A decade of molecular pathogenomic analysis of group A *Streptococcus*

James M. Musser[1] and Samuel A. Shelburne III[2]

[1]Center for Molecular and Translational Human Infectious Diseases Research, The Methodist Hospital Research Institute, and Department of Pathology, The Methodist Hospital, Houston, Texas, USA. [2]Department of Infectious Diseases, Division of Internal Medicine, MD Anderson Cancer Center, Houston, Texas, USA.

Molecular pathogenomic analysis of the human bacterial pathogen group A *Streptococcus* has been conducted for a decade. Much has been learned as a consequence of the confluence of low-cost DNA sequencing, microarray technology, high-throughput proteomics, and enhanced bioinformatics. These technical advances, coupled with the availability of unique bacterial strain collections, have facilitated a systems biology investigative strategy designed to enhance and accelerate our understanding of disease processes. Here, we provide examples of the progress made by exploiting an integrated genome-wide research platform to gain new insight into molecular pathogenesis. The studies have provided many new avenues for basic and translational research.

Introduction

Group A *Streptococcus* (GAS) is a Gram-positive bacterium that causes several diseases in humans, including pharyngitis and/or tonsillitis, skin infections (including impetigo, erysipelas, and other forms of pyoderma), acute rheumatic fever (ARF), scarlet fever, poststreptococcal glomerulonephritis (PSGN), a toxic shock–like syndrome, and necrotizing fasciitis (NF) (1–3). The organism is a human-adapted pathogen, that is, humans are the natural host and there is no animal or environmental reservoir that contributes to the life cycle of GAS. On a global basis, ARF and the rheumatic heart disease that follows are the most common cause of preventable pediatric heart disease worldwide (4). There are 25–35 million cases of GAS pharyngitis per year in the United States, which are responsible for about 1–2 billion dollars per year in direct health care costs, and 600 million cases globally each year (Table 1) (2, 3). The continued high morbidity and mortality caused by GAS in developing nations, the substantial financial burden attributable to GAS-related health care costs in the United States, the development and spread of antibiotic resistance in this pathogen (5–7), and the lack of a licensed human vaccine (8, 9) highlight the need for a fuller understanding of the molecular pathogenesis of GAS. Especially vital is the need for information bearing on the molecular events contributing to pathogen-host interaction, clone emergence, and epidemics.

In recent years, a paradigm shift has occurred in the manner in which microbial pathogenesis problems are studied. The confluence of data derived from genome-sequencing projects and the development of microarray technology permit a genome-wide strategy to be used for pathogenesis research, now commonly referred to as molecular pathogenomics investigation (10). Comparative genome sequencing and other types of analyses of multiple isolates of the same pathogenic bacterial species have revealed unexpectedly large amounts of intraspecies variation in gene content and uncovered strategies used in genome diversification (11–19). Strains of some bacterial species differ in gene content by up to 25% and have extensive allelic variation (11, 13, 15). This striking magnitude of genetic differentiation provides tremendous adaptive flexibility and influences the spectrum of clinical disease caused by distinct clones of the species. For example, the genome of *E. coli* serotype O157:H7, which is responsible for hemolytic uremic syndrome, is approximately 25% larger (i.e., approximately 1 Mb of DNA larger) than that of many other strains of *E. coli* that do not cause significant disease in humans (13).

A molecular pathogenomics approach has been applied to GAS for nearly a decade and has yielded new information about the genetic basis of GAS pathogenesis, clone emergence, and strain genotype–disease phenotype relationships. Herein we highlight key findings that illustrate that molecular pathogenomics can be used to greatly accelerate the rate of obtaining new information about long-standing and previously intractable infectious disease questions. Only by applying hypothesis-driven research and using new technologies will the elucidation of molecular events underlying infectious disease processes proceed with maximum efficiency.

GAS genome sequences

General concepts. The full-genome sequences that are now available for 13 strains of GAS make it one of the more deeply sequenced species of human pathogens (15, 20–28). GAS strains can be classified by either M protein or T protein serotype, depending on the composition of bacterial antigens expressed on their cell surface, and the 13 genome sequences represent serotypes responsible for more than 70% of M protein serotypes that commonly cause GAS pharyngitis and invasive infections in the Western Hemisphere (29–31). The strains selected for sequencing were chosen mainly because they represent abundantly occurring M protein serotypes or because they had a distinct clinical phenotype of special interest, such as high virulence, antimicrobial agent resistance, or association with a particular disease (15, 20–28). Several common themes have emerged from comparison of the genome sequences (15). Each genome is approximately 1.9 Mb in size, with approximately 10% of the overall gene content encoded on variably present exogenous genetic elements such as prophages and integrated conjugative elements (15), with the former accounting for most of the variably present gene content (15, 20–28). Importantly, the marked heterogeneity of gene content is even observed among GAS strains of the same M protein serotype. This extensive variation in gene content stands in stark contrast to humans and many other higher eukaryotic species, which have less than 0.1% variation in gene content among different members of a species. The 13 GAS genome

Table 1
Minimum summary estimates of the global burden of GAS diseases

	Existing cases	Annual new cases	Annual deaths
Severe GAS disease			
RHD	15.6 million	282,000	233,000
History of ARF without carditis, requiring secondary prophylaxis	1.88 million	188,000	—
RHD-related infective endocarditis	—	34,000	8,000
RHD-related stroke	642,000	144,000	108,000
Acute poststreptococcal glomerulonephritis	—	472,000	5,000
Invasive GAS diseases	—	663,000	163,000
Total severe cases	18.1 million	1.78 million	517,000
Superficial GAS disease			
Pyoderma	111 million	—	—
Pharyngitis	—	616 million	—

Table 1 is reproduced with permission from *The Lancet Infectious Diseases* (3). RHD, rheumatic heart disease. —, data not provided in the original reference.

sequences provide a critical resource for initiating investigation into how variation in gene content influences pathogenesis.

Discovery of new virulence factors through whole-genome sequence analysis. As recently as a decade ago, only relatively few GAS virulence factors had been identified. These included M protein, hyaluronic acid capsule, and the extracellular cysteine protease streptococcal pyrogenic exotoxin B (SpeB). Within a few years of the first GAS genome sequence becoming available, no fewer than 13 new proteins that contribute to pathogenesis had been described (32–45). Novel pathways of GAS host-pathogen interaction also had been elucidated. For example, Edwards et al. demonstrated that supernatants derived from GAS cleaved IL-8, resulting in reduced neutrophil activation and migration (37). Using the GAS genome data, the IL-8–degrading activity was shown to be encoded by a previously uninvestigated gene (*spy0416*, using the serotype M1 numbering system for strain SF370), and the proteinase encoded by this gene was subsequently termed "*Streptococcus pyogenes* cell envelope proteinase" (SpyCEP) (46, 47). In a short time, it has been learned that SpyCEP-mediated cleavage of IL-8 decreases neutrophil endothelial transmigration and that degradation of additional chemokines by SpyCEP retards neutrophil activation (46–48). Together, these discoveries indicate a key role for SpyCEP in GAS host-pathogen interaction (46–50). Similar rapid progress, facilitated by the availability of the GAS genome sequences, has been made in identifying and characterizing the function of novel GAS immunoglobulin-degrading enzymes (39, 40, 51), collagen-like proteins (41, 44, 52), and several superantigens (53–55). Given that numerous open reading frames in the GAS genome that encode putative cell surface proteins have yet to be investigated, it seems highly likely to us that heretofore unidentified GAS virulence factors will continue to be discovered in the coming years.

## Expanded understanding of virulence factor regulation pathways

As noted above, the GAS genome encodes a broad range of virulence factors that are critical to the diverse array of infections that the bacterium causes (1). One important impact of GAS genome-wide investigations has been to provide an enhanced molecular understanding of how the pathogen coordinates virulence factor production (56). There are 13 conserved two-component regulatory systems (TCSs) in the completed GAS genomes, only one of which had been studied prior to the availability of the first GAS genome sequence in 2001 (57). TCSs regulate multiple unlinked chromosomal genes and control coordinated expression of genes encoding virulence factors, such as toxins, degradative enzymes, and immune-modulating molecules, in response to environmental stimuli. Since 2001, each of the 13 TCSs has been studied by genome-wide analyses to some degree, many in detail (58–62). Similarly, there are more than 100 putative stand-alone transcriptional regulators encoded within the GAS genome, only a few of which had been identified before the availability of a genome sequence. Of the 100 putative stand-alone transcriptional regulators, 12 have now been investigated in detail, and the transcriptomes that eight of these regulate in vitro have been determined (17, 63–68).

Genome-wide investigations of GAS regulatory pathways have transformed our understanding of pathogenesis. A key finding has been that global gene transcription varies highly depending on the environment, growth conditions, and stage of growth of the bacterium (69–72). For example, Voyich et al. (73) have shown that the GAS transcriptome is substantially altered in response to phagocytosis by human polymorphonuclear leukocytes (PMNs), a key step in invasive GAS disease. Similarly, interaction with pharyngeal epithelial cells and human saliva was found to induce marked alterations in the GAS transcriptome, thereby providing new information about potential strategies used for infection of and persistence in the oropharynx (69, 74). Longitudinal analysis of changes in the GAS transcriptome over time in nonhuman primates has revealed that the temporal pattern of GAS gene transcription in pharyngitis is very closely linked to three distinct phases of infection, namely colonization, acute infection, and asymptomatic carriage (75). During the colonization phase, when GAS CFUs were low, the expression of genes involved in carbohydrate metabolism was greatly increased, suggesting that carbon source acquisition is a key step in initial GAS growth and establishment of infection. Expression of genes encoding GAS virulence factors with known roles in GAS survival, dissemination, inhibition of PMN recruitment, and induction of host cytokines were highly expressed in the acute phase of infection, concomitant with an increase in GAS CFUs and host inflammation. Thus, the in vivo infection data are in strong alignment with ex vivo results (i.e., those obtained using saliva, epithelial cells, and PMNs).

A key finding of repeated GAS transcriptome analyses has been the elucidation of previously unappreciated connections between distinct gene categories. For example, recent work has shown that transcript levels of genes encoding proteins involved in carbohydrate catabolism and genes encoding virulence factors change in concert in response to environmental stimuli (74, 75). Such changes are at least partially due to a regulatory circuit controlled by catabolite control protein A (CcpA) (Figure 1) (59, 74–78). Simi-
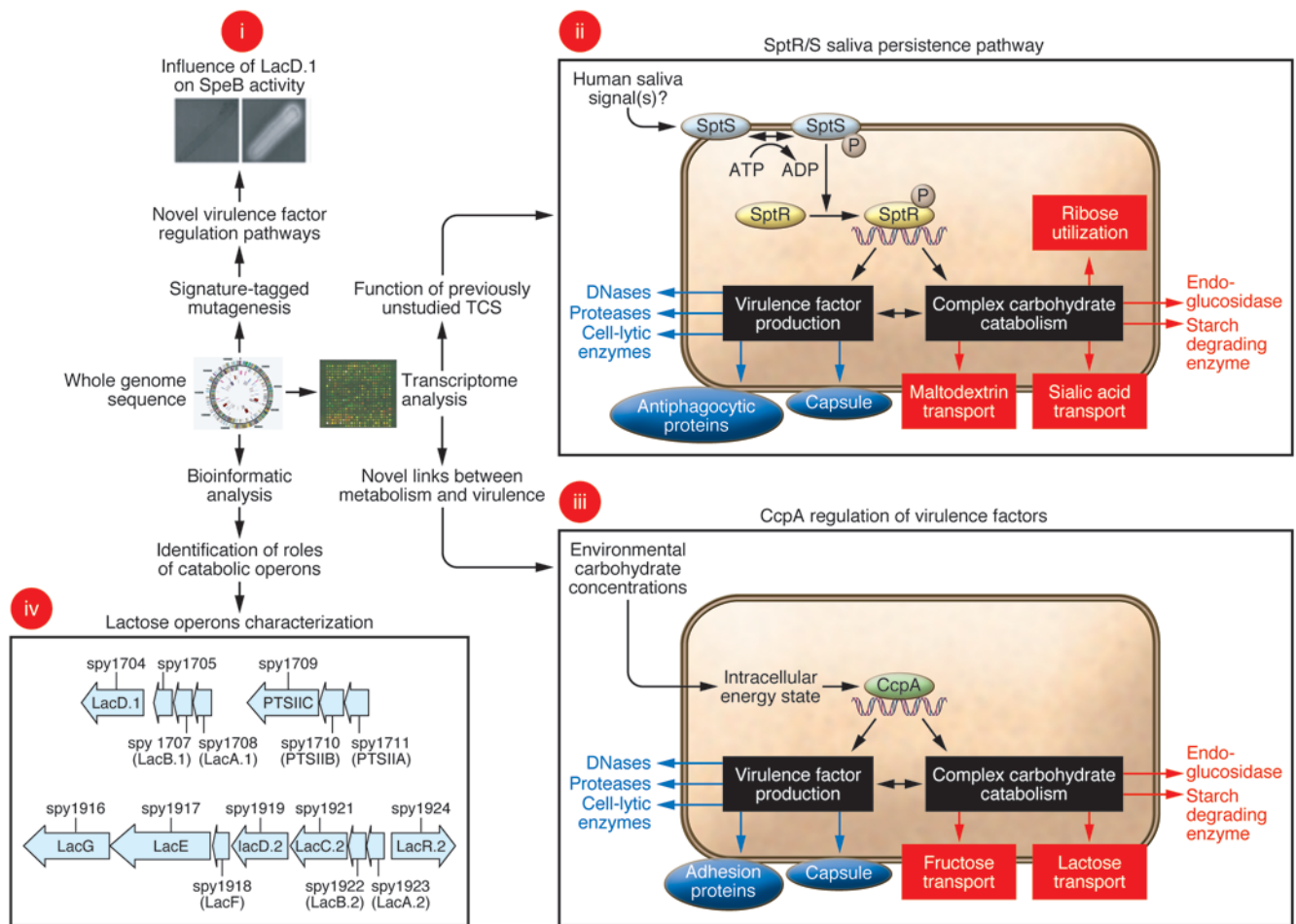
Figure 1
Schematic depicting how genome-wide analyses have revealed previously unsuspected links between carbohydrate metabolism and virulence. Examples of recently discovered relationships between carbohydrate catabolism and virulence include (i) the role of LacD.1 in controlling SpeB production, (ii) identification of the SptR/S saliva persistence pathway, (iii) the discovery of regulation of virulence factors by CcpA, and (iv) the differential role in virulence factor regulation (bottom) and catabolism (top) of the 2 lactose operons. Panel i is reprinted with permission from *EMBO Journal* (79). Panels ii and iii are adapted from *Proceedings of the National Academy of Sciences U. S. A.* (ref. 75 and ref. 77, respectively).

larly, a genome-wide investigation of SpeB regulation led to the discovery that a lactose catabolism enzyme (LacD.1) has evolved to coordinate alterations in GAS virulence factor production as a result of changes in carbon source availability (Figure 1) (79). Analysis of completed GAS genomes led to the discovery that one of the two GAS lactose operons has retained a catabolic role, whereas the other has evolved a regulatory function (Figure 1) (80).

In addition to discovering ties between central metabolic processes and pathogenesis, GAS genome-wide studies have also revealed interactions between metal regulation, oxidative stress, and pathogenesis. Analysis of the completed genomes indicates that GAS encodes two highly conserved metalloregulators, MtsR and PerR, that regulate proteins involved in iron and manganese uptake (81, 82). Animal studies have found that PerR is needed for full GAS virulence in skin and soft tissue infection and in oropharynx infection (65, 82, 83). Surprisingly, the PerR regulon was found to include numerous carbohydrate utilization genes, suggesting links among central metabolic processes, oxidative stress response, and virulence (65). The key role for MtsR in the develop-

ment of NF is discussed below in "Genome-wide dissection of the molecular events underlying GAS epidemics."

Insights into the molecular basis of GAS disease specificity
Investigations into the molecular basis of GAS host site predilection. The two major sites of GAS infection are the human throat and skin (84). It has long been recognized that particular M protein serotypes mainly cause pharyngitis, whereas others predominate in skin infection, leading to the idea of skin-specialist and throat-specialist GAS strains (85, 86). However, the molecular basis for these observations has been unclear (86). Bioinformatic study of GAS genomes led to the identification of an area of the genome that is highly variable between M protein serotypes, referred to as the fibronectin-binding, collagen-binding T antigen (FCT) region (85). Lately, evidence has accumulated that genetic heterogeneity within the FCT region may be a major factor determining why particular GAS strains colonize and infect distinct host regions (87). For example, the FCT region contains genes encoding the recently
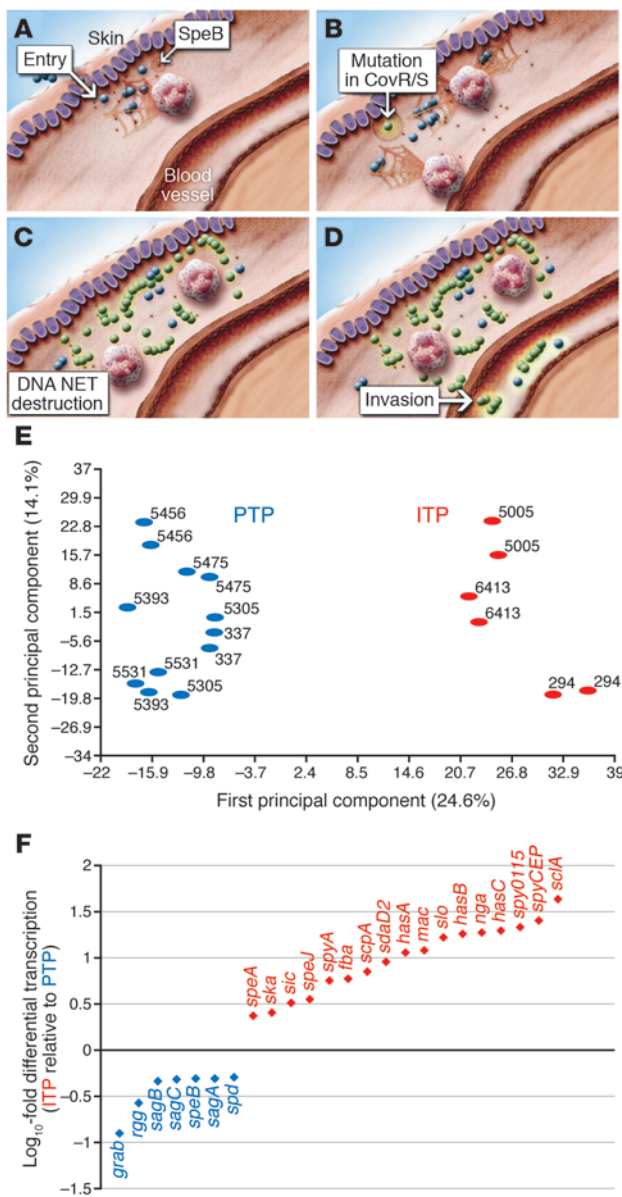
Figure 2
The role of the CovR/S TCS in the transition from mucosal to invasive disease. (**A**) Upon entering subcutaneous tissue, GAS is confronted by neutrophils that produce NETs. (**B**–**D**) Mutation in the CovR/S TCS (**B**) allows for high-level production of a GAS DNase that degrades NETs (**C**), thereby increasing pathogen survival and enhancing the probability of invasive infection (**D**). (**E**) Principal component analysis of expression microarray data demonstrates that mucosal (or pharyngeal) and invasive GAS isolates have strikingly different transcriptomes. (**F**) Examples of genes encoding critical virulence factors that are more highly expressed in invasive isolates (red) or pharyngeal isolates (blue). ITP, invasive transcriptome profile; PTP, pharyngeal transcriptome profile. **A**–**D** are reprinted with permission from *Nature Medicine* (93). **E** and **F** are adapted from *PLoS Pathogens* (92).

complete genomes of GAS isolates recovered from the spleen of mice that had been infected subcutaneously. Compared with the strain used to inoculate the mouse skin, the invasive GAS isolates (i.e., those obtained from the spleen after skin inoculation) had mutated forms of the control of virulence (CovR/S) TCS (91). Mutation of this TCS resulted in derepression of numerous virulence factors critical for combating the host immune system components encountered during bloodstream infection (Figure 2) (91). For example, GAS secretes a potent DNase that is involved in escape from neutrophil extracellular traps (NETs), a host immune defense mechanism generated by dying PMNs (92). The DNase is upregulated in CovR/S mutants, thereby contributing to the development of invasive disease. The clinical relevance of these findings was confirmed by the discovery that many GAS strains causing invasive infections in humans often have function-altering mutations in the genes that encode the components of the CovR/S TCS (91, 93). Thus, it is currently thought that GAS mucosal isolates have an intact CovR/S system that limits GAS virulence factor production, whereas interaction with the host immune system and/or deep tissues selects for strains with CovR/S mutations, leading to a hypervirulent phenotype and the serious manifestations of invasive GAS disease (Figure 2).

Establishment of the key role of lateral gene transfer in the emergence of unusually virulent clones
Several types of molecular events contribute to the evolution and emergence of bacterial strains with enhanced virulence. The most well-understood and by far the most studied process is horizontal gene transfer (HGT), which involves bacterial viruses known as bacteriophages (transduction), plasmids (conjugation), and genomic DNA (transformation) (16, 94). HGT events create new strain genotypes by moving blocs of genetic material — sometimes large pieces of DNA that exceed 40–60 kb in size — between strains. Thus, HGT events represent a quantum evolutionary leap that can increase bacterial fitness by enhancing antimicrobial agent resistance, immune avoidance, and capacity to colonize or infect a new ecological niche. As in eukaryotes, bacterial evolution also occurs by more subtle processes, including accumulation of point mutations and small genomic changes such as those created by slipped-strand mispairing. Several of these molecular processes have contributed to the recent emergence and intercontinental dissemination of a new clone of serotype M1 GAS with distinct virulence properties (22, 95–97). This understanding was revealed by several lines of work, including comparative genome characterization conducted in several laboratories in a span of almost

discovered cell surface pili that are critical to GAS epithelial cell adhesion (32, 33, 36). Thus, strain-to-strain differences in pili composition may contribute to the predisposition of particular GAS M protein serotypes for certain host sites in an analogous fashion to that observed for *E. coli* (88). Elucidation of the molecular basis for why particular GAS strains colonize and infect particular host environments holds the promise for developing novel preventive and therapeutic targets.

New insights into relationships between mucosal and invasive GAS disease. GAS thrives at human mucosal sites and also causes devastating invasive infections. Strains isolated from mucosal sites are genetically indistinguishable from invasive strains by standard assays, such as M protein serotyping and multi-locus sequence typing (89, 90). However, these techniques index only a very small part of the genome, which means that they greatly underestimate the amount of genetic variation present. Using newly developed genome resequencing techniques, Sumby et al. (91) analyzed the
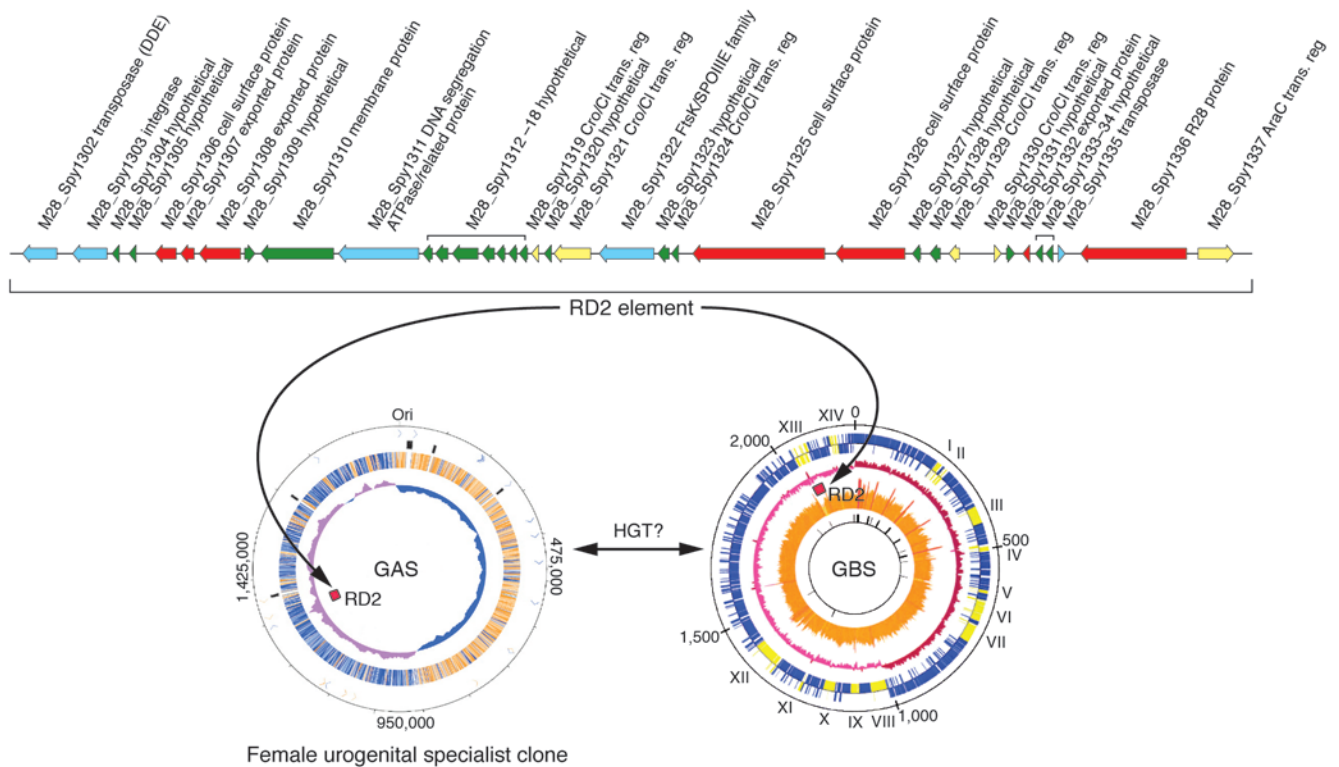
Figure 3
Molecular pathogenomic analysis reveals the contribution of a HGT event to the evolution of a female urogenital specialist GAS clone of serotype M28. The 37-kb region of difference 2 (RD2) element is shared among GBS strains and serotype M28 GAS. An HGT event may have occurred directly between a GAS and GBS strain, or an unknown third-party donor may have participated. Adapted with permission from *The Journal of Infectious Diseases* (23) and *Molecular Microbiology* (125).

20 years (22, 46, 95, 96, 98, 99). Comparative pathogenomic analysis resulted in two particularly important findings (22). First, low- and high-virulence serotype M1 strains differ in bacteriophage content and chromosomal integration site (22). Second, it was unexpectedly also discovered that another HGT event, involving reciprocal recombination of a 36-kb chromosomal region encoding the secreted toxins streptolysin O and NAD+-glycohydrolase, was a critical evolutionary event that shaped the genome of contemporary virulent M1 strains (22). The likely mechanism underlying this event was generalized transduction, a process involving inadvertent packaging of a random chromosomal segment from a donor strain into a bacteriophage capsid head, followed by transfer to a new recipient bacterial strain. Importantly, contemporary virulent M1 strains produce high levels of these two toxins compared with older, less virulent M1 strains, but the underlying molecular mechanism of increased expression is not yet understood.

Host susceptibility to invasive GAS disease
Like other infectious diseases, the clinical manifestations of GAS infection reflect interaction between the bacterium and the host. In contrast to situations in which infection with the pathogen is the prime determinant of disease, such as occurs for *Bacillus anthracis* or ebola virus, most humans are repeatedly exposed to and colonized by GAS without developing clinical symptoms. Moreover, GAS isolates that cause superficial infections such as pharyngitis and impetigo can be genetically closely related to those of the same M protein serotype that cause lethal infections such as NF and

toxic shock syndrome (90, 100). Therefore, it is highly probable that the development of severe clinical manifestations following GAS infection has a strong host susceptibility component. Most investigations of host susceptibility to severe GAS infection have focused on the role of HLA polymorphisms (101, 102). The interaction of GAS superantigens with HLA class II molecules on antigen-presenting cells can result in the activation of up to 25% of all T cells at a given time, although the number of T cells activated by a particular GAS superantigen varies substantially from person to person (103, 104). The protective and deleterious roles of particular HLA alleles in the development of streptococcal toxic shock syndrome were observed in an epidemiologic investigation and confirmed using transgenic mice (101, 102, 105).

In addition to focused research on HLA polymorphisms, recent studies have also begun to dissect the molecular basis of host susceptibility to serious GAS infection in mice using a genome-wide approach (106, 107). Researchers have exploited differences in susceptibility to GAS infection among strains of genetically defined mice to begin to localize protective and deleterious host genetic polymorphisms (106, 107). For example, a heightened inflammatory response to GAS was associated with worse outcomes in a mouse model of infection and linked to alterations in the expression of genes involved in apoptosis, macrophage activation, and prostaglandin synthesis (106). A genome-wide transcriptome analysis of murine macrophages also identified genes encoding proteins involved in prostaglandin synthesis as being upregulated during interaction with GAS, and the use of

inhibitors of prostaglandin synthesis has been associated with severe GAS infection in humans (108, 109). There are significant experimental design and execution barriers to extrapolating the genome-wide approach to enhance understanding of human susceptibility to GAS infections. However, the key point is that the increasing availability of tools such as genome-wide SNP analysis holds significant promise for increasing our understanding of human host determinants of GAS infection.

Serotype M28 and puerperal sepsis: pathogenesis consequences of a chimeric genome

Although group B *Streptococcus* (GBS) is a well-known cause of serious neonatal or maternal infections, GAS can also be responsible for these infections. Genome sequencing and pathogenesis studies have provided unexpected clues as to why serotype M28 GAS strains are repeatedly overrepresented in puerperal sepsis (childbed fever) cases (23).

It was hypothesized (23) that analysis of the genome sequence of a serotype M28 GAS strain causing puerperal sepsis would identify novel genetic elements that contributed to the overrepresentation of strains of this serotype in this infection type. Analysis of the genome sequence of a serotype M28 strain has borne this out, providing the highly unexpected discovery that these GAS strains have a genome that is a chimera, composed of largely GAS genetic material onto which has been molecularly grafted a large piece of foreign DNA shared with GBS strains (23) (Figure 3). The foreign DNA is 37.4 kb in size, was acquired by HGT, and encodes seven secreted proteins that are produced in human infections (35). With very few exceptions, this genetic element is not present in other GAS strains. Importantly, one of the secreted proteins mediates attachment of M28 strains to human urogenital epithelium; another binds to GP340, a large human glycoprotein abundantly present in vaginal and oral secretions (35, 110). Thus, the M28 genome has been shaped by acquisition of a foreign genetic element that assisted in creating a disease-specialist GAS strain. As with other GAS molecular pathogenomics studies, these discoveries have served to catalyze downstream pathogenesis experiments. Studies are ongoing that are designed to provide a deeper understanding of the precise molecular processes involved in this niche adaptation and subsequent infection.

Genome-wide dissection of the molecular events underlying GAS epidemics

GAS has been used as a model system to study the molecular processes contributing to epidemics. Since 1992, more than 350 serotype M3 strains have been recovered in a prospective population-based surveillance study of GAS invasive infections being conducted in Ontario, Canada (27, 111, 112). These strains have caused two temporally distinct epidemic waves, centered in 1995 and 2000 (19). A molecular pathogenomics approach allowed the identification of key contributors to the episodic behavior of the GAS invasive isolates. First, the distinct epidemics were shown to be caused by a heterogeneous array of serotype M3 subclones, rather than recycling of a single clone (19). These distinct clones were characterized by the acquisition or loss of specific prophages that encode known GAS virulence factors. Second, host selective pressure appears to have resulted in a highly successful clone that rose to dominance in the second epidemic wave. DNA sequence analysis, coupled with immunologic studies, identified a four–amino acid duplication in the amino terminus of M protein in the new subclone responsible for

many of the invasive cases in the 2000 epidemic (19). This duplication resulted in alterations in linear B cell epitopes, which produced substantial differences in the ability of human PMNs to phagocytosize and kill strains with the variant M protein (19). This key finding indicated that subtle or relatively minor allelic variation may participate in clone emergence and perpetuation of the epidemics.

To identify microbial genes or specific allelic variants that influence the outcome of host-pathogen interactions, it is possible to use a strategy analogous to that commonly practiced when undertaking genome-wide association studies (GWASs) — using genetic methods such as high-density SNP analysis — on humans with particular disease phenotypes, such as type 2 diabetes mellitus, macular degeneration, and schizophrenia (113). In many regards, GWASs are considerably simpler in bacterial infectious disease studies because, compared with humans, prokaryotic organisms have very small genomes and isogenic mutant strains can be generated and used to confirm the findings from the genetic association study. A novel bacterial GWAS conducted on strains recovered in the Ontario GAS epidemics (17) led to the discovery that a single nucleotide mutation in the gene encoding the MtsR metalloregulatory protein implicated in uptake of iron or manganese results in a decrease in the ability of GAS to cause devastating NF. The mechanism underlying this decreased ability to cause severe disease involves dysregulation of the control circuit responsible for wild-type levels of SpeB (our unpublished observations), a potent broad-spectrum protease that degrades extracellular matrix proteins, inactivates innate immune molecules, and destroys host tissue (114, 115). Inasmuch as single nucleotide mutations are the most abundant cause of genetic variation among members of the same species (13–15, 17, 18, 116), this discovery has broad implications for the confluence of bacterial molecular population genomics and pathogenesis research.

Accelerated vaccine candidate identification and development

In the long term, the key to the control of most human pathogens is the use of efficacious vaccines, and GAS is no exception. Transformation of vaccine research and development by genome-wide analyses is expertly summarized in this Review series by Rinaudo et al. (117). Nevertheless, we believe that a few points related to GAS vaccine research need to be stressed here. Although the quest for a successful human vaccine against GAS pharyngitis and invasive infections has been ongoing for many decades (reviewed in ref. 8), and much progress has been made, a product licensed for use in the United States or elsewhere is still lacking. The many complete genome sequences of GAS that are publicly available have been used to assist the decades-long search for GAS vaccine candidates. For example, Lei et al. (118) analyzed the extracellular proteome and identified several previously undescribed proteins that have been the subject of subsequent vaccine research. Similarly, bioinformatic analysis of multiple GAS genomes identified 16 highly conserved cell surface lipoproteins. Subsequent biochemical, immunological, and mouse model research identified five lipoproteins as potential new vaccine candidates (119). Using a proteomics approach, Cole et al. showed that the GAS cell surface contains numerous immunogenic proteins not previously suspected of being associated with the cell wall, thus expanding the number of potential GAS vaccine candidates (120). The finding of proteins on the GAS cell surface or in the culture supernatant that lack traditional motifs associated with cell surface and secreted

proteins has been a consistent finding of proteomic studies, stressing the need for experimental validation of predictions based on bioinformatic analyses (118, 120–123). Improved modalities for separating GAS cellular versus cell surface constituents may provide additional insights into novel GAS proteins that are critical for host-pathogen interaction and are thus potential vaccine candidates (124). In the most extensive GAS vaccine candidate study to date, Rodriguez-Ortega et al. (50) analyzed the surface-exposed proteome of GAS and identified one new antigen (SpyCEP) that conferred the ability to protect mice against lethal infection.

What does the future hold?

Enormous strides have been made in the last decade in our understanding of GAS-host interactions, and molecular pathogenomics studies have contributed substantially. However, many gaps remain in our knowledge. Of note, many aspects of the host response during GAS infection remain largely unknown, but application of the genome-wide integrative strategies discussed above to the host side of the equation are likely to yield useful data. Analysis of patients with ARF may be particularly useful in this regard, as elucidating the host genetic factors that contribute to susceptibility may aid in the development of novel diagnostics and treatments for this devastating disease. In addition, DNA sequencing costs continue to decrease dramatically, opening the door to many types of projects previously impossible due to financial constraints. For example, it is now reasonable to consider projects that involve genome sequence analysis of many hundreds or more GAS strains, similar to the 1,000–human genome project that is well underway (http://www.1000genomes.org/). Finally, we are hopeful that the new avenues of basic and translational research made possible by molecular pathogenomics studies will ultimately provide strategies for ameliorating human morbidity and mortality caused by GAS.

1. Olsen, R.J., Shelburne, S.A., and Musser, J.M. 2009. Molecular mechanisms underlying group A streptococcal pathogenesis. *Cell Microbiol.* **11**:1–12.
2. Pichichero, M.E. 1998. Group A beta-hemolytic streptococcal infections. *Pediatr. Rev.* **19**:291–302.
3. Carapetis, J.R., Steer, A.C., Mulholland, E.K., and Weber, M. 2005. The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* **5**:685–694.
4. Carapetis, J.R., McDonald, M., and Wilson, N.J. 2005. Acute rheumatic fever. *Lancet.* **366**:155–168.
5. Tanz, R.R., et al. 2004. Community-based surveillance in the United States of macrolide-resistant pediatric pharyngeal group A streptococci during 3 respiratory disease seasons. *Clin. Infect. Dis.* **39**:1794–1801.
6. Robinson, D.A., Sutcliffe, J.A., Tewodros, W., Manoharan, A., and Bessen, D.E. 2006. Evolution and global dissemination of macrolide-resistant group A streptococci. *Antimicrob. Agents Chemother.* **50**:2903–2911.
7. Martin, J.M., Green, M., Barbadora, K.A., and Wald, E.R. 2002. Erythromycin-resistant group A streptococci in schoolchildren in Pittsburgh. *N. Engl. J. Med.* **346**:1200–1206.
8. Reid, S.D., Virtaneva, K., and Musser, J.M. 2003. Group A *Streptococcus* vaccine research: historical synopsis and new insights. In *Bacterial vaccines.* R.W. Ellis and B.R. Brodeur, editors. Landes Bioscience. Austin, Texas, USA. pp. 155–173.
9. Bisno, A.L., Rubin, F.A., Cleary, P.P., and Dale, J.B. 2005. Prospects for a group A streptococcal vaccine: rationale, feasibility, and obstacles--report of a National Institute of Allergy and Infectious Diseases workshop. *Clin. Infect. Dis.* **41**:1150–1156.
10. Musser, J.M., and DeLeo, F.R. 2005. Toward a genome-wide systems biology analysis of host-pathogen interactions in group A *Streptococcus. Am. J. Pathol.* **167**:1461–1472.
11. Whittam, T.S., and Bumbaugh, A.C. 2002. Inferences from whole-genome sequences of bacterial pathogens. *Curr. Opin. Genet. Dev.* **12**:719–725.
12. Blattner, F.R., et al. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.* **277**:1453–1474.
13. Perna, N.T., et al. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature.* **409**:529–533.
14. Welch, R.A., et al. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli. Proc. Natl. Acad. Sci. U. S. A.* **99**:17020–17024.
15. Beres, S.B., and Musser, J.M. 2007. Contribution of exogenous genetic elements to the group A *Streptococcus* metagenome. *PLoS ONE.* **2**:e800.
16. Ochman, H., Lerat, E., and Daubin, V. 2005. Examining bacterial species under the specter of gene transfer and exchange. *Proc. Natl. Acad. Sci. U. S. A.* **102**(Suppl. 1):6595–6599.
17. Beres, S.B., et al. 2006. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A *Streptococcus. Proc. Natl. Acad. Sci. U. S. A.* **103**:7059–7064.
18. Kennedy, A.D., et al. 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. *Proc. Natl. Acad. Sci. U. S. A.* **105**:1327–1332.
19. Beres, S.B., et al. 2004. Genome-wide molecular dissection of serotype M3 group A *Streptococcus* strains causing two epidemics of invasive infections. *Proc. Natl. Acad. Sci. U. S. A.* **101**:11833–11838.
20. McShan, W.M., et al. 2008. Genome sequence of a nephritogenic and highly transformable M49 strain of *Streptococcus pyogenes. J. Bacteriol.* **190**:7773–7785.
21. Holden, M.T., et al. 2007. Complete genome of acute rheumatic fever-associated serotype M5 *Streptococcus pyogenes* strain Manfredo. *J. Bacteriol.* **189**:1473–1477.
22. Sumby, P., et al. 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 group A *Streptococcus* involved multiple horizontal gene transfer events. *J. Infect. Dis.* **192**:771–782.
23. Green, N.M., et al. 2005. Genome sequence of a serotype M28 strain of group A *Streptococcus*: potential new insights into puerperal sepsis and bacterial disease specificity. *J. Infect. Dis.* **192**:760–770.
24. Banks, D.J., et al. 2004. Progress toward characterization of the group A *Streptococcus* metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. *J. Infect. Dis.* **190**:727–738.
25. Nakagawa, I., et al. 2003. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res.* **13**:1042–1055.
26. Smoot, J.C., et al. 2002. Genome sequence and comparative microarray analysis of serotype M18 group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. *Proc. Natl. Acad. Sci. U. S. A.* **99**:4668–4673.
27. Beres, S.B., et al. 2002. Genome sequence of a serotype M3 strain of group A *Streptococcus*: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc. Natl. Acad. Sci. U. S. A.* **99**:10078–10083.
28. Ferretti, J.J., et al. 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes. Proc. Natl. Acad. Sci. U. S. A.* **98**:4658–4663.
29. Shulman, S.T., et al. 2004. Group A streptococcal pharyngitis serotype surveillance in North America, 2000-2002. *Clin. Infect. Dis.* **39**:325–332.
30. O'Loughlin, R.E., et al. 2007. The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. *Clin. Infect. Dis.* **45**:853–862.
31. Luca-Harari, B. et al. 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. *J. Clin. Microbiol.* In press.
32. Abbot, E.L., et al. 2007. Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. *Cell. Microbiol.* **9**:1822–1833.
33. Mora, M., et al. 2005. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl. Acad. Sci. U. S. A.* **102**:15641–15646.
34. Sitkiewicz, I., et al. 2006. Emergence of a bacterial clone with enhanced virulence by acquisition of a phage encoding a secreted phospholipase A2. *Proc. Natl. Acad. Sci. U. S. A.* **103**:16009–16014.
35. Zhang, S., Green, N.M., Sitkiewicz, I., Lefebvre, R.B., and Musser, J.M. 2006. Identification and characterization of an antigen I/II family protein produced by group A *Streptococcus. Infect. Immun.* **74**:4200–4213.
36. Manetti, A.G., et al. 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* **64**:968–983.
37. Edwards, R.J., et al. 2005. Specific C-terminal cleavage and inactivation of interleukin-8 by invasive disease isolates of *Streptococcus pyogenes. J. Infect. Dis.* **192**:783–790.
38. Coye, L.H., and Collins, C.M. 2004. Identification of SpyA, a novel ADP-ribosyltransferase of *Streptococcus pyogenes. Mol. Microbiol.* **54**:89–98.
39. Collin, M., and Olsen, A. 2001. EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J.*

20:3046–3055.

40. von Pawel-Rammingen, U., Johansson, B.P., and Bjorck, L. 2002. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J.* **21**:1607–1615.

41. Rasmussen, M., Eden, A., and Bjorck, L. 2000. SclA, a novel collagen-like surface protein of *Streptococcus pyogenes*. *Infect. Immun.* **68**:6370–6377.

42. Malke, H., Steiner, K., McShan, W.M., and Ferretti, J.J. 2006. Linking the nutritional status of *Streptococcus pyogenes* to alteration of transcriptional gene expression: the action of CodY and RelA. *Int. J. Med. Microbiol.* **296**:259–275.

43. Sumby, P., et al. 2005. Extracellular deoxyribonuclease made by group A *Streptococcus* assists pathogenesis by enhancing evasion of the innate immune response. *Proc. Natl. Acad. Sci. U. S. A.* **102**:1679–1684.

44. Rasmussen, M., and Bjorck, L. 2001. Unique regulation of SclB - a novel collagen-like surface protein of *Streptococcus pyogenes*. *Mol. Microbiol.* **40**:1427–1438.

45. Shelburne, S.A., 3rd, et al. 2006. Maltodextrin utilization plays a key role in the ability of group A *Streptococcus* to colonize the oropharynx. *Infect. Immun.* **74**:4605–4614.

46. Sumby, P., et al. 2008. A chemokine-degrading extracellular protease made by group A *Streptococcus* alters pathogenesis by enhancing evasion of the innate immune response. *Infect. Immun.* **76**:978–985.

47. Hidalgo-Grass, C., et al. 2006. A streptococcal protease that degrades CXC chemokines and impairs bacterial clearance from infected tissues. *EMBO J.* **25**:4628–4637.

48. Zinkernagel, A.S., et al. 2008. The IL-8 protease SpyCEP/ScpC of group A *Streptococcus* promotes resistance to neutrophil killing. *Cell Host Microbe.* **4**:170–178.

49. Sjolinder, H., et al. 2008. The ScpC protease of *Streptococcus pyogenes* affects the outcome of sepsis in a murine model. *Infect. Immun.* **76**:3959–3966.

50. Rodriguez-Ortega, M.J., et al. 2006. Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome. *Nat. Biotechnol.* **24**:191–197.

51. Lei, B., et al. 2002. Opsonophagocytosis-inhibiting Mac protein of group A *Streptococcus*: identification and characteristics of two genetic complexes. *Infect. Immun.* **70**:6880–6890.

52. Caswell, C.C., Lukomska, E., Seo, N.S., Hook, M., and Lukomski, S. 2007. Scl1-dependent internalization of group A *Streptococcus* via direct interactions with the alpha2beta(1) integrin enhances pathogen survival and re-emergence. *Mol. Microbiol.* **64**:1319–1331.

53. Proft, T., Moffatt, S.L., Berkahn, C.J., and Fraser, J.D. 1999. Identification and characterization of novel superantigens from *Streptococcus pyogenes*. *J. Exp. Med.* **189**:89–102.

54. Unnikrishnan, M., et al. 2002. The bacterial superantigen streptococcal mitogenic exotoxin Z is the major immunoactive agent of *Streptococcus pyogenes*. *J. Immunol.* **169**:2561–2569.

55. Proft, T., Webb, P.D., Handley, V., and Fraser, J.D. 2003. Two novel superantigens found in both group A and group C *Streptococcus*. *Infect. Immun.* **71**:1361–1369.

56. Kreikemeyer, B., McIver, K.S., and Podbielski, A. 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. *Trends Microbiol.* **11**:224–232.

57. Levin, J.C., and Wessels, M.R. 1998. Identification of *csrR/csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A *Streptococcus*. *Mol. Microbiol.* **30**:209–219.

58. Dalton, T.L., Collins, J.T., Barnett, T.C., and Scott, J.R. 2006. RscA, a member of the MDR1 family of transporters, is repressed by CovR and required for

growth of *Streptococcus pyogenes* under heat stress. *J. Bacteriol.* **188**:77–85.

59. 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.

60. Voyich, J.M., et al. 2004. Engagement of the pathogen survival response used by group A *Streptococcus* to avert destruction by innate host defense. *J. Immunol.* **173**:1194–1201.

61. Ribardo, D.A., Lambert, T.J., and McIver, K.S. 2004. Role of *Streptococcus pyogenes* two-component response regulators in the temporal control of Mga and the Mga-regulated virulence gene *emm*. *Infect. Immun.* **72**:3668–3673.

62. Sitkiewicz, I., and Musser, J.M. 2006. Expression microarray and mouse virulence analysis of four conserved two-component gene regulatory systems in group A *Streptococcus*. *Infect. Immun.* **74**:1339–1351.

63. Chaussee, M.S., et al. 2002. Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes*. *Infect. Immun.* **70**:762–770.

64. Kreikemeyer, B., et al. 2007. The *Streptococcus pyogenes* serotype M49 Nra-Ralp3 transcriptional regulatory network and its control of virulence factor expression from the novel *eno ralp3 epf sagA* pathogenicity region. *Infect. Immun.* **75**:5698–5710.

65. Gryllos, I., et al. 2008. PerR confers phagocytic killing resistance and allows pharyngeal colonization by group A *Streptococcus*. *PLoS Pathog.* **4**:e1000145.

66. Kinkel, T.L., and McIver, K.S. 2008. CcpA-mediated repression of streptolysin S expression and virulence in the group A *Streptococcus*. *Infect. Immun.* **76**:3451–3463.

67. Roberts, S.A., and Scott, J.R. 2007. RivR and the small RNA RivX: the missing links between the CovR regulatory cascade and the Mga regulon. *Mol. Microbiol.* **66**:1506–1522.

68. Ribardo, D.A., and McIver, K.S. 2006. Defining the Mga regulon: Comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the group A *Streptococcus*. *Mol. Microbiol.* **62**:491–508.

69. Ryan, P.A., Kirk, B.W., Euler, C.W., Schuch, R., and Fischetti, V.A. 2007. Novel algorithms reveal streptococcal transcriptomes and clues about undefined genes. *PLoS Comput. Biol.* **3**:e132.

70. Loughman, J.A., and Caparon, M. 2006. Regulation of SpeB in *Streptococcus pyogenes* by pH and NaCl: a model for in vivo gene expression. *J. Bacteriol.* **188**:399–408.

71. Chaussee, M.A., Dmitriev, A.V., Callegari, E.A., and Chaussee, M.S. 2008. Growth phase-associated changes in the transcriptome and proteome of *Streptococcus pyogenes*. *Arch. Microbiol.* **189**:27–41.

72. Beyer-Sehlmeyer, G., Kreikemeyer, B., Horster, A., and Podbielski, A. 2005. Analysis of the growth phase-associated transcriptome of *Streptococcus pyogenes*. *Int. J. Med. Microbiol.* **295**:161–177.

73. Voyich, J.M., et al. 2003. Genome-wide protective response used by group A *Streptococcus* to evade destruction by human polymorphonuclear leukocytes. *Proc. Natl. Acad. Sci. U. S. A.* **100**:1996–2001.

74. Shelburne, S.A., 3rd, et al. 2008. A direct link between carbohydrate utilization and virulence in the major human pathogen group A *Streptococcus*. *Proc. Natl. Acad. Sci. U. S. A.* **105**:1698–1703.

75. Virtaneva, K., et al. 2005. Longitudinal analysis of the group A *Streptococcus* transcriptome in experimental pharyngitis in cynomolgus macaques. *Proc. Natl. Acad. Sci. U. S. A.* **102**:9014–9019.

76. Graham, M.R., et al. 2006. Analysis of the transcriptome of group A *Streptococcus* in mouse soft tissue infection. *Am. J. Pathol.* **169**:927–942.

77. Shelburne, S.A., 3rd, et al. 2005. Central role of a bacterial two-component gene regulatory system of previously unknown function in pathogen persis-

tence in human saliva. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 16037–16042.

78. Graham, M.R., et al. 2005. Group A *Streptococcus* transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. *Am. J. Pathol.* **166**:455–465.

79. Loughman, J.A., and Caparon, M.G. 2006. A novel adaptation of aldolase regulates virulence in *Streptococcus pyogenes*. *EMBO J.* **25**:5414–5422.

80. Loughman, J.A., and Caparon, M.G. 2007. Comparative functional analysis of the lac operons in *Streptococcus pyogenes*. *Mol. Microbiol.* **64**:269–280.

81. Hanks, T.S., et al. 2006. Differential regulation of iron- and manganese-specific MtsABC and heme-specific HtsABC transporters by the metalloregulator MtsR of group A *Streptococcus*. *Infect. Immun.* **74**:5132–5139.

82. Ricci, S., Janulczyk, R., and Bjorck, L. 2002. The regulator PerR is involved in oxidative stress response and iron homeostasis and is necessary for full virulence of *Streptococcus pyogenes*. *Infect. Immun.* **70**:4968–4976.

83. Brenot, A., King, K.Y., and Caparon, M.G. 2005. The PerR regulon in peroxide resistance and virulence of *Streptococcus pyogenes*. *Mol. Microbiol.* **55**:221–234.

84. Peter, G., and Smith, A.L. 1977. Group A streptococcal infections of the skin and pharynx (first of two parts). *N. Engl. J. Med.* **297**:311–317.

85. Bessen, D.E., and Kalia, A. 2002. Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in *Streptococcus pyogenes*. *Infect. Immun.* **70**:1159–1167.

86. Bessen, D.E., Sotir, C.M., Readdy, T.L., and Hollingshead, S.K. 1996. Genetic correlates of throat and skin isolates of group A streptococci. *J. Infect. Dis.* **173**:896–900.

87. Kratovac, Z., Manoharan, A., Luo, F., Lizano, S., and Bessen, D.E. 2007. Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. *J. Bacteriol.* **189**:1299–1310.

88. Hung, C.S., et al. 2002. Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* **44**:903–915.

89. Haukness, H.A., et al. 2002. The heterogeneity of endemic community pediatric group A streptococcal pharyngeal isolates and their relationship to invasive isolates. *J. Infect. Dis.* **185**:915–920.

90. Hoe, N.P., et al. 2001. Distribution of streptococcal inhibitor of complement variants in pharyngitis and invasive isolates in an epidemic of serotype M1 group A *Streptococcus* infection. *J. Infect. Dis.* **183**:633–639.

91. Sumby, P., Whitney, A.R., Graviss, E.A., DeLeo, F.R., and Musser, J.M. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* **2**:e5.

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

93. Engleberg, N.C., Heath, A., Miller, A., Rivera, C., and DiRita, V.J. 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. *J. Infect. Dis.* **183**:1043–1054.

94. Daubin, V., Moran, N.A., and Ochman, H. 2003. Phylogenetics and the cohesion of bacterial genomes. *Science.* **301**:829–832.

95. Musser, J.M., et al. 1991. *Streptococcus pyogenes* causing toxic-shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. *Proc. Natl. Acad. Sci. U. S. A.* **88**:2668–2672.

96. Musser, J.M., et al. 1995. 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.* **63**:994–1003.

97. Musser, J.M., et al. 1993. Geographic and temporal distribution and molecular characterization of two highly pathogenic clones of *Streptococcus pyogenes* expressing allelic variants of pyrogenic exotoxin A (scarlet fever toxin). *J. Infect. Dis.* **167**:337–346.

98. Aziz, R.K., et al. 2005. Mosaic prophages with horizontally acquired genes account for the emergence and diversification of the globally disseminated M1T1 clone of *Streptococcus pyogenes*. *J. Bacteriol.* **187**:3311–3318.

99. Cleary, P.P., et al. 1992. Clonal basis for resurgence of serious *Streptococcus pyogenes* disease in the 1980s. *Lancet.* **339**:518–521.

100. Chatellier, S., et al. 2000. Genetic relatedness and superantigen expression in group A *Streptococcus* serotype M1 isolates from patients with severe and non-severe invasive diseases. *Infect. Immun.* **68**:3523–3534.

101. Norrby-Teglund, A., Nepom, G.T., and Kotb, M. 2002. Differential presentation of group A streptococcal superantigens by HLA class II DQ and DR alleles. *Eur. J. Immunol.* **32**:2570–2577.

102. Kotb, M., et al. 2002. An immunogenetic and molecular basis for differences in outcomes of invasive group A streptococcal infections. *Nat. Med.* **8**:1398–1404.

103. Sundberg, E., and Jardetzky, T.S. 1999. Structural basis for HLA-DQ binding by the streptococcal superantigen SSA. *Nat. Struct. Biol.* **6**:123–129.

104. Norrby-Teglund, A., et al. 2000. Host variation in cytokine responses to superantigens determine the severity of invasive group A streptococcal infection. *Eur. J. Immunol.* **30**:3247–3255.

105. Nooh, M.M., El-Gengehi, N., Kansal, R., David, C.S., and Kotb, M. 2007. HLA transgenic mice provide evidence for a direct and dominant role of HLA class II variation in modulating the severity of streptococcal sepsis. *J. Immunol.* **178**:3076–3083.

106. Abdeltawab, N.F., et al. 2008. An unbiased systems genetics approach to mapping genetic loci modulating susceptibility to severe streptococcal sepsis. *PLoS Pathog.* **4**:e1000042.

107. Aziz, R.K., et al. 2007. Susceptibility to severe streptococcal sepsis: use of a large set of isogenic mouse lines to study genetic and environmental factors. *Genes Immun.* **8**:404–415.

108. Goldmann, O., et al. 2007. Transcriptome analysis of murine macrophages in response to infection with *Streptococcus pyogenes* reveals an unusual activation program. *Infect. Immun.* **75**:4148–4157.

109. Hamilton, S.M., Bayer, C.R., Stevens, D.L., Lieber, R.L., and Bryant, A.E. 2008. Muscle injury, vimentin expression, and nonsteroidal anti-inflammatory drugs predispose to cryptic group A streptococcal necrotizing infection. *J. Infect. Dis.* **198**:1692–1698.

110. Stalhammar-Carlemalm, M., Areschoug, T., Larsson, C., and Lindahl, G. 1999. The R28 protein of *Streptococcus pyogenes* is related to several group B streptococcal surface proteins, confers protective immunity and promotes binding to human epithelial cells. *Mol. Microbiol.* **33**:208–219.

111. Sharkawy, A., et al. 2002. Severe group A streptococcal soft-tissue infections in Ontario: 1992–1996. *Clin. Infect. Dis.* **34**:454–460.

112. Davies, H.D., et al. 1996. Invasive group A streptococcal infections in Ontario, Canada. Ontario Group A Streptococcal Study Group. *N. Engl. J. Med.* **335**:547–554.

113. McCarthy, M.I., et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9**:356–369.

114. Svensson, M.D., et al. 2000. Role for a secreted cysteine proteinase in the establishment of host tissue tropism by group A streptococci. *Mol. Microbiol.* **38**:242–253.

115. Lukomski, S., et al. 1998. Genetic inactivation of an extracellular cysteine protease (SpeB) expressed by *Streptococcus pyogenes* decreases resistance to phagocytosis and dissemination to organs. *Infect. Immun.* **66**:771–776.

116. Jakobsson, M., et al. 2008. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature.* **451**:998–1003.

117. Rinaudo, C.D., Telford, J.L., Rappuoli, R., and Seib, K.L. 2009. Vaccinology in the genome era. *J. Clin. Invest.* **119**:2515–2525.

118. Lei, B., Mackie, S., Lukomski, S., and Musser, J.M. 2000. Identification and immunogenicity of group A *Streptococcus* culture supernatant proteins. *Infect. Immun.* **68**:6807–6818.

119. Lei, B., Liu, M., Chesney, G.L., and Musser, J.M. 2004. Identification of new candidate vaccine antigens made by *Streptococcus pyogenes*: purification and characterization of 16 putative extracellular lipoproteins. *J. Infect. Dis.* **189**:79–89.

120. Cole, J.N., et al. 2005. Surface analyses and immune reactivities of major cell wall-associated proteins of group A *Streptococcus*. *Infect. Immun.* **73**:3137–3146.

121. Severin, A., et al. 2007. Proteomic analysis and identification of *Streptococcus pyogenes* surface-associated proteins. *J. Bacteriol.* **189**:1514–1522.

122. Nakamura, T., et al. 2004. Two-dimensional gel electrophoresis analysis of the abundance of virulent exoproteins of group A *Streptococcus* caused by environmental changes. *Arch. Microbiol.* **181**:74–81.

123. Zhang, M., et al. 2007. Group A *Streptococcus* cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media. *Proteomics.* **7**:1379–1390.

124. Koller, T., et al. 2008. PlyC, a novel bacteriophage lysin for compartment-dependent proteomics of group A streptococci. *Proteomics.* **8**:140–148.

125. Glaser, P., et al. 2002. Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease. *Mol. Microbiol.* **45**:1499–1513.