Longitudinal analysis of the group A *Streptococcus* transcriptome in experimental pharyngitis in cynomolgus macaques

Kimmo Virtaneva*†, Stephen F. Porcella*†, Morag R. Graham*‡, Robin M. Ireland*, Claire A. Johnson*, Stacy M. Ricklefs*, Imran Babar*, Larye D. Parkins*, Romina A. Romero*§, G. Judson Corn*, Don J. Gardner¶, John R. Bailey¶, Michael J. Parnell¶, and James M. Musser*‖**

*Laboratory of Human Bacterial Pathogenesis and ¶Veterinary Branch, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT 59840; and ‖Center for Human Bacterial Pathogenesis Research, Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030

Communicated by Richard M. Krause, National Institutes of Health, Bethesda, MD, May 10, 2005 (received for review March 3, 2005)

Identification of the genetic events that contribute to host–pathogen interactions is important for understanding the natural history of infectious diseases and developing therapeutics. Transcriptome studies conducted on pathogens have been central to this goal in recent years. However, most of these investigations have focused on specific end points or disease phases, rather than analysis of the entire time course of infection. To gain a more complete understanding of how bacterial gene expression changes over time in a primate host, the transcriptome of group A *Streptococcus* (GAS) was analyzed during an 86-day infection protocol in 20 cynomolgus macaques with experimental pharyngitis. The study used 260 custom Affymetrix (Santa Clara, CA) chips, and data were confirmed by TaqMan analysis. Colonization, acute, and asymptomatic phases of disease were identified. Successful colonization and severe inflammation were significantly correlated with an early onset of superantigen gene expression. The differential expression of two-component regulators *covR* and *spy0680* (*M1_spy0874*) was significantly associated with GAS colony-forming units, inflammation, and phases of disease. Prophage virulence gene expression and prophage induction occurred predominantly during high pathogen cell densities and acute inflammation. We discovered that temporal changes in the GAS transcriptome were integrally linked to the phase of clinical disease and host-defense response. Knowledge of the gene expression patterns characterizing each phase of pathogen–host interaction provides avenues for targeted investigation of proven and putative virulence factors and genes of unknown function and will assist vaccine research.

Affymetrix GeneChip | clinical correlation | colonization | carriage

Identification of the genetic events that contribute to host–pathogen interactions is important for understanding the natural history of infectious diseases and developing therapeutics. Transcriptome studies that analyzed specific end points or disease phases (1–8) have been central to this goal in recent years. Although extensive information has been generated, the conclusions from these important studies are limited to the events occurring at the small number of analyzed time points.

Group A *Streptococcus* (GAS) is a Gram-positive bacterial pathogen that causes extensive human morbidity worldwide (9, 10). This pathogen is responsible for 2 million cases of pharyngitis (PHG) and 15,000 invasive infections annually in the United States. The organism also causes acute rheumatic fever and rheumatic heart disease, which are major public health problems in developing nations. No licensed human vaccine is available to prevent GAS infection, and efficacious therapeutics are needed for treating patients with severe invasive infections. Genome-wide study of GAS has been a fruitful area of research in recent years (11–22). Genome sequences available for six GAS strains and DNA–DNA microarray analyses have revealed the key contribution of bacteriophages to strain diversification, epidemic waves, and pathogenesis (11–16, 21, 23). Similarly, expression microarray analyses conducted on organisms grown *in vitro* have shed light on strategies that are used by GAS to survive and avoid phagocytosis and killing (17–20, 22).

Until very recently, the analysis of GAS-mediated processes contributing to PHG was restricted almost solely to inferences obtainable by serologic study of host antibody responses to relatively few extracellular molecules. Although important information has been obtained from these studies, the inability to directly examine expression of large numbers of GAS genes over time in the posterior pharynx means that we have only a very imprecise understanding of molecular processes contributing to PHG. Several key advances have been made in recent years (24–26), but a far more detailed genetic understanding of the natural history of GAS–host interaction is needed. Thus, we analyzed the transcriptome of GAS during 86 days of upper respiratory tract interaction in 20 cynomolgus macaques. This primate species gets clinical disease that closely mirrors human GAS pharyngitis (PHG) (24, 27) and, hence, is an excellent model. We used a longitudinal experimental protocol designed to identify pathogen genes whose expression was positively or negatively correlated with disease phase, number of colony-forming units (CFUs), and inflammation severity. A serotype M1 strain was used because very large population-based studies have shown that organisms of this antigenic type are frequently the most common cause of PHG and invasive infections (reviewed in ref. 10). The study revealed that temporal changes in the GAS transcriptome were integrally linked to the phase of clinical disease and host-defense response.

Materials and Methods

Detailed descriptions of the procedures that we used are given in *Supporting Materials and Methods*, which is published as supporting information on the PNAS web site.

Serotype M1 GAS Strain. Serotype M1 strain MGAS5005 (GenBank accession no. CP000017), which is genetically representa-

Freely available online through the PNAS open access option.

Abbreviations: GAS, group A *Streptococcus*; CFU, colony-forming unit; SEGS, segmented neutrophils; C-RP, C-reactive protein; TON, tonsillitis; PHG, pharyngitis; PTSAg, pyrogenic toxin superantigen; TCS, two-component gene regulatory systems.

Data deposition: The raw (.cel) files reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database (accession no. GSE2713).

†K.V. and S.F.P. contributed equally to this work.

‡Present address: Canadian Science Centre for Human and Animal Health, 1015 Arlington Street, Winnipeg, MB, Canada R3E 3R2.

§Present address: Department of Public Health, Riverside County Community Health Agency, 4065 County Circle Drive, Suite 403, Riverside, CA 92503.

**To whom correspondence should be addressed. E-mail: musser@bcm.tmc.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0503671102

Table 1. Correlation of GAS gene expression with five medical parameters during three phases of disease

Medical parameter	Days	Infection phase	Correlation	No. of genes	Annotated genes	Annotated genes, %	Representative genes*
GAS CFU	0–4	Colonization	$Z>+2$	125	102	82	*speA2*, *smeZ*, *sdaD2*, *salA regR*, spy0680/0681 (*M1_spy0874/0875*)
			$Z<-2$	147	112	76	*speJ*, *scpA*, *sda*, *ska*, *sclA*, *hlyX*, *epf*
GAS CFU	4–23	Acute	$Z>+2$	182	131	72	*speB*, *sdaD2*, *smeZ*, *cbf*, *fasC*, *vicR*, *covS*, *spy0109* (*M1_spy0128*)
			$Z<-2$	201	138	69	*covR*, *spy0681* (*M1_spy0875*), *crgR*, *ska*, *hlyX*, *cfa*, nox, *spy0720* (*M1_spy0198*)
GAS CFU	23–58	Asymptomatic	$Z>+2$	137	109	80	*mac/ideS*, *sclA*, *isp*, *hlyX*, *lplB*, *spy0385* (*M1_spy0470*), *spy0926* (*M1_spy1212*)
			$Z<-2$	150	118	79	*smeZ*, *inlA*, *emm1*, *nox*, *groEL*, *spy0680* (*M1_spy0874*)
GAS CFU	58–86	Asymptomatic	$Z>+2$	126	94	75	*ska*, *emm1*, *epf*, *ciaH*, *spy0385* (*M1_spy0470*)
			$Z<-2$	153	103	67	*nox*
SEGS	0–23	Acute	$Z>+2$	98	73	74	*isp*, *perR*, *ciaR*, *spy0720* (*M1_spy0198*)
			$Z<-2$	123	89	72	*speB*, *endoS*, *fasA*, *rofA*, *srtA*, *srtI*, *Spy0680/0681* (*M1_spy0874/0875*)
SEGS	23–86	Asymptomatic	$Z>+2$	126	90	71	*smeZ*, *grab*, *isp*, *isp2*, *ciaR*, *spy0720* (*M1_spy0198*)
			$Z<-2$	113	82	73	*sdaD2*, *spy0384* (*SpM1_spy0469*), *spy0680/0681* (*M1_spy0874/0875*)
C-RP	0–23	Acute	$Z>+2$	320	239	75	Four membrane transport regulons, 4 TCS genes, *hlyX*, *smeZ*, *epf*, *emm1*, *scpA*, *cbf*, *spd3*, *mac/ideS*, groEL
			$Z<-2$	333	243	73	Five membrane transport regulons, *speJ*, *sagA*, *sic*, *speB*, *ihk*, *fasX*, *spy0385* (*M1_spy0470*)
C-RP	23–86	Asymptomatic	$Z>+2$	245	189	77	*sdaD2*, *isp*, *scpA*, *inlA*, *zmpS*, *ihk*, *spy0384* (*M1_spy0469*)
			$Z<-2$	221	152	69	*ska*, *emm1*, *grab*, *covS*, *codY*, *lytR*, *cbf*, *spy0109* (*M1_spy0128*)
TON	0–86	Acute infection	$Z>+2$	146	113	77	*cbf*, *lplB*, *spy0384* (*SpM1_spy0469*)
			$Z<-2$	154	108	70	*smeZ*, *ska*, *hlyX*, *isp*, *isp2*, *epf*, *nox*, *ihk*, *hasB*, *perR*, *copZ*
PHG	4–23	Acute	$Z>+2$	142	104	73	*mac/ideS*, *scpA*, *sdaD2*, *hlyX*, *nox*, *fasA*, *oppA*, *copZ*
			$Z<-2$	162	114	70	*smeZ*, *sic*, *grab*, *hasB*, *hlyX*, *groEL*, *perR*, *scar*, *spy0109* (*M1_spy0128*), *spy0385* (*M1_spy0470*)
16 vs. 32 day	n.a.	Transition acute to asymptomatic	$Z>+2$	118	87	74	*speJ*, *groEL*, *amrA*, *lplB*, *spy0926* (*M1_spy1212*)
			$Z<-2$	131	97	74	*hlyX*, *isp*, *fasA*, *srtK*, *endoS*

n.a., Not applicable.

**Streptococcus pyogenes* strain SF370 ORF numbers are listed in parentheses.

tive of serotype M1 isolates obtained from patients with PHG and invasive infections in the United States, Canada, and western Europe, was grown and prepared as described in ref. 24. The genome sequence of this strain was determined recently (28). Strain MGAS5005 makes streptococcal pyrogenic exotoxin A (SpeA) and has been used extensively in animal infection studies, including monkey PHG (24, 27).

Cynomolgus Macaques. We used 20 healthy monkeys matched for age and sex. The animals were GAS-culture-negative, and they had neglible antistreptolysin O (anti-SLO) titers, indicating no recent history of GAS exposure (24, 27). The macaques were subjected to a mock-infection protocol for 5 weeks, rested for 4 weeks, and inoculated intranasally with serotype M1 GAS strain MGAS5005. Samples (throat swabs, blood, and saliva) were collected on days 0, 1, 2, 4, 7, 9, 16, 23, and 32 for all mock- and GAS-inoculated animals, and 32 clinical and laboratory parameters were measured. On day 32, 12 of 20 monkeys remained culture-positive for GAS; hence, measurement of the 32 parameters was continued for all 20 animals on days 45, 58, 72, and 86. All animals were observed for clinical signs of disease during mock and infection periods by the same veterinarian (24, 27).

Affymetrix Custom Array. The Affymetrix custom high-density antisense oligonucleotide array has been described (22).

Bacterial Target Preparation. Bacterial RNA was extracted from throat swabs as described (24) in 96-well format by using a plate centrifugation system (Qiagen, Valencia, CA). More details on the RNA-extraction protocol are provided in *Supporting Materials and Methods*. On average, 4.8 μg of total RNA was recovered per throat swab, with an average A260/280 ratio of 1.8. RNA quality was assessed by 23S and 16S peak clarity and quantity by using a 2100 Bioanalyzer and Pico analysis kit (Agilent Technologies, Palo Alto, CA). Contaminating DNA was removed as described (24). A total of 260 GAS targets, representing 20 monkeys at 13 different time points, was prepared in 96-well format by following the GeneChip *Pseudomonas aeruginosa* genome array expression-analysis protocol (Affymetrix).

Target Hybridization to Custom GeneChip. RNA (1 pg to 50 ng) was purified from each throat-swab specimen, and labeled bacterial cDNA was synthesized and hybridized to 260 Rocky Mountain Laboratory (RML) Affymetrix custom GAS arrays. The amount of labeled RNA used was 1,000-fold less than in previous microarray studies (1–4, 22). Thus, a complex statistical strategy was used to optimize detection of low transcript abundance (see Fig. 4, which is published as supporting information on the PNAS

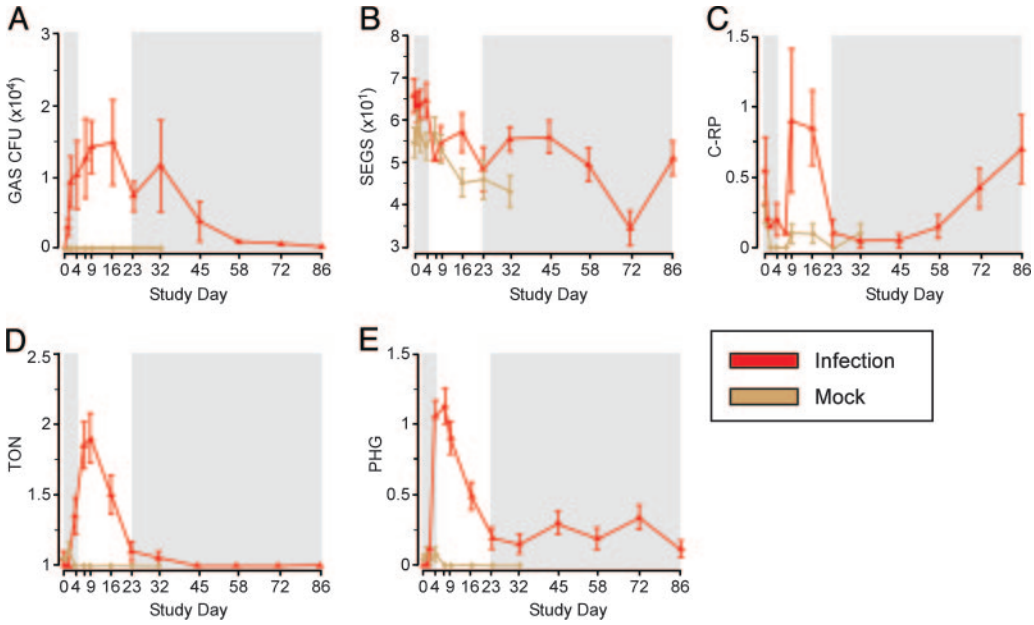

Fig. 1. Five medical parameters used in expression analysis of GAS PHG. Each parameter is depicted by two-way ANOVA plot, with the measurement on the *y* axis and the sample collection day of the study on *x* axis. The infection phase (days 0–86) is marked with a red line, and the mock phase (days 0–32) is labeled brown. SEM is indicated by error bars. (*A*) GAS colony count as measured from tonsillar throat swabs (CFU). (*B*) Percentage of polymorphonuclear leukocytes in peripheral venous blood (SEGS). (*C*) C-RP levels in blood (mg/liter; C-RP). (*D*) TON score on a scale of 1–4. (*E*) PHG severity score on a scale of 0–3 (PHG). Colonization and asymptomatic carriage phases are indicated by gray shading.

web site) and correlate gene expression patterns with clinical findings. Parallel change of the transcript abundance of a GAS gene with a medical parameter resulted in a positive correlation ($+Z$), whereas an inverse transcript level change relative to a medical parameter produced a negative correlation ($-Z$). Representative GAS genes with Z-scores of $+2$ or -2 or higher are shown in Table 1, and they are shown in entirety in Tables 2–7, which are published as supporting information on the PNAS web site. Key results were confirmed by TaqMan RT-PCR analysis, as described extensively in *Supporting Materials and Methods*.

Results and Discussion

Clinical Observations and Laboratory Parameters. The mock- and GAS-infected animals differed significantly (Bonferroni correction, $P < 0.05$) in 13 of the 32 medical parameters that we studied (Table 8, which is published as supporting information on the PNAS web site). GAS CFUs, percentage of segmented neutrophils (SEGS), C-reactive protein (C-RP) level, tonsillitis (TON) score, and PHG score followed a human-like course of clinical disease over time, which is an important observation because these parameters are classic markers of GAS infection (Fig. 1 and Figs. 5 and 6, which are published as supporting information on the PNAS web site).

Three distinct disease phases were observed during the 86-day protocol. Days 0–4, termed the colonization phase, were noteworthy in part because 90% of the monkeys were GAS culture-positive but had only mild tonsil inflammation (Fig. 1 *A* and *D*). Days 5–23 were designated as the acute phase because key infection indicators (such as C-RP, TON, PHG values, and GAS CFUs) peaked during this period (Fig. 1 *A* and *C–E*). Days 23–86 (asymptomatic phase) were characterized by decreased GAS CFUs and declines in 8 of the 13 other medical parameters (Figs. 1, 5*A*, and 6 *B* and *C*).

Overview of Transcriptome Changes. GAS genes with positive or negative expression correlation (Z scores) to changes in CFUs (Fig. 2) and other medical parameters that differed significantly between the mock- and GAS-infected animals (see Fig. 7, which is published as supporting information on the PNAS web site) were grouped into 17 annotation categories and analyzed for trends. Categories for carbohydrate metabolism, information processing, membrane transport, and prophage genes had the greatest changes in numbers throughout the experiment (Figs. 2 and 7). Carbohydrate-metabolism genes with increased expression were most abundant during the colonization phase, when CFU counts were low (Fig. 2). These results indicated that transcription and total numbers of these genes in this category increased when cell densities were low, suggesting that carbohydrate metabolism is a key contributor to initial growth and colonization. When CFU counts were high during this phase, the phage gene category had the greatest number of genes expressed, implying the existence of a relationship between early high cell densities and prophage gene expression.

The carbohydrate-metabolism gene category peaked again in the acute phase of disease, and this rise in numbers of expressed genes correlated with an increased host inflammatory response ($+Z$) (Fig. 7). In addition, information-processing genes and prophage genes had the greatest increase in numbers when their increase in expression correlated with increasing CFU levels on day 16. Increased phage gene expression again was connected to high cell densities, similar to the colonization phase.

In the asymptomatic phase of disease, high numbers of membrane transport genes and information-processing genes were expressed when CFUs were low ($-Z$) (days 23 and 45) (Fig. 2) but only when host inflammatory responses also were high ($+Z$) (Fig. 7). Also, high numbers of prophage genes were expressed during maximum C-RP and SEGS levels (Fig. 7), again emphasizing a connection between inflammation and prophage gene expression.

Pyrogenic Toxin Superantigens (PTSAgs). A key discovery was that genes encoding three GAS PTSAgs (*speA2*, *speJ*, and *smeZ*) were highly expressed in distinct disease phases (Tables 1 and 2), and this expression correlated with inflammation. PTSAgs are po-

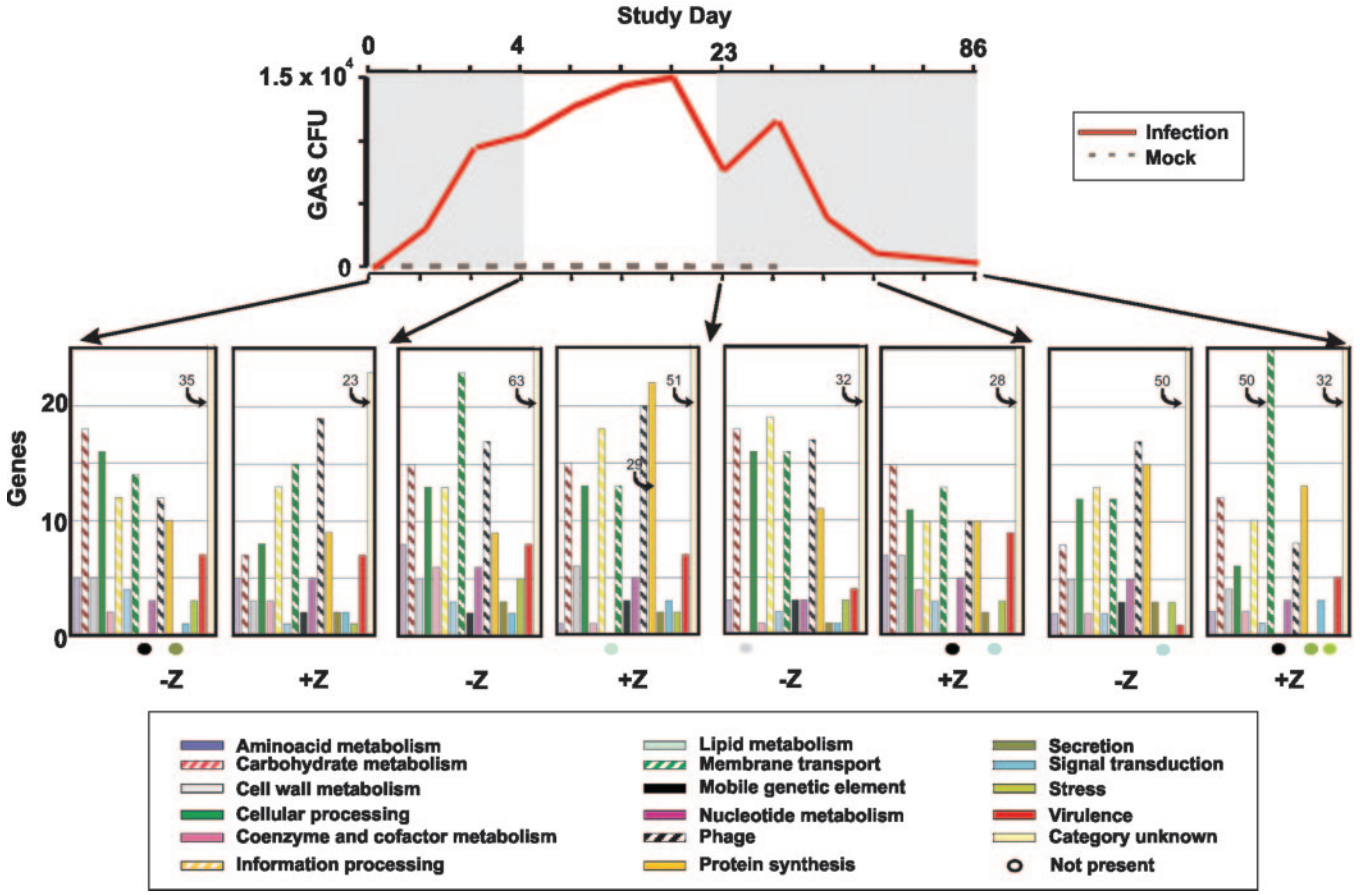

Fig. 2. Functional category analysis of GAS gene expression data. GAS CFU were plotted on the *y* axis, and the study day was plotted on the *x* axis. The infection phase (days 0–86) is indicated by a red line, and the mock phase (days 0–32) is labeled in brown. Number of GAS genes belonging to 17 functional categories were plotted based on their correlation with GAS CFU. Genes having positive (+*Z*) or negative (−*Z*) expression correlation with GAS CFU are shown by histograms at three phases. Colonization and asymptomatic carriage phases are indicated by gray shading.

tent T cell mitogens that stimulate massive release of inflammatory cytokines, a process that can rapidly overwhelm the host. Expression of *speA2*, *speJ*, and *smeZ* was up-regulated in the colonization phase (Tables 1 and 2). *speJ* was expressed the earliest of these three PTSAg (Tables 1 and 2), and consistent with this early expression, increased expression of *speJ* correlated very strongly with low GAS CFUs (−*Z*). Up-regulation of *smeZ* and *speA2* expression quickly followed increased *speJ* expression, but in contrast to *smeZ* and *speA2*, expression of *speJ* correlated (+*Z*) with increasing CFUs (Tables 1 and 2 and Fig. 8 *C* and *D*, which is published as supporting information on the PNAS web site). TaqMan analysis was used to analyze changes more precisely in *speA2* and *smeZ* transcript levels. The *smeZ* transcript level was 24 times higher than the *speA2* level (Fig. 8 *C* and *D*), a very important finding given that SmeZ has been reported to be at least 10 times more potent than other GAS PTSAgs in the ability to stimulate cytokine release by human T cells (29). Together, the data suggest that these three PTSAgs act sequentially to assist colonization, with *speJ* contributing early when GAS densities are low and *smeZ* and *speA2* contributing preferentially to a rapid increase in CFUs.

smeZ expression also correlated (+*Z*) with increasing GAS CFUs in the acute phase of disease (+*Z*) (Table 1). Of note, high *smeZ* expression identified by microarray analysis correlated with peak levels of C-RP (days 9–16), which is a sensitive indicator of inflammation (Table 1). In fact, *smeZ* was the most dominant acute-phase-correlated proinflammatory gene (+*Z*) identified in the study (Table 1). Interestingly, the level of *smeZ* expression did not correlate with peak increases of PHG (days 4–7) and TON (days 7–9) (Tables 1, 5, and 6). We believe that this result suggests a delay between measurable changes in *smeZ* expression and observable downstream clinical effects. Together, the data suggest that SmeZ, is an important mediator of acute-phase inflammation in this model.

The expression of *speJ*, *speA2*, and *smeZ* (as assessed by microarray and confirmed by TaqMan) (Fig. 8 *C* and *D*) was diminished in the asymptomatic phase, resulting in reduced correlation with clinical parameters. The PTSAg-mediated cytokine storm that can result in cell damage, capillary leakage and shock, and altered innate and acquired immune responses may contribute to increasing GAS CFUs counts and, thereby, assist colonization. In addition, expression of *smeZ* and *speA2* appears to be integrally related to peak inflammation. A decrease in expression of these three PTSAgs after the acute phase (Fig. 8 *C* and *D*) correlated with diminished inflammation (Fig. 1), thus potentially contributing to the asymptomatic phase of disease.

Prophages and Prophage Gene Expression. Coculture of GAS with human pharyngeal cells induces phage production (reviewed in ref. 23). However, little is known about induction of GAS prophages *in vivo* or their contribution to pathogenesis. Throughout the course of infection, increased expression of bacteriophage structural genes and their associated proven or putative virulence genes correlated with increasing GAS CFUs (+*Z*) (Fig. 2 and Tables 1 and 2). More specifically, expression of prophage 5005.1 structural genes and *speA2* was positively

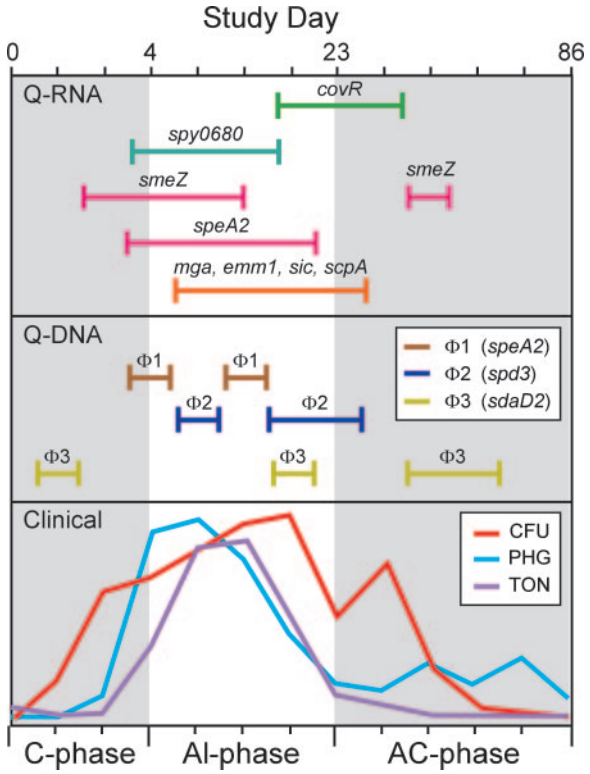

Fig. 3. TaqMan RNA quantitative PCR (Q-PCR) confirmation of expression of GAS SAgs involved in inflammation (*smeZ*, *speA2*, and *speJ*), complement resistance regulon (*mga*, *emm1*, *sic*, and *scpA*), and TCS regulation (*covR* and *spy0680/M1_spy0874*). Induction of strain MGAS5005 prophage (Φ1, Φ2, and Φ3) as confirmed by TaqMan DNA Q-PCR. Bars indicate time points at which the level of RNA or DNA is at least 1.25-fold or greater than mean expression level. Clinical parameters for GAS CFU (red), PHG (blue), and TON (purple) are shown. Colonization and asymptomatic carriage phases are indicated by gray shading.

correlated (+*Z*) with increasing CFUs in the colonization phase (Table 2). Also, expression of the gene (*sdaD2*) encoding an extracellular DNase virulence factor (25) and prophage 5005.3 structural genes correlated (+*Z*) with increased CFUs (Table 2) in the colonization and acute phase of disease. Although expression of prophage 5005.2 structural genes occurred in all three disease phases, expression of the DNase gene (*spd3*) encoded by this bacteriophage correlated only with low C-RP levels in the acute phase (−*Z*) (Table 4).

The observation that prophage gene expression was correlated with increasing CFUs stimulated us to test the hypothesis that the increased phage gene copy number (induction) was correlated with prophage-encoded virulence gene expression. As assessed by DNA TaqMan, PCR prophages 5005.1, 5005.2, and 5005.3 were maximally induced when CFUs were peaking (late colonization and acute phase) (Fig. 3 and Fig. 9, which is published as supporting information on the PNAS web site). Linear regression analysis revealed that all three prophages had a positive association between phage copy number and phage encoded virulence gene expression (*speA2*, *spd3*, and *sdaD2*). These data suggest that as gene copy number increased because of phage replication, virulence gene expression also increased. The increase in DNA copy number of prophage 5005.1 and the concomitant increase in *speA2* gene expression were the most significantly correlated phage-virulence factor pair (correlation, 0.619). Coordinated expression and induction of distinct phages and their extracellular virulence factor genes provides evidence supporting the hypothesis that prophages contribute significantly to the life cycle of GAS *in vivo* by increasing survival and altering the host immune responses in the posterior pharynx.

Two-Component Regulatory System Gene Expression. GAS has 12 conserved two-component gene regulatory systems (TCS) (11–16). Expression of two of these 12 TCS (*covR/S* and *M5005_spy0680/0681/M1_spy0874/0875*) correlated significantly with all three phases of disease. Expression of *covR/S* was nearly absent or very low in the colonization phase (0–4 days), and it gradually increased over time (Tables 1 and 2, and Fig. 8*A*). Importantly, expression of *covR*-regulated virulence genes such as *ska* and *sda* was high in this phase, consistent with the known role of this TCS as a negative regulator (directly or indirectly) of a large number of virulence genes (Table 1) (18, 22). Increased *covR* expression was negatively correlated (−*Z*) with GAS CFUs during the acute phase (Tables 1 and 2). In other words, *covR* expression was high on day 23, when GAS CFUs decreased significantly (Figs. 1 and 8*A*, and Tables 1 and 2). Similarly, the peak of expression of *covR* on day 23 (Fig. 8*A*) correlated with decreased expression of known *covR*-regulated virulence genes, such as *speB* and *mac* (18, 22). *covR* expression was inversely correlated with CFUs and levels of virulence gene expression in the transition from the acute to the asymptomatic phase. Together, these results suggest that there is an integral relationship between *covR/S* expression and GAS CFUs in the posterior pharynx, which is consistent with *covR/S* being a crucial TCS in PHG.

M5005_spy0680/0681(*M1_spy0874/0875*) is a TCS of unknown function that is not regulated directly by *covR/S* (18, 22). In contrast to the *covR/covS* expression and correlation data, expression of *M5005_spy0680/0681*(*M1_spy0874/0875*) was positively correlated (+*Z*) with increasing GAS CFUs during the colonization phase and the asymptomatic phase and negatively correlated (−*Z*) with increasing GAS CFUs in the acute phase (Tables 1 and 2 and Fig. 8*B*). The microarray results and TaqMan confirmation data (Table 2 and Fig. 8 *A* and *B*) suggest that a biphasic or inverted expression relationship exists between these two regulators over time. Together, the data lead us to propose that *covR* and *M5005_spy0680*(*M1_spy0874*) are key TCS regulators acting in concert to aid GAS in establishing or prolonging GAS–host interactions in different phases of disease.

Virulence-Related Genes. Of the 70 ORFs annotated as virulence genes in the GAS genome, 45 had increased expression correlated with either severe (+*Z*) or reduced (−*Z*) PHG in monkeys. For example, increased expression of genes involved in enhanced survival, dissemination, and inhibition of PMN recruitment (*ska*, *sda*, *sdaD2*, and *scpA*), as well as induction of cytokines (*smeZ*, *speA2*, and *speJ*), was correlated with increasing GAS CFUs in the colonization phase (+*Z*) (Tables 1 and 2 and Figs. 8 and 10, which is published as supporting information on the PNAS web site). The three virulence genes that are most highly expressed in the acute phase (*emm1*, *scpA*, and *sic*) are all regulated by *mga* and assist evasion of host defenses (9) (Table 1 and Fig. 10). As GAS CFUs increased, the expression of genes encoding two Ig proteases (SpeB and Mac) that inhibit phagocytosis increased (+*Z*) in the acute and asymptomatic phase, respectively (Tables 1 and 2). SpeB may be an especially important inhibitor of the early host immune response because it degrades most Ig isotypes, including IgM, whereas Mac specifically cleaves IgG and hence, this protein may preferentially target the acquired host immune response (30–32). The expression of genes (*emm1*, *sclA*, and *inlA*) encoding cell-surface-anchored proteins linked to promoting bacterial adhesion to host molecules was correlated with increased CFUs and inflammation in the asymptomatic phase (+*Z*) (Tables 1, 2, and 4). One hemolysin gene (*hlyX*) was down-regulated (Table 1 and Table 9, which is published as supporting information on the PNAS web site), and two genes

encoding proteins associated with acute rheumatic fever, cardiolipin synthetase (*M5005_spy0962/M1_spy1252*), and myosin cross-reactive antigen (*M5005_spy0385/M1_spy0470*) (9, 33) were up-regulated in the asymptomatic phase of infection (Tables 1 and 9 and Fig. 11*C*, which is published as supporting information on the PNAS web site). Together, the data indicate that virulence gene expression has a critical role in all three phases of disease with respect to inflammation and increasing CFU counts.

Concluding Comment. Our longitudinal analysis of the GAS transcriptome over 86 days of host–pathogen interaction in a primate model of human PHG has provided extensive information about temporal changes in gene expression in each phase of clinical disease. A key discovery was that the temporal pattern of GAS gene expression in PHG is tightly linked to three distinct phases of infection. Identification of the pathogen gene expression patterns characterizing each phase of pathogen–host interaction provides avenues of targeted investigation of proven and putative virulence factors and genes of unknown function. Given that every infectious agent has a complex natural history, longitudinal transcriptome analysis of other microbes will undoubtedly provide insight into the genetic events contributing to pathogenesis.

We thank T. G. Schwan, N. Hoe, R. A. Heinzen, and B. J. Hinnebusch for critical review of the manuscript; W. Sheets, D. Dale, R. Larson, R. Rivera, B. Jameson, and D. Myers for assistance with the nonhuman primate experiments; G. Hettrick and A. Mora for graphic assistance; A. Henion and I. Richichi for sample tracking; and R. A. Lempicki, C. Smith, and J. Yang for GeneChip hybridizations. Partek, Inc. assisted with experimental design, and Enodar Biometric Corporation helped with probe-performance analysis (PPA) and correlation analysis. This work was supported in part by National Institutes of Health Grant UO1-AI-060595 (to J.M.M.).

1. Merrell, D. S., Butler, S. M., Qadri, F., Dolganov, N. A., Alam, A., Cohen, M. B., Calderwood, S. B., Schoolnik, G. K. & Camilli, A. (2002) *Nature* **417,** 642–645.
2. Bina, J., Zhu, J., Dziejman, M., Faruque, S., Calderwood, S. & Mekalanos, J. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 2801–2806.
3. Xu, Q., Dziejman, M. & Mekalanos, J. J. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1286–1291.
4. Talaat, A. M., Lyons, R., Howard, S. T. & Johnston, S. A. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 4602–4607.
5. Snyder, J. A., Haugen, B. J., Buckles, E. L., Lockatell, C. V., Johnson, D. E., Donnenberg, M. S., Welch, R. A. & Mobley, H. L. T. (2004) *Infect. Immun.* **72,** 6373–6381.
6. Ylöstalo, J., Randall, A. C., Myers, T. A., Metzger, M., Krogstad, D. J. & Cogswell, F. B. (2005) *J. Infect. Dis.* **191,** 400–409.
7. Shelburne, S. A. & Musser, J. M. (2004) *Curr. Opin. Microbiol.* **7,** 283–289.
8. Boyce, J. D., Cullen, P. A. & Adler, B. (2004) *Emerging Infect. Dis.* **10,** 1357–1362.
9. Cunningham, M. W. (2000) *Clin. Microbiol. Rev.* **13,** 470–511.
10. Musser, J. M. & Krause, R. M. (1998) in *Emerging Infections*, ed. Krause, R. M. (Academic, New York), pp. 185–218.
11. Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A. N., Kenton, S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4658–4663.
12. Smoot, J. C., Barbian, K. D., Van Gompel, J. J., Smoot, L. M., Chaussee, M. S., Sylva, G. L., Sturdevant, D. E., Ricklefs, S. M., Porcella, S. F., Parkins, L. D., *et al.* (2002) *Proc. Natl. Acad. Sci. USA* **99,** 4668–4673.
13. Beres, S. B., Sylva, G. L., Barbian, K. D., Lei, B., Hoff, J. S., Mammarella, N. D., Liu, M.-Y., Smoot, J. C., Porcella, S. F., Parkins, L. D., *et al.* (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10078–10083.
14. Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunga, T., *et al.* (2003) *Genome Res.* **13,** 1042–1055.
15. Banks, D. J., Porcella, S. F., Barbian, K. D., Beres, S. B., Philips, L. E., Voyich, J. M., DeLeo, F. R., Martin, J. M., Somerville, G. A. & Musser, J. M. (2004) *J. Infect. Dis.* **190,** 727–738.
16. Green, N. M., Zhang, S., Porcella, S. F., Barbian, K. D., Beres, S. B., LeFebvre, R. B. & Musser J. M. (2005) *J. Infect. Dis.*, in press.
17. Smoot, L. M., Smoot, J. C., Graham, M. R., Somerville, G. A., Sturdevant, D. E., Lux Migliaccio, C. A., Sylva, G. L. & Musser, J. M. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10416–10421.
18. Graham, M.R., Smoot, L. M., Lux Migliaccio, C. A., Virtaneva, K., Sturdevant D. E., Porcella, S., Federle M. J., Adams, G. J., Scott, J. R. & Musser, J. M. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 13855–13860.
19. Voyich, J. M., Sturdevant, D. E., Braughton, K. R., Kobayashi, S. D., Lei, B., Virtaneva, K., Dorward, D. W., Musser, J. M. & DeLeo, F. R. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1996–2001.
20. Voyich, J. M., Braughton, K. R., Sturdevant, D. E., Vuong, C., Kobayashi, S. D., Porcella, S. F., Otto, M., Musser, J. M. & DeLeo, F. R. (2004) *J. Immunol.* **173,** 1194–1201.
21. Beres, S. B., Sylva, G. L., Sturdevant, D. E., Granville, C. N., Liu, M., Ricklefs, S. M., Whitney, A. R., Parkins, L. D., Hoe, N. P., Adams, G. J., *et al.* (2004) *Proc. Natl. Acad. Sci. USA* **101,** 11833–11838.
22. Graham, M. R., Virtaneva, K., Porcella, S. F., Barry, W. T., Gowen, B. B., Johnson, C. R., Wright, F. A. & Musser, J. M. (2005) *Am. J. Pathol.* **166,** 455–465.
23. Banks, D. J., Beres, S. B. & Musser, J. M. (2005) in *Bacteriophages and Bacterial Pathogens*, eds. Waldor, M., Friedman, D. & Adhya, S. (American Soc. Microbiol., Washington, DC), in press.
24. Virtaneva, K., Graham, M. R., Porcella, S. F., Hoe, N. P., Su, H., Graviss, E. A., Gardner, T. J., Allison, J. E., Lemon, W. J., Bailey, J. R., *et al.* (2003) *Infect. Immun.* **71,** 2199–2207.
25. Ashbaugh, C. D., Moser, T. J., Shearer, M. H., White, G. L., Kennedy, R. C. & Wessels, M. R. (2002) *Cell Microbiol.* **2,** 283–292.
26. Gryllos, I., Cywes, C., Shearer, M. H., Cary, M., Kennedy, R. C. & Wessels, M. R. (2001) *Mol. Microbiol.* **42,** 61–74.
27. Sumby, P., Barbian, K. D., Gardner, D. J., Whitney, A. R., Welty, D. M., Long, R. D., Bailey, J. R., Parnell, M. J., Hoe, N. P., Adams, G. G., *et al.* (2005) *Proc. Natl. Acad. Sci. USA* **102,** 1679–1684.
28. Sumby, P., Porcella, S. P., Madrigal, A. G., Barbian, K. D., Virtaneva, K., Ricklefs, S. M., Sturdevant, D. E., Graham, M. R., Vuopio-Varkila, J., Hoe, N. P. & Musser, J. M. (2005*) J. Infect. Dis.*, in press.
29. Müller-Alouf, H., Proft, T., Zollner, T. M., Gerlach, D., Champagne, E., Desreumaux, P., Fitting, C., Geoffroy-Fauvet, C., Alouf, J. E. & Cavaillon, J. M. (2001) *Infect. Immun.* **69,** 4141–4145.
30. Collin, M. & Olsén, A. (2001) *Infect. Immun.* **69,** 7187–7189.
31. Lei, B., DeLeo, F. R., Hoe, N. P., Graham, M. R., Mackie, S. M., Cole, R. L., Liu, M., Hill, H. R., Low, D. E., Federle, M. J., *et al.* (2001) *Nat. Med.* **7,** 1298–1305.
32. von Pawel-Rammingen, U., Johansson, B. P. & Bjorck, L. (2002) *EMBO J.* **21,** 1607–1615.
33. Dzhuzenova, B. S., Nasonov, E. L., Kovalev, VIu., Lopaeva, O. V., Speranskii, A. I. & Nasonova, V. A. (1992) *Klin. Med.* **70,** 66–71.