Identification of a high-virulence clone of type III Streptococcus agalactiae (group B Streptococcus) causing invasive neonatal disease

(MLSC enzyme genotypes/genetic polymorphism/clones/meningitis)

James M. Musser*, Stephen J. Mattingly†, Roland Quentin‡, Alain Goudeau†, and Robert K. Selander*

*Department of Biology, Mueller Laboratory, Pennsylvania State University, University Park, PA 16802; †Department of Microbiology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78284; and ‡Laboratoires de Bacteriologie et Virologie, Centre Hospitalier Regional et Universitaire de Tours, Hopital Bretonneau, 2, Boulevard Tonnelle 37044, Tours Cedex, France

Contributed by Robert K. Selander, March 20, 1989

ABSTRACT

Chromosomal genotypes of 128 isolates of six serotypes (Ia, Ib, Ic, II, Ic/II, and III) of Streptococcus agalactiae (group B Streptococcus) recovered predominantly from human infants in the United States were characterized by an analysis of electrophoretically demonstrable allelic profiles at 11 metabolic enzyme loci. Nineteen distinctive electrophoretic types (ETs), representing multilocus clonal genotypes, were identified. Mean genetic diversity per locus among ETs of isolates of the same serotype was, on average, nearly equal to that in all 19 ETs. Cluster analysis of the ETs revealed two primary phylogenetic divisions at a genetic distance of 0.65. A single clone (ET 1) represented by 40 isolates expressing type III antigen formed division I. Division II was composed of 18 ETs in three major lineages diverging from one another at distances greater than 0.35 and included strains of all six antigenic classes. The type III organisms in division I produce more extracellular neuraminidase and apparently are more virulent than the type III strains in division II, which are related to strains of other serotypes that cause disease much less frequently. The existence of this unusually virulent clone accounts, in major part, for the high morbidity and mortality associated with infection by type III organisms.

Streptococcus agalactiae (group B Streptococcus) is the most common cause of invasive disease in neonates in the United States (1, 2) and a major etiological agent of economically important bovine mastitis in this country and elsewhere (3). Human infections by this Gram-positive bacterium account for approximately 11,000 cases of meningitis and septicemia each year in the United States and result in considerable morbidity and mortality (2, 4).

Group B organisms are differentiated from other members of the genus Streptococcus by Lancefield serological typing (5) or other methods of identifying group-specific antigens. They are subclassified into four main serotypes, designated as Ia, Ib, Ic, and III, on the basis of cell-wall polysaccharides, and two additional serotypes (Ic and Ic/II) are differentiated by a combination of surface carbohydrate and protein antigen typing (6–8). Although organisms of each of the six serotypes have been recovered from infants with serious infections, strains producing the type III antigen cause more than two-thirds of all neonatal disease (1, 2). It has been suggested that type III strains have high virulence potential and, perhaps, a meningeval or central nervous system tropism (9).

Except for limited data on plasmid content and structure (10), no information has been available concerning the nature and extent of genetic variation in natural populations of S. agalactiae or the genetic relationships of type III isolates recovered from various anatomic sites and from patients with various disease syndromes. Consequently, studies of virulence, the distribution of putative virulence factors, and epidemiology have been conducted almost entirely within a framework based on the serotyping of cell-surface antigens.

It has recently been demonstrated that virulence potential in many pathogenic bacterial species is nonrandomly distributed along phylogenetic lines (11, 12). The present study was conducted to estimate chromosomal genetic diversity and relationships among isolates of type III S. agalactiae as a first step toward identifying representative chromosomal genotypes for molecular and ecological studies of neonatal and maternal pathogenesis. We have discovered that strains synthesizing type III polysaccharide belong to two distantly related evolutionary lineages of clones that apparently differ in ability to invade the central nervous system. A single clone of unusually high virulence is responsible in major part for the high morbidity and mortality caused by type III organisms.

MATERIALS AND METHODS

Bacterial Isolates. A collection of 128 isolates representing five serotypes of S. agalactiae from the United States was examined. The type III isolates (n = 63) were recovered during episodes of clinical disease in neonates in Alaska, California, Colorado, Connecticut, Florida, Hawaii, Illinois, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, New York, Oklahoma, Tennessee, Texas, and Virginia, and from asymptomatic newborns in Texas. Of the 44 isolates of type III from invasive episodes, 24 were cultured from blood, 18 were recovered from cerebrospinal fluid (CSF), and 1 each was grown from joint fluid and pulmonary secretions collected by tracheal aspiration. The 19 isolates from asymptomatic neonates were collected from the umbilicus (16 isolates), rectum (2 isolates), and throat (1 isolate). The sample also included isolates of types I (n = 8), Ib (n = 16), Ic (n = 14), Ic/II (n = 6), and II (n = 11) collected during disease episodes in 15 states, primarily from blood or cerebrospinal fluid. In addition, nine isolates from cases of bovine mastitis in France were examined: types Ia (n = 2), Ib (n = 2), Ic (n = 4), and II (n = 1).

The actual frequencies of recovery of the various serotypes from invasive episodes are not accurately reflected by the numbers of isolates in the collection studied. Non-type-III isolates were greatly overrepresented among the invasive isolates, strictly for purposes of genetic comparison. The human isolates were obtained from collections previously characterized for neuraminidase production and other phenotypes (13–22). No non-type-III carrier isolates were analyzed.

Many of the type III isolates obtained from humans were coded by S.J.M. for blind analysis in the laboratory of R.K.S.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ET, electrophoretic type.
Electrophoresis of Enzymes. Isolates were grown overnight in 200 ml of brain/heart infusion broth (Difco) at 37°C and harvested by centrifugation at 6000 × g for 10 min at 4°C. After suspension in 1.5 ml of 50 mM Tris·HCl buffer containing 5 mM EDTA (pH 7.5), the bacteria were sonicated with a model 200 Sonifer-Cell Disrupter (Branson) equipped with a microtip for 1 min at 50% pulse, with dry ice/methanol cooling. The clear supernatant (lysate) was obtained by centrifugation at 20,000 × g for 20 min at 4°C and stored at −70°C.

Lysates were electrophoresed on starch gels and selectively stained for 11 metabolic enzymes, according to methods described by Selander et al. (23). The enzymes were nucleoside phosphorylase (NSP), phosphoglucose isomerase (PGI), carboxymaltose kinase (CAK), glyceraldehyde-3-phosphate dehydrogenase (G3P), glutamate dehydrogenase (GLD), mutase-2 (HEX), phosphoglucomutase-1 (PM1), phosphoglucomutase-2 (PM2), phenylalanyl-leucine peptidase (PLP), leucyl-glycyl-glycine peptidase (LGG), and adenylyl kinase (ADK). Assays of the lysates for 28 other enzymes in nine different electrophoretic buffer systems did not reveal additional consistently scorable activities.

Distinctive mobility variants (electromorphs) of each enzyme, numbered in order of decreasing rate of anodal migration, were equated with alleles at the corresponding structural gene locus. Because almost all isolates showed activity for all enzymes, the corresponding structural genes are presumed to be located on the chromosome rather than on plasmids. The occasional absence of activity for an enzyme was attributed to the presence of a null allele, designated as zero.

Each isolate was characterized by its combination of alleles at the 11 enzyme loci, and distinctive combinations of electromorphs, corresponding to unique multilocus genotypes, were designated as electrophoretic types (ETs) (23).

Serotyping. Strains from humans were serotyped previously by the method of Lancefield (6, 7). Bovine isolates were coded by J.M.M. for blind serotyping in the laboratory of S.J.M.

Growth of Type III Isolates in a Medium Containing Phosphate. The effect of 200 mM phosphate on growth of type III isolates was determined in a chemically defined medium (FMc; ref. 24), as described elsewhere (22). Briefly, isolates were grown to stationary phase in FMC medium containing 65 mM sodium phosphate, and aliquots of these cultures were inoculated into 10 ml of FMC medium containing 200 mM sodium phosphate. Isolates were scored for inhibition of growth (lag time >5–6 hr) or no inhibition (lack of significant lag phase) (22).

Statistical Analysis. Genetic diversity at an enzyme locus (H) among ETs was calculated from allele frequencies as \( H = 1 - \sum x^2 / n^2 \), where \( x_i \) is the frequency of the \( i \)-th allele and \( n \) is the number of ETs; mean diversity per locus (\( H \)) is the arithmetic average of \( H \) values over all loci (23). Genetic distance between pairs of ETs was expressed as the proportion of enzyme loci at which different alleles were represented (mismatches), and clustering of ETs was performed from a matrix of genetic distances by the average linkage method (23).

RESULTS

Genetic and Genotypic Diversity in Relation to Serotype. In the collection of 128 isolates of S. agalactiae examined, 10 of the 11 enzyme loci assayed were polymorphic for alleles encoding electrophoretically distinguishable variant polypeptides, and one locus (phosphoglucone isomerase) was monomorphic (Table 1). A total of 19 distinctive multilocus genotypes (ETs) was identified, among which mean genetic diversity per locus (\( H \)) was 0.304 (Table 1). Seven ETs were represented by single isolates and 12 ETs were represented by multiple isolates (range, 2 to 40).

Strains of type III S. agalactiae showed the greatest diversity (\( H = 0.455 \)), and those of type Ia were the least variable (\( H = 0.182 \)) (Table 2). The mean numbers of alleles per locus among isolates assigned to the six cell-wall antigen groups were 1.3, 1.8, 2.1, 1.5, 1.4, and 1.9 for types Ia, Ib, Ic, II, Ic/II, and III, respectively.

All serotypes were represented by strains of a variety of multilocus enzyme genotypes, with an average of 5.0 ETs per serotype. The number of ETs per serotype ranged from 3 for type Ia to 9 for type III (Table 3). On average, 98% of the total genetic diversity among all 19 ETs was represented within individual serotypes.

Relationships Among Multilocus Genotypes. The dendrogram in Fig. 1 summarizes estimates of the genetic relationships of the 19 ETs, based on allelic variation at the 11 enzyme loci. The smallest genetic distance (0.09) between ETs corresponds to a single-locus difference.

At a genetic distance of 0.35, there are four lineages of isolates: ET 1 and three groups, labeled A, B, and C, that were represented by 2, 13, and 3 ETs, respectively. ET 1 is separated from clusters A, B, and C at a genetic distance of 0.65, which means that ETs in these two primary divisions of the dendrogram (I and II) differ, on average, at 7 of the 11 loci assayed.

Primary division I is composed of a single ET (ET 1) with 40 isolates, all of which express type III cell-wall polysaccharide.

Cluster A, in division II, contains ETs of three serotypes (Ib, Ic, and III); and all isolates of ETs in this cluster were recovered from human sources.

Cluster B, which is separated from cluster A at a genetic distance of 0.38, consists of 13 ETs represented by isolates of all six serotypes. Isolates of ETs 11 and 12 were recovered...
from cases of human disease and bovine mastitis; and ETs 16 and 17 were each represented by a single isolate from bovine mastitis. Eight of the 13 ETs in cluster B were represented by isolates of two or three antigenic types.

Cluster C diverges from clusters A and B at a genetic distance of 0.48 and includes two ETs, each with two isolates recovered from bovine mastitis.

In summary, there are two strongly differentiated lineages of group B streptococci: division I, with a single ET represented entirely by human isolates expressing type III antigen, and division II, with 18 ETs represented by isolates recovered from human and bovine sources and producing each of the six cell-wall antigenic types.

Lack of Effect of Laboratory Passage on Electromorph Profile. Strain GBS 122 (type III) and its high-phosphate-adapted spontaneous derivative VP6, which was obtained by serial laboratory passage (13), had identical electromorph mobilities for each of the 11 enzymes assayed (data not shown).

Neuraminidase Production. All type III isolates in division I produce high levels of extracellular neuraminidase, with an average activity of 249 nmol/min per mg of cell dry weight (range, 159–355) (14–16). Conversely, all type III isolates of ETs in primary division II are nonproducers of extracellular neuraminidase (less than 10 nmol/min per mg of cell dry weight) (14–16). Most non-type-III strains of ETs in division II are also low producers, synthesizing extracellular neuraminidase with activity less than 110 nmol/min per mg of cell dry weight. Prominent exceptions are three strains of serotypes Ib and Ic assigned to ET 2 and ET 3, in cluster A, which produce, on average, activity of 287 nmol/min per mg of cell dry weight.

Growth of Type III Isolates in FMC Phosphate Medium. Growth of all of the 40 type III isolates (both invasive and noninvasive) of ET 1 in division I was inhibited in FMC medium containing 200 mM phosphate. In contrast, for only 12 of the 23 (52%) type III isolates representing ETs in division II was growth inhibited by this medium. For type III organisms in division II, there was no apparent difference between invasive and carrier strains in growth response in FMC medium containing 200 mM phosphate (Table 3).

### DISCUSSION

#### Rationale for the Analysis.
Because serological methods have been extensively employed to identify strains of *S. agalactiae* and infer their affinities in clinical, epidemiological, physiological, systematic, and other research, it is important to ascertain the degree to which cell-wall polysaccharide types and various other phenotypic characters reflect the genetic structure of populations and the relationships of strains. For the oral streptococci, *Legionella* spp., *Bordetella* spp., *Helicobacter* spp., and other higher organisms, estimates of genetic distance between pairs of strains based on electrophoresis of 15–30 metabolic enzymes have been shown to be strongly correlated (r = 0.8–0.9) with estimates of nucleotide sequence divergence obtained by hybridization of total cellular DNA (11). Although comparable data are not available for estimates of genetic relatedness based on electrophoresis of a panel of only 11 enzymes, it is noteworthy that an analysis of 12 enzymes distinguished the four nominal species of *Shigella* and demonstrated their close genetic affinities with *Escherichia coli* (25). Additionally, Olyhoek et al. (26) found that the major clonal lineages of serotype *A Neisseria meningitidis* could be identified by analysis of electrophoretically demonstrable allelic variation at only 6 loci. Therefore, we have reason to believe that the 11 enzyme loci assayed in this study provide a reasonable basis for estimating both genetic diversity in populations and overall genetic relationships among strains. Examination of allelic variation at more enzyme loci undoubtedly will add more "twigs" to the dendrogram in Fig. 1 but probably will not reveal additional major lineages.

For purposes of reference and discussion, ETs are considered to mark clones.

#### Genetic Diversity Within Serotypes.
For a variety of pathogenic bacterial species, extensive genetic diversity is often
shown among isolates classified on the basis of one or a few surface antigens. For example, in *Haemophilus influenzae* the ability to express serotype b capsule polysaccharide (polyribitolribosylphosphate) was associated with 182 distinct multilocus enzyme genotypes identified in a sample of 1975 type b isolates (27, 28), and the *N*-acylneuraminic acid K1 capsule of *E. coli* was found to be synthesized by strains of 29 genotypes (29). In the present study, multilocus enzyme electrophoresis revealed extensive genetic diversity among isolates of each of the six polysaccharide and protein antigenic types of group B *Streptococcus* and demonstrated that serotypic identity frequently does not indicate close genetic relatedness of strains. Type III isolates are especially heterogeneous and belong to three divergent phylogenetic lineages. Types Ib and Ic isolates occur in all three lineages (A, B, and C) of division II, whereas, types Ia, II, and Ic/II are confined to isolates of ETs in one of the major evolutionary lineages (cluster B) in division II (Fig. 1).

Nine of the 12 ETs with multiple isolates were each represented by two or three different antigenic types, a situation similar to that in *E. coli* (29, 30), *N. meningitidis* (31), and *Haemophilus pleurapneumoniae* (32), in which many ETs are represented by isolates of two or more polysaccharide capsule serotypes.

**Pathogenicity of Type III Strains in Relation to Clonal Structure of Populations.** While our analysis demonstrated that a large variety of genotypes expressing type III polysaccharide antigen may cause invasive disease, it also revealed that isolates of only two multilocus genotypes (ET 1 and ET 12) are responsible for much of the serious type III disease in the United States. In our sample, 95% of the 44 type III isolates recovered from normally sterile body fluids were assigned to these two ETs. Moreover, the strongly virulent type III strains responsible for most disease cases are all members of a single clone (ET 12) that is strongly differentiated from clones of the other two lineages of organisms synthesizing type III antigen and from all other group B streptococci representing six serotypic classes.

The analysis of variation in phenotypic characters within the phylogenetic framework furnished by multilocus enzyme analysis can contribute to an understanding of the evolutionary development of those properties that confer virulence to bacterial clones causing septicemia, meningitis, or other diseases. In this respect, studies of encapsulated *H. influenzae* and *E. coli* have been especially fruitful (27, 28, 33–35). Under clonal population structure, episodes of periodic selection and stochastic extinction generate nonrandom association of phenotypically determinable characters in addition to any associations that may be maintained by epistatic selection (11). Consequently, analyses of adequate sets of different types of polymorphic characters may be expected to yield generally similar patterns of inferred genetic relationships among strains, particularly when the genetic determinants are chromosomal (versus plasmid) in location.

The discovery that unusually virulent isolates of *S. agalactiae* belong to one type III clone suggests that genes mediating pathogenesis are in linkage disequilibrium (nonrandom association) with genes of the multilocus enzyme genotype. Three lines of evidence support this hypothesis.

First, seven of the strains we studied were examined previously for release of type III antigen from the cell surface, a phenotypic trait apparently associated with increased virulence (17, 20). All three of these previously studied isolates that we assigned to ET 1 released higher levels of high molecular weight type III antigen than did the four other isolates, which were of ETs in primary division II.

Second, compared to division II organisms, growth of division I isolates was more frequently inhibited in FMC medium containing 200 mM phosphate (*P* < 0.001). On the basis of an earlier analysis of 15 invasive and carrier type III isolates, one of us (S.J.M.) hypothesized that the growth response in this medium reflected fundamental differences between virulent and avirulent (carrier) isolates (22). However, our results demonstrate that the growth response in FMC phosphate medium varies not as a function of clinical source but with the multilocus enzyme genotype (Table 3). Hence, we attribute the differences reported earlier for invasive versus carrier isolates to inadequate sample size; they clearly do not reflect a consistent phenotypic separation of strains from disease and carriage sources.

Third, the level and composition of extracellular products expressed by strains assigned to the two primary divisions differ. The extracellular product secreted by type III strains (sialidase) and proteases has been studied by a number of investigators because these enzymes may be virulence factors (13–16, 36). Although neuraminidase is produced by strains of all serotypes (14–16), type III strains from infants with invasive infections more frequently produce high levels of this enzyme than do type III strains recovered from asymptptomatically colonized infants (15, 16). All isolates of ET 1 express high levels of the enzyme, whereas most strains of ETs in division II synthesize little or no extracellular neuraminidase. Similarly, type III isolates in division I elaborate, on average, approximately 6-fold more extracellular protease (13), synthesize larger amounts of extracellular type-specific antigen (18), and are more lethal for mice (18) than type III strains in division II. Our results support the interpretation that these phenotypes are causally related to the high virulence potential of isolates in division I, but additional studies, such as those employing insertionally inactivated mutants (37, 38), will be required to rigorously test this hypothesis.

In summary, for type III group B streptococci, there is a strong relationship between subspecific genetic structure and the production of several extracellular substances implicated in virulence. Examination of additional phenotypic characteristics and ecological properties should reveal other differences between organisms in division I and division II.

**Analogies with Encapsulated *H. influenzae*.** There are several striking analogies between our findings for *S. agalactiae* and results reported for encapsulated strains of *H. influenzae*. In *H. influenzae*, both serotype a and serotype b capsule polysaccharides are expressed by strains of two highly divergent phylogenetic divisions, and the chromosomal genotypes of the two evolutionary lineages differ significantly in virulence potential (27, 28). Similarly, in *S. agalactiae*, type III isolates of only one of the two primary divisions are commonly recovered from invasive infections. For encapsulated *H. influenzae*, there is strong evidence that horizontal transfer and recombination of a segment of DNA coding for the b-specific polysaccharide accounts for the expression of type b antigen in two distantly related phylogenetic lines (28). By analogy, it is probable that the occurrence of serotype-III expression in widely divergent clonal lineages of *S. agalactiae* also reflects one or more episodes of horizontal transfer and recombination of genes necessary for polysaccharide antigen synthesis. It is also possible that transfer of one or more genes mediating other virulence factors has produced the highly virulent type III clone in division I. The absence of allelic variation among division I isolates, as revealed by our analysis, strongly suggests that ET 1 is a recently newly evolved clone. In this regard, it may be significant that the incidence of group B streptococcal neonatal disease in the United States apparently has increased since the 1960s (2, 39).

A corollary of the hypothesis of recent origin of ET 1 is that the clonal composition of populations will vary geographically, because the spread of a new clone over wide distances cannot occur instantaneously. Thus, it is of interest that there is considerable regional and continental variation in the incidence of neonatal disease caused by this pathogen (40).
Concluding Comments. This study adds to a growing body of information demonstrating that knowledge of subspecific structure and evolutionary relationships provides important insights relating to bacterial pathogenesis. The extensive genetic diversity among type III isolates of S. agalactiae has important implications for studies of the basic biology and epidemiology of this pathogen. Moreover, the discovery that isolates of a single clone account for much of the morbidity and mortality caused by type III organisms should stimulate further comparative physiological and genetic studies employing the evolutionary genetic framework constructed by our analysis.

We thank Sheila A. Plock and Krista A. Fletcher for technical assistance. This research was supported by Grants AI-24144 and AI-22380 from the National Institutes of Health to R.K.S. and S.J.M., respectively.