This is a pre- or post-print of an article published in
Rato, M.G., Nerlich, A., Bergmann, R., Bexiga, R., Nunes, S.F., Vilela, C.L., Santos-Sanches, I., Chhatwal, G.S.
Virulence gene pool detected in bovine group C *Streptococcus dysgalactiae* subsp. *dysgalactiae* using a group A *Streptococcus pyogenes* virulence microarray

**Running title**

GAS virulence genes in bovine GCS

**Authors**

Márcia G. Rato¹, Andreas Nerlich², René Bergmann², Ricardo Bexiga³, Sandro F. Nunes³†, Cristina L. Vilela³, Ilda Santos-Sanches¹* and Gursharan S. Chhatwal²

¹- Centro de Recursos Microbiológicos. Departamento de Ciências da Vida. Faculdade de Ciências e Tecnologia. Universidade Nova de Lisboa. Quinta da Torre, 2829-516 Caparica, Portugal; 2- Dept. of Medical Microbiology, Helmholtz Centre for Infection Research, Inhoffenstraße 7, 38124, Braunschweig, Germany; 3- Centro de Investigação Interdisciplinar em Sanidade Animal, Faculdade de Medicina Veterinária, Av. da Universidade Técnica, 1300-477 Lisboa, Portugal.

†presently at Cambridge University, Cambridge, United Kingdom

* corresponding author: Centro de Recursos Microbiológicos, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica; e-mail: isanches@fct.unl.pt; Telephone/Fax: +351-212948530
Abstract

A custom designed microarray containing 220 virulence genes of *S. pyogenes* (GAS) was used to test group C *Streptococcus dysgalactiae* subsp. *dysgalactiae* (GCS) field strains causative of bovine mastitis and group C or group G *Streptococcus dysgalactiae* subsp. *equisimilis* (GCS/GGS) isolates from human infections - the latter used for comparative purposes - for the presence of virulence genes. All bovine and all human isolates carried a fraction of the 220 genes (23% and 39%, respectively). The virulence genes encoding streptolysin-S, glyceraldehyde-3-phosphate dehydrogenase, PAM and the collagen-like protein SclB were detected in the majority of both bovine and human isolates (94-100%). Virulence factors, usually carried by human beta-hemolytic streptococcal pathogens, such as streptokinase, laminin-binding protein and C5a peptidase precursor were detected in all human isolates, but not in bovine isolates. On the other hand, GAS bacteriophage-associated virulence genes encoding superantigens, DNase and/or streptodornase were detected in bovine isolates (72%), but not in the human isolates. Determinants located in non-bacteriophage-related mobile elements, such as the gene coding R28 were detected in all the bovine and human isolates. Several virulence genes, including genes of bacteriophage origin, were shown to be expressed by RT-PCR. Phylogenetic analysis of superantigen gene sequences revealed a high (>98%) identity among genes of bovine GCS, of the horse pathogen *S. equi* subsp. *equi* and of the human pathogen GAS. Our findings indicate that alpha-hemolytic bovine GCS, an important mastitis pathogen and considered to be a non-human pathogen, carries important virulence factors responsible for virulence and pathogenesis in humans.
Introduction

The alpha-hemolytic Lancefield group C *Streptococcus dysgalactiae* subsp. *dysgalactiae* (GCS) is a pathogen frequently associated with clinical and subclinical bovine mastitis, a disease that causes major economic losses in the dairy industry (67, 52). Virulence determinants have been identified in this pathogen, such as surface proteins which specifically interact with plasma or extracellular matrix proteins of the host, such as alpha-2-macroglobulin, plasminogen, albumin, fibrinogen, fibronectin, vitronectin and collagen (47, 35, 30, 64) and genes coding for proteins assumed to play a role in mastitis, such as alpha-2-macroglobulin-, immunoglobulin G-, or immunoglobulin A- binding protein Mig (25, 56), alpha 2-macroglobulin-, or immunoglobulin G- binding protein Mag (24), and fibrinogen- binding M-like protein (65).

Recently, *S. dysgalactiae* subsp. *dysgalactiae* has been reported to be associated with toxic shock like syndrome in cattle (9), suppurative polyarthritis in lambs (28), bacteremia in dogs (66), systemic granulomatous inflammatory disease, and severe septicaemia in fish (16) and in ascending upper limb cellulitis in humans in contact with raw fish (27). The presence of streptococcal pyrogenic exotoxin G gene (*spegg*) and streptolysin S structural gene (*sagA*), which have been associated with invasive disease in the exclusively human pathogen *Streptococcus pyogenes* (group A *Streptococcus*- GAS), has been documented in fish isolates of *S. dysgalactiae* subsp. *dysgalactiae* (1). We have previously reported the presence of GAS phage-encoded virulence genes among alpha-hemolytic *S. dysgalactiae* subsp. *dysgalactiae* isolates from bovine mastitis, namely the streptococcal pyrogenic exotoxin genes *speK*, *speC*, *speL*, *speM* and phage DNase1 gene *spd1*, and other genes
located on mobile genetic elements (48). So far no more information is available regarding
the presence of GAS virulence genes among *S. dysgalactiae* subsp. *dysgalactiae* strains and
nothing is known regarding the presence of GAS prophages in *S. dysgalactiae* subsp.
*dysgalactiae*. However, exchange of lysogenic phages among GAS and other human and
animal species, particularly the group C *Streptococcus dysgalactiae* subsp. *equisimilis* (a
pathogen that colonizes and infects humans with a clinical spectrum of diseases resembling
GAS), *Streptococcus equi* subsp. *equi* (exclusively horse pathogen) and *S. equi* subsp.
*zooepidemicus* (a zoonotic pathogen) has been previously reported (68, 19).

The aim of the present work was to use a microarray of genes encoding GAS virulence
factors (41) to have a better insight into the virulence gene pool (either chromosomal or
mobile genetic elements encoded) shared between GAS and alpha-hemolytic *S.
dysgalactiae* subsp. *dysgalactiae* isolates, associated with bovine mastitis, in comparison
with beta-hemolytic *S. dysgalactiae* subsp. *equisimilis* isolates associated with human
disease.

**Materials and Methods**

**Bacterial isolates and identification:** A total of 18 alpha-hemolytic *S. dysgalactiae* subsp.
dysgalactiae field isolates of Lancefield group C (GCS), causative of bovine subclinical
mastitis in dairy herds in Portugal, were used in the present study. Detailed information
regarding these field isolates, including identification and molecular typing data has been
described previously (48). In addition, six non duplicated beta-hemolytic *S. dysgalactiae*
subsp. *equisimilis* isolates of Lancefield group G (GGS) (n=5) and group C (GCS) (n=1),
causing pharyngitis (n=5) and invasive disease (n=1), in humans collected in Portugal, were
included in the study for purpose of comparison. Identification of *S. dysgalactiae* subsp. *equisimilis* isolates was based on colony morphology, hemolysis in blood agar plates, Lancefield grouping using the Streptex kit (Remel Europe Ltd, Dartford, England) and by PCR amplification of the 16S rRNA gene and sequencing (71). Two GCS alpha-hemolytic *S. dysgalactiae* subsp. *dysgalactiae* strains were analysed for detection of selected virulence genes. One of these strains was associated with toxic shock-like syndrome in cattle (9). The other strain caused ascending upper limb cellulitis in humans in contact with raw fish (27). We confirmed the identification of both strains by PCR amplification of the 16S rRNA gene and sequencing (71) and of the *sodA* gene and sequencing (1). Sequences were analysed by using the BioEdit sequence alignment editor (17) and compared with sequences from the National Center for Biotechnology Information (NCBI) database (Bethesda, MD, USA) by using the Blast alignment tool (www.ncbi.nlm.nih.gov/BLAST).

**Pulsed-field gel electrophoresis (PFGE) profiles for clonal characterization:** The description of the bovine *S. dysgalactiae* subsp. *dysgalactiae* GCS field strains (n=18) including typing by PFGE has been described previously (48). Typing of the human *S. dysgalactiae* subsp. *equisimilis* GGS/GCS isolates (n=6) by PFGE was carried out in this study as described for the bovine pathogen *S. uberis* (49). Smal-digested DNA banding patterns obtained by PFGE were analysed visually according to previous established criteria (61).

**Microarray design and hybridization:** The microarray has been described previously (41) and was used with minor modifications. In brief, the array consists of 70mer oligonucleotides from the conserved regions of all ‘classical’ GAS virulence factors and
orthologous of virulence factors found in other bacterial species as well as putative virulence genes present in the M1, M3, and M18 genome. In total 220 virulence factor/extracellular protein genes, ten house-keeping genes (positive controls), and ten negative controls were randomly spotted in six locations on the chip.

Genomic DNA was extracted using circconicum beads in combination with the Qiagen DNeasy kit (Qiagen). Genomic DNA was partially digested with AluI yielding fragments of a size between 500-1000 base pairs and labeled with biotin. The labeled DNA was purified (PCR Purification Kit; Qiagen) and labeling efficiency was verified by gel electrophoresis. Array hybridization was performed at 42°C for 16 hours, followed by incubation with Streptavidin-Cy5 using a SlideBooster SB800 (Advalytix). Fluorescence signals were obtained with a DNA Microarray Scanner G2565CA (Agilent Technologies) at 633 nm excitation and quantified using ImaGene software (BioDiscovery).

**Microarray data processing:** The raw data were corrected for background and transformed to log scale. A two component normal mixture model (39) was fitted to the corrected data by a maximum likelihood method adapted from the mclust package (12). A discriminant function was calculated to represent the propensity of a gene for being present or absent. Discriminant values were stored in a signal probability matrix and colored for presentation purposes using the following scheme: black indicates state 0 (not present), green indicates state 1 (present), and yellow indicates indecisive measurement. Dendrograms were calculated based on the signal probability matrix by using euclidean distance and the average linkage method by using the R statistics package (www.r-project.org).
Confirmation of the array data by PCR screening: In order to confirm the results obtained in the GAS array, PCR was carried out on several genes (speH, speC, speA, speL, speK, speI, speM, sdn, ssa, smeZ, sla, drs, prtf2, speG, ska, dppA, lbp, scpA, emm, isp, SpyM3_1736, slo, nga, spegg and sagA). The primer sequences, gene description and amplification length of each reaction are described in Table 1. Samples without DNA and strains lacking (negative) or carrying (positive) specific genes were used as controls in the PCR.

Sequence data and phylogenetic analysis of bacteriophage-associated virulence superantigen genes: Sequences of the superantigen encoded genes speC, speK, speL and speM of the bovine isolates under study, with high identity among them, were chosen to generate an alignment of DNA sequences of the alleles of those genes and homologue sequences deposited in the National Center for Biotechnology Information (NCBI) (Bethesda, MD, USA) database, in particular sequences of the speC, speK, speL and speM of S. pyogenes, seeL and seeM of S. equi subsp. equi, szeL and szeM of S. equi subsp. zooepidemicus and sdm of S. dysgalactiae subsp. dysgalactiae. The alignment (380 bp) was used to construct a phylogenetic tree by using MEGA version 4 software (60). The p-distance parameter and neighbour-joining method were used. Bootstrap values were calculated from 1000 replicates. Deduced amino acid sequences from these bovine alleles were compared with similar sequences from NCBI database by using the Blast alignment tool (www.ncbi.nlm.nih.gov/BLAST). Sequencing was performed by STAB-Vida (Lisbon, Portugal) using the same primers used for amplification. Nucleotide sequences were analysed by using the BioEdit sequence alignment editor (17).
emm-typing: Determination of emm gene (coding the M protein) types was performed as described elsewhere (http://www.cdc.gov/ncidod/biotech/strep/M-ProteinGene_typing.htm).

Gene expression assays by reverse-transcriptase PCR (RT-PCR): For RNA extraction all isolates were grown in Todd-Hewitt broth (Oxoid Limited, Basingstoke, England) supplemented with 1% yeast extract (BD, Franklin Lakes, USA), at 37°C until the mid-exponential phase (optical density of 0.5 at 600 nm) and the NucleoSpin RNAII kit (Macherey-Nagel, Dueren, Germany) was used according to the manufacturer’s instructions, followed by the addition of 2U/µL of DNaseI (Applied Biosystems/Ambion). RNA quality was confirmed in a 1% agarose gel electrophoresis and images captured using the Gel Doc XR system and the Quantity One 1-D Analysis Software (Bio-Rad). To confirm that no remaining DNA was present in RNA samples, PCR assays were performed targeting the housekeeping genes rpsL and rpsB, and the following genes under study: speM, speK, speL, speC, spd1, sdn, slo scpA, ska, nga, lmb, isp, dppA, emm and SpyM3_1736, using RNA as a template. Reverse transcriptase (RT) reactions for cDNA synthesis were performed using the SuperScript First-Strand Synthesis System for RT-PCR kit (Invitrogen, New Zealand). PCR targeting those genes under study was carried out again using cDNA as a template.

Results
Identification by 16S rDNA gene sequence analysis: The identification of the bovine *S. dysgalactiae* subsp. *dysgalactiae* was confirmed using the 16S rDNA gene sequencing as described (48). The 16S rDNA gene sequences of all human GCS/GGS isolates from the present study showed 99-100% identity to 16S rDNA gene sequences of *S. dysgalactiae* subsp. *equisimilis* deposited in the NCBI (Bethesda, MD, USA) database. Taking into account the phenotypic characteristics of the isolates together with the 16S rDNA gene sequence analysis we could confirm the six human beta-hemolytic GGS/GCS isolates included in this study as *S. dysgalactiae* subsp. *equisimilis*.

PFGE profiles: The 18 bovine *S. dysgalactiae* subsp. *dysgalactiae* isolates had 15 PFGE patterns as shown before (48). The six *S. dysgalactiae* subsp. *equisimilis* isolates had six different PFGE patterns with more than six band differences and were therefore considered unrelated according to established criteria (61).

emn-typing: None of the 18 alpha-hemolytic group C *S. dysgalactiae* subsp. *dysgalactiae* bovine isolates were typed by emm-typing since no amplification for this gene could be obtained. The *emm* gene types of the six human *S. dysgalactiae* subsp. *equisimilis* isolates were as follows: stC839 (GCS isolate; n=1), stG485, stG480, stG6792, and stG4831 (GGS isolates; n=5).

Microarray data: Of the 220 GAS virulence genes on the array (41), 44 genes (20%) were present in all bovine GCS (*S. dysgalactiae* subsp. *dysgalactiae*) and in all human GGS/GCS (*S. dysgalatiae* subsp. *equisimilis*) isolates, whereas 66 genes (30%) were not present in any of the isolates tested. The remaining genes (50%) showed variable
distributions among isolates of both origins. Relevant genes present in at least one isolate (bovine or human) and genes present in all the bovine isolates (with exception of hypothetical proteins) are shown in Table 2. Nine genes (ska, dppA, lmb, scpA, emm, isp, nga, slo, SpyM3_1736) were present in all the human isolates and absent in the bovine isolates, whereas only one gene (SpyM3_0345), encoding an uncharacterized protein, was detected in all bovine isolates and absent in all human isolates.

Among the total of 37 bacteriophage related genes on the array, only two (a hypothetical phage protein and a putative endolysin) were present in all the bovine isolates, both of them were also detected in all the human isolates, in addition to a putative lysin. None of the 13 bacteriophage encoded virulence genes (speC, speJ, speI, speH, ssa, mf4, slaA, speA3 speK, speL, speM, spd1, sdn) were detected in human GGS/GCS isolates, whereas at least one of the following six genes (speC, speK, speL, speM, spd1 and sdn) were detected in 72% (n=13) of the bovine GCS isolates.

Other GAS genes on the array, located in variable and recombinatory loci of the GAS genomes (40, 41), were detected in both human and bovine isolates. These include: Spy2009 gene coding a transposase and sic (at the mga chromosomal location), spyM18_2055 gene coding an amidase (at the spa chromosomal location), malE (at the maltose chromosomal location), prtF15 and sfbl (at the fibronectin-collagen-T antigen encoding region - FCT); citE (at the cit chromosomal location), and epf (at the sagA chromosomal location). Interestingly, the gene coding for the adhesin R28 carried by a transposon was present in all bovine and human isolates.

**Confirmation of the array data by PCR screening:** All PCR results confirmed the GAS microarray data, except for speM, speK and speG genes. By using other primers (designed
in this study) targeting the speM, speK of *S. pyogenes* and spegg of *S. dysgalactiae* subsp. *equisimilis* (73), eight bovine isolates were found that carried speK, one bovine isolate carried speM, one bovine isolate carried speK and speM and five human isolates carried spegg.

**Sequence data and phylogenetic analysis of bacteriophage-associated virulence superantigen genes:** Three speK alleles (*speK*-B-1, *speK*-B-2, *speK*-B-3) were identified among nine bovine isolates, two speM alleles (*speM*-B-1, *speM*-B-2), were identified in two bovine isolates, one speC allele (*speC*-B) was identified in six bovine isolates and one speL allele (*speL*-B) was identified in four bovine isolates. The phylogenetic tree based on sequences of those alleles, and of homologue sequences available at the NCBI database showed four major groups - each one comprising one of the four spe genes (*speK*, *speM*, *speC* or *speL*) of bovine *S. dysgalactiae* subsp. *dysgalactiae* isolates and of *S. pyogenes* as shown in Figure 1. The tree also shows that speM and speK diverged more recently as these two groups showed higher identities among them than with the speL or speC groups.

Amino acid sequences deduced from the bovine GCS alleles showed always 98-99% identity with the homologue GAS pyrogenic exotoxins gene sequences from NCBI database, with the exception of bovine SpeL, which showed higher identity with SeeM from *S. equi* subsp. *equi* (99%) than with SpeL from *S. pyogenes* (96%). In particular, and interestingly, SpeM from bovine isolates in the present study showed 99% identity with SpeM from GAS and 100% identity with Sdm from *S. dysgalactiae* subsp. *dysgalactiae* (sequences deposited in NCBI). Alignments of amino acid sequences of bovine GCS alleles from the present study and GAS superantigens available in NCBI database are shown in Figure 2.
Gene expression assays by reverse-transcriptase PCR (RT-PCR): Transcriptional analysis revealed that the bacteriophage-associated virulence genes speM, speK, speL, speC, spd1 and sdn detected in bovine GCS isolates were transcribed, and that the virulence genes ska, dppA, lmb, scpA, emm, isp, SpyM3_1736, slo and nga detected only in human isolates, were also transcribed.

Virulence gene profiling of invasive Streptococcus dysgalactiae subsp. dysgalactiae: The two invasive GCS alpha-hemolytic S. dysgalactiae subsp. dysgalactiae strains associated with toxic shock-like syndrome in cattle and with ascending upper limb cellulitis in humans who were in contact with raw fish, were analysed for detection of the following 18 virulence genes: sagA, sla, sdn, spd1, speI, speC, speA, speB, speK, speF, speM, speH, speL, speJ, ssa, smeZ, spegg, and prtf1 (see Table 1).

Both strains were negative for all the genes with the exception of sagA. The sagA gene sequence of the animal strain showed 95% of identity with the sagA gene sequences of S. dysgalactiae subsp. dysgalactiae and S. dysgalactiae subsp. equisimilis available in NCBI database. The sagA gene sequence of the human strain showed 100% of identity with the sagA gene sequence of S. dysgalactiae subsp. dysgalactiae and 99% of identity with the sagA gene sequence of S. dysgalactiae subsp. equisimilis available in NCBI database.

Discussion

Assessment of genes associated with mobile genetic elements (MGE) among Streptococcus dysgalactiae (subspecies dysgalactiae and equisimilis): An array containing
220 virulence genes from *S. pyogenes* (Group A *Streptococcus*, GAS) was used to analyze the virulence gene pool among *S. dysgalactiae* subsp. *dysgalactiae* (Group C *Streptococcus*, GCS), associated with bovine mastitis, and among *S. dysgalactiae* subsp. *equisimilis* (Group C or Group G *Streptococcus*, GCS/GGS), associated with human pharyngitis and invasive disease episodes. We had previously reported the presence of GAS virulence genes in bovine GCS (48), which motivated us to further analyse the bovine strains in search for the presence of other GAS virulence genes (either chromosomal or encoded by mobile genetic elements). The genes included in the array used in this study are from M1, M3 and M18 GAS genomes, of which particularly M1 and M3 are usually associated with severe disease in Europe and North America (54, 33). The results showed that genes were unevenly distributed among isolates of different host origins. The bovine GCS and human GAS shared 23% of all genes and both the human GCS/GGS and GAS shared 39% of all genes. A higher content of GAS virulence genes in human GGC/GGS was expected since both species share the same tissue niche in humans and cause a similar spectrum of diseases (11, 22, 58). Nevertheless, and most interestingly is that if we restrict the comparison to bacteriophage genes on the array, we observed that 65% of the phage-related genes were present in at least one bovine GCS isolate, whereas 35% were detected in at least one human GCS/GGS isolate.

None of the 13 bacteriophage virulence related genes (*speC, speI, speH, ssa, mf4, slaA, speA3 speK, speL, speM, spd1, sdn*), encoding streptococcal phospholipase, streptodornase, DNase and superantigens, all associated with GAS disease, were detected in the human GCS/GGS isolates. However, six of those 13 genes were detected in the bovine GCS isolates and we have observed that these GAS phage genes (*speK, speL, speM, speC, spd1 and sdn*) present in bovine GCS are expressed in vitro. Specifically, the distribution of
genes among the bovine GCS ranged from one phage-gene (sdn) present in three isolates, to five genes (speC-speK-speL-speM-spdl) present in one isolate. Also, the observed linkage of genes in a same bovine GBS strain, in particular speC-spdl (from M1 phage), speM-speL (from M18 phage) and speK (from M3 phage) indicates poly-lysogeny, similarly to what has been described for GAS (3). However, the lack of association of sets of genes encoded by complete phage genomes in our bovine isolates strongly suggests a recombinatory mosaic nature of phages as observed for GAS (3).

Also interesting was the observation that “non-bacteriophage associated” genes previously shown to be located in recombinatorial and mutational hotspots of GAS genome and thus considered to be associated with lateral gene transfer (40, 41) were detected in both the bovine GCS and human GCS/GGS isolates from the present study (see Table 2). These genes belong to four of the five large chromosomal regions described to have variable loci in GAS genome (41). One of the chromosomal regions includes the FCT locus (FCT, Fibronectin, Collagen binding, and T antigen) (63), which is considered one of the major location of adhesins in GAS, such as the fibronectin-binding proteins PrtF15 and SfbI among others, which were detected among all bovine and human isolates of the present study. Another chromosomal region (sagA) includes the gene epf (present in almost all bovine or human isolates) encoding extracellular matrix binding protein, which is another adhesin. Another region includes the maltose transport and cit operons, which are separated in the GAS genome by different complete bacteriophages, and include the genes malE and citE (41), which were detected in all bovine and human isolates. The fourth region includes the Spy2009 (coding a transposase), sic, spa, scl and SpyM18_2055 (coding an amidase) genes present in all bovine and human isolates, with the exception of the sic gene which
was present in only half of the human isolates). These genes are located in the mga and spa loci.

Furthermore, the gene coding the cell surface anchored adhesin R28, carried by integrative conjugative elements in GAS which resemble genetic elements of Group B *Streptococcus agalactiae* (GBS) (57) was detected in all bovine and human isolates. Also, the C5a peptidase precursor *scpA* gene, as well as the *lbp* gene encoding the laminin-binding protein, both known to be carried by a composite transposon of GBS (13) were detected in all human isolates and not in bovine isolates. Together our data highlight the importance of MGE mediating lateral gene transfer among different streptococcal species including bovine GCS.

**Nonrandom distribution of GAS virulence genes in bovine GCS isolates:** Genes of GAS encoding adhesins such as glyceraldehyde-3-phosphate dehydrogenase, a putative enolase, PrtF15, R28, a putative internalin A precursor, and a putative fibronectin-binding protein-like protein A (5, 10, 26, 57, 69), detected in all bovine GCS *S. dysgalactiae*, subsp. *dysgalactiae* isolates, strongly suggest that these might represent important virulence factors in this particular subspecies. In particular, the *emm* gene, encoding the antiphagocytic M protein was not present in the bovine isolates, although 94% of these isolates carried the gene encoding PAM, a member of the M protein family. Binding to extra cellular matrix proteins by GAS, such as fibronectin, can initiate mechanisms in the host cell that trigger internalization of the pathogen into host cells (43, 45).

Bovine GCS are known to be able to invade, persist, and induce cellular damage in bovine mammary epithelial cells (2).
Also interestingly, streptolysin S, strongly associated with invasive disease in GAS and associated with the beta-hemolytic phenotype of GAS and GCS/GGS (S. dysgalactiae subsp. equisimilis) (4, 21), was detected in all the bovine isolates, which are alpha-hemolytic. Furthermore, presence of streptolysin S gene (sagA) in alpha-hemolytic strains of group G (S. dysgalactiae subsp. equisimilis) and group C (S. dysgalactiae subsp. dysgalactiae), from human and animal origin, respectively, has been reported before (1, 72).

Bovine GCS is also known to grow well in mammary secretions, either during lactation or in secretions from dry animals (44, 56), which may be necessary to survive and establish in the specific environmental niche that is the bovine mammary gland. GAS genes related with housekeeping functions were also detected in the bovine isolates from the present study, such as arginine deiminase, putative metal binding protein of ABC transporter, putative PTS system IIB component, maltose/maltodextrin-binding protein, and putative ferric uptake regulator (7, 29, 32, 53, 50).

The gene encoding 6-phospho-beta-galactosidase, an enzyme of the glycoside hydrolase family 1 (6) and associated with the capacity of carbohydrate utilization, was detected in only 39% of the bovine isolates (and absent in all the human isolates). In order to ascertain if growth curves (using the same media and conditions of growth) of the bovine isolates from the present study were comparable or not, an automated Growth curve analysis system (BioScreen C®, Piscataway, NJ) was used. We have observed differences in growth curves among isolates from the present study, which do not seem to correlate with the presence/absence of the 6-phospho-beta-galactosidase gene (data not shown). Variable growth pattern of strains in bovine mammary secretions was previously described (44) and
may be related with the presence/absence of genes associated with the capacity of carbohydrate utilization, but not probably of this particular gene.

**Virulence genes detected only in the human GCS/GGS isolates:** In epidemiologically and genetically unrelated *S. dysgalactiae* subsp. *equisimilis* strains from human disease (pharyngitis and invasive), we detected nine genes that were not detected in the bovine isolates. Out of these nine genes, the following five were previously described in human *S. dysgalactiae* subsp. *equisimilis*: *lbp* (encoding the adhesion Lmb, a laminin-binding protein) (62), *ska* (encoding the plasminogen-activating Ska-streptokinase A protein) (37), *slo* (encoding the cytolytic streptolysin O toxin) (55) and *emm* (encoding the M protein) as well as *scpA* (encoding a C5a peptidase) - the two latter acting on the complement pathway of the host inhibiting bacterial opsonisation and phagocytosis (20, 8). Both the *lbp* and the *scpA* genes have been found in human GBS and are usually absent in bovine GBS isolates (14). These two last genes and also *ska* are known to be carried by a bovine pathogen *S. uberis*, although Lmb is not required for attachment of *S. uberis* to host epithelial cells, and the Ska locus is devoid of plasminogen activator coding sequences (70). The four remaining genes were found for the first time, in our study, in human *dysgalactiae* subsp. *equisimilis*: the *nga* gene (encoding a GAS extracellular enzyme NAD glycohyderolase) associated with cytotoxicity in host cells (34), the *isp* gene (encoding an immunogenic secreted protein) which has an unknown function although it is known to be expressed by GAS in the human host and generates antibody response (38), the *dppA* gene (encoding a dipeptide permease complex) which is a membrane-associated transporter for dipeptides in GAS and regulated by the multigene transcriptional regulator Mga (46) and the *SpyM3_1736* gene (coding a putative ABC transporter protein). ABC transporters are commonly
associated with virulence in GAS (29, 53). By gene expression assays (RT-PCR) we observed that all these nine genes from GAS were expressed in human GCS/GGS isolates. These observations together with the absence of these nine genes in the bovine GCS from this study suggest that they may be more important in human-host streptococci than in animal-host streptococci.

**Sequence data and phylogenetic analysis of superantigen genes of bovine group C S. dysgalactiae subsp. dysgalactiae:** Phylogenetic analysis of the alleles of phage-related superantigen genes of the bovine S. dysgalactiae subsp. dysgalactiae under study together with homologue gene sequences of other species: S. pyogenes, S. equi subsp. equi and S. equi subsp. zooepidemicus and of S. dysgalactiae, available in NCBI database revealed four groups or clades in the phylogenetic tree (Fig. 1). The bovine alleles were distributed by the four clades according to the S. pyogenes alleles. In the speK clade, the bovine alleles are organized together and separately from the seeL and szeL genes of S. equi subsp. equi and S. equi subsp. zooepidemicus, and the speK gene of S. pyogenes, which was expected since the streptococcal pathogens S. equi subsp. equi, S. equi subsp. zooepidemicus and S. pyogenes are known to share a common phage pool (19). In contrast, within the speL clade, the sequences of seeM and szeM genes of S. equi subsp. equi and S. equi subsp. zooepidemicus, respectively, were organized closer to the bovine speL-B allele from the present study, and separately to the speL sequence of S. pyogenes, suggesting a common phage content among the animal species. Also, considering the speC clade we may speculate that the same or similar phage(s) are shared between the bovine S. dysgalactiae subsp. dysgalactiae and the human S. pyogenes.
The reason why seeL and szeL sequences were not located in speL clade, and seeM and szeM sequences were not in the speM clade may be due to the nomenclature given to these genes of S. equi subsp. equi and S. equi subsp. zooepidemicus genes.

Bovine GCS SpeM amino acid sequence (with a length of 177 amino acid residues), showed 100% identity with S. dysgalactiae-derived mitogen (Sdm) sequence deposited in the NCBI database. The Sdm (encoded by the sdm gene) was the only superantigen described so far in S. dysgalactiae subsp. dysgalactiae with mitogenic activity (42). In agreement with our findings, these authors (42) also noticed high identities between sdm (from S. dysgalactiae subsp. dysgalactiae) and speM of S. pyogenes. Probably sdm and speM are the same gene, with a different nomenclature.

Also interesting is the fact that bovine GCS SpeL amino acid sequence from the present study showed higher identity with SeeM (99%) from S. equi subsp. equi than with SpeL (96%) from S. pyogenes, according to the NCBI database.

Our findings underline the role of GAS phages (which are known to be shared with S. equi subsp. equi and S. equi subsp. zooepidemicus) in the genetic diversity of bovine S. dysgalactiae subsp. dysgalactiae.

In conclusion, the presence of GAS virulence genes, particularly genes encoded by MGE, either random or non-randomly distributed among strains of bovine GCS may contribute to the increased virulence potential of these strains, namely the possibility of dissemination to different tissue of the host and to take advantage of new niches. As we have pointed out before, S. dysgalactiae subsp. dysgalactiae should not be disregarded as an infection agent in humans. In fact, this subspecies was associated with invasive disease in humans (27) and
here shown to carry the *S. pyogenes* streptolysin S gene, further suggesting *S. dysgalactiae* subsp. *dysgalactiae* as an emerging zoonotic pathogen.

**Acknowledgments**

This study was supported by projects POCTI/ESP/48407/2002 (Fundação para a Ciência e Tecnologia, Portugal)- FEDER (Fundo Europeu de Desenvolvimento Regional), PROC 60839 (Fundação Calouste Gulbenkian, Portugal), Project ref.46 (Centro de Investigação Interdisciplinar em Sanidade Animal/Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Portugal), and CREM (Centro de Recursos Microbiológicos, Portugal). M.G.R. was supported by PhD grant SFRH/BD/32513/2006 (Fundação para a Ciência e Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior).

We gratefully acknowledge Dr. Sonia Chérnier (Institut national de Santé Animale, Montreal, Canada) and Dr. Koh Tze Hsien (Singapore General Hospital, Singapore) for providing the two invasive GCS alpha-hemolytic *S. dysgalactiae* subsp. *dysgalactiae* strains included in this work.

**References**


Page 23


31. Lintges, M., S. Arlt, P. Uciechowski, B. Plümäkers, R. R. Reinert, A. Al-Lahham,
R. Lüttractive, and L. Rink. 2007. A new closed-tube multiplex real-time PCR to
detect eleven superantigens of Streptococcus pyogenes identifies a strain without

encoding IIAB(Man)L in Streptococcus salivarius is part of a tetracistronic operon
encoding a phosphoenolpyruvate: mannose/glucose phosphotransferase system.
Microbiology. 146:677-685.

Creti, K. Ekelund, M. Koliou, P. T. Tassios, M. van der Linden, M. Straut, J.
Vuopio-Varkila, A. Bouvet, A. Efstratiou, C. Schalén, B. Henriques-Normark,
Strep-EURO Study Group and A. Jasir. 2009. Clinical and microbiological
47:1155-1165.

(CMT): A Functional Equivalent of Type III Secretion in Gram-Positive Bacteria.
Cell. 104:143–152.

35. Mamo, W., G. Fröman, A. Sundás, and T. Wadström. 1987. Binding of
fibronectin, fibrinogen and type II collagen to streptococci isolated from bovine


44. Oliver, S. P. 1991. Growth of *Staphylococcus aureus* and *Streptococcus* species in bovine mammary secretions during the nonlactating and peripartum periods following intramammary infusion of lipopolysaccharide at cessation of milking. Zentralbl Veterinarmed B. **38:**538-544.


70. Ward, P. N., M. T. Holden, J. A. Leigh, N. Lennard, A. Bignell, A. Barron, L.
    Maskell, M. Kehoe, C. G. Dowson, N. Chanter, A. M. Whatmore, S. D. Bentley,
    and J. Parkhill. 2009. Evidence for niche adaptation in the genome of the bovine
    pathogen *Streptococcus uberis*. BMC Genomics. 10:54-71.


    Li, K. C. Lam, and K. Y. Yuen. 2003. Analysis of a viridans group strain reveals a
    case of bacteremia due to lancefield group G alpha-hemolytic *Streptococcus
    dysgalactiae* subsp *equisimilis* in a patient with pyomyositis and reactive arthritis. J
    Clin Microbiol. 41:613-618.

73. Zhao, J., T. Hayashi, S. Saarinen, A. C. Papageorgiou, H. Kato, K. Imanishi, T.
    expression, and characterization of the superantigen streptococcal pyrogenic exotoxin

**Figure legends**

**FIG.1** Phylogenetic analysis of superantigen gene sequences of bovine group C
*Streptococcus* from the present study and of sequences of *S. pyogenes*, *S. equi* subsp. *equi,*
*S. equi* subsp. *zooepidemicus* and *S. dysgalactiae* subsp. *dysgalactiae* available in National
Center for Biotechnology Information (NCBI) database (Bethesda, MD, USA). Bovine

**FIG.2** Alignments of amino acid sequences deduced from bovine group C streptococci superantigen alleles (*Spe*C-B, *Spe*K-B-1, *Spe*K-B-2, *Spe*K-B-3, *Spe*L-B, *Spe*M-B-1, *Spe*M-B-2) from the present study and of *Spe*C, *Spe*K, *Spe*L and *Spe*M sequences of *S. pyogenes* strains (MGAS8232 and MGAS6180) available in the National Center for Biotechnology (Bethesda, MD, USA) database, showing amino acid differences in *boldface* and *underlined*. Identities of 99% (A), 98-99% (B), 96% (C) and 99% (D) were observed among sequences. Nucleotide sequences length from which the alignments were originally created were of: 409 bp (*spe*C); 649 bp (*spe*K); 471 bp (*spe*L); and 504 bp (*spe*M).

**Tables**

**Table 1.** Genes and PCR primers used for screening of virulence determinants among Group C *Streptococcus dysgalactiae* subsp. *dysgalactiae* and Group C and G *Streptococcus dysgalactiae* subsp. *equisimilis*. 


<table>
<thead>
<tr>
<th>Gene description (origin)</th>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Product (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>antiphagocytic M protein (chromosome)</td>
<td>emm (forward)</td>
<td>TATTCGCTTAGAAAAATTA</td>
<td>variable</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>emm (reverse)</td>
<td>GCAAGTTCCTCCAGCTTTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>streptolysin S (chromosome)</td>
<td>sagA (forward)</td>
<td>TACCTCAAATATTATTAGCTACT</td>
<td>487</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>sagA (reverse)</td>
<td>GATGATACCCCGATAAGGATAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin H precursor (phage)</td>
<td>speH (forward)</td>
<td>AGATTGGAATACACAGG</td>
<td>416</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>speH (reverse)</td>
<td>CTATTCTCTCGTTATTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin F (chromosome)</td>
<td>speF (forward)</td>
<td>TACCTGGAATCAAGACG</td>
<td>782</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>speF (reverse)</td>
<td>GTAATTTAATGGGTGAGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin B (chromosome)</td>
<td>speB (forward)</td>
<td>TTCTAGGATACCTTACCAGC</td>
<td>300</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>speB (reverse)</td>
<td>ATTTTAGCAGTTGAGTAGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin C (chromosome)</td>
<td>speC (forward)</td>
<td>GCAGGGTAAATTTTTCACGACACA</td>
<td>407</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>speC (reverse)</td>
<td>TGTGCAATTTCGATTCTGCAGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin A (chromosome)</td>
<td>speA (forward)</td>
<td>ATGGAAAACAATAAAAAAGTATTG</td>
<td>755</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>speA (reverse)</td>
<td>TTACTGGTGGTTAGTAGACCTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin L (chromosome)</td>
<td>speL (forward)</td>
<td>CTGGTAGGATGTGTCTGCGGAAGAG</td>
<td>605</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>speL (reverse)</td>
<td>AGCACCCGCTTCTTCTCGCCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin K precursor (phage)</td>
<td>speK (forward)</td>
<td>TACAAATGATGTGAGAAATCCAAGGAACATATATGCT</td>
<td>656</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>speK (reverse)</td>
<td>CAAAGTGACTCTTACCTATCAATCGTTT</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin I precursor (phage)</td>
<td>speI (forward)</td>
<td>ATGAAGGTCGCCATTTC</td>
<td>516</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>speI (reverse)</td>
<td>TCCTCTGTCTGACCTATGCTCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin M precursor (phage)</td>
<td>speM (forward)</td>
<td>CCAATATGAAGATAAAATCAACGGGACCTT</td>
<td>600</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>speM (reverse)</td>
<td>CAAAGTGACTCTTACTATCATATCGTTT</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>pyrogenic exotoxin J (chromosome)</td>
<td>speJ (forward)</td>
<td>ATCTTCTATGGGTACG</td>
<td>535</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>speJ (reverse)</td>
<td>TTTCATGTTTATGCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>streptodornase (phage)</td>
<td>sdn (forward)</td>
<td>ACCCCATCGGAAGATAAACGC</td>
<td>489</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>sdn (reverse)</td>
<td>AACGTTCAACAGGCCCTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>streptococcal Dnase 1 (phage)</td>
<td>spd1 (forward)</td>
<td>ACCCTGAGATTGCTGTCAT</td>
<td>400</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>spd1 (reverse)</td>
<td>ACTGTGACGCAGCTAGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>streptococcal superantigen SSA (phage)</td>
<td>ssa (forward)</td>
<td>TCCACAGGTCAGCCTTTACTAG</td>
<td>502</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>ssa (reverse)</td>
<td>TGATCAAAATATTGCTTCCAGGTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene/Protein Name</td>
<td>Forward Primer</td>
<td>Reverse Primer</td>
<td>Length (bp)</td>
<td>Coverage (bp)</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------------</td>
<td>---------------------------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>phospholipase A2 (phage)</td>
<td>sla (forward)</td>
<td>sla (reverse)</td>
<td>440</td>
<td>36</td>
</tr>
<tr>
<td>mitogenic exotoxin Z (chromosome)</td>
<td>smeZ (forward)</td>
<td>smeZ (reverse)</td>
<td>400</td>
<td>Provided by G. S. Chhatwal</td>
</tr>
<tr>
<td>pyrogenic exotoxin G of <em>S. pyogenes</em> (chromosome)</td>
<td>speG (forward)</td>
<td>speG (reverse)</td>
<td>389</td>
<td>This study</td>
</tr>
<tr>
<td>pyrogenic exotoxin G of <em>S. dysgalactiae</em> (chromosome)</td>
<td>spegg (forward)</td>
<td>spegg (reverse)</td>
<td>420</td>
<td>This study</td>
</tr>
<tr>
<td>Drs protein (chromosome)</td>
<td>drs (forward)</td>
<td>drs (reverse)</td>
<td>760</td>
<td>18</td>
</tr>
<tr>
<td>fibronectin binding protein</td>
<td>prtf1 (forward)</td>
<td>prtf1 (reverse)</td>
<td>930</td>
<td>59</td>
</tr>
<tr>
<td>Prtf1 (chromosome)</td>
<td>prtf2 (forward)</td>
<td>prtf2 (reverse)</td>
<td>2000</td>
<td>Provided by G. S. Chhatwal</td>
</tr>
<tr>
<td>fibronectin binding protein</td>
<td>slo (forward)</td>
<td>slo (reverse)</td>
<td>408</td>
<td>This study</td>
</tr>
<tr>
<td>Prtf2 (chromosome)</td>
<td>nga (forward)</td>
<td>nga (reverse)</td>
<td>375</td>
<td>This study</td>
</tr>
<tr>
<td>streptolysin O (chromosome)</td>
<td>scpA (forward)</td>
<td>scpA (reverse)</td>
<td>591</td>
<td>Provided by G. S. Chhatwal</td>
</tr>
<tr>
<td>C5A peptidase precursor (transposon)</td>
<td>ska (forward)</td>
<td>ska (reverse)</td>
<td>598</td>
<td>Provided by G. S. Chhatwal</td>
</tr>
<tr>
<td>streptokinase A precursor (chromosome)</td>
<td>lmb (forward)</td>
<td>lmb (reverse)</td>
<td>408</td>
<td>This study</td>
</tr>
<tr>
<td>NAD-glycohydrolase precursor (chromosome)</td>
<td>isp.1 (forward)</td>
<td>isp.1 (reverse)</td>
<td>375</td>
<td>This study</td>
</tr>
<tr>
<td>laminin-binding protein (transposon)</td>
<td>dppA (forward)</td>
<td>dppA (reverse)</td>
<td>429</td>
<td>This study</td>
</tr>
<tr>
<td>immunogenic secreted protein (chromosome)</td>
<td>SpyM3_1736 (forward)</td>
<td>SpyM3_1736 (reverse)</td>
<td>1045</td>
<td>This study</td>
</tr>
<tr>
<td>surface lipoprotein DppA (chromosome)</td>
<td>SpyM3_1736 (forward)</td>
<td>SpyM3_1736 (reverse)</td>
<td>392</td>
<td>This study</td>
</tr>
<tr>
<td>putative ATB-binding cassette</td>
<td>SpyM3_1736 (forward)</td>
<td>SpyM3_1736 (reverse)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
transporter protein (chromosome) SpyM3_1736 (reverse) GGTGTCACTCTAGTTTACCTTT

1 (http://www.cdc.gov/ncidod/biotech/strep/M-ProteinGene_typing.htm)

Table 2. Distribution of Group A Streptococcus pyogenes virulence factors of the array in bovine mastitis Group C Streptococcus dysgalactiae subsp. dysgalactiae isolates and human non-invasive and invasive group C/ G Streptococcus dysgalactiae subsp. equisimilis. Only genes present in at least one isolate (bovine or human) and genes present in all bovine group C Streptococcus dysgalactiae subsp. dysgalactiae isolates are listed. Hypothetical proteins are not listed in table.

<table>
<thead>
<tr>
<th>Human Group A Streptococcus pyogenes virulence class/ gene</th>
<th>Distribution (%/ No. of isolates)</th>
<th>Bovine Group C Streptococcus dysgalactiae subsp. dysgalactiae (n=18)</th>
<th>Human Group C/ G Streptococcus dysgalactiae subsp. equisimilis (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adhesins:</td>
<td></td>
<td>Bovine Group C Streptococcus dysgalactiae subsp. dysgalactiae (n=18)</td>
<td>Human Group C/ G Streptococcus dysgalactiae subsp. equisimilis (n=6)</td>
</tr>
<tr>
<td>antiphagocytic M protein (emm)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>laminin-binding protein (lmb) a</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>collagen-like surface protein (scl) a</td>
<td></td>
<td>89 (n=16)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>PAM</td>
<td></td>
<td>94 (n=17)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. adhesion protein (adcA)</td>
<td></td>
<td>94 (n=17)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. extracellular matrix binding protein (epf) a</td>
<td></td>
<td>94 (n=17)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. pullulanase (pulA)</td>
<td></td>
<td>94 (n=17)</td>
<td>83 (n=5)</td>
</tr>
<tr>
<td>collagen-like protein ScIB (scl/B)</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>glyceraldehyde-3-phosphate dehydrogenase (plr)</td>
<td></td>
<td>100 (n=18)</td>
<td>83 (n=5)</td>
</tr>
<tr>
<td>Protein Category</td>
<td>Gene/Protein Name</td>
<td>Presence/Expression</td>
<td>Reference Samples</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>R28</td>
<td>put. choline binding protein (spyM3_0025)</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td></td>
<td>put. collagen-like protein (spy1054)</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td></td>
<td>put. enolase (eno)</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td></td>
<td>put. internalin A precursor (inlA)</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>Fibronectin-binding proteins:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SfbX</td>
<td></td>
<td>6 (n=1)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>SfbI</td>
<td></td>
<td>89 (n=16)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>prtF15</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. fibronectin-binding protein-like protein A (spyM3_0652)</td>
<td>100 (n=18)</td>
<td>67 (n=4)</td>
<td></td>
</tr>
<tr>
<td>Proteases:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5A peptidase precursor (scpA)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. C3-degrading proteinase (cppA)</td>
<td></td>
<td>22 (n=4)</td>
<td>0 (n=0)</td>
</tr>
<tr>
<td>put. exfoliative toxin (spyM3_0632)</td>
<td></td>
<td>94 (n=17)</td>
<td>67 (n=4)</td>
</tr>
<tr>
<td>put. serine protease (degP)</td>
<td></td>
<td>94 (n=17)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. protease (spyM3_0418)</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>Toxins:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>streptolysin O (slo)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. hemolysin III (hlyIII)</td>
<td></td>
<td>83 (n=15)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. hemolysin (hlyX)</td>
<td></td>
<td>94 (n=17)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>streptolysin S associated protein (sagA)</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. DNA-entry nuclease (endA)</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. hemolysin (hlyA)</td>
<td></td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>Other virulence factors:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dipeptide permease complex (dppA)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>immunogenic secreted protein precursor (isp)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>inhibitor of complement-mediated lysis (sic)</td>
<td></td>
<td>0 (n=0)</td>
<td>50 (n=3)</td>
</tr>
<tr>
<td>nicotine-adename-dinucleotide (NAD)-glycohydrolase (nga)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>pyrogenic exotoxin G precursor (speG)</td>
<td></td>
<td>0 (n=0)</td>
<td>83 (n=5)</td>
</tr>
<tr>
<td>streptokinase A precursor (ska)</td>
<td></td>
<td>0 (n=0)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>transposase (spy2009)</td>
<td></td>
<td>22 (n=4)</td>
<td>83 (n=5)</td>
</tr>
<tr>
<td>put. hyaluronidase (hyl)</td>
<td></td>
<td>28 (n=5)</td>
<td>100 (n=6)</td>
</tr>
</tbody>
</table>
put. 6-phospho-beta-galactosidase (*lacG*) 39 (n=7) 0 (n=0)
immunogenic secreted protein precursor homologue (*isp.2*) 67 (n=12) 100 (n=6)
extracellular hyaluronate lyase (*hylA*) 78 (n=14) 50 (n=3)
streptococcal protective antigen (*spa*) 78 (n=14) 50 (n=3)
put. penicillin-binding protein 1A (*pbp1A*) 83 (n=15) 33 (n=2)
put. glutathione peroxidise (*spyM3_0428*) 89 (n=16) 100 (n=6)
put. GTP-binding protein LepA (*lepA*) 89 (n=16) 100 (n=6)
arginine deiminase (*arcA*) 100 (n=18) 100 (n=6)
maltose/maltodextrin-binding protein (*malE*) a 100 (n=18) 100 (n=6)
67 kDa Myosin-crossreactive streptococcal antigen (*spyM3_0332*) 100 (n=18) 100 (n=6)
put. acid phosphatase (*lppC*) 100 (n=18) 100 (n=6)
put. amidase (*spyM18_2055*) a 100 (n=18) 100 (n=6)
put. carbamate kinase (*arcC*) 100 (n=18) 100 (n=6)
put. citrate lyase beta subunit (*citE*) a 100 (n=18) 50 (n=3)
put. cytoplasmic membrane protein (*lemA*) 100 (n=18) 100 (n=6)
put. dipeptidase (*pepD*) 100 (n=18) 100 (n=6)
put. dipeptidase (*spyM3_1763*) 100 (n=18) 100 (n=6)
put. ferric uptake regulator (*spf*) 100 (n=18) 100 (n=6)
put. fructose-1-phosphate kinase (*fruK*) 100 (n=18) 100 (n=6)
put. lipoprotein (*atmB*) 100 (n=18) 100 (n=6)
put. manganese-dependent inorganic pyrophosphatase (*spyM3_0278*) 100 (n=18) 100 (n=6)
put. metal binding protein of ABC transporter (*mtsA*) 100 (n=18) 100 (n=6)
put. peptidyl-prolyl cis-trans isomerase (*cypB*) 100 (n=18) 100 (n=6)
put. protease maturation protein (*prsA*) 100 (n=18) 50 (n=3)
put. proton-translocating ATPase subunit b (*spyM3_0495*) 100 (n=18) 100 (n=6)
put. PTS system IIB component (*spyM3_1679*) 100 (n=18) 100 (n=6)
put. sugar transporter sugar binding lipoprotein 100 (n=18) 100 (n=6)
put. two-component sensor histidine kinase (*yesM*) 100 (n=18) 100 (n=6)
pur. uridine kinase (*udk*) 100 (n=18) 100 (n=6)

**Bacteriophage-encoded recognized virulence factors:**

streptodornase, phage associated (*sdn*) 22 (n=4) 0 (n=0)
pur. exotoxin L precursor, phage associated (*speL*) 22 (n=4) 0 (n=0)
<table>
<thead>
<tr>
<th>Genes</th>
<th>Yes (%)</th>
<th>No (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pyrogenic exotoxin C precursor, phage associated (speC)</td>
<td>33 (n=6)</td>
<td>0 (n=0)</td>
</tr>
<tr>
<td>put. DNase, phage associated (spd1)</td>
<td>33 (n=6)</td>
<td>0 (n=0)</td>
</tr>
<tr>
<td>streptococcal pyrogenic K exotoxin, phage associated (speK)</td>
<td>50 (n=9)</td>
<td>0 (n=0)</td>
</tr>
<tr>
<td>put. exotoxin M precursor, phage associated (speM)</td>
<td>56 (n=10)</td>
<td>0 (n=0)</td>
</tr>
<tr>
<td><strong>Other bacteriophage-encoded factors:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>put. deoxyribonuclease</td>
<td>33 (n=6)</td>
<td>50 (n=3)</td>
</tr>
<tr>
<td>put. lysin - phage associated</td>
<td>89 (n=16)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>hypothetical phage protein</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
<tr>
<td>put. endolysin, phage associated</td>
<td>100 (n=18)</td>
<td>100 (n=6)</td>
</tr>
</tbody>
</table>

*a* Genes associated with lateral gene transfer (non-bacteriophage related)
FIG. 1 Phylogenetic analysis of superantigen gene sequences of bovine group C Streptococcus from the present study and of sequences of S. pyogenes, S. equi subsp. equi, S. equi subsp. zooepidemicus and S. dysgalactiae subsp. dysgalactiae available in National Center for Biotechnology Information (NCBI) database (Bethesda, MD, USA). Bovine group C Streptococcus sequences from the present study are designated in the tree in italic. Other sequences are designated as follows: SESE - S. equi subsp. equi (seeL and seeM); SESZ - S. equi subsp. zooepidemicus (szeL and szeM); SDSD - S. dysgalactiae subsp. dysgalactiae (sdm); GAS - S. pyogenes (speK, speM, speL and speC). Accession numbers (of bovine alleles, seeL, seeM, szeL, szeM and sdm) and
Gene ID (of speK, speM, speL and speC) are included in front of gene name between brackets.
A)  

25

SpeC-B  KDYVINSMS
SpeC-MGAS8232  KDYVINSMS

B)  

15  95  105  165  195

SpeK-B-1  YAFRXDEEI  FAIVKSYDVY  CKEQFNVIDG  RKSLMGDSKI  GGTHTNLFT
SpeK-B-2  YAFRXDEEI  FAIVKSYDVY  CKEQFNVIDG  RKSLMGDSKI  GGTHTNLFT
SpeK-B-3  YAFRXDEEI  FAIVKSYDVY  CKEQFNVIDG  RKSLMGDSKI  GGTHTNLFT
SpeK-MGAS6180  YAFRXDEEI  FAIVKSYDVY  CKEQFNVIDG  RKSLMGDSKI  GGTHTNLFT

C)  

25  35  45  65  75  125  155

SpeL-B  RDTIKKNKP  VEGKRGTQVI  IDAQHTKVWW  SNLYPSVEGK  FNVDQDIF  DGHKKDDTSK  KSPMN
SpeL-MGAS6323  RDTIKKNKP  VEGKRGTQVI  IDAQHTKVWW  SNLYPSVEGK  FNVDQDIF  DGHKKDDTSK  KSPMN

D)  

45  75

SpeM-B-1  RFKEEKVDI  GLIKTSREK
SpeM-B-2  RFKEEKVDI  GLIKTSREK
SpeM-MGAS8232  RFKEEKVDI  GLIKTSREK
FIG. 2 Alignments of amino acid sequences deduced from bovine group C streptococci superantigen alleles (SpeC-B, SpeK-B-1, SpeK-B-2, SpeK-B-3, SpeL-B, SpeM-B-1, SpeM-B-2) from the present study and of SpeC, SpeK, SpeL and SpeM sequences of *S. pyogenes* strains (MGAS8232 and MGAS6180) available in the National Center for Biotechnology (Bethesda, MD, USA) database, showing amino acid differences in **boldface** and **underlined**. Identities of 99% (A), 98-99% (B), 96% (C) and 99% (D) were observed among sequences. Nucleotide sequences length from which the alignments were originally created were of: 409 bp (*speC*); 649 bp (*speK*); 471 bp (*speL*); and 504 bp (*speM*).