Contrasting Molecular Epidemiology of Group A Streptococci Causing Tropical and Nontropical Infections of the Skin and Throat

Debra E. Bessen,1,a Jonathan R. Carapetis,2,b Bernard Beall,1 Rebecca Katz,1 Megan Hibble,2 Bart J. Currie,2 Tracy Collingridge,3 Marc W. Izzo,1 Dominick A. Scaramuzzino,1 and Kadaba S. Sriprakash2

Disease caused by group A streptococci (GAS) in tropical regions often takes the form of impetigo, whereas pharyngitis tends to predominate in temperate zones. GAS derived from asymptomatic throat infections and pyoderma lesions of rural Aboriginal Australians were evaluated for phylogenetic distant emm genes, which represent ecological markers for tissue site preference. On the basis of the percentage of total isolates from a given tissue, emm pattern A–C organisms exhibited a stronger predilection for the throat, whereas pattern D organisms preferred the skin. Only 16% of isolates collected by active surveillance displayed pattern A–C, which reflects the low incidence of oropharyngeal infection. Importantly, most (70%) pattern A–C organisms were isolated from skin sores, despite their innate tendency to infect the throat. Combined with findings from nontropical populations, analysis of the data supports the hypothesis that GAS tissue preferences are genetically predetermined and that host risk factors for infection strongly influence the differential reproduction of individual clones.

Group A streptococci (GAS; Streptococcus pyogenes) are among the most prevalent bacterial pathogens that afflict humans. These organisms are usually transmitted from person to person through respiratory droplets or by close contact. They are not known to have a well-established secondary host or environmental reservoir; therefore, GAS appear to occupy a narrow ecological niche.

GAS has medical importance as a causative agent of toxic shock syndrome and necrotizing fasciitis; it also acts as a trigger of autoimmunity (e.g., rheumatic fever). Despite these, GAS usually causes only a mild illness (pharyngitis or impetigo) that results from infection at the nasopharyngeal mucosa or superficial layers of the skin [1–3]. The relative incidence of GAS pharyngeal infection and impetigo can vary markedly among different host populations, in accordance with both season and locale [4, 5]. In the temperate regions of North America and Europe, pharyngitis is common during the winter months, and impetigo, although less common, is most often encountered during warmer weather. In many tropical regions, such as the Northern Territory of Australia, GAS impetigo is far more prevalent than pharyngeal infection, and there are no discrete seasonal peaks in incidence of disease [6].

The M proteins of GAS form surface fibrils and provide the basis of widely used epidemiological typing schemes that employ serological methods (M type) or nucleotide sequence analysis of the M protein gene (emm type); the 2 methods are highly concordant [7]. Decades of epidemiological studies in the United States, Europe, and the Caribbean demonstrate that most M types associated with pharyngitis (M1, M3, M5, M6, M12, M18, M19, and M24) are rarely found in impetigo lesions. On the other hand, many skin isolates could not be readily M typed, and alternative serotyping schemes were often used, such as T typing and OF typing. These findings led to the widely recognized concept that there are distinct GAS populations of “skin types” and “throat types” [3, 4, 8–12]. With the recent advent of emm sequence typing, it has become apparent that many isolates from tropical regions are of newly recognized emm types that had been M nontypeable by serological means [13–16].

Although >150 distinct emm and emm-like genes are now recognized, their evolutionary history can be traced to 4 major phylogenetic lineages (designated as subfamilies [SF]) [17]. The content and relative chromosomal arrangements of the 4 emm SF genes are found to exist in only 5 basic patterns, A–E (figure

Received 11 May 2000; revised 22 June 2000; electronically published 8 September 2000.

Presented in part: 14th Lancefield International Symposium for Streptococci and Streptococcal Diseases, Auckland, New Zealand, October 1999 (abstract O8.8).

Informed consent was obtained from all subjects (parents or guardians in the case of minors), and the surveys were performed with institutional review board approval.

Financial support: American Heart Association (Grant-in-Aid) and National Institutes of Health (AI-28944 and GM-60793 to D.E.B.), National Health and Medical Research Council and Australian Rotary Health Research Foundation (J.R.C.), Wilbur Downs Travel Fellowship (R.K.), Australian National Heart Foundation (B.J.C.), and Channel 7 Children’s Medical Research Foundation (K.S.S.).

D.E.B. is an Established Investigator of the American Heart Association.

Present affiliation: Melbourne University Department of Paediatrics, Royal Children’s Hospital, Parkville, Victoria, Australia.

Reprints or correspondence: Dr. Debra Bessen, Yale University School of Medicine, Dept. of Epidemiology and Public Health, 60 College St., Box 208034, New Haven, CT 06520 (debra.bessen@yale.edu).

The Journal of Infectious Diseases 2000;182:1109–16
© 2000 by the Infectious Diseases Society of America. All rights reserved.
0022-1899/2000/18204-0014$02.00
shown. These 5 emm patterns (A–E) account for all emm gene arrangements among the group A streptococci (GAS) isolates under study in this report. Intergenic distances are typically 0.2–0.25 kb. The 5’ portion of the central emm gene that contains the determinants for emm sequence type (i.e., emm type) is indicated. Approximate positions for oligonucleotide primer sites, used for polymerase chain reaction–based mapping (table 1), are shown; occasionally the G3 priming site is in the SF3 emm of pattern E strains, rather than in the SF2 emm gene (data not shown).

Figure 1. The emm chromosomal patterns. Content and relative arrangement of emm subfamily (SF) genes (SF1–SF4), on the basis of phylogenetic divergence at the 3’ ends of emm and emm-like genes, is shown. These 5 emm patterns (A–E) account for all emm gene arrangements among the group A streptococci (GAS) isolates under study in this report. Intergenic distances are typically 0.2–0.25 kb. The 5’ portion of the central emm gene that contains the determinants for emm sequence type (i.e., emm type) is indicated. Approximate positions for oligonucleotide primer sites, used for polymerase chain reaction–based mapping (table 1), are shown; occasionally the G3 priming site is in the SF3 emm of pattern E strains, rather than in the SF2 emm gene (data not shown).

1), with very few exceptions [18, 19]. In previous studies on emm pattern distribution among epidemiologically defined GAS isolates (mostly from nontropical regions), emm pattern A–C isolates were found to be disproportionately associated with the nasopharynx, whereas emm pattern D strains were most often isolated from impetigo lesions [19]. Organisms of a third pattern group, emm pattern E, were readily found at both tissue sites. In an experimental model for impetigo, whereby GAS are topically applied to slightly damaged human skin that is engrafted onto mice with the scid mutation, pattern D strains display a higher overall virulence than pattern A–C strains [20]. The findings of these studies suggest that emm pattern A–C is a genetic marker for preferential infection at the throat, whereas emm pattern D is a marker for strains with a greater tendency to infect the skin, although the associations with each tissue are not absolute.

To test the broader applicability of emm patterns as markers for principal tissue reservoir preference, organisms isolated from the throat and skin of subjects living in a tropical community were analyzed for emm pattern. This is the first report of a large-scale analysis of tropical-derived GAS strains for emm pattern content. To our knowledge, this is also the first molecular genetic analysis of GAS throat and skin isolates that have been collected by intensive, population-based active surveillance.

Methods

Bacteria and subjects. Periodic population-based screening of a remote, rural Aboriginal island community over a 25-month period from 1994 to 1996 was conducted as part of a scabies control program [21]. Screening for β-hemolytic streptococci was performed on 7 separate occasions, with a minimum duration of 4 weeks between any 2 screening visits, and included a total of 508 subject visits that involved 224 individual subjects (105 children and 119 adults). For the initial visit, all available residents were targeted for screening, whereas all children and some readily accessible adults were targeted on later visits (the residents are a mobile population). Seventy-four percent of adults were seen at only 1 or 2 visits, whereas 64% of the children were examined at ≥3 visits. We did not preselect subjects on the basis of health status, nor were subjects selected on the basis of the presence or absence of visible skin lesions. Despite multiple shipments of GAS from Australia to the United States, not all organisms remained viable, and these were excluded from the molecular analysis.

Throat swabs were taken from all subjects; there were no cases of symptomatic or clinical pharyngitis when throat swabs were taken. Swabs were transferred directly to blood agar plates in the field to avoid possible loss of viability through handling. A total of 17 GAS-positive throat swabs (n = 508 throat screened) were collected from the community members, and all 16 isolates that arrived at Yale in viable form are included in this report. No attempts were made to distinguish among asymptomatic throat isolates as causing either carriage or clinically inapparent infection on the basis of serological evidence, because, in the absence of seroconversion from an earlier GAS-free time (data that are unavailable), serological measurements are not absolutely definitive for infection. In particular, serological markers of recent GAS infection are uniformly elevated in children in rural Australian Aboriginal communities [22], presumably because of high prevalence rates of pyoderma; therefore, serology is not useful for diagnosis of asymptomatic pharyngitis in this population.

Skin swabs were obtained only from pyoderma lesions (<3 swabs per affected person), thereby representing true infections. Swabs of normal skin were not taken, because previous surveys in this population revealed low GAS isolation rates from normal skin (data not shown), and GAS on normal skin in populations with high prevalences of pyoderma usually represent those found in pyoderma lesions [23]. In some instances, pyoderma swabs were obtained in the absence of throat swabs either because they were not sought or because the patient did not give consent; these subject visits are excluded from this analysis. The skin lesions represented a range of infections, classified as wet/pus, crust, and flat/dry. Recovery of GAS was highest from the wet/pus and crusted groups (97%) and lowest from the flat/dry group (44%). All the lesions were nonbullous impetigo; there were no cases of bullous impetigo.

Of the 265 impetigo lesions from 508 subject visits that were screened for β-hemolytic streptococci, GAS was isolated from 150 lesions, group G streptococci from 6 lesions, and group C streptococci from 1 lesion (staphylococcal isolation rates were not determined). From each skin swab, ≤4 colonies were analyzed, but
isolates from tropical Australia were subjected to central (st; also referred to as in this report include 16 throat samples and 125 skin samples (141 total).

Molecular typing methods: emm type determination. All 141 isolates from tropical Australia were subjected to emm sequence typing [24]. The emm sequence typing is based on the 5'-end of the central emm gene (figure 1). Newly identified emm sequence types (st; also referred to as emm type) stCK401, stCK249, and stNS554 were assigned GenBank accession numbers AF183963–AF183965, respectively. All GAS displaying new emm types were reconfirmed for group A carbohydrate by a latex agglutination test, because other β-hemolytic streptococci (i.e., groups C and G) can harbor structurally related emm-like genes. Furthermore, they were confirmed as S. pyogenes on the basis of a positive test for pyrrolidonyl arylamidase production and a negative test for acetoin production (Voges-Proskauer test), thereby distinguishing them from the S. milleri group. A unique emm type is defined as having <95% sequence identity to any other known emm type over 160 base pairs near the 5'-end [24]. Indels (i.e., insertions or deletions) >4 codons or frameshift mutations relative to the reference emm typing strain were not encountered.

Molecular typing methods: emm pattern determination. The emm chromosomal pattern was determined by polymerase chain reaction (PCR)-based mapping, as described elsewhere [17, 18, 25], and is reported as emm pattern groupings A–C, D, or E. In brief, PCR-based mapping of the emm chromosomal region involves the generation of overlapping PCR products by use of shared priming sites to construct a linear map. Oligonucleotide primers were originally designed on the basis of extensive sequence analysis of multiple emm genes, and the method has been substantiated by Southern hybridization analysis [18, 25–27].

In this report, an improved, more efficient strategy for emm pattern determination was applied to isolates from Australia. To generate a linear map, 8 oligonucleotide primers (all analyzed elsewhere, as cited above) were paired for a total of 7 PCR amplifications (table 1). The primers we used are as follows: UP-2, 5'-TCTGGATCCCCACTCTCCCCAAACAAGTTGCG-3'; SF1-A, 5'-C TCTTAGGGTCAGCTAAGCCTGATTGTTG-3'; G3-F, 5'-CGAGAAATGAAAACGGTTTACAGAACC-3'; IG-F, 5'-CTGGGCCTTTACTCCTTCAATTAC-3'; SF1-R, 5'-GTGCTTTGACCTTTACCTGGAACAGCTT-3'; SF2-R, 5'-GGTACCGTGGTGACACTTT-3'; SF3-R, 5'-GCTGGTTTGAGCAGCGGCTTACC-3'; and G3-R, 5'-GGTCTCTGATACGGCTTCTTCTCTC-3'.

This strategy was successful for 78 of 82 Australian strains attempted; patterns for 4 isolates were established by including SF4-A paired with SF3-R (positive for D and E).

Despite these improvements in the emm patterning scheme, it remains a highly labor-intensive method. In addition, of >300 isolates subjected to emm pattern determination in several studies [25, 28, 29] (this report and unpublished data), only 2 isolates were found to differ in emm pattern from other isolates of the shared emm type. Therefore, for the 59 remaining strains, emm pattern was inferred if >3 isolates of the same emm type from the study community had already undergone the full set of PCR amplifications, because the isolates had closed epidemiological relationships. Confirmatory reactions were performed for inferred emm patterns A–C (UP-2/SF1-R), D (SF4-A/SF1-R), and E (SF4-A/SF2-R); confirmatory reactions were 100% concordant with the emm pattern predictions.

Statistics. Statistical significance was calculated with tests for independence; we used the χ² test with Yates correction for sample size and, for calculations where the sample size was deemed too small, Fisher’s exact text (2-tailed) was recommended and is also reported (Epi Info, version 6.02; Centers for Disease Control and Prevention, Atlanta). Relative risks were calculated as risk ratios and are presented with 95% confidence intervals (CIs).

Results

Population-based surveillance in an Aboriginal community. In the rural and remote island community that we studied, located in the Top End of Northern Territory, Australia, culture swabs were obtained from Aboriginal subjects at the oropharynx and pyoderma lesions on 7 occasions over 25 months during the mid-1990s. Although not all residents of the community were screened at each visit, the repeated screening of many individual residents, particularly children, is likely to provide a population-based representation of the epidemiology of skin infection and throat carriage/asymptomatic infection in this community. Among the 508 subject visits, GAS-positive skin infections (n = 150) outnumbered GAS-positive throat swabs (n = 17) by 8.8-fold. In children, the most comprehensively surveyed group, this ratio increased to 15 (n = 136 GAS-positive skin swabs and n = 9 GAS-positive throat swabs). Only 17 (3.4%) of 508 throat swabs were positive for GAS. All throat cultures appear to represent either an asymptomatic carrier state or clinically inapparent infection.

For this study, emm pattern A–C strains are estimated to account for only 16% of the total GAS among the host population, whereas 43% and 40% of the GAS-positive cultures were emm patterns D and E, respectively. The findings indicate that GAS infection within the study community is largely attributable to emm pattern D and E strains.
Of the 17 throat-derived and 150 skin-derived isolates of GAS, 16 and 125, respectively, arrived at Yale University (New Haven, CT) in viable form and are included in the molecular analyses of this report. Of the 16 GAS from oropharyngeal swabs that were evaluated, 7 (44%) were *emm* pattern A–C strains (figure 2A). In contrast, pattern A–C isolates represent 16 (13%) of the GAS recovered from pyoderma lesions. The relative risk of GAS isolates from the throat belonging to pattern A–C, compared with those from pyoderma lesions, was 3.42 (95% CI, 1.66–7.02). Thus, although the absolute number of *emm* pattern A–C strains was greater in skin lesions than in throat swabs, *emm* pattern A–C appeared to display a >3-fold higher affinity for the throat than for skin lesions. In contrast, pattern D isolates represent 3 (19%) and 58 (46%) of the total GAS derived from the throat and skin, respectively; these values equate to a relative risk for skin over throat infection of 2.47 (95% CI, 0.88–6.98). Importantly, the difference in the distribution of *emm* pattern A–C versus D strains at the throat versus skin was highly significant (n = 7, pattern A–C throat; n = 16, pattern A–C skin; n = 3, pattern D throat; and n = 58, pattern D skin; P < .005, Fisher’s exact test).

Although the number of throat isolates is small, the data are consistent with the hypothesis generated through earlier studies of primarily nontropical populations [19, 29], which states that *emm* pattern A–C strains constitute a large proportion of throat isolates relative to pattern D strains and thereby exhibit an innate tendency or predilection to be present at the throat. In contrast, *emm* pattern D strains exhibit a greater innate tendency to infect the skin.

Although *emm* pattern A–C was the most abundant *emm* pattern grouping isolated from the oropharynx of subjects in the study community, the actual number of pattern A–C GAS obtained from skin lesions exceeded that recovered from the throat by ~2.3-fold (figure 2B). Pattern D and E isolates were recovered from skin lesions more often than from throat swabs by magnitudes of ~19-fold and ~8.5-fold, respectively. Thus, for this tropical community, where GAS impetigo is highly endemic and pharyngitis is very rare, skin lesions appear to be the principal tissue reservoir for all GAS strains, irrespective of *emm* pattern.

In several instances, multiple strains (as defined by distinct *emm* types) were isolated from a single impetiginous lesion or from different lesions on the same subject at the same time [30]. Therefore, it was of interest to determine whether *emm* pattern A–C strains derived from impetigo lesions were more likely to be found in association with a coinfection by a nonpattern A–C strain. Eight coinfected skin lesions were identified, each harboring GAS of 2 distinct *emm* types and representing a total of 3 pattern A–C isolates and 13 pattern D or E isolates. Thus, for the 16 skin lesions containing a pattern A–C isolate (figure 2B), 3 (18.8%) were coinhabited by a pattern D or E strain. No statistically significant difference in coinfection by multiple strains versus a lack of coinfection was found between pattern A–C and either pattern D or E isolates (Fisher’s exact test).

A GAS impetigo infection can spread to the upper respiratory tract (URT) and give rise to a secondary acquisition within an interval averaging 2–3 weeks’ duration, although colonization of the throat may be only transient [4, 23]. Also, certain M types have a propensity to infect both the throat and skin [8]. In the Aboriginal community that we studied, the 141 GAS isolates were derived from a total of 78 individuals. Thirty-four of 78 GAS-positive subjects yielded ≥2 isolates at ≥1 visits. However, only 6 subjects had ≥1 isolate of the same *emm* type; isolates from these 6 subjects are included in this report because they satisfy the inclusion criteria. For 2 of 6 subjects, GAS of the same *emm* type were isolated from both the throat and a skin lesion on the same day. For the remaining 4 subjects, 2 GAS isolates of the same *emm* type were recovered at least 10

Figure 2. The *emm* patterns of group A streptococci (GAS) isolated from the throat and skin of subjects in a rural Aboriginal community. A. Percentage of isolates from a given tissue site represented by *emm* patterns A–C, D, and E. B. No. of isolates represented by *emm* patterns A–C, D, and E. Data plotted are for *emm* pattern A–C isolates from the throat (n = 7) and skin (n = 16), pattern D at the throat (n = 3) and skin (n = 58), and pattern E at the throat (n = 6) and skin (n = 51).
weeks apart. The failure to find evidence for high levels of spread from one tissue to a second site for organisms sharing the same emm type may be the consequence of the long interval between serial screenings (minimum of 4 weeks). Alternatively, an absence of risk factors for throat infection might reduce the overall rate of secondary spread.

**Genetic diversity and lack of geographic barriers to global spread.** One possible explanation for the predominance of emm pattern D and E isolates in tropical Australia is that there is a geographic barrier to migration of the microorganism or its emm genes. For all emm patterns, the 5’ ends of the central emm genes display a high level of sequence heterogeneity; they encode the determinants of a serological typing scheme and are also the targets of protective immunity. emm sequence typing is highly discriminatory and provides a reasonably good substitute for serological typing [24]. By multilocus typing methods, emm type is highly concordant with “clone,” although there are a few notable exceptions whereby one emm type is associated with divergent genetic backgrounds (M. C. Enright [Oxford University, Oxford, UK], B. G. Spratt [Oxford University], and D. E. Bessen [Yale University], personal communication) [24, 28, 29, 31–33]. There are many instances, however, in which a single emm type is associated with multiple sof types, and the genetic backgrounds of these pairs remain to be fully explored.

Of the 141 GAS under study that were isolated from Aboriginal subjects in the study community, 31 distinct emm types were identified (figure 3). The emm type diversity index, defined as the number of distinct emm types divided by the total number of isolates, is ~2-fold greater for emm pattern D and E isolates, compared with that of pattern A–C strains. In fact, 96% of pattern A–C isolates were attributable to organisms representing only 2 distinct emm types (emm14 and emmst3765).

Geographic barriers leading to reduced migration are often the basis for an uneven spatial distribution of organisms that share a recent common ancestor. Of the 31 emm types recovered from study community members, 28 emm types, representing 93.6% of the total isolates, have been reported elsewhere for GAS isolates, recovered outside of Oceania or Southeast Asia [15, 28, 34]. Thus, as many as 9.7% (3/31) of the emm types may have recently emerged within the Top End of the Northern Territory or neighboring countries (emmstCK249, emmstCK401, and emmstNSS54). However, most emm types found among rural Australian Aborigines are associated with isolates from individuals on other continents, located far from the Top End of the Northern Territory, which suggests that other factors lead to the geographic partitioning of GAS strains.

**Discussion**

On the basis largely of studies of nontropical host populations, it was hypothesized that emm pattern A–C strains of GAS are more likely to cause throat infection, whereas emm pattern D strains have an innate tendency to cause impetigo [19]. The findings of the population-based survey of GAS in the Australian Aboriginal community that we studied are generally consistent with this hypothesis. The number of pyodermal infections in community members outnumbered throat carriage/infection by ~9-fold. This is reflected in the total number of GAS isolates defined by each emm pattern, whereby 61 pattern D isolates were collected, compared with only 23 pattern A–C isolates (a ratio of ~3:1). Thus, in a community where pyoderma is common and throat infection is rare, pattern D isolates predominated over pattern A–C isolates.

Surveillance in the study community was limited to a 25-month period. However, a total of 31 distinct emm types was
recovered during this time. Therefore, the possibility that the data set was biased by the introduction of only a few clones can be safely ruled out. Although some emm types became more dominant relative to others, the host population was exposed to a minimum of 31 distinct clones of GAS.

A large body of literature on epidemiological surveys conducted in the United States and Europe throughout the latter half of the 20th century supports the concept that distinct subsets of GAS throat and skin strains exist [3, 4, 8–10, 12]. Among GAS reported to have no obvious tissue site preference are the opacity factor-positive M types, which are represented by emm pattern E [35, 36]. Because GAS pharyngitis has been far more prevalent than impetigo in temperate regions, the vast majority are identified as having M types typically associated with emm patterns A–C or E [11, 19, 28, 29, 37, 38].

Although active surveillance in the tropical Aboriginal community demonstrates that ~85% of all GAS isolates are patterns D or E, similar population-based studies have not been conducted for noninvasive infections in nontropical host populations within the recent past. Invasive GAS disease in Connecticut (a population-based survey conducted March–August 1995) is overwhelmingly caused by emm pattern A–C and E strains (<5% pattern D). However, whether this distribution is a strict reflection of the total strains in the community is not known [29]. Clearly, future investigations should aim to broaden the scope of population-based surveillance for superficial GAS infections.

Geographic barriers to strain migration, by themselves, do not provide an adequate explanation for the predominance of emm pattern D and E strains among rural Aboriginal Australians. An alternative explanation is the existence of conditions that preferentially favor the transmission of emm patterns D and E strains over emm pattern A–C strains. Tropical climate, poor living conditions, and high rates of scabies infection can increase the risk for developing skin infection in the Aboriginal population [4, 5]. Like GAS, the scabies mite primarily infests the superficial layers of the skin, just below the stratum corneum, and has been implicated as a risk factor in the development of GAS pyoderma in both Trinidad and tropical Australia [6, 9].

Crowding is a major risk factor for nasopharyngeal infection [4]. Perhaps the increased incidence of GAS pharyngitis during colder weather in the United States and Europe is explained by the increased crowding (i.e., increased host density) that results from additional time spent indoors. However, rural Aboriginal communities are characterized by overcrowded housing, so the reasons for the low incidence of pharyngitis are not entirely clear. Conceivably, if the age-specific incidence for impetigo peaks during early childhood [4], multiple infections might lead to the development of non-type-specific protective immunity that in turn could serve as a protective factor against URT infections later in life. However, we currently lack evidence to support this idea.

Host genetic susceptibility is another potential risk factor that might conceivably influence the association between the human host and GAS of particular emm patterns. Nonetheless, even if the Australian Aboriginal subjects had a heightened genetic susceptibility to infection by emm pattern D isolates (or resistance to patterns A–C), it would not obviate the biological significance of the strong link observed between an abundance of emm pattern D isolates and a high incidence of pyoderma, rather than throat infection. In theory, a host genetic susceptibility could undermine the strength of the relationship between the tropics and emm pattern D, yet this idea seems inconsistent with the global trend of high pyoderma prevalence in the tropics [4–6, 9, 36].

Despite the low incidence of GAS recovery from the throat, pattern A–C throat isolates outnumbered pattern D throat isolates by >2-fold in the study community. Thus, for the individual subjects in whom risk factors for throat infection are present, pattern A–C strains predominate, reflecting an innate tendency of these organisms to infect this tissue. However, the magnitude of the emm pattern-based preference for the throat is weaker for this collection of tropical isolates than for previously reported nontropical isolates [19, 28]. This finding may be partly explained by differences in sampling: most throat isolates reported in previous studies were associated with clinical symptoms of disease, whereas none of the GAS isolated from the throats of Aboriginal Australians were associated with overt disease. Another contributing factor may be the high endemicity of tropical skin infection in community members, coupled with the preferred directionality for secondary infection—spread from the skin to the throat (rather than throat to skin) [23]—resulting in a preponderance of skin strains colonizing the URT. Yet another factor may be the strong seasonal influences on GAS infections in temperate regions. Seasonally dependent risk factors can affect transmission rates by leading to changes in the density of susceptible hosts. When seasonally dependent risk factors disappear, the density of susceptible hosts will drop, and, if it falls below a critical threshold, transmission will cease [39, 40]. Thus, in nontropical zones, the large temporal distance brought about by discrete seasonal peaks of pharyngitis and impetigo might act to further diminish the association of pattern A–C strains with the skin and of pattern D strains with the throat.

The seemingly apparent paucity of risk factors for pharyngitis among the Aboriginal host population is exemplified by the finding that pattern A–C isolates represent almost half of all throat isolates in the study community (albeit as an asymptomatic infection), yet they are present in 2.3-fold greater numbers in impetigo lesions (albeit as a minority population). Thus, pattern A–C strains are present in the host population and they constitute the predominant pattern at the throat, yet throat infection rates remain extremely low. With the seeming lack of risk factors for throat infection but in the presence of very high levels of skin infection, nearly all clones are obtained in higher numbers from the skin, despite their innate biological preferences. The successful infection of the skin by 2 particular pattern A–C strains (emm14 and emmst3765), largely in the ab-
sence of coinfection by typical skin strains (i.e., patterns D or E), is a finding that suggests the need for more extensive investigation. Their evolutionary relationships to other pattern A–C strains, compared with pattern D or E strains, may shed light on their biological behavior.

In addition to preferential infection of the URT, emm pattern A–C strains predominate among GAS associated with acute rheumatic fever in nontropical populations [19]. Yet, in the Australian Aboriginal population, which has the highest reported incidence of acute rheumatic fever in the world [6], throat carriage rates of GAS are typically low, and, as reported here, pattern A–C strains are uncommon and more often isolated from skin lesions. M protein serotypes of GAS associated with well-documented outbreaks of rheumatic fever in the United States and Europe are rarely found in the Aboriginal population [5, 6, 41]. This paradox is difficult to reconcile with the widely accepted premise that rheumatic fever is never triggered by a GAS skin infection [1–3]. It remains to be determined whether the few pattern A–C isolates that infect the throats of Aboriginal subjects are sufficient to account for all cases of rheumatic fever. It may also be that putative rheumatogenic factors are not limited to pattern A–C strains and that a clinically inapparent infection by a pattern D or E strain can readily trigger this autoimmune disease. Contrary to the widely held view, it is conceivable that an oropharyngeal site for GAS infection is not an absolute prerequisite for the subsequent development of rheumatic fever [42]. In light of the findings of this report, it seems important to reconsider the fundamental basis underlying the link between rheumatic fever and URT infection only.

The findings of this study may have important implications for the design of prospective vaccines against GAS infection. Strategies aimed at providing mucosal immunity at the URT rather than at systemic immunity may not adequately protect against skin infection. M type-specific vaccines that are based solely on emm pattern A–C organisms may not elicit protection against most GAS pyoderma infections. The finding of 31 distinct emm types within a single rural community over a short time dim the prospects for successful implementation of M type–based vaccines that are tailor-made for different host populations.

The association of emm pattern A–C organisms with the throat and emm pattern D organisms with skin infection is weaker in this tropical community than had been reported for nontropical regions [19]. Nonetheless, the prevailing trends are emergent on several orders of scale: the overall prevalence of GAS throat versus skin disease in a community, the total number of isolates recovered from a community that are defined for each emm pattern, and the relative proportion of each emm pattern at a given tissue site. Thus, the association of particular strains of GAS with specific tissue appears to be, in large part, genetically predetermined by the bacterium. It follows that key epidemiological features of GAS disease are profoundly shaped by the presence or absence of putative risk factors for human infection. It is the environmental and host risk factors that provide the selective pressures, which, in turn, differentially affect the reproductive capacity and ensuing spread of organisms belonging to each of the genetically distinct emm pattern groupings.

Acknowledgments

We thank Eric Peterson and Michelle Benitez at Yale University for excellent technical assistance.

References


