Genotypic characterization of Malaysian human isolates of *Streptococcus pneumoniae* from carriage and clinical sources

Nik Noorul Shakira Mohamed Shahrin, Siti Norbaya Masri, Niazlin Mohd Taib, Syafinaz Amin Nordin, Farida Jamal, Mohd Nasir Mohd Desa

**A R T I C L E   I N F O**

Article history:
Received 25 June 2014
Received in revised form 8 October 2014
Accepted 20 October 2014

Keywords:
*Streptococcus pneumoniae*
Serotype
Virulence
Pili
Vaccine
PBP-RFLP
BOX-PCR

**A B S T R A C T**

This study characterized carriage and clinical pneumococcal isolates for serotypes, penicillin susceptibility, virulence genes and restriction fragment length polymorphism (RFLP) pattern of penicillin binding protein (PBP) genes. DNA fingerprint of isolates was generated by BOX-PCR. Majority of serotypes were 23F followed by 19F, 19A and 6A. Twenty-four percent of isolates were penicillin non-susceptible (PNSP). All of the targeted virulence genes were detected in all isolates with the exception of pili; 20.6% (*n* = 22) for PI-1 and 14.0% (*n* = 15) for PI-2. Of the 13 isolates which carried both PI-1 and PI-2, 10 were of clinical origin. Digested *pbp*-DNA produced three PBP-RFLP profiles for *pbp1a* (A1 to A3), six profiles for *pbp2b* (B1 to B6) and seven for *pbp2x* (X1 to X7) mostly in PNSPs. Based on BOX-PCR analysis, the majority of isolates were genetically diverse with a small number of potentially related isolates carrying pili genes. No obvious genotypic association was observed pertaining to carriage and clinical origin of isolates.

© 2014 Elsevier Ltd. All rights reserved.

**1. Introduction**

*Streptococcus pneumoniae* (pneumococcus) causes invasive pneumococcal diseases (IPD) and non-invasive pneumococcal diseases (non-IPD) with significant morbidity and mortality worldwide [1]. There are more than 90 serotypes of *S. pneumoniae* which differ in virulence, geographical distribution and extent of drug resistance. Current pneumococcal vaccines such as pneumococcal conjugate vaccines (PCV) provide a limited coverage due to serotype-specific protection. Currently there are PCV7, PCV10 and PCV13 where the number represents the number of included serotype, respectively [1,2].

Penicillin has been the drug of choice for pneumococcal disease treatment since the 1940s. However, treatment for pneumococcal diseases is complicated by the increasing rate of resistance to penicillin and other β-lactams antibiotics [2,3]. Resistance to penicillin and β-lactams in *S. pneumoniae* arises from alterations in penicillin-binding proteins (PBPs) which are located in the bacterial cell wall. Among the PBPs, modifications of *pbp1a*, *pbp2b* and *pbp2x* contribute to their decreased affinity to those antibiotics, which leads to development of penicillin and β-lactams resistance in pneumococci [4].

---

**E-mail addresses:** mnasir@upm.edu.my, nasirdes@yahoo.com (M.N.M. Desa).

http://dx.doi.org/10.1016/j.cimid.2014.10.005
0147-9571/© 2014 Elsevier Ltd. All rights reserved.
**S. pneumoniae** possesses several virulence properties such as autolysin A (LytA), pneumococcal surface adhesin A (PsAA), pneumococcal surface protein A (PspA), pneumolysin (Ply), choline binding protein A (CbpA), choline binding protein G (CbpG), pneumococcal adherence and virulence factor A (PavA) and the multimeric filaments; the pili [5–7]. These proteins and enzymes have been proposed to have distinct roles in pneumococcal pathogenesis in different host niches [8]. There are two different pilus islets (Pils) that encode the structural and biosynthetic genes for two antigenically different types of pili, PI-1 [5] and PI-2 [9]. The PI-1 was reported to play a role in virulence in murine models of pneumococcal infection [5] and PI-2 was proposed to play a role in pneumococcal adhesion to human respiratory cells [10].

BOX-PCR typing method targets a dispersed repetitive motif in the genome of *S. pneumoniae*. There are approximately 25 BOX repetitive elements in non-coding regions dispersed throughout the pneumococcal genome with three discriminate regions: boxA (59 bp), boxB (45 bp) and boxC (50 bp) [11]. The multicycopy BOX elements have been recognized to be useful for rapid molecular discrimination of pneumococcal strains and have been successfully employed elsewhere [12–16].

Analysis on the interrelation of serotypes, antibiotic susceptibility, virulence gene distribution and genomic DNA fingerprint patterns in *S. pneumoniae* from different sources of isolation may provide useful epidemiological perspective and genetic linkage of the studied pneumococcal population. Comparison of such data from various localities may as well indicate the geographical variation of the organism. Subsequently, the findings can be used as guidelines in managing the pneumococcal prophylactic measures either locally or globally [2,3,13,15]. In Malaysia, *S. pneumoniae* is also an important disease-causing pathogen and yet data on local isolates, particularly those at genetic level are still lacking. In this study, we characterized a collection of human pneumococcal isolates, derived from both community and hospital, for their genotypic properties covering serotypes, penicillin susceptibility, virulence genes availability, PBP genes profile (*pbp1a, pbp2b* and *pbp2x*) and BOX-PCR based DNA typing. Comparative analysis was conducted to find potential correlations.

## 2. Material and methods

### 2.1. Ethics statement

The Universiti Putra Malaysia (UPM) Medical Research Ethics Committee found no objectionable ethical issues of this study. Study did not involve human subjects except some clinical and demographic data associated with the isolates that are no longer traceable to the sampled individuals.

### 2.2. Pneumococcal isolates

One hundred and seven viable *S. pneumoniae* isolates were available in our stock culture collection consisting isolates from both carriage (community) and clinical (hospital) sources. Carriage isolates (*n* = 50) were collected from anterior nares of healthy children of five years or younger in 2010 as described previously [17]. Clinical isolates (*n* = 57) were isolated from 1989 to 2012 from patients of five years or younger (*n* = 6), five to twelve years (*n* = 5), 13 to 50 years (*n* = 18) and above 50 years (*n* = 12). Sixteen clinical isolates came from patients of unknown age. The clinical isolates were randomly obtained from various isolation sites from admitted patients at a number of hospitals in Malaysia. Thirty-seven were from sterile anatomical sites (blood, cerebrospinal fluid (CSF) and pleural fluid) and the rest were various including sputum, eye, ear, pus, throat, nasopharyngeal fluid and others (Fig. 1). All 107 isolates were obtained from a different individual.

### 2.3. Pneumococcal identification and serotyping

All isolates were re-identified as *S. pneumoniae* by standard bacteriological methods consisting of colonial morphology, α-hemolysis on 5% Sheep Blood Agar, Gram stain, catalase reaction, bile solubility and susceptibility to erythromycin, erythromycin, clindamycin, and optochin. Forty-one of the carriage isolates were previously serotyped by multiplex PCR in our previous study [18]. The remaining 66 isolates were serotyped using similar PCR-approach in this study with an additional primer set (set F) for detecting serotypes/serogroup 12F, 22F, 15A/F and 8 (Table 1).

### 2.4. Penicillin susceptibility testing

The minimum Inhibitory Concentration (MIC) value for penicillin was determined by Epsilometer Test (E-Test) method (AB Biodisk, Solna, Sweden) following procedures recommended by the Clinical and Laboratory Standard Institute (CLSI) guidelines [19]. The MIC results obtained were compared to the stated CLSI interpretive criteria; for clinical isolates from CSF, the criterion for meningitis group with parenteral penicillin was referred; MIC ≤0.06 μg/ml is susceptible and MIC ≥0.12 μg/ml is resistant. As for the rest of clinical isolates, the criterion for non-meningitis with parenteral penicillin was referred; MIC ≤2, 4 and ≥8 μg/ml for susceptible, intermediate and resistant, respectively. Meanwhile, for carriage isolates, the criterion for oral penicillin was used, with MIC ≤0.06, 0.12–1, and ≥2 μg/ml for susceptible, intermediate, and resistant, respectively. *S. pneumoniae* ATCC 49619 with known zone of inhibition size and MIC was used as the reference strain throughout the testing.

### 2.5. Genomic DNA extraction

Pure bacterial cells, grown overnight on Columbia Agar supplemented with 5% sheep blood incubated in the presence of 5% CO2 at 37 °C were prepared for the DNA extraction. GeneAll® Exgene™ (GeneAll Biotechnology Co. Ltd, Korea) kit was used for the extraction in accordance to the instructions provided by the manufacturer.
Fig. 1. Dendrogram of BOX-PCR DNA fingerprint cluster analysis of 107 Streptococcus pneumoniae isolates. The dendrogram was generated by BioNumerics using the unweighted pair-group arithmetic mean method (UPGMA). The percentage of similarity is measured by the top scale. The dashed vertical line indicates 80% dendrogramatic similarity level.
<table>
<thead>
<tr>
<th>Target gene</th>
<th>Forward primer sequence (5’ to 3’)</th>
<th>Reverse primer sequence (5’ to 3’)</th>
<th>Amplicon size (bp)</th>
<th>Annealing temperature (°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ply</td>
<td>ATTTCCCTGAAAGCCTAGCAACCA</td>
<td>GAATTCCTGGCTTCTCTCAAGTC</td>
<td>348</td>
<td>53</td>
<td>[40]</td>
</tr>
<tr>
<td>lytA</td>
<td>CAGCGTGTTGAACCGTATTGGA</td>
<td>TGTTGTTTATGCTTTAGCTGAAAA</td>
<td>173</td>
<td>53</td>
<td>[41]</td>
</tr>
<tr>
<td>pspA</td>
<td>CCGGATCGAGGAGGTCTATGTTGGGT</td>
<td>CCAATAGGTTTCTGTTGCGTCTCCACC</td>
<td>800–1000</td>
<td>55</td>
<td>[42]</td>
</tr>
<tr>
<td>pspB</td>
<td>CTCTCCTCAGGTTAGCCTTGGTG</td>
<td>CCAGAAGAATGAAAGGAGA</td>
<td>838</td>
<td>53</td>
<td>[43]</td>
</tr>
<tr>
<td>pspC</td>
<td>TGAAATCTCTGGTCTTGGT</td>
<td>TAAGTTCCCTTTCTTACTACAG</td>
<td>1177</td>
<td>55</td>
<td>[34]</td>
</tr>
<tr>
<td>sipA</td>
<td>CTGCTAGGAGGATCTCCTATCTATC</td>
<td>CTACAGGGCGTCTTGCTTG</td>
<td>550</td>
<td>55</td>
<td>[38]</td>
</tr>
<tr>
<td>PBp genotyping</td>
<td>CTCCTGCTCAGGTTAGCCTTGG</td>
<td>2400</td>
<td>55</td>
<td>[20]</td>
<td></td>
</tr>
<tr>
<td>pbp1a</td>
<td>CCGATCGAGGATCTCCTATCTATC</td>
<td>GAATTCCTGGCTTCTCTCAAGTC</td>
<td>348</td>
<td>53</td>
<td>[40]</td>
</tr>
<tr>
<td>pbp2b</td>
<td>GAGTCTCTAAATGAGTTAGGAGGAGGAGGAGG</td>
<td>CCAGAAGAATGAAAGGAGA</td>
<td>838</td>
<td>53</td>
<td>[43]</td>
</tr>
<tr>
<td>pbp2x</td>
<td>CTCCTGCTCAGGTTAGCCTTGG</td>
<td>2400</td>
<td>55</td>
<td>[20]</td>
<td></td>
</tr>
</tbody>
</table>

* Serotyping utilized primer sets A–F; sets A–E are as published [18] and only set F is shown.

2.6. Detection of S. pneumoniae virulence/pili genes

Detection of S. pneumoniae virulence (lytA, ppy, pspA, pspA, cbpA, cbpC, and pava) and pili (rlhA for P-I and sipA for PI-2) genes were carried out using published primers (Table 1) in BioRad MyCycler™ Thermal Cycler (BioRad, USA). The reactions involved 5 μl of i-DNA PCR mastermix (i-DNA Biotechnology, Singapore), 0.5 μl of DNA template (10–20 ng), 0.5 μl (10 μM) of each primer and 18.5 μl of PCR grade water. PCR were performed as follows: initial denaturation step at 94 °C for 5 min, 30 amplification cycles of denaturation at 94 °C for 30 s, annealing temperature (according to respective primers, Table 1) for 30 s and elongation temperature at 72 °C for 30 s to 1 min. The amplification was completed with a final elongation at 72 °C for 5 min. All PCR products were electrophoresed on 2% agarose gel stained with GelRed™ (Biotium®, US) for 1 h at 90 V. Genetic polymorphism of the pbp1a, pbp2b and pbp2x was investigated by RFLP analysis. PCR products were digested using FastDigest® Hinfl kit (Fermentas, USA) based on the instructions provided. Digested DNA was electrophoresed at 90 V for 60 min on 2% agarose gel incorporated with GelRed™ (Biotium®, US). The 100 bp plus DNA ladder marker (Vivantis, Malaysia) was included in every electrophoresis run for molecular weight reference. DNA fingerprints of the genes were examined and photographed under UV illuminator (GelDoc 1000 System, Bio-Rad, USA). The RFLP profiles produced from pbp1a, pbp2b and pbp2x were coded as “A,” “B” and “X,” respectively.

2.7. Genotyping of PBP genes by restriction fragment length polymorphism (RFLP) analysis

Genes pbp1a, pbp2b and pbp2x were amplified using previously described primers [20] (Table 1). Optimized PCR reactions were performed in a total volume of 25 μl. Reaction mix contained 10 μl of Expand High Fidelity Buffer (10 x) with 15 mM MgCl2 (Roche, Germany), 0.5 μl of DNA template (10–20 ng), 0.5 μl (10 μM) of each primer (forward and reverse) for respective genes, 0.8 μl of PCR Grade Nucleotide Mix (10 mM of each dNTP) (Roche, Germany), 1.0 μl of Expand High Fidelity Enzyme Mix (Roche, Germany) and 11.7 μl of PCR grade water. Amplification parameters in BioRad MyCycler™ Thermal Cycler (BioRad, USA) consisted initial denaturation at 94 °C for 5 min followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 2 min and elongation at 72 °C for 3 min. Final elongation was performed at 72 °C for 7 min. Five microliter of PCR products were electrophoresed on 1.0% agarose gel stained with GelRed™ (Biotium®, US) for 1 h at 90 V. Genetic polymorphism of the pbp1a, pbp2b and pbp2x was investigated by RFLP analysis. PCR products were digested using FastDigest® Hinfl kit (Fermentas, USA) based on the instructions provided. Digested DNA was electrophoresed at 90 V for 60 min on 2% agarose gel incorporated with GelRed™ (Biotium®, US). The 100 bp plus DNA ladder marker (Vivantis, Malaysia) was included in every electrophoresis run for molecular weight reference. DNA fingerprints of the genes were examined and photographed under UV illuminator (GelDoc 1000 System, Bio-Rad, USA). The RFLP profiles produced from pbp1a, pbp2b and pbp2x were coded as “A,” “B” and “X,” respectively.

2.8. BOX-PCR

BOX-PCR was performed in a total volume of 25 μl using 22-mer primer, BOX-A1R 5’-CTACGCGCAAGCGCACGCTAGC-3’ [13] in BioRad MyCycler™ Thermal Cycler (BioRad, USA). PCR reactions included 5 μl of i-DNA PCR mastermix (i-DNA Biotechnology, Singapore), 1.0 μl of DNA template (10–20 ng), 0.6 μl (10 μM) of BOX-A1R primer and 13.4 μl of PCR grade water. PCR parameters were as follows; an initial denaturation step at 94 °C for 4 min, followed by 30 cycles of denaturation at 94 °C for 1 min; annealing at 50 °C for 2 min and extension at 72 °C for 8 min. A single final extension was performed at 72 °C for 5 min. PCR products were electrophoresed on 1.5% (w/v) agarose stained with GelRed™ (Biotium®, US) for 70 min at 70 V with 100 bp plus DNA ladder marker (Vivantis, Malaysia) for reference.
The generated DNA Fingerprints were analyzed by visual inspection and using Bionumerics software version 7.1 (Applied Maths, Kortrijk, Belgium). A dendrogram was generated by Dice coefficient. Patterns were clustered with the unweighted pair group method with arithmetic mean (UPGMA). A similarity of 80% was used to define a cut-off for the BOX-PCR types.

2.9. Statistical analysis

The distribution of the source of isolation, vaccine serotypes and genes (PI-1/PI-2) of the isolates was analyzed using Pearson Chi-square test (two-tailed) in a $2 \times 2$ contingency table. Statistical significance was set at a $P$ value $\leq 0.050$.

3. Results

3.1. Serotyping and penicillin susceptibility

Of the 107 isolates, eleven different serotypes with four different serogroups were obtained, where further differentiation in serotypes could not be confirmed using the multiplex method. A further four isolates were unable to be serotyped as they did not amplify products for any of the molecular targets and were thus designated as non-typeable (NT) in this study [21–23]. All NT isolates were from the carriage group. Based on our collection, the most dominant serotype was $23F$ ($n = 25$) followed by serotypes $19F$ ($n = 18$), $19A$ ($n = 15$), $6A$ ($n = 15$), $15B/C$ ($n = 7$), $6B$ ($5$), $14$ ($n = 4$), $1$ ($n = 4$), $11A/D$ ($n = 3$), $7A/F$ ($n = 2$), $23A$ ($n = 1$), $18A/B/C$ ($n = 1$), $17F$ ($n = 1$), $4$ ($n = 1$) and $12F$ ($n = 1$). Eighty-one isolates (75.7%) were susceptible to penicillin, grouped as penicillin susceptible $S. pneumoniae$ (PSSP), $10$ (9.3%) were penicillin intermediate $S. pneumoniae$ (PISP) and $16$ (15.0%) were penicillin resistant $S. pneumoniae$ (PRSP). The PISPs and PRSPs were collectively referred as penicillin non-susceptible $S. pneumoniae$ (PNSP). The occurrence of PNSP was highest among serotype $6A$ (13 of 15), followed by $23F$ (5 of 25), $19F$ (3 of 18), NT (3 of 4) and $6B$ (2 of 5). Based on source of isolation, $88.5\%\ (n = 23)$ of the PNSP isolates were from the carriage group and the remaining $11.5\%\ (n = 3)$ were clinical isolates.

3.2. Presence and distribution of virulence/pili genes

In all 107 isolates, $lytA$, $ply$, $psaA$, $pspA$, $cbpA$, $cbpG$, and $pavA$ were detected by PCR. However, only $20.6\%$ of the isolates ($n = 22$) amplified the targets for PI-1 and $14.0\%\ (n = 15)$ for PI-2. The PI-1 was detected in $S. pneumoniae$ isolated from year $1989\ (n = 2)$, $1996\ (n = 2)$, $2010\ (n = 16)$ and $2012\ (n = 2)$ with $14$ of clinical origin and eight carriage isolates. The majority of isolates with PI-2 were from year $2010\ (n = 12)$ and the remaining were from year $2012\ (n = 3)$. Of these, $11$ were clinical isolates and $4$ were of carriage. Thirteen isolates was observed to carry both PI-1 and PI-2, mostly from year $2010\ (n = 11)$ and the rest from $2012\ (n = 2)$; $10$ isolates were of clinical origin while three isolates were from the carriage group.

The most dominant serotype to carry $rlfA$ was $19F\ (n = 13)$, followed by serotypes $23F\ (n = 3)$, $6B\ (n = 4)$, $4\ (n = 1)$ and NT ($n = 1$). The majority of $19F$ isolates also possessed the $sipA$ gene ($n = 13$), which was also observed in serotype $11A\ (n = 1)$ and $7A/F\ (n = 1)$. All isolates with both PI-1 and PI-2 were from serotype $19F\ (n = 13$). Pearson Chi-square analysis indicates a significant relationship between the PI-1 and/or PI-2 presence within the PCV7- and PCV10-serotypes ($P \leq 0.050$) but not PCV13 serotypes ($P > 0.050$) (Table 2). No significant association was suggested between the presence of PI-1 and PI-2 with the source of isolation as well as penicillin susceptibility of the isolates, respectively ($P > 0.050$).

3.3. PBP-RFLP profiles

Digested $pbp$-DNA produced three PBP-RFLP profiles for $pbp1a$ (assigned as A1 to A3), six profiles for $pbp2b$ (assigned as B1 to B6) and seven profiles for $pbp2x$ (assigned as X1 to X7). Overall, the PNSP isolates produced variable profiles of RFLP for $pbp1a$, $pbp2b$ and $pbp2x$ which is consistent with the altered $pbp$-genes. The profiles observed were A2 and A3 for $pbp1a$, B2 to B6 for $pbp2b$ and X3 to X7 for $pbp2x$. The $pbp2x$ showed most diverse profiles dispersed among isolates with various penicillin MIC values as compared to $pbp1a$ and $pbp2b$ genes. Generally, the majority of isolates with penicillin MIC values less than $1.0 \mu g/ml$ had uniform PBP-RFLP profiles ($pbp1a$; profile A1, $pbp2b$; B1, $pbp2x$; X1 or X2). Based on PBP-RFLP combination profile, profile A2B4X5 was commonly observed among PI-1 and PI-2 positive isolates. Meanwhile, two NT isolates were found to share the unique $pbp2b$ profile B6 that was not observed in other isolates. Another unique $pbp2b$ profile, B5, was observed in one isolate with penicillin MIC value $0.38 \mu g/ml$. PSSP isolates from year 1989 shared the same PBP-RFLP profile with PSSP isolates from recent years; 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2010 and 2012 (for $pbp1a$ and $pbp2b$; A1 and B1 profiles, and for $pbp2x$; X1 and X2 profiles), indicating that $pbp$ genes were conserved in PSSP over the years. As for PNSP, most of the isolates with a penicillin MIC of greater than $1.5 \mu g/ml$ exhibited a PBP-RFLP combination of A3B2X3.

3.4. BOX-PCR

The 107 $S. pneumoniae$ isolates were assigned into 17 clusters by BOX-PCR analysis (Fig. 1). Eighty-four of the isolates grouped with at least one other isolate and were assigned cluster labels of BOX-PCR types I to XVII. The remaining 23 isolates produced unrelated BOX-PCR profiles. Interestingly isolates with penicillin MIC of $1.0 \mu g/ml$ or greater were largely clustered in either BOX-PCR type III or VII, while isolates with PI-1 and PI-2 were observed to be more clustered in BOX-PCR type VII. BOX-PCR type III, the second largest cluster with $87.2\%$ similarity, included 11 isolates mostly from year 2010 and of carriage origin. Serotypes $6A$ and $23F$ interchangeably predominated but those with higher MIC regardless of the serotypes in this cluster had a similar PBP-RFLP pattern of A3B2X3, which were not found in other clusters except in BOX-PCR type V. Meanwhile, BOX-PCR type VII, the largest cluster with $81.6\%$ dendrogramatic similarity, consisted 16 clinical isolates of predominantly serotype $19F$ isolated in year 1994.
Table 2
Comparative analysis of source of isolation, vaccine serotypes, PI-1 and PI-2 availability.

<table>
<thead>
<tr>
<th>Source</th>
<th>Vaccine serotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Carriage</td>
</tr>
<tr>
<td></td>
<td>PCV 7</td>
</tr>
<tr>
<td>PI-1 (rhlA)</td>
<td>PCR positive</td>
</tr>
<tr>
<td></td>
<td>PCR negative</td>
</tr>
<tr>
<td></td>
<td>P value</td>
</tr>
<tr>
<td>PI-2 (sipA)</td>
<td>PCR positive</td>
</tr>
<tr>
<td></td>
<td>PCR negative</td>
</tr>
<tr>
<td></td>
<td>P value</td>
</tr>
<tr>
<td>PI-1 &amp; PI-2 (rhlA &amp; sipA)</td>
<td>PCR positive</td>
</tr>
<tr>
<td></td>
<td>PCR negative</td>
</tr>
<tr>
<td></td>
<td>P value</td>
</tr>
</tbody>
</table>

* Significant P values in bold.

until year 2012. Nevertheless, only five isolates of serotype 19F in this cluster showed a similar PBP-RFLP profile of A2B4X5, which were not found in other clusters except in one isolate in cluster XI. On the other hand, almost all serotypes 19F in this cluster carried both PI-1 and PI-2 gene. BOX-PCR type I was the third largest cluster but consisted of PSSP carrying a PBP-RFLP profile (A1B1X2) similar to that of the other PSSP’s in other small clusters and unrelated isolates.

4. Discussion

The top four serotypes observed in this study: 23F, 19F, 19A and 6A showed an agreement with findings by other studies conducted in Malaysia and other Asian countries [2,24,25]. With regard to penicillin susceptibility, carriage isolates contributed as the majority of the PNSP group with serotype 6A being the most dominant. However, the frequency of PNSP and serotype in this study may not truly reflect the current scenario in this country due to inclusion of old clinical isolates that spanned from the 1980s. Nevertheless, at a certain extent, this study may indicate the persistence of penicillin resistance and certain serotypes in this country.

The susceptibility interpretive criterion used in this study was meant for clinical purposes by taking into account the isolation sites and diseases apart from the MIC values alone. As far as in vitro cellular analysis is concerned, PBP genes in this study showed alterations at all except as MIC 1.0 μg/ml. This could potentially lead to some phenotypic change of isolates in responding to the penicillin concentration as compared to that at a lower MIC. These penicillin resistant markers might be conserved over the years as a similar PBP-RFLP profile was observed in isolates in year 2001 as well as in year 2010. In this study, PNSPs carrying the altered PBP genes was found to be more common in carriage isolates possibly due to the potentially prolonged colonization of isolates which could allow more extensive selection for PNSP to thrive at the carriage sites. In general, PNSPs have also been isolated in hospital settings particularly in non-invasive clinical isolates [2].

Comparison in relation to the various anatomical sites of isolation, temporal variation and ages of patients were not done due to unfeasible distribution to allow a fair categorical analysis. The former only involved clinical isolates which were isolated in various years and largely penicillin susceptible based on the CLSI’s definition. Meanwhile, carriage group were all isolated from anterior nares in a single year of 2010. As for ages, all carriage isolates were from the same age group (≤5 years old), while only six clinical isolates from children of a known similar age group. The remainder of the clinical isolates was from a range of age groups; and some of unknown ages. Nevertheless, although the carriage isolates were all from young healthy children, they may potentially have some age or site-related features which could associate them with certain genotypic features different from those of the clinical group, which were largely from teens, adults and elders of various isolation sites. Unfortunately, no conclusions can be made as there was no obvious association seen in this study.

On the other hand, data on the virulence and pili gene distribution in this study may support the current approach in focusing on proteins as vaccine candidates for pneumococcal disease prevention. Previously, there are two types of licensed vaccines; the pneumococcal capsular polysaccharide vaccine (PPV) and PCV. The PPV has minimal efficacy in children of less than two years of age, among the elders and in immune-deficient patients such as HIV-infected persons [26]. Thus, several PCVs were introduced after the PPV to overcome these limitations; PCV7, PCV10 and PCV13 targeting serotypes commonly associated with diseases [2]. However, an increasing trend of acute otitis media (AOM) and IPD incidence caused by non-PCV serotypes of S. pneumoniae have been reported [27–29]. For the latter, the use of PCV may have also resulted in selective pressure for the emergence and invasive proliferation of non-PCV serotypes [21,22]. In developing countries where pneumococcal diseases were reported to be high, the usage of PCVs is limited due to their high cost [17]. Thus, although PCV-serotypes were observed in this study, it is not known whether these isolates were related to the clonal dissemination of pneumococcal isolates that emerged due to the usage of the PCVs in developed countries. An example is serotype 19A (included in PCV13) which was reported to be frequently resistant to penicillin in many parts of the world especially in the United States [30]. However, isolates with serotype
19A in this study were penicillin susceptible indicating that the isolates in this study could be limited to a local setting.

As an alternative vaccination strategy, some common pneumococcal virulence proteins, including pili have been studied in limited strains to evaluate their potential to be used as future improved vaccine candidates [5,7,31]. Our findings suggest that pneumococcal isolates have largely similar regimens of virulence properties regardless of the origin, serotype and penicillin susceptibility of the isolates. This indicates a good remark for the respective proteins as potential vaccine candidates with a wider coverage. On the other hand, presence of PI-1 and PI-2 was limited to only a small proportion of the isolates. Similarly, other studies also showed a limited presence of PI-1 and PI-2. For example, PI-1 has been reported to be present in only 14.4% to 35.2% in pneumococcal isolates [10,32–36]. In this study, the PI-2 was detected in, but not restricted to, serotypes 19F, 23F, 6B and 8 (included in PCV7, PCV10 and PCV13) as also observed in other studies [33,34,37]. Whereas the second type of pili, PI-2 has been reported elsewhere to be present in 16% to 21% of pneumococcal isolates, and frequently from serotypes 1, 2, 7F, 19A, and 19F [9,38]. In this study, the PI-2 was detected among three serotypes/serogroup; 19F, 11A and 7A/F. Interestingly, the majority of the isolates with both PI-1 and PI-2 gene in this study were from clinical sources.

In the genomic DNA typing analysis, our data revealed that the PBP-RFLP profiles, serotypes and penicillin susceptibility patterns were not fully conserved among isolates within one major BOX-PCR cluster. Instead, those carrying both pili determinants were mostly grouped together in a major cluster. The former differences could be attributed to random mutations, horizontal transfer of PBP genes, capsular genetic switch and exchange of other genetic determinants between those isolates and other pneumococcal strains or viridans streptococci as reported in other studies [4,15,17]. Such events may have taken place over time to result in genetic profile diversity among isolates sharing similar BOX-PCR fingerprints specifically those in cluster VII in this study. Our findings are in agreement with other studies [15,16] that isolates in similar BOX-PCR cluster may not necessarily exhibit a homogeneity in the various genotypes.

In conclusion, taking the isolates as a whole, this study shows consistencies with other studies for serotypes prevalence and altered pbp genes in PNSP. Nonetheless, the relationship against their isolation sources remains unknown. In regards to BOX-PCR analysis, the majority of isolates in this study appear genetically distinct, with only a small number of potentially related isolates carrying the pili genes. Additionally, the gene targets for LytA, Ply, PsaA, PspA, CbpA, CbpG and PavA appear to be widely conserved among the pneumococci in this study, indicating that these targets may be worthwhile future vaccine candidates. As for pili, although isolates with both PI-1 and PI-2 are mostly present in clinical isolates and possibly related to certain serotypes, their number is low to claim their importance in clinical setting. More studies involving isolates from wider geographical sites are warranted to establish the findings.

**Acknowledgments**

This study was funded by a Research University Grant Scheme (No. 04-02-12-1764RU) of Universiti Putra Malaysia and partly by a Fundamental Research Grant Scheme (No. 01-03-11-987FR) of the Ministry of Education, Malaysia.

**References**


