



Genomic Surveillance of *Streptococcus pyogenes* Strains Causing Invasive Disease, United States, 2016–2017

Yuan Li^{1*}, Joy Rivers¹, Sandra Mathis¹, Zhongya Li¹, Srinivasan Velusamy¹, Srinivas A. Nanduri¹, Chris A. Van Beneden¹, Paula Snippes-Vagnone², Ruth Lynfield², Lesley McGee¹, Sopio Chochua¹, Benjamin J. Metcalf¹ and Bernard Beall¹

¹ Respiratory Diseases Branch, Division of Bacterial Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, United States Department of Health and Human Services, Atlanta, GA, United States, ² Minnesota Department of Health, St Paul, MN, United States

OPEN ACCESS

Edited by:

Dieter Jahn,
Technische Universität Braunschweig,
Germany

Reviewed by:

D. Patric Nitsche-Schmitz,
Waldorf School Cologne, Germany
Phil Giffard,
Charles Darwin University, Australia

*Correspondence:

Yuan Li
yqh8@cdc.gov

Specialty section:

This article was submitted to
Infectious Diseases,
a section of the journal
Frontiers in Microbiology

Received: 16 March 2020

Accepted: 16 June 2020

Published: 24 July 2020

Citation:

Li Y, Rivers J, Mathis S, Li Z,
Velusamy S, Nanduri SA,
Van Beneden CA,
Snippes-Vagnone P, Lynfield R,
McGee L, Chochua S, Metcalf BJ and
Beall B (2020) Genomic Surveillance
of *Streptococcus pyogenes* Strains
Causing Invasive Disease,
United States, 2016–2017.
Front. Microbiol. 11:1547.
doi: 10.3389/fmicb.2020.01547

Background: *Streptococcus pyogenes* is a major cause of severe, invasive infections in humans. The bacterial pathogen harbors a wide array of virulence factors and exhibits high genomic diversity. Rapid changes of circulating strains in a community are common. Understanding the current prevalence and dynamics of *S. pyogenes* lineages could inform vaccine development and disease control strategies.

Methods: We used whole-genome sequencing (WGS) to characterize all invasive *S. pyogenes* isolates obtained through the United States Center for Disease Control and Prevention's Active Bacterial Core surveillance (ABCs) in 2016 and 2017. We determined the distribution of strain features, including *emm* type, antibiotic resistance determinants, and selected virulence factors. Changes in strain feature distribution between years 2016 and 2017 were evaluated. Phylogenetic analysis was used to identify expanding lineages within *emm* type.

Results: Seventy-one *emm* types were identified from 3873 isolates characterized. The *emm* types targeted by a 30-valent M protein-based vaccine accounted for 3230 (89%) isolates. The relative frequencies of *emm* types collected during the 2 years were similar. While all isolates were penicillin-susceptible, erythromycin-resistant isolates increased from 273 (16% of 2016 isolates) to 432 (23% of 2017 isolates), mainly driven by increase of the *erm*-positive *emm* types 92 and 83. The prevalence of 24 virulence factors, including 11 streptococcal pyrogenic toxins, ranged from 6 to 90%. In each of three *emm* types (*emm* 49, 82, and 92), a subgroup of isolates significantly expanded between 2016 and 2017 compared to isolates outside of the subgroup (*P*-values < 0.0001). Specific genomic sequence changes were associated with these expanded lineages.

Conclusions: While the overall population structure of invasive *S. pyogenes* isolates in the United States remained stable, some lineages, including several that were antibiotic-resistant, increased between 2016 and 2017. Continued genomic surveillance can help monitor and characterize bacterial features associated with emerging strains from invasive infections.

Keywords: group A streptococcus, genomics, surveillance studies, vaccines, antibiotic resistance, virulence

INTRODUCTION

Group A *Streptococcus* (GAS) is a Gram-positive bacterium that causes a variety of diseases in humans. Invasive GAS (iGAS) infections are associated with high case fatality ratios. Clinical manifestations of iGAS disease include cellulitis, bacteremia, pneumonia, necrotizing fasciitis, and streptococcal toxic shock syndrome. Outbreaks of iGAS disease can occur and have been reported in residents of long-term care facility (Deutscher et al., 2011; Dooling et al., 2013; Ahmed et al., 2018; Nanduri et al., 2019), in people experiencing homelessness (Bundle et al., 2017; Adebajo et al., 2018; Mosites et al., 2018), and in people who inject drugs (Lamagni et al., 2008; Kwiatkowska et al., 2018). Frequency of predominant GAS strains in a community often vary over time, presumably due to changes in population-level immunity against certain strains (Vlaminckx et al., 2005; Stockmann et al., 2012). The introduction of a novel or rare strain into a population has repeatedly been demonstrated to cause a rise in *Streptococcus pyogenes* infections (Tyrrell et al., 2010; Turner et al., 2015; Mosites et al., 2018), and key genomic changes within relatively uncommon strains that affect virulence factor structure or expression have been associated with subsequent global expansion (Turner et al., 2015; Beres et al., 2016). Although stable between 2000 and 2012, the incidence of invasive GAS infections in the United States (Nelson et al., 2016; CDC, 2019) nearly doubled from a baseline of 3.4–4.3 per 100,000 population to 7.26 cases per 100,000 persons in 2017 (CDC, 2019). Understanding iGAS disease trends and associated strain features through surveillance increases the understanding of population dynamics of GAS and can guide development of effective prevention strategies.

The GAS genome encodes a wide array of virulence factors. The M protein, encoded by the *emm* gene, is a major virulence factor involved in resisting phagocytosis (Fischetti, 1989) and elicits type-specific antibodies; this antigen is the foundation of the 30-valent M protein-based vaccine currently under development (Dale et al., 2011). The high *emm* gene sequence diversity within the *S. pyogenes* species formed the basis for the most-widely used strain subtyping method, *emm*-typing, with over 240 *emm* types documented to date (Bessen et al., 2018). The *erm* family antibiotic resistance genes in *S. pyogenes* are carried by several different transposon elements and shows evidence of intra-species spreading (Brenciani et al., 2007). Numerous exotoxin genes in *S. pyogenes* encode superantigens that stimulate uncontrolled T-cell responses underlying streptococcal toxic shock syndrome (Fraser and Proft, 2008). In the United States, approximately 40–60% of *S. pyogenes*

isolates contain an active *sof* gene (Chochua et al., 2017). The presence of *sof* can be predicted from *emm* gene sequence type (Facklam et al., 2002; Smeesters et al., 2010) and genetic organization of *emm* and neighboring “*emm*-like” genes (Bessen and Hollingshead, 1994). These two broad categories of GAS, distinguished by *sof*, *emm* pattern (Sanderson-Smith et al., 2014), or *emm*-locus gene organization (Bessen et al., 1996) differ in disease seasonality (Chochua et al., 2017; Smeesters et al., 2017), distributions between tropical/non-tropical locations, and exhibit different tissue preferences (Bessen et al., 2018). Recent developments in whole-genome sequencing (WGS) technology allow a fairly detailed characterization of *S. pyogenes* strain features, including *emm* type, antibiotic resistance genes, exotoxin genes, and other virulence factors for a population (Chochua et al., 2017). Use of WGS to characterize iGAS isolates collected from longitudinal surveillance will provide invaluable information on population dynamics, vaccine target prevalence, virulence factor distributions, and strain features associated with emerging lineages. In this study we aimed to determine the distribution of critical *S. pyogenes* strain features in a well-defined United States population and monitor changes from 2016 to 2017.

MATERIALS AND METHODS

Surveillance and Isolate Collection

All iGAS isolates were identified through the Active Bacterial Core surveillance (ABCs) as previously described (Chochua et al., 2017). ABCs is active, laboratory- and population-based surveillance of invasive bacterial infections as part of the Emerging Infections Program Network (EIP) of the United States Centers for Disease Control and Prevention (CDC). The iGAS surveillance areas represented approximately 34.2 million residents in 10 sites in 2017 (CDC, 2019). ABCs staff routinely contact microbiology laboratories serving hospitals at each site to identify cases of iGAS disease. A case is defined as illness with isolation of GAS from a normally sterile site or from a wound culture accompanied by necrotizing fasciitis or streptococcal toxic shock syndrome in a resident of a surveillance site. All available GAS isolates are sent to CDC’s *Streptococcus* laboratory for characterization.

In this study we characterized case isolates identified between January 01, 2016 and December 31, 2017. We limited our analyses to *S. pyogenes* (99.4% of GAS isolates); non-*S. pyogenes* GAS isolates (e.g., group

A *Streptococcus dysgalactiae* subspecies *equisimilis* have been previously described and were excluded (Chochua et al., 2019).

Whole-Genome Sequencing Analysis

Whole-genome sequencing was performed according to the method previously reported (Chochua et al., 2017). Briefly, each isolate was colony-purified, and the genomic DNA sample was extracted from an overnight broth culture. A multiplexed sequencing library was prepared from equal mixing of 48 genomic DNA samples. The library was sequenced on the Illumina MiSeq platform to produce 250-bp, paired-end reads with expected coverage depth of $>50\times$ for each isolate. Resulting sequencing reads were processed by a validated bioinformatics pipeline (Chochua et al., 2017) to obtain draft whole genome assembly (contigs), to identify subtype [*emm* type and multi-locus sequence type (ST)], to infer antimicrobial susceptibility, and to determine strain features for each isolate. The bioinformatics pipeline analysis output of each isolate was manually curated, and the curated results were entered into a MySQL database management system. The MySQL database was queried on January 10, 2019 to retrieve WGS analysis results of the study isolates. Bioinformatics scripts and associated sequence databases are available at <https://github.com/BenJamesMetcalf>.

Estimation of Vaccine Coverage

A 30-valent M protein-based vaccine candidate (Dale et al., 2011) composed of antigenic peptides from M proteins of 30 different *emm* types is a leading vaccine candidate currently in human trials (Vekemans et al., 2019). An *S. pyogenes* isolate was considered covered by the M protein-based vaccine candidate if the isolate belonged to one of the 30 *emm* types. Another antigen, the M-related protein (Mrp), has been proposed as a next generation GAS vaccine (Courtney et al., 2017); the addition of the additional Mrp antigen to M protein-based vaccines would broaden global coverage. Mrp antibodies are shown to be acquired following natural infection in the same age-related manner as M protein antibodies. An *S. pyogenes* isolate was considered as covered by the M-related proteins-based vaccine candidate if the isolate was positive for the *mrp* gene target (Chochua et al., 2017).

Antimicrobial Resistance Determinants

The WGS-analysis pipeline screened for the presence of streptococcal antimicrobial resistance determinants (gene alleles), which included 21 targets in an in-house sequence database (Chochua et al., 2017), 1913 targets in the Resfinder sequence database (Zankari et al., 2012), and 1653 targets in the ARG-ANNOT sequence database (Gupta et al., 2014). The presence of the antimicrobial resistance genes or mutations was used in a “rules-based” resistance classification of a strain to perform a WGS-based antimicrobial susceptibility test (AST) as previously described (Chochua et al., 2017). WGS-based AST quality control was performed by comparing to conventional phenotypic,

culture-based AST (see **Supplementary Method** for detail). A full list of antibiotic resistance targets observed in the study isolates is shown in **Supplementary Table S1**. The four common antimicrobial resistance targets analyzed in this study were:

ERM: presence of *erm* family genes associated with macrolide and clindamycin resistance.

TET: presence of *tet* family genes associated with tetracycline resistance.

PARC_GYRA : presence of *parC/gyrA* mutations associated with fluoroquinolone resistance.

MEF: presence of *mef* family genes associated with macrolide resistance.

Additionally, a PBP2x typing scheme was incorporated that detects substitutions within the transpeptidase region of PBP2x²¹.

Phylogenetic Analysis

Phylogenetic analysis was performed separately for each individual *emm* type that contained 20 or more isolates. Draft assemblies of isolates were first aligned by a fast core-genome multi-aligner implemented in Parsnp v1.2 (Treangen et al., 2014) with the -x option to filtering out regions of recombination. The core genome alignment generated by Parsnp was used to build a phylogenetic tree using FastTree v2.1.10 (Price et al., 2009). All monophyletic groups in the phylogenetic tree were identified using the subtree function implemented in the R ape package (R.C.Team, 2015). Isolates belonging to a monophyletic group were defined as a subgroup. To identify an expanded subgroup, the relative frequency of the subgroup in 2016 was compared to that in 2017 using a Fisher’s exact test and a *p*-value less than the Bonferroni-corrected threshold for multiple testing was considered significant. If two subgroups were both significantly expanded and one subgroup contained all isolates of the other subgroup, then only the subgroup with smaller *p*-value was retained for further analysis.

To identify genome sequence features that were characteristic for a subgroup, draft assemblies were converted into *k*-mers using the fsm-lite software, which enabled frequency-based string mining (Lees et al., 2016). A characteristic *k*-mer of subgroup X was defined as a *k*-mer that was present in $>90\%$ of isolates within subgroup X but in $<10\%$ of isolates outside subgroup X, or present in $<10\%$ of isolates within subgroup X but in $>90\%$ of isolates outside subgroup X. Annotation of characteristic *k*-mers were performed by mapping the *k*-mer sequences against reference, complete GAS genomes using miniMap (Li, 2016).

Statistical Analysis

Test of equal proportions in two groups was performed using the Fisher’s exact test. *P*-values were adjusted for multiple comparisons using the Bonferroni method where indicated. The strength of the association between *emm* type and strain features was quantified by the Cramer’s *V* statistic. All

statistical analysis was performed in the R software v3.2.2 (R.C.Team, 2015).

RESULTS

Distribution of *emm* Types and STs

The ABCs identified 4,425 iGAS cases between January 01, 2016 and December 31, 2017. GAS isolates from 3,895 (88%) cases were available and characterized by WGS. We included 3,873 (99.4%) isolates identified as *S. pyogenes*; 22 non-*S. pyogenes* GAS isolates (all group A *Streptococcus dysgalactiae* subspecies *equisimilis*) were excluded. **Table 1** shows the distribution of strain features inferred from WGS. A total of 71 *emm* types were identified among the 3,864 isolates (9 isolates were non-typeable). The number of isolates within each *emm* type ranged from 1 (for 20 *emm* types) to 566 (*emm1*). The 10 most common *emm* types (in decreasing order: 1, 89, 49, 12, 3, 82, 92, 28, 4, and 11) each contained ≥ 58 isolates and collectively accounted for 2,773 isolates (72%). More than 97% of isolates ($n = 3,773$) corresponded to the 31 most common *emm* types that each represented at least 10 isolates (**Figure 1A**).

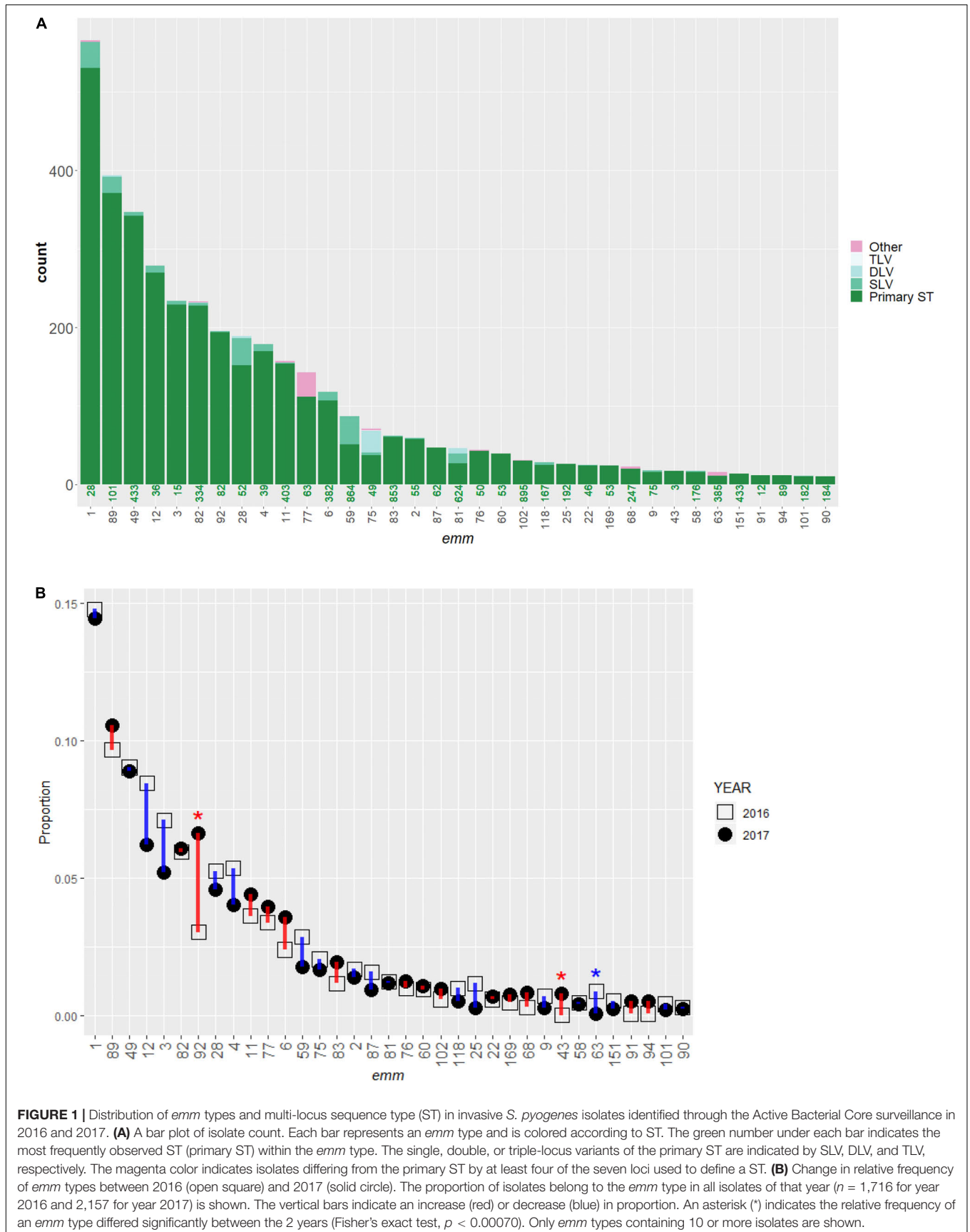
A total of 143 STs were assigned for 3,845 isolates while STs for the other 28 isolates (0.72%) were unassigned at the time of database query. Isolates within a given *emm* type were almost exclusively represented by a primary ST (the most frequently observed ST within an *emm* type) and its single, double, or triple-locus variants (**Figure 1A**). Isolates with an ST that differed from the within-*emm* type primary ST by four or more loci (**Figure 1A**, red bars) were only seen in 64 isolates (1.7%), approximately half of which ($n = 31$) were type *emm77*. Each of the 10 most common STs (in decreasing order: 28, 101, 433, 36, 334, 15, 82, 39, 403, and 52) were also the primary ST of one of the respective 10 most common *emm* types (1, 89, 49, 12, 3, 82, 92, 28, 4, and 11). The 10 and 20 most common STs represented 2664 (69%) and 3081 (85%) isolates, respectively. The *S. pyogenes* ST was strongly correlated with the *emm* type (Cramer's $V = 0.93$). Within each ST, the number of *emm* types observed was either one (in 135 STs), or two (in 8 STs).

In 2016 and 2017, 1,716, and 2,157 isolates were identified, respectively. The relative frequency of common *emm* types in 2016 was similar to that in 2017 (**Figure 1B**) with three significant exceptions: *emm92*, 63, and 43 (Fisher's exact test, P -values < 0.001). The proportion of *emm92* isolates increased from 3.0% ($n = 52$) in 2016 to 6.7% ($n = 143$) in 2017 and occurred

TABLE 1 | Prevalence of surface structure and virulence related strain features in invasive *S. pyogenes* isolates identified through the Active Bacterial Core surveillance system in the United States in 2016 and 2017 ($n = 3873$).

Strain feature ^a	Description	Count (proportion)
SPE_G	Streptococcal pyrogenic exotoxin G	3,500(90.4%)
FBAA	Fibronectin-binding protein of group A streptococci type A	3,204(82.7%)
SME_Z	Streptococcal mitogenic exotoxin	3,123(80.6%)
D330G	NADase D330G substitution associated with increased NADase activity	3,038(78.4%)
CAPSULE	<i>hasA</i> hyaluronic acid synthetase operon determinant	2,650(68.4%)
MRP	<i>emm</i> -like <i>mrp</i> virulence gene	2,609(67.4%)
ENN	<i>emm</i> -like <i>enn</i> virulence gene	2,565(66.2%)
SOF	Serum opacity factor	2,515(64.9%)
PRTF2	<i>S. pyogenes</i> fibronectin-binding protein	2,477(64%)
SFB1	<i>S. pyogenes</i> fibronectin-binding protein I	2,424(62.6%)
SPE_C	Streptococcal pyrogenic exotoxin C	2,019(52.1%)
PNGA3	Pnga 3 – Clade 3 up-regulated promoter of the <i>nga</i> operon	1,865(48.2%)
SPE_J	Streptococcal pyrogenic exotoxin J	1,190(30.7%)
SPE_H	Streptococcal pyrogenic exotoxin H	1,003(25.9%)
SPE_I	Streptococcal pyrogenic exotoxin I	954(24.6%)
SDA1	Virulence associated DNase	890(23%)
SIC	Streptococcal inhibitor of complement	826(21.3%)
SPE_A	Streptococcal pyrogenic exotoxin A	788(20.3%)
SS_A	Streptococcal Superantigen A	510(13.2%)
SPE_K	Streptococcal pyrogenic exotoxin K	496(12.8%)
R28	Homolog of a group B streptococcal adhesin	365(9.4%)
SPE_L	Streptococcal pyrogenic exotoxin L	269(6.9%)
SPE_M	Streptococcal pyrogenic exotoxin M	263(6.8%)
ROCA	<i>rocA</i> null mutation conserved within type <i>emm3</i> strains	237(6.1%)

^aA strain feature is either positive (presence) or negative (absence) in an isolate.



primarily in three states: Colorado, Tennessee, and New Mexico (**Supplementary Figure S1A**). The proportion of *emm63* isolates decreased from 0.9% ($n = 15$) to <0.1% ($n = 1$); the decrease was ubiquitous across all five states where *emm63* was present in 2016 (**Supplementary Figure S1B**). In four states the proportion of *emm43* isolates increased from 0% ($n = 0$) to 0.8% ($n = 17$) (**Supplementary Figure S1C**).

Estimated Vaccine Coverage

Twenty-four of the 30 *emm* types covered by the M protein-based vaccine candidate were found in 3,411 (88%) of the 3,873 study isolates. WGS identified the presence of the Mrp target in 2,609 (67%) isolates (**Table 1**), which included nearly all (409 of 462) isolates not targeted by the 30-valent M protein-based vaccine candidate. The combination of antigens found in both vaccine candidates could elicit protection from infection due to 98.6% of invasive *S. pyogenes* isolates in this study.

Major Genetic Determinants to Antimicrobial Resistance

Among the 3,873 isolates, the most frequently identified resistance determinants were the TET target in 876 (22.6%) isolates, the ERM target in 711 (18.4%) isolates, the MEF target in 36 (0.9%) isolates, and the GYRA_PARC target in 49 (1.4%) isolates (**Figure 2A** and **Supplementary Table S1**). The ERM target identified in the study isolates included *ermB* ($n = 219$), *ermT* ($n = 276$), *ermTR* ($n = 211$), and unassigned *erm* genes ($n = 5$, **Supplementary Table S1**). Seven ERM-positive isolates carried additional changes associated with erythromycin/clindamycin resistance (**Supplementary Table S1**), including 23S rRNA mutations ($n = 3$), *lsa* gene ($n = 2$), and *lsac* gene ($n = 2$). Most of the *gyrA/parC* gene mutations encoded the PARC-S79F ($n = 42$) substitution associated with intermediate levofloxacin resistance (MIC = 4 $\mu\text{g}/\text{mL}$). Only four isolates carried mutations in both *gyrA* and *parC* and were levofloxacin-resistant (MIC ≥ 8 $\mu\text{g}/\text{mL}$, **Supplementary Table S1**).

No PBP2x sequence types or other determinants associated with penicillin non-susceptibility were identified (**Supplementary Table S1**). Approximately (79% 3,220/4,095) of isolates were of the reference sequence type pBP2x-1, similar to the percentage (76.1%) found during 2015. The second most common type (PBP2x-3) was observed among 396 isolates, all of which were type *emm89*. Although 31 additional PBP2x types were each observed among 1–64 isolates, none were associated with non-susceptibility to beta lactam antibiotics, however, it is possible that our conventional testing panel for six different beta lactam antibiotics did not contain antimicrobial concentrations low enough to detect low level changes above basal wild type MICs. One previously described PBP2x substitution reported to be associated with very small increases of MICs for penicillin and ampicillin, P601L (PMID: 31996443), was observed among 10 isolates. Four isolates (3 *emm75* and 1 *emm1*) contained PBP2x-7 (P601L). Six *emm89* isolates contained PBP2x-16 (S562T + P601L), from 2015 isolates in our previous study we note that we have encountered a third type, PBP2x-28 (M593V,

P601L) from a single type *emm87* isolate (Fraser and Proft, 2008). No gene alleles associated with reduced susceptibility to vancomycin, chloramphenicol, linezolid, trimethoprim-sulfamethoxazole, or synergid were found (**Supplementary Table S1**). Only one isolate was non-susceptible to rifampin (MIC = 6 $\mu\text{g}/\text{mL}$) associated with an *rpoB* mutation resulting in the RPOB-D476E change (**Supplementary Table S1**).

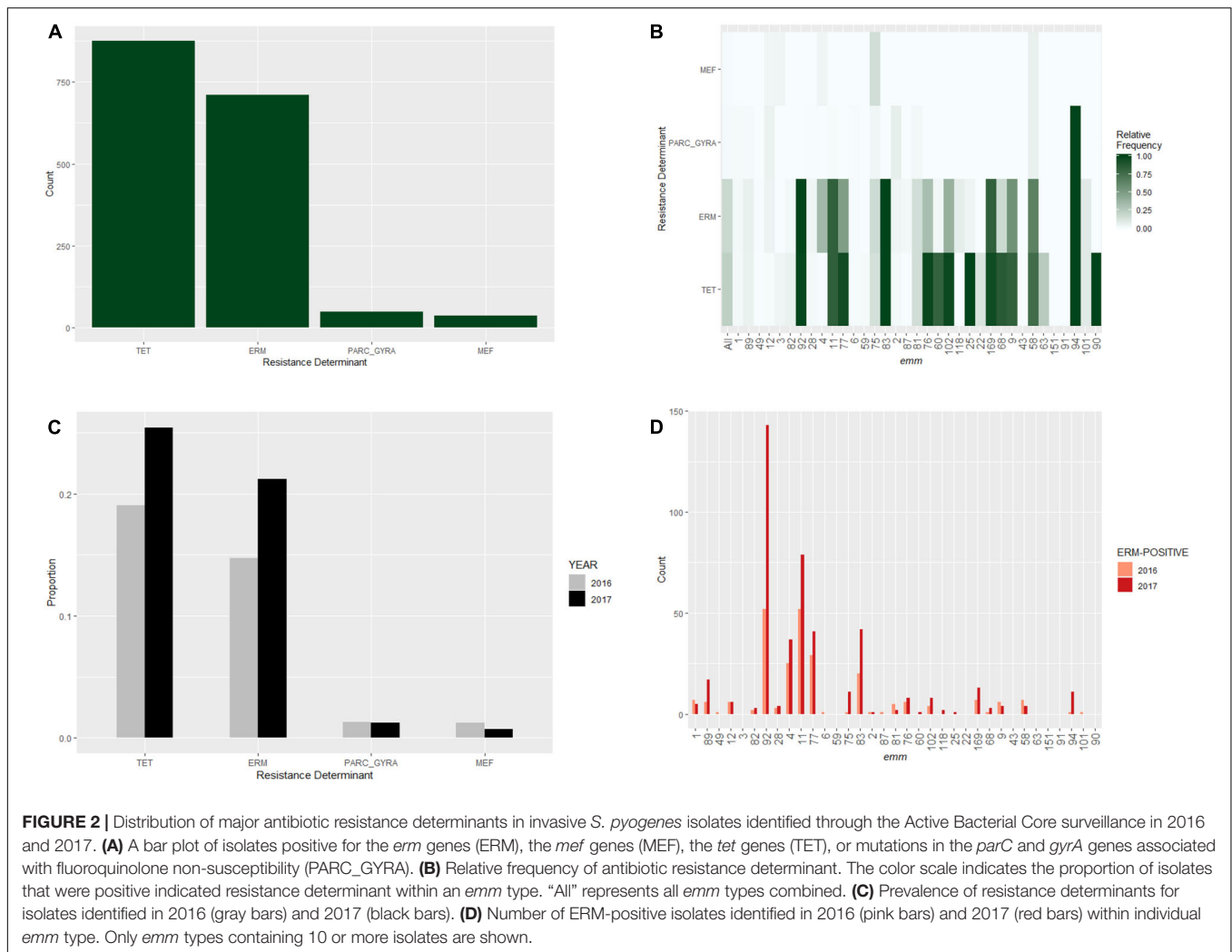
Major antimicrobial resistance determinants showed strong association with relatively infrequent *emm* types (**Figure 2B**). Greater than 80% of isolates belonging to *emm* types 92, 11, 83, 169, and 94 had an erythromycin resistance determinant (**Figure 2B**). All 12 isolates belonging to *emm94* carried the PARC-S79F substitution which is a levofloxacin resistance determinant (**Figure 2B**). Major resistance determinants were rarely seen in the six most common *emm* types (1, 89, 49, 12, 3, and 82; **Figure 2B**).

We next investigated whether the relative frequency of common resistance determinants differed between years (**Figure 2C**). The proportion of ERM-positive isolates in 2017 (21%) was significantly higher than that in 2016 (15%; Fisher's exact test, P -value = 0.001). A similar increase was observed for the proportion of TET-positive isolates (25% vs. 19%; $P = 0.001$). In contrast, no significant between-year difference in the relative frequency of MEF-positive or *GyrA/ParC* substitution mutants was found (P -values > 0.05). The prevalence of erythromycin-resistance (isolates positive for ERM or MEF target) increased from 16% of 2016 isolates to 22% of 2017 isolates. The largest increase in ERM-target positive isolates was due to the increases of the predominantly resistant *emm* types 92 (*ermT*), 83 (*ermTR*), and 11 (*ermTR*) (**Figure 2D**).

Surface Structure and Virulence Related Strain Features

The WGS-analysis pipeline screened for the presence of 13 surface structure and virulence factor gene targets and 11 streptococcal pyrogenic exotoxin gene targets in each isolate (**Table 1**). The prevalence of different strain features among the 3,873 isolates ranged from 237 (6.1%) for the *rocA* null mutation seen only in *emm3* strains to 3,500 (90%) for streptococcal pyrogenic exotoxin G (**Table 1**). Strain features showed strong association with *emm* type (Cramer's V range 0.76–0.95; **Figure 3B**). Approximately one-third of the study isolates ($n = 1223$) were predicted to lack the hyaluronic acid capsule by virtue of lacking the *hasA* target [e.g., being nearly all *emm4* and *emm22* isolates (Flores et al., 2012)], or by containing inactivating frameshifts within *hasA* (*emm28* is an example where all isolates contained a conserved frameshift within *hasA*) (Turner et al., 2019). It is interesting to note the fairly common occurrence of what appear to be sporadically occurring null mutations within *hasA* that occur within many predominantly encapsulated GAS (data not shown).

None of the 24 strain features showed significant between-year difference in the relative frequency (**Figure 3C**, Fisher's exact test, p -values > 0.05 after Bonferroni correction). A close examination of the yearly change in strain feature relative frequency within individual *emm*-type identified several notable



variations (Figure 3D). For example, within *emm11*, the proportion of isolates positive for the virulence associated DNase (SDA1) increased from 3.2% (2 of all 62 *emm11* isolates in 2016) to 20.2% (17 of all 84 *emm11* isolates in 2017). Within *emm6*, the proportion of isolates positive for the streptococcal pyrogenic exotoxin I (SPE_I) increased from 12% (5/41) to 30% (22/74). Within *emm3*, the proportion of isolates positive for the streptococcal pyrogenic exotoxin C (SPE_C) decreased from 17% (21/122) to 7% (8/110).

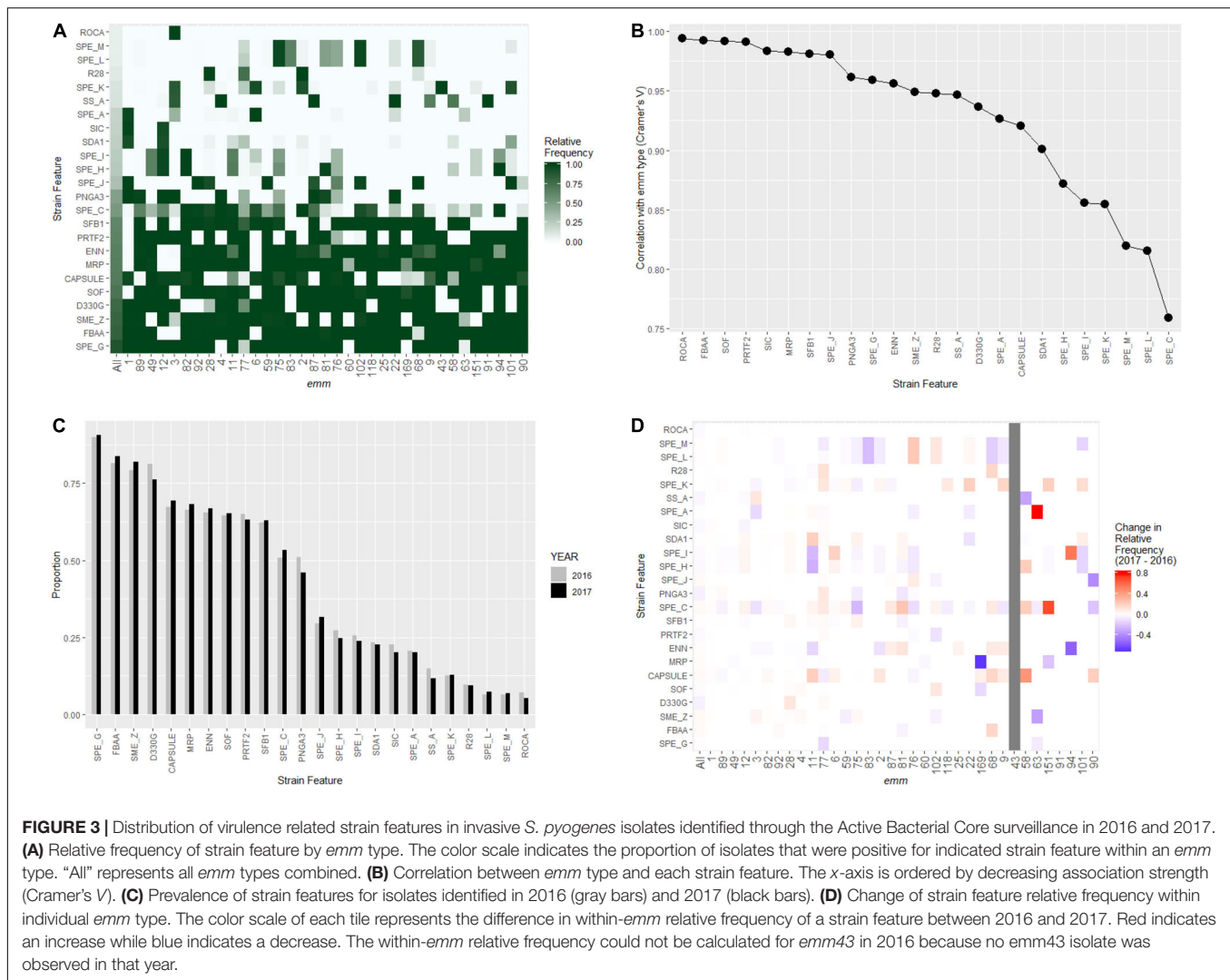
Expanded Lineages Within *emm* Type

We used a phylogenetic approach to investigate whether the relative frequency of a subgroup (lineage) within an *emm* type increased from 2016 to 2017. A subgroup was defined as a single branch that consisted of a common ancestor (the ancestral node) and all its lineal descendants on an *emm*-specific core genome tree. A total of 3,236 subgroups were identified from the 26 *emm* types which contained 20 or more isolates. Three subgroups showed a significant expansion from 2016 to 2017 (Figure 4), including subgroup32 in *emm92* (Figures 4A,B), subgroup25

in *emm82* (Figures 4C,D), and subgroup171 in *emm49* (Figures 4E,F).

Based on all characteristic *k*-mers identified, the subgroup32 of *emm92* differed from non-subgroup32 *emm92* by a sequence variation at a position corresponding to position 995380 of the reference genome NC_003737.2 (M1 SF370 complete genome). Nucleotide sequence at this position was C for all the subgroup32 isolates and A for all the non-subgroup32 isolates. The C to A sequence variation represented a modifier variant upstream of a putative pyridoxamine kinase gene (RefSeq:NP_269347.1). Within *emm92*, the proportion of subgroup32 isolates increased from 0.39 (20 of 52 isolates) in 2016 to 0.76 (109 of 143 isolates) in 2017 (Figure 4B; p -value = 2.1×10^{-6} , Fisher's Exact test; Bonferroni-corrected threshold: $0.05/3236 = 1.5 \times 10^{-5}$).

The subgroup25 of *emm82* differed from non-subgroup25 of *emm82* by a sequence variation in the fibronectin binding protein gene (*prtF2*, YP_279582.1), among others (Supplementary Table S2). Nucleotide sequence at *prtF2* position 751 was A for all the subgroup25 isolates and G for all the non-subgroup25 isolates. The A to G sequence variation resulted in a Thr to



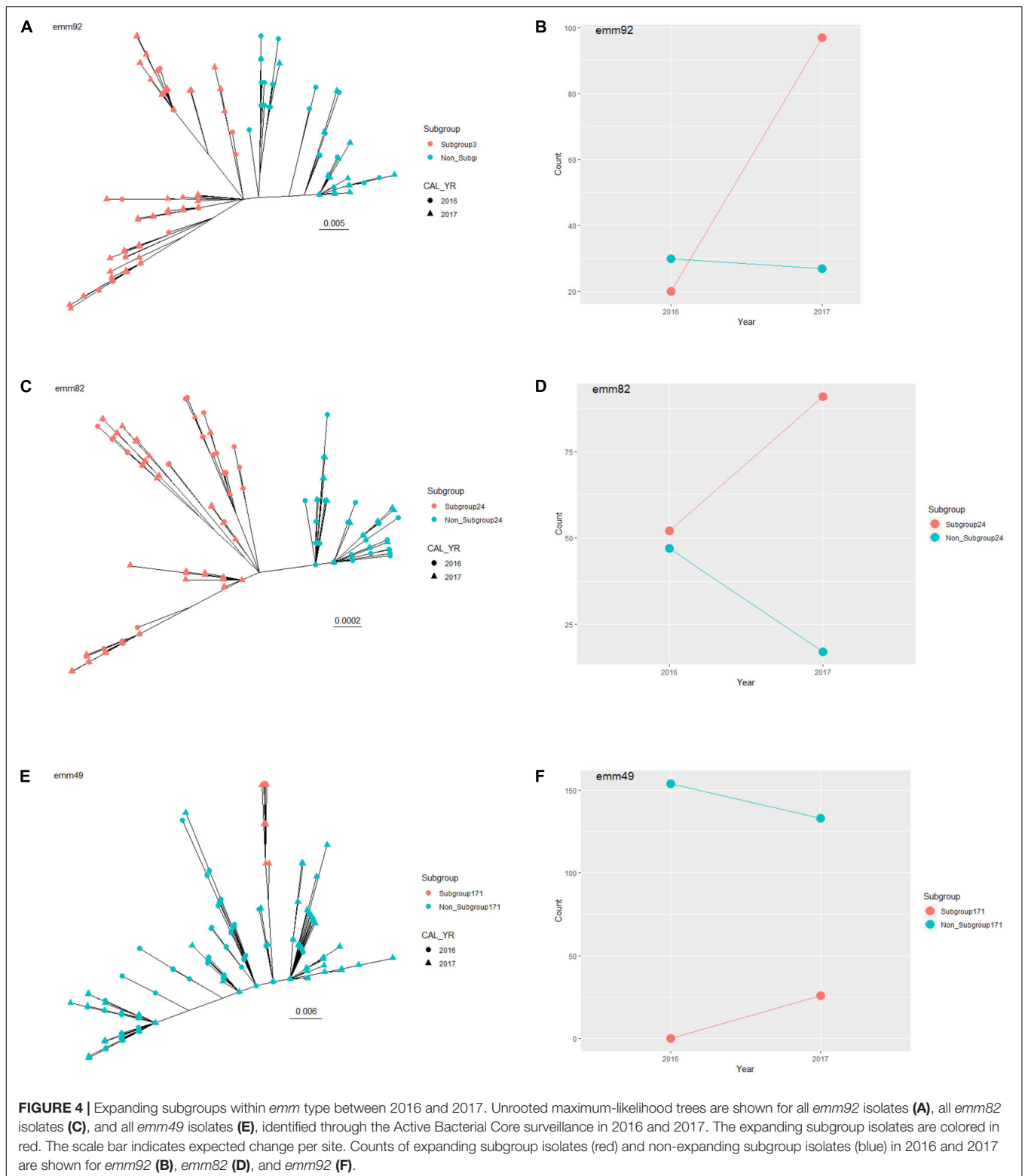
Ala alteration at amino acid position 251 of the prtF2 protein. Within *emm*82, the proportion of subgroup25 isolates increased from 0.53 (52/102) in 2016 to 0.81 (106/131) in 2017 (**Figure 4D**; p -value = 1.4×10^{-6}).

The subgroup171 of *emm*49 differed from non-subgroup171 of *emm*49 by a sequence variation in the branched-chain amino acid aminotransferase gene (*bcaT*, YP_598381.1), among others (**Supplementary Table S3**). All the subgroup171 isolates had a nucleotide sequence of A at *bcaT* gene position 97 while all the non-subgroup171 isolates had a sequence of G. This A to G sequence variation resulted in a Ser to Gly alteration at amino acid position 33 of the BcaT protein. Within *emm* 49, the proportion of subgroup171 isolates increased from 0 (0/155) in 2016 to 0.16 (30/192) in 2017 (**Figure 4E**; p -value = 1.4×10^{-6}).

DISCUSSION

In this study we characterized the genomic features of 3873 *S. pyogenes* isolates recovered from a population-based

surveillance system for invasive GAS infections operating in 10 geographically disparate United States sites in 2016–2017. For the vast majority of the GAS isolates, *emm* type appeared to be a very good approximation of clonal complex, defined by a founder ST and all STs sharing four or more identical ST loci with the founder ST (Feil et al., 2004). Presence of an *emm* type in two or more clonal complexes remained rare except for *emm*77, unlike the much more complex serotype-clonal relationship observed in another human-specific pathogen, *Streptococcus pneumoniae* (Beall et al., 2018). The overall frequency distribution of *emm* types in both years was similar (**Figure 1B**), which reflected a lack of substantial population structure shift among common *emm* types during this 2-year time period. Significant changes in relative frequency was observed only for *emm* types that are normally rare within ABCs (*emm* 92, 43, and 63), which could result from either directional selection forces acting upon these types or a temporary neutral drift that tended to influence types with smaller effective population size more profoundly. The GAS vaccine candidates based on M- and M related-protein showed high theoretical coverage, consistent with results



from previous studies (Dale et al., 2011; Chochua et al., 2017; Courtney et al., 2017). The hypervariable region of the M protein was shown to contain epitopes that both elicit protective

antibodies against GAS infection and minimally evoke human tissue cross-reactive antibodies (Dale et al., 1993, 2011; Dale and Walker, 2020). In a recent phase I clinical study, the

30-valent M protein-based vaccine was shown to be safe and immunogenic without eliciting autoimmunity in adults (Pastural et al., 2020).

The WGS-based monitoring system based on PBP2x transpeptidase sequence types automatically flags all GAS strains with *pbp2x* mutations in the routine surveillance (Chochua et al., 2017) as well as in outbreak investigations (Mosites et al., 2018; Vannice et al., 2019). Recently, two nearly isogenic strains were flagged by the CDC's bioinformatics pipeline and showed substantially elevated ampicillin and amoxicillin MICs (0.125–0.25 $\mu\text{g/ml}$), that are the highest GAS MICs recorded to date for these antibiotics (Vannice et al., 2019). Both isolates carried a same point mutation within the *pbp2x* leading to a PBP2x T553K substitution (Vannice et al., 2019) and would not have been phenotypically tested if not for the PBP2x point mutation that was detected through our established pipeline (Chochua et al., 2017). The WGS-based monitoring system supports a more sensitive level of surveillance on reduced beta lactam susceptibility and allows further investigations on potential selective advantages conferred by such mutations. PBP2x substitutions found in the recent studies (Musser et al., 2020; Olsen et al., 2020) of *emm* 1, 28, and 89 strains that showed modestly elevated beta lactam MICs, are also represented within this sampling. As shown this year 2016–2017 strain survey and the previous survey of 2015 ABCs GAS (Chochua et al., 2017), the P601L substitution was found among additional *emm* types and within three different PBP types, however, our MIC determinations were not low enough to discern these relatively subtle MIC increases. We now plan to examine our PBP2x type variants carefully to determine exact phenotypic changes that have occurred and to continue to carefully monitor these variants. Although it does appear that *pbp2x* point mutants regularly appear in our surveillance as a consequence of beta lactam selective pressure, present data suggests that there are strict constraints upon sequence variation (Hayes et al., 2020).

Two of the major antibiotic resistance determinants, the ERM and TET targets, were present in approximately 20% of isolates, yet were not evenly distributed among *emm* types (Figure 2B). Instead, the presence of ERM and TET targets were confined to a specific group of lineages and were rarely seen within the most common *emm* types in this study. This observation suggested that horizontal transfer of these resistant genes across lineages was uncommon (Varaldo et al., 2009). Although these genes were carried by various mobile genetic elements (data not shown), it is possible that these genetic elements confer detrimental fitness costs when transferred into normally non-resistant lineages. The proportion of erythromycin resistant isolates increased substantially between 2016 and 2017. The increase appeared to be caused by a general expansion of ERM-positive lineages relative to the ERM-negative lineages, with *emm* types 92 and 83 taking the leading role (Figure 2D). While changes in antibiotic consumption in certain populations have been associated with fluctuation in macrolide-resistance rates in *S. pyogenes*, (Murase and Yoshida, 1971; Bergman et al., 2004; Hsueh et al., 2005) it would be unlikely to explain a 7% increase in ERM-positive isolates within 1 year. The increasing frequency

of ERM-positive isolates that are also resistant to clindamycin, are a concern to public health. Clindamycin is regularly used to treat *S. pyogenes* infections in combination with β -lactams for patients with severe invasive disease (Allen and Moore, 2010; Stevens et al., 2014), and both clindamycin and erythromycin are alternative antibiotics for patients allergic to β -lactams (Bisno et al., 2002).

The WGS-based strain characterization enabled comprehensive detection of several previously characterized surface structure and virulence related features. A specific combination of these virulence features was usually unique to an *emm* type or a group of *emm* types and could be a contributing factor to the observed variation of virulence among different *emm* types. For example, isolates that were both PNGA3-positive and SOF-negative were solely represented by *emm* types 1, 12, and 3 while other SOF-negative *emm* types in Figure 3A (6, 84, 43, 191, and 101) were PNGA3-negative. These three *emm* types have been shown to be associated with higher risk of death due to iGAS infection (O'Loughlin et al., 2007; Thigpen et al., 2007; Nelson et al., 2016). WGS further enabled defining strain lineages within *emm* type based on genome-wide sequence variations. Three such lineages exhibited remarkable within-*emm* expansion between 2016 and 2017. These expansions could likely reflect development of local disease clusters or outbreaks that were initiated by ancestors of the individual lineages by chance. Alternatively, a within-*emm* expansion could suggest a selective advantage conferred by a novel bacterial feature other than *emm* type or strain features tightly associated with *emm* type. For example, all *emm*92 isolates were erythromycin-resistant but only subgroup32 of *emm*92 isolates were expanding, suggesting neither the *emm* type nor erythromycin resistance was sufficient to explain the different dynamics. In fact, genomic sequence variations specific to each expanding lineage were identified in this study. The information could aid in both tracking future spread of a lineage and unveiling mechanisms underlying the emergence of a successful clade. In retrospect, monitoring of specific *emm* types in this manner could reveal key changes likely to be critical for strain success. For example, recombinational changes resulted in an altered the surface protein antigenic profile and increased toxin production by the *emm*89 clonal complex. These changes led to the dramatic global emergence during the past 15 years (Beres et al., 2016). We have recently noted the appearance of 11 isolates of the recently reported M1_{UK} exotoxin over-producing clone, identifiable through a small set of characteristic base substitutions, that has dramatically emerged as an important cause of scarlet fever in the United Kingdom (Lynskey et al., 2019; Li et al., 2020). We will prospectively monitor this strain carefully within ABCs to determine if it becomes an important cause of invasive infections.

This study had several limitations. First, *emm* type distributions are known to vary substantially among different geographic locations (Steer et al., 2009). The broad category of "sof-positive *emm* types" when found in temperate or tropical countries often display completely unrelated lineages to their counterparts in North America or Western European countries (McGregor et al., 2004; Sakota et al., 2006; Bessen et al., 2018).

Therefore, it remained to be demonstrated whether the theoretical vaccine coverage and the strain feature frequencies would be similar in regions outside the ABCs catchment areas. Second, the study period (24-month) was relatively short, which made it difficult to determine whether any lineage dynamic observed represented a temporary, stochastic fluctuation or a long-term, mechanistic change in strain virulence. Continued genomic surveillance is critical to understanding the impact of molecular events associated with emerging lineages causing iGAS disease in order to design more effective control measures. Third, in this study we mainly performed genomic characterization of invasive GAS isolates. Phenotypic testing of strain virulence and toxin productions was not performed, although genome data on antimicrobial susceptibility have been validated. In addition, the observations of this study were limited to invasive infections in the United States. Future studies are needed to compare these observations to characterization of non-invasive infections collected concurrently in time and place, and to gain better understanding of the dynamics of strain transmission within and between the different infection types. Non-invasive GAS infections also cause a large burden to human health. While the absolute reduction in non-invasive GAS infections from a vaccine will be great, impact of any decline in the considerable morbidity and mortality due to invasive infections will be non-trivial and important drivers of GAS vaccine policy. Additional experimental studies and phenotypic testing are needed to evaluate fitness changes associated with a strain. The study only examined a selected group of previously reported virulence factors. More in-depth sub-analyses of additional genetic factors with influence on virulence are needed in future studies.

In conclusion, findings of the study suggested that the overall population structure of invasive *S. pyogenes* in the United States was stable between 2016 and 2017, but expanding lineages, including antibiotic resistant ones, were evident. The stability of the population structure within ABCs over time has far-reaching implications. For example, the increasing resistance within ABCs was largely due to types that are included within the current experimental 30 valent vaccine. Surveillance of this nature will serve to distinguish non-significant alterations associated with emerging strains from the important ones. Further, the high-quality genomic sequences reported here provides an excellent foundation for other independent analyses in the field of GAS genomics.

REFERENCES

- Adebanjo, T., Mosites, E., Van Beneden, C. A., Onukwube, J., Blum, M., Harper, M., et al. (2018). Risk factors for group A *Streptococcus* colonization during an outbreak among people experiencing homelessness in Anchorage, Alaska, 2017. *Clin. Infect. Dis.* 67, 1784–1787. doi: 10.1093/cid/ciy429
- Ahmed, S. S., Diebold, K. E., Brandvold, J. M., Ewaidah, S. S., Black, S., Ogundimu, A., et al. (2018). The role of wound care in 2 group A streptococcal outbreaks in a Chicago skilled nursing facility, 2015–2016. *Open Forum Infect. Dis.* 5:ofy145. doi: 10.1093/ofid/ofy145

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI BioProject PRJNA395240.

ETHICS STATEMENT

The ABCs case reporting and isolate collection were reviewed in accordance with Centers for Disease Control and Prevention human research protection procedures and were determined to be non-research public health surveillance. The ABCs sites reviewed the protocol and obtained institutional review board approval where required.

AUTHOR CONTRIBUTIONS

YL designed the study, performed the analyses provided, and wrote the initial manuscript draft. SC performed, directed, and evaluated all whole genome sequencing data. BM developed the bioinformatics pipeline and provided periodic updates. JR, SM, and ZL performed whole genome sequencing. PS-V and RL provided phenotypic MIC data and analysis for all isolates from the Minnesota surveillance site. SV maintained *emm* typing test as well as performed conventional testing of selected isolates. SN and CV assisted in population-based epidemiologic data analysis and writing. BB and LM oversaw all laboratory operations. BB guided bioinformatics pipeline development. All authors reviewed the manuscript and provided constructive feedback.

FUNDING

The CDC provided financial support for the Active Bacterial Core surveillance program through cooperative agreements with the Active Bacterial Core surveillance sites. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.01547/full#supplementary-material>

- Allen, U., and Moore, D. (2010). Invasive group A streptococcal disease: management and chemoprophylaxis. *Can. J. Infect. Dis. Med. Microbiol.* 21, 115–118. doi: 10.1155/2010/585187
- Beall, B., Chochua, S., Gertz, R. E. Jr., Li, Y., Li, Z., McGee, L., et al. (2018). A population-based descriptive atlas of invasive pneumococcal strains recovered within the U.S. during 2015–2016. *Front. Microbiol.* 9:2670. doi: 10.3389/fmicb.2018.02670
- Beres, S. B., Kachroo, P., Nasser, W., Olsen, R. J., Zhu, L., Flores, A. R., et al. (2016). Transcriptome remodeling contributes to epidemic disease caused by

- the human pathogen *Streptococcus pyogenes*. *mBio* 7:e00403-16. doi: 10.1128/mBio.00403-16
- Bergman, M., Huikko, S., Pihlajamäki, M., Laippala, P., Palva, E., Huovinen, P., et al. (2004). Effect of macrolide consumption on erythromycin resistance in *Streptococcus pyogenes* in Finland in 1997–2001. *Clin. Infect. Dis.* 38, 1251–1256. doi: 10.1086/383309
- Bessen, D. E., and Hollingshead, S. K. (1994). Allelic polymorphism of emm loci provides evidence for horizontal gene spread in group A streptococci. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3280–3284. doi: 10.1073/pnas.91.8.3280
- Bessen, D. E., Smeesters, P. R., and Beall, B. W. (2018). Molecular epidemiology, ecology, and evolution of group A streptococci. *Microbiol. Spectr.* 6, 1–33. doi: 10.1128/microbiolspec.CPP3-0009-2018
- Bessen, D. E., Sotir, C. M., Readdy, T. L., and Hollingshead, S. K. (1996). Genetic correlates of throat and skin isolates of group A streptococci. *J. Infect. Dis.* 173, 896–900. doi: 10.1093/infdis/173.4.896
- Bisno, A. L., Gerber, M. A., Gwaltney, J. M. Jr., Kaplan, E. L., Schwartz, R. H., and Infectious Diseases Society of America (2002). Practice guidelines for the diagnosis and management of group A streptococcal pharyngitis. *Clin. Infect. Dis.* 35, 113–125. doi: 10.1086/340949
- Brenciani, A., Bacciaglia, A., Vecchi, M., Vitali, L. A., Varaldo, P. E., and Giovanetti, E. (2007). Genetic elements carrying erm(B) in *Streptococcus pyogenes* and association with tet(M) tetracycline resistance gene. *Antimicrob. Agents Chemother.* 51, 1209–1216. doi: 10.1128/AAC.01484-06
- Bundle, N., Bubba, L., Coelho, J., Kwiatkowska, R., Cloke, R., King, S., et al. (2017). Ongoing outbreak of invasive and non-invasive disease due to group A *Streptococcus* (GAS) type emm66 among homeless and people who inject drugs in England and Wales, January to December 2016. *Euro. Surveill.* 22:30446. doi: 10.2807/1560-7917.ES.2017.22.3.30446
- CDC (2019). Available online at: <https://www.cdc.gov/abcs/reports-findings/sur-reports.html> (accessed July 15, 2020).
- Chochua, S., Metcalf, B. J., Li, Z., Rivers, J., Mathis, S., Jackson, D., et al. (2017). Population and whole genome sequence based characterization of invasive group A streptococci recovered in the United States during 2015. *mBio* 8:e01422-17. doi: 10.1128/mBio.01422-17
- Chochua, S., Rivers, J., Mathis, S., Li, Z., Velusamy, S., McGee, L., et al. (2019). Emergent invasive group A *Streptococcus dysgalactiae* subsp. *equisimilis*, United States, 2015–2018. *Emerg. Infect. Dis.* 25, 1543–1547. doi: 10.3201/eid2508.181758
- Courtney, H. S., Niedermeyer, S. E., Penfound, T. A., Hohn, C. M., Greeley, A., and Dale, J. B. (2017). Trivalent M-related protein as a component of next generation group A streptococcal vaccines. *Clin. Exp. Vaccine Res.* 6, 45–49. doi: 10.7774/cevr.2017.6.1.45
- Dale, J. B., and Walker, M. J. (2020). Update on group A streptococcal vaccine development. *Curr. Opin. Infect. Dis.* 33, 244–250. doi: 10.1097/QCO.0000000000000644
- Dale, J. B., Chiang, E. Y., and Lederer, J. W. (1993). Recombinant tetravalent group A streptococcal M protein vaccine. *J. Immunol.* 151, 2188–2194.
- Dale, J. B., Penfound, T. A., Chiang, E. Y., and Walton, W. J. (2011). New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. *Vaccine* 29, 8175–8178. doi: 10.1016/j.vaccine.2011.09.005
- Deutscher, M., Schillie, S., Gould, C., Baumbach, J., Mueller, M., Avery, C., et al. (2011). Investigation of a group A streptococcal outbreak among residents of a long-term acute care hospital. *Clin. Infect. Dis.* 52, 988–994. doi: 10.1093/cid/cir084
- Dooring, K. L., Crist, M. B., Nguyen, D. B., Bass, J., Lorentzson, L., Toews, K. A., et al. (2013). Investigation of a prolonged Group A Streptococcal outbreak among residents of a skilled nursing facility, Georgia, 2009–2012. *Clin. Infect. Dis.* 57, 1562–1567. doi: 10.1093/cid/cit558
- Facklam, R. F., Martin, D. R., Lovgren, M., Johnson, D. R., Efstratiou, A., Thompson, T. A., et al. (2002). Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: emm103 to emm124. *Clin. Infect. Dis.* 34, 28–38. doi: 10.1086/324621
- Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., and Spratt, B. G. (2004). eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J. Bacteriol.* 186, 1518–1530. doi: 10.1128/jb.186.5.1518-1530.2004
- Fischetti, V. A. (1989). Streptococcal M protein: molecular design and biological behavior. *Clin. Microbiol. Rev.* 2, 285–314. doi: 10.1128/cmr.2.3.285
- Flores, A. R., Jewell, B. E., Fittipaldi, N., Beres, S. B., and Musser, J. M. (2012). Human disease isolates of serotype m4 and m22 group A streptococcus lack genes required for hyaluronic acid capsule biosynthesis. *mBio* 3:e00413-12. doi: 10.1128/mBio.00413-12
- Fraser, J. D., and Proft, T. (2008). The bacterial superantigen and superantigen-like proteins. *Immunol. Rev.* 225, 226–243. doi: 10.1111/j.1600-065X.2008.00681.x
- Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., et al. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* 58, 212–220. doi: 10.1128/AAC.01310-13
- Hayes, A., Lacey, J. A., Morris, J. M., Davies, M. R., and Tong, S. Y. C. (2020). Restricted sequence variation in *Streptococcus pyogenes* penicillin binding proteins. *mSphere* 5:e00090-20. doi: 10.1128/mSphere.00090-20
- Hsueh, P. R., Shyr, J. M., and Wu, J. J. (2005). Decreased erythromycin use after antimicrobial reimbursement restriction for undocumented bacterial upper respiratory tract infections significantly reduced erythromycin resistance in *Streptococcus pyogenes* in Taiwan. *Clin. Infect. Dis.* 40, 903–905. doi: 10.1086/427946
- Kwiatkowska, R. M., Manley, P., Sims, B., Lamagni, T., Ready, D., Coelho, J., et al. (2018). Outbreak of group A *Streptococcus* emm94.0 affecting people who inject drugs in southwest England, April 2017. *Am. J. Infect. Control* 46, 238–240. doi: 10.1016/j.ajic.2017.08.011
- Lamagni, T. L., Neal, S., Keshishian, C., Hope, V., George, R., Duckworth, G., et al. (2008). Epidemic of severe *Streptococcus pyogenes* infections in injecting drug users in the UK, 2003–2004. *Clin. Microbiol. Infect.* 14, 1002–1009. doi: 10.1111/j.1469-0691.2008.02076.x
- Lees, J. A., Vehkala, M., Valimäki, N., Harris, S. R., Chewapreecha, C., Croucher, N. J., et al. (2016). Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. *Nat. Commun.* 7:12797. doi: 10.1038/ncomms12797
- Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32, 2103–2110. doi: 10.1093/bioinformatics/btw152
- Li, Y., Nanduri, S. A., Van Beneden, C. A., and Beall, B. W. (2020). M1UK lineage in invasive group A streptococcus isolates from the USA. *Lancet Infect. Dis.* 20, 538–539. doi: 10.1016/S1473-3099(20)30279-6
- Lynskey, N. N., Jauneikaite, E., Li, H. K., Zhi, X., Turner, C. E., Mosavie, M., et al. (2019). Emergence of dominant toxigenic M1T1 *Streptococcus pyogenes* clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. *Lancet Infect. Dis.* 19, 1209–1218. doi: 10.1016/S1473-3099(19)30446-3
- McGregor, K. F., Bilek, N., Bennett, A., Kalia, A., Beall, B., Carapetis, J. R., et al. (2004). Group A streptococci from a remote community have novel multilocus genotypes but share emm types and housekeeping alleles with isolates from worldwide sources. *J. Infect. Dis.* 189, 717–723. doi: 10.1086/381452
- Mosites, E., Frick, A., Gounder, P., Castrodale, L., Li, Y., Rudolph, K., et al. (2018). Outbreak of invasive infections from subtype emm26.3 Group A *Streptococcus* among homeless adults—Anchorage, Alaska, 2016–2017. *Clin. Infect. Dis.* 66, 1068–1074. doi: 10.1093/cid/cix921
- Murase, T., and Yoshida, S. (1971). Effect of hypotension on levels of antidiuretic hormone in plasma in dogs. *Endocrinol. Jpn.* 18, 215–219. doi: 10.1507/endocrj1954.18.215
- Musser, J. M., Beres, S. B., Zhu, L., Olsen, R. J., Vuopio, J., Hyyryläinen, H. L., et al. (2020). Reduced in vitro susceptibility of *Streptococcus pyogenes* to beta-Lactam antibiotics associated with mutations in the pbp2x gene is geographically widespread. *J. Clin. Microbiol.* 58:e01993-19. doi: 10.1128/JCM.01993-19
- Nanduri, S. A., Metcalf, B. J., Arwady, M. A., Edens, C., Lavin, M. A., Morgan, J., et al. (2019). Prolonged and large outbreak of invasive group A *Streptococcus* disease within a nursing home: repeated intrafacility transmission of a single strain. *Clin. Microbiol. Infect.* 25, e1–e248. doi: 10.1016/j.cmi.2018.04.034
- Nelson, G. E., Pondo, T., Toews, K. A., Farley, M. M., Lindegren, M. L., Lynfield, R., et al. (2016). Epidemiology of invasive group A Streptococcal infections in the United States, 2005–2012. *Clin. Infect. Dis.* 63, 478–486. doi: 10.1093/cid/ciw248

- O'Loughlin, R. E., Roberson, A., Cieslak, P. R., Lynfield, R., Gershman, K., Craig, A., et al. (2007). The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. *Clin. Infect. Dis.* 45, 853–862. doi: 10.1086/521264
- Olsen, R. J., Zhu, L., and Musser, J. M. (2020). A single amino acid replacement in penicillin binding protein 2X in *Streptococcus pyogenes* significantly increases fitness upon Subtherapeutic Benzylpenicillin treatment in a mouse model of necrotizing myositis. *Am. J. Pathol.* 1–7. doi: 10.1016/j.ajpath.2020.04.014
- Pastural, E., McNeil, S. A., MacKinnon-Cameron, D., Ye, L., Langley, J. M., Stewart, R., et al. (2020). Safety and immunogenicity of a 30-valent M protein-based group A streptococcal vaccine in healthy adult volunteers: a randomized, controlled phase I study. *Vaccine* 38, 1384–1392. doi: 10.1016/j.vaccine.2019.12.005
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077
- R.C.Team (2015). *R: A Language and Environment for Statistical Computing*. Available online at: <https://www.R-project.org> (accessed February 29, 2020).
- Sakota, V., Fry, A. M., Lietman, T. M., Facklam, R. R., Li, Z., and Beall, B. (2006). Genetically diverse group A streptococci from children in far-western Nepal share high genetic relatedness with isolates from other countries. *J. Clin. Microbiol.* 44, 2160–2166. doi: 10.1128/JCM.02456-05
- Sanderson-Smith, M., De Oliveira, D. M., Guglielmini, J., McMillan, D. J., Vu, T., Holien, J. K., et al. (2014). A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. *J. Infect. Dis.* 210, 1325–1338. doi: 10.1093/infdis/jiu260
- Smeesters, P. R., Laho, D., Beall, B., Steer, A. C., and Van Beneden, C. A. (2017). Seasonal, geographic, and temporal trends of emm clusters associated with invasive group A streptococcal infections in US multistate surveillance. *Clin. Infect. Dis.* 64, 694–695. doi: 10.1093/cid/ciw807
- Smeesters, P. R., McMillan, D. J., and Sriprakash, K. S. (2010). The streptococcal M protein: a highly versatile molecule. *Trends Microbiol.* 18, 275–282. doi: 10.1016/j.tim.2010.02.007
- Steer, A. C., Law, I., Matatolu, L., Beall, B. W., and Carapetis, J. R. (2009). Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect. Dis.* 9, 611–616. doi: 10.1016/S1473-3099(09)70178-1
- Stevens, D. L., Bisno, A. L., Chambers, H. F., Dellinger, E. P., Goldstein, E. J., Gorbach, S. L., et al. (2014). Practice guidelines for the diagnosis and management of skin and soft tissue infections: 2014 update by the infectious diseases society of America. *Clin. Infect. Dis.* 59, e10–e52. doi: 10.1093/cid/ciu444
- Stockmann, C., Ampofo, K., Hersh, A. L., Blaschke, A. J., Kendall, B. A., Korgenski, K., et al. (2012). Evolving epidemiologic characteristics of invasive group A streptococcal disease in Utah, 2002-2010. *Clin. Infect. Dis.* 55, 479–487. doi: 10.1093/cid/cis422
- Thigpen, M. C., Richards, C. L. Jr., Lynfield, R., Barrett, N. L., Harrison, L. H., Arnold, K. E., et al. (2007). Invasive group A streptococcal infection in older adults in long-term care facilities and the community, United States, 1998-2003. *Emerg. Infect. Dis.* 13, 1852–1859. doi: 10.3201/eid1312.070303
- Treangen, T. J., Ondov, B. D., Koren, S., and Phillippy, A. M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 15:524. doi: 10.1186/s13059-014-0524-x
- Turner, C. E., Abbott, J., Lamagni, T., Holden, M. T., David, S., Jones, M. D., et al. (2015). Emergence of a new highly successful acapsular group A *Streptococcus* clade of genotype emm89 in the United Kingdom. *mBio* 6:e00622-15. doi: 10.1128/mBio.00622-15
- Turner, C. E., Holden, M. T. G., Blane, B., Horner, C., Peacock, S. J., Sriskandan, S. (2019). Available online at: <https://www.biorxiv.org/content/10.1101/684167v1.full> (accessed July 15, 2020).
- Tyrrell, G. J., Lovgren, M., St Jean, T., Hoang, L., Patrick, D. M., Horsman, G., et al. (2010). Epidemic of group A *Streptococcus* M/emm59 causing invasive disease in Canada. *Clin. Infect. Dis.* 51, 1290–1297. doi: 10.1086/657068
- Vannice, K., Ricaldi, J., Nanduri, S., Fang, F. C., Lynch, J., Bryson-Cahn, C., et al. (2019). *Streptococcus pyogenes* pbp2x mutation confers reduced susceptibility to beta-lactam antibiotics. *Clin. Infect. Dis.* 71, 201–204. doi: 10.1093/cid/ciz1000
- Varaldo, P. E., Montanari, M. P., and Giovanetti, E. (2009). Genetic elements responsible for erythromycin resistance in streptococci. *Antimicrob. Agents Chemother.* 53, 343–353. doi: 10.1128/AAC.00781-08
- Vekemans, J., Gouvea-Reis, F., Kim, J. H., Excler, J. L., Smeesters, P. R., O'Brien, K. L., et al. (2019). The path to group A *Streptococcus* vaccines: world health organization research and development technology roadmap and preferred product characteristics. *Clin. Infect. Dis.* 69, 877–883. doi: 10.1093/cid/ciy1143
- Vlaminckx, B. J., van Pelt, W., Schouls, L. M., van Silfhout, A., Mascini, E. M., Elzenaar, C. P., et al. (2005). Long-term surveillance of invasive group A streptococcal disease in The Netherlands, 1994-2003. *Clin. Microbiol. Infect.* 11, 226–231. doi: 10.1111/j.1469-0691.2004.01068.x
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261 doi: 10.1093/jac/dks261

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Rivers, Mathis, Li, Velusamy, Nanduri, Van Beneden, Snippet-Vagnone, Lynfield, McGee, Chochua, Metcalf and Beall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.