Brief Definitive Report

Characterization and Clonal Distribution of Four Alleles of the *speA* Gene Encoding Pyrogenic Exotoxin A (Scarlet Fever Toxin) in *Streptococcus pyogenes*

By Kimberlyn Nelson,* Patrick M. Schlievert,‡ Robert K. Selander,* and James M. Musser*§

*From the *Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania 16802; the ‡Department of Microbiology, University of Minnesota, Minneapolis, Minnesota 55455; and the §Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania 19104*

Summary

Streptococcus pyogenes strains producing pyrogenic exotoxin A (scarlet fever toxin) have recently caused episodes of streptococcal toxic-shock–like syndrome (TSLS). We exploited knowledge of genetic diversity and relationships among exotoxin A–producing patient strains provided by multilocus enzyme electrophoresis to select strains for comparative sequencing of toxin genes. Our analysis identified four alleles of *speA* in natural populations, one of which (*speA1*) occurs in many distinct clonal lineages and is probably old. Two other alleles (*speA2* and *speA3*), characterized solely by single amino acid substitutions, were each identified in single clones that together have caused the majority of TSLS episodes. It is unlikely that these alleles have had a long association with *S. pyogenes* clones. A fourth allele (*speA4*) also is present in a single phylogenetic lineage and is 9% divergent from the other three toxin alleles. An absence of synonomous (silent) nucleotide changes in *speA2* and *speA3* is unusual and suggests that the allelic variation is not selectively neutral, which implies that the toxins are not functionally equivalent. These results may be important in helping to understand the recent increase in frequency and severity of disease caused by *S. pyogenes*.

Unusually severe infections caused by *Streptococcus pyogenes* have been reported with increasing frequency in the United States in recent years (1, 2). Many patients have soft tissue infections, such as cellulitis or fasciitis, but some individuals have experienced deeper infections, including osteomyelitis, myometritis, sepsis, and meningitis. This has led to the description of a streptococcal toxic-shock–like syndrome (TSLS).

Severe invasive *S. pyogenes* infections have also been reported in Canada (3), the United Kingdom (4), Denmark (5), Sweden (6), and Norway (7). There is evidence that most or all of the increase in frequency of severe episodes in the United Kingdom, certain other European countries, and the United States occurred in the 1980s (4, 6, 8), and has been caused by strains that express M protein serotypes 1 or 3 (4, 8, 9). Taken together, the data indicate that the frequency of unusually severe infections caused by strains of *S. pyogenes* expressing only a few of the >80 recognized M protein serotypes has increased on an intercontinental scale in recent years.

A recent analysis by multilocus enzyme electrophoresis of 103 epidemiologically unassociated *S. pyogenes* isolates cultured from patients with severe infections in the United States revealed that nearly half the disease episodes and more than two-thirds of the cases of TSLS were caused by strains of two related clonal genotypes (10). The great majority of isolates of these two clones, which have been designated as electrophoretic types 1 and 2 (ET 1 and ET 2), produced M proteins of serotype 1 and 3, respectively (10). The same study also confirmed an earlier report (11) of a relatively high frequency of expression of streptococcal pyrogenic exotoxin A (SPE A) (scarlet fever toxin) among isolates from TSLS patients. The demonstration of this association both within clones and among distantly related clones supported the hypothesis that SPE A, which is encoded by the *speA* gene, is a causal factor in the pathogenesis of TSLS. This observation was particularly important because none of 80 *S. pyogenes* strains isolated from a variety of clinical sources in the United States between 1976 and 1986 produced SPE A (1), whereas many isolates recovered earlier in this century from patients with severe scarlet fever expressed this exotoxin.

SPE A, which is bacteriophage encoded, is a member of a family of "superantigen" toxins synthesized by *S. pyogenes*

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/91/11/1271/04 $2.00
Volume 174 November 1991 1271–1274

and *Staphylococcus aureus* that exhibit significant levels of amino acid homology and share functional activities, including pyrogenicity, enhancement of susceptibility to endotoxic shock, and enhancement of delayed hypersensitivity to induced skin rashes (12). These toxins also stimulate proliferation of murine and human lymphocytes bearing particular TCR VB gene products and induce clonal anergy in T cells, either by clonal deletion or functional inactivation (12).

Because epidemiologic and toxicologic evidence indicates that expression of SPE A is a causal factor in the etiology of many cases of TSLS (10), scarlet fever (13), and rheumatic fever (14), we undertook the present project to study the molecular population genetics of the *speA* gene in greater detail. Here we report the nucleotide sequences and clonal distribution of four alleles of *speA* in strains of *S. pyogenes* recovered from patients with severe invasive diseases.

Materials and Methods

Bacterial Isolates. A sample of 10 strains of *S. pyogenes* (Table 1) was drawn from a collection of 108 isolates recovered from patients with severe streptococcal diseases in the United States in the period 1986–1990 (10). The strains are representative of seven distinctive clones that carry the *speA* gene (10). In addition, we studied nine isolates recovered in the 1980s from patients living in Yugoslavia (n = 2), the United Kingdom (n = 2), France (n = 2), Germany (n = 2), and New Zealand (n = 1); and a patient isolate recovered in Germany in the 1970s.

All but three of the strains (MGAS 251, MGAS 256, and MGAS 262) expressed levels of SPE A detectable by Ouchterlony immunodiffusion with hyperimmune rabbit antiserum prepared against purified exotoxin A (11).

Sequencing of speA. The *speA* gene of each of the 20 isolates was amplified by PCR, with synthetic oligonucleotides, as described previously (10). The portion sequenced (708 bp) represents 94% of *speA* and encodes virtually the entire mature protein.

Estimating Genetic Relationships among Clones. Methods of estimating genetic relationships among *S. pyogenes* clones by multilocus enzyme electrophoresis have been described (10). The ET designations are cognate with those used previously (10).

Results and Discussion

Clonal Relationships among Multilocus Enzyme Genotypes. The electrophoretic types of the 10 United States isolates were presented earlier (10). The multilocus enzyme allele profiles of strains MGAS 480 (Yugoslavia), MGAS 492 (United Kingdom), and MGAS 496 (Germany) were identical to that recorded previously for isolates of ET 1, and the allele profiles of MGAS 485 (Yugoslavia), MGAS 491 (United Kingdom), and MGAS 495 and MGAS 624 (Germany) were identical to that of ET 2 (10). The profiles of MGAS 493 and MGAS 494 (France) were identical and were designated as ET 14a because they differ only slightly from the profile of ET 14 (10). MGAS 500, which was recovered from a patient in New Zealand with scarlet fever, was assigned to ET 3a (Table 1). Estimates of the genetic relationships among the isolates are shown in Fig. 1.

Alleles speA1, speA2, and speA3. The sequences of the PCR-amplified 708-bp segment of the *speA* gene from five isolates (MGAS 500, MGAS 327, MGAS 165, MGAS 167, and MGAS 156), representing ET 3a, ET 4, ET 5, ET 16, and ET 20, respectively, were identical to a published sequence (15) of *speA* and will be referred to as the *speA1* allele (Table 1). This allele also occurred in strains MGAS 493 and MGAS 494, representing ET 14a. The sequences of all seven isolates of ET 1 (four of which expressed serotype M1 protein) were invariant and differed from the published sequence by a single transition (G→A) at nucleotide site 328, which results in a Gly → Ser substitution at amino acid position 110 (*speA2*). Similarly, the *speA* sequences of all five isolates of ET 2 were identical and differed from the published sequence by one base change (G→A) at nucleotide 316, which specifies a Val→Ile replacement at amino acid 106 (*speA3*).

Interestingly, these mutations occur in a segment consisting of five amino acids (residues 86–90 in the numbering system used in reference 12) that are highly conserved in the aligned sequences of staphylococcal enterotoxin serotype A (SEA), staphylococcal enterotoxin serotype B (SEB), SEC_1, SEC_3, SED, SEE, and streptococcal pyrogenic exotoxin C. For example, the valine residue at amino acid 106 (15) in the deduced sequence of the SPE A1 protein is fully conserved in these seven exotoxins, and a glycine residue occurs at amino acid 110 (15) in all sequences except that of SED (12). It is also noteworthy that the segment of SPE A containing the variations we have detected is located immediately adjacent to a region containing cysteine residues that are involved in formation of a disulfide loop that is believed to be required for mitogenicity of SPE A and other bacterial superantigens (16). Grossman et al. (16) have suggested that an intact disulfide linkage is required for interaction of toxin with the TCR, but not with MHC class II molecules. It is, therefore, reasonable to conclude that the three toxin variants may display qualitative or quantitative heterogeneity in one or more of the functions ascribed to SPE A.

Allele speA4. In striking contrast to the relative homogeneity of the *speA* sequences in the other 19 isolates, the sequence of the single isolate of ET 26 (MGAS 262) showed a large number of nucleotide substitutions. Compared with *speA1*, there were 67 nucleotide differences (91% nucleotide identity) that produce 26 amino acid changes (89% amino acid similarity) (data not shown). Although the nucleotide changes are distributed in a seemingly random pattern throughout the structural gene, almost all of them occur in regions encoding amino acids that are not highly conserved when SPE A is aligned with the SEA through SEE, and SPE C (12).

Distribution of speA Alleles among Clonal Lineages. In the small sample of strains examined, all seven strains of ET 1 (four of which expressed serotype M1 protein) had the *speA2* allele, and all five ET 2 strains carried *speA3* and expressed protein serotype M3 (Table 1). In contrast, the *speA1* allele was identified in strains of six distinct clonal lineages, some of which are highly divergent in overall chromosomal genomic character (Fig. 1). The *speA4* allele occurred in a single strain that was the most divergent multilocus enzyme genotype (ET 26) in a large sample of *speA*-containing strains (10).

Table 1. *Properties of 20 Isolates of S. pyogenes Representing Nine ETs*

ET	Strain	M protein serotype	*speA* allele	Locality
1	250	NS*	*speA2*	California
1	251	NS	*speA2*	California
1	256	NS	*speA2*	California
1	285	M1	*speA2*	Colorado
1	480	M1	*speA2*	Yugoslavia
1	492	M1	*speA2*	United Kingdom
1	496	M1	*speA2*	Germany
2	158	M3	*speA3*	Nebraska
2	485	M3	*speA3*	Yugoslavia
2	491	M3	*speA3*	United Kingdom
2	495	M3	*speA3*	Germany
2	624	M3	*speA3*	Germany
3a	500	M49	*speA1*	New Zealand
4	327	M2	*speA1*	Arizona
5	165	NT‡	*speA1*	Minnesota
14a	493	M3	*speA1*	France
14a	494	M3	*speA1*	France
16	167	M66	*speA1*	Texas
20	156	M18	*speA1*	Nebraska
26	262	NT	*speA4*	California

* NS, not studied.
‡ NT, nontypable.

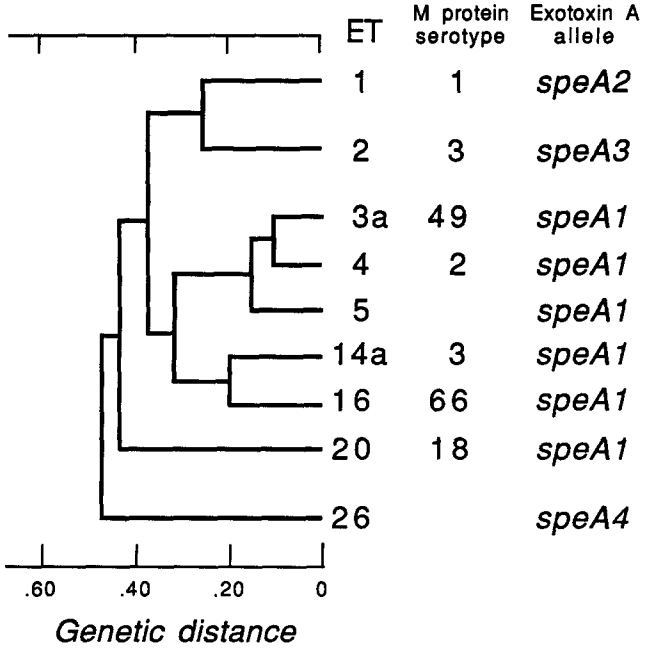

Figure 1. Dendrogram of 9 *S. pyogenes* ETs. Each isolate was characterized by its combination of alleles at 12 enzyme loci, and distinctive multilocus enzyme genotypes were designated as ETs. The dendrogram was generated from a matrix of genetic distances between pairs of ETs by the average linkage method (10). Strains are designated by the ET numbers given in Table 1. The DNA sequence data for *speA2, speA3*, and *speA4* are available from EMBL/GenBank/DDBJ under accession numbers X61554–X61573.

We have previously hypothesized (10) that the presence of the identical *speA* sequence in strains with highly differentiated chromosomal genotypes is a consequence of bacteriophage-mediated horizontal transfer and recombination. The occurrence of the *speA1* allele on several diverse chromosomal backgrounds and the restriction, in our sample, of *speA2* and *speA3* to single clonal lineages can be interpreted as evidence that the latter two alleles are evolutionarily younger than *speA1*. An alternative, but not mutually exclusive, hypothesis is that the *speA1* allele has a greater dispersion velocity in natural populations. Thus, for example, this allele might be laterally transferred at a relatively high rate owing to dis-

tinctive biological characteristics of the bacteriophage that carries it.

Because only a relatively small sample of isolates has thus far been analyzed for allelic variation in *speA*, our observations should not be interpreted to mean that expression of the M1 protein serotype is confined to isolates of ET 1, or that all toxinogenic M1 isolates carry the *speA2* allele. In this regard, we note that allelic heterogeneity in metabolic enzyme genes has been observed among isolates expressing the same M protein serotype (see Fig. 1; and J. M. Musser, unpublished data). Therefore, additional large-scale studies of allelic variation in *speA* among toxinogenic strains of known multilocus enzyme genotype will be required to fully delineate the strength of clone-toxin allele-M protein serotype associations present in natural populations of *S. pyogenes*. These studies are in progress.

Inasmuch as synonomous (silent) nucleotide changes outnumbered nonsynonomous changes by a factor of ~20 in a sample of 25 different bacterial genes (17), it is a curious circumstance that the only mutations distinguishing the *speA2* and the *speA3* alleles from the common *speA1* allele result in amino acid replacements. It is also noteworthy that both mutations occur in the same gene segment, adjacent to a region that encodes a group of amino acids that is highly conserved in other bacterial superantigens (12). In addition, the ratio (1.3:1) of synonomous to nonsynonomous changes in *speA4* also is unusually low. Taken together, these observations suggest that there are functional correlates of the allelic variation and that the alleles have been subject to natural selection. This hypothesis is currently being tested by comparative analysis of purified toxin functions in murine and human systems and in animal models of streptococcal pathogenesis. Studies of this type may be particularly important in helping to explain temporal variation in the frequency and/or severity of disease caused by this pathogen.

We thank J. M. Gearhart, M. Hughes, and S. A. Plock for technical assistance. Dr. Joseph J. Ferretti generously supplied three of the strains examined.

This research was supported by grants AI-24332 (to R. K. Selander) and HL-36611 (to P. M. Schlievert) from the National Institutes of Health.

Address correspondence to James M. Musser, Department of Pathology, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030.

Received for publication 3 June 1991 and in revised form 24 July 1991.

References

1. Stevens, D.L., M.H. Tanner, J. Winship, R. Swarts, K.M. Ries, P.M. Schlievert, and E. Kaplan. 1989. Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. *N. Engl. J. Med.* 321:1
2. Givner, L.B., J.S. Abramson, and B. Wasilauskas. 1991. Apparent increase in the incidence of invasive group A beta-hemolytic streptococcal disease in children. *J. Pediatr.* 118:341.
3. Burrows, K.J., S.A. Halperin, M. Swift, and R. Bortolussi. 1990. Changing pattern of group A streptococcal (GAS) bacteremia in children. *Interscience Conference on Antimicrobial Agents and Chemotherapy.* 651. (Abstr.)
4. Gaworzewska, E., and G. Colman. 1988. Changes in the pattern of infection caused by *Streptococcus pyogenes*. *Epidemiol. Infect.* 100:257.
5. Henrichsen, J., S. Hoffman, N.P. Rothgardt, and T. Ronne. 1988. Infection with group A streptococci. Statens Seruminstitut, Copenhagen. *Epi-Nyt.* 50.
6. Stromberg, A., V. Romanus, and L.G. Burman. 1991. Outbreak of group A streptococcal bacteremia in Sweden: an epidemiologic and clinical study. *J. Infect. Dis.* 164:595.
7. Martin, P.R., and E.A. Hoiby. 1990. Streptococcal serogroup A epidemic in Norway 1987-1988. *Scand. J. Infect. Dis.* 22:421.
8. Schwartz, B., R. Facklam, and R. Breiman. 1990. The changing epidemiology of group A streptococcal infections in the U.S.: association with changes in serotype. *Interscience Conference on Antimicrobial Agents and Chemotherapy.* 10. (Abstr.)
9. Kohler, W. 1990. Streptococcal toxic shock syndrome. *Zbl. Bakt.* 272:257.
10. Musser, J.M., A.R. Hauser, M.H. Kim, P.M. Schlievert, K. Nelson, and R.K. Selander. 1991. *Streptococcus pyogenes* causing toxic-shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. *Proc. Natl. Acad. Sci. USA.* 88:2668.
11. Lee, P.K., and P.M. Schlievert. 1989. Quantification and toxicity of group A streptococcal pyrogenic exotoxins in an animal model of toxic shock syndrome-like illness. *J. Clin. Microbiol.* 27:1890.
12. Marrack, P., and J. Kappler. 1990. The staphylococcal enterotoxins and their relatives. *Science (Wash. DC).* 248:705.
13. Yu, C.-E., and J.J. Ferretti. 1989. Molecular epidemiologic analysis of the type A streptococcal exotoxin (erythrogenic toxin) gene (*speA*) in clinical *Streptococcus pyogenes* strains. *Infect. Immun.* 57:3715.
14. Gray, E.D., L.W. Wannamaker, E.M. Ayoub, A. El Kholy, and A.H. Abdin. 1981. Cellular immune responses to extracellular streptococcal products in rheumatic heart disease. *J. Clin. Invest.* 68:665.
15. Weeks, C.R., and J.J. Ferretti. 1986. Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin) gene from *Streptococcus pyogenes* bacteriophage T12. *Infect. Immun.* 52:144.
16. Grossman, D., R.G. Cook, J.T. Sparrow, J.A. Mollick, and R.R. Rich. 1990. Dissociation of the stimulatory activities of staphylococcal enterotoxins for T cells and monocytes. *J. Exp. Med.* 172:1831.
17. Ochman, H., and A.C. Wilson. 1987. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* 26:74.