00335
Schmitz, F. J., Beyer, A., Charpentier, E., Normark, B. H., Schade, M., Fluit, A. C., et al. (2003). Toxin-gene profile heterogeneity among endemic invasive European group A streptococcal isolates. J. Infect. Dis. 188, 1578–1586. doi: 10.1086/379230
Schwarz, S., Kehrenberg, C., Doublet, B., and Cloeckaert, A. (2004). Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS Microbiol. Rev.* 28, 519–542. doi: 10.1016/j.femsre.2004.04.001

Serbanescu, M. A., Cordova, M., Krastel, K., Flick, R., Beloglazova, N., Latos, A., et al. (2015). Role of the *Streptococcus mutans* CRISPR-Cas systems in immunity and cell physiology. *J. Bacteriol.* 197, 749–761. doi: 10.1128/JB.02333-14

Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* 82, 237–266. doi: 10.1146/annurev-biochem-072911-172315

Stone, E., Campbell, K., Grant, I., and McAuliffe, O. (2019). Understanding and Exploiting Phage-Host Interactions. *Viruses* 11:567. doi: 10.3390/v11060567

Suzuki, H., Lefébure, T., Hubisz, M. J., Pavinski Bitar, P., Lang, P., Siepel, A., et al. (2011). Comparative genomic analysis of the *Streptococcus dysgalactiae* species group: gene content, molecular adaptation, and promoter evolution. *Genome Biol. Evol.* 3, 168–185. doi: 10.1093/gbe/evr006

Takahashi, T., Kaneko, M., Mori, Y., Tsuruji, M., Kikuchi, N., and Hiramune, T. (1997). Phylogenetic analysis of *Staphylococcus* based on the 16S rDNA sequence and assignment of clinical isolates from animals. *J. Vet. Med. Sci.* 59, 775–783. doi: 10.1292/jvms.59.775

Trieu-Cuot, P., de Cespédès, G., Bentorcha, F., Delbos, F., Gaspar, E., and Horaud, T. (1993). Study of heterogeneity of chloramphenicol acetyltransferase (CAT) genes in streptococci and enterococci by polymerase chain reaction: characterization of a new CAT determinant. *Antimicrob. Agents Chemother.* 37, 2593–2598. doi: 10.1128/aac.37.12.2593

Watson, B., Staals, R., and Fineran, P. C. (2018). CRISPR-Cas-mediated phage resistance enhances horizontal gene transfer by transduction. *mBio* 9, e02406–17. doi: 10.1128/mBio.02406-17

Wong, S. S., and Yuen, K. Y. (2012). *Streptococcus pyogenes* and re-emergence of scarlet fever as a public health problem. *Emerg. Microbes Infect.* 1:e2. doi: 10.1038/emi.2012.9

Woodford, N. (2005). Biological counterstrike: antibiotic resistance mechanisms of Gram-positive cocci. *Clin. Microbiol. Infect.* 11, 2–21. doi: 10.1111/j.1469-0691.2005.01140.x

Yamada, S., Shibasaki, M., Murase, K., Watanabe, T., Aikawa, C., Nozawa, T., et al. (2019). Phylogenetic relationship of prophages is affected by CRISPR selection in Group A *Streptococcus*. *BMC Microbiol.* 19:24. doi: 10.1186/s12866-019-1393-y

Zadoks, R. N., Middleton, J. R., McDougall, S., Katholm, J., and Schukken, Y. H. (2011). Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. *J. Mam. Gland Biol. Neoplas.* 16, 357–372. doi: 10.1007/s10911-011-9236-y

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.