

RESEARCH ARTICLE

Open Access

Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS)

Yumi Shimomura¹, Kayo Okumura^{1,4}, Somay Yamagata Murayama², Junji Yagi³, Kimiko Ubukata², Teruo Kirikae¹, Tohru Miyoshi-Akiyama^{1*}

Abstract

Background: *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) causes invasive streptococcal infections, including streptococcal toxic shock syndrome (STSS), as does Lancefield group A *Streptococcus pyogenes* (GAS). We sequenced the entire genome of SDSE strain GGS_124 isolated from a patient with STSS.

Results: We found that GGS_124 consisted of a circular genome of 2,106,340 bp. Comparative analyses among bacterial genomes indicated that GGS_124 was most closely related to GAS. GGS_124 and GAS, but not other streptococci, shared a number of virulence factor genes, including genes encoding streptolysin O, NADase, and streptokinase A, distantly related to SIC (DRS), suggesting the importance of these factors in the development of invasive disease. GGS_124 contained 3 prophages, with one containing a virulence factor gene for streptodornase. All 3 prophages were significantly similar to GAS prophages that carry virulence factor genes, indicating that these prophages had transferred these genes between pathogens. SDSE was found to contain a gene encoding a superantigen, streptococcal exotoxin type G, but lacked several genes present in GAS that encode virulence factors, such as other superantigens, cysteine protease *speB*, and hyaluronan synthase operon *hasABC*. Similar to GGS_124, the SDSE strains contained larger numbers of clustered, regularly interspaced, short palindromic repeats (CRISPR) spacers than did GAS, suggesting that horizontal gene transfer via streptococcal phages between SDSE and GAS is somewhat restricted, although they share phage species.

Conclusion: Genome wide comparisons of SDSE with GAS indicate that SDSE is closely and quantitatively related to GAS. SDSE, however, lacks several virulence factors of GAS, including superantigens, SPE-B and the *hasABC* operon. CRISPR spacers may limit the horizontal transfer of phage encoded GAS virulence genes into SDSE. These findings may provide clues for dissecting the pathological roles of the virulence factors in SDSE and GAS that cause STSS.

Background

Since Lancefield group G streptococci (GGS) have been considered components of the normal flora of the human throat, skin, and genitourinary tract, the contributions of GGS to streptococcal disease have often been overlooked [1]. Over the last decade, however, infections by pathogenic GGS have been reported, including life-

threatening invasive β -hemolytic streptococcal disease [1-7], making it necessary to expand our knowledge of the pathogenesis of GGS infections, especially invasive infections. Several species of streptococci can carry group C and G antigens, including *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE), *S. canis*, *S. dysgalactiae* subsp. *dysgalactiae*, *S. equi* subsp. *equi* (SESE), *S. equi* subsp. *zooepidemicus* (SESZ), and *S. anginosus* group bacteria. SDSE, which consists of Lancefield group G and C bacteria, in a ratio of about 4:1 [3,8,9], has been isolated from patients at higher frequency than

* Correspondence: takiyam@ri.ncgm.go.jp

¹Department of Infectious Diseases, National Center for Global Health and Medicine, 1-21-1, Toyama, Shinjuku-ku, Tokyo 162-8655, Japan
Full list of author information is available at the end of the article

other GGS and GCS species. For example, of 313 strains of GCS and GGS isolated from clinical samples in Southern India between 2006 and 2007, 254 (81.1%) were SDSE [9], as were 80% of the 266 invasive non-A and non-B β -hemolytic streptococcal isolates in the USA [3]. The spectrum and clinical courses of SDSE infection, including pharyngitis, cellulitis, infective arthritis, vertebral osteomyelitis, and streptococcal toxic shock syndrome (STSS), show substantial overlap with those of GAS [10-16]. Despite the increased clinical importance of SDSE, however, the entire SDSE genome has not yet been sequenced. Knowledge of its entire genome sequence is fundamental to gain insights into the pathogenicity of SDSE. We describe here the entire genome sequence of a Lancefield group G SDSE strain, GGS_124, which had been isolated from a patient with STSS.

Results

Selection of an SDSE isolate for genome sequencing

We chose a clinical isolate of SDSE, strain GGS_124, for genome sequence determination for several reasons. First, GGS_124 belongs to Lancefield group G, to which most clinical isolates of SDSE also belong [3,8,9]. Second, GGS_124 caused STSS in a patient. Third, GGS_124 was the most virulent strain among 8 group G SDSE isolates, as determined by their LD₅₀ values in a mouse infection model (Table 1).

Overview of the SDSE GGS_124 genome sequence

We found that, similar to other streptococcal genomes, the SDSE GGS_124 genome consists of a single circular chromosome of 2,106,340 bp (Additional file 1) and has a G+C content of 39.6% (Figure 1). Based on its location in the intergenic region upstream of the *dnaA* gene (SDEG_0001), the GC skew, and the clustering of *dnaA* box motifs, the start point of the SDSE GGS_124 genome was assigned to the putative origin of replication

(*oriC*). An AT-rich 13-mer (AGTCTGTTTTTTT), located in the intergenic region upstream of the *dnaA* gene [17], was selected as the starting point for nucleotide numbering. The GGS_124 genome was shown to contain 2095 predicted coding sequences (CDS), which account for 1.83 Mbp (86.9%) of the genome. In addition, this genome was shown to harbor 3 prophage-like elements, designated Φ GGS_124.1, Φ GGS_124.2, and Φ GGS_124.3. Moreover, there were 27 insertion sequence (IS) elements throughout the genome.

Genome sequence homology analysis of GGS_124 with the other 11 sequenced streptococcal species and subspecies showed that GGS_124 was closest in sequence to GAS, with 72% similarity (Additional file 1). GGS_124 was less similar to SESZ and SESE, with 65% and 64% coverage. Although *S. agalactiae* is the closest relative of SDSE based on 16S rRNA analysis, the *S. agalactiae* strains were less similar to SDSE than GAS based on the genome wide comparison (Additional file 1). In addition, we constructed a phylogenetic tree of all sequenced *Streptococcus* species based on the neighbor-joining method (Additional file 2). Although neighbor-joining methods are less accurate than the other methods such as most likelihood methods, SDSE is clustered with the GAS strains as their closest relative.

We compared the gene organization of GGS_124 with that of GAS by genomic rearrangement analyses (Figure 2 and Additional file 3). GAS could be divided roughly into 2 groups according to the orientation of the genes [18,19]. Therefore, SSI-1 and MGAS315, both of which are M3 serotype strains and have opposite gene orientations from each other, were chosen for the analysis. We found that the GGS_124 genome was organizationally more similar to that of GAS strain MGAS315 than GAS strain SSI-1 (Figure 2). Interestingly, the colinearity of GGS_124 and *S. uberis* genomes was quite remarkable but the percent amino acid identity was lower than that of the GAS strains (Additional file 3). The gene structure of GGS_124 was similar to the structures of GAS strain SSI-1, SESZ strain MGCS10565, and SESE strain 4047, although the GGS_124 genome contains large-scale genomic rearrangements. The GGS_124 genome differed markedly in gene organization from the genome of GBS strain A909.

When we compared genes from GGS_124 and two relatively homologous species, GAS (MGAS315) and SESZ (MGCS10565) (Figure 3), we found that these three streptococcal genomes contain more than 1,200 orthologous genes, accounting for 59% of the total CDSs of GGS_124. GGS_124 shares 282 genes with MGAS315 and 153 genes with MGCS10565. Moreover, 71.6% of the genes of GGS_124 were homologous to GAS genes, with 88.5% amino acid identity, whereas 66.5% of GGS_124 genes were homologous to MGCS10565

Table 1 *emm* types and mouse LD₅₀ values of 8 SDSE isolates used in this study

Strain	Origin	Symptom STSS/Non- STSS	LD ₅₀ value (CFU/head)	<i>emm</i> type
GGS_124	human	STSS	2.1 × 10 ⁶	stG480.0
168	human		4.6 × 10 ⁶	stG480.0
GGS_117	human	STSS	5.6 × 10 ⁶	stG4974.1
170	human		5.6 × 10 ⁶	stC36.0
164	human		1.9 × 10 ⁷	stG485
GGS_118	human	STSS	2.0 × 10 ⁷	stG67920
169	human		4.4 × 10 ⁷	stG11
163	human		4.5 × 10 ⁷	stG643

LD₅₀ values of the isolates were determined as described in MATERIALS AND METHODS.

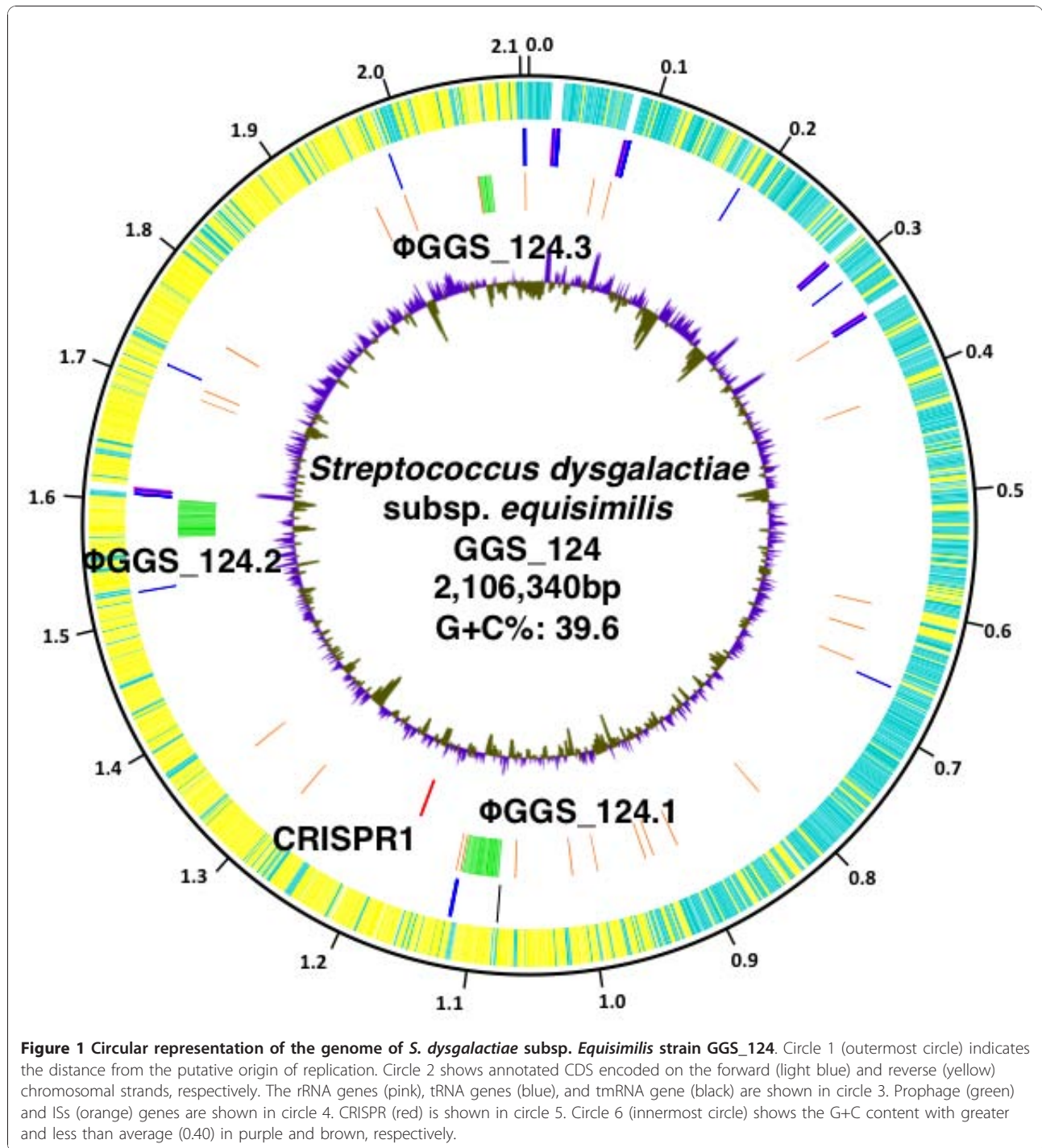


Figure 1 Circular representation of the genome of *S. dysgalactiae* subsp. *Equisimilis* strain GGS_124. Circle 1 (outermost circle) indicates the distance from the putative origin of replication. Circle 2 shows annotated CDS encoded on the forward (light blue) and reverse (yellow) chromosomal strands, respectively. The rRNA genes (pink), tRNA genes (blue), and tmRNA gene (black) are shown in circle 3. Prophage (green) and ISs (orange) genes are shown in circle 4. CRISPR (red) is shown in circle 5. Circle 6 (innermost circle) shows the G+C content with greater and less than average (0.40) in purple and brown, respectively.

genes, with 79.9% amino acid identity. These findings indicate that SDSE is closely related to GAS in both nucleotide and amino acid sequences.

We also analyzed the distribution of genes shown to be more homologous to genes derived from bacteria other than GAS (Additional file 4). We found that 299 genes showed higher similarity to genes from *Streptococci* other

than GAS and 92 genes showed higher similarity to genes from a genus other than *Streptococcus*. In addition, we identified 11 genes that did not show significant homology to any genes in the databases. These genes were scattered throughout the entire GGS_124 genome, suggesting that they had not been acquired by massive genome recombination.

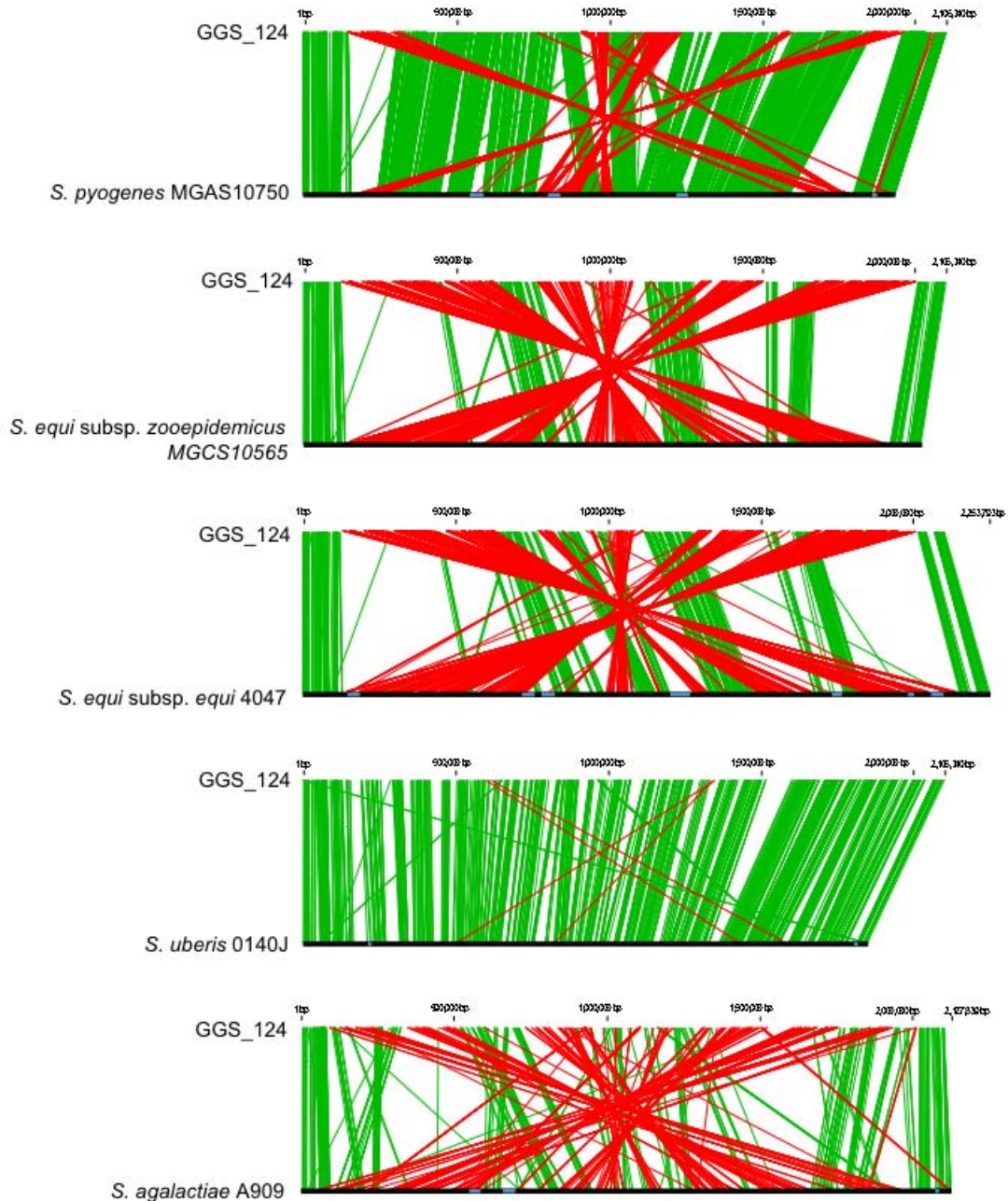


Figure 2 Genome rearrangement maps of *S. dysgalactiae* subsp. *equisimilis* GGS_124 with five species in the pyogenic group.

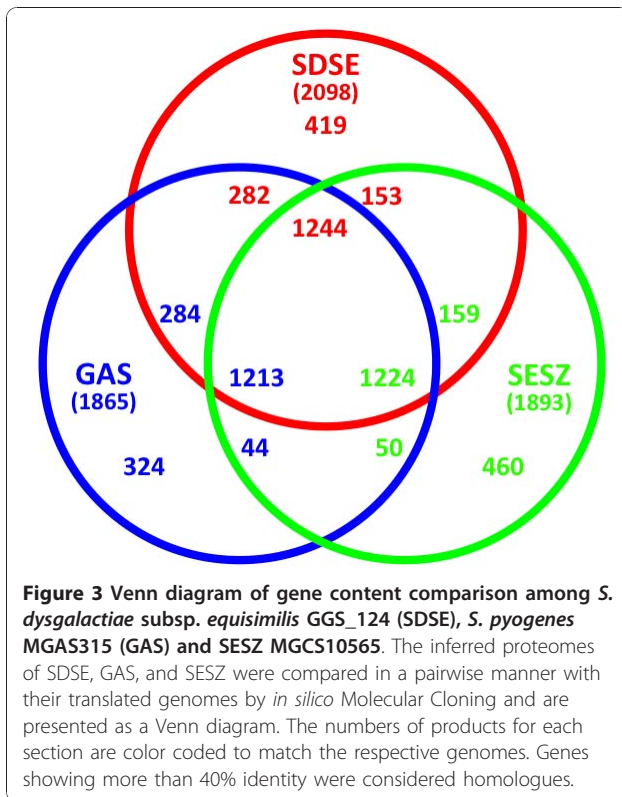
Sequences were aligned from the predicted replication origin of each genome. The colored bars separating each genome (red and green) represent similarity matches identified by *in silico* Molecular Cloning. Links shown in green match in the same orientation, while those in red match in the reverse orientation. Prophages are highlighted as pale blue boxes.

Putative prophages and CRISPR/Cas

We found that all three prophage-like elements of GGS_124 were homologous to previously sequenced GAS prophages, and that they were integrated at sites similar to those of GAS strains, with the same upstream and downstream genes (Figure 4).

(i) Prophage GGS_124.1

We found that the Φ GGS_124.1 prophage is 35,593 bp in length with a G+C content of 38.04% and carries 60 CDS. Ninety-seven percent of the CDS in Φ GGS_124.1 have homologues, with more than 40% identity to GAS prophages, suggesting that Φ GGS_124.1 is a chimeric



phage. This prophage was inserted at the predicted bacteriophage T12att site, which has been shown to be a gene that encodes a serine tRNA and is located between the CDS of SDEG_1100 and SDEG_1161 [20]. Six GAS strains, MGAS10394, MGAS315, MGAS5005, MGAS6180, MGAS8232, and SSI-1, have prophage elements: Φ 10394.3, which carries *speK* and the streptococcal phospholipase A2 gene (*sla*); Φ 315.2, which carries *ssa*; Φ 5005.1, which carries *speA*; Φ 6180.1, which carries the *speC* and Dnase (*spd*) genes; Φ 8232.3, which carries *speL* and *speM*; and SPsP5, which carries *speC*, respectively [18,21-25]. In addition, Φ GGG_124.1 was found to contain a prophage-associated virulence factor gene for deoxyribonuclease (*sdC*).

(iii) Prophage GGS_124.2

We found that the Φ GGG_124.2 prophage is 35,814 bp in length, with a G+C content of 38.20% and 61 CDS. Ninety-five percent of the CDS in Φ GGG_124.1 have homology with genes in GAS prophages, making it likely that Φ GGG_124.2 is chimeric phage. The chromosomal phage attachment site (*attB*) and the Φ GGG_124.2 phage-encoded attachment site (*attP*) were not found, but the products of *attP/attB* recombination, *attL* and *attR*, with the same sequences as those of GAS prophages SPsP2 and Φ 315.5 were identified. The genome context around the integration site for Φ GGG_124.2 was found to be conserved at the phage integration sites of 4 GAS

strains, MGAS10394, MGAS315, SSI-1, and Manfredo, which contain the prophage elements Φ 10394.6, carrying *sdn*; Φ 315.5, carrying *speA*; SPsP2, carrying *speA*; and Φ Man.1, carrying the DNase gene *mf3*, respectively [18,19,21,22]. No known prophage-associated virulence factor genes were found in Φ GGG_124.2.

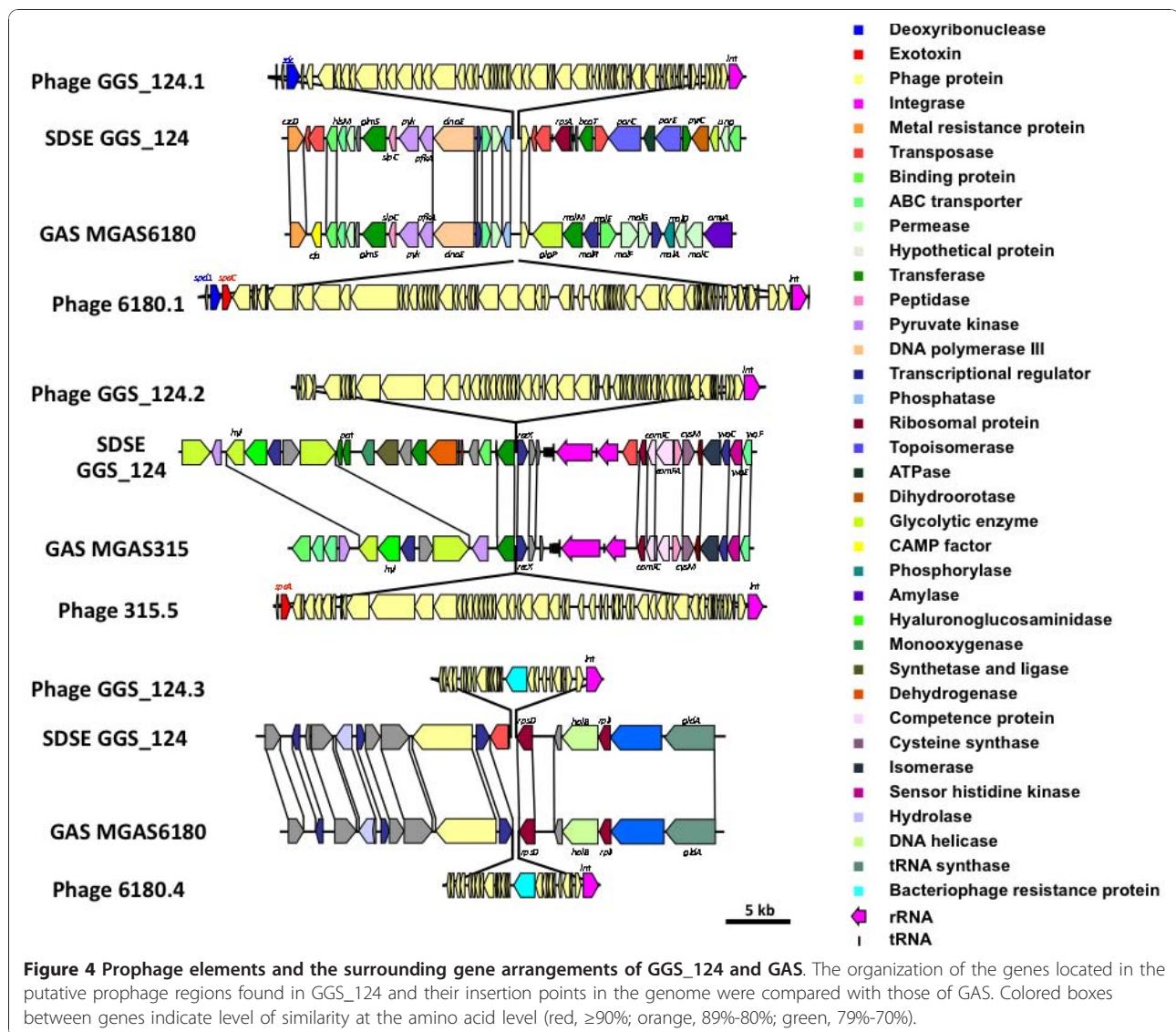
(iii) Prophage GGS_124.3

We also found a prophage remnant, Φ GGG_124.3, which was about 12.6 kb length and homologous to the previously sequenced GAS prophage remnants Φ 6180.4 and Φ 10270.5 with a nucleotide identity of 73%. Φ GGG_124.3 and the two GAS phage remnants were found to be located between genes encoding a putative transcriptional regulator protein and the 30S ribosomal protein. In strain GGS_124, two truncated transposase proteins, SDEG_2117 and SDEG_2118, were found to be inserted upstream of GGS_124.3. No virulence factor genes are present in GGS_124.3.

Prokaryotes possess a system that mediates resistance to infection by foreign DNA, such as viruses [26,27]. When bacteria are exposed to phages, short fragments derived from phage DNA are integrated into clusters of regularly interspaced short palindromic repeat (CRISPR) regions of the bacterial genome as spacers [27]. CRISPR RNA transcripts and CRISPR-associated proteins (Cas), act in complexes to interfere with virus proliferation [26]. This system has also been observed in GAS [20], SESZ [22,28], *S. mutans* [29], and *S. thermophilus* [30]. GGS_124 harbors a CRISPR/Cas system consisting of an array of genes, *can1*, *cas1*, *cas2*, and *csn2*, and CRISPR (Figure 5). The GGS_124 CRISPR has 19 direct repeats of 36 bp each and 18 spacer sequences 30 or 32 bp in length; 6 of these sequences are homologous to GAS prophage sequences, with more than 80% coverage (Additional file 5). When we analyzed the number of CRISPR spacers in an additional 7 SDSE isolates (Table 2), we found that the mean number of CRISPR spacers was higher in SDSE (18.0 ± 3.3 spacers) than in GAS strains (4.0 ± 1.0 spacers; range, 0 to 9) (Table 2). These results suggest that prophage infection of SDSE is somewhat restricted, resulting in a smaller number of virulence factors located in the prophage regions of SDSE. Alternatively, SDSE may be in contact with phages more frequently, with integrated phages having a fitness cost for SDSE.

Virulence factors encoded by the GGS_124 genome

An analysis of 58 SDSE strains isolated from human infections using targeted microarrays containing 216 GAS virulence genes composed of 70mer oligonucleotides showed that about 50% of the GAS virulence genes represented in the microarray were present in SDSE [31]. Three molecular markers, *speB*, the intergenic region upstream of the *scpG* gene and *virPCR*, have



been shown helpful in discriminating between GAS and SDSE [32]. Based on homology analyses with known bacterial virulence factors, such as pore-forming toxins, a superantigen, proteases, FCT-like regions, adhesins, hyaluronidase, and a nuclease, we identified 71 putative virulence factor genes in the GGS_124 genome; their details are shown in Additional file 6. Comparison of the virulence factors in GGS_124 with those of other streptococcal species indicated that the putative virulence factors most similar to those of GGS_124 were found in GAS. In contrast, superantigen, SPE-B and the *has* operon of GAS are not conserved in GGS_124.

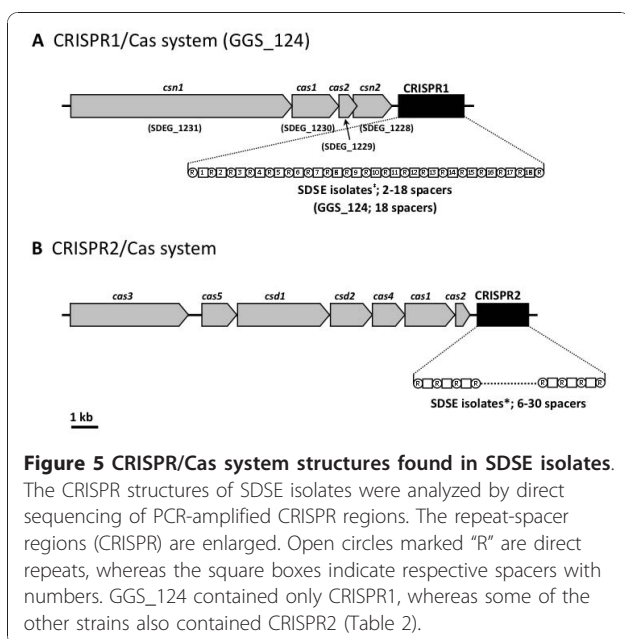
(i). Pore-forming toxins

GGS_124 has several putative hemolysins, including HlyX (SDEG_0427), HlyIII (SDEG_1015), and HlyA1 (SDEG_1483), which have also been detected in GAS, SESZ, SESE, *S. uberis*, and GBS. GGS_124 also has

genes encoding streptolysin S (*sagA*) (SDEG_0705) and its biosynthesis proteins (*sagBCDEFGHI*) (SDEG_0706 to 0713), which are also present in GAS [33], SESZ, and SESE [22,28,34]. In addition, GGS_124 carries a gene for streptolysin O (SLO) (SDEG_2027), which is essential for GAS virulence and is required for the organism to escape from the endosome into the cytosol following invasion of host cells [35].

(ii). Superantigen

GGS_124 possesses only one superantigen gene, exotoxin G variant 4 (*spegg4*), which is homologous to GAS streptococcal exotoxin G (SpeG), with 79% amino acid identity (Additional file 6). In a previous analysis of the superantigenic activities of the *spegg4* product in human peripheral blood mononuclear cells [36], we found that its mitogenic activity was about 1% that of SpeG from GAS. Other genome-encoded superantigen



genes for mitogenic exotoxin Z (*smeZ*), which are present in GAS [37], were not found in the GGS_124 genome.

(iii) Proteases

We found that a putative proteinase (SDEG_1906) and streptococcal C5a peptidase (*scpB*) (SDEG_0933) [38] were conserved among GGS_124 and 5 closely related species. GGS_124 also has a gene with homology (42% amino acid identity) to exfoliative toxin A of *Staphylococcus aureus* strain Mu50 (SAV1173), which causes staphylococcal scalded skin syndrome [39]. GGS_124 also carries a gene for streptokinase (SDEG_0233), similar to streptokinase A of GAS, with 88% amino acid identity (Additional file 6). This protein complexes with plasminogen to form an activator, which has serine protease activity and cleaves free plasminogen, leading to activation of the zymogen [40]. Strikingly, GGS_124 lacks streptococcal cysteine protease (SpeB), an erythrogenic toxin produced by GAS with cysteine protease activity [41]. The GGS_124 genome lacks approximately 7 kb of

Table 2 Presence of Cas genes and the number of spacers in CRISPR1/Cas and CRISPR2/Cas

Species	Strain	CRISPR1/Cas			CRISPR2/Cas		
		Cas genes	No. of spacers	Acc. No.	Cas genes	No. of spacers	Acc. No.
<i>Streptococcus dysgalactiae</i> subsp. <i>equisimilis</i>	GGG_124	+	18	AP010935.1	-	0	-
	168	+	2	AB553332	+	13	AB553333
	GGG_117	+	8	AB553338	+	12	AB553339
	170	+	9	AB553336	+	10	AB553337
	164	+	17	AB553343	+	6	AB553331
	GGG_118	+	8	AB553342	+	13	AB553341
	169	+	7	AB553334	+	30	AB553335
	163	+	3	AB553340	N. D.	N. D.	
<i>Streptococcus pyogenes</i>	MGAS8232	-	0	AE009949.1	-	0	OAE009949.1
	MGAS10394	-	0	CP000003.1	-	0	CP000003.1
	MGAS10750	+	0	CP000262.1	+	5	CP000262.1
	Manfredo	-	0	AM295007.1	-	0	AM295007.1
	MGAS10270	+	2	CP000260.1	+	3	CP000260.1
	MGAS315	+	0	AE014074.1	-	0	AE014074.1
	MGAS5005	+	3	CP000017.1	+	4	CP000017.1
	MGAS9429	+	0	CP000259.1	+	7	CP000259.1
	MGAS2096	+	0	CP000261.1	+	6	CP000261.1
	SF370	+	6	AE004092.1	+	3	AE004092.1
	SSI-1	+	0	BA000034.2	-	0	BA000034.2
	MGAS6180	+	4	CP000056.1	+	1	CP000056.1
	NZ131	+	4	CP000829.1	+	5	CP000829.1
	<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	MGCS10565	+	17	CP001129.1	+	9
H70		-	0	FM204884.1	+	18	FM204884.1
<i>Streptococcus equi</i> subsp. <i>equi</i>	4047	-	0	FM204883.1	-	0	FM204883.1

N.D.: No amplicon was obtained in PCR analyses.

the GAS strain MGAS315 sequence, including genes encoding SpeB (SpyM3_1742), the transcriptional regulator RopB (SpyM3_1744), and mitogenic factor 25K precursor (SpyM3_1745). Since several transposase and related genes (SDEG_0212, 0206, 0205, 0201, 0194) are located in the corresponding region, it is highly likely that the region that included *speB* was present in the common ancestor of GAS and SDSE but was not retained by SDSE after speciation.

We found that *speB* was not present in GGS_124, in agreement with the results of a microarray study, which showed that all of the 58 examined strains of group C and G SDSE isolated from patients lacked the *speB* gene [31,32]. We therefore examined whether SDSE strains have protease activity similar to that of SpeB (Additional file 7). We did not detect any SpeB-like protease activity in strains GGS_124 or GGS_118, which had been isolated from two patients with STSS. In contrast, a GAS strain produced a proteinase that was sensitive to E-64, which inhibits cysteine proteases, including SpeB.

(iv) FCT-like regions

Recently, GAS and GBS were shown to express pili, which are synthesized by proteins encoded by genes in FCT regions [1,42,43]. GGS_124 harbors 2 FCT-like regions, which are probable operons expressing different pilus-like structures (Figure 6). One of these contains genes encoding the transcriptional regulator RofA (SDEG_0156), two putative fimbrial structural subunit proteins (SDEG_0157 and SDEG_0158), two sortases (SDEG_0159 and

SDEG_0160), and a putative fibronectin binding protein (SDEG_0161). It is similar to the FCT-6 region, which is conserved among M2 GAS, GBS, and SESZ [22,42,44]. The second region contains genes encoding a putative transcriptional regulator (SDEG_1782), a defective collagen binding protein (SDEG_1781), a signal peptidase I (SDEG_1780), a backbone protein (SDEG_1779), and an ancillary protein (SDEG_1778). It is similar to the FCT-3 region, which was found in M3, M5, M18, and M49 GAS [43].

(v) Adhesins

GGS_124 possesses genes that encode putative adhesion proteins, including proteins similar to putative fibronectin binding proteins (SDEG_0161, 1263, and 1984), pullulanase (SDEG_0237), phosphopyruvate hydratase (SDEG_0704), laminin binding protein (SDEG_0935), internalin protein (SDEG_1372), and collagen binding protein (SDEG_1781), all of which bind to the extracellular matrix (Additional file 6). GGS_124 also possesses genes encoding immunoglobulin G binding protein (SDEG_1358) [45] and multifunctional streptococcal plasmin receptor (Plr)/streptococcal surface dehydrogenase (SDH)/glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which binds to complement component C5a (SDEG_1936) [46] (Additional file 6), although the product of SDEG_1936 lacks a signal peptide.

(vi) Hyaluronan capsule synthesis

Hyaluronic acid (HA), synthesized via the *hasABC* operon, is considered a pleiotropic virulence factor involved in

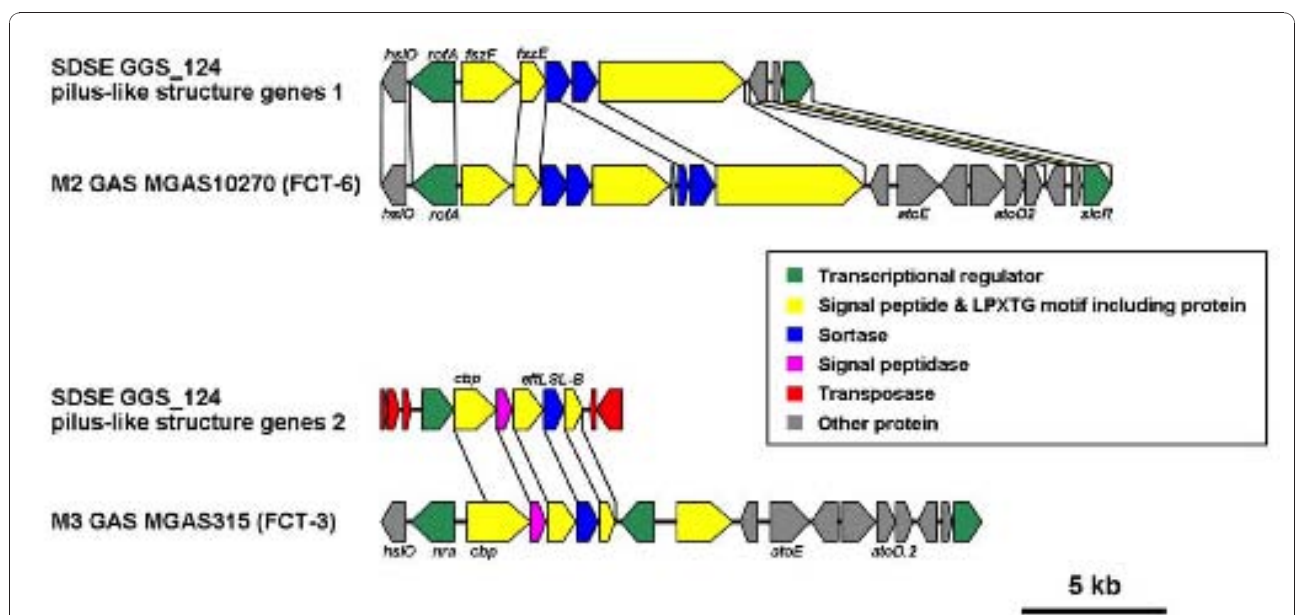


Figure 6 Organization of genes encoding pilus-like structure proteins in *S. dysgalactiae* subsp. *equisimilis* GGS_124. The organizations of genes encoding pilus-like structure proteins in GGS_124 were compared with those in GAS and SESZ. Colored boxes between genes indicate similarity at the amino acid level (red, ≥90%; orange, 90%-80%; yellow-green, 79%-70%; green, 69%-60%; purple, 59%-40%; gray, <40%).

many aspects of GAS infection [47]. GGS_124, however, does not contain an *hasABC* operon, in contrast to the genomes of GAS, SESZ, SESE, and *S. uberis*. Rather, GGS_124 possesses only one gene, encoding glycosyl transferase (SDEG_0628), which shows a low level of similarity to *hasA* of GAS (20% amino acid similarity). Although one SDSE strain has been shown to possess a hyaluronan synthase (AF023876.1), very similar to the product of *hasA* [48], GAS gene microarray analysis of 58 SDSE strains isolated from human infections showed that all harbored only *hasC* [31]. We found that GGS_124 also harbors only *hasC* (SDEG_1980) (Additional file 6), making it unlikely that SDSE produces HA *via* the *hasABC* operon.

(vii) Hyaluronidase

GGS_124 possesses a gene in a non-prophage region of the genome that encodes a putative hyaluronidase (SDEG_0654), with 66% identity to *hylB* in SESZ (Additional file 6). Hyaluronidase in GAS is thought to contribute to the spread of bacteria in tissues and may allow GAS to utilize host hyaluronic acid or its own capsule as an energy source [49]. The hyaluronidase in GGS_124 may have a function similar to that in GAS.

(viii) Nucleases

GGS_124 possesses 5 genes that encode putative nucleases with a secretion signal peptide: genome-encoded streptodornase (SDEG_0541), extracellular nuclease (SDEG_0714), DNA-entry nuclease (SDEG_0732), cell surface 5'-nucleotidase (SDEG_0825), and prophage-associated deoxyribonuclease (SDEG_1103), all of which are predicted to code for a secretion signal peptide. Two of them, SDEG_0714 and SDEG_0825, code for potential cell wall anchor motifs, LPKAG and LPMAG, respectively (Additional file 6). The putative streptodornase SDEG_0541 and DNA-entry nuclease SDEG_0732 are homologous to phage-encoded extracellular streptodornase D Sda1 of GAS (PHA01790) [50] and DNA-entry nuclease EndA of *S. pneumoniae* TIGR4 (SP_1964) [51], respectively (Additional file 6). Sda1 and EndA have been found to degrade neutrophil extracellular traps (NETs) [50,51], which are composed of granule proteins and chromatin released by neutrophils and can catch and kill surrounding bacteria [52]. The putative extracellular DNase SDEG_0714 is similar to M1 GAS cell-wall-located DNase SpnA (Spy0747), which has been reported important for virulence, *e.g.*, dispersion in host tissue [53] (Additional file 6).

(ix) Other virulence factors

GGS_124 possesses genes encoding the multifunctional M protein (*stg480.0*) (SDEG_0230). The M protein of GAS shows antiphagocytic and adhesin activities, whereas the adhesion function of the GGS M protein may be due to a collagen binding motif [54,55]. Since the product of *Stg480.0* lacks this motif, the M protein of GGS_124 may not act as an adhesin.

Streptococcal inhibitor of complement (SIC) and distantly related to SIC (DRS) are some extent of homology. DRS binds complement factors but does not inhibit complement mediated cell lysis [56,57], whereas SIC inhibits complement mediated cell lysis [58]. GGS_124 harbors a putative DRS gene (SDEG_0932), which consists of a signal sequence, two repeat regions, and a proline-rich region typical of DRS, and is homologous to the Drs12.04 protein of GAS strain *emm12* with 48% amino acid identity [59] (Additional file 6).

The GGS_124 genome harbors a gene encoding a collagen-like protein (SDEG_1113), similar to the collagen-like repeat phage protein of SESE 4047 (SEQ_0837), with 41% amino acid identity. Streptococcal collagen-like proteins (Scl) are cell-surface molecules of GAS with domains containing tracks of repeated Gly-Xaa-Yaa sequences that form a mammalian collagen-like triple-helical structure. These proteins mediate the internalization of GAS into human cells upon binding of Scl to the human collagen receptor integrin [60]. The GGS_124 gene encoding collagen-like protein does not contain a signal peptide or LPXTG motif, suggesting that the gene product may not be expressed on the cell surface. In contrast, GGS_124 does not harbor genes encoding proteins similar to the other collagen-like proteins (*sclA* and *sclB*) in GAS.

NAD glycohydrolase (SDEG_2029), which is located in the NADase-streptolysin O operon of the GAS genome [61], was found to be conserved in the same operon in GGS_124 (Additional file 6). This enzyme is expressed after streptolysin O is injected into host cells and accelerates cell death [61,62].

(x) Distribution of virulence factors among Streptococci

We also assessed the presence or absence of representative virulence factors among sequenced streptococcal species, including GAS (MGAS315), SESE (4047), SESZ (MGCS10565), GBS (A909), and *S. uberis* (0140J) (Additional file 8). Among 30 virulence factors, most of those located in the core genome, but not those located in streptococcal phages, are conserved in GGS_124, except for *speB*. In contrast, other *streptococci* lack genes encoding streptolysin O, NAD glycohydrolase and DRS (or SIC), suggesting the importance of these proteins in the pathogenicity of SDSE and GAS in humans, causing STSS.

Putative virulence factors unique to SDSE

We identified 20 gene products in GGS_124 containing signal peptides and LPXTG cell wall surface anchor motifs that showed low similarity to known proteins. Using PCR, we analyzed the distributions of these putative virulence factors in 8 SDSE isolates (Table 3 arranged according to their decreasing lethality in mice). A putative T-antigen-like protein structural subunit

Table 3 Putative virulence factors found in GGS_124 and their prevalence in the SDSE isolates based on the results of PCR analyses

<i>S. dysgalactiae</i> subsp. <i>equisimilis</i> GGS_124					Best hit strain		PCR analysis								
Locus tag	Product name	Length (aa)	LPXTG motif	% Identity	Strain	Product name	Reference sequence	GGS_124	168	GGS_117	170	164	GGS_118	169	163
	fimbrial subunit				<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	fimbrial subunit									
SDEG_0157	Protein	645	IPNTG	40.48	MGCS10565	protein	YP_002124169.1	Yes	x	x	x	x	x	x	x
	T-antigen-like fimbrial structural subunit				<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	T-antigen-like fimbrial structural subunit									
SDEG_0158	protein	315	IPKTG	49.84	MGCS10565	protein	YP_002124168.1	Yes	Yes	Yes	x	x	x	x	x
SDEG_0180	hypothetical protein	184		-	No hit	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	hypothetical membrane associated				<i>Streptococcus pyogenes</i>	hypothetical membrane-associated									
SDEG_0267	protein	242		44.35	MGAS2096	Protein	YP_601272.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SDEG_0574	cell surface serine endopeptidase	216		49.36	<i>Streptococcus Agalactiae</i>	CspA	AAN85092.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	cell wall surface				<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	cell wall surface									
SDEG_0805	anchor family protein	486	LPKAG	43.73	MGCS10565	anchor family protein	YP_002123384.1	Yes	Yes	x	x	Yes	x	x	Yes
	histidine triad protein				<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	histidine triad protein									
SDEG_0918	histidine triad protein	153		51.35	MGCS10565	HtpA	YP_002123384.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	complement inhibitor				<i>Streptococcus pyogenes</i>	complement inhibitor									
SDEG_0932	protein	226		47.92	MGAS2096	protein	YP_601343.1	Yes	x	x	x	x	Yes	x	x
	nisin resistance				<i>Streptococcus agalactiae</i> 2603V/R	nisin resistance									
SDEG_0979	protein, putative	322		42.17	<i>Streptococcus agalactiae</i> 2603V/R	Protein	NP_687984.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SDEG_1327	YaeC family protein	280		49.47	<i>Streptococcus agalactiae</i> 2603V/R	YaeC family protein	NP_687791.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	probable surface antigen negative regulator				<i>Streptococcus suis</i> 98HAH33	probable surface antigen negative regulator Par	YP_001200806.1	Yes	x	Yes	Yes	Yes	Yes	x	Yes
SDEG_1429	regulator	185		52.22	<i>Streptococcus pyogenes</i>	regulator Par									

Table 3 Putative virulence factors found in GGS_124 and their prevalence in the SDSE isolates based on the results of PCR analyses (Continued)

SDEG_1480	hypothetical protein	299	LPVTG	33.33	MGAS6180	hypothetical protein	YP_280631.1	Yes	x	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SDEG_1511	hypothetical protein	546		27.59	<i>Streptococcus suis</i>	hypothetical protein	ABQ42885.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
					<i>Streptococcus uberis</i>											
SDEG_1573	adhesion protein	661	LPKTG	38.8	UT888	adhesion protein	ABB52003.1	Yes	Yes	x	x	x	x	x	x	Yes
SDEG_1601	hypothetical protein	249		-	No hit	-	-	Yes	x	Yes	x	x	Yes	x	x	
					<i>Streptococcus pyogenes</i>											
SDEG_1773	hypothetical protein	210	FPSTG	37.96	M1 GAS	hypothetical protein	NP_268519.1	Yes	Yes	Yes	x	x	Yes		Yes	Yes
					<i>Streptococcus sanguinis</i>											
SDEG_1969	hypothetical protein	234		32.77	SK36	hypothetical protein	YP_001035903.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	protein F2-like fibronectin binding				<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>	fibronectin binding										
SDEG_1984	Protein	528	LPATG	40.46	ATCC 35246	protein	ABC87919.1	Yes	x	x	Yes	Yes	x		Yes	Yes
SDEG_2022	hypothetical protein	106		-	No hit	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SDEG_2141	hypothetical protein	175		-	No hit	-	-	Yes	x	x	x	Yes	x	x	x	Yes

Genes encoding proteins, which contain predicted signal peptide based on SignalP <http://www.cbs.dtu.dk/services/SignalP> and the LPXTG motif based on HMMER http://bamics3.cmbi.kun.nl/jos/sortase_substrates/help.html and were not highly homologous to putative virulence factors of GGS_124. Distribution of the putative virulence factors of GGS_124 in the 16 SDSE strains was analyzed by PCR, and the results are presented as SDSE strains in order of decreasing lethality in mice. Yes: PCR analysis positive; x: PCR analysis negative.

(SDEG_0158), encoded in the FCT-6-like region in GGS_124, was detected in the 3 most virulent strains, GGS_124, GGS_168, and GGS_117. In both GGS_168 and GGS_117, the FCT-6-like region was not detected by PCR (data not shown). Interestingly, the SDEG_1601 gene encoding a functionally unknown hypothetical protein was amplified in isolates that cause STSS. However, no virulence factors associated with mouse lethality or Lancefield groups were found.

Relatedness of *emm* type and pathogenicity in SDSE

Particular M (or *emm*) types of GAS have been associated with certain streptococcal diseases [63]. In investigating the pathogenic traits of M4, M12, M1, and M3 GAS clinical isolates in a murine model [64], we found that murine lethality was closely associated with M type. The M1 and M3 types of GAS, which are isolated at high frequency from patients with STSS, showed higher virulence in mice than did M12 and M4. Using this mouse model, we assessed the virulence of the *emm* types of 8 group G SDSE isolates to analyze whether most frequent *emm*-type of isolates in epidemiological studies (see below) are more virulent in mice (Table 1). We found that the lethality in mice of these isolates was not associated with their isolation from patients with STSS or with the frequently isolated *emm*-type in humans such as stG10.0 and stG643.0 (see below). To determine whether the *emm* type of SDSE was associated with pathogenicity in humans, we reviewed the epidemiological data regarding the isolation frequencies of GCS and GGS in relation to their *emm* types [3,6,7,65,66] (Additional file 9). We found that stG10.0 was the most frequent *emm* type in Portugal and Japan, stG643.0 was the most frequent in western Norway, and stG6.0 was the most frequent in the USA. When we calculated the degree of correspondence among the isolation rates of *emm* types in each area using the Kendall tau rank correlation coefficient, we found no significant correlations among all regions (data not shown), suggesting that there is no linkage between *emm* type of SDSE and infectivity in humans or mice.

Discussion

Comparative analysis of *Streptococcus* 16S rRNA sequences had indicated that SDSE was more closely related to GBS than to GAS [67,68]. In contrast, we found that SDSE, which belongs to Lancefield groups C and G [3,65], is more closely related to GAS than to other sequenced streptococci based on genome wide and gene level comparisons

SDSE is known to cause diseases very similar to those caused by GAS, such as pharyngitis, cellulitis, infective arthritis, vertebral osteomyelitis, and STSS [10-16]. This similarity may be due, at least in part, to their

conservation of a large number of genes for virulence factors. Figure 7 shows a summary of putative virulence factors and proposed virulence functions in SDSE. SDSE shares most of the virulence factor genes of GAS, including streptolysin O, streptokinase, FCT-like regions, NADase, and DRS. However, GGS_124 and probably almost all other SDSEs lack SpeB, superantigens except for SpeG, and hyaluronan synthesis *via hasABC* [31]. As *spegg4*, which has about 1% of the mitogenic activity of GAS-derived SpeG, is the only gene encoding a protein homologous to superantigens, and most SDSE isolates do not harbor superantigen-like genes other than *spegg* [31,69], it is highly unlikely that superantigens play a significant role in the pathogenesis of SDSE infection in humans.

Particular M (or *emm*) types of GAS have been associated with certain streptococcal diseases [63]. We therefore analyzed the virulence in a mouse model of several SDSE strains bearing different *emm*-types, but we were unable to find a significant correlation between *emm*-type of SDSE and virulence. Further studies are needed to provide further insight into the linkage between *emm* type of SDSE and infectivity in humans.

Streptococcal phages are considered critical in horizontal gene transfer, especially in the transport of virulence factors [28,70], in *Streptococci*. Three prophage elements in GGS_124 were found to be homologous to GAS prophages in both nucleotide and amino acid sequences. The positions of prophage insertion were also conserved between GGS_124 and previously sequenced GAS prophages, suggesting that SDSE and GAS share the same phage species, and that horizontal gene transfer between SDSE and GAS has occurred. However, GGS_124 does not contain prophages that encode genes for superantigens, Sla, or MF. These phage encoded genes were also missing from all strains previously analyzed by a GAS microarray [31]. Sdn is an exception, since it was detected in 4 of these 58 SDSE strains [31]. These results showed that SDSE may have some resistance to infection by GAS phages carrying genes encoding virulence factors.

Prokaryotes possess the CRISPR/Cas system, which mediates resistance to infection by foreign DNA, such as viruses [26,27]. GGS_124 has a CRISPR/Cas system, designated CRISPR1/Cas, whereas the other SDSE isolates analyzed in this study had another CRISPR/Cas system, designated CRISPR2/Cas. We found that SDSE strains usually have a higher total number of spacers than GAS, suggesting that prophage infection of SDSE was restricted to some extent, resulting in a smaller number of virulence factors located in the prophage regions of SDSE. Similar restrictions were observed in SESZ when compared with SESE. For example, the SESE 4047 genome, which contains no CRISPR, contains genes encoding virulence factors in prophage regions. In contrast, SESZ MGCS10565

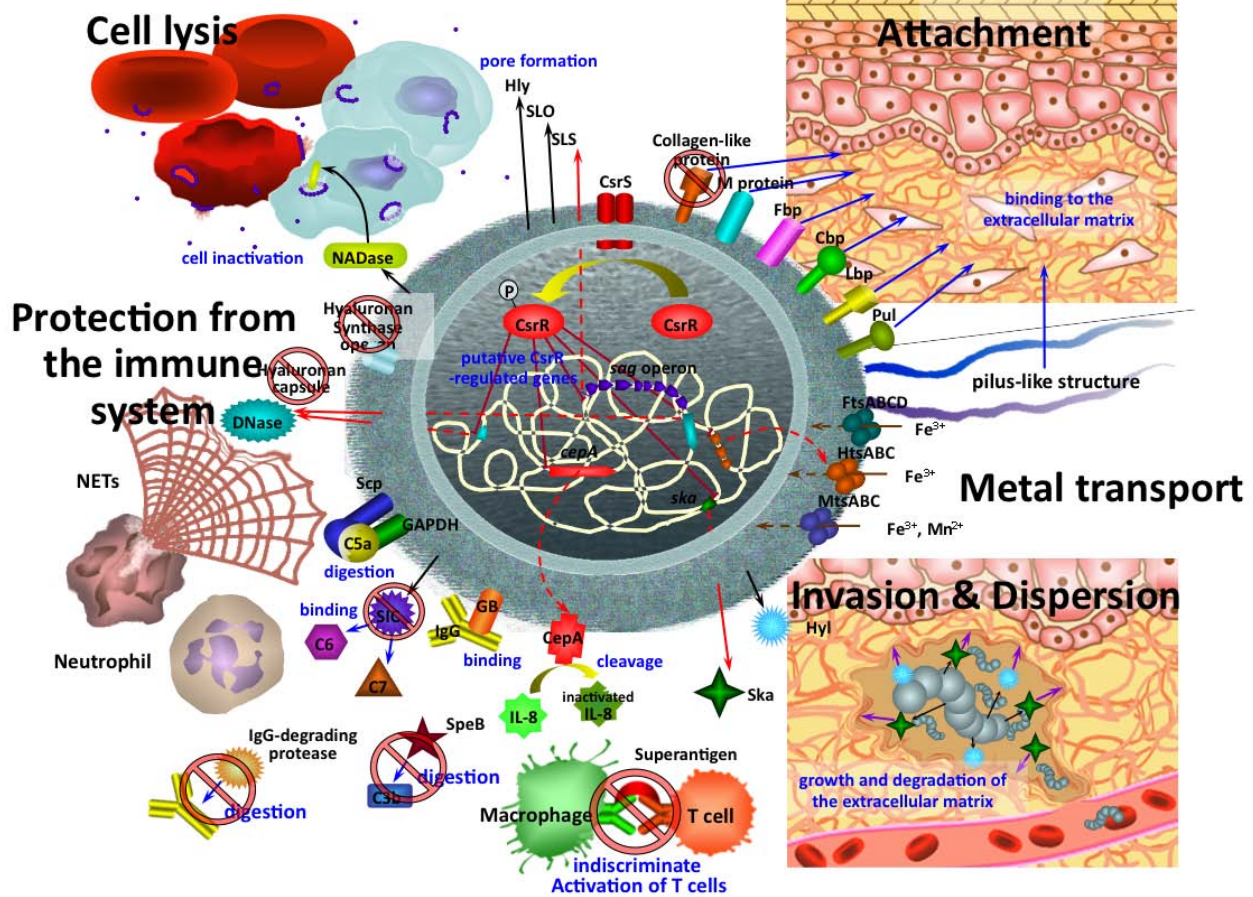


Figure 7 Putative virulence factors and posited virulence function of SDSE. Cell-surface proteins, extracellular secreted proteins, metal transporters, and the two-component regulator CsrR/CsrS, which affect the expression of approximately 10% of all genes in a GAS strain [78,79], are shown. The putative virulence factors on the cell surface of SDSE are adhesins, including M protein, fibronectin binding protein (Fbp), collagen binding protein (Cbp), laminin binding protein (Lbp), and pullulanase (Pul) [80]. Factors that protect bacteria from the host immune system are shown, including cell envelope proteinase A (CepA), which cleaves within the interleukin-8 (IL-8) C-terminal α -helix [38]; immunoglobulin G binding protein (GB) [49]; glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and C5a peptidase (Scp) [46]. Putative pathogenic roles of the virulence factors of SDSE, including hyaluronidase (Hyl), streptokinase (Ska), extracellular nuclease and streptodornase (DNase), which digest neutrophil extracellular traps (NETs) released from dead neutrophils [50], and the pore-forming proteins, such as hemolysin (Hly), streptolysin S (SLS), streptolysin O (SLO), and NAD glycohydrolase (NADase), are indicated. FtsABCD and HtsABC are ferrichrome transporters and MtsABC is a metal transporter. Black arrows show protein secretion, red arrows show expression of genes regulated by CsrR, blue arrows show protein attachment to the extracellular matrix, brown arrows show metal transport from the extracellular environment into the cell, and purple arrows show degradation of extracellular matrix by secreted Hyl or Ska. The factors marked with a Stop sign, which are major virulence factors of GAS, do not function in SDSE.

and H70, which contain 26 and 18 spacers, respectively, do not carry any prophages. Thus, the CRISPR system in streptococci sharing prophages may play a substantial role in the spread of virulence factors among species. Alternatively, these virulence factors may not benefit to SDSE during carriage or disease, such that the integration of these specific phages is not selected for.

Conclusions

We have shown here that SDSE likely acts as a pathogen, based on its genome sequence and close relationship with

GAS. As the frequency of isolation of SDSE from patients has increased, it should not be overlooked as a source of infection.

Methods

Bacterial strains

Three *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) isolates, GGS_124, GGS_117, and GGS_118, were obtained from 3 patients with STSS, and 5 SDSE isolates were isolated from 5 patients with non-STSS (Table 1). All SDSE isolates were classified as Lancefield

group G. *Streptococcus pyogenes* (GAS). NIH9 [71] was used as the SpeB-producing strain.

emm typing

The *emm* types were classified according to a protocol for *emm* typing of the Centers for Disease Control and Prevention (CDC) http://www.cdc.gov/ncidod/biotech/strep/protocol_emm-type.htm.

SDSE infection in mice

All animal experiments were performed according to the guidelines of the Ethics Review Committees of Animal Experiments of Tokyo Women's Medical University and the National Center for Global Health and Medicine. Virulence in mice was determined as described [64]. Briefly, LD₅₀ values were determined by intraperitoneal (*i.p.*) injection of each SDSE strain into 5-6-7 week old female ddY mice.

Genome sequencing

We obtained draft contig data of GGS_124 from a commercial service (454 Life Sciences, Branford, CT), and the gaps between the contigs were tiled by PCR after closure to validate assembly using specific primers or by primer walking and an ABI 3100 genetic analyzer (Applied Biosystems Inc., Foster City, CA). The percentage of QV40+ bases, an index of the quality of data from pyro-sequencing, was 99.87%. The genome sequence of GGS_124 has been deposited in the DDBJ database (accession no. AP010935).

Genome annotation and bioinformatics

Transfer RNAs (tRNAs), transfer-messenger RNA (tmRNA), and rRNA sequences were predicted using ARAGORN [72]. Coding sequences (CDS) were predicted using *in silico* Molecular Cloning (In Silico Biology Inc., Yokohama City, Kanagawa, Japan) for selection of optimal start sites. Predicted genes and intergenic regions were compared using the NCBI sequence database <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, and predicted CDS and start sites were adjusted accordingly. A genome-wide homology search was performed using the discontinuous megaBLAST algorithm <http://blast.ncbi.nlm.nih.gov/Blast.cgi> with a word size of 11 and rewards and penalties (2, -3) that optimize for alignments of about 85% identity. The phylogenetic tree of all sequenced *Streptococcus* species was constructed based on CVTree [http://tlife.fudan.edu.cn/cvtree/\[73\]](http://tlife.fudan.edu.cn/cvtree/[73]). Secretion signal peptides were predicted using SignalP [http://www.cbs.dtu.dk/services/SignalP\[74\]](http://www.cbs.dtu.dk/services/SignalP[74]). Sortases and cell wall sorting signals were predicted with available hidden Markov models using HMMER [http://bamics3.cmbi.kun.nl/jos/sortase_substrates/help.html\[75\]](http://bamics3.cmbi.kun.nl/jos/sortase_substrates/help.html[75]). Clustered, regularly interspaced, short palindromic repeat

(CRISPR) spacers were detected using CRISPR Finder [http://crispr.u-psud.fr/\[76\]](http://crispr.u-psud.fr/[76]). The phylogenetic tree of all sequenced *Streptococcus* was constructed based on the CVTree [http://tlife.fudan.edu.cn/cvtree/\[73\]](http://tlife.fudan.edu.cn/cvtree/[73])

Detection and sequencing of CRISPRs and CRISPR-associated protein (Cas) genes in the 8 SDSE strains

CRISPRs and Cas genes in the SDSE strains were detected by conventional PCR using ExTaq (Takara Bio Inc., Otsu, Shiga, Japan) and the primers listed in Additional file 10. Template DNA was isolated as described [64]. The PCR cycling conditions were 94°C for 2 min followed by 30 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 4 min, and a final extension at 72°C for 6 min. The amplified CRISPR fragments were sequenced using an ABI 3100 genetic analyzer (Applied Biosystems Inc.). The CRISPR spacers were identified using CRISPR Finder [http://crispr.u-psud.fr/\[76\]](http://crispr.u-psud.fr/[76]). The sequences of CRISPR and the spacers have been deposited in the DDBJ database (accession numbers are listed in Table 2).

Determination of SpeB-like protease activity

SpeB-like protease activity was assayed as described [77], with slight modifications. Briefly, DTT was added to bacterial culture supernatants to a final concentration of 10 mM and incubated at 37°C for 30 min. An equal amount of 2 g/L azocasein was added, and incubation was continued for an additional 30 min at 37°C. Trichloroacetic acid was added to a final concentration of 5%, and incubation was continued for 15 min at 4°C. After centrifugation, an equal amount of 5 M NaOH was added to the supernatant, and its absorbance at 450 nm was measured.

Distribution of putative virulence factors unique to GGS_124 among the other SDSE

The distribution of putative virulence factors of GGS_124 among SDSE strains was analyzed by conventional PCR using ExTaq (Takara Bio Inc.) and the primers listed in Additional file 10. Template DNA extraction and PCR were performed as described [64].

Additional material

Additional file 1: Overview and comparison of the genome sequences of GGS_124 and genus *Streptococcus* available in databases as of January 2010.

Additional file 2: Unrooted phylogenetic tree of genus *Streptococcus*, including *S. dysgalactiae* subsp. *equisimilis* GGS_124. The phylogenetic tree of all sequenced *Streptococcus* was constructed based on CVTree [http://tlife.fudan.edu.cn/cvtree/\[73\]](http://tlife.fudan.edu.cn/cvtree/[73]), which constructs whole genome based phylogenetic trees without sequence alignment by using a Composition Vector (CV) approach. The genetic distances between the major nodes are shown.

Additional file 3: Genome rearrangement map of SDSE strain GGS_124 relative to GAS strains MGAS315 and strain SSI-1, and *S. uberis* 0140J. The genes were aligned from the predicted replication

origin of each genome. The colored bars separating each genome represent similarity matches identified by *in silico* molecular cloning. BLASTP comparisons of CDS with GAS MGAS315 and SSI-1 and *S. uberis* 0140J are shown as amino acid identities of $\geq 90\%$ (red), 89%-80% (orange), 79%-70% (yellow), 69%-60% (green), 59%-50% (light blue), and 49%-40% (dark blue). Prophages are highlighted as green boxes.

Additional file 4: GGS_124 genes showing higher similarity to genes from bacteria other than GAS, or no similarity to genes in the databases

Additional file 5: Features of CRISPR found in the GGS_124 genome and phages derived from GAS containing sequences homologous to GGS_124 spacers

Additional file 6: Putative virulence factors found in GGS_124 and their homologous genes in other streptococcal species

Additional file 7: Determination of SpeB-like protease activity in the SDSE isolates. SDSE (GGS_124 and GGS_118) and GAS (NIH9) were cultured in BHI in the presence or absence of E-64, and the culture supernatants were analyzed for protease activity using azocasein as a substrate. The background activity of BHI is also shown.

Additional file 8: Distribution of the virulence factors found in SDSE among other streptococci

Additional file 9: Comparison of emm-type of GCS and GGS isolates from humans reported from 4 countries. Shown are emm-specific differences of SDSE in invasive and noninvasive infections from 1998 to 2004 in Portugal [65], the emm types of 128 strains of SDSE collected from 11 medical institutions in Japan from September 2003 to October 2005 [7], the emm types of 64 GCS and GGS isolates associated with noninvasive disease in western Norway from February 2005 to March 2006 [66], and the emm types of 212 invasive SDSE isolates collected in Atlanta, Georgia, from July 2002 to June 2004 and in the San Francisco Bay Area of California from January 2003 to December 2004 in the USA [3]. Each stack was ordered from higher (top) to lower isolation frequency (bottom).

Additional file 10: List of oligonucleotide primers used in this study.

Acknowledgements

The authors thank Mr. Hanasaki and Mr. Shimada for assisting in gap filling of the GGS_124 genome and animal experiments. This work was partly supported by a grant for Research on Emerging and Reemerging Infectious Diseases (H19 Shinkouh-02). T. M. A. was supported by International Cooperation Research Grant (21A129) from the Ministry of Health, Labor, and Welfare of Japan.

Author details

¹Department of Infectious Diseases, National Center for Global Health and Medicine, 1-21-1, Toyama, Shinjuku-ku, Tokyo 162-8655, Japan. ²Graduate School of Infection Control Science, Kitasato University, 5-9-1, Shirokane, Minato-ku, Tokyo 108-8641, Japan. ³Department of Microbiology and Immunology, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan. ⁴Department of Animal and Food Hygiene, Obihiro University of Agriculture and Veterinary Medicine, 2-11 Inada-cho, Obihiro, Hokkaido 080-8555, Japan.

Authors' contributions

YS and TMA performed the molecular genetic studies, participated in sequence alignment and drafted the manuscript. KO performed the molecular genetic studies and participated in sequence alignment. SYM and KU analyzed the genome sequences. JY, YS and TMA performed animal experiments. TK drafted the manuscript. UK, JY and TK conceived of the study, participated in its design and coordination, and helped to draft the manuscript. All authors read and approved the final manuscript.

Received: 7 June 2010 Accepted: 11 January 2011
Published: 11 January 2011

References

1. Lindbaek M, Hoiby EA, Lermark G, Steinsholt IM, Hjortdahl P: **Clinical symptoms and signs in sore throat patients with large colony variant beta-haemolytic streptococci groups C or G versus group A.** *Br J Gen Pract* 2005, **55**:615-619.
2. Bramhachari PV, Kaul SY, McMillan DJ, Shaila MS, Karmarkar MG, Sriprakash KS: **Disease burden due to Streptococcus dysgalactiae subsp. equisimilis (group G and C streptococci) is higher than that due to Streptococcus pyogenes among Mumbai school children.** *J Med Microbiol* 2010, **59**:220-223.
3. Broyles LN, Van Beneden C, Beall B, Facklam R, Shewmaker PL, Malpiedi P, Daily P, Reingold A, Farley MM: **Population-based study of invasive disease due to beta-hemolytic streptococci of groups other than A and B.** *Clin Infect Dis* 2009, **48**:706-712.
4. Ekelund K, Skinhøj P, Madsen J, Konradsen HB: **Invasive group A, B, C and G streptococcal infections in Denmark 1999-2002: epidemiological and clinical aspects.** *Clin Microbiol Infect* 2005, **11**:569-576.
5. Lopardo HA, Vidal P, Sparo M, Jeric P, Centron D, Facklam RR, Paganini H, Pagniez NG, Lovgren M, Beall B: **Six-month multicenter study on invasive infections due to Streptococcus pyogenes and Streptococcus dysgalactiae subsp. equisimilis in Argentina.** *J Clin Microbiol* 2005, **43**:802-807.
6. Sunaoshi K, Murayama SY, Adachi K, Yagoshi M, Okuzumi K, Chiba N, Morozumi M, Ubukata K: **Molecular emm genotyping and antibiotic susceptibility of Streptococcus dysgalactiae subsp. equisimilis isolated from invasive and non-invasive infections.** *J Med Microbiol* 2010, **59**:82-88.
7. Sunaoshi KS, Aburahashi H, Kobayashi R, Yamamoto Y, Okuzumi K, Yoshida A, Misawa Y, Adachi K, Ubukata K: **Emm typing by genetic identification of Streptococcus dysgalactiae subsp. equisimilis and susceptibility to oral antibiotics.** *Kansenshogaku Zasshi* 2006, **80**:488-495.
8. Takahashi T, Ubukata K, Watanabe H: **Invasive infection caused by Streptococcus dysgalactiae subsp. equisimilis: characteristics of strains and clinical features.** *J Infect Chemother* 2010.
9. Reissmann S, Friedrichs C, Rajkumari R, Itzek A, Fulde M, Rodloff AC, Brahmadrath KN, Chhatwal GS, Nitsche-Schmitz DP: **Contribution of Streptococcus anginosus to infections caused by groups C and G streptococci, southern India.** *Emerg Infect Dis* 2010, **16**:656-663.
10. Dawson ED, Taylor AW, Smagala JA, Rowlen KL: **Molecular detection of Streptococcus pyogenes and Streptococcus dysgalactiae subsp. equisimilis.** *Mol Biotechnol* 2009, **42**:117-127.
11. Hilmarsdottir I, Valsdottir F: **Molecular typing of Beta-hemolytic streptococci from two patients with lower-limb cellulitis: identical isolates from toe web and blood specimens.** *J Clin Microbiol* 2007, **45**:3131-3132.
12. Horii T, Izumida S, Takeuchi K, Tada T, Ishikawa J, Tsuboi K: **Acute peritonitis and salpingitis associated with streptococcal toxic shock syndrome caused by Lancefield group G alpha-haemolytic Streptococcus dysgalactiae subsp. equisimilis.** *J Med Microbiol* 2006, **55**:953-956.
13. Kumar A, Sandoe J, Kumar N: **Three cases of vertebral osteomyelitis caused by Streptococcus dysgalactiae subsp. equisimilis.** *J Med Microbiol* 2005, **54**:1103-1105.
14. Misawa Y, Okugawa S, Ubukata K, Okuzumi K, Okada M, Moriya K, Koike K: **A case of severe necrotizing cellulitis caused by group G Streptococcus dysgalactiae subsp. equisimilis.** *Kansenshogaku Zasshi* 2006, **80**:436-439.
15. Moser C, Andresen K, Kjerulf A, Salamon S, Kemp M, Christensen JJ: **Infective Arthritis: Bacterial 23S rRNA Gene Sequencing as a Supplementary Diagnostic Method.** *Open Microbiol J* 2008, **2**:85-88.
16. Savini V, Catavittello C, Talia M, Manna A, Pompetti F, Di Bonaventura G, Di Giuseppe N, Febbo F, Balbinot A, Di Zaccaro S, Esattore F, D'Antonio D: **Beta-lactam failure in treatment of two group G Streptococcus dysgalactiae subsp. equisimilis Pharyngitis patients.** *J Clin Microbiol* 2008, **46**:814-816.
17. Mott ML, Berger JM: **DNA replication initiation: mechanisms and regulation in bacteria.** *Nat Rev Microbiol* 2007, **5**:343-354.
18. Nakagawa I, Kurokawa K, Yamashita A, Nakata M, Tomiyasu Y, Okahashi N, Kawabata S, Yamazaki K, Shiba T, Yasunaga T, Hayashi H, Hattori M, Hamada S: **Genome sequence of an M3 strain of Streptococcus pyogenes reveals a large-scale genomic rearrangement in invasive**

- strains and new insights into phage evolution. *Genome Res* 2003, **13**:1042-1055.
19. Holden MT, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, Dowd L, Feltwell T, Hamlin N, Holroyd S, Jagels K, Moule S, Mungall K, Quail MA, Price C, Rabinowitsch E, Sharp S, Skelton J, Whitehead S, Barrell BG, Kehoe M, Parkhill J: **Complete genome of acute rheumatic fever-associated serotype M5 Streptococcus pyogenes strain manfredo.** *J Bacteriol* 2007, **189**:1473-1477.
 20. McShan WM, Ferretti JJ, Karasawa T, Suvorov AN, Lin S, Qin B, Jia H, Kenton S, Najar F, Wu H, Scott J, Roe BA, Savic DJ: **Genome sequence of a nephritogenic and highly transformable M49 strain of Streptococcus pyogenes.** *J Bacteriol* 2008, **190**:7773-7785.
 21. Banks DJ, Porcella SF, Barbian KD, Beres SB, Phillips LE, Voyich JM, DeLeo FR, Martin JM, Somerville GA, Musser JM: **Progress toward characterization of the group A Streptococcus metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain.** *J Infect Dis* 2004, **190**:727-738.
 22. Beres SB, Sesso R, Pinto SW, Hoe NP, Porcella SF, Deleo FR, Musser JM: **Genome sequence of a Lancefield group C Streptococcus zooepidemicus strain causing epidemic nephritis: new information about an old disease.** *PLoS One* 2008, **3**:e3026.
 23. Sumbly P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM: **Evolutionary origin and emergence of a highly successful clone of serotype M1 group A Streptococcus involved multiple horizontal gene transfer events.** *J Infect Dis* 2005, **192**:771-782.
 24. Green NM, Zhang S, Porcella SF, Nagiec MJ, Barbian KD, Beres SB, LeFebvre RB, Musser JM: **Genome sequence of a serotype M28 strain of group A streptococcus: potential new insights into puerperal sepsis and bacterial disease specificity.** *J Infect Dis* 2005, **192**:760-770.
 25. Smoot JC, Barbian KD, Van Gompel JJ, Smoot LM, Chaussee MS, Sylva GL, Sturdevant DE, Ricklefs SM, Porcella SF, Parkins LD, Beres SB, Campbell DS, Smith TM, Zhang Q, Kapur V, Daly JA, Veasy LG, Musser JM: **Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.** *Proc Natl Acad Sci USA* 2002, **99**:4668-4673.
 26. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J: **Small CRISPR RNAs guide antiviral defense in prokaryotes.** *Science* 2008, **321**:960-964.
 27. Marraffini LA, Sontheimer EJ: **CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.** *Science* 2008, **322**:1843-1845.
 28. Holden MT, Heather Z, Paillot R, Steward KF, Webb K, Ainslie F, Jourdan T, Bason NC, Holroyd NE, Mungall K, Quail MA, Sanders M, Simmonds M, Willey D, Brooks K, Aanensen DM, Spratt BG, Jolley KA, Maiden MC, Kehoe M, Chanter N, Bentley SD, Robinson C, Maskell DJ, Parkhill J, Waller AS: **Genomic evidence for the evolution of Streptococcus equi: host restriction, increased virulence, and genetic exchange with human pathogens.** *PLoS Pathog* 2009, **5**:e1000346.
 29. van der Ploeg JR: **Analysis of CRISPR in Streptococcus mutans suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages.** *Microbiology* 2009, **155**:1966-1976.
 30. Horvath P, Barrangou R: **CRISPR/Cas, the immune system of bacteria and archaea.** *Science* 2010, **327**:167-170.
 31. Davies MR, McMillan DJ, Beiko RG, Barroso V, Geffers R, Sriprakash KS, Chhatwal GS: **Virulence profiling of Streptococcus dysgalactiae subspecies equisimilis isolated from infected humans reveals 2 distinct genetic lineages that do not segregate with their phenotypes or propensity to cause diseases.** *Clin Infect Dis* 2007, **44**:1442-1454.
 32. McMillan DJ, Vu T, Bramhachari PV, Kaul SY, Bouvet A, Shaila MS, Karmarkar MG, Sriprakash KS: **Molecular markers for discriminating Streptococcus pyogenes and S. dysgalactiae subspecies equisimilis.** *Eur J Clin Microbiol Infect Dis* 2010, **29**:585-589.
 33. Betschel SD, Borgia SM, Barg NL, Low DE, De Azavedo JC: **Reduced virulence of group A streptococcal Tn916 mutants that do not produce streptolysin S.** *Infect Immun* 1998, **66**:1671-1679.
 34. Flanagan J, Collin N, Timoney J, Mitchell T, Mumford JA, Chanter N: **Characterization of the haemolytic activity of Streptococcus equi.** *Microb Pathog* 1998, **24**:211-221.
 35. Nakagawa I, Amano A, Mizushima N, Yamamoto A, Yamaguchi H, Kamimoto T, Nara A, Funao J, Nakata M, Tsuda K, Hamada S, Yoshimori T: **Autophagy defends cells against invading group A Streptococcus.** *Science* 2004, **306**:1037-1040.
 36. Zhao J, Hayashi T, Saarinen S, Papageorgiou AC, Kato H, Imanishi K, Kirikae T, Abe R, Uchiyama T, Miyoshi-Akiyama T: **Cloning, expression, and characterization of the superantigen streptococcal pyrogenic exotoxin G from Streptococcus dysgalactiae.** *Infect Immun* 2007, **75**:1721-1729.
 37. Kamezawa Y, Nakahara T, Nakano S, Abe Y, Nozaki-Renard J, Isono T: **Streptococcal mitogenic exotoxin Z, a novel acidic superantigenic toxin produced by a T1 strain of Streptococcus pyogenes.** *Infect Immun* 1997, **65**:3828-3833.
 38. Edwards RJ, Taylor GW, Ferguson M, Murray S, Rendell N, Wrigley A, Bai Z, Boyle J, Finney SJ, Jones A, Russell HH, Turner C, Cohen J, Faulkner L, Sriskandan S: **Specific C-terminal cleavage and inactivation of interleukin-8 by invasive disease isolates of Streptococcus pyogenes.** *J Infect Dis* 2005, **192**:783-790.
 39. Nishifuji K, Sugai M, Amagai M: **Staphylococcal exfoliative toxins: "molecular scissors" of bacteria that attack the cutaneous defense barrier in mammals.** *J Dermatol Sci* 2008, **49**:21-31.
 40. Lahteenmaki K, Kuusela P, Korhonen TK: **Bacterial plasminogen activators and receptors.** *FEMS Microbiol Rev* 2001, **25**:531-552.
 41. Chaussee MS, Gerlach D, Yu CE, Ferretti JJ: **Inactivation of the streptococcal erythrogenic toxin B gene (speB) in Streptococcus pyogenes.** *Infect Immun* 1993, **61**:3719-3723.
 42. Falugi F, Zingaretti C, Pinto V, Mariani M, Amodeo L, Manetti AG, Capo S, Musser JM, Orefici G, Margarit I, Telford JL, Grandi G, Mora M: **Sequence variation in group A Streptococcus pili and association of pilus backbone types with lancefield T serotypes.** *J Infect Dis* 2008, **198**:1834-1841.
 43. Mora M, Bensi G, Capo S, Falugi F, Zingaretti C, Manetti AG, Maggi T, Taddei AR, Grandi G, Telford JL: **Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens.** *Proc Natl Acad Sci USA* 2005, **102**:15641-15646.
 44. Lauer P, Rinaudo CD, Soriani M, Margarit I, Maione D, Rosini R, Taddei AR, Mora M, Rappuoli R, Grandi G, Telford JL: **Genome analysis reveals pili in Group B Streptococcus.** *Science* 2005, **309**:105.
 45. Bjorck L, Kronvall G: **Purification and some properties of streptococcal protein G, a novel IgG-binding reagent.** *J Immunol* 1984, **133**:969-974.
 46. Terao Y, Yamaguchi M, Hamada S, Kawabata S: **Multifunctional glyceraldehyde-3-phosphate dehydrogenase of Streptococcus pyogenes is essential for evasion from neutrophils.** *J Biol Chem* 2006, **281**:14215-14223.
 47. Stollerman GH, Dale JB: **The importance of the group A streptococcus capsule in the pathogenesis of human infections: a historical perspective.** *Clin Infect Dis* 2008, **46**:1038-1045.
 48. Kumari K, Weigel PH: **Molecular cloning, expression, and characterization of the authentic hyaluronan synthase from group C Streptococcus equisimilis.** *J Biol Chem* 1997, **272**:32539-32546.
 49. Starr CR, Engleberg NC: **Role of hyaluronidase in subcutaneous spread and growth of group A streptococcus.** *Infect Immun* 2006, **74**:40-48.
 50. Buchanan JT, Simpson AJ, Aziz RK, Liu GY, Kristian SA, Kotb M, Feramisco J, Nizet V: **DNase expression allows the pathogen group A Streptococcus to escape killing in neutrophil extracellular traps.** *Curr Biol* 2006, **16**:396-400.
 51. Beiter K, Wartha F, Albiger B, Normark S, Zychlinsky A, Henriques-Normark B: **An endonuclease allows Streptococcus pneumoniae to escape from neutrophil extracellular traps.** *Curr Biol* 2006, **16**:401-407.
 52. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A: **Neutrophil extracellular traps kill bacteria.** *Science* 2004, **303**:1532-1535.
 53. Hasegawa T, Minami M, Okamoto A, Tatsuno I, Isaka M, Ohta M: **Characterization of a virulence-associated and cell-wall-located DNase of Streptococcus pyogenes.** *Microbiology* 2010, **156**:184-190.
 54. Nitsche DP, Johansson HM, Frick IM, Morgelin M: **Streptococcal protein FOG, a novel matrix adhesion interacting with collagen I in vivo.** *J Biol Chem* 2006, **281**:1670-1679.
 55. Barroso V, Rohde M, Davies MR, Gillen CM, Nitsche-Schmitz DP, Dinkla K, Chhatwal GS: **Identification of active variants of PARF in human pathogenic group C and group G streptococci leads to an amended description of its consensus motif.** *Int J Med Microbiol* 2009, **299**:547-553.
 56. Binks M, Sriprakash KS: **Characterization of a complement-binding protein, DRS, from strains of Streptococcus pyogenes containing the emm12 and emm55 genes.** *Infect Immun* 2004, **72**:3981-3986.

57. Sagar V, Kumar R, Ganguly NK, Menon T, Chakraborti A: **DRS is far less divergent than streptococcal inhibitor of complement of group A streptococcus.** *J Bacteriol* 2007, **189**:2933-2935.
58. Fernie-King BA, Seilly DJ, Davies A, Lachmann PJ: **Streptococcal inhibitor of complement inhibits two Additional components of the mucosal innate immune system: secretory leukocyte proteinase inhibitor and lysozyme.** *Infect Immun* 2002, **70**:4908-4916.
59. Hartas J, Sriprakash KS: **Streptococcus pyogenes strains containing emm12 and emm55 possess a novel gene coding for distantly related SIC protein.** *Microb Pathog* 1999, **26**:25-33.
60. Caswell CC, Barczyk M, Keene DR, Lukomska E, Gullberg DE, Lukomski S: **Identification of the first prokaryotic collagen sequence motif that mediates binding to human collagen receptors, integrins alpha2beta1 and alpha11beta1.** *J Biol Chem* 2008, **283**:36168-36175.
61. Kimoto H, Fujii Y, Yokota Y, Taketo A: **Molecular characterization of NADase-streptolysin O operon of hemolytic streptococci.** *Biochim Biophys Acta* 2005, **1681**:134-149.
62. Madden JC, Ruiz N, Caparon M: **Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria.** *Cell* 2001, **104**:143-152.
63. Cunningham MW: **Pathogenesis of group A streptococcal infections.** *Clin Microbiol Rev* 2000, **13**:470-511.
64. Miyoshi-Akiyama T, Zhao J, Kikuchi K, Kato H, Suzuki R, Endoh M, Uchiyama T: **Quantitative and qualitative comparison of virulence traits, including murine lethality, among different M types of group A streptococci.** *J Infect Dis* 2003, **187**:1876-1887.
65. Pinho MD, Melo-Cristino J, Ramirez M: **Clonal relationships between invasive and noninvasive Lancefield group C and G streptococci and emm-specific differences in invasiveness.** *J Clin Microbiol* 2006, **44**:841-846.
66. Kittang BR, Langeland N, Mylvaganam H: **Distribution of emm types and subtypes among noninvasive group A, C and G streptococcal isolates in western Norway.** *APMIS* 2008, **116**:457-464.
67. Bentley RW, Leigh JA, Collins MD: **Intragenomic structure of Streptococcus based on comparative analysis of small-subunit rRNA sequences.** *Int J Syst Bacteriol* 1991, **41**:487-494.
68. Chatellier S, Harel J, Zhang Y, Gottschalk M, Higgins R, Devriese LA, Brousseau R: **Phylogenetic diversity of Streptococcus suis strains of various serotypes as revealed by 16S rRNA gene sequence comparison.** *Int J Syst Bacteriol* 1998, **48 Pt 2**: 581-589.
69. Sachse S, Seidel P, Gerlach D, Gunther E, Rodel J, Straube E, Schmidt KH: **Superantigen-like gene(s) in human pathogenic Streptococcus dysgalactiae, subsp equisimilis: genomic localisation of the gene encoding streptococcal pyrogenic exotoxin G (speG(dys)).** *FEMS Immunol Med Microbiol* 2002, **34**:159-167.
70. Ikebe T, Wada A, Inagaki Y, Sugama K, Suzuki R, Tanaka D, Tamaru A, Fujinaga Y, Abe Y, Shimizu Y, Watanabe H: **Dissemination of the phage-associated novel superantigen gene speL in recent invasive and noninvasive Streptococcus pyogenes M3/T3 isolates in Japan.** *Infect Immun* 2002, **70**:3227-3233.
71. Miyoshi-Akiyama T, Ikebe T, Watanabe H, Uchiyama T, Kirikae T, Kawamura Y: **Use of DNA arrays to identify a mutation in the negative regulator, csrR, responsible for the high virulence of a naturally occurring type M3 group A streptococcus clinical isolate.** *J Infect Dis* 2006, **193**:1677-1684.
72. Laslett D, Canback B: **ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences.** *Nucleic Acids Res* 2004, **32**:11-16.
73. Xu Z, Hao B: **CVTree update: a newly designed phylogenetic study platform using composition vectors and whole genomes.** *Nucleic Acids Res* 2009, **37**:W174-8.
74. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
75. Boekhorst J, de Been MW, Kleerebezem M, Siezen RJ: **Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs.** *J Bacteriol* 2005, **187**:4928-4934.
76. Grissa I, Vergnaud G, Pourcel C: **CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats.** *Nucleic Acids Res* 2007, **35**:W52-7.
77. Kuo CF, Wu JJ, Lin KY, Tsai PJ, Lee SC, Jin YT, Lei HY, Lin YS: **Role of streptococcal pyrogenic exotoxin B in the mouse model of group A streptococcal infection.** *Infect Immun* 1998, **66**:3931-3935.
78. Gryllos I, Grifantini R, Colaprico A, Jiang S, Deforce E, Hakansson A, Telford JL, Grandi G, Wessels MR: **Mg(2+) signalling defines the group A streptococcal CsrRS (CovRS) regulon.** *Mol Microbiol* 2007, **65**:671-683.
79. Graham MR, Virtaneva K, Porcella SF, Gardner DJ, Long RD, Welty DM, Barry WT, Johnson CA, Parkins LD, Wright FA, Musser JM: **Analysis of the transcriptome of group A Streptococcus in mouse soft tissue infection.** *Am J Pathol* 2006, **169**:927-942.
80. Hytonen J, Haataja S, Finne J: **Streptococcus pyogenes glycoprotein-binding streptadhesin activity is mediated by a surface-associated carbohydrate-degrading enzyme, pullulanase.** *Infect Immun* 2003, **71**:784-793.

doi:10.1186/1471-2164-12-17

Cite this article as: Shimomura et al.: Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS). *BMC Genomics* 2011 **12**:17.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

