
Article:
Reglinski, M., Sriskandan, S. and Turner, C. orcid.org/0000-0002-4458-9748 (2019) Identification of two new core chromosome-encoded superantigens in Streptococcus pyogenes; speQ and speR. Journal of Infection, 78 (5). pp. 358-363. ISSN 0163-4453

https://doi.org/10.1016/j.jinf.2019.02.005

Journal of Infection 78 (2019) 358–363

Contents lists available at ScienceDirect

Journal of Infection

journal homepage: www.elsevier.com/locate/jinf

Identification of two new core chromosome-encoded superantigens in *Streptococcus pyogenes; speQ* and *speR*

Mark Reglinski[a], Shiranee Sriskandan[a], Claire E Turner[a,b,*]

[a] Department of Infectious Diseases, Imperial College London, London W12 0NN, UK
[b] Department of Molecular Biology & Biotechnology, The Florey Institute, University of Sheffield, Sheffield S10 2TN, UK

ARTICLE INFO

Article history:
Accepted 17 February 2019
Available online 20 February 2019

Keywords:
Streptococcus pyogenes
Group A *Streptococcus*
Superantigens
Scarlet fever
Toxic shock syndrome
Mitogen

SUMMARY

Superantigens are ubiquitous within the *Streptococcus pyogenes* genome, which suggests that superantigen-mediated T-cell activation provides a significant selective advantage. *S. pyogenes* can carry a variable complement of the 11 known superantigens. We have identified two novel *S. pyogenes* superantigens, denoted *speQ* and *speR*, adjacent to each other in the core-chromosome of isolates belonging to eleven different *emm*-types. Although distinct from other superantigens, *speQ* and *speR* were most closely related to *speK* and *speJ*, respectively. Recombinant SPEQ and SPER were mitogenic towards human peripheral blood mononuclear cells at ng/ml concentrations, and SPER was found to be more mitogenic than SPEQ.

Background

The human pathogen, *Streptococcus pyogenes,* produces numerous virulence factors, including the extracellular superantigen toxins which are associated with the development of streptococcal toxic shock syndrome and scarlet fever. Superantigens are able to cross-link the major histocompatibility complex (MHC) class II with the T cell receptor, bypassing the normal antigen presenting process and resulting in a high level of T cell activation, cytokine release and inflammation.[1] Although superantigens are thought to promote invasive disease, associated with high mortality rates, they may have a localized role in productive infection of the nasopharynx.[2]

Genetically and structurally related superantigens are also produced by *Staphylococcus aureus* and some other streptococcal species including the group C/G streptococci. There are 14 known streptococcal superantigens, 11 of which have been found in *S. pyogenes; speA, speC, speG-speM, smeZ* and *ssa.* All, except *speJ* and *ssa*, have been found in other streptococcal species as well. Three superantigen genes, *szeN, szeP* and *szeF* have been found only in *S. equi subsp zooepidemicus*.[3] Commons et al. later suggested renaming these to *speN, speP* and *speO*, respectively to standardize the nomenclature across all streptococci.[1]

The superantigen genes *speG, speJ* and *smeZ* are encoded on the core chromosome but are not ubiquitous among *S. pyogenes* isolates. The other eight identified *S. pyogenes* superantigens are associated with prophages which have the potential to be mobile, introducing variability among isolates. As there is variability in the complement of superantigens carried by *S. pyogenes* isolates, along with mobility and sharing across other streptococcal species, there may be streptococcal superantigens that are yet to be identified.

Whilst testing for the presence of the 11 known superantigens in whole genome sequence (WGS) data from two *S. pyogenes* isolates from the pre-antibiotic era, we identified two new adjacent potential superantigen genes. We subsequently confirmed the presence of both genes in WGS from modern *S. pyogenes* isolates of different *emm* genotypes. The two new potential superantigen genes were not associated with recognised mobile genetic elements but were limited to certain *emm*-types. We have termed the genes *speQ* and *speR*, to follow the proposed nomenclature, and confirmed that they are indeed mitogenic towards human mononuclear cells.

Methods

Bacterial strains and growth conditions

Two *S. pyogenes emm*60 isolates (H865 and H870) were first isolated in 1938 from puerperal sepsis patients at Queen Charlottes Hospital London. *S. pyogenes* isolates were cultured on Columbia blood agar plates or statically in Todd Hewitt broth at 37 °C with

* Corresponding author: Department of Molecular Biology & Biotechnology, The Florey Institute, University of Sheffield, Sheffield S10 2TN, UK.
E-mail address: c.e.turner@sheffield.ac.uk (C.E. Turner).

https://doi.org/10.1016/j.jinf.2019.02.005

5% CO_2. *E. coli* were cultured in LB at 37 °C with 225 rpm agitation and supplemented with 100ug/ml ampicillin where appropriate.

Gene identification

The *speQ* and *speR* genes were first identified from whole genome sequence (WGS) data of *emm*60 isolates H865 and H870 (short read archive ERR485817 and ERR485821, respectively). The presence and sequence of *speQR* were confirmed by PCR with primers spanning from the upstream *ideS* gene and the downstream hypothetical gene (*speQR*-region primers listed in Supplementary Table 1), and Sanger sequencing.

Whole genome sequence analysis

Publicly available genome sequence fastq data for other *S. pyogenes* strains representing 86 different *emm*-types[4–6] were obtained and assembled *de novo* for identifying the presence or absence of *speQ* and *speR*.

Fastq reads were assembled de novo using Velvet[7] or as previously described.[6] Assembly statistics are available at Mendeley Data (doi:10.17632/b89yzfcxp8.1 and doi:10.17632/f2d39nsfwk.1) or in the original study.[6] The genomic region spanning the potential *speQR* locus was extracted from de novo assemblies and examined for the presence or absence of complete superantigen genes, but in some cases the quality of assembled sequence data was too low for adequate confirmation of the complete *speQR* locus or the allelic sequence.

Multi-locus sequence (MLST) data were obtained from de-novo assemblies of some UK data.[8] The presence of the 11 known streptococcal superantigens was determined by mapping of the short read sequence data to a pseudosequence of concatenated genes and confirmed through BLAST analysis of de-novo assemblies. Other MLST, *emm*-genotype and superantigen data were obtained from the original studies. We excluded isolates where the *emm* type could not be definitively assigned. This included WGS from Kapatai et al.[5] where *emm*-type determined by WGS was reported to be different to the original Sanger sequenced *emm*-type, and *emm*-negative isolates from Chochua et al.[4]

For phylogenetic analysis, fastq data was mapped to the completed *S. pyogenes emm*89 reference genome H293 (HG316453.2)[9] and single nucleotide polymorphisms extracted from the core genome using SNP-sites[10] to generate a maximum likelihood phylogeny with RAxML.[11]

Different *speQR* DNA sequence alleles were identified and submitted to Genbank (accession numbers: BK010649- BK010666, BK010692, BK010693).

Recombinant protein expression

BamHI-ended coding sequences for *speQ.1* and *speR.1* were amplified from H865 gDNA using the primers listed in Supplementary Table 1 and cloned into the overexpression vector pET-19b (Novagen). Recombinant proteins were expressed in One Shot BL21(DE3) Chemically Competent *E. coli* (Life Technologies) and purified to apparent homogeneity using the Ni-NTA purification system (Novagen). Protein concentrations were measured using the Pierce Coomassie Plus (Bradford) Assay Kit. 1 µg aliquots of each protein were separated by SDS-PAGE and visualized by Instant-Blue staining (Expedeon). Recombinant IdeS was produced as previously described[12] and purified alongside recombinant SPEQ and SPER. Recombinant SPEC was purchased from Toxin Technology, Inc (Sarasota, Florida).

SPEQ and SPER antisera

Antisera towards SPEQ and SPER were raised by immunizing mice intramuscularly with 10 µg of recombinant protein, emulsified 1:1 with Freund's complete adjuvant, and booster immunizations at 21 and 35 days in Freund's incomplete adjuvant. Blood was collected on day 42 and the resulting antiserum was pooled.

Human mononuclear cell proliferation assay

Healthy donor human mononuclear cells (MNCs) were purified as previously described[13] and diluted to 1×10^6 cells/ml in RPMI media (Life Technologies) (+10% FCS). Cells were seeded into 96 well plates at a concentration of 2×10^5/well and incubated with decreasing concentrations of recombinant protein for 48 h. Cell proliferation was measured with a Colorimetric Cell Proliferation BrdU ELISA (Roche).

Immunoblotting

S. pyogenes culture supernatants collected at different time points were concentrated 16-fold by TCA precipitation. Separated proteins were transferred to PVDF membrane (Hybond-LFP, GE Healthcare) which were blocked with 5% non-fat milk (Sigma Aldrich) in PBS and probed with a 1:1000 dilution of mouse antiserum raised against SPEQ or SPER. Bound antibodies were detected using a 1:80,000 dilution of HRP-conjugated goat anti-mouse IgG (Abcam).

RNA extraction and PCR

RNA was extracted from early, mid and late logarithmic growth phases of the two *emm*60 strains (H865 and H870) using a hot acidic phenol method as previously described.[14] RNA samples were treated with Turbo DNA-*free* (Ambion) DNase and 1 µg was converted into cDNA using Transcriptor reverse transcriptase (Roche) and random hexaoligos (RT+ samples). To control for contaminating genomic DNA equivalent reverse transcriptase negative (RT−) samples were generated with another 1 µg but the Transcriptor reverse transcriptase was excluded from the reaction. PCR for *speQR* co-transcription was performed on 200 ng of the RT+ cDNA and RT− samples using *speQR* primers (Supplementary Table 1) and visualized on an agarose gel.

Ethics

Normal human donor cells were acquired from an approved sub-collection of the Imperial College Tissue Bank. All murine procedures were approved by the local ethical review process at Imperial College London and conducted in accordance with the relevant, UK Home Office approved, project license.

Results

Identification of two new potential superantigen genes; speQ *and* speR

We sequenced the genomes of two viable *emm*60 isolates, originally collected in 1938 from two puerperal sepsis patients, and analyzed the genomes for the presence of superantigens. We could not detect the presence of any of the known streptococcal superantigens by short read sequence mapping or BLAST analysis of de novo assembled genomes. The analysis did, however, indicate the presence of sequence in the genomes of both *emm*60 isolates with partial homology to *speK*. We identified this homologous sequence to be within one of two hypothetical genes located immediately

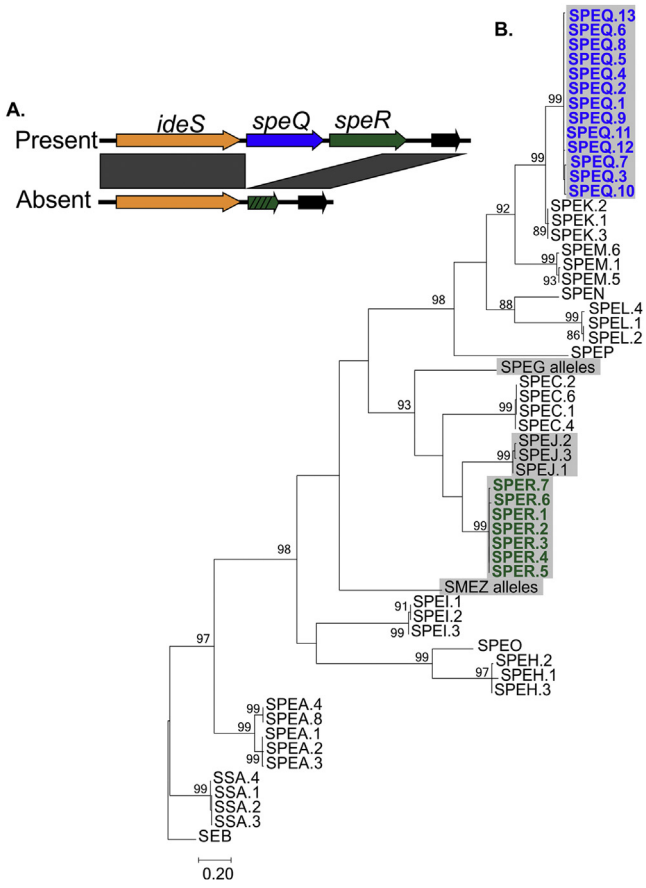

Fig. 1. Schematic representation of the chromosomal location of *speQ* and *speR* and similarity to other superantigens. (A) In *emm*-types where the two superantigens genes *speQ* (blue) and *speR* (green) were present in the chromosome they were located immediately downstream of the gene *ideS* (orange) encoding for the immunoglobulin cleaving protease, and upstream of a gene encoding for a hypothetical protein (black). In *emm*-types where the full length superantigen genes were absent in the chromosome, a C-terminal ~364 base pair fragment of *speR* was present downstream of *ideS*. (B) Available sequences of the superantigen alleles determined by Commons et al.[1] were obtained and translated to amino acids. The signal sequences were identified by SignalP[15] and excluded before alignment with the predicted mature protein sequences of SPEQ (blue) and SPER (green) alleles. Core chromosome (i.e. not associated with known prophage elements) *S. pyogenes* superantigens are shaded grey. The *Staphylococcus aureus* superantigen SEB sequence was also included and used to root the maximum likelihood tree. The 28 SPEG alleles and 39 SMEZ alleles were condensed to a single branch for illustration purposes. Scale represents amino acid substitutions per site. Bootstrap values greater than 80% are shown.

downstream of the gene *ideS* (also known as *mac*) encoding for an immunoglobulin cleaving protease (Fig. 1(A)). BLAST indicated that these genes were closely related to other streptococcal superantigens and carried the typical superantigen C terminal β-grasp domain.[1] We therefore predicted that these would be superantigen genes and denoted them *speQ* and *speR*. PCR and Sanger sequencing confirmed the WGS data. BLASTn and BLASTp of completed available *S. pyogenes* genomes also identified *speQ* and *speR* in an *emm*87 strain NGAS743 (DI45_05770 and DI45_05775, respectively; Genbank CP007560.1).[16] In isolates where full length *speQ* and *speR* genes were absent, a C-terminal fragment of *speR* was present immediately downstream of *ideS* (Fig. 1(A)). We also performed BLASTp analysis of the entire NCBI database, excluding *S. pyogenes*, but did not identify SPEQ or SPER in any other available genomes including other streptococcal species.

Phylogenetic analysis of the amino acid sequences of SPEQ, SPER, and all other available superantigen alleles from all streptococcal species[1] demonstrated that, although phylogenetically distinct, SPEQ is closely related to the prophage-associated SPEK sharing 84% amino acid identity, and SPER is most closely related to the chromosomal SPEJ sharing 64% amino acid identity (Fig. 1(B)). Comparisons were made between SPEQ, SPEK, SPEJ and SPER to identify two superantigen signature amino acid motifs (Supplementary Figure 1).[1,3] SPER, like SPEJ, had the motif Y-G-G-(LIV)-T-x_4-N (Prosite PS00277) but only a partial match for this was identified in SPEQ and SPEK. All four superantigens had the motif K-x_2-(LIVF)-x_4-(LIVF)-D-x_3-R-x_2-L-x_5-(LIV)-Y (Prosite PS00278) and a C-terminal zinc binding domain (HxD).

To determine the presence of *speQR* in other *S. pyogenes* genotypes, publicly available WGS fastq data were obtained from the short read archive for UK isolates[5,6] and USA isolates[4] totaling 4262 genomes tested covering 86 different genotypes (Supplementary Table 2). Complete *speQ* and *speR* were identified in the assembled genome sequence of isolates belonging to the *emm*-types *emm*9, 15, 18, 42, 53, 58, 60, 77, 87, 94 and 169 (Supplementary Table 3). However, not all isolates belonging to some of these genotypes carried the complete *speQR* locus, which was unexpected given the lack of association with mobile genetic elements. Only one out of 41 *emm*18 (USA isolate 20154046) had complete *speQR*, as did 21/24 *emm*58 isolates and 49/72 *emm*77 isolates. The presence or absence of complete *speQR* in these genotypes appeared to be associated with divergent lineages and multi-locus sequence types (MLST) within these *emm*-types (Supplementary Figure 2, Supplementary Table 3), indicative of the same *emm* gene on completely different genetic backgrounds. In contrast, all *emm*94 isolates were MLST-89, but 2/50 did not carry the complete *speQR* allele and formed a separate sub-lineage. As these isolates were still relatively closely related there may have been a horizontal gene transfer event of the *speQR* region.

The majority of isolate genomes that were positive for *speQR* were also positive for at least one other superantigen gene (Supplementary Table 3). The exceptions to this were 4/5 *emm*60, 1/72 *emm*77, and 2/2 *emm*169 isolates where no superantigen genes other than *speQR* were detected.[4]

From the WGS analysis, thirteen DNA alleles for *speQ* and seven DNA alleles for *speR* were identified. The variation was limited to single nucleotide polymorphisms, except a region in *speQ* which varied in the number of a 15 bp/5aa repeat (Supplementary Figure 3). This 15 bp/5aa region repeated twice, four times and five times in three alleles, *speQ.2, speQ.4* and *speQ.5*, respectively; these alleles were found only in genotype *emm*9 isolates. Based on amino acid sequence, SPEQ.1, SPEQ.6, SPEQ.8, SPEQ.11 and SPEQ.12 were identical. SPEQ.3, SPEQ.7, SPEQ.9 and SPEQ.13 each differ from SPEQ.1 by one amino acid residue and SPEQ.10 differs by two amino acid residues. For SPER, SPER.1, SPER.2, SPER.4, and SPER.5 were identical by amino acid sequence, but SPER.3, SPER.6 and SPER.7 each differ by one amino acid.

Recombinant SPEQ and SPER induced proliferation of human mononuclear cells

To determine if SPEQ and SPER were capable of inducing proliferation of human T cells, we recombinantly expressed both proteins in *E. coli* (Fig. 2(A)). These recombinant toxins represented gene alleles *speQ.1* and *speR.1*. Purified toxins were then used to stimulate human mononuclear cells (MNCs) and proliferation was measured by BrdU uptake (Fig. 2(B)). Both SPEQ and SPER induced proliferation, although a 10-fold greater concentration of SPEQ than SPER was required to generate an equivalent response. Proliferation after stimulation with another streptococcal superantigen, SPEC, required 100-1000-fold lower concentration than SPEQ and SPER. As a control, the non-mitogenic IdeS was recombinantly expressed and purified in the same manner as SPEQ and SPER but failed to stimulate any proliferation, as expected.

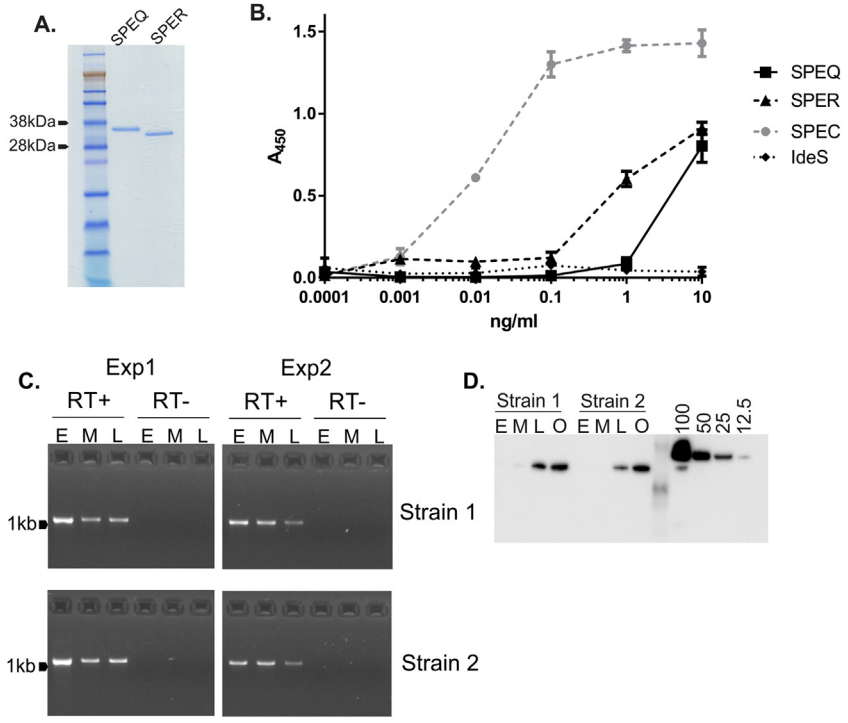

Fig. 2. Activity and expression of SPEQ/R. (A) Instant blue stained SDS-PAGE gel of purified recombinantly-expressed SPEQ and SPER which were then used to stimulate human mononuclear cells (MNCs). (B) Proliferation of MNCs, measured by BrdU assay, required 10 fold more SPEQ (solid line, squares) than SPER (dashed line, triangles). The streptococcal superantigen SPEC (grey dashed line, circles) induced proliferation at a concentration ~100 fold lower than SPER. The protein IdeS (dotted line, diamonds) which was expressed and purified in the same manner as SPEQ and SPER, induced no proliferation of MNCs at any concentration used. Data represent mean and SD from a single MNC donor representative of data from two other donors. (C) RNA was extracted from two strains of *S. pyogenes emm*60 cultured to early (E), mid (M) and late (L) logarithmic phase and converted to cDNA. This was repeated on two separate occasions (Exp1 and Exp2) and gave very similar results. PCR was performed on 200 ng of cDNA using primers that spanned *speQ* and *speR* and, for both strains and in both experiments, a band was detected at ~1 kb following PCR on the RT+ samples, confirming co-transcription of *speQ* and *speR*. No bands were present in samples where the reverse transcriptase had been excluded from the RNA to cDNA reverse transcription reaction (RT−), suggesting no contamination of genomic DNA. (D) SPEQ was detected by Western blot in the culture supernatant of two *emm*60 strains (Strain 1 and Strain 2) at late logarithmic phase (L) and following overnight culture (O), but not at early (E) and only faintly at mid (M) logarithmic phase. A standard curve of 100, 50, 25 and 12.5 ng of rSPEQ was included to provide quantification. Culture supernatants were concentrated 20-fold.

Both speQ *and* speR *are expressed by* S. pyogenes *during culture*

To confirm *speQ* and *speR* expression by *S. pyogenes*, transcription and protein expression were measured. RNA was extracted at early, mid and late-logarithmic phases of growth of two *emm*60 strains and converted to cDNA for PCR. Primers that spanned across both *speQ* and *speR* confirmed the two genes are co-transcribed (Fig. 2(C)). Although only semi-quantitative, transcription appeared greatest at early and mid-logarithmic phases of growth.

Culture supernatants from the same two strains of *emm*60 *S. pyogenes* were probed by Western blot for SPEQ and SPER using an antibody raised in mice against recombinant proteins. SPEQ could be detected at late-logarithmic phase and following overnight culture in both *emm*60 strains (Fig. 2(D)). Using rSPEQ at known concentrations the estimated concentration of SPEQ was ~90–127 ng/ml in late-logarithmic phase culture and increased to ~155–163 ng/ml by overnight culture. We were, however, unable to detect SPER using the rSPER murine antibody in either strain at any growth phase, which was unexpected given the co-transcription.

Discussion

We identified two potential superantigen genes present in the chromosomes of two 1930s *S. pyogenes emm*60 isolates and subsequently identified the same genes in isolates belonging to 10 other *emm*-types in modern international isolates. We termed these genes *speQ* and *speR* and confirmed that they were capable of inducing proliferation of human cells.

We tested the genomes over 4000 different isolates representing 86 *emm*-types and detected *speQR* in strains belonging to *emm*9, 15, 18, 42, 53, 58, 60, 77, 87, 94 and 169, although both *speQR* positive and negative lineages existed within these genotypes (Supplementary Figure 2). Both *emm*77 and *emm*87 have been reported as common causes of invasive disease in various countries.[17]

Like the majority of other superantigens, both *speQ* and *speR* carry two of the three classic superantigen motifs. The third was absent in *speQ*, as also observed in the closest relative *speK*, although present in *speR* and may relate to the different mitogenic potential; 10-fold more SPEQ than SPER was required to generate an equivalent mitogenic response. The mitogenic activity of both SPEQ and SPER was 10–100 fold lower than that of SPEC. This may limit contribution of SPEQ/R to virulence in the presence of much more potent superantigens. The majority of isolates whose genomes tested positive for *speQ/R* also carried at least one other superantigen genes.

Across the entire collection of 1441 USA isolates (previously all typed for the 11 known superantigen genes), the prevalence of *speQR* was 6%, similar to *speL* (5%) and *speM* (6%).[4] The most commonly found superantigen gene within the USA collection was *speG* (93%) followed by *smeZ* (91%), *speC* (51%), *speJ* (41%), *speA* (26%), *speH* (25%), *speI* (23%), *ssa* (10%), *speK* (9%).[4] Of those that were positive for *speQR*, the prevalence of *smeZ* was still high

(94%) and similar for *speK* (11%) and *speC* (40%), but fewer were positive for *speG* (54%) as well as *speA* (4%), *speH* (4%), *speI* (1%), *speJ* (28%), and more were positive for *ssa* (41%), *speL* (16%) and *speM* (16%). This may, however, reflect an association of superantigen complement with *emm*-type. At least one superantigen gene was detected in all 1441 USA isolates, except for four (of 5) *emm*60 isolates and one (of 54) *emm*77 isolate.[4] We identified that these five 'superantigen negative' isolates were positive for *speQR*, consistent with our initial finding that two 1930s *emm*60 were only positive for *speQR* and no other known superantigens.

Interestingly, in all isolates where full length *speQ* and *speR* genes were absent, immediately downstream of *ideS* was a C-terminal fragment of *speR* (Fig. 1(A)). This suggests that *speQR* genes were present in the most recent common ancestor of all *emm*-types but were lost as the *emm*-types diverged. Despite being chromosomally encoded, *speQ* and *speR* were less frequent among contemporary sequenced isolates than the other chromosomally encoded superantigens *speG, speJ* and *smeZ*. Quite why *speQ* and *speR* have persisted in the *emm* types that we identified is unclear; this may reflect a requirement of *S. pyogenes* to express at least one superantigen.

Although superantigens are implicated in the pathogenesis of scarlet fever and streptococcal toxic shock, it is widely recognized that the production of superantigens must play a role in *S. pyogenes* fitness.[18,19] This has been borne out in epidemiological studies that show alterations of superantigen gene content to be implicated in emergence or expansion of new lineages in the population,[20,21] and in animal models where superantigens are shown to be necessary for successful infection.[2,19]

Superantigens crosslink the major histocompatibility molecule (MHC) class II on antigen presenting cells with the T cell receptor. The binding of the T cell receptor occurs through the variable (V) β-region, with a specificity towards different Vβ variants.[18] As each superantigen can activate a different repertoire of Vβ, the more superantigens expressed by *S. pyogenes* the greater the heterogeneity of T cell expansion. Further work is required to determine the Vβ preference of SPEQ and SPER but it may be that even in the presence of more potent superantigens, they still contribute to pathogenesis through the expansion of a different T cell repertoire.

Genetic comparison of *speQ* and *speR* with the other known streptococcal superantigens indicated that they were closely related to the prophage-associated *speK* and the core chromosomal *speJ*, respectively. As SPEQ shares 84% amino acid identity with SPEK, it is possible that *speK* originated from *speQ* that was picked up by a bacteriophage. We did not identify any *speQR* orthologues in the available genomes of other streptococcal species.

Although we identified thirteen different alleles of *speQ,* the most common was *speQ.1,* found in 81% (174/216) of *speQR*-positive isolate genomes where we could confirm the allele. It was also the original allele we identified in the *emm*60 isolates from the 1930s. The alleles *speQ.2, speQ.4* and *speQ.5* that vary by a repeat region were restricted to the *emm*9 genotype. It is unclear as to the significance of this or the impact the repeat region may have on mitogenic activity of these alleles. The original 1930s *emm*60 *speR* allele, *speR.1* was found in 40% (82/216) but the most common was *speR.2*, found in 56% (120/216) of *speQR*-positive isolate genomes; these two alleles should, however, encode identical proteins.

Co-transcription of *speQR* was detected with greater expression in the early stages of exponential growth. Surprisingly, we were unable to detect SPER protein in *S. pyogenes* culture supernatant, although we were able to detect SPEQ at levels similar to SPEC expressed by *emm*3 *S. pyogenes*.[22] Other superantigens have been shown to be sensitive to SPEB degradation,[23] and this may be the case for SPER. It is also possible that the antibody we raised, while able to detect recombinant SPER, was unable to detect native SPER.

We have shown that there are now 16 streptococcal superantigen genes; each one may play a role in promoting *S. pyogenes* virulence. While the benefits of superantigen production remain incompletely understood, the ubiquitous presence of superantigen genes within the *S. pyogenes* genome suggests that they do play a significant role in *S. pyogenes* disease.

Conflicts of interest

None

Acknowledgments

This work was supported by the NIHR Biomedical Research Centre awarded to Imperial College NHS Healthcare Trust, an MRC Confidence in Concept award, the UKCRC (National Centre for Infection Prevention and Management) (SS, MR), and an Imperial College JRF (CET).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jinf.2019.02.005.

References

1. Commons RJ, Smeesters PR, Proft T, Fraser JD, Robins-Browne R, Curtis N. Streptococcal superantigens: categorization and clinical associations. *Trends Mol Med* 2014;**20**:48–62.
2. Zeppa JJ, Kasper KJ, Mohorovic I, Mazzuca DM, Haeryfar SMM, McCormick JK. Nasopharyngeal infection by *Streptococcus pyogenes* requires superantigen-responsive Vbeta-specific T cells. *Proc Natl Acad Sci USA* 2017;**114**:10226–31.
3. Paillot R, Darby AC, Robinson C, Wright NL, Steward KF, Anderson E, et al. Identification of three novel superantigen-encoding genes in *Streptococcus equi subsp. zooepidemicus*. szeF, szeN, *and* szeP. *Infect Immun* 2010;**78**:4817–27.
4. Chochua S, Metcalf BJ, Li Z, Rivers J, Mathis S, Jackson D, et al. Population and whole genome sequence based characterization of invasive group A streptococci recovered in the United States during 2015. *MBio* 2017;**8**:e01422-17.
5. Kapatai G, Coelho J, Platt S, Chalker VJ. Whole genome sequencing of group A *Streptococcus*: development and evaluation of an automated pipeline for *emm* gene typing. *PeerJ* 2017;**5**:e3226.
6. Turner CE, Bedford L, Brown NM, Judge K, Török ME, Parkhill J, et al. Community outbreaks of group A *Streptococcus* revealed by genome sequencing. *Sci Rep* 2017;**7**:8554.
7. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;**18**:821–9.
8. Page AJT B, Keane JA. Multilocus sequence typing by blast from de novo assemblies against PubMLST. *J Open Source Softw* 2016;**1**:118.
9. Turner CE, Abbott J, Lamagni T, Holden MT, David S, Jones MD, et al. Emergence of a new highly successful acapsular group a *Streptococcus* clade of genotype *emm*89 in the United Kingdom. *MBio* 2015;**6**:e00622.
10. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2016;**2**:e000056.
11. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–13.
12. Reglinski M, Lynskey NN, Sriskandan S. Modification of the classical Lancefield assay of group A streptococcal killing to reduce inter-donor variation. *J Microbiol Methods* 2016;**124**:69–71.
13. Unnikrishnan M, Altmann DM, Proft T, Whaid F, Cohen J, Fraser JD, et al. The bacterial superantigen streptococcal mitogenic exotoxin Z is the major immunoactive agent of *Streptococcus pyogenes*. *J Immunol* 2002;**169**:2561–9.
14. Turner CE, Kurupati P, Jones MD, Edwards RJ, Sriskandan S. Emerging role of the interleukin-8 cleaving enzyme SpyCEP in clinical *Streptococcus pyogenes* infection. *J Infect Dis* 2009;**200**:555–63.
15. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011;**8**:785–6.
16. Athey TB, Teatero S, Sieswerda LE, Gubbay JB, Marchand-Austin A, Li A, et al. High incidence of invasive group A *Streptococcus* disease caused by strains of uncommon *emm* types in Thunder Bay, Ontario, Canada. *J Clin Microbiol* 2016;**54**:83–92.
17. Gherardi G, Vitali LA, Creti R. Prevelent *emm* types among invasive GAS in Europe and North America since year 2000. *Front Public Health* 2018;**6**:59.
18. Proft T, Fraser JD Streptococcal superantigens: biological properties and potential role in disease. *Streptococcus pyogenes*. Ferretti JJ, Stevens DL, Fischetti VA, editors. Oklahoma City: Basic Biology to Clinical Manifestations [Internet]; 2016.

19. Kasper KJ, Zeppa JJ, Wakabayashi AT, Xu SX, Mazzuca DM, Welch I, et al. Bacterial superantigens promote acute nasopharyngeal infection by Streptococcus pyogenes in a human MHC Class II-dependent manner. *PLoS Pathog* 2014;**10**:e1004155.
20. Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long SW, et al. Evolutionary pathway to increased virulence and epidemic group A Streptococcus disease derived from 3615 genome sequences. *PNAS* 2014;**111**:E1768–76.
21. Al-Shahib A, Underwood A, Afshar B, Turner CE, Lamagni T, Sriskandan S, et al. Emergence of a novel lineage containing a prophage in *emm*/M3 group A *Streptococcus* associated with upsurge in invasive disease in the UK. *Microb Genom* 2016;**2**:e000059.
22. Afshar B, Turner CE, Lamagni TL, Smith KC, Al-Shahib A, Underwood A, et al. Enhanced nasopharyngeal infection and shedding associated with an epidemic lineage of *emm*3 group A *Streptococcus*. *Virulence* 2017;**8**:1390–400.
23. Kansal RG, Nizet V, Jeng A, Chuang WJ, Kotb M. Selective modulation of superantigen-induced responses by streptococcal cysteine protease. *J Infect Dis* 2003;**187**:398–407.