Remarkable Genome Stability among *emm1* Group A *Streptococcus* in Belgium over 19 Years

Jasmine Coppens[1], Basil Britto Xavier[1], Katherine Loens[2], Christine Lammens[1], Margareta Ieven[1,2], Veerle Matheeussen[2], Herman Goossens[1,2], and Surbhi Malhotra-Kumar[1,*]

[1]Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp, Belgium

[2]Belgian Reference Centre for Group A Streptococcus, Antwerp University Hospital, Antwerp, Belgium

*Corresponding author: E-mail: surbhi.malhotra@uantwerpen.be.

Accepted: April 23, 2019

Data deposition: This project has been deposited at NCBI under the accession PRJNA497706.

Abstract

During the last two decades, there has been a public health concern of severe invasive infections caused by Group A *Streptococcus* (GAS) of the *emm1* genotype. This study investigated the dynamics of *emm1* GAS during 1994–2013 in Belgium. *emm1* GAS isolated from blood, tissue, and wounds of patients with invasive infections ($n = 23$, S1–S23), and from patients with uncomplicated pharyngitis ($n = 15$, NS1–NS15) were subjected to whole-genome mapping (WGM; *kpn*) (Opgen). Whole-genome sequencing was performed on 25 strains (WGS; S1–S23 and NS6–NS7) (Illumina Inc.). Belgian GAS belonged to the M1T1 clone typified by the 36-kb chromosomal region encoding extracellular toxins, NAD+-glycohydrolase and streptolysin O. Strains from 1994–1999 clustered together with published strains (MGAS5005 and M1476). From 2001 onward, invasive GAS showed higher genomic divergence in the accessory genome and harbored on average 7% prophage content. Low evolutionary rate (2.49E-008; $P > 0.05$) was observed in this study, indicating a highly stable genome. The studied invasive and pharyngitis isolates were no genetically distinct populations based on the WGM and core genome phylogeny analyses. Two copies of the *speJ* superantigen were present in the 1999 and 2010 study strains ($n = 3$), one being chromosomal and one being truncated and associated with phage remnants.

This study showed that *emm1* GAS in Belgium, compared with Canada and UK M1 strains, were highly conserved by harboring a remarkable genome stability over a 19-year period with variations observed in the accessory genome.

Key words: WGS, speJ, GAS, WGM, cgMLST, wgMLST.

Introduction

The human pathogen Group A *Streptococcus* (GAS) causes infections ranging from pharyngitis to invasive infections, with the most severe manifestation being streptococcal toxic shock syndrome (STSS) (Carapetis et al. 2005). To cause this diversity of infections, GAS possesses a large number of virulence factors. Among these, the cell wall-associated M protein is a major antiphagocytic virulence factor (Bisno et al. 2003) and is widely used for genotyping GAS. The observed resurgence of severe invasive GAS infections over the past 30 years is correlated with a single, globally disseminated M protein-encoding gene *(emm)* 1 GAS subclone, M1T1, characterized by a 36 kb bacteriophage harboring virulence factors extracellular streptodornase D (*Sda1*; also known as streptococcal deoxyribonucleases), that aids the bacterial spread and deep tissue invasion, and immune stimulatory exotoxin type A (SpeA) (Sumby et al. 2005; Aziz and Kotb 2008; Cole et al. 2011). Another important virulence factor is the streptococcal inhibitor of complement (*sic*), which has been shown to interfere with the host-protective cytolytic complement cascade and neutralizes the effect of several antimicrobial proteins (Fernie-King et al. 2001) and is also used for *emm* GAS subtyping (Akesson et al. 1996).

Furthermore, the bacterial superantigens (Sags) are an important family of secreted virulence factors, including SpeA, SpeC, SpeG, SpeH, SpeI, SpeJ SpeK, SpeL, SpeM, Ssa, and SmeZ (Mccormick et al. 2001). These streptococcal pyrogenic exotoxins play an essential role in the pathogenesis of STSS. Most of them are phage encoded ($n = 8$; SpeA, SpeC, SpeH, SpeI, SpeK, SpeL, SpeM, and Ssa), but some are

chromosome-encoded ($n = 3$; SpeG, SpeJ, and SmeZ) (Nakagawa et al. 2003; Proft et al. 2003; Commons et al. 2008). Also, a higher acquisition of foreign DNA might be linked to clustered regularly interspaced short palindromic repeats (CRISPR)/Cas systems as it is the case in enterococci (Palmer and Gilmore 2010).

Commonly used techniques to study GAS clonality and outbreaks are multilocus sequence typing (MLST), and more recently whole-genome sequencing (WGS). In this study, we track the dynamics of *emm1* GAS population in Belgium over 19 years using genome-wide analysis by employing whole-genome mapping (WGM) and WGS.

Materials and Methods

Study Design and Strain Characterization

The study aimed to track the genomic evolution of *emm1* Belgian GAS strains and to understand the differences between invasive and noninvasive strains. GAS was recovered from patients with uncomplicated pharyngitis or invasive infections from laboratories and hospitals spread all over Belgium during 1994–2013. The reference center collected more than 12,976 GAS strains during this period. A total of 23 invasive GAS strains (labeled as S) were randomly selected and obtained from blood ($n = 13$), and tissue (wounds/throat swabs obtained from patients with skin and soft-tissue infections [SSTIs; $n = 3$] body sites or were from the undefined origin [$n = 4$]) (supplementary table 1, Supplementary Material online). Among these 23 patients with invasive infections, 47% ($n = 11$; S1, S2, S6, S13, S14, S15, S16, S17, S18, S20, and S22) had STSS. The samples were defined as STSS based on clinical diagnosis. Fifteen noninvasive strains (NS) were also randomly selected and cultured from throat swabs collected from patients with uncomplicated pharyngitis during 1999–2010. All strains were identified to the species level by MALDI-TOF MS (Bruker Daltonik GmbH, Bremen, Germany), and genotyped as *emm1* by polymerase chain reaction (PCR) (https://www.cdc.gov/streplab/groupa-strep/emm-background.html).

Antimicrobial Susceptibility Testing

Disc susceptibility testing was performed for erythromycin (15 μg), penicillin (10 μg), ciprofloxacin (5 μg), and clindamycin (2 μg) (Rosco Diagnostica, Denmark) and interpretation was done according to CLSI M100, 2018 (CLSI, 2018). The study strains were screened for the presence of macrolide resistance genes (*erm(A)*, *erm(B)*, and *mef(A/E)*) by multiplex PCR as described before (Malhotra-Kumar et al. 2005).

Whole-Genome Mapping

Thirty-eight strains were mapped on the Argus WGM (OpGen Inc., Gaithersburg) according to manufacturer's protocols. Briefly, GAS colonies grown overnight at 37 °C on blood agar plates were used for isolation of high molecular weight (HMW) DNA using Argus HMW DNA isolation kit. The assembly of restricted DNA molecules and identification of *KpnI* restriction sites was performed using MapManager software (OpGen Inc.). Visualization, editing (adjusting the starting point) and analysis of maps was performed by excluding fragment sizes smaller than 5 kb from the analysis and relative tolerance of 5%, combined with an absolute tolerance of 2,500 bp, along with the unweighted pair-group method with arithmetic mean (UPGMA) clustering with similarity coefficient cutoff of 95% and absolute number of unmatched fragments using BioNumerics v7.5 (Applied Maths, Belgium) as described earlier (Sabirova et al. 2016).

Whole-Genome Sequencing

WGS of 25 strains (invasive isolates in 1994–2013, $n = 23$; and noninvasive isolates in 2007 and 2008, $n = 2$) was performed using the NexteraXT sample preparation kit, 2×150 bp paired-end sequencing using Miseq (Illumina Inc.). Processed raw sequences were de novo assembled using SPAdes v3.12.0 (Bankevich et al. 2012). In-silico analyses were done for screening of CRISPR-*cas* and prophages using the online servers http://crispr.i2bc.paris-saclay.fr and www.phaster.ca, respectively (Grissa et al. 2007; Arndt et al. 2016). Comparative genome analysis and gene by gene analysis of the 25 sequenced strains was performed using 10 previously published *emm1* GAS strain sequences (supplementary table 1, Supplementary Material online).

To identify core genome single nucleotide polymorphism (SNP) differences and for phylogenetic analysis, all sequenced strains along with the reference *emm1* MGAS5005 were analyzed using Parsnp v1.2 (Treangen et al. 2014). The genetic relatedness analysis of 25 strains was based on concatenated core genome SNPs, and a total of 332 core genome SNP positions were identified, and a midpoint rooted tree was used. The evolutionary analysis was conducted using MEGA X (Kumar et al. 2016). The evolutionary history was inferred by using the Maximum Likelihood method and Kimura 2-parameter model, selected as the best model (Kimura 1980; Tao et al. 2018). The tree with the highest log-likelihood (−2127.94) was used. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining 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. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories [+*G*, parameter = 200.0000]). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 0.00% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Temporal Estimation (TempEst) v1.5.1 was used to conduct the linear

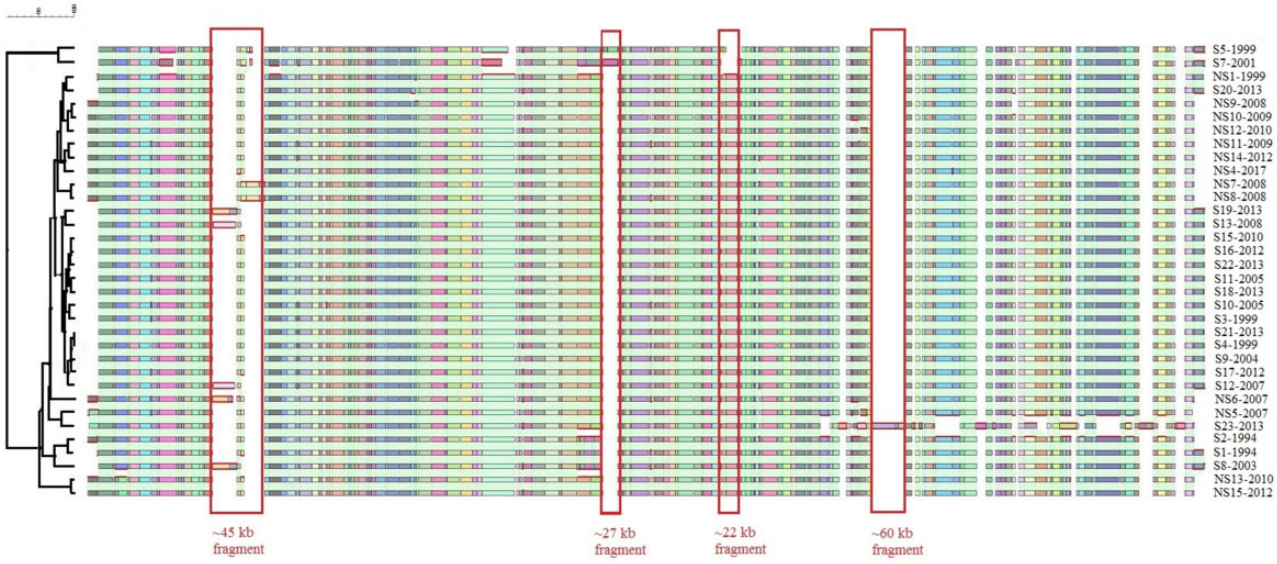

Fig. 1—Comparison of whole-genome maps of invasive and noninvasive *emm1* GAS from 1994 to 2013 WGMs using majority of UPGMA in BioNumerics. Green shaded areas indicate identical restriction patterns among the maps and red horizontal marks represent variations. A similarity coefficient of 95% was utilized, which assigned the strains to two clusters.

regression of maximum likelihood root-to-tip distances against the year of strain isolation (http://tree.bio.ed.ac.uk/software/tempest/).

The independent evolutionary rates among study strains over 19 years were determined using the branch length and the rate of correlation (Corrtest) (Tao et al. 2018).

Next, we generated core and whole-genome allelic profiles (scheme) of the study strains along with ten public available *emm1* strains, and compared the core genome MLST (cgMLST) and whole-genome MLST (wgMLST) allelic loci differences (ChewBBACA) and visualized these using PHYLOViZ (Ribeiro-Gonçalves et al. 2016; Silva et al. 2018). We utilized the adjusted Wallace coefficient to compare partitions and Simpson's diversity index to compare the discriminatory power of the different techniques utilized here (Carrico et al. 2006).

Results

emm1 GAS in Belgium Exhibit a Low Evolutionary Rate and a Highly Stable Core Genome

The antibiotic susceptibility pattern was similar among all tested strains. All strains were susceptible for penicillin, ciprofloxacin, and clindamycin. NS2 (1999), NS3 (2005), S14 (2009), and S6 (1999) were resistant to erythromycin and harbored *mef(A)*, whereas all the other strains were susceptible to erythromycin.

We also utilized WGM to further delve into the inter- and intra-strain variations with the focus on large genome rearrangements. However, WGMs of all 38 strains showed an overall similarity of 87% and these formed two clusters based on UPGMA. The strains S5 (1999) and S7 (2001) formed one cluster (C1), with an overall similarity of >91%, and the remaining strains clustered (C2) with an overall similarity of ~92% (fig. 1).

Further analysis by WGS on 25 strains identified that all strains are sequence type (ST) 28, which is the most dominant ST type isolated in humans. One strain (S13, 2008), harbored a nonsynonymous SNP in the *emm* gene (G352A), but this mutation was not part of hypervariable region (27–267 bp) of the *emm* typing. Because all sequenced strains belonged to the same ST type, we analyzed the *sic* gene for genetic diversity and identified 21 *sic* variants among the 25 strains, of which 16 were novel. Variations between *sic* alleles mainly mapped to the central region of the protein (from 34 to 126 AA) that resulted in different predicted protein lengths (282–353 AA) (supplementary fig. 1, Supplementary Material online). Mutations in the two-component regulatory system (covRS) which have been shown to strongly upregulate many virulence factors, including Sic, triggering the transition of M1T1 GAS from localized to systemic infection (Cole et al. 2011), were completely absent from the strains analyzed here.

The 25 sequenced strains showed an overall nucleotide similarity of 86% and compared with reference strain (MGAS5005; CP000017), 384 core genome SNPs were identified in the study strains. Compared with S1, the oldest strain (1994) in our collection here, study strains showed SNPs ranging from 30 to 58. Based on the core genome SNP phylogeny, these formed two major clusters: C1 with S2 (1994), S3

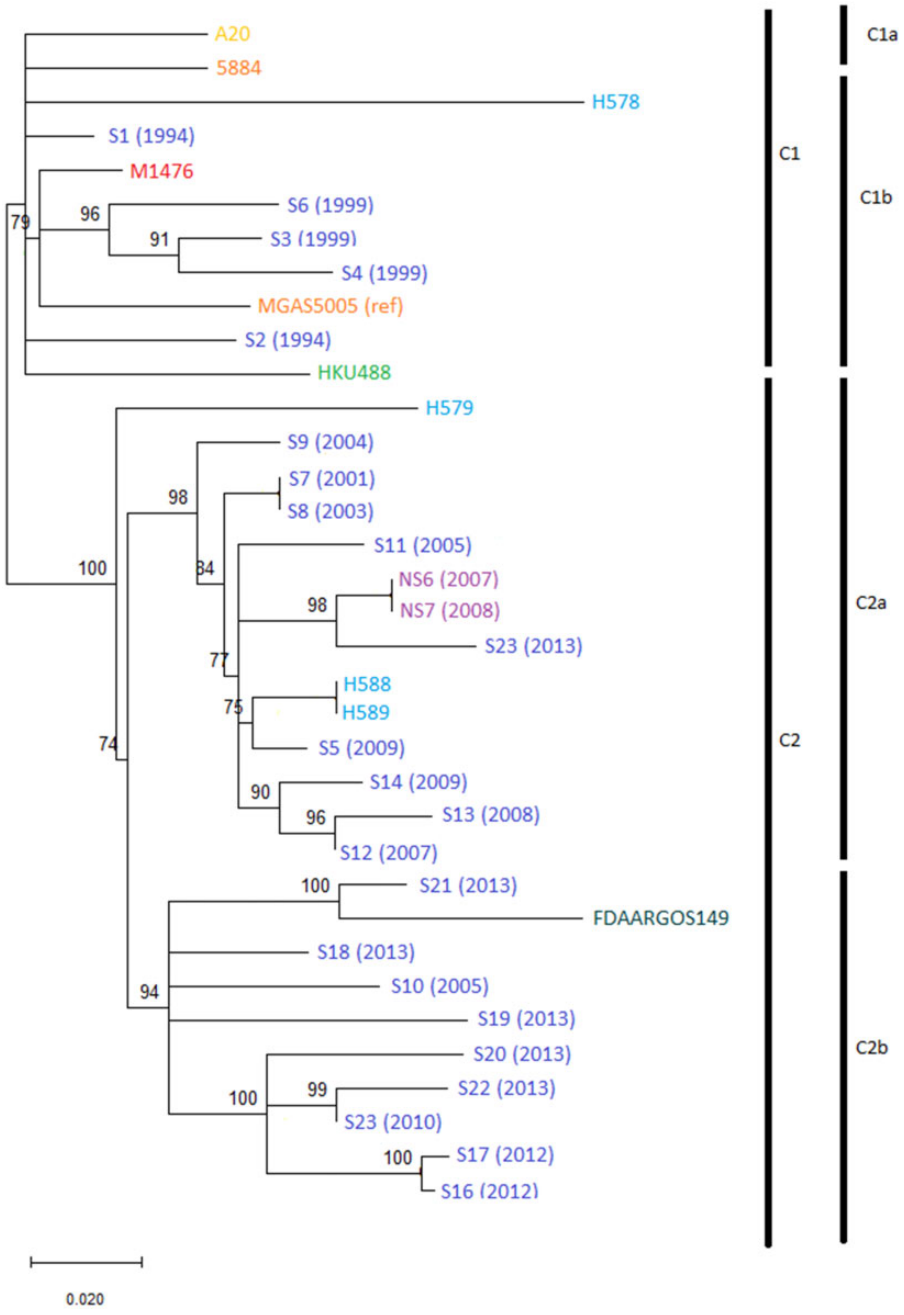

Fig. 2—Core genome SNP based phylogeny using ParSNP, including 25 *emm1* GAS strains and reference strains (MGAS5005, FDAARGOS149, HKU488, 5448, M1476, A20, H578, H579, H588, and H589). For the study strains, year of isolation is given and the color indicates the country of isolation, including dark blue: invasive study strains from Belgium; purple: noninvasive study strains from Belgium; light blue: United Kingdom; light green: Hong Kong; dark green: unknown; orange: Canada; and red: Japan. The bootstrap values (1,000 replicates) indicated on each node.

(1999), S4 (1999), and S6(1999), all isolated in the 1990s, and C2 with strains from 2001 to 2013. C1 showed a maximum of 33 within-cluster SNPs and clustered together with published strains MGAS5005, A20, 5448, H578, and M1476 (fig. 2). Although C2 showed a maximum of 27 SNP differences overall, based on bootstrap values, these formed two subclusters (C2a and C2b), of which subcluster C2a grouped with H588 and H589 and subcluster C2b with FDAARGOS-149 (fig. 2). Using these core genome-based phylogenetic analyses, we showed the genetic relatedness between strains isolated in the United Kingdom and Belgium. This finding showed the potential that transfers happened through people

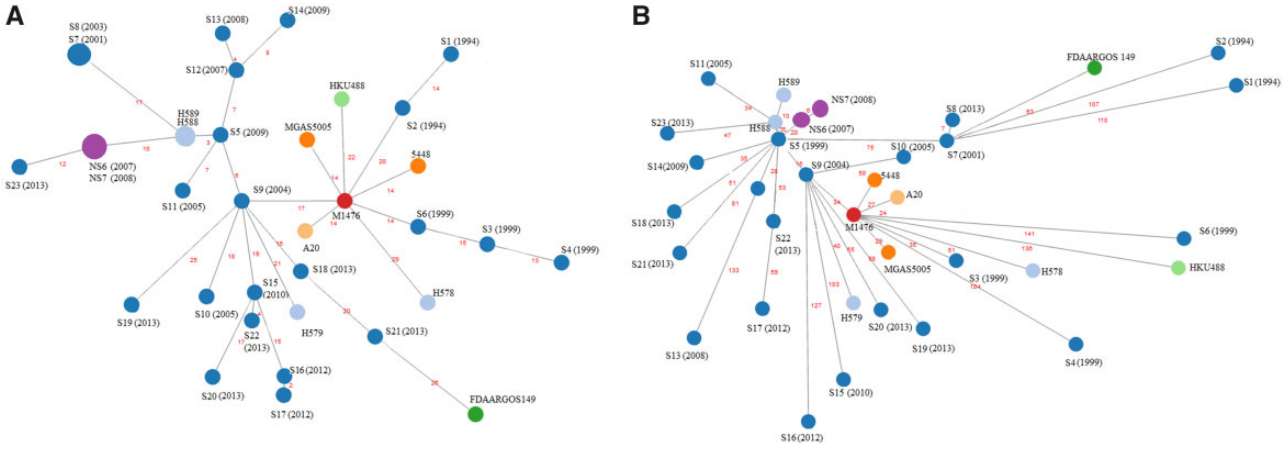

Fig. 3—cgMLST (*A*) and wgMLST (*B*) visualized using PHYLOViZ. All sequenced strains are included as well as several references strains (MGAS5005, FDAARGOS149, HKU488, 5448, M1476, A20, H578, H579, H588, and H589). For the study strains, year of isolation is given and the color indicates the country of isolation, including dark blue: invasive study strains from Belgium; purple: noninvasive study strains from Belgium; light blue: United Kingdom; light green: Hong Kong; dark green: unknown; orange: Canada; and red: Japan. The allelic differences are given in red.

transmissions although this is based on a study with a limited number of samples. This analyses further showed a year-based clustering of strains not only in 1999 (S3, S4, and S6) but also in 2012 (S16 and S17) (fig. 2). Also, the two noninvasive isolates from 2007 to 2008 (NS6–NS7) grouped in the C2 cluster, and were highly similar to each other. Interestingly, NS6–NS7 showed a higher similarity with an invasive strain from 2013 (S23, 6-SNP difference) than with those from 2007 to 2008 (S12–S13, 16-SNP differences) (fig. 2).

Next, we utilized a gene by gene approach to determine cgMLST and wgMLST allelic loci differences. In total 1,735 allelic loci were identified utilizing study and *emm1* reference strains from a public database. Of which 1,164 allelic loci were defined as the core genome, and utilized for the cgMLST scheme. For wgMLST, 1,164 allelic loci of the core genome and 571 of the accessory genome were used to create a wgMLST scheme. Maximum allelic differences among study strains (including reference strains) in the core genome were 29 (2.49%) (fig. 3*A*). Strains S1 and S2, both isolated in 1994, clustered closely together, and a similar trend was observed for strains S3, S4, and S6 (1999), and for S16 and S17 (2012) by cgMLST. The two noninvasive strains isolated from 2007 to 2008 (NS6–NS7) grouped closely with S23 from 2013, with a maximum of 12 allelic differences between the noninvasive 2007–2008 strains (NS6 and NS7) and S23. When comparing the number of SNPs and cgMLST allelic differences, the absolute number of differences was 58 and 29, respectively.

Core genome SNP acquisition over 19 years among the study strains was very low (2.18×10^{-4} core genome substitutions/site/year) and was in concordance with the independent evolutionary rates (Corrtest score of 2.49E-008, $P > 0.05$). Taken together, these data indicate a very low rate of core genome evolution in the GAS studied here.

Belgian GAS Belongs to the M1T1 Subclone and Fluctuation of the Accessory Genome Over Time Was Determined by wgMLST

Substitutions in the accessory genome were more frequent than those observed in the core genome of our study strains (Castillo-Ramírez et al. 2011). Of the 1,735 wgMLST allelic loci obtained from analyzed strains, 571 loci (32.9%) comprised the accessory genome (fig. 3*B*). The two noninvasive strains still clustered closely together with a maximum of six allelic differences. On the other hand, the 1994 strains S1 and S2 showed a maximum difference of 142 that is more diverse compared with the cgMLST. The same trend was observed in 1999 (S4 and S6: max. difference 174) and 2013.

The 36-kb region located between *purA* and *nadC* (MGAS5005 146,418–182,435 bp)—harboring the bacteriophage-encoded virulence factors Sda1 and SpeA—is a marker of the globally disseminated GAS M1T1 subclone (Nasser et al. 2014), and was present in all our sequenced strains. This is known to be a highly variable region as previous studies showed that 64% of the genome-wide SNPs in the M1T1 genome were localized in this 36-kb region (Nasser et al. 2014). Although the per-base-pair SNP acquisition rate in this region was indeed higher in our strains with the SNPs in this region ranging from 7 to 30, these made up only about 10% of the genome-wide SNPs ($n = 279$) observed in our study strains.

The accessory genome, comprising of prophages and its remnants, accounted for an average of 7% of the entire genome among the study strains ($n = 25$). Strain S15 (2010)

showed the lowest prophage content (5%) and S11 the highest (8.7%). Interestingly, prophage Φ315 and P9 were identified in all invasive GAS strains and in the two noninvasive strains sequenced from 2007 (NS6) to 2008 (NS7) (supplementary table 2A, Supplementary Material online). Also, the accessory elements tend to vary between the strains causing similar invasive diseases.

In total four different large genomic indels (~22–60 kb) were identified compared with the oldest strain (S1, 1994) (fig. 1). The invasive S8 (2003), S12 (2007), S13 (2008), and S19 (2013), and noninvasive NS6–NS8 (2007–2008) strains all showed presence of a ~45 kb *Streptococcus* phage 2,096. In-silico analysis of the WGMs also determined a ~22 kb deletion comprising of genes associated with lactose metabolism and multicopper oxidase pathway (*lacA, lacB*, *copA*, and *copZ*) as well as acetyltransferase and phosphotransferase (PTS) systems, and a galactose-specific IIA component that were lost in S5 (1999) and S7 (2001).

Diversity of Mobile Elements Harboring Superantigens and Presence of CRISPR-Cas Elements in M1T1 GAS

All sequenced invasive GAS strains harbored *speA*, *speG*, and *smeZ* (supplementary table 2B, Supplementary Material online). *speJ* (396,893–397,591 bp in S6, 1999) was detected in all invasive GAS except S17 (2012). In addition, several strains (S5, 1999; S6, 1999; S15, 2010) also harbored a truncated *speJ* (718,291–718,500 bp in S6, 1999) flanked by phage remnants.

Lastly, screening of CRISPR/Cas elements in the study and reference strains found a similar number of CRISPR loci (two confirmed and three questionable CRISPR elements) in the study M1T1 GAS, MGAS5005, SF370, and M1 476.

Discussion

emm1 GAS Isolated in Belgium Have a Highly Stable and Conserved Genome

We studied the genomic evolution of *emm1* GAS causing infections in Belgium during 1994–2013. Based on core genome SNPs, cgMLST, wgMLST, and evolutionary rate analyses, we show a very low rate of genomic evolution of *emm1* GAS strains in Belgium over nearly two decades. This low rate is very well demonstrated by the low number of SNP acquisitions noted in our strains in the 36-kb region which typifies these strains as M1T1 and which has been shown to be one of the highly variable regions in the *emm1* genome. The high relatedness among GAS strains made it impossible to demonstrate the discriminatory power of the different typing methods utilized here, and thus difficult to identify the optimal typing method for GAS.

The invasive and noninvasive isolates analyzed here were genetically similar populations based on the core genome and WGM analyses. Although recent studies have highlighted differences between invasive and noninvasive strains in the accessory genome, primarily in exotoxins and superantigens encoding genes (Unnikrishnan et al. 2002; Fernandes et al. 2017), this study did not find any gene-based tropism among invasive GAS causing specific infections or between invasive and noninvasive GAS.

Nonetheless, we did identify genomic changes over time within the M1T1 population. The earlier invasive GAS, isolated in 1994 and 1999, clustered together and with several *emm1* reference strains that we utilized, including MGAS5005, which represent the globally disseminated M1T1 clone (Sumby et al. 2005). From 2001 onward, the invasive strains sequenced here showed higher diversity primarily due to variations in the accessory genome as observed with the wgMLST data. Similar studies in *S*taphylococcus *aureus* have shown that gene gains and losses in the accessory genome have been of paramount importance for microevolutionary scales (decades) (Graña-Miraglia et al. 2017; Frisch et al. 2018). Changes in our strains were mostly due to acquisition of mobile elements such as transposons and prophages; the 1994–1999 strains showed a prophage content of 6.76% (range: 5.14–7.49%) of the entire genome that increased to an average 7.67% (range: 6.43–8.93%) in strains isolated from 2001 onward. Despite this increase, the prophage content of the Belgian M1T1 strains was much lower than that reported by previous studies (~12%) (Banks et al. 2002; Beres et al. 2002; Brüssow et al. 2004). This prompted us to analyze our strains for the presence of CRISPR-Cas elements, which are known to defense against plasmids and phage acquisition (Palmer and Gilmore 2010). However, despite the increase in mobile elements found in our strains from 2001 onward, the number of CRISPR elements identified remained stable.

Interestingly, we found a second copy of the *speJ* superantigen, being truncated, and associated with phage remnants. Whether *speJ* might have transferred historically into the *emm1* GAS genome via a temperate phage that was later lost has been a matter of debate (Proft and Fraser 2016), with a study showing a potential association of *speJ* with a bacteriophage and other studies showing it to be chromosomal (Friães et al. 2013; Proft and Fraser 2016).

In conclusion, this is one of the few studies tracking the long-term dynamics of *emm1* GAS isolated in the same geographical region. Although the number of strains was limited, we highlight the unique characteristics of the Belgian M1T1 subclone that differs remarkably from the M1T1 GAS observed in the other countries, including Canada and the United Kingdom, as a high level of genomic conservation and the low evolutionary rates were observed in this study.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank The Belgian National Reference Center for invasive GAS for providing strains, Sabine Chapelle for excellent technical assistance and Liesbet Van Heirstraeten for proofreading the manuscript. The National Reference Center is partially supported by the Belgian Ministry of Social affairs through a fund within the Health Insurance System. This work was partly financially supported by EU grant PREPARE (Platform foR European Preparedness Against (Re-)Emerging Epidemics, EU-FP7, # 602525), and by OpGen Inc., Gaithersburg, the United States in the framework of the European Public Health Initiative (EUPHi). J.C. was supported by the University of Antwerp doctoral assistant funds. B.B.X was supported by University of Antwerp Research funds (BOF-DOCPRO 2012-27450).

Author Contributions

J.C. carried out the MLST, optical mapping experiments, participated in the sequence analyses, and drafted the manuscript. B.B.X. performed bioinformatic analysis and helped to draft the manuscript. K.L., C.L., M.I., V.M., H.G., and S.M.K. conceived the study and revised the manuscript. All authors read and approved the final manuscript.

Literature Cited

Akesson P, Sjöholm AG, Björck L. 1996. Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. J Biol Chem. 271:1081–1088.

Arndt D, et al. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44(W1):W16–W21.

Aziz RK, Kotb M. 2008. Rise and persistence of global M1T1 clone of *Streptococcus pyogenes*. Emerg Infect Dis. 14(10):1511–1517.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Banks DJ, Beres SB, Musser JM. 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. Trends Microbiol. 10(11):515–521.

Beres SB, et al. 2002. Genome sequence of a serotype M3 strain of group A *Streptococcus*: phage-encoded toxins, the high-virulence phenotype, and clone emergence. Proc Natl Acad Sci U S A. 99(15):10078–10083.

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

Brüssow H, Canchaya C, Hardt W. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev. 68(3):560–602.

Canton R, Mazzariol A, Morosini M-I, Baquero F, Cornaglia G. 2005. Telithromycin activity is reduced by efflux in *Streptococcus pyogenes*. J Antimicrob Chemother. 55:489–495.

Carapetis JR, Steer AC, Mulholland EK, Weber M. 2005. The global burden of Group A *Streptococcus* disease. Lancet Infect Dis. 5(11):685–694.

Carrico JA, et al. 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol. 44(7):2524–2532.

Castillo-Ramírez S, et al. 2011. The impact of recombination on d*N*/d*S* within recently emerged bacterial clones. PLoS Pathog. 7(7):e1002129.

CLSI. 2018. CLSI document M100. Performance standards for antimicrobial susceptibility testing. 28th ed. Wayne (PA): Clinical and Laboratory Standards Institute.

Cole JN, Barnett TC, Nizet V, Walker MJ. 2011. Molecular insight into invasive group A streptococcal disease. Nat Rev Microbiol. 9(10):724–736.

Commons R, et al. 2008. Superantigen genes in group A streptococcal isolates and their relationship with emm types. J Med Microbiol. October:1238–1246.

Fernandes GR, et al. 2017. Genomic comparison among lethal invasive strains of *Streptococcus pyogenes* serotype M1. Front Microbiol. 23;8:1993.

Fernie-King BA, et al. 2001. Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes. Immunology 103(3):390–398. http://www.ncbi.nlm.nih.gov/pubmed/11454069.

Friães A, Pinto FR, Silva-Costa C, Ramirez M, Melo-Cristino J. 2013. Superantigen gene complement of *Streptococcus pyogenes*—relationship with other typing methods and short-term stability. Eur J Clin Microbiol Infect Dis. 32(1):115–125.

Frisch MB, et al. 2018. Invasive methicillin-resistant *Staphylococcus aureus* USA500 strains from the U.S. emerging infections program constitute three geographically distinct lineages. mSphere 3.

Graña-Miraglia L, et al. 2017. Rapid gene turnover as a significant source of genetic variation in a recently seeded population of a healthcare-associated pathogen. Front Microbiol. 8:1817.

Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35(Web Server):W52–W57.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16(2):111–120.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 33(7):1870–1874.

Malhotra-Kumar S, Lammens C, Piessens J, Goossens H. 2005. Multiplex PCR for simultaneous detection of macrolide and tetracycline resistance determinants in *Streptococci*. Antimicrob Agents Chemother. 49(11):4798–4899.

Mccormick JK, Yarwood JM, Schlievert PM. 2001. Toxic shock syndrome and bacterial superantigens: an update. Annu Rev Microbiol. 55:77–104.

Nakagawa I, et al. 2003. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. Genome Res. 13:1042–1055.

Nasser W, et al. 2014. Evolutionary pathway to increased virulence and epidemic group A *Streptococcus* disease derived from 3,615 genome sequences. Proc Natl Acad Sci U S A. 111(17):E1768–E1776.

Palmer KL, Gilmore MS. 2010. Multidrug-resistant enterococci lack CRISPR-cas. MBio 1:e00227.

Proft T, Fraser JD. 2016. Streptococcal superantigens: biological properties and potential role in disease NBK 333435.

Proft T, Sriskandan S, Yang L, Fraser JD. 2003. Superantigens and streptococcal toxic shock syndrome. Emerg Infect Dis. 9(10):1211–1218.

Ribeiro-Gonçalves B, Francisco AP, Vaz C, Ramirez M, Carriço JA. 2016. PHYLOViZ online: web-based tool for visualization, phylogenetic

inference, analysis and sharing of minimum spanning trees. Nucleic Acids Res. 44(W1):W246–W251.
Sabirova JS, et al. 2016. Whole-genome typing and characterization of blaVIM19-harbouring ST383 *Klebsiella pneumoniae* by PFGE, whole-genome mapping and WGS. J Antimicrob Chemother. 71(6):1501–1509.
Silva M, et al. 2018. chewBBACA: a complete suite for gene-by-gene schema creation and strain identification. Microb Genom. 4 (3).
Sumby P, et al. 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 Group A *Streptococcus* involved multiple horizontal gene transfer events. J Infect Dis. 192(5):771–782.
Tao Q, Tamuar K, Battistuzzi F, Kumar S. 2018. Pervasive correlation of molecular evolutionary rates in the tree of life. bioRxiv. 346635.
Tenover FC, et al. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol. 33(9):2233–2239.
Treangen TJ, Ondov BD, Koren S, Phillippy AM. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 15(11):524.
Unnikrishnan M, et al. 2002. The bacterial superantigen streptococcal mitogenic exotoxin Z is the major immunoactive agent of *Streptococcus pyogenes*. J Immunol. 169(5):2561–2569.

Associate editor: Esperanza Martinez-Romero