INFECTION AND IMMUNITY, May 2005, p. 2805–2811 Vol. 73, No. 5
0019-9567/05/$08.00+0 doi:10.1128/IAI.73.5.2805–2811.2005

Genomic Subtraction Followed by Dot Blot Screening of *Streptococcus pneumoniae* Clinical and Carriage Isolates Identifies Genetic Differences Associated with Strains That Cause Otitis Media

Melinda M. Pettigrew[1]* and Kristopher P. Fennie[2]

Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut,[1] *and Yale University School of Nursing, New Haven, Connecticut*[2]

Received 13 October 2004/Returned for modification 30 November 2004/Accepted 3 January 2005

Streptococcus pneumoniae **strains are the leading cause of bacterial otitis media, yet little is known about specific bacterial factors important for this disease. We utilized a molecular epidemiological approach involving genomic subtraction of the *S. pneumoniae* serogroup 19 middle ear strain 5093 against the laboratory strain R6. Resulting subtraction PCR (sPCR) products were used to screen a panel of 93 middle ear, 90 blood, 35 carriage, and 58 cerebrospinal fluid isolates from young children to identify genes found more frequently among middle ear isolates. Probe P41, similar to a hypothetical protein of *Brucella melitensis*, occurred among 41% of middle ear isolates and was found 2.8 (95% confidence interval [CI], 1.32 to 6.5), 3.3 (95% CI, 1.9 to 5.7), and 1.8 (95% CI, 1.1 to 3.0) times more frequently among middle ear strains than carriage, blood, or meningitis strains, respectively. sPCR fragment H10, similar to an unknown *Streptococcus agalactiae* protein, was present in 31% of middle ear isolates and occurred 3.6 (95% CI, 1.2 to 11.2), 2.8 (95% CI, 1.5 to 5.4), and 2.6 (95% CI, 1.2 to 5.5) times more often among middle ear isolates than carriage, blood, or meningitis strains, respectively. These studies have identified two genes of potential importance in otitis media virulence. Further studies are warranted to outline the precise role of these genes in otitis media pathogenesis.**

Streptococcus pneumoniae is a gram-positive diplococcus that colonizes the human upper respiratory tract and causes disease under certain circumstances. Among young children, it is a common cause of invasive bacterial infections, such as pneumonia, septicemia, and meningitis, and is responsible for 30 to 50% of otitis media cases (2, 20, 26). The annual burden of pneumococcal otitis media in the United States among children younger than 5 years of age is estimated at 7 million cases (26). Research suggests that acute otitis media caused by *S. pneumoniae* is clinically more severe than acute otitis media caused by other bacterial pathogens (15, 20, 22, 27). Rates of antibiotic resistance are increasing among *S. pneumoniae* isolates (6, 25), and the *S. pneumoniae* conjugate vaccine is of limited effectiveness for otitis media (8).

Despite the importance of *S. pneumoniae* as an agent of otitis media, little is known about the specific bacterial virulence factors important for invasion of the middle ear space. Research indicates that *S. pneumoniae* strains differ in their ability to cause disease. For example, serogroups 19, 6, 23, 14, 3, and 18 are the most likely to cause otitis media and account for over 73% of middle ear isolates (4). A chinchilla model of otitis media suggested that capsule type influences otitis media pathogenesis, as a serotype 3 strain was shown to produce more attenuated otitis media than a type 23B strain (10). The genetic background of these strains was not known, and the expression of additional genes may have influenced the severity of the middle ear infection in this study. A study of 672 penicillin-resistant pneumococcal isolates showed that certain genotypes, as defined by pulsed-field gel electrophoresis, were more prevalent among middle ear isolates than among isolates from other specimen sources (29). These studies lend support to the hypothesis that certain genetic subsets of pneumococcal strains are more likely to cause otitis media. In contrast, another study compared the frequency of serotypes and clones that cause otitis media with the frequency of the serotypes and clones carried among healthy Finnish children and concluded that most pneumococcal carriage serotypes and clones are equally capable of causing otitis media (13).

We used a molecular epidemiological (31) approach involving genomic subtraction followed by a dot blot hybridization screening of a panel of pneumococcal isolates to identify genes that might play a role in otitis media. These experiments were based on the hypotheses that *S. pneumoniae* isolates differ in their ability to cause otitis media, that these differences in pathogenic potential are based on genetic differences among strains beyond capsule type, and that genes found more frequently among middle ear isolates than in carriage, blood, and cerebrospinal fluid isolates offer a selective advantage for invasion of the middle ear space.

MATERIALS AND METHODS

Otitis media isolates. Two hundred pneumococcal middle ear isolates were obtained from Edward O. Mason at the Baylor College of Medicine in Houston, Texas. These isolates were collected as part of an ongoing 9-year surveillance study of pneumococcal infections in children (the U.S. Pediatric Multicenter Pneumococcal Surveillance Group). Isolates of *S. pneumoniae* were obtained from middle ear fluid by swab of spontaneous drainage, by myringotomy, by tympanocentesis, or at surgery for placement of tympanostomy tubes at one of

* Corresponding author. Mailing address: Yale University School of Medicine, Department of Epidemiology and Public Health, 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.

five participating hospitals (Texas Children's Hospital in Houston, Children's Hospital of Pittsburgh, Children's Hospital—San Diego, Columbus Children's Hospital, and Arkansas Children's Hospital). Ninety-three of these were randomly selected for the present study. Meningitis and blood isolates from children less than 5 years of age were obtained from the Active Bacterial Core Surveillance Emerging Infections Program network at the Centers for Disease Control and Prevention. *S. pneumoniae* throat isolates were collected in a study of *Haemophilus influenzae* and *S. pneumoniae* colonization among healthy children less than 3 years of age attending 16 licensed day care centers in Washtenaw County, Michigan (9).

Selection and description of strains used for differential cloning by subtractive hybridization. In order to identify genes associated with pneumococcal otitis media, we conducted genomic subtraction of the serogroup 19 middle ear strain 5093 against the laboratory strain R6. The middle ear strain 5093 was originally selected as the tester strain because serogroup 19 was the most common group among our otitis media isolates (35 of 93 strains were group 19). Furthermore, pulsed-field gel electrophoresis demonstrated that this strain represented a common electrophoretic type among group 19 strains (data not shown). The middle ear strain 5093 was later typed by multilocus sequence typing using methods described previously by others (7), and it has been entered into the MLST database (http://www.mlst.net). The laboratory reference strain R6 was chosen as the driver because it is a nonencapsulated laboratory strain and is known to have lost many genes, some of which were likely critical for virulence. Furthermore, the genomic sequence has been fully determined (16), thus facilitating our ability to verify that the identified subtraction products were 5093 specific.

Differential cloning by subtractive hybridization. Subtractive hybridization was conducted using a commercially available kit from Clontech (PCR-Select bacterial genomic subtraction kit; Palo Alto, CA), which is based on the suppressive subtractive hybridization method (5, 12). Briefly, this method involves the ligation of primers to the strain of interest (tester) followed by hybridization of the tester DNA with a reference strain (driver). The PCR is then used to amplify the unhybridized, tester-specific sequences. Genomic DNA from the tester (5093) and driver (R6) were isolated using a Wizard genomic DNA isolation kit according to the manufacturer's instructions (Promega, Madison, WI). The pooled secondary PCR products identified by subtractive hybridization were cloned into the vector pCR2.1-TOPO (Invitrogen, Carlsbad, CA).

Analysis of sPCR inserts. Subtraction PCR (sPCR) probes were amplified from 192 subtraction clones using T7 and M13 reverse primers (35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 2 min). The PCR products were purified using a QIAquick PCR purification kit, and samples were sent for DNA sequence analysis to the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University. DNA sequences were compared to those in the NCBI database (http://www.ncbi.nih.gov/BLAST/) in order to estimate similarity with published sequences, to verify that sequences were tester specific, and to identify duplicate clones.

Screening of *S. pneumoniae* isolates. The presence or absence of each unique, tester-specific sPCR fragment was evaluated within the *S. pneumoniae* collection by dot blot hybridization as follows: each *S. pneumoniae* isolate was streaked out on Trypticase soy agar plates with 5% sheep blood and incubated at 37°C with 5% CO_2 overnight. A colony from each plate was used to inoculate a 96-well plate containing 800 μl of Todd-Hewitt broth per well and incubated overnight at 37°C. The 96-well plates were then centrifuged at 3,000 rpm in a bench top centrifuge with a horizontal rotor at 1,000 × *g* for 20 min. The supernatant was discarded, and the pellets were resuspended in 800 μl of lysis buffer (0.4 M NaOH, 10 mM EDTA). The plates were incubated at 80°C for 20 min. Eighty μl of DNA lysate from each well was blotted onto Hybond N^+ membranes (Amersham Pharmacia Biotech, Piscataway, N.J.) using a Bio-Dot microfiltration apparatus (Bio-Rad, CA). Blotting was followed by a wash step using 80 μl 0.4 M NaOH per well. Blots were allowed to air dry, and the DNA was cross-linked to the membranes. The tester strains 5093 (positive control) and R6 (negative control) were placed on each membrane. Three blots were made for each hybridization experiment, one containing 93 middle ear isolates plus controls, one containing 90 blood isolates and controls, and the third containing 58 cerebrospinal fluid (CSF) isolates and 35 carriage isolates plus controls.

Each unique, 5093-specific sPCR fragment was labeled with alkaline phosphatase, and dot blots were hybridized at 65°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.). Each sPCR probe was used to separately screen two sets of three blots. Strain samples that produced discrepant hybridization results with a particular sPCR probe were retested by Southern blot hybridization.

Cloning of DNA sequences surrounding otitis media-associated sPCR probes. The regions flanking sPCR probes P41 and H10 were obtained by PCR using the commercially available Universal Genome Walker kit (Clontech, Palo Altos, CA) according to the manufacturer's instructions. Briefly, 5093 genomic DNA was digested separately using each of four different blunt-ended restriction enzymes. These separate pools of DNA were independently ligated to adaptors to create four different Genome Walker libraries. In order to obtain upstream and downstream sequences, a two-step touch-down PCR was conducted using an adaptor primer (provided with the kit) and a gene-specific primer (designed based on the sPCR sequence of either P41 or H10). Gene-specific primers were as follows: P41Up1, 5′ GTTTTCAAACCATATTGCAAATCCAAACC 3′; P41Dn1, 5′ CTCTCTCCCTGTAATTAATCAACCTGCT 3′; H10Up1, 5′ GTC CCTATTTCTAAATAATTCGGTGATAC 3′; and H10Dn1, 5′ TGCCACGAA TTTATTTCCCAATAATTCTG 3′. PCR conditions were 7 cycles at 94°C for 25 s and 70°C for 4 min, 35 cycles of 94°C for 25 s and 65°C for 4 min, and a final 4-min extension step at 65°C. Each PCR mixture was diluted 1:50, and then 1 μl was used for a second round of PCR using a nested gene-specific primer and a nested adaptor-specific primer. Nested gene-specific primers were P41Up2, 5′-CCAACTATATAAGTAATATTCATATCTTTG-3′; P41Dn2, 5′-TCGATAAA AAATACAATGAGAATCCACATC-3′; H10 Up2, 5′-GCAATATCTGATAT ACATGGTCACCTAG-3′; and H10Dn2, 5′-ATCGTTAATTTTCGTATAATA CTCGTTTC-3′. Nested PCR conditions were 5 cycles at 94°C for 25 s and 70°C for 4 min, 22 cycles of 94°C for 25 s and 65°C for 4 min, and a final 4-min extension step at 65°C. The resulting PCR DNA fragments were cloned into pCR2.1-TOPO (Invitrogen, Carlsbad, CA) and transformed into DH5α cells, and the plasmids were sequenced at W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University. Sequences were edited and assembled using Lasergene Navigator software from DNASTAR.

Statistical analyses. Differences in the proportions of each sPCR probe among each pneumococcal collection were calculated by χ^2 test ($\alpha = 0.05$). A Bonferroni adjustment was used to adjust for multiple comparisons (44 unique sPCR probes); an association was considered significant if the *P* value was less than 0.0011 (0.05/44). Prevalence ratios were calculated as the ratio of the proportion of middle ear isolates with the sPCR fragment of interest to the proportion of either carriage, blood, or meningitis isolates with the sPCR fragment of interest (reference group). Only those probes with prevalence ratios that were significant, as determined by 95% confidence intervals (CIs), were considered for further analysis. Differences in the proportions of the P41 and H10 probe among each pneumococcal collection, stratified into serogroup 19 and non-serogroup 19 strains, were calculated by Fisher's exact test. Statistical calculations were done using SAS version 8.0 (SAS Institute, Cary, NC).

Nucleotide sequence accession number. The middle ear strain 5093 sequence was entered into the MLST database as sequence type ST-1396.

RESULTS

Fifty-two unique, middle ear strain-specific clones were obtained from the subtraction procedure and used to screen the pneumococcal strain collection. Despite a lack of DNA sequence similarity, eight sPCR fragments cross-hybridized with the driver strain R6 and were not examined further. The sizes of and potential matches in the NCBI database for the remaining 44 sPCR fragments are given in Table 1. Twelve of the 44 sPCR probes were similar to known pneumococcal genes or to pneumococcal phage or hypothetical proteins. Two probes, P125 and H174, were similar to capsule biosynthesis genes. Three of the probes had amino acid similarity to proteins associated with antibiotic resistance (P40, H115, and H129). Seventeen of the 45 sPCR fragments (38%) are similar to either hypothetical proteins or proteins of unknown function.

The distribution of each sPCR fragment among *S. pneumoniae* middle ear, carriage, blood, and CSF isolates is given in Table 2. Probes H147 and P164, with 30% amino acid identity to tetanus toxin C fragment, and probe H147, with 57% similarity to a hypothetical *Arabidopsis* protein, hybridized only to the tester strain 5093. The distribution of 11 of the probes differed significantly across the collection, as measured by a χ^2 test. Individual prevalence ratios and 95% CIs comparing the

TABLE 1. sPCR genomic subtraction probes and their size, percent amino acid identity, and potential function

sPCR probe	Size (bp)	% Identity	BLAST match and potential function
Known *S. pneumoniae* proteins			
P10	905	99	Cell wall surface anchor family protein *S. pneumoniae* TIGR4
P125	409	87	UDP-glycosyltransferase *S. pneumoniae*
H110	283	98	Phosphotransferase system, IIA component *S. pneumoniae* TIGR4
H115	153	100	ABC transporter *S. pneumoniae*
H174	281	81	Protein-tyrosine phosphatase CpsB *S. pneumoniae*
H182	850	82	Type 1 restriction modification *S. pneumoniae* TIGR4
Phage-related proteins			
P20	410	88	Integrase/recombinase phage family *S. pneumoniae* TIGR4
P23	397	99	ISI380 Spn1 transposase *S. pneumoniae* TIGR4
H21	597	50	Transposase *L. mesenteroides*
H50	707	82	cI repressor bacteriophage EJ-1
H51	464	46	Integrase/recombinase, phage integrase family *S. pneumoniae* TIGR4
Known proteins from other bacterial species			
P38	324	34	Seven TM receptor *Caenorhabditis elegans*
P40	594	38	ABC multidrug/lipid transporter *Thermoanaerobacter tergcongensis*
P46	613	32	Putative primase *Staphylococcus aureus* MW2
P53	446	34	Putative ligase *Oryza sativa*
P65	1,270	32	COG4566: response regulator *Burkholderia fungorum*
P67	409	100	Tn*916* FtsK/SpoE *S. agalactiae*
P91	324	41	Putative transcriptional regulatory protein *Clostridium tetani* E88
P124	652	26	Protein phosphatase-1 regulatory subunit *Plasmodium yoelii*
P155	1154	26	Putative collagen binding protein *Streptococcus pyogenes* MGAS8232
P164	760	30	Tetanus neurotoxin C fragment
H62	548	35	DNA primase *Lactococcus lactis*
H87	374	93	COG0741: soluble lytic murein transglycosylase *Enterococcus faecium*
H121	627	89	Tn*916*, FtsK/SpoIIIE family protein *S. agalactiae* 2603V/R
H129	248	94	Tetracycline resistance protein TetM *Enterococcus faecalis*
H139	511	28	Fibrinogen-binding protein Efb *S. aureus*
H141	447	50	5′-*N*-acetyltransferase *Oceanobacillus iheyensis* HTE831
Hypothetical proteins or those of unknown function			
P3	246	62	Hypothetical protein *S. pyogenes*
P6	371	59	Hypothetical protein SP1134 *S. pneumoniae* TIGR4
P24	690	26	Unknown environmental sequence
P41	681	37	Hypothetical protein BMEI11681 *Brucella melitensis*
P44	524	36	Unknown environmental sequence
P48	180	92	Open reading frame 16 enterococcus
P57	625	40	Hypothetical conserved protein *Magnetospirillum magnetotacticum*
P94	319	48	Hypothetical protein *S. pyogenes*
P132	316	38	Hypothetical protein PG0565 *Porphyromonas gingivalis* W83
P152	288	81	Hypothetical protein *S. pneumoniae* TIGR4
H10	763	75	Unknown *S. agalactiae* NEM316
H71	321	94	Hypothetical protein *S. agalactiae* 2603V/R
H116	725	69	Unknown *S. agalactiae* NEM316
H118	773	24	Hypothetical protein OB3378 *O. iheyensis* HTE831
H142	292	88	Conserved hypothetical protein SP1612 *S. pneumoniae* TIGR4
H147	383	57	Hypothetical protein At2g16180 *Arabidopsis thaliana*
H165	399	29	Hypothetical protein *Plasmodium falciparum* 3D7

middle ear isolates with either throat, blood, or CSF isolates were calculated for each of the probes to identify sPCR fragments that occurred at a high frequency among middle ear isolates when compared separately to carriage or invasive isolates. Thirty sPCR probes hybridized more frequently to middle ear isolates than to blood isolates (Table 3). Ten probes were found more frequently among middle ear isolates than among CSF isolates, and two probes hybridized more frequently to middle ear isolates than to carriage isolates. Probes H10 and P41 were selected as having potential importance in otitis media pathogenesis, because these two probes occurred more frequently among middle ear strains than among carriage, blood, or CSF isolates. The prevalence ratios for these two probes are given in Table 4.

The distribution of pneumococcal capsule types differs among carriage, otitis media, and invasive isolates. Probes P41 and H10 were isolated from a serogroup 19 strain, and group 19 was the most common serogroup among our middle ear isolates (Table 5). To determine whether the greater prevalence of P41 and H10 among otitis media strains was due simply to an association of these probes with serogroup 19 strains, we examined the distribution of these two probes within the pneumococcal strain collections stratified into group 19 strains and non-serogroup 19 strains. The frequency distri-

TABLE 2. Distribution of sPCR probes among middle ear, carriage, blood, and CSF *S. pneumoniae* isolates

sPCR probe	No. (%) for:[b]			
	Middle ear	Carriage	Blood	CSF
P10	28 (30)	12 (34)	16 (18)	19 (33)
P125	31 (33)	7 (20)	18 (20)	21 (36)
H110	43 (46)	20 (57)	41 (46)	26 (45)
H115	25 (27)	10 (29)	7 (8)	17 (29)
H174	30 (32)	7 (20)	13 (14)	12 (21)
H182	60 (65)	30 (86)	61 (68)	50 (86)
P20	41 (44)	10 (29)	48 (53)	33 (57)
P23	60 (65)	17 (49)	59 (66)	34 (59)
H21[a]	23 (25)	6 (17)	2 (2)	8 (14)
H50	10 (11)	5 (14)	2 (2)	7 (12)
H51	29 (31)	13 (37)	11 (12)	20 (34)
P38	28 (30)	6 (17)	9 (10)	19 (33)
P40	15 (6)	5 (14)	2 (2)	7 (12)
P46	18 (19)	6 (17)	2 (6)	7 (21)
P53	18 (19)	6 (17)	3 (3)	12 (21)
P65[a]	32 (34)	19 (54)	19 (21)	28 (48)
P67[a]	26 (28)	5 (15)	3 (3)	7 (12)
P91	48 (52)	14 (40)	22 (24)	22 (38)
P124	23 (25)	8 (23)	19 (21)	12 (20)
P155[a]	20 (22)	3 (9)	1 (1)	3 (5)
P164	1 (1)	0 (0)	0 (0)	0 (0)
H62	16 (17)	6 (17)	3 (3)	5 (9)
H87	17 (18)	5 (14)	2 (2)	4 (7)
H121	22 (24)	9 (26)	5 (6)	7 (12)
H129[a]	22 (24)	8 (23)	4 (4)	6 (10)
H139	10 (11)	5 (14)	2 (2)	7 (12)
H141	22 (24)	9 (26)	14 (16)	19 (33)
P3	22 (24)	9 (26)	11 (12)	14 (24)
P6	10 (11)	4 (11)	3 (3)	7 (12)
P24	21 (23)	4 (11)	8 (9)	2 (3)
P41[a]	44 (47)	6 (17)	13 (14)	15 (26)
P44[a]	16 (17)	3 (9)	2 (2)	1 (2)
P48[a]	26 (28)	8 (23)	5 (6)	7 (12)
P57[a]	16 (17)	6 (17)	2 (2)	1 (2)
P94	19 (20)	6 (17)	10 (11)	18 (31)
P132	20 (22)	5 (14)	13 (14)	14 (24)
P152	13 (14)	6 (17)	9 (10)	11 (19)
H10[a]	29 (31)	3 (9)	10 (11)	7 (12)
H71[a]	29 (31)	9 (26)	4 (4)	9 (16)
H116	27 (29)	9 (26)	9 (10)	12 (21)
H118	23 (25)	9 (26)	9 (10)	8 (14)
H142	25 (27)	11 (31)	21 (23)	18 (31)
H147	1 (1)	0 (0)	0 (0)	0 (0)
H165	25 (27)	9 (26)	7 (8)	10 (17)

[a] sPCR probes with a significantly different distribution across *S. pneumoniae* collections as determined by χ^2 test after Bonferroni adjustment for multiple comparisons ($\alpha = 0.0011$).

[b] The number of isolates positive in the dot blot for the given probe and the percent positive among the strains tested in that group of pneumococcal isolates are shown. For middle ear isolates, $n = 93$; for carriage isolates, $n = 35$; for blood isolates, $n = 90$; and for CSF isolates, $n = 58$.

TABLE 3. sPCR probes occurring more frequently among middle ear isolates than among carriage, blood, or CSF *S. pneumoniae* isolates[a]

sPCR probe	Result for isolate indicated		
	Carriage	Blood	CSF
P10			
P125		X	
H110			
H115		X	
H174		X	
H182			
P20			
P23			
H21		X	
H50		X	
H51		X	
P38		X	
P40		X	
P46		X	
P53		X	
P65		X	
P67		X	
P91		X	
P124			
P155		X	X
P164			
H62		X	
H87		X	X
H121		X	
H129		X	X
H139		X	
H141			
P3			
P6		X	
P24		X	X
P41	X	X	X
P44		X	X
P48		X	X
P57		X	X
P94			
P132			
P152			
H10	X	X	X
H71		X	X
H116		X	
H118		X	
H142			
H147			
H165		X	

[a] Prevalence ratios and 95% CIs were calculated separately for each sPCR fragment to compare the prevalence among *S. pneumoniae* middle ear strains using either carriage, blood, or meningitis strains as the reference group. An X indicates that this probe hybridized significantly more frequently to middle ear isolates than did the reference group. For middle ear isolates, $n = 93$; for carriage isolates, $n = 35$; for blood isolates, $n = 90$; and for CSF isolates, $n = 58$.

butions of these probes among the stratified pneumococcal collections are shown in Table 6.

Because of the association of probes H10 and P41 with middle ear strains, we conducted a genome walk up- and downstream from probes H10 and P41 in an attempt to obtain the surrounding DNA sequences. The genome walk using H10-specific primers produced a 1,577-bp fragment of DNA (GenBank accession number AY845429) encompassing the H10 sPCR fragment. H10 contains a 339-bp gene with 74% identity to an unknown protein from *Streptococcus agalactiae* NEM316 (11). H10 is in between a gene with 47% amino acid identity across 117 amino acids to FtsK-like DNA segregation ATPase of *Bacillus subtilis* and 41% similarity to a putative serine/threonine phosphatase of *Bacillus cereus*. The genome walk with P41-specific primers produced a 2,035-bp fragment of DNA. Sequence analysis indicates that P41 has a 1,017-bp open reading frame, and the translated gene is 339 amino acids

TABLE 4. Prevalence ratios and 95% CIs for the prevalence of P41 and H10 in *S. pneumoniae* middle ear strains using either carriage, blood, or CSF strains as the reference group

Collection	Prevalence ratio (95% CI) for:	
	P41	H10
Carriage	2.8 (1.3–6.5)	3.6 (1.2–11.2)
Blood	3.3 (1.9–5.7)	2.8 (1.5–5.4)
CSF	1.8 (1.1–3.0)	2.6 (1.2–5.5)

TABLE 5. Distribution of capsule serogroups among *S. pneumoniae* strain collections

Collection	No. (%) for serogroup indicated[a]							
	4	6	9	14	18	19	23	Other
Middle ear (n = 91)	0 (0)	14 (15)	4 (4)	10 (11)	1 (1)	35 (38)	11 (12)	16 (18)
Carriage (n = 34)	0 (0)	14 (41)	1 (3)	2 (6)	1 (3)	5 (15)	3 (9)	8 (24)
Blood (n = 90)	3 (3)	13 (14)	7 (8)	4 (4)	7 (8)	14 (16)	6 (7)	36 (40)
CSF (n = 55)	1 (2)	15 (27)	3 (5)	11 (20)	7 (13)	6 (11)	8 (15)	4 (7)
Total (n = 270)[b]	4 (1)	56 (21)	15 (6)	27 (10)	16 (6)	60 (22)	28 (10)	64 (24)

[a] The percentages are those of strains of each serotype within each strain collection.
[b] Serogroup data are missing for six of the *S. pneumoniae* isolates used in this study.

with low similarity (37% identity across 81 amino acids) to hypothetical protein BMEI1681 of *Brucella melitensis* strain 16 M (GenBank accession number AY845429). P41 is situated between a gene with 96% amino acid identity to a ribosomal protein L11 methyltransferase of *S. pneumoniae* strain TIGR4 and a partial match with 61% amino acid identity to a hypothetical protein of *Leuconostoc mesenteroides*.

DISCUSSION

Genomic subtraction of the type 19 middle ear strain 5093 against the laboratory strain R6 and a hybridization screen of pneumococcal disease and carriage isolates identified two sPCR fragments that hybridized significantly more often to middle ear isolates than to carriage or invasive isolates from young children. Probe P41 occurred among 41% of the middle ear isolates and was found more frequently among these isolates than in carriage, blood, and meningitis strains. sPCR fragment H10 was present in 31% of middle ear isolates and also occurred more often among middle ear isolates than in carriage, blood, and CSF isolates. These sPCR probes are absent in the sequenced R6 and TIGR4 reference strains. Our results suggest that otitis media is not simply a disease of opportunity that results from the overgrowth of colonizing strains following an viral infection, but that special bacterial characteristics may be required for carriage strains to invade the middle ear.

Hanage et al. compared the frequency of serotypes and clones that cause otitis media with the frequency of the serotypes and clones carried in healthy Finnish children to determine whether all carriage isolates are equally capable of causing otitis media (13). The authors found two serotypes, 19F and 23F, significantly associated with otitis media. The association with otitis media was not based solely on capsule, because three multilocus sequence types (MLST) expressed capsule 23F yet differed in their propensity to cause otitis media, thus indicating the importance of genetic factors other than capsule in otitis media pathogenesis. However, these authors concluded that most pneumococcal carriage serotypes and clones were equally capable of causing otitis media, because otitis media clones and serotypes were found in a relative frequency that was proportional to their prevalence in carriage studies. This finding does not contradict our results, because the pneumococcal clones were defined by MLST, which measures allelic variation among seven housekeeping genes. Individual sequence types that express different capsule types were described (13), suggesting that the total genomic content of clones, as measured by MLST, varies due to the horizontal transfer of individual genes. Thus, pneumococcal carriage and otitis media clones could appear similar as determined by capsule and MLST type but differ in the content of specific virulence-associated genes.

The survival of pneumococcal strains within different ecological niches of the body is likely to involve distinct adaptations. Signature-tagged mutagenesis has identified putative virulence factors specific for pneumonia (28) and for colonization of mucosal surfaces (14). Differential fluorescence induction analysis has also identified tissue-specific putative virulence factors important for the invasion of different tissue sites (23). A putative serine protease, HtrA, has been shown to be important in nasopharyngeal colonization (30). Recently, 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 (19). Given the identification of virulence factors important for pneumonia and colonization, it is reasonable to think that specific genes would enhance the ability to invade the middle ear space.

The combination of the *S. pneumoniae* capsule type and the genetic background of the strain is important for determining virulence (1, 3, 18). Furthermore, the influence of the combination of capsule and genetic background in pathogenesis differs depending on the site of infection (17). Interestingly, genes necessary for capsule biosynthesis, which are known to be critical for virulence, have not been identified in signature-tagged mutagenesis experiments (14, 21, 28). This is probably because acapsular mutants do not grow well in vitro. Our experiments identified sPCR fragments with similarity to genes important for capsule synthesis (Table 1, sPCR fragments H174 and P125). Identification of these sPCR fragments highlights the power of this technique to identify genes that would be missed by mutagenesis screens because they are necessary

TABLE 6. Distribution of probe P41 and probe H10 among *S. pneumoniae* strain collections stratified by serogroup 19 or all other serogroups combined

Probe	Serogroup	No. (%) for isolates from:			
		Middle ear	Carriage	Blood	CSF
P41	19 (n = 60)[a]	23 (66)	2 (40)	3 (21)	5 (83)
	Other (n = 210)	21 (37)	4 (14)	10 (13)	9 (18)
H10	19 (n = 60)	17 (49)	1 (20)	2 (14)	1 (17)
	Other (n = 210)	11 (20)	2 (7)	8 (11)	6 (12)

[a] Among serogroup 19 strains, the distribution of P41 varied significantly across *S. pneumoniae* collections as determined by Fisher's Exact test at $\alpha = 0.05$.

for in vitro growth. Our results also identified the *tetM* gene encoding antibiotic resistance (sPCR fragment H129), which was found at higher frequency among day care and middle ear isolates, consistent with the higher rates of antibiotic resistance seen with these populations (6).

Thirty sPCR probes occurred significantly more frequently among middle ear isolates than among blood isolates. In comparison, 10 and 2 probes hybridized more frequently to middle ear isolates than to CSF and carriage isolates, respectively. Qualitatively, these results suggest that, as a group, our middle ear strains are genetically most similar to carriage isolates and more different from blood isolates than meningitis isolates. This result is intriguing, given that an experimental meningitis model with gerbils showed that *S. pneumoniae* strains can cause otitis media and then invade the central nervous system without a detectable bacteremic state (24).

One limitation of our study is that our strains were collected in different geographic regions, and pneumococcal serotypes are known to vary between different areas. Serogroup 19 was the most common serogroup among our middle ear strains. It is possible that P41 and H10 are associated with serogroup 19 strains, and that serogroup 19 strains were more prevalent in the region from which our middle ear strains were collected. In this case, P41 and H10 could be markers for serogroup 19 strains instead of markers for otitis media virulence. We attempted to control for this by examining the prevalence of our probe among serogroup 19 strains and non-serogroup 19 strains separately. Among serogroup 19 strains, P41 occurred in 83% of CSF, 66% of middle ear, 40% of carriage, and 20% of blood isolates. The high prevalence of P41 among group 19 CSF isolates may indicate that it is important for meningitis pathogenesis among these strains or may be due to the low number of group 19 CSF strains. Among non-19 serogroups, P41 occurred at the highest frequency in middle ear strains. Probe H10 occurred in 49% of group 19 middle ear strains and 20% of non-serogroup 19 middle ear strains. The distribution trend of P41 and H10 among the pneumococcal strains collections supports our hypothesis that these probes are associated with otitis media. However, with the exception of P41 in serogroup 19 strains, the differences in probe distribution between strain collections were not statistically significant. This is likely due in part to the low numbers of isolates within each group after stratification.

sPCR probes P41 and H10 contain DNA sequences with similarity to proteins of unknown function, and it is therefore difficult to speculate about their precise role in colonization of the middle ear space. Furthermore, these genes may not be directly involved in otitis media virulence but may instead serve as a marker for other genes linked to these on the chromosome. For example, H10 lies next to a putative serine/threonine phosphatase that could be important for otitis media pathogenesis. Strain differences in virulence are also likely due in part to variation in protein expression between strains or differences in expression patterns by disease site (i.e., different genes may be expressed in the throat versus the middle ear). These issues are not addressed in the present study. Nevertheless, our approach, involving genomic subtraction of otitis media isolates from the laboratory strain R6 followed by a hybridization screen of pneumococcal isolates, has identified two genes of potential importance in otitis media and placed them in perspective regarding their relative importance within a population of pneumococcal strains. Future studies of the prevalence of these genes among a larger collection of isolates, RNA expression studies, and mechanistic studies using an otitis media animal model will shed additional light on the role of these genes in otitis media pathogenesis.

ACKNOWLEDGMENTS

We thank Edward Mason, Betsy Foxman, Janet Gilsdorf, and the Active Bacterial Core Surveillance Emerging Infections Program network at the Centers for Disease Control and Prevention for use of their *S. pneumoniae* strains. We also thank Janet Gilsdorf for critical reading of the manuscript.

Funding for this research was provided to M.M.P. by the National Institute on Deafness and Other Communication Disorders (R21 DC006260).

REFERENCES

1. **Blue, C. E., and T. J. Mitchell.** 2003. Contribution of a response regulator to the virulence of *Streptococcus pneumoniae* is strain dependent. Infect. Immun. **71:**4405–4413.
2. **Bluestone, C. D., J. S. Stephenson, and L. M. Martin.** 1992. Ten-year review of otitis media pathogens. Pediatr. Infect. Dis. J. **11:**S7–S11.
3. **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:**73–78.
4. **Dagan, R., R. Melamed, M. Muallen, L. Piglansky, D. Greenberg, O. Abramson, P. M. Mendelman, N. Bohidar, and P. Yagupsky.** 1996. Reduction of nasopharyngeal carriage of pneumococci during the second year of life by a heptavalent conjugate pneumococcal vaccine. J. Infect. Dis. **174:**1271–1278.
5. **Diatchenko, L., Y. F. Lau, A. P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E. D. Sverdlov, and P. D. Siebert.** 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA **93:**6025–6030.
6. **Doern, G. V., A. B. Brueggemann, H. Huynh, E. Wingert, and P. Rhomberg.** 1999. Antimicrobial resistance with *Streptococcus pneumoniae* in the United States, 1997–98. Emerg. Infect. Dis. **5:**757–765.
7. **Enright, M., and B. Spratt.** 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology **144:**3049–3060.
8. **Eskola, J., T. Kilpi, A. Palmu, J. Jokinen, J. Haapakoski, E. Herva, A. Takala, H. Kayhty, P. Karma, R. Kohberger, G. Siber, P. H. Makela, et al.** 2001. Efficacy of a pneumococcal conjugate vaccine against acute otitis media. N. Engl. J. Med. **344:**403–409.
9. **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.
10. **Giebink, G. S.** 1977. Comparison of otitis media due to *Streptococcus pneumoniae* types 3 and 23 in the chinchilla model. J. Infect. Dis. **136S:**191–195.
11. **Glaser P., C. Rusniok, C. Buchrieser, F. Chevalier, L. Frangeul, T. Msadek, M. Zouine, E. Couve, L. Lalioui, C. Poyart, P. Trieu-Cuot, and F. Kunst.** 2002. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. Mol. Microbiol. **45:**1499–1513.
12. **Gurskaya, N. G., L. Diatchenko, A. Chenchik, P. D. Siebert, G. L. Khaspekov, K. A. Lukyanov, L. L. Vagner, O. D. Ermolaeva, S. A. Lukyanov, and E. D. Sverdlov.** 1996. Equalizing cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemagglutinin and phorbol 12-myristate 13-acetate. Anal. Biochem. **240:**90–97.
13. **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.
14. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389–1405.
15. **Heikkinen, T., F. Ghaffar, A. O. Okorodudu, and T. Chonmaitree.** 1998. Serum interluekin-6 in bacterial and nonbacterial acute otitis media. Pediatrics **102:**296–299.
16. **Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-**

Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709–5717.

17. **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.
18. **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.
19. **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.
20. **Kilpi, T., E. Herva, T. Kaijalainen, R. Syrjanen, and A. K. Takala.** 2001. Bacteriology of acute otitis media in a cohort of Finnish children followed for the first two years of life. Pediatr. Infect. Dis. J. **20:**654–662.
21. **Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison, and D. W. Holden.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:** 555–571.
22. **Leibovitz, E., S. Raiz, L. Piglansky, D. Greenberg, P. Yagupsky, D. M. Fliss, A. Leiberman, and R. Dagan.** 1998. Resistance pattern of middle ear fluid isolates in acute otitis media recently treated with antibiotics. Pediatr. Infect. Dis. J. **17:**463–469.
23. **Marra, A., J. Asundi, M. Bartilson, S. Lawson, S. Fang, J. Christine, C. Wiesner, D. Brigham, 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.
24. **Marra, A., and D. Brigham.** 2001. *Streptococcus pneumoniae* causes experimental meningitis following intranasal and otitis media infections via a nonhematogenous route. J. Infect. Dis. **69:**7318–7325.
25. **McCormick, A. W., C. G. Whitney, M. M. Farley, R. Lynfield, L. H. Harrison, N. M. Bennett, W. Schaffner, A. Reingold, J. Hadler, P. Cieslak, M. H. Samore, and M. Lipsitch.** 2003. Geographic diversity and temporal trends of antimicrobial resistance in *Streptococcus pneumoniae* in the United States. Nat. Med. **9:**424–430.
26. **Overturf, G. D., and the Committee on Infectious Diseases.** 2000. American Academy of Pediatrics: technical report: prevention of pneumococcal infections, including the use of pneumococcal conjugate and polysaccharide vaccines and antibiotic prophylaxis. Pediatrics **106:**367–376.
27. **Pitkaranta, A., A. Virolainen, J. Jero, E. Arruda, and F. G. Hayden.** 1998. Detection of rhinovirus, respiratory syncytial virus, and coronavirus infections in acute otitis media by reverse transcriptase polymerase chain reaction. Pediatrics **102:**291–295.
28. **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.
29. **Richter, S. S.** 2002. The molecular epidemiology of penicillin-resistant *Streptococcus pneumoniae* in the United States, 1994–2000. Clin. Infect. Dis. **34:** 330–339.
30. **Sebert, M. E., L. M. Palmer, M. Rosenberg, and J. N. Weiser.** 2002. Microarray-based identification of *htrA*, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization. Infect. Immun. **70:**4059–4067.
31. **Zhang, L., B. Foxman, S. D. Manning, P. Tallman, and C. F. Marrs.** 2000. Molecular epidemiologic approaches to urinary tract infection gene discovery in uropathogenic *Escherichia coli*. Infect. Immun. **68:**2009–2015.

Editor: A. D. O'Brien