Analysis of Nontypeable *Haemophilus influenzae* Phase-Variable Genes During Experimental Human Nasopharyngeal Colonization

Jessica Poole,1,a Eric Foster,2,a Kathryn Chaloner,2 Jason Hunt,1 Michael P. Jennings,7 Thomas Bair,3 Kevin Knudtson,4 Erik Christensen,1 Robert S. Munson Jr,4 Patricia L. Winokur,5,a and Michael A. Apicella1

1Department of Microbiology, 2Department of Biostatistics, 3Bioinformatics Facility, 4DNA Facility, and 5Department of Internal Medicine, University of Iowa, Iowa City; 6Research Institute at Nationwide Children’s Hospital, Ohio State University, Columbus; and 7Institute for Glycomics, Griffith University, Gold Coast, Australia

(See the editorial commentary by Inzana on pages 713–6.)

**Background.** Studies of nontypeable *Haemophilus influenzae* (NTHi) have demonstrated that a number of genes associated with infectivity have long repeat regions associated with phase variation in expression of the respective gene. The purpose of this study was to determine the genes that underwent phase variation during a 6-day period of experimental human nasopharyngeal colonization.

**Methods.** Strain NTHi 2019StrR1 was used to colonize the nasopharynx of human subjects in a study of experimental colonization. Thirteen phase-variable genes were analyzed in NTHi 2019StrR1. Samples of NTHi 2019StrR1 were cultured from subjects during the 6-day colonization period. We used capillary electrophoresis and Roche 454 pyrosequencing to determine the number of repeats in each gene from each sample.

**Results.** A significant number of samples switched licA and igaB from phase off in the inoculated strain to phase on during the 4-day period of observation. lex2A also showed variability as compared to baseline, but the differences were not significant. The remaining genes showed no evidence of phase variation.

**Conclusions.** Our studies suggest that the phase-on genotypes of licA and igaB are important for early human nasopharynx colonization. lex2A showed a trend from phase off to phase on, suggesting a potentially important role in the colonization process.

**Keywords.** *Haemophilus influenzae*; phase variation; licA; igaB; lex2A nasopharyngeal colonization.

Phase variation is described as a random process by which a clonal population of microbes can present heterogeneous phenotypes as a result of a reversible genetic event [1, 2]. The process can involve several mechanisms, including slipped-strand mispairing (SSM), site-specific recombination, and epigenetic regulation mediated by DNA methylation [2]. SSM is a mechanism found in human pathogens, including pathogenic Neisseria, Bordetella pertussis, *Haemophilus influenzae*, and *Helicobacter pylori*. The resulting frame shift causes on-off changes in gene expression [3]. The process can occur at rates as high as $2\times10^{-4}$ or $3\times10^{-4}$ colony forming units [3]. Nontypeable *H. influenzae* (NTHi), which colonizes human mucosal surfaces and can cause disease, has genes with long tandem repeats that undergo high rates of phase variation [3]. In 1989, Weiser et al identified a tetranucleotide repeat region (5′-CAAT-3′) in licA that was responsible for phase-variable expression of that gene in *H. influenzae* type b [4]. Subsequent work showed that a number of genes in the *H. influenzae* chromosome had similar characteristics and underwent phase variation. These genes are associated with iron regulation, membrane protein expression, and lipooligosaccharide (LOS) biosynthesis [5–9]. The opportunity to directly study the effect of experimental human colonization on changes in the phase-on,
phase-off status of NTHi phase-varying genes was presented to us. This article describes the results of that study.

MATERIALS AND METHODS

Strains
NTHi 2019 is a clinical isolate obtained from a patient with chronic obstructive pulmonary disease (COPD) and chronic bronchitis in 1985. NTHi 2019StrR1 was created from NTHi 2019 by using a specific chromosomal point mutation at nucleotide 128, resulting in a change in amino acid 43 from lysine to arginine in the ribosomal RNA gene, rpsL, that encodes the S12 polypeptide to encode streptomycin resistance. NTHi 2019StrR1 was susceptible to amoxicillin, doxycycline, trimethoprim-sulfamethoxazole, cefuroxime-axetil, ceftriaxone, and ciprofloxacin. This organism was resistant to streptomycin (minimum inhibitory concentration, ≥256 μg/mL). This strain was used as the inoculation strain for a National Institutes of Health–and Food and Drug Administration–approved human experimental colonization study (unpublished data). Forty-eight samples of NTHi 2019StrR1 were recovered from 15 subjects who participated in the 2 studies of NTHi human experimental nasopharyngeal colonization over a 6-day colonization period (unpublished data). Pharyngeal swab samples were streaked on brain heart infusion broth for 6 hours as previously described [4].

DNA Isolation and Polymerase Chain Reaction (PCR)
DNA was isolated from each of the 48 samples, using the Masterpure DNA purification kit from Epicentre Biotechnologies (Madison, WI). As a zero-time baseline, DNA was isolated from randomly selected NTHi 2019StrR1 freezer stock vials used to inoculate the subjects. DNA from 5 separate stocks was used in the pyrosequencing studies and in the capillary electrophoresis studies. PCR was performed using primers obtained from Integrated DNA Technologies (Iowa City, IA). Primers were labeled with 6-carboxyfluorescein for capillary electrophoresis sequencing or with MID codes to identify each subject sample for pyrosequencing shown in Supplementary Table 1.

Capillary Electrophoresis to Determine Nucleotide Repeat Length
Initial studies used a modified version of the capillary electrophoresis method developed by Fox et al to determine the length of the tandem repeats in PCR products for licA, lic3A, and igaB, using Peak Scanner software (Applied Biosystems International) [11]. The primers used in the capillary electrophoresis studies are shown in Supplementary Table 2.

Amplicon Sequencing
In the second half of the study, changes in the number of repeats of 12 of the 13 phase variable genes with tandem repeats were analyzed simultaneously by pyrosequencing, using the GS FLX Titanium System (Roche, Branford, CT) according to the manufacturer’s recommended protocols. Pyrosequencing was performed by the University of Iowa DNA Facility and involved a 4-region slide, in which amplicon fragments of isolated DNA from each day were placed into a separate region on the chip. Sequencing libraries, prepared using the GS FLX Titanium Rapid Library Preparation Kit (Roche), were started with a modified fragment end repair step. The modified fragment end repair reaction did not include the T4 polymerase, and only half the recommended volumes of the polynucleotide kinase and Taq polymerase were used to minimize exonuclease digestion of the MIDs that had been incorporated into the 5’ ends of the genome-specific amplification primers. Following ligation of the RL adaptors, a modified AMPure XP cleanup kit was used that excluded the sizing solution and used a 1.6:1 (v/v) ratio to DNA to target fragments longer than 200 bp. Libraries were clonally amplified by the GS FLX Titanium MV emulsion PCR kit (Lib-L; Roche), using a 0.4 copy per bead ratio. Beads with amplified libraries were loaded onto GS FLX Titanium PicoTiterPlate set up with 4-region dividers to accommodate the sample pools.

Informatics Analysis of the Expansion of the Various Repeats
After a run was completed on the 454, the sffile tool (Roche v2.3) was used to split the sffiles into separate fasta files based on bar codes and assign them to patient identification numbers on the basis of plate quadrant. Then, sequences with exact matches to the 10 bp before and after the expected repeat regions, considering both forward and reverse reads, were identified. By use of the program countRepeats.py (available at: https://bitbucket.org/tbair/published/src), the exact positions of the start and the end of the repeats were identified, and the number of repeats was identified. Sequences that did not have an exact multiple of the repeat because of sequencing errors or nontypical expansion were identified. Statistics on the minimum, maximum, and average number of repeats were calculated, as well as the number of sequences that were in or out of phase. These data were then incorporated into downstream statistical analysis.

Fluorescent Antibody Studies
Fluorescent antibody analysis was performed using monoclonal antibody 12D9, which is specific for the phosphorylcholine epitope, on strains grown in supplemented brain heart infusion broth for 6 hours as previously described [4].
Statistical Analysis

For measurements on each gene, the proportion of sequences that were phase on was calculated. For proportions that were 0 or 1, an adjustment of $10^{-2}$ was used. For each gene, a baseline confidence interval for the proportion of genes that were phase on was calculated using control DNA samples from the inoculating bacteria. The proportions were transformed to a log odds (logit) scale, and a normal distribution was fit to the logit values, using maximum likelihood and a continuity adjustment to the normal approximation that accounted for the different number of reads for each measurement. A mean and a 95% confidence interval for the mean were calculated in the logit scale and back transformed to an interval for mean baseline proportion.

The in vivo proportions for each gene were also transformed to the logit scale and were analyzed using a linear mixed model. The inherent correlation between measurements on the same subjects was accounted for by using a random effect for subject. The days after challenge were then incorporated as fixed effects. The 95% confidence interval for the mean logit for each day after challenge was calculated and also transformed back to the proportion scale. If this interval did not overlap with the baseline interval, the change from baseline was considered to be statistically significant ($P < .05$). The difference between technical replicates was also examined within the appropriate linear mixed models.

RESULTS

On the basis of sequence analysis of the NTHi strain 2019 genome, we identified 14 genes with ≥12 tandem repeats within their open reading frames (one gene, hgpB, was not studied). Seven of these genes were involved in LOS biosynthesis, 5 are involved in iron acquisition from hemoglobin, 1 encoded DNA methyltransferase, and 1 encoded immunoglobulin A (IgA) protease. Studies were initiated on 13 of these genes (Table 1) in strains recovered from 15 subjects during the 2 challenge studies (a multidose study and a 2-dose study) to determine whether there was on-off switching in these phase-variable genes during colonization. In our initial studies, 2 phase-variable genes, licA and lic3A, were selected for analysis by capillary electrophoresis of samples from the multidose study. Figure 1A demonstrates that licA was phase on in only a mean of 2% of the inoculate strain population and that, by day 3 after nasopharyngeal inoculation, licA phase on had increased to a mean of approximately 15%, with this level maintained in subsequent samples. In contrast, Figure 1B shows that, for lic3A, a mean of approximately 95% of the population was phase on in the inoculation samples, and this percentage did not change significantly over the next 4 days in the organisms recovered from the patients.

To study whether phase variation was occurring in the samples from both colonization studies, we used 454 pyrosequencing to simultaneously analyze all of these genes except igaB from the 48 samples recovered from the 15 subjects who participated in both studies. The samples were isolated over the last 4 days of the 6-day colonization period. To accomplish this, we used primers that were bar coded for each patient, primers that were phase on had increased to a mean of approximately 18.8% on day 3 and to 25% on day 6. Fluorescent antibody studies were performed on the sample isolated on day 6 from subject 10 (Figure 3A and 3B). Genotyping indicated that >60% of the organisms in this sample population had shifted to phase on. These studies used monoclonal antibody 12D9, which is specific for the phosphorylcholine epitope on the LOS. As can be seen, there was a marked increase in fluorescence of the bacteria in the sample recovered on day 6 after

---

**Table 1. Phase-Variable Genes Analyzed in this Study, Repeat Sequence, and Number of Repeats, Based on Genome Sequencing of Nontypeable Haemophilus influenzae**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Annotation</th>
<th>Repeat Sequence</th>
<th>No. of Repeats (Phase On)</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choline kinase</td>
<td>licA</td>
<td>CAAT</td>
<td>25, 27</td>
<td>KC607490</td>
</tr>
<tr>
<td>Glycosyltransferase</td>
<td>lic2A</td>
<td>CAAT</td>
<td>16, 17</td>
<td>KC607492</td>
</tr>
<tr>
<td>Sialyltransferase</td>
<td>lic3A</td>
<td>CAAT</td>
<td>15, 16</td>
<td>KC607491</td>
</tr>
<tr>
<td>Glycosyltransferase</td>
<td>lgtC</td>
<td>GACA</td>
<td>4</td>
<td>KC607500</td>
</tr>
<tr>
<td>Putative glycosyltransferase</td>
<td>tgt2</td>
<td>CCAA</td>
<td>12, 13</td>
<td>KC607499</td>
</tr>
<tr>
<td>LOS biosynthesis</td>
<td>lex2A</td>
<td>GCAA</td>
<td>21</td>
<td>KC607502</td>
</tr>
<tr>
<td>Hemoglobin-haptoglobin binding protein</td>
<td>hbpC</td>
<td>CCAA</td>
<td>46, 47</td>
<td>KC607495</td>
</tr>
<tr>
<td></td>
<td>hbpB2</td>
<td>CCAA</td>
<td>48, 49</td>
<td>KC607497</td>
</tr>
<tr>
<td></td>
<td>hbpB3</td>
<td>CCAA</td>
<td>31, 33</td>
<td>KC607496</td>
</tr>
<tr>
<td></td>
<td>hbpC2</td>
<td>CCAA</td>
<td>40, 41</td>
<td>KC607494</td>
</tr>
<tr>
<td>Sialyltransferase</td>
<td>siaA2</td>
<td>CAAT</td>
<td>20, 21</td>
<td>KC607493</td>
</tr>
<tr>
<td>Mod methylase TIIIR/M system</td>
<td>IgaA protease</td>
<td>igaB AAATTC11</td>
<td>11</td>
<td>KC607498</td>
</tr>
</tbody>
</table>

*a* Three possible starts in 2 reading frames; gene is out of frame, so add or subtract 1 repeat to have an in-frame start.

*b* Four possible starts in 2 reading frames; gene is in frame.

*c* Two possible starts in 2 reading frames; gene is in frame.

*d* Three possible starts in 2 reading frames; gene is out of frame.
inoculation. *lex2A* showed more variability in the percentage of the population in which it was phase on as compared to baseline, but the change in mean values was not significant (Figure 4). Analysis of the remaining phase-variable genes showed minimal changes between the challenge strain and the sample recovered from days 3 through 6. Pyrosequencing studies of *lic3A* and the remaining 10 genes showed that the phase status changed in only a few instances during the period of colonization, compared with the inoculation strain (Table 2).

The phase variation of *igaB* was studied by capillary electrophoresis alone. Samples from all 15 subjects were studied. We found that, over the 6-day colonization period, *igaB* was initially predominantly phase off, with a shift to phase on in a substantial percentage of the population (Figure 5).

### Table 2. Percentage of Phase-Variable Genes in Nontypeable *Haemophilus influenzae* 2019StrR<sup>+</sup> Isolates That Were Phase On During the Last 4 Days of the 6-Day Experimental Colonization Period

<table>
<thead>
<tr>
<th>Gene</th>
<th>Inoculum Percentage On (95% CI)</th>
<th>Day 3 Percentage On (95% CI)</th>
<th>Day 4 Percentage On (95% CI)</th>
<th>Day 5 Percentage On (95% CI)</th>
<th>Day 6 Percentage On (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>hgpC</em></td>
<td>46 (33.3–59.1)</td>
<td>5.8 (1.3–22.2)</td>
<td>43.4 (14.2–78.0)</td>
<td>8.9 (2.4–28.0)</td>
<td>38.6 (13.0–72.6)</td>
</tr>
<tr>
<td><em>hgpB</em>2</td>
<td>30.4 (21.2–41.4)</td>
<td>37.6 (21.0–57.7)</td>
<td>40.1 (23.8–59.1)</td>
<td>21.0 (11.4–35.4)</td>
<td>36.4 (21.0–55.2)</td>
</tr>
<tr>
<td><em>hgpB</em>3</td>
<td>33.8 (23.7–45.8)</td>
<td>24.1 (11.7–43.2)</td>
<td>39.2 (21.9–59.8)</td>
<td>24.2 (12.8–40.8)</td>
<td>26.5 (13.9–44.6)</td>
</tr>
<tr>
<td><em>hgpC</em>2</td>
<td>35.4 (24.8–47.5)</td>
<td>11.5 (3.3–33)</td>
<td>31.3 (11.6–61.3)</td>
<td>23.8 (8.7–50.4)</td>
<td>16.7 (5.6–40.5)</td>
</tr>
<tr>
<td><em>lex2A</em></td>
<td>14.3 (9.4–21.1)</td>
<td>20.5 (14.8–27.8)</td>
<td>27.4 (20.4–35.7)</td>
<td>24.0 (17.9–31.3)</td>
<td>16.5 (11.9–22.4)</td>
</tr>
<tr>
<td><em>lic2A</em></td>
<td>1.3 (0.8–2.1)</td>
<td>0.7 (0.3–1.8)</td>
<td>0.9 (0.4–2)</td>
<td>2.1 (0.9–4.6)</td>
<td>3.0 (1.3–6.7)</td>
</tr>
<tr>
<td><em>lic3A</em></td>
<td>90.7 (86.1–93.9)</td>
<td>80.3 (65.2–89.9)</td>
<td>88.5 (78.3–94.30)</td>
<td>93.0 (86.3–96.5)</td>
<td>85.0 (72.2–92.4)</td>
</tr>
<tr>
<td><em>licA</em></td>
<td>3.8 (2.4–6)</td>
<td>18.8 (7.1–40.9)</td>
<td>13.4 (5.0–31.4)</td>
<td>18.5 (7.3–39.6)</td>
<td>25.2 (10.3–49.6)</td>
</tr>
<tr>
<td>Mod-like</td>
<td>7.4 (4.7–11.3)</td>
<td>2.6 (1.2–5.8)</td>
<td>2.5 (1.2–5.8)</td>
<td>1.6 (0.8–3.3)</td>
<td>6.3 (3–12.7)</td>
</tr>
<tr>
<td><em>lgtC</em></td>
<td>6.8 (4.3–10.8)</td>
<td>4.9 (2.2–11)</td>
<td>8.7 (4–17.8)</td>
<td>3.0 (1.4–6.4)</td>
<td>5.5 (2.5–11.4)</td>
</tr>
<tr>
<td><em>pgt2</em></td>
<td>0.4 (0.2–9)</td>
<td>0.1 (0.0–2)</td>
<td>0.2 (0.1–4)</td>
<td>0.1 (0.0–2)</td>
<td>0.1 (0.0–2)</td>
</tr>
<tr>
<td><em>siaA</em>2</td>
<td>2.5 (1.6–3.5)</td>
<td>0.6 (1.3–4.4)</td>
<td>1.3 (0.6–2.8)</td>
<td>1.3 (0.6–2.8)</td>
<td>1.3 (0.6–2.8)</td>
</tr>
<tr>
<td><em>igaB</em></td>
<td>3.8 (2.0–7.3)</td>
<td>8.3 (4.7–14.2)</td>
<td>11.3 (6.8–18.3)</td>
<td>10.3 (6.1–16.9)</td>
<td>12.5 (7.5–20.2)</td>
</tr>
</tbody>
</table>
In summary, both licA and igaB showed a significant shift from phase off to phase on during a 6-day colonization period, compared with the inoculation strain. The other 12 genes examined did not show significant differences in phase variation, compared with the inoculation strain.

**DISCUSSION**

NTHi is a commensal bacterium of the human upper airway that can cause diseases such as otitis media, acute sinusitis, and bronchitis [12]. It possesses a number of genes with multiple tandem repeats, which undergo phase variation at high rates [3]. Phase variation is the rapid, reversible switching of gene expression that creates a heterogeneous phenotypic population in a clonal population. Site-specific recombination, SSM, and epigenetic regulation by DNA methylation are all mechanisms that mediate phase variation [2]. The mechanism for phase variation in *H. influenzae* is SSM. SSM is an intrahelical event and involves the mispairing of complementary bases at a site of a short tandem repeat by the local denaturation and displacement of DNA strands during replication. This leads to an insertion or deletion in the short tandem repeats [13]. The resulting frame shift in the open reading frame of a gene causes premature termination of translation if the gene is out of frame for expression (ie, phase off) or synthesis of a complete protein if in frame (ie, phase on) [3]. Alteration of repeats in promoter regions can also influence gene expression by altering transcription efficiency.
In this study, we looked at 13 phase-variable genes in 48 NTHi 2019StrK1 samples recovered over 4 days during an experimental human nasopharyngeal colonization study. These genes are involved in LOS biosynthesis, membrane protein expression, IgA protease production, iron regulation, and epigenetic regulation [5–9, 14–17]. The inoculation strain had only one of these genes predominantly phase on, the sialyltransferase licA3A. The 4 iron regulation genes showed an intermediate pattern, with approximately 30–40% of the population being phase on at inoculation, and this persisted at that level throughout the study. The LOS biosynthesis gene, lex2A, had a mean phase-on frequency of 14% in the inoculation strain, and, by days 4 and 5 of the observation period, approximately 23–25% population had switched to phase on. At day 4, the 95% confidence intervals of the lex2A phase-on percentage was barely overlapping the confidence intervals of that for the inoculation strain. licA and igaB were the only genes that showed a significant change in the samples recovered during the course of the infection, as compared to baseline. This change began with the first isolation of studies on day 3 for licA and continued until day 6. The average phase on percentage for licA in the inoculating strain was approximately 2%. This increased to 19% on day 3, to 13% on day 4, to 19% on day 5, and to 25% on day 6. The phase-on status of igaB samples rose progressively over the colonization period and, by day 6, was significantly different than that for the inoculation strain (Figure 5). The relationship between phase variation of licA and alternative start sites has been extensively studied by Weiser et al [4] and High et al [18].

LOS is a major antigenic component of the NTHi surface [15]. H. influenzae expresses host carbohydrate mimics within the oligosaccharide region of the LOS. These include phosphorylcholine (ChoP), N-acetyl-lactosamine, pK antigen, and sialylactosamine [12]. H. influenzae acquires choline from its environment, as choline is a major constituent of eukaryotic membrane lipids [19, 20]. Choline is a structure present on other human respiratory pathogens, including Neisseria meningitidis, Streptococcus pneumoniae, and Pseudomonas aeruginosa. The phosphorylcholine epitope undergoes phase variation in H. influenzae, on pili of N. meningitidis and N. gonorrhoeae, and on a protein in P. aeruginosa [20, 21]. licA in H. influenzae has a tetranucleotide repeat sequence of 5′-(CAAT)-3′ within its open reading frame, which is part of a 4-gene operon involved in deposition of phosphorylcholine on the LOS [8, 20, 21].

H. influenzae licA phase variation studies in animal models support our findings of a significant selection for licA on during the course of the human nasopharyngeal colonization [8]. This would support our finding of a significant selection for licA on during the course of the human nasopharyngeal colonization. Weiser et al showed that, in the infant rat model, there is a gradual selection for licA on (ChoP+) over 10 days in the nasopharynx and that constitutively licA-off (ChoP−) mutants were cleared faster from the nasopharynx, compared with wild-type controls [17]. Phase variation of ChoP has also been investigated in the chinchilla model. It was found that, during nasopharyngeal colonization, there was a switch to ChoP+ in as early as 3 days and that chinchillas inoculated with ChoP− phenotype had a greater level of NTHi 2019 per milliliter than chinchillas inoculated with a mostly ChoP− population of 2019. This suggested better colonization ability by the ChoP+ variants in chinchillas. Furthermore, there was a higher concentration of bacteria from the middle ear of chinchillas with the ChoP− phenotype. ChoP− variants were also able to colonize the middle ear, but the infection was less severe than the infection displayed by ChoP+ variants [15].

The expression of ChoP has been linked to susceptibility to serum killing mediated by C-reactive protein (CRP). With H. influenzae, CRP-mediated killing is done through activation of complement and bacteriolysis. This is different than with pneumococcus, where CRP-complement interaction mediates opsonization by phagocytes. The position of the phosphorylcholine on LOS affects the susceptibility of H. influenzae to CRP-mediated killing [22, 23]. CRP is present in the human respiratory tract. It was shown that CRP is present in both inflamed (concentration, 0.17–42 μg/mL) and uninflamed
(concentration, <0.05 to 0.88 μg/mL) human respiratory secretions. The human epithelial cells of the respiratory tract were discovered to be capable of expressing CRP, but complement is required for killing of NTHi [24]. This may suggest that the levels of functionally active complement are not high enough in the airway to make NTHi favor the ChoP– phenotype. The ability to turn off licA could be more of a factor in systemic disease. licA in H. influenzae type B was examined in the infant rat model, with comparisons in bacteria from the nasopharynx, bloodstream, and cerebrospinal fluid obtained 48 hours after infection. There was a different selection for licA on in the nasopharynx, with almost equal amounts on and off, compared with the bloodstream and cerebrospinal fluid, in which the gene was mostly off [8]. A licA mutant that is constitutively on has been constructed, and it showed a reduced ability to cause invasive disease and increased susceptibility to the bactericidal activity of serum with CRP [25]. From this, there can be a selection for licA, depending on where the bacteria are growing, that is potentially due to CRP.

The ability of NTHi to adhere to and invade the epithelial lining in the upper airway is an important step in colonization and persistence. Subsets of LOS glycoforms containing ChoP play an important role in adherence and entry into human bronchial epithelial cells, providing further evidence of why a selection for licA might be expected. Infected monolayers were analyzed by confocal microscopy and showed that ChoP+-NTHi cells colocalized with the platelet-activating factor (PAF) receptor [26]. Further evidence demonstrated that NTHi binding to the PAF receptor initiates receptor coupling to a PTX-sensitive heterotrimeric G protein complex followed by cell entry [12].

The role of LOS phosphorylcholine in colonization and pathogenesis has also been investigated with Histophilus somni in cattle. H. somni is among the primary pathogens responsible for respiratory disease and other systemic infections in cattle. H. somni LOS also undergoes antigenic phase variation. Elsewai-fi et al found that ChoP expression was associated with colonization in the respiratory tract of compromised calves, whereas systemic dissemination was associated with the loss of ChoP expression [27]. ChoP was not required for adherence to bovine turbinate cells but could play a role in pathogenesis of H. somni by binding to the PAF receptor [27, 28].

The igaB gene encodes a type I IgA1 protease that cleaves human IgA1 in the hinge region of the α-heavy chain and is present in essentially all of the H. influenzae genomes [29, 30]. Some strains of H. influenzae isolated from patients with COPD have acquired another IgA protease gene. The gene, designated igaB, may have been acquired from N. meningitidis, based on flanking transposase sequences and sequence homology [29]. In strain 2019, the igaB gene is phase variable, and igaB was initially predominantly phase off. There was a progressive change to phase on over the course of the experiment; this difference was statistically significant at day 6. IgA1 is the most abundant immunoglobulin associated with human upper respiratory tract mucosal surfaces, so iga (igaA or igaB) would be important for colonization and persistence [30]. Differences in IgA1 protease levels between clinical and carriage strains have been observed, and levels were significantly higher in the clinical strains [30]. Human milk lactoferrin can inactivate the IgA1 protease by extracting the preprotein from the cell membranes [31]. The igaB variable region works well to differentiate H. influenzae and Haemophilus haemolyticus [32]. This could be important clinically because β-hemolysis by H. haemolyticus is a poor indicator of species separation [33].

Our study showed that there was greater variability in lexA as compared to the inoculated strain than with other genes investigated. The difference was not significant, but a definite population of strains had shifted from phase off to phase on. There is evidence that lexA is somehow necessary for the phenotypic expression of lexB. Mutant constructs of lexA and lexB have been made. A lexB mutant produces a truncated LOS, because lexB is required to add the second β-glucose in the oligosaccharide extension on the first heptose. lexA mutants gave the same LOS phenotype, observed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and colony immunoblot analyses, as the lexB mutants. This showed that lexA must be in frame for transcription of lexB [34].

In summary, we examined 13 phase-variable genes from samples obtained from an experimental human colonization study for a shift in their phase-on and phase-off status over the 6-day course of colonization. licA and igaB, which encode phosphorylcholine kinase and IgA protease, respectively, showed a significant difference from baseline over the course of the study. This change was from almost completely phase off to a significant portion of the population being phase on. This correlated with other studies showing there is a selection for licA on (ChoP+) and igaB on in the airway for H. influenzae to colonize and persist during infections [15, 20].

Supplementary Data

Supplementary materials are available at The Journal of Infectious Diseases online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyrighted. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgment. We thank Ms Margaret Ketterer for excellent technical assistance.

Financial support. This work was supported by the National Institutes of Health (grants AI024616, AI30040, and NCATS UL1 TR00442) and a Helen C. Levitt Endowed Annual Visiting Professorship (to M. P. J.).

Potential conflicts of interest. All authors: No reported conflicts.
All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References