

# Emergent Invasive Group A *Streptococcus dysgalactiae* subsp. *equisimilis*, United States, 2015–2018

Sopio Chochua, Joy Rivers, Sandra Mathis, Zhongya Li, Srinivasan Velusamy, Lesley McGee, Chris Van Beneden, Yuan Li, Benjamin J. Metcalf, Bernard Beall

The term group A *Streptococcus* is considered synonymous for the species *Streptococcus pyogenes*. We describe an emergent invasive *S. dysgalactiae* subspecies *equisimilis* lineage that obtained the group A antigen through a single ancestral recombination event between a group C *S. dysgalactiae* subsp. *equisimilis* strain and a group A *S. pyogenes* strain.

The Centers for Disease Control and Prevention's Active Bacterial Core surveillance (ABCs) performs population-based surveillance of invasive group A *Streptococcus* (GAS) infections. Isolates collected from a population of ≈34 million persons are subjected to whole-genome sequence (WGS)-based characterization. We recently detected group A carbohydrate-positive *S. dysgalactiae* subsp. *equisimilis* (SE) isolates employing the *gacI* (1) sequence query within our bioinformatics pipeline (2). GAS is considered synonymous with *S. pyogenes*, rare occurrences of group A SE have been noted (3,4).

## The Study

During January 1, 2015–November 1, 2018, a total of 5,480 ABCs GAS isolates were subjected to WGS. We identified 35 atypical *gacI*-positive isolates; each yielded 1 of the M protein gene (*emm*) subtypes *stG245.0*, *stG485.0*, or *stG652.0* commonly associated with SE (4–6). These 35 isolates lacked multilocus sequence types (MLSTs) inclusive of known *S. pyogenes* allelic designations. Lancefield grouping (7) and MLST (<https://pubmlst.org/sdysgalactiae>) (6) revealed the 35 isolates were serologically group A and MLST sequence type (ST) 128 (GAS/ST128/SE). We received 13 additional SE isolates recovered through ABCs GAS surveillance during this period that were found to be non-group A isolates (9 group G, 2 group C, and 2 group

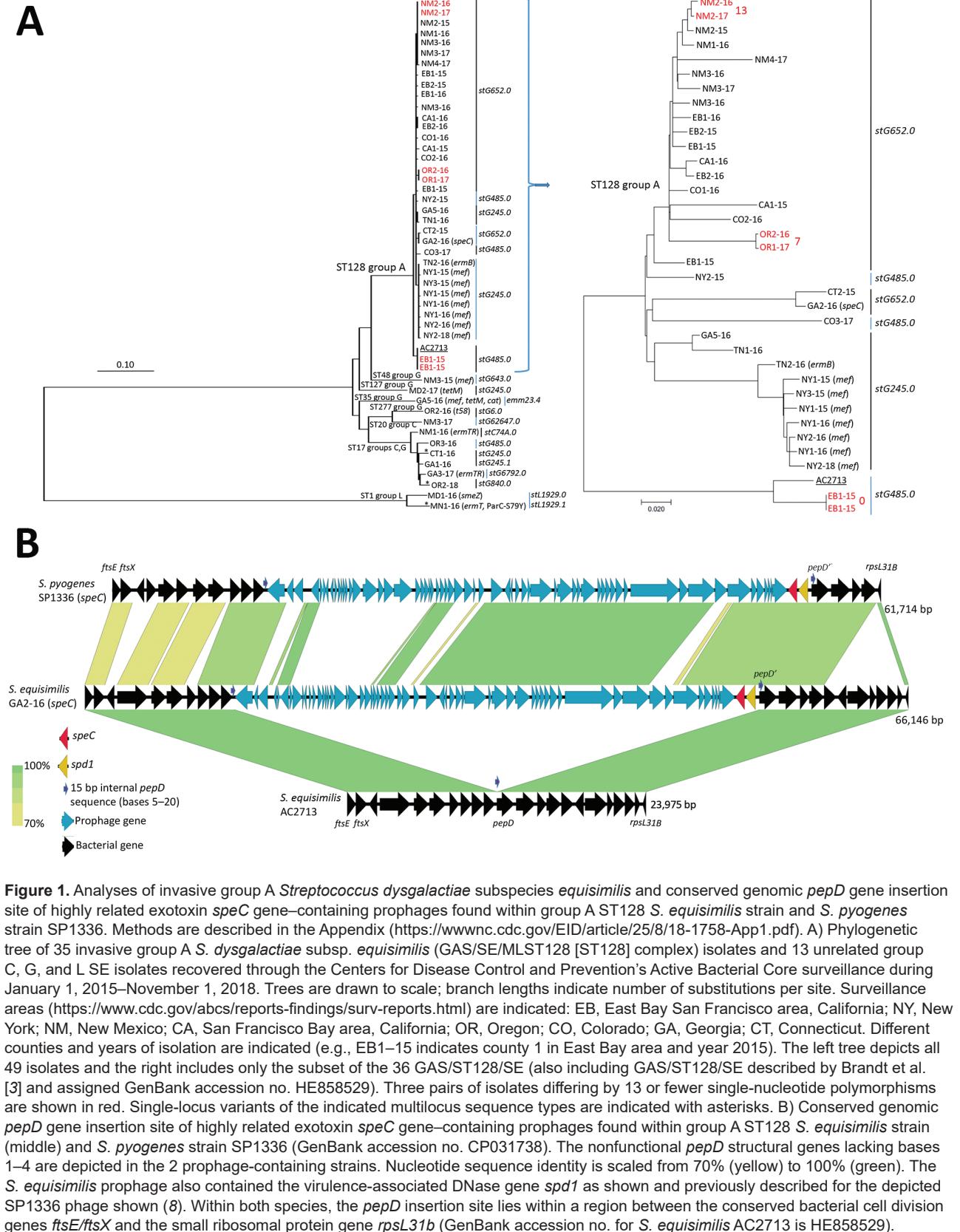
L) with MLSTs unrelated to ST128 (Figure 1). According to our normal protocol, these 13 non-group A SE isolates and 2 group G *S. canis* isolates that we also received were removed from the ABCs GAS database.

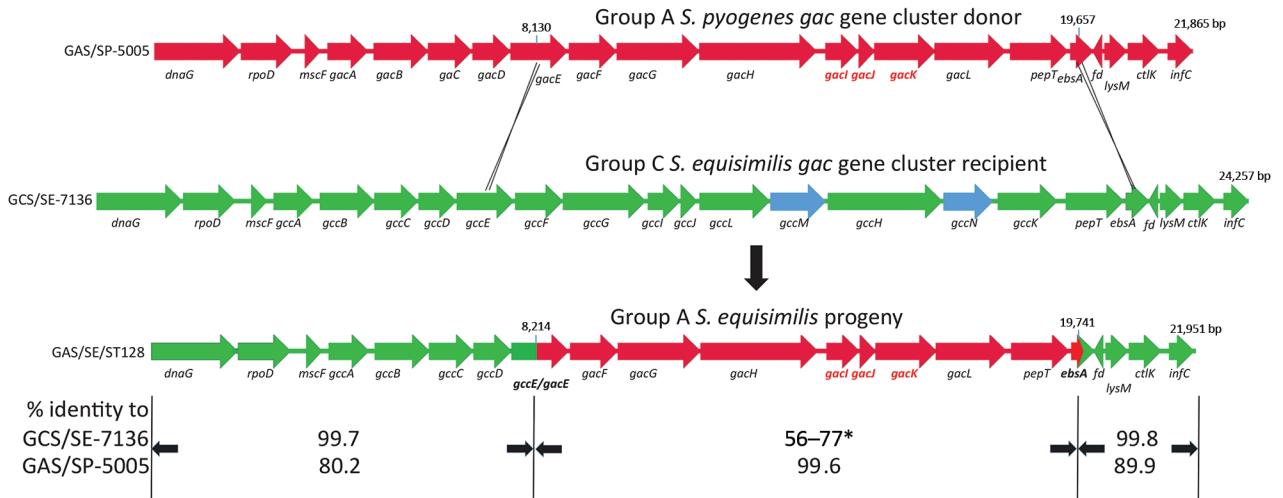
The Lancefield group A carbohydrate consists of a polyrihamnose chain with an immunodominant *N*-acetylglucosamine side chain (9) that functions in GAS pathogenesis (1). The group C carbohydrate also has a polyrihamnose backbone; however, its immunodominant side chain is the disaccharide *N*-acetylglucosaminosyl-*N*-acetylglucosamine (9). Genomic comparison of the 12 gene group A carbohydrate synthetic cluster *gacA-L* (1) from *S. pyogenes* with the corresponding regions of the 35 GAS/ST128/SE revealed an upstream crossover point within the *S. pyogenes* *gacE* ABC transporter gene and a downstream crossover point within *ebaA* (Figure 2). The ancestral recipient SE strain was implicated as group C *S. equisimilis* (GCS/SE) by virtue of the near-identical sequence of the 1,363-bp sequence within GAS/ST128/SE encompassing sections of *gacD* and *gacE* homologs (designated as *gccD* and *gccE*) (Figure 2) with GCS and the marked divergence of this 1,363-bp sequence from group G SE (data not shown). This sequence is immediately adjacent to the upstream crossover point shown between *S. pyogenes* and GCS/SE (SP-5005 and SE-7136; Figure 2). We also found these same crossover points within the group carbohydrate gene cluster of the available genomic sequence from the previously described invasive GAS/SE strain AC-2713 recovered in 1999 (3). Subsequent genomic analysis revealed AC-2713 to be ST128 and *emm* type *stG485.0*. Phylogenetic analysis revealed that AC-2713 differed by 126 single-nucleotide polymorphisms from a pair of genetically indistinguishable GAS/ST128/SE recovered within the East Bay area of San Francisco, California, USA (Figure 1). These 2 isolates were from recurrent invasive GAS infections within the same patient that occurred 1.5 months apart.

Comparison of the *S. pyogenes* *gacA-L* cluster with the corresponding *gcc* loci from group C SE strains (SE-7136; Figure 2) revealed that GCS/SE genes shared homology with all 12 *gacA-L* genes (56%–89% sequence identity). The weakest conservation was observed between the *gac/gccIJK* genes (56%–69% identity), consistent with the requirement of *gacIJK* for the group A immunodominant

Author affiliation: Centers for Disease Control and Prevention, Atlanta, Georgia, USA

DOI: <https://doi.org/10.3201/eid2508.181758>





**Figure 2.** Ancestral recombination event depicting *Streptococcus pyogenes* group A carbohydrate gene donor (GAS/SP-5005; GenBank accession no. NC007297), group C *S. dysgalactiae* subsp. *equisimilis* recipient (GCS/SE7136; GenBank accession no. NCTC7136), and progeny group A *S. dysgalactiae* subsp. *equisimilis* progeny (GAS/SE/ST128) described in study of emergent invasive group A *Streptococcus dysgalactiae* subspecies *equisimilis*, United States, 2015–2018. The deduced crossover points between the group A gene cluster (red) donor and group C (green) recipient strains are shown. The 3 genes required for inclusion of the immunodominant *N*-acetylglucosamine side chain within the group A carbohydrate (*gacI*, *gacJ*, and *gacK*) are shown in red. The coordinates of the fragment transferred that is highly conserved between the donor and the progeny are indicated. The length of the 3 genomic regions are indicated. The *gacE/gccE* and *ebsA* genes are shown as green/red hybrids. The extra *gcc* cluster genes not conserved within the *gac* cluster are shown in blue. The relative sequence identities of the 3 different regions of progeny (bottom) *gac* cluster genes with the group A *S. pyogenes* donor (top) and group C *S. equisimilis* recipient (middle) are indicated. The middle segment (asterisk) indicates a range of 56%–77% sequence identity between each of the 8 structural genes (*gacF*–*pepT*) that were received intact from the *S. pyogenes* donor. The *gac* cluster genes are described in more detail in van Sorge et al. (1). Gene assignments are as follows: *dnaG*, DNA primase; *rpoD*, major RNA polymerase sigma factor; *mscF*, metal sulfur complex assembly factor; *gacA*–*L*, group A carbohydrate biosynthetic genes (putative functions described in van Sorge et al. [1]); *gccA*–*N*, group C carbohydrate biosynthetic genes. *gccA*–*L* are functional homologs of *gacA*–*L*. *gccM* and *gccN* putatively encode an additional glycosyl transferase and UDP-monosaccharide 4-epimerase, respectively; *ebsA*, pore-forming protein; *fd*, ferredoxin (complement strand); *ctkK*, cytidylate kinase; *infC*, translation initiation factor IF-3.

*N*-acetylglucosamine side chain but not for synthesis of the polyrhamnose core (1). Two additional genes, designated *gccM* (glycotransferase gene) and *gccN* (UDP-monosaccharide epimerase gene), were evident within the *gcc* gene cluster. In the ancestral recombination event, an 11,527-bp GAS (*S. pyogenes*) chromosomal segment composed of the *gacE* 3' portion, along with the *gacF*–*L* genes and a 5J portion of *ebsA*, replaced the corresponding 13,813 bp of a GCS/ST128/SE strain, resulting in the recombinant GAS/ST128/SE lineage (Figure 2). This fragment encompasses the intact 7-gene *gacF*–*gacL* segment; each gene shared 99.4%–99.7% sequence identity with counterparts in *S. pyogenes*. The evident functionality of the hybrid *gac/gcc* cluster within the GAS/ST128/SE progeny lineage is consistent with identical roles of the first 3 genes of the cluster (*gac/gccA*–*C*) in the biosynthesis of the polyrhamnose core (1) that is present within the groups A, C, and G carbohydrates (9). Each of these 3 genes are also required for *S. pyogenes* viability (1).

The occurrence of multiple *emm* types within the same MLST is common in SE (5,6) and differs from *emm*/MLST

associations within *S. pyogenes*, where an MLST is nearly always definitive of a single *emm* type (2,10). The presence of 3 different *emm* types and 8 macrolide-resistant isolates within GAS/ST128/SE (Figure 1) is indicative of a long-standing successful lineage. A single isolate of this lineage was positive for the exotoxin gene *speC* (Figure 1) that was carried on a prophage highly similar to a previously described *speC*-positive *S. pyogenes* strain (8). The relative genomic positions of the prophages are exactly conserved between the 2 species, inserted within the *pepD* gene in the genomic region that lies between the bacterial cell division genes *ftsE/ftsX* and the ribosomal protein gene *rpsL31B* (Figure 1). The number of single-nucleotide polymorphism differences between individual GAS/ST128/SE core genomes ranged from 0 to 613 (Figure 1). The GAS/ST128/SE strain AC-2713 recovered 20 years ago (3) is also indicative of a long-established lineage.

The 34 GAS/ST128/SE isolates for which information was available (32 from blood, 1 from a joint, and 1 from a surgical wound) recovered in ABCs since January 1, 2015, were recovered from older adults (age range 22–93

years; mean age 63 years) from 8 ABCs sites; most (85%) patients were men. Most patients had underlying medical conditions (data not shown), including 16 with diabetes, 15 with cellulitis (including 1 who had necrotizing fasciitis), 8 with pneumonia, and 6 with septic shock. One patient with bacteremia died.

## Conclusions

ABCs identifies invasive infections caused by GAS without identification of isolates to the species level. Since 2015, when we implemented WGS as our primary platform for GAS characterization, we have identified rarely occurring non-*S. pyogenes* isolates through our bioinformatics pipeline automated MLST function rather than previously employed phenotypic testing. Of  $\approx 16,000$  GAS isolates recovered from ABCs during 1994–2014, only 11 had *emm* types characteristic of SE. All 11 were collected during 2011–2014 and were of the 3 *emm* types found among the 35 GAS/ST128/SE isolates from this study. Genomic analysis verified the GAS/ST128/SE lineage of these 11 older isolates (data not shown). Finding 35 additional invasive isolates of this lineage recovered during January 1, 2015–November 1, 2018, through ABCs suggests a level of expansion attributable to strain adaptation and fitness or to a more susceptible population. Attempts to identify circulating ST128/SE strains of the original group C have been unsuccessful, including an examination of a population-based sampling of SE (5).

Because group A SE is suspected to be rare, these findings raise the question of whether invasive disease attributable to SE of groups C, G, and L is also increasing. A 2-year population-based study of  $\beta$ -hemolytic streptococcal disease attributable to Lancefield groups other than A and B within 2 ABCs sites during 2002–2004 revealed that 80% of such isolates were SE (11), with clinical manifestations and targeted susceptible populations similar to *S. pyogenes*. Incidence of invasive disease attributable to non-group A SE during this period was estimated at 2.5 cases/100,000 population, similar to the incidence of GAS infections (2.89 cases/100,000 population) in these same 2 sites. The incidence of overall invasive GAS disease in the United States has also markedly increased during recent years, from 3.4 cases/100,000 population in 2012 to 7.2 cases/100,000 population in 2017 (<https://www.cdc.gov/abc/reports-findings/survreports/gas17.html>).

## Acknowledgments

We are indebted to all of the hospitals and laboratories participating in the ABCs component of the Emerging Infections Programs network, a collaboration of the CDC, state health departments, and universities. We thank the following persons from the following programs and institutions for their

contributions to the establishment and maintenance of the ABCs system: California Emerging Infections Program, J. Watt, A. Reingold, G. Rothrock, and M. Apostol; Colorado Emerging Infections Program, R. Herlihy, N. Alden, and H. Maguire; Connecticut Emerging Infections Program, S. Petit, M. Cartter, C. Marquez, and D. Noel; Georgia Emerging Infections Program, M. Farley, S. Thomas, A. Tunali, and D. Goodenough; Maryland Emerging Infections Program, L. Harrison, T. Carter, R. Hollick, K. Holmes, and R. Myers; Minnesota Emerging Infections Program, R. Lynfield, K. Como-Sabetti, R. Danila, and P. Vangnone; New Mexico Emerging Infections Program, C. Smelser, S. Torres, S. Khanlian, R. Mansmann, and K. Angeles; New York Emerging Infections Program, N. Bennett, A. Muse, L. Dufort, G. Dumyati, S. McGuire, and N. Spina; Oregon Emerging Infections Program, P. Cieslak, A. Thomas, and T. Poissant; Tennessee Emerging Infections Program, W. Schaffner, B. Barnes, K. Leib, and A. Elie; CDC, M. Arvay, J. Onukwube, O.M. Almendares, and H.R. Pham.

This study used the *S. dysgalactiae* subsp. *equisimilis* MLST website (<http://pubmlst.org/sdysgalactiae>) at the University of Oxford. The development of this site has been funded by the Wellcome Trust.

Major funding for this work was provided through support from CDC's Emerging Infection Program and the CDC Advanced Molecular Detection initiative.

## About the Author

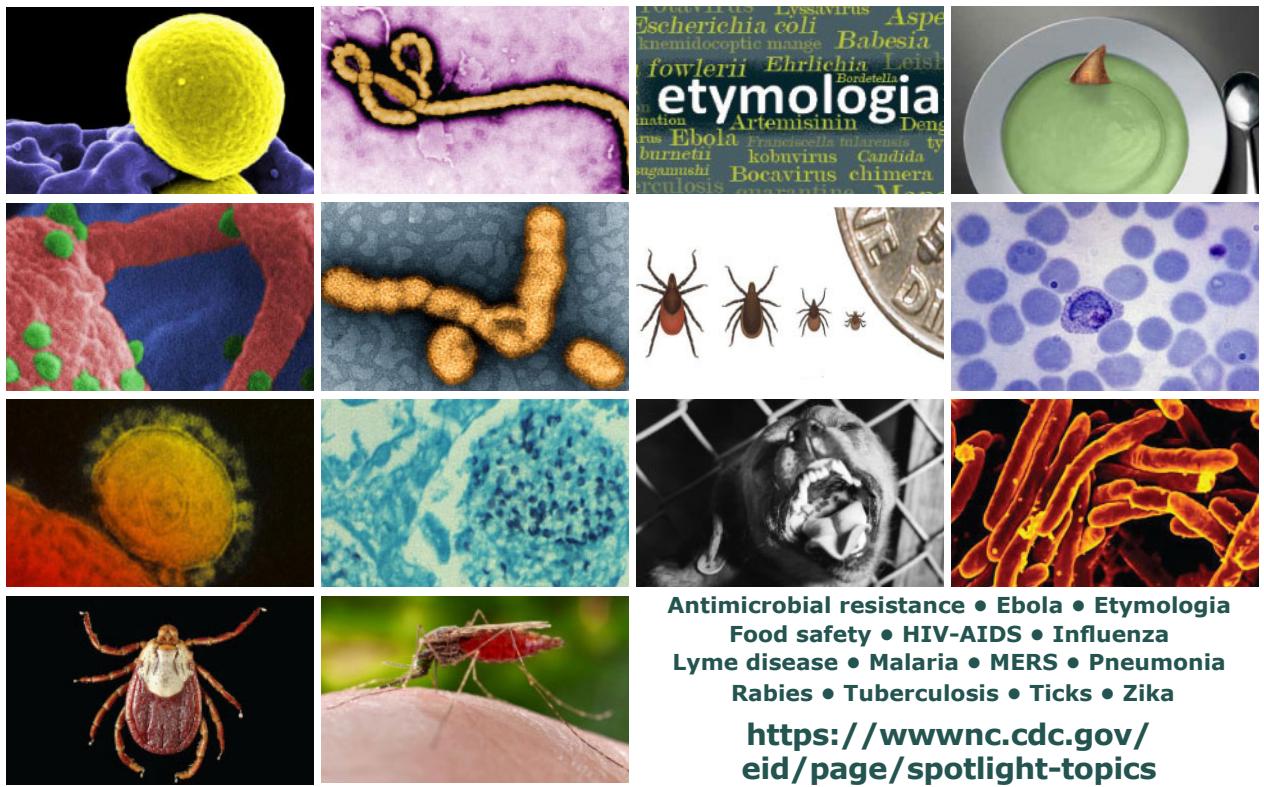
Dr. Chochua is a researcher in the Streptococcus Laboratory, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA. Her research interests are next-generation sequencing, characterization of clinical streptococcal isolates, antimicrobial resistance, genetic adaptations, and outbreak responses.

## References

- van Sorge NM, Cole JN, Kuipers K, Henningham A, Aziz RK, Kasirer-Friede A, et al. The classical Lancefield antigen of group A *Streptococcus* is a virulence determinant with implications for vaccine design. *Cell Host Microbe*. 2014;15:729–40. <http://dx.doi.org/10.1016/j.chom.2014.05.009>
- Chochua S, Metcalf BJ, Li Z, Rivers J, Mathis S, Jackson D, et al. Population and whole genome sequence based characterization of invasive group A streptococci recovered in the United States during 2015. *MBio*. 2017;8:e01422–17. <http://dx.doi.org/10.1128/mBio.01422-17>
- Brandt CM, Haase G, Schnitzler N, Zbinden R, Lütticken R. Characterization of blood culture isolates of *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *J Clin Microbiol*. 1999;37:4194–7.
- Tanaka D, Isobe J, Watahiki M, Nagai Y, Katsukawa C, Kawahara R, et al.; Working Group for Group A Streptococci in Japan. Genetic features of clinical isolates of *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *J Clin Microbiol*. 2008;46:1526–9. <http://dx.doi.org/10.1128/JCM.02188-07>

- Ahmad Y, Gertz RE Jr, Li Z, Sakota V, Broyles LN, Van Beneden C, et al. Genetic relationships deduced from *emm* and multilocus sequence typing of invasive *Streptococcus dysgalactiae* subsp. *equisimilis* and *S. canis* recovered from isolates collected in the United States. *J Clin Microbiol.* 2009; 47:2046–54. <http://dx.doi.org/10.1128/JCM.00246-09>
  - McMillan DJ, Bessen DE, Pinho M, Ford C, Hall GS, Melo-Cristino J, et al. Population genetics of *Streptococcus dysgalactiae* subspecies *equisimilis* reveals widely dispersed clones and extensive recombination. *PLoS One.* 2010;5:e11741. <http://dx.doi.org/10.1371/journal.pone.0011741>
  - Lancefield RC. The antigenic complex of *Streptococcus haemolyticus*: I. Demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. *J Exp Med.* 1928; 47:91–103. <http://dx.doi.org/10.1084/jem.47.1.91>
  - Walker MJ, Brouwer S, Forde BM, Worthing KA, McIntyre L, Sundac L, et al. Detection of epidemic scarlet fever group A *Streptococcus* in Australia. *Clin Infect Dis.* 2019; Epub ahead of print. <http://dx.doi.org/10.1093/cid/ciz099>
  - Coligan JE, Kindt TJ, Krause RM. Structure of the streptococcal groups A, A-variant and C carbohydrates. *Immunochemistry.* 1978;15:755–60. [http://dx.doi.org/10.1016/0161-5890\(78\)90105-0](http://dx.doi.org/10.1016/0161-5890(78)90105-0)
  - 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;69:2416–27. <http://dx.doi.org/10.1128/IAI.69.4.2416-2427.2001>
  - Broyles LN, Van Beneden C, Beall B, Facklam R, Shewmaker PL, Malpiedi P, et al. Population-based study of invasive disease due to beta-hemolytic streptococci of groups other than A and B. *Clin Infect Dis.* 2009;48:706–12. <http://dx.doi.org/10.1086/597035>
- Address for correspondence: Bernard Beall, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, Mailstop C02, Atlanta, GA 30329-4027, USA; email: [bbeall@cdc.gov](mailto:bbeall@cdc.gov)

# Emerging Infectious Diseases Spotlight Topics



**etymologia**

Antimicrobial resistance • Ebola • Etymologia  
Food safety • HIV-AIDS • Influenza  
Lyme disease • Malaria • MERS • Pneumonia  
Rabies • Tuberculosis • Ticks • Zika

<https://wwwnc.cdc.gov/eid/page/spotlight-topics>

EID's spotlight topics highlight the latest articles and information on emerging infectious disease topics in our global community.

# Emergent Invasive Group A *Streptococcus dysgalactiae* subsp. *equisimilis*, United States, 2015–2018

## Appendix

### Construction of Maximum-Likelihood Trees

We constructed maximum-likelihood trees (Figure 1, <https://wwwnc.cdc.gov/EID/article/25/8/18-1758-F1.htm>) by using kSNP3.0 (1) with a kmer size of 19 and used MEGA 7 for evolutionary analysis (2). We inferred the evolutionary history by using the maximum-likelihood method on the basis of the general time reversible model (3). The trees with the highest log likelihood (−124033.37 for left tree and −13853.47 for right tree) are shown. We obtained initial trees for the heuristic search automatically by applying neighbor-join and BioNJ algorithms to a matrix of pairwise distances estimated using the maximum composite likelihood approach and then selecting the topology with superior log likelihood value (2). Multilocus sequence type, *emm* type determination, and grouping traits were determined as described previously (4–6). By using the CDC *S. pyogenes* whole-genome sequencing bioinformatics pipeline (4), we found that all but 7 of the strains shown were positive for the *sfbl* fibronectin binding protein gene (data not shown). Three of the 7 *sfbl*-negative strains included AC2713 and the 2 identical EB1–5 isolates. All strains except for the 2 group L strains (left tree) were positive for the query for the *S. pyogenes nga* determinant encoding the extracellular NAD<sup>+</sup> glycohydrolase virulence factor (data not shown). As shown, 13 strains were positive for *erm*- or *mef*- encoded macrolide-resistance determinants, including a cluster of 8 GAS/ST128 isolates of *emm* subtype *stG245.0*. A single group *ermT*-positive group L strain had the ParC S79Y substitution associated with fluoroquinolone-resistance. In addition, a single group G/ST277 strain was positive for the pilus determinant T58. Single isolates were positive for the CDC pipeline queries for exotoxin genes *smeZ* (group L in left tree) and *speC* (GAS/ST128/*stG652.0*). All strains were β-hemolytic because of expression of the previously

characterized homologue of the *S. pyogenes* streptolysin S determinant (data not shown) (7). A total of 19,701 positions (for 49 isolates of left tree) and 2,108 (36 isolates of right tree) are included in the final datasets. The distance reference is indicative of 394 single nucleotide polymorphisms based on 19,701 positions within the 48 genomes (left tree) or 42 positions based upon 2,108 positions within 36 genomes (right tree). Genome fastQ accession numbers are provided for the 48 isolates (Appendix Table) and are associated with laboratory identifiers as well as the information provided from CDC *S. pyogenes* bioinformatics pipeline as described in Chochua et al. (4) and Appendix (Construction of Maximum-Likelihood Trees).

**Appendix Table.** SRA accession numbers, laboratory identifiers, and other features of *Streptococcus dysgalactiae* subsp. *equisimilis* study isolates

Biosample accession no.	LABID*	Group†	MLST	County, year‡	<i>emm</i> subtype§	Features detected from <i>S. pyogenes</i> WGS bioinformatics
						pipeline¶
SAMN09848964	20173110	A	128	NM2, 2016	stG652.0	<i>sfb1, nga</i>
SAMN09848983	20173524	A	128	NM2, 2017	stG652.0	<i>sfb1, nga</i>
SAMN07153383	20154682	A	128	NM2, 2015	stG652.0	<i>sfb1, nga</i>
SAMN08690859	20164183	A	128	NM1, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08691925	20171692	A	128	NM3, 2016	stG652.0	<i>sfb1, nga</i>
SAMN09849571	20182387	A	128	NM3, 2017	stG652.0	<i>sfb1, nga</i>
SAMN09849325	20181492	A	128	NM4, 2017	stG652.0	<i>sfb1, nga</i>
SAMN07153783	20156439	A	128	EB1, 2015	stG652.0	<i>sfb1, nga</i>
SAMN07154238	20162102	A	128	EB2, 2015	stG652.0	<i>sfb1, nga</i>
SAMN08691109	20165290	A	128	EB1, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08691691	20170764	A	128	NM3, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08691143	20165326	A	128	CA1, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08692346	20173686	A	128	EB2, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08691545	20170128	A	128	CO1, 2016	stG652.0	<i>sfb1, nga</i>
SAMN07154260	20162127	A	128	CA1, 2015	stG652.0	<i>sfb1, nga</i>
SAMN08691583	20170168	A	128	CO2, 2016	stG652.0	<i>sfb1, nga</i>
SAMN08691937	20171717	A	128	OR2, 2016	stG652.0	<i>sfb1, nga</i>
SAMN09849664	20182677	A	128	OR1, 2017	stG652.0	<i>sfb1, nga</i>
SAMN07153796	20156452	A	128	EB1, 2015	stG652.0	<i>sfb1, nga</i>
SAMN07154052	20160976	A	128	NY2, 2015	stG485.0	<i>sfb1, nga</i>
SAMN08690593	20162522	A	128	GA5, 2016	stG245.0	<i>sfb1, nga</i>
SAMN08691696	20170850	A	128	TN1, 2016	stG245.0	<i>sfb1, nga</i>
SAMN07154389	20163069	A	128	CT2, 2015	stG652.0	<i>sfb1, nga</i>
SAMN08691781	20171173	A	128	GA2, 2016	stG652.0	<i>sfb1, nga, speC</i>
SAMN09849027	20176966	A	128	CO3, 2017	stG485.0	<i>sfb1, nga</i>
SAMN08691278	20165952	A	128	TN2, 2016	stG245.0	<i>sfb1, nga, ermB</i>
SAMN07153678	20156157	A	128	NY1, 2015	stG245.0	<i>sfb1, nga, mef</i>
SAMN07153302	20154376	A	128	NY3, 2015	stG245.0	<i>sfb1, nga, mef</i>
SAMN07153560	20155412	A	128	NY1, 2015	stG245.0	<i>sfb1, nga, mef</i>
SAMN09848910	20163223	A	128	NY1, 2016	stG245.0	<i>sfb1, nga, mef</i>
SAMN08691739	20170914	A	128	NY1, 2016	stG245.0	<i>sfb1, nga, mef</i>
SAMN08691478	20166757	A	128	NY2, 2016	stG245.0	<i>sfb1, nga, mef</i>
SAMN10342858	20185475	A	128	NY2, 2018	stG245.0	<i>sfb1, nga, mef</i>
SAMN07154253	20162120	A	128	EB1, 2015	stG485.0	<i>nga</i>
SAMN07154273	20162140	A	128	EB1, 2015	stG485.0	<i>nga</i>
SAMN09848906	20161329	G	48	NM3, 2015	stG643.0	<i>nga, mef</i>
SAMN09849191	20180052	G	127	MD2, 2017	stG245.0	<i>nga, tetM</i>
SAMN10342859	20162531	G	35	GA5, 2016	<i>emm23.4</i>	<i>sfb1, nga, mef, tetM, cat</i>
SAMN10342860	20170556	G	277	OR2, 2016	<i>stG6.0</i>	<i>nga, tee58</i>
SAMN08692801	20175240	C	20	NM3, 2017	<i>stG62647.0</i>	<i>nga</i>
SAMN08691915	20171682	C	17	NM1, 2016	<i>stC74A.0</i>	<i>sfb1, nga, ermTR</i>
SAMN09848934	20170560	G	17	OR3, 2016	stG485.0	<i>sfb1, nga</i>
SAMN08691738	20170893	G	206	CT1, 2016	stG245.0	<i>sfb1, nga</i>
SAMN08690595	20162524	G	17	GA1, 2016	stG245.1	<i>sfb1, nga</i>
SAMN09849033	20177043	G	17	GA3, 2017	stG6792.0	<i>sfb1, nga</i>
SAMN10342861	20185314	G	470	OR2, 2018	stG840.0	<i>sfb1, nga</i>
SAMN10342862	20161734	L	1	MD1, 2016	stL1921.0	<i>sfb1, smeZ</i>
SAMN10342863	20163969	L	467	MN1, 2016	stL1929.1	<i>sfb1, ermT, ParC-S79Y</i>

\*These isolates are listed in same order (top to bottom) as in Figure 1, panel A (<https://wwwnc.cdc.gov/EID/article/25/8/18-1758-F1.htm>).

†The group A strains were positive for the group A carbohydrate cluster gene 1 (*gacI*). The group carbohydrate for these and the strains of groups C, G, and L were also determined serologically.

‡County and year of isolation. For example, NM2, 2016 refers to New Mexico county 2 and isolation during 2016.

§*emm* subtypes provided from CDC M protein gene database at [ftp://ftp.cdc.gov/pub/infectious\\_diseases/biotech/tsemm](ftp://ftp.cdc.gov/pub/infectious_diseases/biotech/tsemm).

¶Additional features obtained from CDC whole genome sequencing bioinformatics pipeline as described in Chochua et al. (4) and this Appendix (Construction of Maximum-Likelihood Trees).

## References

1. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics*. 2015;31:2877–8. [PubMed](#) <http://dx.doi.org/10.1093/bioinformatics/btv271>
2. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4. [PubMed](#) <http://dx.doi.org/10.1093/molbev/msw054>
3. Nei M, Kumar S. *Molecular evolution and phylogenetics*. New York: Oxford University Press; 2000. p. 150–75.
4. Chochua S, Metcalf BJ, Li Z, Rivers J, Mathis S, Jackson D, et al. Population and whole genome sequence based characterization of invasive group A streptococci recovered in the United States during 2015. *MBio*. 2017;8:e01422–17. [PubMed](#) <http://dx.doi.org/10.1128/mBio.01422-17>
5. Lancefield RC. The antigenic complex of *Streptococcus haemolyticus*: I. Demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. *J Exp Med*. 1928;47:91–103. [PubMed](#) <http://dx.doi.org/10.1084/jem.47.1.91>
6. Broyles LN, Van Beneden C, Beall B, Facklam R, Shewmaker PL, Malpiedi P, et al. Population-based study of invasive disease due to beta-hemolytic streptococci of groups other than A and B. *Clin Infect Dis*. 2009;48:706–12. [PubMed](#) <http://dx.doi.org/10.1086/597035>
7. Humar D, Datta V, Bast DJ, Beall B, De Azavedo JC, Nizet V. Streptolysin S and necrotising infections produced by group G streptococcus. *Lancet*. 2002;359:124–9. [PubMed](#) [http://dx.doi.org/10.1016/S0140-6736\(02\)07371-3](http://dx.doi.org/10.1016/S0140-6736(02)07371-3)