Comparison of Results Generated by Serotyping, Pulsed-Field Restriction Analysis, Ribotyping, and Repetitive-Sequence PCR Used To Characterize Penicillin-Resistant Pneumococci from the United States

W. MICHAEL DUNNE, JR.,[1]* KAREN SUE KEHL,[2] CAROL A. HOLLAND-STALEY,[3]
ANGELA B. BRUEGGEMANN,[4] MICHAEL A. PFALLER,[4] AND GARY V. DOERN[4]

*Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri[1];
Department of Pathology, Medical College of Wisconsin, Milwaukee, Wisconsin[2]; Division of
Infectious Diseases, Henry Ford Hospital, Detroit, Michigan[3]; and Department
of Pathology, University of Iowa College of Medicine, Iowa City, Iowa[4]*

One hundred forty-seven isolates of *Streptococcus pneumoniae* with high-level penicillin resistance collected during a national surveillance program in the United States were characterized by serotyping, pulsed-field restriction analysis, ribotyping, and repetitive-sequence (BOX element) PCR. The results generated by each method were compared by frequency of association to examine whether relationships existed between the various typing methods and statistically to determine association with the geographic source of the isolate or the age of the patient from whom the isolate was obtained. When the data were examined by pairwise analysis of individual strain classifications produced by each typing method, no statistically significant relationships between strain type, geographic location, or patient age were identified, suggesting that distinct clones of penicillin-resistant *S. pneumoniae* have been widely distributed throughout the United States. However, we did observed shared expression of two or three typing markers at a high frequency (>50%) among clusters of strains, indicating a certain level of concordance between the various typing methods used to classify penicillin-resistant *S. pneumoniae*.

Thirty years ago, the recovery of penicillin-resistant *Streptococcus pneumoniae* (PRSP) strains from patients residing in the United States was considered an anomaly. Over the past decade, however, there has been a dramatic rise in the rate of isolation of PRSP. At present, it is estimated that nearly one-third of all clinical isolates of *S. pneumoniae* from the United States demonstrate intermediate or high levels of resistance to penicillin (3). In essence, the isolation of penicillin-susceptible phenotypes of pneumococci in this country is rapidly approaching the point at which it will be the exception rather than the rule.

In a recent study, Doern and colleagues (4) found that 103 (70.1%) of 147 PRSP isolates with high-level resistance collected from 30 surveillance centers across the United States during 1994 and 1995 clustered within 9 of 38 possible pulsed-field gel electrophoresis (PFGE) types. Four of the PFGE types accounted for 76 (51.7%) of the 147 PRSP isolates, while only six serotypes were represented among the entire collection. Collectively, these results suggest that the majority of PRSP isolates in the United States are represented by a relatively limited number of clonal groups. However, that study found no relationship between PFGE type, patient age, and the geographic location of the isolate.

During the course of the discussion, the authors raised a number of provocative questions regarding the molecular typing of PRSP, not the least of which emphasizes that the use of a single typing method to establish the genetic relatedness of PRSP strains is completely arbitrary. Indeed, a variety of genotypic and phenotypic markers have been employed in previous studies to examine the epidemiological relationship of pneumococcal isolates (1, 4, 5, 8, 12, 14, 15). Clearly, the discriminatory power of any typing method (the ability to accurately distinguish unrelated strains) dictates the number of PRSP clones recognized. The inherent danger with the availability of multiple typing methods for the epidemiological investigation of PRSP (or any other organism for that matter) is that strain relatedness established by the use of one method might be missed by another. Further, because the art of molecular epidemiology is still in its infancy, the overall correlation of different typing methods for any given organism is generally unknown. In this analysis, we examine results obtained by using several typing methods when applied to the epidemiological investigation of PRSP. These methods included serotyping, pulsed-field restriction analysis (PFGE), ribotyping, and repetitive-sequence PCR (Rep-PCR). The discriminatory value of each of these methods was evaluated, as were relationships between cluster assignments.

MATERIALS AND METHODS

Bacterial strains. One hundred forty-seven isolates of PRSP were included in the evaluation. This collection was characterized previously (4) and included all of the strains with high-level penicillin resistance (MIC, ≥2 µg/ml) that had been identified during the course of a national surveillance program in the United States conducted from 1994 to 1995. Isolates were maintained at −70°C on porous beads prior to testing. The patient population from whom the isolates
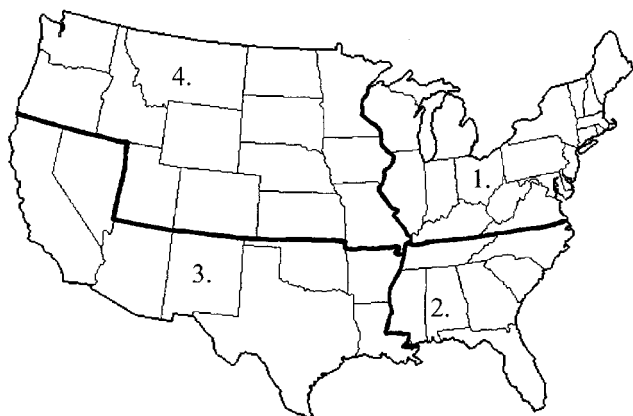
FIG. 1. Regional definitions used for the analysis of clonal relationships between geographic location and epitypes of 147 strains of PRSP. Isolate locations were divided into region 1 (northeast, $n = 55$), region 2 (southeast, $n = 27$), region 3 (southwest, $n = 30$), and region 4 (northwest, $n = 35$).

were obtained had a mean age of 30.1 years (range, 2 months to 94 years old). The isolates were cultured from patients with a variety of community-acquired pneumococcal infections, including acute exacerbation of chronic bronchitis, pneumonia, otitis media, sinusitis, meningitis, and bacteremia.

Typing methods. Determination of capsular serotype was accomplished by slide agglutination using antisera obtained from the Statens Seruminstitut, Copenhagen, Denmark. PFGE typing was performed as previously described (4) using the method of LeFevre et al. (8). Briefly, *Sma*I restriction fragments of chromosomal DNA were separated on 1% agarose gels with a CHEF-DR II instrument (Bio-Rad, Richmond, Calif.) and stained with ethidium bromide. Strains were determined to be of the same clone when no more than a three-band difference was observed between individual PFGE profiles determined by visual inspection (13). Rep-PCR of PRSP was performed according to the method of van Belkum et al. (14) using primers directed against the BOX repeat element. The amplified products were separated on 1% agarose gels, stained with ethidium bromide, and then scanned and digitalized using the Eagle Eye II Still Video (Stratagene, La Jolla, Calif.), and a phylogenetic tree was constructed using GelCompar software (Applied Maths, Kortrijk, Belgium). A similarity matrix was generated using the band-based Dice similarity coefficient (2) from stored gel patterns. Banding patterns were compared using the unweighted pair group method with average linkages as previously described (14). An 85% homology cutoff level was selected to cluster strains into homology groups. Ribotyping was performed as previously described using the RiboPrinter (7). Briefly, this system automates the ribotyping process by orchestrating bacterial lysis, restriction enzyme digestion of bacterial DNA using *Eco*RI, separation of fragments by gel electrophoresis, and Southern blot hybridization. To examine the relationship between typing results and geographic location, strains were classified by site of origin into four regional groups (Fig. 1) with the following distribution: region 1 (northeast), $n = 55$; region 2 (southeast), $n = 27$; region 3 (southwest), $n = 30$; and region 4 (northwest), $n = 35$.

The term epitype is used here to refer to the classification of a strain into a homology group (single strain or cluster) based on the results of any of the typing methods. For example, PFGE type 1 and ribotype 22 are both considered epitypes.

Data analysis. Using the SPSS statistical program, a pairwise analysis of the database parameters, consisting of the Rep-PCR type, PFGE type, ribotype, serotype, patient age, and location for all 147 PRSP isolates, was performed using the Pearson chi-square test.

RESULTS

PFGE. As described previously, PFGE identified a total of 38 distinct profiles for the 147 PRSP strains. Of these, however, most strains ($n = 103$; 70.1%) were assigned to one of nine PFGE types (Table 1). Further, 51.7% of the total were grouped into one of four PFGE types (types 1, 2, 10, and 11).

Serotype. Serotyping established the identification of six different capsular types within the collection of PRSP isolates (Table 1). Of these, the majority (55.8%) of strains were either serotype 23F ($n = 48$) or 6B ($n = 34$). The remaining strains were divided among serotypes 19F, 9A, 14, and 6A, while 14 of the isolates were nontypeable.

Ribotype. Ribotyping of the PRSP isolates produced 31 distinct patterns. Nearly 81% of the strains ($n = 119$) belonged to one of eight ribotypes (Table 1), but ribotypes 22 and 1 collectively accounted for 38% of the 147 isolates.

Rep-PCR. Rep-PCR typing of the PRSP isolates using the BOX repeat element produced 51 distinct fingerprints which could be consolidated into 33 Rep-PCR types at the 85% homology level. Of those 33 types, 8 (Rep-PCR types 1, 3, 4, 8, 10, 11, 22, and 24) represented 69% of the total number of isolates.

Location. Of the total collection of PRSP isolates, 55 isolates originated from medical centers located in region 1 (northeast), while 27, 30, and 35 isolates were recovered from patients treated in regions 2 (southeast), 3 (southwest), and 4 (northwest), respectively (Fig. 1).

Relationship between epitypes, location, and age. There were no statistically significant associations identified between

TABLE 1. Most frequently identified serotypes, PFGE types, ribotypes, and Rep-PCR types among 147 strains of PRSP

Classification	Type	No. of isolates
Serotype	23F	48
	6B	34
	19F	20
	9A	15
	14	14
	6A	2
	NT[a]	14
PFGE	1	33
	11	20
	2	13
	10	10
	3	7
	5	6
	6	5
	7	5
	13	4
Ribotype	22	36
	1	20
	30	13
	4	12
	10	12
	12	10
	8	8
	9	8
	14	3
	19	3
Rep-PCR	3	31
	1	16
	4	16
	11	12
	24	9
	8	7
	10	5
	22	5
	5	4
	12	4
	19	4
	20	4

[a] NT, nontypeable.

TABLE 2. Percentages of strains within individual PFGE types in which a second epitype was shared at a frequency of ≥50%[a]

PFGE type (n)	Shared epitype (n)	% of strains
1 (33)	Ribotype 22 (29)	87.8
1 (33)	Serotype 23F (28)	84.8
2 (12)	Ribotype 30 (11)	91.7
2 (12)	Serotype 23F (11)	91.7
3 (7)	Ribotype 12 (7)	100
3 (7)	Serotype 19F (7)	100
5 (6)	Ribotype 4 (4)	66.7
5 (6)	Serotype 6B (6)	100
6 (5)	Ribotype 1 (5)	100
6 (5)	Serotype 6B (5)	100
7 (5)	Ribotype 1 (4)	80
7 (5)	Serotype 6B (4)	80
10 (10)	Ribotype 8 (7)	70
10 (10)	Nontypeable (10)	100
11 (20)	Serotype 9A (14)	70

[a] Only PFGE types having five or more strains were considered.

any one epitype and patient age and geographic location. When the relationships between epitypes were examined statistically, the P values for most comparisons were significant, primarily due to the sparseness of the raw data for certain epitypes and because the null hypothesis presumed the unrelatedness of epitypes. To circumvent this, we examined the frequency with which strains of PRSP shared epitypes by examining (i) only those epitypes consisting of five or more strains and (ii) those markers shared by >50% of the strains within an individual epitype.

Association of PFGE type with other epitypes. Since PFGE is considered by some to represent the "gold standard" of molecular typing methods, we first examined the coexpression of different epitypes sorted by PFGE types (Table 2). Of particular interest, 87.8% of PFGE type 1 strains were identified as ribotype 22, and 84.8% of PFGE type 1 strains expressed serotype 23F capsular antigen. Overall, 75.8% of all PFGE type 1 strains were both ribotype 22 and serotype 23F. Other shared markers among PFGE types include PFGE type 2, ribotype 30, and serotype 23F (11 of 12 strains); PFGE type 3, serotype 19, and ribotype 12 (7 of 7 strains); and PFGE type 6, serotype 6B, and ribotype 1 (5 of 5 strains). Interestingly, all PFGE type 10 strains were nontypeable, and no PFGE types

TABLE 3. Percentages of strains within individual ribotypes in which a second epitype was shared at a frequency of ≥50%[a]

Ribotype (n)	Shared epitype (n)	% of strains
1 (20)	Serotype 6B (19)	95
4 (12)	Serotype 6B (11)	91.7
8 (8)	PFGE 10 (7)	87.5
8 (8)	Nontypeable (8)	100
9 (8)	PFGE 11 (8)	100
9 (8)	Serotype 9A (6)	75
9 (8)	Rep-PCR 4 (5)	62.5
10 (12)	PFGE 11 (9)	81.8
10 (12)	Serotype 9A (6)	50
10 (12)	Rep-PCR 3 (6)	50
12 (10)	Serotype 19F (10)	100
12 (10)	PFGE 3	70

[a] Only ribotypes consisting of five or more strains were considered.

TABLE 4. Percentages of strains within individual serotypes in which a second epitype was shared at a frequency of ≥50%[a]

Serotype (n)	Shared epitype (n)	% of strains
6B (34)	Ribotype 19 (19)	55.9
9A (15)	PFGE 11 (14)	93.3
9A (15)	Rep-PCR 3 (8)	53.3
19F (20)	Ribotype 12 (10)	50
23F (48)	Ribotype 22 (27)	56.3
23F (48)	PFGE 1 (28)	58.3
Nontypeable (14)	PFGE 10 (10)	71.4
Nontypeable (14)	Ribotype 8 (8)	57.1

[a] Only serotypes consisting of five or more strains were considered.

shared a Rep-PCR type at a frequency exceeding 50% of the strains.

Association of ribotype with other epitypes. When PSRP isolates were sorted by ribotype, the following associations were noted. Ninety-five and 91.7% of ribotypes 1 and 4 expressed serotype 6B antigen, respectively, while all ribotype 12 strains were serotype 19F (Table 3). None of the ribotype 8 strains produced detectable capsular antigen, while all ribotype 9 strains were concomitantly classified as PFGE type 11. Only two associations were noted between ribotype and Rep-PCR type at a frequency at or exceeding 50% of strains: ribotype 9 and Rep-PCR type 4 (62.5%) and ribotype 10 and Rep-PCR type 3 (50%).

Association of serotype with other epitypes. While several epitypes were observed in association with particular serotypes at or above the 50% threshold, only the combination of serotype 9A and PFGE type 11 exceeded 90% of strains (Table 4).

Association of Rep-PCR type with other epitypes. Thirteen of 16 Rep-PCR type 1 strains (81.3%) also produced 23F capsular antigen, while 6 of 7 Rep-PCR type 8 strains (85.7%) were also identified as PFGE type 2 strains (Table 5). A variety of other associations were identified, none of which exceeded 80% of the Rep-PCR type strains. Although not representing more than 50% of any one epitype, seven strains in the collection shared the combination of Rep-PCR type 1, ribotype 22, serotype 23F, and PFGE type 1.

DISCUSSION

It is clear from this investigation that the use of different typing methods for the classification of PRSP can generate a

TABLE 5. Percentages of strains within individual Rep-PCR types in which a second epitype was shared at a frequency of >50%[a]

Rep-PCR type (n)	Shared epitype (n)	% of strains
1 (16)	PFGE 1 (10)	62.5
1 (16)	Serotype 23F (13)	81.3
1 (16)	Ribotype 22 (10)	62.5
4 (16)	PFGE 11 (8)	50
8 (7)	PFGE 2 (6)	85.7
8 (7)	Serotype 23F (5)	71.4
8 (7)	Ribotype 30 (4)	57.1
10 (5)	Ribotype 22 (3)	60
11 (12)	Serotype 6B (9)	75
24 (9)	Ribotype 1 (6)	75
24 (9)	Serotype 6B (7)	77.8

[a] Only Rep-PCR types consisting of five or more strains were considered.

variety of epidemiological profiles. This is true not only from the perspective of the varying discriminatory power of each method but also for the ability of any individual method to cluster strains into distinct homology groups or epitypes. It should come as no surprise that the correlation among typing methods in this study was highly variable. Each method exploits a different target for the purpose of strain differentiation. Ribotyping, PFGE, and Rep-PCR are dependent upon the genomic distribution of ribosomal gene sequences, restriction enzyme cleavage sites, and palindromic repeat sequences, respectively, while serotyping requires that all of the genetic, enzymatic, and transport components necessary to produce an exopolysaccharide are expressed and functional.

We know from a previous study reported by van Belkum et al. (14) that 28 isolates of *S. pneumoniae* could be differentiated into as few as 7 epitypes by ribotyping or as many as 21 by Rep-PCR using the BOX repeat element as the primer target. Similar differences in the discriminatory powers of various phenotypic and genotypic markers have been also been reported for *Shigella sonnei, Burkholderia cepacia,* and *Serratia marcescens* by Peter Liu and colleagues (9–11). However, the discriminatory index or resolving power of any typing method is somewhat arbitrary and can be adjusted by the investigator to meet certain criteria. This phenomenon has been observed by Versalovic et al. (15), who examined the relationship of multilocus enzyme electrophoresis (MLEE), serotyping, and Rep-PCR (using a different repetitive sequence target) for typing 46 PRSP strains isolated in the greater metropolitan Houston, Texas, area. Each method had a different discriminatory index, with MLEE providing the greatest ability to distinguish strains ($n = 31$), followed by Rep-PCR ($n = 16$) and serotyping ($n = 4$). At first glance, and using the greatest discriminatory power of MLEE, there appeared to be one major association between typing methods, which consisted of MLEE type 1, serotype 6, and Rep-PCR type A and included 20 strains or 43.4% of the total. However, when the MLEE types were compiled into clusters of lineages at a genetic distance of 0.35 (65% relatedness), seven distinct lineages became apparent. Based on this new definition, a single supertype of 39 isolates (84.8%) represented by strains in the MLEE lineage A, serotype 6, and Rep-PCR type A was identified.

The interesting aspect of this investigation is not that there are differences in the clustering of PRSP strains by four independent typing methods but that certain epitypes seem to associate despite the unique nature of their determinants. Most notable is the coexpression of multiple epitypes at a high frequency by clusters of PRSP strains, including (i) ribotype 22 and serotype 23F by 25 of 33 (75.8%) PFGE type 1 strains, (ii) ribotype 30 and serotype 23F by 11 of 12 (92%) PFGE type 2 strains, (iii) serotype 19F and ribotype 12 by 7 of 7 PFGE type 3 strains, (iv) serotype 6B and ribotype 1 by 5 of 5 PFGE type 6 strains, and (v) PFGE type 10 and nontypeable serotype by 7 of 8 (87.5%) ribotype 8 strains. All told, these clusters represent nearly 40% of all PRSP isolates in this collection and can add considerable discriminatory power to the definition of clonal groups when examining the regional distribution of PRSP in the future. It is also likely that additional high-frequency combinations will be discovered when greater numbers of PRSP strains are examined. The infrequency of associations of either PFGE type, ribotype, or serotype and the groupings

established by Rep-PCR typing likely reflects the discrimination index of the latter. In this study, for example, 33 Rep-PCR types were identified using an 85% homology level as the cutoff for relatedness. Had a 70% homology level been selected, 17 Rep-PCR types would have been identified, and 14 high-frequency associations (with PFGE types 1 and 2, 10, and 11; ribotypes 8, 9, 10, 12, 22, and 30; and serotypes 14, 23F, 9A, and nontypeable) would have been recognized. The discriminatory index of a typing method could therefore be focused to a level that provides a more compatible picture. Conversely, the high discriminatory index of the Rep-PCR method might reflect the frequent exchange or relocation of the BOX repeat element in the pneumococcal genome relative to changes or deletions of other targets for typing, thus providing a useful technique to measure short-term events, rather than long-term epidemiological events as would be the case with ribotyping (6). Rep-PCR typing, therefore, could be used to subtype strains within a particular cluster or clonal group. For example, the cluster of 25 strains defined by the combination of PFGE type 1, ribotype 22, and serotype 23F could be further subdivided into eight distinct Rep-PCR types if one was attempting the establish the clonality of an outbreak of PRSP caused by this epitype.

While statistical analysis of the relationship between epitypes has proven problematic, the lack of a statistical correlation between age, location, and any single epitype in this study is interesting and might reflect the thorough distribution of PRSP clones across the United States over time. However, ongoing chronological studies in regions where the age-associated clonality of multidrug-resistant, invasive pneumococci has already been established (12) would be required to draw firm conclusions regarding the dilution of PRSP epitypes.

REFERENCES

1. **Coffey, T. J., M. Daniels, L. K. McDougal, C. G. Dowson, F. C. Tenover, and B. G. Spratt.** 1995. Genetic analysis of clinical isolates of *Streptococcus pneumoniae* with high-level resistance to expanded-spectrum cephalosporins. Antimicrob. Agents Chemother. **39:**1306–1313.
2. **Dice, L. R.** 1945. Measures of the amount of ecological association between species. Ecology **26:**297–302.
3. **Doern, G. V., M. A. Pfaller, K. Kugler, J. Freeman, and R. N. Jones.** 1998. Prevalence of antimicrobial resistance among respiratory tract isolates of *Streptococcus pneumoniae* in North America: 1997 results from the SENTRY Antimicrobial Surveillance Program. Clin. Infect. Dis. **27:**764–770.
4. **Doern, G. V., A. B. Brueggemann, M. Blocker, M. Dunne, H. P. Holley, Jr., K. S. Kehl, J. Duval, K. Kugler, S. Putnam, A. Rauch, and M. A. Pfaller.** 1998. Clonal relationships among high-level penicillin-resistant *Streptococcus pneumoniae* in the United States. Clin. Infect. Dis. **27:**757–761.
5. **Hall, L. M. C.** 1998. Application of molecular typing to the epidemiology of *Streptococcus pneumoniae.* J. Clin. Pathol. **51:**270–274.
6. **Hermans, P. W. M., M. Sluijter, T. Hoogenboezem, H. Heersma, A. van Belkum, and R. de Groot.** 1995. Comparative study of five different DNA fingerprint techniques for molecular typing of *Streptococcus pneumoniae* strains. J. Clin. Microbiol. **33:**1606–1612.
7. **Hollis, R. J., J. L. Bruce, S. J. Fritschel, and M. A. Pfaller.** 1999. Comparative evaluation of an automated ribotyping instrument versus pulsed-field gel electrophoresis for epidemiological investigation of clinical isolates of bacteria. Diagn. Microbiol. Infect. Dis. **34:**263–268.
8. **LeFevre, J. C., G. Gaucon, A. M. Sicard, and A. M. Gasc.** 1993. DNA fingerprinting of *Streptococcus pneumoniae* strains by pulsed-field gel electrophoresis. J. Clin. Microbiol. **31:**2724–2728.
9. **Liu, P. Y.-F., Y.-J. Lau, B.-S. Hu, J.-M. Shir, M.-H. Cheung, Z.-Y. Shi, and W.-S. Tsai.** 1994. Use of PCR to study epidemiology of *Serratia marcescens*

isolates in nosocomial infection. J. Clin. Microbiol. **32:**1935–1938.

10. **Liu, P. Y.-F., Y.-J. Lau, B.-S. Hu, J.-M. Shyr, Z.-Y. Shi, W.-S. Tsai, Y.-H. Lin, and C.-Y. Tseng.** 1995. Analysis of clonal relationships among isolates of *Shigella sonnei* by different molecular typing methods. J. Clin. Microbiol. **33:**1779–1783.

11. **Liu, P. Y.-F., Z.-Y. Shi, Y.-J. Lau, B.-S. Hu, J.-M. Shyr, W.-S. Tsai, Y.-H. Lin, and C.-Y. Tseng.** 1995. Comparison of different PCR approaches for characterization of *Burkholderia (Pseudomonas) cepacia* isolates. J. Clin. Microbiol. **33:**3304–3307.

12. **Rudolph, K. M., M. J. Crain, A. J. Parkinson, and M. C. Roberts.** 1999. Characterization of a multidrug-resistant clone of invasive *Streptococcus pneumoniae* serotype 6B in Alaska using pulsed-field gel electrophoresis and PspA serotyping. J. Infect. Dis. **180:**1577–1583.

13. **Tenover, F. C., R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. A. Murray, D. H. Persing, and B. Swaminathan.** 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J. Clin. Microbiol. **33:**2233–2239.

14. **van Belkum, A., M. Sluijter, R. de Groot, H. Verbrugh, and P. W. M. Hermans.** 1996. Novel BOX repeat PCR assay for high-resolution typing of *Streptococcus pneumoniae* strains. J. Clin. Microbiol. **34:**1176–1179.

15. **Versalovic, J., V. Kapur, E. O. Mason, Jr., U. Shah, T. Koeuth, J. R. Lupski, and J. M. Musser.** 1993. Penicillin-resistant Streptococcus pneumoniae strains recovered in Houston: identification and molecular characterization of multiple clones. J. Infect. Dis. **167:**850–856.