Microarray-Based Identification of *htrA*, a *Streptococcus pneumoniae* Gene That Is Regulated by the CiaRH Two-Component System and Contributes to Nasopharyngeal Colonization

M. E. Sebert,[1] L. M. Palmer,[2] M. Rosenberg,[2] and J. N. Weiser[3]*

Department of Pediatrics, Children's Hospital of Philadelphia,[1] and Department of Microbiology, University of Pennsylvania,[3] Philadelphia, Pennsylvania 19104, and Antimicrobials and Host Defense CEDD, GlaxoSmithKline Pharmaceuticals, Collegeville, Pennsylvania 19426[2]

Nasopharyngeal carriage is the reservoir from which most disease with *Streptococcus pneumoniae* arises. Survival as a commensal in this environment is likely to require a set of adaptations distinct from those needed to cause disease, some of which may be mediated by two-component signal transduction systems (TCSTS). We examined the contributions of nine pneumococcal TCSTS to the process of nasopharyngeal colonization by using an infant rat model. Whereas deletions in all but one of these systems have been associated previously with a high degree of attenuation in a murine model of pneumonia, only the CiaRH system was necessary for efficient carriage. Transcriptional analysis by using microarray hybridization identified a locus consisting of two adjacent genes, *htrA* and *spoJ*, that was specifically and strongly downregulated in a Δ*ciaRH*-null mutant. A *S. pneumoniae* strain lacking the *htrA* gene encoding a putative serine protease, but not one lacking *spoJ*, showed decreased fitness in a competitive model of colonization, a finding consistent with this gene mediating a portion of the carriage deficit observed with the Δ*ciaRH* strain.

Streptococcus pneumoniae exists in nature both as a commensal on the mucosal surface of the human upper respiratory tract and as a pathogen causing disease in a wide range of sites, including the lung, middle ear, sinuses, blood, and meninges. Survival in these different host environments is likely to require adaptive responses, some of which may be mediated by two-component signal transduction systems (TCSTS). These signaling systems (reviewed in reference 27) are found in a wide range of bacteria where they have been shown to regulate diverse processes, including chemotaxis, nutrient utilization, surface adhesion, and the switch between aerobic and anaerobic metabolism. TCSTS consist of a sensor kinase that autophosphorylates in response to an environmental stimulus and a cognate response regulator to which an activated phosphate is transferred and that then mediates a downstream response, often acting as a DNA-binding protein to cause changes in gene expression.

Two genomic surveys of *S. pneumoniae* have identified 13 pairs of genes encoding putative TCSTS as well as a single unpaired putative response regulator (12, 29). The ability of null mutants in these systems to cause disease has been previously examined in two studies of virulence. In a murine model of pneumonia developing after intranasal inoculation in which all of the TCSTS except ComDE were tested, deletions in eight of the systems resulted in an attenuation of at least 3 orders of magnitude (29). In contrast, a second study testing a collection of single-crossover mutants in a different genetic background showed no evidence of decreased virulence for any of the strains in a murine model of intraperitoneal infection in which

the requirements of adaptation to the mucosal surface of the airway were circumvented (12). Despite the importance of nasal carriage both as the first step in the pathogenesis of pneumococcal disease and in the persistence of the organism in the community, the role of *S. pneumoniae* two-component systems in mucosal colonization of the nasopharynx has not been described.

Although experimental characterization of the pneumococcal TCSTS has been an active area of recent investigation, the current understanding of the physiological role of TCSTS in *S. pneumoniae* is incomplete. Best described among these systems is ComDE, which activates the genetic competence system in response to competence-stimulating peptide (CSP) (3, 21). A second system, CiaRH, appears to operate upstream of the ComDE pathway and has been shown to modulate competence expression but is also associated with other characteristics such as resistance to the β-lactam antibiotic cefotaxime (6, 8). Other pneumococcal TCSTS have been linked to a coordinated cell-death pathway (17, 18) and to the regulation of a bacteriocin-like peptide locus (5).

Because the nasopharynx is the site where the pneumococcus resides as a commensal and serves as the primary reservoir from which cases of pneumococcal disease arise, we sought to determine the role of *S. pneumoniae* TCSTS and of the genes whose expression they control in sensing and adapting to this environment. We selected for this study the eight TCSTS in which mutants had yielded the most marked attenuation in the previously reported pneumonia model (29), as well as the cognate histidine kinase of the one response regulator that had been reported to be essential. Our study examined the ability of null mutants in these *S. pneumoniae* TCSTS to persist during nasopharyngeal carriage by using an infant rat model of colonization where carriage is maintained for up to several weeks without the development of either localized or invasive

---

* Corresponding author. Mailing address: 402A Johnson Pavilion, Department of Microbiology, University of Pennsylvania, Philadelphia, PA 19104-6076. Phone: (215) 573-3511. Fax: (215) 898-9557. E-mail: weiser@mail.med.upenn.edu.

disease. After the identification of one TCSTS required for nasopharyngeal colonization, microarray gene expression data were used to identify a set of genes whose expression was altered in a mutant lacking this TCSTS. Finally, mutational analysis was used to provide independent confirmation of the role of one of these genes in the colonization process.

MATERIALS AND METHODS

Bacterial strains and mutagenesis. Experiments were performed with *S. pneumoniae* 0100993, a type 3 clinical isolate from the United Kingdom (29). Mutants in the pneumococcal TCSTS were generated by allelic replacement with a constitutively expressed *ermAM* cassette as previously described (29). In order to ensure that the deletion mutants and parent strain were otherwise isogenic, chromosomal DNA was extracted from these strains by the method of Pearce et al. (20) and was transformed back into the parent 0100993 strain by a previously described modification (32) of the method of Lacks and Hotchkiss (11) by using selection for erythromycin resistance. Strain 0100993-SR contains a spontaneous mutation conferring streptomycin resistance that was backtransformed into the *S. pneumoniae* 0100993 background.

In vitro growth rates were measured in C+Y medium (pH 6.8) (11) from absorbance readings obtained at 620 nm for individual strains and from colony counts on selective media for mixed cultures.

Nasopharyngeal colonization assays. Bacteria were prepared for inoculation by growth in C+Y medium (pH 6.8) to an optical density at 620 nm (OD_{620}) of 0.3. Randomized litters consisting each of 10 to 12 newborn Sprague-Dawley rats (Taconic, Germantown, N.Y.) were inoculated intranasally with 4×10^6 to 1.6×10^7 CFU of each strain, and nasopharyngeal colonization was monitored as previously described (31). Each animal's nasopharynx was washed by instillation of 20 to 40 μl of phosphate-buffered saline into the left naris, followed by the recovery of fluid expelled from the opposite naris after passage through the nasopharynx. Nasal washes were performed over a 10-day period after inoculation, and the density of organisms in the recovered fluid was determined by colony counts on serial dilutions of the washes plated on tryptic soy agar (Becton Dickinson, Sparks, Md.) containing catalase (3,000 U; Worthington Biochemicals, Freehold, N.J.), as well as neomycin (20 μg/ml), to limit the growth of other organisms. The data for each litter are represented as geometric means ±1 standard error of the mean. All procedures were performed in accordance with institutional animal care guidelines.

For competitive assays of nasopharyngeal colonization, the two strains were grown separately in vitro to an OD_{620} of 0.3 and then mixed 1:1 prior to intranasal inoculation. Serial nasal washings were then plated on selective media containing either erythromycin (1 μg/ml) or streptomycin (200 μg/ml) in addition to neomycin (20 μg/ml) to obtain colony counts for each strain individually.

RNA purification and labeling. RNA was extracted from cultures grown in TH-Y broth to OD_{600} of 0.3 by a modification of the method of Chuang et al. (4). Briefly, bacteria were collected and washed by centrifugation at 4°C and resuspended in ice-cold phosphate-buffered saline before lysis with an equal volume of medium containing 0.4 M NaCl, 40 mM EDTA, 1% 2-mercaptoethanol, 1% sodium dodecyl sulfate, and 20 mM Tris-HCl (pH 7.5) plus a 0.17 volume of buffer-saturated phenol. The mixture was boiled 40 s and centrifuged to remove cellular debris. The supernatant was extracted four times with 25:24:1 phenol-chloroform-isoamyl alcohol (IAA) and once with 24:1 chloroform-IAA. RNA was precipitated with isopropyl alcohol, and the pellet was washed twice with 70% ethanol. RNA was then treated with DNase I (GenHunter, Nashville, Tenn.) according to the manufacturer's protocol, reextracted twice with phenol-chloroform-IAA and once with chloroform-IAA, and then reprecipitated with sodium acetate and ethanol before use.

cDNA was prepared by reverse transcription (RT) with 2 μg of total RNA for each hybridization. RNA was mixed with 1 μg of random nonamers (1 μg/μl) and water to a volume of 11 μl, heated to 70°C for 10 min, and chilled on ice 1 min. Then, 2 μl of 10× PCR buffer (Invitrogen, Carlsbad, Calif.), 2 μl of 25 mM MgCl₂, 2 μl of 0.1 M dithiothreitol, 1 μl of a deoxynucleoside triphosphate mix (containing 10 mM concentrations each of dATP, dGTP, and dTTP, plus 5 mM dCTP), and 1 μl of Cy3-dCTP (Amersham Pharmacia Biotech, Piscataway, N.J.) were added, and the mixture was preheated to 42°C for 5 min. The sample was then incubated with 1 μl of (200 U) SuperScript II reverse transcriptase (Invitrogen) at 42°C for 120 min, with the addition of an additional 1 μl of reverse transcriptase after 60 min. The sample was then heated to 70°C for 15 min prior to digestion with 1 μl of (2 U) RNase H at 37°C for 30 min. cDNA was purified by two successive passes through a Qiagen (Valencia, Calif.) QiaQuick PCR purification column according to the manufacturer's protocol.

Microarray analysis. *S. pneumoniae* open reading frames (ORFs) were identified from partial genomic sequence, available as of 1998, including proprietary sequence data generated for strain 0100993 by SmithKline Beecham and for strain R6 by Incyte, and publicly accessible sequence data for the strain TIGR4 from The Institute for Genomic Research. The identified ORFs were filtered to reduce redundancy, resulting in a set of 1,974 ORFs. Primer pairs were designed to produce 150- to 600-bp products by using the Primer3 program from the Whitehead Institute (http://www-genome.wi.mit.edu/). A total of 1,855 PCRs produced products of the predicted size. Comparison of these ORFs with the published TIGR4 pneumococcal sequence (28) reveals that 1,712 of the primer pairs correspond to 1,481 ORFs in the published sequence. Of these ORFs, 36 were also represented by 59 additional PCR products amplified from different regions of the predicted gene sequences in order to monitor the consistency of the results. A total of 110 spots containing heterologous eukaryotic DNA, yeast tRNA, and bacteriophage DNA were used as negative controls.

Microarrays were generated by using PCR products resuspended in 6 M sodium thiocyanate and deposited onto silanized glass slides by a Molecular Dynamics generation III Arrayer (Sunnyvale, Calif.). Array elements were printed in duplicate on each slide. Slides were baked in an 80°C vacuum oven overnight and rinsed in 100% isopropanol for 10 min, followed by boiling in water for 5 min before use. The arrays were first prehybridized with 70 μl of a solution containing 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.1% SDS, 0.25 μg of yeast tRNA/μl, 0.25 μg of salmon testes DNA/μl, and 0.25 μg of calf thymus DNA/μl for at least 4 h at 55°C. Cy3-labeled cDNA from RT of 2 μg of total RNA was suspended in 40 μl of prehybridization solution, boiled 5 min, and hybridized to the array for 16 to 20 h at 60°C. Slides were washed to a maximum stringency of 0.2× SSC at 55°C before being scanned with a confocal fluorescence scanner (Molecular Dynamics). Array images were analyzed by using the Autogene software package (BioDiscovery, Marina Del Rey, Calif.).

Hybridizations were done individually by using a single-color labeling scheme, and fluorescence values were normalized to the median fluorescence intensity of the entire set of ORFs on the array. Microarray data were collected for two independent RNA samples for each TCSTS mutant and for three independent wild-type RNA samples.

TaqMan quantitative RT-PCR. Total RNA isolated from *S. pneumoniae* strains 0100993 and Δ*ciaRH* for the microarray analysis was further processed by additional DNA removal and RT with commercial kits according to the manufacturers' instructions (DNA-*free* [Ambion, Austin, Tex.] and SuperScript Preamplification System for First Strand cDNA Synthesis [Life Technologies, Gaithersburg, Md.]). For each RNA sample, duplicate RT reactions were performed, as well as a control without reverse transcriptase, in order to determine the levels of DNA contamination.

PCRs were set up in triplicate by using TaqMan PCR Master Mix (Applied Biosystems, Foster City, Calif.) according to the instructions provided. Real-time sequence-specific detection and relative quantitation were performed with the ABI PE Sequence Detection System. Relative quantitation of *S. pneumoniae* *htrA* and *spoJ* cDNA, normalized to *S. pneumoniae* *era* as an endogenous control, allowed standardization of sample to sample variations in starting cDNA concentrations. Forward and reverse primers (5′-GCCATCGGTAGCCCGTT-3′ and 5′-TTTAAGGATACATTTCTATTGAGACTGGA-3′, respectively) were designed to amplify an 86-bp fragment of the *S. pneumoniae* *htrA* gene and (5′-TTCAACCGATTATTGTTCGTCAA-3′ and 5′-GTGAAGCCCGATAGC GTCTC-3′, respectively) a 78-bp fragment of the *S. pneumoniae* *spoJ* gene. For the endogenous control, forward and reverse primers (5′-GATTATCGAGCGT CTCAAGGCT-3′ and 5′-GTCTGGATGGACCTTATCGATTTT-3′, respectively) were designed to amplify a 76-bp fragment of the *S. pneumoniae* gene *era*. The corresponding probes—5′-[FAM]-ACGATACCTTGAGTGACAGTATTT GCATATTTCAGAACC-[TAMRA]-3′, 5′-[FAM]-CTCCTGCAAGGATTTCA TAACCAATAACAGGAG-[TAMRA]-3′, and 5′-[FAM]-TCACCACCAAAA TCACAGGAACCTTGG-[TAMRA]-3′—complementary to *S. pneumoniae* *htrA*, *spoJ*, and *era*, respectively, were obtained from Applied Biosystems. Serial dilutions of *S. pneumoniae* 0100993 chromosomal DNA were employed for each probe to generate standard curves.

Computer-assisted sequence analysis. Protein sequence homology searches were performed against the nonredundant protein database by using the BLASTP algorithm available on the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/BLAST/). Alignments of the pneumococcal HtrA and SpoJ predicted proteins with homologues from other bacterial species were performed by using the CLUSTALW algorithm implemented in the MacVector software package (Oxford Molecular). Protein domain searches were carried out by using the SMART algorithm (26). Signal sequences
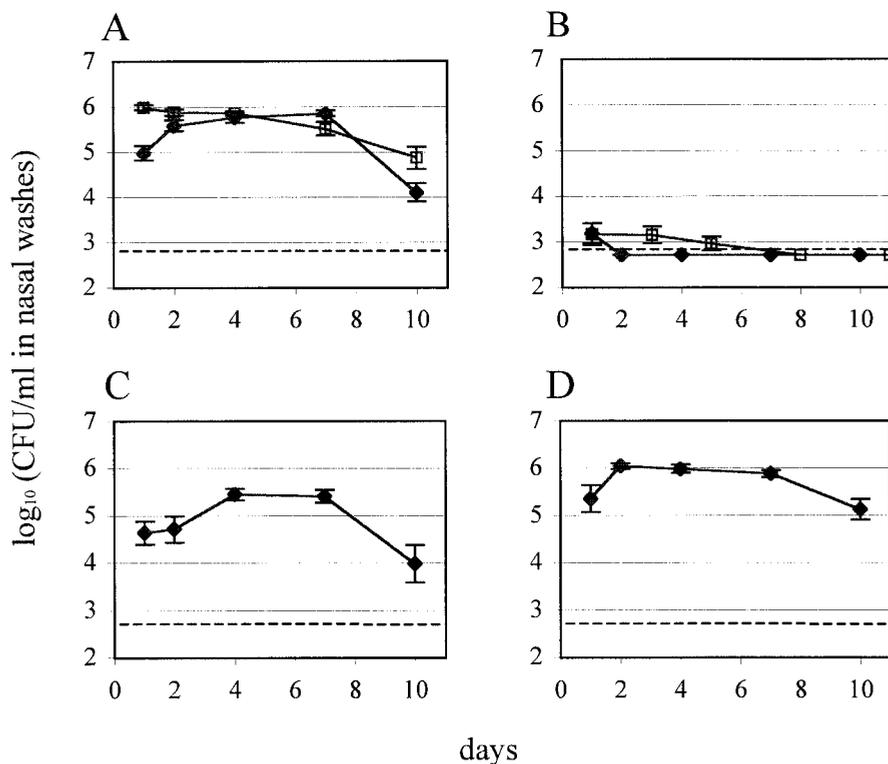
FIG. 1. Recovery of *S. pneumoniae* strains from infant rat nasal washes. (A) 0100993 parent strain (♦) and 0100993-SR (□); (B) Δ*ciaRH* (these two symbols represent two independent experiments); (C) Δ*htrA*; and (D) Δ*spoJ*. Data shown for 0100993 are geometric mean values from independent inoculations of four litters; data for 0100993-SR are derived from inoculation of two litters; and data for all other strains represent inoculations of single, randomized litters consisting of 10 to 12 individuals each. Error bars indicate ±1 standard error of the mean. Dashed lines indicate the lower limit of detection of the colonization assay.

were identified by using the algorithm of von Heijne (30) implemented by PSORT (http://psort.nibb.ac.jp).

RESULTS

Nasopharyngeal carriage by *S. pneumoniae* TCSTS mutants. The ability to establish and maintain nasopharyngeal colonization in the infant rat was assessed for the following eleven *S. pneumoniae* mutants carrying deletions in nine two-component signal transduction systems, as well as for the 0100993 parent strain: Δ*ciaRH*, Δ478HK, Δ480RR, Δ481RR, Δ484RR, Δ486HK, Δ486RR, Δ489RR, Δ492HK, Δ539HK, and Δ539RR. These represent all of the pneumococcal two-component systems that had previously been shown to be substantially attenuated ($\geq 10^3$-fold) in a murine model of pneumonia (29) when mutants in the same genetic background were tested. The Δ492HK strain was included in this survey because, although not attenuated in the pneumonia model, the cognate response regulator of this system has been found to be essential under standard growth conditions (29).

After a period of rising colony counts during the establishment of colonization, the wild-type strain 0100993 reached a maximum density of colonization by day 4 that persisted until day 7 and during which time ca. 10^6 CFU/ml of nasal wash were consistently recovered (Fig. 1A). This plateau was followed by a decline of variable magnitude in the density of colonization on day 10. Of the TCSTS deletion mutants tested,

only the Δ*ciaRH* strain, containing a mutation spanning both the histidine kinase and response regulator genes, showed a significant deficit in colonization (Fig. 1B). Inoculation with this strain resulted in the initial recovery of a small number of colonies, which by the second day after inoculation was below the detection threshold of the assay. Strains carrying mutations in the other eight two-component systems examined showed colonization at levels equal to or greater than the wild type and which persisted for an equivalent duration (data not shown).

The in vitro doubling time of the Δ*ciaRH* strain (36.1 ± 0.4 min) was not significantly different from that of *S. pneumoniae* 0100993 (34.9 ± 2.3 min). Therefore, the attenuation of the Δ*ciaRH* strain in nasal carriage could not be explained by a constitutive defect in growth, although this observation does not exclude the possibility of a selective growth defect in vivo.

Microarray analysis of TCSTS mutants. cDNA microarray analysis of large-scale gene expression patterns in the TCSTS mutant strains was used in an attempt to identify the molecular basis for the colonization deficit of the Δ*ciaRH* strain. In addition to Δ*ciaRH* and the wild-type strain, microarray hybridizations were performed with the Δ478HK, Δ486HK, Δ486RR, Δ492HK, Δ539HK, and Δ539RR strains, none of which showed a colonization deficit.

While the large majority of genes displayed comparable expression levels in the wild type and Δ*ciaRH* deletion mutant, a subset of genes showed significantly altered transcript levels in

TABLE 1. ΔciaRH downregulated and upregulated ORFs

ORF type and ΔciaRH/wt ratio[a]	P[b]	Specific[c]	TIRG4 ORF[d]	TIRG4 annotation[e]
Downregulated ORFs				
0.04 (−1.43)	1.3×10^{-2}	Y	SP2239	Serine protease (htrA[f])
0.09 (−1.06)	3.6×10^{-2}	Y	SP2240	SpoJ protein[g] (spoJ[f])
0.11 (−0.94)	5.2×10^{-3}	Y	SP0798	DNA-binding response regulator CiaR (ciaR)
0.20 (−0.70)	5.5×10^{-3}	N	SP1468	Pyridoxine biosynthesis protein
0.25 (−0.61)	2.7×10^{-3}	N	SP1467	Conserved hypothetical protein
0.30 (−0.53)	3.3×10^{-3}	N	SP1742	Conserved hypothetical protein
0.31 (−0.51)	2.7×10^{-2}	N	SP1771	Glycosyltransferase, family 2/glycosyltransferase family 8
0.32 (−0.49)	3.0×10^{-2}	N	SP1546	Conserved domain protein
0.34 (−0.46)	3.8×10^{-2}	N	SP1027	Conserved hypothetical protein
0.34 (−0.46)	4.5×10^{-2}	Y	SP0799	Sensor histidine kinase CiaH (ciaH)
0.35 (−0.46)	2.1×10^{-3}	N	SP2106	Glycogen phosphorylase family protein
0.35 (−0.46)	1.4×10^{-2}	N	SP0463	Cell wall surface anchor family protein
0.35 (−0.45)	2.6×10^{-2}	Y	SP1268	Choline transporter (licB)
0.35 (−0.45)	1.9×10^{-2}	Y	SP0797	Aminopeptidase N (pepN)[h]
0.36 (−0.44)	6.2×10^{-3}	Y	SP1744	Iojap-related protein
0.37 (−0.44)	7.0×10^{-3}	N	SP1790	ATPase, AAA family
0.40 (−0.40)	4.9×10^{-2}	Y	SP1567	Endoribonuclease L-PSP
0.40 (−0.40)	4.5×10^{-2}	N	SP1555	Dihydrodipicolinate reductase (dapB)
0.41 (−0.39)	7.6×10^{-3}	N	SP0105	L-Serine dehydratase, iron-sulfur-dependent, alpha subunit (sdhA)
0.43 (−0.37)	3.1×10^{-2}	Y	SP0081	Glycosyl transferase, family 2, authentic point mutation
0.43 (−0.37)	3.3×10^{-2}	Y	SP1357	ABC transporter, ATP-binding/permease protein
0.46 (−0.34)	3.5×10^{-2}	Y	SP0855	Topoisomerase IV, subunit A (parC)
0.46 (−0.33)	4.1×10^{-2}	Y	SP0479	Potassium uptake protein, Trk family
0.50 (−0.30)	1.5×10^{-2}	N	SP1212	tRNA pseudouridine synthase B (truB)
Upregulated ORFs				
39.6 (1.60)	1.5×10^{-4}	N	SP0955	Competence protein CelB (celB)
33.0 (1.52)	3.7×10^{-2}	N	SP2047	Conserved domain protein
33.0 (1.52)	6.0×10^{-3}	N	SP2051	Competence protein CglC (cglC)
24.5 (1.39)	3.6×10^{-2}	N	SP2053	Competence protein CglA (cglA)
23.7 (1.38)	4.3×10^{-3}	Y	SP1266	DNA processing protein DprA, putative
19.9 (1.30)	4.1×10^{-2}	N	SP2052	Competence protein CglB (cglB)
19.4 (1.29)	8.7×10^{-3}	N	SP2048	Conserved hypothetical protein
19.4 (1.29)	5.7×10^{-5}	Y	SP2208	Helicase, putative
19.3 (1.29)	1.8×10^{-2}	Y	SP0954	Competence protein CelA (celA)
18.2 (1.26)	2.3×10^{-2}	Y	SP1908	Single-stranded DNA-binding protein, authentic point mutation (ssbB)
10.7 (1.03)	2.9×10^{-2}	Y	SP1088	DNA repair protein RadC (radC)
10.4 (1.02)	1.0×10^{-2}	Y	SP2201	Choline-binding protein D (cbpD)
4.6 (0.66)	3.3×10^{-2}	Y	SP1810	Hypothetical protein
3.9 (0.59)	4.6×10^{-2}	N	SP1941	Competence/damage-inducible protein CinA
3.3 (0.52)	2.1×10^{-2}	Y	SP0791	Oxidoreductase, aldo/keto-reductase family
3.3 (0.52)	3.0×10^{-2}	N	SP2206	Ribosomal subunit interface protein (yfiA)
3.3 (0.51)	4.5×10^{-3}	N	SP0143	Conserved domain protein
2.6 (0.41)	1.1×10^{-3}	Y	SP1937	Autolysin (lytA)[i]
2.5 (0.40)	5.0×10^{-2}	Y	SP0577	PTS system, beta-glucoside-specific IIABC components
2.4 (0.39)	4.4×10^{-2}	Y	SP1051	Conserved hypothetical protein
2.2 (0.34)	2.8×10^{-2}	N	SP2156	SPFH domain/band 7 family
2.0 (0.31)	2.5×10^{-2}	Y	SP2013	Conserved hypothetical protein

[a] ΔciaRH/0100993 wild-type (wt) gene expression ratios as determined by microarray hybridization. The $\log_{10}$(ΔciaRH/wild-type) gene expression ratios are given in parentheses to facilitate comparison with Fig. 2.

[b] P values from individual t tests that reflect the confidence assigned to measured differences in gene expression between the ΔciaRH mutant and the wild-type strain 0100993 are given.

[c] Indicates whether the altered expression of the specified gene in the ΔciaRH mutant was considered to be specific to the ΔciaRH strain as determined by comparison to the expression profiles of the Δ478HK, Δ486HK, Δ486RR, Δ492HK, Δ539HK, and Δ539RR strains as detailed in the text.

[d] Designations of the ORFs in the TIGR4 sequence (28) matching the microarray probes showing differences in expression between ΔciaRH and 0100993.

[e] Published annotations of the TIGR4 ORFs.

[f] Gene designations htrA and spoJ used in this article.

[g] Annotated in TIGR4 sequence as "spspoJ protein."

[h] This ORF terminates 111 nucleotides upstream of ciaRH and is transcribed in the same direction in the TIGR4 sequence (28). Its decreased expression in the ΔciaRH strain raises the possibility that CiaR controls expression of an extended operon containing both pepN and the ciaRH locus itself.

[i] Increased lytA expression was seen with two independent probes for this gene present on the microarray. The expression ratio presented is the geometric mean of the two observations (3.1 and 2.2) while the P value shown is the product of the two independent P values (2.9×10^{-2} and 3.7×10^{-2}).

the ΔciaRH strain. This group consisted of 24 genes that showed decreased expression and 22 genes that showed increased expression in this mutant (Table 1). These 46 genes fulfilled criteria chosen to identify genes that showed significant differences in expression in replicate experiments and for

which the magnitudes of the expression differences were substantial. These criteria required both a $\log_{10}$(mutant strain/wild type) expression ratio of greater than $|0.3|$, corresponding to at least a twofold change in expression, and an unpaired t test P value of mutant versus wild-type $\log_{10}$(relative fluorescence)

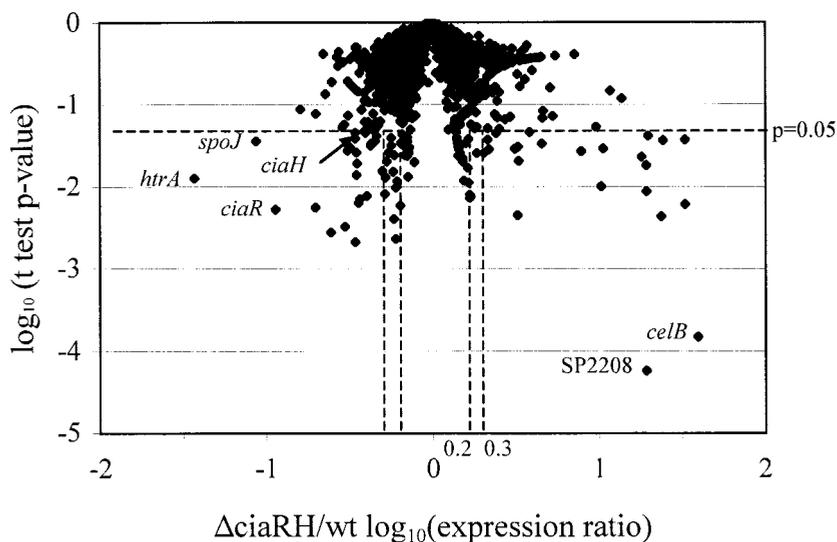FIG. 2. Magnitude and significance of gene expression changes in the Δ*ciaRH* mutant compared to 0100993. The *x* axis represents the Δ*ciaRH*/0100993 expression ratio on a logarithmic scale such that the dashed lines at 0.3 and −0.3 correspond to the twofold induction and repression criteria used in defining potential Δ*ciaRH*-regulated genes. The dashed lines at 0.2 and −0.2 indicate criteria used in defining the set of genes excluded based on potential regulation in other TCSTS examined. The *y* axis represents the degree of confidence assigned to differences in gene expression between the Δ*ciaRH* mutant and strain 0100993 given by *P* values from individual *t* tests. The horizontal dashed line indicates the position on the logarithmic scale of the $P < 0.05$ criterion used in defining potentially regulated genes.

values of <0.05. Figure 2 shows the distribution of these values for the comparison of the Δ*ciaRH* gene expression profile with that of the wild-type strain. This set of putative CiaRH-regulated genes was narrowed by subtracting from this set those genes that were also affected in any of six other TCSTS mutants that had not shown attenuation in colonization. A less-stringent definition of potentially regulated genes in other TC-STS was adopted by relaxing the $\log_{10}$(null mutant/wild type) expression ratio criterion to greater than |0.2| in order to ensure that the remaining Δ*ciaRH*-specific genes did not show borderline expression changes in other mutant strains. Based on these criteria, the set of 46 potential Δ*ciaRH*-regulated genes was narrowed to 24 genes whose regulation was specific to the CiaRH system. This set included 12 genes that were downregulated 2.2- to 27.0-fold and 12 genes that were up-regulated 2.0- to 23.7-fold in the Δ*ciaRH* null mutant (indicated by "Y" in the third column of Table 1). The low expression values measured for *ciaR* and *ciaH* are the anticipated result of deletion of the majority of the ORFs for both of these genes in the Δ*ciaRH* strain, which would eliminate most of the cDNA complementary to these microarray probes even if transcription of this locus were unaltered. These probes consequently served as positive controls for the ability of the system to detect changes in transcript abundance.

Identification and evaluation of a locus downregulated in Δ*ciaRH*. The selection of candidate genes to explain the colonization deficit of the Δ*ciaRH* strain from among those listed in Table 1 was based on the hypotheses that such genes would most likely be specific to the Δ*ciaRH* mutant relative to the other TCSTS mutants tested and that a decrease in expression, rather than an increase, was likely to cause the loss of the ability to colonize. Because the set of Δ*ciaRH*-upregulated genes was dominated by competence-associated genes in which altered expression generally was not restricted to the Δ*ciaRH*

mutant, these genes seemed less likely to explain the Δ*ciaRH* colonization deficit. Among the set of 12 genes specifically downregulated by the Δ*ciaRH* mutation, SP2239 and SP2240 (hereafter referred to as *htrA* and *spoJ*, respectively, on the basis of homologies outlined below) were the most strongly affected and stood out as candidates to be responsible for at least part of the observed deficit in nasopharyngeal colonization. *htrA* and *spoJ* were downregulated in the Δ*ciaRH* strain by 27.0- and 11.5-fold, respectively, and did not show significant changes in expression in the other TCSTS mutants analyzed (Fig. 3). Independent confirmation of the repression of *htrA* and *spoJ* in the Δ*ciaRH* strain was obtained by using real-time
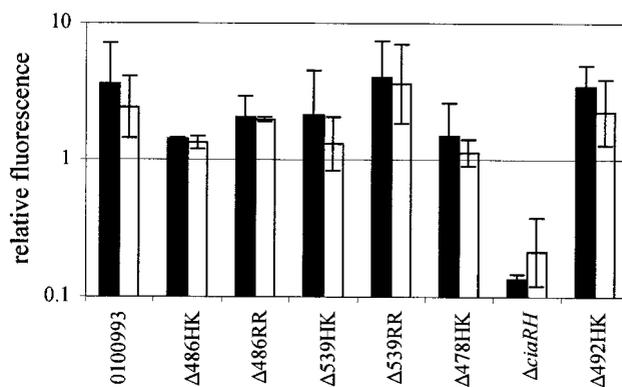

FIG. 3. Expression of *htrA* (■) and *spoJ* (□) in the 0100993 wild-type strain, the Δ*ciaRH* deletion mutant, and the six other TCSTS mutants for which expression data was obtained. Values were determined by microarray hybridization and are given as the geometric mean fluorescence normalized to the median fluorescence level of the entire set of probes on each array. Error bars represent ±1 standard deviation.
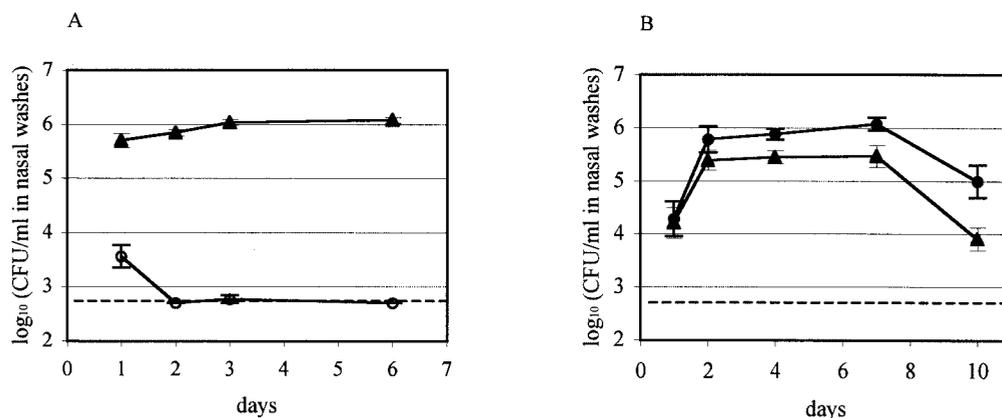
A

B

FIG. 4. Competitive nasopharyngeal colonization assays of the 0100993-SR transformed parent strain (A and B, ▲) versus the Δ*htrA* (A, ○) or Δ*spoJ* (B, ●) null mutants. The data shown for each pair of strains are derived from competitive inoculations of one litter consisting of 10 to 12 individuals. A second independent competition experiment of Δ*htrA* against 0100993-SR gave similar results. Dashed lines indicate the lower limit of detection of the colonization assay.

quantitative RT-PCR, which showed these genes to be downregulated 37.2- and 10.2-fold, respectively, compared to the wild type. The next most strongly downregulated gene identified as specific to Δ*ciaRH*, excluding the genes *ciaR* and *ciaH* bearing deletions, was *licB*, whose expression was decreased only 2.9-fold.

In both the regional sequence of the R6 derivative strain 801 published by Gasc et al. (7) and in the serotype 4 strain TIGR genome sequence (28), *htrA* and *spoJ* are adjacent on the pneumococcal chromosome and are transcribed in the same direction, with 60 bases between the end of *htrA* and the beginning of *spoJ*. Translation of the pneumococcal *htrA* ORF yields a protein that shares 25% identity and 39% similarity with the *Escherichia coli* periplasmic serine protease HtrA (14) and 52% identity and 68% similarity with the *Lactococcus lactis* surface-expressed serine protease HtrA (24) over the entire lengths of the alignments. The predicted pneumococcal HtrA protein contains a trypsin-family serine protease motif (amino acids 86 to 277) and a PDZ domain sequence (amino acids 279 to 374), which may participate in protein-protein interactions. *htrA* is predicted to encode an N-terminal signal peptide. SpoJ in pneumococcus is homologous to the sporulation protein Spo0J of *Bacillus subtilis* (9, 16), sharing 40% identity and 61% similarity at the amino acid level.

Because the *htrA-spoJ* locus was strongly downregulated in the Δ*ciaRH* strain, deletion of these genes offered the means to investigate their contributions to the Δ*ciaRH* colonization deficit. Separate null mutants in *htrA* and *spoJ* were generated by insertion-deletion mutagenesis by using an *ermAM* cassette and were tested in the infant rat model of nasopharyngeal carriage. The Δ*htrA* strain was recovered in nasal washings throughout the 2-week time course of the experiment, although colony counts for this strain were significantly lower than for the wild-type strain 0100993 on days 2 and 7 (Fig. 1C). The Δ*spoJ* strain, in contrast, showed no deficit in its ability to colonize the infant rat nasopharynx (Fig. 1D).

A competitive model of colonization involving coinoculation of the defined mutant and an isogenic streptomycin-resistant *S. pneumoniae* 0100993 derivative, 0100993-SR, was employed to investigate the possibility of a partial loss of colonization fit-

ness caused by the Δ*htrA* mutation not sufficient to eliminate carriage entirely. The fitness of 0100993-SR was first assessed in the single-strain inoculation model, in which it established the same plateau density of colonization as the wild-type strain (Fig. 1A). This streptomycin-resistant strain was then employed as the control in a dual-inoculation model of carriage. After inoculation of Δ*htrA* and 0100993-SR in a ratio of 1:1.1 as determined by colony counts of the inoculum on selective media, 0100993-SR rapidly outcompeted the Δ*htrA* strain such that within 2 days recovery of erythromycin-resistant Δ*htrA* colonies had decreased to the lower limit of detection of the colonization assay (Fig. 4A). In contrast, when 0100993-SR and Δ*spoJ* were inoculated in a ratio of 1:1.7, the Δ*spoJ* strain achieved colonization densities as high or higher than the 0100993-SR transformed control (Fig. 4B). This maintenance of carriage by the Δ*spoJ* mutant demonstrated that the competitive colonization deficit of the Δ*htrA* strain was specific to the *htrA* gene rather than being due either to a polar effect on a downstream gene or to a site-independent effect of carrying the *ermAM* cassette. Measurement of growth rates in vitro demonstrated that this competitive effect of the Δ*htrA* mutation was specific to the in vivo situation. In mixed culture in vitro, the growth rates of 0100993-SR and the Δ*htrA* strain were indistinguishable, having doubling times of 34.2 ± 0.3 and 34.7 ± 0.3 min, respectively.

DISCUSSION

We have investigated the ability of null mutants in nine *S. pneumoniae* TCSTS to colonize the infant rat nasopharynx and have identified one of these systems, CiaRH, as essential for the maintenance of nasal carriage. The other TCSTS deletion strains evaluated established robust colonization despite seven of these TCSTS having shown attenuation of at least 10^3-fold in a murine model of pneumonia (29). Although host species-specific factors cannot be excluded, this difference in fitness for colonization of the upper respiratory tract compared to invasion of the lower respiratory tract suggests that these seven TCSTS mediate adaptations specific to the latter environment. That such a large number of pneumococcal TCSTS appear not
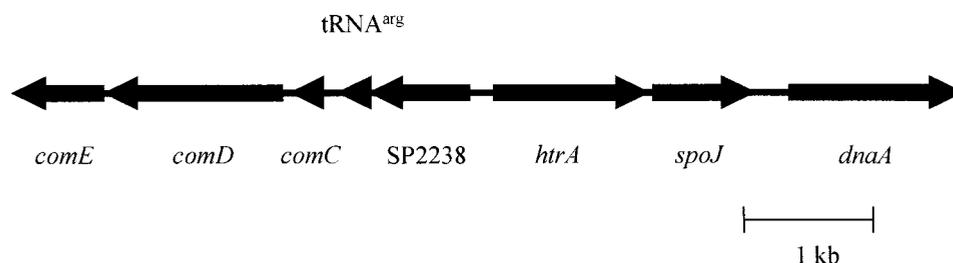
tRNA^arg

FIG. 5. Organization of the region of the pneumococcal chromosomal surrounding the *htrA-spoJ* locus as determined from the TIGR4 sequence (28). Arrows indicate ORFs as well as a tRNA^arg sequence.

to play an active role in its commensal colonization of the nasopharyngeal mucosal surface is consistent with *S. pneumoniae* having evolved principally to survive in this niche and requiring active adaptations away from this basal state as an infection progresses.

Our findings were based on analyzing defined TCSTS mutants generated by transformation. With the exception of Δ*ciaRH*, there was a tendency for these mutants to colonize more efficiently than the parent strain during the initial days postinoculation. The observation that the control streptomycin-resistant strain derived from the wild type by transformation was also able to establish high density nasopharyngeal colonization more rapidly than its parent strain (Fig. 1A, day 1) suggests that the process of transformation itself might have a small positive effect on colonization fitness. The transformation process has previously been reported to select for a less heavily encapsulated subpopulation of *S. pneumoniae* (32) that might be better suited for carriage.

Our results showing no difference in the growth rates in vitro between the Δ*ciaRH* strain and the wild type differs from that reported by Lange et al. (12) who, using different media, described a mid-exponential-growth-phase delay with a Δ*ciaRH* strain before growth was resumed at the wild-type rate. The same study also demonstrated a more prolonged mid-exponential-growth-phase delay for their RR04 strain carrying a response regulator mutation corresponding to our Δ481RR strain, which in our study was able to colonize the infant rat nasopharynx to high densities. We therefore do not believe that the inability of the Δ*ciaRH* mutant to colonize can be attributed to a constitutive growth defect.

Microarray hybridization experiments revealed that genes previously reported to be involved in later stages of the induction of competence for genetic transformation (22, 23, 25) predominated among those upregulated in response to the Δ*ciaRH* mutation. This observation is consistent with studies that have shown that loss-of-function mutations in the CiaRH system relieve the repression of the competence system that is seen ordinarily with microaerobiasis and at nonoptimal bacterial densities, under which conditions the wild type does not display competence (6, 15). Although deletions of *comD*, which eliminate competence (21), have been reported to attenuate systemic virulence (1), it is uncertain whether the tendency toward increased competence of the Δ*ciaRH* mutant is related to this strain's inability to colonize the nasopharynx. Our finding that the Δ*ciaRH* mutant is unable to persist in nasal carriage suggests that expression of the competence phenotype either directly interferes with colonization or is regu-

lated reciprocally with other adaptations necessary for colonization.

Comparison of the transcriptional changes induced by the Δ*ciaRH* mutation with the changes associated with mutations in other pneumococcal TCSTS genes by means of a subtractive algorithm enabled us to identify 24 genes that were specifically affected in the colonization-attenuated Δ*ciaRH* strain relative to other strains that had shown robust colonization. The algorithm's inclusion criterion of genes having individual t test P values of <0.05 associated with their expression differences was chosen to minimize false negative results while recognizing that the resulting group of genes would contain false-positives due to the multiple hypothesis testing inherent in analyses of microarray data sets. This subtractive procedure revealed that, while upregulation of many late competence genes was a pattern present in several other TCSTS mutants, downregulation of *htrA* and *spoJ* was seen only with the Δ*ciaRH* mutant.

Competitive nasal carriage assays demonstrated that deletion of *htrA*, but not of the downstream gene *spoJ*, led to a decrease in fitness for colonization and therefore was likely to contribute to the attenuated colonization of the Δ*ciaRH* strain in which *htrA* expression was repressed. Sequence comparisons suggest that the pneumococcal HtrA protein belongs to a family of serine proteases with a wide phylogenetic distribution in both gram-negative and gram-positive bacteria where they are expressed in the periplasmic space or on the bacterial surface, respectively. Studies have not yet addressed the potential proteolytic activity of this protein in *S. pneumoniae* nor have potential substrates been identified.

The strongly decreased expression of the pneumococcal *htrA* gene in the Δ*ciaRH* mutant raises the possibility that HtrA might participate in the inhibitory control CiaRH exerts over the competence pathway. The most readily apparent mechanism for a surface protease to exert such an effect would be through degradation of the secreted CSP at the cell surface. The proximity in the TIGR4 pneumococcal genome sequence of *htrA* to the *comCDE* locus (Fig. 5) encoding the CSP precursor and the ComDE two-component system is also consistent with the possibility of a functional connection between HtrA and the competence system.

Pneumococcal HtrA may alternatively perform a function similar to those attributed to other members of the HtrA family. These proteases have been implicated in the degradation of denatured periplasmic or surface proteins, and mutations eliminating HtrA in other bacterial species have been shown to cause increased sensitivity to thermal, oxidative or osmotic stress (reviewed in reference 19). Strains carrying de-

letions of *htrA* have been found previously to have reduced virulence in murine models of infection with *Salmonella enterica* serovar Typhimurium (2, 10) and *Yersinia enterocolitica* (13). Whether pneumococcal HtrA functions in a similar role with regard to stress in vitro or invasive infection in vivo remains a subject for investigation.

Because the Δ*htrA* strain, unlike Δ*ciaRH*, was able to establish colonization in a single-strain carriage model, *htrA* clearly cannot be the only gene regulated by CiaRH that contributes to fitness for nasopharyngeal carriage. The list of genes presented in Table 1 provides a starting point for identifying such additional contributing loci. Beyond these genes that showed differential expression in our microarray analysis, other loci may participate in the adaptations orchestrated by CiaRH that are necessary for colonization. Since our array was designed based on preliminary sequence data and did not contain probes to 34% of the ORFs found in the completed TIGR4 genome sequence, genes of interest may have been overlooked because they were not represented on the array. Additionally, regulated genes that are expressed even in their activated states at levels near or below the threshold of detection of the microarray system would not have been identified by this procedure. Finally, any genes activated by CiaR during colonization, but not under the in vitro conditions used to produce RNA for the microarray experiments, would not have been detected by this screening method. Techniques such as characterization of CiaR-binding sites and *htrA* upstream regulatory elements, followed by genomic searches for similar sequences, may help to elucidate further the genetic and physiologic adaptations mediated by CiaRH that are required for persistence of the pneumococcus in its commensal state.

This study has identified the *S. pneumoniae* CiaRH two-component signal transduction system as uniquely necessary for nasopharyngeal colonization among the nine TCSTS tested. The attenuation of a Δ*ciaRH* null mutant in separate models of mucosal carriage and pneumonia now establishes this signaling pathway as a potential target for disrupting both pneumococcal colonization and invasive disease. Using microarray hybridization to screen for genes having altered expression levels in a Δ*ciaRH* mutant, the colonization deficit of this mutant has been attributed in part to the downregulation of the putative surface-localized serine protease HtrA. This study demonstrates the utility of combining microarray expression data with conventional phenotypic analyses in the study of bacterial signal transduction pathways and the role of these systems in pathogenesis. Additional studies will be necessary to dissect further the pathways under the control of the CiaRH system that are required for nasopharyngeal colonization by *S. pneumoniae*.

ACKNOWLEDGMENTS

obtained from The Institute for Genomic Research website (http://www.tigr.org).

REFERENCES

1. **Bartilson, M., A. Marra, J. Christine, J. S. Asundi, W. P. Schneider, and A. E. Hromockyj.** 2001. Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide. Mol. Microbiol. **39:**126–135.
2. **Chatfield, S. N., K. Strahan, D. Pickard, I. G. Charles, C. E. Hormaeche, and G. Dougan.** 1992. Evaluation of *Salmonella typhimurium* strains harbouring defined mutations in *htrA* and *aroA* in the murine salmonellosis model. Microb. Pathog. **12:**145–151.
3. **Cheng, Q., E. A. Campbell, A. M. Naughton, S. Johnson, and H. R. Masure.** 1997. The *com* locus controls genetic transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **23:**683–692.
4. **Chuang, S. E., D. L. Daniels, and F. R. Blattner.** 1993. Global regulation of gene expression in *Escherichia coli*. J. Bacteriol. **175:**2026–2036.
5. **de Saizieu, A., C. Gardès, N. Flint, C. Wagner, M. Kamber, T. J. Mitchell, W. Keck, K. E. Amrein, and R. Lange.** 2000. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. J. Bacteriol. **182:**4696–4703.
6. **Echenique, J. R., C.-R. Sabine, and M.-C. Trombe.** 2000. Competence regulation by oxygen in *Streptococcus pneumoniae*: involvement of *ciaRH* and *comCDE*. Mol. Microbiol. **36:**688–696.
7. **Gasc, A.-M., P. Giammarinaro, S. Richter, and M. Sicard.** 1998. Organization around the *dnaA* gene of *Streptococcus pneumoniae*. Microbiol. **144:**433–439.
8. **Guenzi, E., A. M. Gasc, M. A. Sicard, and R. Hakenbeck.** 1994. A two-component signal-transducing system is involved in competence and penicillin susceptibility in laboratory mutants of *Streptococcus pneumoniae*. Mol. Microbiol. **12:**505–515.
9. **Ireton, K., N. W. Gunther IV, and A. D. Grossman.** 1994. *spo0J* is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. J. Bacteriol. **176:**5320–5329.
10. **Johnson, K., I. Charles, G. Dougan, D. Pickard, P. O'Gaora, G. Costa, T. Ali, I. Miller, and C. Hormaeche.** 1991. The role of a stress-response protein in *Salmonella typhimurium* virulence. Mol. Microbiol. **5:**401–407.
11. **Lacks, S., and R. D. Hotchkiss.** 1960. A study of the genetic material determining an enzyme activity in pneumococcus. Biochim. Biophys. Acta **39:**508–517.
12. **Lange, R., C. Wagner, A. de Saizieu, N. Flint, J. Molnos, M. Stieger, P. Caspers, M. Kamber, W. Keck, and K. E. Amrein.** 1999. Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. Gene **237:**223–234.
13. **Li, S.-R., N. Dorrell, P. H. Everest, G. Dougan, and B. W. Wren.** 1996. Construction and characterization of a *Yersinia enterocolitica* O:8 high-temperature requirement (*htrA*) isogenic mutant. Infect. Immun. **64:**2088–2094.
14. **Lipinska, B., S. Sharma, and C. Georgopoulos.** 1988. Sequence analysis and regulation of the *htrA* gene of *Escherichia coli*: a σ³²-independent mechanism of heat-inducible transcription. Nucleic Acids Res. **16:**10053–10067.
15. **Martin, B., M. Prudhomme, G. Alloing, C. Granadel, and J.-P. Claverys.** 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **38:**867–878.
16. **Mysliwiec, T. H., J. Errington, A. B. Vaidya, and M. G. Bramucci.** 1991. The *Bacillus subtilis spo0J* gene: evidence for involvement in catabolite repression of sporulation. J. Bacteriol. **173:**1911–1919.
17. **Novak, R., E. Charpentier, J. S. Braun, and E. Tuomanen.** 2000. Signal transduction by a death signal peptide: uncovering the mechanism of bacterial killing by penicillin. Mol. Cell **5:**49–57.
18. **Novak, R., B. Henriques, E. Charpentier, S. Normark, and E. Tuomanen.** 1999. Emergence of vancomycin tolerance in *Streptococcus pneumoniae*. Nature **399:**590–593.
19. **Pallen, M. J., and B. W. Wren.** 1997. The HtrA family of serine proteases. Mol. Microbiol. **26:**209–221.
20. **Pearce, B. J., Y. B. Yin, and H. R. Masure.** 1993. Genetic identification of exported proteins in *Streptococcus pneumoniae*. Mol. Microbiol. **9:**1037–1050.
21. **Pestova, E. V., L. S. Håvarstein, and D. A. Morrison.** 1996. Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. Mol. Microbiol. **21:**853–862.
22. **Pestova, E. V., and D. A. Morrison.** 1998. Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a *lacZ* reporter insertion vector. J. Bacteriol. **180:**2701–2710.
23. **Peterson, S., R. T. Cline, H. Tettelin, V. Sharov, and D. A. Morrison.** 2000. Gene expression analysis of the *Streptococcus pneumoniae* competence regulons by use of DNA microarrays. J. Bacteriol. **182:**6192–6202.
24. **Poquet, I., V. Saint, E. Seznec, N. Simoes, A. Bolotin, and A. Gruss.** 2000. HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing. Mol. Microbiol. **35:**1042–1051.
25. **Rimini, R., B. Jansson, G. Feger, T. C. Roberts, M. de Francesco, A. Gozzi,**

F. Faggioni, E. Domenici, D. M. Wallace, N. Frandsen, and A. Polissi. 2000. Global analysis of transcription kinetics during competence development in *Streptococcus pneumoniae* using high-density DNA arrays. Mol. Microbiol. **36:**1279–1292.

26. Schultz, J., R. R. Copley, T. Doerks, C. P. Ponting, and P. Bork. 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. **28:**231–234.

27. Stock, A. M., V. L. Robinson, and P. N. Goudreau. 2000. Two-component signal transduction. Annu. Rev. Biochem. **69:**183–215.

28. Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498–506.

29. Throup, J. P., K. K. Koretke, A. P. Bryant, K. A. Ingraham, A. F. Chalker, Y. Ge, A. Marra, N. G. Wallis, J. R. Brown, D. J. Holmes, M. Rosenberg, and M. K. Burnham. 2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. Mol. Microbiol. **35:**566–576.

30. von Heijne, G. 1986. A new method for predicting signal cleavage sites. Nucleic Acids Res. **14:**4683–4690.

31. Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure. 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect. Immun. **62:**2582–2589.

32. Weiser, J. N., and M. Kapoor. 1999. Effect of intrastrain variation in amount of capsular polysaccharide on the genetic transformation of *Streptococcus pneumoniae*: implications for virulence studies of encapsulated strains. Infect. Immun. **67:**3690–3692.

Editor: E. I. Tuomanen