Short Communication

Ethnic Differences in the Prevalence, Clinical Outcome and cag Pathogenicity Island (cagPAI) Virulence Gene Profiles of *Helicobacter pylori* Strains from Malaysia

Hamat, R. A.¹*, Nor Amalina, E.², Malina, O.¹, Zamberi, S.¹, Alfizah, H.³, Rizal, A. M.⁴, Aminuddin, A.⁵ and Ramelah, M.⁶

¹Department of Medical Microbiology and Parasitology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
²School of Medicine, University Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
³Department of Medical Microbiology and Immunology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia
⁴Department of Community Health, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia
⁵Faculty of Medicine, Universiti Teknologi MARA Selayang Campus, Jalan Prima Selayang 7, 68100 Batu Caves, Selangor, Malaysia
⁶Collaborative Innovation Centre, Universiti Kebangsaan Malaysia, Bangi, 43600 Bangi, Selangor, Malaysia

ABSTRACT

Different *Helicobacter pylori* genes may be well conserved within different ethnic groups and could give rise to different clinical outcomes. In this study, we demonstrated a low prevalence of *H. pylori* infection (19.2%) which is in concordance with the current trend demonstrated locally and abroad. The Indians had the highest prevalence of *H. pylori* infection among other ethnic groups (Malays= 8.6 %, Chinese= 24.3 %, Indians= 33.9%). cagM and cagT were the most predominant genes found (63.4% for each), followed by cagA (62.2 %), cagE (48.2%), cag6-7 (46.3%), cag10 (42.1%), cag13 (4.9%) and IS605 (3.7%). No significant association was found between *H. pylori* infection and *H. pylori* genes with ethnic groups or clinical outcomes. Indians who had a combination of cagA/E/M genes of *H. pylori* were likely to be associated with 21-time of having non-ulcer dyspepsia (NUD) than peptic ulcer disease (PUD). Therefore, these genes may serve as useful markers in predicting the clinical presentation of a *H. pylori* infection among...
Indians in our studied population. Hence, this preliminary data might explain why Indians have a low prevalence of gastric cancer and peptic ulcer disease despite having persistently high prevalence of \textit{H. pylori} infection for many decades (“Indian enigma”) in Malaysian patients.

\textit{Keywords}: \textit{Helicobacter pylori}, \textit{cag} pathogenicity island, virulence genes, peptic ulcer diseases, non-peptic ulcer dyspepsia, ethnicity, Malaysia

\section*{INTRODUCTION}

\textit{Helicobacter pylori} is present in more than 50\% of the world’s population (Peterson, 1992). This unique bacterial species is extremely diverse in its genomic structures. This panmicticism is believed to be important in colonization and infection in the human gastro-duodenal system. A variety of diseases including gastritis, gastric ulcers (GU), duodenal ulcers (DU) and gastric cancer have been attributed to \textit{H. pylori}. The prevalence rates of \textit{H. pylori} vary between populations or sub-populations (ethnic groups) within the same geographic locations (Suerbaum & Michetti, 2002). Data from endoscopic surveys and sero-prevalence studies amongst the three Malaysian ethnic groups revealed that Indians have the highest prevalence of \textit{H. pylori} infection (68.9\% - 75.0\%) compared to Chinese (45.0\% - 60.6\%), while Malays consistently have the lowest prevalence rate (8.0\% - 43.3\%). Ironically, in terms of the severity of disease caused by \textit{H. pylori}, gastric cancer occurs more frequently in Chinese (68.0\%) rather than Indians (16.5\%) and Malays (15.5\%)(Haron \textit{et al.}, 1994; Kang \textit{et al.}, 1990). In addition, it has been documented that peptic ulcer disease (PUD) is rather low among Indians compared to Chinese (Owen \textit{et al.}, 2001; Goh, 1997). This paradox which is known as the “Indian enigma” (high prevalence of \textit{H. pylori} but low percentage of gastric cancer and PUD) may be explained by differential virulence of the \textit{H. pylori} strains or different susceptibility genotypes in the Indian ethnic group. \textit{cagPAI}, a 40-kb region of first chromosomal DNA is the major genetic determinant of \textit{H. pylori} virulence that consists of 27 genes and it is divided into two segments, \textit{cag} I and \textit{cag} II, by an insertion sequence known as \textit{IS605} (Censini \textit{et al.}, 1996). There is now increasing evidence for the existence of several novel bacterial pathogenicity markers on the \textit{cag} (cytotoxin-associated gene) pathogenicity island (\textit{cagPAI}) that may play important roles in the disease invoking potential of \textit{H. pylori} (Mattar \textit{et al.}, 2007; Van Doorn \textit{et al.}, 1998). Meanwhile, the overall genetic variability of \textit{H. pylori} may also be held responsible for the great diversity of clinical outcomes (Proença-Módena \textit{et al.}, 2007; Maeda \textit{et al.}, 1999). However, this has not been well explored locally in Malaysia and detailed data is lacking. Therefore, we conducted a study to investigate the role of several selected genes in the \textit{cagPAI} (\textit{cagA}, \textit{cagE}, \textit{cagM}, \textit{cagT} and \textit{cag6-7}, \textit{cag10}, \textit{cag13} represented the \textit{cag} I and \textit{cag} II segments, respectively and \textit{IS605}) versus ethnicity and clinical outcomes. To our knowledge, this is the largest set of \textit{cagPAI} genes that has
been studied so far in *H. pylori* from clinical isolates in Malaysia.

**MATERIALS AND METHODS**

A total of 855 gastric biopsies were obtained from dyspeptic patients who had undergone endoscopy from May 2004 to May 2007 in Universiti Kebangsaan Malaysia Medical Center. These patients were enrolled by a purposive sampling. Selection of patients and study protocols were followed according to a previous local study (Ramelah *et al.*, 2005). The study protocol was approved by the Medical Ethics Committee of Universiti Kebangsaan Malaysia (FF-075-2003). *H. pylori* culture was performed as previously described with modifications (Proença-Módena *et al.*, 2007). Briefly, specimens for culture were immediately transported to the laboratory in Stuart transport medium (Oxoid, UK). Biopsies were sub-cultured onto Columbia agar base medium (Oxoid Ltd., Basingstoke, Hampshire, England) containing Dent’s supplement and 7% lysed horse blood. The cultures were incubated under microaerophilic conditions (5–6% O₂, 8–10% CO₂, 80–85% N₂, and a relative humidity of at least 95%) at 37°C for a minimum period of 5 days. The bacteria were identified as *H. pylori* based on the typical Gram stain morphology and biochemical tests such as urease, oxidase and catalase reactions. A pool of bacterial colonies obtained from each single plate was used for DNA extraction. Two *H. pylori* reference strains: American type culture collection (ATCC) 700824 (strain J99) and ATCC 43256 were used in this study. Chromosomal DNA of each strain was prepared using the High Pure PCR Template Preparation kit (Roche, Mannheim, Germany) in accordance with the manufacturer’s instructions.

PCR analysis for cagPAI was performed to amplify *cagA*, *cagE*, *cagM*, *cagT*, *cag6-7*, *cag10*, *cag13* and *IS605* genes. The primers and PCR conditions used for each gene have been reported previously by Maeda *et al.* (1999). Briefly, the PCR was performed in a total reaction volume of 25µl containing 2.5µl 10 X PCR buffer, 1µl of each primer (each at 10pmol/µl), 2 mM MgCl₂; 200µM of dNTP, 5 U of super Taq (Super Taq DNA Polymerase Mbiotech) and 1µL (10ng) of genomic DNA from the culture lysates. DNA extracts of *H. pylori* ATCC 700824 and *H. pylori* ATCC 43256 were used as positive controls while negative controls (without DNA) were added to each PCR run. PCR products were visualized by electrophoresis in 1.2% agarose gel, stained with ethidium bromide, and examined under UV illumination. Standard 1-kb DNA ladders (Fermentas Life Science, Hanover, MD) were used as molecular size markers.

Chi-squared and Fisher’s exact tests were used to determine the difference in the distribution of *H. pylori* genes and the correlation between genes, clinical disease and ethnicity. *p* values of less than 0.05 were considered statistically significant.

**RESULTS AND DISCUSSION**

In this study, we reported a lower overall prevalence of *H. pylori* amongst our studied population (19.2%; 164 out of 855) compared
to 49% and 22% by Goh et al. (1997) and Mahadeva et al. (2005), respectively. This is not surprising as high standard of living conditions and improved personal hygiene could contribute to major decline of *H. pylori* infection in Asian populations (Tan & Goh, 2008). Interestingly, our finding corroborates data from a study conducted in a different locality, i.e. the north eastern part of Peninsular Malaysia which was 19.3% (Raj et al., 2001). With regard to ethnic differences, a similar pattern was observed; Indians had the highest prevalence rate of *H. pylori* infection (33.9%; 42/124) compared to Chinese (24.3%; 98/404) and Malays (8.6%; 28/327) (data not shown). This is very much consistent with previous local findings as well (Mahadeva et al., 2005; Goh et al., 1997; Kang et al., 1990) and could be explained by the role of genetic heterogeneity in susceptible hosts or other co-factors that could possibly account for the inter-ethnic variations (Alm & Trust, 1999).

For instance, *cagA* gene in *cagPAI* region of *H. pylori* strains was responsible for more severe pathological changes and clinical outcomes in Western countries (Crabtree et al., 1993; Xiang et al., 1993). Surprisingly, similar correlations were not found in some Asian countries (Zheng et al., 2000; Maeda et al., 1998). *cagA* gene was reported to be insufficient in inducing some disease processes in *H. pylori* infection in China (Hu et al., 1995). Tan et al. (2005) reported 84% of their patients who had *cagA* gene were not associated with clinical outcomes. In addition, no significant correlation was observed in the prevalence of this gene among the Malays (76.6%), Chinese (86.4%) and Indians (86.8%). In our previous report, we could not reveal the association between either the *cagA* or 3′ end region of *cagA* gene with gastroduodenal diseases. However, Chinese (96%) in whom have been known to have a higher risk of peptic ulcer disease and gastric cancer had the highest prevalence of *cagA* subtype A strains compared to Malays (72%) and Indians (69%). The difference was statistically significant (p<0.0005) (Ramelah et al., 2005). In a recent work by our group, Eastern CagA strains were significantly predominance among the Chinese (82.9%) compared to Malays (11.4%) and Indians (1.4%). However, no association could be made between Eastern CagA strains and the clinical outcome (Mohamed et al., 2009). It has been known that Eastern CagA strains confer higher tyrosine CagA phosphorylation activity than Western strains, which is related to their pathogenic potentials (Higashi et al., 2002a). As previous works were only focussing on *cagA* and its variants, and could not demonstrate the association with clinical outcome, we hypothesized that other potential virulence genes on *cagPAI* could also be involved in *H. pylori* infection among our multi-ethnic population. Maeda et al. (1999) evaluated 15 genes on the *cagPAI* and concluded that *cagA* gene could not be used as a virulence marker in Japanese *H. pylori* strains. Recently, a systematic mutagenesis study of individual genes on *cagPAI* by Fischer et al. (2001)
confirmed that 17 and 14 out of 27 genes were involved in *H. pylori* pathogenesis by encoding a type IV secretion system for the translocation of *cag* toxicity protein A (CagA) and for the full induction of interleukin 8 (IL-8). Among the eight genes, *H. pylori cagM* and *cagT* genes were both predominantly found in our study population (63.4% for each) (TABLE 1). To the best of our knowledge, this is the first report on the prevalence of *cagM* and *cagT* genes in Malaysia; nonetheless, we could not reveal any significant association in the prevalence of *cagM* and *cagT* genes as well as other cagPAI genes with clinical outcome (TABLE 2). Goh et al. (2007) suggested that other virulent genes could account for the paradoxical findings of *H. pylori* infection

### TABLE 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Ethnic group</th>
<th>Malay n = 28 (%)</th>
<th>Chinese n = 94 (%)</th>
<th>Indian n = 42 (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td><em>CagA</em> (n=102)</td>
<td>Positive</td>
<td>16 (57.1)</td>
<td>12 (42.9)</td>
<td>62 (66.0)</td>
<td>32 (34.0)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>CagE</em> (n=102)</td>
<td>Positive</td>
<td>13 (46.4)</td>
<td>15 (53.6)</td>
<td>44 (46.8)</td>
<td>50 (53.2)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>CagM</em> (n=102)</td>
<td>Positive</td>
<td>16 (57.1)</td>
<td>12 (42.9)</td>
<td>59 (62.8)</td>
<td>35 (37.2)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>CagT</em> (n=102)</td>
<td>Positive</td>
<td>16 (57.1)</td>
<td>12 (42.9)</td>
<td>62 (66.0)</td>
<td>32 (34.0)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cag6-7</em> (n=102)</td>
<td>Positive</td>
<td>8 (28.6)</td>
<td>20 (71.4)</td>
<td>59 (62.8)</td>
<td>35 (37.2)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cag10</em> (n=102)</td>
<td>Positive</td>
<td>12 (42.9)</td>
<td>16 (57.1)</td>
<td>38 (40.4)</td>
<td>56 (59.6)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cag13</em> (n=102)</td>
<td>Positive</td>
<td>2 (7.1)</td>
<td>26 (92.9)</td>
<td>3 (3.2)</td>
<td>91 (96.8)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>IS605</em> (n=102)</td>
<td>Positive</td>
<td>0 (0.0)</td>
<td>28 (100.0)</td>
<td>3 (3.2)</td>
<td>91 (96.8)</td>
</tr>
</tbody>
</table>

*Chinese vs. Malay (Pearson Chi-square, 2 x 2 table; χ² = 10.189, p = 0.001)
Chinese vs. Indian (Pearson Chi-square, 2 x 2 table; χ² = 19.842, p = 0.000)
and clinical outcome in our multi-ethnic groups. They found that only 23.3% of \textit{H. pylori}-positive Indians had gastric cancer compared to 82.3% and 60.5% in Malays and Chinese, respectively. Mohamed \textit{et al.} (2009) reported Western CagA strains were significantly found in Indian patients (43.5%), which could probably explained lower incidence of gastric cancer or PUD in this ethnic group. In our study, when each ethnic group was analyzed, the association was only statistically significant in the proportion of positive \textit{cagA/E/M} genes amongst Indians who had NUD than PUD (p<0.047) (TABLE 3). In addition, Indians with the combination of \textit{cagA/E/M} genes were likely (twenty-one times) associated with NUD than PUD. These findings are rather surprising as the presence of these markers has been associated with high risk of developing PUD and bad clinical outcome (Mattar \textit{et al.}, 2007; Censini \textit{et al.}, 1996). \textit{cagA} gene in particular was predominantly found in patients with DU in both European and Polynesian in Auckland (83% and 86%, respectively) (Campbell \textit{et al.}, 1997). Meanwhile, \textit{cagE} gene was proposed to be a more reliable virulence marker than \textit{cagA} gene (Ikenoue \textit{et al.}, 2001; Audibert \textit{et al.}, 2001; Maeda \textit{et al.}, 1999) but Tan \textit{et al.} (2005) reported that Chinese (39.0%) had significantly lower prevalence rate of \textit{cagE} gene than Malay (70.0%) and Indian (81.6%) patients which appears \textit{cagE} gene is not a good marker. \textit{cagE} and \textit{cagM} genes are responsible for activating the transcription factor NF-κB, which mediates IL-8 secretion (Glocker \textit{et al.}, 1998). However, our findings corroborate data with a study conducted in Taiwan which revealed no association between the presence of \textit{cagA}, \textit{cagE} and \textit{cagM} genes with the type of disease and/or the histological findings in their patients (Sheu \textit{et al.}, 2002). Furthermore, the involvement of these genes and others in \textit{H. pylori} pathogenesis has also been disputed (Hsu P-I, \textit{et al.}, 2002; Segal \textit{et al.}, 1999). In

\begin{table}
\centering
\begin{tabular}{llllll}
\hline
Gene & PUD & NUD & Adjusted OR & p value \textsuperscript{a} & \\
& n=35 (%) & n=129 (%) & (95 % CI) & \\
\hline
\textit{cagA} (n=102) & 19 (54.3) & 83 (62.2) & 0.658 (0.309-1.402) & 0.327 & \\
\textit{cagE} (n=79) & 14 (40.0) & 65 (48.2) & 0.656 (0.307-1.403) & 0.341 & \\
\textit{cagM} (n=104) & 22 (62.9) & 82 (63.6) & 0.970 (0.447-2.103) & 1.000 & \\
\textit{cagT} (n=104) & 21 (60.0) & 83 (64.3) & 0.831 (0.386-1.789) & 0.694 & \\
\textit{cag} (n=76) & 17 (48.6) & 59 (45.7) & 1.121 (0.530-2.367) & 0.849 & \\
\textit{cag10} (n=69) & 16 (45.7) & 53 (42.1) & 1.208 (0.569-2.561) & 0.623 & \\
\textit{cag13} (n=8) & 2 (5.7) & 6 (4.7) & 1.242 (0.240-6.442) & 0.679 & \\
\textit{IS605} (n=6) & 1 (2.9) & 5 (3.9) & 0.157 (0.082-6.455) & 1.000 & \\
\hline
\end{tabular}
\caption{Relationship between \textit{Helicobacter pylori} genes (\textit{cagA}, \textit{cagE}, \textit{cagM}, \textit{cagT}, \textit{cag}6-7, \textit{cag10}, \textit{cag13} and \textit{IS605}) and clinical outcomes in 164 patients with \textit{H. pylori} infection}
\end{table}

Note: PUD = peptic ulcer disease; NUD = non-ulcer dyspepsia; OR = odds ratio; CI = confidence interval

\textsuperscript{a}Fisher’s Exact test; p value < 0.05 is considered significant
addition, the prevalence of cag6-7 gene in this study was significantly higher amongst the Chinese (62.8 %) compared to the Malays (28.6 %) (TABLE 1). The difference in the prevalence of cag6-7 gene was highly statistically significant between Chinese and Indians ($p=0.000$). Interestingly, a study in Japan revealed that 93.1% (27/29) of cancer patients had cag6-7 gene although the significance of this is still uncertain (Deguchi et al., 2004), as this finding was not analyzed statistically.

Our study had several limitations. All cagPAI genes could not be chosen for the study due to budget constraints. The number of Indian patients was relatively small despite the significant findings observed in this ethnic group. However, the present study was conducted as a pilot study involving the largest number of H. pylori genes in different ethnic groups in Malaysia. Thus, large prospective or multi-centered studies are needed to investigate the pathogenic impact of H. pylori virulence genes amongst different ethnic groups in relation to its clinical relevance.

**CONCLUSION**

The decline in the prevalence of H. pylori infection has not only been observed in several developed countries but also in Malaysia. Detection of cagA/E/M genes will partially resolve the “Indian enigma”: it might be that the presence of different H. pylori genes/genotypes that might explain why Indians have a lower risk of developing severe disease outcomes despite having the highest prevalence rate of H. pylori infection. This could have clinical implication when initiating anti-H. pylori therapy in the multi-ethnic population in Malaysia.

**ACKNOWLEDGMENTS**

This study was supported by a grant from the Ministry of Science, Technology and Innovation of Malaysia (No. 06-02-04-0907-PR0073/05-2). Our special thanks

---

**TABLE 3**

Relationship between *Helicobacter pylori* genes (cagA, cagE, cagM, cagT, cag6-7 and cag10) and clinical outcomes in 42 Indian patients with *H. pylori* infection

<table>
<thead>
<tr>
<th>Gene(s)</th>
<th>PUD n=7 (%)</th>
<th>NUD n=35 (%)</th>
<th>Adjusted OR (95 % CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cagA (n=24)</td>
<td>1 (4.2)</td>
<td>23 (95.8)</td>
<td>0.087 (0.009-0.808)</td>
<td>0.031&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>cagE (n=22)</td>
<td>1 (4.5)</td>
<td>21 (95.5)</td>
<td>0.111 (0.012-1.025)</td>
<td>0.041&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>cagM (n=29)</td>
<td>2 (6.9)</td>
<td>27 (93.1)</td>
<td>0.119 (0.019-0.731)</td>
<td>0.021&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>cagT (n=26)</td>
<td>3 (11.5)</td>
<td>23 (88.5)</td>
<td>0.391 (0.075-2.041)</td>
<td>0.397</td>
</tr>
<tr>
<td>cag6-7 (n=9)</td>
<td>1 (11.1)</td>
<td>8 (88.9)</td>
<td>0.562 (0.059-5.386)</td>
<td>1.000</td>
</tr>
<tr>
<td>cag10 (n=19)</td>
<td>1 (5.3)</td>
<td>18 (94.7)</td>
<td>0.157 (0.051-1.447)</td>
<td>0.105</td>
</tr>
<tr>
<td>cagA/E/M (n=21)</td>
<td>1 (4.8)</td>
<td>20 (95.2)</td>
<td>0.125 (0.014-1.151)</td>
<td>0.047&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Note: PUD = peptic ulcer disease; NUD = non-ulcer dyspepsia; OR = odds ratio; CI = confidence interval

Statistical analysis was not done for cag13 and IS605 genes as number of samples was too small
<sup>a,b,c,d</sup> by Fischer’s exact tests; p values <0.05 are considered significant

---

ETHNIC DIFFERENCES IN THE PREVALENCE, CLINICAL OUTCOME AND CAG PATHOGENICITY ISLAND (CAGPAI) VIRULENCE GENE PROFILES OF *H. pylori*
go to the staff of the Endoscopy and Histopathology Unit, Universiti Kebangsaan Malaysia Medical Center for their laboratory assistance. We also thank The Dean of Faculty of Medicine and Health Sciences, Universiti Putra Malaysia for her support as well.

REFERENCES


Ethnic Differences in the Prevalence, Clinical Outcome and cag Pathogenicity Island (cagPAI) Virulence Gene Profiles of H. pylori


