Virulence

ISSN: 2150-5594 (Print) 2150-5608 (Online) Journal homepage: https://www.tandfonline.com/loi/kvir20

The M protein of group A Streptococcus is a key virulence factor and a clinically relevant strain identification marker

David Metzgar & Antonella Zampolli

To cite this article: David Metzgar & Antonella Zampolli (2011) The M protein of group A Streptococcus is a key virulence factor and a clinically relevant strain identification marker, Virulence, 2:5, 402-412, DOI: 10.4161/viru.2.5.16342

To link to this article: https://doi.org/10.4161/viru.2.5.16342

Virulence 2:5, 402-412; September/October 2011; © 2011 Landes Bioscience

The M protein of group A Streptococcus is a key virulence factor and a clinically relevant strain identification marker

David Metzgar[1,*] and Antonella Zampolli[2]

[1]Ibis Biosciences (a Subsidiary of Abbott, Inc.); Carlsbad, CA USA; [2]The Scripps Research Institute; La Jolla, CA USA

Key words: group A Streptococcus, acute rheumatic fever, rheumatic heart disease, PSGN, ARF, RHD, streptococcal toxic shock syndrome, epidemic, public health

The M protein coats group A Streptococci (GAS) and acts as the primary antigen and determinant of type-specific immunity. M is essential for GAS virulence, providing antiphagocytic functions critical to survival in human tissues and fluids. Specific regions of M protein also serve as shared antigens, and cross-reactivity between these epitopes and human proteins may be the source of autoimmune sequelae such as rheumatic heart disease. The M protein is hypervariable and has long served as the primary target for epidemiological typing of GAS. Though other markers or genotyping methods may be necessary to increase strain resolution when clones of a given M type differ in clinically critical ways, M typing remains the most directly informative and well-documented method for tracking outbreaks of GAS, predicting clinical outcomes during those outbreaks, and measuring the general threat presented by GAS at any given time and place.

Introduction

β-hemolytic group A Streptococcus (GAS or *Streptococcus pyogenes*) is a ubiquitous pathogen, occurring often in both asymptomatic carrier states and active infections of its only hosts, humans. The most common infections, which usually occur in multiple episodes during childhood, express themselves as pharyngitis (strep throat) or impetigo (blistering skin infection).[1] In developed regions, distinct groups of serotypes (M types) are responsible for these two primary categories of infection.[1-7] GAS can cause a wide variety of more serious conditions including bacteremia, diverse focal infections, puerperal sepsis and necrotizing fasciitis. Some patients experience severe responses to streptococcal toxins, expressed as scarlet fever and streptococcal toxic shock syndrome (STSS).[1] Acute infection may also lead to post-infection autoimmune conditions including acute rheumatic fever (ARF), rheumatic heart disease and post-streptococcal glomerulonephritis (PSGN, an inflammatory kidney condition). Prior to the discovery of penicillin, these severe forms of infection and post-infection autoimmune disease were leading causes of morbidity and mortality in human populations.[2,8]

Infection with specific serotypes (M types) is more likely to result in severe reactions, rheumatic sequelae, and specific clinical outcomes such as puerperal sepsis and necrotizing fasciitis. Some aspects of these type-specific outcomes derive directly from virulence properties of the M protein, and hence are readily predicted through determination of the M protein serotype via serology or sequence analysis of the *emm* gene. The tendency to induce acute rheumatic fever is possibly the best-studied example.[9-11] Other determinants of virulence are encoded by a variety of chromosomal genes, many of which are associated with highly mobile phage elements. For these, the genetic links to virulence are far less clear and likely polygenic. The mobility of these elements generates significant diversity within and between serotypes.[12-14] In this review we describe the known clinical correlates of M type, including virulence properties and antibiotic resistance phenotypes. We discuss the evolutionary mechanisms that contribute to M type diversity, distribution, and intraserotypic strain variability, and we discuss the existing evidence supporting the concept that the M protein itself contributes directly to key virulence properties, including the propensity for specific M types to induce acute rheumatic fever. All of these aspects of the M protein contribute to its unique value as a strain typing marker.

M Protein as a Primary Antigen and Serological Marker

The M protein was described in 1927 by Rebecca Lancefield,[15] who demonstrated it to be a primary source of strain-specific immunity. Having experienced an acute GAS infection, individuals develop very strong lasting immunity to the infecting serotype.[6] This immunity prevents further suppurative (pus-forming) infections, though the same M type may persist or recolonize in asymptomatic carrier states.[8,16] The M protein, essentially a long α-helical coiled coil anchored in the bacterial membrane and extending from the surface of the cell (**Fig. 1**), coats the surface of GAS.[16]

M proteins are distinguished by their ability to produce an apoproteinase that cleaves apoprotein A1, also termed Serum

*Correspondence to: David Metzgar; Email: davidleemetzgar@yahoo.com
Submitted: 05/04/11; Revised: 07/22/11; Accepted: 07/25/11
http://dx.doi.org/10.4161/viru.2.5.16342

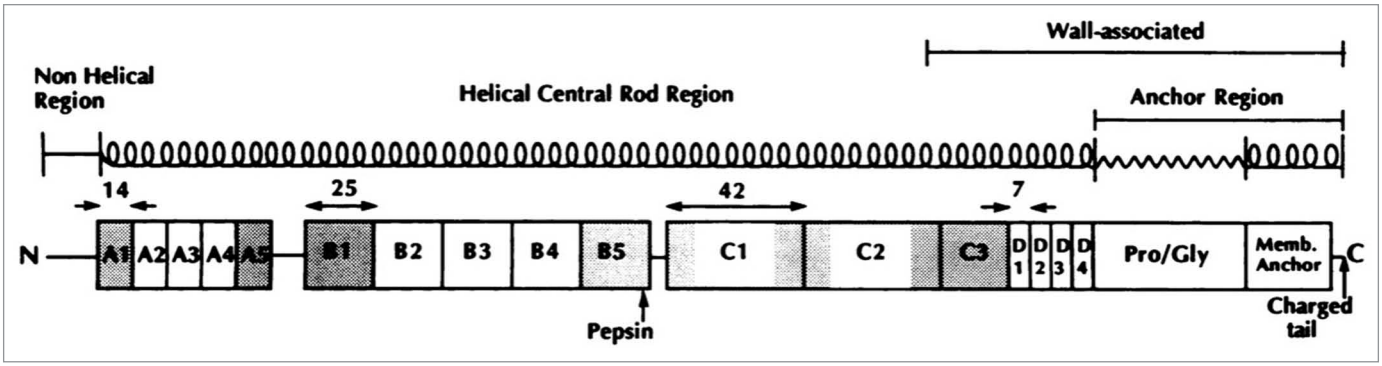

Figure 1. Structure of M6. M6 includes all the building blocks of the M protein family, though the presence and repeat number of elements can vary widely between M types. The length of the non-helical portion at the N-terminus is variable, and is completely missing in M49;[7] the number and size of the A, B and C blocks is also variable, and the A and B block are missing in some types. Numbers above each block (associated with horizontal range arrows) indicate the number of amino acids per block repeat unit. Pro/Gly indicates a Proline-Glycine rich region (reproduced with permission from Fischetti, 1989).[16]

Opacity Factor (sof);[1] they are also distinguished as belonging to class I by reactivity with antibodies that recognize a domain conserved in the C repeats of class I M proteins, while class II is distinguished by a lack of reactivity.[9] The two classifications correlate, with class I M proteins falling in the sof- group and class II in the sof+ group.

The N-terminal region, farthest from the cell and most likely to be bound by antibodies, contains an extremely variable domain associated with serotype identity. Regions closer to the cell membrane are less diverse, but also significantly variable and antigenic (**Fig. 1**). Four patterns are recognized for the class I/sof group, termed A, B, C (often referred to collectively as A–C) and D. M proteins in these groups have a non helical N-terminal region and, closer to the C terminus, different combinations of A, B, C and D repeats, followed by a wall-anchoring domain. The D pattern is distinguished by having an average length of 355 residues, while the A, B and C patterns average 444 residues. The E pattern M proteins belong to the class II/sof$^+$ group, and have a shorter wall-anchoring domain.[6,9,16] M types characterized by emm genes with A–C patterns are generally associated with pharyngitis, those with D patterns are generally associated with impetigo and those with E patterns are generalists and may be associated with both types of infection.[6,9,16]

GAS strains are most commonly discriminated by identification of the M protein (M typing). This was historically accomplished by serological analysis of a purified M protein fragment,[15] a technique recently replaced by sequence analysis targeting the fragment of the *emm* gene that codes for the primary antigenic region of M protein (*emm* typing).[17] The two methods, serological and genetic, are extremely well correlated. Standardized reference databases of *emm* sequences from both previously recognized M types and novel *emm* alleles have been established to allow rapid and unambiguous translation of *emm* sequence data into M type information, and to allow the identification and description of novel *emm* types.[17] As these two techniques differ primarily in the method of data collection rather than the resulting identifications, we will refer to them collectively in this review as "M typing."

M Protein as a Virulence Factor

M proteins are primary virulence factors for GAS strains.[9,16,18] They offer protection through diversity, providing immunologically distinct surface coats to different serotypes. Novel serotypes thereby avoid antibodies raised by hosts in response to previous infections.[4,15,16,19] As more individuals are exposed to a specific serotype, that serotype becomes less fit due to the inability to transmit successfully between the decreased numbers of remaining susceptible individuals, a phenomenon known as herd immunity.[20-23] As herd immunity to a specific serotype increases, other serotypes become favored through a relative increase in transmission and hence reproduction. Because this process generally favors unique (new or recently rare) serological variants, evolutionary pressure is thought to drive the process of diversification in primary antigens and to drive cyclical turnovers in the dominance of existing serotypes. The resulting variability makes primary antigens very informative strain markers for many pathogens, and the strong pressures driving cyclic strain turnovers (sometimes termed dynamic epidemiology),[24] must be considered as potential causal factors in any investigation of changing virulence or rates of antibiotic resistance among GAS isolates.[21]

M proteins offer active protection against phagocytosis and thereby allow the pathogen to persist in infected tissues.[9,16,19] M protein coated strains can resist opsonization by binding inhibitory regulators of the complement system (factor H, factor H-like 1, C4-binding protein, CD46 complement regulatory protein) and fibrinogen, which specifically inhibits binding of C3b, evading both the classic and the alternative pathways of complement activation. M proteins also contribute to the phagocytosis resistance of GAS by non-immune binding of the Fc region of IgG.[1,6]

The ability of M proteins belonging to the D pattern to bind plasminogen contributes critically to their virulence. Plasmin activation can initiate or increase fibrinolysis, potentially counteracting the host's coagulatory response at the site of infection and favoring the spreading of bacteria.[25]

Table 1. Examples of M types associated with specific clinical and epidemiological presentations

Clinical presentation	Associated M types	References
Pharyngitis	M1, M3, M5, M6, M12, M14, M17, M19, M24	1
Acute rheumatic fever (ARF)	M1, M3, M5, M6, M11, M12, M14, M17, M18, M19, M24, M27, M29, M30, M32, M41	9, 18, 40
Epidemic ARF	M5, M18	39
Geographically widespread epidemics	M1	2, 12
Fatality	M1, M3, M12, M28	32, 34
Necrotizing fasciitis	M1, M3, M28	34, 36, 37
Streptococcal toxic shock syndrome (STSS)	M1, M3	12, 35, 38
Impetigo	M33, M41, M42, M52, M53, M70	1
Puerperal sepsis	M28	34, 36, 37
Acute glomerulonephritis	M1, M4, M12, M49, M55, M57, M60	18
Meningitis	M1, M12	36

M proteins can activate coagulation and thrombosis. Several M proteins can activate the intrinsic pathway of coagulation on the bacterial surface by binding kininogen and generating bradykinin, while M1 and M3 have been shown to trigger the extrinsic pathway of coagulation on human monocytes and endothelial cells.[26] M1 and M5 bound to fibrinogen can recruit inactive platelets: in the presence of antibodies against bacterial antigens that also interact with the platelet Fc receptor, platelets are activated and thrombosis can ensue.[27,28]

Interaction with leukocytes appears to be at the core of the M protein pro-inflammatory role. Fibrinogen bound to soluble M1 protein is recognized by neutrophils via β2 integrins, resulting in activation and release of mediators of vascular leakage, a key occurrence in STSS. M1 also induces production of interleukin (IL)-6, IL-1 β and Tumor Necrosis Factor-α in monocytes via Toll-like receptors.[29] Fibrinogen-binding M proteins can activate platelets, and these in turn can activate neutrophils and monocytes, amplifying the pro-inflammatory effect.[28]

Some M proteins, such as M5, may act as superantigens,[18,19] inducing massive T cell proliferation and cytokine release, potentially contributing to scarlet fever and STSS; some belonging to the rheumatogenic kind (M5, M6, M12, M24) can elicit antibodies that cross-react with mammalian proteins, among them myosin, tropomyosin, laminin and keratin.[16,19]

The ability to induce the formation of host-reactive antibodies has led some to believe that M protein plays a direct role in the development of post-infection autoimmune disorders including acute rheumatic fever (ARF) and post-streptococcal glomerulonephritis (PSGN).[9,16,19] This hypothesis is mechanistically supported by physicochemical similarity and significant amino acid sequence homology between specific M proteins and specific host proteins. Different M proteins share homology with different host fibrillar proteins, and this diversity likely contributes to the variability seen between M types in terms of their potential to cause different types of nonsuppurative sequelae.[9,16,18,19]

A specific variant of the C block repeat region (Class I C) is the likely source of the shared antigenic factor known as MAP I. The presence of MAP I is highly correlated with rheumatogenicity (the ability to induce acute rheumatic fever).[9] A few MAP I strains are also associated with pyoderma (suppurative skin infection). Strains associated with one or more outbreaks of ARF in developed nations (including M1, M3, M5, M6, M11, M12, M14, M17, M18, M19, M24, M27, M29, M30, M32 and M41) share both the expression of MAP I antigen (as demonstrated by serology) and conserved epitopes in the C block repeat, while nonrheumatogenic strains do not.[9] Similarities between this shared antigen and human heart epitopes such as myosin and tropomyiosin are a possible source of rheumatogenic propensity,[1,9] and sequence analysis of this region might offer direct prediction of rheumatogenic potential. The MAP I antigen could possibly be used to generate a broad-spectrum vaccine offering specific protection against ARF (thus avoiding the problem of variable relationships between specific serotypes and ARF in remote locations where associated morbidity is high[30]), but such a vaccine might itself illicit autoimmune reactions and rheumatic conditions. Specific variants of the B block of the M protein share amino acid homology with other antigens (MHC class II).[31]

Clinical Relevance of M Type

Across the US and Europe, M type is closely correlated with specific clinical phenotypes. Many studies show that serotype diversity is significantly decreased among invasive isolates, as compared with the pharyngitis-associated strain collections. Moreover, the types that are most dominant in invasive categories, and hence those responsible for the observed decrease in diversity, are almost always the same—most notably M1 and M3.[8,12,18,32-36] Similar dynamics, including decreased diversity and the dominance of specific types, are apparent in studies of more specific clinical phenomena such as fatality, associated with M1, M3 and M12,[32] puerperal sepsis, associated with M28,[34,36,37] STSS, associated with M1 and M3,[12,35,38] epidemic ARF, associated with M5 and M18,[39] and geographically widespread epidemic behavior, most commonly associated with M1.[2,12] Despite genetic changes in these serotypes over time, including alterations of virulence and antibiotic resistance factors, these associations generally remain stable. This indicates that the M protein itself, or proteins coded by genes closely linked to the *emm* gene,

are the most likely sources of these specific clinical presentations. Examples of these associations are shown in **Table 1**.

The proportional dominance of rheumatogenic M types among noninvasive pharyngitis-associated isolates correlates with the rate of ARF in the US population.[40] As the proportion of rheumatogenic types (including M1, M3, M5, M6, M14, M18, M19, M24, M27 and M29) decreased from 49.7% in 1961–68 to 17.9% in 2000–2004, rheumatic fever essentially disappeared.[18,40] Similar relationships between proportional dominance of specific serotypes (particularly M1 and M3) and the rate of invasive GAS infection has also been documented, even in the absence of differential serotype dominance among invasive and noninvasive cases.[41] Periods of dominance of these strains correlate with regional increases in invasive GAS disease, suggesting that "appearance of M1 and M3 strains in patients with pharyngitis… may serve as sentinels to herald the appearance of invasive GAS disease in a region or community."[41]

Some phenotypes correlate very closely with specific M types. A study in Denmark comparing invasive isolates to uncomplicated pharyngitis strains showed that M1, M3 and M28 were dominant among invasive isolates.[34] Among these, M28 was specifically associated with puerperal sepsis, which the authors attributed to a unique collection of virulence factors apparently acquired from group B streptococci.[34] M28 has also been associated with puerperal sepsis in Norway.[36] A study in France found M28 in similar association with puerperal sepsis as well as other types of invasive infection, including necrotizing fasciitis, genital peritonitis, bacteremia, endocarditis and meningitis.[37] This strain was resistant to many antibiotics, including bacitracin. GAS is otherwise sensitive to bacitracin, but other β-hemolytic streptococci (such as group B strains) are resistant.[37] Cumulatively, this suggests a virulent and resistant strain to which it is well worth paying specific attention—best accomplished through M type-specific surveillance. Other correlations between M type and antibiotic resistance phenotypes are far less robust, as discussed in the following section.

Correlations between M Type and Antibiotic Resistance

Despite decades of primary treatment of GAS with penicillin[42] and the preemptive, long-term use of penicillin for GAS prophylaxis in high risk populations,[43,44] this pathogen has yet to evolve strong resistance.[43,45] Therefore, penicillin remains the treatment of choice and is generally effective. However, many people are allergic to penicillin and for these subpopulations macrolides and other antibiotics provide critical alternatives.[43] Specific GAS strains exhibit resistance to almost all categories of antibiotics aside from penicillin, though most strains exhibit only one or a few of the possible specific resistance phenotypes.[44] Diagnostic and surveillance testing for resistance is important for effective and timely treatment of GAS, particularly in severe and life-threatening cases of invasive disease.

Antibiotic resistance elements in GAS are of diverse origin and are highly mobile, many being transferred readily from strain to strain (and likely from other species to GAS) by phage-mediated recombination. Tetracycline resistance offers an excellent example.[46,47] Studies suggest the M type distributions of many resistance traits are not well conserved over time and space.[22,34,36,46-48] However, a recent US study showed that even these highly variable traits segregate significantly with M type over shorter periods of time.[44]

Macrolide resistance offers an excellent example of a phenotype that is more highly correlated with M type and for which M typing offers a significant degree of predictive power. A strong association between erythromycin (or other macrolide) resistance has been shown for M4, M6, M12 and M75—members of this small group are responsible for much of the observed macrolide resistance in surveys of GAS isolates from the US,[44,49,50] Belgium,[51] Greece[52] and Spain.[53,54] In general, the majority of isolates of these types are found to be resistant, while only scattered isolates of other serotypes exhibit the same resistance phenotype—with some local exceptions, such as the dominance of M2 among erythromycin resistant isolates in Italy.[55] In some cases, specific clones of a given serotype have been responsible for the spread of multiple resistance phenotypes derived from specific resistance cassettes—for example the spread of an erythromycin and tetracycline resistant clone of M77 [carrying the tet(O) determinant] in Italy, Norway and Denmark.[56,57] For most antibiotic resistance phenotypes, M typing can offer rapid results useful for estimating the risk of resistance if there is sufficient recent data from neighboring areas,[44] but greater power is offered by more specific clonal identification techniques such as MLST, PFGE or RFLP. Given the mobility of the genetic elements responsible for much of GAS resistance, even high resolution multilocus strain typing methods are only likely to correlate with specific resistance phenotypes over limited temporal and spatial distance.

Some rare forms of antibiotic resistance are very tightly correlated with specific M types—for example, ofloxacin resistance reveals tight linkage to M6 in both the US[44] and Norway.[36] Almost all M6 isolates studied in these reports were resistant to ofloxacin, while all other M types were susceptible. Another example of tight correlation between rare antibiotic resistance phenotypes and specific M types comes from a study in France which showed a unique resistance to bacitracin (paired with wide resistance to other diverse antibiotics) among M28 isolates.[37] Another study showed that isolates of a variety of M types collected from apparently hospital acquired GAS infections (HAI GAS) expressed more (and more diverse) antibiotic resistance phenotypes than non-hospital acquired strains, and those strains tended to express resistance phenotypes not otherwise associated with their M types.[21] The number of HAI GAS strains studied was very small, but the results were suggestive of antibiotic resistance element transfer in the hospital environment. Further study of M type diversity and antibiotic resistance distribution among apparent nosocomial GAS infections is warranted to address the hypothesis that GAS may be accumulating resistance elements and being selected for resistance in such circumstances.

Just as changes in serotype dominance may drive changes in the rate of invasive GAS disease,[22] the rise and fall of antibiotic resistance at specific surveillance sites appears to be the result of large scale strain turnover events, likely driven by herd immunity

rather than local antibiotic use patterns.[21] This reiterates the importance of considering M type distributions and the forces driving changes in those distributions when attempting to understand the etiology of changes in resistance rates or rates of specific clinical outcomes.

Geographic Variability in M Type-Specific Genetic and Phenotypic Correlates

The developed world shares strains in which the associations between M type and tissue tropism, virulence and nonsuppurative sequelae are recognized and robust. Surveillance efforts in less developed areas, and specifically in rural tropical communities with little outside contact, have revealed populations of GAS in which many of these associations are completely different.[5] Distributions in Africa and on Pacific islands are more heterogeneous (lacking in distinctly dominant strains), exhibit very different M type distributions, and include strains that bear M types common in developed nations but are associated with completely different clinical correlates.[5,30] In the developed world, there are well-documented associations between specific M types and tissue tropism (**Table 1**). There appears to be a stable separation between serotypes that cause pharyngitis and serotypes that cause impetigo, suggesting strong genetic linkage between *emm* genes and the genetic determinants of tissue tropism. In developed regions, serious sequelae such as ARF, along with invasive and fatal acute infections, are almost exclusively associated with pharyngitis-associated types. These associations do not hold true in island populations of Australian aborigines,[5,58] in villages in Thailand[59,60] or in Ethiopia.[61] This is likely reflective of relative isolation rather than environmental selection, as separate island communities in close proximity (and presumably with very similar environments, but little travel between them) carry very different strains.[58] The diversity at the source of these different populations likely arose from recombination over long time periods, as even the correlations between M type and MLST alleles (which are extremely robust and consistent throughout the developed world) differ between nearby islands harboring aboriginal populations.[58] It is likely that recombination between the antigenic fragment of M and the regions of M associated with rheumatic sequelae (and possibly with superantigenicity) is responsible for the decoupling of serotype and clinical outcome seen in these exotic isolates. Sequence analysis of the full *emm* gene of such strains would be very interesting and informative.

It is worth noting that these isolated rural populations carry far higher burdens of GAS-related morbidity and mortality than do the populations of developed nations.[30] Proposed M protein-specific vaccines targeting the most rheumatogenic or invasive serotypes found in the US and Europe would likely do little good in these communities.[30] It is also worth considering the risk that, if we eradicate our own virulent GAS strains with vaccines designed against the strains that represent the greatest threat in the developed world, it may open the door for the emergence of new M types with new virulence properties from less-developed areas. The bulk of this review focuses on associations seen in the relatively homogeneous developed regions of the world where historical data are dense enough to offer significant conclusions, but the reader should keep in mind the fact that the small amount of data we have from undeveloped and developing regions suggests that unique type-specific associations exist in such areas.

Secondary Determinants of Virulence

The focus of GAS typing efforts has shifted occasionally away from M type toward other markers of virulence. The resurgence of ARF in the 1980s is an example. The majority of strains collected in association with ARF cases and contacts at this time were mucoid (that is, they revealed a mucoid colony morphology upon initial plate culture), causing many people to focus on mucoid phenotype and question the relevance of M type. However, the fact that most of these outbreaks were associated with mucoid strains of a few serotypes—mainly M1, M3 and M18[18,62]—points to the difficulties involved in discriminating between potential contributing factors (here, M type and mucoid phenotype).

A study in the late 1980s showed that M1 strains were more likely to generate mucoid colonies if collected from patients with rheumatic fever (80%) and invasive infections (22%) than if from uncomplicated pharyngitis (6%), and this general association held true across several other M types.[62] More recent data showed an association between a mucoid strain of M3 and apparently high virulence,[63] but again this observation was limited to a single serotype, a serotype which has been specifically associated with high virulence for decades. These sorts of observations could be the result of strains circulating with unique virulence factors directly related to their mucoid phenotype (as several authors have suggested), or it could be the result of a regulatory or mutational switch that causes strains to change from a nonmucoid to a mucoid phenotype in the course of invasive infection (something akin to, or resulting from, the *covS* mutational switch from noninvasive to invasive phenotypes discussed later, in the section on GAS evolution). It could also result from the spread of an invasive strain carrying the genotypic determinants of a mucoid phenotype in which invasivity has nothing to do with the mucoid phenotype, but is merely linked to it genetically. It has been noted that the mucoid phenotype may be rapidly lost when strains are cultured in the laboratory, suggesting mucoid appearance is an unstable phenotype.[2] The important points here are that phenotypes (like mucoid morphology, and in contrast to genotypic traits like *emm* type) are not necessarily stably heritable strain markers. No morphological phenotype has ever demonstrated the robust predictive power offered by M typing.

The Spe (streptococcal pyrogenic exotoxin) proteins A–D, F–H and J, and other proteins including streptococcal superantigen (SSA) and the streptococcal mitogenic exotoxins SMEZ and Z-2 are a diverse group of virulence factors capable of inducing proliferation of specific T cell populations and causing release of excessive amounts of inflammatory cytokines.[1] The B block of the M protein may, in certain serotypes such as M5, also act as a superantigen and possibly contribute to inflammatory responses.[31] All of these proteins contribute to the virulence of GAS, most specifically to severe acute reactions such as scarlet

fever and streptococcal toxic shock syndrome. Many of these have also been implicated in ARF and other post-infection autoimmune sequelae, though this connection is far less clear for all except the M protein.[1] Many virulence phenotypes have been attributed to specific pyrogenic exotoxins, but in many cases the observed associations are confounded by association with specific M types, similar to the situation described for mucoid phenotypes. While these proteins certainly have an impact on virulence (especially severe acute reactions), it may well be a combinatorial effect dependent on both the presence and expression level of multiple exotoxins as well as on the M type background of the strain. Some of the genes coding for these proteins are phage-mobilized, including *speA* and *speC*, while others are not, including *speB*.[1]

An increase in invasive, life-threatening septic GAS infections was seen in both the US and the UK between the 1970s and 1980s. This was associated with an increase in the prevalence of well-recognized virulent M types, primarily M1, M3 and M18.[18] Some studies of this resurgence in the US identified an association between the presence of the *speA* gene and both streptococcal toxic shock syndrome and fatality.[18,64,65] A specific clone of M1 was generally associated with this increase in invasive disease, specifically sepsis, and some argued that the strain's invasivity might be related to the presence of *speA*, as the emergent clone carried *speA* while previously dominant M1 strains did not.[65] A similar rise in invasive, septic GAS disease was observed during the same time period in the UK, involving the same M types.[18,66] In contrast to what was seen in the US, no strains producing SpeA were identified from patients with severe or fatal acute infections,[18,66] suggesting that the increase in invasivity was related to an increase in the dominance of virulent M types rather than the presence or absence of *speA*.[18]

SpeA has also been implicated in ARF, but (for example), while M5 and M18 are both strongly associated with ARF,[67] the M18 strain associated with ARF carries *speA*[39] while M5 does not.[44,67] The author of a review covering these subjects noted that, at least among developed nations, the M types most clearly associated with ARF are well-recognized and have remained consistent since the advent of serotyping assays.[18] These include M1, M3, M5, M6 and M18. This author provided the simplest possible conclusion from this wealth of data: "M protein is the chief virulence factor of group A streptococcus."[18]

While secondary factors may play a very significant role in GAS virulence, none have the robust associations with virulence phenotypes as M type. This ultimate conclusion is supported by numerous symptom correlations which are consistent with respect to M type but apparently independent of the presence or absence of other virulence factors (as discussed throughout this review). The only robust correlations between GAS genotypes and specific clinical outcomes are those that derive from M type.

A decade of whole-genome comparative sequence analysis studies and transcriptome and proteome analyses (genome-wide analysis of quantitative gene transcription and protein translation patterns, comparing different genotypes and/or different environmental influences), paired with strain-specific epidemiological analyses, have revealed many new mechanisms and sources of pathogenic traits in GAS, as summarized in a recent review in reference 20. This systems-based approach is known as molecular pathogenomics, and will ultimately provide the path to a full understanding of the molecular basis of the different clinical behavior of distinct GAS strains.

The mechanisms of pathogenesis revealed through this approach are as diverse as they are numerous. The recent clonal expansion of a uniquely virulent strain of M1 appears to be driven by bacteriophage transduction events leading to changes in bacteriophage content and organization as well as reciprocal recombination of a large region carrying well-recognized virulence factors including streptolysin O and NADase.[14] The transferred region, including both the virulence factors and associated regulatory elements, appears to have come from an M12 strain.[14]

Subtle changes in antigenic regions, including changes in M that alter strain-specific immunity in humans but which do not alter the serotype as defined by traditionally accepted culture-based serological analyses, may alter the efficacy of strain-specific host immunity. This could allow previously common strains to re-emerge and create new epidemics, as may be the case for M3. A recently emergent and virulent M3 strain experienced a small (four amino acid) duplication in the M protein itself, associated with the resurgence of this M type during an epidemic of GAS disease in 2000.[20,68] This alteration caused changes in host immune responses, suggesting significant antigenic novelty. Offering an even more striking example of the potential effects of dispersed and subtle mutations, a genome wide (mutation) association study (GWAS) offered evidence that a single nucleotide polymorphism in a metalloregulatory protein is associated with a change in the ability of GAS to cause necrotizing fasciitis, likely through alteration of the regulation of SpeB expression.[20,69]

Whole genome sequencing of an apparently clonal M18 strain associated with several US outbreaks of acute rheumatic fever in the 1980s and 1990s, revealed many virulence factors.[39] To offer an idea of the complexity of determining the etiological basis of differential virulence and outcomes, M18 was found to carry 178 open reading frames (most of them likely functional genes) not found in the more common serotype M1, which has been associated with sporadic cases of ARF but not the high attack rate epidemic ARF characteristic of M18.[39]

Complete sequencing of an M5 strain, also strongly associated with ARF (strain Manfredo), also revealed many previously implicated virulence factors. However, only one coding sequence, a single gene encoding an uncharacterized surface-anchored protein (SpyM50104), was shared with M18 but not with the 10 other serotypes (which are not associated with epidemic ARF) for which whole genome sequence data are available.[67] The authors do not attempt to claim that this protein is associated with rheumatogenicity, as little is known about its biochemical properties. It is important to remember that both of these strains also share the class I C repeat features of the M protein, itself a likely direct source of rheumatogenicity, and they also share thousands of single base mutations not shared with less rheumatogenic types.[67]

Comparative genomics suggests that M28 strains associated with puerperal sepsis (child-bed fever, usually associated with

group B streptococci) may express this phenotype due to the acquisition of a large fragment of group B genomic sequence.[20,34,37]

Host factors certainly contribute greatly to GAS virulence and clinical outcome, adding significant complexity to any effort to deconvolute the source of differential presentation. The best-characterized host factor is probably that of infection by Varicella Zoster Virus (VZV or human herpesvirus 3).[70,71] Recent VZV infection increases the risk of invasive GAS by a factor of 58.[71] Other factors, such as cytokine response patterns and the presence of antistrep or antitoxin antibodies, likely also play a role.[70] Past infection and the presence of type-specific antibodies to a specific strain greatly limit the severity and potential invasivity of future colonization by the same strain.

Environmental factors (such as poverty and crowding) may also play a role—one study noted that American Indians in Arizona suffered 10 times the rate of invasive GAS disease than other residents,[35] while another noted that ARF has persisted in low-income black and Hispanic communities in the United States during times when it was very rare in other populations.[18] The high rates of ARF in undeveloped aboriginal communities also suggest a strong environmental component.[30] Differences in host response to superantigens may also play a role in severity of symptoms.[1,20] Investigation of host-specific factors led to the discovery that specific HLA alleles are involved in the development of streptococcal toxic shock syndrome[20] and ARF,[18] suggesting that direct prediction of specific case outcomes might require analysis of not only the complete genome of the infecting GAS strain, but also analysis of the host genome.

GAS Evolution

Mobile phages and phage-like elements encoding virulence factors and antibiotic resistance elements account for the majority of variation in gene content between M types and almost all of the variability within M types.[4,12-14,20,72] The genes mobilized by these elements include pyrogenic superantigens (*speA* and *speC*), DNases, phospholipase A2 (similar to snake venom toxin), *mefA* (a macrolide efflux factor) and many others.[4,12,48] Many of these have been associated with specific phenotypes, but none have proven to be solely responsible for, or uniquely associated with, specific symptoms, sequelae or outcomes.[10,12] It is possible that these virulence factors act in groups to produce specific effects on hosts and their effects may be modulated by the M protein with which they are associated, as the M protein is clearly a critical virulence factor itself.[16,18] Strains causing particularly virulent outbreaks often display unique combinations or physical organizations of prophages and associated virulence factors. While specific phage-associated virulence factors are not always associated with invasivity, the total number of these factors may be related to invasive potential.[48] The phage-associated virulence factor content and organization of specific M types, like M3, have evolved actively since the advent of GAS surveillance and collection,[12] yet the most important type-specific correlations (such as the correlation between M3 and invasivity/fatality) have remained intact.

Phage-associated virulence factors may play a role in pathogenicity, but that role appears to be secondary to that of either the M protein itself or with as-yet undocumented virulence factors that are more closely linked to M type. These could potentially include chromosomally encoded virulence factors such as streptolysins, C5a peptidase and extracellular cysteine protease.[12]

If both phage content and genomic position contribute to virulence, then only complete genome sequencing has a chance of capturing all the relevant information needed to predict virulence precisely, and even that would only be useful if we understood both the independent and combinatorial contributions of the many known and suspected virulence factors carried by GAS. Until then, the best we can do is identify existing and emerging virulent clones through serotype-specific surveillance, using signatures such as M type and MLST paired with records of the associations between specific genotypes and clinically relevant phenotypes (including virulence, invasivity, rheumatogenicity and antibiotic resistance). These associations will change over time, but the evidence suggests they change slowly enough that M typing (combined, when necessary, with secondary typing such as MLST to discriminate phenotypically divergent clones) can allow a very significant degree of predictive power.

Some aspects of GAS pathogenicity and disease specificity are modulated by local mutation, rather than by presence or absence of specific genes. Frameshift mutations in the *covS* ("control of virulence") gene generate widespread changes in virulence gene expression (the levels of expression of at least 23 potential virulence factors change significantly). In turn, this results in rapid switching from a noninvasive pharyngeal phenotype to an invasive phenotype.[73] Frameshift mutations capable of generating this change can occur over more than 500 bp of the *covS* open reading frame,[73] suggesting that monitoring this locus for all relevant mutations would be difficult. Moreover, this switch can occur readily in the course of a single infection, and hence the presence of the noninvasive allele does not insure that an infection won't become invasive. However, as *covS* is a regulatory switch as opposed to a direct virulence factor, it can be assumed that only an M type with inherent invasive potential can become invasive on the basis of this switch. As a hypervariable switching mechanism, it is neither practical nor predictive to directly monitor the *covS* gene. It does, however, offer one partial explanation for why M types with a demonstrated potential for invasivity are commonly detected in noninvasive infections.

The genetic factors directly affecting GAS virulence are far too numerous to rapidly analyze with contemporary technologies, and our knowledge of those factors is rudimentary at best. It seems wise to encourage the development of rapid strain-typing systems (such as *emm*-typing and multi-locus genotyping methods[44]) for use in clinical diagnosis, where they can provide significant predictive power and aid in immediately relevant decisions (such as choice of antibiotics and analysis of the risk of transmission and invasive disease). It would be counterproductive to abandon them simply because they are imprecise or incomplete.

Despite observed changes in the genetic background of specific serotypes (such as M1, M3 and M5, as discussed here), these same types are consistently associated with outbreaks and severe invasive disease. The fluctuating dominance and precise clinical correlates of these strains are certainly affected by mutation,

horizontal gene transfer and changes in herd immunity among their human hosts, but the key virulence features that set these strains apart from other serotypes has remained the same over the course of several decades. Among all the diverse factors that contribute to the virulence of GAS, M protein is one of the most important and certainly the best characterized and understood. In the absence of whole-genome sequencing in a diagnostic setting (and the ability to fully interpret the resulting data in a clinically relevant manner), M typing offers a very attractive level of predictive power for clinical outcome, public health risk and antibiotic resistance. As a primary virulence factor contributing to superantigenic and rheumatogenic properties, it also provides direct predictive power for specific aspects of virulence.

Population Biology and Dynamic Epidemiology of GAS

As an obligate human pathogen, the population dynamics of GAS are subject to the primary influences of both type-specific immunity (from either past exposure or vaccination) and interventions such as antibiotic treatment.[4] Because the M protein is the primary target of type-specific human immunity and because GAS is a human pathogen, GAS strain distributions are likely forged by processes of herd immunity affecting specific serotypes and resulting strain replacements (as discussed in the section "M protein as a virulence factor").[4,21] Specific strains are favored if the population has low specific immunity to those strains (by nature of their recent rarity), and those strains can spread and replace previously dominant strains. Increasing dominance increases the rate of specific immunity in the host population, in turn favoring other strains. This process is cyclic, and favors the maintenance of a broad variety of diverse serotypes. It is quite easy to mistake the dynamic strain turnovers driven by herd immunity for selection events driven by other processes.

In 2002–2004 (and earlier and more transiently in 1999–2000), M75 became dominant across the US among isolates collected from military recruits with pharyngitis.[21,74] In 2004–2007, the same serotype emerged and became quite dominant in Finland, as seen in studies of invasive GAS isolates.[22] Isolates from the US were resistant to macrolides,[21] while isolates from Finland were sensitive.[22] Resistant M75 was also observed in Spain.[53] In the US, macrolide resistant M75 became dominant at approximately the same time at multiple geographically distant recruit training centers.[21] Some of these sites actively used macrolides for prophylaxis and/or treatment of GAS in penicillin-allergic recruits, while others did not use macrolides.[21,44] The observed pattern of temporally correlated increases in dominance of a specific strain, regardless of antibiotic use or antibiotic resistance, suggests that the rise of M75 represented a strain turnover event driven by low herd immunity to a specific serotype, rather than by antibiotic use or resistance factors.

An invasive macrolide- and tetracycline-resistant strain of M77 was observed in Italy and Norway in the 1990s and early 2000s, representing a high proportion of isolates from invasive cases.[56] The same M type was observed in the US in the late 90s,[74] and later in the following decade[21,44] when it was also associated with at least one fatal case of invasive disease. The US strain, in contrast to the European strain, was macrolide-sensitive. This is a common theme—strain turnover events appear to be driven by global factors such as herd immunity and variable resistance elements and phage-mobilized virulence factors are carried along by genetic hitchhiking in cases where they are linked with a favored (or disfavored) M type. This dynamic suggests that, while attention should be paid to the rates and distributions of antibiotic resistance elements, efforts to explain their distributions on the basis of local antibiotic use must take into account M type distributions on larger geographic scales. Epidemiologists should recognize the convincing evidence that large-scale strain turnovers are primarily driven by type-specific selection (such as herd immunity) rather than by the presence or absence of resistance. It seems likely, however, that local antibiotic use in the regions where epidemic clones originate do directly influence the antibiotic resistance levels exhibited by emerging strains.

Horizontal transfer and recombination driven by phages are driving forces of GAS evolution.[4,13] These processes alter the association between M alleles and other virulence factors, generating intra-serotype diversity. Hence, all members of any given M type are not genetically identical, nor do they necessarily share all phenotypic traits.[13] As the authors of one paper note, "…insufficient information is available to estimate the rate at which such recombination events occur in nature. A high rate in evolutionary terms could nevertheless be sufficiently low to allow the M type to provide a useful marker over relatively short periods."[13] While this may or may not be true with respect to antibiotic resistance, which clearly derives from proteins other than M, plenty of data exists to support the geographically broad and temporally stable associations between specific M types and specific clinical outcomes. This may be the result of genomic stability over the studied time frames, or it may be the result of serotype-specific clinical properties deriving directly from the M protein. In situations where it becomes necessary to distinguish strains of any given M type, multi-locus sequence typing (MLST) can reliably provide higher resolution[4] with a corresponding increase in typing effort (MLST requires sequencing seven housekeeping gene segments, for example).[75] MLST types have been carefully validated, shown to present stable associations with M types over the span of decades and across multiple continents and MLST types are defined in a curated online database with descriptions of standardized techniques for data collection and analysis.[36,75]

The Past, Present and Future of GAS Typing and Molecular Epidemiology

Some GAS experts have suggested that the widespread use of M typing for GAS strain discrimination led to the incorrect notion that shared serotype was indicative of identical genotype, and hence phenotype.[72,76] The discovery that serotypes may include multiple genetically and phenotypically divergent strains does not obviate the value of M typing. To the contrary—despite great effort, we have yet to find other loci that offer as much discriminatory power as does M, and no other virulence

factors appear to contribute as directly to measureable phenotypes such as invasivity and rheumatogenicity. The associations between these traits and specific M types have remained stable over half a century and are likely derived at least in part from M protein itself. Efforts to identify specific clones of particularly virulent serotypes such as M1, M3 and M28 have ultimately led to the conclusion that, while there are multiple genetically diverse strains of these M types in circulation, none of them can be convincingly proven to be more closely associated with severe disease than others.[23]

Traits unrelated to the M protein, including some antibiotic resistance phenotypes, are correlated with M type, at least over time periods of a decade or more within continental geographic scales.[21,36,44,55,56] M type can provide a rapid and widely accessible method for indirectly (and admittedly, partially) predicting outcome, optimizing antibiotic therapy and estimating public health risk. While M itself contributes to some of these phenotypes, the other genetic elements responsible for virulence and resistance are numerous, diverse and affected by each other, by host factors, and by variable regulatory elements (both cis and trans). Phage mobility and other factors lead to significant strain diversity within serotypes.[12-14,72] All these factors make direct and unambiguous prediction of phenotype from genetic analysis essentially impossible. That leaves us with indirect (or partially indirect) strain identifiers, and M is the best marker we have (in three distinct ways—it is the most diverse single-locus strain marker, it contributes directly to important phenotypes like rheumatogenicity, and it has the richest history of correlative data supporting inference of M-independent phenotypes like antibiotic resistance). The medical community should be aware of the predictive potential of rapid strain typing, but also of the imperfect and sometimes ambiguous nature of phenotype inference from marker genotypes. Furthermore, they should be aware of the critical role that reference lab surveillance plays in maintaining and improving our knowledge of the temporally and geographically varying relationships between M type and phenotype, and the critical role medical providers play in that process through reporting of severe illness and the contribution of associated samples, isolates and antibiotic efficacy data to surveillance agencies.

References

1. Cunningham MW. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev 2000; 13:470-511; PMID:10885988; http://dx.doi.org/10.1128/CMR.13.3.470-511.2000.
2. Stollerman GH. Rheumatic fever in the 21st century. Clin Infect Dis 2001; 33:806-14; PMID:11512086; http://dx.doi.org/10.1086/322665.
3. Anthony BF, Kaplan EL, Wannamaker LW, Chapman SS. The dynamics of streptococcal infections in a defined population of children: serotypes associated with skin and respiratory infections. Am J Epidemiol 1976; 104:652-66; PMID:793381.
4. Bessen DE. Population biology of the human restricted pathogen, *Streptococcus pyogenes*. Infect Genet Evol 2009; 9:581-93; PMID:19460325; http://dx.doi.org/10.1016/j.meegid.2009.03.002.
5. McDonald M, Currie BJ, Carapetis JR. Acute rheumatic fever: a chink in the chain that links the heart to the throat? Lancet Infect Dis 2004; 4:240-5; PMID:15050943; http://dx.doi.org/10.1016/S1473-3099(04)00975-2.
6. Oehmcke S, Shannon O, Mörgelin M, Herwald H. Streptococcal M proteins and their role as virulence determinants. Clin Chim Acta 2010; 411:1172-80; PMID:20452338; http://dx.doi.org/10.1016/j.cca.2010.04.032.
7. Smeesters PR, McMillan DJ, Sriprakash KS. The streptococcal M protein: a highly versatile molecule. Trends Microbiol 2010; 18:275-82; PMID:20347595; http://dx.doi.org/10.1016/j.tim.2010.02.007.
8. Schwartz B, Facklam RR, Breiman RF. Changing epidemiology of group A streptococcal infection in the USA. Lancet 1990; 336:1167-71; PMID:1978035; http://dx.doi.org/10.1016/0140-6736(90)92777-F.
9. Bessen D, Jones KF, Fischetti VA. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. J Exp Med 1989; 169:269-83; PMID:2642529; http://dx.doi.org/10.1084/jem.169.1.269.
10. Stollerman GH. Rheumatic fever. Lancet 1997; 349:935-42; PMID:9093263; http://dx.doi.org/10.1016/S0140-6736(96)06364-7.
11. Bisno AL. The concept of rheumatogenic and nonrheumatogenic group A streptococci. In: Read SE, Zabriskie JB, Eds. Streptococcal Diseases and the Immune Response. New York, NY: Academic Press Inc. 1980; 789-903.
12. Banks DJ, Beres SB, Musser JM. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. Trends Microbiol 2002; 10:515-21; PMID:12419616; http://dx.doi.org/10.1016/S0966-842X(02)02461-7.
13. Kehoe MA, Kapur V, Whatmore AM, Musser JM. Horizontal gene transfer among group A streptococci: implications for pathogenesis and epidemiology. Trends Microbiol 1996; 4:436-43; PMID:8950813; http://dx.doi.org/10.1016/0966-842X(96)10058-5.
14. Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, et al. Evolutionary origin and emergence of a highly successful clone of serotype M1 group a Streptococcus involved multiple horizontal gene transfer events. J Infect Dis 2005; 192:771-82; PMID:16088826; http://dx.doi.org/10.1086/432514.
15. Lancefield RC. The antigenic complex of *Streptococcus hemolyticus*. I. Demonstration of a type-specific substance in extracts of *Streptococcus hemolyticus.* J Exp Med 1928; 47:91-103; PMID:19869404; http://dx.doi.org/10.1084/jem.47.1.91.
16. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. Clin Microbiol Rev 1989; 2:285-314; PMID:2670192.
17. Facklam R, Beall B, Efstratiou A, Fischetti V, Johnson D, Kaplan E, et al. *emm* typing and validation of provisional M types for group A streptococci. Emerg Infect Dis 1999; 5:247-53; PMID:10221877; http://dx.doi.org/10.3201/eid0502.990209.
18. Bisno AL. Group A streptococcal infections and acute rheumatic fever. N Engl J Med 1991; 325:783-93; PMID:1870652; http://dx.doi.org/10.1056/NEJM199109123251106.
19. Robinson JH, Kehoe MA. Group A streptococcal M proteins: virulence factors and protective antigens. Immunol Today 1992; 13:362-7; PMID:1281632; http://dx.doi.org/10.1016/0167-5699(92)90173-5.
20. Musser JM, Shelburne SA, 3rd. A decade of molecular pathogenomic analysis of group A Streptococcus. J Clin Invest 2009; 119:2455-63; PMID:19729843; http://dx.doi.org/10.1172/JCI38095.
21. Metzgar D, McDonough EM, Hansen CJ, Blaesing CR, Baynes D, Hawksworth AW, et al. Local changes in rates of group A Streptococcus disease and antibiotic resistance are associated with geographically widespread strain turnover events. Virulence 2010; 1:247-53; PMID:21178449; http://dx.doi.org/10.4161/viru.1.4.11979.
22. Siljander T, Lyytikäinen O, Vähäkuopus S, Snellman M, Jalava J, Vuopio J. Epidemiology, outcome and *emm* types of invasive group A streptococcal infections in Finland. Eur J Clin Microbiol Infect Dis 2010; 29:1229-35; PMID:20563620; http://dx.doi.org/10.1007/s10096-010-0989-9.
23. Johnson DR, Wotton JT, Shet A, Kaplan EL. A comparison of group A streptococci from invasive and uncomplicated infections: are virulent clones responsible for serious streptococcal infections? J Infect Dis 2002; 185:1586-95; PMID:12023764; http://dx.doi.org/10.1086/340576.
24. Kaplan EL, Wotton JT, Johnson DR. Dynamic epidemiology of group A streptococcal serotypes associated with pharyngitis. Lancet 2001; 358:1334-7; PMID:11684215; http://dx.doi.org/10.1016/S0140-6736(01)06415-7.
25. Sun H, Ringdhal U, Homeister JW, Fay WP, Engleberg NC, Yang AY, et al. Plasminogen is a critical host pathogenicity factor for group A streptococcal infection. Science 2004; 305:1283-6; PMID:15333838; http://dx.doi.org/10.1126/science.1101245.
26. Påhlman LI, Malmström E, Mörgelin M, Herwald H. M protein from *Streptococcus pyogenes* induces tissue factor expression and pro-coagulant activity in human monocytes. Microbiology 2007; 153:2458-64; PMID:17660410; http://dx.doi.org/10.1099/mic.0.2006/003285-0.
27. Sjöbring U, Ringdahl U, Ruggeri ZM. Induction of platelet thrombi by bacteria and antibodies. Blood 2002; 100:4470-7; PMID:12393652; http://dx.doi.org/10.1182/blood-2002-01-0069.
28. Shannon O, Hertzén E, Norrby-Teglund A, Mörgelin M, Sjöbring U, Björck L. Severe streptococcal infection is associated with M protein-induced platelet activation and thrombus formation. Mol Microbiol 2007; 65:1147-57; PMID:17662041; http://dx.doi.org/10.1111/j.1365-2958.2007.05841.x
29. Påhlman LI, Mörgelin M, Eckert J, Johansson L, Russell W, Riesbeck K, et al. Streptococcal M protein: a multipotent and powerful inducer of inflammation. J Immunol 2006; 177:1221-8; PMID:16818781.
30. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global *emm* type distribution of group A streptococci: systematic review and implications for vaccine development. Lancet Infect Dis 2009; 9:611-6; PMID:19778763; http://dx.doi.org/10.1016/S1473-3099(09)70178-1.

31. Wang B, Schlievert PM, Gaber AO, Kotb M. Localization of an immunologically functional region of the streptococcal superantigen pepsin-extracted fragment of type 5 M protein. J Immunol 1993; 151:1419-29; PMID:8335938.
32. O'Loughlin RE, Roberson A, Cieslak PR, Lynfield R, Gershman K, Craig A, et al. The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States 2000–2004. Clin Infect Dis 2007; 45:853-62; PMID:17806049; http://dx.doi.org/10.1086/521264.
33. Ekelund K, Darenberg J, Norrby-Teglund A, Hoffmann S, Bang D, Skinhøj P, et al. Variations in *emm* type among group A streptococcal isolates causing invasive or noninvasive infections in a nationwide study. J Clin Microbiol 2005; 43:3101-9; PMID:16000420; http://dx.doi.org/10.1128/JCM.43.7.3101-9.2005.
34. Luca-Harari B, Ekelund K, van der Linden M, Staum-Kaltoft M, Hammerum AM, Jasir A. Clinical and epidemiological aspects of invasive *Streptococcus pyogenes* infections in Denmark during 2003 and 2004. J Clin Microbiol 2008; 46:79-86; PMID:17959766; http://dx.doi.org/10.1128/JCM.01626-07.
35. Hoge CW, Schwartz B, Talkington DF, Breiman RF, MacNeill EM, Englender SJ. The changing epidemiology of invasive group A streptococcal infections and the emergence of streptococcal toxic shock-like syndrome. A retrospective population-based study. JAMA 1993; 269:384-9; PMID:8418346; http://dx.doi.org/10.1001/jama.269.3.384.
36. Meisal R, Andreasson IK, Høiby EA, Aaberge IS, Michaelsen TE, Caugant DA. *Streptococcus pyogenes* isolates causing severe infections in Norway in 2006 to 2007: *emm* types, multilocus sequence types and superantigen profiles. J Clin Microbiol 2010; 48:842-51; PMID:20042624; http://dx.doi.org/10.1128/JCM.01312-09.
37. Mihaila-Amrouche L, Bouvet A, Loubinoux J. Clonal spread of *emm* type 28 isolates of *Streptococcus pyogenes* that are multiresistant to antibiotics. J Clin Microbiol 2004; 42:3844-6; PMID:15297545; http://dx.doi.org/10.1128/JCM.42.8.3844-6.2004.
38. Szczypa K, Sadowy E, Izdebski R, Strakova L, Hryniewicz W. Group A streptococci from invasive-disease episodes in Poland are remarkably divergent at the molecular level. J Clin Microbiol 2006; 44:3975-9; PMID:16957033; http://dx.doi.org/10.1128/JCM.01163-06.
39. Smoot JC, Barbian KD, Van Gompel JJ, Smoot LM, Chaussee MS, Sylva GL, et al. Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks. Proc Natl Acad Sci USA 2002; 99:4668-73; PMID:11917108; http://dx.doi.org/10.1073/pnas.062526099.
40. Shulman ST, Stollerman G, Beall B, Dale JB, Tanz RR. Temporal changes in streptococcal M protein types and the near-disappearance of acute rheumatic fever in the United States. Clin Infect Dis 2006; 42:441-7; PMID:16421785; http://dx.doi.org/10.1086/499812.
41. Kiska DL, Thiede B, Caracciolo J, Jordan M, Johnson D, Kaplan EL, et al. Invasive group A streptococcal infections in North Carolina: epidemiology, clinical features and genetic and serotype analysis of causative organisms. J Infect Dis 1997; 176:992-1000; PMID:9333158; http://dx.doi.org/10.1086/516540.
42. Wannamaker LW, Denny FW, Perry WD, Rammelkamp CH Jr, Eckhardt GC, Houser HB, et al. The effect of penicillin prophylaxis on streptococcal disease rates and the carrier state. N Engl J Med 1953; 249:1-7; PMID:13063671; http://dx.doi.org/10.1056/NEJM195307022490101.
43. Russell KL. Respiratory infections in military recruits. In: Lenhart MK, Lounsbury DE, North RB Jr, Eds. Textbooks of Military Medicine: Recruit Medicine. Washington, DC: Borden Institute (Walter Reed Army Medical Center) 2006; 1:227-53.
44. Metzgar D, Baynes D, Hansen CJ, McDonough EA, Cabrera DR, Ellorin MM, et al. Inference of antibiotic resistance and virulence among diverse group A Streptococcus strains using *emm* sequencing and multilocus genotyping methods. PLoS ONE 2009; 4:6897; PMID:19730733; http://dx.doi.org/10.1371/journal.pone.0006897.
45. Horn DL, Zabriskie JB, Austrian R, Cleary PP, Ferretti JJ, Fischetti VA, et al. Why have group A streptococci remained susceptible to penicillin? Report on a symposium. Clin Infect Dis 1998; 26:1341-5; PMID:9636860; http://dx.doi.org/10.1086/516375.
46. Ayer V, Tewodros W, Manoharan A, Skariah S, Luo F, Bessen DE. Tetracycline resistance in group a streptococci: emergence on a global scale and influence on multiple-drug resistance. Antimicrob Agents Chemother 2007; 51:1865-8; PMID:17307980; http://dx.doi.org/10.1128/AAC.01341-06.
47. Giovanetti E, Brenciani A, Lupidi R, Roberts MC, Varaldo PE. Presence of the tet(O) gene in erythromycin- and tetracycline-resistant strains of *Streptococcus pyogenes* and linkage with either the mef(A) or the erm(A) gene. Antimicrob Agents Chemother 2003; 47:2844-9; PMID:12936983; http://dx.doi.org/10.1128/AAC.47.9.2844-9.2003.
48. Vitali LA, D'Ercole S, Petrelli D, Di Luca MC, Rombini S, Prenna M, et al. Distribution of phage-associated virulence genes in pharyngeal group a streptococcal strains isolated in Italy. J Clin Microbiol 2009; 47:1575-7; PMID:19279174; http://dx.doi.org/10.1128/JCM.00925-08.
49. Kaplan EL, Johnson DR, Del Rosario MC, Horn DL. Susceptibility of group A beta-hemolytic streptococci to thirteen antibiotics: examination of 301 strains isolated in the United States between 1994 and 1997. Pediatr Infect Dis J 1999; 18:1069-72; PMID:10608626; http://dx.doi.org/10.1097/00006454-199912000-00008.
50. Martin JM, Green M, Barbadora KA, Wald ER. Erythromycin-resistant group A streptococci in schoolchildren in Pittsburg. N Engl J Med 2002; 346:1200-6; PMID:11961148; http://dx.doi.org/10.1056/NEJMoa013169.
51. Descheemaeker P, Chapelle S, Lammens C, Hauchecorne M, Wijdooghe M, Vandamme P, et al. Macrolide resistance and erythromycin resistance determinants among Belgian *Streptococcus pyogenes* and *Streptococcus pneumoniae* isolates. J Antimicrob Chemother 2000; 45:167-73; PMID:10660498; http://dx.doi.org/10.1093/jac/45.2.167.
52. Grivea IN, Al-Lahham A, Katopodis GD, Syrogiannopoulos GA, Reinert RR. Resistance to erythromycin and telithromycin in *Streptococcus pyogenes* isolates obtained between 1999 and 2002 from Greek children with tonsillopharyngitis: phenotypic and genotypic analysis. Antimicrob Agents Chemother 2006; 50:256-61; PMID:16377695; http://dx.doi.org/10.1128/AAC.50.1.256-61.2006.
53. Albertí S, García-Rey C, Domínguez MA, Aguilar L, Cercenado E, Gobernado M, et al. Survey of *emm* gene sequences from pharyngeal *Streptococcus pyogenes* isolates collected in Spain and their relationship with erythromycin susceptibility. J Clin Microbiol 2003; 41:2385-90; PMID:12791853; http://dx.doi.org/10.1128/JCM.41.6.2385-90.2003.
54. Perez-Trallero E, Marimo'n JM, Montes M, Orden B, de Pablos M. Clonal differences among erythromycin-resistant *Streptococcus pyogenes* in Spain. Emerg Infect Dis 1999; 5:235-40; PMID:10221875; http://dx.doi.org/10.3201/eid0502.990207.
55. Creti R, Gherardi G, Imperi M, von Hunolstein C, Baldassarri L, Pataracchia M, et al. Association of group A streptococcal *emm* types with virulence traits and macrolide-resistance genes is independent of the source of isolation. J Med Microbiol 2005; 54:913-7; PMID:16157543; http://dx.doi.org/10.1099/jmm.0.46035-0.
56. Palmieri C, Vecchi M, Littauer P, Sundsfjord A, Varaldo PE, Facinelli B. Clonal spread of macrolide- and tetracycline-resistant [erm(A) tet(O)] *emm*77 *Streptococcus pyogenes* isolates in Italy and Norway. Antimicrob Agents Chemother 2006; 50:4229-30; PMID:17043129; http://dx.doi.org/10.1128/AAC.00943-06.
57. Nielsen HU, Hammerum AM, Ekelund K, Bang D, Pallesen LV, Frimodt-Møller N. Tetracycline and macrolide co-resistance in *Streptococcus pyogenes*: co-selection as a reason for increase in macrolide-resistant *S. pyogenes*? Microb Drug Resist 2004; 10:231-8; PMID:15383167.
58. McGregor KF, Bilek N, Bennett A, Kalia A, Beall B, Carapetis JR, et al. Group A streptococci from a remote community have novel multilocus genotypes but share *emm* types and housekeeping alleles with isolates from worldwide sources. J Infect Dis 2004; 189:717-23; PMID:14767827; http://dx.doi.org/10.1086/381452.
59. Pruksakorn S, Sittisombut N, Phornphutkul C, Pruksachatkunakorn C, Good MF, Brandt E. Epidemiological analysis of non-M-typeable group A Streptococcus isolates from a Thai population in northern Thailand. J Clin Microbiol 2000; 38:1250-4; PMID:10699034.
60. Kaplan EL, Johnson DR, Nanthapisud P, Sirilertpanrana S, Chumdermpadetsuk S. A comparison of group A streptococcal serotypes isolated from the upper respiratory tract in the USA and Thailand: implications. Bull World Health Organ 1992; 70:433-7; PMID:1394774.
61. Tewodros W, Kronvall G. M protein gene (*emm* type) analysis of group A beta-hemolytic streptococci from Ethiopia reveals unique patterns. J Clin Microbiol 2005; 43:4369-76; PMID:16145079; http://dx.doi.org/10.1128/JCM.43.9.4369-76.2005.
62. Johnson DR, Stevens DL, Kaplan EL. Epidemiologic analysis of group A streptococcal serotypes associated with severe systemic infections, rheumatic fever or uncomplicated pharyngitis. J Infect Dis 1992; 166:374-82; PMID:1634809; http://dx.doi.org/10.1093/infdis/166.2.374.
63. Tamayo E, Montes M, García-Medina G, García-Arenzana JM, Pérez-Trallero E. Spread of a highly mucoid *Streptococcus pyogenes* emm3/ST15 clone. BMC Infect Dis 2010; 10:233; PMID:20687911; http://dx.doi.org/10.1186/1471-2334-10-233.
64. Stevens DL, Tanner MH, Winship J, Swarts R, Ries KM, Schlievert PM, et al. Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. N Engl J Med 1989; 321:1-7; PMID:2659990; http://dx.doi.org/10.1056/NEJM198907063210101.
65. Cleary PP, Kaplan EL, Handley JP, Wlazlo A, Kim MH, Hauser AR, et al. Clonal basis for resurgence of serious *Streptococcus pyogenes* disease in the 1980s. Lancet 1992; 339:518-21; PMID:1346879; http://dx.doi.org/10.1016/0140-6736(92)90339-5.
66. Gaworzewska ET, Hallas G. Group A streptococcal infections and a toxic shock-like syndrome. N Engl J Med 1989; 321:1545-7; PMID:2682245; http://dx.doi.org/10.1056/NEJM198911303212212.
67. Holden MT, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, et al. Complete genome of acute rheumatic fever-associated serotype M5 *Streptococcus pyogenes* strain Manfredo. J Bacteriol 2007; 189:1473-7; PMID:17012393; http://dx.doi.org/10.1128/JB.01227-06.
68. Beres SB, Sylva GL, Sturdevant DE, Granville CN, Liu M, Ricklefs SM, et al. Genome-wide molecular dissection of serotype M3 group A Streptococcus strains causing two epidemics of invasive infections. Proc Natl Acad Sci USA 2004; 101:11833-8; PMID:15282372; http://dx.doi.org/10.1073/pnas.0404163101.
69. Beres SB, Richter EW, Nagiec MJ, Sumby P, Porcella SF, DeLeo FR, et al. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A Streptococcus. Proc Natl Acad Sci USA 2006; 103:7059-64; PMID:16636287; http://dx.doi.org/10.1073/pnas.0510279103.

70. American Academy of Pediatrics. Committee on Infectious Diseases. Severe invasive group A streptococcal infections: a subject review. Pediatrics 1998; 101:136-40; PMID:11345977; http://dx.doi.org/10.1542/peds.101.1.136.
71. Laupland KB, Davies HD, Low DE, Schwartz B, Green K, McGeer A. Invasive group A streptococcal disease in children and association with varicella-zoster virus infection. Ontario Group A Streptococcal Study Group. Pediatrics 2000; 105:60; PMID:10799624; http://dx.doi.org/10.1542/peds.105.5.e60.
72. Musser JM, Kapur V, Szeto J, Pan X, Swanson DS, Martin DR. Genetic diversity and relationships among *Streptococcus pyogenes* strains expressing serotype M1 protein: recent intercontinental spread of a subclone causing episodes of invasive disease. Infect Immun 1995; 63:994-1003; PMID:7868273.
73. Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog 2006; 2:5; PMID:16446783; http://dx.doi.org/10.1371/journal.ppat.0020005.
74. Barrozo CP, Russell KL, Smith TC, Hawksworth AW, Ryan MAK, Gray GC. National Department of Defense surveillance data for antibiotic resistance and *emm* gene types of clinical group A streptococcal isolates from eight basic training military sites. J Clin Microbiol 2003; 41:4808-11; PMID:14532227; http://dx.doi.org/10.1128/JCM.41.10.4808-11.2003.
75. Enright MC, Spratt BG, Kalia A, Cross JH, Bessen DE. Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between *emm* type and clone. Infect Immun 2001; 69:2416-27; PMID:11254602; http://dx.doi.org/10.1128/IAI.69.4.2416-27.2001.
76. Reid SD, Hoe NP, Smoot LM, Musser JM, Group A. Streptococcus: allelic variation, population genetics and host-pathogen interactions. J Clin Invest 2001; 107:393-9; PMID:11181637; http://dx.doi.org/10.1172/JCI11972.