Rapid Emergence of a New Clone Impacts the Population at Risk and Increases the Incidence of Type emm89 Group A Streptococcus Invasive Disease


1Public Health Ontario, Toronto, Canada; 2Department of Laboratory Medicine and Pathobiology, University of Toronto, Canada; 3Mount Sinai Hospital, Toronto, Canada; and 4Department of Pathology and Genomic Medicine, Houston Methodist Hospital, Texas

Background. Invasive group A Streptococcus (iGAS) disease caused by type emm89 strains has been increasing worldwide, driven by the emergence of an epidemic clonal variant (clade 3 emm89). The clinical characteristics of patients with emm89 iGAS disease, and in particular with clade 3 emm89 iGAS disease, are poorly described.

Methods. We used population-based iGAS surveillance data collected in metropolitan Toronto, Ontario, Canada during the period 2000–2014. We sequenced the genomes of 105 emm89 isolates representing all emm89 iGAS disease cases in the area during the period and 138 temporally matched emm89 iGAS isolates collected elsewhere in Ontario.

Results. Clades 1 and 2 and clade O, a newly discovered emm89 genetic variant, caused most cases of emm89 iGAS disease in metropolitan Toronto before 2008. After rapid emergence of new clade 3, previously circulating clades were purged from the population and the incidence of emm89 iGAS disease significantly increased from 0.14 per 100 000 in 2000–2007 to 0.22 per 100 000 in 2008–2014. Overall, emm89 organisms caused significantly more arthritis but less necrotizing fasciitis than strains of the more common type emm1. Other clinical presentations were soft tissue and severe respiratory tract infections. Clinical outcomes did not differ significantly between emm89 clades overall. However, clade 3 emm89 iGAS disease was more common in youth and middle-aged individuals.

Conclusions. The rapid shift in emm89 iGAS strain genetics in metropolitan Toronto has resulted in a significant increase in the incidence of emm89 iGAS disease, with noticeably higher rates of clade 3 disease in younger patients.

Keywords. emerging strain genotype; group A Streptococcus; invasive disease; populations at risk; whole-genome sequencing.

Group A Streptococcus ([GAS] also known as Streptococcus pyogenes) causes a broad range of diseases, from pharyngitis and superficial skin infections to life-threatening necrotizing fasciitis and streptococcal toxic shock syndrome (STSS) [1]. After a decline in the mid-20th century, re-emergence of invasive GAS (iGAS) infections has led to disease incidences estimated at 1.5–4.3 cases per 100 000 in high-income countries [2–4]. Group A Streptococcus isolates can be differentiated into more than 200 types by deoxyribonucleic acid (DNA) sequence analysis of the region of the emm gene encoding the variable amino terminus portion of the M protein (a virulence factor with antiphagocytic properties) [5–7]. Invasive GAS disease burden is associated with a relatively small number of emm types in many high-income countries [8–10].

Investigative approaches that include whole-genome sequencing (WGS) of population-based samples have resulted in several important discoveries that have enhanced our understanding of iGAS diseases [11–14]. For example, analysis of >3500 emm1 GAS genomes has uncovered allele variation due to recombination in the promoter spacer region of genes nga and slo encoding secreted toxins NAD+-glycohydrolase (NADase) and streptolysin O (SLO), respectively [15, 16]. Enhanced expression of these key virulence factors by a recombinant epidemic emm1 clone led to the global spread of severe emm1 iGAS infections in the 1980s [16, 17]. Similar WGS-based approaches have begun to identify contributors to the sudden increase in emm89 iGAS disease reported in recent years in several countries [16, 18–21]. Analysis of genome-wide single-nucleotide polymorphisms (SNPs) defined 3 genetically distinct emm89 clades, which differ in virulence in mouse and nonhuman primate models of infection [15, 16, 22]. The more virulent emm89 clonal clade (named “clade 3”) has the same polymorphisms in the promoter spacer region of the nga and slo gene cluster (also resulting in enhanced NADase and SLO toxin production).
activity) as epidemic emm1 strains [16]. Furthermore, clade 3 emm89 strains lack the has genes required for the biosynthesis of the hyaluronic acid capsule typical of GAS [22]. Replacement of historic clades 1 and 2 by clade 3 temporally coincides with the increase in emm89 iGAS disease that has been documented in the United States and some European countries [16, 18].

Early reports described associations of emm89 GAS with skin infections [23], healthy children and young adult patients developing gastrointestinal symptoms and peritonitis [24, 25], and adults with necrotizing fasciitis and STSS [26]. Although our understanding of the biology of emm89 organisms has greatly advanced [15, 16, 18–20, 27], emm89 iGAS disease remains relatively poorly characterized clinically. Moreover, the clinical presentations and populations at risk for iGAS disease caused by the various emm89 clades have not been thoroughly investigated. To better understand these public health issues, we analyzed comprehensive population-based surveillance data for emm89 iGAS disease collected in metropolitan Toronto in the period 2000–2014, in combination with WGS analysis of the emm89 isolates responsible for these infections.

METHODS

Clinical Data, Isolates, and Laboratory Methods

The Toronto Invasive Bacterial Diseases Network (TIBDN) is a collaboration of all hospitals, microbiology laboratories, and public health units serving the city of Toronto and Peel region, Ontario, Canada (hereafter designated "metropolitan Toronto"; population 4.2 million in 2014). The TIBDN uses standardized methods and forms to collect clinical data including demographic information, disease manifestations, and underlying medical conditions from all patients with iGAS disease in the geographical area. Data used in this study are from January 1, 2000 to December 31, 2014; collection and usage was approved by the Research Ethics Boards of participating TIBDN institutions. Invasive GAS cases were defined as those in which acute illness occurred in association with the isolation of GAS from a normally sterile site (blood, cerebrospinal, pleural, peritoneal, pericardial, or joint fluid [including bursa], bone, aspirates, and tissue specimens or swabs obtained during surgery). Streptococcal toxic shock syndrome and necrotizing fasciitis were defined as previously described [27, 28].

We included 1 strain from each of the 105 emm89 iGAS disease cases recorded in metropolitan Toronto during the period (Table S1). We also analyzed clinical data and isolates from 138 additional emm89 iGAS cases that occurred in the province of Ontario outside of metropolitan Toronto (Table S1), recovered by hospitals that report and submit iGAS isolates to TIBDN on a voluntary basis. Clinical data were limited for these cases. Isolates were cultured at 37°C with 5% CO2 on Columbia blood agar plates containing 5% sheep blood or in Todd-Hewitt broth supplemented with 0.2% yeast extract. Isolates were confirmed to be GAS by β-hemolysis on sheep blood agar, grouping of carbohydrate antigen, large colony size, and bacitracin susceptibility [29].

Molecular Typing and Whole-Genome Sequencing of Group A Streptococcus Strains

Deoxyribonucleic acid was prepared from overnight cultures using the QIAamp DNA minikit (QIAGEN, Toronto, ON, Canada). emm typing was performed by polymerase chain reaction (PCR) and Sanger sequencing, as described previously [30]. We followed the procedures described by Nasser et al [17] to (1) prepare genomic libraries and perform paired-end Illumina genome sequencing of isolates, (2) identify polymorphisms against reference genomes, and (3) establish core-genome SNP-based phylogenies. We confirmed polymorphisms in the nga/slo promoter spacer region of all strains by PCR amplification and Sanger sequencing using previously reported primers [16]. Presence or absence of the 3-gene has locus was confirmed by PCR, as described previously [31].

Statistical Analysis

Statistical analysis was performed using SAS (version 9.3). Contingency tables were tested with 2-tailed χ² or Fisher’s exact tests, as appropriate. Differences in disease incidence and length of hospitalization between emm89 clades were evaluated with Poisson regression analysis and the Kruskal-Wallis test, respectively. P values <.05 were considered statistically significant.

RESULTS

The Incidence of emm89 Invasive Group A Streptococcus Disease Significantly Increased in Metropolitan Toronto Commensurate With Emergence of Clade 3 Strains

One hundred five of the 1596 (7%) iGAS disease cases in metropolitan Toronto between 2000 and 2014 were caused by emm89 strains. During this period, the overall incidence of iGAS infections remained relatively stable. In contrast, the incidence of emm89 GAS infections increased significantly from 0.14 per 100 000 in 2000–2007 to 0.22 per 100 000 in 2008–2014 (χ², P = .021) (Table S2). To test the hypothesis that this increase in incidence was due to the emergence of clade 3 emm89 strains, we used core-genome-SNP analysis to determine the phylogenetic relationships of the 105 emm89 iGAS isolates. Results revealed a genetically diverse emm89 population, including strains belonging to previously described clades 1, 2 and 3, which were genetically closely related to reference strains of those clades isolated elsewhere (Figure 1A). We also identified a fourth, distinct phylogenetic group, which we have named clade O. Strains of this clade were genetically distantly related to clades 1, 2, or 3 strains (Figure 1A), further supporting the notion that genetic diversity among emm89 GAS is vastly higher than that observed in some other emm types [15]. We next studied the temporal distribution of the strains. Clade 1 strains (17% of the emm89 iGAS isolates) predominated in the earlier years of this investigation (Figure 1B). Clade 2 strains (11% of the emm89 iGAS
isolates) were recovered primarily from 2005–2008. Clade O strains (14% of the total emm89 isolates) were found only in 2005–2007. The circulating emm89 strain population shifted dramatically in 2008, due to rapid expansion of clade 3 (57% of the total emm89 isolates), and 94% of those recovered from 2008–2014; Figure 1B), which effectively purged historic clades 1, 2, and O from the population. To investigate whether a similar replacement of historic clades by clade 3 occurred in the rest of Ontario, we sequenced the genomes of 138 emm89 isolates voluntarily submitted during the same period (2000–2014) from areas of the province outside of metropolitan Toronto (Figure S1A). A similar increase in the number of emm89 iGAS cases caused by clade 3 strains was observed in the rest of the province beginning in 2007 (Figure S1B and S1C).

**Patients With Invasive Group A Streptococcus Disease Caused by Strains of the Different emm89 Clades Differed in Comorbidities and Risk Factors**

At least 1 underlying illness potentially predisposing to iGAS disease was present in 76% of emm89-infected patients, and 37% of patients had more than 1 underlying illness. Although there was no significant difference in the frequency of underlying illness between the different emm89 clades, the distribution of underlying illness and risk factors was slightly different. Patients infected with clade 2 had higher rates of diabetes and cancer than patients infected with the other 3 clades (Table 1). This might be correlated with the observation that clade 2-infected patients had higher rates of STSS, intensive care unit (ICU) admission, and case fatality (Table 2). In contrast, we did not observe statistically significant differences in the length of hospitalization (median, 10 days) between patients infected with the different clades. More patients infected with clade 1 were admitted from nursing homes, whereas more patients infected with clade 2 and clade 3 strains were admitted from home than other clades (Table 1). Intravenous (IV) drug use and alcohol abuse was reported by 4% and 11% of emm89-infected patients, respectively, and all cases of self-reported IV drug use were associated with clades 1 or O strains (Table 1). None of these cases were clustered in space or time associated with homelessness, although, as a group, clade O strains occurred more frequently among homeless patients (Table 1). Although we did not observe clade 3 emm89 infections in metropolitan Toronto among IV drug users, 14 clade 3 patients from the rest of Ontario had a history of IV drug use (Table S3).
Overall, Clade 3 emm89 Strains Caused Invasive Group A Streptococcus Disease in Younger Patients

The 105 emm89 iGAS disease cases from metropolitan Toronto occurred in patients aged 8 months to 96 years old, 53% of whom were female. The median age at infection was 51 years old for females and 48 years old for males. Males and females and patients of different ages had slightly different disease manifestations (Figure 2). Adult males had more arthritis, and soft tissue infections were more common among the elderly. The greatest incidence of infection occurred in those patients >75 years old, and this age

Table 2. Clinical Presentation and Outcomes of Patients With emm89 and emm1 iGAS Infections in Metropolitan Toronto, 2000–2014

<table>
<thead>
<tr>
<th>Clinical Presentation/Outcome</th>
<th>Clade O No. (%)</th>
<th>Clade 1 No. (%)</th>
<th>Clade 2 No. (%)</th>
<th>Clade 3 No. (%)</th>
<th>emm89Total No. (%)</th>
<th>emm1 No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>15 (100)</td>
<td>18 (100)</td>
<td>12 (100)</td>
<td>60 (100)</td>
<td>105 (100)</td>
<td>397 (100)</td>
</tr>
<tr>
<td>Arthritis</td>
<td>2 (13)</td>
<td>4 (22)</td>
<td>2 (17)</td>
<td>9 (15)</td>
<td>17 (16)</td>
<td>32 (8)</td>
</tr>
<tr>
<td>Bacteremia without focus</td>
<td>0 (0)</td>
<td>3 (17)</td>
<td>3 (25)</td>
<td>7 (12)</td>
<td>13 (12)</td>
<td>61 (15)</td>
</tr>
<tr>
<td>Soft Tissue Infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Necrotizing fascitis</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (3)</td>
<td>2 (2)</td>
<td>36 (9)</td>
</tr>
<tr>
<td>Other soft tissue</td>
<td>10 (67)</td>
<td>4 (22)</td>
<td>1 (8)</td>
<td>20 (33)</td>
<td>35 (33)</td>
<td>138 (35)</td>
</tr>
<tr>
<td>Respiratory Tract Infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower respiratory</td>
<td>2 (13)</td>
<td>3 (17)</td>
<td>5 (42)</td>
<td>8 (13)</td>
<td>18 (17)</td>
<td>79 (20)</td>
</tr>
<tr>
<td>Upper respiratory</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td>1 (8)</td>
<td>7 (12)</td>
<td>10 (9)</td>
<td>20 (5)</td>
</tr>
<tr>
<td>Peripartum infection</td>
<td>0 (0)</td>
<td>1 (6)</td>
<td>0 (0)</td>
<td>4 (7)</td>
<td>5 (5)</td>
<td>8 (2)</td>
</tr>
<tr>
<td>Other*</td>
<td>0 (0)</td>
<td>2 (11)</td>
<td>0 (0)</td>
<td>3 (5)</td>
<td>5 (5)</td>
<td>23 (6)</td>
</tr>
<tr>
<td>STSS</td>
<td>1 (7)</td>
<td>5 (28)</td>
<td>5 (42)</td>
<td>14 (23)</td>
<td>25 (24)</td>
<td>117 (29)</td>
</tr>
<tr>
<td>Case fatality</td>
<td>1 (7)</td>
<td>4 (22)</td>
<td>5 (42)</td>
<td>10 (17)</td>
<td>20 (19)</td>
<td>81 (20)</td>
</tr>
<tr>
<td>ICU admission</td>
<td>4 (27)</td>
<td>6 (33)</td>
<td>5 (42)</td>
<td>16 (27)</td>
<td>31 (30)</td>
<td>158 (40)</td>
</tr>
</tbody>
</table>

Abbreviations: ICU, intensive care unit; iGAS, invasive group A Streptococcus; STSS, streptococcal toxic shock syndrome.

*Other includes peritoneal infection, gynecological infection not associated with pregnancy.

*Case fatality was defined as death that could be attributed to GAS infection within 30 days of positive culture.

*P < .05 for the comparison with all emm1 patients.
The group had the highest case fatality rate (Table S4). The median age of patients infected with historic clades 1, 2, and O was 51.4 years (range 6–96 years), whereas among clade 3-infected patients the median age was 49.3 years (8 months–90 years). Relative to the historic clades, the incidence of clade 3 infections was higher across all age groups, and particularly among those <19 years old, and middle-aged patients (those in their 30s and 40s) (Table S4). Although our data do not allow the calculation of precise incidence rates in areas of Ontario outside of metropolitan Toronto, we observed that in these areas the largest proportion of patients infected with clade 3 strains was in the 30–39 age group. Relative to the historic clades, significantly more clade 3 strains were isolated from patients aged 20–49 (Figure S1D). Taking all isolates from Ontario together, clade 3 strains infected proportionally more patients aged 20–49 and fewer patients aged 75+, relative to historic clades (Table S5).

Clinical Characteristics of \textit{emm89} Invasive Group A \textit{Streptococcus} Infections

The most common manifestation of \textit{emm89} iGAS disease was soft tissue infection, followed by lower respiratory infections and arthritis (Table 2). Streptococcal toxic shock syndrome occurred in 24% of \textit{emm89}-infected patients. The overall case fatality rate was 19%. To assess the features of \textit{emm89} disease, we compared the cohort of \textit{emm89} patients to that of patients with \textit{emm1} iGAS disease in the same population (Table 2). The most common disease manifestation among \textit{emm1} strains was also soft tissue infection, followed by lower respiratory infections and bacteremia without focus. Significantly more cases of arthritis and significantly fewer cases of necrotizing fasciitis were observed among patients with \textit{emm89} iGAS compared with those with \textit{emm1} disease (Table 2). There were no significant differences between clades with respect to the sites of isolation, although \textit{emm89} iGAS strains were less frequently isolated from blood than \textit{emm1} (Table 3). Overall, 76% of \textit{emm89}-infected patients and 70% of \textit{emm1}-infected patients in metropolitan Toronto had at least 1 underlying illness. A significantly greater proportion of \textit{emm89}-infected patients had diabetes compared with \textit{emm1}-infected patients (Table 1).
DISCUSSION

Systems biology approaches combining in-depth genomic strain characterization with in vitro testing and experimental infection of nonhuman primates has unambiguously demonstrated that the recently emerged genetic clade 3 emm89 GAS, which overexpresses the cytolytic toxins NADase and streptolysin O, is the main driver of the rapid worldwide increase in emm89 iGAS disease [15, 16, 18, 22]. Here, using similar WGS-based approaches and temporal analysis, we show that in metropolitan Toronto the emergence and rapid expansion of clade 3 clonal progeny is responsible for the significant increase in incidence of emm89 iGAS disease observed since 2008 in a context of relatively stable total iGAS disease burden. Emergence and rapid expansion of clade 3 in metropolitan Toronto was concurrent with rapid decline and apparent extinction of previously circulating emm89 clades 1, and 2, and O.

The paucity of reports examining the clinical features of emm89 iGAS disease is at odds with the magnitude of the increase in emm89 iGAS disease reported here and in several other countries [15, 16, 18, 20, 32]. To begin to address this circumstance, we assessed comprehensively the clinical features of emm89 iGAS disease using population-based surveillance data collected over 14 years. One of our findings was that emm89 iGAS disease is characterized by the very frequent occurrence of soft tissue, severe respiratory infections, and arthritis. The prevalence of arthritis among emm89 patients was significantly higher than among type emm1 strains in metropolitan Toronto and similar to that reported previously in Europe [23, 33]. In contrast, emm89 organisms caused significantly less necrotizing fasciitis than emm1 strains. We next investigated the hypothesis that, similar to what has been described in animal models [16, 22], clade 3 emm89 causes more severe disease in human patients than previously circulating clades. Although clade 3 caused slightly more necrotizing fasciitis, there were no significant differences in clinical diagnosis between patients infected with the different clades. Similar to a previous report [18], we also did not observe significant differences in 7- or 30-day mortality between clade 3 and other emm89 clade types. We did not detect significant differences between emm89 clades for pre-disposing conditions such as chronic illness or immune suppression. Although significantly fewer clade 3-infected patients were IV drug users in metropolitan Toronto, data from the rest of Ontario showed that IV drug usage was significantly associated with clade 3 strains.

We next examined whether clade 3 emergence and rapid expansion correlated with enhanced ability of the strains to cause disease in different groups of individuals than previously circulating clades, and we made several interesting observations. First, although studies of iGAS incidence by age are frequently bimodal, with rates peaking in children <2 years old and highest rates in those >85 years of age [26, 27, 34], we identified very few cases of historic clades 1, 2, and O in children <5 years old and in those aged 5–19, age groups in which clade 3 emm89 iGAS disease was more frequent. In addition, clade 3 cases were significantly more common in patients aged 20–49 years old, who less frequently reported underlying comorbidities. Thus, since the emergence of clade 3, there is a trend towards a profile of emm89 GAS infected patient that is younger, slightly healthier, and with a lower rate of underlying illness.

CONCLUSIONS

Clonal replacement among iGAS is a well described phenomenon [35, 36]. Given that the replacement of historic emm89 clades by clade 3 has occurred simultaneously in several unrelated countries, which have different models of access to healthcare, and whose populations are dissimilar in many social and human factors [16, 18], we speculate that emergence of this clone is dependent primarily on bacterial factors such as enhanced ability to persist and infect naive hosts rather than on host factors. The use of nonhuman primate models of infection has shown that enhanced NADase and SLO toxin activity has endowed clade 3 strains with increased fitness in the upper respiratory tract [15, 16]. It is interesting to note that emm89 pharyngitis also strikingly increased in frequency in 2007 in Ontario [21], likely related to the emergence of clade 3. We speculate that this enhanced ability to persist in
the upper respiratory tract may be one key contributor to the rapid dissemination of clade 3 and its ability to readily infect younger, healthier patients. Data presented here, data from the Centers for Disease Control and Prevention Active Bacterial Core surveillance, and our unpublished data for Ontario for the period 2015–2016 suggest that emm89 iGAS cases continue to occur in high numbers in North America. Very recent results from Finland have shown that diversifying clonal variants (subclades) of clade 3 emm89 have appeared that are associated with increased mortality [37]. Continued monitoring of changes in emm89 genetic diversity and associated iGAS disease is warranted.

**Supplementary Data**

Supplementary materials are available at Open Forum Infectious Diseases online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

**Acknowledgments**

We thank Chin-Yu (Amanda) Hong, Agron Plevneshi, and Wallis Rudnick (Toronto Invasive Bacterial Diseases Network) for database curation. We are grateful to Kelsie Jagt and Adriana Peci (Public Health Ontario) for statistical consultation. We also thank John McLaughlin (Public Health Ontario) for critical reading of an earlier version of this manuscript.

**Financial support.** This work was funded by the Fondren Foundation (to J. M. M.) and Public Health Ontario (internal grant PIF-2015-007; to N. E.).

**Potential conflicts of interest.** All authors: No reported conflicts.

**References**


