Unique Genomic Arrangements in an Invasive Serotype M23 Strain of *Streptococcus pyogenes* Identify Genes That Induce Hypervirulence

Yunjuan Bao,[a,b] Zhong Liang,[a] Claire Booyjzsen,[a] Jeffrey A. Mayfield,[a] Yang Li,[b] Shaun W. Lee,[c] Victoria A. Ploplis,[a] Hui Song,[b] Francis J. Castellino[a]

W. M. Keck Center for Transgene Research and Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana, USA[a]; Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China[b]; Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA[c]

The first genome sequence of a group A *Streptococcus pyogenes* serotype M23 (*emm23*) strain (M23ND), isolated from an invasive human infection, has been completed. The genome of this opacity factor-negative (SOF⁻) strain is composed of a circular chromosome of 1,846,477 bp. Gene profiling showed that this strain contained six phage-encoded and 24 chromosomally inherited well-known virulence factors, as well as 11 pseudogenes. The bacterium has acquired four large prophage elements, ΦM23ND.1 to ΦM23ND.4, harboring genes encoding streptococcal superantigen (*ssa*), streptococcal pyrogenic exotoxins (*speC*, *speH*, and *speI*), and DNases (*spd1* and *spd3*), with phage integrase genes being present at one flank of each phage insertion, suggesting that the phages were integrated by horizontal gene transfer. Comparative analyses revealed unique large-scale genomic rearrangements that result in genomic rearrangements that differ from those of previously sequenced GAS strains. These rearrangements resulted in an imbalanced genomic architecture and translocations of chromosomal virulence genes. The *covS* sensor in M23ND was identified as a pseudogene, resulting in the attenuation of *speB* function and increased expression of the genes for the chromosomal virulence factors multiple-gene activator (*mga*), M protein (*emm23*), C5a peptidase (*scpA*), fibronectin-binding proteins (*sfbI* and *fbp54*), streptolysin O (*slo*), hyaluronic acid capsule (*hasA*), streptokinase (*ska*), and DNases (*spd* and *spd3*), which were verified by PCR. These genes are responsible for facilitating host epithelial cell binding and and/or immune evasion, thus further contributing to the virulence of M23ND. In conclusion, strain M23ND has become highly pathogenic as the result of a combination of multiple genetic factors, particularly gene composition and mutations, prophage integrations, unique genomic rearrangements, and regulated expression of critical virulence factors.

Streptococcus pyogenes (group A streptococcus [GAS]) is a pathogenic low-G+C-content beta-hemolytic Gram-positive bacterium (1). GAS is responsible for ~700 million infections worldwide per year (2), ranging from simple pharyngitis and impetigo to more invasive life-threatening infections that include necrotizing fasciitis and streptococcal toxic shock syndrome (3). Acute rheumatic fever (ARF) and acute glomerulonephritis are among the more serious nonsuppurative sequelae that can result from GAS infection (4).

Currently, >200 strains of *S. pyogenes* have been identified by *emm* genotyping (5). The genomes of 20 strains that span a range of serotypes have been completely sequenced and assembled. Of the many known virulence factors of this bacterial strain, e.g., C5a peptidase (*speB*), the hyaluronic acid (*hya*) capsule, streptolysin S (*sls*), and streptolysin O (*slo*), the M protein, which is encoded by the *emm* gene, is one of the most important features of this group of bacteria (6). The M protein is composed of multiple N-terminal A and B modules, which are highly variable among M types, along with well-conserved C-terminal C and D modules. The N-terminal A region is the most variable of these domains (7), thus rendering it suitable for use for distinct serotyping. In this regard, GAS strains are accordingly serologically classified as different M types on the basis of the first ~50 amino acid residues of this hypervariable N terminus (8).

In addition to the *emm* gene, up to two additional subfamily *emm*-like genes, *fcR* and *enn*, that encode the Fc binding regions of IgG and IgA, respectively, have been found in a cluster in different GAS strains. This *emm* subfamily of genes can be identified through the nucleotide sequences of their 3' peptidoglycan-spanning domains (9, 10). Not all of these genes are present in every

GAS strain, and their presence and chromosomal arrangement have been used to further map GAS strains as chromosomal patterns A to E, with the intent of correlating these genomic patterns with tissue tropism and virulence (11). Patterns A to C are associated with pharyngeal disease, pattern D is associated with skin disease, and pattern E is associated with both (12). In addition, two additional classes of GAS strains have been designated. Class 1 is a serum opacity factor (*sof* and *sfbII*)-negative (SOF⁻) strain that ordinarily contains the novel *fbpB* genotype of a fibronectin-binding protein (FBP), *prtF2* (13–15). Additionally, the genomes of some of these strains contain the FBP *sfbI* gene, which is an important determinant for epithelial cell (EpC) binding and invasion (16). These strains are also linked to a surface-exposed antigen in the C-repeat region of the M protein, which interacts with ARF antibodies (17). Class II SOF (*sof*, or *sfbII*)-positive (SOF⁺) GAS strains normally contain the *fbpB* genotype of *prtF2* (6, 14). Further, the genomes of many of these strains also contain the FBP genes *sfbI* (*prtF1*) (18) and *sfbX*, the latter of which is located immediately downstream of the *sof* (*sbfII*) bicistronic tandem

(19). SOF$^+$ strains generally do not immunoreact with ARF antibodies to the class I C-repeat surface of the M protein (17). Many class I and class 2 strains contain a gene for another FBP, *fbpA*, which is regulated by the one-component multiple-gene activator (Mga) and exists as a member of its *cis* regulon (14). Most SOF$^-$ and SOF$^+$ strains display FBP54 (*fbp54*), which is also a fibronectin/fibrinogen binding protein (20, 21).

The M protein is most prominently regulated by the first component (Mga) of the *mga* regulon, formerly known as *vir* or *mry* (22, 23). Mga is maximally expressed during the logarithmic growth phase (LP) in response to changing environmental conditions, e.g., temperature, pH, CO_2 levels, and/or iron concentration (24). The largest *mga cis*regulon is present in pattern D strains (15) and consists of the following sequential order of genes: *mga-fcR-emm-enn-scpA* (C5a peptidase)-*fbpA*. Expression of all of these genes is controlled by Mga (15, 25, 26). Mga also regulates in *trans* a number of other GAS proteins, e.g., streptococcal inhibitor of complement (*sic*), and the bicistronic FBPs *sof-sbfX* (19, 27). *mga* expression is, in turn, regulated by itself (Mga), as well as by the transcriptional regulator genes *rgg* (*ropB*) and *rofA* (*nra*) (23).

Thus, GAS produces numerous proteins and regulatory systems that are strain specific and help the bacteria circumvent the host innate immune system. Of the known virulence-enhancing extracellular proteins produced by GAS, many are capable of triggering a severe nonspecific host immune response (28). Because of the large number of strains and the broad variation in their properties and characteristics, a large variance in virulence is displayed phenotypically as different levels of severity of infection and invasiveness. Therefore, sequencing of the full genomes of GAS strains, particularly those harboring different M types, is integral to understanding the nature of this organism. This, in turn, is critical to developing effective strategies for the host against this bacterium. Most importantly, such knowledge will allow us to understand GAS evolution and development, which will allow the assessment and prediction of important patterns in GAS phylogeny. Toward this end, we have sequenced the full genome of *S. pyogenes* strain M23ND, the first serotype M23 strain to be reported in this manner. The circular genome of this strain possesses ∼1.85 Mbp and appears to have a high rate of genetic recombination. As a unique isolate with many unusual properties, M23ND offers an excellent opportunity to examine some of the more variable and subtle characteristics of GAS that are associated with severe GAS infections.

MATERIALS AND METHODS

GAS strain. *S. pyogenes* strain ATCC 21059 is a serotype M23 GAS strain. This bacterium was isolated as strain Sv in 1965 from a patient with a case of severe streptococcal disease (29). We refer to this fully sequenced strain as M23ND.

Strain handling. The GAS isolate was cultivated from glycerol stock cultures that had been grown on blood agar at 37°C in 5% CO_2 for 24 h. Genomic DNA (gDNA) was extracted using a mini-DNA kit (Qiagen, Valencia, CA).

Genome sequencing and gene annotation. The entire genome of M23ND was sequenced using an Illumina Miseq sequencer (Illumina, CA) with read lengths of 150 bp on both strands and 454 pyrosequencing (Roche 454 Life Science, Basel, Switzerland). A high-quality draft genome assembly which comprised eight scaffolds was obtained. The gaps were closed using PCR primer walking, and the complete circular genome was derived. The protein-coding sequences were predicted using the Glimmer (version 3.02b) program (30), the rRNA sequences were predicted

using the RNAmmer server (31), and tRNA sequences were detected by use of the tRNAScan-SE server (31). Genome annotation was performed using the automated RAST annotation server (32) and manual curation.

Comparative analysis of GAS strains. Genome sequences for the 20 fully sequenced GAS strains currently available were downloaded from the NCBI genome database (33).

The genome sequence of *S. pyogenes* M23ND was compared to the 20 other whole-genome GAS sequences publicly available using images generated by the BLAST Ring Image Generator (BRIG) (34). The comparative genomic architecture of the GAS genomes was determined by BLASTn analysis and graphically represented using the Artemis comparison tool (ACT) (35). Visualization of phage locations was implemented using the Geneious (version 7.0.6) program (Biomatters, Auckland, New Zealand).

Phylogenetic analyses. Phylogenetic analyses were conducted on the basis of multilocus sequence typing (MLST) (36) of seven housekeeping marker genes, *viz.*, those for glucose kinase (*gki*), glutamine transporter protein (*gtr*), glutamate racemase (*murI*), DNA mismatch repair (MRR) protein (*mutS*), transketolase (*recP*), xanthine phosphoribosyltransferase (*xpt*), and acetyl coenzyme A acetyltransferase (*yqiL*), as well as for the detection of single nucleotide polymorphisms (SNPs) in virulence genes. The construction and visualization of phylogenetic trees or networks were implemented in the SplitsTree program (37).

Quantitative reverse transcription-PCR (qRT-PCR). Total gDNA was isolated from a selected single colony (of 10) of M23ND, after growth rates and the presence of some known genes were checked. The sample was cultured overnight at 37°C with 5% CO_2 in Todd-Hewitt broth (BD Biosciences, San Jose, CA) supplemented with 10% yeast extract (THY). The gDNA was isolated after treatment of the cells with lysozyme-proteinase K and cell lysis buffer (100 mM Tris, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, pH 8.5) and extracted with phenol-chloroform-isoamyl alcohol (25:24:1, vol/vol/vol). The gDNA was precipitated with isopropanol and washed with 70% ethanol. Approximately 100 ng of gDNA was used in a 30-μl PCR mixture for gene validation.

Total RNA was isolated from single colonies grown to logarithmic growth phase (LP; A_{600} = 0.5 to 0.6) or stationary growth phase (SP; A_{600} > ∼1.0) at 37°C with 5% CO_2 in THY medium as described earlier. Two treatments with DNase I were performed to eliminate any DNA contamination (15). Approximately 50 ng total RNA was used in a 30-μl reverse transcription-PCR (RT-PCR) mixture for gene transcription validation. A sample without reverse transcriptase added was used as the negative control. Primers (see Table S1 in the supplemental material) specific for the coding sequences were designed to obtain a PCR amplicon of ∼200 to 500 bp for both gDNA detection and mRNA transcription. Relative gene expression levels were analyzed by the $2^{-\Delta\Delta C_T}$ threshold cycle (C_T) method (15). Triplicate measurements of the C_T number were collected for each chosen gene relative to the C_T number for the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) housekeeping gene for both mutant *covS* and wild-type *covS* strains grown to LP and SP. The ratio of the level of expression of each gene relative to the level of expression of the mutant *covS* strain grown to LP was calculated.

Nucleotide sequence accession number. The genomic sequence of M23ND has been deposited in the NCBI GenBank under accession number CP008695.

RESULTS

General background of GAS strain M23ND. GAS strain M23ND was originally isolated from a case of severe streptococcal infection for its cytotoxic anticancer properties by virtue of cytolytic streptolysin S (SLS) production (U.S. patents 3,477,914 and 4,328,218 [38, 39]). At that time it was taxonomically classified as *Streptococcus haemolyticus* and designated strain Sv (ATCC 21059) (29, 40, 41). Subsequently, strain Sv was serologically identified as having a type 23 M protein, which was confirmed by comparison of its amino acid sequence with that of M23 isolate Memphis deposited
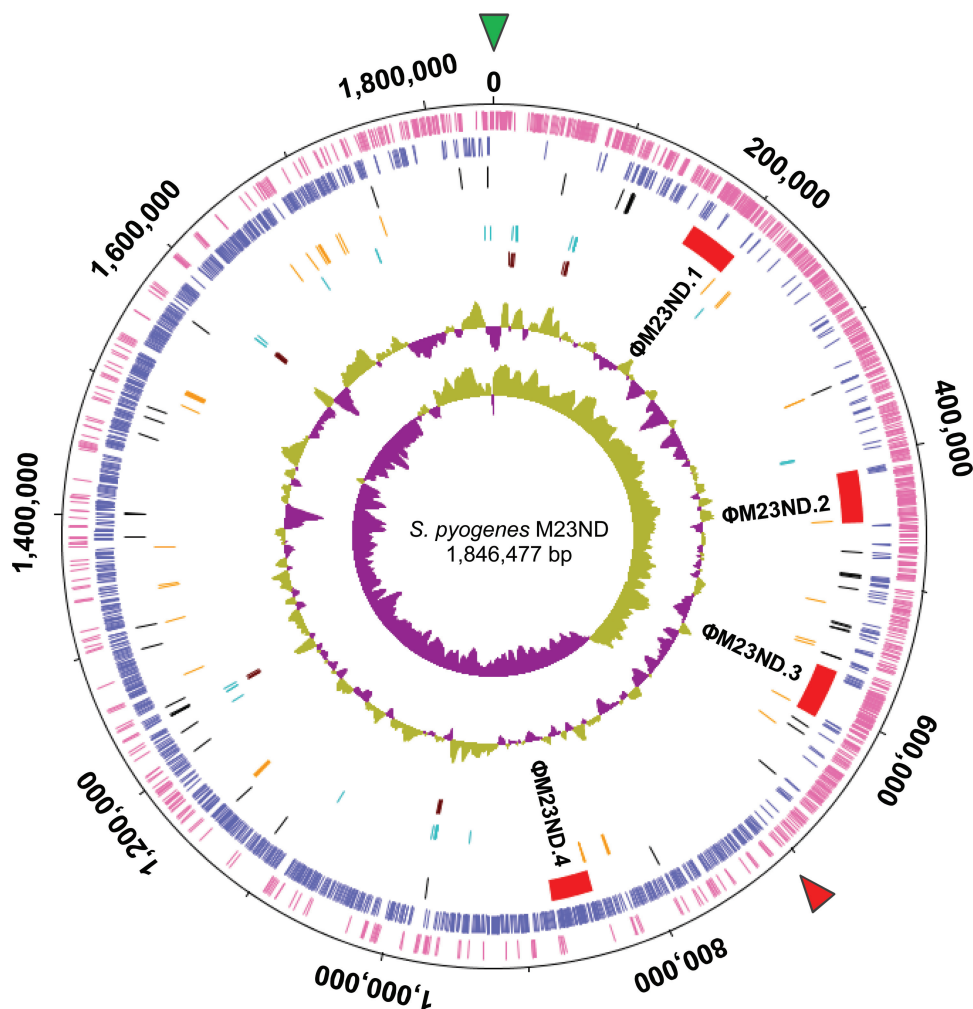
FIG 1 Circular representation of the 1,846,477-bp genome of *S. pyogenes* strain M23ND. Data are shown from the outermost to the innermost circles. Circles 1 and 2 display annotated coding sequences for the reverse (pink) and forward (blue) DNA strands, respectively. Circle 3 shows the locations of four phage elements, ΦM23ND.1 to ΦM23ND.4 (red boxes), and short mobile elements (black lines). Circle 4 illustrates the virulence genes identified in M23ND (orange). Circles 5 and 6 represent the locations of tRNA (turquoise) and rRNA (brown) genes, respectively. The genome contains 57 tRNA genes and 5 rRNA operons. Circles 6 and 7 display the GC content and GC skew ([G − C]/[G + C]). The triangles on the outermost circumference indicate the positions of the replication origin (*ori*; green triangle) at bp 0 and the replication terminus (*ter*; red triangle) at bp 702443. The genome contains 1,851 ORFs, 231 of which are present in phage regions. Approximately 1,389 genes have assigned functions.

in GenBank (GenBank accession number U11953). Additionally, the *mga* regulon of the Sv strain was classified as the small variety. The 4,791-bp incomplete sequence nonetheless allowed identification of the gene sequence to be 5′-*mga*-*emm*-*scpA*-3′, and thus, Sv represents a pattern A strain (12).

Overall features of genome sequence of GAS strain M23ND. The genome of *S. pyogenes* M23ND consists of a single circular chromosome of 1,846,477 bp with an average G+C content of 38.6% (Fig. 1). The genome contains 1,851 predicted open reading frames (ORFs) and 11 pseudogenes which, in total, cover 86.3% of the entire genome sequence. This M23 strain contains 44 unique and unidentified ORFs on the basis of a minimum 85% homology by BLASTn comparison. Approximately 12% of the ORFs are contained in four prophage segments which account for 8.5% of the genome. The ORFs within the prophage regions encode multiple known or putative virulence factors that include toxins, superantigens, mitogenic factors, and fibronectin-binding proteins.

Comparative study of GAS strain M23ND and characterization of phage elements contributing to its genetic diversity. Comparative genome mapping (Fig. 2) illustrates the genomic profile of M23ND in comparison with the profiles of the 20 other previously fully sequenced GAS strains, the properties of which are summarized in Table 1. This comparison showed that the sequences are conserved throughout the genome, except at the points of insertion of short mobile genetic elements and at four large prophage regions where the genome sequences exhibit mosaic profiles across the GAS strains compared. The short mobile elements of M23ND have orthologs in only some of the other strains. However, the average homologies for the prophage sequences aligned between different strains (91 to 94%) are lower than those for nonprophage sequences (>98%). It generally appears that mobile genetic elements, including prophages, are major sources of genetic diversity among different GAS strains.

In order to more fully assess the contributions of M23ND prophage elements to genetic diversity, we identified the locations
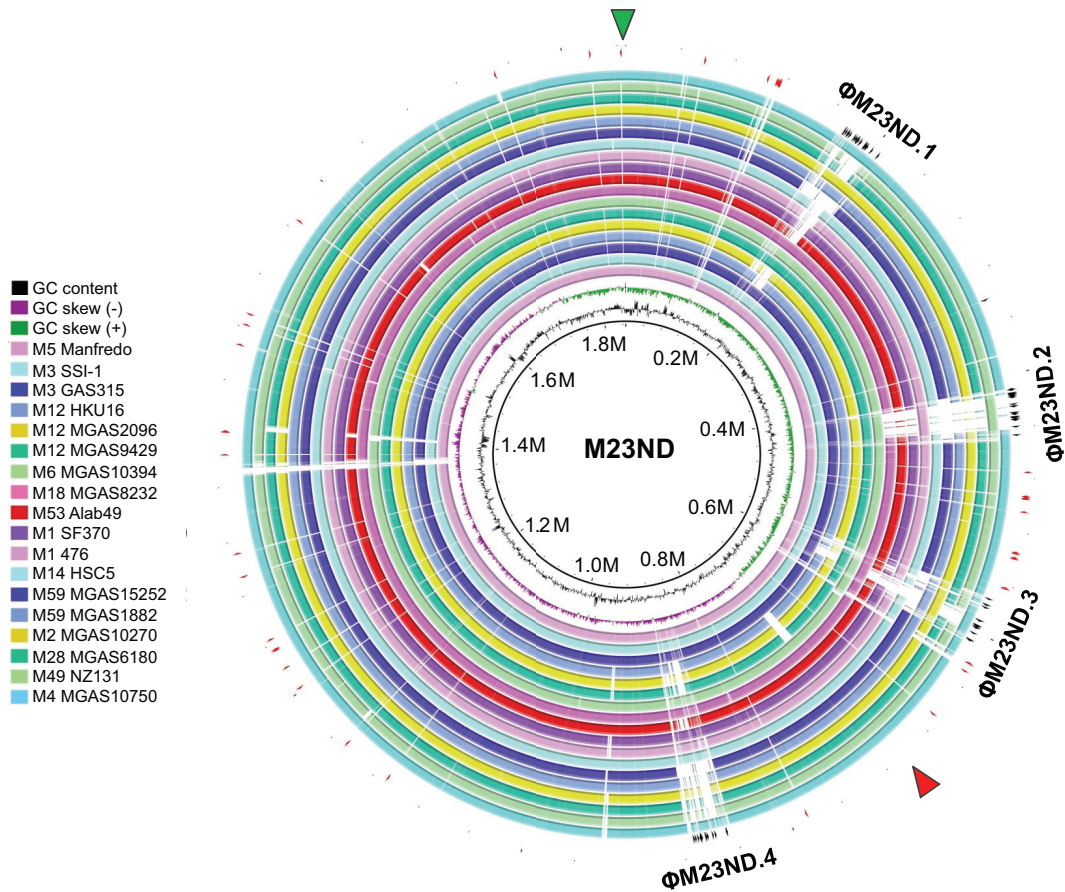
FIG 2 Global comparisons of fully sequenced *S. pyogenes* genomes. The sequences of the 20 previously sequenced GAS genomes were obtained from the NCBI database and were correlated with a basic local alignment with *S. pyogenes* strain M23ND (innermost circle). The areas of similarity and divergence within the sequences are contrasted, with areas with white gaps indicating the regions of the highest variance. The profiles of M1 strain 5005 and M1 strain A20 are similar to the profile of M1 strain 476 and are therefore not included. Phage proteins and short mobile genetic elements are indicated by black and red arrows, respectively, in the outer circle. The strains whose sequences were compared to the sequence of M23ND, which is positioned on the central ring, are M1 strain SF370, M1 strain 476, M2 strain 10270, M3 strain 315, M3 strain SSI-1, M4 strain 10750, M5 strain Manfredo, M6 strain 10394, M12 strain 9429, M12 strain 2096, M12 strain HKU16, M14 strain HSC5, M18 strain 8232, M28 strain 6180, M49 strain NZ131, M53 strain Alab49, M59 strain 1882, and M59 strain 15252. Numbers in the innermost circle are positions. M, million bases.

and lengths of the prophage segments based on gene annotations and BLAST comparisons (Table 2 and Fig. 3A). There are four large prophage-related elements, designated ΦM23ND.1 to ΦM23ND.4, with individual lengths ranging from 34 to 42 kbp. Three of the four prophages are distributed within the clockwise half circle proximal to the origin of replication (*ori*) site of the genome. Comparisons with other sequenced *S. pyogenes* strains show that prophages are generally inserted in the second replichore. However, for M23ND, three of the four prophages are inserted in the first replichore, thus highlighting the imbalanced architecture of the M23ND strain. Four phage integrase genes are situated on the terminal ends of each of these prophages and are located at bp 168633 to 170048, 409285 to 410331, 572756 to 573884, and 875643 to 876785. Only the last prophage carries most genes in the forward direction (Fig. 1).

In addition, the mismatch repair (MRR) genes *mutS* (bp 1789917 to 1792472) and *mutL* (bp 1792601 to 1794583), present on a polycistronic gene cluster and controlled by a single promoter, are uninterrupted in M23ND, such that both genes are expressed throughout its growth cycle. This assists in maintenance of the genetic material of

this strain and limits the mutability of the already virulent and established M23ND. This is unlike the less virulent SF370 strain, which contains a prophage (ΦSF370.4) insertion between *mutS* and *mutL* that inactivates *mutL* in a growth-dependent manner and allows early-growth-phase mutations to occur, perhaps randomly generating more virulent strains (42).

The prophages in M23ND have genomic integration locations different from those of their orthologs in other strains (Fig. 3A). For example, prophage ΦM23ND.1 is very similar to ΦManfredo.4 (M5), ΦMGAS10394.3 (M6), ΦMGAS1882.2 (M18), and those in several other strains, but this prophage is inserted at different genomic locations in these strains. Prophage ΦM23ND.3 shares high homology with ΦM23ND.1 but is integrated at two distinct sites. This implies that these two prophages arose from a common ancestor and likely underwent horizontal gene transfer within or between strains. These intra- and interstrain recombination events suggest the important role of horizontal gene transfer in shaping the genetic diversity of *S. pyogenes* evolution (43, 44).

An immune defense system, clustered regulatory interspaced short palindromic repeats (CRISPRs), has recently been identified

TABLE 1 Characteristics of fully sequenced *S. pyogenes* genomes[a]

Strain	M type	RefSeq accession no.	Clinical source	Genome size (bp)	No. of: Prophages	Genes	Proteins	% GC content
476	1	NC_020540	STSS	1,831,128	3	1,644	1,571	38.5
5005	1	NC_007297	CSF	1,838,554	3	1,950	1,865	38.5
A20	1	NC_018936	Blood, NF	1,837,281	3	1,915	1,828	38.5
SF370	1	NC_002737	Wound infection	1,852,441	4	1,810	1,696	38.5
10270	2	NC_008022	Pharynges	1,928,252	5	2,068	1,987	38.4
315	3	NC_004070	STTS	1,900,521	6	1,951	1,865	38.6
SSI-1	3	NC_004606	STTS	1,894,275	6	1,931	1,859	38.6
10750	4	NC_008024	Pharynges	1,937,111	4	2,059	1,978	38.3
Manfredo	5	NC_009332	ARF	1,841,271	5	1,906	1,745	38.6
10394	6	NC_006086	Pharynges	1,899,877	8	1,971	1,886	38.7
2096	12	NC_008023	Glomerulonephritis	1,860,355	2	1,979	1,898	38.7
9429	12	NC_008021	Infection	1,836,467	3	1,962	1,877	38.5
HKU16	12	NZ_AFRY01000001	Blood, scarlet fever	1,908,100	3	1,950	1,865	38.4
HSC5	14	NC_021807	Infection	1,818,351	3	1,829	1,744	38.5
8232	18	NC_003485	ARF	1,895,017	5	1,924	1,839	38.5
M23ND	**23**		**Invasive infection**	**1,846,477**	**4**	**1,851**	**1,620**	**38.6**
6180	28	NC_007296	Invasive infection	1,897,573	4	1977	1894	38.4
NZ131	49	NC_011375	Glomerulonephritis	1,815,785	3	1,788	1,700	38.6
Alab49	53	NC_017596	Impetigo lesion	1,827,308	4	1,866	1,773	38.6
1882	59	NC_017053	Infection	1,781,029	2	1,792	1,691	38.5
15252	59	NC_017040	Invasive infection	1,750,832	2	1,757	1,662	38.5

[a] Boldface data represent data for strain M23ND from the current study. CSF, cerebrospinal fluid; NF, necrotizing fasciitis; STTS, streptococcal toxic shock; ARF, acute rheumatic fever.

in many bacterial strains, including *S. pyogenes*, to be a protective mechanism against exogenous prophages or plasmid DNAs (45–47) via inhibition of the acquisition of phage elements (48). Considering the prevalence of phages in the genome of M23ND, it was of interest to determine whether the CRISPR system was involved in the protection system of this strain. Remarkably, we did not identify CRISPR sequences or the CRISPR-associated (*cas*) genes in M23ND. Further examination of genome sequences of the other 20 sequenced GAS strains showed that several other strains from different serotypes do not contain the CRISPR sequences, including M5 strain Manfredo, M6 strain 10394, M18 strain 8232, M53 strain Alab49, M3 strain SSI-1, M3 strain 315, and M14 strain

HSC5. These strains also do not contain the *cas* genes, except for the two serotype M3 strains, SSI-1 and 315. The absence of the CRISPR system in GAS strains represented the clustering of the phylogenetically related strains (see additional details below); i.e., M23ND was closest to M5 strain Manfredo and M6 strain 10394 and shared with them a common ancestor, M18 strain 8232; M53 was closest to the two M3 strains (strains SSI-1 and 315), and they are located in the same evolutionary branch; M14 strain HSC5 was itself in a separate branch. This indicated that CRISPRs may represent a type of evolutionary product inherited via selective pressure. The absence of CRISPRs in some GAS strains and their specific amounts of phage elements could be a balanced result of the

TABLE 2 Phage elements in M23ND and their orthologs in other GAS strains

GAS strain	Ortholog in phage element[a]: ΦM23ND.1 (168433–209722)	ΦM23ND.2 (409085–450619)	ΦM23ND.3 (572556–612004)	ΦM23ND.4 (842696–876985)
M1 SF370				ΦSF370.3
M2 10270	Φ10270.1		Φ10270.1	
M3 315	Φ315.2	Φ315.6	Φ315.2	Φ315.3
M3 SSI-1	ΦSSI.5	ΦSSI.1	ΦSSI.5	ΦSSI.4
M4 10750	Φ10750.1		Φ10750.1	Φ10750.3
M5 Manfredo	ΦManfredo.4	ΦManfredo.3	ΦManfredo.4	ΦManfredo.2
M6 10394	Φ10394.3	Φ10394.1	Φ10394.3	Φ10394.5
M12 2096		Φ2096.3		
M12 9429	Φ9429.1	Φ9429.3	Φ9429.1	
M12 HKU16		ΦHKU16.2		
M14 HSC5				ΦHSC5.2
M18 8232	Φ8232.2		Φ8232.2	Φ8232.5
M49 NZ131		ΦNZ131.2		
M59 1882	Φ1882.1		Φ1882.1	

[a] Phage element ΦM23ND.1 carries the gene for a mitogenic factor (*spd1*) and the gene for exotoxin C (*speC*), ΦM23ND.2 carries the gene for superantigen A (*ssa*), ΦM23ND.3 carries the gene for a mitogenic factor (*spd3*), and ΦM23ND.4 carries the genes for exotoxin types I and H (*speI* and *speH*). The genomic locations of phage elements in M23ND are denoted in parentheses.
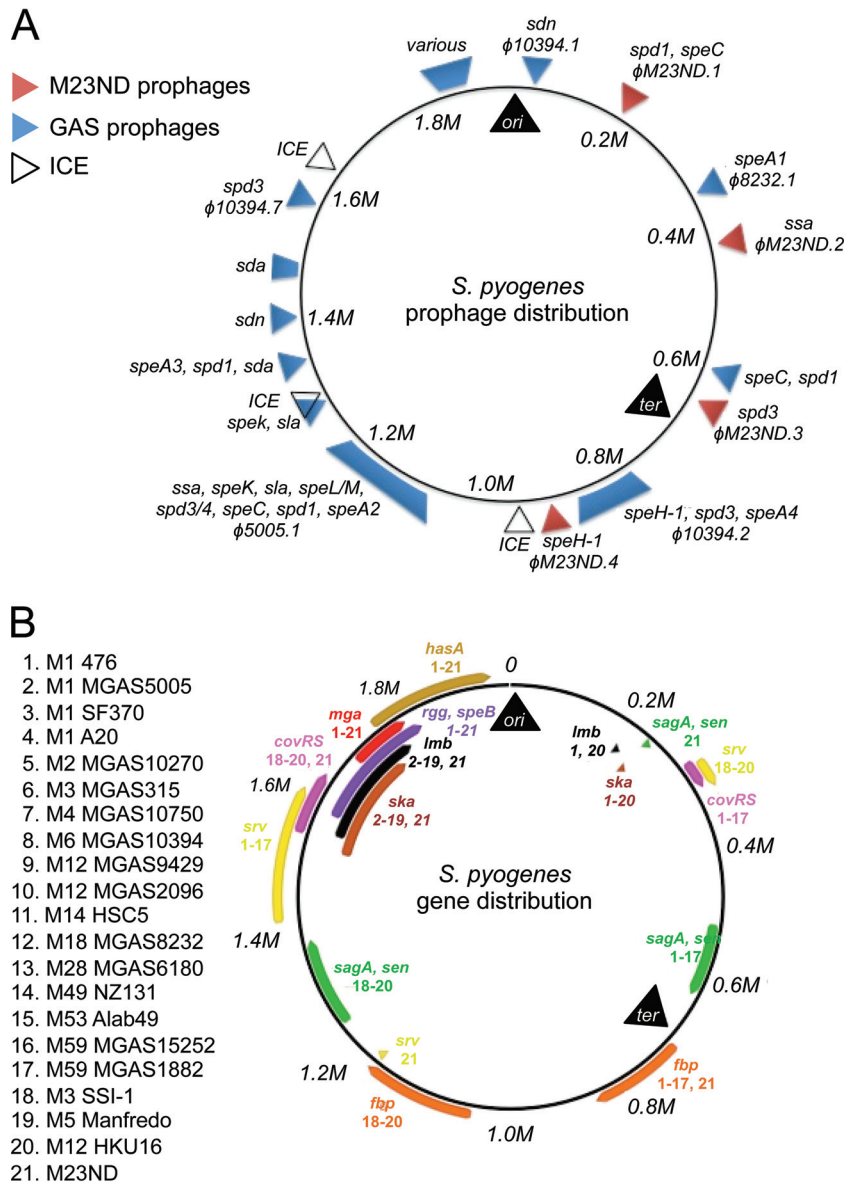
FIG 3 DNA characterizations of genomic sequences of *S. pyogenes*. (A) A circular visualization of the locations of phage elements across M23ND (red) in comparison with the locations in the 20 previously fully sequenced GAS genomes available through NCBI (blue). The integration sites of phages are generally clustered at several regions, but the locations of orthologous phages are nonconserved. For example, prophage ΦM23ND.1 is closely similar to ΦManfredo.4 (M5), ΦMGAS10394.3 (M6), ΦMGAS1882.2 (M18), and several other prophages, but these prophages are inserted at distinct sites, suggesting that phage recombination via horizontal gene transfer plays an important role in genetic diversity. Black triangles on the circumference, positions of the replication origin (*ori*) at bp 0 and terminus (*ter*) at bp 702433. Inversions around *ori* (white triangle) are clearly observed, especially for *srv*, and inversions around *ter* (white triangle) are readily seen in the cases of *sen*, *sagA*, and *fbp* (*fbp54*). ICE, integrating conjugative elements. (B) A circular visualization of the comparative locations of the virulence factors of interest within the genomes of the 21 fully sequenced GAS genomes. Elongated bars, regions in which each gene can be found across all of the 21 fully sequenced and assembled GAS genomes, unless the gene appears elsewhere, e.g., *covRS* for SSI-1, Manfredo, HKU16, and M23ND or *sen*, *srv*, and *sagA* for M23ND.

advantages of enhanced virulence and the disadvantages of excess foreign toxins from phage acquisition.

Sources of genomic diversity in *S. pyogenes* M23ND: extensive unique genomic arrangements and imbalanced genomic architecture. Further inspection of the comparative genomic architectures between the fully sequenced GAS strains (see Fig. S1 in the supplemental material) showed that strain M23ND exhibited unique genomic rearrangements compared with those of M5 strain Manfredo, M12 strain HKU16, and M3 strain SSI-1, which

contain similar central inversions across the replication origin (*ori*) and *dif*-like replication terminus (*ter*) axes (49). M23ND differs from these three strains, in that the *dif*-like replication terminus (*ter*) site was found to be located within the half circle of the genome (at bp 702433) and the inversions and translocations are asymmetrical across the *ori-ter* axes. Additionally, M23ND exhibits gene translocations over and above those found in M3 strain SSI-1, M5 strain Manfredo, and M12 strain HKU16. For example, the critical gene *sagA*, which encodes SLS, responsible for the beta-
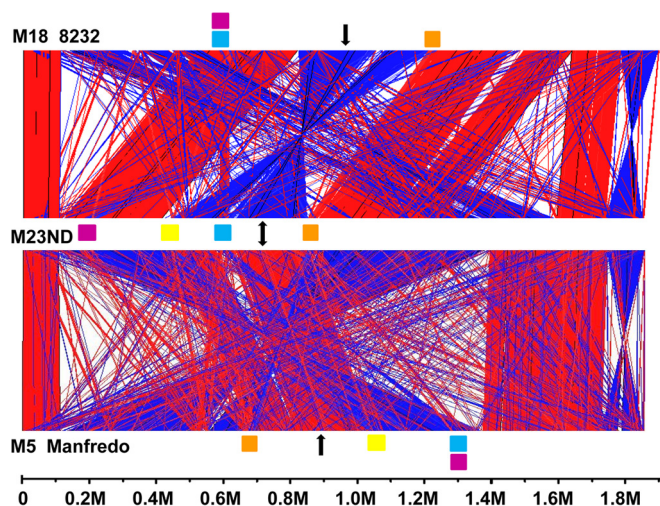
FIG 4 Comparisons of the whole genome of *S. pyogenes* M23ND with the whole genomes of two phylogenetic neighbors, M5 strain Manfredo and M18 strain 8232. Red and blue lines, forward and reverse alignments, respectively; colored boxes, prophage elements; arrows, replication terminus in each genome. Large segmental inversions and translocations are observed in M23ND relative to the genomic architectures of its neighbors. These rearrangements result in an imbalanced global genomic architecture, where the replication terminus (*ter*) and three prophage insertions were located within the same replichore of the chromosome. M23ND also contains an additional short inversion in the final 100 kb of the sequence. Prophage elements are located on the breakpoints of the rearrangements or within the rearrangements themselves.

hemolytic characteristic of GAS, is found in flanking locations of the *ter* site on the various fully sequenced genomes, due to inversion around the *ori-ter* axes (Fig. 3B). It is symmetrical for all genomes, except for that of M23ND, due to the nonsymmetrical *ter* with respect to *ori*. A similar situation exists for *sen*, a gene essential for GAS survival. For *srv*, a transcriptional regulator of virulence, it would appear that inversions around a symmetrical *ori-ter* placed this gene on the opposite replichore for M3 strain SSI-1, M5 strain Manfredo, and M12 strain HKU16, and then another inversion in the last three genomes around *ter* of M23ND uniquely placed *srv* in M23ND. These examples highlight the unique extent of gene translocations in strain M23ND. Of particular note, the M23ND genome also contains an extra short inversion within the last 100 kb of the chromosome (see Fig. S1 in the supplemental material). This feature has not been seen previously in any other sequenced GAS genome. These data suggest that M23ND may share an ancestor with M3 strain SSI-1, M5 strain Manfredo, and M12 strain HKU16 but experienced distinct rearrangement events and, thus, underwent a disparate evolutionary path.

By comparing the genome sequence of M23ND with the genome sequences of M5 strain Manfredo and the representative M18 strain 8232, the genomes of which do not exhibit significant rearrangements, we found that the prophage segments are located at the breakpoints or within the rearranged genomic regions (Fig. 4). This characteristic has also been reported previously for another GAS strain, M3 strain SSI-1 (50). It was proposed that the genomic rearrangements were triggered by prophage recombinations to balance the global genomic architecture (50). However, in M23ND, the clustering of all four prophages and the *ter* site within a single half

circle of the M23ND genome has disrupted the balance of the global genomic architecture. The replichores in the clockwise and counter-clockwise directions are unequal in length. This may point to a positive correlation between an unbalanced replichore architecture and the severity of invasive infection induced by M23ND. Previous research using *Escherichia coli* found that such an imbalance can affect bacterial growth and fitness (51, 52). However, it is difficult to establish a firm relationship between this genotype and clinical phenotypes displayed in GAS, due to limited examples of imbalanced replichores in the currently available fully sequenced genomes.

Profiling of the prophage-carried virulence genes of *S. pyogenes*. Virulence factors in GAS genomes are major contributors to the pathogenesis of *S. pyogenes*. A variety of these factors have been identified in previous GAS studies to be phage carried or chromosomally inherited. We investigated a total of 38 virulence factors of interest to our studies, of which 30 were found to be present in M23ND (see Table S2 in the supplemental material). We profiled all of these virulence genes in M23ND and compared them with those in other GAS strains to assess their patterns in genomic distributions.

M23ND contains six well-established virulence genes carried by prophage elements. Specifically, the *ssa* gene and genes for exotoxin types I and H (*speI* and *speH*) were carried by ΦM23ND.2 and ΦM23ND.4, respectively, while the gene for exotoxin C (*speC*) was carried by ΦM23ND.1. The gens for two mitogenic factors, *spd1* and *spd3*, were carried by two related prophages, ΦM23ND.1 and ΦM23ND.3, respectively. The acquisition of such toxins and pyogenic genes has been reported to correlate with severe invasive infections and epidemic outbreaks of *S. pyogenes* (53). Very likely, the carriage of the combination of several virulence factors (*ssa*, *speC*, *speI*, and *speH*) and endonucleases (*spd1* and *spd3*) in strain M23ND is a major factor responsible for the virulence of this strain (54). M23ND is one of few fully sequenced strains, in addition to strains of the M3, M4, and M12 lineages, to carry *ssa*.

Despite the properties of the phage-carried virulence factors, their spatial distribution across species is discordant with that of the carrier prophages (Fig. 3B; see also Table S2 in the supplemental material). Generally, each virulence factor could be carried by divergent prophage elements integrated at diverse sites across different GAS strains. In order to extend this finding to other phage-carried virulence factors, *viz.*, *sda*, *speA*, *speK*, *slaA*, *speL*, and *speM*, we mapped the locations of all of the established virulence genes carried by phages across the 21 fully sequenced GAS strains (Fig. 5). It was found that the virulence genes are scattered nearly randomly throughout the chromosomes, although strains of the same M type tend to cluster these genes in similar areas of the chromosome. However, there are currently too few fully sequenced and assembled *S. pyogenes* genomes to make any definitive conclusions regarding this point. These phenomena reflect the complex recombinative evolution of prophage and virulence genes carried by phages in GAS strains and underscore their dominant contributions to genetic diversity.

Profiling of chromosomal *S. pyogenes* virulence genes. The genome of strain M23ND also carries 24 chromosomally inherited established virulence genes (see Table S2 in the supplemental material). In contrast to the virulence genes carried by phages, the chromosomal genes are present in nearly all of the sequenced GAS strains, with the exception of the absence of *speG* in M4 strain
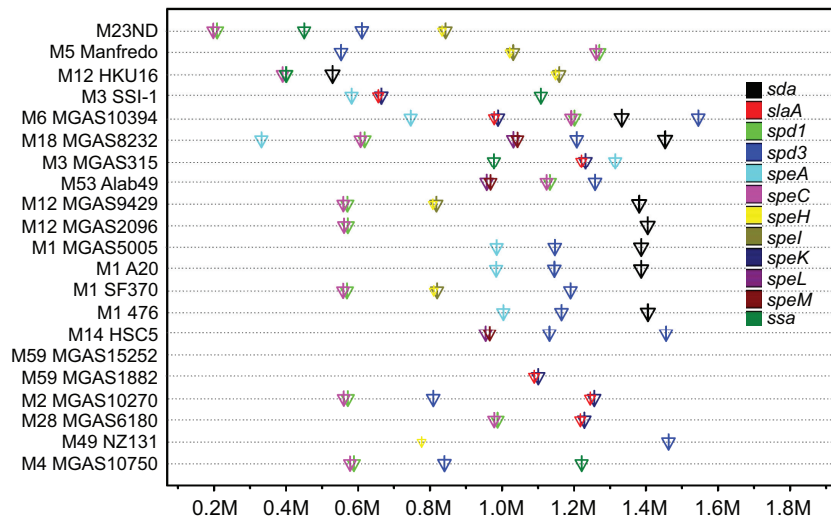
FIG 5 Profiling of phage-carried virulence factors across *S. pyogenes* genomes. Specific genes are represented by colored triangles, and the length of each gene is scaled by the triangle size. M23ND contains six known phage-carried virulence factors incorporated into four prophage regions, namely, *speC*, *spd1*, *ssa*, *spd3*, *speI*, and *speH*. Profiling of these six virulence factors, together with others across the 21 GAS sequences, showed that they are randomly distributed throughout the chromosome.

10750, *smeZ* in M2 strain 10270, M49 strain NZ131, *hasA* in M4 strain 10750, *slaA* in M18 strain 8232, and *endoS* in M49 strain NZ131. The major deviations in this regard are *sfbI*, *sic*, and *speJ*, which are carried by 10/21, 4/21, and 7/21 sequenced GAS strains, respectively. *sfbI* is a novel virulence gene product in M23ND with, at most, 66% nucleotide sequence homology across orthologs in other strains. Specifically, while this gene is present in almost half of the fully sequenced GAS strains, 34% to 85% of the sequence of *sfbI* in M23ND is unique to this strain. This gene is critically involved in streptococcal adherence to host epithelial cells (EpCs) via fibronectin. A BLASTp analysis of the protein sequence of SfbI in M23ND shows that it contains a fibronectin-binding protein signal sequence and five fibronectin-binding repeat domains positioned between amino acids 410 and 583 near the COOH terminus. SfbI in M23ND is, accordingly, a major candidate for promoting bacterial internalization into host EpCs, as well as invasion and infection specificity (55).

We discovered that the genomic locations of chromosomally inherited virulence factors (Fig. 6) and regulatory genes (Fig. 7) are more conserved than those of genes carried by phages across the GAS strains, with the exception of M23ND, M5 strain Manfredo, M12 strain HKU16, and M3 strain SSI-1. The streptococcal pyrogenic exotoxin genes, *viz.*, *speG*, *speJ*, and *smeZ*, along with other key genes, *viz.*, *sagA*, *slo*, *sfbI*, *ska*, and *sen*, and genes for extracellular toxins, *viz.*, *spyA*, *prtS*, *ideS*, *lmb*, *sclA*, *hylA*, *nga*, *graB*, *dltA*, *dltC*, *cfa*, and *plr* (Fig. 6A and B), are translocated away from these conserved sites. A similar situation exists with regulatory genes, *viz.*, *covRS*, *mga*, *rgg*, and *srv* (Fig. 7), along with the M23ND M protein gene, *emm23*, which coexists within the *mga cis* regulon. A close examination of the distribution of these genes revealed that the translocations correspond to the large genomic rearrangements outlined above (Fig. 4). Large segmental inversions and translocations carried the genes to their present locations.

Further examination of the sequences of the chromosomally inherited virulence genes via multiple-sequence alignment showed

that the orthologous genes share high similarity (>95%) between different GAS strains, except for genes such as *sfbI* and *sic*, which are not present in all GAS strains, and the binding-related genes, *viz.*, *graB*, *ska*, *endoS*, and *ideS*, which contain long divergent regions between GAS strains. The exceptional genes containing divergent regions encode mainly binding-related proteins, and the divergent regions fall into the binding domains within the parent protein, e.g., *ska*. Furthermore, the divergent sequences can be grouped into a few clusters independent of the serotypes of the GAS strains. It is likely that this divergence of the virulence genes was induced from horizontal gene exchange and was functionally related to binding specificity in adaptation to particular host challenges. In order to assess the adaptation roles played by the highly conserved virulence genes, we performed single nucleotide polymorphism (SNP) detection for those genes. We determined a frequency of 39 SNPs/kb, which is compatible with the 41 SNPs/kb observed for the seven commonly used housekeeping genes (36). This indicates that the evolutionary drive for enhanced adaptation of GAS strains may not originate from point mutations but may originate from horizontal gene transfer, a feature that is strikingly similar to that of genes carried by phages. This conclusion is also supported by the phylogenetic network constructed from SNP detection for the conserved virulence genes (see Fig. S2A in the supplemental material). The network was topologically similar to the background evolutionary structure inferred from the pairwise whole-genome comparisons (see Fig. S2B in the supplemental material) and based on multilocus sequence typing (MLST) of seven housekeeping genes (see Fig. S2C in the supplemental material). M23ND is most closely related to serotype M5 strain Manfredo and serotype M6 strain 10394, which may also share an ancestor with M18 strain 8232, having diverged more recently.

Pseudogene candidates for hypervirulence of M23ND. Gene inactivations can play a major role in the growth and invasive properties of GAS, and some occur during the course of infection to maintain GAS virulence during different challenges by the host (56). Two pseudogenes of particular relevance have been identified in the MD23 genome. One gene, the superantigen-encoding
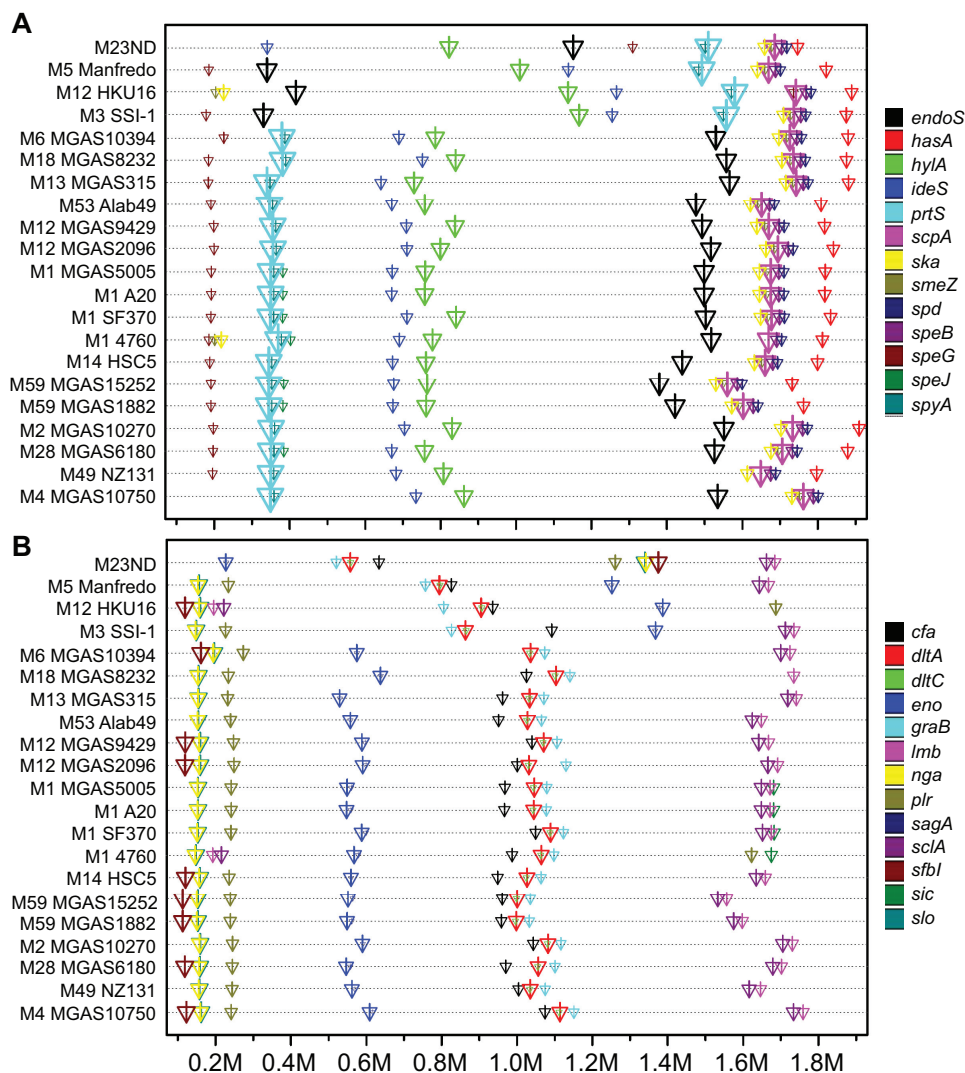
FIG 6 Profiling of chromosomally inherited virulence factors across *S. pyogenes* genomes. Genes are represented by colored triangles, and the length of each gene is scaled by the triangle width. These genes are present in almost all of the 21 fully sequenced GAS genomes, except for fibronectin-binding protein (*sfbI*), streptococcal inhibitor of complement (*sic*), and pyrogenic exotoxin J (*speJ*), which are carried by 10, 4, and 7 GAS strains, respectively. The overall genomic locations of chromosomal virulence factors are conserved across different GAS strains, with the exception of M23ND, M5 strain Manfredo, M12 strain HKU16, and M3 strain SSI-1, where the gene locations are obscured by large-scale genomic arrangements, including translocations and inversions. A total of 26 known virulence factors present in M23ND include *endoS*, *hasA*, *hylA*, *ideS*, *prtS*, *scpA*, *ska*, *smeZ*, *spd*, *speB*, *speG*, *speJ*, and *spyA* (A) and *cfa*, *dltA*, *dltC*, *eno*, *graB*, *lmb*, *nga*, *plr*, *sagA*, *sclA*, *sfbI*, *sic*, and *slo* (B).

speH gene, carried by prophage ΦM23ND.4 and regulated by bacterial *rgg* (57), resulted from a 1-bp frameshift deletion in the ORF and formed a hypothetically dysfunctional gene. Another critically important pseudogene, the CovS component of the CovRS two-component regulatory system, was found to contain a 5-nucleotide deletion in strain M23ND, resulting in attenuated *speB* expression and increased virulence and invasiveness in mice (58). Another pseudogene, *trx*, is essential for inhibiting O_2-mediated killing of *Mycobacterium*. Inactivation of this gene can be harmful to bacteria but is unlikely to be essential for GAS elimination, since GAS is killed by O_2-independent mechanisms (59).

Other M23ND pseudogenes of lesser-studied relevance are mainly involved with catabolism, biosynthesis, or signaling. These genes encode an ammonium transporter (*amount*), a chloride channel protein, glutamine 5-kinase, 2-(5′-triphosphoribosyl)-

3′-dephospho-coenzyme A synthetase, asparaginyl-tRNA synthetase-related protein, lanthionine biosynthesis protein, a mobile element protein, and a protein of unknown function. While we cannot know whether these genes were inactivated during evolution and/or during infection to allow GAS survival at different points of invasion, they are clearly compatible with the survival of GAS at the final invasion stage.

Identification of the genetic properties that contribute to pyogenic invasion and virulence. A previous study reported the association between mutations in the *covRS* two-component regulatory system of GAS strains with increased virulence (60). These mutations resulted in decreased expression of the gene for cysteine protease, *speB*, and upregulation of multiple virulence genes (61). To explore the mechanism of severe invasion and virulence of the currently studied strain, M23ND, we examined the genomic mu-
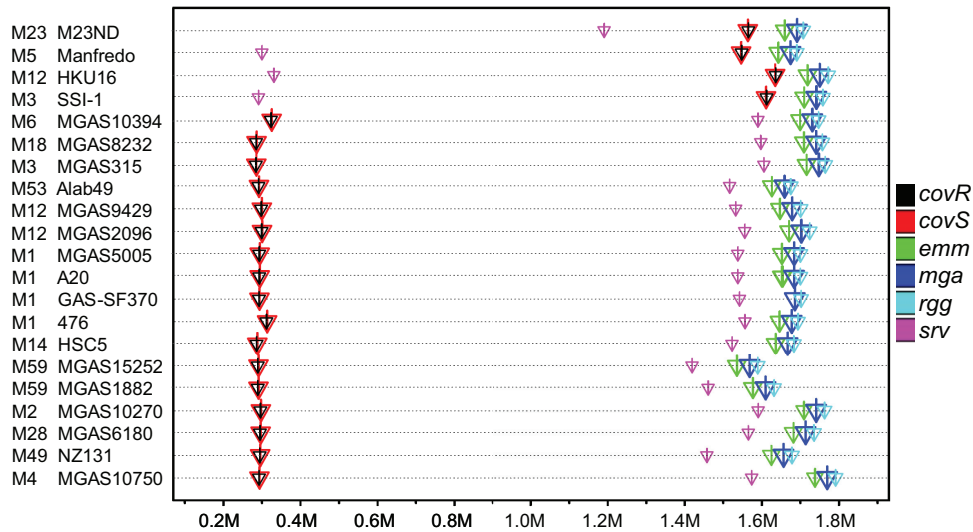
FIG 7 Profiling of six regulatory genes across *S. pyogenes* genomes. Genes are represented by colored triangles, and the length of each gene is scaled by the triangle width. All 21 currently available genome sequences were analyzed. The genes examined included *emm*, *mga*, *covR*, *covS*, *rgg*, and *srv*. The results for the M-protein gene, *emm*, have been artificially shifted in order to avoid overlap with those for *mga*. The locations of regulatory genes are highly conserved across all GAS strains, except for *srv* in M23ND. This gene is displaced from a position of approximately Mbp 1.5 to approximately Mbp 1.2, induced by the translocation of a large fragment within the genome. Similar translocations are evident for other genes in M23ND, e.g., *sagA* and *sfbI*.

tations in *covRS* and *in vitro* expression of the related virulence genes. We found that *covS* is indeed a pseudogene in M23ND and presents a highly attenuated expression of *speB*, determined by Western immunoblotting (data not shown). Similarly, inactivating mutations in *covS* and the corresponding attenuated expression of *speB* were also reported in GAS strains of other serotypes, such as a serotype M1 invasive isolate (56, 62), a highly virulent M3 isolate (63), M53 (15, 56) and M81 (64) invasive strains, and strain M23ND, evaluated in the current study. SpeB plays a complex role in virulence, initially being a factor that enhances invasion of the bacteria in EpCs (65), but during later stages this protease can have detrimental effects, since it can catalyze the degradation of other important virulence factors, such as M protein and FBP (66). Thus, under optimal conditions, this protease would be upregulated during initial infection phases and downregulated after initial invasion. The CovRS system is capable of accomplishing this task by rapidly generating inactivating mutations in CovRS (56).

In addition to the point mutations, the invasiveness of GAS requires expression of a combination of virulence genes which function at different infectious stages of infection. Invasion is initiated by adherence to EpCs mediated by surface-binding proteins, such as FBP, laminin-binding protein, collagen-binding protein, and fibrinogen-binding protein. After invasion, the GAS strains developed mechanisms to escape host innate immune systems by expressing genes for the DNases, hyaluronic acid capsule synthesis genes (*hasABC*), streptolysin O (*slo*), NAD glycohydrolase (*nga*), and pyrogenic exotoxin B (*speB*), among others. Invasive GAS strains also contain multicistronic regulatory systems, which, via mutations, can upregulate or downregulate important virulence determinants. The regulation of the genes for these systems during the course of infection allows invasive GAS strains to resist killing and persist at the infection sites (67).

Expression of critical virulence genes in M23ND. While the gene composition of M23ND was identified through genome se-

quencing, we verified the presence of these genes of interest to our work by RT-PCR (data not shown). Next, qRT-PCR was employed to assess the expression properties of these genes during LP and SP in both a clinical mutant *covS* strain and the isogenic wild-type *covS* strain. Expression of the cysteine protease SpeB is well-known to be regulated by the growth phase and by the CovRS system. As a verification that M23ND functioned similarly to many other GAS cell lines, we found, in agreement with previous work on other strains, that the *speB* mRNA is produced at the highest level during SP growth and with the CovRS system intact (Fig. 8A).

As examples of the data obtained with other genes, we found that genes facilitating host surface binding or initial immune invasion showed qualitatively higher mRNA levels during LP, including the genes for the M protein (*emm23*) and the C5a peptidase (*scpA*) (Fig. 8B), as well as the mRNA of a fibronectin-binding protein (*sfbI*), the pore forming protein, streptolysin O (*slo*), and the capsule-encoding gene *hasA* (Fig. 8C). An opposite result was found for the chromosomal DNase *spd* gene (Fig. 8C), which was slightly upregulated during SP (Fig. 8C). This result is logical, since *spd*, carried on the chromosome, and *spd1* and *spd3*, carried on the phages, are the only extracellular DNases present in M23ND and may be required at a later stage of infection when DNA nets encapsulate the bacteria. While some of these genes are universally present in different serotypes of GAS strains, e.g., *mga*, *emm*, *scpA*, *ska*, *fbp54*, *plr*, and *spd*, due to their essential role in GAS pathogenesis, the existence of several genes are serotype specific, e.g., *spd1*, *spd3*, *sfbI*, *enn*, and *fbpA* (the last two are absent in M23ND), or linked to tissue tropism, e.g., *ska* and *sfbI*. Some critical genes are regulated by the two-component regulator CovRS, and the ability of CovRS to become inactivated during the course of infection is a process that is particularly relevant with *speB*, *hasA*, and *slo* expression in M23ND to enhance its virulence (Fig. 8A and C).

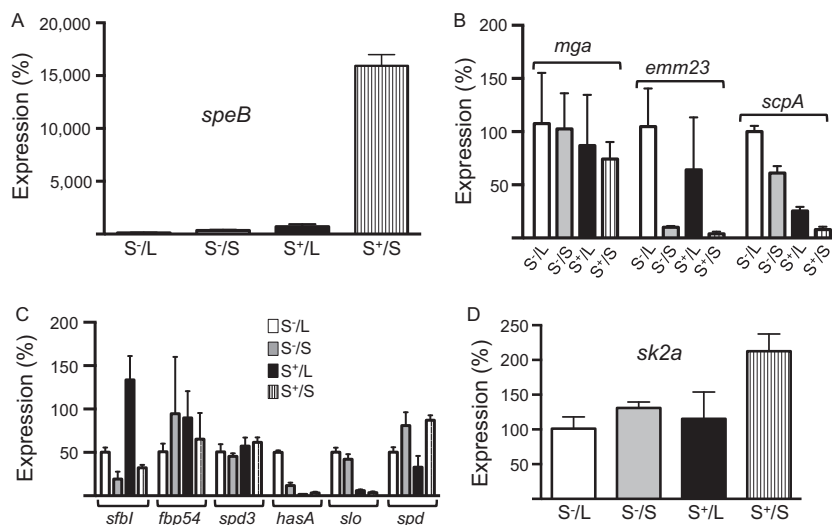Of special interest to our work, the *ska* gene, which is typically

FIG 8 Gene expression in full-length and truncated CovS M23ND strains. mRNAs isolated from the two strains (CovS$^+$ [S$^+$] and CovS$^-$ [S$^-$] strains) during LP (L; $A_{600} = 0.6$) and SP (S; $A_{600} > 1.0$) were analyzed for virulence factor gene expression using qRT-PCR. Primers specific for each gene (see Table S1 in the supplemental material) were used to measure the levels of transcription relative to the level of transcription of the GAPDH housekeeping gene. (A) Relative expression of the gene for extracellular cysteine protease, *speB*. (B) Gene expression in the multigene regulon (Mga) of GAS, which, in the case of M23ND, contains a minimal gene content, viz., the genes for M protein (*emm23*) and C5a peptidase (*scpA*). (C) qRT-PCR analyses of cell surface and secreted virulence factors involved in fibronectin binding (*sfbI* and *fbp54*), hyaluronic acid capsule biosynthesis (*hasA*), DNase activity (*spd* and *spd3*), and host cell lysis (*slo*). (D) qRT-PCR analysis of host-plasminogen activator *sk2a*, showing very similar expression in both strains and throughout both LP and SP growth.

under strong CovRS regulation, is nearly equally expressed in both CovR$^+$ CovS$^-$ and CovR$^+$ CovS$^+$ M23ND strains (Fig. 8D). In addition, the mRNAs of both *emm23* and *scpA* appear to deviate from strict Mga-mediated expression, as the levels of both transcripts are attenuated during SP, while the level of *mga* remains nearly constant (Fig. 8B). This finding indicates that factors other than Mga influence the expression of *emm23* and *scpA*. Lastly, the mRNA of the fibronectin-binding protein *sfbI* is regulated in a growth- and CovRS-dependent manner, since the CovR$^+$ CovS$^+$ strain produces significantly more *sfbI* transcript during LP, supporting its role in initial adherence and colonization. All of these observations suggest unique gene regulation in M23ND, consistent with the variation in genetic organization observed. While greater differences in expression of several of these genes are seen during LP and SP and in CovR$^+$ CovS$^+$ and CovR$^+$ CovS$^-$ AP53 cells (15) than in M23ND cells, any strain-dependent variations are likely to contain substantial contributions from the particular genomic architectural features of the strains.

DISCUSSION

The first complete genome of a serotype M23 GAS strain, M23ND, was sequenced and compared to the genomes of 20 other fully sequenced *S. pyogenes* strains currently available at NCBI. Our nucleotide sequence analysis showed that the genome contained four externally integrated prophage elements that carried six virulence genes, including the genes for four streptococcal pyrogenic exotoxins (*ssa*, *speC*, *speI*, and *speH*) and two endonucleases (*spd1* and *spd3*). The acquisition of virulence factors via prophage integration plays an important role in the pathogenesis of GAS isolates. In the present study, we propose that the combined recombination of the six virulence genes carried on the phages is one of the major contributing factors responsible for the severity of *S. pyogenes* M23ND infection. A comparative study revealed large-scale genomic rearrangements, unique to M23ND, that resulted in

genomes with arrangements different from those of previously sequenced GAS strains. However, the rearranged genomic architecture is imbalanced, yielding two unequal replichores. It is possible that this resultant imbalance contributes to the invasive nature of M23ND.

We also identified several known chromosomally inherited virulence factors, mainly related to host cell surface adherence and host immune system interactions. This indicates that a multitude of adaptive virulence factors has evolved to allow *S. pyogenes*, in general, and M23ND, in particular, to become versatile human pathogens. The genomic locations of these bacterial chromosomal genes are generally conserved across various GAS strains but are translocated in M23ND, resulting in global genomic rearrangement patterns unique to the genome of this strain. Translocations in chromosomal virulence genes may provide an alternative hypothesis for the enhanced adaptation of streptococci to particular environmental pressures via altered gene regulation patterns. However, whether the translocation of virulence genes is related to pathogenicity in *S. pyogenes* is still unclear. The small number of fully assembled *S. pyogenes* genomes prevents extensive and detailed assessments of the effects of gene translocations on virulence. Bacterial virulence is a complex process with many end-points. Some of these genes may be needed for certain stages of an infection, while a combination of genes is likely essential for severe virulence and death of the host. Therefore, we believe that point mutations in virulence genes should be the dominant factor for altering gene expression. Point mutations in regulatory virulence genes or genes for extracellular toxins provoke survival advantages and enhanced fitness in environments via the modulation of gene expression.

Phylogenetic analyses of housekeeping genes, based on MLST, showed that M23ND most likely shared a common ancestor (M18 strain 8232) with M5 strain Manfredo and M6

strain 10394. However, the genetic development, based on phage-carried virulence factors, reflects a complex evolutionary path for M23ND mediated by extensive horizontal gene transfer.

Ultimately, the pathogenesis and invasiveness of *S. pyogenes* isolate M23ND are unlikely to have been induced by a single factor. It is reasonable to conclude that the demonstrated lethality is a combined consequence of multiple factors, including the acquisition of virulence genes from prophages, global genomic rearrangements, and mutations in critical virulence factors. In addition, adaptive gene expression is also an important factor for invasive infection and environmental fitness.

It should necessarily be considered that all strains of GAS that have been analyzed normally originate from a host and are oftentimes further genetically manipulated and further passaged in the laboratory setting. These practices result in gene interactions between host and bacteria that may alter the genetic characteristics of the bacteria, and the nature of the originating infecting agent may differ genetically from that of the isolate. The most obvious example of this is the inactivation of *covR* and *covS* during the course of infection (56) and the profound consequences of *covRS* inactivations toward virulence, especially in regard to *speB* production. *speB*, produced in CovR$^+$ CovS$^+$ cells, is beneficial to the dissemination of the initial infection, since *speB* catalyzes the degradation of extracellular matrix components, e.g., fibronectin and vitronectin (68); activates proinflammatory proteins, e.g., interleukin-1β (69); degrades IgG and IgA (70); and inactivates complement factors (71), thereby circumventing the host immune response. This protease can also cleave EpC junction proteins, thereby facilitating GAS translocation across the epithelium (72). However, in late infection, the presence of *speB* may not be beneficial to bacterial dissemination since this protease inactivates GAS virulence factors. Thus, it is an advantage for GAS to downregulate *speB* expression at late infection stages. Since mutants with inactivated *speB* have not been found to date during the course of infection, with mutagenic inactivation of *covR*, *covS*, and/or another *speB* regulator, *rgg* (73), these loci serve as mutagenic loci during infection to downregulate the expression of *speB*. This highly tuned phase-switch mechanism with *covRS* would initiate hyperinvasive disease (74), to a major extent via SpeB regulation, by maximizing conditions for initial GAS localized protease base invasion into tissue and at later stages by rapidly attenuating SpeB production and preserving GAS virulence factors that are needed for dissemination.

Another manner in which M23ND assembles a virulent proteolytic surface is via binding of host plasminogen (hPg) and host plasmin (hPm). Upon examining the amino acid sequence of the major predicted hPg/hPm binding protein, M23, we propose that M23 should not bind hPg/hPm directly but should employ its ability to bind host fibrinogen (hFg), which will then allow hPg/hPm binding. We have previously established that a coinheritance of isoforms of SK and the mode of binding of hPg occurs (75–77). The amino acid sequence of SK secreted by M23ND is the SK2a form, which maximally activates hPg bound to GAS via hFg/M protein. Thus, the principles of coinheritance of the forms of SK and M protein are verified with M23ND.

In conclusion, on a gene content level, the hypervirulence of M23ND is consistent with the *covS* mutation found, the presence of the prophage superantigen gene *ssa*, and the expression of the critical fibronectin-binding protein gene *sfbI*. These genes, plus the gene architectural features described throughout the report, explain the presence of this form of the bacterium in the hyperinfectious human isolate from which it was discovered.

REFERENCES

1. **Lancefield RC.** 1928. The antigenic complex of Streptococcus haemolyticus. I. Demonstration of a type-specific substance in extracts of Streptococcus haemolyticus. J. Exp. Med. **47**:91–103.
2. **Carapetis JR, Steer AC, Mulholland EK, Weber M.** 2005. The global burden of group A streptococcal diseases. Lancet Infect. Dis. **5**:685–694. http://dx.doi.org/10.1016/S1473-3099(05)70267-X.
3. **Schwartz B, Facklam RR, Breiman RF.** 1990. Changing epidemiology of group A streptococcal infection in the USA. Lancet **336**:1167–1171. http://dx.doi.org/10.1016/0140-6736(90)92777-F.
4. **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.
5. **Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, Henningham A, Steer AC, Bessen DE, Dale JB, Curtis N, Beall BW, Walker MJ, Parker MW, Carapetis JR, Van Melderen L, Sriprakash KS, Smeesters PR, the M Protein Study Group.** 5 May 2014. A systematic and functional classification of Streptococcus pyogenes that serves as a new tool for molecular typing and vaccine development. J. Infect. Dis. http://dx.doi.org/10.1093/infdis/jiu260.
6. **Smeesters PR, McMillan DJ, Sriprakash KS.** 2010. The streptococcal M protein: a highly versatile molecule. Trends Microbiol. **18**:275–282. http://dx.doi.org/10.1016/j.tim.2010.02.007.
7. **Fischetti VA.** 1991. Streptococcal M protein. Sci. Am. **264**:58–65. http://dx.doi.org/10.1038/scientificamerican0691-58.
8. **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.
9. **Hollingshead SK, Readdy TL, Yung DL, Bessen DE.** 1993. Structural heterogeneity of the emm gene cluster in group A streptococci. Mol. Microbiol. **8**:707–717. http://dx.doi.org/10.1111/j.1365-2958.1993.tb01614.x.
10. **Podbielski A.** 1993. Three different types of organization of the vir regulon in group A streptococci. Mol. Gen. Genet. **237**:287–300.
11. **Bessen DE, Sotir CM, Readdy TL, Hollingshead SK.** 1996. Genetic correlates of throat and skin isolates of group A streptococci. J. Infect. Dis. **173**:896–900. http://dx.doi.org/10.1093/infdis/173.4.896.
12. **Bessen DE, Izzo MW, Fiorentino TR, Caringal RM, Hollingshead SK, Beall B.** 1999. Genetic linkage of exotoxin alleles and emm gene markers for tissue tropism in group A streptococci. J. Infect. Dis. **179**:627–636. http://dx.doi.org/10.1086/314631.
13. **Terao Y, Kawabata S, Kunitomo E, Murakami J, Nakagawa I, Hamada S.** 2001. Fba, a novel fibronectin-binding protein from Streptococcus pyogenes, promotes bacterial entry into epithelial cells, and the fba gene is positively transcribed under the Mga regulator. Mol. Microbiol. **42**:75–86. http://dx.doi.org/10.1046/j.1365-2958.2001.02579.x.
14. **Ramachandran V, McArthur JD, Behm CE, Gutzeit C, Dowton M, Fagan PK, Towers R, Currie B, Sriprakash KS, Walker MJ.** 2004. Two distinct genotypes of prtF2, encoding a fibronectin binding protein, and evolution of the gene family in Streptococcus pyogenes. J. Bacteriol. **186**:7601–7609. http://dx.doi.org/10.1128/JB.186.22.7601-7609.2004.
15. **Liang Z, Zhang Y, Agrahari G, Chandrahas V, Glinton K, Donahue DL, Balsara RD, Ploplis VA, Castellino FJ.** 2013. A natural inactivating mutation in the CovS component of the CovRS regulatory operon in a pattern D streptococcal pyogenes strain influences virulence-associated genes. J. Biol. Chem. **288**:6561–6573. http://dx.doi.org/10.1074/jbc.M112.442657.
16. **Molinari G, Rohde M, Guzmán CA, Chhatwal GS.** 2000. Two distinct pathways for the invasion of Streptococcus pyogenes in non-phagocytic cells. Cell. Microbiol. **2**:145–154. http://dx.doi.org/10.1046/j.1462-5822.2000.00040.x.

17. **Bessen D, Jones KF, Fischetti VA.** 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. J. Exp. Med. **169:**269–283. http://dx.doi.org/10.1084/jem.169.1.269.

18. **Guzman CA, Talay SR, Molinari G, Medina E, Chhatwal GS.** 1999. Protective immune response against Streptococcus pyogenes in mice after intranasal vaccination with the fibronectin-binding protein SfbI. J. Infect. Dis. **179:**901–906. http://dx.doi.org/10.1086/314655.

19. **Jeng A, Sakota V, Li Z, Datta V, Beall B, Nizet V.** 2003. Molecular genetic analysis of a group A Streptococcus operon encoding serum opacity factor and a novel fibronectin-binding protein, SfbX. J. Bacteriol. **185:**1208–1217. http://dx.doi.org/10.1128/JB.185.4.1208-1217.2003.

20. **Courtney HS, Li Y, Dale JB, Hasty DL.** 1994. Cloning, sequencing, and expression of a fibronectin/fibrinogen-binding protein from group A streptococci. Infect. Immun. **62:**3937–3946.

21. **Courtney HS, Dale JB, Hasty DI.** 1996. Differential effects of the streptococcal fibronectin-binding protein, FBP54, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. Infect. Immun. **64:**2415–2419.

22. **Perez-Casal J, Caparon MG, Scott JR.** 1991. Mry, a trans-acting positive regulator of the M protein gene of Streptococcus pyogenes with similarity to the receptor proteins of two-component regulatory systems. J. Bacteriol. **173:**2617–2624.

23. **Hondorp ER, McIver KS.** 2007. The Mga virulence regulon: infection where the grass is greener. Mol. Microbiol. **66:**1056–1065. http://dx.doi.org/10.1111/j.1365-2958.2007.06006.x.

24. **Kreikemeyer B, KS McIver, Podbielski A.** 2003. Virulence factor regulation and regulatory networks in Streptococcus pyogenes and their impact on pathogen-host interactions. Trends Microbiol. **11:**224–232. http://dx.doi.org/10.1016/S0966-842X(03)00098-2.

25. **Simpson WJ, LaPenta D, Chen C, Cleary PP.** 1990. Coregulation of type 12 M protein and streptococcal C5a peptidase genes in group A streptococci: evidence for a virulence regulon controlled by the virR locus. J. Bacteriol. **172:**696–700.

26. **Stenberg L, O'Toole P, Lindahl G.** 1992. Many group A streptococcal strains express two different immunoglobulin-binding proteins, encoded by closely linked genes: characterization of the proteins expressed by four strains of different M-type. Mol. Microbiol. **6:**1185–1194. http://dx.doi.org/10.1111/j.1365-2958.1992.tb01557.x.

27. **Almengor AC, Walters MS, McIver KS.** 2006. Mga is sufficient to activate transcription in vitro of sof-sfbX and other Mga-regulated virulence genes in the group A streptococcus. J. Bacteriol. **188:**2038–2047. http://dx.doi.org/10.1128/JB.188.6.2038-2047.2006.

28. **Ferretti JJ, McShan WM, Ajdic D, Savic DJ, Savic G, Lyon K, Primeaux C, Sezate S, Suvorov AN, Kenton S, Lai HS, Lin SP, Qian Y, Jia HG, Najar FZ, Ren Q, Zhu H, Song L, White J, Yuan X, Clifton SW, Roe BA, McLaughlin R.** 2001. Complete genome sequence of an M1 strain of Streptococcus pyogenes. Proc. Natl. Acad. Sci. U. S. A. **98:**4658–4663. http://dx.doi.org/10.1073/pnas.071559398.

29. **Okamoto H, Minami M, Shoin S, Koshimura S, Shimizu R.** 1966. Experimental anticancer studies. XXXI. On the streptococcal preparation having potent anticancer activity. Jpn. J. Exp. Med. **36:**175–186.

30. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.

31. **Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

32. **Aziz RK, Kotb M.** 2008. Rise and persistence of global M1T1 clone of Streptococcus pyogenes. Emerg. Infect. Dis. **14:**1511–1517. http://dx.doi.org/10.3201/eid1410.071660.

33. **Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I.** 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res. **42:**D553–D559. http://dx.doi.org/10.1093/nar/gkt1274.

34. **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.

35. **Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423. http://dx.doi.org/10.1093/bioinformatics/bti553.

36. **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

37. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23:**254–267. http://dx.doi.org/10.1093/molbev/msj030.

38. **Okamoto H, Koshimura S.** November 1969. Treating method of Streptococcus hemolyticus and preparation containing the said microorganism. US patent 3,477,914.

39. **Sotomura M, Iwamoto S, Sawada T, Inoue S, Suzuki A, Ikeda Y.** May 1982. Method for the treatment of cells of Streptococcus pyogenes. US patent 4,328,218.

40. **Okamoto H, Shoin S, Minami M, Koshimura S, Shimizu R.** 1966. Experimental anticancer studies. XXX. Factors influencing the streptolysin S-forming ability of streptococci having anticancer activity. Jpn. J. Exp. Med. **36:**161–174.

41. **Hong K.** 2000. Characterization of group A streptococcal strains Sv and Su: determination of emm gene typing and presence of small vir regulon. Res. Microbiol. **151:**29–36. http://dx.doi.org/10.1016/S0923-2508(00)00130-3.

42. **Scott J, Thompson-Mayberry P, Lahmamsi S, King CJ, McShan WM.** 2008. Phage-associated mutator phenotype in group A streptococcus. J. Bacteriol. **190:**6290–6301. http://dx.doi.org/10.1128/JB.01569-07.

43. **Botstein D.** 1980. A theory of modular evolution for bacteriophages. Ann. N. Y. Acad. Sci. **354:**484–490. http://dx.doi.org/10.1111/j.1749-6632.1980.tb27987.x.

44. **McShan WM, Ferretti JJ.** 1997. Genetic diversity in temperate bacteriophages of Streptococcus pyogenes: identification of a second attachment site for phages carrying the erythrogenic toxin A gene. J. Bacteriol. **179:**6509–6511.

45. **Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.** 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science **315:**1709–1712. http://dx.doi.org/10.1126/science.1138140.

46. **Marraffini LA, Sontheimer EJ.** 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science **322:**1843–1845. http://dx.doi.org/10.1126/science.1165771.

47. **Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J.** 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science **321:**960–964. http://dx.doi.org/10.1126/science.1159689.

48. **Nozawa T, Furukawa N, Aikawa C, Watanabe T, Haobam B, Kurokawa K, Maruyama F, Nakagawa I.** 2011. CRISPR inhibition of prophage acquisition in Streptococcus pyogenes. PLoS One **6:**e19543. http://dx.doi.org/10.1371/journal.pone.0019543.

49. **Duggin IG, Dubarry N, Bell SD.** 2011. Replication termination and chromosome dimer resolution in the archaeon Sulfolobus solfataricus. EMBO J. **30:**145–153. http://dx.doi.org/10.1038/emboj.2010.301.

50. **Nakagawa I, Kurokawa K, Yamashita A, Nakata M, Tomiyasu Y, Okahashi N, Kawabata S, Yamazaki K, Shiba T, Yasunaga T, Hayashi H, Hattori M, Hamada S.** 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. http://dx.doi.org/10.1101/gr.1096703.

51. **Esnault E, Valens M, Espeli O, Boccard F.** 2007. Chromosome structuring limits genome plasticity in Escherichia coli. PLoS Genet. **3:**e226. http://dx.doi.org/10.1371/journal.pgen.0030226.

52. **Matthews TD, Maloy S.** 2010. Fitness effects of replichore imbalance in Salmonella enterica. J. Bacteriol. **192:**6086–6088. http://dx.doi.org/10.1128/JB.00649-10.

53. **Musser JM, Nelson K, Selander RK, Gerlach D, Huang JC, Kapur V, Kanjilal S.** 1993. Temporal variation in bacterial disease frequency: molecular population genetic analysis of scarlet fever epidemics in Ottawa and in eastern Germany. J. Infect. Dis. **167:**759–762. http://dx.doi.org/10.1093/infdis/167.3.759.

54. **Aziz RK, Edwards RA, Taylor WW, Low DE, McGeer A, Kotb M.** 2005. Mosaic prophages with horizontally acquired genes account for the emergence and diversification of the globally disseminated M1T1 clone of Streptococcus pyogenes. J. Bacteriol. **187:**3311–3318. http://dx.doi.org/10.1128/JB.187.10.3311-3318.2005.

55. **Hanski E, Caparon M.** 1992. Protein F, a fibronectin-binding protein, is an adhesin of the group A streptococcus Streptococcus pyogenes. Proc. Natl. Acad. Sci. U. S. A. **89:**6172–6176. http://dx.doi.org/10.1073/pnas.89.13.6172.

56. **Mayfield JA, Liang Z, Agrahari G, Lee SW, Donahue DL, Ploplis VA,**

Castellino FJ. 2014. Mutations in the control of virulence sensor gene from Streptococcus pyogenes after infection in mice lead to clonal bacterial variants with altered gene regulatory activity and virulence. PLoS One 9:e100698. http://dx.doi.org/10.1371/journal.pone.0100698.

57. Anbalagan S, Dmitriev A, McShan WM, Dunman PM, Chaussee MS. 2012. Growth phase-dependent modulation of Rgg binding specificity in Streptococcus pyogenes. J. Bacteriol. 194:3961–3971. http://dx.doi.org/10.1128/JB.06709-11.

58. Maamary PG, Ben Zakour NL, Cole JN, Hollands A, Aziz RK, Barnett TC, Cork AJ, Henningham A, Sanderson-Smith M, McArthur JD, Venturini C, Gillen CM, Kirk JK, Johnson DR, Taylor WL, Kaplan EL, Kotb M, Nizet V, Beatson SA, Walker MJ. 2012. Tracing the evolutionary history of the pandemic group A streptococcal M1T1 clone. FASEB J. 26:4675–4684. http://dx.doi.org/10.1096/fj.12-212142.

59. Wieles B, Ottenhoff TH, Steenwijk TM, Franken KL, de Vries RR, Langermans JA. 1997. Increased intracellular survival of Mycobacterium smegmatis containing the Mycobacterium leprae thioredoxin-thioredoxin reductase gene. Infect. Immun. 65:2537–2541.

60. Cole JN, Barnett TC, Nizet V, Walker MJ. 2011. Molecular insight into invasive group A streptococcal disease. Nat. Rev. Microbiol. 9:724–736. http://dx.doi.org/10.1038/nrmicro2648.

61. Maamary PG, Sanderson-Smith ML, Aziz RK, Hollands A, Cole JN, McKay FC, McArthur JD, Kirk JK, Cork AJ, Keefe RJ, Kansal RG, Sun H, Taylor WL, Chhatwal GS, Ginsburg D, Nizet V, Kotb M, Walker MJ. 2010. Parameters governing invasive disease propensity of non-M1 serotype group A streptococci. J. Innate Immun. 2:596–606. http://dx.doi.org/10.1159/000317640.

62. Cole JN, McArthur JD, McKay FC, Sanderson-Smith ML, Cork AJ, Ranson M, Rohde M, Itzek A, Sun H, Ginsburg D, Kotb M, Nizet V, Chhatwal GS, Walker MJ. 2006. Trigger for group A streptococcal M1T1 invasive disease. FASEB J. 20:1745–1747. http://dx.doi.org/10.1096/fj.06-5804fje.

63. Miyoshi-Akiyama T, Ikebe T, Watanabe H, Uchiyama T, Kirikae T, Kawamura Y. 2006. Use of DNA arrays to identify a mutation in the negative regulator, csrR, responsible for the high virulence of a naturally occurring type M3 group A streptococcus clinical isolate. J. Infect. Dis. 193:1677–1684. http://dx.doi.org/10.1086/504263.

64. Garcia AF, Abe LM, Erdem G, Cortez CL, Kurahara D, Yamaga K. 2010. An insert in the covS gene distinguishes a pharyngeal and a blood isolate of Streptococcus pyogenes found in the same individual. Microbiology 156:3085–3095. http://dx.doi.org/10.1099/mic.0.042614-0.

65. Tsai PJ, Kuo CF, Lin KY, Lin YS, Lei HY, Chen FF, Wang JR, Wu JJ. 1998. Effect of group A streptococcal cysteine protease on invasion of epithelial cells. Infect. Immun. 66:1460–1466.

66. Wei L, Pandiripally V, Gregory E, Clymer M, Cue D. 2005. Impact of the SpeB protease on binding of the complement regulatory proteins factor H and factor H-like protein 1 by Streptococcus pyogenes. Infect. Immun. 73:2040–2050. http://dx.doi.org/10.1128/IAI.73.4.2040-2050.2005.

67. Agrahari G, Liang Z, Mayfield JA, Balsara RD, Ploplis VA, Castellino FJ. 2013. Complement-mediated opsonization of invasive group A Streptococcus pyogenes strain AP53 is regulated by the bacterial two-component cluster of virulence responder/sensor (CovRS) system. J. Biol. Chem. 288:27494–27504. http://dx.doi.org/10.1074/jbc.M113.494864.

68. Kapur V, Topouzis S, Majesky MW, Li LL, Hamrick MR, Hamill RJ, Patti JM, Musser JM. 1993. A conserved Streptococcus pyogenes extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. Microb. Pathog. 15:327–346. http://dx.doi.org/10.1006/mpat.1993.1083.

69. Kapur V, Majesky MW, Li LL, Black RA, Musser JM. 1993. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from Streptococcus pyogenes. Proc. Natl. Acad. Sci. U. S. A. 90:7676–7680. http://dx.doi.org/10.1073/pnas.90.16.7676.

70. Collin M, Olsen A. 2001. Effect of SpeB and EndoS from Streptococcus pyogenes on human immunoglobulins. Infect. Immun. 69:7187–7189. http://dx.doi.org/10.1128/IAI.69.11.7187-7189.2001.

71. Honda-Ogawa M, Ogawa T, Terao Y, Sumitomo T, Nakata M, Ikebe K, Maeda Y, Kawabata S. 2013. Cysteine proteinase from Streptococcus pyogenes enables evasion of innate immunity via degradation of complement factors. J. Biol. Chem. 288:15854–15864. http://dx.doi.org/10.1074/jbc.M113.469106.

72. Sumitomo T, Nakata M, Higashino M, Terao Y, Kawabata S. 2013. Group A streptococcal cysteine protease cleaves epithelial junctions and contributes to bacterial translocation. J. Biol. Chem. 288:13317–13324. http://dx.doi.org/10.1074/jbc.M113.459875.

73. Chaussee MS, Ajdic D, Ferretti JJ. 1999. The rgg gene of Streptococcus pyogenes NZ131 positively influences extracellular SPE B production. Infect. Immun. 67:1715–1722.

74. Cole JN, Pence MA, von Köckritz-Blickwede M, Hollands A, Gallo RL, Walker MJ, Nizet V. 2010. M protein and hyaluronic acid capsule are essential for in vivo selection of covRS mutations characteristic of invasive serotype M1T1 group A streptococcus. mBio 1(4):e00191-10. http://dx.doi.org/10.1128/mBio.00191-10.

75. Zhang Y, Liang Z, Hsueh HT, Ploplis VA, Castellino FJ. 2012. Characterization of streptokinases from group A streptococci reveals a strong functional relationship that supports the coinheritance of plasminogen-binding M-protein and cluster 2b streptokinase. J. Biol. Chem. 287:42093–42103. http://dx.doi.org/10.1074/jbc.M112.417808.

76. Zhang Y, Liang Z, Glinton K, Ploplis VA, Castellino FJ. 2013. Functional differences between Streptococcus pyogenes cluster 1 and cluster 2b streptokinases are determined by their β-domains. FEBS Lett. 587:1304–1309. http://dx.doi.org/10.1016/j.febslet.2013.02.033.

77. Zhang Y, Mayfield J, Ploplis VA, Castellino FJ. 2014. The β-domain of cluster 2b streptokinase is a major determinant for the regulation of its plasminogen activation activity by cellular plasminogen receptors. Biochem. Biophys. Res. Commun. 444:595–598. http://dx.doi.org/10.1016/j.bbrc.2014.01.109.