Genome Analysis of *Streptococcus pyogenes* Associated with Pharyngitis and Skin Infections

Joe Ibrahim[1], Jonathan A. Eisen[2], Guillaume Jospin[2], David A. Coil[2], Georges Khazen[3], Sima Tokajian[1]*

1 Department of Natural Sciences, Lebanese American University, School of Arts and Sciences, Byblos, Lebanon, 2 University of California Davis Genome Center, Davis, California, United States of America, 3 Department of Computer Science and Mathematics, Lebanese American University, School of Arts and Sciences, Byblos, Lebanon

* stokajian@lau.edu.lb

Abstract

Streptococcus pyogenes is a very important human pathogen, commonly associated with skin or throat infections but can also cause life-threatening situations including sepsis, streptococcal toxic shock syndrome, and necrotizing fasciitis. Various studies involving typing and molecular characterization of *S. pyogenes* have been published to date; however next-generation sequencing (NGS) studies provide a comprehensive collection of an organism's genetic variation. In this study, the genomes of nine *S. pyogenes* isolates associated with pharyngitis and skin infection were sequenced and studied for the presence of virulence genes, resistance elements, prophages, genomic recombination, and other genomic features. Additionally, a comparative phylogenetic analysis of the isolates with global clones highlighted their possible evolutionary lineage and their site of infection. The genomes were found to also house a multitude of features including gene regulation systems, virulence factors and antimicrobial resistance mechanisms.

Introduction

Streptococcus pyogenes, also referred to as Group A Streptococcus (GAS) for harboring Lancefield group A antigen, is a clinically important human pathogen [1]. Despite its limited prevalence in modern times as compared to other pathogens, the myriad of infections it causes are still commonly lethal [2]. Streptococcal infections range from localized throat infections such as tonsillitis or pharyngitis, to invasive infections such as sepsis, necrotizing fasciitis and streptococcal toxic shock syndrome (STSS) [1, 2]. Most of the infections are seen in children aged four to seven years [3], with penicillin still being effectively used for treatment. Drug resistant clones however, have been increasingly reported globally [4, 5], and were attributed to environmental and intracellular resistance mechanisms [6]. Treatment complications usually arise with the use of other antimicrobial agents when the patient is allergic to penicillin.

A number of genome-encoded virulence factors such as pili, M proteins, leukocidins, streptolysins, complement inhibiting proteins, immunoglobulin-degrading enzymes, and

superantigens have been detected in *S. pyogenes*, [1, 2], in addition to efflux pumps and leuko-cyte evasion strategies [7]. Interestingly horizontal gene transfer (HGT) and prophage integra-tion are common amongst *S. pyogenes* genomes giving them plasticity and genomic variation [7]. These factors collectively confer additional virulence and resistance capabilities and alter the regulation of existing genes [8]. The *emm* gene, encoding the M protein contains con-served, semi-conserved, and hypervariable regions, and is thus used as an epidemiological marker for GAS [9]. *emm* occurrence amongst GAS strains is linked to geographic localization [10].

Globally, the GAS disease burden is not yet fully quantified, 2005 estimates showed that a minimum of 18.1 million people were suffering from invasive diseases, with another 111 mil-lion cases of streptococcal pyoderma and 616 million cases of pharyngitis recorded. 1.78 mil-lion additional incident cases were estimated to occur each year [11]. In the Middle East and in Lebanon in particular, studies linked to the emergence and epidemiology of *S. pyogenes* are still very limited [11]. Previous studies, focused at establishing *emm* and pulse field gel electro-phoresis (PFGE) based typing schemes [12, 13]. *emm*1, 22, and 28, which were the most com-mon types, were reported as being susceptible to penicillin and vancomycin but resistant to erythromycin and clindamycin [12]. Karaky et al. in a more recent study detected 33 *emm* types and subtypes with the dominant being *emm*1, *emm*22, *emm*28, *emm*88 and *emm*4 and with 10% of the isolates being resistant to erythromycin and 3% resistant to erythromycin and clindamycin [13]. In this study we sequenced the genomes of nine *S. pyogenes* isolates repre-senting the most commonly recovered *emm* types to build on the previous findings and to elu-cidate the molecular epidemiology, genomics, and phylogenomics of this important human pathogen.

Materials and Methods

Ethical Approval

Ethical approval was not required as clinical isolates were collected and stored as part of rou-tine clinical care. Clinical isolates and patient records/information were anonymous and de-identified prior to analysis

Bacterial isolates and genomic DNA extraction

This study was conducted on nine *S. pyogenes* bacterial isolates previously collected from the American University of Beirut Medical Center (AUB-MC). The samples were recovered from throat and pus swabs of patients with streptococcal infections during the period from August 2010 to November 2011, and were chosen to cover the common *emm* types in the country (Table 1). The isolates were cultured overnight on Trypticase Soy Agar (TSA) (Bio-Rad, USA) medium. DNA was extracted using the Nucleospin Tissue kit (Macherey-Nagel, Germany) fol-lowing the manufacturer's instructions.

Genome sequencing

DNA extracted from each *S. pyogenes* isolate (50-ng/sample) was prepared for sequencing with the use of the Nextera XT DNA Sample Prep Kit (Illumina). Clean up was performed using the AMPure XP PCR purification beads (Agencourt, Brea, CA, USA). The resulting individual DNA libraries with fragment sizes ranging from 500–1000 bp were quantified by quantitative PCR on a CFX96 (Bio-Rad, USA) in triplicate at two concentrations, 1:1000 and 1:2000, using the Kapa library quantification kit (Kapa Biosystems, Woburn, MA, USA). Based on the indi-vidual library concentrations, equimolar pools of the indexed libraries were prepared at a

Table 1. Epidemiological Information on clinical *S. pyogenes* isolates. The relationship between *emm* types [14], MLST types, and diseases [15] in addition to site of isolation. emm pattern A-C is usually linked to upper respiratory tract infections (URT), pattern D is linked to skin infections while pattern E represents a generalist group.

Sample Name	Sex	Age	Origin	Specimen	Disease	ST-Type	*emm* typing in [12]	Virulence Factors in [13]	*emm* Pattern	Tissue preference/ Associated Disease
SP1	F	48	Lebanon	Swab, throat	Pharyngitis	ST-36	12	Cys. Prot. B—SpeL	A-C	Throat/URT
SP2	F	47	Lebanon	Swab, pus	Dermatitis	ST-304	108	Cys. Prot. B—SpeL	D	Skin/-
SP3	M	7	Lebanon	Swab, throat	Pharyngitis	ST-101	89	Cys. Prot. B—SpeL	E	No preference/URT
SP4	M	8	Lebanon	Swab, throat	Pharyngitis	ST-52	28	Cys. Prot. B—SpeL—ssa	E	No preference/ Invasive
SP5	M	35	Lebanon	Swab, pus	Dermatitis	ST-28	1	SpeA—Cys. Prot. B—SpeL	A-C	Throat/Invasive
SP6	F	23	Lebanon	Swab, throat	Pharyngitis	ST-101	89	SpeG	E	No preference/URT
SP7	M	4	Lebanon	Swab, throat	Pharyngitis	ST-46	22	SpeA—Cys. Prot. B—ssa—smeZ	E	No preference/Invasive
SP8	F	1	Lebanon	Swab, throat	Pharyngitis	ST-109	85	SpeA—SpeH—ssa	D	Skin/Skin
SP10	M	7	Lebanon	Swab, throat	Pharyngitis	ST-167	118	Cys. Prot. B—SpeH—SpeK—smeZ	E	No preference/-

doi:10.1371/journal.pone.0168177.t001

concentration of at least 1 nM using 10 mM Tris-HCl (pH 8.0) and 0.05% Tween 20. Pooled paired-end libraries were sequenced to a read length of at least 250 bp.

Genome assembly

De novo assembly of the sequenced genomes was done using A5 assembler (v. 20130627) with default assembly parameters. This pipeline automates the processes of data cleaning, error correction, contig assembly, scaffolding, and quality control [16, 17].

Genome annotation and gene detection

The assembled genomes were annotated using the RAST server (http://rast.nmpdr.org) that uses subsystems technology to assign gene function. RAST was also used for the identification of protein encoding genes, rRNAs, and tRNAs [18, 19]. The SEED viewer service from RAST in combination with the web tools provided by the Center for Genomic Epidemiology (CGE) website (www.genomicepidemiology.org) were used to generate a detailed list of the genetic elements in question. The ResFinder 2.1 web server was used to identify acquired antimicrobial resistance genes present on the bacterial genome [20]. Due to ResFinder's inability to confirm the functional integrity, and levels of gene expression and resistance arising to acquired mutations in housekeeping genes, a hybrid resistance profile was generated using the ResFinder hits in addition to phenotypical data from previously published studies. The VirulenceFinder 1.2 service from the same website was used to identify additional virulence-specific genes [21]. The PathogenFinder 1.1 service from CGE, was used to obtain an overview of the genomic pathogenic gene families [22].

Multi-locus sequence typing (MLST)

MLST typing of the isolates was carried out using CGE's MLST 1.7 server to detect sequence polymorphisms within the *gki*, *gtr*, *muri*, *muts*, *recp*, *xpt*, and *yqil* genes [23].

Phage and mobile element detection

Phage detection was done using the publicly available Phage Search Tool (PHAST) (http://phast.wishartlab.com/index.html) [24]. This tool provides the closest identity match for detected phages in addition to their site of integration. Putative insertion elements were double checked using BLASTx with an identity threshold of 80%. Putative phage insertion sequences were then annotated using RAST in order to determine the genes they encode.

Phylogenetic analysis

To determine the phylogenetic relatedness a concatenated marker gene maximum-likelihood tree was constructed using a number of *S. pyogenes* reference genomes chosen based on BLAST similarity results, clonal complexes and sequence types (MGAS6180 CP000056, MGAS10394 CP000003, M1476 AP012491, M1GAS SF370 NC002737, NZ131 CP000829, A20 CP003901, 7F7 PRJNA238516, and *S. pneumoniae* R6 AE007317 as an outlier strain). The genomes were first processed with PhyloSift [25], the tree was then constructed using FastTree [26], visualized and edited with Dendroscope [27]. Pairwise alignment and visualization of our selected genomes with the respective reference strains was achieved through the Mauve aligner [28] using defaults settings.

Nucleotide Sequence Accession Numbers

The whole-genome shotgun projects have been deposited at DDBJ/EMBL/GenBank with accession numbers AYPA00000000 (SP1-LAU), AWOZ00000000 (SP2-LAU), AWPA00000000 (SP3-LAU), AWPB00000000 (SP4-LAU), AWPC00000000 (SP5-LAU), AWPD00000000 (SP6-LAU), AWPE00000000 (SP7-LAU), AWPF00000000 (SP8-LAU), and AWPG00000000 (SP10-LAU) [29].

Results and Discussion

Sequencing resulted in an average of 2,503,465 paired-end reads per isolate, with the average being 1,976,732 high-quality reads following quality trimming and error correction. The average sequence coverage of the whole genomes was 293X and the minimum sequence coverage was 85X for the SP6 isolate. The average N50 for the assemblies was 196,439 bp with the lowest being 98,990 bp for SP7. The initial assemblies resulted in an average of 77 contigs per isolate all of which greater or equal to 500 bp in length. During scaffolding, some contigs were merged based on short overlaps and read-pair information, yielding a reduced final average of 70 contigs per isolate. The complete details and statistics of the sequencing and assembly results for each isolate are shown in Tables 2 & 3 [29].

The average genome size was 1.83 Mbp and the G+C content ranged between 38.2% and 38.5% (Table 2) with an average of 38% both falling within the general ranges for the species [7]. An average of 1844 open reading frames (ORFs) was detected in the sequenced genomes,

Table 2. Genome assembly statistics.

	SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8	SP10
Genome size (bp)	1,925,871	1,727,943	1,745,842	1,906,369	1,813,544	1,733,546	1,953,601	1,917,411	1,771,196
Number of contigs (> = 500 bp)	155	29	21	28	26	19	158	186	12
Average contig size (bp)	12,425	59,584	83,135	68,085	69,752	91,239	12,365	10,309	147,600
Longest contig size (bp)	343,437	653,370	659,470	249,942	752,626	674,975	207,328	250,346	781,011
GC content (mol %)	38.4	38.4	38.4	38.2	38.4	38.4	38.4	38.5	38.4
N50 (bp)	209,002	280,940	167,926	205,587	308,655	116,374	98,990	104,600	275,797

doi:10.1371/journal.pone.0168177.t002

Table 3. Gene prediction and annotation summary of the nine *S. pyogenes* isolates.

	SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8	SP10
# of predicted genes	1991	1694	1725	1905	1832	1698	2041	1971	1740
# of pseudogenes	22	26	27	24	18	30	31	37	34
# of predicted potein-coding genes	1907	1669	1711	1923	1824	1692	1937	1866	1740
# of tRNA genes	56	65	56	54	57	54	56	56	56
# of rRNA genes	7	7	6	6	9	4	5	10	4
# of subsystems	317	313	311	317	319	314	316	322	317

doi:10.1371/journal.pone.0168177.t003

1752 of which encoding proteins, 57 tRNAs and 6 rRNAs. The average gene size was 868 bp with coding sequences covering around 85% of the genomes (Table 3). This was consistent with the averages obtained within the used reference strains, and in general amongst *S. pyogenes* strains [30]. An average of 315 discrete biological subsystems were identified (Table 3) the majority of which were related to nutrient metabolism, virulence factor and bacterial cell wall synthesis subsystems (Fig 1). Only nine amino acid synthesis subsystems were detected, reflecting the fastidious growth requirements of *S. pyogenes* [31]. The inherent lack of biosynthetic pathways however, is offset by the relative abundance of membrane transport systems (37 subsystems)—that scavenge resources from the environment—including around 10 putative ABC transporters, used specifically for peptide uptake. Virulence factors and defense mechanisms make up on average 10% of the genome, with an equal percentage allocated to cell wall and capsule proteins. Considering the small size of the streptococcal genome, virulence related genes take up a significant part, which goes in line with *S. pyogenes* being a strict human pathogen [1]. The various sequence types (ST) identified based on MLST database search were: ST-36, ST-304, ST-101, ST-52, ST-28, ST-46, ST-109, and ST-167. ST were then correlated to *emm* types generated in an earlier study [12] as well as to the disease associated
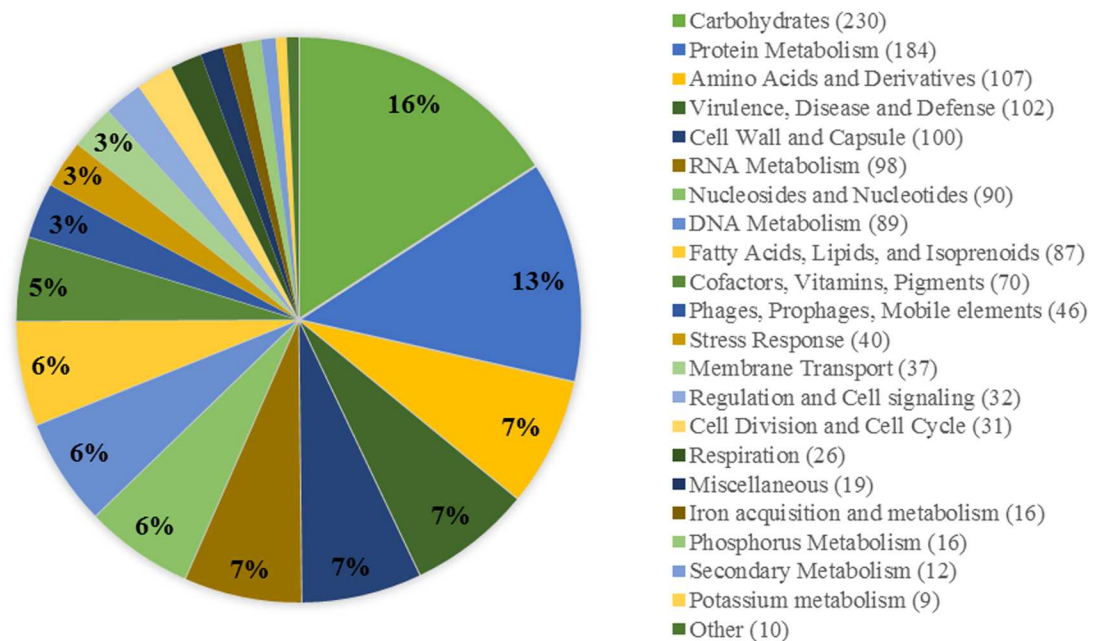

Legend:
- Carbohydrates (230)
- Protein Metabolism (184)
- Amino Acids and Derivatives (107)
- Virulence, Disease and Defense (102)
- Cell Wall and Capsule (100)
- RNA Metabolism (98)
- Nucleosides and Nucleotides (90)
- DNA Metabolism (89)
- Fatty Acids, Lipids, and Isoprenoids (87)
- Cofactors, Vitamins, Pigments (70)
- Phages, Prophages, Mobile elements (46)
- Stress Response (40)
- Membrane Transport (37)
- Regulation and Cell signaling (32)
- Cell Division and Cell Cycle (31)
- Respiration (26)
- Miscellaneous (19)
- Iron acquisition and metabolism (16)
- Phosphorus Metabolism (16)
- Secondary Metabolism (12)
- Potassium metabolism (9)
- Other (10)

Fig 1. Average subsystem category distribution in the streptococcal genome. The pie chart shows the subsystem related genes as a percentage of the whole genomic content. Numbers next to the label entries indicate the number of predicted genes involved in a particular subsystem.

doi:10.1371/journal.pone.0168177.g001

with these typing patterns (Table 2) and were found to be mostly compatible with the site of isolation.

Gene Regulation

Sigma factors (σ factors), used for bacterial transcription initiation, can be highly variable. *Bacillus* spp. have up to 18 σ factors while *Haemophilus influenzae* has only three [32]. A total of three to four factors were detected in the studied isolates conforming to the number of σ factors found in small bacterial genomes (1–4 σ factors) [33]. σ^{70} (*ropD*) factor, the main σ factor of the species, and ComX were detected in all the isolates in this study. ComX, a homologue of the σ factor present in *S. pneumoniae*, transcriptionally regulates competence specific genes serving in DNA uptake and integration [34], and is encoded by two ORFs [35]. Sequence of an additional putative σ factor was detected in three of the isolates (SP2, SP3 and SP7), which could be linked to a minor heat competence transcriptional regulator found in *Escherichia coli*, σ^{24} [36]. The acquisition of σ^{24} might be an adaptation to help it survive, as *S. pyogenes* often encounters high temperatures inside the human host due to inflammatory responses and immune defenses [31].

Comparable to other microorganism, *S. pyogenes* responds to environmental changes by choosing the appropriate transcriptional signals [33] in addition to a variety of stress response proteins, specifically proteases and highly conserved stress regulation genes [31]. Osmoregulation is maintained by a two-component system comprising aquaporin Z and two osmotically inducible outer membrane proteins, OmpA and OsmY, the genes of which were detected in all the isolates. Choline and betaine uptake and biosynthesis genes (*opuAA*, *opuAB*, *opuAC*, *proV*, *proX* and *chA*) were also identified. These are highly conserved elements encoding proteins that can act as both osmoprotectants and energy sources [37]. Similar to other lactic acid producing streptococcal species, acid stress response in *S. pyogenes*, is achieved by proton translocation through the F_0F_1 ATP synthase [38]. Eight genes encoding structural parts of the ATP synthase as well as cold and heat shock response proteins were identified. CspA (*cspA*), a cold shock response protein belonging to the Csp family of proteins, and which allows *S. pyogenes* to cope with temperature reductions, was also detected [39]. The heat shock response family of proteins was detected as part of the *dnaK* gene cluster that houses a range of protein families including ribosomal methyltransferases, and chaperones (DnaJ, DnaK). The *dnaK* gene cluster is still poorly characterized in streptococci, with the exception of its high conservation among firmicutes [40]. The main role of *dnaK* gene cluster is to prevent the aggregation of heat denatured proteins inside the cell following heat shock [41]. Its activation has been reported inside macrophages, during an infection with *Salmonella enterica* [41], which consequently could be related to pathogenesis.

Gene regulation in *S. pyogenes* is stringently controlled mainly in response to environmental changes, and often linked to the type of infection superficial versus invasive [42]. To date, studies have identified around 13 two-component regulatory systems in *S. pyogenes*, in addition to more than a hundred putative independent transcriptional regulators [43, 2]. Perhaps one of the most important regulons detected in all the sequenced genomes is the control of virulence regulatory system (covRS), otherwise known as the capsule synthesis regulon (csrRS). CovRS is a two-component system made up of a membrane-bound sensor kinase (CovS), and a DNA binding response regulator (CovR) [44]. The regulon governs streptococcal virulence in response to environmental conditions such as pH, temperature, and ion concentrations [2]. This controls the expression of 10 to 15% of the streptococcal genes either directly or indirectly [45]. CovRS positively regulates the expression of several virulence factors including the streptococcal cysteine protease SpeB, and the hyaluronic acid capsule synthesis proteins [42].

Unidirectional mutations in the covRS regulon have been reported to alter the expression of virulence factors affecting the progression of streptococcal invasive diseases [46]. The multiple-gene regulator (Mga) controls the expression of many virulence genes, including the M family of proteins, in response to environmental carbohydrate availability [47]. Mga regulation represents phase and metabolite-dependent regulation in *S. pyogenes*, and the regulon is essential for the organism when shifting from the colonization to deep tissue invasion [48].

Chromosomal replication is another important aspect of the microbial cell cycle; it is initiated by the binding of the DnaA protein on specific DnaA boxes located in the replication origin [oriC] [49]. With the increased availability in complete genome sequences, the proper identification of replication origins [oriCs] and their characterization has become essential for the analysis of bacterial whole genomes [50]. For our purpose, we used the web-based Ori-Finder system as described by Gao & Zhang [50] to identify oriCs and DNA boxes in our assembled genomes and the results subsequently BLASTed against the DoriC database to confirm the reliability of the prediction [51]. All of the isolates with the exception of SP7 and SP8, exhibited a 154 nucleotide oriC sequence similar to that found in the M1GAS SF370 strain. The average oriC AT content was 0.7078 and it housed 3 DnaA boxes; TGTGGAAAA, TTATCCACA and TTATCCACT. The OriCs of SP2 and SP4 were very similar to that of the MGAS6180 and A20 strain respectively. The SP8 oriC was found to be highly similar to that of the NZ131 strain being 232 nucleotides in length and having a 0.6897 AT content. Interestingly, a 173 nucleotide long oriC was identified in SP7 which was 63.01% composed of AT residues and interspersed by five DnaA boxes; three TTATCCACAs and two TGTGAATAA, a result that was uncommon amongst *S. pyogenes* as the species mostly possesses either three or seven DnaA boxes in the replication origin region an example of the latter being strains MGAS15252 and MGAS1882 [51]. Originally, DnaA boxes were shown to control microbial initiation mass [52], whereas more recent studies have outlined the organizational importance of the oriC region in influencing bacterial proliferation and by that invasion and pathogenesis [53]. Some researchers believe that this organization is conserved among bacterial species while others do not agree [54]. Our findings were in agreement with the former claim as the oriCs we detected in our biotypes were very similar to those found in the reference strains, all of which generally exhibited species-wide homology.

Virulence Factors

S. pyogenes possesses an arsenal of virulence factors that targets and impairs the immune system [2]. Many of the genes constituting the streptococcal virulome were identified in the sequenced genomes (Table 4). Most of the virulence factors detected in this study, were scattered throughout the genome and were are not strictly part of a pathogenicity island, which was in harmony with previous findings [31]. For the most part, these elements are conserved amongst strains, apart from the ones carried by phages and prophages [55]. Genes coding for some exotoxins, including the streptokinase, hyaluronate lyase, and nicotine adenine dinucleotide glycohydrolase (NADGH), were highly conserved amongst the isolates (Table 4); these normally induce apoptosis in neutrophils and macrophages [2]. Interestingly, the cyclic AMP (cAMP) factor gene *cfa* was detected in all the sequenced genomes and reference strains. *cfa* was originally thought to be exclusive to Group B Streptococci (GBS) [56], however its presence was confirmed in GAS following the original sequencing work of Ferretti et al. [31]. The cAMP factor is an extracellular protein that causes synergistic lysis of host erythrocytes [57], with its role in pathogenesis not being fully understood. Numerous related streptococcal species are naturally competent for transformation via a pathway yet to be described [31], hence the presence of the *cfa* gene could be the outcome of mobilization and horizontal gene transfer (HGT).

Table 4. List of genes attributable to virulence traits in the streptococcal genomes.

Virulence Trait	SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8	SP10	Gene(s) with potential for conferring virulence traits
Antiphagocytic M protein	+	+	+	+	+	+	+	+	+	*emm*, *enn*X, *fbp*, *iga*R
Streptokinase	+	+	+	+	+	+	+	+	+	ska
CAMP factor	+	+	+	+	+	+	+	+	+	*cfa*
Streptolysin O	+	+	+	+	+	+	+	+	+	*slo*
Streptolysin S	+	+	+	+	+	+	+	+	+	*sag*B, C, D, E, F, H, I, asn-ORF, ABC transporter
Hyaluronate lyase precursor	+	+	+	+	+	+	+	+	+	*hyl*
Hyaluronan synthase	+	+	-	+	+	-	-	+	+	*has*A
Streptococcal pyrogenic exotoxin A*	-	-	-	-	+	-	-	-	-	*spe*A
Cysteine Protease B*	+	+	+	+	+	+	+	+	+	*spe*B
Streptococcal pyrogenic exotoxin C*	-	-	-	-	-	-	+	-	-	*spe*C
Streptococcal pyrogenic exotoxin G	+	+	+	+	+	+	+	+	+	*spe*G
Streptococcal pyrogenic exotoxin H*	-	-	-	-	-	-	-	-	-	*spe*H
Streptococcal pyrogenic exotoxin I*	+	-	-	-	-	-	-	+	-	*spe*I
Streptococcal pyrogenic exotoxin J	-	+	-	-	+	-	-	-	-	*spe*J
Streptococcal pyrogenic exotoxin K*	+	-	+	+	-	+	+	-	+	*spe*K
Streptococcal pyrogenic exotoxin L*	-	-	-	+	-	-	-	-	-	*spe*L
Streptococcal pyrogenic exotoxin M*	-	-	-	-	-	-	-	-	-	*spe*M
Streptococcal mitogenic exotoxin Z	+	+	+	+	+	+	+	+	+	*smeZ*
Streptococcal superantigen A*	+	-	-	-	-	-	+	+	-	*ssa*A
C5a peptidase	+	+	+	+	+	+	+	+	+	*scp*A
Secreted endo-beta-N-acetylglucosaminidase	+	+	+	+	+	+	+	+	+	*ndo*S
Streptococcal inhibitor of complement	-	-	-	-	+	-	-	-	-	*sic*
Exotoxin nucleases	+	+	+	+	+	+	+	+	+	*spd*1, 2, 3, 4, *sda*
Immunoglobulin-binding protease	+	+	+	+	+	+	+	+	+	*ide*S
Adhesins and invasins	7	4	6	7	3	6	9	7	6	*fba*, fibronectin binding proteins, serum opacity factors
Collagen-like surface proteins	+	+	+	+	+	+	+	+	+	*scl*A, B

. + and–indicate the presence or absence of genes respectively, numbers indicate the number of pertaining genes detected, while * indicates phage-encoded superantigens.

Slo and *sls*, encoding streptolysin S and streptolysin O, were also detected in all of the isolates (Table 4). SLO and SLS are leukocidins that disrupt host cell membranes and induce apoptosis in phagocytes [58]. SLS is virtually secreted by all GAS [59]; it increases streptococcal resistance to phagocytosis by distorting neutrophil membranes [60], in addition to disrupting erythrocytes, lymphocytes, and even platelets [61]. Individual *S. pyogenes* genomes can greatly differ in their genetic composition, mainly due to the acquisition of exogenous genetic elements either through HGT, or bacteriophage integration [62]. The *sic* gene, encoding the streptococcal inhibitor of complement and uniquely detected in M1 and M57 GAS strains [63], was detected in SP5 (*emm*1) (Table 4) and in the M1 GAS reference strain. This conforms our findings to the work done by Bahnan et al. [12]. Host neutralizing antibodies exert selective pressure usually leading to high polymorphism in the *sic* gene even within the same *emm* type [64], however this was not observed in our study with both the M1 GAS reference and SP5 having homologous *sic* gene sequences.

The endo-beta-N-acetylglucosaminidase of streptococci, EndoS (*ndo*S), is a secreted immunoglobulin degrading enzyme that enhances the organism's ability to resist opsonophagocytosis,

and is encoded by the 3 Kbp *ndo*S gene [2]. *ndo*S was detected in all biotypes (Table 4) with its size ranging from 2943 bp to 3015bp, these size differences can be attributed to possible insertions or deletions. Individual BLASTing of the sequences, revealed even distribution between the two protein isoforms EndoS and EndoS F2, which are similar enzymes but with different oligosaccharide specificity [65]. Both however, are distinct from the EndoS2 that was previously only detected in M49 GAS strains [66], another fact in agreement with the findings of Bahnan et al. [12]. The heterologous expression of the EndoS in GAS strains other than the M1 serotype however, should not be easily dismissed as it has been shown to enhance virulence in murine models of invasive streptococcal infections [67]. Genes encoding the streptodornase B (*spd*1), an exotoxin nuclease, were detected in all of the isolates as well as in the chosen reference strains (Table 4); conserved in size and in genomic location. The gene was part of an extended SpeB-SpeF regulon and is flanked by a transcriptional regulator and a low temperature response protein. These gene sequences shared homology with their counterparts on bacteriophages, specifically the *S. pyogenes* phages 315.3, 315.6, and phage 9. Despite being traditionally phage encoded [68], the uniformity of the location of these sequences in the genomes indicates that they could be fixed remnants of phage integration events over time. Streptodornase D, also encoded on phages 315.3 and 315.6, was detected in six of the studied isolates (Table 4), while absent in the chosen reference strains. However, SEED similarity matches linked it to MGAS5005 and MGAS9429 strains, further elucidating the possibility of HGT. The immunoglobulin-binding protease, IdeS (*ide*S), was detected in all the biotypes (Table 4). BLAST revealed the protein products to be of the Mac-1 isoform with none being related to Mac-2 [69]. Despite impairing neutrophil mediated phagocytosis [70], IdeS is not essential for phagocyte resistance [71], which could explain the lack of an evolutionary pressure to drive diversity among the products of this gene.

Three to nine fibronectin binding (FBP) and collagen-like surface protein genes were additionally detected in the tested isolates (Table 4), these were present in three major genomic locations one gene was found downstream of the C5a peptidase gene, another set (two to three genes) was found downstream of the streptopain/streptopain inhibitor gene set, and the remaining genes were located in FCT region. These genes have different characteristics depending on *emm* type [72]. The FCT region, considered "a hot intergenomic recombinatorial site", is a ≈11 to 16 Kbp chromosomal region flanked by highly conserved genes, a chaperon and a hypothetical protein, between which a unique combination of conserved and semi-conserved loci was present [73, 74]. The organization of this region was determined manually by examining the gene loci in the isolates and comparing it with similar loci on known references. To date, nine distinct FCT variants have been reported [74]. All the biotypes had the FCT-3 arrangement with the exception of SP8, which belonged to the FCT-1 (Fig 2). FCT-1 is usually rare amongst GAS isolates and can be can be detected in M1 and M6 strains such as the SF370 and MGAS10394. After further sequence annotation the SP8 FCT-1 configuration was determined to be similar to that present in the M6 serotype containing the distinctive T-antigen backbone protein and the ancillary protein FctX. The FCT-3 configuration was comparable to that present in M3 and M49 serotypes with the putative chaperon protein SipA [74]. Additional putative genes were detected as part of the locus, these were mainly encoding sortases and signal peptidases. Clinically, FBP contributes to the adhesion of *S. pyogenes* cells to host cells, and enhances resistance to phagocytosis by averting the C3 convertase from depositing on the bacterial cell [75]. These proteins are vital to the pathogenesis of *S. pyogenes*; mutations were associated with a decrease in the epithelial cell adhesion efficiency by up to 10% in murine models [76, 77]. Even in GAS strains where high-affinity plasminogen binding proteins are not expressed, but FBP is present instead, plasminogen-mediated virulence can still be activated through a stable cell-associated enzymatic activity that lyses fibrin clots [78]. *Prt*F1
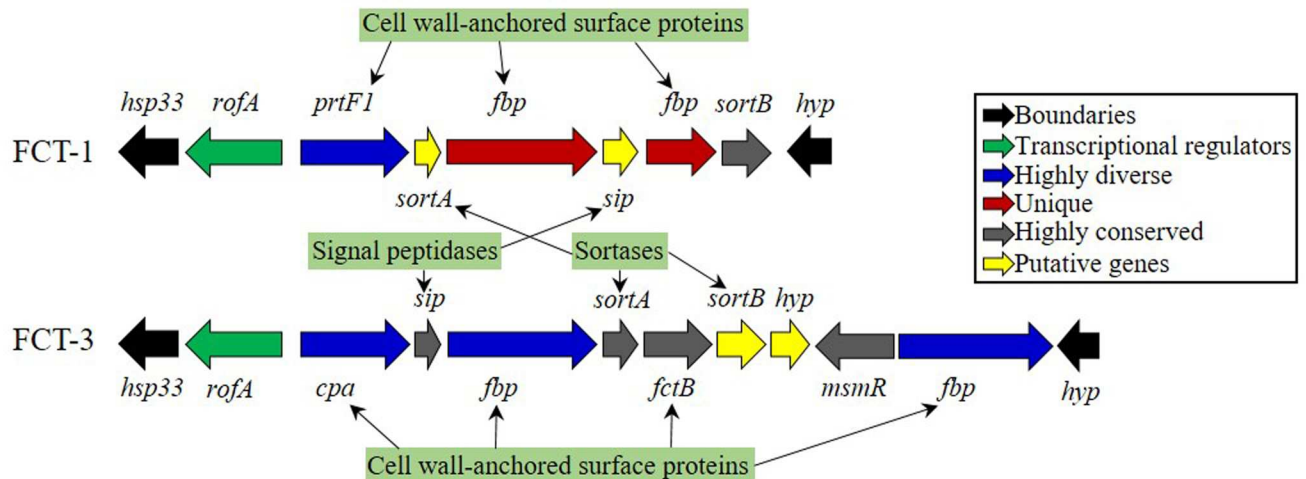
Fig 2. Fibronectin-Collagen-T-antigen (FCT) region organization of FCT types 1 and 3. FCT-1 (SP8) was identified in only one of the isolates whereas FCT-3 was seen in all of the rest. The region is flanked by a chaperon (*hsp*33) and a hypothetical protein (*hyp*). In between are mostly cell wall-anchored proteins namely fibronectin binding proteins (*fpb* namely *prt*F1/2) and collagen binding proteins (*cfa*). Sortases (*sort*) that modify surface proteins are also found in addition to signal peptidases (*sip*). A combination of unique, highly diverse, and highly conserved loci determines the FCT type.

doi:10.1371/journal.pone.0168177.g002

was identified in all the genomes, while *cpa* in six of them; both gene products could prime skin and throat infections [74] as is the case with our isolates.

The bacterial C5a peptidase (*scp*A), was detected in all isolates with an approximate size of 3550 bp (Table 4), it was identified strictly as part of the Mga regulon [79]. The locus extending from the *sme*Z gene through the *mf1* gene to the Mga regulon, is one of the best studied loci in the streptococcal genome for its virulence factors and inherent regulatory components [31]. This regulon is an independent, ubiquitous, multiple-gene regulator that controls the expression of numerous virulence genes namely those encoding the M family of proteins (*emm*, *mrp*, *arp*, and *enn*) that contribute to adhesion, invasion, and host immune evasion [47, 80]. It also governs the transcription of other non-M protein genes (*sic*, *scp*A, *scl*A) [81]. Mga regulates the expression of about 10% of the streptococcal genome [80] (Hondorp and McIver, 2007). In this work, regions associated with the Mga regulon showed high variability in genetic content; these included multiple M protein genes in addition to the C5a peptidase. BLAST results of the individual gene sequences showed the presence of the *emm* gene in all biotypes complemented by a gene encoding a FBP in six of the isolates (SP2, SP3, SP6, SP7, SP8 and SP10) as well as in the MGAS6180 reference strain. An immunoglobulin A receptor (*iga*R) was also detected in six of the isolates (SP2, SP3, SP4, SP7, SP8 and SP10). SP6 additionally carried the *enn*X gene in the operon, which was homologous to the one present in the reference strain NZ131 (Table 4). The genetic diversity of the Mga regulon detected here is in accordance with previous findings [31, 72, 80]. This diversity correlates with both tissue tropism and disease manifestation [72, 80]; usually, strains presenting one M family protein are limited to throat infections, and those exhibiting three or more proteins are implicated in more invasive infections [82]. Six of our isolates (SP2, SP3, SP4, SP6, SP7 and SP10) had M family protein genes, demonstrating their potential in causing invasive infections.

Clinically, many consider superantigens to be more significant than *emm* types in disease manifestation [83]. These pyrogenic exotoxins are perhaps the most important of all the streptococcal antigens [84] due to their ability to over stimulate the human immune system and contribute to tissue inflammation [85]. Excluding SpeG, SpeJ, and SmeZ, all superantigens are

phage encoded [86]. Particular serotypes, such as M types 1, 3, 5, 6, 14, 18, 19, and 24 are normally associated with throat infections and rheumatic fever, while other such as M types 2, 49, 57, 59, 60, and 61 are linked to pyoderma and which are associated with pyoderma and acute glomerulonephritis [87, 88]. In this study, seven distinct Sag profiles were identified, with SpeGKZ being the most prevalent (three out of nine) (Table 4). The conserved gene product SpeB, along with SpeG and SmeZ were detected in all the isolates as well as the used reference strains. SpeH on the other hand, has been reported to have low prevalence [89, 90] and was not detected in any of the isolates or reference strains. This superantigen, detected in M12 MGAS9429 and M1 GAS160490 strains for example, is considered a variable characteristic in different strains and is not attributed to distinct *emm* types [91]. Our findings were consistent with the worldwide dissemination of the species [90], and with previous findings [9]. The superantigen profile of SP1 which (SpeGIKZ) matched that of the M5 Manfredo strain that is associated with rheumatic fever [92]. The SP2 profile (SpeGJZ) can be associated to skin infections as it housed the same superantigens as the M59 MGAS15252 strain, a serotype isolated from pyodermal infections [93]. SP3, SP6, and SP10 (SpeGKZ) were similar in superantigens to the puerperal sepsis strain M28 MGAS6180 only lacking SpeJ [94]. SP4 was the only isolate to house SpeL (Table 4), this antigen was first detected in M3 serotype isolates causing toxic shock-like syndrome (TSLS) cases, and it transfer to the *S. pyogenes species* was attributed to phage HGT [95]. The SAg profiles of SP1, SP3, and SP6 also matched *emm* types 12 and 89 isolates, which are linked to puerperal sepsis and cellulitis [9]. Interestingly, SP5 of *emm* type 1 was the only isolate having SpeA, which also present in the M1 MGAS5005 and M1 GAS160490 reference strains. SpeA is a characteristic unique to *emm*1 types and associated with severe infections such as kidney failure [96, 97]. SP7's profile matched that of M3 MGAS315, indicating the implication of these two strain's in throat infections. SpeI, which has a very low prevalence [9], was found to be present in SP8 in addition to SpeG, Z and SSA. This contrasted previous findings that revealed the co-occurrence of SpeI and SpeH, on the same bacteriophage (φ370.2) [31], our findings were more harmonious with those of Commons et al. [86] who reported that SpeI may have been lost during integration in the streptococcal genome, or that it was encoded by a gene carried on an entirely different phage, both of which can explain the low prevalence of this SAg [90, 9]. SP8 was similar to the M4 MGAS10750 strain associated with pharyngitis and scarlet fever [98], with the addition of SpeI. SmeZ normally exhibits a mosaic structure and a wide allelic variation amongst *S. pyogenes* strains in an attempt to escape antibody neutralisation [89]. Three variant types were detected in our isolates, SP1 and SP8 were of the smez-3 type, SP3, 4, 6 and 7 were of the smez-8 type, while SP2, 5 and 10 were smez-12. These alleles can be correlated (but not limited) to single *emm* types, including *emm* 12, *emm* 59 and *emm* 89 respectively [89]. M59 MGAS15252 and M12 MGAS2096 strains were also found to have the corresponding SmeZ types 3 and 8 respectively. None of the variants were found to have the nonfunctional single base deletion, leaving SMEZ to propagate GAS pathogenesis.

Antimicrobial Resistance

Penicillin is still the standard drug for the management of streptococcal infections due to the organisms' lack of natural resistance against it, with macrolides being alternatively used in patients with penicillin hypersensitivity [99, 2]. This led to the development of resistant strains either through efflux pumps (*mef*A) or through the ribosomal methylases (*erm*) [100]. All of the isolates undertaken exhibited an ABC transporter membrane-spanning permease linked to macrolide resistance (Table 5). These results matched with the ones previously reported by Karaky et al. [13], and supported the increase in macrolide resistance among *S. pyogenes*

strains [12,13]. In the Middle East, macrolide resistance is at 10%-23%, which is significantly higher than in developed countries such as Germany (2.6%), Belgium (3.3%), and Spain (7.6%), where a restrictive use of macrolides was instituted [101, 102, 103].

Fluoroquinolone resistance is another feature of GAS that has been reported by several studies [104, 105, 106], it occurs at two levels: a low-level resistance due to mutations in the quinolone-resistance-determining region (QRDR) in the topoisomerase IV (*par*C and *par*E) [102], and a less-frequently prevalent high level [107] caused by additional mutations in *gyr*A and *gyr*B genes of DNA gyrase [106]. In this study both levels of fluoroquinolone resistance were detected (Table 5). Our findings are concordant with the global increase in fluoroquinolone resistance, with resistance rates over a two-year span in Belgium and Spain, rising from 4.3% and 1.9% to 21.6% and 30.8% respectively [106, 104]. In contrast, resistance in the studied isolates to fluoroquinolones was not limited to a set of *emm* types.

Tetracycline resistance in *S. pyogenes* is usually conferred through ribosomal protection genes namely *tet*M and *tet*O [108], and the elongation factors EF-G and EF-T [109]. SP7 and SP8 carried tetracycline resistance factors, specifically *tet*M and EF-G (Table 5), suggesting a multi-drug resistant European clone lineage [110]. The same isolates also exhibited a Multi antimicrobial extrusion protein (Na(+)/drug antiporter) efflux pump conferring tetracycline resistance. Although not very common amongst streptococci [111], this pump is possibly analogous to the one detected in *S. aureus* [112]. It is noteworthy in this respect that tetracycline resistance conferring genes are often associated with macrolide resistance genes on the same mobile element, hence the unregulated use of tetracycline and macrolides may reciprocally augment resistance to both agents [113, 114].

Bacteriophages, mobile elements and genomic recombination

In Gram-positive organisms, HGT transduction via bacteriophages often causes the most important genomic alterations and confers pathogenic traits [115]. The streptococcal genome in particular is highly dynamic owning to the numerous phage integration sites and transposable elements that can make up to 7–14% of the total genome [31, 116]. Phage-encoded genes are largely responsible for the pathogenesis and invasiveness of *S. pyogenes* [2]. Moreover, phages widely contribute to the diversity of *S. pyogenes* strains; approximately 90% of the streptococcal genomic content is shared even among different serotypes with the exclusion of unique prophage-bound sequences [117]. Numerous confirmed and putative prophages and prophage-like elements were detected (Table 6). The φ315.x family of *S. pyogenes* specific phages was the most prevalent type followed by the P9 bacteriophage. These phage loci were found encoding

Table 5. List of gene products identified in the isolates conferring resistance to respective antimicrobial agents.

Gene Product Features	SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8	SP10	Antimicrobial Resistance
DNA gyrase subunit A	+	+	+	+	+	+	+	+	+	Fluoroquinolones
DNA gyrase subunit B	+	+	+	+	+	+	+	+	+	
Topoisomerase IV subunit B	+	+	+	+	+	+	+	+	+	
Topoisomerase IV subunit C	+	+	+	+	+	+	+	+	+	
ABC transporter membrane-spanning permease	+	+	+	+	+	+	+	+	+	Macrolides
Translation elongation factor G	-	-	-	-	-	-	+	+	-	Tetracyclines
Tetracycline resistance protein TetM	-	-	-	-	-	-	+	+	-	
Multidrug resistance efflux pump PmrA	+	+	+	+	+	+	+	+	+	MDR
Multi antimicrobial extrusion protein (Na(+)/drug antiporter)	-	-	-	-	-	-	+	+	-	

+ and–indicate the presence or absence of genes respectively.

doi:10.1371/journal.pone.0168177.t005

Table 6. Identity of putative phages and phage elements detected on the *S. pyogenes* genomes.

Sample Name	Phage Name	Position	No. of CDS	Size (Kbp)	GC %
SP1	P9 Phage*	543481–587686	58	44.2	39.5%
	φ315.2*△	1235237–1280343	59	45.1	38.6%
	φ315.2*	1753586–1838804	104	85.2	37.8%
	φ315.2*△	1836641–1850391	22	13.7	39.2%
	Temperate phage φNIH1.1*△	1856910–1921492	85	64.5	39.0%
SP2	φ315.2*△	36431–86156	70	49.7	36.9%
SP3	Enterococcus phage EFC-1	826891–851153	24	24.2	38.9%
	Bacillus phage Grass	1229568–1251639	22	22	39.7%
	φ315.4*	1702210–1746332	58	44.1	39.4%
SP4	Enterococcus phage EFC-1	1051712–1083210	27	31.4	36.8%
	φ315.6*△	1157140–1191141	45	34	39.2%
	φ315.2*	1678710–1710884	35	32.1	35.7%
	φ315.4*△	1815357–1861474	56	46.1	39.3%
	Temperate phage φNIH1.1*△	1865571–1906405	36	40.8	39.7%
SP5	φ315.3*	814075–836688	32	22.6	37.6%
	φ315.3*	940132–963402	31	23.2	38.5%
	P9 Phage*△	1206221–1259975	65	53.7	39.6%
	φ315.2*	1658995–1687543	24	28.5	36.3%
SP6	Temperate phage φNIH1.1*	485708–527064	52	41.3	39.0%
	Bacillus phage G	1041449–1074731	19	33.2	39.5%
	Shigella phage SfIV	1723507–1733053	11	9.5	42.1%
SP7	φ315.6*△	1113085–1154775	56	41.6	38.5%
	φ315.3*	1539782–1557690	27	17.9	36.4%
	φ315.4*△	1596065–1609564	22	13.5	36.7%
	Bacillus phage BCJA1c	1773969–1809715	32	35.7	35.8%
	φ315.4*	1804621–1885835	105	81.2	38.7%
	Bacillus phage G	1894717–1910959	25	16.2	42.2%
SP8	Temperate phage φNIH1.1*	580207–619535	53	39.3	38.3%
	φ315.2*△	1683466–1730855	68	47.3	38.0%
	P9 Phage*	1880672–1891879	23	11.2	40.0%
SP10	φ315.3*△	1401342–1434194	47	32.8	37.8%

* indicates Streptococcal specific phages
△ indicates confirmed complete phage sequences.

streptodornases A and B and a hyaluronidase, which are major virulence factors. It is noteworthy that the same set of prophages were integrated at different positions within the tested isolates, indicating genomic rearrangements [72]. The φ315.x phages encoded phage holin genes, these are associated with hot phage recombinatorial regions further adding to the variety introduced by phage integration events [72]. Interestingly, the same phage sequences (φ 315.2) were detected three times in the SP1 genome (Table 6). This is attributed to the fact that phage integration is an ongoing process, with some of phage sequences being probably remnants of old integration events [94], and/or the outcome of new acquisitions. Additionally, similar phages can code for different gene sets and thus still contribute to streptococcal genomic diversity even with multiple infections of the same phage. Non species-specific phages were also identified in our isolates, namely *Bacillus spp*. and *Enterococcus spp*. phages, these stress again at phage DNA uptake being an important route through which *S. pyogenes* acquires genetic determinants to

propagate and thrive in the host. Given that the species is a strict human pathogen, the actual mechanisms through which phages were acquired are not well understood. Previous work attributed this phenomenon to the presence of small signaling molecules in phages, analogous to those present in mammalian cells. Such molecules act as inducers, activating phage induction only when the environmental conditions are optimal to accommodate a competent recipient host organism such as *S. pyogenes* [118].

Phylogeny

Coupling WGS with phylogenetic analysis has been shown to yield a high discriminatory power when dealing with closely related isolates, and has allowed for more robust epidemiological analyses [119, 120]. Here phylogenetic analysis was performed based on 40 gene families, revealing three major clusters [25] (Fig 3). As reported by Wu et al., the gene families used for the analysis are universally present across bacterial taxa, have a low copy number variation across taxa, and can be used to produce robust phylogenetic trees reflecting as much as possible the evolution of the species from which the genes have originated [121]. The genomic plasticity of *S. pyogenes* is manifested in the first phylogenetic branch through the clustering of SP7
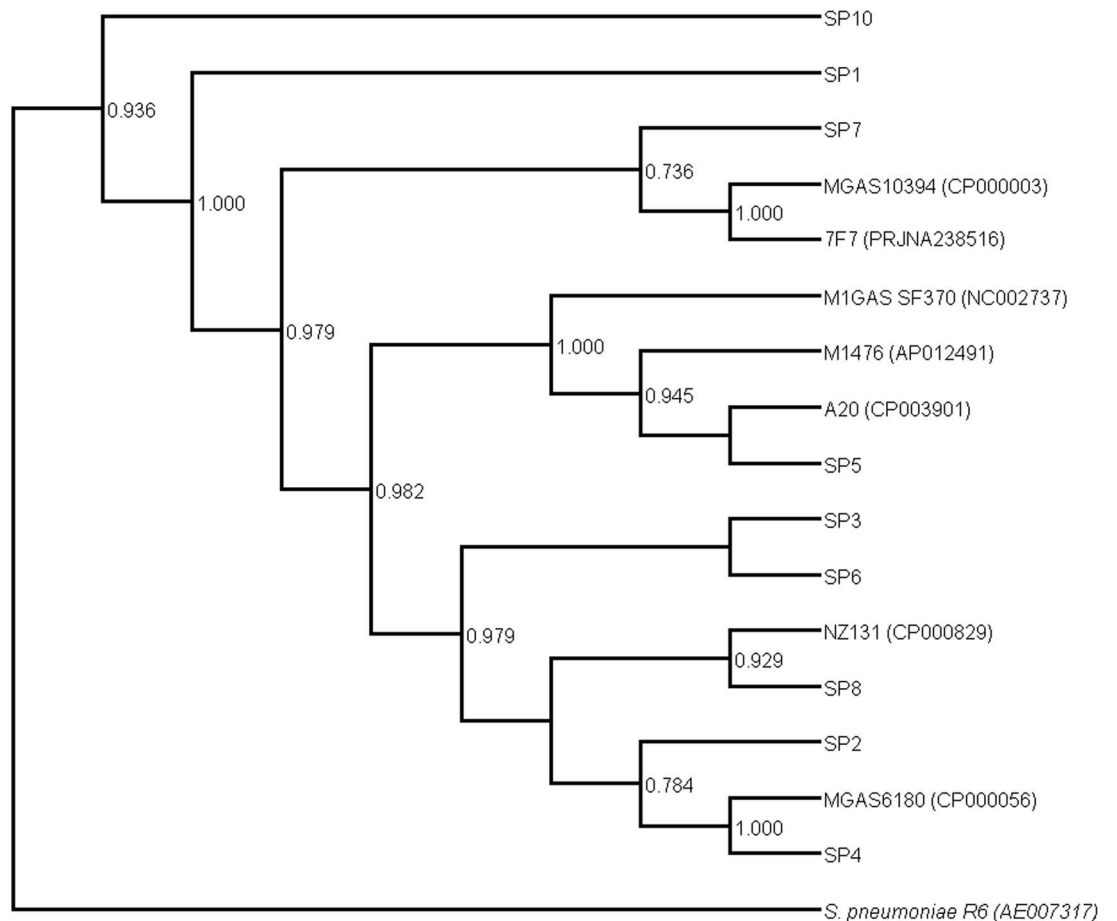
Fig 3. Phylogenetic tree of SP isolates and corresponding reference strains. The tree was constructed using 40 conserved coding marker genes of nine isolate and eight reference strain genomes. Three major clusters can be observed with bootstrap values on the nodes. The *S. pneumoniae* R6 strain is used as an outgroup for a more robust visualization.

doi:10.1371/journal.pone.0168177.g003
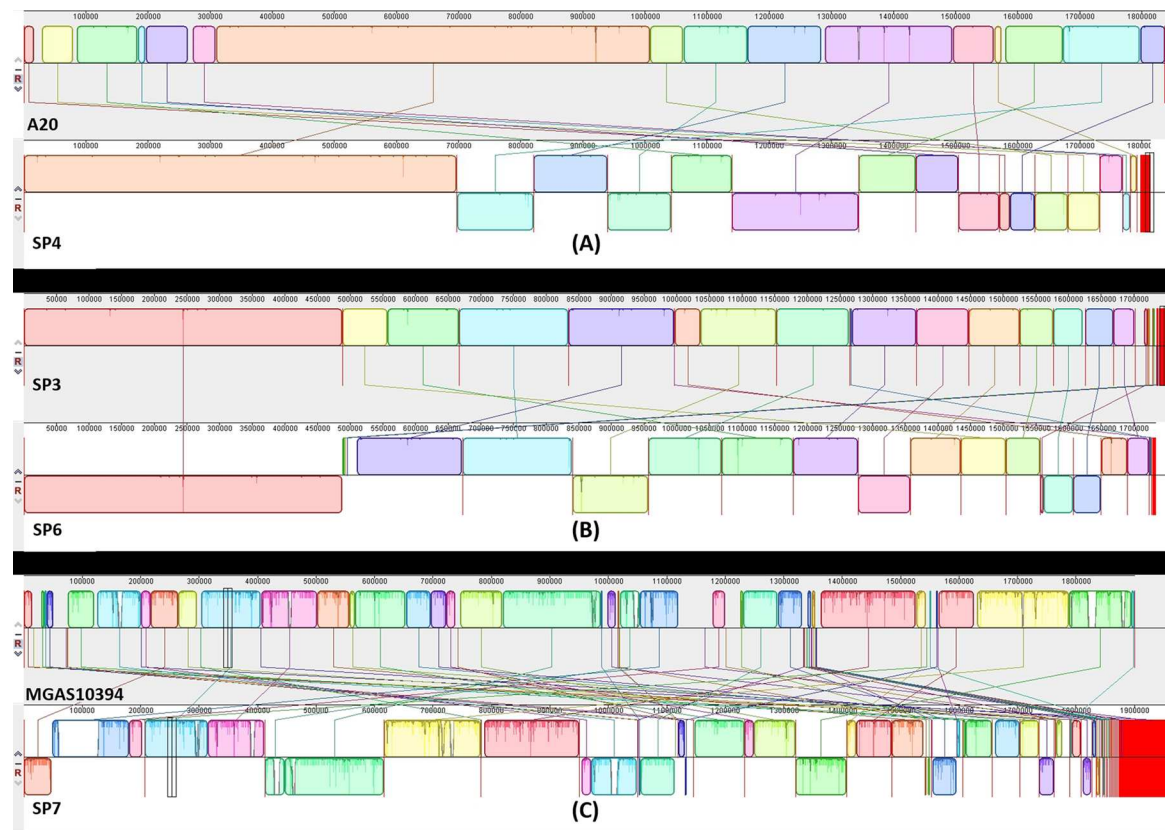
Fig 4. Genome comparison. Pairwise alignment of genomes from sub branches SP4, SP6, and SP7 with A20 (CP003901), SP3, and MGAS10394 (CP000003) respectively using the Mauve progressive alignment [25]. The colored blocks indicate homologous corresponding regions between the genomes that are internally free of rearrangement. Seismic lines inside blocks indicate the degree of similarity between alignments, while the red blocks indicate non-aligned sequences.

with clonally and serotypically different reference strains (Table 1, Figs 3 and 4C). Sequence data of the MGAS10394 genome has revealed the presence of eight prophage-like elements [122], a fact in agreement with our findings of six phage-related elements at least on the SP7 genome (Table 6). Additionally, SP7 was one of the few genomes in the studied isolates to have the SpeA superantigen (Table 4), another characteristic of MGAS10394 [122]. Moreover, MGAS10394 was originally isolated from the throat of a patient with pharyngitis, and determined to be macrolide resistant [123], sharing both with SP7 isolate (Table 5).

In the second cluster, SP5 was grouped with the reference strains A20, M1476 and M1GAS SF370 indicating that these genomes are closely related given the fact that they all belong to the same clonal complex and *emm* type (*emm1*). In this context, it is noteworthy that SP5 was clinically isolated from a skin infection, and that the covRS system and the cysteine protease B superantigen were detected in its genome (Table 4). These findings are in harmony with the original study characterizing the A20 strain [124], again asserting a high similarity between the two genomes. Pairwise alignment of the two genomes also showed a high degree of homology with large parts of the of the two genome being preserved and homologous (Fig 4A).

In the third cluster, SP2, SP3, SP4, SP6 and SP8 were grouped closely with NZ131 and MGAS6180 indicating a high similarity and also reflective of the preliminary BLAST results. Within this branch, three additional subgroups can be distinguished. Being both clonal and of the same *emm* type (*emm89*/ST-101), SP3 and SP6 grouped very closely together and aligned

with high sequence similarity (Figs 3 and 4B). Looking more closely at their genomes (Tables 4, 5 and 6), a very high similarity between genomic features can be seen, this sheds light on their close phylogenetic proximity. SP8 clustered with NZ131 whereas SP2 and SP4 grouped with the MGAS6180 strain with SP4 being much closer to the reference. Both SP4 and MGAS6180 are of the same *emm* and sequence types (*emm1* and ST-52) (Table 1). Interestingly, MGAS6180 is linked to puerperal sepsis with its virulence and pathogenesis attributed to of foreign gene uptake [94]. SP1 and SP10 did not group closely with any other isolate or reference strain (Fig 3), indicating considerable phylogenetic differences from their counterparts. Despite being associated with pharyngitis and skin infections, the distinct clustering of our isolates with different invasive strains, indicated potential invasiveness.

A recent study by Sanderson-Smith et al. [125], described a novel systematic and functional classification of *S. pyogenes* isolates into based on DNA sequencing of the full *emm* gene, and typed over 1000 GAS isolates in phylogenetic *emm*-clusters. Our phylogenetic analysis is in general agreement with the findings of this novel typing method; SP7 along with reference strains MGAS10394 and 7F7, clustered together and belonged to the D4 *emm*-cluster, SP3, 4, 6 grouped with MGAS6180 strain in the E4 *emm*-cluster, while SP5 clustered with the M1 serotype group of references M1GAS SF370, MGA1467, A20, collectively part of the A-C3 *emm*-cluster. SP1, phylogenetically clustering distantly from the rest, was also grouped in the A-C4 cluster *emm*-cluster phylogenetically further away from the previously mentioned clusters [125]. These findings further confirm that our isolates do not belong to a single phylogenetic cluster, but on the contrary, capture the range of serotypes that are implicated in throat and skin infections and that are present in the Lebanese community.

Conclusion

Despite modern advancements in medicine, the disease burden of *S. pyogenes* remains very real especially in less developed countries. Its plethora of virulence factors and superantigens and its resistance to antimicrobial agents can quickly turn superficial infections into life threatening ones. The heterogeneity of stress response elements present on the streptococcal genome allows the organism to dynamically resist harsh conditions both inside and outside the host. The additional multiple gene regulation systems and their strict control over virulence factors in response to environmental conditions outline infections as being a highly transitional event closely dependent on external host stimuli. The high genomic plasticity characteristic of the species adds to the complexity of the identification process.

This study generated the draft genomes sequences of nine *S. pyogenes* clinical isolates and investigated their genomic content, invasive potential, antibiotic resistance, and examined their epidemiological origin through comparative phylogenetic methods. Regulation, signaling, and stress response elements were successfully identified in all of the isolates. Their overall virulence related elements being directly associated with throat and skin infections, without excluding their potential invasiveness. The hypervariable FCT region was found to be of Type-3 in all of the isolates excluding SP8, which exhibited an FCT-1 organization. The isolates showed large variation in the M protein family as well as in the FCT regulon. Seven distinct superantigen profiles were detected, showing common virulence traits between our isolates and those associated with pharyngitis and skin infections. Additionally, macrolide, fluoroquinolone, and rare resistance to tetracyclines were documented. Bacteriophages were detected in all of the isolates conferring superantigens as well as genes generally associated with invasive streptococcal diseases and resistance, a fact further confirmed by the phylogenetic study that linked our isolates to worldwide invasive-type clones. The phylogenetic study, in addition to the virulence profiles, showed a clear association between genome, site of isolation, and disease

type. Moreover, we concluded that these strains did not form a single phylogenetic cluster, but instead captured the genomic diversity found in the reference strains.

In all, the detailed genome analysis will undoubtedly provide new insights on *S. pyogenes*. The expected merits of this study are not merely theoretical, as GAS remains an important cause of diseases in Lebanon and in the world, studying its functional genomics could help in better understanding the molecular mechanisms and epidemiology of its pathogenesis. Our findings add value to epidemiological studies of *S. pyogenes*, and provide the first such study for Lebanese isolates.

Author Contributions

Conceptualization: JI ST JE GJ DC.

Data curation: JI ST JE GJ GK DC.

Formal analysis: JI ST JE GJ GK DC.

Funding acquisition: ST JE DC.

Investigation: JI ST.

Methodology: JI ST JE GJ DC.

Project administration: ST DC.

Resources: ST JE DC.

Software: JI ST GJ GK DC.

Supervision: ST.

Validation: JI ST DC.

Visualization: JI ST JE GJ GK DC.

Writing – original draft: JI ST.

Writing – review & editing: JI ST JE GJ GK DC.

References

1. Cunningham MW. Pathogenesis of Group A Streptococcal Infections. Clin Microbiol Rev. 2000; 13(3), 470–511. PMID: 10885988

2. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, et al. Disease Manifestations and Pathogenic Mechanisms of Group A Streptococcus. Clin Microbiol Rev. 2014; 27(2): 264–301. doi: 10.1128/CMR.00101-13 PMID: 24696436

3. Twisselmann B. Epidemiology, treatment, and control of infection with Streptococcus pyogenes in Germany. Euro Surveill. 2000; 4(46): piid = 1490.

4. Albrich CW, Monnet LD, Harbarth S. Antibiotic Selection Pressure and Resistance in Streptococcus pneumoniae and Streptococcus pyogenes. Emerg Infect Diseases. 2004; 10(3):514–517.

5. Davies RM, Holden TM, Coupland P, Chen JHK, Venturini C, Barnett TC, et al. Emergence of scarlet fever Streptococcus pyogenes emm12 clones in Hong Kong is associated with toxin acquisition and multidrug resistance. Nature Genet. 2015; 47: 84–87. doi: 10.1038/ng.3147 PMID: 25401300

6. Pichichero ME, Casey JR. Systematic review of factors contributing to penicillin treatment failure in Streptococcus pyogenes pharyngitis. Otolaryngol Head Neck Surg. 2007; 137, 851–857. doi: 10.1016/j.otohns.2007.07.033 PMID: 18036409

7. Wong S, Yuen KY. Streptococcus pyogenes and reemergence of scarlet fever as a public health concern. Emerg Microbes Infect. 2012; 1(2): e2.

8. Vojtek I, Pirzada ZA, Henriques-Normark B, Mastny M, Janapatla RP, Charpentier E. Lysogenic transfer of group A Streptococcus superantigen gene among streptococci. J Infect Dis. 2008; 197(2): 225–234. doi: 10.1086/524687 PMID: 18179387

9. Luca-Harari B, Darenberg J, Neal S, Siljander T, Strakova L, Tanna A, et al. Clinical and microbiological characteristics of severe Streptococcus pyogenes disease in Europe. J Clin Microbiol. 2009; 47: 1155–1165. doi: 10.1128/JCM.02155-08 PMID: 19158266

10. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: Systematic review and implications for vaccine development. Lancet Infect Dis. 2009; 9(10): 611–616. doi: 10.1016/S1473-3099(09)70178-1 PMID: 19778763

11. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. Lancet Infect Dis. 2005; 5(11): 685–694. doi: 10.1016/S1473-3099(05)70267-X PMID: 16253886

12. Bahnan W, Hashwa F, Araj G, Tokajian S. emm typing, antibiotic resistance and PFGE analysis of Streptococcus pyogenes in Lebanon. J Med Microbiol. 2011; 60(1): 98–101.

13. Karaky NM, Araj GF, Tokajian ST. Molecular characterization of Streptococcus pyogenes group A isolates from a tertiary hospital in Lebanon. J Med Microbiol. 2014; 63(9): 1197–1204.

14. McGregor KF, Spratt BG, Kalia A, Bennett A, Bilek N, Beall B, Bessen DE. Multilocus sequence typing of Streptococcus pyogenes representing most known emm types and distinctions among subpopulation genetic structures. J Bacteriol, 2004; 186(13), 4285–4294. doi: 10.1128/JB.186.13.4285-4294.2004 PMID: 15205431

15. Enright MC, Spratt BG, Kalia A, Cross JH, Bessen DE. Multilocus sequence typing of Streptococcus pyogenes and the relationships between emm type and clone. Infect Immun, 2001; 69(4), 2416–2427. doi: 10.1128/IAI.69.4.2416-2427.2001 PMID: 11254602

16. Tritt A, Eisen J, Facciotti M, Darling A. An Integrated Pipeline for de Novo Assembly of Microbial Genomes. PLoS ONE. 2012; 7(9): e42304. doi: 10.1371/journal.pone.0042304 PMID: 23028432

17. Coil D, Jospin G, Darling A. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. Bioinfo. 2015; 31: 587–589.

18. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: Rapid annotations using subsystems technology. BMC Genomics. 2008; 9: 75. doi: 10.1186/1471-2164-9-75 PMID: 18261238

19. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014; 42, 206–214.

20. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. Antimicrob Agents Chemother. 2012; 67(11): 2640–2644.

21. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol. 2014; 52(5): 1501–1510. doi: 10.1128/JCM.03617-13 PMID: 24574290

22. Cosentino S, Larsen M, Aarestrup FM, Lund O. PathogenFinder-distinguishing friend from foe using bacterial whole genome sequence data. PLoS ONE. 2013; 8(10): e77302. doi: 10.1371/journal.pone.0077302 PMID: 24204795

23. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total genome sequenced bacteria. J Clin Microbiol. 2012; 50(4): 1355–1361. doi: 10.1128/JCM.06094-11 PMID: 22238442

24. Zhou Y, Liang Y, Lynch HK, Dennis JJ, Wishart DS. PHAST: A fast phage search tool. Nucleic Acids Res. 2011; 39(2): 347–352.

25. Darling AE, Jospin G, Lowe E, Matsen FA IV, Bik HM, Eisen JA. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ, 2014; 2:e243. doi: 10.7717/peerj.243 PMID: 24482762

26. Price MN, Dehal PS, Arkin AP. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE, 2010; 5(3):e9490. doi: 10.1371/journal.pone.0009490 PMID: 20224823

27. Huson DH, Scornavacca C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. Syst Biol, 2012; 0(0):1–7.

28. Darling AE, Mau B, Perna NT. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE, 2010; 5(6): e11147. doi: 10.1371/journal.pone.0011147 PMID: 20593022

29. Tokajian S, Eisen JA, Jospin G, Coil DA. Draft Genome Sequences of Streptococcus pyogenes Strains Associated with Throat and Skin Infections in Lebanon. Genome Announc. 2014; 2(3): e00358–14. doi: 10.1128/genomeA.00358-14 PMID: 24831139

30. National Center for Biotechnology Information. Genome. Available from: http://www.ncbi.nlm.nih.gov/genome/ [Accessed 7th March 2015].

31. Ferretti JJ, McShan WM, Ajdic D, Savic DJ. Savic G, Lyon K, et al. Complete genome sequence of an M1 strain of Streptococcus pyogenes. PNAS. 2001; 98(8): 4658–4663. doi: 10.1073/pnas.071559398 PMID: 11296296

32. Kunst F, Ogasawara N, Moszer I, Albertini A, Alloni G, Azevedo V, et al. The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature. 1997; 390(6657): 249–256. doi: 10.1038/36786 PMID: 9384377

33. Feklístov A, Sharon BD, Darst SA, Gross CA. Bacterial sigma factors: A historical, structural, and genomic perspective. Annu Rev Microbiol. 2014; 68: 357–376. doi: 10.1146/annurev-micro-092412-155737 PMID: 25002089

34. Luo P, Morrison DA. Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in Streptococcus pneumoniae. J Bacteriol. 2003; 185(1): 349–358. doi: 10.1128/JB.185.1.349-358.2003 PMID: 12486073

35. Opdyke JA, Scott JR, Moran CP. A secondary RNA polymerase sigma factor from Streptococcus pyogenes. Mol. Microbiol. 2001; 42(2), 495–502. PMID: 11703670

36. Las Peñas AD, Connolly L, Gross CA. Sigma E is an essential sigma factor in Escherichia coli. J Bacteriol. 1997; 179(21): 6862–6864. PMID: 9352942

37. Wargo MJ, Szwergold BS, Hogan DA. Identification of two gene clusters and a transcriptional regulator required for Pseudomonas aeruginosa glycine betaine catabolism. J Bacteriol. 2008; 190(8): 2690–2699. doi: 10.1128/JB.01393-07 PMID: 17951379

38. Suzuki T, Tagami J, Hanada H. Role of F1F0-ATPase in the growth of Streptococcus mutans GS5. Appl Microbiol Biotechnol. 2000; 88 (4): 555–562.

39. Phadtare S. Recent developments in bacterial cold-shock response. Curr. Issues Mol. Biol. 2004; 6 (2), 125–136. PMID: 15119823

40. Homuth G, Masuda S, Mogk A, Kobayashi Y, Schumann W. The dnaK operon of Bacillus subtilis is heptacistronic. J Bacteriol. 1997; 179(4): 1153–1164. PMID: 9023197

41. Takaya A, Tomoyasu T, Matsui H, Yamamoto T. (2004). The DnaK/DnaJ chaperone machinery of Salmonella enterica serovar Typhimurium is essential for invasion of epithelial cells and survival within macrophages, leading to systemic infection. Infect Immun. 2004; 72(3): 1364–1373. doi: 10.1128/IAI.72.3.1364-1373.2004 PMID: 14977940

42. Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog. 2006; 2(1): e5. doi: 10.1371/journal.ppat.0020005 PMID: 16446783

43. Musser JM, Shelburne SA. A decade of molecular pathogenomic analysis of group A Streptococcus. J. Clin. Invest. 2009; 119(9), 2455–2463. doi: 10.1172/JCI38095 PMID: 19729843

44. Dalton TL, Scott JR. CovS inactivates CovR and is required for growth under conditions of general stress in Streptococcus pyogenes. J Bacteriol. 2004; 186(12): 3928–3937. doi: 10.1128/JB.186.12.3928-3937.2004 PMID: 15175307

45. Trevino J, Perez N, Ramirez-Pena E, Liu Z, Shelburne SA, Musser JM, et al. CovC simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A Streptococcus virulence factor-encoding genes. Infect Immun. 2009; 77(8): 3141–3149. doi: 10.1128/IAI.01560-08 PMID: 19451242

46. Walker MJ, Hollands A, Sanderson-Smith ML, Cole JN, Kirk JK, Henningham A, et al. DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat Med. 2007; 13 (8): 981–985. doi: 10.1038/nm1612 PMID: 17632528

47. Ribardo DA, McIver KS. Defining the Mga regulon: Comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the group A Streptococcus. Mol Microbiol. 2006; 62(2): 491–508. doi: 10.1111/j.1365-2958.2006.05381.x PMID: 16965517

48. Hondorp ER, Hou SC, Hause LL, Gera K, Lee CE, McIver CS. PTS phosphorylation of Mga modulates regulon expression and virulence in the group A Streptococcus. Mol Microbiol. 2013; 88(6): 1176–1193. doi: 10.1111/mmi.12250 PMID: 23651410

49. Crooke E, Thresher R, Hwang DS, Griffith J, Kornberg A. Replicatively active complexes of DnaA protein and the Escherichia coli chromosomal origin observed in the electron microscope. J. Mol. Biol. 1993; 233, 16–24. doi: 10.1006/jmbi.1993.1481 PMID: 8377183

50. Gao F, Zhang CT. Ori-Finder: A web-based system for finding oriC s in unannotated bacterial genomes. BMC Bioinformatics, 2008; 9, 79. doi: 10.1186/1471-2105-9-79 PMID: 18237442

51. Gao F, Luo H, Zhang CT. DoriC 5.0: an updated database of oriC regions in both bacterial and archaeal genomes. Nucleic Acids Res. 2013; 41, 90–93.

52. Christensen BB, Atlung T, Hansen FG. DnaA boxes are important elements in setting the initiation mass of Escherichia coli. J Bacteriol. 1999; 181(9), 2683–2688. PMID: 10217754

53. Robinson NP, Bell SD. Origins of DNA replication in the three domains of life. FEBS J. 2005; 272(15), 3757–3766. doi: 10.1111/j.1742-4658.2005.04768.x PMID: 16045748

54. Richter S, Hess WR, Krause M, Messer W. Unique organization of the dnaA region from Prochlorococcus marinus CCMP1375, a marine cyanobacterium. Mol. Gen. Genet. 1998; 257(5), 534–541. PMID: 9563839

55. Inagaki Y, Myouga F, Kawabata H, Yamai S, Watanabe H. (2000). Genomic differences in Streptococcus pyogenes serotype M3 between recent isolates associated with toxic shock-like syndrome and past clinical isolates. J Infect Dis. 2000; 181(3): 975–983. doi: 10.1086/315299 PMID: 10720520

56. Gase K, Ferretti FF, Primeaux C, McShan WM. Identification, cloning, and expression of the CAMP factor gene (cfa) of group A streptococci. Infect Immun. 1999; 67(9): 4725–4731. PMID: 10456923

57. Podbielski A, Blankenstein O, Lütticken R. Molecular characterization of the cfb gene encoding group B streptococcal CAMP-factor. Med. Microbiol. Immun. 1994; 183(5), 239–256.

58. Timmer AM, Timmer JC, Pence MA, Hsu LC, Ghochani M, Frey TG, et al. Streptolysin O promotes group A Streptococcus immune evasion by accelerated macrophage apoptosis. J Biol Chem. 2009; 284(2): 862–871. doi: 10.1074/jbc.M804632200 PMID: 19001420

59. Nizet V, Beall B, Bast DJ, Datta V, Kilburn L, Low DE, et al. Genetic locus for streptolysin S production by group A Streptococcus. Infect. Immun. 2000; 68(7), 4245–4254. PMID: 10858242

60. Miyoshi-Akiyama T, Takamatsu D, Koyanagi M, Zhao J, Imanishi K, Uchiyama T. Cytocidal effect of Streptococcus pyogenes on mouse neutrophils in vivo and the critical role of streptolysin S. J. Infect. Dis. 2005; 192(1), 107–116. doi: 10.1086/430617 PMID: 15942900

61. Ofek I, Bergner-Rabinowitz S, Ginsburg I. Oxygen-stable hemolysins of group A Streptococci VIII. Leukotoxic and antiphagocytic effects of Sstreptolysins S and O. Infect. Immun. 1972; 6(4), 459–464. PMID: 4634952

62. Beres SB, Musser JM. Contribution of exogenous genetic elements to the group A Streptococcus metagenome. PLoS ONE. 2007; 2(8): e800. doi: 10.1371/journal.pone.0000800 PMID: 17726530

63. Akesson P, Sjoholm AG, Bjorck L. Protein SIC, a novel extracellular protein of Streptococcus pyogenes interfering with complement function. J Biol Chem. 1996; 271(2): 1081–1088. PMID: 8557634

64. Hoe NP, Vuopio-Varkila J, Vaara M, Grigsby D, De Lorenzo D, Fu YX, et al. Distribution of streptococcal inhibitor of complement variants in pharyngitis and invasive isolates in an epidemic of serotype M1 group A Streptococcus infection. J Infect Dis. 2001; 183(4): 633–639. doi: 10.1086/318543 PMID: 11170990

65. Plummer TH, Tarentino AL. Purification of the oligosaccharide-cleaving enzymes of Flavobacterium meningosepticum. Glycobiology. 1991; 1(3), 257–263. PMID: 1794038

66. Sjogren J, Struwe WB, Cosgrave E, Rudd PM, Stervander M, Allhorn M, et al. EndoS2 is a unique and conserved enzyme of serotype M49 group A Streptococcus that hydrolyzes N-linked glycans on IgG and alpha1-acid glycoprotein. Biochem J. 2013; 455(1): 107–118. doi: 10.1042/BJ20130126 PMID: 23865566

67. Sjogren J, Okumura CY, Collin M, Nizet, Hollands A. Study of the IgG endoglycosidase EndoS in group A streptococcal phagocyte resistance and virulence. BMC Microbiology. 2011; 11(120).

68. Aziz RK, Ismail SA, Park HW, Kotb M. Post-proteomic identification of a novel phage-encoded streptodornase, Sda1, in invasive M1T1 Streptococcus pyogenes. Mol Microbiol. 2004; 54(1): 184–97. doi: 10.1111/j.1365-2958.2004.04255.x PMID: 15458415

69. Agniswamy J, Lei B, Musser JM, Sun PD. Insight of host immune evasion mediated by two variants of group A Streptococcus Mac protein. J Biol Chem. 2004; 279(50): 52789–52796. doi: 10.1074/jbc.M410698200 PMID: 15466462

70. Lei B, DeLeo FR, Hoe NP, Graham MR, Mackie SM, Cole RL, et al. Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. Nat Med. 2001; 7(12): 1298–1305. doi: 10.1038/nm1201-1298 PMID: 11726969

71. Okumura CY, Anderson EL, Dohrmann S, Tran DN, Olson J, von Pawel-Rammingen U, et al. IgG protease Mac/IdeS is not essential for phagocyte resistance or mouse virulence of M1T1 group A Streptococcus. mBio, 2013; 4(4), e00499–13. doi: 10.1128/mBio.00499-13 PMID: 23900173

72. Nakagawa I, Kurokawa K, Yamashita A, Nakata M, Tomiyasu Y, Okahashi N, et al. 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 Research. 2003; 13(6A), 1042–1055. doi: 10.1101/gr.1096703 PMID: 12799345

73. Bessen DE, Kalia A. Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in Streptococcus pyogenes. Infect Immun. 2002; 70(3): 1159–1167. doi: 10.1128/IAI.70.3.1159-1167.2002 PMID: 11854196

74. Kratovac Z, Manoharan A, Luo F, Lizano S, Bessen DE. Population genetics and linkage analysis of loci within the FCT region of Streptococcus pyogenes. J Bacteriol. 2007; 189(4): 1299–1310. doi: 10.1128/JB.01301-06 PMID: 17028269

75. Horstmann RD, Sievertsen HJ, Knobloch J, Fischetti VA. Antiphagocytic activity of streptococcal M protein: Selective binding of complement control protein factor H. PNAS. 1988; 85(5): 1657–1661. PMID: 2964038

76. Schulze K, Olive C, Ebensen T, Guzman CA. Intranasal vaccination with SfbI or M protein-derived peptides conjugated to diphtheria toxoid confers protective immunity against a lethal challenge with Streptococcus pyogenes. Vaccine. 2006; 24(35–36): 6088–6095. doi: 10.1016/j.vaccine.2006.05.024 PMID: 16828529

77. Lizano S, Luo F, Bessen DE. Role of Streptococcal T Antigens in Superficial Skin Infection. J Bacteriol. 2007; 189(4): 1426–1434. doi: 10.1128/JB.01179-06 PMID: 17012387

78. Wang H, Lottenberg R, Boyle MD. Analysis of the interaction of group A streptococci with fibrinogen, streptokinase and plasminogen. Microb Pathog. 1995; 18(3): 153–166. PMID: 7565010

79. Chen CC, Cleary PP. Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus pyogenes. J Biol Chem. 1990; 265(6): 3161–3167. PMID: 2406246

80. Hondorp ER, McIver KS. The Mga virulence regulon: Infection where the grass is greener. Mol Microbiol. 2007; 66(5): 1056–65. doi: 10.1111/j.1365-2958.2007.06006.x PMID: 18001346

81. McIver KS, Heath AS, Green BD, Scott JR. Specific binding of the activator Mga to promoter sequences of the emm and scpA genes in the group A streptococcus. J Bacteriol. 1995; 177, 6619–6624. PMID: 7592441

82. Hollingshead SK, Readdy TL, Yung DL, Bessen DE. Structural heterogeneity of the emm gene cluster in group A streptococci. Mol Microbiol. 1993; 8(4): 707–717. PMID: 8332063

83. Rantala S, Vahakuopus S, Siljander T, Vuopio J, Huhtala H, Vuento R, et al. Streptococcus pyogenes bacteraemia, emm types and superantigen profiles. Eur J Clin Microbiol Infect Dis. 2012; 31(5), 859–865. doi: 10.1007/s10096-011-1385-9 PMID: 21877175

84. Spaulding AR, Salgado-Pabón W, Kohler PL, Horswill AR, Leung DY, Schlievert PM. Staphylococcal and streptococcal superantigen exotoxins. Clin Microbiol Rev. 2013; 26(3): 422–247. doi: 10.1128/CMR.00104-12 PMID: 23824366

85. Stevens DL, Tanner MH, Winship J, Swarts R, Ries KM, Schlievert PM, Kaplan E. Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. N Engl J Med. 1989; 321(1): 1–7. doi: 10.1056/NEJM198907063210101 PMID: 2659990

86. Commons RJ, Smeesters PR, Proft T, Fraser JD, Robins-Browne R, Curtis N. Streptococcal superantigens: Categorization and clinical associations. Trends Mol Med. 2014; 20(1): 48–62. doi: 10.1016/j.molmed.2013.10.004 PMID: 24210845

87. Bisno AL. The concept of rheumatogenic and non-rheumatogenic group A streptococci. In: Read SE, Zabriskie JB, editors. Streptococcal diseases and the immune response. New York, N.Y: Academic Press Inc.; 1980. pp. 789–803.

88. Stollerman GH. Rheumatic fever. Lancet. 1997; 349(9056): 935–942. doi: 10.1016/S0140-6736(96)06364-7 PMID: 9093263

89. Proft T, Moffatt SL, Weller KD, Paterson A, Martin D, Fraser JD. The streptococcal superantigen SMEZ exhibits wide allelic variation, mosaic structure, and significant antigenic variation. J Exp Med. 2000; 191(10): 1765–1776. PMID: 10811869

90. Commons RJ, Rogers S, Gooding T, Danchin M, Carapetis J, Robins-Browne R. Superantigen genes in group A streptococcal isolates and their relationship with emm types. J Med Microbiol. 2008; 57(10): 1238–1246.

91. McCormick JK, Pragman AA, Stolpa JC, Leung DY, Schlievert PM. Functional Characterization of Streptococcal Pyrogenic Exotoxin J, a Novel Superantigen. Infect Immun. 2001; 69(3): 1381–1388. doi: 10.1128/IAI.69.3.1381-1388.2001 PMID: 11179302

92. Holden MTG, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, et al. Complete genome of acute rheumatic fever-associated serotype M5 Streptococcus pyogenes strain Manfredo. J. Bacteriol. 2007; 189(4): 1473–1477. doi: 10.1128/JB.01227-06 PMID: 17012393

93. Fittipaldi N, Beres SB, Olsen RJ, Kapur V, Shea PR, Watkins ME, et al. Full-genome dissection of an epidemic of severe invasive disease caused by a hypervirulent, recently emerged clone of group A Streptococcus. Am J Pathol. 2012; 180(4): 1522–1534. doi: 10.1016/j.ajpath.2011.12.037 PMID: 22330677

94.   Green NM, Zhang S, Porcella SF, Nagiec MJ, Barbian KD, Beres SB, et al. Genome sequence of a serotype M28 strain of group A Streptococcus: Potential new insights into puerperal sepsis and bacterial disease specificity. Int J Infect Dis. 2005; 192(5): 760–70.

95.   Ikebe T, Wada A, Inagaki Y, Sugama K, Suzuki R, Tanaka D, et al. 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(6): 3227–3233. doi: 10.1128/IAI.70.6.3227-3233.2002 PMID: 12011018

96.   Luca-Harari B, Ekelund K, van der Linden M, Staum-Kaltoft M, Hammerum AM, Jasir A. Clinical and epidemiological aspects of invasive Streptococcus pyogenes infections in Denmark during 2003 and 2004. J Med Microbiol. 2008; 46(1): 79–86.

97.   Chan JC, Chu YW, Chu MY, Cheung TK, Lo JY. Epidemiological analysis of Streptococcus pyogenes infections in Hong Kong. Pathology. 2009; 41(7): 681–686. doi: 10.3109/00313020903257723 PMID: 20001348

98.   Beres SB, Richter EW, Nagiec MJ, Sumby P, Porcella SF, DeLeo FR, et al. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A Streptococcus. Proc Natl Acad Sci U S A. 2006; 103(18): 7059–64. doi: 10.1073/pnas.0510279103 PMID: 16636287

99.   Imohl M, Reinert RR, Ocklenburg C, van der Linden M. Epidemiology of invasive Streptococcus pyogenes disease in Germany during 2003–2007. Immunol Med Mic. 2010; 58(3): 389–396.

100. Seppala H, Skurnik M, Soini H, Roberts MC, Huovinen P. A novel erythromycin resistance methylase gene (ermTR) in Streptococcus pyogenes. Antimicrob Agents Chemother. 1998; 42(2): 257–262. PMID: 9527769

101. Farmand S, Henneke P, Hufnagel M, Berner R. Significant decline in the erythromycin resistance of group A Streptococcus isolates at a German paediatric tertiary care centre. Eur J Clin Microbiol. 2012; 31(5): 707–710.

102. Van Heirstraeten L, Leten G, Lammens C, Goossens H, Malhotra-Kumar S. Increase in fluoroquinolone non-susceptibility among clinical Streptococcus pyogenes in Belgium during 2007–10. J Antimicrob Chemother. 2012; 67(11): 2602–2605. doi: 10.1093/jac/dks281 PMID: 22815354

103. Ardanuy C, Domenech A, Rolo D, Calatayud L, Tubau F, Ayats J, et al. Molecular characterization of macrolide- and multidrug-resistant Streptococcus pyogenes isolated from adult patients in Barcelona, Spain (1993–2008). J Antimicrob Chemother. 2010; 65(4): 634–643. doi: 10.1093/jac/dkq006 PMID: 20118164

104. Wajima T, Murayama SY, Sunaoshi K, Nakayama E, Sunakawa K, Ubukata K. Distribution of emm type and antibiotic susceptibility of group A streptococci causing invasive and noninvasive disease. J Med Microbiol. 2008; 57(11): 1383–1388.

105. Yan SS, Schreckenberger PC, Zheng X, Nelson NA, Harrington SM, Tjhio J, et al. (2008). An intrinsic pattern of reduced susceptibility to fluoroquinolones in pediatric isolates of Streptococcus pyogenes. Diagn Microbiol Infect Dis. 2008; 62(2): 205–209. doi: 10.1016/j.diagmicrobio.2008.04.018 PMID: 18554840

106. Montes M, Tamayo E, Orden B, Larruskain J, Perez-Trallero E. Prevalence and clonal characterization of Streptococcus pyogenes clinical isolates with reduced fluoroquinolone susceptibility in Spain. J. Antimicrob. Chemother. 2010; 54(1), 93–97.

107. Perez-Trallero E, Tamayo E, Montes M, Garcia-Arenzana JM, Iriart V. In vitro activities of retapamulin and 16 other antimicrobial agents against recently obtained Streptococcus pyogenes isolates. Antimicrob. Agents Chemother. 2011; 55(5), 2406–2408. doi: 10.1128/AAC.01665-10 PMID: 21343452

108. Chopra I, Roberts M. Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev. 2001; 65(2): 232–260. doi: 10.1128/MMBR.65.2.232-260.2001 PMID: 11381101

109. Li W, Atkinson GC, Thakor NS, Allas U, Lu CC, Chan KY, et al. Mechanism of tetracycline resistance by ribosomal protection protein Tet(O). Nat Commun. 2013; 4: 1477. doi: 10.1038/ncomms2470 PMID: 23403578

110. Mihaila-Amrouche L, Bouvet A, Loubinoux J. Clonal spread of emm type 28 isolates of Streptococcus pyogenes that are multiresistant to antibiotics. J Clin Microbiol. 2004; 42(8), 3844–3846. doi: 10.1128/JCM.42.8.3844-3846.2004 PMID: 15297545

111. Malhotra-Kumar S, Lammens C, Piessens J, Goossens H. Multiplex PCR for simultaneous detection of macrolide and tetracycline resistance determinants in streptococci. Antimicrob Agents Chemother. 2005; 49(11): 4798–4800. doi: 10.1128/AAC.49.11.4798-4800.2005 PMID: 16251336

112. McAleese F, Petersen P, Ruzin A, Dunman PM, Murphy E, Projan SJ, et al. A novel MATE family efflux pump contributes to the reduced susceptibility of laboratory-derived Staphylococcus aureus mutants to tigecycline. Antimicrob Agents Chemother. 2005; 49(5): 1865–1871. doi: 10.1128/AAC.49.5.1865-1871.2005 PMID: 15855508

113.   Brenciani A, Bacciaglia A, Vignaroli C, Pugnaloni A, Varaldo PE, Giovanetti E. Phim46.1, the main Streptococcus pyogenes element carrying mef(A) and tet(O) genes. Antimicrob Agents Chemother. 2010; 54(1): 221–229. doi: 10.1128/AAC.00499-09 PMID: 19858262

114.   Giovanetti E, Brencian A, Tiberi E, Bacciaglia A, Varaldo PE. ICESp2905, the erm(TR)-tet(O) element of Streptococcus pyogenes, is formed by two independent integrative and conjugative elements. Antimicrob Agents Chemother. 2012; 56(1): 591–594. doi: 10.1128/AAC.05352-11 PMID: 21986826

115.   Canchaya C, Fournous G, Brussow H. The impact of prophages on bacterial chromosomes. Mol Microbiol. 2004; 53(1): 9–18. doi: 10.1111/j.1365-2958.2004.04113.x PMID: 15225299

116.   Bisno AL, Brito MO, Collins CM. Molecular basis of group A streptococcal virulence. Lancet Infect Dis. 2003; 3(4): 191–200. PMID: 12679262

117.   Musser JM, DeLeo FR. Toward a genome-wide systems biology analysis of host-pathogen interactions in group A Streptococcus. Am. J. Pathol. 2005; 167(6), 1461–1472. doi: 10.1016/S0002-9440 (10)61232-1 PMID: 16314461

118.   Fischetti VA. In vivo acquisition of prophage in Streptococcus pyogenes. Trends in Microbiol. 2007; 15 (7): 297–300.

119.   Lewis T, Loman NJ, Bingle L, Jumaa P, Weinstock GM. Mortiboy D, et al. (2010). High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak. J Hosp Infect, 2010; 75(1), 37–41. doi: 10.1016/j.jhin.2010.01.012 PMID: 20299126

120.   Ben Zakour NL, Venturini C, Beatson SA, Walker JM. Analysis of a Streptococcus pyogenes puerperal sepsis cluster by use of whole-genome sequencing. J Clin Microbiol. 2012; 50(7), 2224–2228. doi: 10.1128/JCM.00675-12 PMID: 22518858

121.   Wu D, Jospin G, Eisen JA. Systematic identification of gene families for use as markers for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. PLoS ONE. 2013; 8(10):e77033 doi: 10.1371/journal.pone.0077033 PMID: 24146954

122.   Banks DJ, Porcella SF, Barbian KD, Beres SB, Philips LE, Voyich JM, et al. Progress toward characterization of the group A Streptococcus metagenome: Complete genome sequence of a macrolide-resistant serotype M6 strain. J Infect Dis. 2004; 190(4), 727–738. doi: 10.1086/422697 PMID: 15272401

123.   Banks DJ, Porcella SF, Barbian KD, Martin JM, Musser JM. Structure and distribution of an unusual chimeric genetic element encoding macrolide resistance in phylogenetically diverse clones of group A Streptococcus. J Infect Dis, 2003; 188(12), 1898–1908. doi: 10.1086/379897 PMID: 14673771

124.   Zheng PX, Chung KT, Chiang-Ni C, Wang SY, Tsai PJ, Chuang WJ, et al. Complete Genome Sequence of emm1 Streptococcus pyogenes A20, a Strain with an Intact Two-Component System, CovRS, Isolated from a Patient with Necrotizing Fasciitis. Genome Announc. 2013; 1(1): e00149–12. doi: 10.1128/genomeA.00149-12 PMID: 23405303

125.   Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, et al. A Systematic and Functional Classification of Streptococcus pyogenes That Serves as a New Tool for Molecular Typing and Vaccine Development. J Infect Dis. 2014; 210(8): 1325–1338. doi: 10.1093/infdis/jiu260 PMID: 24799598