Clonal Analysis of Colonizing Group B *Streptococcus*, Serotype IV, an Emerging Pathogen in the United States

Michelle J. Diedrick, Aurea E. Flores, Sharon L. Hillier, Roberta Creti, and Patricia Ferrieri

Departments of Laboratory Medicine and Pathology and Pediatrics, University of Minnesota Medical School, Minneapolis, Minnesota; Department of Obstetrics, Gynecology, and Reproductive Sciences, Magee Women’s Hospital and University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; and Dipartimento di Malattie Infettive, Parasitarie ed Immunomediate, Istituto Superiore di Sanità, Rome, Italy

Received 10 February 2010/Returned for modification 30 March 2010/Accepted 24 June 2010

Colonizing group B *Streptococcus* (GBS) capsular polysaccharide (CPS) type IV isolates were recovered from vaginal and rectal samples obtained from 97 (8.4%) nonpregnant women of 1,160 women enrolled in a U.S. multicenter GBS vaccine study from 2004 to 2008. Since this rate was much higher than the rate of prevalence of 0.4 to 0.6% that we found in previous studies, the isolates were analyzed by using surface protein profile identification, pulsed-field gel electrophoresis (PFGE), and multilocus sequence typing (MLST) to characterize them and identify trends in DNA clonality and divergence. Of the 101 type IV isolates studied, 53 expressed α and group B protective surface (BPS) proteins, 27 expressed BPS only, 20 expressed α only, and 1 had no detectable surface proteins. The isolates spanned three PFGE macrorestriction profile groups, groups 37, 38, and 39, of which group 37 was predominant. The isolates in group 37 expressed the α and BPS proteins, while those in groups 38 and 39 expressed the α protein only, with two exceptions. MLST studies of selective isolates from the four protein profile groups showed that isolates expressing α,BPS or BPS only were of a new sequence type, sequence type 452, while those expressing α only or no proteins were mainly of a new sequence type, sequence type 459. Overall, our study revealed a limited diversity in surface proteins, MLST types, and DNA macrorestriction profiles for type IV GBS. There appeared to be an association between the MLST types and protein expression profiles. The increased prevalence of type IV GBS colonization suggested the possibility that this serotype may emerge as a GBS pathogen.

Group B *Streptococcus* (GBS) (*Streptococcus agalactiae*) is a leading cause of neonatal infection in the United States, with maternal vaginal or rectal colonization often resulting in the transmission of GBS to the infant during the perinatal period (8, 23). GBS isolates are classified according to nine capsular polysaccharide (CPS) types: types Ia, Ib, and II to VIII and the recently proposed type IX (9, 15, 21, 23, 46, 52). Isolates that do not express any of the known CPS types are designated non-typable (NT) (2, 6, 21, 40). In addition to CPS, GBS may express one or more surface-localized proteins, including the α and β components of the c protein (24); the alpha-like R proteins, specifically R1, R4(Rib), and R1,R4 (also known as Alp3) (14, 17, 19, 30, 40); and the group B protective surface (BPS) protein (12). Certain protein profiles are associated with each capsular polysaccharide CPS type (2), for example, the c(α only) protein with types Ia and II, c(α + β) with type Ib, and R4(Rib) with type III (2, 14). BPS, expressed by fewer than 3% of colonizing isolates, can be found alone or with another protein in type Ia, II, and V isolates (12, 14).

In the United States, the predominant serotypes over the past 2 decades, constituting 70 to 75% of all GBS isolates, have been type Ia, type III, and the more recently emerged type V (14, 15, 20, 52). The remaining isolates consisted primarily of types Ib and II, with types IV, VI, VII, and VIII making up a small fraction of the isolates. We found type IV to represent between 0.4 and 0.6% of colonizing GBS isolates (14, 15), but only rare type IV isolates were found in invasive GBS disease during that same time period (14, 43, 52).

In contrast to the previously low percentage of type IV isolates reported for the United States, recent studies in the United Arab Emirates, Turkey, and Zimbabwe showed large proportions of type IV isolates among their GBS isolates. In the United Arab Emirates, type IV was the predominant serotype among colonized pregnant women, representing 26.3% of the GBS isolates (1). In eastern Turkey, it was the second most common serotype, at 8.3%, among colonizing isolates (10), and in Zimbabwe, it was the fourth most common serotype, comprising 5.1% of GBS isolates from colonized pregnant women and 4.0% of all GBS isolates from various sites, including blood and cerebrospinal fluid (CSF), from hospitalized patients (36).

Immunization studies of humans (3, 28) and protection studies with mice (37) have shown the potential of vaccines against the common GBS serotypes to prevent invasive neonatal GBS disease through the vaccination of pregnant women (3, 28). The GBS strains described here are from a phase II randomized, double-blinded clinical trial of a GBS serotype III-tetanus toxoid (CPS III-TT) vaccine to prevent the vaginal acquisition of GBS type III in nonpregnant women in three areas of the United States: Pittsburgh (PA), Georgia, and Texas (S. Hillier, unpublished data). Because we found type IV isolates for almost 10% of these patients, we examined the type IV isolates for surface proteins and clonality.

Pulsed-field gel electrophoresis (PFGE) was used in this
analysis because it is a widely used method that can further characterize GBS isolates within particular CPS type and/or protein profile groups (2, 4, 6, 48). Multilocus sequence typing (MLST) was performed in order to assess the general relatedness of strains within and across laboratories (25, 50). Together, the discriminatory power of PFGE and the objectivity of MLST gave insight into the GBS type IV population genetic structure and the identification of emerging clones (2, 5, 13, 18, 19).

MATERIALS AND METHODS

Bacterial isolates. GBS isolates from vaginal and rectal cultures of 1,160 nonpregnant women in the vaccine trial were received at the University of Minnesota GBS Molecular Reference Laboratory from January 2004 to May 2008. Of the 1,160 patients, 797 visited the clinics for a screening visit only. The remaining 363 patients made between 1 and 11 visits over a period of 18 months. Vaginal and rectal samples were taken and cultured. GBS isolates were isolated, extracted, and serotyped for each visit, with three GBS colonies picked per positive culture.

Storage and growth. GBS isolates were subcultured onto 5% sheep blood agar plates (BAPs) and grown at 37°C overnight in Todd-Hewitt broth (BD, Sparks, MD) for serotyping or for 48 h on BAPs for PFGE analysis (5). The isolates were stored at −30°C in Todd-Hewitt broth with 2% defibrinated sheep blood (14, 17, 40).

Serotyping and detection of surface proteins. The capsular polysaccharide type and protein profile were determined for each isolate by using HCl extraction and Ouchterlony double immunodiffusion with rabbit antiserum monospecific for CPS type or surface protein as we have previously described (14, 17, 24, 40). Proteins under study in our laboratory included c(α), c(β), R1, R4(Rib), R1, R4 (known as Alp3), and BPS (47). HCl extracts were concentrated 3-fold and retested when a reaction was weak or inconclusive. Prototype strains for the surface proteins were the ones used in our previous studies; strain 3139 (S. Bergner-Rabinowitz, Israel), which did not express any of these proteins, was the CPS type IV prototype (14, 17, 24).

DNA preparation. At least one isolate per patient colonized with GBS type IV was studied by PFGE. DNA samples were prepared by using a rapid PFGE procedure described previously by our laboratory (5) and digested with the low-frequency restriction enzyme SmaI. The DNA fragments were resolved by using PFGE. Included in each gel was a lambda molecular size standard and isolate 89-022 (l(h+β)), the internal standard described previously by our laboratory (2, 5, 6).

DNA macrorestriction analysis. After staining with ethidium bromide, gels were photographed under UV light with a Kodak DC290 digital camera. Kodak ID digital image analysis software was used to determine the molecular sizes of the bands. The band patterns were then compared to numbered prototypes by using our previously described method (2, 6), adapted from the one described previously by Tenover et al. (48). An isolate with a band pattern identical to that of a prototype was given the same number designation, and those with one to two, three, four, and five band differences were considered subtypes and were given, in addition, the designations a, b, c, and d, respectively. A band difference was counted as the presence or absence of a band (6). If an isolate displayed six or more band differences, it was considered unrelated to the prototype (2, 6). The first isolate with a new PFGE profile was designated the prototype to which all subsequent isolates were compared.

MLST. Representative isolates studied by MLST were selected from the major surface protein profiles and PFGE macrorestriction profile groups found among these isolates. Isolates were chosen within each group to represent a wide time period. Extraction of DNA and partial sequencing of seven housekeeping genes were carried out as previously described (25). Alleles were given number classifications by comparing them to existing sequences in the online MLST database (http://pubmlst.org/sagalactiae/). The determination of sequence types (STs) based on the unique set of alleles for each isolate and examination of relationships between STs were also accomplished by using the database.

RESULTS

Of the 1,160 patients in the multicenter vaccine study, 97 (8.4%) were colonized at least once with a type IV isolate. For 93 of the 97 patients, only one isolate per patient was included in this analysis because the protein profiles and PFGE profiles were the same for each culture site and visit. The remaining four patients were counted twice, one time for each unique protein profile or PFGE profile identified among their type IV isolates. Of the 101 type IV isolates included in the study, 63 were from vaginal cultures and 38 were from rectal cultures; 50 were from patients with screening visits only and 51 were from patients who had 1 to 10 GBS-positive visits. For these 51 patients, the average number of culture-positive visits was 5, and 17 of the 51 patients (33%) had seven or more culture-positive visits. The majority of the patients were from Pittsburgh (n = 78; 80.4%), 16 (16.5%) were from Georgia, and 3 (3.1%) were from Texas. These numbers reflected the distribution of patients with all GBS isolates in the vaccine study: 857 (73.9%) from Pittsburgh, 243 (20.9%) from Georgia, and 60 (5.2%) from Texas (data not shown).

When digested by SmaI, the type IV isolates in this study gave clear, consistent band patterns, each comprised of 7 to 10 bands. All type IV isolates differed by at least six bands compared to our existing PFGE prototypes 1 to 36, so a new PFGE profile group, group 37, was established, followed by profile groups 38 and 39. PFGE profile group 37 was the predominant profile among type IV isolates in each geographical region (range of 62.5% to 84.0%), with Pittsburgh representing the region with the highest frequency. PFGE group 39 was the next predominant profile (range of 11.1% to 37.5%), with clinic sites in Georgia having the highest frequency (data not shown).

The examples of PFGE profiles from type IV isolates in Fig. 1 illustrated the wide diversity of profiles seen among IV/α only isolates in comparison to profiles of IV/α,BPS or IV/BPS isolates. In Fig. 1A, all of the IV/α,BPS and IV/BPS isolates exhibiting PFGE profile 37 were from different patients. Of the IV/α,BPS and IV/BPS isolates, 89% displayed a PFGE profile identical to the one shown here. In Fig. 1B, the IV/α only isolates, from seven different patients, displayed five unique band patterns. A variety of band patterns within PFGE profile groups 38 and 39 was typical of the IV/α only isolates studied. Studied by MLST, the isolates with BPS only or α,BPS in lanes 2, 4, 6, and 7 of Fig. 1A were of ST 452, while isolates expressing α only in lanes 3 and 6 to 9 of Fig. 1B were of ST 459.

The distribution of the 101 type IV isolates in this study by protein profile, PFGE profile, and sequence type is presented in Table 1. Four different surface protein profiles were observed. Of all isolates, 53 (52.5%) expressed both the α and BPS proteins. The second most common surface protein profile was BPS only (27 isolates; 26.7%), followed by α only (20 isolates; 19.8%). No proteins were detected for one isolate (1.0%). The high prevalence of the BPS protein in the type IV isolates is worth noting, since this protein was less common among isolates of other serotypes (data not shown).

As shown in Table 1, these isolates were distributed into two sequence types and three PFGE groups. PFGE profile group 37, the most common macrorestriction group, contained 81 (80.2%) of the 101 isolates; 4 isolates (4.0%) were in group 38, and 16 isolates (15.8%) were in group 39. The 16 isolates studied by MLST were found to be either ST 452 or 459 and were cleanly divided into the two groups according to their PFGE profile groups, with all 10 isolates in group 37 corresponding to ST 452 and all six isolates in groups 38 and 39.
exhibiting ST 459. Neither ST had been described before this study. ST 452 was found to have the allelic profile 5-25-4-3-2-1 exhibiting ST 459. Neither ST had been described before this study. ST 452 was found to have the allelic profile 5-25-4-3-2-1-3, a single-locus variant (SLV) of STs 24 and 468, the latter being a new ST from an invasive type IV isolate from Minnesota (P. Ferrieri and R. Creti, unpublished data). Sequence being a new ST from an invasive type IV isolate from Minnesota, from 0.7% in 1993 to 1999 to 1.8% in 2000 to 2008 (16). Although this is a small increase, it is significant since in the last 8 years, type IV has been found mostly in infants with early-onset or late-onset disease compared to previous years, where type IV invasive disease was seen primarily in adults (16).

The results from the present study of type IV isolates by DNA macrorestriction and MLST suggested that, in general, there was an association between the sequence type, CPS/protein profile, and PFGE patterns. This is supported by findings that except for one isolate, all ST 459 isolates expressed the α protein only and were in PFGE profile groups 38 and 39, and, with one exception, ST 452 isolates expressed the α,BPS or BPS protein and were in PFGE profile group 37.

The protein profile appeared to be associated with PFGE profile diversity, evidenced by the wider variety of PFGE profiles for IV/α-only isolates than for those with α,BPS or BPS. However, this diversity of PFGE profiles among isolates with identical to the group 39 prototype strain, while the remaining 10 isolates (62.5%) were distributed throughout subgroups 39a to 39d. PFGE group 38 was too small for meaningful comparisons.

**DISCUSSION**

It was of particular interest to study the vaginal/rectal type IV isolates from this population because the 8.4% rate of isolation among this group of patients is more than a 10-fold increase compared to rates found in our previous studies of colonizing GBS isolates from mother-infant pairs and non-pregnant women in various regions of the United States from 1993 to 2002 (14, 15, 23). Based on this evolution of a relatively high prevalence of type IV colonization, it is possible that type IV may emerge as an important pathogen, much like type V in the 1990s (11, 20, 52). Further evidence of this possibility is the increase that we have found in the prevalence of type IV invasive GBS disease in Minnesota, from 0.7% in 1993 to 1999 to 1.8% in 2000 to 2008 (16). Although this is a small increase, it is significant since in the last 8 years, type IV has been found mostly in infants with early-onset or late-onset disease compared to previous years, where type IV invasive disease was seen primarily in adults (16).

The results from the present study of type IV isolates by DNA macrorestriction and MLST suggested that, in general, there was an association between the sequence type, CPS/protein profile, and PFGE patterns. This is supported by findings that except for one isolate, all ST 459 isolates expressed the α protein only and were in PFGE profile groups 38 and 39, and, with one exception, ST 452 isolates expressed the α,BPS or BPS protein and were in PFGE profile group 37.

The protein profile appeared to be associated with PFGE profile diversity, evidenced by the wider variety of PFGE profiles for IV/α-only isolates than for those with α,BPS or BPS. However, this diversity of PFGE profiles among isolates with

| Protein profile | PFGE profile group | No. of isolates | No. of isolates analyzed by MLST | Sequence type (no. of isolates)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α,BPS</td>
<td>37</td>
<td>45</td>
<td>4</td>
<td>452 (4)</td>
</tr>
<tr>
<td></td>
<td>37a</td>
<td>6</td>
<td>2</td>
<td>452 (2)</td>
</tr>
<tr>
<td></td>
<td>37b</td>
<td>2</td>
<td>1</td>
<td>452 (1)</td>
</tr>
<tr>
<td>BPS only</td>
<td>37</td>
<td>26</td>
<td>2</td>
<td>452 (2)</td>
</tr>
<tr>
<td></td>
<td>37a</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>α only</td>
<td>37</td>
<td>1</td>
<td>1</td>
<td>452 (1)</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>3</td>
<td>2</td>
<td>459 (2)</td>
</tr>
<tr>
<td></td>
<td>38c</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>5</td>
<td>2</td>
<td>459 (2)</td>
</tr>
<tr>
<td></td>
<td>39a</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39b</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39c</td>
<td>4</td>
<td>1</td>
<td>459 (1)</td>
</tr>
<tr>
<td></td>
<td>39d</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

| None           | 39                 | 1               | 1                              | 459 (1)                     |

a The numbers in parentheses represent the numbers of isolates with the specified sequence type.

**FIG. 1.** Small macrorestriction analysis by PFGE of 15 CPS type IV isolates with α-only, BPS, and α,BPS proteins. Lane designations are along the top of the gel, and PFGE profiles are along the bottom. “Std” refers to the lambda molecular size standard. (A) Seven type IV isolates with BPS and α,BPS protein profiles. Lane 1, lambda standard; lanes 2 to 5, IV/BPS; lanes 6 to 8, IV/α,BPS; lane 9, Ib/c(α+β) (internal standard). Lane 6 is the type 37 prototype. Isolates in lanes 2, 4, 6, and 7 were ST 452. (B) Eight CPS type IV isolates with the α protein. Lane 1, lambda standard; lanes 2 to 9, type IV/α only); lane 10, type Ib/c(α+β) (internal standard). Lanes 6 and 7 are duplicates of the type 39 prototype; lane 8 is the type 38 prototype. Isolates in lanes 3 and 6 to 9 were of ST 459.

**TABLE 1.** Protein profile, PFGE profile, and sequence type distribution of 101 CPS type IV isolates from 97 patients in a 2004-2008 GBS vaccine study in the United States

| Protein profile | PFGE profile group | No. of isolates | No. of isolates analyzed by MLST | Sequence type (no. of isolates)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α,BPS</td>
<td>37</td>
<td>45</td>
<td>4</td>
<td>452 (4)</td>
</tr>
<tr>
<td></td>
<td>37a</td>
<td>6</td>
<td>2</td>
<td>452 (2)</td>
</tr>
<tr>
<td></td>
<td>37b</td>
<td>2</td>
<td>1</td>
<td>452 (1)</td>
</tr>
<tr>
<td>BPS only</td>
<td>37</td>
<td>26</td>
<td>2</td>
<td>452 (2)</td>
</tr>
<tr>
<td></td>
<td>37a</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>α only</td>
<td>37</td>
<td>1</td>
<td>1</td>
<td>452 (1)</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>3</td>
<td>2</td>
<td>459 (2)</td>
</tr>
<tr>
<td></td>
<td>38c</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>5</td>
<td>2</td>
<td>459 (2)</td>
</tr>
<tr>
<td></td>
<td>39a</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39b</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>39c</td>
<td>4</td>
<td>1</td>
<td>459 (1)</td>
</tr>
<tr>
<td></td>
<td>39d</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

| None           | 39                 | 1               | 1                              | 459 (1)                     |

a The numbers in parentheses represent the numbers of isolates with the specified sequence type.
the α protein only may be attributed to additional factors, such as the length of time since the original clone appeared, genetic evolution of the individual serotypes (4, 11, 13), or lateral gene transfer (49).

Prior to this study, type IV isolates had been examined by PFGE only on a very small scale. Two Italian studies analyzed type IV isolates using PFGE, one finding three distinct PFGE groups among six isolates (42) and the other finding three groups among three isolates (19). A study in France examined two type IV isolates and found them to be in the same PFGE group (41). In Norway, six distinct band patterns were found among six invasive type IV isolates (44), but a grouping method was not applied. Due to technical limitations of comparing these previously reported PFGE profiles, we could not ascertain any resemblance to our DNA profiles.

The new sequence types 452 and 459, exhibited by the type IV isolates in this study, have allelic similarities to STs in clonal complexes (CCs) 23 and 1, respectively. CC 23 has typically included, among others, STs 23 and 24, which exhibit 4 and 6 allelic similarities to ST 452, respectively (19, 34). CC 1 has included STs 1, 2, and 196, which have 4, 5, and 6 similarities to ST 459, respectively (7, 19, 29). CC 1 and CC 23 have been found to be among the major clonal complexes associated with human GBS invasive disease and colonization (26, 32).

Several sequence types of CPS type IV isolates have been reported in the literature from 2005 to 2009, the most common being sequence type 196 (19, 22, 29, 32, 35, 51). Among the type IV invasive isolates from Minnesota collected between 1996 and 2007 were STs 196 and 291 as well as the new sequence types 452 and 459 described here plus a new sequence type, ST 468 (P. Ferrieri and R. Creti, unpublished data). In contrast to STs of more recently isolated type IV isolates, type IV reference strain 3139 from the 1970s was designated ST 2 (Ferrieri and Creti, unpublished).

In comparing the type IV isolates from our current study to those examined by other studies in our laboratory from 1995 to 2000, we found that they differed in surface protein expression and PFGE profiles. Prior to this study, our type IV isolates expressed the R1,R4 (Alp3) protein and, like the much older type V isolates (2). Compared to PFGE profile 4, the new PFGE profiles 37, 38, and 39 differed by more than six bands, as measured by the size of the DNA fragments.

The type IV isolates in this study were discovered to belong to three PFGE groups with three different protein profiles very soon after the first appearance of CPS type IV in three geographically distant regions in the United States. This finding is in contrast to our previous observations with type V, PFGE profile group 4, isolates, whose DNA macrorestriction patterns very gradually shifted, over a period of several years, from being identical to the O profile described previously by others (11) to evolving into subgroups of profile 4 (2). Despite differences, the type IV isolates presented here were highly clonal, evidenced by the finding that all 80 type IV isolates with α,BPS or BPS belonged to PFGE group 37, associated with ST 452. Although 20 of the 21 isolates with α-only or no proteins showed greater diversity, they spanned only two closely related PFGE groups (groups 38 and 39), associated with ST 459.

PFGE and MLST, although sometimes viewed as inter-changeable, were each shown to have unique benefits. PFGE appeared to have greater discriminatory power than MLST, and it could be a valuable molecular epidemiological tool in the investigation of clusters of type IV disease or colonization. Using MLST, it was found that the type IV isolates had two unique sequence types, while PFGE demonstrated that many slight DNA profile variations existed within each sequence type. However, MLST was a more objective method that allowed us to compare our results with those of other laboratories and to see allelic relationships among sequence types.

GBS conjugate vaccines that are being developed to prevent invasive disease may protect only against serotypes Ia, Ib, II, III, and V. Although a pentavalent vaccine that includes the more common serotypes is expected to provide protection for a high percentage of GBS cases (3, 20, 38), the global presence of type IV and the possibility of shifts in serotype predominance due to serotype replacement, capsular switch, or changes in bacterial virulence factors (1, 9, 31, 40, 49) necessitates ongoing surveillance of types IV and VI to VIII and the newly proposed type IX (46). Because of the emergence of invasive disease due to type IV in Minnesota, Boston, and other regions of the United States (16, 27, 39, 45), monitoring of type IV colonization and disease in the population is warranted. If the prevalence of type IV in invasive disease significantly increases, this polysaccharide and/or associated GBS virulence proteins should be included in vaccine design (33, 49).

ACKNOWLEDGMENT

This work was supported by contract N01-AI-25495 from the National Institutes of Health, Bethesda, MD.

REFERENCES


