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

Nathan C. LaCross,¹ Carl F. Marrs,¹ Mayuri Patel,² Sara A. Sandstedt,¹ and Janet R. Gilsdorf^{1,2*}

Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan,¹ and Department of Pediatrics and Communicable Diseases, University of Michigan Medical School, Ann Arbor, Michigan²

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. Dhooze 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 de-

scribed (8, 10), and the presence of *iga*, which encodes immunoglobulin A1 protease, was determined by PCR amplification of an 855-bp conserved fragment (F, TGAATAACGAGGGG CAATATAAC; R, TCACCGCACTTAATCACTGAAT) 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, 165, 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, and three STs differed by six of seven loci. This lack of shared alleles between strains implies that the majority of STs collected are relatively unrelated. Seven of the nine STs were present only during a single sampling period. Furthermore, no ST was shared between the two children, despite attendance at the same day care center, although other children from the center exhibited strain sharing as defined by pulsed-field gel electrophoresis (21).

The population recombination rate, ρ , 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 ρ 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

* Corresponding author. Mailing address: Department of Pediatrics and Communicable Diseases, L2225 Women's Hospital, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0244. Phone: (734) 763-2440. Fax: (734) 936-7635. E-mail: gilsdorf@umich.edu.

[▽] Published ahead of print on 24 September 2008.

TABLE 1. Characteristics of *Haemophilus* isolates^a

Sample	Child 22						Child 26					
	Isolate	ST	<i>iga</i>	P6	Hemolysis	Porphyrin	Isolate	ST	<i>iga</i>	P6	Hemolysis	Porphyrin
1	22.1–21	176	+	+	–	–	26.1–21	359	+	+	–	–
	22.1–22	176	+	+	–	–	26.1–22	359	+	+	–	–
	22.1–23	176	+	+	–	–	26.1–23	355	+	+	–	–
	22.1–24	2	+	+	–	–						
	22.1–25	NA	–	–	+	+						
2	22.2–21	176	+	+	–	–	26.2–25	142	+	+	–	–
	22.2–22	2	+	+	–	–	26.2–21	NA	–	+	–	–
	22.2–23	2	+	+	–	–	26.2–22	NA	–	+	–	–
	22.2–25	2	+	+	–	–	26.2–23	NA	–	+	–	–
	22.2–24	NA	–	–	+	–	26.2–24	NA	–	+	–	–
							26.2–26	NA	–	+	–	–
							26.2–27	NA	–	+	–	–
							26.2–28	NA	–	+	–	–
3	22.3–21	NA	–	–	+	+	26.3–23	357	+	+	–	–
	22.3–22	NA	–	–	–	+	26.3–27	358	+	+	–	–
							26.3–22	NA	+	–	+	–
							26.3–24	NA	–	+	–	–
							26.3–28	NA	–	+	–	–
							26.3–25	NA	–	–	–	–
							26.3–26	NA	–	–	–	–
4	22.4–21	165	+	+	–	–	26.4–24	146	+	+	–	–
	22.4–22	165	+	+	–	–	26.4–21	NA	–	+	–	–
	22.4–23	165	+	+	–	–	26.4–22	NA	–	+	–	–
	22.4–24	165	+	+	–	–	26.4–29	NA	–	+	–	–
	22.4–25	165	+	+	–	–	26.4–210	NA	–	–	+	–
	22.4–26	165	+	+	–	–	26.4–23	NA	–	–	–	–
							26.4–26	NA	–	–	–	–
							26.4–25	NA	–	–	–	–
							26.4–27	NA	–	–	–	–
							26.4–28	NA	–	–	–	–

^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.

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 2. Population recombination rate for each MLST locus

Gene	Recombination rate (ρ) ^a	<i>P</i> value ^b
<i>adk</i>	6	0.106
<i>atpG</i>	9	0.516
<i>frdB</i>	14	0
<i>fucK</i>	0	0.896
<i>mdh</i>	94	0.096
<i>pgi</i>	59	0
<i>recA</i>	0	0

^a ρ , 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.

FIG. 2. Maximum parsimony majority-rule consensus tree constructed by Erwin et al. (4) from 4,545 equally most parsimonious trees, using all *H. influenzae* STs in the MLST database. The nine STs identified in this study are marked, and the number in parentheses indicates the child from which that ST was collected.

lates and, using an unweighted-pair group method using average linkages dendrogram, demonstrated that 8 of 11 strains displaying the same biotype (V, which produces indole and has ornithine decarboxylase, but not urease, activity) (9) formed a distinct monophyletic group and that strains containing the *hmw*-related sequence were separated from those strains lacking *hmw* (3). In comparing MLST to a typing strategy based on the 16S rRNA gene sequence, Sacchi et al. (19) found that both methods corroborate the clonal nature of typeable *H. influenzae* and the lack of clonality of NTHi isolates.

The results of the present study demonstrate the high level of genetic diversity found in NTHi isolates that colonize the pharynx of healthy children, even within single individuals, and indicate that results obtained from small samples (e.g., this study) mirror those from much larger collections (e.g., the MLST database). The diversity observed for this study using MLST corroborates that seen by pulsed-field gel electrophore-

sis analysis of the same strains (21) and further discriminates between *H. influenzae* and related species, as well as comprehensively examines genetic diversity in terms of relatedness (with eBURST and phylogenetics) and population recombination rates.

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 survey of the genetic differences present among NTHi bacteria residing within the nasopharynx. 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

- Bou, R., A. Dominguez, D. Fontanals, I. Sanfeliu, I. Pons, J. Renau, V. Pineda, E. Lobera, C. Latorre, M. Majo, and L. Salleras for the working group on invasive disease caused by *Haemophilus influenzae*. 2000. Prevalence of *Haemophilus influenzae* pharyngeal carriers in the school population of Catalonia. *Eur. J. Epidemiol.* **16**:521–526.
- Dhooge, I., M. Vaneechoutte, G. Claeys, G. Verschraegen, and P. Van Cauwenberge. 2000. Turnover of *Haemophilus influenzae* isolates in otitis-prone children. *Int. J. Pediatr. Otorhinolaryngol.* **54**:7–12.
- Erwin, A. L., K. L. Nelson, T. Mhlanga-Mutangadza, P. J. Bonthuis, J. L. Geelhood, G. Morlin, W. C. Unrath, J. Campos, D. W. Crook, M. M. Farley, F. W. Henderson, R. F. Jacobs, K. Muhlemann, S. W. Satola, L. van Alphen, M. Golomb, and A. L. Smith. 2005. Characterization of genetic and phenotypic diversity of invasive nontypeable *Haemophilus influenzae*. *Infect. Immun.* **73**:5853–5863.
- Erwin, A. L., S. A. Sandstedt, P. J. Bonthuis, J. L. Geelhood, K. L. Nelson, W. C. Unrath, M. A. Diggle, M. J. Theodore, C. R. Pleatman, E. A. Mothershead, C. T. Sacchi, L. W. Mayer, J. R. Gilsdorf, and A. L. Smith. 2008. Analysis of genetic relatedness of *Haemophilus influenzae* isolates by multilocus sequence typing. *J. Bacteriol.* **190**:1473–1483.
- Farjo, R. S., B. Foxman, M. J. Patel, L. Zhang, M. M. Pettigrew, S. I. McCoy, C. F. Marrs, and J. R. Gilsdorf. 2004. Diversity and sharing of *Haemophilus influenzae* strains colonizing healthy children attending day-care centers. *Pediatr. Infect. Dis. J.* **23**:41–46.
- Feil, E. J. 2004. Small change: keeping pace with microevolution. *Nat. Rev. Microbiol.* **2**:483–495.
- Feil, E. J., B. C. Li, D. M. Aanensen, W. P. Hanage, and B. G. Spratt. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J. Bacteriol.* **186**:1518–1530.
- Juliao, P. C., C. F. Marrs, J. Xie, and J. R. Gilsdorf. 2007. Histidine auxotrophy in commensal and disease-causing nontypeable *Haemophilus influenzae*. *J. Bacteriol.* **189**:4994–5001.
- Kilian, M. 2005. Genus *Haemophilus*, p. 883–904. In G. M. Garrity, D. J. Brenner, N. R. Krieg, and J. T. Staley (ed.), *Bergey's manual of systematic bacteriology*, 2nd ed., vol. 2. Springer-Verlag, New York, NY.
- McCrea, K. W., J. Xie, N. LaCross, M. Patel, D. Mukundan, T. F. Murphy, C. F. Marrs, and J. R. Gilsdorf. 2008. Relationships of nontypeable *Haemophilus influenzae* strains to hemolytic and nonhemolytic *Haemophilus haemolyticus* strains. *J. Clin. Microbiol.* **46**:406–416.
- McVean, G., P. Awadalla, and P. Fearnhead. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics* **160**:1231–1241.
- Meats, E., E. J. Feil, S. Stringer, A. J. Cody, R. Goldstein, J. S. Kroll, T.

- Popovic, and B. G. Spratt.** 2003. Characterization of encapsulated and non-encapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. *J. Clin. Microbiol.* **41**:1623–1636.
13. **Mukundan, D., Z. Ecevit, M. Patel, C. F. Marrs, and J. R. Gilsdorf.** 2007. Pharyngeal colonization dynamics of *Haemophilus influenzae* and *Haemophilus haemolyticus* in healthy adult carriers. *J. Clin. Microbiol.* **45**:3207–3217.
 14. **Murphy, T. F., A. L. Brauer, S. Sethi, M. Kilian, X. Cai, and A. J. Lesse.** 2007. *Haemophilus haemolyticus*: a human respiratory tract commensal to be distinguished from *Haemophilus influenzae*. *J. Infect. Dis.* **195**:81–89.
 15. **Murphy, T. F., C. Kirkham, and D. J. Sikkema.** 1992. Neonatal, urogenital isolates of biotype 4 nontypeable *Haemophilus influenzae* express a variant P6 outer membrane protein molecule. *Infect. Immun.* **60**:2016–2022.
 16. **Murphy, T. F., S. Sethi, K. L. Klingman, A. B. Brueggemann, and G. V. Doern.** 1999. Simultaneous respiratory tract colonization by multiple strains of nontypeable *Haemophilus influenzae* in chronic obstructive pulmonary disease: implications for antibiotic therapy. *J. Infect. Dis.* **180**:404–409.
 17. **Musser, J. M., J. S. Kroll, E. R. Moxon, and R. K. Selander.** 1988. Clonal population structure of encapsulated *Haemophilus influenzae*. *Infect. Immun.* **56**:1837–1845.
 18. **Perez-Losada, M., E. B. Browne, A. Madsen, T. Wirth, R. P. Viscidi, and K. A. Crandall.** 2006. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infect. Genet. Evol.* **6**:97–112.
 19. **Sacchi, C. T., D. Alber, P. Dull, E. A. Mothershed, A. M. Whitney, G. A. Barnett, T. Popovic, and L. W. Mayer.** 2005. High level of sequence diversity in the 16S rRNA genes of *Haemophilus influenzae* isolates is useful for molecular subtyping. *J. Clin. Microbiol.* **43**:3734–3742.
 20. **Samuelson, A., A. Freijd, J. Jonasson, and A. A. Lindberg.** 1995. Turnover of nonencapsulated *Haemophilus influenzae* in the nasopharynxes of otitis-prone children. *J. Clin. Microbiol.* **33**:2027–2031.
 21. **St. Sauver, J., C. F. Marrs, B. Foxman, P. Somsel, R. Madera, and J. R. Gilsdorf.** 2000. Risk factors for otitis media and carriage of multiple strains of *Haemophilus influenzae* and *Streptococcus pneumoniae*. *Emerg. Infect. Dis.* **6**:622–630.