Variation in the Presence of Neuraminidase Genes among *Streptococcus pneumoniae* Isolates with Identical Sequence Types

Melinda M. Pettigrew,[1]* Kristopher P. Fennie,[2] Matthew P. York,[1]
Janeen Daniels,[1]† and Faryal Ghaffar[3]

*Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven,
Connecticut[1]; Yale University School of Nursing, New Haven, Connecticut[2]; and
University of Texas Southwestern Medical Center, Dallas, Texas[3]*

***Streptococcus pneumoniae* frequently colonizes the upper respiratory tract of young children and is an important cause of otitis media and invasive disease. Carriage is more common than disease, yet the genetic factors that predispose a given clone for disease are not known. The relationship between capsule type, genetic background, and virulence is complex, and important questions remain regarding how pneumococcal clones differ in their ability to cause disease. Pneumococcal neuraminidase cleaves sialic acid-containing substrates and is thought to be important for pneumococcal virulence. We describe the distribution of multilocus sequence types (ST), capsule type, and neuraminidase genes among 342 carriage, middle ear, blood, and cerebrospinal fluid (CSF) pneumococcal strains from young children. We found 149 STs among our *S. pneumoniae* isolates. *nanA* was present in all strains, while *nanB* and *nanC* were present in 96% and 51% of isolates, respectively. The distribution of *nanC* varied among the strain collections from different tissue sources ($P = 0.03$). The prevalence of *nanC* was 1.41 (95% confidence interval, 1.11, 1.79) times higher among CSF isolates than among carriage isolates. We identified isolates of the same ST that differed in the presence of *nanB* and *nanC*. These studies demonstrate that virulence determinants, other than capsule loci, vary among strains of identical ST. Our studies suggest that the presence of *nanC* may be important for tissue-specific virulence. Studies that both incorporate MLST and take into account additional virulence determinants will provide a greater understanding of the pneumococcal virulence potential.**

Streptococcus pneumoniae colonizes the nasopharynx of up to 55% of healthy young children (3). Asymptomatic carriage is far more common than disease, yet these bacteria are important causes of otitis media, pneumonia, bacteremia, and meningitis in young children. Despite advances in our understanding of the epidemiology and pathogenesis of pneumococcal infections, the precise genetic factors that predispose a given pneumococcal clone for disease versus carriage remain unknown.

Pneumococci individually express one of approximately 90 extracellular and structurally distinct capsular polysaccharides. The chemical composition of the *S. pneumoniae* capsule is generally considered the most important virulence factor (24, 26). *S. pneumoniae* strains differ in their abilities to cause disease; the vast majority of infections are due to 20 of the 90 serotypes (36). Animal studies indicate that the capsule type must be considered in light of the genetic background of the strain (8, 21). The influence of the combination of capsule and genetic background in pathogenesis differs depending on the site of infection (20). The balance of the relationship between capsule type and serotype in determining virulence potential

has important implications for the success of vaccination strategies (25, 31).

Several population-based studies have used multilocus sequence typing (MLST) of *S. pneumoniae* strains to explore the relationship between capsule type and genetic background in determining virulence potential. MLST analysis of invasive and carriage isolates showed similarities in the invasiveness of different genetic clones of the same serotype, suggesting that capsular type is more important that genotype in defining the virulence of a clone (4). Another study concluded that clonal properties in addition to capsule were important in determining invasive disease potential (37).

While MLST provides a powerful tool to define clonal groupings (9), it evaluates differences based on the sequence of relatively conserved housekeeping genes. MLST does not capture the total genomic variation within individual pneumococcal isolates. Comparison of the R6, TIGR4, and G54 genomes indicated that 86, 177, and 178 genes, respectively, were unique compared to the other two strains (43). The *S. pneumoniae* genome may contain a core complement of conserved genes with additional regions that exhibit high diversity (15). MLST studies have described individual sequence types (STs) that express different capsule types (19, 35). Given the high level of diversity among pneumococcal strains, which is believed to arise through horizontal transfer (7, 12), we hypothesized that pneumococcal isolates could appear identical by MLST type but differ in the content of additional virulence-associated genes.

---

* Corresponding author. Mailing address: Department of Epidemiology and Public Health, Yale University School of Medicine, 60 College Street, P.O. Box 208034, New Haven, CT 06520-8034. Phone: (203) 785-5220. Fax: (203) 785-6130. E-mail: melinda.pettigrew@yale.edu.
† Present address: Spelman College, Atlanta, Ga.

Two neuraminidases, encoded by *nanA* and *nanB*, have been described for *S. pneumoniae* (2, 6, 39). A *nanB* homolog, *nanC* (SP1326) (43), has also been identified, but its expression and activity have not been described. Pneumococcal neuraminidase cleaves sialic acid-containing substrates and is thought to promote pneumococcal colonization by exposing host cell receptors (44). NanA may help promote colonization through desialylation of host proteins that mediate bacterial clearance, such as lactoferrin or immunoglobulin A2 (23). NanA also has been shown to desialylate lipopolysaccharides of *Neisseria meningitidis* and *Haemophilus influenzae* strains (39). The desialylation of lipopolysaccharide may give pneumococci a competitive advantage over *N. meningitidis* and *H. influenzae*, which reside in the same host niche, by making them more susceptible to complement-mediated clearance.

We describe the distribution of ST, capsule type, and neuraminidase genes among a large collection of carriage, middle ear, blood, and cerebrospinal fluid *S. pneumoniae* strains from young children in the United States. Furthermore, we compare the distribution of *nanB* and *nanC* among pneumococcal isolates from different tissue sources.

MATERIALS AND METHODS

Bacterial strains. Blood and cerebrospinal fluid *S. pneumoniae* strains from children under 5 were obtained from the Active Bacterial Core Surveillance Emerging Infections Program network at the Centers for Disease Control and Prevention. This program conducts surveillance to determine the incidence and epidemiologic characteristics of invasive disease in California, Colorado, Connecticut, Georgia, Maryland, Minnesota, New York, Oregon, and Tennessee. The 100 pneumococcal otitis media isolates used in this study were obtained from Edward O. Mason at Baylor College of Medicine in Houston, Texas (45). These strains were collected from children attending eight U.S. hospitals between 1995 and 2001. *S. pneumoniae* carriage strains were obtained from two separate sources. Nasopharyngeal isolates were obtained from healthy infants in a prospective study of pneumococcal carriage among conjugate vaccine recipients in Texas between 2000 and 2001 (13). *S. pneumoniae* throat isolates from healthy children were collected in 2001 during a study of the prevalence of *S. pneumoniae* and *H. influenzae* colonization and risk factors for carriage among children less than 3 years of age attending 16 licensed day care centers in Washtenaw County, MI (10).

Individual groups, from whom these strains were obtained, serotyped the isolates by using the Neufeld-Quelling reaction (1). Serotypes of pneumococcal strains implicated in capsule switches were retested using a sequential multiplex PCR assay (30). Factor-specific data were not available for all of the isolates. In these instances, the isolates were classified by their serogroup. Isolates identified as 4, 9V, 6B, 14, 19F, 18C, and 23F were classified as vaccine serotypes. Serotypes belonging to the same serogroup as vaccine serotypes were considered vaccine related (6A, 9A, 9L, 9N, 18B, 19A, 19B, 19C, 23A, and 23B). Serotypes not classified as vaccine or vaccine-related serotypes were considered nonvaccine serotypes. Pneumococcal blood and cerebrospinal fluid isolates were grouped as invasive strains. While carriage strains from healthy children and middle ear isolates may have the potential to cause invasive disease, we considered these isolates "noninvasive" for the purposes of our analyses.

Multilocus sequence typing. The profiles of the pneumococcal strain were obtained by PCR amplification of internal fragments of seven housekeeping genes using previously published methods (9). Individual *S. pneumoniae* strains were grown overnight on trypticase soy agar plates with 5% sheep blood and incubated at 37°C with 5% CO_2. A colony from each plate was used to inoculate a 96-well plate containing 100 μl of Tris-EDTA buffer and boiled for 10 min. One microliter of the crude boil preparation lysate was used for PCR using high-fidelity Advantage2 *Taq* polymerase (Clontech, Palo Alto, CA). PCR cycling conditions were as follows: 35 cycles of 94°C for 30 s, 52°C for 30 s, and 72°C for 45 s. The primers used for MLST analysis are the same as those used for other MLST typing studies, with the exception that M13F (5′TGTAAAACGACGGC CAGT 3′) and M13R (5′-AGGAAACAGCTATGACCAT-3′) primer tails were added to each of the gene-specific primers. The addition of primer tails increases the ease of sequencing in a 96-well format because each of the alleles can be

sequenced with the same forward and reverse primer. Importantly, the primer tails do not compromise the accuracy of the sequence data.

Ninety-six-well plates of PCR products were sequenced by Genaissance Pharmaceuticals in New Haven, CT. DNA sequences from the seven housekeeping loci were trimmed to the precise region using EditSeq software (Lasergene Navigator). Assignment of STs and clonal complexes (CC) was performed using the eBURST program (http://spneumoniae.mlst.net/) (11). The founding genotype for each CC was identified as the genotype that differs from the highest number of other genotypes in the complex at only one locus out of seven.

Detection of neuraminidase genes. Five-hundred-base-pair portions of *nanA*, *nanB*, and *nanC* genes were PCR amplified using the *nanA* forward primer (5′-ATAGACGTGCGCAAAATACAGAATCA-3′) and reverse primer (5′GT CGAACTCCAAGCCAATAACTCCT-3′), *nanB* forward primer (5′-ACTACG AGGTGTTAATCGTGAAGG-3′) and reverse primer (5′-CCAATACCCGCA GGCATAACATC-3′), and the *nanC* forward primer (5′-TGGGGTAAGTAC AAACAAGAGG-3′) and reverse primer (5′-CTAATGGTACTGGCGAAAAT CA-3′). PCR conditions were 35 cycles of denaturation at 98°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 1 min when screening for *nanA* and *nanB*. Thirty-five cycles of denaturation at 98°C for 30 s, annealing at 51°C for 30 s, and extension at 72°C for 1 min were used to amplify *nanC*. A subset of 20 *nanA*, *nanB*, or *nanC* product was sequenced to verify the accuracy of PCR amplification.

One isolate was negative for *nanA*, 52 isolates were negative for *nanB*, and 184 were negative for *nanC* by PCR. Each of the PCR-negative isolates was further examined by dot blot or Southern hybridization using previously described methods (32). Briefly, a *nanA*-, *nanB*-, or *nanC*-specific probe was generated by PCR amplification of TIGR4 genomic DNA using the gene-specific primer pairs listed above. Each probe was labeled with alkaline phosphatase, and blots were hybridized at 60°C overnight. The Gene Images AlkPhos Direct Labeling and ECF chemifluorescence detection system (Amersham Biosciences, Piscataway, N.J.) was used for labeling, hybridization, washes, and signal detection according to the manufacturer's instructions. Blots were exposed to Hyperfilm ECL film (Amersham Biosciences, Piscataway, N.J.).

Characterization of atypical *nanB* PCR products. DNA sequencing was conducted for unusually large *nanB* fragments of about 1,000 bp. Large *nanB* fragments were cloned into the plasmid TOPO vector pCR2.1 (Invitrogen), and the plasmid was used to transform One Shot *Escherichia coli* cells. Plasmids were purified using the QIAprep Miniprep kit from QIAGEN. Plasmid products were sequenced at Genaissance, and the resulting sequence was used to search the NCBI GenBank database.

Data analysis. Statistical analyses were done using SAS version 9.1 (SAS Institute, Cary, NC). We described the distribution of *S. pneumoniae* STs, clonal groups, and neuraminidase genes in our collection using simple descriptive statistics. Nine pneumococcal strains contained large *nanB* fragments with an insertion. The insertion is within the coding region of the *nanB* gene; therefore, we did not count these strains in our prevalence estimates of *nanB*. We conducted log linear modeling using binomial maximum likelihood to determine predictors of invasiveness and to compare the prevalence of *nanC* in middle ear, blood, or cerebrospinal fluid (CSF) isolates with the prevalence among carriage strains. This method was chosen over logistic regression because it allows direct estimation of prevalence ratios (14, 41). Contingency table analysis was carried out to examine the effects of *nanC* on invasiveness and tissue site while controlling for ST and vaccine serotypes. An extended Cochran-Mantel Haenszel test was used to determine significance.

RESULTS

Multilocus sequence typing. The 342 pediatric clinical and carriage *S. pneumoniae* strains in our collection were of 149 different STs (Table 1). Sixty-seven new STs were identified in these studies and were deposited in the MLST database. The newly identified STs were isolated from carriage (20 strains), middle ear (22 strains), blood (9 strains), and cerebrospinal fluid (16 strains). eBURST analysis identified 14 CCs with 3 or more STs, 28 different CCs with 2 or more STs, and 54 singletons (data not shown). The largest clonal complex contained 26 isolates and 13 STs with ST 439 as the predicted founder.

Of the 50 STs containing more than one isolate, 11 contained isolates that expressed different capsule types. When comparing our strains with those entered in the international

TABLE 1. Heterogeneity within serotypes of 342 pneumococcal isolates from young children

Serogroup	Total no. of isolates	No. of STs	ST(s) (no. of isolates)
1	3	2	227 (2), 305 (1)
2	1	1	1519[a] (1)
3	13	3	180 (11), 1119 (1), 1511[a] (1)
4	3	3	244 (1), 867 (1), 899 (1)
5	1	1	1522[a] (1)
6[b]	13	11	376 (2), 90 (2), 138 (1), 460 (1), 639 (1), 1060 (1), 1165 (1), 1379 (1), 1398[a] (1), 1517[a] (1), 1518[a] (1)
6A	25	8	460 (8), 376 (7), 473[c] (5), 62[c] (1), 90 (1), 690 (1), 1190 (1), 1402[a] (1)
6B	38	24	1092 (6), 315 (3), 385 (3), 138 (2), 146 (2), 384 (2), 1212 (2), 1391 (2), 90 (1), 147 (1), 176 (1), 439[c] (1), 473[c] (1), 641 (1), 678[d] (1), 1060 (1), 1177 (1), 1394[a] (1), 1509[a] (1), 1536[a] (1), 1538[a] (1), 1539[a] (1), 1845[a] (1), 1891[a] (1)
7	7	3	191 (5), 1395[a] (1), 1535[a] (1)
9A	1	1	643 (1)
9L	1	1	66 (1)
9N	3	2	66 (1), 632 (2)
9V	12	6	156[c] (6), 643 (2), 66 (1), 166 (1), 1269 (1), 1520[a] (1)
10	2	2	816 (1), 1513[a] (1)
11A	10	5	62[c] (6), 1510[a] (1), 1530[a] (1), 1841[a] (1), 1889[a] (1)
12F	7	3	220 (5), 218 (1), 246 (1)
14	35	15	13 (17), 124 (4), 156[c] (2), 15 (1), 67[c] (1), 125 (1), 132 (1), 236[c] (1), 656 (1), 671 (1), 689 (1), 1508[a] (1), 1528[a] (1), 1529[a] (1), 1537[a] (1)
15	7	2	199[c] (6), 1757[a] (1)
16	5	2	1390[a] (3), 659 (2)
18B	3	1	113 (4)
18C	16	7	113 (5), 1073 (5), 1524[a] (2), 496 (1), 638 (1), 1515[a] (1), 1516[a] (1)
19A	13	5	199[c] (9), 37 (1), 153[d] (1), 193 (1), 274 (1)
19B	1	1	1843[a] (1)
19F	55	31	236[c] (10), 81[c] (6), 271 (4), 251 (2), 257 (2), 425 (2), 473 (2), 1512[a] (2), 1521[a] (2), 1843[a] (2), 43 (1), 67[c] (1), 83[c] (1), 177 (1), 193 (1), 199[c] (1), 320 (1), 651 (1), 1258 (1), 1388[a] (1), 1392[a] (1), 1396 (1), 1459 (1), 1523 (1), 1531[a] (1), 1533[a] (1), 1534[a] (1), 1842[a] (1), 1846[a] (1), 1890[a] (1), 1891[a] (1)
22F	4	1	433 (4)
23A	2	1	1839[a] (2)
23B	2	2	1448 (1), 1847[a] (1)
23F	34	17	37[c] (10), 81[c] (8), 1389[a] (2), 33 (1), 36 (1), 83[c] (1), 355 (1), 439[c] (1), 626 (1), 629 (1), 802 (1), 1092 (1), 1393[a] (1), 1397[a] (1), 1532[a] (1), 1837[a] (1), 1838[a] (1)
24F	1	1	72 (1)
29	6	1	558[c] (6)
31	3	1	568 (3)
33F	2	2	636 (1), 1514[a] (1)
34	1	1	547 (1)
35	6	4	558[c] (3), 65 (1), 452 (1), 498 (1)
38	3	1	393 (3)
NT[e]	3	3	568 (1), 1186 (1), 1844[a] (1)

[a] These STs were newly identified within these studies and were deposited within the MLST database.

[b] Factor data were not available for all of the pneumococcal strains. Strains without type-specific data are listed according to their serogroup.

[c] STs identified in this study that express more than one capsule type.

[d] STs identified in this study that display a capsular switch when compared to strains entered in the MLST database.

[e] NT, nontypeable.

database (www.MLST.net), two of our isolates expressed a capsule type that differed from that of isolates of the same ST. One hundred ninety-three isolates were vaccine serotypes, and 136 were nonvaccine serotypes (Table 1). Seventy-five percent (257) of the isolates were vaccine serogroups. Factor-specific data were not available for 13 of our serogroup 6 isolates. Among the 329 isolates with serotype-specific data, the proportion that were vaccine serotypes differed significantly by body site ($P < 0.0001$). The proportion of vaccine serotypes was highest among CSF isolates (82%). Vaccine serotypes made up 51% of carriage isolates.

Vaccine serotype 6B contained a large proportion of unique STs; out of 38 isolates, we identified 24 different STs (63% unique). In contrast, serotype 6A contained 8 different STs among 25 isolates (32% unique). The vaccine serotype 19F contained 31 different STs among 55 isolates (56% unique), whereas the vaccine-related serotype 19A had 13 isolates that fell into 5 STs (38%). The diversity of STs within the carriage, middle ear, blood, and CSF collections was high; these data are presented in Table 2.

Distribution of neuraminidase genes. We examined the distribution of *nanA*, *nanB*, and *nanC* among our pneumococcal isolates. The *nanA* gene was present in 100% of our *S. pneumoniae* strains. Table 3 shows the distribution of *nanB* and *nanC* genes among our pneumococcal isolates. The expected size of the *nanB* PCR product was 500 bp. Nine strains (4 carriage, 3 middle ear, and 2 CSF) contained a *nanB* PCR product of more than 1.4 kb. These PCR fragments were sequenced and contained a 1,236-bp insertion with 99% nucleotide identity to an IS239 transposase (GenBank accession no. DQ351287) of *S. pneumoniae*. Because these nine strains contained an altered *nanB* gene, we excluded them from our prevalence estimates for *nanB*. The *nanB* gene was present in the majority of our remaining isolates (96%). *nanC* was the least prevalent of the neuraminidase genes and was present in slightly more than half of our *S. pneumoniae* strains.

Of the possible combinations of neuraminidase genes, the distribution among our isolates ($n = 329$) is as follows: 5 (1%) had *nanA* only, 160 (49%) had *nanA* and *nanB* only, 10 (3%) had *nanA* and *nanC* only, and 154 (47%) had *nanA*, *nanB*, and *nanC* genes.

We examined the distribution of *nanB* and *nanC* among our strains by ST. We identified 6 STs that contained 10 or more isolates (STs 13, 37, 81, 180, 199, and 236). *nanB* was present in 100% of the isolates of each of these STs. Three isolates in our collection were ST 393 and did not have *nanB*. STs 452, 1119, 1379, and 1533 were each represented by one isolate that did not contain *nanB*. Two STs were discordant for the presence of *nanB*: ST 193 contained two isolates, one of which had *nanB*, and ST 473 contained eight isolates, one of which had *nanB*. All of the ST 13 and 81 isolates contained *nanC*. In contrast, *nanC* was absent in all isolates of STs 37 and 180. Of the 50 STs that contained more than one isolate, 17 were discordant for the presence of *nanC* (Table 4). These results indicate that isolates of identical ST differ in the presence of *nanB* and *nanC*.

Associations of *nan* genes with isolates from different tissue sources. The distribution of *nanB* did not vary significantly among our carriage, middle ear, blood, and CSF collections (Table 3). The distribution of *nanC* varied significantly among strains from different tissue sources ($P = 0.03$). We also examined the prevalence of *nanC* among invasive and noninvasive strains using a log linear model. The prevalence of *nanC* was 1.32 times higher (95% confidence interval [CI], 1.01, 1.72)

TABLE 2. *S. pneumoniae* sequence types by isolate source

Isolate source	No. of isolates (no. of STs)	ST(s) (no. of isolates)
Blood	82 (50)[a]	199 (8), 220 (5), 376 (4), 433 (4), 81 (3), 113 (3), 393 (3), 13 (2), 37 (2), 62 (2), 191 (2), 227 (2), 632 (2), 659 (2), 1073 (2), 1390 (2), 36 (1), 43 (1), 66 (1), 72 (1), 124 (1), 132 (1), 156 (1), 176 (1), 180 (1), 193 (1), 218 (1), 244 (1), 251 (1), 385 (1), 425 (1), 452 (1), 498 (1), 558 (1), 636 (1), 638 (1), 641 (1), 643 (1), 678 (1), 690 (1), 899 (1), 1092 (1), 1269 (1), 1397 (1), 1402 (1), 1514 (1), 1516 (1), 1521 (1), 1522 (1), 1535 (1)
CSF	57 (36)	13 (7), 37 (3), 113 (3), 124 (2), 138 (2), 156 (2), 191 (2), 376 (2), 473 (2), 1212 (2), 1389 (2), 1391 (2), 1512 (2), 1524 (2), 146 (1), 15 (1), 81 (1), 90 (1), 166 (1), 305 (1), 355 (1), 460 (1), 547 (1), 643 (1), 689 (1), 867 (1), 1060 (1), 1073 (1), 1390 (1), 1392 (1), 1393 (1), 1394 (1), 1515 (1), 1536 (1), 1537 (1), 1892 (1)
Middle ear	100 (59)	236 (9), 13 (7), 81 (7), 180 (6), 271 (4), 156 (4), 62 (3), 199 (3), 460 (3), 67 (2), 83 (2), 90 (2), 376 (2), 37 (1), 113 (1), 124 (1), 125 (1), 138 (1), 146 (1), 153 (1), 177 (1), 191 (1), 193 (1), 246 (1), 274 (1), 384 (1), 425 (1), 558 (1), 629 (1), 639 (1), 671 (1), 802 (1), 1060 (1), 1119 (1), 1165 (1), 1177 (1), 1258 (1), 1396 (1), 1398 (1), 1508 (1), 1509 (1), 1510 (1), 1511 (1), 1513 (1), 1517 (1), 1518 (1), 1519 (1), 1520 (1), 1521 (1), 1523 (1), 1528 (1), 1529 (1), 1530 (1), 1531 (1), 1532 (1), 1533 (1), 1534 (1), 1890 (1), 1891 (1)
Carriage	103 (56)	558 (7), 473 (6), 37 (5), 199 (5), 460 (5), 1092 (5), 180 (4), 568 (4), 81 (3), 315 (3), 1843 (3), 62 (2), 66 (2), 236 (2), 257 (2), 385 (2), 439 (2), 1073 (2), 1839 (2), 13 (1), 33 (1), 65 (1), 90 (1), 113 (1), 147 (1), 156 (1), 251 (1), 320 (1), 376 (1), 384 (1), 496 (1), 626 (1), 643 (1), 651 (1), 656 (1), 816 (1), 1186 (1), 1190 (1), 1379 (1), 1388 (1), 1395 (1), 1448 (1), 1459 (1), 1538 (1), 1539 (1), 1757 (1), 1837 (1), 1838 (1), 1840 (1), 1841 (1), 1842 (1), 1844 (1), 1845 (1), 1846 (1), 1847 (1), 1889 (1)

[a] A total of 149 STs were identified in these studies. These numbers do not add up to 149 because several STs appeared among isolates from multiple sources.

among invasive isolates than among noninvasive isolates. We calculated individual prevalence ratios of *nanC* among middle ear, blood, and CSF using carriage isolates as the reference group. The prevalence ratios (PRs) and 95% confidence intervals were not statistically significant when comparing either

TABLE 3. Prevalence of *nanB* and *nanC* among *S. pneumoniae* isolates

Isolate group	No. of isolates with gene/no. total (%)	
	nanB[a,b]	*nanC*[c,d]
Blood	76/81 (94)	41/81 (51)
Cerebrospinal fluid	53/55 (96)	38/56 (68)
Total invasive	129/136 (95)	79/137 (58)
Middle ear	94/96 (98)	49/99 (50)
Carriage	93/99 (94)	45/103 (44)
Total noninvasive	187/195 (96)	94/202 (47)
Total	316/331 (96)	173/339 (51)

[a] Total no. does not equal 342. Nine isolates were excluded due to insertion elements within the *nanB* gene. Two additional strains were excluded because we were unable to confirm the presence of *nanB* by Southern or dot blot hybridization.
[b] For distribution among sites, $P = 0.47$; for distribution by invasiveness, $P = 0.65$.
[c] Three isolates were excluded because we were unable to confirm the presence of *nanC* by Southern or dot blot hybridization.
[d] For distribution among sites, $P = 0.03$; for distribution by invasiveness, $P = 0.04$.

middle ear isolates to carriage isolates (PR, 1.12; 95% CI, 0.85, 1.48; $n = 202$) or for blood isolates compared to carriage isolates (PR, 1.16; 95% CI, 0.84, 1.62; $n = 184$ isolates). The prevalence of *nanC* was 1.41 times higher (95% CI, 1.11, 1.79) among CSF isolates than among carriage isolates ($n = 159$ strains).

The greater prevalence of *nanC* among CSF isolates could be due to an association of *nanC* with certain serotypes or STs that were more common among CSF strains. We were unable to detect an association between *nanC* and CSF isolates when controlling for individual serotypes. We observed a trend towards statistical significance when controlling for individual STs ($P = 0.09$). The lack of statistical significance was likely due to a lack of power resulting from the large number of STs and serotypes. We used contingency table analysis controlling for vaccine serotypes. After controlling for vaccine serotypes, *nanC* was still associated with CSF isolates compared to carriage isolates ($P = 0.02$).

TABLE 4. STs containing more than one isolate that are discordant for *nanC*[a]

ST	No. of isolates	No. (%) of isolates with *nanC*
124	4	2 (50)
138	3	1 (33)
156	8	1 (13)
199	16	3 (19)
236	11	1 (9)
376	9	8 (89)
385	3	2 (67)
460	9	1 (11)
473	8	3 (38)
558	9	3 (33)
568	4	3 (75)
1092	6	1 (17)
1390	3	1 (33)

[a] STs 67, 251, 439, and 1060 each contained two isolates, one with *nanC* and one without.

DISCUSSION

S. pneumoniae strains cause a significant amount of morbidity and mortality in young children (40). Conjugate polysaccharide vaccines are highly effective in preventing invasive disease due to vaccine serotypes. However, concerns have been raised that the protection afforded by the vaccine will be compromised over time by the spread of nonvaccine serotypes or strains that have emerged due to capsular switching (25, 31, 42). The extent of this problem will depend largely on the ability of nonvaccine strains to cause disease. If replacing strains are equally virulent, we may see an increase in invasive disease due to nonvaccine serotypes or strains that have emerged due to capsular switching (16, 34). Recent studies indicate that this is occurring (5). If pneumococcal clones have a high level of virulence only in conjunction with a specific vaccine capsule type, then capsule switches to nonvaccine serotypes will not lead to an increase in invasive disease.

Studies evaluating the relationship between capsule type and the genetic background of a given strain have often relied on MLST data (4, 16, 17). MLST is a powerful technique but does not capture the total genetic variation within individual pneumococcal isolates. While it is generally known that pneumococcal strains undergo a high level of horizontal transfer, this study indicates that pneumococcal clones (as defined by ST) not only undergo serotype switching but also vary in the presence or absence of other important virulence-associated loci.

The survival of pneumococcal strains within different ecological niches of the body is likely to involve distinct adaptations (27, 38). The distribution of the pneumococcal metalloprotease, encoded by *zmpC*, is variable and has been associated with pneumonia (28). Signature-tagged mutagenesis has identified putative virulence factors specific for pneumonia (33) and for colonization of mucosal surfaces (18). An ATP-binding cassette transporter, the Ami-AliA/AliB permease, has been shown to be important for colonization and not invasive disease in a mouse model of infection (22). In addition to tissue-specific virulence factors, new research indicates that specific virulence factors are important for transition between tissue sites (29).

The high prevalence of *nan*A suggests that it is an essential gene for colonization and pathogenesis in all pneumococcal strains. *S. pneumoniae* isolates containing *nanC* were overrepresented among our invasive isolates. The prevalence of *nanC* was 41% higher among CSF isolates than among carriage isolates. Serotypes 1, 4, 8, and 12F have been associated with invasive disease (37). We did not have any serotype 8 strains in our study. *nanC* was present in all of the serotype 4 and 12F strains. We identified three serotype 1 strains, two from blood that both had *nanC* and one CSF isolate that lacked *nanC*. STs 113, 124, 138, and 191 were identified within our strain collection and have been associated with invasiveness (4, 17). *nanC* was present in each of the ST 113 (*n* = 17) isolates. In contrast, none of the ST 191 (*n* = 5) isolates contained *nanC*. Among the ST 124 isolates, *nanC* was present in two CSF isolates and absent in one blood and one middle ear isolate. ST 138 was also discordant for *nanC*. The gene was present in one CSF isolate and absent in one CSF and one middle ear isolate.

It is possible that *nanC* is not directly involved in tissue-specific virulence and is simply associated with STs and/or serogroups that are more likely to cause meningitis. While we cannot definitively rule out this possibility, *nanC* remained significantly associated with CSF isolates after controlling for vaccine serotypes. These results suggest that the presence of *nanC* contributes to the ability of pneumococci to cause meningitis. Future experiments involving the evaluation of *nanC* mutants in animal models should be performed to define more precisely the role of *nanC* in pneumococcal pathogenesis.

The fact that our strain collections were not all sampled from the same population must be considered as a limitation of our study. Large *S. pneumoniae* collections that include carriage, middle ear, and invasive pneumococcal isolates sampled from the same geographic location within the United States during the same point in time are difficult to obtain, and this study demonstrates the importance of establishing such collections.

Another limitation of our study is that the carriage, middle ear, and invasive strain collections were serotyped in different laboratories. Thirteen of the 149 STs (9%) identified in our study displayed capsule types that either differed from other isolates of the same ST within our study or in comparison to the international MLST database. We confirmed the serotypes of strains exhibiting a capsule switch using a multiplex PCR approach (30). A recent paper indicates that levels of capsule switching in pneumococci are likely higher than has previously been suspected (19). Our results were similar; these investigators identified capsule switches in 11 out of 109 STs (10%) in their collection of clinical isolates from Scotland.

In summary, the results of this study repeat the growing theme of significant genetic variability among pneumococci. These studies demonstrate that virulence determinants, other than capsule loci, vary among strains of the same ST. In addition, these results affirm that *nanA* is essential for virulence and that having *nanC* may predispose strains for invasion of the CSF. Studies that incorporate MLST and take into account the presence or absence of additional virulence determinants will provide a greater understanding of the virulence potential of pneumococcal strains.

REFERENCES

1. **Austrian, R.** 1976. The quellung reaction, a neglected microbiological technique. Mt. Sinai J. Med. **43:**669–705.
2. **Berry, A. M., R. A. Lock, and J. C. Paton.** 1996. Cloning and characterization of *nanB*, a second *Streptococcus pneumoniae* neuraminidase gene, and purification of the NanB enzyme from recombinant *Escherichia coli*. J. Bacteriol. **178:**4854–4860.
3. **Bogaert, D., A. van Belkum, M. Sluijter, A. Luijendijk, R. de Groot, H. C. Rumke, H. A. Verbrugh, and P. W. Hermans.** 2004. Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. Lancet **363:**1871–1872.
4. **Brueggemann, A. B., D. Griffiths, E. Meats, T. Peto, D. W. Crook, and B. G. Spratt.** 2003. Clonal relationships between invasive and carriage *Streptococcus pneumoniae*. J. Infect. Dis. **187:**1424–1432.

5. **Byington, C. L., M. H. Samore, G. J. Stoddard, S. Barlow, J. Daly, K. J. Korgenski, S. Firth, D. Glover, J. Jensen, E. O. Mason, C. K. Shutt, and A. T. Pavia.** 2005. Temporal trends in invasive disease due to *Streptococcus pneumoniae* among children in the intermountain West: emergence of nonvaccine serogroups. Clin. Infect. Dis. **41:**21–29.

6. **Camara, M., G. J. Boulnois, P. W. Andrew, and T. J. Mitchell.** 1994. A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. Infect. Immun. **62:**3688–3695.

7. **Claverys, J. P., M. Prudhomme, I. Mortier-Barriere, and B. Martin.** 2000. Adaptation to the environment: *Streptococcus pneumoniae*, a paradigm for recombination-mediated genetic plasticity? Mol. Microbiol. **35:**251–259.

8. **Coffey, T. J., M. C. Enright, M. Daniels, R. Morona, W. Hryniewicz, J. C. Paton, and B. G. Spratt.** 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*. Mol. Microbiol. **27:**73–83.

9. **Enright, M. C., and B. G. Spratt.** 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology **144:**3049–3060.

10. **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.

11. **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.

12. **Feil, E. J., J. M. Smith, M. C. Enright, and B. G. Spratt.** 2000. Estimating recombinational parameters in *Streptococcus pneumoniae* from multilocus sequence typing data. Genetics **154:**1439–1450.

13. **Ghaffar, F., T. Barton, J. Lozano, L. S. Muniz, P. Hicks, V. Gan, N. Ahmad, and G. H. McCracken.** 2004. Effect of the 7-valent pneumococcal vaccine on nasopharyngeal colonization by *Streptococcus pneumoniae* in the first 2 years of life. Clin. Infect. Dis. **39:**930–938.

14. **Greenland, S.** 2004. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am. J. Epidemiol. **160:**301–305.

15. **Hakenbeck, R., N. Balmelle, B. Weber, C. Gardes, W. Keck, and A. de Saizieu.** 2001. Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. Infect. Immun. **69:**2477–2486.

16. **Hanage, W. P., K. Auranen, R. Syrjanen, E. Herva, P. H. Makela, T. Kilpi, and B. G. Spratt.** 2004. Ability of pneumococcal serotypes and clones to cause acute otitis media: implications for the prevention of otitis media by conjugate vaccines. Infect. Immun. **72:**76–81.

17. **Hanage, W. P., T. Kaijalainen, R. Syrjanen, K. Auranen, M. Leinonen, P. H. Makela, and B. G. Spratt.** 2005. Invasiveness of serotypes and clones of *Streptococcus pneumoniae* among children in Finland. Infect. Immun. **73:**431–435.

18. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389–1405.

19. **Jefferies, J. M. C., A. Smith, S. C. Clarke, C. G. Dowson, and T. Mitchell.** 2004. Genetic analysis of diverse disease causing pneumococci indicates high levels of diversity within serotypes and capsule switching. J. Clin. Microbiol. **42:**5681–5688.

20. **Kadioglu, A., S. Taylor, F. Iannelli, G. Pozzi, T. J. Mitchell, and P. W. Andrew.** 2002. Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is affected by pneumolysin deficiency and differences in capsule type. Infect. Immun. **70:**2886–2890.

21. **Kelly, T., J. P. Dillard, and J. Yother.** 1994. Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*. Infect. Immun. **62:**1813–1819.

22. **Kerr, A. R., P. V. Adrian, S. Estevao, R. de Groot, G. Alloing, J. P. Claverys, T. J. Mitchell, and P. W. Hermans.** 2004. The Ami-AliA/AliB permease of *Streptococcus pneumoniae* is involved in nasopharyngeal colonization but not in invasive disease. Infect. Immun. **72:**3902–3906.

23. **King, S. J., K. R. Hippe, J. M. Gould, D. Bae, S. Peterson, R. T. Cline, C. Fasching, E. N. Janoff, and J. N. Weiser.** 2004. Phase variable desialylation of host proteins that bind to *Streptococcus pneumoniae* in vivo and protect the airway. Mol. Microbiol. **54:**159–171.

24. **Knecht, J., G. Schiffman, and R. Austrian.** 1970. Some biological properties of pneumococcus type 37 and the chemistry of its capsular polysaccharide. J. Exp. Med. **132:**475–487.

25. **Lipsitch, M.** 1999. Bacterial vaccines and serotype replacement: lessons from *Haemophilus influenzae* and prospects for *Streptococcus pneumoniae*. Emerg. Infect. Dis. **5:**336–345.

26. **Magee, A. D., and J. Yother.** 2001. Requirement for capsule in colonization by *Streptococcus pneumoniae*. Infect. Immun. **69:**3755–3761.

27. **Marra, A., J. Asundi, M. Bartilson, S. Lawson, S. Fang, J. Christine, C. Wiesner, Brighanm, W. P. Schneider, and A. E. Hromockyj.** 2002. Differential fluorescence induction analysis of *Streptococcus pneumoniae* identifies genes involved in pathogenesis. Infect. Immun. **70:**1422–1433.

28. **Oggioni, M. R., G. Memmi, T. Maggi, D. Chiavolini, F. Iannelli, and G. Pozzi.** 2003. Pneumococcal zinc metalloprotease ZmpC cleaves human metalloproteinase 9 and is a virulence factor in experimental pneumonia. Mol. Microbiol. **49:**795–805.

29. **Orihuela, C. J., G. Gao, K. P. Francis, and E. I. Tuomanen.** 2004. Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. J. Infect. Dis. **190:**1661–1669.

30. **Pai, R., R. E. Gertz, and B. Beall.** 2006. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. J. Clin. Microbiol. **44:**124–131.

31. **Pelton, S. I.** 2000. Acute otitis media in the era of effective pneumococcal conjugate vaccine: will new pathogens emerge? Vaccine **19:**S96–S99.

32. **Pettigrew, M. M., and K. P. Fennie.** 2005. Genomic subtraction followed by dot blot screening of *Streptococcus pneumoniae* clinical and carriage isolates identifies genetic differences associated with strains that cause otitis media. Infect. Immun. **73:**2805–2811.

33. **Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari, and D. Simon.** 1998. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:**5620–5629.

34. **Posfay-Barbe, K. M., and E. R. Wald.** 2004. Pneumococcal vaccines: do they prevent infection and how? Curr. Opin. Infect. Dis. **17:**177–184.

35. **Ramirez, M., and A. Tomasz.** 1999. Acquisition of new capsular genes among clinical isolates of antibiotic-resistant *Streptococcus pneumoniae*. Microb. Drug Resist. **5:**241–246.

36. **Robinson, K. A., W. Baughman, G. Rothrock, N. L. Barrett, M. Pass, C. Lexau, B. Damaske, K. Stefonek, B. Barnes, J. Patterson, E. R. Zell, A. Schuchat, C. G. Whitney, and Active Bacterial Core Surveillance /Emerging Infections Program Network.** 2001. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995–1998: opportunities for prevention in the conjugate vaccine era. JAMA **285:**1729–1735.

37. **Sandgren, A., K. Sjostrom, B. Olsson Liljequist, B. Christensson, A. Samuelsson, G. Kronvall, and B. Henriques Normark.** 2004. Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*. J. Infect. Dis. **189:**785–796.

38. **Sebert, M. E., L. M. Palmer, M. Rosenberg, and J. N. Weiser.** 2002. Microarray-based identification of *htr*A, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization. Infect. Immun. **70:**4059–4067.

39. **Shakhnovich, E. A., S. J. King, and J. N. Weiser.** 2002. Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. Infect. Immun. **70:**7161–7164.

40. **Shinefield, H. R., and S. Black.** 2000. Efficacy of pneumococcal conjugate vaccines in large-scale field trials. Pediatr. Infect. Dis. J. **19:**394–397.

41. **Spiegelman, D., and E. Hertzmark.** 2005. Easy SAS calculations for risk or prevalence ratios and differences. Am. J. Epidemiol. **162:**199–200.

42. **Spratt, B. G., and B. M. Greenwood.** 2000. Prevention of pneumococcal disease by vaccination: does serotype replacement matter? Lancet **356:**1210–1211.

43. **Tettelin, H., and S. K. Hollingshead.** 2004. Comparative genomics of *Streptococcus pneumoniae*: intrastrain diversity and genome plasticity, p. 15–29. *In* E. Toumanen, T. J. Mitchell, D. A. Morrison, and B. G. Spratt (ed.), The pneumococcus. ASM Press, Washington, D.C.

44. **Tong, H. H., M. A. James, I. Grants, G. Liu, G. Shi, and T. F. DeMaria.** 2000. Comparison of structural changes of cell surface carbohydrates in the Eustachian tube of epithelium of chinchillas infected with *Streptococcus pneumoniae* neuraminidase-deficient mutant or its isogenic parent strain. Microb. Pathog. **31:**309–317.

45. **Wald, E. R., E. O. Mason, Jr., J. S. Bradley, W. J. Barson, S. L. Kaplan, and the U.S. Pediatric Multicenter Pneumococcal Surveillance Group.** 2001. Acute otitis media caused by *Streptococcus pneumoniae* in children's hospitals between 1994 and 1997. Pediatr. Infect. Dis. J. **20:**34–39.

Editor: J. N. Weiser