

Pneumococcal *pspA* Sequence Types of Prevalent Multiresistant Pneumococcal Strains in the United States and of Internationally Disseminated Clones

BERNARD BEALL,^{1*} GIOVANNI GHERARDI,^{1†} RICHARD R. FACKLAM,¹
AND SUSAN K. HOLLINGSHEAD²

*Respiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta, Georgia 30333,¹ and
Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35205²*

Received 26 April 2000/Returned for modification 1 July 2000/Accepted 1 August 2000

In a recent genotypic survey of β -lactam-resistant pneumococci recovered in different areas of United States during 1997, eight clonal types that each represented 3 to 40 isolates accounted for 134 of 144 isolates (G. Gherardi, C. Whitney, R. Facklam, and B. Beall, *J. Infect. Dis.* 181:216–229, 2000). We determined the degree of *pspA* gene diversity among these 134 isolates and for 11 previously characterized internationally disseminated multiresistant strains. Thirty-four different *pspA* restriction profiles were determined for an amplicon encompassing the variable portion of the structural gene that encodes the surface-exposed domain of PspA and a variable-length proline-rich putative cell wall-associated domain. These restriction profiles closely correlated with those of 33 different *pspA* sequence types of an approximately 230-residue region corresponding to residues 182 to 410 of the strain Rx1 PspA. These residues encompass a 100-residue clade-defining region known to contain cross-protective epitopes for which 17 sequence types were found. Distinct, conserved *pspA* sequence types were found for the majority of strains within seven of the eight U.S. clonal types assessed, while one pulsed-field gel electrophoresis type was represented by isolates of three distinct PspA clades. Sequence typing of *pspA* provides an added level of specificity in the subtyping of isolates and is a necessary first step in determining the components needed in a PspA vaccine which could elicit effective cross-protective coverage.

Streptococcus pneumoniae is a major cause of morbidity and mortality worldwide and is the leading cause of bacterial pneumonia, meningitis, bacteremia, and otitis media. Pneumococcal polyvalent purified capsular polysaccharide vaccines are serotype specific in their protection, contain only 23 of the 90 known capsular types, and are unable to elicit effective immunologic responses in children younger than 2 years of age and many elderly individuals (6, 13). Although conjugate vaccines should be much more effective in these individuals, they are still limited to offering protection against specific capsular types. For these reasons, there is much interest in developing pneumococcal protein vaccines. One protein vaccine candidate, PspA, is an antigenically variable surface virulence factor (16) that interferes with complement-mediated clearance of pneumococci during bacteremia (1, 2). Antibodies to PspA protect mice against lethal systemic infection, and this protection is often cross-protective for strains of different capsular serotypes expressing distinct PspA molecule types (1, 2, 17, 19). Immunization of humans with a single recombinant PspA stimulated antibodies broadly cross-reactive to heterologous PspA molecules (23). Although it is known that PspA is broadly cross-reactive, cross-reactivity does not necessarily correspond to cross-protection, and the total array of common circulating PspA types should be determined before the degree of cross-protection afforded by specific sequence types can be assessed. It is particularly important to assess PspA variability among strains with serotypes not targeted by impending conjugate

vaccines and circulating strains that have become highly resistant to antibiotics used for treatment. Highlighting the importance of the latter group is the recent report that high-level penicillin resistance was an independent parameter of mortality in pneumococcal bacteremia in a population with high human immunodeficiency virus seroprevalence (27). For the study presented here, we examined the *pspA* sequences of multiresistant, genetically characterized U.S. isolates recovered in 1997 and 11 previously characterized internationally disseminated multiresistant strains.

MATERIALS AND METHODS

Isolates. Eight genetically related sets comprised 134 of 144 multiresistant isolates that were recovered in 1997 and that were described previously (9). These isolates were obtained from the Emerging Infections Program/Active Bacterial Core Surveillance (available at <http://www.cdc.gov/ncidod/dbmd/abcs>) conducted in seven states within the United States.

Previously characterized clones. The following previously genetically characterized strains were provided by L. McDougal and F. Tenover and were designated according to the Pneumococcal Molecular Epidemiology Network (15). These included clones Spain^{23F}-1 (strain SP193, ATCC 700669) (4, 21), Spain^{6B}-2 (GM17, SP194, ATCC 700670) (22), France^{9V}-3 (strain SP195, ATCC 700671; highly related to strain 665) (4, 24), Tennessee^{23F}-4 (strain SP196, ATCC 51916) (20, 24), England¹⁴-9 (strain SP200, ATCC 700676) (10), Spain¹⁴-5 (strain VH14, SP197, ATCC 700672) (5), Hungary^{19A}-6 (SP220, ATCC 700673) (21), South Africa^{19A}-7 (SP198, strain 17619, ATCC 700674) (25), South Africa^{6B}-8 (strain SP199, ATCC 700675) (24), Slovakia¹⁴-10 (strain 91-006571, SP221, ATCC 700677) (8), and Slovakia^{19A}-11 (strain 91-0006571, SP222, ATCC 700678) (8).

PCR, restriction profiling, and sequencing. Primers F (GCCAGCGTCGC TATCTTAGGGGCTGG) and R (GAACCATCAGTATTGTAGAAGTAC CA) were used for PCR and were derived by comparisons of conserved regions between the two sequences corresponding to GenBank accession numbers U89711 (28) and A41971 (9). Primer R was used to obtain the sequences of 550 to 780 bases used for analysis. Restriction analysis of the amplicon generated with primers F and R was accomplished as described previously for penicillin-binding protein gene amplicons (9) following digestion of 10 μ l of unpurified PCR products with 2 U each of the enzymes *Rsa*I, *Hinf*I, and *Hae*III. Amplicons showing identical triple-digest patterns were additionally compared by *Dde*I

* Corresponding author. Mailing address: Centers for Disease Control and Prevention, Mailstop C02, 1600 Clifton Rd., NE, Atlanta, GA 30333. Phone: (404) 639-1237. Fax: (404) 639-3123. E-mail: beb0@cdc.gov.

† Present address: Libera Università Campus Bio-Medico, 00155, Rome, Italy.

TABLE 1. *pspA* RFLP types and sequence accession numbers for 1997 multiresistant isolates and internationally disseminated clones that were genotyped

Reference strain (PFGE type)	No. of isolates	PFGE subtypes ^a	Serotype (no. of isolates)	<i>pspA</i> GenBank accession no. ^b	<i>pspA</i> RFLP type ^c	<i>pspA</i> clade (closest match/% identity) ^d
Spain ^{23F} -1(A1)	16	A2–A3, A5, A6	23F (11), 19F (5)	AF254259	a1	3 (BG8090/97.1)
	2	A2	23F (2)	AF255542	a2	3 (BG8090/97.1)
	3	A2	23F (2), 19F(1)	AF255543	a3	3 (BG8090/92.2)
	1	A2	19F (1)	AF254256	a4	3 (BG8090/96.1)
	1	A4	23F (1)	AF288751	a5	5 (ATCC6303/98.3)
France ^{9V} -3 (B1)	33	B2–B5, B8–B15, B21, B22	9V (16), 9A (10), 14 (9)	AF252286	b1	3 (AC122/100)
	1	B6	9V (1)	AF252286	b4	3 (AC122/100)
	2	B2	9V (2)	AF253404	b2	3 (AC122/100)
	1	B7	9V (1)	AF253405	b3	3 (AC122/99)
	1	B13	14 (1)	AF253406	c1	1 (AC94/97)
England ¹⁴⁻⁹ (C1)	31	C1–C11	14 (30)	AF253406	c1	1 (AC94/97)
Strain 121 (D1)	1	D1	6B (1)	AF255548	d3	3 (BG8090/99)
	2	D4	23F (1), 6B (1)	AF255547	d1	3 (BG8090/99)
	1	D7	6A (1)	AF255547	d2	3 (BG8090/99)
	2	D3 (1), D5 (1)	6B (2)	AF255549	d4	1 (AC94/93)
	1	D2	6B (1)	AF255550	d5	1 (BG6692/100)
	1	D8	6B (1)	AF255551	d6	1 (BG6692/100)
	2	D6, D9	23F (1), 6B (1)	AF255552	d7	2 (DBL5/96)
	Tennessee ^{23F} -4 (E1)	6	E1	23F	AF253407	e1
3		E2, E5	23F (6)	AF255900	e2	1 (BG6843/92)
3		E2, E7, E8	23F (3)	AF255901	e3	1 (BG6843/92)
3		E2	23F (3)	AF255902	e4	1 (BG6843/92)
3		E2	23F (3)	AF255903	e5	1 (BG6843/92)
1		E3	23F (1)	AF255904	e6	1 (BG6843/92)
1		E7	23F (1)	AF255905	e7	1 (BG6843/92)
1		E4	23F (1)	AF255906	e8	1 (BG6843/92)
1		E6	23F (1)	AF255907	e9	1 (BG6843/92)
Strain 127 (G1)	2	G1	19F (3)	AF25554	g1	3 (AC122/100)
	1	G2	19F (1)	AF255545	g2	3 (AC122/100)
Strain 69 (H1)	4	H1	14 (4)	AF255908	h1	3 (AC122/96.1)
Spain ^{6B} -2 (K1)	4	K1–K5	6B	AF254257	k1	1 (DBL1/100)
South Africa ^{6B} -8 (M)		M1	6B	AF254258	m1	1 (BG6692/98)
Slovakia ^{19A} -11 (N)		N1	19A	AF254255	n1	5 (ATCC6303/94.9)
Hungary ^{19A} -6 (O)		O1	19A	AF254254	o1	5 (ATCC6303/99.2)
South Africa ^{19A} -7 (P)		P1	19A	AF253408	p1	4 (BG7817/99)
Spain ¹⁴ -5 (α)		α 1	14	AF253406	c1	1 (AC94/97)
Slovakia ¹⁴ -10 (β)		β 1	14	AF255546	c1	1 (AC94/97)

^a Isolates were genotyped previously (9). For each group those isolates that differed by only one to six bands from subtype 1 were assigned a common type (A, B, C, etc.). For each type, isolates with a difference from subtype 1 of more than six bands of were considered unrelated isolates and were assigned a different PFGE type. Reference strains 121, 127, and 69 represent new pneumococcal PFGE types (D, G, and H, respectively) (9) and are included in the number of 1997 U.S. isolates given in the second column.

^b Each different accession number represents a unique *pspA* sequence of 550 to 775 bases encompassing the clade-defining region and variable-length proline-rich repeat region.

^c The following *pspA* RFLP types shared identical clade-defining regions: a1 and a2; b1, b2, b4, g1, and g2; b3 and c1; d1 to d3; d5 and d6; and e1 to e9.

^d Percent identity to the clade-defining region of closest match from reference 11.

digestion in the same manner to ensure that there was true representation of highly similar amplicons. Amplicons from each of the 11 internationally dispersed clones were also profiled by restriction analysis and sequenced. Independent amplicons from at least two 1997 isolates from each identical restriction profile set were sequenced. When possible, *pspA* amplicons from isolates that differed in serotype, geographic source, and pulsed-field gel electrophoresis (PFGE) profile from within each identical *pspA* restriction profile set were used for sequence analysis.

DNA sequence analysis. The translation of the primer R-generated sequence complement was used for amino acid sequence comparisons by using the Wisconsin Package, version 10.0, software (Genetics Computer Group, Madison, Wis.).

Nucleotide sequence accession numbers. The GenBank accession numbers for the partial *pspA* sequences encompassing roughly codons 180 to 410 of the Rx*A* *pspA* gene (19) are listed in Table 1.

RESULTS AND DISCUSSION

Primers were selected based upon sequences conserved between the *pspA* DNA sequence of strain EF5668 (19) and the PspA amino acid sequence of JY2008 (28), which were the only two PspA sequences accessible from GenBank at the time

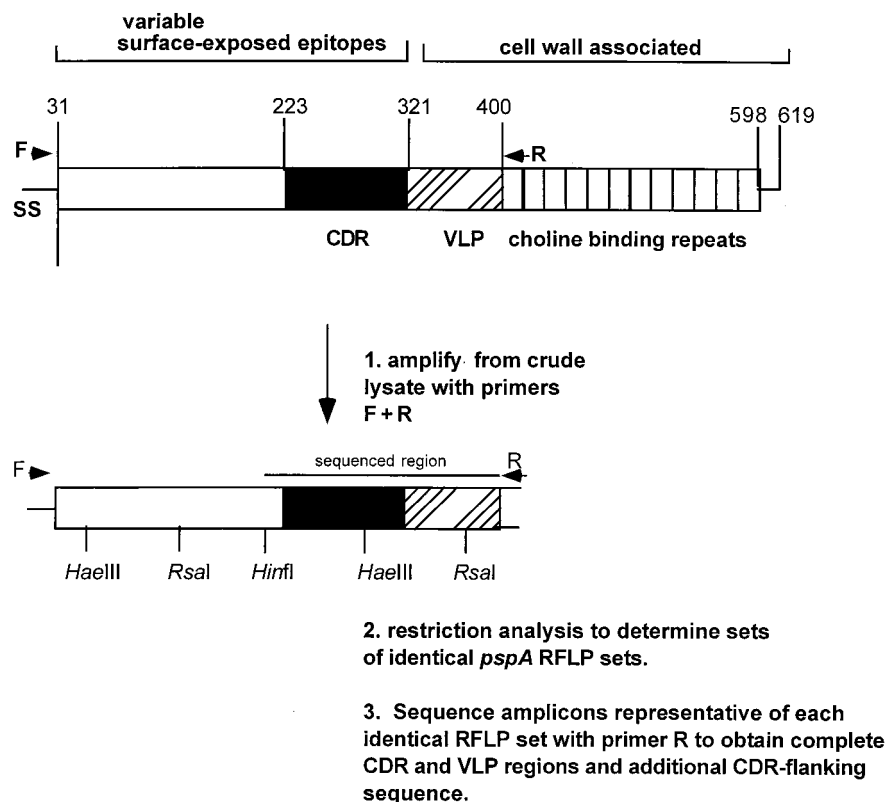


FIG. 1. General features of PspA and method used for sequence typing of *pspA* amplicons. The *pspA* codon positions are taken from references 3 and 11 and are based upon the Rx1 *pspA* sequence (27). Abbreviations: SS, signal sequence; CDR, clade-defining region containing cross-protective epitopes; VLP, variable-length proline-rich repeat region.

that this study was initiated. Two primers were selected for amplification based on their ability to generate a single 1,000- to 1,500-bp amplicon from all pneumococcal strains tested. Primer F annealed to the conserved signal sequence codons 12 to 20, while primer R annealed to codons 411 to 419 and 465 to 473 of the strain EF5668 and strain JY2008 *pspA* genes, respectively (Fig. 1). Thus, this amplicon carries the sequence that encodes the α -helical surface-exposed region with most antigenic epitopes and the closely situated variable-length proline-alanine repeat region (VLP in Fig. 1) (3, 11, 18, 28). The surface-exposed region encompasses the 100-residue *pspA* clade-defining region which includes protection-eliciting epitopes (3, 11).

Previously, sequence-based divisions among diverse PspA proteins classified these proteins into three families that were further subdivided into six clades (3, 11). The clade-defining regions of PspA proteins in different clades vary at >20% of amino acid positions (3). Family 1 is composed of clades 1 and 2, and family 2 is composed of clades 3 to 5. Family 3 has only one highly divergent member. Although both clades and families are easily recognized sequence-based divisions of PspA proteins, it is the family-level division that appears to be the most important serologically. Rabbit antisera made to recombinant PspA proteins could reliably distinguish PspA families but could not always distinguish PspA clades within the same family (23; S. Hollingshead and D. E. Briles, unpublished data).

Figure 1 summarizes the method of *pspA* amplicon-based subtyping used. Table 1 summarizes the *pspA* amplicon sequencing and restriction pattern results. The PspA sequences

of only families 1 and 2, consisting of clades 1 to 5, were found among the multiresistant pneumococcal strains used for this study. Seventy-three of the 134 1997 U.S. isolates (54%), representing five distinct PFGE types, were found to be of PspA clade 3, and the sequences of their clade 3-defining regions diverged from those of previously sequenced clade 3-defining regions by no more than 7.8%. Fifty-eight isolates (43%), representing three PFGE types, were found to contain clade 1 PspA sequences, with their clade-defining regions having 92 to 100% identity to previously sequenced clade 1-defining regions. Only one clade 2 PspA sequence was found among these 1997 U.S. isolates, from two independent PFGE type D isolates comprising two different serotypes (subtypes D2 and D9 in Fig. 2). Only one clade 5 isolate was found among the 1997 U.S. isolates and consisted of a PFGE type A isolate (subtype A4 in Fig. 2). The sequences of the clade 5-defining regions of two internationally disseminated clones (PFGE types N1 and O1 in Table 1) were found to have 95 to 99% amino sequence identity to the clade-defining region of clade 5 strain ATCC 6303 (11). Finally, a single clade 4 PspA sequence was found from the internationally disseminated strain South Africa^{19A}-7 (PFGE type P1 in Table 1). The relation of clades 3 to 5 (family 2) compared to clades 1 and 2 (family 1) is evident in Fig. 3, in which the clade-defining regions representing all of the sequences examined in this study are aligned.

As reported previously (3, 11), the 650- to 775-bp region of *pspA* sequenced was remarkably variable, with the more distantly related sequences (e.g., clade 1 versus clade 3) sharing only about 40% sequence identity at either the DNA or the protein level. For isolates of PFGE types A, B, C, H, and K,



FIG. 2. PspA clades of previously determined PFGE types (9) of multiresistant, invasive pneumococcal isolates. The dendrogram gives underestimated Dice coefficients because closely migrating bands could not be properly resolved with the program used. Visually identical PFGE profiles gave average Dice coefficients of ~0.86 to 0.93. UPGMA, unweighted pair group method with arithmetic averages. The strains with the closest matches in their PspA clade-defining regions (11) are listed along with their *pspA* GenBank accession numbers.

one specific *pspA* amplicon restriction fragment length polymorphism (RFLP) pattern represented the majority of isolates (89 of 101 [88.1%]) obtained within a PFGE type (Table 1). For isolates of PFGE types A, B, C, and K, the major *pspA*

sequence and RFLP type was additionally shared with those of previously characterized internationally disseminated antibiotic-resistant strains of the same PFGE types. With only three exceptions, each specific RFLP type was associated with a

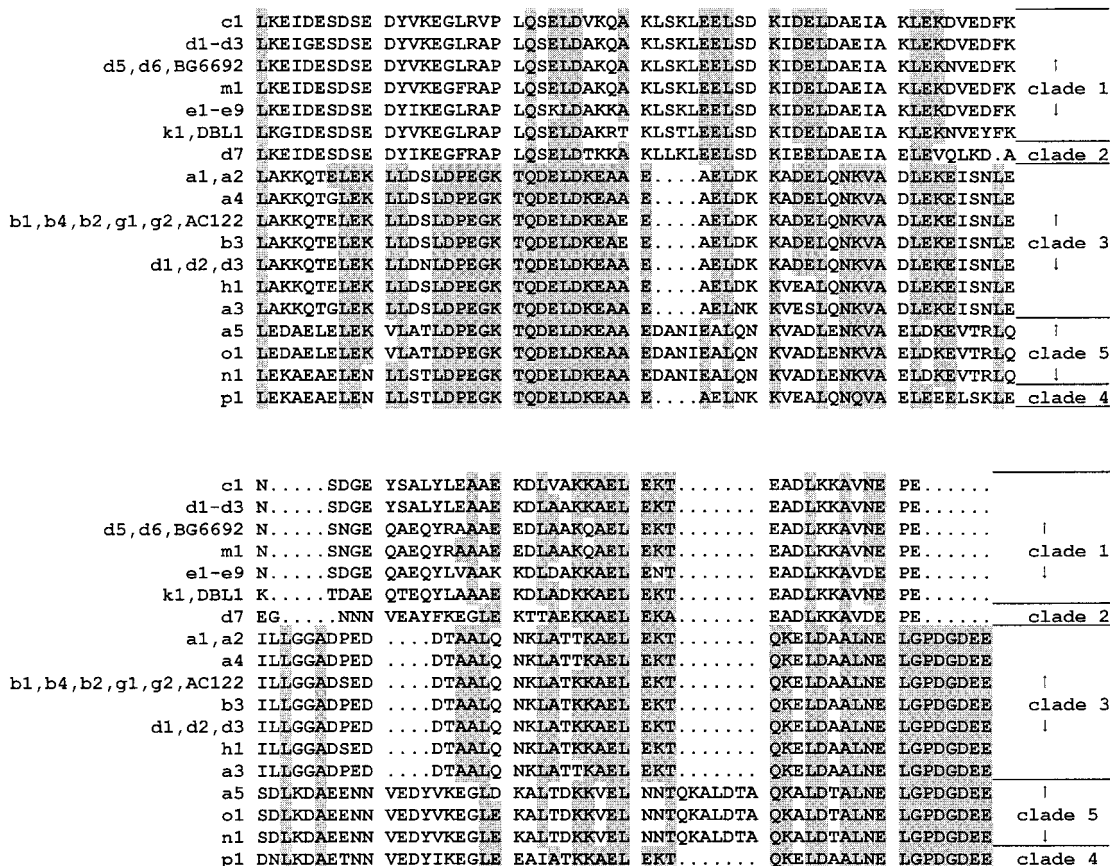


FIG. 3. Alignment of the PspA clade-defining regions that contain cross-protective epitopes from different *pspA* amplicon RFLP types. BG6692 and DBL1 refer to PspA protein sequences from reference 11. The PileUp program was used as described previously (11). Positions shared by nine or more PspA sequences are shaded.

specific *pspA* sequence type. Closely related RFLP patterns b1 and b4 both corresponded to the GenBank accession number AF252286 sequence, and similarly, RFLP patterns d1 and d2 both corresponded to the GenBank accession number AF255547 sequence. One strain (Slovakia¹⁴⁻¹⁰) contained only one conservative substitution in PspA relative to the other isolates of *pspA* RFLP type c1. The sequence differences between sets of highly related amplicons sharing the same clade-defining region were most often the result of differences in the number of repetitive sequences in the variable-length proline-rich domain. For example, the accession number AF25429 and AF255542 sequences (corresponding to RFLP types a1 and a2, respectively) differ only by the deletion of a PAPAPKPEQ direct repeat (data not shown), and this 27-bp deletion appears to account for the difference in the amplicon restriction profile (Fig. 4; compare lanes 2 and 3). Similarly, the nine sequences corresponding to RFLP types e1 to e9, respectively, shared identical clade-defining regions, and the principal difference between them was in the number of proline-rich repeats in the variable-length proline-rich region (Fig. 5).

Only two *pspA* sequence types were shared among strains that did not appear to be highly related according to the results of PFGE. The PspA sequence with GenBank accession number AF253406 (768 bases) was shared exclusively among PFGE type C clinical isolates: PFGE type C isolate England¹⁴⁻⁹, PFGE type α isolate Spain¹⁴⁻⁵, and a PFGE type B isolate. The GenBank accession number AF255546 sequence, which differed from the GenBank accession number AF253406 se-

quence by only one conservative substitution (data not shown), was also found in the distinct PFGE type β strain Slovakia¹⁴⁻¹⁰. It is interesting that all of these strains with the *pspA* sequence corresponding to GenBank accession number AF253406 (and the single strain with the GenBank accession number AF255546 sequence), although they appeared to be nonrelated according to the results of PFGE and penicillin-binding protein gene-*dhf* restriction profiling (data not shown), were all serotype 14 (Table 1). Further analysis of the degree of relatedness between these strains should be facilitated by multilocus sequence typing (7).

With the exceptions of PFGE type D isolates, the majority of isolates within a given PFGE type exhibited identity or a high degree of homology within their *pspA* sequences, which is evident in the alignment of the approximately 100-residue clade-defining region (Fig. 3). The PFGE type D isolates displayed the widest range of variability in the *pspA* clade-defining region, with isolates found to be of clade 3 (four isolates), clade 1 (four isolates), and clade 2 (two isolates). The clade 5 PspA protein from the isolate of PFGE subtype A4 is quite divergent from the clade 3 PspA proteins shared by the other 23 PFGE type A isolates (including Spain^{23F-1}). This is consistent with the fact that PFGE subtype A4 is the most divergent PFGE subtype within type A (Fig. 2). This sequence (GenBank accession number AF25426 [RFLP type a5]) differed from the sequence found for Hungary^{19A-6} (RFLP type o1) by only 2 codon substitutions over a 227-codon overlap and by only 1 conservative substitution in the clade-defining region (Fig. 3).

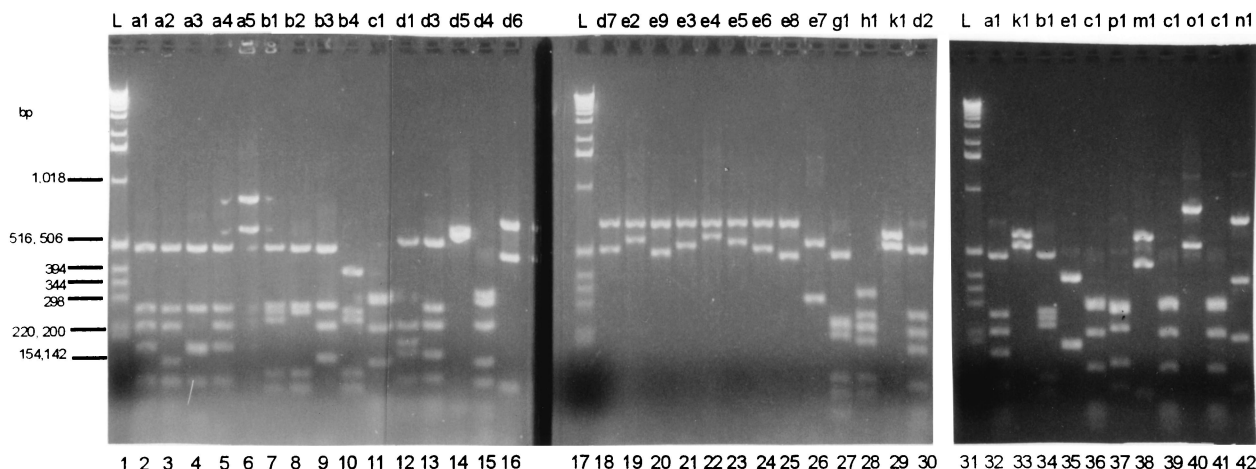


FIG. 4. RFLP patterns of *pspA* amplicons. Amplicons generated with primers F and R were subjected to triple digestion with *Hae*III, *Hin*II, and *Rsa*I prior to agarose gel electrophoresis. The RFLP type designations at the top are from Table 1. Lanes 1, 17, and 31, size ladder (L); lanes 2 to 16 and 18 to 30, 1997 multiresistant U.S. isolates; lanes 32 to 42, internationally disseminated multiresistant strains: lane 32, Spain^{23F}-1; lane 33, Spain^{6B}-2; lane 34, France^{9V}-3; lane 35, Tennessee^{23F}-4; lane 36, Spain¹⁴-5; lane 37, S. Africa^{19A}-7; lane 38, S. Africa^{6B}-8; lane 39, England¹⁴-9; lane 40, Hungary^{19A}-6; lane 41, Slovakia¹⁴-10; lane 42, Slovakia^{19A}-11.

Additionally, 1 of the 40 PFGE type B isolates had the major PFGE type C *pspA* sequence (GenBank accession number AF253406) and RFLP type c1 (Table 1).

Isolates of PFGE type D are a more broadly divergent group on the basis of the deeper branching points in the dendrogram (Fig. 2), which is consistent with PFGE type D having three different PspA clades. The deeper branching points could possibly indicate an older age of that clonal type, which would allow more time for horizontal *pspA* transfer events. However, isolates of PFGE type B have deeper branch points than isolates of the PFGE types other than D, and yet 39 of the 40 PFGE type B isolates (except for subtype B13) have very conserved *pspA* sequences.

Comparisons of approximately 160- to 230-codon regions of *pspA* consisting of the clade-defining region plus the flanking sequence that includes 20 to 50 codons on the N-terminal side and the variable-length proline-rich domain on the C-terminal side, in combination with RFLP analysis of the *pspA* amplicon, provides a potential tool for distinguishing between highly related isolates (Fig. 3 and 4; Table 1). A total of 145 multire-

sistant pneumococci were found to exhibit 34 *pspA* amplicon restriction profiles that were predictive of specific sequence types. In a similar manner, restriction profiling of *pspA* amplicons was previously found to be useful in revealing related subsets of *pspA* (A. Brooks-Walter, L. S. McDaniel, S. K. Hollingshead, M. J. Crain, and D. E. Briles, Abstr. 35th Intersci. Conf. Antimicrob. Agents Chemother., abstr. B12, p. 28, 1995) and was used to assist in demonstrating the clonal origin of serotype 9L isolates from western Canada (26). All of the isolates in the present study were found to contain family 1 or family 2 *pspA* alleles that encoded PspA proteins primarily of clade 1 and clade 3, respectively. The most prevalent PFGE types of resistant pneumococci from within the United States were found to have *pspA* amplicons and sequence types identical to those for highly related internationally dispersed multiresistant strains. PspA has recently been shown to have a mosaic structure that presumably results from extensive immunologic selection for intraspecies recombination events (11, 12). Although different serotypes and clonal backgrounds appear to vary in their degrees of virulence (14), it is unknown if or how different PspA types affect pneumococcal virulence. For epidemiologic studies and for PspA-based vaccine considerations, it may possibly prove to be important to track pneumococcal strains with regard to the sequence types of this variable protein.

ACKNOWLEDGMENTS

We thank the Georgia, California, Connecticut, Minnesota, Oregon, Maryland, New York, and Tennessee Emerging Infections Program Network/Active Bacterial Core Surveillance for isolates. We are grateful to Linda McDougal and Fred Tenover for sharing characterized pneumococcal clones and information. Zhongya Li provided excellent technical assistance. We thank Ruth Franklin and Deloise Jackson for the bulk of the serotyping results.

G.G. was a recipient of a career award from the Italian Fondazione Cariverona Progetto Sanità.

REFERENCES

1. Briles, D. E., J. D. King, M. L. Gray, L. S. McDaniel, E. Swialto, and K. A. Benton. 1996. PspA, a protection-eliciting pneumococcal protein: immunogenicity of isolated native PspA in mice. *Vaccine* 14:858-867.
2. Briles, D. E., R. C. Tart, H.-Y. Wu, B. A. Ralph, M. W. Russell, and L. S. McDaniel. 1996. Systemic and mucosal protective immunity to pneumococcal surface protein A. *Ann. N. Y. Acad. Sci.* 797:118-126.
3. Briles, D. E., S. Hollingshead, E. Swiatlo, J. P. Dillard, P. Smith, K. A.

E5, E7	DEPETPAPAP	APKPAPAPAP	TPEA.....PAPA	PKPAPAPKPA
E2	DEPETPAPAP	APKPAPAPAP	TPEA.....	PAPAPKPAPA	PKPAPAPKPA
E3	DEPETPAPAP	APKPAPAPAP	TPEA.....	PAPAPKPA
E4	DEPETPAPAP	APKPAPAPAP	TPEAPAPAPK	PAPAPKPAPA	PKPAPAPKPA
E6	DEPETPAPAP	APKPAPAPAP	TPEA.....
E8	DEPETPAPAP	APKPAPAPAP	TPEA.....
E9	DEPETPAPAP	APKPAPAPAP	TPEA.....PA
E1	DEPET.....
E5, E7	PAPKPAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E2	PAPKPAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E3	PAPKPAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E4	PAPKPAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E6PAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E8	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E9	PAPKPAPAPK	PAPAPKPAPA	PKPAPAPKPA	PAPKPAPAPK	PAPAPKPETP
E1	PAPAPKPAPA	PAPAPTP.EA	PAPAPKPETP
E5, E7	KTGWKQENGM				
E2	KTGWKQENGM				
E3	KTGWKQENGM				
E4	KTGWKQEKVC				
E6	KTGWKQENGM				
E8	KTGWKQENGM				
E9	KTGWKQEKVC				
E1	KTGWKQEKVC				

FIG. 5. Alignment of PFGE type E (subtypes E1 to E9) PspA variable-length proline-rich regions.

- Benton, B. A. Ralph, A. Brooks-Walter, M. J. Crain, S. K. Hollingshead, and L. S. McDaniel. 1997. PspA and PspC: their potential for use as pneumococcal vaccines. *Microb. Drug. Resist.* **3**:401-408.
4. Coffey, T. J., C. G. Dowson, C. G., Daniels, M., Zhou, J., Martin, C., Spratt, B. G., and J. M. Musser. 1991. Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of *Streptococcus pneumoniae*. *Mol. Microbiol.* **5**:2255-2260.
 5. Coffey, T. J., S. Berron, M. Daniels, M. E. Garcia-Leoni, E. Cercenado, E. Bouza, A. Fenoll, and B. G. Spratt. 1996. Multiply antibiotic-resistant *Streptococcus pneumoniae* recovered from Spanish hospitals (1988-1994): novel major clones of serotypes 14, 19F and 15F. *Microbiology* **142**:2747-2757.
 6. Cowan, M. J., A. J. Ammann, D. W. Wara, V. M. Howie, L. Schultz, N. Doyle, and M. Kaplan. 1978. Pneumococcal polysaccharide immunization in infants and children. *Pediatrics* **62**:721-727.
 7. Enright, M. C., and B. G. Spratt. 1998. A multilocus typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**:3049-3060.
 8. Figueiredo, A. M., R. Austrian, P. Urbaskova, L. A. Teixeira, and A. Tomasz. 1995. Novel penicillin-resistant clones of *Streptococcus pneumoniae* in the Czech Republic and in Slovakia. *Microb. Drug Resist.* **1**:71-78.
 9. Gherardi, G., C. Whitney, R. Facklam, and B. Beall. 2000. Major sets of genetically related β -lactam resistant pneumococci in the United States based upon PFGE and *pbp1a-pbp2b-pbp2x-dhf* restriction profiles. *J. Infect. Dis.* **181**:216-229.
 10. Hall, L. M., R. A. Whitley, B. Duke, R. C. George, and A. Efstratiou. 1996. Genetic relatedness within and between serotypes of *Streptococcus pneumoniae* from the United Kingdom: analysis of multilocus enzyme electrophoresis, pulsed-field gel electrophoresis, and antimicrobial resistance patterns. *J. Clin. Microbiol.* **34**:853-859.
 11. Hollingshead, S. K., R. Becker, and D. E. Briles. 2000. University of PspA: mosaic genes and evidence for past recombination in *Streptococcus pneumoniae*. *Infect. Immun.* **68**:5889-5900.
 12. Hollingshead, S. K., D. Bessen, and D. E. Briles. 1998. Archeological footprints of horizontal gene transfer: mosaic cell surface proteins of *Streptococcus pyogenes* and *Streptococcus pneumoniae*. In C. Kado and M. Syvanen (ed.), *Horizontal gene transfer: implications and consequences*, Chapman & Hall, New York, N.Y.
 13. Jernigan, D. B., M. S. Cetron, and R. F. Breiman. 1996. Defining the public health impact of drug resistant *Streptococcus pneumoniae*: report of a working group. *Morb. Mortal. Wkly. Rep.* **45**:1-20.
 14. 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.
 15. Klugman, K. 1998. Pneumococcal molecular epidemiology network. *ASM News* **64**:371.
 16. McDaniel, L. S., J. Yother, M. Vijayakumar, L. McGarry, W. R. Guild, and D. E. Briles. 1987. Use of insertional inactivation to facilitate studies of biological properties of pneumococcal surface protein A (PspA). *J. Exp. Med.* **165**:381-394.
 17. McDaniel, L. S., J. S. Sheffield, P. Delucchi, and D. E. Briles. 1991. PspA, a surface protein of *Streptococcus pneumoniae*, is capable of eliciting protection against pneumococci of more than one capsular type. *Infect. Immun.* **59**:222-228.
 18. McDaniel, L. S., B. A. Ralph, D. O. McDaniel, and D. E. Briles. 1994. Location of protection-eliciting epitopes on PspA of *Streptococcus pneumoniae* between amino acid residues 192 and 260. *Microb. Pathog.* **17**:323-337.
 19. McDaniel, L. S., D. O. McDaniel, S. K. Hollingshead, and D. E. Briles. 1998. Comparison of the PspA sequence from *Streptococcus pneumoniae* EF5668 to the previously identified PspA sequence from strain Rx1 and ability of PspA from EF5668 to elicit protection against pneumococci of different capsular types. *Infect. Immun.* **66**:4748-4754.
 20. McDougal, L. K., J. K. Rasheed, J. W. Biddle, and F. C. Tenover. 1995. Identification of multiple clones of extended-spectrum cephalosporin-resistant *Streptococcus pneumoniae* isolates in the United States. *Antimicrob. Agents Chemother.* **39**:2282-2288.
 21. Muñoz, R., T. J. Coffey, M. Daniels, C. G. Dowson, G. Laible, J. Casal, R. Hakenbeck, M. Jacobs, J. M. Musser, and B. G. Spratt. 1991. Intercontinental spread of a multiresistant clone of serotype 23F *Streptococcus pneumoniae*. *J. Infect. Dis.* **164**:302-306.
 22. Muñoz, R., J. M. Musser, M. Crain, D. E. Briles, A. Morton, A. J. Parkinson, U. Sorenson, and A. Tomasz. 1992. Geographic distribution of penicillin-resistant clones of *Streptococcus pneumoniae*: characterization by penicillin-binding protein profile, surface protein A typing, and multilocus enzyme analysis. *Clin. Infect. Dis.* **15**:112-118.
 23. Nabors, G. S., P. A. Braun, D. J. Herrmann, M. L. Heise, D. J. Pyle, S. Gravenstein, M. Schilling, L. M. Ferguson, S. K. Hollingshead, E. E. Briles, et al. 2000. Immunization of healthy adults with a single recombinant pneumococcal surface protein A (PspA) stimulated cross-reactive antibodies to heterologous PspA molecules. *Vaccine* **18**:1743-1754.
 24. Sibold, C., J. Wang, J. Henrichsen, and R. Hakenbeck. 1992. Genetic relationships of penicillin-susceptible and -resistant *Streptococcus pneumoniae* strains isolated on different continents. *Infect. Immun.* **60**:4119-4126.
 25. Smith, A. M., and K. P. Klugman. 1997. Three predominant clones identified within penicillin-resistant South African isolates of *Streptococcus pneumoniae*. *Microb. Drug Resist.* **3**:385-389.
 26. Swiatlo, E., M. J. Crain, L. S. McDaniel, A. Brooks-Walter, T. J. Coffey, B. G. Spratt, D. A. Morrison, and D. E. Briles. 1996. DNA polymorphisms and variant penicillin-binding proteins as evidence that relatively penicillin-resistant pneumococci in western Canada are clonally related. *J. Infect. Dis.* **174**:884-888.
 27. Turett, G. S., S. Blum, B. A. Fazal, J. E. Justman, and E. E. Telzak. 1999. Penicillin resistance and other predictors of mortality in pneumococcal bacteremia in a population with high human immunodeficiency virus seroprevalence. *Clin. Infect. Dis.* **29**:321-327.
 28. Yother, J., and D. E. Briles. 1992. Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis. *J. Bacteriol.* **174**:601-609.