ORIGINAL ARTICLE

Distribution of virulence genes and the molecular epidemiology of Streptococcus pyogenes clinical isolates by emm and multilocus sequence typing methods

Siti Nur Adila Hamzah, MSc¹, Mohd Nasir Mohd Desa, PhD², Azmiza Syawani Jasni, PhD⁴, Niazlin Mohd Taib, Mpath¹, Siti Norbayu Masri, Mpath¹, Rukman Awang Hamat, Mpath¹

¹Department of Medical Microbiology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Selangor, Malaysia, ²Department of Biomedical Science, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

ABSTRACT

Background: Streptococcus pyogenes has a variety of virulence factors and the predominant invasive strains differ according to specific emm types and geographical orientation. Although emm typing is commonly used as the gold standard method for the molecular characterisation, multilocus sequence typing (MLST) has become an important tool for comparing the genetic profiles globally. This study aimed to screen selected virulence genes from invasive and non-invasive clinical samples and to characterise the molecular epidemiology by emm typing and MLST methods.

Materials and Methods: A total of 42 S. pyogenes isolates from invasive and non-invasive samples collected from two different tertiary hospitals were investigated for the distribution of virulence factors and their molecular epidemiology by emm and multilocus sequence typing methods. Detection of five virulence genes (speA, speB, speJ, ssa and sdaB) was performed using multiplex polymerase chain reaction (PCR) using the standard primers and established protocol. Phylogenetic tree branches were constructed from sequence analysis utilised by neighbour joining method generated from seven housekeeping genes using MEGA X software.

Results: Multiplex PCR analysis revealed that sdaB/speF (78.6%) and speB (61.9%) were the predominant virulence genes. Regardless of the type of invasiveness, diverse distribution of emm types/subtypes was noted which comprised of 27 different emm types/subtypes. The predominant emm types/subtypes were emm63 and emm18 with each gene accounted for 11.8% whereas 12% for each gene was noted for emm28, emm97.4 and emm91. The MLST revealed that the main sequence type (ST) in invasive samples was ST402 (17.7%) while ST473 and ST318 (12% for each ST) were the major types in non-invasive samples. Out of 18 virulotypes, Virulotype A (five genes, 55.6%) and Virulotype B (two genes, 27.8%) were the major virulotypes found in this study. Phylogenetic analysis indicated the presence of seven different clusters of S. pyogenes. Interestingly, Cluster VI showed that selected emm/ST types such as emm71/ST318 (n=2), emm70.1/ST318 (n=1), emm44/ST31 (n=1) and emm18/ST442 (n=1) have clustered within a common group (Virulotype A) for both hospitals studied.

Conclusion: The present study showed that group A streptococci (GAS) are genetically diverse and possess virulence genes regardless of their invasiveness. Majority of the GAS exhibited no restricted pattern of virulotypes except for a few distinct clusters. Therefore, it can be concluded that virulotyping is partially useful for characterising a heterogeneous population of GAS in hospitals.

KEYWORDS:

Emm type; Multilocus sequence typing; Streptococcus pyogenes; virulence genes

INTRODUCTION

Group A Streptococcus (GAS) or better known as Streptococcus pyogenes possesses various virulence factors that are involved in severe life-threatening infections. Invasive S. pyogenes infections have contributed to high lethality rates ranging from 10 to 30%, resulting in more than 600,000 deaths worldwide. Virulence factors such as superantigens (SAgs), adhesins, proteases and leukocidins play a major role in the pathogenesis of S. pyogenes diseases. For instance, adhesins are groups of proteins which are responsible for the initial attachment of S. pyogenes to epithelial cells.

In the presence of other virulence factors such as M protein and streptokinase, speB gene can contribute to GAS pathogenicity. Interestingly, specific serotypes of highly virulent S. pyogenes strains are recognised to express more than one fibronectin-binding proteins. The M1 and M49 serotypes were found to produce a cellular surface fibronectin-binding protein (FboA), which has a pivotal role in S. pyogenes invasion into deeper tissues. Superantigens such as streptococcal pyrogenic exotoxins (SpeA, SpeC and SpeG to SpeM) and the streptococcal superantigen (SSA) have been shown to mediate massive uncontrolled inflammatory reactions by overstimulating the host inflammatory cells mainly T-cell lymphocytes in severe invasive S. pyogenes infections. These of some these toxin genes are exclusively...
expressed in certain S. pyogenes serotypes. The M1 serotype that has a specific combination of speA and smeZ genes were found to increase the pathogenicity of this strain.17

To date, studies on the distribution of S. pyogenes serotypes and their virulence genes are limited in Malaysia.13 Moreover, previous findings on S. pyogenes serotypes and virulence factors were characterised based on M-typing using specific antisera which is obsolete currently as the reagents are not widely available.18 Besides, some of the S. pyogenes strains could not be typed by the antisera, thus limits the use of M-typing as the typing method of choice.4,19 The emm typing method has now been accepted as the gold standard since M-typing is encoded by the 5′ end of the hypervariable region of emm gene, and this gene can be sequenced for the typing purposes.11 Meanwhile, S. pyogenes genetic lineages can be characterised and globally compared using the multilocus sequence typing (MLST) method.20 This method utilises seven highly conserved house-keeping genes and different nucleotide sequences can be determined to characterise different sequence types.21 Previous data showed that certain emm types are associated with different types of S. pyogenes disease manifestations and invasiveness and the epidemiology of emm types is geographically oriented.22,23 Therefore, to better understand the epidemiology of S. pyogenes infections in Malaysia, this pilot study aimed to characterise the selected virulence genes of S. pyogenes from various clinical specimens via emm typing and MLST methods.

**MATERIALS AND METHODS**

The present study was conducted in 2018 by utilising the previous collection of 42 S. pyogenes isolates which were obtained from two different tertiary hospitals from the year 2014 to 2015. The distance between these two hospitals is approximately 30km apart. The isolates were stored at -70°C and re-identification of the isolates was performed using the Gram staining method, bacitracin susceptibility (Oxoid, Basingstoke, United Kingdom), PYR test (Oxoid, Basingstoke, United Kingdom), latex agglutination (Oxoid, Basingstoke, United Kingdom), and species-specific polymer chain reaction (PCR) method.15 The isolates were collected from blood (n=10), pus (n=22), tissue (n=7), wound (n=2) and throat (n=1). The sources for the isolates were categorised into invasive and non-invasive samples based on Creti et al.15 The approval to conduct the study was obtained from the Ethics Committee for Research involving Human Subjects of Universiti Putra Malaysia with the reference number of UPM/TNCPI/RMC/1.4.18.2.

**Multiplex polymerase chain reaction for the detection of selected virulence genes**

Fresh bacterial colonies growing on Columbia agar enriched with 5% of sheep blood (Isolab Sdn. Bhd, Shah Alam, Selangor, Malaysia) were used for DNA extraction using the HiYield Genomic DNA kit (Real Biotech Corporation, Taipei, Taiwan) in accordance to the manufacturer's instructions. DNA extracts were stored at -30°C for further use.

Determination of five virulence genes (speA, speB, spef, sdaB, and ssa) was carried out using a multiplex PCR kit (Qiagen, Germantown, USA) using few sets of primers (Table 1) according to the multiplex PCR protocol described in a previous study.17 An aliquot of DNA template (1µL) was transferred into an Eppendorf tube containing 12.5µL Qiagen® Multiplex PCR master mix with HotStarTaq DNA polymerase (Qiagen, Germantown, USA), 0.5µL of each virulence gene primer pair and 4.5µL multiplex PCR buffer. DNA amplification was performed using a Bio-Rad thermal cycler (Bio-Rad Laboratories Ltd., Watford, Hertfordshire, UK) with an initial denaturation process at 95°C for 3 min, followed by 35 cycles (30s at 94°C, 90 s at 57.4°C, 90s at 72°C) and the final extension process at 72°C for 10min. The PCR products were analysed with gel electrophoresis using 2% agarose gel containing 1× TBE buffer and 1µL of gel stain (Nanogenone Solutions Sdn Bhd, Batu Caves, Selangor, Malaysia). The gel was then viewed using a gel documenting system Alpha image TM 2200 (Alpha Innotech Cooperation, San Liandro, USA). The multiplex PCR products of five virulence genes were sequenced and blasted with the GenBank sequences for the similarity index.

**The emm typing of S. pyogenes**

The emm typing was performed according to the recommended protocol by the Centers for Disease Control and Prevention (CDC) (http://www.cdc.gov/ncidod/biotech/strep/protocols.htm). Bacterial DNA of all the isolates were prepared and PCR technique was used for DNA amplification using a Bio-Rad thermal cycler (Bio-Rad, California, USA). A set of forward and reverse primers were used as follows: 5′-TATT(CG)GCTTAGAAATTA-3′ and 5′-GCCTAATCCAATGTTT-3′, respectively. The PCR cycling conditions used were as follows: 94°C for 15 s, 46°C for 30 s, and 72°C for 75 s for the first 10 cycles, and then 94°C for 15 s, 46°C for 30 s and 72°C for 75 s with a 10s increment for each of the subsequent 19 cycles). DNA purification and sequencing were performed by the 1st Base Laboratory Sdn. Bhd., Seri Kembangan, Malaysia. Received sequences were then edited using Bioedit software version 7.0 (https://bioedit.software.informer.com/7.0/) and compared with reference sequences using the BLAST algorithm.

**Multilocus sequence typing (MLST) and phylogenetic study**

The MLST was performed by sequencing seven housekeeping genes (gki, gtr, muri, mutS, recP, xpt, and yqiL) for all the isolates according to the established protocol with slight modifications.24 A total of 4µL DNA template was mixed with 25µL of green master mix (Vivantis Technologies Sdn. Bhd., Subang Jaya, Selangor, Malaysia), 1.0µL of each housekeeping primer pair, and 19µL of PCR buffer. The PCR amplification was performed using a Bio-Rad thermal cycler (Bio-Rad, California, USA). An initial denaturation process was conducted at 95°C for 5 min, followed by 35 cycles (45 s at 55°C, 90 s at 72°C) and the final extension process at 72°C for 1 min. The sequencing analysis of DNA was conducted by MytacG Sdn. Bhd, Kajang, Selangor (Malaysia) and the data was submitted to the NCBI website for determination of gene similarity percentage using the BLAST analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequence alignment was then analysed using a MEGA X software (https://www.megasoftware.net/home). In each locus, the different sequence was assigned by a distinct allele number, generating a seven-integer allelic profile for each isolate. Isolates with similar allelic profiles were assigned to the similar sequence type (ST). A complete database of alleles,
The gel electrophoresis patterns of multiplex PCR products are shown in Figure 1. Table III shows the distribution of emm types, subtypes and sequence type (ST) according to the type of samples in this study.

RESULTS

Among the 42 S. pyogenes isolates, the predominant virulence genes detected were as follows: sdaB (78.6%), speB (61.9%), speJ (52.4%), speA (47.6%) and ssa (42.9%). The distribution of five virulence genes in the invasive and non-invasive samples is shown in Table II. The gel electrophoresis patterns of multiplex PCR products are shown in Figure 1.

DISCUSSION

Streptococcus pyogenes is a human pathogen that is responsible for multiple infections globally. The pathogenic properties of S. pyogenes are often associated with the production of virulence factors such as superantigens, proteinases and adhesins. In particular, streptococcal erythrogenic exotoxins (SPEs) are involved in a massive inflammatory response and tissue destruction. Thus, it is very important to determine the potential of S. pyogenes strains and categorise them according to the MLST method for global comparison. Moreover, superantigen profiles of S. pyogenes strains are associated with specific emm types and their association differs across different countries.
Distribution of virulence genes and the molecular epidemiology of Streptococcus pyogenes clinical isolates

Fig. 1: Agarose gel electrophoresis patterns showing multiplex PCR amplification product for the S. pyogenes virulence genes. Lane M (Nanogene Solutions Sdn. Bhd., Batu Caves, Selangor, Malaysia): DNA molecular size marker (100bp). The specific virulence genes are labelled with arrows. Lane 1 to 10 represents isolates sampled from Hospital Serdang; Lane 11 to 15 represents isolates sampled from Hospital Kuala Lumpur.

Fig. 2: Phylogenetic relationship among S. pyogenes strains inferred with the neighbor-joining method using concatenated sequences of seven housekeeping genes used for MLST. The taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Cluster V and VI showed S. pyogenes strains with two restricted virulotypes (A and L) and only Virulotype A, respectively in two different hospitals. HS=Hospital Serdang; HKL=Hospital Kuala Lumpur. Virulotype A=five genes; Virulotype D and N=four genes; Virulotype E, F, H, J, L and Q=three genes; Virulotype B, I, K and O=two genes; Virulotype C, G, M and R (one gene); Virulotype P=no gene detected.
In the present study, all *S. pyogenes* isolates had at least one virulence gene regardless of their source. Comparatively, all five virulence genes (*sdaB, speB, speF, speA* and *ssa*) were commonly detected in *S. pyogenes* isolates collected from non-invasive samples but not the invasive samples, which could be due to a slightly higher number of non-invasive samples in the present study (Table I). Nonetheless, high-frequency rates of similar virulence genes have also been reported among *S. pyogenes* strains from non-invasive samples in previous studies.\(^{21,24}\) This is not surprising as the distribution of virulence genes varies due to the differences in the acquisition of the genes which can be chromosomally encoded or mediated by mobile genetic elements.\(^{22}\) Both *streptococcal pyrogenic exotoxin A* (*speA*) (65.0%) and *streptococcal superantigen* (*ssa*) (88.9%) genes are highly transmissible via horizontal gene transfer through mobile genetic elements.\(^{21}\) The emergence of phage-encoded exotoxins would create an unusual virulent clone among certain population due to selective pressures for bacterial fitness.\(^{24}\) Moreover, the lack of specific immunity among the study population towards such strains may pose a substantial risk of a streptococcal outbreak in future. Similar findings have been reported where high frequency of *speA* gene was detected in non-invasive cases (pharyngitis) in Taiwan and Norway.\(^{1,21,27}\) However, lower frequency rates have been reported in recent studies in Pakistan (38%) and India (25.1%).\(^{24,28}\) The *SpeA*, *SpeF* and *SSA* toxins have been commonly associated with severe *S. pyogenes* infections such as toxic shock-like syndrome, multiorgan failures and scarlet fever.\(^{25}\) *SpeB*, a cysteine protease is a potent pro-inflammatory inducer and is commonly associated with necrotising fasciitis.\(^{21}\) In the present study, a slightly low frequency of *speB* gene (61.9%) was noted from the analysis of *S. pyogenes* isolates. The *speB* gene was shown to be highly conserved in *S. pyogenes* isolates (100%) in some studies while few studies have reported lower frequency rates ranging from 0 to 60%.\(^{21,24}\) The differences in the prevalence rates of *speB* gene could possibly be explained by the different strains in certain geographical regions, the number and types of samples and the methods used for *speB* gene detection. Interestingly, streptococcal DNase B (*sdbB*) gene was frequently detected among non-invasive than invasive samples in this study (54.5% versus 45.5%) and only absence in isolates collected from HKL. This finding can be associated with the common observation in phylogenetetic tree. The *sdbB* gene has been designated as a *streptococcal pyrogenic exotoxin F* (*speF*) gene that could directly damage pulmonary endothelial cells in a mouse model.\(^{21}\) This chromosomally encoded gene causes increased permeability of lung blood vessels which is a risk factor for acute respiratory distress syndrome (ARDS) cases.\(^{21,25}\) The high-frequency rate of *sdbB/speF* gene (100%) among invasive strains was reported in a few studies.\(^{21,24}\) Inversely, lower frequency rates of *sdbB/speF* gene were reported in recent studies.\(^{24,25}\)

A diverse distribution of *emm* types/subtypes among *S. pyogenes* was observed, in which 27 *emm* types/subtypes were detected (Table III). No new *emm* types/subtypes were detected. The results obtained align with other findings which reported no dominancy of a single *emm* type/subtype.\(^{24,25}\) Moreover, it has been documented that *S. pyogenes* strains in developing countries have diverse *emm* types compared to developed countries.\(^{24}\) The *emm* types/subtypes were also widely distributed among invasive and non-invasive samples in the present study which is inconsistent with findings from other studies where certain *emm* types (*emm1, 3, 6, 12 and 89*) were specifically found in invasive strains.\(^{1,2}\) In general, the frequencies of five virulence genes were higher in non-invasive than invasive isolates (Table I) and serious attention is needed as phage-encoded superantigens in non-invasive *S. pyogenes* strains can easily be transferred to invasive *S. pyogenes* strains and other non-pathogenic streptococci via horizontal gene transfer.\(^{24}\) Nevertheless, the findings from this study are in contrast with other previous findings where a higher percentage of virulence genes was noted in invasive than non-invasive samples.\(^{7,24}\)

In the present study, few prominent clusters with restricted patterns of virulotypes were observed by phylogenetic analysis (Figure 2). The findings are in accordance with a study that demonstrated certain *emm* types shared the common toxin gene profiles.\(^{24}\) Balaji and colleagues reported that different *emm* types had distinct toxin-gene profiles but a phylogenetic analysis was not investigated among their isolates.\(^{24}\) The *spe* genes provide as marker for horizontal gene movements and encode the function of exotoxin in GAS pathogenesis. Surveillance that includes invasive (tissue sources) and non-invasive (pus) GAS isolates is important to distinguish between virulence properties and the prevalence of a particular GAS strain in the general population and to evaluate epidemiological changes in GAS diseases.\(^{24}\)

Specific dominant virulotypes (A and L) were prevalent in both hospitals, and close monitoring via molecular typing methods is urgently required. Toxin gene profiling (virulotyping) has been proposed in many studies to support other molecular typing methods such as the MLST for genotypic determination of *S. pyogenes*.\(^{1,2}\) Besides, it was reported that several *emm* types have specific toxin gene profiles which were reflected by the spread of specific invasive clones in European countries.\(^{7,24}\) The *emm1/ST28* isolates that exhibited virulotype A and L in this study could pose a risk of clonal transmission in both hospitals. Hypervirulent characteristic of *emm1/ST28* strain is typically associated with high fatality rates among invasive streptococcal cases in some countries.\(^{26,27}\) The different *emm*/ST types that exhibited virulotype A in this study could be explained by the restricted or lack of transfer of phage-encoded superantigen genes within this clone. It is still unknown whether a single virulotype detected in a clone could be due to the underlying biological factors or the selective advantage of *S. pyogenes* strains with certain *emm*/ST types. Nonetheless, different patterns of toxin gene profiling within multiple *emm* types are usually reflected by ongoing horizontal transfer of phage-encoded superantigen genes over time.\(^{21}\) Thus, continuous molecular surveillance is needed to identify the emergence of novel lineages of *S. pyogenes* local strains.

This study has several limitations. The total sample was too small in the present study for the statistical analysis to be carried out resulting in findings from this study could not be generalised to other Malaysian hospitals. However, as this study in the first report on virulence characteristics of *S. pyogenes* in Malaysia, it is very important to investigate the current
molecular epidemiology of the local strains for controlling the potential clonal spread of S. pyogenes isolates in hospitals across the country.

CONCLUSION

S. pyogenes is genetically diversified, and apart from the emm typing and MLST methods, virulotyping is essential to characterise the heterogeneous nature of S. pyogenes strains. Future study with a larger sample number of S. pyogenes isolates from different sources is needed to support the findings from this study. Continuous monitoring of S. pyogenes via molecular methods is warranted in the future. Thus, potential nosocomial outbreaks of invasive clones can be controlled accordingly.

ACKNOWLEDGMENT

The authors would like to thank the pathology department from Hospital Serdang and Hospital Kuala Lumpur for their cooperation during sample collection. This work was funded by the Universiti Putra Malaysia Research Grant (GP-IPS/2018/9656600).

REFERENCES


