Group A Streptococcal Pharyngitis Serotype Surveillance in North America, 2000–2002

Stanford T. Shulman,[1] Robert R. Tanz,[2] William Kabat,[1] Kathleen Kabat,[1] Emily Cederlund,[1] Devendra Patel,[1] Zhongya Li,[3] Varja Sakota,[3] James B. Dale,[4,5] Bernard Beall,[3] and the US Streptococcal Pharyngitis Surveillance Group[a]

Divisions of [1]Infectious Diseases and [2]General Academic Pediatrics, Children's Memorial Hospital, Department of Pediatrics, Northwestern University, Feinberg School of Medicine, Chicago, Illinois; [3]Centers for Disease Control and Prevention, Atlanta, Georgia; and [4]Veterans Affairs Medical Center and [5]University of Tennessee Health Science Center, Memphis, Tennessee

Geographic and interseasonal heterogeneity of pharyngeal group A streptococcal (GAS) genotypes (*emm* types) is poorly characterized. We evaluated *emm* type and subtype distribution among pediatric pharyngitis isolates obtained from 9 sites in the United States during 2000–2001 (year 1) and from 10 sites in the United States and 1 site in Canada during 2001–2002 (year 2). The 7 predominant types were the same in both years, although their order changed. *emm* 12, 1, and 28 accounted for 49.2% of year 1 isolates, and *emm* 1, 12, and 4 accounted for 47.1% of year 2 isolates; 6 types accounted for 72.1% in year 1 and 69.4% in year 2. From year 1 to year 2, the proportions of *emm* 12 and 28 decreased and *emm* 1 and 6 increased. Striking intersite and interseasonal variations in the distribution of predominant *emm* types were observed. We conclude that the most-predominant GAS genotypes were similar for each year despite fluctuations, that intersite and intrasite variations in the distribution of *emm* types were apparent, and that *emm* type surveillance is needed as M protein vaccine development proceeds.

Since the development of group A streptococcal (GAS) M typing by Swift et al. in 1943 [1], several studies have evaluated the distribution of various M types among clinical isolates. In 1957, Quinn et al. [2] studied hemolytic streptococci in school children, and Hope-Simpson [3] characterized pharyngeal *Streptococcus pyogenes* in Gloucestershire, United Kingdom, from 1962 through 1975. These studies demonstrated the predominance of certain M and T typable strains (notably, M1, 6, 12, and 28 and T3, 4, and 12). Stollerman [4] reported that 81% of M-typable acute pharyngitis isolates at Children's Memorial Hospital (CMH; Chicago, IL) from 1956 through 1961 were M types 12, 6, 1, 5, and 3, with 18 other types accounting for the remaining M-typable organisms. These early studies were hampered by a limited array of typing sera; many isolates were nontypeable. Molecular genotyping of *emm* genes encoding the type-specific portion of M proteins has led to identification of an increasing number of *emm* types, with >160 distinct *emm* genotypes now identified [5, 6]. We and others have shown that many types cocirculate within communities [7–9]. Immunity to GAS strains is predominantly M type–specific, mediated by opsonophagocytic antibody to the hypervariable 35–50 amino terminal residues of M protein [10].

The serotype distribution of invasive GAS strains is well characterized [11, 12], but there has not been systematic acquisition and characterization of recent pediatric pharyngeal isolates from representative sites in the United States. Progress in M protein–based vaccine development highlights the importance of assessing the distribution of pharyngeal GAS types and the extent of geographic and temporal variation in their distribution [13, 14]. In addition, there has not been analysis of the

Received 8 December 2003; accepted 18 February 2004; electronically published 15 July 2004.

Presented in part: Annual Meeting of the Infectious Diseases Society of America, October 2002, Chicago, Illinois (abstract 110); Annual Meeting of the Pediatric Academic Society, May 2003, Seattle, Washington (abstract 1880).

[a] Members of the United States Streptococcal Pharyngitis Surveillance Group are listed at the end of the text.

Reprints or correspondence: Dr. Stanford T. Shulman, Div. of Infectious Diseases, 2300 Children's Plaza, Box 20, Chicago, IL 60614 (sshulman@northwestern.edu).

Clinical Infectious Diseases 2004;39:325–32

1058-4838/2004/3903-0006$15.00

variation within the serotype-specific M protein domains of pharyngeal GAS strains.

We report our characterization of pediatric pharyngeal GAS isolates collected prospectively and systematically at North American sites during 2000–2001 and 2001–2002. Isolates were genotyped and subtyped, and their *emm* type distribution was compared with the recent Active Bacterial Core surveillance collection of invasive GAS isolates [12].

METHODS

Subjects and specimens. We established the US Streptococcal Pharyngitis Surveillance Group at 9 diverse pediatric sites during 2000–2001 (year 1) and 10 US and 1 Canadian site during 2001–2002 (year 2) (table 1). All sites were primary care offices except the California site, which is a laboratory that services linked offices, and the Montreal site, which is a children's hospital.

From a convenience sample of children aged 3–18 years with acute pharyngitis, we collected ~1000 GAS isolates from 15 November 2000 to 15 May 2001 and ~1000 GAS isolates from 1 December 2001 to 15 May 2002. Each site had 2–4 collection intervals each year. Swabs or streaked plates, identified only by the age and sex of the patient and the date of acquisition, were shipped overnight to the Infectious Diseases Laboratory at CMH. The CMH institutional review board considered this an exempt study, because specimens were obtained for routine care, results could not be traced to individuals, and data were not utilized for clinical management.

Confirmation of group A streptococci. Swab samples were plated promptly, and individual β-hemolytic colonies were subcultured and grouped (PathoDX) to confirm the presence of Lancefield group A organisms.

Genotyping of group A streptococci. Single colonies were incubated overnight at 35° C, and the lawn of streptococci was harvested and suspended. Aliquots were frozen at −70°C in a 50% glycerol/50% Mueller-Hinton solution. DNA extracts were prepared as follows: aliquots were heated at 70°C for 10 min, centrifuged, and pellets were incubated in 50 μL of 10 mmol/L Tris and 1 mmol/L EDTA (pH, 8.0) with 2 μL hyaluronidase (45,000 U/mL) and 10 μL mutanolysin (3000 U/mL) for 30 min at 37°C, boiled for 10 min, and frozen at −30°C. DNA extracts were shipped to the Streptococcal Research Laboratory at the Centers for Disease Control and Prevention (CDC; Atlanta, GA).

emm *Sequence typing.* With use of established protocols [5], isolates were *emm* sequenced. An *emm* sequence type shares >95% DNA sequence identity over a 160–bp 5′ region, with single in-frame deletions and/or insertions of up to 7 codons or single-frame shifts affecting no more than 7 codons permissible [5]. Base changes within the 150-bp region encoding the predicted mature M protein N terminus relative to CDC type strains are assigned subtype designations (e.g., *emm* 3.4, with the reference strain *emm* 3.0). The majority of CDC reference strains for *emm* types 1–60 were originally obtained from R. Lancefield (Rockefeller University). Type and subtype sequences are available for downloading and search with the Basic Local Alignment Search Tool (Blast) online (at http://www.cdc.gov/ncidod/biotech/strep/strepblast.htm). We deduced predicted signal sequence cleavage sites using the tools on the Center for Biological Sequence Analysis Web site (http://www.cbs.dtu.dk/services/SignalP/) exactly as instructed and using sequences trained on gram-positive bacteria. Most sequences examined included ~700 bps; however, the data presented here focus on 150-base type–specific sequences encompassing approximately bases 67–216 of the *emm* amplicon.

Table 1. Geographic distribution of group A streptococcal isolates in study year 1 (2000–2001) and year 2 (2001–2002).

	No. of isolates	
Site	Year 1	Year 2
Chicago, IL	117	115
Sioux Falls, SD	91	93
Bristol, CT	114	97
Gainesville, FL	105	89
Panorama City, CA	112	94
Federal Way, WA	54	50
Sellersville, PA	111	102
Austin, TX	134	102
Little Rock, AR	104	92
Sandy, UT	…	103
Montreal, QU	…	96
Total	942	1033

Data analysis. Data were entered into an Excel spread sheet file (Microsoft). Descriptive statistics characterized age and sex of the subject from whom the sample was obtained. χ^2 Analysis was performed for categorical variables. The relative proportions of the 10 most prevalent serotypes from each study site were assessed graphically to identify possible temporal and geographic trends.

An invasive index was calculated for each common *emm* type, defined as the proportion of invasive isolates in the 2000–2001 CDC surveillance collection [12] divided by the proportion of pharyngeal isolates obtained in 2000–2002 associated with that *emm* type in the present study.

RESULTS

In year 1, 1007 of 1070 specimens yielded GAS isolates. Eleven isolates from a tenth site were excluded because of the small number of isolates obtained from that site. Twenty-one isolates were lost. A total of 975 confirmed GAS isolates from 9 sites were characterized fully; 942 isolates from children 3–18 years

old (range, 54–134 isolates per site; median, 111 isolates) (table 1) constituted the year 1 set; 52.4% of the isolates were obtained from male subjects. The mean age (±SD) of the subjects was 8.1 ± 3.5 years.

In year 2, 1076 of 1144 specimens from 11 sites yielded GAS isolates. Of these, 1054 of 1076 isolates were from children 3–18 years old (range, 50–115 isolates per site; median, 96 isolates). Twenty-one of these isolates were lost, and 1033 isolates were fully characterized as the year 2 set; 52.6% of the isolates were from male subjects. The mean age (±SD) of the subjects was 7.8 ± 3.3 years.

Isolate heterogeneity. *emm* Type distribution of study isolates is presented in table 2, with cumulative frequencies shown in figure 1. Overall, 29 *emm* types were recovered in year 1, including 2 new types (deletion variants of *emm* 12 and 1). Three types (*emm* 12, 1, and 28) accounted for 463 (49.2%) of 942 isolates, and these 3 types plus *emm* 4, 3, and 2 accounted for 679 (72.1%) of 942 isolates. The 10 most prevalent types accounted for 845 (89.7%) of 942 isolates. Nineteen other *emm* types—8 (including the 2 deletion variants) isolated once each—accounted for the remainder (table 2).

In year 2, 31 *emm* types were identified (including 1 new type), with *emm* 1, 12, and 4 accounting for 487 (47.1%) of 1033 isolates; including *emm* 28, 3, and 2, the 6 most prevalent *emm* types accounted for 69.4% of the isolates. Ten types accounted for 85.9% of isolates, a proportion similar to that found in year 1; 21 other types, including 10 that were isolated once each, accounted for the remainder.

Only 1 isolate exhibited an *emm* gene defect predicted to result in a nonfunctional M protein: 1 of 2 *emm* 41.2 genes was predicted to encode a truncated M protein of 63 amino acids because of a 14-base insertion immediately after codon 54.

Year-to-year fluctuation in national proportions of the most common *emm* types was noted (figure 1 and table 2). For example, from year 1 to year 2, *emm* 12 and 28 isolates decreased by 21% ($P = .02$) and 38% ($P < .001$), respectively, and *emm* 1 and 6 increased by 44% and 17%, respectively ($P < .001$ for each). In addition, *emm* 75 isolates tripled in frequency and displaced *emm* 22 as one of the 10 most frequently isolated types.

Geographic diversity. Intersite variation in degree of serotype diversity was evident, with the percentage of isolates accounted for by the 6 genotypes that were most prevalent nationally in year 1 ranging from 59.0% (in Illinois) to 81.5% (in Washington) at individual sites, and with the 10 genotypes that were most prevalent accounting for a percentage of isolates that ranged from 74.4% (in Illinois) to 98.1% (in Washington). Similarly, year 2 data show the percentage of isolates accounted for by the 7 most prevalent genotypes ranged from 56.5% (in Arkansas) to 88.2% (in Texas); the percentage of isolates accounted for by the 10 most prevalent genotypes ranged from 81.3% (in California and Quebec) to 97.1% (in Texas).

Geographic variability in distribution of individual *emm* types in year 1 and year 2 was apparent for each prevalent type (table 3). For example, in year 1, the percentage of isolates accounted for by *emm* 12, the most common *emm* type nationally, ranged from 4.4% (in South Dakota) to 27.9% (in Arkansas); the same was true for other prevalent types. The same phenomenon was observed in year 2 (table 3); little cor-

Table 2. Distribution of *emm* types among isolates recovered from patients with acute pharyngitis in study year 1 (2000–2001) and year 2 (2001–2002).

	No. (%) of isolates	
emm Type	Year 1	Year 2
12	194 (20.6)	169 (16.3)
1	146 (15.5)	231 (22.3)
28	123 (13.1)	84 (8.1)
4	78 (8.3)	87 (8.4)
3	73 (7.7)	82 (7.9)
2	65 (6.9)	64 (6.2)
6	50 (5.3)	64 (6.2)
89	46 (4.9)	44 (4.3)
77	39 (4.1)	37 (3.6)
22	31 (3.3)	25 (2.4)
44/61	22 (2.3)	20 (1.9)
75	18 (1.9)	58 (5.6)
5	14 (1.5)	15 (1.4)
58	7 (0.7)	1 (0.1)
82	7 (0.7)	0 (0)
11	7 (0.7)	13 (1.3)
18	3 (0.3)	1 (0.1)
73	3 (0.3)	3 (0.3)
87	3 (0.3)	1 (0.1)
83	2 (0.2)	0 (0)
102	2 (0.2)	2 (0.2)
9	1 (0.1)	0 (0)
118	1 (0.1)	11 (1.0)
78	1 (0.1)	6 (0.6)
94	0 (0)	4 (0.4)
114	0 (0)	2 (0.2)
41	0 (0)	2 (0.2)
33	1 (0.1)	0 (0)
14	1 (0.1)	0 (0)
96	0 (0)	1 (0.1)
29	0 (0)	1 (0.1)
53	0 (0)	1 (0.1)
92	1 (0.1)	1 (0.1)
68	1 (0.1)	1 (0.1)
57	0 (0)	1 (0.1)
New types	2 (0.2)	1 (0.1)
Total	942 (100)	1033 (100)

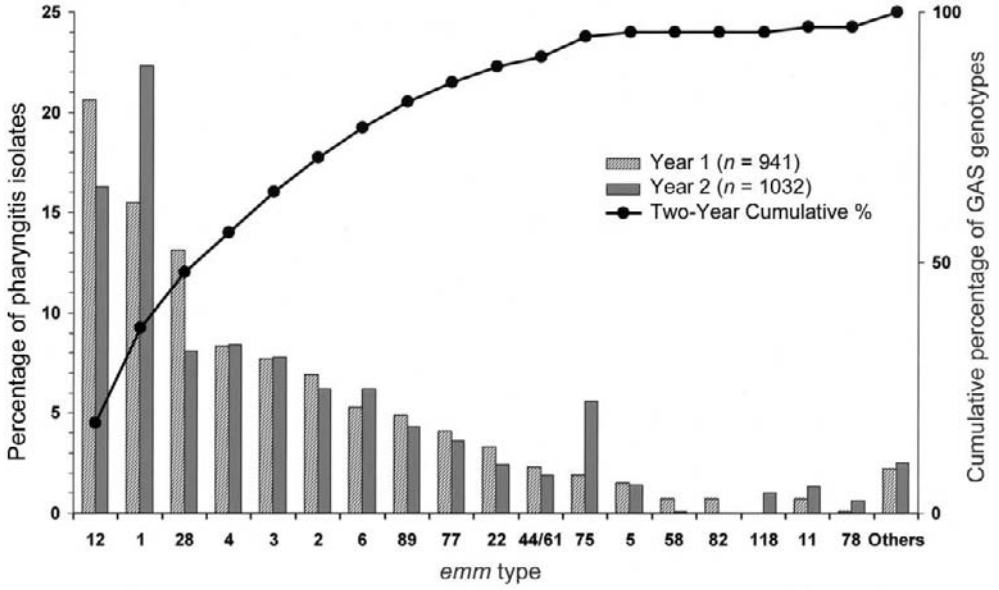

Figure 1. Cumulative percentage of group A streptococcal (GAS) genotypes for years 1 and 2 *(solid line)* and percentage of isolates recovered from patients with acute pharyngitis in year 1 *(cross-hatched bars)* and year 2 *(solid bars),* by *emm* type.

relation between study years is apparent. Figure 2 shows interyear variability of the most prevalent genotypes at 3 selected sites.

In addition, several less common genotypes in year 1 were maldistributed: 11 (50%) of 22 *emm* 44/61 isolates were from Illinois and 6 (27%) were from Texas; 4 (57%) of 7 *emm* 11 isolates were from California; 5 (71%) of 7 *emm* 82 isolates were from Connecticut; 2 (67%) of 3 *emm* 87 isolates were from Illinois; and 13 (72%) of 18 *emm* 75 isolates were from Illinois, Texas, and Arkansas. In year 2, 5 (45%) of 11 *emm* 118 isolates were from Pennsylvania and 3 were from Utah; 8 (40%) of 20 *emm* 44/61 strains were from Connecticut; 11 (25%) of 44 *emm* 89 isolates were from Pennsylvania; 6 (100%) of 6 *emm* 78 isolates were from Arkansas; and 2 (50%) of 4 *emm* 94 isolates were from California.

Subtypes. The array of subtypes based on sequences encoding amino acids 1–50 relative to CDC reference strains (i.e., reference sequences designated as *emm* 1.0, etc.) is presented in figure 3 (online only). Remarkably, *emm* 4.3 was the only 150-base subtype sequence differing from the CDC reference

Table 3. Geographic variability of common *emm* types of isolates recovered from patients with acute pharyngitis in study year 1 (2000–2001) and year 2 (2001–2002).

	Percentage of isolates of stated type, by study year and site					
	Year 1			Year 2		
emm Type	Mean value for all sites	Lowest value (site[s])	Highest value (site)	Mean value for all sites	Lowest value (site[s])	Highest value (site)
12	20.6	4.4 (SD)	27.9 (AR)	16.3	2.9 (UT)	27.1 (CA)
1	15.5	8.9 (TX)	29.8 (AR)	22.3	10.4 (CA)	41.1 (FL)
28	13.1	4.8 (AR)	22.4 (TX)	8.1	1.1 (AR)	15.1 (SD)
4	8.3	3.6 (PA)	24.2 (SD)	8.4	1.0 (CT)	14.8 (IL)
3	7.7	0.8 (IL)	18.5 (WA)	7.9	0 (IL, PA)	19.6 (TX)
2	6.9	1.0 (AR)	14.2 (TX)	6.2	2.1 (CT)	11.9 (PA)
6	5.3	1.9 (FL)	8.0 (CA)	6.2	0 (PA)	16.5 (CT)
89	4.9	1.1 (SD)	12.5 (AR)	4.3	0 (WA)	10.9 (PA)
77	4.1	0 (AR)	14.3 (SD)	3.6	0 (CT, IL)	12 (WA)
22	3.3	0 (SD)	9.5 (FL)	2.4	0 (PA, TX)	7.3 (CA)
5	1.5	0 (5 sites)	8.1 (PA)	1.4	0 (5 sites)	4.2 (QU)

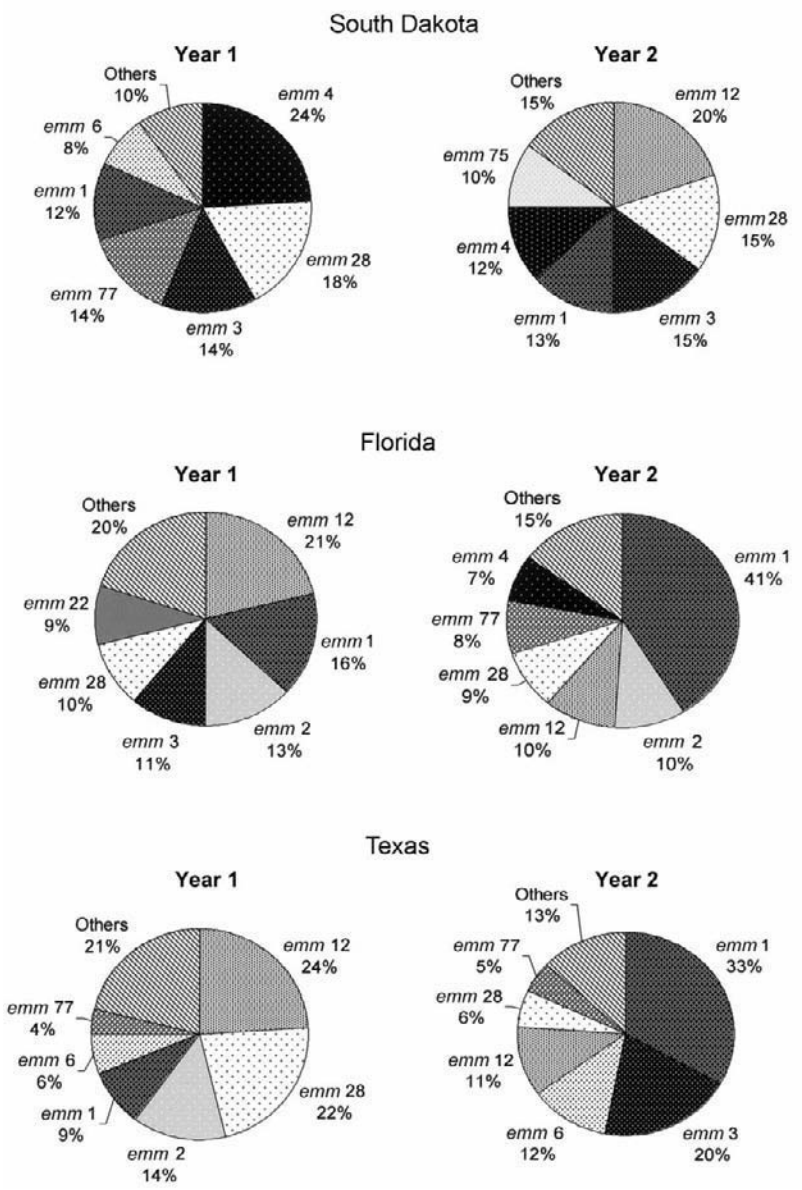

Figure 2. The distribution of group A streptococcal genotypes in years 1 and 2 at 3 selected sites (South Dakota, Florida, and Texas). It is apparent that there are substantial intersite differences and substantial year-to-year variation in which genotypes are most prevalent.

type solely by a synonymous substitution; all other nucleotide substitutions resulted in amino acid substitutions. In year 1, study isolates of 10 *emm* types displayed no differences within residues 1–50 or type-specific region variants, compared with CDC reference strains, including *emm* 2, 22, 28, 4, 44/61, 75, 82, 33, 92, and 78, while 26 subtypes occurred among *emm* types 1, 5, 9, 11, 12, 14, 18, 58, 68, 73, 77, 83, 87, 89, and 102, mostly single missense mutations of reference-type strains (figure 3; online only). The most common of these subtypes were 32 *emm* 12 variants and 14 *emm* 89 variants, including 22 *emm* 12.2 isolates (all from Connecticut) and 14 *emm* 89.1 isolates. In year 2, no type-specific region amino acid sequence variants were found within 17 *emm* types, including *emm* types 1, 2, 4 (including 4 *emm* 4.3 isolates with synonymous substitution), 18, 44, 53, 58, 73, 75, 77, 78, 87, 92, 94, 96, 114, and 118.

In marked contrast, no sequence identical to the M3 Lancefield reference strain *emm* sequence was found among 155 *emm* 3 study isolates. Almost all year 1 *emm* 3 isolates represented subtypes *emm* 3.1 (54 isolates), *emm* 3.4 (11 isolates), and *emm* 3.2 (5 isolates), and most year 2 *emm* 3 isolates were *emm* 3.1 (54 isolates), *emm* 3.4 (13 isolates), and *emm* 3.12 (8 isolates).

The CDC has accumulated a large database of *emm* gene sequences, primarily from invasive GAS isolates (available at http://www.cdc.gov/ncidod/biotech/strep/strepindex.htm),

particularly from Active Bacterial Core surveillance (available at http://www.cdc.gov/abcs). Subtypes found both among pharyngeal isolates in the present study and among 1061 invasive isolates obtained during 1999 and 2000 [12] are shown in figure 3 (online only); 39 of 63 variant subtypes shown in this figure were initially discovered during this pharyngitis surveillance.

Geographic associations of several variant subtypes were noteworthy; for example, 17 of 21 *emm* 89.1 isolates were from Arkansas, 9 of 10 year 2 *emm* 6.7 isolates were from Utah, 8 of 8 *emm* 6.8 isolates were from South Dakota, 5 of 5 year 1 *emm* 3.2 isolates were from Connecticut, and 4 of 4 year 2 *emm* 6.9 isolates were from Texas.

Except for *emm* 5 and *emm* 6 (accounting for 14 and 50 year 1 isolates, respectively, and 15 and 64 year 2 isolates, respectively), *emm* types associated with ⩾4 isolates were characterized primarily by a single 50-codon type-specific sequence. Type *emm* 5 exhibited 5 subtypes in year 1 and 6 subtypes (including *emm* 5.0) in year 2; much *emm* 5 variation results from homologous excisions of tandem repeats such that the 3′ boundary of the hypervariable type-specific region is altered [12] (data not shown). Similarly, *emm* 6 exhibited 8 subtypes in year 1 and 9 subtypes in year 2.

New types. Three new *emm* types (stIL103, stIL62, and stPA57), from Illinois and Pennsylvania, were identified. Type stIL103 is very similar to *emm* 1, differing by a single intragenic deletion of 54 base pairs, resulting in deletion of M1 residues 7–24 and changing M1 Leu25 to isoleucine (figure 3; online only). Similarly, stIL62 differs from type *emm* 12 by a 42-base deletion within the type-specific sequence, leading to deletion of 14 codons. Type stPA57 was truly unique; it was most similar to *emm* 5 subtypes, but it had only ~50%–60% amino acid sequence identity (figure 3; online only).

Sex and age. No sex differences in genotype distribution were apparent. However, age was related to the distribution of several genotypes, as described elsewhere [15].

Invasive index. Table 4 shows the invasive index for 17 prevalent GAS genotypes, calculated from the relative proportion of pharyngeal isolates (reported in the present study) and invasive isolates for each genotype [12]. Certain *emm* types are less commonly associated with invasive disease than with pharyngitis (*emm* 2, 4, 6, 12, and 44/61), and others are more frequently associated with invasive infections (*emm* 3, 11, 73, 82, and 114). Of note, *emm* 1, the single most common invasive genotype, accounts for virtually identical proportions of invasive and pharyngeal isolates, with an invasive index of 0.96.

Table 4. Invasive index of group A streptococcal genotypes.

emm Type	Percentage of isolates of stated type: Pharyngitis[a] (n = 1975)	Percentage of isolates of stated type: Invasive[b] (n = 1061)	Invasive index value
1	18.9	18.2	0.96
12	18.4	8.4	0.46
28	10.6	7.9	0.75
4	8.4	2.4	0.29
3	7.8	10.2	1.31
2	6.6	2.4	0.36
6	5.8	1.5	0.26
89	4.6	5.5	1.20
77	3.8	3.6	0.95
22	2.8	2.4	0.86
44/61	2.1	1.0	0.48
5	1.4	1.7	1.21
75	3.8	2.9	0.76
82	0.4	5.9	14.75
11	1.0	3.4	3.40
114	0.1	2.6	26.00
73	0.3	2.2	7.33

NOTE. The invasive index is calculated by dividing the percentage of invasive isolates by the percentage of pharyngitis isolates. Note that the invasive index is calculated from the pharyngitis data in the present study (representing 10 widely scattered North American sites, each of which contributed approximately equal numbers of isolates) and from the invasive isolate data from the 9 states comprising the Active Bacterial Core surveillance sites (in which more populous states contribute proportionally more isolates).

[a] Data presented are for study year 1 and year 2, as reported in the present study.

[b] Data presented are for isolates reported in [12].

DISCUSSION

In this study, we identified features of pharyngeal GAS isolates with implications for streptococcal vaccine development. Relatively few *emm* types accounted for the majority of isolates at geographically diverse surveillance sites. The 10 most prevalent genotypes accounted for 86%–90% of isolates. However, during 2 respiratory seasons, 38 genotypes were identified, including 3 new genotypes. Three genotypes (*emm* 12, 1, and 28) represented 49% of isolates in year 1, and *emm* 1, 12, and 4 accounted for 47% of isolates in year 2. Six genotypes accounted for 72% of isolates in year 1 and 69% in year 2. That the surveillance site with the least heterogeneity during year 1 (South Dakota) yielded 10 *emm* types emphasizes the complex epidemiology of streptococcal pharyngitis. During year 2, less-striking heterogeneity was observed, but all sites yielded ⩾12 *emm* types. Considerable geographic variation in the prevalence of particular *emm* types was noted, with substantial site-to-site differences in prevalence of *emm* types. The epidemiologic complexity is highlighted further by an association between the age of subjects and the predominance of certain *emm* types [15]. In addition, intrasite differences in *emm* type distribution between study years were striking in some instances (figure 2). Year-to-year changes in genotypic prevalence at individual sites is consistent with the predominantly type-specific nature of immunity to group A streptococci, as was recently emphasized [16].

Despite considerable geographic and temporal variation in *emm* type distribution, coverage by an effective multivalent M-protein vaccine would be substantial. The 26-valent M-protein vaccine that is in clinical trials includes immunogens for 86% of pharyngeal isolates and a highly conserved epitope (streptococcal protective antigen) that may protect against additional types [13]. This vaccine includes 35–50 amino acid N-terminal segments of *emm* types 1–3, 5–6, 11–14, 18, 19, 22, 24, 28, 29, 33, 43, 59, 75–77, 89, 92, 101, and 114. Of note, rabbits immunized with the 26-valent vaccine developed serum bactericidal activity against 5 of 7 randomly selected *emm* 4 isolates from the present study, even though M4 is not included in the vaccine [13].

It is noteworthy that several classic rheumatogenic types, such as *emm* 5 and 18, were infrequent among acute-pharyngitis isolates in this study. *emm* 18 was isolated only 4 times, with no isolates from the Utah site, where acute rheumatic fever is common and thought to be related to mucoid *emm* 18 [17]. Recent data from Utah indicate that M18 strains became rare during 2001–2002, coinciding with year 2 of our surveillance [18]. *emm* 5 Isolates, also classically considered rheumatogenic, accounted for 1.5% of our isolates.

Most assessments of GAS type distribution focus on invasive or sterile site isolates [11, 19]. Studies evaluating *emm* type distribution among pharyngitis isolates generally are either limited in scope or limited geographically. A recent pharyngitis study at 8 US basic training sites indicated predominance of *emm* 29 (virtually all isolates of which were at 1 base), 3, 6, 44/61, 2, 75, and 1, with few *emm* 12 and 28 isolates [20].

Comparison of our pharyngitis isolates to sterile site isolates obtained from 1995–1999 and from 2000–2001 yields findings of interest [11, 12]. Sterile site isolates are more heterogeneous, with the 3 most prevalent types (*emm* 1, 3, and 28) accounting for 37.5% of isolates (compared with 49% for pharyngitis isolates) and the 6 most prevalent accounting for 52.2% (compared with ~70% for pharyngitis isolates). The 10 most prevalent types comprise ~90% of pharyngitis isolates but only 64% [11] to 68.6% [12] of invasive isolates. There is considerable *emm* type overlap between pharyngitis and sterile-site isolates, with *emm* 1, 12, and 28 among the 4 most prevalent types for both. This reinforces the concept that children with pharyngitis are an important reservoir for GAS strains with invasive potential [7]. Nevertheless, interesting differences in the relative frequencies of certain *emm* types exist between pharyngitis and sterile-site isolates, as expressed by the invasive index (table 4). Among the 12 most common types, *emm* 2, 4, 6, 44/61, and 12 each had a low invasive index value and accounted for smaller proportions of invasive isolates than did pharyngeal isolates; *emm* 1, 3, 5, 11, 73, 82, and 89 had higher invasive index values. Differences in virulence factors, including in exotoxins, may account for variations in invasiveness. When considering the invasive index values, it should be recognized that, although both surveillance systems were geographically diverse, the specific geographic areas and the sampling methods differed. This possibly could influence the observed differences in invasive index values, but it seems unlikely that it could change the findings substantially.

Although much heterogeneity has been observed among the >160 *emm* sequence types, the present study demonstrates remarkable sequence conservation within most of the common types. Most of these recent isolates (i.e., those within *emm* types 1–89) share identical M type–specific sequences with corresponding reference strains that were isolated 20–70 years ago [5, 6, 21]. M subtype variation is relatively minor (figure 3; online only). Because M protein is subject to selective pressure by host immune responses, undefined structural restrictions on variation of the type-specific region must exist.

Our study is unique in that we present data on large numbers of pharyngitis isolates that were systematically acquired from geographically diverse pediatric sites (including all US regions and 1 Canadian city), that were collected throughout the streptococcal season, and that were all *emm* typed [11]. The results demonstrate that, despite considerable *emm* type heterogeneity, the majority of pharyngitis isolates are concentrated within relatively few M serotypes. An effective M protein–based vaccine encompassing the most prevalent serotypes may impact very substantially on streptococcal pharyngitis and its complications.

THE US STREPTOCOCCAL PHARYNGITIS SURVEILLANCE GROUP

Dr. Richard Nachman (Chicago, IL), Dr. Edward Rothstein (Sellersville, PA), Dr. Thomas Zavelson (Gainesville, FL), Dr. Cathy Hofer and Dr. Chris Tiongson (Sioux Falls, SD), Dr. S. Michael Marcy and Dr. Emily Chang (Panorama City, CA), Dr. Jon Almquist (Federal Way, WA), Dr. Del Hodder (Bristol, CT), Dr. John Ledbetter and Dr. Tracey Stewart (Little Rock, AR), Dr. Paige Suffrendini (Austin, TX), Dr. Julie Gustin (Sandy, UT), and Dr. Jane McDonald (Montreal, Quebec).

Acknowledgments

Financial support. ID Biomedical and Children's Memorial Institute for Education and Research (Chicago, IL).

Conflict of interest. J. B. D. receives research support from and owns stock in ID Biomedical. All other authors: No conflict.

References

1. Swift HF, Wilson AT, Lancefield RC. Typing group A hemolytic streptococci by M precipitin reactions in capillary pipettes. J Exp Med **1943**; 78:127.
2. Quinn RW, Denny FW, Riley HD. Natural occurrence of hemolytic streptococci in normal school children. Am J Public Health **1957**; 47: 995.

3. Hope-Simpson RE. *Streptococcus pyogenes* in the throat: a study in a small population, 1962–1975. J Hyg Lond **1981**; 87:109–29.
4. Stollerman GH. Prospects for a vaccine against group A streptococci: the problem of the immunology of M proteins. Arth Rheum **1967**; 10: 245–55.
5. Whatmore AM, Kapur V, Sullivan DJ, Musser JM, Kehoe MA. Non-congruent relationships between variation in *emm* gene sequences and the population genetic structure of group A streptococci. Mol Microbiol **1994**; 14:619–31.
6. Beall B, Facklam RR, Thompson T. Sequencing *emm*-specific PCR products for routine and accurate typing of group A streptococci. J Clin Microbiol **1996**; 34:953–8.
7. Haukness HA, Tanz RR, Thomson RB Jr, et al. The heterogeneity of endemic community pediatric group A streptococcal pharyngeal isolates and their relationship to invasive isolates. J Infect Dis **2002**; 185: 915–20.
8. Ameen AS, Nsanze H, Dawson KP, et al. Serotypes of group A streptococci isolated from healthy schoolchildren in the United Arab Emirates. Bull WHO **1997**; 75:355–9.
9. Dicuonzo GG, Gherardi G, Lorino S, et al. Group A streptococcal genotypes from pediatric throat isolates in Rome, Italy. J Clin Microbiol **2001**; 39:1687–90.
10. Lancefield RC. Current knowledge of the type-specific M antigens of group A streptococci. J Immunol **1962**; 89:307–13.
11. O'Brien KL, Beall B, Barrett NL, et al. Epidemiology of invasive group A streptococcus disease in the United States, 1995–1999. Clin Infect Dis **2002**; 35:268–78.
12. Li Z, Sakota V, Jackson D, Franklin AR, Beall B. The array of M protein gene subtypes in 1061 recent invasive group A streptococcal isolates recovered from the Active Bacterial Core surveillance. J Infect Dis **2003**; 188:1587–92.
13. Hu MC, Walls MA, Stroop SD, et al. Immunogenicity of a 26-valent group A streptococcal vaccine. Infect Immun **2002**; 70:2171–7.
14. Dale JB. Multivalent group A streptococcal vaccine designed to optimize the immunogenicity of six tandem M-protein fragments. Vaccine **1999**; 17:193–200.
15. Jaggi P, Tanz RR, Beall B, Shulman ST. Age-related differences in pharyngeal group A streptococcal (GAS) M type distribution [abstract 1227]. In: Program and abstracts of the 2003 Pediatric Academic Societies' Annual Meeting (Seattle, WA). Pediatr Res **2003**; 53:216A.
16. Dale JB, Shulman ST. Dynamic epidemiology of group A streptococcal serotypes [letter]. Lancet **2002**; 359:889.
17. Smoot JC, Korgenski EK, Daly JA, et al. Molecular analysis of group A *Streptococcus* type *emm* 18 isolates temporally associated with acute rheumatic fever outbreaks in Salt Lake City, Utah. J Clin Microbiol **2002**; 40:1805–10.
18. Veasy LG, Tani LY, Daly JA, et al. Temporal association of mucoid strains of *Streptococcus pyogenes* with a continuing high incidence of rheumatic fever in Utah [abstract 1881]. In: Program and abstracts of the 2003 Pediatric Academic Societies' Annual Meeting (Seattle, WA). Pediatr Res **2003**; 53:329A.
19. Schwartz B, Facklam RR, Breiman RF. Changing epidemiology of group A streptococcal infection in the USA. Lancet **1990**; 336:1167–71.
20. Barrozo CP, Russell KL, Smith TC, et al. National Department of Defense surveillance data for antibiotic resistance and *emm* gene types of clinical group A streptococcal isolates from eight basic training military sites. J Clin Microbiol **2003**; 41:4808–11.
21. Facklam RR, Martin DR, Lovgren M, et al. Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: *emm* 103 to *emm* 124. Clin Infect Dis **2002**; 34:28–38.