

Prevalence and Distribution of the *hmw* and *hia* Genes and the HMW and Hia Adhesins among Genetically Diverse Strains of Nontypeable *Haemophilus influenzae*

JOSEPH W. ST. GEME III,^{1*} VINI V. KUMAR,¹ DAVID CUTTER,¹ AND STEPHEN J. BARENKAMP²

Edward Mallinckrodt Department of Pediatrics and Department of Molecular Microbiology, Washington University School of Medicine, and Division of Infectious Diseases, St. Louis Children's Hospital, St. Louis, Missouri 63110,¹ and Department of Pediatrics, St. Louis University School of Medicine, and The Pediatric Research Institute, Cardinal Glennon Children's Hospital, St. Louis, Missouri 63104²

Received 20 May 1997/Returned for modification 28 July 1997/Accepted 8 October 1997

Nontypeable *Haemophilus influenzae* is a common cause of human disease and initiates infection by colonizing the upper respiratory tract. In previous work we identified high-molecular-weight adhesins referred to as HMW1 and HMW2, expressed by nontypeable strain 12, and determined that most strains of nontypeable *H. influenzae* express one or two antigenically related proteins. More recently, we determined that some strains lack HMW1- and HMW2-like proteins and instead express an adhesin called Hia. In the present study, we determined the prevalence and distribution of the *hmw* and *hia* genes in a collection of 59 nontypeable strains previously characterized in terms of genetic relatedness. Based on Southern analysis, 47 strains contained sequences homologous to the *hmw1* and *hmw2* genes and nine strains contained homologs to *hia*. No strain harbored both *hmw* and *hia*, and three strains harbored neither. Although the *hmw* and *hia* genes failed to define distinct genetic divisions, the *hmw*-deficient strains formed small clusters or lineages within the larger population structure. Additional analysis established that the *IS1016* insertion element was uniformly absent from strains containing *hmw* sequences but was present in two-thirds of the *hmw*-deficient strains. As *IS1016* is associated with the capsule locus (*cap*) in most encapsulated strains of *H. influenzae*, we speculate that *hmw*-deficient nontypeable strains evolved more recently from an encapsulated ancestor.

Haemophilus influenzae is a gram-negative bacterium that commonly inhabits the human upper respiratory tract. Isolates of *H. influenzae* are subdivided into encapsulated and nonencapsulated forms (11). Encapsulated strains express one of six structurally and antigenically distinct capsular polysaccharides, designated serotypes a to f (11). Nonencapsulated strains are defined on the basis of their failure to agglutinate with typing antisera against the known *H. influenzae* capsular structures and are referred to as nontypeable (11).

Analysis by multilocus enzyme electrophoresis indicates that encapsulated strains are clonal and can be segregated into genetically related clusters, which are grouped into two major phylogenetic divisions (9, 10). Division I includes clusters of serotypes a, b, c, d, and e strains, and division II includes clusters of serotypes a, b, and f strains. All encapsulated strains of *H. influenzae* have genes necessary for encapsulation (*cap* genes). These genes are organized into functionally distinct regions, with serotype-specific DNA flanked by DNA common to all serotypes (7). In division I strains, the *cap* gene cluster lies between direct repeats of an insertion element referred to as *IS1016* (5). Division II strains also contain one or more copies of *IS1016*, but in these strains *IS1016* is not associated with *cap* genes (5).

The population structure of nontypeable *H. influenzae* is less well defined. Musser et al. examined 65 epidemiologically distinct isolates by multilocus enzyme electrophoresis and found considerable heterogeneity, with each isolate corresponding to

a unique electromorph type (8). Furthermore, comparison with 177 type b isolates revealed no overlap, suggesting that nontypeable strains are not simply phenotypic variants of common type b strains (8). More recently we studied a series of 123 pharyngeal isolates of nontypeable *H. influenzae* collected from healthy 3-year-old Finnish children (19). Among these isolates, one was a capsule-deficient type b strain. Interestingly, of the remaining 122 isolates, 31% hybridized with a probe containing the type b *cap* locus. These findings suggested that a subgroup of nontypeable strains might be more recent descendants of an encapsulated ancestor.

In previous work directed at understanding the mechanism by which strains of nontypeable *H. influenzae* colonize the respiratory tract, we identified two high-molecular-weight adhesins referred to as HMW1 and HMW2 (2, 18). Both of these proteins are expressed by nontypeable *H. influenzae* 12 and show significant amino acid sequence similarity to one another (2). Preliminary studies suggested that most nontypeable strains express one or two proteins that are antigenically related to HMW1 and HMW2. Further analysis revealed that strains lacking HMW1- and HMW2-like proteins remain capable of efficient attachment to cultured epithelial cells (3). Consistent with this observation, we recently identified another nontypeable *H. influenzae* high-molecular-weight adhesin referred to as Hia (3). Based on Southern hybridization studies, encapsulated strains of *H. influenzae* lack *hmw1*- and *hmw2*-like genes but uniformly possess an allelic variant of *hia* (16). In serotype b strains, the counterpart to *hia* is associated with expression of short, thin surface appendages referred to as fibrils and is called *hsf* (13, 16).

In the present study, we sought to determine the prevalence of the *hmw* and *hia* genes in a collection of nontypeable *H. influenzae* strains. In addition, we addressed the hypothesis

* Corresponding author. Mailing address: Department of Molecular Microbiology, Washington University School of Medicine, 660 South Euclid Ave., Box 8230, St. Louis, MO 63110. Phone: (314) 362-5401. Fax: (314) 362-1232. E-mail: stgeme@borcim.wustl.edu.

that strains can be separated into genetically distinct divisions according to the adhesin they possess. Finally, we examined the possibility that strains lacking *hmw* genes are more likely to contain *cap*-specific sequences or *IS1016* elements and might be more recent derivatives of an encapsulated line.

In performing this study, we exploited the collection of nontypeable *H. influenzae* strains previously characterized by Musser et al. and defined in terms of genetic relatedness (8). Four biotype 4 strains originally isolated from the blood of newborn infants or women with obstetrical infections were excluded because DNA hybridization studies indicate that they represent a separate *Haemophilus* species (12), a fifth strain was excluded because it is now known to express a serotype e capsule (14), and a sixth strain was missing from the collection. The remaining 59 strains included in our analysis were recovered between 1937 and 1985 during episodes of otitis media, bacteremia, or sinusitis in children throughout the United States and Canada (8). Control strains included nontypeable *H. influenzae* 12 and 11 and *H. influenzae* type b strain Eagan. Strain 12 is the strain from which the *hmw1* and *hmw2* genes were originally cloned (2), and strain 11 is the strain from which *hia* was originally cloned (3). Strain Eagan contains an allelic variant of *hia* (referred to as *hsf*) and an intact type b *cap* locus (13, 15).

In initial experiments, we examined the 59 strains by Southern analysis, probing with a 3.2-kb *SpeI-EcoRI* fragment that corresponds to the promoter and 5' coding sequence of the *H. influenzae* 12 *HMW1* structural gene (*hmw1A*). Hybridization conditions were as previously described (17). As summarized in Fig. 1, 47 strains demonstrated specific high-molecular-weight bands of hybridization. In some cases, only a single band was discernible, while in others two bands were clearly present. Of note, in strain 12 the chromosomal fragments harboring *hmw1A* and *hmw2A* comigrated, giving the impression of a single band of hybridization. Figure 2A depicts a representative sampling of these strains. Nontypeable *H. influenzae* 11, known to lack *hmw* genes, served as a negative control.

Next we probed chromosomal DNA from the 59 strains with a 1.6-kb *SspI-StyI* fragment that corresponds to an intragenic region of the *H. influenzae* 11 *hia* locus. Nontypeable *H. influenzae* 12, lacking an *hia* homolog, served as a negative control, while strain 11 was the positive control. As summarized in Fig. 1 and illustrated in Fig. 2B, nine strains demonstrated specific bands of hybridization. Interestingly, the 47 strains with *hmw* sequences and the nine strains with an *hia* homolog were mutually exclusive but failed to segregate into genetically distinct divisions (Fig. 1). Nevertheless, the strains containing an *hia* homolog formed small clusters within the larger population structure.

To gain additional insight into the evolutionary relationship between typeable and nontypeable strains of *H. influenzae*, we performed Southern analysis with pUO38, a pBR322 derivative that contains a complete set of the *cap* genes, including one copy of *IS1016*, from a phylogenetic division I *H. influenzae* type b strain (6). We refer to this set of genes, which includes sequences common to all encapsulated *H. influenzae* strains, as the *cap* locus. Of the 47 strains with *hmw* genes, none demonstrated hybridization with pUO38 (Fig. 1). In contrast, 8 of the remaining 12 strains, including 6 of the 9 strains with an *hia* gene, showed a single band of hybridization (Fig. 3). Further analysis revealed that in all strains, hybridization with *IS1016* accounted entirely for the hybridization seen with pUO38 (not shown). *H. influenzae* type b strain Eagan served as a control for these blots and gave the predicted pattern of hybridizing bands, including bands of 2.1, 2.7, 4.4, 9.0, 10.7, and 16 kb with pUO38 and bands of 9.0, 10.7, and 16 kb with

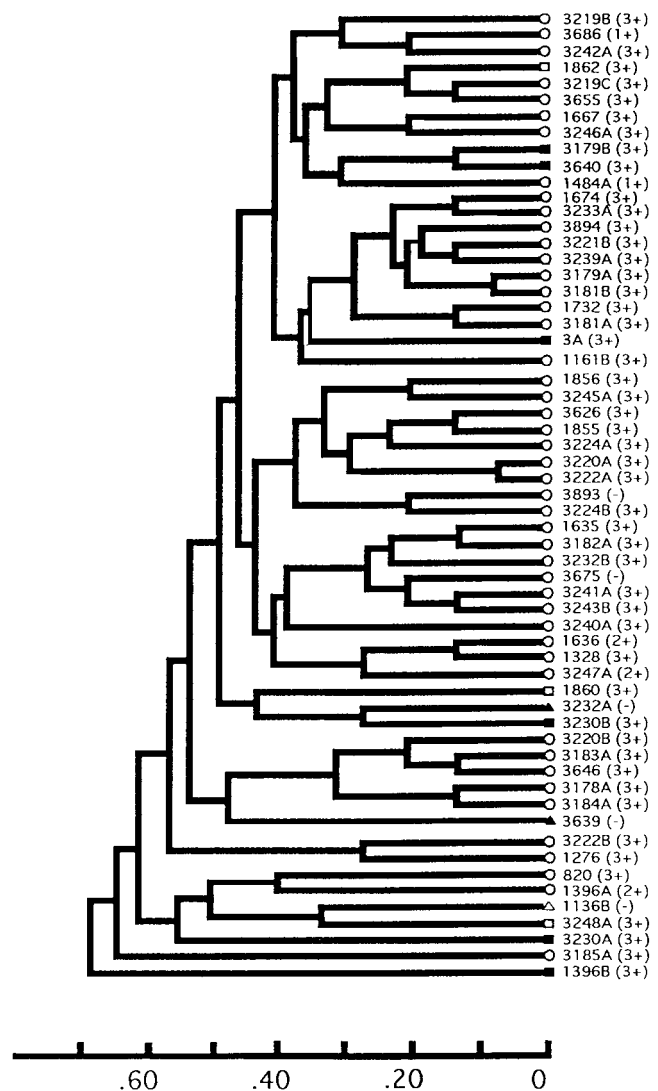


FIG. 1. Genetic relationships between the 59 strains examined in this study. The dendrogram was generated previously by Musser et al. (8) by the average-linkage method of clustering from a matrix of coefficients of genetic distance, based on 15 metabolic enzymes. Circles indicate strains that hybridized with *hmw* but not *hia*, squares indicate strains that hybridized with *hia* but not *hmw*, and triangles indicate strains that failed to hybridize with either *hia* or *hmw*. Blackened symbols indicate strains that hybridized with pUO38 and *IS1016*. The numbers and letters to the right of the symbols are strain designations. Numbers in parentheses refer to level of adherence.

IS1016 (5, 19). Hybridization studies performed with radiolabeled pBR322 were negative for all strains (not shown).

To determine whether the presence of *hmw* and *hia* sequences correlated with specific protein expression, whole-cell lysates of all 59 strains were examined by Western immunoblot analysis with a rabbit polyclonal antiserum reactive with HMW1 and HMW2 and a rabbit polyclonal antiserum raised against Hia. Forty-five of the 47 strains with *hmw* genes expressed one or two proteins that reacted with serum 25G (2), with the reactive proteins ranging in size between ~105 and ~160 kDa. Figure 4A shows a representative set of these strains. Among the 12 strains that failed to hybridize with *hmw1A*, none expressed proteins reactive with serum 25G. Control strains for these immunoblots included strain 12 (ex-

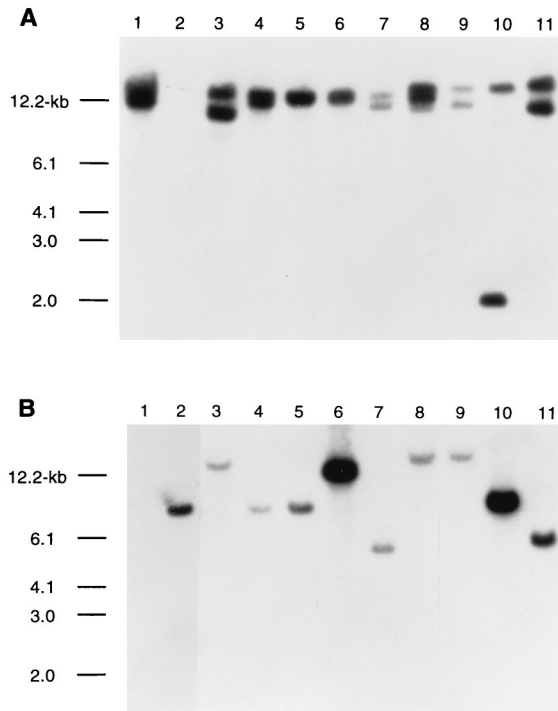


FIG. 2. Southern blotting performed with probes for *hmw* sequences and *hia*. Chromosomal DNA was digested with *Bgl*II, separated by agarose electrophoresis, and subjected to Southern analysis. (A) Southern analysis with a probe corresponding to the promoter and 5' coding sequence of *hmw1A*. Strains by lane: 1, 12; 2, 11; 3, 3219B; 4, 3686; 5, 3242A; 6, 3219C; 7, 3655; 8, 1667; 9, 3246A; 10, 1484A; 11, 1674. Less chromosomal DNA was inadvertently loaded in lanes 7 and 9. (B) Southern analysis with a probe that represents an intragenic fragment of *hia*. Strains by lane: 1, 12; 2, 11; 3, 1862; 4, 3179B; 5, 3640; 6, 3A; 7, 1860; 8, 3230B; 9, 3248A; 10, 3230A; 11, 1396B. Comparable quantities of DNA were loaded in each lane.

pressing HMW1 and HMW2) and strain 11 (lacking HMW proteins).

Analysis of the 59 strains with serum 36B (3), directed against *Hia*, revealed reactivity with all 9 strains containing an *hia* gene but with none of the remaining 50 isolates. Interestingly, in six of the nine reactive strains, multiple high-molecular-weight bands were detected, suggesting formation of mul-

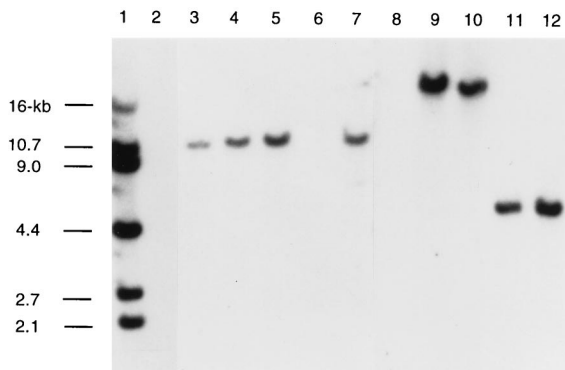


FIG. 3. Southern blotting performed with radiolabeled pUO38. Chromosomal DNA was digested with *Eco*RI, separated by agarose electrophoresis, and subjected to Southern analysis. Strains by lane: 1, Eagan; 2, 1862; 3, 3179B; 4, 3640; 5, 3A; 6, 1860; 7, 3230B; 8, 3248A; 9, 3230A; 10, 1396B; 11, 3232A; 12, 3639.

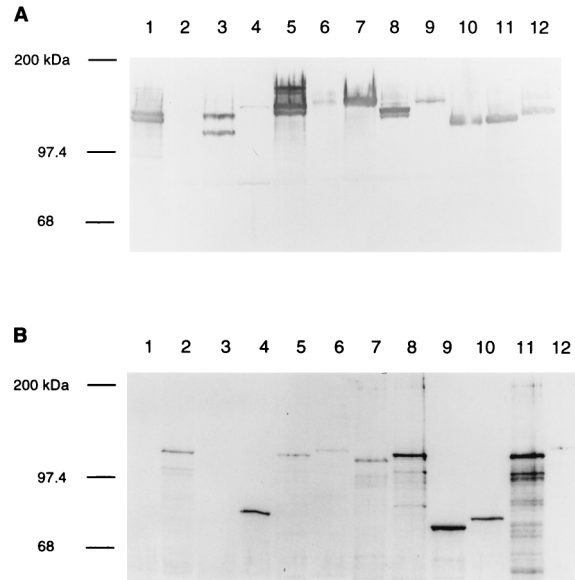


FIG. 4. Western immunoblotting performed with serum 25G against HMW1 and HMW2 or serum 36B against *Hia*. Whole-cell lysates were prepared, and proteins were separated on a 7.5% polyacrylamide gel. (A) Western analysis with serum 25G. Strains by lane: 1, 12; 2, 11; 3, 3219B; 4, 3686; 5, 3242A; 6, 3219C; 7, 3655; 8, 1667; 9, 3246A; 10, 1484A; 11, 1674; 12, 3894. (B) Western analysis with serum 36B. Strains by lane: 1, 12; 2, 11; 3, 11*hia::kan*; 4, 1862; 5, 3179B; 6, 3640; 7, 3A; 8, 1860; 9, 3230B; 10, 3248A; 11, 3230A; 12, 1396B.

timers or aggregates (not shown). Formic acid treatment (4) resulted in elimination of the high-molecular-weight bands and generation of a predominant band ranging in size between ~80 and ~120 kDa (Fig. 4B). Control strains for these blots again included strain 12 and strain 11, along with a strain 11 mutant deficient in expression of *Hia*.

To examine whether expression of HMW or *Hia* proteins correlated with a capacity for adherence to human epithelial cells, we performed adherence assays with Chang epithelial cells. HMW1, HMW2, and *Hia* originally were identified as adhesins based on the ability to promote attachment to Chang cells. In the present study, adherence was measured by staining samples with Giemsa and then examining with light microscopy and counting the number of bacteria associated with 25 cells. Adherence was considered 3+ if the average number of bacteria per cell was greater than 25, 2+ if the average number was 10 to 25, and 1+ if the average number was 1 to 10. Strains were considered nonadherent if the average number of bacteria per cell was less than one. As summarized in Fig. 1, 45 of the 47 strains with *hmw* genes, including all 45 containing proteins reactive with serum 25G, demonstrated appreciable *in vitro* adherence. In 43 of these strains, the level of adherence was moderate to high (2+ to 3+). All nine of the strains with an *hia* homolog demonstrated high-level adherence (3+). Interestingly, the three strains that were negative by Southern blotting with both *hmw1A* and *hia* were nonadherent.

To summarize, Southern analysis of the 59 strains included in this study revealed three subsets, including 80% with an *hmw1A* or *hmw2A* homolog, 15% lacking *hmw1A* and *hmw2A* genes but harboring an *hia* homolog, and 5% lacking both *hmw* and *hia* sequences. Strains containing an *hmw* gene were uniformly devoid of IS1016, an insertion element associated with the *cap* locus in division I encapsulated *H. influenzae*. In contrast, nearly 70% of the *hmw*-deficient strains possessed one copy of this element.

Based on our earlier observations that a subset of nontypeable strains contain the *hia* gene, which is allelic to the *hsf* locus found in all encapsulated *H. influenzae* strains (3, 16), and that 31% of pharyngeal isolates of nontypeable *H. influenzae* from healthy 3-year-old Finnish children demonstrated hybridization with the *cap* locus (19), we hypothesized that strains possessing an *hia* homolog would contain *cap* sequences as well, reflecting more recent evolution from an encapsulated ancestor. Indeed, in the present study we found that six of nine strains that hybridized with *hia* also hybridized with pUO38. Further analysis revealed that in all six of these strains, the hybridization with pUO38 was due entirely to the presence of an *IS1016* element. It is possible that these strains experienced a deletion of capsule genes due to recombination between duplicated copies of *IS1016*, leaving one intact copy of *IS1016*. Such precedent exists in the form of the longtime laboratory strain Rd, which originally contained capsule genes flanked by *IS1016* elements and expressed a serotype d capsule but now contains a single copy of *IS1016* with no capsule genes and is nonencapsulated (5).

Examination of the genetic relatedness between the 47 strains with *hmw* sequences and the 12 strains without such sequences revealed the absence of separable divisions. Nevertheless, the *hmw*-deficient strains were found to form clusters within the larger population structure. For example, strains 3179B and 3640, both of which hybridized with *hia* and *IS1016*, are more closely related to each other than they are to other strains in the collection. Similarly, strains 1860 (containing *hia* [*hia*+]) but lacking *IS1016* [*IS1016*-]), 3232A (*hia*-, *IS1016*+), and 3230B (*hia*+, *IS1016*+) form a second cluster of closely related strains, and strains 1136B (*hia*-, *IS1016*-), 3248A (*hia*+, *IS1016*-), and 3230A (*hia*+, *IS1016*+) form a third such cluster. Consistent with the evidence that strains 3179B and 3640 are closely related and the possibility that they have evolved from the same encapsulated ancestor, in both strains an ~10-kb *Eco*RI fragment hybridizes with *IS1016*.

Division I and division II encapsulated *H. influenzae* strains are separated from one another by a genetic distance of 0.66 (9, 10). As shown in Fig. 1, the first 58 nontypeable strains in the collection we studied are separated by genetic distances that range between 0.08 and 0.64. Furthermore, the analysis by Musser et al. revealed that strains in positions 1 to 3 (strains 3219B, 3686, and 3242A [Fig. 1]) lie between type b strains belonging to clonal groups A1a and A2a, and strains in positions 4 to 22 (including strains 1862, 3179B, 3640, and 3A) lie between type b strains in clonal groups A2a and B1b (8, 9). Thus, 58 of the 59 nontypeable strains we examined, including 11 of the 12 *hmw*-deficient strains, are likely to be more closely related to division I encapsulated strains than they are to division II strains. One conclusion consistent with our observations and the dendrogram that exists for encapsulated strains is that strains 1862, 3179B, and 3640 all evolved from a type b ancestor. Similarly, superimposition of the dendrogram for encapsulated *H. influenzae* suggests that strain 3A might have evolved from a cluster B type b, type d, or type a strain; strains 1860, 3232A, and 3230B might have evolved from a cluster C type b or a cluster D type c strain; and strain 3639 might have evolved from a cluster E type c strain. The origins of strains 1136B, 3248A, 3230A, and 1396B are less clear, except that strain 1396B is more closely related to division II encapsulated strains, suggesting possible evolution from a type a, type b, or type f strain.

To examine the correlation between the presence of *hmw* sequences and expression of HMW protein, we performed Western immunoblot analysis with an antiserum raised against HMW1 and reactive with both HMW1 and HMW2. All but

two of the 47 strains with an *hmw1A* or *hmw2A* homolog reacted with the antiserum. Of the two nonreactive strains, neither demonstrated appreciable adherence to Chang epithelial cells, thus confirming the absence of significant protein expression. One possibility suggested from a recent study by Barenkamp (1) is that protein expression in these strains was turned off during the course of infection, perhaps related to antibody directed against the HMW proteins, which are known to be immunogenic (2). Whether these strains have intact *hmw* loci is presently under investigation.

To conclude, our results indicate that 95% of nontypeable *H. influenzae* strains contain either *hmw* or *hia* sequences. Furthermore, it appears that *hmw*-deficient strains might have evolved more recently from an encapsulated progenitor. One possible model is that the primordial *H. influenzae* strain was nonencapsulated and spawned two separate but overlapping lineages. The first such lineage acquired *hmw* genes and remained nonencapsulated, while the second acquired *hia* and *cap* genes and became encapsulated. Subsequent mutations within the *cap* locus resulted in evolution of a relatively restricted set of nonencapsulated strains. Studies of additional strain collections may contribute to a better understanding of the prevalence of the *hmw* and *hia* genes and provide further insights into the evolutionary relationships between nontypeable and encapsulated forms.

We thank J. M. Musser for critical reading of the manuscript and for useful discussions.

This work was supported by Public Health Service grant 1R01 DC-02873-01 from the National Institute on Deafness and Other Communication Disorders and by an American Heart Association grant-in-aid to J.W.S.

REFERENCES

- Barenkamp, S. J. 1996. Immunization with high-molecular-weight adhesion proteins of nontypeable *Haemophilus influenzae* modifies experimental otitis media in chinchillas. *Infect. Immun.* **64**:1246-1251.
- Barenkamp, S. J., and E. Leininger. 1992. Cloning, expression, and DNA sequence analysis of genes encoding nontypeable *Haemophilus influenzae* high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of *Bordetella pertussis*. *Infect. Immun.* **60**:1302-1313.
- Barenkamp, S. J., and J. W. St. Geme III. 1996. Identification of a second family of high-molecular-weight adhesion proteins expressed by non-typeable *Haemophilus influenzae*. *Mol. Microbiol.* **19**:1215-1223.
- Klingman, K. L., and T. F. Murphy. 1994. Purification and characterization of a high-molecular-weight outer membrane protein of *Moraxella (Branhamella) catarrhalis*. *Infect. Immun.* **62**:1150-1155.
- Kroll, J. S., B. M. Loynds, and E. R. Moxon. 1991. The *Haemophilus influenzae* capsulation gene cluster: a compound transposon. *Mol. Microbiol.* **5**:1549-1560.
- Kroll, J. S., and E. R. Moxon. 1988. Capsulation and gene copy number at the *cap* locus of *Haemophilus influenzae* type b. *J. Bacteriol.* **170**:859-864.
- Kroll, J. S., S. Zamze, B. Loynds, and E. R. Moxon. 1989. Common organization of chromosomal loci for the production of different capsular polysaccharides in *Haemophilus influenzae*. *J. Bacteriol.* **171**:3343-3347.
- Musser, J. M., S. J. Barenkamp, D. M. Granoff, and R. K. Selander. 1986. Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae*. *Infect. Immun.* **52**:183-191.
- Musser, J. M., J. S. Kroll, D. M. Granoff, E. R. Moxon, B. R. Brodeur, J. Campos, H. Dabernat, W. Frederiksen, J. Hamel, G. Hammond, E. A. Hoiby, K. E. Jonsdottir, M. Kabeer, I. Kallings, W. N. Khan, M. Kilian, K. Knowles, H. J. Koornhof, B. Law, K. I. Li, J. Montgomery, P. E. Pattison, J.-C. Piffaretti, A. K. Takala, M. L. Thong, R. A. Wall, J. I. Ward, and R. K. Selander. 1990. Global genetic structure and molecular epidemiology of encapsulated *Haemophilus influenzae*. *Rev. Infect. Dis.* **12**:75-111.
- Musser, J. M., J. S. Kroll, E. R. Moxon, and R. K. Selander. 1988. Evolutionary genetics of the encapsulated strains of *Haemophilus influenzae*. *Proc. Natl. Acad. Sci. USA* **85**:7758-7762.
- Pittman, M. 1931. Variation and type specificity in the bacterial species *Haemophilus influenzae*. *J. Exp. Med.* **53**:471-493.
- Quentin, R., A. Goudeau, R. J. Wallace, Jr., A. L. Smith, R. K. Selander, and J. M. Musser. 1990. Urogenital, maternal, and neonatal isolates of *Haemophilus influenzae*: identification of unusually virulent serologically nontypeable clone families and evidence for a new *Haemophilus* species. *J. Gen. Microbiol.* **136**:1203-1209.
- St. Geme, J. W., III, and D. Cutter. 1995. Evidence that surface fibrils

- expressed by *Haemophilus influenzae* type b promote attachment to human epithelial cells. *Mol. Microbiol.* **15**:77–85.
14. **St. Geme, J. W., III, and D. Cutter.** Unpublished data.
 15. **St. Geme, J. W., III, and D. Cutter.** 1996. Influence of pili, fibrils, and capsule on *in vitro* adherence by *Haemophilus influenzae* type b. *Mol. Microbiol.* **21**:21–31.
 16. **St. Geme, J. W., III, D. Cutter, and S. J. Barenkamp.** 1996. Characterization of the genetic locus encoding *Haemophilus influenzae* type b surface fibrils. *J. Bacteriol.* **178**:6281–6287.
 17. **St. Geme, J. W., III, and S. Falkow.** 1991. Loss of capsule expression by *Haemophilus influenzae* type b results in enhanced adherence to and invasion of human cells. *Infect. Immun.* **59**:1325–1333.
 18. **St. Geme, J. W., III, S. Falkow, and S. J. Barenkamp.** 1993. High-molecular-weight proteins of nontypeable *Haemophilus influenzae* mediate attachment to human epithelial cells. *Proc. Natl. Acad. Sci. USA* **90**:2875–2879.
 19. **St. Geme, J. W., III, A. Takala, E. Esko, and S. Falkow.** 1994. Evidence for capsule gene sequences among pharyngeal isolates of nontypeable *Haemophilus influenzae*. *J. Infect. Dis.* **169**:337–342.

Editor: J. G. Cannon