High Genetic Diversity of Nontypeable Haemophilus influenzae Isolates from Two Children Attending a Day Care Center

Nathan C. LaCross, Carl F. Marrs, Mayuri Patel, Sara A. Sandstedt, and Janet R. Gilsdorf

Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, and Department of Pediatrics and Communicable Diseases, University of Michigan Medical School, Ann Arbor, Michigan

Received 16 May 2008/Returned for modification 11 July 2008/Accepted 13 September 2008

Twenty-one nontypeable Haemophilus influenzae (NTHi) isolates from the throats of two healthy children were genotyped by multilocus sequence typing. Nine unique sequence types (STs) were identified. These STs were scattered throughout the phylogenetic tree of reported NTHi STs, demonstrating the high level of NTHi diversity found in colonized children.

Nontypeable Haemophilus influenzae (NTHi) bacteria live almost exclusively in the human pharynx (9). The NTHi carriage rate among children varies between 25 and 81%; this wide distribution has been associated with differences in proximity to other children (e.g., attending a day care center), antibiotic use, and exposure to secondhand smoke (1, 5, 21). In addition to asymptomatic carriage, NTHi organisms may cause a variety of respiratory infections, including otitis media, sinusitis, bronchitis, and pneumonia. Recent studies have shown that NTHi colonization is an active, dynamic process. Dhoooge et al. (2) and Samuelson et al. (20) analyzed banding patterns of arbitrarily primed PCR fragments and restriction fragments, respectively, from colonizing NTHi strains obtained over 1 to 2 months from the same individuals and demonstrated rapid strain turnover. Additional studies used similar molecular methods to demonstrate that children (5, 21) and adults (13, 16) are often simultaneously colonized with multiple strains of NTHi.

The aim of this study was to further explore the apparent diversity of colonizing NTHi by characterizing 46 putative NTHi isolates from the throats of two healthy children (identification no. 22 and 26) attending the same day care center and reported in a prior study (21), using multilocus sequence typing (MLST), a DNA sequence-based typing system. Four throat swabs had been collected from each child, once a week for 3 weeks. Each swab was streaked on two chocolate agar plates, and five colonies were selected from each plate for a maximum of 10 colonies from each child at each collection period (21).

Isolates from the two children were confirmed H. influenzae strains by the X-factor (heme) requirement, using the porphyrin test (5); the absence of hemolysis, using horse blood agar plates (Remel, Lenexa, KS); and P6 outer membrane protein characterization by immunoblot assay using the 7F3 monoclonal antibody (kindly provided by Timothy Murphy) (13), which binds an epitope of P6 that is highly specific to H. influenzae bacteria (15). Genomic DNA was prepared as previously described (8, 10), and the presence of iga, which encodes immunoglobulin A1 protease, was determined by PCR amplification of an 855-bp conserved fragment (F, TGAATAACGAGGGGCAATATAC; R, TCACCAGACTTAATCACTGAAT) and subsequent visualization on 1% agarose gels stained with ethidium bromide. For the purposes of this study, isolates that possessed iga, reacted to the 7F3 antibody, were nonhemolytic, and did not produce porphyrin were designated NTHi isolates. This is similar to the definition used by Murphy et al. (14) to differentiate between H. influenzae and nonhemolytic variants of H. haemolyticus. Based on the above criteria, 21 of the 46 isolates were identified as NTHi (Table 1).

MLST was used to genotype the 21 NTHi strains, as described by Meats et al. (12). A total of nine unique sequence types (STs) were identified: three from child 22 and six from child 26 (Table 1). All three STs from child 22 and two STs from child 26 had been previously identified and entered into the MLST database (STs 2, 142, 146, 166, and 176); the remaining four STs were novel and have been entered into the MLST database (STs 355, 357, 358, and 359). Seven of the nine STs differed from all other STs in this study by at least four of seven loci. The population recombination rate, \( \rho \), of each of the seven MLST loci was estimated using the standard likelihood coalescent approach implemented by LDHat version 2.1 software (11). Estimates of \( \rho \) varied considerably across the different loci, ranging from a low of zero for the fucK and recA genes to a high of 94 for the mdh gene (Table 2). These data are similar to the values found by Pérez-Losada et al. (18) when all H. influenzae strains in the MLST database were analyzed (as of January 2004) and indicate that the rate of recombination is not uniform across the NTHi genome.

Analysis of the MLST data using eBURST version 3 software was performed as described previously (7). Sequence types 357 and 358 from child 26 are single-locus variants and...
form a small group, while the remaining STs are singletons (Fig. 1A), which mirrors the pattern seen in an eBURST analysis of all 321 NTHi strains residing in the MLST database (Fig. 1B). In contrast, eBURST analysis of all 260 type b strains in the MLST database reveals that one large clonal complex predominates (Fig. 1C). Like NTHi strains, the type b strains represent a very wide range of isolation dates and geographic locations. Overall, this suggests that NTHi strains are considerably more diverse than type b strains and show a less clonal pattern of descent, as has been suggested previously using different analytic techniques (17, 19).

The relationship between the NTHi STs found in this study was also examined phylogenetically. Erwin et al. (4) constructed a maximum parsimony majority-rule consensus tree from 4,545 equally most parsimonious trees, using the concatenated MLST sequences from all H. influenzae sequence types in the MLST database, including the nine STs found in this study (Fig. 2). These nine NTHi STs are scattered in the tree, implying that they are more related to other STs within the MLST database than to each other.

The advent of DNA sequence-based typing methods has aided the study of bacterial diversity. While MLST samples only a minute fraction of the genome and provides no information regarding differential gene content or arrangement, it is less resource intensive than full-genome sequencing and provides useful information, such as analyses of sequence divergence, recombination, phylogenetics, and population structure (6), beyond that available from non-sequence-based methods. Previous studies have demonstrated the utility of MLST in assessing diversity and relationships between isolates of H. influenzae. Erwin et al. (4) used MLST to type a number of commensal, otitis media-associated, and invasive NTHi iso-

---

**TABLE 1. Characteristics of Haemophilus isolates**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Isolate</th>
<th>ST</th>
<th>iga</th>
<th>P6</th>
<th>Hemolysis</th>
<th>Porphyrin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.1–21</td>
<td>176</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>2</td>
<td>22.2–22</td>
<td>176</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>3</td>
<td>22.3–21</td>
<td>176</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>4</td>
<td>22.4–21</td>
<td>176</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

a STs, as determined by MLST, were assigned to those isolates with characteristics indicative of H. influenzae, as follows: iga, presence (+) of the iga gene as determined by PCR; P6, reactivity (+) to the H. influenzae strain-specific 7F3 monoclonal antibody epitope; Hemolysis, hemolytic activity (+) during growth on horse blood agar plates; Porphyrin, presence (+) of fluorescence under UV light after incubation with δ-aminolevulinic acid. NA, not applicable.

**TABLE 2. Population recombination rate for each MLST locus**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Recombination rate (p)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>adk</td>
<td>6</td>
<td>0.106</td>
</tr>
<tr>
<td>aprG</td>
<td>9</td>
<td>0.516</td>
</tr>
<tr>
<td>frdB</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>fucK</td>
<td>0</td>
<td>0.896</td>
</tr>
<tr>
<td>mdh</td>
<td>94</td>
<td>0.096</td>
</tr>
<tr>
<td>pg1</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>recA</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

a p, population recombination rate estimated by LDhat software.
b P values were estimated using the likelihood permutation test in LDhat software.
FIG. 1. eBURST analysis of the nine STs from children 22 and 26 (A), all NTHi STs in the MLST database (STs from child 22 are circled, while STs from child 26 are in squares) and all type b STs in the MLST database (B), (C) are shown. STs differing from another ST by only one of seven loci are connected by lines and form a group. The size of the circles is proportional to the abundance of the corresponding STs in the data set, and the relative placement of unconnected STs is random.
The week-to-week genetic variability of the NTHi strains in the present study suggests the existence of a pool of NTHi strains within each child and the inadequacy of the bacterial isolation technique to identify all STs present during a single sampling procedure. Alternatively, this larger pool may be within the community, and the lack of persistence and relatedness of the collected isolates reflects the rapid turnover of strains within each child. A third, though relatively unlikely, option is that the high diversity observed for this study is the result of extremely rapid evolution. The true scenario may be a combination of these possibilities, in which individuals quickly gain and lose rapidly evolving NTHi strains, and standard collection techniques are incapable of providing an accurate rate survey of the genetic differences present among NTHi bacteria residing within the nasopharynges. Furthermore, the genetic diversity observed using MLST may result in an underestimation of whole-genome diversity as the housekeeping genes used for the analysis evolve relatively slowly. Thus, strains with identical MLST STs may vary considerably in other, faster-evolving gene regions.

This study was funded by grants DC05840 and HL083893 to J.R.G.

REFERENCES


