Evolutionary constraints shaping *Streptococcus pyogenes*–host interactions

Reid V. Wilkening[1,3] and Michael J. Federle[2,3]

[1]Dept. of Microbiology and Immunology; [2]Dept. of Medicinal Chemistry and Pharmacognosy; [3]Center for Biomolecular Sciences
University of Illinois at Chicago, Chicago, IL, USA 60607

Correspondence: mfederle@uic.edu
Phone: 312-413-0213

Keywords: pathogen evolution; bottleneck; mucosa; *in vivo* selection; sterile-site infection; epidemic serotypes

Abstract

Research on the Gram-positive human-restricted pathogen *Streptococcus pyogenes* (Group A *Streptococcus*, GAS) has long focused on invasive illness, the most severe manifestations of GAS infection. Recent advances in descriptions of molecular mechanisms of GAS virulence, coupled with massive sequencing efforts of isolate genomes, have allowed the field to better understand the molecular and evolutionary changes leading to pandemic strains. These findings suggest it necessary to rethink the dogma involving GAS pathogenesis, and that the most productive avenues for research going forward may be investigations into GAS in its 'normal' habitat, the nasopharynx, and its ability to either live with its host in an asymptomatic lifestyle or as an agent of superficial infections. This review will consider these advances, focusing on the natural history of GAS, the evolution of pandemic strains and novel roles for several key virulence factors that may allow the field to better understand their physiological role.

The status of the *Streptococcus pyogenes* field

Streptococcus pyogenes (Group A Streptococcus, GAS) is a Gram-positive, human-restricted bacterium accounting for a diversity of diseases, ranging from 600 million annual cases of uncomplicated pharyngitis to more severe, life-threatening invasive illnesses such as necrotizing fasciitis and pneumonia [1]. GAS is also carried asymptomatically in 5-20% of school-age children and 25% of adults with household contact of infected school-age children [2–4] . Despite representing a minority of cases, invasive illness has long been the focal point of GAS research (recently reviewed in [5]). However, in recent years the field has shifted attention from these extreme cases towards understanding the genetic factors and molecular mechanisms that make GAS such a successful human colonizer and pathogen. Furthermore, massive sequencing efforts have given us evolutionary and molecular explanations for the success of new serotypes of GAS and have focused our attention on the natural history of GAS.

It is the goal of this review to integrate the expanding fields of GAS evolutionary biology with recent advances in understanding the function of virulence factors and virulence factor regulation. We will first define the natural history of GAS by considering the

challenges posed by the primary environment, the nasopharynx. We will next reassess the drivers of invasive GAS illness, before highlighting the genetic changes that have given rise to the M1 and M89 pandemic strains of GAS. Finally, we will consider recent advances in elucidating the function of virulence factors. This review is by no means an exhaustive exploration of these topics, collectively representing many decades of work. Rather, this review seeks to provide a new framework in which to deliberate recent advances in the field, as well as identifying holes in our knowledge. Together, these topics highlight the complex role of GAS virulence factors and regulatory networks, and the field's evolving understanding of how GAS is such a successful human pathogen.

Defining the natural history of GAS

The defining feature of GAS, which sets it apart from many other streptococcal pyogenic species, is its restriction to solely colonize humans. Outside of the human host, there is no known reservoir for GAS[2]. This narrow range of host adaptation is unusual and may have driven speciation, as most pyogenic streptococcal species can infect many mammals [6]. Indeed, the speciation of GAS was in part possible due to acquiring the M-protein island, 35 genes that are found in all GAS genomes including the secreted pyogenic exotoxin SpeB and its regulator RopB. Lefébure *et al.* propose that integration of SpeB and RopB may have played a role in limiting GAS virulence, and that this additional regulatory network may have aided in adapting GAS to the nasopharynx [6].

GAS primarily colonizes the nasopharynx, a challenging environment for bacteria that serves as a bottleneck to infection. First, GAS must pass through a mucus layer to reach and attach to the epithelial cells of the nasopharynx. Confounding observations regarding hyaluronate capsule production indicate benefits and costs in producing the polysaccharide. While the hyaluronic capsule may aid in passing though mucus, the negatively charged capsule likely repels surface interactions in the context of negative charges present on the surface of epithelial cells and could provide a reason why hyper-encapsulated strains are impaired in host colonization [7–9]. GAS hyaluronic acid does contribute to cellular interactions by directly binding cell-surface protein CD44, but it has been hypothesized that hyper-encapsulation may impede surface attachment by

masking other bacterial adhesins [8,10]. Whereas loss of capsule has been implicated in improved carriage, at least in an M3 isolate [11], Lynskey *et al* found the opposite to be true in a hyper-encapsulated M18 serotype, which appears to have an advantage in whole blood survival and murine nasopharyngeal carriage, demonstrating that mucoid M18 isolates contain a truncation of the Regulator of CovR protein *rocA*. This truncation lead to decreased *covR* expression, which in turn limited repression of the capsule biosynthesis *has* operon[12]. These data provide mechanistic support for the observations put forth by Smoot *et al.* wherein they report that hypermucoid M18 isolates were largely responsible for outbreaks of rheumatic fever occurring in the mountain states between 1985-1987 and once again between 1997 and 1999 [13]. Careful regulation of hyaluronic capsule production by GAS is likely to impact success during colonization.

Secondly, during colonization, GAS is challenged in acquiring nutrients. There are few free carbohydrates present in the nasal epithelium, and glucose is actively removed from the airway surface fluid [14,15]. Recent Tn-seq studies testing *S. pneumoniae* colonization demonstrated that more genes are essential for viability in the upper airway than in the lungs, providing experimental evidence that the nasopharynx is a challenging niche, and GAS colonization studies in non-human primates found that genes for mannose and maltodextrin metabolism are upregulated during nasopharyngeal colonization, suggesting that alternative carbon sources are essential [16,17].

Finally, GAS must avoid immune clearance, and the pathogen's genome encodes an abundance of virulence factors used to disrupt host surveillance, to kill or to interfere with innate cellular responses, and to withstand direct assaults on bacterial viability. However, virulence factors are energetically costly to produce, are advantageous only in specific circumstances, and if misregulated could have counter-productive consequences if immune-responses become stimulated [18,19]. Evolutionary biology predicts that complex virulence-regulatory and quorum-sensing pathways, such as those observed in GAS, may play key roles in bypassing bottlenecks such as the

nasopharynx as well as in driving GAS speciation [6,20]. These mechanisms allow rapid and dynamic alterations in gene expression without requiring a fixed mutation in the genome. Using both experimental and mathematical modeling, Lysenko and colleagues demonstrated the value of capsule phase variation in the presence and absence of *Haemophilus influenzae* coinfection on *Streptococcus pneumoniae.* While capsule expression imposes a fitness cost during monoculture, it is necessary for survival during co-culture. Should capsule expression or repression become genetically fixed, *S. pneumoniae* loses the ability to adapt to new environments[21]. Kono *et al* recently extended these findings by experimentally demonstrating that *S. pneumoniae* faces a significant bottleneck passing from the nasopharynx into the bloodstream and in transiting between hosts [22]. Numerous studies have found that hypervirulent strains of GAS, caused largely by mutations in regulatory pathways, are attenuated for transmission to new hosts (discussed below), and similar in-host virulence adaptations have been made in both Gram-positive (e.g. *Staphylococcus aureus)* and -negative bacteria (e.g. *Pseudomonas aeruginosa*) [8,18,23–26]. However, as most animal models of pathogenesis bypass natural routes of infection, they fail to recapitulate the primary modes of GAS colonization, infection and transmission in and from the human nasopharynx. Therefore, the role of the nasopharyngeal environment as a bottleneck to infection and selective force remains largely unassessed.

To understand how GAS successfully colonizes and establishes infections in new hosts, it will be imperative that the field applies animal models that most accurately mimic natural routes of infection. Furthermore, as GAS demonstrates very different disease phenotypes in sterile and non-sterile compartments, it is necessary to critically assess findings in the literature based on models utilized. Perhaps the best animal model, experimental infections in non-human primates, are prohibitively expensive for many researchers and ethically challenging. Murine models of GAS infection are the most commonly utilized, yet most research is conducted with inoculation methods that bypass the integument (i.e. using intramuscular, intravenous or intraperitoneal infections), potentially biasing results. Upper respiratory tract infections in mice can be equally problematic. Alam *et al*. used luciferase-expressing GAS to demonstrate that great care

must be taken in nasal dosing to ensure that GAS remain contained in the nasal-associated lymphoid tissue (NALT), rather than being aspirated into the lungs [27]. The use of humanized mice (such as those expressing human HLA) appear to better model natural routes of infection, however these mice are also cost-prohibitive for routine use [28]. Finally, a murine vaginal colonization model may successfully recapitulate the carriage state of GAS, as Watson *et al.* have found GAS can be asymptomatically carried for extended experiments, allowing assessment of host effector actions on the healthy mucosa [29] There have been numerous other model systems developed for use with GAS. Each model has distinct advantages and disadvantages, and it is essential that the field carefully consider the model used when assessing published findings. For further discussion we point the reader to [30].

Reconsidering invasive illness

With the global emergence of what has now been identified as the pandemic M1 clade, epidemiologists also observed an increase in invasive GAS illness, including necrotizing fasciitis [31,32]. For simplicity, any invasive illnesses will be considered a sterile-site (SS) infection. These severe infections and their causative strains and mechanisms of pathogenesis rapidly became the central focus of GAS research, and SS infections have remained the primary animal model to assess any GAS attribute. In recent years, there has been an increasing push to reconsider whether SS diseases provide a long-term benefit in the evolution of GAS, or are the result of maladaptive, short-term solutions to *in vivo* selections. For an excellent discussion of the evolutionary support for this hypothesis, please refer to [33].

Genetic factors leading to invasive illness

Analysis of isolates recovered from SS infections revealed several dramatic phenotypic differences from strains recovered from superficial infections. SS isolates tend to be hyper-encapsulated and display low or no expression of the secreted cysteine protease SpeB [34,35]. Genetic studies have revealed that these phenotypes are primarily associated with mutations abrogating function of the *covRS* (*csrRS*) two-component system [36–39]. The *covRS* pathway is responsible for regulating between 10 and 15%

of the GAS genome and when intact serves to repress many virulence factors, such as hyaluronic acid capsule synthesis, the DNAse Sda1, and the IL-8 protease SpyCEP [40–44]. Mutations in *covRS* abolish virulence factor governance by diminishing CovR's ability to repress transcription, leading to high levels of gene expression. Conversely, this also leads to reduced expression of SpeB. SpeB degrades both host factors as well as GAS surface-attached and secreted virulence factors like M protein and Sda1, and as SpeB activity is easily assessed experimentally, loss of SpeB production is a common proxy for CovRS integrity[35,45]. The efficacy of the innate immune system, primarily mediated by neutrophils, to target and clear GAS from infected tissues provides a strong selective pressure on the bacteria to counteract neutrophil assault. A mutation occurring in *covRS* presents a means to express virulence factors without restraint providing an immediate benefit and the best likelihood to surmount phagocytosis and neutrophil extracellular trap (NET) entrapment [37,45,46]. There remains some debate as to which CovRS-controlled virulence factor provides the most impact on survival in this situation; however, the answer is likely to be dependent on any given strain's allelic repertoire of virulence and regulatory genes and unique circumstance of the host-tissue environment, and we point the reader to these citations for further reading [37,45–49]

Fitness costs of invasion

Despite the severe consequences of invasive illness to the host, SS infections appear to be an evolutionary dead end for the bacteria, an assertion that is now supported experimentally, epidemiologically and by means of evolutionary studies. A 2006 study of 220 cases of GAS in the Netherlands found no association between *emm* type and invasive disease (refuting the hypothesis that the M1 serotype is particularly invasive). Rather, this study concluded that an uptick in invasive disease, found to coincide with the expansion of pandemic M1 GAS, was likely due to the success of M1 in causing superficial infections, and that invasive illness was statistically in line with disease burden [50,51]. Similar epidemiological findings have been made from studies in the UK with regards to the M89 clade 3 pandemic strain (discussed below) [52]. In 2010, Hollands and colleagues supported these findings experimentally by

demonstrating that the invasive M1T1 isolate MGAS5448 was impaired in binding epithelial cells, keratinocytes, fibronectin or intact mouse epithelium. They proposed that these strains' characteristic of hyper-encapsulation, which is a prominent factor in making these isolates so successful at avoiding phagocytosis, and thus so deadly, once in SS may limit CD44-mediated adherence [8] The work by Hollands *et al.* stands in contrast to the numerous other reports that found M1T1 SS isolates to be hypervirulent *in vivo* (examples include: [35,47,53–55]). However, infection methodologies used in these studies primarily bypassed barriers of the integument by conducting direct bloodstream or subcutaneous skin inoculation, effectively circumventing what may be the most essential bottleneck to GAS colonization, and therefore artificially biasing the fitness value of the phenotype.

Alam *et al*. extended this hypothesis by assessing murine NALT carriage of *emm1, emm2, emm75* and *emm81* isolates. When comparing nasopharyngeal carriage, and spread of a wild-type M75 isolate to an isogenic Δ*covRS* mutant, they found that Δ*covRS* mutant carriage was significantly attenuated and was less successful at colonizing uninfected, cohoused mice. These findings were not due to decreased airborne shed or *in vitro* fitness [56]. In a study of recent pharyngeal and SS *emm1* and *emm12* isolates, Feng *et al.* showed there was no greater propensity of invasive isolates to produce spontaneous *speB*-null derivatives than the pharyngeal isolates. All *speB*-null clones from either isolate group were found to have developed a *covS* mutation leading to loss of *speB* expression [48]. These findings support the hypothesis that *in vivo* selection drives the loss of *covR* regulation and the transition to invasive illness, rather than the hypothesis that unique subsets of GAS strains or serotypes possess inherent invasive phenotypes [48]. Such in-host phenotypic variation has been characterized in other organisms, including *S. aureus,* where a longitudinal study was able to follow a single isolate from carriage to invasive disease [57], and for *S. pneumoniae,* where meningeal infections have been found to stem from single cocci expressing the pilus protein RrgA [58].

Lastly, the ongoing expansion and availability of sequencing has allowed an exploration into the evolutionary drivers of invasive illness. Sequencing the *covRS* locus of 191 GAS isolates from Portugal and other temperate climates demonstrated that

covR, *covS* and *ropB* (the regulator of *speB*) are under stabilizing selection, i.e. that these gene functions are evolutionarily conserved as assessed by decreasing genetic diversity [34]. Similar conclusions have been reached studies conducted in Taiwan and Japan [38,39]. These findings support the assertion that the transition to an invasive phenotype (i.e. *covRS*-null) likely occurs in a unique host, such mutant strains do not persist in the community, and that these mutations are non-adaptive. It is essential to consider both the *covRS* status of strains used in studies, as well as the model systems used, when seeking to investigate the role of factors that might impact fitness in the host. For example, Alam *et al.* found that not only does the strain of bacteria alter carriage duration, the strain of mouse matters as well. In comparing *emm75* carriage in five strains of mice, percent of infected mice ranged from 0 to 100% at 72 hours [56]. This finding should serve as an important reminder of the vast variation seen in using a non-native host for GAS studies, as well as the importance of using natural routes of infection (i.e. nasopharyngeal) when assessing GAS virulence.

The evolution of epidemic GAS

Historically, the predominant GAS serotypes have been observed to wax and wane over time, with trends in serotype prevalence emerging both geographically and temporally [59,60]. Here we will consider and compare the genetic events and molecular alterations that allowed two serotypes, M1 and M89 clade 3, to attain global success. In the mid-1980s, the M1 serotype of GAS emerged and quickly became the dominant, globally circulating strain. Entering the 21st century, the M89 serotype rapidly rose to prominence, establishing itself as one of the top five *emm* types observed globally [52]. The ever-expanding availability and declining cost of whole genome sequencing (WGS) has allowed for novel studies into the evolutionary history of these highly successful strains of GAS and has enabled researchers to parse the events leading to M1 and M89 clade 3 emergence (Figure 1).

M1 GAS

The M1 serotype is the most thoroughly investigated of GAS, as it has long been associated with invasive disease. Currently it is hypothesized that M1 GAS is no more

naturally invasive than other strains, but rather infection is more likely to progress to invasive infection due to the prevalence of this strain in the population, indicative of its success as a human colonizer. Nasser *et al*. expand on the work of Sumby *et al.* by sequencing 3,615 global M1 isolates and by comparing the complete sequences of the pre-epidemic M1 reference strain SF370 to the post-epidemic strain MGAS5005, they concluded that three horizontal gene transfer (HGT) events were essential in the epidemic conversion [43]. First, the lysogenic infections with phages carrying the superantigen *speA* and the DNAse *sdaD2* spread through the M1 population. Initially, two alleles of *speA* were in circulation, *speA1* and *speA2* (Gly110Ser). However, epidemic strains of M1 GAS only carry the *speA2* allele which bind HLA-DQ with greater affinity *in vitro* than the *speA1* allele. However, despite improved binding, SpeA1 and SpeA2 demonstrate equivalent mitogenicity *in vitro*[61,62]. Nasser *et al.* propose that in the mid-1970s, a single nonsynonymous mutation of *speA1* gave rise to *speA2*, citing the fact all (and only) pandemic M1 strains express *speA2.* In the late 1960s, prophage 5005.3, bearing *sdaD2* was first detected in the M1 population. Finally, an HGT event, predicted to have occurred in 1983 with the *emm12* serotype, led to the transfer of a 36kb region that led to upregulation of the virulence factors *slo* and *nga.* The predicted timing of this HGT coincides well with the emergence of M1 as an epidemic strain [63]. These findings stand in contrast to work by Maamary *et al.,* who used genome-sequence data of ten M1 isolates to conclude that the *emm12* translocation occurred prior to the acquisition of the *speA2* allele, which subsequently led to pandemic spread. However, Maamary *et al.* were unable to find a phenotypic role for *speA* in their models, proposing that *in vivo, speA* may provide immunosuppression at the site of infection [32]. Regardless of the order of acquisition of these genetic elements, subsequent analysis of the M1 isolate MGAS2221 *slo* and *nga* promoters by Zhu *et al.* identified several SNPs when compared to the SF370 reference strain. Site-directed mutagenesis of the *slo/nga* promoter and subsequent phenotypic characterization confirmed that a –22 T to G and –18 C to T conversion was necessary and sufficient to confer the observed upregulation of both virulence factors [64].

M89 Clade 3 GAS

M89 clade 3 GAS has recently emerged as one of the top five most frequent global serotypes of GAS [52]. M89 isolates can be sorted into 1 of 3 clades. Clades 1 and 2 are encapsulated yet have failed become pandemic. Intriguingly, the epidemic variant of M89 GAS has been determined to be acapsular, containing a deletion of the biosynthetic *hasABC* genes [52,65]. The hyaluronic acid capsule of GAS has long been proposed as an essential virulence factor, therefore the global emergence of an unencapsulated strain merited further investigation [66–68]. Turner *et al.* performed WGS on 131 M89 isolates collected in the UK between 2004 and 2013. Bioinformatic analysis demonstrated that 83 of the 131 isolates appear to be members of what was later classified by Zhu *et al.* as pandemic clade 3, and comparative genomics revealed six regions containing 88% of all identified SNPs [64]. The authors chose to investigate two regions of polymorphism more closely, as region 2 impacted the *slo/nga* locus, while region 6 encompassed the missing *has* operon. Phenotypic assays revealed that clade 3 isolates had increased Slo and NADase activity, while making no capsule. Further analysis of the *slo/nga* locus found that it contained 99% DNA identity with the modern M1 and M12 locus, leading the authors to suggest that a HGT from either M1 or M12 have led to the global emergence of M89 clade 3 GAS. Whole-blood survival studies revealed no difference in abilities between clade 2 and clade 3 isolates, suggesting that increased *nga/slo* expression by clade 3 may be sufficient to overcome the lack of capsule [52]. These observations have been corroborated by findings from Zhu *et al.,* who performed promoter swaps in the *slo/nga* locus between clade1/2 and clade 3 isolates. As demonstrated with M1 *slo/nga* expression studies, the M89 clade 3 polymorphisms were necessary and sufficient to increase expression of *slo* and *nga* [65]. Finally, Friães *et al*. performed multi-locus sequence typing (MLST) on 886 of the strains sequenced by Zhu *et al.* and Turner *et al*, in addition to 125 M89 isolates from Portugal, with the goal of assessing whether global M89 pandemic isolates had been derived in parallel or from one common lineage [34]. Importantly, they note that pre-epidemic UK isolates belong to ST101, while pre-epidemic isolates from other geographic sites were of ST407 and ST408, both single-locus variants of ST101. Based on this observation, coupled with convergence of superantigen types in the pandemic strain, Friaes *et al*. concluded that the HGT event(s) necessary to yield the pandemic

clade 3 isolates occurred “a limited number of times or even on a single occasion,” yet has rapidly outcompeted other M89 strains to attain global spread [34]. Beres *et al.* draw a divergent conclusion based on their assessment of 1,200 whole-genome sequences, concluding that clade 1 isolates represent an ancestral lineage, from which clade 2 and clade 3 have emerged. Other unique features of serotype M89 merit further investigation [69].

The importance of slo/nga expression in bypassing the nasopharyngeal bottleneck

These studies into the emergence of the pandemic variants of both M1 and M89 GAS provide us with a unique window into the most essential *in vivo* factors that make GAS a successful pathogen. Though it remains unclear whether the *slo/nga* locus in M89 GAS came from pandemic M1 or from the original M12 donor via HGT, it is possible to conclude that increased expression of these virulence factors provides a significant fitness advantage. In the case of the M89 isolates, this advantage is significant enough to compensate for loss of capsule synthesis, an unexpected phenotype. It remains possible that loss of capsule may improve fomite or mucosal adherence, though this hypothesis has not been experimentally demonstrated [52]. It is interesting to note that in surveillance studies from the UK, Portugal and Finland there has been no observed association between the pandemic M89 variant and an uptick in fatalities from invasive illness, despite the increased expression of *slo* and *nga* [34,52,70]. However, Feng *et al.* have recently found that 89% of invasive M89 isolates they assessed were *speB*-null, while all five pharyngitis-associated isolates were *speB*+, suggesting that in some hosts there is still an *in vivo* selective pressure for loss of CovRS regulation [48]. Little is known about the actions of SLO or NADase during colonization or superficial infection (see discussion below). Though much remains unknown, the dual emergence of pandemic M1 and M89 GAS via the same mechanism suggests a primary importance of SLO and NADase in GAS virulence. A better understanding of the bottleneck(s) facing GAS during nasopharyngeal infection may reveal novel therapeutic strategies, such as targeted anti-virulence therapies. It is also worth noting the importance of HGT in the emergence of new pandemic strains, as genetic malleability likely plays a role in the evolution of pandemic isolates. For example, Beres *et al*. found that there is significant

366 genetic diversity amongst *emm89* isolates (mean genetic distance= 610 SNPs) as
367 compared to *emm1* isolates (mean genetic distance = 106 SNPs) [69]. Though natural
368 competence has been extensively characterized in other streptococcal species, the
369 mechanism by which HGT occurs in GAS is unknown, as it has never been observed
370 under laboratory conditions [71].
371
372 **Newly discovered roles for virulence factors**
373 GAS maintains a diverse arsenal of virulence factors that can be applied both
374 offensively and defensively in the human host. Though these factors have long been
375 studied both in molecular and animal models, the field continues to uncover new
376 understandings of the roles for these potent effector molecules that may reflect and
377 explain more accurately their functions *in vivo.* Though GAS is a human specialist, the
378 observed high rates of asymptomatic carriage in the population indicates that GAS is a
379 facultative pathogen rather than an obligate pathogen [72]. Facultative pathogens are
380 faced with different selective pressures, and the role of virulence factors may not be
381 conserved between sterile and non-sterile infection sites. The selective forces on
382 virulence factors can take one of two paths, either by direct selection, where the
383 virulence factor is needed to exploit a host niche or to transmit to a new host, or by
384 coincidental selection, where the benefit of a virulence factor is derived from infection at
385 other sites. By primarily studying invasive illness, the GAS field may have
386 unintentionally focused efforts on coincidental effects of virulence factors. Here we
387 present three recent examples from the literature of virulence factors with novel
388 functions that may better describe their role in the context of direct selective pressures
389 faced by GAS.
390
391 *SpeA:* The streptococcal superantigens (SAgs) are a diverse yet enigmatic family of
392 effectors (at least 14 different SAgs have been identified) that are commonly encoded
393 on mobile genetic elements. SAgs are the effectors of streptococcal toxic shock
394 syndrome (STSS) and scarlet fever through a mechanism of non-specific T cell
395 activation by binding to both MHC-II and T cell receptor β-chains, causing massive T
396 cell activation [73]. Recently, Kasper *et al.* presented data supporting the alternative

hypothesis that the primary function of SAgs is acting in a *locally* immunosuppressive fashion. They demonstrated that C57BL/6 mice expressing HLA-DR4 and HLA-DQ8 carried 100- and 10,000- fold more SpeA+ GAS in their nasal mucosa than wild-type mice, respectively. A derivative of MGAS8232 in which all six SAgs were deleted (ΔSAg) was attenuated in the humanized mouse model. ΔSAg bacteria with cis complementation of *speA* regained the ability to colonize humanized mice, and SpeA toxoid immunization provided mice with protection in the intranasal infection model [28]. As STSS is a rare manifestation of GAS illness and has a high mortality rate, it seems unlikely that STSS provides an evolutionary advantage to GAS, and an *in vivo* role of *speA* as a virulence factor has been elusive, as a Δ*speA* mutant performed equally in several assays when compared to wild type by Maamary *et al.* [32,74]. Yet, SAgs are present in many epidemiologically-determined successful GAS isolates and appear to have been an important step in emergence of the M1 pandemic strain. This work by Kasper *et al.* provides an intriguing alternative role for SAgs that merits further consideration.

Streptokinase (ska): The secreted effector streptokinase converts human plasminogen to fibrinolytic plasmin, allowing GAS to degrade fibrin clots. The accumulation of streptokinase on the surface of *speB*-null GAS has been proposed to be essential in the promotion of invasive illness [37]. It was previously shown that streptokinase could cleave Factor XII to active Factor XIIa [75]. In a recent publication, Wollein Woldetoft *et al*. have put forth a novel role for streptokinase in allowing colonization of the nasopharynx [76]. They first observed that human saliva was capable of activating both the extrinsic and intrinsic clotting cascades in human plasma, an exudate secreted in response to bacterial infection. Δ*ska* GAS mixed with saliva and plasma were quickly and permanently trapped in fibrin clots, while cells capable of secreting streptokinase could escape. They propose that this clotting reaction may be a mechanism to trap GAS on the mucosal surface to promote bacterial clearance, with streptokinase primarily acting to allow GAS to escape this immune mechanism [76]. The efficacy of these proposed mechanisms remains to be evaluated *in vivo*.

Streptolysin O (slo) and S. pyogenes NAD glycohydrolase (nga): As previously mentioned, both Streptolysin O (SLO) and the NADase Nga (also known as SPN) are essential to the success of both M1 and M89 epidemic strains of GAS [52,64]. However, the *in vivo* role of both virulence factors continues to be defined. □-hemolytic SLO is proposed to have several novel functions beyond acting as a pore-forming cholesterol dependent cytolysin. As far back as 1972, Andersen and Van Epps demonstrated that SLO could suppress neutrophil chemotaxis [77]. More recently, Uchiyama *et al.* have expanded these findings, demonstrating that at sublytic concentrations, SLO can limit neutrophil responsiveness by abrogating oxidative burst, NET formation, degranulation, IL-8 release and neutrophil migration. By limiting neutrophil response, the authors propose that SLO may be essential in establishing an infection of the nasopharynx [78]. Logsdon *et al.* found that SLO may also limit internalization of GAS by keratinocytes by blocking clatharin dependent internalization. This may limit GAS access to a proposed intracellular niche [79]. Other groups have focused on the function of SLO in delivering the effector NADase. O'Seaghdha and Wessels have recently published the finding that SLO and NADase co-expression improves intracellular invasion and survival of GAS, putting forth a model where SLO helps promote uptake into autophagosomes. In the absence of NADase activity, lysosomal fusion promotes efficient killing of GAS, while the co-expression of NADase allows prolonged intracellular survival [80]. Similar observations were made that SLO/NADase activity improves GAS survival in macrophages, even allowing cytosolic replication [81,82]. In contrast, Chandrasekaran and Caparon have investigated the divergent intracellular activities of Nga with active and inactive NADase enzymatic domains. Though Nga has been studied primarily in its role as an NAD glycohydrolase, many strains of GAS encode an Nga lacking NADase activity. Chandrasekaran and Capron demonstrate that regardless of Nga NADase activity, translocation of SPN in an SLO-dependent fashion leads to target cell death but via divergent mechanisms [83]. It is possible that the variant activities of Nga prove advantageous in GAS strains with varying tropisms. The continued refinement and elucidation the *in vivo* activities of SLO and Nga will better enable the field to understand the role of these host effectors in the natural lifecycle of GAS, and understand how their expression by pandemic strains provides a selective advantage.

Concluding Remarks

Though it seems unlikely GAS can be considered part of the normal flora, the GAS field has come a long way from intensive focus on the bacteria as a cause of invasive infection. This review highlights the importance of considering the natural history of both GAS and other human-associated pathogens when studying virulence factor function and regulation, while highlighting questions that are primed for investigation (see Outstanding Questions). By understanding the pressures driving adaptive behaviors that allow the emergence of successful isolates, the field will be better equipped to interrogate the functions of GAS's numerous virulence factors. Furthermore, this review highlights the need to develop better models for GAS infection, especially in terms of interaction with the healthy nasopharynx. Current models are both heterogeneous and limited in their ability to fully replicate GAS infection in humans. A unified model would benefit GAS research, as dramatic variation in both model systems and infecting strains used limit the fields' ability to directly compare results. Finally, with a solid understanding of GAS behaviors, we can begin to more efficiently assess the host side of the equation, with the goal of identifying patients at risk for more severe manifestations of GAS illness.

Glossary:

Carriage: For the purposes of this review, we will define carriage as duration of disease burden. This can apply to either asymptomatic colonization or to duration of superficial disease.

Bottleneck: An environmentally-mediated point which dramatically decreases population size.

Nasal Associated Lymphoid Tissue (NALT): An area of organized lymphoid tissue found in the nares of the mouse. Considered to be the region most physiologically akin to the human tonsils.

Acknowledgments

Support for these efforts were provided by The American Heart Association Predoctoral fellowship 15PRE22710027 for RVW, the Burroughs Wellcome Fund Investigators in the Pathogenesis of Infectious Disease for MJF, and NIH grant AI091779.

References

1 Carapetis, J.R. *et al.* (2005) The global burden of group A streptococcal diseases. *Lancet Infect Dis* 5, 685–94

2 Bessen, D. (2009) Population biology of the human restricted pathogen, Streptococcus pyogenes. *Infection, Genetics and Evolution* 9, 581–593

3 DeMuri, G.P. and Wald, E.R. (2014) The Group A Streptococcal Carrier State Reviewed: Still an Enigma. *J Pediatric Infect Dis Soc* 3, 336–42

4 Martin, J. (2016) The Streptococcus pyogenes carrier state. In Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Stevens, D.L., Ferretti, J.J. University of Oklahoma Health Sciences Center

5 Stevens, D.L. and Bryant, A.E. (2016) Severe Group A Streptococcal Infections. In Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Stevens, D.L., Ferretti, J.J. University of Oklahoma Health Sciences Center

6 Lefébure, T. *et al.* (2012) Gene repertoire evolution of Streptococcus pyogenes inferred from phylogenomic analysis with Streptococcus canis and Streptococcus dysgalactiae. *PloS one.* 7, doi:10.1371/journal.pone.0037607

7 Siegel, S.J. and Weiser, J.N. (2015) Mechanisms of Bacterial Colonization of the Respiratory Tract. *Annu. Rev. Microbiol.* 69, 425–444

8 Hollands, A. *et al.* (2010) Genetic Switch to Hypervirulence Reduces

Colonization Phenotypes of the Globally Disseminated Group A Streptococcus M1T1 Clone. *J Inect. Dis.* 202,11-19

9 Nelson, A.L. *et al.* (2007) Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.* 75, 83–90

10 Cywes, C. *et al.* (2000) CD44 as a receptor for colonization of the pharynx by group A Streptococcus. *J. Clin. Invest.* 106, 995–1002

11 Flores, A.R. *et al.* (2014) Asymptomatic carriage of group A streptococcus is associated with elimination of capsule production. *Infect. Immun.* 82, 3958-3967

12 Lynskey, N. *et al.* (2013) RocA Truncation Underpins Hyper-Encapsulation, Carriage Longevity and Transmissibility of Serotype M18 Group A Streptococci. *PLoS Pathogens* 9, DOI:10.1371.ppat.1003842

13 Smoot, J.C. *et al.* (2002) Molecular analysis of group A Streptococcus type emm18 isolates temporally associated with acute rheumatic fever outbreaks in Salt Lake City, Utah. *J. Clin. Microbiol.* 40, 1805–1810

14 Pezzulo, A. *et al.* (2011) Glucose Depletion in the Airway Surface Liquid Is Essential for Sterility of the Airways. *Plos One* 6, DOI:10.1371/journal.pone.0016166

15 Philips, B. *et al.* (2003) Factors determining the appearance of glucose in upper and lower respiratory tract secretions. *Intens Care Med* 29, 2204–2210

16 Van Opijnen, T. and Camilli, A. (2012) A fine scale phenotype–genotype virulence map of a bacterial pathogen. *Genome Res* 22, 2541–2551

17 Virtaneva, K. *et al.* (2005) Longitudinal analysis of the group A Streptococcus transcriptome in experimental pharyngitis in cynomolgus macaques. *P Natl Acad Sci Usa* 102, 9014–9019

18 Laabei, M. *et al.* (2015) Evolutionary Trade-Offs Underlie the Multi-faceted Virulence of Staphylococcus aureus. *PLOS Biology* 13, DOI:10.1371/journal.pbio.1002229

19 Andreoni *et al.* (2014) The IL-8 protease SpyCEP is detrimental for Group A Streptococcus host-cells interaction and biofilm formation. *Front Microbiol* DOI: 10.3389/fmicb.2014.00339

20 Abel, S. *et al.* (2015) Analysis of Bottlenecks in Experimental Models of Infection.

PLoS Pathog. 11, DOI:10.1371/journal.ppat.1004823

21 Lysenko, E. *et al.* (2009) Within-Host Competition Drives Selection for the Capsule Virulence Determinant of Streptococcus pneumoniae. *Curr Biol* 20, 1222–1226

22 Kono, M. *et al.* (2016) Single Cell Bottlenecks in the Pathogenesis of Streptococcus pneumoniae. *Plos Pathog* 12, DOI:10.1371/journal.ppat.1005887

23 Das, S. *et al.* (2016) Natural mutations in a Staphylococcus aureus virulence regulator attenuate cytotoxicity but permit bacteremia and abscess formation. *P Natl Acad Sci USA.* 113, E3101–E3110

24 Bragonzi, A. *et al.* (2009) Pseudomonas aeruginosa Microevolution during Cystic Fibrosis Lung Infection Establishes Clones with Adapted Virulence. *Am J Resp Crit Care* 180, 138–145

25 Lorè, N. *et al.* (2012) Cystic Fibrosis-Niche Adaptation of Pseudomonas aeruginosa Reduces Virulence in Multiple Infection Hosts. *Plos One* 7, DOI:10.1371/journal.pon.0035648

26 Gellatly, S.L. and Hancock, R.E. (2013) Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. *Pathog Dis* 67, 159–173

27 Alam, F.M. *et al.* (2013) Non-Invasive Monitoring of Streptococcus pyogenes Vaccine Efficacy Using Biophotonic Imaging. *PLoS ONE* 8, DOI:10.1371/journal.pone.0082123

28 Kasper, K. *et al.* (2014) Bacterial Superantigens Promote Acute Nasopharyngeal Infection by Streptococcus pyogenes in a Human MHC Class II-Dependent Manner. *Plos Pathog* 10, DOI:10.1371/journal.ppat.1004155

29 Watson, M.E. *et al.* (2013) Murine vaginal colonization model for investigating asymptomatic mucosal carriage of Streptococcus pyogenes. *Infect. Immun.* 81, 1606–1617

30 Watson, M.E. *et al.* Animal Models of Streptococcus pyogenes Infection. In Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Stevens, D.L., Ferretti, J.J. University of Oklahoma Health Sciences Center

31 Aziz, R. and Kotb, M. (2008) Rise and Persistence of Global M1T1 Clone of Streptococcus pyogenes. *Emerg Infect Dis* 14, 1511–1517

32 Maamary, P. *et al.* (2012) Tracing the evolutionary history of the pandemic group A streptococcal M1T1 clone. *Faseb J* 26, 4675–4684

33 Wollein Waldetoft, K. and Råberg, L. (2014) To harm or not to harm? On the evolution and expression of virulence in group A streptococci. *Trends Microbiol.* 22, 7–13

34 Friães, A. *et al.* (2015) Consequences of the variability of the CovRS and RopB regulators among Streptococcus pyogenes causing human infections. *Sci Rep* 5,DOI:10.1038/srep12057

35 Kansal *et al.* (2000) Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. *Infect. Immun.* 68, 6362–6369

36 Treviño, J. *et al.* (2009) CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A Streptococcus virulence factor-encoding genes. *Infect. Immun.* 77, 3141–3149

37 Cole, J.N. *et al.* (2011) Molecular insight into invasive group A streptococcal disease. *Nat. Rev. Microbiol.* 9, 724–736

38 Ikebe, T. *et al.* (2010) Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. *PLoS Path.* 6, DOI:10.1371/journal.ppat.1000832

39 Lin, J.-N.N. *et al.* (2014) Association between polymorphisms in the csrRS two-component regulatory system and invasive group A streptococcal infection. *Eur. J. Clin. Microbiol. Infect. Dis.* 33, 735–743

40 Federle, M.J. *et al.* (1999) A response regulator that represses transcription of several virulence operons in the group A streptococcus. *J. Bacteriol.* 181, 3649–3657

41 Engleberg, C.N. *et al.* (2001) Spontaneous mutations in the CsrRS two-component regulatory system of Streptococcus pyogenes result in enhanced virulence in a murine model of skin and soft tissue infection. *Journal of Infectious Diseases* 183, 1043–1054

42 Graham, M.R. *et al.* (2002) Virulence control in group A Streptococcus by a two-

component gene regulatory system: global expression profiling and in vivo infection modeling. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13855–13860

43 Sumby, P. *et al.* (2006) Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Path.* 2,DOI:10.1371/journal.ppat.0020005

44 Churchward, G. (2007) The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. *Mol. Microbiol.* 64, 34–41

45 Walker, M. *et al.* (2007) DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. *Nat Med* 13, 981–985

46 Li, Y. *et al.* (2013) Within host selection is limited by effective population of Streptococcus pneumoniae during nasopharyngeal colonization. *Infect. Immun.* 82, 1579-1590

47 Li, J. *et al.* (2014) Neutrophils Select Hypervirulent CovRS Mutants of M1T1 Group A Streptococcus During Subcutaneous Infection of Mice. *Infect. Immun.* 82, 4534-4543

48 Feng, W. *et al.* (2016) Contemporary Pharyngeal and Invasive emm1 and Invasive emm12 Group A Streptococcus Isolates Exhibit Similar In Vivo Selection for CovRS Mutants in Mice. *Plos One* 11, DOI:10.1371/journal.pone.0162742

49 Liu, G. *et al.* (2015) The Mga Regulon but Not Deoxyribonuclease Sda1 of Invasive M1T1 Group A Streptococcus Contributes to In Vivo Selection of CovRS Mutations and Resistance to Innate Immune Killing Mechanisms. *Infect Immun* 83, 4293–4303

50 Rogers, S. *et al.* (2007) Strain Prevalence, Rather than Innate Virulence Potential, Is the Major Factor Responsible for an Increase in Serious Group A Streptococcus Infections. *Journal of Infectious Diseases* 195, 1625–1633

51 McMillan, D. *et al.* (2006) Genes for the Majority of Group A Streptococcal Virulence Factors and Extracellular Surface Proteins Do Not Confer an Increased Propensity to Cause Invasive Disease. *Clin Infect Dis* 43, 884–891

52 Turner, C.E. *et al.* (2015) Emergence of a New Highly Successful Acapsular Group A Streptococcus Clade of Genotype emm89 in the United Kingdom. *MBio* 6, DOI:10.1128/mBio.00622-15

53 Olsen, R.J. *et al.* (2015) The majority of 9,729 group A streptococcus strains causing disease secrete SpeB cysteine protease: pathogenesis implications. *Infect. Immun.* 83, 4750–4758

54 Bao, Y.-J.J. *et al.* (2015) CovRS Regulated Transcriptome Analysis of a Hypervirulent M23 Strain of Group A Streptococcus pyogenes Provides New Insights on Virulence Determinants. *J. Bacteriol.* 197, 3191-3205

55 Mayfield, J.A. *et al.* (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, DOI:10.1371/journal/pone.100698

56 Alam, F.M. *et al.* (2013) Inactivation of the CovR/S virulence regulator impairs infection in an improved murine model of Streptococcus pyogenes naso-pharyngeal infection. *PLoS ONE* 8, DOI:10.1371/journal.pone.0061655

57 Young, B. *et al.* (2012) Evolutionary dynamics of Staphylococcus aureus during progression from carriage to disease. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4550–4555

58 Iovino, F. *et al.* (2016) Pneumococcal meningitis is promoted by single cocci expressing pilus adhesin RrgA. *J. Clin. Invest.* 126, 2821–2826

59 Gaworzewska and Colman (1988) Changes in the pattern of infection caused by Streptococcus pyogenes. *Epidemiol Infect* 100, 257–269

60 Steer, A. *et al.* (2009) Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect Dis* 9, 611–616

61 Nelson, K. *et al.* (1991) Characterization and clonal distribution of four alleles of the speA gene encoding pyrogenic exotoxin A (scarlet fever toxin) in Streptococcus pyogenes. *J. Exp. Med.* 174, 1271–1274

62 Kline, J.B. and Collins, C.M. (1996) Analysis of the superantigenic activity of mutant and allelic forms of streptococcal pyrogenic exotoxin A. *Infect. Immun.* 64, 861–869

63 Nasser *et al.* (2014) Evolutionary pathway to increased virulence and epidemic group A Streptococcus disease derived from 3,615 genome sequences. *Proc.*

Natl. Acad. Sci. U.S.A. 111, E1768-E1776

64 Zhu, L. *et al.* (2015) A molecular trigger for intercontinental epidemics of group A Streptococcus. *J. Clin. Invest.* DOI: 10.1172/JCI82478

65 Zhu, L. *et al.* (2015) Trading Capsule for Increased Cytotoxin Production: Contribution to Virulence of a Newly Emerged Clade of emm89 Streptococcus pyogenes. *Mbio.* 6, DOI:10.1128/mBio.01378-15

66 Cole, J.N. *et al.* (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, DOI:10.1128/mBio.00191-10

67 Schrager *et al.* (1998) Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A Streptococcus to CD44 on human keratinocytes. *J. Clin. Invest.* 101, 1708-1716

68 Heath, A. *et al.* (1999) A two-component regulatory system, CsrR-CsrS, represses expression of three Streptococcus pyogenes virulence factors, hyaluronic acid capsule, streptolysin S, and pyrogenic exotoxin B. *Infect. Immun.* 67, 5298–5305

69 Beres, S. *et al.* (2016) Transcriptome Remodeling Contributes to Epidemic Disease Caused by the Human Pathogen Streptococcus pyogenes. *mBio* 7, DOI:10.1128/mBio.00403-16.

70 Latronico, F. *et al.* (2016) Genomic tracks behind spread of bacteremic Group A Streptococcus type emm89 in Finland, 2004-2014. *J. Infect. Dis.* 214, 1987-1995

71 Mashburn-Warren, L and Morrison, DA (2012) The cryptic competence pathway in Streptococcus pyogenes is controlled by a peptide pheromone. *J. Bacteriol.* 194. 4589-4600

72 Brown, S.P. *et al.* Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends Microbiol.* 20, 336-342

73 Commons, R. *et al.* (2014) Streptococcal superantigens: categorization and clinical associations. *Trends Mol. Med.* 20, 48–62

74 Proft, T. and Fraser, J.D. (2016) Streptococcal Superantigens: Biological properties and potential role in disease. In Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Stevens, D.L., Ferretti, J.J. University

of Oklahoma Health Sciences Center

75 Nitzsche, R *et al.* (2015) Streptococcus pyogenes triggers activation of the human contact system by streptokinase. *Infect. Immun.* 83, 3035-3042

76 Waldetoft, K. *et al.* (2016) Saliva-Induced Clotting Captures Streptococci: Novel Roles for Coagulation and Fibrinolysis in Host Defense and Immune Evasion. *Infect Immun* 84, 2813–2823

77 Andersen, B.R. and Epps, D.E. Van (1972) Suppression of chemotatic activity of human neutrophils by streptolysin O. *J. Infect. Dis.* 125, 353–359

78 Uchiyama, S. *et al.* (2015) Streptolysin o rapidly impairs neutrophil oxidative burst and antibacterial responses to group a Streptococcus. *Front. Immunol.* 6, DOI:10.3389/fim-mu.2015.00581

79 Logsdon, L. *et al.* (2011) Streptolysin O inhibits clathrin-dependent internalization of group A Streptococcus. *mBio* 2, DOI:10.1128/mBio.00332-10

80 O'Seaghdha, M. and Wessels, M. (2013) Streptolysin O and its Co-Toxin NAD-glycohydrolase Protect Group A Streptococcus from Xenophagic Killing. *PLoS Path* 9, DOI:10.1371/journal.ppat.1003394

81 Bastiat-Sempe, B. *et al.* (2014) Streptolysin O and NAD-Glycohydrolase Prevent Phagolysosome Acidification and Promote Group A Streptococcus Survival in Macrophages. *MBio* 5, DOI:10.1128/mBio.01690-14

82 O'Neill, A.M. *et al.* (2016) Cytosolic Replication of Group A Streptococcus in Human Macrophages. *MBio* 7, DOI:10.1128/mBio.00020-16

83 Chandrasekaran, S. and Caparon, M.G. (2015) The Streptococcus pyogenes NAD(+) Glycohydrolase Modulates Epithelial Cell PARylation and HMGB1 Release. *Cell. Microbiol.* 17, 1376-1390

Figure 1. Recent emergence of two pandemic GAS serotypes. Acquisition of *sdaD2* by phage lysogenization, and subsequently acquisition of *speA1*, which by single nucleotide polymorphism evolved to *speA2*. This was followed by a horizontal genetic transfer event(s) leading to enhanced expression of Nga and SLO (marked **nga*/*slo*),and led to the modern M1 epidemic serotype. Similarly, for serotype M89, three clades have emerged from a common ancestor. Clade 3 has six SNP-containing regions that differ from clades 1 and 2. The two regions that have been best studied encompass the *has* capsule biosynthesis genes and the *nga*/*slo* locus.