JCM
Journals.ASM.org

# Matrix-Assisted Laser Desorption Ionization–Time of Flight and Comparative Genomic Analysis of M-18 Group A *Streptococcus* Strains Associated with an Acute Rheumatic Fever Outbreak in Northeast Italy in 2012 and 2013

**Paolo Gaibani,[a] Erika Scaltriti,[b] Claudio Foschi,[a] Enrico Baggio,[a] Maria Vittoria Tamburini,[a] Roberta Creti,[c] Maria Grazia Pascucci,[d] Marco Fagioni,[e] Simone Ambretti,[a] Francesco Comandatore,[f] Stefano Pongolini,[b] Maria Paola Landini[a]**

Operative Unit of Clinical Microbiology, St. Orsola–Malpighi University Hospital, Bologna, Italy[a]; Sezione Diagnostica di Parma, Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Parma, Italy[b]; Dipartimento di Malattie Infettive, Parassitarie ed Immunomediate, Istituto Superiore di Sanità, Rome, Italy[c]; Public Health Authority Emilia Romagna, Rome, Italy[d]; Bruker Daltonics s.r.l., Macerata, Italy[e]; Dipartimento di Scienze Veterinarie e Sanità Pubblica, Università degli Studi di Milano, Milan, Italy[f]

**Acute rheumatic fever (ARF) is a postsuppurative sequela caused by *Streptococcus pyogenes* infections affecting school-age children. We describe here the occurrence of an ARF outbreak that occurred in Bologna province, northeastern Italy, between November 2012 and May 2013. Molecular analysis revealed that ARF-related group A *Streptococcus* (GAS) strains belonged to the M-18 serotype, including subtypes *emm18.29* and *emm18.32*. All M-18 GAS strains shared the same antigenic profile, including SpeA, SpeB, SpeC, SpeL, SpeM, and SmeZ. Matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) analysis revealed that M-18 GAS strains grouped separately from other serotypes, suggesting a different *S. pyogenes* lineage. Single nucleotide polymorphisms and phylogenetic analysis based on whole-genome sequencing showed that *emm18.29* and *emm18.32* GAS strains clustered in two distinct groups, highlighting genetic variations between these subtypes. Comparative analysis revealed a similar genome architecture between *emm18.29* and *emm18.32* strains that differed from noninvasive *emm18.0* strains. The major sources of differences between M-18 genomes were attributable to the prophage elements. Prophage regions contained several virulence factors that could have contributed to the pathogenic potential of *emm18.29* and *emm18.32* strains. Notably, phage ΦSPBO.1 carried erythrogenic toxin A gene (*speA1*) in six ARF-related M-18 GAS strains but not in *emm18.0* strains. In addition, a phage-encoded hyaluronidase gene (*hylP.2*) presented different variants among M-18 GAS strains by showing internal deletions located in the α-helical and TSβH regions. In conclusion, our study yielded insights into the genome structure of M-18 GAS strains responsible for the ARF outbreak in Italy, thus expanding our knowledge of this serotype.**

*Streptococcus pyogenes*, group A streptococcus (GAS), is a Gram-positive bacterium responsible for a wide spectrum of diseases ranging from moderate or mild infections to severe invasive diseases such as necrotizing fasciitis and toxic shock-like syndrome (TSLS). Several GAS infections can cause severe postinfectious sequelae, including acute poststreptococcal glomerulonephritis, acute rheumatic fever (ARF), and rheumatic heart disease (1).

ARF is a systemic disorder resulting from an autoimmune disease following a GAS infection that usually occurs in children between 5 and 15 years of age (2). During the last several decades, the incidence of ARF cases has significantly declined in the United States and Western Europe, whereas it remains high in Eastern Europe, Asia, and Australia (3). However, the resurgence of ARF in several geographical areas, including United States, is a matter of concern (4).

*S. pyogenes* possesses different virulence factors such as the M protein and superantigens (SAgs) that contribute to the pathogenesis of GAS infection (1). On the basis of the high variability of the M protein among GAS strains, the 5′-terminal sequence of the *emm* gene (*emm* typing) is considered a reliable molecular marker commonly used for epidemiological studies (5).

Previous studies indicated that *emm* types 1, 3, 5, 6, 18, 19, 24, and 29 have been isolated from ARF cases, suggesting a "rheumatogenic" role of certain serotypes (2). Epidemiological study of different ARF outbreaks in the United States revealed a strict association with serotype M-18 GAS (6).

Recently, matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry (MS) has been introduced in microbiological laboratories for prompt highly accurate identification and classification of bacterial species (7). Several studies have now demonstrated the ability of MALDI-TOF MS to

**Received** 10 December 2014 **Returned for modification** 12 January 2015
**Accepted** 20 February 2015

**Accepted manuscript posted online** 4 March 2015

**Citation** Gaibani P, Scaltriti E, Foschi C, Baggio E, Tamburini MV, Creti R, Pascucci MG, Fagioni M, Ambretti S, Comandatore F, Pongolini S, Landini MP. 2015. Matrix-assisted laser desorption ionization–time of flight and comparative genomic analysis of M-18 group A *Streptococcus* strains associated with an acute rheumatic fever outbreak in Northeast Italy in 2012 and 2013. J Clin Microbiol 53:1562–1572. doi:10.1128/JCM.03465-14.

**Editor:** S. S. Richter

Address correspondence to Paolo Gaibani, paolo.gaibani@unibo.it.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JCM.03465-14.



doi:10.1128/JCM.03465-14

**TABLE 1** Superantigens and *emm* type of GAS strains collected during an ARF outbreak

| *emm* subtype | No. of isolates expressing an SAg gene | | | | | | | | | | | Antigenic profile | No. of isolates |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *speA* | *speC* | *speG* | *speH* | *speI* | *speJ* | *speK* | *speL* | *speM* | *ssa* | *smeZ* | | |
| *6.4* | | | 32 | | | | 32 | | | | 32 | 1 | 32 |
| *28.0* | | | 9 | | | 9 | | | | | 9 | 2 | 9 |
| *18.29*[a] | 6 | 6 | 6 | | | | | 6 | 6 | | 6 | 3 | 6 |
| *18.32* | 3 | 3 | 3 | | | | | 3 | 3 | | 3 | 3 | 3 |
| *5.3* | | | 4 | | | | | | | | 4 | 4 | 4 |
| *5.6* | | | 1 | | | | | | | | 1 | 4 | 1 |
| *5.18* | | | 1 | | | | | | | | 1 | 4 | 1 |
| *1.0* | 3 | | 4 | | | 4 | | | | | 4 | 5 | 4 |
| *44.0* | | | 2 | | | 2 | | 2 | 2 | | 2 | 6 | 2 |
| *89.0* | | | 3 | | | | | | | | 3 | 4 | 3 |
| *3.88* | 1 | | 1 | | | | | | | | 1 | 7 | 1 |
| *3.1* | | | 1 | | | | 1 | | | | 1 | 1 | 1 |
| *9.0* | | | 1 | | | | | | | | 1 | 4 | 1 |
| *102.3* | | | 1 | | | | | | | | 1 | 4 | 1 |
| *12.0* | | | 1 | 1 | 1 | | 1 | | | | 1 | 8 | 1 |
| Total | 13 | 9 | 70 | 1 | 1 | 15 | 34 | 11 | 11 | | 70 | | 70 |

[a] Two consecutive GAS strains were isolated from the first ARF case.

type and distinguish a wide range of bacterial species at subspecies or strain level (8, 9).

On the other hand, whole-genome sequencing is a consolidated procedure for epidemiological and evolutionary purposes (10–12). This technique is highly sensitive and can identify single nucleotide polymorphisms (SNPs) throughout the genome (13). In particular, the resolution of strains genetically indistinguishable by other molecular techniques (i.e., multilocus sequence typing [MLST] and pulsed-field gel electrophoresis) made whole-genome sequencing technology a powerful tool for the epidemiological investigations of related high clonal bacterial isolates (12, 13).

We report here an epidemiological investigation based on both whole-genome sequencing and MALDI-TOF MS on serotype M-18 GAS strains collected from primary school-age children in Bologna province during an ARF outbreak in early 2013.

## MATERIALS AND METHODS

**Epidemiological investigation.** In February 2013, a notification of an ARF case following hospital admission was reported in an 11-year-old otherwise healthy Caucasian boy resident in Bologna province, Emilia-Romagna region. In this area, six ARF cases had been diagnosed in the previous 3 months. After the last notification of ARF diagnosis, the Regional Health Agency instituted active epidemiological surveillance to monitor all potential ARF contact cases following World Health Organization (WHO) criteria (14). The active surveillance protocol required both clinical evaluation and culture screening of all classmates of ARF cases. At the same time, 14 GAS strains isolated from school-age children with pharyngotonsillitis in the Bologna metropolitan area were collected. Two months later, diagnosis of ARF was notified in a 4-year-old Caucasian boy resident in the same province. Subsequently, 14 GAS-positive samples were collected from classmates of the second ARF case and from 34 symptomatic children resident in the same area. The date of isolation, mucoid trait, *emm* type, antimicrobial resistance, superantigen genes, and epidemiological linkage for each strain are listed in Table S1 in the supplemental material.

**Bacterial isolation and identification.** 70 GAS strains were isolated from throat swabs collected at the bacteriology laboratory of St. Orsola–Malpighi Hospital, Bologna, except for two 30-year-old *emm18* GAS strains that had been deposited in the bacterial collection bank of Istituto Superiore di Sanità (ISS). Bacteria were initially identified using standard methods and confirmed by MALDI-TOF 3.1 RTC (Bruker Daltonics, GmbH, Germany) according to the manufacturer's instructions. Antimicrobial susceptibility to penicillin, ampicillin, tetracycline, chloramphenicol, erythromycin, and clindamycin was tested by MicroScan semiautomated system (Siemens, Germany), and the results were interpreted according to EUCAST criteria (15). GAS isolates were examined for the presence of a mucoid phenotype by culture visualization and categorized as either mucoid or nonmucoid.

***emm* typing and SAg genes.** Genomic DNA from 70 GAS strains were extracted from pure cell bacterial culture by using a manual DNeasy Blood & Tissue kit (Qiagen, Basel, Switzerland) according to the manufacturer's protocol. PCR amplification of the *emm* gene was performed as previously described (16). In order to assign the specific *emm* type and subtype, the first 240 nucleotides of each sequence were compared to the *S. pyogenes emm* database available at the CDC website (http://www.cdc.gov/streplab/index.html). SAg genes were analyzed by multiplex PCR assays, as previously described (17). The exotoxin genes *speB* and *speF* were used as PCR internal controls. The presence of the SAg genes (*speA*, *speC*, *speG*, *speH*, *speI*, *speJ*, *speK*, *speL*, *speM*, *ssa*, and *smeZ*) was confirmed by single PCRs (Table 1).

**MLST.** To determine the genetic relationship between the 11 *S. pyogenes* isolates belonging to the serotype M-18, multilocus sequence typing (MLST) based on seven housekeeping genes (*gki*, *gtr*, *murI*, *mutS*, *recP*, *xpt*, and *yiqL*) was performed (18). The allele numbers and relative sequence types were assigned by using the *S. pyogenes* MLST database (http://spyogenes.mlst.net).

**MALDI-TOF MS sample preparation and analysis.** Sample preparation for MALDI-TOF MS was performed as previously described with minor modifications (8). Briefly, colonies of fresh overnight culture derived from 49 GAS isolates were resuspended at 1 McFarland, and 1 ml of bacterial suspension was centrifuged at 5,000 × *g* for 5 min. Pellets were suspended with 300 μl of distilled water and 900 μl of absolute ethanol and pelleted again. The supernatants were then discharged, and cells were suspended in 20 μl of formic acid (70%) and 20 μl of acetonitrile. A whole-cell suspension was centrifuged at 12,000 × *g* for 5 min, and 1-μl portions of the supernatants were placed on a MALDI sample slide (Bruker-Daltonics, Bremen, Germany) and dried at room temperature. The sample was then overlaid with 1 μl of matrix solution (α-cyano-4-

hydroxycinnamic acid in 50% acetonitrile and 2.5% trifluoroacetic acid) and dried at room temperature. A MALDI-TOF MS measurement was performed with a Bruker MicroFlex MALDI-TOF MS (Bruker-Daltonics) using FlexControl software and a DH5-α *Escherichia coli* protein extract (Bruker-Daltonics) was deposited on the calibration spot of the sample slide for external calibration. Spectra collected in the positive-ion mode within a mass range of 2,000 to 20,000 Da were analyzed using a Bruker Biotyper (Bruker-Daltonics) automation control and the Bruker Biotyper 3.1 software and library (a database with 5,627 entries). The clustering analysis of the GAS strains was performed by generation of the dendrogram based on the different serotypes collected in the present study. In detail, 49 strains representing 11 different serotypes included: M-1 ($n = 4$), M-3 ($n = 2$), M-5 ($n = 6$), M-6 ($n = 9$), M-9 ($n = 1$), M-12 ($n = 1$), M-18 ($n = 11$), M-28 ($n = 9$), M-44 ($n = 2$), M-89 ($n = 3$), and M-102 ($n = 1$). The main spectra (MSPs) of each strain were generated from 10 technical replicates prior to manual visualization inspections using FlexAnalysis 3.4 software. The relationship between MSPs obtained from each strain was visualized in a score-oriented dendrogram using the average linkage algorithm implemented in the MALDI Biotyper 3.1 software.

**Whole-genome sequencing and comparative genomics.** Whole-genome sequencing was conducted on the two *S. pyogenes* isolates from the ARF case and nine M-18 GAS strains included in the present study (see Table S1 in the supplemental materials). Libraries were prepared with a Nextera XT sample preparation kit (Illumina), and sequencing was performed on the Illumina MiSeq platform (Illumina, San Diego, CA) with a 2×250 paired-end run. All read sets were evaluated for sequence quality and read-pair length using FastQC (19) and then assembled with MIRA 4.0 using a *de novo* assembly mode (20).

For comparative studies, SNPs were identified with an in-house Perl pipeline based on the Mauve software (21). In this approach, the genome of *S. pyogenes*, group A strain SP665Q, was used as a reference, and the other 11 genome assemblies included in the present study were aligned against it. All of the alignments were then merged, and the coordinates of all nucleotide variations were detected on the basis of the annotated reference strain assembly (SP665Q). All of the variations were then organized in a matrix to assess the presence or absence pattern in specific subsets of strains (e.g., *emm18.29* and *emm18.32*). The core SNPs, defined as the nondegenerate SNPs present in all 12 genomes and flanked by conserved positions, were extracted and finally subjected to Bayesian phylogenetic analysis with MrBayes (22). The Bayesian analysis was run on the GTR substitution model for 2,000,000 generations with a chain sampled every 1,000th generation. The final parameter values and trees are summarized after 25% of the posterior sample was discarded. The Bayesian tree is displayed and edited using FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree).

Informative SNPs (i.e., present in at least two strains) were extracted from core SNPs using an in-house script. In addition, among the strains with *emm* type 18.29, all genes presenting at least one core SNP were selected and compared to the virulence factors of pathogenic bacteria (VFPB) database (23), using a blast search with a $10^{-5}$ value cutoff.

Prophages were detected and analyzed using the free web tool PHAST (PHAge Search Tool) (24). The comparative sequence circular maps of whole genomes and concatenated prophages of each strain of *S. pyogenes* M18GAS were generated using BRIG (25).

**Accession numbers.** The sequences of the 11 M-18 GAS genomes were deposited at EMBL/EBI under the following accession numbers: M18GASBO1065 (CDGV01000001 to CDGV01000182), M18GASBO665 (CDGO01000001to CDGO01000182), M18GASBO9 (CDGY01000001 to CDGY01000207), M18GASBO8 (CDGW01000001 to CDGW01000183), M18GASBO7 (CDGX01000001 to CDGX01000199), M18GASBO6 (CDGM01000001 to CDGM01000081), M18GASBO5 (CDGQ01000001 to CDGQ01000096) M18GASBO4 (CDGN01000001 to CDGN01000158), M18GASBO3 (CDGS01000001 to CDGS01000294), M18GASBO2 (CDHA01000001 to CDHA01000222), and M18GASBO1 (CDHB01000001 to CDHB01000235) (study project PRJEB7108).

## RESULTS

**Characterization of GAS isolates.** Eight ARF cases were recorded in the Bologna province between November 2012 and May 2013. All patients were aged between 4 and 12 years. In recent years, the number of ARF cases in this province has ranged from one to four per year. At time of diagnosis, seven of eight ARF cases were culture negative, and only one GAS strain was isolated from a patient resident in Bologna province.

During the surveillance period, 70 GAS isolates were collected (see Table S1 in the supplemental material). In detail, two consecutive isolates were collected from the first patient with ARF (i.e., SPBO1 and SPBO2), six were isolated from classmates of the first case and 14 were isolated from symptomatic children resident in the Bologna area at the time of ARF diagnosis. Two months later, a second collection comprised 14 GAS isolates collected from classmates of the second ARF case and 34 GAS isolates obtained from symptomatic school-age children resident in the same area. However, no GAS was isolated from the second case of ARF.

Analysis of the *emm* sequence revealed that the GAS isolates obtained from patient with ARF was *emm18.29*, as two of six of the GAS isolates were obtained from contacts (Fig. 1A). At the same time, GAS isolated from symptomatic children of the community were: *emm28.0* (3 cases), *emm18.29* (2 cases), *emm89.0*, *emm1.0*, *emm56.0*, *emm9.0*, *emm6.4*, *emm102.3*, *emm12.0*, *emm5.3*, and *emm3.8* (one case each), as shown in Fig. 1B.

Molecular investigation conducted among the 14 GAS collected from the classmates of the second ARF case showed that two isolates (14.3%) were *emm18.32*, four (28.5%) were *emm28.0*, and eight (57.2%) were *emm6.4*, as shown Fig. 1C. Among the 35 GAS isolates collected from symptomatic school-age children in the community, eight different *emm* types were identified: *emm6.4* (23 cases), *emm28.0* (2 cases), *emm44.0* (3 cases), *emm1.0* (2 cases), and *emm5.3*, *emm18.32*, *emm89.0*, *emm5.1*, and *emm3.1* (one case each), as shown in Fig. 1D.

Our data indicate that *emm18.29* was the dominant serotype among the isolates collected from the class of the first case, whereas *emm18.32* spread in the class of the second case. To investigate the relationship between the specific *emm* type and the SAg profile, GAS isolates were evaluated by molecular analysis for the different SAg genes (see Table S1 in the supplemental material). Molecular analysis revealed eight different antigenic profiles among 70 GAS isolates. The chromosomally encoded *speB*, *speG*, and *smeZ* genes were present in all isolates. In addition, all *emm18* strains, including *emm18.29* and *emm18.32* and the two *emm18.0* strains derived from the ISS bank collection (SP665Q and SP1065Q), presented a characteristic SAg profile showing *speA*, *speC*, *speL*, and *speM* genes. Overall, the isolates with the same *emm* type shared a common SAg profile, with the exception of the *emm3* isolates. MLST analysis showed that all *emm18* GAS isolates belonged to sequence type 42.

To determine the association of the mucoid phenotype with *emm* type, all GAS isolates were analyzed by culture visualization. Among GAS isolates, the M-102, M-9, and M-18 GAS strains showed the highest incidence (100, 100, and 87.5%, respectively) of mucoid traits, whereas all *emm1* isolates showed a nonmucoid colony type (see Table S1 in the supplemental material). Antimicrobial susceptibility testing showed that all isolates were susceptible to penicillin, ampicillin, clindamycin, chloramphenicol, tetracycline, and clindamycin, whereas only two GAS strains

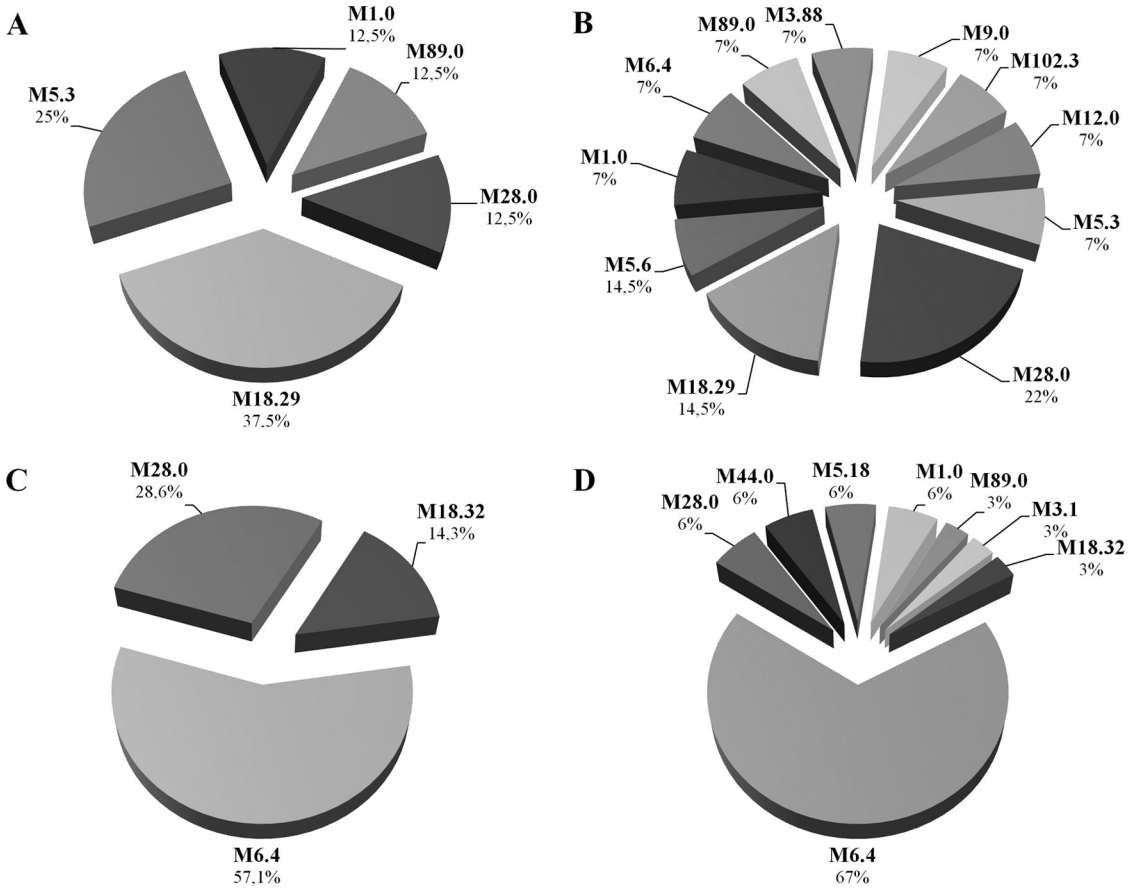


**FIG 1** Distribution of *emm* types among *S. pyogenes* strains collected during an ARF outbreak occurred in Bologna province. (A) GAS strains isolated from the ARF case and class contacts during the first episode. (B) GAS isolated from class children during the second ARF episode. (C) Distribution of *emm* types among GAS isolates collected from the community at the time of the first ARF case. (D) Distribution of *emm* types among GAS collected from school-age children of the community at time of the ARF episode.

belonging to the *emm89.0* and *emm5.3* serotypes were resistant to erythromycin.

**MALDI-TOF and clustering analysis.** MALDI-TOF MS analysis of different *S. pyogenes* strains showed that 37 of 50 (74.0%) isolates clustered in accordance with the serotype group, as shown in Fig. 2. In detail, the main spectra generated by MALDI-TOF MS analysis demonstrated a high overall discriminatory power of the strains belonging to the serotypes M-18, M-28, and M-3. However, different clustering groups were observed for M-1, M-5, M-6, M-89, and M-44 strains by showing a different protein mass spectral profiling among isolates belonging to the same serotype (Fig. 2). Notably, the score-oriented dendrogram showed that all isolates belonging to serotype M-18 formed a separate clustering group that was clearly distinguishable from other serotypes. In addition, a different cluster grouping within the M-18 serotype was observed between *emm18.0* and the *emm18.32*/*emm18.29* subtypes, with a critical distance of 500 (see Fig. 2), respectively corresponding to the *S. pyogenes* strains isolated 30 years ago and the isolates involved in the ARF outbreak in Bologna province during 2013. However, MALDI-TOF MS analysis was not able to distinguish among *emm18.29* and *emm18.32* subtypes.

**Whole-genome sequencing and phylogenetic analysis of M-18 GAS isolates.** The draft genomes of *emm18.29*, *emm18.32*, and *emm18.0* GAS isolates were assembled into average 185 contigs with a G+C content of 38.6% for a total of 1,929,545 bp (Fig. 3). Genome annotations predicted a total of 1,881 open reading frames.

To investigate the relationship among M-18 GAS isolates, a whole-genome analysis was performed on the basis of core SNPs identified with the Mauve-based approach (593 core SNPs). Phylogenetic analysis indicated that two main clusters were highlighted with high posterior probabilities: the first cluster included the *emm18.29* strains, while the second included the *emm18.32* strains. In addition, the two e*mm18.0* GAS isolates derived from the ISS bank collection (SP665Q and SP1065Q) showed a closely relationship to the GAS8232 strain. Phylogenetic analysis indicated that *emm18.0* strains and GAS8232 differed by 198 informative SNPs and were nearer to *emm18.32* strains than to *emm18.29* strains (Fig. 4). Comparison of SNPs between M-18 strains isolated during the ARF outbreak in Bologna province showed that the two consecutive GAS isolates from the ARF case (SPBO1 and SPBO2) differed by 14 SNPs (0 informative SNPs) and were closely related to strains collected from classmates (SPBO3 and SPBO4) and from the community (SPBO5 and SPBO6), as shown in Fig. 4. Comparison of *emm18.29* genomes disclosed 216 different SNPs and 5 informative SNPs in this group. In addition, all of the *emm18.32* GAS strains collected from children during the sec-

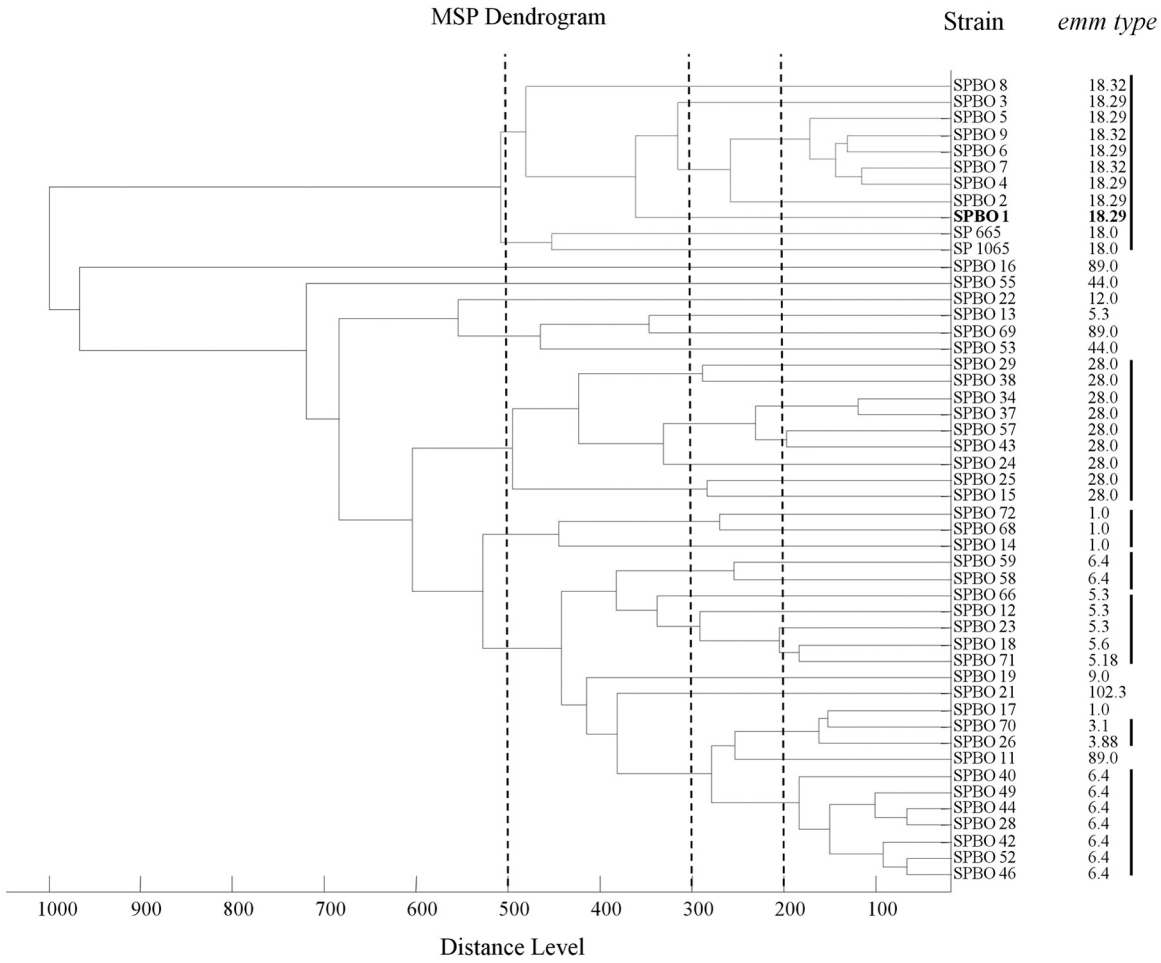


**FIG 2** Score-oriented dendrogram based on the main spectra (MSP) of 50 GAS strains obtained by Bruker MALDI-TOF MS and analyzed using Biotyper 3.1 software. Correlation with the 11 different *emm* types (M-1, M-3, M-5, M-6, M-9, M-12, M-18, M-28, M-44, M-89, and M-102) shown. Dotted lines define a similarity cutoff value of 500, 300, and 200 used for clustering groups of serotypes M-18, M-1, M-3, M-5, M-28, and M-6.

ond ARF episode were closely related (53 different SNPs and 0 informative SNPs) (Fig. 4).

Further comparison of all M-18 genomes identified 73 and 207 SNPs exclusive to *emm18.29* and *emm18.32* clusters, respectively (see Table S2 in the supplemental material). Analysis of all unique SNPs of *emm18.29* cluster showed that 36.9% ($n = 27$) were synonymous, 58.9% ($n = 43$) were nonsynonymous, and 4.2% ($n = 3$) were located in intergenic regions. At the same time, the SNPs of *emm18.32* cluster were 41.6% ($n = 86$) synonymous, 50.2% ($n = 104$) nonsynonymous, and 8.2% ($n = 17$) within intergenic regions. Analysis of synonymous and nonsynonymous substitutions in the coding regions of the core genome revealed that SNPs were associated with different GAS virulent factors (see Table S2 in the supplemental material) present in the VFPB database (23). Notably, 46 SNPs substitutions occurred in prophage elements, including *hyaluronidase* and *gp58*-like genes in both *emm18.29* and *emm18.32* clusters (see Table S2 in the supplemental material). In detail, five SNPs in the hyaluronidase (*hylP*) genes among the two cluster subtypes were synonymous and six were nonsynonymous, most of them located in the N-terminal region of *hylP* gene.

**Comparison of phage elements in M-18 GAS strains.** Analysis of M-18 GAS chromosomes revealed that six regions contained prophage elements (ΦSPBO.1, ΦSPBO.2, ΦSPBO3, ΦSPBO4, ΦSPBO.5, and ΦSPBO.6) ranging from 7.6 to 75.8 kb (Table 2). The genome distribution of the prophages across the GAS chromosome showed that all M-18 strains shared a common localization of these elements, as shown in Fig. 3. Comparison to other GAS genomes revealed that five prophag elementse (ΦSPBO.1, ΦSPBO2, ΦSPBO3, ΦSPBO4, and ΦSPBO.5) showed similar chromosome locations to that of the GAS8232 genome (11), suggesting conservative site integrations of these regions across M-18 GAS strains (Fig. 3). Moreover, examination of prophages elements showed a similar genetic architecture with GAS8232 (M-18) and MGAS315 (M-3) strains, as shown in Fig. 5 (11, 26).

Genomes of M-18 GAS strains contained prophage regions harboring several virulence factors, including exotoxin type A (SpeA), exotoxin type C (SpeC), exotoxin type L (SpeL), exotoxin type M (SpeM), mitogenic factor (DNase), and streptodornase (Sdn) (Table 2). Comparison of the prophage elements showed that the ΦSPBO.2, ΦSPBO.3, and ΦSPBO.5 regions were shared among all M-18 GAS strains. These three phage regions contained different virulence factors such as genes encoding SpeC, mitogenic factor, SpeL/SpeM, hyaluronidase, and streptodornase. In-

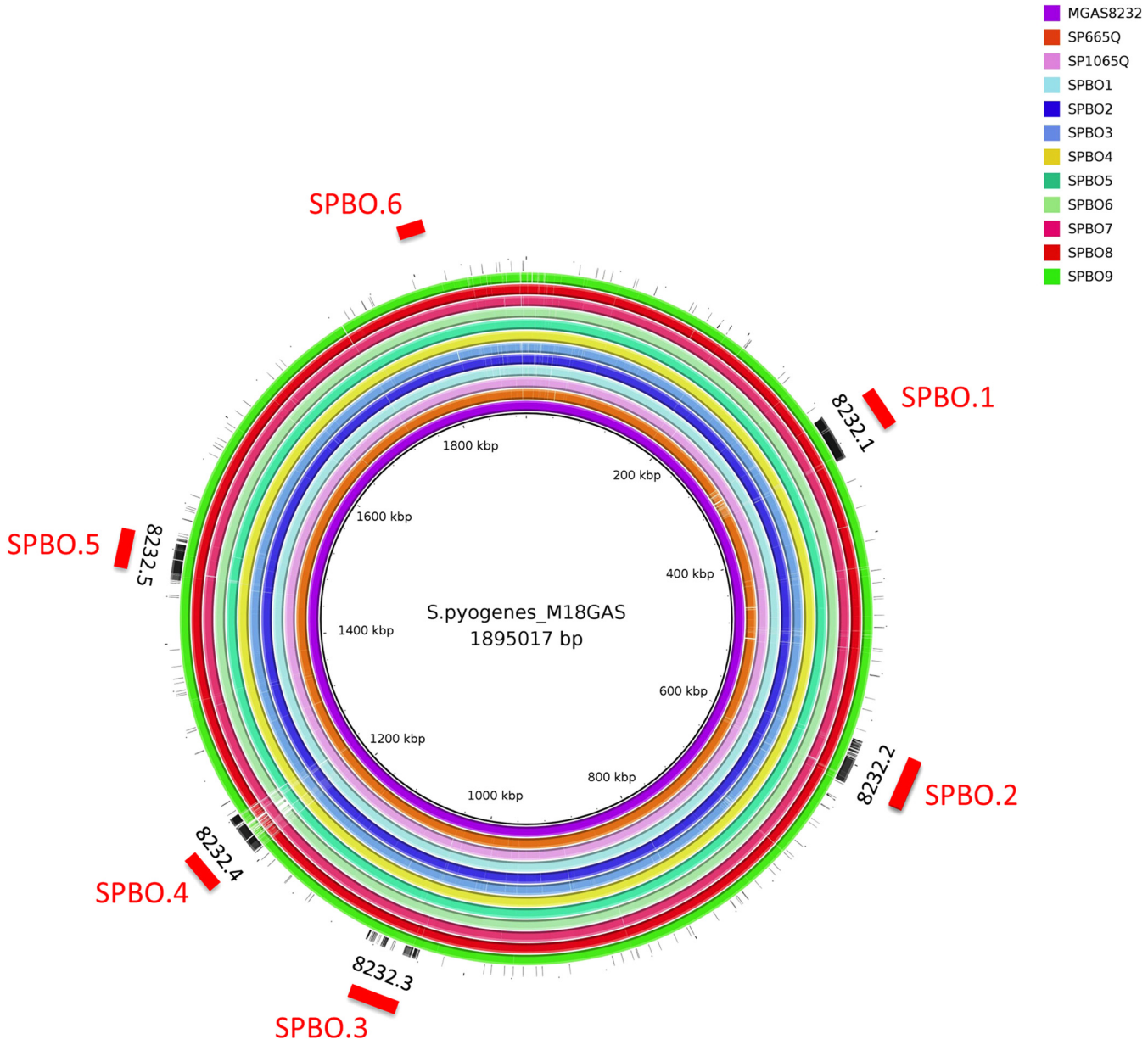


**FIG 3** Comparison of the group A *Streptococcus* serotype M18 chromosomes. The circular representation shows a genome comparison from center to periphery, respectively, of strains GAS8232, SPBO1, SPBO2, SPBO3, SPBO4, SPBO5, SPBO6, SPBO7, SPBO8, SPBO9, SP665Q, and SP1065Q (see the legend for the color associations). The regions of differences within GAS genomes are indicated with white gaps. The genomic localizations of the prophage elements shared with GASM8232 are indicated as black boxes outside the circular GAS chromosome maps. The locations of the prophage regions (ΦSPBO.1, ΦSPBO.2, ΦSPBO.3, ΦSPBO.4, ΦSPBO.5, and ΦSPBO.6) of M-18 GAS isolates collected from Bologna province are indicated as red boxes.

terestingly, closely association between phages encoded SpeC and SpeL/speM were observed in seven M-18 GAS strains (Table 2).

Moreover, the ΦSPBO.1 region was present in 10 of 11 of the M-18 GAS strains, lacking in the SP665Q isolate (*emm18.0* subtype). Analysis of virulence factors showed that ΦSPBO.1 region has variants of the *speA* gene (*speA1*) in six of nine *emm18.29* and *emm18.32* strains but was absent in *emm18.0* strains (Table 2). In addition, phage ΦSPBO.4 was absent in *emm18.32* strains, whereas it was present in all *emm18.0* and *emm18.29* strains. This region contained a gene encoding mitogenic factor. Interestingly, each phage region had a hyaluronidase gene (Table 2). Similar findings were observed in a previous study demonstrating that all phages in M-3 GAS strains contained a hyaluronidase gene (26).

**Comparison of phage-encoded hyaluronidase (*hylP.2*) gene among M-18 GAS strains.** To investigate the genetic variability in the *hylP.2* gene, the gene-encoded HylP.2 protein of M-18 strains was compared to previously described GAS isolates. Our analysis indicated that the *hylP.2* gene was located in phage ΦSPBO.1 in all *emm18.29*, *emm18.32*, and *emm18.0* GAS strains (Table 2). Analysis of the *hylP.2* gene derived from the *emm18.29*, *emm18.32*, and *emm18.0* strains showed a high rate of homology with the M18

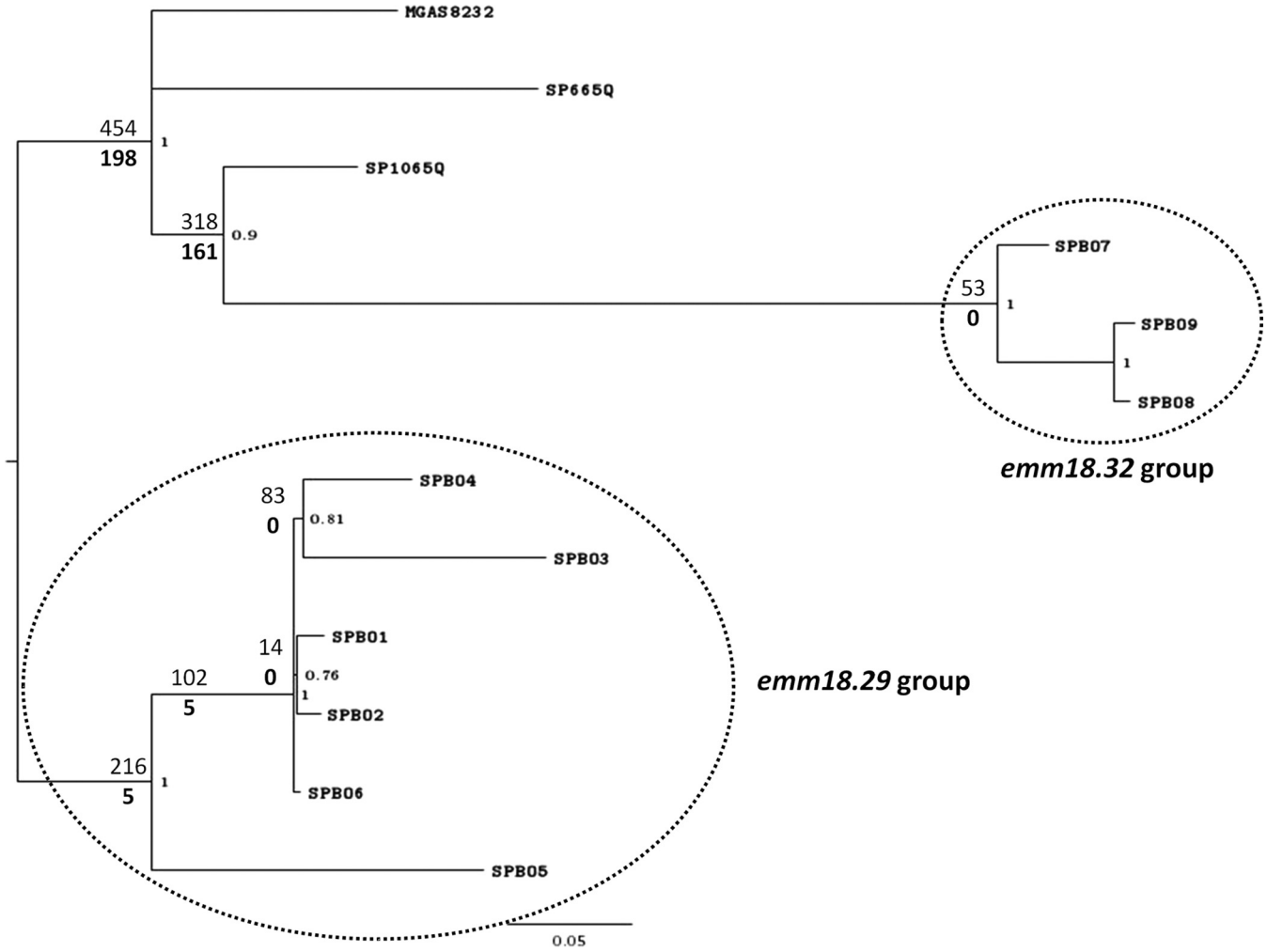


**FIG 4** Bayesian tree-based on core SNPs identified with the Mauve-based approach. In each node of the tree, posterior probabilities (>0.7) are indicated on the right of the node, while the numbers of different and informative SNPs located, respectively, up and down the branch are on the left of the node.

(MGAS8232) and M5 (Manfredo) GAS strains, two strains isolated from ARF cases (11, 27). Comparison of deduced amino acid sequences with M5-GAS Manfredo strain demonstrated that HylP.2 were C terminally truncated in all M-18 GAS isolates (including *emm18.29*, *emm18.32*, *emm18.0*, and GAS8232 strains) (Fig. 6). Of note, three *emm18.29* isolates (SPBO1, SPBO3, and SPBO6) showed an internal deletion located between the N-terminal and the TSβH domains. At the same time, the deduced amino acid sequence of *hylP.2* gene in the SPBO8 isolate was truncated in the TSβH region (Fig. 6). Comparison analysis conducted on the *hylP.2* gene revealed different clustering groups according to the corresponding subtypes (data not shown).

**TABLE 2** Prophage elements in M-18 GAS strains

| | Prophage element size (kb) in M-18 GAS strain[a] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *emm18.29* | | | | | | *emm18.32* | | | *emm18.0* | | |
| Phage | SPBO1 | SPBO2 | SPBO3 | SPBO4 | SPBO5 | SPBO6 | SPBO7 | SPBO8 | SPBO9 | SP665Q | SP1065Q | Virulence factor(s) |
| ΦSPBO.1 | 56.4* | 67.2 | 52.5 | 67.4* | 64.9* | 61.2* | 68.7 | 75.8* | 57.6* | - | 66.1 | *speA*, hyaluronidase (*hylP.2*) |
| ΦSPBO.2 | 32.5 | 38.5 | 36.9 | 39.5† | 38.3† | 37.2† | 38 | 38.6† | 37.3† | 36.4† | 39.6† | *speC* mitogenic factor, hyaluronidase |
| ΦSPBO.3 | 47.8 | 68.5 | 59.4 | 57‡ | 58.4‡ | 59.4‡ | 57.3 | 57.4‡ | 59.1‡ | 57.6‡ | 67.1‡ | *speL*, *speM*, hyaluronidase |
| ΦSPBO.4 | 33 | 31.8 | 30.6 | 25.2 | 30.6 | 32.5 | | | | 43.6 | 41.1 | Hyaluronidase, mitogenic factor |
| ΦSPBO.5 | 60 | 47.3 | 39.3 | 45.7 | 47.2 | 51 | 48.5 | 42.6 | 48 | 51.3 | 46.6 | Hyaluronidase, streptodornase |
| ΦSPBO.6 | 14.1 | 12.8 | 12.5 | 19.5 | 47.2 | 18.4 | 17 | | 10.5 | | 7.6 | Hyaluronidase |

[a] *, Prophage-containing variant of *speA* (*speA1*); †, prophage-containing *speC*; ‡, prophage-containing *speL* and *speM*.

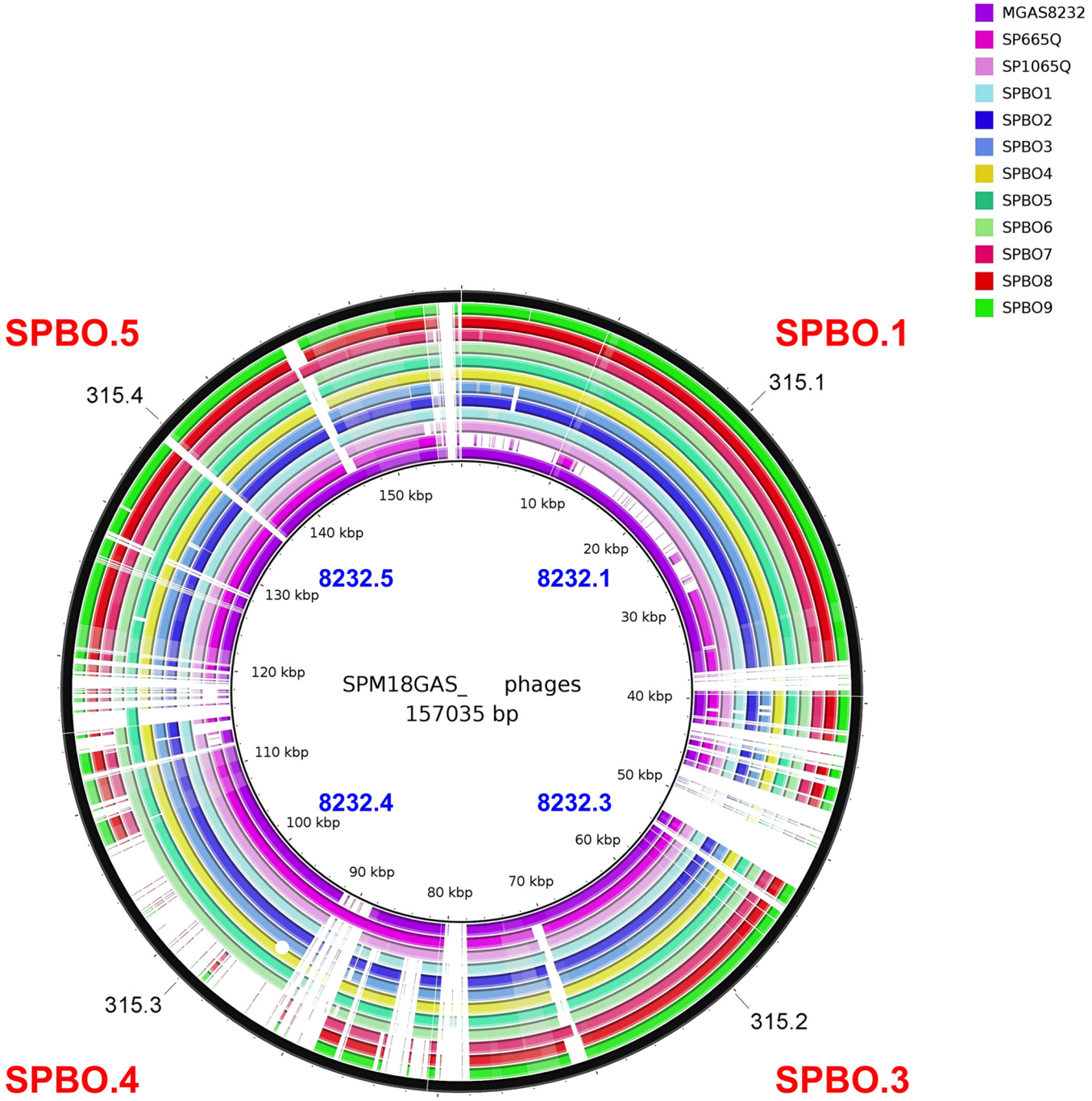


**FIG 5** Circular representation of concatenated prophage elements integrated in the genome of the M-18 GAS strains compared to M-3 GAS prophages. The outermost black circle represent the concatenated M-3 GAS prophages (Φ315.1, Φ315.2, Φ315.3, and Φ315.4). The prophages of M-18 GAS strains collected in the present study (SPBO1, SPBO2, SPBO3, SPBO4, SPBO5, SPBO6, SPBO7, SPBO8, SPBO9, SP1065Q, and SP665Q) and reference strain (M18GAS8232) are indicate in red and blue, respectively. The areas of similarity and divergence are contrasted with white gapped areas indicating regions of highest variance.

## DISCUSSION

Since late 2012, an outbreak of acute rheumatic fever (ARF) was observed in the Bologna province, Northeast Italy, where the annual frequency of ARF in resident children has ranged from 1 to 4 cases per year. From November 2012 to May 2013, eight cases of ARF were recorded, showing a significant increase of ARF in this area. Molecular analysis conducted among GAS collected from both contacts and the community during the outbreak indicated that M-18 represented one of the most prevalent serotype within classes of two unrelated ARF episodes. Our findings showed that two distinct subtypes, i.e., *emm18.29* and *emm18.32*, were observed to spread separately across the two classes of ARF cases. Analysis of GAS isolates from children in the community showed that the majority of the isolates were subtype *emm6.4*. Phenotypic

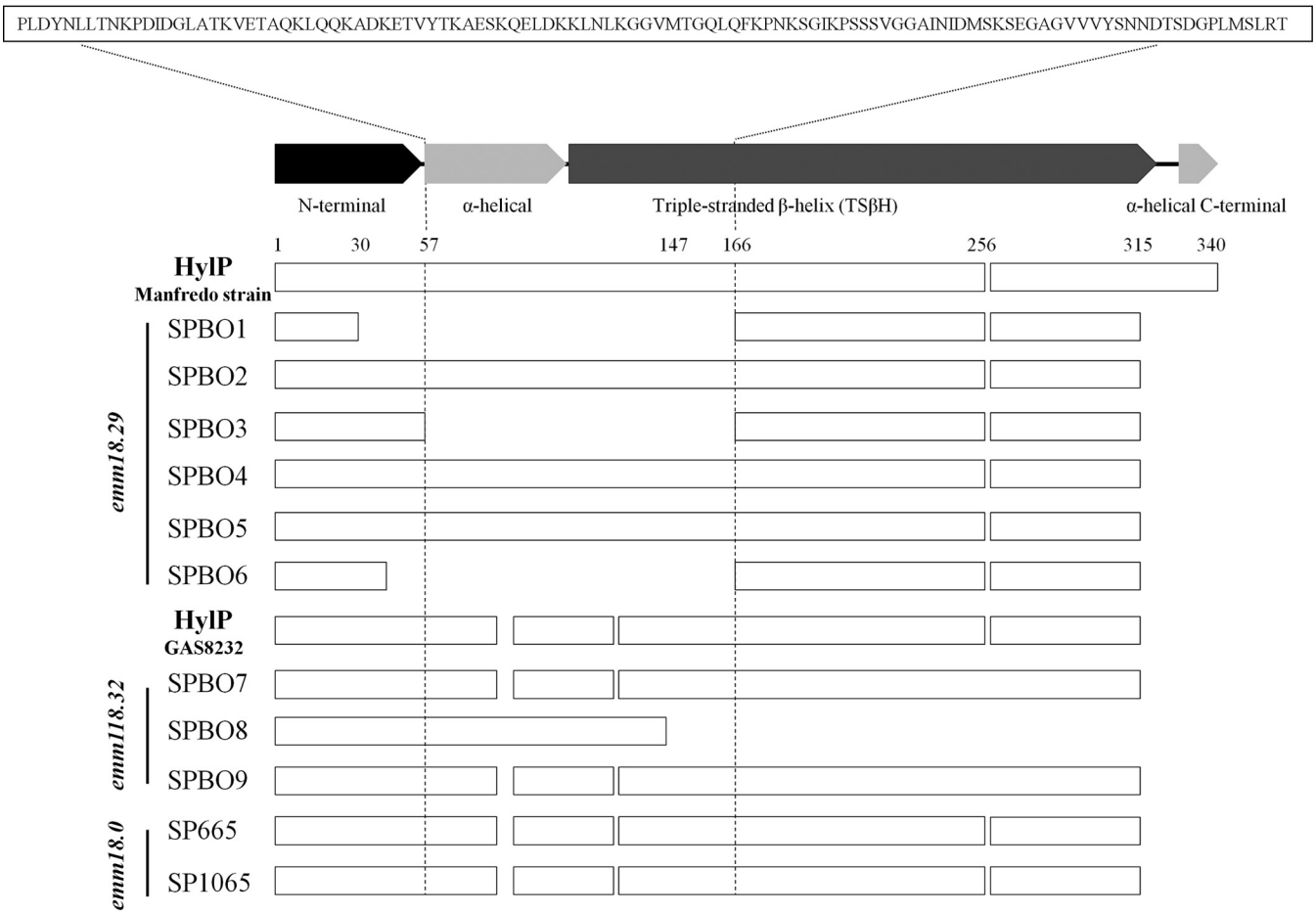


**FIG 6** Alignment of the phage-encoded hyaluronidase gene (*hylP.2*) derived from M-18 GAS (*emm18.29*, *emm18.32*, and *emm18.0* subtypes). The *hylP.2* sequences from ARF-related M-18 (GAS8232) and M-5 (Manfredo) strains are shown. Dotted lines show the common deleted regions between α-helical and TSβH domains between SPBO1, SPBO3, and SPBO6 strains. The deduced amino acid sequence of a 327-nucleotide deletion in *emm18.29* strains is shown.

analysis of GAS showed that the M-18 GAS strains presented a high frequency of mucoid strains, confirming previous findings (5). Indeed, GAS mucoid isolates have been observed to correlate with invasive infections and the pathogenesis of rheumatic fever (1). Previous studies clearly demonstrated that higher capsule production has been observed in M-18 strains responsible for multiple ARF outbreaks (11). Although our study showed that GAS M-18 was responsible of an ARF case and that this serotype has spread in the two classes, we cannot exclude the possibility that others strains (i.e., serotypes) have been responsible of others ARF cases.

The present study evaluated the ability of MALDI-TOF MS to correctly identify GAS and to distinguish among the different serotypes. MALDI-TOF analysis demonstrated an excellent discrimination of M-18 GAS strains by showing a separate cluster. Overall, cluster analysis by MALDI-TOF showed a good concordance (74%) with *emm* typing methods. In detail, a 100% concordance was observed with serotypes M-28 (9 of 9) and M-3 (2 of 2), whereas 75 and 83.4% concordances were found with M-1 (3 of 4) and M-5 (5 of 6), respectively. The discrepancies observed within these serotypes were attributable to two isolates, SPBO13 (*emm5.3*) and SPBO17 (*emm1.0*), grouped separately from strains belonging to the same serotypes. These results are in accordance with a previous study demonstrating the potential of the MALDI-TOF Biotyper system for GAS clustering analysis, thus showing a high discriminatory power among different serotypes (9). However, a poor discriminatory classification was observed for M-89 serotype. Based on these findings and for its rapidity and low cost analysis, we suggest the MALDI-TOF technique could be used successfully for the identification of the M-18 serotype among GAS strains.

In recent years, whole-genome sequencing has been applied for epidemiological purposes by showing a more accurate resolution than classical genotypic methods (13). Whole-genome sequencing has been extensively used to explore the genetic organization of bacterial genomes and to compare the rearrangements between closely related strains (10, 12, 28). The present study described the complete genomes of nine M-18 GAS strains isolated during an ARF outbreak in northeastern part of Italy and compared them to two M-18 *S. pyogenes* collected from noninvasive infections in the same area 30 years ago. SNP phylogenetic analysis revealed that *emm18.29* and *emm18.32* subtypes segregated in two separate clusters, whereas *emm18.0* GAS strains did not cluster in a distinct group. Genome polymorphism analysis showed that isolates from ARF cases and from community and class contacts were closely

related and showed a low number of informative SNPs both in *emm18.29* and *emm18.32* strains.

Our analysis revealed that M-18 strains possessed several virulence genes, including *speA*, *speC*, *speL*, *speM*, *smeZ*, *mitogenic factor*, and *hyaluronidase*, most of them located in prophage elements. Integrated prophages represent one of the most divergent tracts among GAS genomes and the majority of genetic variations among M-18 GAS strains (29).

Our results indicated that prophage elements are located in the same genomic locations among different M-18 strains collected in the present study and in the MGAS8232 reference strain. In addition, genomic organization revealed that three prophage elements (ΦSPBO.2, ΦSPBO.3, and ΦSPBO.5) were common among M-18 GAS strains. Alignment of the prophage sequences showed a similar architecture between subtypes *emm18.29*, *emm18.32*, and *emm18.0*, thus revealing a similar genomic architecture of M-18 GAS strains. It has been established that hypervirulent GAS strains acquire virulence factors via prophage integrations (29). Recently, Bao et al. reported that prophage integrations represent one of the multiple genetic factors related to the pathogenic role of the M-23ND GAS strain (30). Our findings showed that three prophages (ΦSPBO.1, ΦSPBO.3, and ΦSPBO.5) present in the M-18 GAS strains were similar with phages of M-3 serotypes (Φ315.1, Φ315.2, and Φ315.4) that have been previously associated with the emergence of virulent M-3 subclones (26) by different sequential acquisition.

We showed that the streptococcal pyrogenic exotoxin A gene (*speA1*) was located in the prophage ΦSPBO.1 region in six of nine *emm18.29* and *emm18.32* strains, whereas this gene was absent in *emm18.0* strains. Also, we observed that the *speC* was present in all M-18 GAS strains possessing *speL* and *speM*, thus showing a strict correlation between these prophage-encoded SAg genes. Based on these findings, we hypothesize that a different combination of phage-encoded virulence factors could be related to the virulence of *emm18.29* and *emm18.32* strains isolated during focal ARF that differed from noninvasive *emm18.0* strains collected from the same area 30 years ago.

Analysis of the phage-encoded virulence factors demonstrated that phage-encoded hyaluronidase showed a higher number of synonymous and nonsynonymous substitutions than other genes within the M-18 GAS genomes. Previous studies reported that *hylP* and *hylP.2* genes were present with different alleles among different serotypes from both invasive and noninvasive GAS isolates (31, 32). However, M-18 GAS strains have been observed to possess a unique *hylP.2* gene structure among different isolates (33). Our findings demonstrated that the *hylP.2* gene possesses an internal deletion located between the N-terminal and TSβH regions in different M-18 GAS strains. Therefore, the truncated gene structures observed in several M-18 GAS isolates could be related to a different or nonfunctional activity of the HylP.2 protein. Previous study showed that inactivation in hyaluronate lyase (HylA) restored full encapsulation in partially encapsulated M-4 GAS strains, thus demonstrating the mutually exclusive interaction between the hyaluronan capsule and active hyaluronidase (32). In addition, Schommer et al. demonstrated in a mouse model that the difference in capsule size was regulated by bacterial hyaluronidase and that the high molecular mass of the hyaluronan capsule influences GAS virulence by facilitating deep tissue infections (34). Based on our findings, we hypothesize that inactivation of HylP.2 could determine a different encapsulation (i.e., capsule sizing) of M-18 GAS strains, thus resulting in a more virulent clone. Therefore, the internal deletion in the *hylP.2* gene observed in different isolates could reflect a different virulence potential among the M-18 GAS strains.

In conclusion, we investigated the molecular and epidemiological linkage between GAS strains isolated during an ARF outbreak in Bologna province in early 2013. Our study explored the genome sequence of M-18 GAS strains, thus providing a better understanding of the genetic architecture of the M-18 serotype and expanding our knowledge of the genetic elements related to the GAS infections.

## ACKNOWLEDGMENT

This study was supported by funds from the Emilia-Romagna region.

## REFERENCES

1. **Cunningham MW.** 2014. Rheumatic fever, autoimmunity, and molecular mimicry: the streptococcal connection. Int Rev Immunol **33:**314–329. http://dx.doi.org/10.3109/08830185.2014.917411.
2. **Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V.** 2014. Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. Clin Microbiol Rev **27:**264–301. http://dx.doi.org/10.1128/CMR.00101-13.
3. **Tibazarwa KB, Volmink JA, Mayosi BM.** 2008. Incidence of acute rheumatic fever in the world: a systematic review of population-based studies. Heart **94:**1534–1540. http://dx.doi.org/10.1136/hrt.2007.141309.
4. **Wolfe RR.** 2000. Incidence of acute rheumatic fever: a persistent dilemma. Pediatrics **105:**1375. http://dx.doi.org/10.1542/peds.105.6.1375.
5. **Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR.** 2009. Global *emm* type distribution of group A streptococci: systematic review and implications for vaccine development. Lancet Infect Dis **9:**611–616. http://dx.doi.org/10.1016/S1473-3099(09)70178-1.
6. **Smoot JC, Korgenski EK, Daly JA, Veasy LG, Musser JM.** 2002. Molecular analysis of group A *Streptococcus* type *emm18* isolates temporally associated with acute rheumatic fever outbreaks in Salt Lake City, Utah. J Clin Microbiol **40:**1805–1810. http://dx.doi.org/10.1128/JCM.40.5.1805-1810.2002.
7. **Sauer S, Kliem M.** 2010. Mass spectrometry tools for the classification and identification of bacteria. Nat Rev Microbiol **8:**74–82. http://dx.doi.org/10.1038/nrmicro2243.
8. **Mencacci A, Monari C, Leli C, Merlini L, De Carolis E, Vella A, Cacioni M, Buzi S, Nardelli E, Bistoni F, Sanguinetti M, Vecchiarelli A.** 2013. Typing of nosocomial outbreaks of *Acinetobacter baumannii* by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol **51:**603–606. http://dx.doi.org/10.1128/JCM.01811-12.
9. **Wang J, Zhou N, Xu B, Hao H, Kang L, Zheng Y, Jiang Y, Jiang H.** 2012. Identification and cluster analysis of *Streptococcus pyogenes* by MALDI-TOF mass spectrometry. PLoS One **7:**e47152. http://dx.doi.org/10.1371/journal.pone.0047152.
10. **Sassera D, Comandatore F, Gaibani P, D'Auria G, Mariconti M, Landini MP, Sambri V, Marone A.** 2014. Comparative genomics of closely related strains of *Klebsiella pneumoniae* reveals genes possibly involved in colistin resistance. Ann Microbiol **64:**887–890. http://dx.doi.org/10.1007/s13213-013-0727-5.
11. **Smoot JC, Barbian KD, Van Gompel JJ, Smoot LM, Chaussee MS, Sylva GL, Sturdevant DE, Ricklefs SM, Porcella SF, Parkins LD, Beres SB, Campbell DS, Smith TM, Zhang Q, Kapur V, Daly JA, Veasy LG, Musser JM.** 2002. Genome sequence and comparative microarray analysis of serotype M18 group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. Proc Natl Acad Sci U S A **99:**4668–4673. http://dx.doi.org/10.1073/pnas.062526099.
12. **Gaiarsa S, Comandatore F, Gaibani P, Corbella M, Dalla Valle C, Epis S, Scaltriti E, Carretto E, Farina C, Labonia M, Landini MP, Pongolini S, Sambri V, Bandi C, Marone P, Sassera D.** 2014. Genomic epidemiology of *Klebsiella pneumoniae*: the Italian scenario, and novel insights into the origin and global evolution of resistance to carbapenem antibiotics. Antimicrob Agents Chemother **59:**389–396. http://dx.doi.org/10.1128/AAC.04224-14.

13. **Aziz RK, Nizet V.** 2010. Pathogen microevolution in high resolution. Sci Transl Med **2:**16ps4. http://dx.doi.org/10.1126/scitranslmed.3000713.
14. **Carapetis JR, Parr J, Cherian T.** 2006. Standardization of epidemiologic protocols for surveillance of poststreptococcal sequelae: acute rheumatic fever, rheumatic heart disease and acute poststreptococcal glomerulonephritis. January. National Institutes of Health/National Institute of Allergy and Infctious Disease, Bethesda, MD. http://www.niaid.nih.gov/topics/strepthroat/documents/groupasequelae.pdf.
15. **European Committee on Antimicrobial Susceptibility Testing.** 2014. Breakpoint tables for interpretation of MICs and zone diameters, version 4.0. EUCAST, Basel, Switzerland.
16. **Beall B, Facklam R, Thompson T.** 1996. Sequencing *emm*-specific PCR products for routine and accurate typing of group A streptococci. J Clin Microbiol **34:**953–958.
17. **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:**115–125. http://dx.doi.org/10.1007/s10096-012-1726-3.
18. **McGregor KF, Spratt BG, Kalia A, Bennett A, Bilek N, Beall B, Bessen DE.** 2004. Multilocus sequence typing of *Streptococcus pyogenes* representing most known *emm* types and distinctions among subpopulation genetic structures. J Bacteriol **186:**4285–4294. http://dx.doi.org/10.1128/JB.186.13.4285-4294.2004.
19. **Andrews S.** 2014. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics/Babraham Institute, Cambridge, United Kingdom. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
20. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information. Comput Sci Biol **1999:** 45–56.
21. **Darling AE, Mau B, Perna NT.** 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss, and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.
22. **Huelsenbeck JP, Ronquist F.** 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755. http://dx.doi.org/10.1093/bioinformatics/17.8.754.
23. **Chen LH, Xiong ZH, Sun LL, Yang J, Jin Q.** 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res **40:**D641–D645. http://dx.doi.org/10.1093/nar/gkr989.
24. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39:**347–352. http://dx.doi.org/10.1093/nar/gkq749.
25. **Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA.** 2011. Blast Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics **12:**402. http://dx.doi.org/10.1186/1471-2164-12-402.
26. **Beres SB, Sylva GL, Barbian KD, Lei B, Hoff JS, Mammarella ND, Liu MY, Smoot JC, Porcella SF, Parkins LD, Campbell DS, Smith TM, McCormick JK, Leung DY, Schlievert PM, Musser JM.** 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:**10078–10083. http://dx.doi.org/10.1073/pnas.152299499.
27. **Holden MT, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, Dowd L, Feltwell T, Hamlin N, Holroyd S, Jagels K, Moule S, Mungall K, Quail MA, Price C, Rabbinowitsch E, Sharp S, Skelton J, Whitehead S, Barrell BG, Kehoe M, Parkhill J.** 2007. Complete genome of acute rheumatic fever-associated serotype M5 *Streptococcus pyogenes* strain Manfredo. J Bacteriol **189:**1473–1477. http://dx.doi.org/10.1128/JB.01227-06.
28. **Parkhill J, Wren BW.** 2011. Bacterial epidemiology and biology lessons from genome sequencing. Genome Biol **12:**230. http://dx.doi.org/10.1186/gb-2011-12-10-230.
29. **Canchaya C, Proux C, Fournous G, Bruttin A, Brüssow H.** 2003. Prophage genomics. Microbiol Mol Biol Rev **67:**238–276. (Erratum, **67:** 473, 2003.) http://dx.doi.org/10.1128/MMBR.67.2.238-276.2003.
30. **Bao Y, Liang Z, Booyjzsen C, Mayfield JA, Li Y, Lee SW, Ploplis VA, Song H, Castellino FJ.** 2014. Unique genomic arrangements in an invasive serotype M23 strain of *Streptococcus pyogenes* identify genes that induce hypervirulence. J Bacteriol **196:**4089–4102. http://dx.doi.org/10.1128/JB.02131-14.
31. **Marciel AM, Kapur V, Musser JM.** 1997. Molecular population genetic analysis of a *Streptococcus pyogenes* bacteriophage-encoded hyaluronidase gene: recombination contributes to allelic variation. Microb Pathog **22:** 209–217. http://dx.doi.org/10.1006/mpat.1996.9999.
32. **Mylvaganam H, Bjorvatn B, Hofstad T, Osland A.** 2001. Molecular characterization and allelic distribution of the phage-mediated hyaluronidase genes *hylP* and *hylP2* among group A streptococci from western Norway. Microb Pathog **30:**311. http://dx.doi.org/10.1006/mpat.2001.0445.
33. **Henningham A, Yamaguchi M, Aziz RK, Kuipers K, Buffalo CZ, Dahesh S, Choudhury B, Van Vleet J, Yamaguchi Y, Seymour LM, Ben Zakour NL, He L, Smith HV, Grimwood K, Beatson SA, Ghosh P, Walker MJ, Nizet V, Cole JN.** 2014. Mutual exclusivity of hyaluronan and hyaluronidase in invasive group A *Streptococcus*. J Biol Chem **289:** 32303–32315. http://dx.doi.org/10.1074/jbc.M114.602847.
34. **Schommer NN, Muto J, Nizet V, Gallo RL.** 2014. Hyaluronan breakdown contributes to immune defense against group A *Streptococcus*. J Biol Chem **289:**26914–26921. http://dx.doi.org/10.1074/jbc.M114.575621.