INFECTION AND IMMUNITY, May 2004, p. 2457–2461 Vol. 72, No. 5
0019-9567/04/$08.00+0 DOI: 10.1128/IAI.72.5.2457–2461.2004

MINIREVIEW

Haemophilus influenzae: Genetic Variability and Natural Selection To Identify Virulence Factors

Janet R. Gilsdorf,[1,2,3]* Carl F. Marrs,[2,3] and Betsy Foxman[2,3]

Department of Pediatrics and Communicable Diseases, University of Michigan Medical School,[1] and Department of Epidemiology[2] and Center for Molecular and Clinical Epidemiology of Infectious Diseases,[3] University of Michigan School of Public Health, Ann Arbor, Michigan

The evolutionary processes of natural selection govern the nucleotide sequences of bacterial genes such that during replication over many generations, bacterial virulence factor genes change and their products become diverse. This diversity, which both facilitates and reflects an organism's ability to survive in a variety of ecologic niches and under different environmental conditions, occurs through two general strategies—variation in gene expression, which is driven by mechanisms governing gene transcription or translation, and variation in gene content, which is driven by either vertical or horizontal evolution. In this regard, vertical evolution refers to the passage of genetic material from parent to offspring through cell division, with its attendant mutations resulting from mistakes in replication such as point mutations, gene inversions, or spontaneous deletions. Horizontal evolution in bacterial cells occurs by the acquisition of new genetic material from transformation of native DNA, transduction by phages, or conjugation by plasmids; this new genetic material is then passed on to subsequent generations through vertical evolution.

Until recently, bacterial pathogens were characterized solely by their phenotypic characteristics, which describe only gross strain-to-strain differences. While extremely useful, these techniques are limited in their ability to identify unique members of widely variable bacterial populations. Recent advances in bacterial genomics, furthered by the availability of complete genomic sequences from a growing number of organisms, suggest that some "clonal" designations may be misguided, given the high level of genetic diversity of bacterial strains from the same species. For example, sequence comparisons have shown up to 25% differences in gene content among strains of *Neisseria meningitidis*, *Helicobacter pylori*, and *Escherichia coli* (5).

While individual fitness characteristics of bacteria foster the survival of individual organisms, the population dynamics of bacteria encompass fitness characteristics that foster the survival of the group. As in all populations, not every member of a bacterial population needs to succeed in all possible environments; rather, the sum of the specialized fitnesses of individual bacteria ensures the survival of the population in variable environments. Thus, gene products required for bacterial survival in one environmental niche may not be required in another niche. Over time and under the influence of natural selection, the gene contents of organisms from the same species living in different niches will be altered to reflect the necessity for certain genes and the dispensability of others. Bacterial factors that are highly diverse are most susceptible to this process of selection. In this paper we describe how the results of evolutionary processes, as reflected by bacterial population characteristics, may be used to identify potential bacterial virulence factors. We use *Haemophilus influenzae*, whose known virulence factors are highly variable, as an example.

* Corresponding author. Mailing address: L2225 Women's Hospital, Box 0244, University of Michigan, Ann Arbor, MI 48109-0244. Phone: (734) 763-2440. Fax: (734) 936-7635. E-mail: gilsdorf@umich.edu.

H. INFLUENZAE AND DIVERSITY

H. influenzae, a gram negative coccobacillus whose environmental niche is primarily restricted to the human respiratory tract, is classified on the basis of production of a polysaccharide capsule: strain types a through f produce antigenically distinct capsules, and nontypeable strains produce no capsule. In addition to colonizing the nasopharynges of healthy humans, *H. influenzae* causes respiratory infections such as acute otitis media, sinusitis, bronchitis, and pneumonia, as well as invasive blood-borne infections such as meningitis, septic arthritis, and cellulitis. Measurements of the gene contents of two *H. influenzae* strains has revealed that the clinical type b strain Eagan possesses 270 kb of additional genomic material relative to strain Rd, a type d strain that has lost its capsule and is highly laboratory adapted (9). More-detailed analysis has shown that strain Eagan possesses gene regions not found in Rd, including the *cap* region genes, encoding the type b capsular polysaccharide; the *hif* gene cluster, encoding the pilus adhesin; the *hmw1* and *hmw2* gene clusters, encoding the high-molecular-weight adhesions (18); and the tryptophanase gene cluster (25). Although, to date, the entire genome sequence of only one *H. influenzae* strain (Rd) is available (16), genetic sequence analysis of a nontypeable *H. influenzae* (NTHi) strain is in progress (http://www.microbial-pathogenesis.org) and, when complete, will add important information on the extent of *H. influenzae* genomic diversity.

Because phenotypic typing is too limited to optimally describe bacterial diversity, several genotyping systems have been developed for *H. influenzae*, including enterobacterial repetitive intergenic consensus typing (3, 19, 31, 41), pulsed-field gel electrophoresis (19, 31, 38, 43), and ribotyping (31, 46). Recently, Meats et al. (28) used multilocus sequence typing to genotype *H. influenzae*; they identified 19 sequence types among 26 nontypeable strains and only 12 sequence types among 50

type b strains. The average number of alleles for the seven enzymes tested was 27, with a range of 22 to 36. In addition, repetitive-elements PCR-based genotyping systems using the *H. influenzae* uptake signal sequences and intergenic dyad sequences have been developed (7; M. Patel, I. Z. Ecevit, J. R. Gilsdorf, and C. F. Marrs, submitted for publication). These studies all demonstrate significant genetic diversity among strains, with nontypeable strains showing more diversity than typeable strains.

In adapting to its limited ecologic niches in humans, *H. influenzae* most likely relies on environmental sensing strategies, such as the two-component regulatory systems, less commonly than enteric organisms do (47); the genomic sequence of *H. influenza*e strain Rd contains 4 sets of sensor-regulator gene pairs, whereas *E. coli* has 40 (16). Rather, *H. influenzae* relies on fitness selection and clonal expansion of those population members that express the phenotypic characteristics important for survival in a particular human niche.

The sources of selective pressure that drive phenotypic, and thus genotypic, diversity among various populations of *H. influenzae* include host immune responses (8), nutritional resources in the various microenvironments of the human respiratory mucosa (45) or bloodstream (20), and host factors necessary for bacterial adherence (34), cellular penetration (34), and host-to-host transmission (4). The following genetic mechanisms that alter either *H. influenzae* gene expression or gene content generate *H. influenzae* strains with fitness characteristics optimal for survival under various selective pressures, thus allowing the bacteria to adapt to rapidly changing microenvironments.

ALTERED GENE EXPRESSION: PHASE VARIATION

Phase variation is a strategy frequently used by *H. influenzae* to alter the gene expression of virulence factors through slipped-strand mispairing, which is mediated by short DNA repeats (typically 2 to 7 nucleotides each) in either the coding regions or the upstream promoter regions of virulence genes. Spontaneous gain or loss of repeat units in these unstable regions either results in a translational frameshift or alters the distance spanned by the promoter, thus turning gene expression off or on. The *H. influenzae* pilus gene *hifA* (44) and adhesin genes *hmw1A* and *-2A* (13) utilize repeat units of 2 and 7 bp, respectively, in their promoter regions to vary the efficiency of transcription. In addition, five lipooligosaccharide (LOS)-modifying genes (*lic1A*, *lic2A*, *lic3A*, *lgtC*, and *lex2A*), four iron acquisition genes with homology to hemoglobin binding proteins, and the *yopA* homologue demonstrate tetrad repeats of various lengths in their coding regions; the number of repeats correlates with off or on expression of the genes (47). Weiser and Pan (48) have shown that Galα1-4Gal, which is modified at the terminal galactose by a putative galactosyltransferase encoded by *lic2*, may be expressed more commonly by strains infecting patients with pneumonia than by strains from colonized individuals without respiratory tract infections, suggesting that LOS-modifying genes may be host site specific in their expression.

ALTERED GENE EXPRESSION: UNSTABLE MRNA

A recent study of *H. influenzae* biotype IV strains (a subset of NTHi associated with genital and neonatal infections) demonstrates their failure to express pili in spite of the presence of complete *hif* operons and promoter regions (6). The authors conclude that a deletion of 10 of the 44- or 45-nucleotide repetitive palindromic extragenic sequences in the genital strains relative to a piliated *H. influenzae* type b (Hib) strain generated unstable mRNA, resulting in transcription but not translation of *hifB* through *hifE*. Thus, *H. influenzae* appears to have at least two mechanisms by which pilus expression may be modulated, resulting in populations of organisms with variable piliation.

ALTERED GENE CONTENT: POINT MUTATIONS

Sequence analysis of *H. influenzae* genes that encode several surface proteins reveals multiple point mutations, especially throughout their variable regions. These single-nucleotide polymorphisms in genes encoding pathogenic factors reflect spontaneous point mutations that occur because of replication mistakes during cell division (vertical evolution) or because of acquisition by transformation and homologous recombination of gene regions containing very limited gene differences (horizontal evolution). In the highly variable *H. influenzae* proteins P5 (15), P1 (30), and HifA (10), nonsynonymous (amino acid-changing) mutations are seen more commonly than synonymous (silent) mutations, suggesting that mutations resulting in altered protein structure offer a selective advantage to the bacteria.

ALTERED GENE CONTENT: INSERTIONS AND DELETIONS

Comparisons of gene sequences from multiple *H. influenzae* strains reveal many insertions and deletions in genes or gene regions encoding virulence factors (18). The *hif* gene region encoding *H. influenzae* pili, for example, shows remarkable strain-to-strain heterogeneity (17, 29, 36), with insertions, deletions, and duplications of genes, open reading frames, and intergenic regions. Among 20 *H. influenzae* isolates, nine different gene arrangements were found in this region (29).

While gene insertions may result from rearrangements within the bacterial chromosome, in *H. influenzae* most insertions appear to occur through introduction of new DNA by lateral gene transfer through horizontal evolution. As a result, several *H. influenzae* genes encoding virulence factors, such as *hifA* and *hifE* (which encode the major structural subunits and adhesive molecules, respectively, of pilus adhesins) (10, 27), *iga* (which encodes immunoglobulin A1 protease) (33), *hap* (which encodes an *H. influenzae* adhesive/penetration protein) (21), *rfaF* (encoding an LOS heptosyltransferase), and *lic1D* (encoding LOS phosphocholine transferase) (11) exhibit intragenic mosaicism, an apparent mixing and matching of gene regions acquired from other *H. influenzae* strains that differ in the genetic configurations of these regions. Meats et al. (28) observed that allelic differences among seven multilocus sequence typing loci occurred more commonly from recombina-

TABLE 1. Prevalences of virulence genes in *H. influenzae* isolates from the middle-ear cavities of children with acute otitis media and from the throats or nasopharynges of healthy children[a]

Gene	Gene product	No. of isolates with gene/total isolates (%) from:		*P*	Source or reference
		Middle ear	Throat or nasopharynx		
lic2A	LOS glycosyltransferase	47/48 (98)	69/90 (77)	0.001	35
lic2B	LOS glycosyltransferase	25/48 (52)	13/90 (14)	0.000026	35
hif cluster	Adhesin	10/50 (20)	23/62 (37)	0.38	I. Z. Ecevit, unpublished data
hmw	Adhesin	5/9 (55)	7/30 (23)	0.03	I. Z. Ecevit, unpublished data

[a] The *H. influenzae* strains were collected over many years from widely diverse geographic locations. The *hif* cluster and *hmw* genes were detected by dot blot DNA hybridization using DNA probes designed to hybridize with *hifB* and *hifC*, *hmw1A*, *hmw2A*, and *hmwC*.

tion than from point mutation and that mosaic gene structure was found in four of the seven loci.

H. influenzae, like other human mucosal pathogens such as *Neisseria gonorrhoeae*, *N. meningitidis*, *Streptococcus pneumoniae*, *Moraxella catarrhalis*, and *H. pylori*, are naturally competent and possess specific genetic mechanisms for taking up nonepisomal DNA from their environment (14). In *H. influenzae*, horizontal gene transfer through transformation and homologous recombination is facilitated by the presence of the so-called uptake sequences, 5′-AAGTGCGGT-3′, of which 1,465 copies are randomly distributed throughout the *H. influenzae* strain Rd genome; if extended uptake sequences and singly mutated sites are considered, 2,229 uptake sequence sites occupying approximately 3.5% of the genome are seen (40). As evidence of their potential role in facilitating the horizontal transfer of *H. influenzae* virulence factors, direct repeats of uptake sequences flank the *H. influenzae* tryptophanase gene cluster (25).

The role of insertion elements (IS) and phage elements in the horizontal transfer of virulence genes is shown by the 711-bp IS*1016* direct repeats that flank the *cap* (polysaccharide) region (22), the 32-bp direct repeats flanking the *hif-1* gene cluster present in certain *H. influenzae* biogroup aegyptius strains (36), the 59-bp direct repeats flanking the *hif* region (29), and the duplicated inverted repeats flanking the *hif* region (29, 36). Furthermore, as described for *Staphylococcus aureus* (39), phage- and transposon-like elements may catalyze deletional events in *H. influenzae*. Comparison of *H. influenzae* strains by subtractive hybridization has revealed a relatively high number of phage-like genes that are present in one strain and not in the subtracted strain (32, 42), supporting the notion that many differences in gene content between *H. influenzae* strains may be associated with phage-related horizontal transfer into, or deletion out of, a strain.

ALTERED GENE CONTENT: DUPLICATIONS

Gene duplications have been shown to result in alterations in *H. influenzae* pathogenicity. The *cap* locus, which encodes *H. influenzae* capsule production by strain types a through f, is present as a tandem duplication. A deletion of the carbohydrate transport gene *bexA* in one copy of the *cap* locus in type b strains appears to stabilize the duplication (12, 24). The resulting increased production of the type b capsule contributes to the increased virulence of type b strains. Recombination events resulting in additional copies of the *cap* locus generate Hib variants with increased pathogenicity proportional to the number of copies present and the amount of capsule produced (12, 23). Because of the presence of a second copy of the *hif* locus, some *H. influenzae* biogroup aegyptius strains (including strains isolated from patients with Brazilian purpuric fever) display abundant piliation, reflecting the decreased probability that, at any point in time, both copies of the phase-variable pilus promoters are present in the off orientation (35, 37).

ALTERED GENE CONTENT: EVIDENCE FOR EVOLUTION-DRIVEN DIFFERENCES IN VIRULENCE FACTORS

Through the mechanisms described above, diverse populations of *H. influenzae* isolated from different human body sites show variable expression of virulence factors due to altered gene content or gene expression. For example, although pathogenic *H. influenzae* strains colonize the nasopharynx prior to establishing infection at another body site, the population of *H. influenzae* causing acute otitis media (derived from a subset of the throat population) may exhibit characteristics not required for survival in the throat. Examination of *H. influenzae* isolates from the middle-ear cavities of children with acute otitis media and from the throats of healthy children demonstrates variable prevalences of genes encoding several virulence factors (15a, 32), suggesting that certain factors may improve *H. influenzae* fitness in specific ecologic niches, whereas in other niches these factors may be lost through evolution with no deleterious effect on the ability of *H. influenzae* to survive (Table 1). These results suggest that *H. influenzae* strains possessing the *lic2A*, *lic2B*, and *hmw* genes have a fitness advantage in the middle ear, whereas *H. influenzae* strains possessing the *hif* cluster have a fitness advantage in the throat or nasopharynx. These observations suggest that the *lic2A*, *lic2B*, and *hmw* gene products are virulence factors for *H. influenzae* acute otitis media and that the *hif* gene products are virulence factors for respiratory tract colonization. The precise role these genes play in infection or colonization remains unclear.

USE OF NATURAL SELECTION TO IDENTIFY *H. INFLUENZAE* VIRULENCE FACTORS

In contrast to genetic diversity resulting from variations in gene expression (in which genes are present but not transcribed or translated under certain environmental conditions), diversity resulting from variations in gene content occurs either because genetic material no longer essential to the survival or

pathogenicity of an organism in a specific environment is lost (with no impact on survival) or because genetic material that offers a fitness advantage in a specific environment has been acquired. Various comparative genomics techniques have been used to identify genomic differences among bacteria occupying unique environmental niches. In the case of *H. influenzae*, several strategies based on gene content, including genomic analysis and mapping by in vitro mutagenesis (GAMBIT) (1, 2), have identified genes essential for survival under various environmental conditions. In addition, differential fluorescence induction has been used to identify *H. influenzae* genes expressed in vivo in a chinchilla model of acute otitis media (26). Signature-tagged mutagenesis has also been used to identify genes required for *H. influenzae* survival in an infant rat model of *H. influenzae* bacteremia (20). Among the 25 *H. influenzae* genes essential for invasive disease in the rat were 4 previously described *H. influenzae* virulence factors: *hia* (adhesin), *licA* (biosynthesis of LOS), IS*1016* flanking the capsule locus, and a gene homologous to PBP-7. The genes identified by these techniques require further study of their specific roles in the pathogenesis of *H. influenzae* infection and disease.

Two studies have used subtractive hybridization to detect potential virulence factors. This technique identifies genes of one *H. influenzae* strain (the tester strain, representing a population of bacteria associated with illness or a specific epidemiologic niche) that are not present in another strain (the driver strain, which is not associated with that illness or epidemiologic niche). Smoot et al. (42) subtracted the genome of a noninvasive *H. influenzae* biogroup aegyptius isolate from that of Brazilian purpuric fever strain F3031. Twenty-one tester-specific clones (present in the BPF strain and absent in the noninvasive *H. influenzae* strain) were identified, including those possessing *iga* (encoding immunoglobulin protease), *hmcD* (encoding hemocin), various phage-related genes, and hypothetical genes of unknown function. From this study no conclusions could be drawn about the relative prevalence of these genes in the two bacterial populations represented by the tester and driver strains, thus limiting the ability to understand their real roles in *H. influenzae* virulence.

Pettigrew et al. (32) subtracted the genome of the fully sequenced *H. influenzae* laboratory strain Rd (driver strain) from that of an NTHi strain isolated from a child with acute otitis media (tester strain). Of the 36 tester (otitis media strain)-specific DNA fragments, 17 were phage-related proteins, underscoring the importance of horizontal transfer of genetic material in genetic variability between strains. In addition, clones containing 5 known *H. influenzae* genes were identified, including a methyltransferase-encoding gene of *H. influenzae* strain Rd, a gene encoding a LOS biosynthesis enzyme, the LOS biosynthesis gene *lic2B*, and the *H. influenzae* adhesin genes *hmw1* and *hmw2*, as well as 14 clones containing genetic material with homology to 13 genes of other microbes. To assess the potential relevance of these genes to the pathogenesis of acute otitis media, the prevalence of these tester-specific genetic regions among *H. influenzae* otitis media isolates was compared to their prevalence in *H. influenzae* isolates from throat specimens. *lic2B* was found in 14% of NTHi throat isolates from healthy children, 39% of NTHi throat isolates from ill children with respiratory tract infections, 52% of NTHi middle-ear isolates from children, and 88% of invasive Hib isolates. Additional studies utilizing traditional methods to define virulence factors will be required in order to understand the precise role of these factors in the pathogenesis of *H. influenzae* infection and disease.

In summary, the convergence of the investigative techniques used in epidemiological and microbiological studies offers a largely untapped, fruitful field for the identification of bacterial virulence factors by utilizing the power of natural selection that has resulted in diverse bacterial populations.

ACKNOWLEDGMENTS

This paper was supported by National Institutes of Health grants DC 005840 and AI 25630 and by a grant from the Deafness Research Foundation (all to J.R.G.).

REFERENCES

1. **Akerley, B. J., E. J. Rubin, A. Camilli, D. J. Lampe, H. M. Robertson, and J. J. Mekalanos.** 1998. Systematic identification of essential genes by in vitro mariner mutagenesis. Proc. Natl. Acad. Sci. USA **95:**8927–8932.
2. **Akerley, B. J., E. J. Rubin, V. Novick, K. Amaya, N. Judson, and J. J. Mekalanos.** 2002. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. Proc. Natl. Acad. Sci. USA **99:**966–971.
3. **Belkum, A., B. Duim, A. Regelink, L. Moller, W. Quint, and L. VanAlphen.** 1994. Genomic DNA fingerprinting of clinical *Haemophilus influenzae* isolates by polymerase chain reaction amplification: comparison with major outer-membrane protein and restriction fragment length polymorphism analysis. J. Med. Microbiol. **41:**63–68.
4. **Bergstrom, C. T., P. McElhany, and L. A. Real.** 1999. Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc. Natl. Acad. Sci. USA **96:**5095–5100.
5. **Boucher, Y., C. L. Nesbo, and W. F. Doolittle.** 2001. Microbial genomes: dealing with diversity. Curr. Opin. Microbiol. **4:**285–289.
6. **Bruant, G., N. Gousset, R. Quentin, and A. Rosenau.** 2002. Fimbrial *ghf* gene cluster of genital strains of *Haemophilus* spp. Infect. Immun. **70:**5438–5445.
7. **Bruant, G., S. Watt, R. Quentin, and A. Rosenau.** 2003. Typing of nonencapsulated *Haemophilus* strains by repetitive-element sequence-based PCR using intergenic dyad sequences. J. Clin. Microbiol. **41:**3473–3480.
8. **Brunham, R. C., F. A. Plummer, and R. S. Stephens.** 1993. Bacterial antigenic variation, host immune response, and pathogen-host coevolution. Infect. Immun. **61:**2273–2276.
9. **Butler, P. D., and E. R. Moxon.** 1990. A physical map of the genome of *Haemophilus influenzae* type b. J. Gen. Microbiol. **136:**2333–2342.
10. **Clemans, D. L., C. F. Marrs, M. Patel, M. Duncan, and J. R. Gilsdorf.** 1998. Comparative analysis of *Haemophilus influenzae hifA* (pilin) genes. Infect. Immun. **66:**656–663.
11. **Cody, A. J., D. Field, E. J. Feil, S. Stringer, M. E. Deadman, A. G. Tsolaki, B. Gratz, V. Bouchet, R. Goldstein, D. W. Hood, and E. R. Moxon.** 2003. High rates of recombination on otitis media isolates of non-typeable *Haemophilus influenzae.* Inf. Gen. Evol. **3:**57–66.
12. **Corn, P. G., J. Anders, A. K. Takala, H. Kayhty, and S. K. Hoiseth.** 1993. Genes involved in *Haemophilus influenzae* type b capsule expression are frequently amplified. J. Infect. Dis. **167:**356–364.
13. **Dawid, S., S. J. Barenkamp, and J. W. St Geme III.** 1999. Variation in expression of the *Haemophilus influenzae* HMW adhesins: a prokaryotic system reminiscent of eukaryotes. Proc. Natl. Acad. Sci. USA **96:**1077–1082.
14. **Deitsch, K. W., E. R. Moxon, and T. E. Wellems.** 1997. Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. Microbiol. Mol. Biol. Rev. **61:**281–293.
15. **Duim, B., L. Vogel, W. Puijk, H. M. Jansen, R. H. Meloen, J. Dankert, and L. van Alphen.** 1996. Fine mapping of outer membrane protein P2 antigenic sites which vary during persistent infection by *Haemophilus influenzae.* Infect. Immun. **64:**4673–4679.

15a. **Ecevit, I. Z., K. W. McCrea, M. M. Pettigrew, A. Sen, C. F. Marrs, and J. R. Gilsdorf.** Prevalence of *hifB*, *hifC*, *hmw1A*, hmw2A, *hmwC*, and *hia* genes in *Haemophilus influenzae* isolates. J. Clin. Microbiol., in press.

16. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenny, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Philips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae.* Science **269:**496–512.
17. **Geluk, F., P. P. Eijk, S. M. van Ham, H. M. Jansen, and L. van Alphen.** 1998.

The fimbria gene cluster of nonencapsulated *Haemophilus influenzae.* Infect. Immun. **66:**406–417.
18. **Gilsdorf, J. R.** 1998. Antigenic diversity and gene polymorphisms in *Haemophilus influenzae.* Infect. Immun. **66:**5053–5059.
19. **Gomez-De-Leon, P., J. I. Santos, J. Caballero, D. Gomez, L. E. Espinosa, I. Moreno, D. Pinero, and A. Cravioto.** 2000. Genomic variability of *Haemophilus influenzae* isolated from Mexican children determined by using enterobacterial repetitive intergenic consensus sequences and PCR. J. Clin. Microbiol. **38:**2504–2511.
20. **Herbert, M. A., S. Hayes, M. E. Deadman, C. M. Tang, D. W. Hood, and E. R. Moxon.** 2002. Signature tagged mutagenesis of *Haemophilus influenzae* identifies genes required for *in vivo* survival. Microb. Pathog. **33:**211–223.
21. **Kilian, M., K. Poulsen, and H. Lomholt.** 2002. Evolution of the paralogous *hap* and *iga* genes in *Haemophilus influenzae*: evidence for a conserved *hap* pseudogene associated with microcolony formation in the recently diverged *Haemophilus aegyptius* and *H. influenzae* biogroup aegyptius. Mol. Microbiol. **46:**1367–1380.
22. **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.
23. **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.
24. **Kroll, J. S., E. R. Moxon, and B. M. Loynds.** 1993. An ancestral mutation enhancing the fitness and increasing the virulence of *Haemophilus influenzae* type b. J. Infect. Dis. **168:**172–176.
25. **Martin, K., G. Morlin, A. Smith, A. Nordyke, A. Eisenstark, and M. Golomb.** 1998. The tryptophanase gene cluster of *Haemophilus influenzae* type b: evidence for horizontal gene transfer. J. Bacteriol. **180:**107–118.
26. **Mason, K. M., R. S. Munson, Jr., and L. O. Bakaletz.** 2003. Nontypeable *Haemophilus influenza*e gene expression induced in vivo in a chinchilla model of otitis media. Infect. Immun. **71:**3454–3462.
27. **McCrea, K. W., J. M. St. Sauver, C. F. Marrs, D. Clemans, and J. R. Gilsdorf.** 1998. Immunologic and structural relationships of the minor pilus subunits among *Haemophilus influenzae* isolates. Infect. Immun. **66:**4788–4796.
28. **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 noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J. Clin. Microbiol. **41:**1623–1636.
29. **Mhlanga-Mutangadura, T., G. Morlin, A. L. Smith, A. Eisenstark, and M. Golomb.** 1998. Evolution of the major pilus gene cluster of *Haemophilus influenzae.* J. Bacteriol. **180:**4693–4703.
30. **Munson, R., Jr., S. Grass, M. Einhorn, C. Bailey, and C. Newell.** 1989. Comparative analysis of the structures of the outer membrane protein P1 genes from major clones of *Haemophilus influenzae* type b. Infect. Immun. **57:**3300–3305.
31. **Pettigrew, M. M., B. Foxman, Z. Ecevit, C. F. Marrs, and J. Gilsdorf.** 2002. Use of pulsed-field gel electrophoresis, enterobacterial repetitive intergenic consensus typing, and automated ribotyping to assess genomic variability among strains of nontypeable *Haemophilus influenzae.* J. Clin. Microbiol. **40:**660–662.
32. **Pettigrew, M. M., B. Foxman, C. F. Marrs, and J. R. Gilsdorf.** 2002. Identification of the lipooligosaccharide biosynthesis gene *lic2B* as a putative virulence factor in strains of nontypeable *Haemophilus influenzae* that cause otitis media. Infect. Immun. **70:**3551–3556.
33. **Poulsen, K., J. Reinholdt, and M. Kilian.** 1992. A comparative genetic study of serologically distinct *Haemophilus influenzae* type 1 immunoglobulin A1 proteases. J. Bacteriol. **174:**2913–2921.
34. **Rao, V. K., G. P. Krasan, D. R. Hendrixson, S. Dawid, and J. W. St. Geme III.** 1999. Molecular determinants of the pathogenesis of disease due to non-typeable *Haemophilus influenzae.* FEMS Microbiol. Rev. **23:**99–129.
35. **Read, T. D., M. Dowdell, S. W. Satola, and M. M. Farley.** 1996. Duplication of pilus gene complexes of *Haemophilus influenzae* biogroup aegyptius. J. Bacteriol. **178:**6564–6570.
36. **Read, T. D., S. W. Satola, and M. M. Farley.** 2000. Nucleotide sequence analysis of hypervariable junctions of *Haemophilus influenzae* pilus gene clusters. Infect. Immun. **68:**6896–6902.
37. **Read, T. D., S. W. Satola, J. A. Opdyke, and M. M. Farley.** 1998. Copy number of pilus gene clusters in *Haemophilus influenzae* and variation of the *hifE* pilin gene. Infect. Immun. **66:**1622–1631.
38. **Saito, M., A. Umeda, and S. Yoshida.** 1999. Subtyping of *Haemophilus influenzae* strains by pulsed-field gel electrophoresis. J. Clin. Microbiol. **37:**2142–2147.
39. **Sassetti, C., and E. J. Rubin.** 2002. Genomic analyses of microbial virulence. Curr. Opin. Microbiol. **5:**27–32.
40. **Smith, H. O., J. F. Tomb, B. A. Dougherty, R. D. Fleischmann, and J. C. Venter.** 1995. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. Science **269:**538–540.
41. **Smith-Vaughan, H. C., K. S. Sriprakash, A. J. Leach, J. D. Mathews, and D. J. Kemp.** 1998. Low genetic diversity of *Haemophilus influenzae* type b compared to nonencapsulated *H. influenzae* in a population in which *H. influenzae* is highly endemic. Infect. Immun. **66:**3403–3409.
42. **Smoot, L. M., D. D. Franke, G. McGillivary, and L. A. Actis.** 2002. Genomic analysis of the F3031 Brazilian purpuric fever clone of *Haemophilus influenzae* biogroup aegyptius by PCR-based subtractive hybridization. Infect. Immun. **70:**2694–2699.
43. **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.
44. **van Ham, S. M., L. van Alphen, F. R. Mooi, and J. P. van Putten.** 1993. Phase variation of *H. influenzae* fimbriae: transcriptional control of two divergent genes through a variable combined promoter region. Cell **73:**1187–1196.
45. **van Ulsen, P., M. van Schilfgaarde, J. Dankert, H. Jansen, and L. van Alphen.** 2002. Genes of non-typeable *Haemophilus influenza*e expressed during interaction with human epithelial cell lines. Mol. Microbiol. **45:**485–500.
46. **Wang, C. C., L. K. Siu, M. K. Chen, Y. L. Yu, F. M. Lin, M. Ho, and M. L. Chu.** 2001. Use of automated riboprinter and pulsed-field gel electrophoresis for epidemiological studies of invasive *Haemophilus influenzae* in Taiwan. J. Med. Microbiol. **50:**277–283.
47. **Weiser, J.** 2000. The generation of diversity by *Haemophilus influenzae.* Trends Microbiol. **8:**433–435.
48. **Weiser, J. N., and N. Pan.** 1998. Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide. Mol. Microbiol. **30:**767–775.

Editor: J. B. Kaper