INFECTION AND IMMUNITY, Mar. 1996, p. 919–925
0019-9567/96/$04.00+0

Vol. 64, No. 3

Population Structure of *Streptococcus agalactiae* Reveals an Association between Specific Evolutionary Lineages and Putative Virulence Factors but Not Disease

MAJBRITT HAUGE, CHRISTINA JESPERSGAARD, KNUD POULSEN, AND MOGENS KILIAN*

Department of Medical Microbiology and Immunology, University of Aarhus, Aarhus, Denmark

Received 28 August 1995/Returned for modification 8 November 1995/Accepted 22 December 1995

To evaluate the genetic diversity and relationships in a collection of 85 Danish strains of *Streptococcus agalactiae* (group B streptococcus) we have performed restriction fragment length polymorphism analysis on *Eco*RI- and *Msp*I-digested whole-cell DNA using as probes rRNA, DNA fragments representing the genes encoding hyaluronidase, C5a-peptidase, alpha-antigen, and beta-antigen as well as two randomly selected genomic DNA fragments for which the coding potential is unknown. In addition, we have assayed for expression of hyaluronidase activity and beta-antigen. Combined analyses of our data and those previously obtained by multilocus enzyme electrophoresis and serotyping revealed a population separating into six major lineages that correlate with individual serotypes. The significant linkage disequilibrium of alleles indicates that the *S. agalactiae* population examined is predominantly clonal. Notably, strains expressing the serotype III capsule divide into two distant evolutionary lineages, of which one lacks expression of hyaluronidase activity. Six North American isolates of serotype III clustered together with multiple Danish serotype III strains, showing that the combinations of characters on which the phylogenetic tree was based are conserved worldwide. Occurrence of beta-antigen correlated with a specific version of the alpha-antigen gene and was exclusively associated with a single major phylogenetic lineage. Comparisons with the clinical history of the strains revealed no evidence of differences in pathogenic potential among the six major genetic divisions.

Streptococcus agalactiae is an important human pathogen. In the United States and most other developed countries, this bacterium is the leading cause of sepsis and meningitis in neonates and pregnant women and invasive *S. agalactiae* disease is a major problem also in nonpregnant adults, especially those who are elderly and those who have chronic disease (7). Like most other bacterial pathogens, *S. agalactiae* is also found in healthy carriers, and it is not known why some individuals develop disease whereas the majority remain unaffected. Differences in the immunologic status of the host as well as variation in virulence potential among the *S. agalactiae* strains may determine the outcome of colonization.

S. agalactiae possesses a variety of potential virulence factors including a polysaccharide capsule occurring in seven distinct structural and antigenic types (Ia, Ib, II, III, IV, V, VI), secretion of the enzymes hyaluronidase (formerly incorrectly identified as neuraminidase [21]) and C5a-peptidase, and the ability to bind human immunoglobulin A via the Fc portion to the surface-associated beta-antigen protein. Only little is known of the exact function of these factors in the process of infection. The serotype-specific capsular polysaccharides are essential for pathogenesis (22), and serotype III is prevalent among strains isolated from infants with invasive infections (1, 6). Similarly, elevated levels of hyaluronidase activity have been found to be associated with serotype III strains causing neonatal disease (17).

Some strains, mainly of serotypes Ia, Ib, and II, express the alpha- and beta-antigens which belong to the C protein complex (2). Similar to the capsule, the alpha-antigen may protect the bacterium against opsonophagocytic killing in the absence of antibody (13). The alpha- and beta-antigens as well as the capsular polysaccharides elicit protective immunity in mice (10, 15, 22).

For a number of bacterial pathogens, population genetic analysis based on multilocus enzyme electrophoresis (MLEE) combined with information on site of isolation has been a valuable tool for evaluating the variation in virulence potential among strains (4, 18, 20). Using this approach, Musser et al. (19) identified an apparently high-virulence clonal type of *S. agalactiae* serotype III strains causing neonatal disease. In the collection of North American strains analyzed the genetic relationships revealed by MLEE did not correlate with serotype of the strains and they concluded that serotypic identity does not indicate close genetic relatedness. Overall similar results were obtained by Helmig et al. (8) for a collection of *S. agalactiae* strains isolated in Denmark, though the clonality and genetic distances differed somewhat from those observed for the North American collection. In addition, Helmig et al. found evidence for a phylogenetic division of *S. agalactiae* into two groups, one comprising virtually all isolates from diseased infants and with a clonal population structure like traditional mucosal pathogens and one containing the majority of carrier isolates and with a genetic structure resembling that of opportunistic pathogens characterized by a large number of clonal types.

In the present study we have further examined the collection of *S. agalactiae* studied by Helmig and coworkers and used additional typing methods to characterize the strains at the DNA level with the primary aim of identifying virulence-associated properties. The results reveal genetic variation and relationships not observed previously by MLEE typing alone. We observe a strong correlation between the different characteristics including serotype and do not find conclusive evidence for differences in pathogenic potential among the major genetic divisions.

* Corresponding author. Mailing address: Department of Medical Microbiology and Immunology, The Bartholin Building, University of Aarhus, DK 8000 Aarhus C, Denmark. Phone: 45 89 42 17 35. Fax: 45 86 19 61 28.

MATERIALS AND METHODS

Bacterial strains. The collection of *S. agalactiae* examined has been previously described and analyzed using serotyping and MLEE (8). In the present study only a single representative of multiple and apparently identical isolates from the same subject or from mother-child pairs was included. In addition, seven strains of the original collection were lost. The remaining 91 human isolates of *S. agalactiae* included 85 isolates from Denmark and 6 isolates from the United States obtained from J. M. Musser, Houston, Tex. The Danish isolates were collected between 1988 and 1992 from infants with neonatal sepsis or meningitis or from their mothers (37 strains) and from healthy pregnant women (48 strains). The six North-American strains were three representatives of the "high-virulence" serotype III clonal type (ET 1) and three strains of a "low-virulence" serotype III clonal type (ET 12) defined by Musser et al. (19).

RFLP analyses. Whole cell DNA was isolated as described previously (9) by a method involving treatment of the bacteria harvested from a Todd-Hewitt broth culture with lysozyme, lysis with sodium dodecyl sulfate (SDS) and digestion with proteinase K, extractions with phenol/chloroform, and eventual ethanol precipitation. The DNA was dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5), and the quality and quantity were examined by agarose gel electrophoresis. Restriction endonuclease digestion of approximately 2 μg of whole-cell DNA with *Eco*RI and *Msp*I was performed according to the manufacturer's recommendations (Boehringer Mannheim, Mannheim, Germany). Following treatment with 0.05 μg of RNase (DNase free, Boehringer Mannheim) electrophoresis was performed in 1% agarose gels for 16 h at 2 V/cm in TAE buffer (0.04 M Tris/acetate, 0.002 M EDTA, pH 8.0). The restriction fragments were then transferred and fixed onto Nytran nylon membranes (Schleicher & Schuell, Keene, N.H.) and hybridizations were performed at 65°C as described previously (23) except that the membranes were soaked in 1% (vol/vol) Triton X-100 before prehybridization and that 0.1% sodium pyrophosphate was included in all hybridization buffers. The final posthybridization wash was at 65°C in 0.5× SET (0.075 M NaCl, 0.25 mM EDTA, 10 mM Tris-HCl, pH 7.0)–0.1% SDS–0.1% sodium pyrophosphate. The same membranes were used for several hybridizations each time after stripping by soaking in boiling 0.1% SDS.

The probe used for ribotyping was 16S and 23S rRNA from *Escherichia coli* (Boehringer Mannheim) labeled with ^{32}P-dATP (Amersham International, Amersham, England) by randomly primed reverse transcription using murine leukemia virus (MuLV) reverse transcriptase (Boehringer Mannheim) as described previously (23). The two anonymous probes termed ap-Pst0.8 and ap-Pst4.0 were DNA fragments of 0.8 and 4.0 kb, respectively, randomly selected from a library of *Pst*I-digested genomic DNA from *S. agalactiae* 3165, serotype III, cloned into the plasmid vector pBluescript KS (Stratagene, La Jolla, Calif.) and propagated in *E. coli* XLI Blue (Stratagene). The two probes used for restriction fragment length polymorphism (RFLP) analysis of the hyaluronidase gene area were fragments of genomic DNA from *S. agalactiae* 3076, serotype Ib, amplified by PCR and cloned into plasmid pCR using the TA-Cloning Kit (R & D Systems Europe Ltd., Abingdon, United Kingdom). The PCR primers were chosen according to the published nucleotide sequence of the hyaluronidase gene (12), and the fragments represented base pairs (bp) 51 to 1490 and 1462 to 3009, respectively. The two probes were hybridized separately to the membranes. The RFLP typing of the C5a-peptidase gene area was done with a 3.5 kb fragment of genomic DNA from *S. agalactiae* 941, serotype Ia, amplified by PCR using the ForW and RevW primers as described previously (4). The identity of the PCR product was confirmed by restriction analysis with *Acc*I (Boehringer Mannheim) resulting in a pattern in agreement with the published restriction map of the C5a-peptidase gene (6). Two alpha-antigen probes were used for hybridization. Both were DNA fragments of genomic DNA from *S. agalactiae* 3076 amplified by PCR and cloned by using the TA-Cloning Kit. The two sets of primers were chosen according to the published nucleotide sequence to represent the 5′ end and virtually the entire alpha-antigen gene including the tandem repeat region, i.e., bp 249 to 756 and 249 to 3018, respectively, in Michel et al. (16). The beta-antigen gene was detected by two probes in the RFLP analysis. The two fragments of genomic DNA from *S. agalactiae* 950, serotype Ib, were amplified by PCR and cloned by using the TA-Cloning Kit. The primers were chosen according to the published nucleotide sequence of the beta-antigen gene (10) and represented bp 432 to 2115 and 1914 to 3778, respectively. The two probes were hybridized separately. The identities of the hyaluronidase and alpha- and beta-antigen PCR fragments cloned into the pCR vector were confirmed by partial sequencing by using the Sequenase Version 2.0 DNA Sequensing Kit (Amersham), and vector sequences were removed by *Eco*RI digestion followed by agarose gel electrophoresis and elution using the Gene Clean Kit (BIO 101, Vista, Calif.). All the DNA fragments used as probes were labelled with ^{32}P-dATP by nick translation (23).

Hyaluronidase activity assay. Detection of hyaluronidase was performed on brain heart infusion agar containing hyaluronidase and albumin according to the method described by Smith and Willett (24). The approximate diameter of the clear zones around colonies developing after 10 min of incubation with 2 N acetic acid was measured.

SDS-PAGE and Western blots. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of cell-bound proteins was performed using 4 to 20% gradient gels. Approximately 30 mg (wet weight) of cells was boiled for 10 min. in 200 μl of sample buffer (0.2 M Tris-HCl [pH 8.8], 2% SDS, 10% sucrose, 5% mercaptoethanol, 0.005% pyronine, 0.001% bromphenol blue), and 10 μl of the supernatant was applied to the gel. After electrophoresis the proteins in the gel were stained with Coomassie brilliant blue dye or alternatively they were electroblotted onto nitrocellulose filters. N-terminal amino acid sequencing was performed by Lars Sottrup-Jensen, Department of Molecular Biology, University of Aarhus, on the predominant band that varied in occurrence among the strains. A rabbit antiserum was raised against beta-antigen by immunizing with the band cut out of the gel after staining with 0.5 M KCl and homogenized in PBS buffer (0.01 M Na phosphate, 0.15 M NaCl, pH 7.4). The Western blots (immunoblots) were incubated with a 1:1,000 dilution of the antiserum followed by alkaline phosphatase-conjugated swine antirabbit antibody (diluted 1:3,000; Dako) and subsequently visualized with indoxyl phosphate-nitroblue tetrazolium solution.

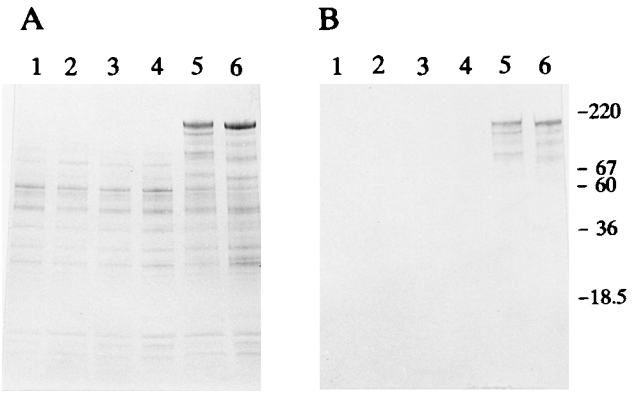


FIG. 1. Gel electrophoresis and immunoblotting of *S. agalactiae* cell-bound proteins. (A) The proteins were prepared by boiling the cells in SDS (as described in the text) from representatives of clonal types 3 (lane 1), 8 (lane 2), 10 (lane 3), 8 (lane 4), 20 (lane 5), and 23 (lane 6) and analyzed by SDS-PAGE. (B) A similar gel was blotted onto a nitrocellulose membrane and probed with a polyclonal antiserum raised against the beta-antigen. Molecular mass markers in kilodaltons are indicated to the right.

Clustering analysis. Analysis of the data obtained for the 91 strains was performed using the programs ETDIV and ETCLUS by T. S. Whittam, Department of Biology, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park. Null alleles from the MLEE analyses were treated as missing data. For the RFLP data, different patterns of each of the gene regions analyzed were treated as different alleles. The clonal types were each characterized by a distinct combination of alleles of the 23 characteristics included in the cluster analysis.

RESULTS

SDS-PAGE of cell-bound proteins. To estimate the overall phenotypic diversity in the collection of 85 Danish and 6 North American *S. agalactiae* isolates, we examined the variation in size of cell-bound proteins. The profiles of soluble proteins upon boiling the cells in SDS were very similar except for a predominant protein with an apparent molecular mass of 120 kDa (Fig. 1) which was present only in some strains (see below). N-terminal sequencing of 20 amino acid residues of the protein in this band of strain 3076, serotype Ib, revealed that it represented the beta-antigen.

Ribotypes. To estimate the genetic relationships at the DNA level in the population, we performed ribotyping of the strains. Southern blots of total cellular DNA digested with the restriction enzymes *Eco*RI and *Msp*I were probed with reverse-transcribed 16S and 23S rRNA from *E. coli*, and the patterns of hybridization were visually compared. Five different patterns characterized by 4 to 6 bands were distinguished in the *Eco*RI blots, and seven patterns characterized by 9 to 12 bands were detected among the *Msp*I blots (Fig. 2). Details on these data are available upon request. By combining the results obtained with the two enzymes the number of ribotypes detected among the 91 strains amounted to 12 (termed 1 through 12 in Fig. 3).

RFLP analyses using anonymous probes. To further esti-

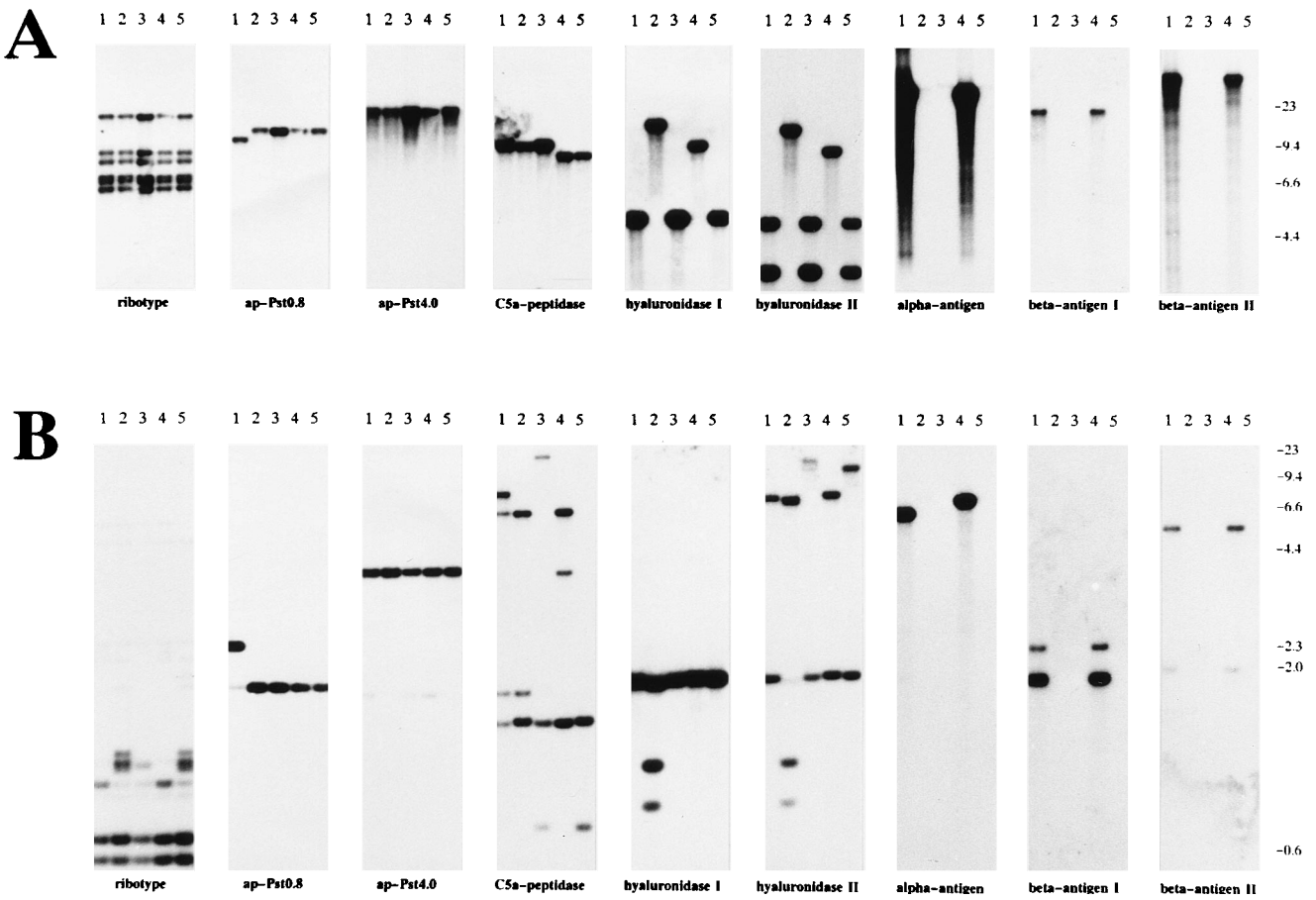


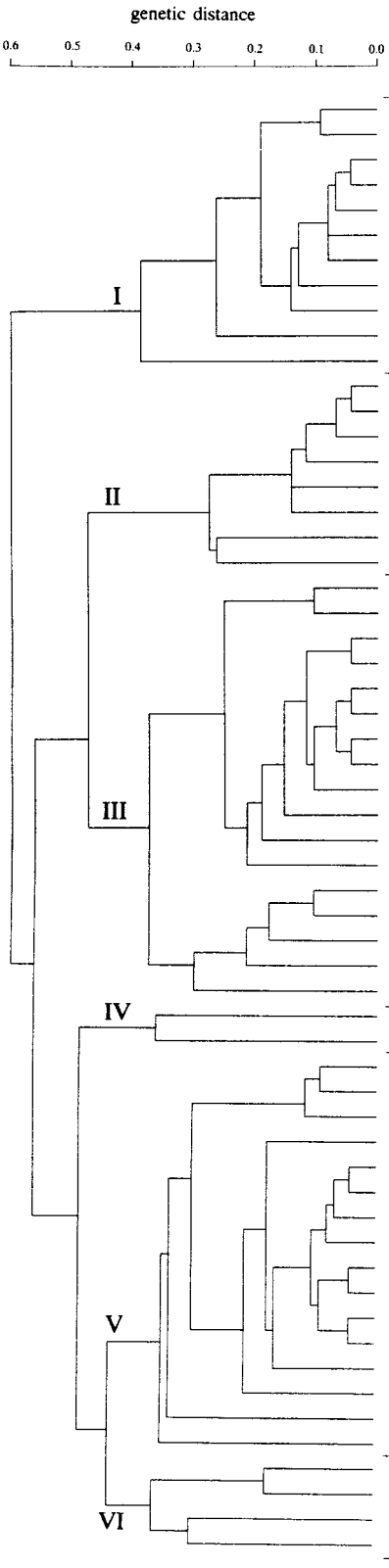
FIG. 2. Examples of *Eco*RI and *Msp*I RFLP typing of whole-cell DNA from *S. agalactiae*. The figure shows the autoradiograms of Southern blots of *Eco*RI-restricted DNA (panel A) and *Msp*I-restricted DNA (panel B) from strains representing clonal types 34 (lane 1), 4 (lane 2), 44 (lane 3), 27 (lane 4), and 38 (lane 5) hybridized sequentially with each of the nine probes rRNA (ribotype); ap-Pst0.8; ap-Pst4.0; and DNA fragments representing the C5a peptidase gene, the 5′ half of the hyaluronidase gene (hyaluronidase I), the 3′ half of the hyaluronidase gene (hyaluronidase II), the alpha-antigen gene, the 5′ part of the beta-antigen gene (beta-antigen I), and the 3′ part of the beta-antigen gene (beta-antigen II) as described in the text. Molecular mass markers in kilobases are indicated to the right of the autoradiograms.

mate the genetic variation and relationships in the population we hybridized the Southern blots with two anonymous DNA probes, ap-Pst0.8 and ap-Pst4.0, randomly selected from a genomic library of *S. agalactiae* 3165. All strains hybridized with the two probes (Fig. 2). Differences in the hybridization patterns divided the 91 strains into 12 (A through L) and 6 (a through f) different ap-Pst0.8 and ap-Pst 4.0 RFLP types, respectively. The distribution of the individual RFLP types in the strain material is shown in Fig. 3.

Hyaluronidase activity and polymorphism of the hyaluronidase gene. To investigate if expression of hyaluronidase, which is a possible virulence factor of *S. agalactiae*, or if distinct versions of the hyaluronidase gene were correlated with virulence we assayed each strain for hyaluronidase activity and hybridized the Southern blots with two DNA probes representing each half of the hyaluronidase gene. Hyaluronidase activity was detected in 67 of the 85 Danish isolates including all serotype Ia, Ib, II, and V strains (Fig. 3). The 18 hyaluronidase negative strains were all of serotype III except for one closely associated nonserotypable strain. The three high-virulence North-American serotype III strains were all positive for hyaluronidase, whereas the three low-virulence serotype III strains lacked the activity.

Hybridization to the Southern blots of *Eco*RI- and *Msp*I-digested whole-cell DNA with the two DNA probes representing each half of the hyaluronidase gene divided the strains into 10 distinct hyaluronidase RFLP types (designated 1 through 10) (Fig. 2). The probes hybridized to all strains including those that had no detectable hyaluronidase activity, indicating that absence of activity is not due to lack of the gene. Hyaluronidase RFLP type 2 was associated with lack of hyaluronidase activity (Fig. 3). The amount of hyaluronidase measured as the size of clearing zone in the activity assay varied among the strains (data not shown). Interestingly, we observed a correlation between apparently reduced hyaluronidase activity and hyaluronidase RFLP type 3 (Fig. 3), of which most were serotype Ib. None of the hyaluronidase RFLP types appeared to be significantly correlated with disease (Fig. 3). Notably, the presence of hyaluronidase activity was apparently not crucial for pathogenicity as 24% of the Danish disease isolates belonged to RFLP type 2 which is associated with lack of enzymatic activity.

C5a-peptidase RFLP typing. To examine whether certain versions of the gene encoding C5a-peptidase might be correlated with virulence, we probed the Southern blots with the C5a-peptidase gene (Fig. 2). All 91 strains except one hybridized with the probe. A total of 18 distinct C5a-peptidase RFLP types, termed 1 through 18, were identified (Fig. 3).



clonal type	number of strains	source	serotype	ribotype	ap-Pst0.8 RFLP type	ap-Pst4.0 RFLP type	hyaluronidase expression	hyaluronidase RFLP type	C5a-peptidase RFLP type	alpha-antigen gene type	beta-antigen expression	beta-antigen gene type
1	1	1C	III	4	B	c	-	2	2	0	-	0
2	1	1C	III	4	B	c	-	2	2	0	-	0
3	8	2C,4P,2L	III	4	B	c	-	2	2	0	-	0
4	2	1P,1L	III	4	A	c	-	2	2	0	-	0
5	1	1C	III	4	L	c	-	2	2	0	-	0
6	1	1C	III	4	B	c	-	2	2	0	-	0
7	1	1P	III	9	B	c	-	2	2	0	-	0
8	2	1C,1P	III	4	B	f	-	2	2	0	-	0
9	1	1P	III	4	B	c	-	2	2	0	-	0
10	1	1C	III	4	A	b	-	8	2	0	-	0
11	1	1P	III	4	K	c	-	2	2	0	-	0
12	1	1P	Ia	1	E	a	+	4	1	0	-	0
13	10	5C,5P	Ia	1	A	a	+	4	1	0	-	0
14	1	1C	Ia	1	B	a	+	4	1	0	-	0
15	1	1P	Ia	1	A	a	+	10	1	0	-	0
16	1	1P	Ia	1	A	a	+	4	17	0	-	0
17	1	1C	NT	6	A	a	+	4	1	0	-	0
18	1	1C	Ia	1	A	a	+	5	1	0	-	0
19	1	1C	Ia	5	D	d	+	5	16	0	-	0
20	1	1C	V	2	A	b	+	3	5	0	+	+
21	1	1C	V	2	A	b	+	3	4	0	+	+
22	1	1P	Ib	2	A	b	+	3	1	+	-	+
23	1	1C	Ib	2	A	b	+	3	1	+	+	+
24	3	3C	Ib	2	A	b	+	3	1	+	+	+
25	1	1P	Ib	2	C	b	+	3	1	+	+	+
26	1	1C	Ib	2	A	b	+	3	1	+	+	+
27	3	2C,1P	Ib	2	A	b	+	3	1	+	+	+
28	1	1P	Ib	2	A	b	+	3	1	+	-	+
29	1	1P	Ia	7	A	b	+	3	1	+	+	+
30	1	1C	Ib	8	J	b	+	3	17	+	+	+
31	1	1P	III	2	A	b	+	4	1	+	-	+
32	1	1C	II	2	H	b	+	6	8	+	+	+
33	1	1C	II	2	F	b	+	6	6	+	+	+
34	1	1C	II	2	G	b	+	6	7	+	+	+
35	1	1C	V	2	F	e	+	6	6	+	+	+
36	1	1P	II	7	I	b	+	6	2	+	-	0
37	1	1C	II	4	A	c	+	5	3	0	-	0
38	1	1C	NT	12	A	c	+	5	15	0	-	0
39	1	1P	III	3	A	b	+	1	3	0	-	0
40	1	1C	III	3	A	b	+	1	3	0	-	0
41	1	1C	III	3	C	b	+	1	3	0	-	0
42	1	1P	III	3	A	b	+	1	0	0	-	0
43	8	1C,6P,1H	III	3	A	b	+	1	3	0	-	0
44	2	1P,1H	III	3	A	b	+	1	9	0	-	0
45	1	1P	III	3	A	b	+	1	3	0	-	0
46	1	1C	NT	3	A	b	+	1	3	0	-	0
47	1	1C	III	3	C	b	+	1	10	0	-	0
48	2	1C,1P	III	3	C	b	+	1	3	0	-	0
49	1	1C	III	3	C	b	+	1	9	0	-	0
50	3	2C,1P	III	3	C	b	+	1	9	0	-	0
51	1	1C	III	3	C	b	+	1	11	0	-	0
52	1	1H	III	2	C	b	+	7	18	0	-	0
53	1	1C	II	3	C	b	+	1	12	0	-	0
54	1	1P	III	3	A	b	+	1	3	0	-	0
55	1	1C	NT	10	A	b	+	5	5	0	-	0
56	1	1C	NT	10	I	b	+	5	3	0	-	0
57	1	1C	NT	11	A	b	-	9	14	0	-	0
58	1	1C	Ia	11	A	b	+	5	13	0	-	0

Alpha- and beta-antigen analysis. To examine the diversity of the genes encoding the C-protein alpha- and beta-antigens we hybridized the Southern blots with DNA probes representing part of the genes encoding these two proteins. Furthermore, we assayed for expression of the beta-antigen. For 19 of the 91 strains, including all strains of serotype Ib, probing with the 5′-terminal part of the alpha-antigen gene resulted in a single strongly hybridizing fragment in both the *Eco*RI and the *Msp*I digestions (Fig. 2). In both digests the size of the fragment varied slightly, but this variation was not consistently scored. Virtually all strains showed additional weakly hybridizing fragments upon prolonged exposure of the autoradiograms, indicating that the *S. agalactiae* genome contains additional sequences with significant homology to the alpha-antigen gene probe (Fig. 2). Hybridization of the *Eco*RI digests with the second alpha-antigen gene probe, which included the tandem repeat region, revealed hybridizing fragments in all strains that hybridized strongly with the probe representing only the 5′-terminal of the gene. However, the larger probe, in addition, detected highly homologous sequences in DNA from the majority of serotype Ia strains, in a single serotype III strain, and in two nontypable strains (clonal types 12 to 17, 51, 57 and 58 in Fig. 3). The intensity of hybridization differed significantly between strains presumably as a result of different numbers of repeats in the alpha-antigen gene in individual strains. As a consequence of this variation and the fact that results obtained with the two probes are not independent, only hybridization results obtained with the short probe are included in Fig. 3 and the analyses forming the basis of the dendrogram (see below).

The beta-antigen RFLP analysis was performed using two probes representing each half of the beta-antigen gene. Among the 91 strains analyzed 20 hybridized with both probes, whereas the remaining 71 strains showed no hybridization. Although the pattern of hybridization varied, we scored only for presence versus absence of hybridization. Of the 20 strains with a detectable beta-antigen gene, 17 were found to express the protein. Expression of beta-antigen was detected by Western blots of cell-bound proteins, incubated with antisera raised against beta-antigen (Fig. 1).

Genetic relationships in the *S. agalactiae* population analyzed. Preliminary analysis of the data obtained in this study revealed a strong correlation between individual characters including serotype. This finding contrasts with the scattering of serotypes in the dendrogram obtained for the same strain collection by Helmig et al. (8) based on allelic variation in 11 enzyme loci. This discrepancy prompted us to reexamine the data of Helmig et al. For the enzyme ADK, a large number of strains lacked enzyme activity, which, in the statistical analysis was treated as a true null allele. This is most likely a misinterpretation, as lack of ADK activity did not correlate with any of the mutually correlating characteristics detected in the present study. In addition, the allelic variation in two pairs of enzymes (PM1 and PM2, PLP and LGG) was identical, suggesting that these pairs may represent identical proteins.

We performed cluster analysis on the *S. agalactiae* strain collection combining the MLEE and serotype data of Helmig et al. (8) (excluding data on the enzymes ADK, PM1, and PLP) with the data obtained in the present study including data for each of the restriction enzymes *Eco*RI and *Msp*I on ribotyping, anonymous probe RFLP typing, each of the two hyaluronidase probe RFLP typings, and C5a-peptidase RFLP typing as well as presence or absence of the alpha- and beta-antigen genes, the former as detected with the 5′-terminal probe only. In contrast to occasional lacking enzyme activities in the MLEE analyses, alpha- and beta-antigen gene types 0 were included as true null alleles. Hyaluronidase expression was not included as a separate characteristic, as RFLP types of the corresponding gene reflected presence or absence of enzyme activity. Hence, a total of 23 characteristics was included in the cluster analysis.

A total of 58 distinct combinations of alleles, each representing a clonal type, were found among the 91 *S. agalactiae* strains, with each comprising from 1 to 10 strains (Fig. 3). The resulting dendrogram, which is a computer-generated genealogical tree assuming no recombination between strains, is shown in Fig. 3. At a genetic distance of approximately 0.4, i.e. the level at which strains joining in the dendrogram, in average, differ in 40% of the characteristics analyzed, the clonal types clustered into six divisions (termed I through VI in Fig. 3). The six divisions correlated strongly with serotype, ribotype, and hyaluronidase gene type and expression. Presence of gene sequences hybridizing with the alpha- and beta-antigen gene probes as well as expression of beta-antigen was exclusively associated with strains in division III. The strong correlation between individual characters indicates a clonal structure of the population. However, few exceptions, e.g., ap-Pst4.0 RFLP type b of clonal type 10, serotype III of clonal type 31, and ribotype 4 of clonal type 37, suggest that recombination between strains may occur.

Serotype III strains clustered into two distinct divisions linking at a genetic distance of 0.6. While serotype III strains in division V showed hyaluronidase activity, with one exception (clonal type 52) associated with RFLP type 1, those of division I lacked hyaluronidase activity and showed hyaluronidase RFLP type 2 (except for clonal type 10). The two divisions also showed distinct and exclusive ribotypes, i.e., 3 and 4, respectively, and different C5a-peptidase RFLP patterns. RFLP patterns detected with the two anonymous probes ap-Pst0.8 and ap-Pst4.0 showed less correlation with other characters. Thus, ap-Pst0.8 RFLP type A was found among all serotypes as well as among all prevalent ribotypes. With few exceptions, RFLP type B was associated with serotype III combined with ribotype 4, and RFLP type C with serotype III combined with ribotype 3.

With the exception of divisions IV and VI, which contained only two and four strains, respectively, all divisions composed both carrier isolates and patient isolates. No significant difference in relative proportion of carrier and patient isolates was observed in these divisions, and no single characteristic preferentially was associated with disease or carriage (Fig. 3).

The six North American isolates representing high- and low-virulence clonal types, respectively, identified by Musser et al.

FIG. 3. Genetic relationships among the 91 strains of *S. agalactiae* analyzed. The dendrogram was constructed from the data on MLEE, ribotyping, RFLP typings, and presence or absence of the alpha- and beta-antigen genes as described in the text. Each distinct combination of these characteristics was given a clonal type numbered sequentially from top to bottom. Six major lineages, termed I through VI, separating at a genetic distance of 0.4 are indicated to the left. The columns to the right indicate the number of strains of each clonal type, the source of isolation (C, carrier isolate; P, patient isolate; L, low-virulence American isolate; H, high-virulence American isolate) and results for serotype and *Eco*RI and *Msp*I RFLP types by using, as probes, rRNA (ribotype), two anonymous DNA fragments (ap-Pst0.8 and ap-Pst4.0), the hyaluronidase gene, the C5a-peptidase gene, the beta-antigen gene, and the alpha-antigen gene, as well as results on expression of hyaluronidase and beta-antigen. Alpha-antigen gene types + and 0 indicate presence and absence, respectively, of hybridization. See text for additional nomenclature of the typings.

(19), clustered in the two separate divisions (I and V) of serotype III strains together with Danish isolates.

DISCUSSION

During recent years population genetic studies of many species of human and veterinary bacterial pathogens have provided valuable frameworks for the understanding of ecology, pathogenicity, and epidemiology of these bacteria. Moreover, analysis of data from such studies often have provided insight into their phylogeny and the impact of recombination on the evolution of individual species and of single gene loci.

Most studies have used variation in the electrophoretic mobility of selected intracellular housekeeping enzymes as markers of genetic diversity and relationships. The study described here was based on two previous studies of *S. agalactiae* which, based on MLEE analysis of the same 11 gene loci, appeared to disclose potentially important differences in the pathogenic potential among subpopulations of this species. In an attempt to identify virulence factors explaining such differences, we combined MLEE data from one of these studies with data on capsular serotype and allelic variation in genes encoding rRNA, selected cell wall proteins, and the putative virulence factors hyaluronidase and C5a-peptidase, as well as two anonymous parts of the chromosome.

The dendrogram in Fig. 3, which is based on these 23 independent characteristics, divided the 58 clonal types identified among the 91 strains into six major clusters closely associated with individual serotypes. Omission of serotype data or excluding single strains from the analysis did not influence the overall structure of the dendrogram (not shown). Combined with the observed strong association between virtually all characteristics including serotype and ribotype (Fig. 3), this clearly indicates that the computer-generated dendrogram is very robust. Thus, it may be concluded that the individual serotypes of *S. agalactiae*, with the possible exception of serotype V, are genuine genetic entities. The few exceptions (e.g., clonal types 29, 31, and 53) presumably reflect rare recombination events in an otherwise predominantly clonal population. This is in sharp contrast to the findings mentioned in two studies of *S. agalactiae* based on MLEE (8, 17). Both these studies found that expression of the individual serotypic polysaccharides was evenly distributed over the populations analyzed and that several electrophoretic types included strains of different serotypes. This emphasizes that MLEE analyses based on the examination of a limited number of polymorphic enzymes may lead to erroneous conclusions. In these cases, the low number of enzymes examined, combined with the inclusion of lacking enzyme activities as null alleles and the observation that two enzymes analyzed appear to be identical (see above), apparently have resulted in a distorted picture of the genetic relationships of the strains examined.

Strains expressing the serotype III capsular polysaccharide account for approximately 60% of invasive neonatal infections, although they represent only about one-third of strains isolated from asymptomatically colonized individuals (1, 6). Such epidemiologic observations have suggested that the type III capsule serves as a virulence factor, a hypothesis that has been supported by direct experimental evidence obtained in a neonatal rat model (22). Serotype III was also significantly associated with disease in our strain material ($P = 0.026$). Our study furthermore confirms the observations by Musser et al. (19) and Helmig et al. (8) that strains expressing type III polysaccharide belong to two distantly related evolutionary lineages of clones. While, in the study by Musser et al., one of the lineages was found to consist of a single clone (ET1), which joined the remaining part of the population at a genetic distance of about 0.65, the same homogeneity was not found in the present study. Our cluster V (Fig. 3) contains three representatives of Musser's ET1, which, according to MLEE analysis of 11 gene loci, were identical (8, 19). Our analyses revealed that two of the three strains (indicated by "H" in Fig. 3) were indeed closely related whereas the third differed by several characteristics, including ribotype, RFLP type of the hyaluronidase and C5a-peptidase genes, and one of the anonymous parts of the chromosome.

Two of the ET1 strains of Musser et al. (19) were similar to 17 Danish isolates assigned to a relatively tight cluster (clonal types 43 to 50). This cluster may, in practical terms, be considered a clone, if a clone is defined as a set of genetically similar cells originating from a common ancestor and evolving without chromosomal recombination (14). Consequently, this cluster appears to represent a clone that successfully spreads throughout the world and causes disease in some individuals. The same seems to apply to the serotype III strains in the cluster around clonal type 3 and the serotype Ia strains in clonal type 13.

Musser et al. (19) concluded that their ET1 represents an unusually virulent clone that accounts, in major part, for the high morbidity and mortality associated with infections by type III organisms in the United States. This conclusion was based on the fact that 34 of 40 isolates of ET1 were isolated from serious neonatal infections in contrast ($P < 0.001$) to only 8 out of 23 from type III strains in the other division. If we compare the clinical origin of strains included in the two tight clusters of type III strains in our dendrogram (clonal types 3 to 9 versus 43 to 50) we do not find a significant difference in the distribution of disease and carrier isolates and, consequently, cannot conclude that the two divisions of type III strains differ in pathogenic potential. The strain material of our study represents a consecutive sample of isolates from Denmark. In contrast, Musser and coworkers' strains were selected from several previous studies of virulence-associated properties, which may have introduced a sampling bias. The possibility cannot be excluded, however, that the difference between the two studies is due to geographic or temporal differences in the clonal composition of *S. agalactiae* populations.

The possibility that the two evolutionary divisions of type III strains differ in virulence is particularly important, as such a difference might facilitate identification of virulence-associated properties or alleles of genes. Notably, the two divisions differ by hyaluronidase activity (erroneously identified as neuraminidase in some previous studies). While all strains had chromosomal sequences hybridizing with the hyaluronidase gene probe, type III strains of division I lacked hyaluronidase activity measured as described. On the basis of the above discussion we conclude that expression of hyaluronidase activity is not a decisive virulence factor of *S. agalactiae.*

Other putative virulence factors of *S. agalactiae* include the C5a-peptidase and the immunoglobulin A-binding beta-antigen. All strains in our collection contained chromosomal sequences hybridizing with the C5a-peptidase gene probe in agreement with previous findings based on the examination of a smaller collection of isolates (5). The diversity of this gene locus (Fig. 3) may reflect differences in the expression of the peptidase activity and, thus, in the ability to evade the C5a-mediated attraction of polymorphonuclear leukocytes (3). Further studies are required to elucidate this possibility.

In full agreement with the clonal population structure, the immunoglobulin A-binding beta-antigen as well as the alpha-antigen gene were exclusively associated with distinct evolutionary lineages encompassing primarily strains of the sero-

types Ia, Ib, and II (divisions III, II, and III, respectively [Fig. 3]). Although these divisions contain disease isolates, the absence of these antigens from other clusters containing serotype III strains suggest that the C proteins detected in this study are not major virulence determinants. However, with regard to the alpha-antigen, the observed differences in size and signal of hybridizing DNA fragments as well as the different results obtained with the two probes clearly demonstrate a significant degree of diversity, part of which may result from variations in the number of tandem repeats in the gene (16). Possible relationships between the observed differences in the alpha-antigen gene and expression of the protein remain to be examined. In addition, it is conceivable that the recently described Rib protein in serotype III strains represents the alpha-antigen analog in that serotype (25). Further studies are required to disclose the phylogenetic and functional similarities of these cell wall proteins.

This study shows that each of the serotypes of *S. agalactiae* represents a genuine genetic entity, and it confirms the phylogenetic dichotomy of strains expressing the serotype III capsular polysaccharide. However, we find no evidence for differences between the pathogenic potential of serotype III strains in the two evolutionary lineages. Although the serotype III capsule is significantly associated with disease, neither this particular capsular structure nor any other single factor appears to explain the pathogenicity of this species.

ACKNOWLEDGMENTS

This work was supported by the Danish Medical Research Council grant no. 12-1615. Majbritt Hauge received a stipend from Novo Nordisk A/S, Denmark.

We gratefully acknowledge Lars Sottrup-Jensen, Department of Molecular Biology, University of Aarhus, for performing the N-terminal amino acid sequencing.

REFERENCES

1. **Baker, C. J., and F. F. Barrett.** 1974. Group B streptococcal infection in infants: the importance of the various serotypes. J. Am. Med. Assoc. **230:** 1158–1160.
2. **Bevanger, L.** 1983. Ibc proteins as serotype markers of group B streptococci. Acta Path. Microbiol. Immunol. Scand. Sect. B **91:**231–234.
3. **Bohnsack, J. F., K. W. Mollison, A. M. Buko, J. C. Ashworth, and H. R. Hill.** 1991. Group B streptococci inactivate complement component C5a by enzymic cleavage at the C-terminus. Biochem. J. **273:**635–640.
4. **Caugant, D. A., L. O. Frøholm, K. Bøvre, E. Holten, C. E. Frasch, L. F. Mocca, W. D. Zollinger, and R. K. Selander.** 1986. Intercontinental spread of a genetically distinctive complex of clones of *Neisseria meningitidis* causing epidemic disease. Proc. Natl. Acad. Sci. USA **83:**4927–4931.
5. **Cleary, P. P., J. Handley, A. N. Suvorov, A. Podbielski, and P. Ferrieri.** 1992. Similarity between the group B and A streptococcal C5a peptidase genes. Infect. Immun. **60:**4239–4244.
6. **Dillon, H. C., Jr., S. Khare, and B. M. Gray.** 1987. Group B streptococcal carriage and disease: a 6-year prospective study. J. Pediatr. **110:**31–36.
7. **Farley, M. M., R. C. Harvey, T. Stull, J. D. Smith, A. Schuchat, J. D. Wenger, and D. S. Stephens.** 1993. A population-based assessment of invasive disease due to group B streptococcus in nonpregnant adults. N. Engl. J. Med. **328:** 1807–1811.
8. **Helmig, R., N. Uldbjerg, J. Boris, and M. Kilian.** 1993. Clonal analysis of *Streptococcus agalactiae* isolated from infants with neonatal sepsis or meningitis and their mothers and from healthy pregnant women. J. Infect. Dis. **168:**904–909.
9. **Hohwy, J., and M. Kilian.** 1995. Clonal diversity of the *Streptococcus mitis* biovar 1 population in the human oral cavity and pharynx. Oral Microbiol. Immunol. **10:**19–25.
10. **Jerlstrom, P. G., G. S. Chhatwal, and K. N. Timmis.** 1991. The IgA-binding β antigen of the c protein complex of group B streptococci: sequence determination of its gene and detection of two binding regions. Mol. Microbiol. **5:**843–849.
11. **Lancefield, R. C., M. McCarty, and W. N. Everly.** 1975. Multiple mouse-protective antibodies directed against group B streptococci. Special reference to antibodies effective against protein antigens. J. Exp. Med. **142:** 165–179.
12. **Lin, B., S. K. Hollingshead, J. E. Coligan, M. L. Egan, J. R. Baker, and D. G. Pritchard.** 1994. Cloning and expression of the gene for group B streptococcal hyaluronate lyase. J. Biol. Chem. **269:**30113–30116.
13. **Madoff, L. C., J. L. Michel, and D. L. Kasper.** 1991. A monoclonal antibody identifies a protective C-protein alpha-antigen epitope in group B streptococci. Infect. Immun. **59:**204–210.
14. **Maynard Smith, J.** 1995. Do bacteria have population genetics?, p. 1–12. *In* S. Baumberg, J. P. W. Young, E. M. H. Wellington, and J. R. Saunders (ed.), Population genetics of bacteria. Cambridge University Press, Cambridge.
15. **Michel, J. L., L. C. Madoff, D. E. Kling, D. L. Kasper, and F. M. Ausubel.** 1991. Cloned alpha and beta C-protein antigens of group B streptococci elicit protective immunity. Infect. Immun. **59:**2023–2028.
16. **Michel, J. L., L. C. Madoff, K. Olson, D. E. Kling, D. L. Kasper, and F. M. Ausubel.** 1992. Large, identical, tandem repeating units in the C protein alpha antigen gene, *bca*, of group B streptococci. Proc. Natl. Acad. Sci. USA **89:**10060–10064.
17. **Milligan, T. W., C. J. Baker, D. C. Straus, and S. J. Mattingly.** 1978. Association of elevated levels of extracellular neuraminidase with clinical isolates of type III group B streptococci. Infect. Immun. **21:**738–746.
18. **Musser, J. M., J. S. Kroll, E. E. Moxon, and R. K. Selander.** 1988. Evolutionary genetics of the encapsulated strains of *Haemophilus influenzae*. Proc. Natl. Acad. Sci. USA **85:**7758–7762.
19. **Musser, J. M., S. J. Mattingly, R. Quentin, A. Goudeau, and R. K. Selander.** 1989. Identification of a high-virulence clone of *Streptococcus agalactiae* (group B *Streptococcus*) causing invasive neonatal disease. Proc. Natl. Acad. Sci. USA **86:**4731–4735.
20. **Musser, J. M., P. M. Schlievert, A. W. Chow, P. Evans, B. N. Kreiswirth, V. T. Rosdahl, A. S. Naidu, W. Witte, and R. K. Selander.** 1990. A single clone of *Staphylococcus aureus* causes the majority of cases of toxic shock syndrome. Proc. Natl. Acad. Sci. USA **87:**225–229.
21. **Pritchard, D. G., and B. Lin.** 1993. Group B streptococcal neuraminidase is actually a hyaluronidase. Infect. Immun. **61:**3234–3239.
22. **Rubens, C. E., M. R. Wessels, L. M. Heggen, and D. L. Kasper.** 1987. Transposon mutagenesis of type III group B *Streptococcus*: correlation of capsule expression with virulence. Proc. Natl. Acad. Sci. USA **84:**7208–7212.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Smith, R. F., and N. P. Willett.** 1968. Rapid plate method for screening hyaluronidase and chondroitin sulfatase-producing microorganisms. Appl. Microbiol. **16:**1434–1436.
25. **Stålhammar-Carlemalm, M., Stenberg, L. and G. Lindahl.** 1993. Protein Rib: A novel group B streptococcal surface protein that confers protective immunity and is expressed by most strains causing invasive infections. J. Exp. Med. **177:**1593–1603.

Editor: V. A. Fischetti