Significant association of the dupA gene of Helicobacter pylori with duodenal ulcer development in a South-east Indian population

Jawed Alam,1 Sankar Maiti,2 Prachetash Ghosh,1 Ronita De,1 Abhijit Chowdhury,3 Suryasnata Das,4 Ragini Macaden,4 Harshad Devarbhavi,4 T. Ramamurthy1 and Asish K. Mukhopadhyay1

1National Institute of Cholera and Enteric Diseases, Kolkata 700010, India
2IISER (Indian Institute of Science Education and Research), Kolkata, India
3Centre for Liver Research, School of Digestive and Liver Diseases, Institute of Post Graduate Medical Education & Research, Kolkata, India
4St John’s Medical College Hospital, Bangalore, India

A novel virulence factor, duodenal ulcer-promoting gene A (dupA), in Helicobacter pylori has been found to be associated with disease in certain populations but not in others. This study analysed a South-east Indian population as part of the debate about the relevance of dupA for the prediction of clinical outcomes. A total of 140 H. pylori strains isolated from duodenal ulcer (DU) (n=83) and non-ulcer dyspepsia (NUD) patients (n=57) were screened by PCR and dot-blot hybridization to determine the presence of the ORFs jhp0917 and jhp0918. Part of jhp0917–jhp0918 was sequenced to search for the C/T insertion that characterizes dupA and the levels of dupA transcripts were also assessed. The PCR and dot-blot results indicated the presence of jhp0917 and jhp0918 in 37.3 % (31/83) and 12.2 % (7/57) of H. pylori strains isolated from DU and NUD patients, respectively. Sequencing analysis showed insertion of a C at nt 1386 in the region of jhp0917, forming the dupA gene in 35 strains. RT-PCR analysis detected the dupA transcript in 28 of these 35 strains. The expression level of the dupA transcript varied from strain to strain, as shown by real-time PCR. The results demonstrated that analysis based on PCR only for dupA may produce an erroneous interpretation. The prevalence of dupA was significantly greater among strains isolated from patients with DU than from patients with NUD in this population (P=0.001, odds ratio=4.26, confidence interval=1.60–11.74). Based on these findings, dupA can be considered a biomarker for DU patients in India. The reported discrepancies for this putative virulence marker in different populations may be due to the genome plasticity of H. pylori.

INTRODUCTION

Helicobacter pylori is a Gram-negative, spiral pathogen that infects >50 % of the world’s population (Brown, 2000). Infection with H. pylori plays an important role in the development of peptic ulcer disease, distal gastric carcinoma (GC) and gastric mucosa-associated lymphoid tissue lymphoma (Mégraud & Lamouliatte, 1992; Parsonnet et al., 1991; Wotherspoon et al., 1991). In India, ~65–70 % of the population are infected with H. pylori (Graham et al., 1991; Singh et al., 2002). Overall, 15–20 % of infected patients develop gastric or duodenal ulcer (DU) and <1 % develop gastric adenocarcinoma. H. pylori infection is more prevalent in developing countries, and its incidence is decreasing in western countries (Czinn, 2005). The decisive factor(s) in H. pylori-mediated infection is still unclear. However, involvement of several virulence factors of the bacteria, host genetics and environmental influences are believed to determine the outcome of infection. Among the host factors, pro-inflammatory cytokine gene polymorphisms have been associated with DU and GC (El-Omar et al., 2000; Machado et al., 2001; Rocha et al., 2005).

Several bacterial virulence genes such as vacA, cagA, babA and oipA of H. pylori have been investigated to understand their association with gastroduodenal disease (Covacci et al., 1993; Atherton et al., 1995; Yamaoka et al., 1999,
One possible problem that has complicated identification of definite disease-specific \textit{H. pylori} virulence factors is the considerable geographical diversity in the prevalence of \textit{H. pylori} virulence factors. Cytotoxin-associated gene (\textit{cag}) was the first reported gene that varies in \textit{H. pylori} strains and is considered a marker for the presence of the \textit{cag} pathogenicity island (\textit{cag-PAI}), which includes a number of other genes associated with increased virulence (Broulet \textit{et al.}, 2001; Censini \textit{et al.}, 1996; Rahman \textit{et al.}, 2003). However, none of the above-mentioned virulence factors have exhibited any discriminating roles in the development of peptic ulcer versus GC. In addition to \textit{cag-PAI}, comparison of the whole genome of two unrelated \textit{H. pylori} strains (J99 and 26695) (Alm \textit{et al.}, 1999; Tomb \textit{et al.}, 1997) indicated the presence of a hypervariable region called the ‘plasticity zone’ with low G+C content, along with strain-specific ORFs. This plasticity region is 45 kb and continuous in strain J99, and 68 kb and discontinuous in strain 26695. Compared with 38 ORFs of the plasticity zone (\textit{jhp0914–jhp0951}) in strain J99, 33 were found to be absent in strain 26695 (Kersulyte \textit{et al.}, 2003; Occhialini \textit{et al.}, 2000; Pacheco \textit{et al.}, 2008; Yakoob \textit{et al.}, 2010; Yamaoka, 2008). Recently, a novel DU-promoting gene (\textit{dupA}) was described, which consists of two ORFs, \textit{jhp0917} and \textit{jhp0918}, and forms one continuous gene by the insertion of a T or C nucleotide after nt 1385 of \textit{jhp0917} in the 3' region (Lu \textit{et al.}, 2005). This gene (which is homologous to \textit{virB4}, encoding a component of the type IV secretion system) is located in the plasticity region and is associated with increased risk of DU and protection against gastric atrophy, intestinal metaplasia and GC in Japan and Korea (Lu \textit{et al.}, 2005).

However, the role of \textit{dupA} as a virulence marker is still controversial. Some researchers have supported the interpretations of Lu \textit{et al.} (2005) but others did not find any association. Hussein \textit{et al.} (2008) reported that the \textit{dupA} gene is associated with peptic ulcer but did not find any negative association with GC in the Iraqi population. In Chinese and north Indian populations, a significant association of \textit{dupA} with DU has been established (Arachchi \textit{et al.}, 2007; Zhang \textit{et al.}, 2008). In contrast, Argent \textit{et al.} (2007) showed no association of \textit{dupA} with DU in populations from Belgium, South Africa, China and the USA. Douraghi \textit{et al.} (2008) showed no association of \textit{dupA} with any clinical outcome in an Iranian population, and Schmidt \textit{et al.} (2009) identified no consistent association between \textit{dupA} and DU or GC across Swedish, Australian and ethnic Chinese, Indian and Malaysian populations residing in Singapore and Malaysia. A meta-analysis study by Shiota \textit{et al.} (2010) showed that the presence of \textit{dupA} is significantly associated with DU, whilst another systematic review confirmed that \textit{dupA} is associated with gastroduodenal diseases (Hussein, 2010).

There are also indications of significant geographical differences among strains. Indian \textit{H. pylori} strains are genetically distinct from East Asian and Western strains (Mukhopadhyay \textit{et al.}, 2000). Moreover, our recent study showed that the presence of strains with an intact \textit{cag-PAI} was found more frequently in Kolkata than in southern India, indicating regional variations in \textit{H. pylori} gene pools (Patra \textit{et al.}, 2011). These considerations and our interest in the dynamics of genetic traits associated with \textit{H. pylori} infection and disease association motivated us to conduct the present study to investigate the prevalence of the \textit{dupA} gene of \textit{H. pylori} in DU and non-ulcer dyspepsia (NUD) patients isolated from a south-east Indian population and also to determine the association of \textit{dupA} with clinical outcome in a different setting.

**METHODS**

**Collection of biopsy samples.** A total of 221 adult subjects of both genders (aged between 20 and 65 years) with upper gastrointestinal disorders underwent endoscopy at the hospital of the Institute of Post Graduate Medical Education and Research, Kolkata, and St John’s Medical College Hospital, Bangalore, India, during 2006–2008. A detailed history of the patient was taken and a physical examination of each subject was carried out prior to endoscopy. The objective of the study was explained to every individual and informed consent was obtained from each individual under protocols approved by the ethical committees of the respective institutes based on the Helsinki Declaration. During endoscopy, two biopsies, one from the antrum and the other from the corpus of the stomach, were obtained from each subject. Biopsy samples in 0.6 ml \textit{Brucella} broth (Difco Laboratories) with 15% glycerol were transported to the laboratory under ice-cold conditions and stored at −70 °C until culture.

**\textit{H. pylori} culture.** In the laboratory, the \textit{Brucella} broth containing the specimen was vortexed for 2 min and 200 \mu l of the mixture was streaked onto Petri plates containing brain–heart infusion agar (Difco Laboratories) supplemented with 7% sheep blood, 0.4% IsoVitalex, amphotericin B (8 \mu g ml\(^{-1}\)), trimethoprim (5 \mu g ml\(^{-1}\)), vancomycin (6 \mu g ml\(^{-1}\)) and nalidixic acid (8 \mu g ml\(^{-1}\)) (all from Sigma Chemicals). The plates were incubated for 3 days at 37 °C in a double gas incubator (Heraeus Instruments) in an atmosphere of 5% O\(_2\), 10% CO\(_2\) and 85% N\(_2\). \textit{H. pylori} colonies were identified by their typical colony morphology, appearance on Gram staining and positive reactions in urease, catalase and oxidase tests, along with a urease PCR. The bacteria were subcultured at 37 °C on the above medium and under the same microaerophilic condition.

**Extraction of genomic DNA.** Cells were harvested from the culture plates and washed with PBS (pH 8.0), followed by centrifugation at 955 \textit{g} for 1 min. The pelleted cells were resuspended in 540 \mu l TE buffer (pH 8), 60 \mu l 10% SDS (Sigma) and 9 \mu l proteinase K (20 mg ml\(^{-1}\)) (Invitrogen). The mixture was incubated at 50 °C for 1 h, followed by addition of 100 \mu l 5 M NaCl and 80 \mu l 10% CTAB solution, and then incubated again at 65 °C for 10 min. DNA was extracted using a standard phenol–chloroform method (Ausubel \textit{et al.}, 1993).

**PCR amplification.** PCR amplification was performed in a final volume of 20 \mu l containing template DNA (2–20 ng), 2 \mu l 10 \times buffer (Roche), 2.5 mM dNTPs (Roche) and 10 pmol of the appropriate primers in the presence of 1 U Taq DNA polymerase (Roche). The cycling program was: initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 1 min, with a final extension at 72 °C for 7 min. Genomic DNA from strains 999...
and 26695 was included as positive and negative controls, respectively. The PCR products were analysed by 1.5 % agarose gels (containing 0.5 μg ethidium bromide mL⁻¹) in 1 x TAE buffer. The gels were scanned under UV light and analysed with Quantity One software (Bio-Rad). The amplicon size was confirmed using DNA molecular size standards.

**Dot-blot hybridization.** Dot blots were performed using the DNA extracted from all the strains to avoid false-negative results in the PCR assay due to variations in the primer annealing sites. Approximately 50 ng purified DNA in 2 μl was spotted individually onto a Hybond-N⁺ membrane (Amersham Biosciences). The membrane was gently placed on 2 ml denaturation solution for 5 min, with care being taken not to submerge the membrane in the solution. The membrane was subsequently treated with 2 ml neutralization solution for 5 min. Following neutralization, the membrane was air dried and UV cross-linked (Bio-Rad). Parts of the ORFs (jhp0918F/jhp0918R, respectively) (Table 1). The amplified fragments were labelled with alkaline phosphatase (DupAsetIF/DupAsetIR, respectively) (Table 1) to check the integrity of the cDNA. Real-time PCR was carried out in a final volume of 12 μl with StepOnePlus Real-Time PCR System (Applied Biosystems) with 75 ng cDNA, 10 pmol each primer (DupAsetIF/DupAsetIR, RpsTF/RpsTR; Table 1) and Power SYBR Green master mix (Applied Biosystems). Cycling conditions were 10 min at 95 °C for 5 min, followed by 40 cycles of 95 °C for 30 s and 62 °C for 30 s. The threshold cycle number (Ct) of triplicate reactions was determined using StepOne software version 2.1, and the mean Ct of triplicate reactions was determined. The internal control gene rpsT was amplified simultaneously in separate reaction tubes under the same conditions. The levels of expression of the dupA gene in relation to clinical outcome. For univariate analysis, a χ² test was used. A value of P<0.05 was considered statistically significant.

**Statistical analysis.** A univariate analysis was performed to determine the risk of having the dupA gene in relation to clinical outcome. For univariate analysis, a χ² test was used. A value of P<0.05 was considered statistically significant.

**Table 1.** Primers used in this study for amplification and sequencing

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>Sequence (5'→3')</th>
<th>Amplicon (bp)</th>
<th>Annealing temp. (°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jhp0917F/Jhp0917R</td>
<td>TGGTTTCTACTGACAGAGGCG/AAACCGGCTACAGGAAATCCTCCC</td>
<td>307</td>
<td>55</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td>Jhp0918F/Jhp0918R</td>
<td>CCTATATGCTAAGGCGCCTC/AGCCTGAAAGCTTGTTAAGG</td>
<td>276</td>
<td>55</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td>UreBF/UreBR</td>
<td>CGTCCCGGAATAGCTGCCATAGT/GTATGCTGCTGCAAGGCTTAA</td>
<td>480</td>
<td>50</td>
<td>This study</td>
</tr>
<tr>
<td>Set5F/Set5R</td>
<td>CTAGCGGACAGAATTTTAATGAT/GCTAATTTTGTGACGATATT</td>
<td>350</td>
<td>56</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAsetIR</td>
<td>CGTGATCACAATTTAGTGTCTT/CCAGAACTGCTGATCTTCTTTCT</td>
<td>214</td>
<td>54</td>
<td>Gomes et al. (2008)</td>
</tr>
<tr>
<td>RpsTF/RpsTR</td>
<td>GGCAAATCATGATCCGGAGA/GGTTCAATCAGAGGTTTTTCT</td>
<td>217</td>
<td>55</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAsetIR</td>
<td>GGGGATTGTGCGCTGACGTG/GCAAGGTTGCTGGTGAGT</td>
<td>805</td>
<td>58</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/Jhp0918R</td>
<td>GGGGATTGTGCGCTGACGTG/GCAAGGTTGCTGGTGAGT</td>
<td>1022</td>
<td>55</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAR</td>
<td>CGTGATCACAATTTAGTGTCTT/CCAGAACTGCTGATCTTCTTTCT</td>
<td>563</td>
<td>50</td>
<td>This study</td>
</tr>
<tr>
<td>CagA5cF/CagA3cR</td>
<td>GTGTAAGAGGCTGTGCTGCCG/GGGTTGATGATATTTCCATAA</td>
<td>350</td>
<td>55</td>
<td>Chattopadhyay et al. (2004)</td>
</tr>
</tbody>
</table>

_and 26695 was included as positive and negative controls, respectively. The PCR products were analysed by 1.5 % agarose gels (containing 0.5 μg ethidium bromide mL⁻¹) in 1 x TAE buffer. The gels were scanned under UV light and analysed with Quantity One software (Bio-Rad). The amplicon size was confirmed using DNA molecular size standards._

**Dot-blot hybridization.** Dot blots were performed using the DNA extracted from all the strains to avoid false-negative results in the PCR assay due to variations in the primer annealing sites. Approximately 50 ng purified DNA in 2 μl was spotted individually onto a Hybond-N⁺ membrane (Amersham Biosciences). The membrane was gently placed on 2 ml denaturation solution for 5 min, with care being taken not to submerge the membrane in the solution. The membrane was subsequently treated with 2 ml neutralization solution for 5 min. Following neutralization, the membrane was air dried and UV cross-linked (Bio-Rad). Parts of the ORFs (jhp0917F/jhp0917R and jhp0918F/jhp0918R, respectively) (Table 1). The amplified fragments were purified using a QIAquick PCR Purification kit. The purified PCR product was quantified on a 1 % agarose gel where the intensity of the band was compared with a λ HindIII digest. The PCR-purified products were sequenced using a Big Dye Terminator version 3.1 Cycle Sequencing kit on an ABI PRISM 3100 Genetic Analyser (Applied Biosystems).

**Gene expression assay by RT-PCR.** Total RNA of _H. pylori_ was isolated using TRizol reagent (Invitrogen) according to the manufacturer’s protocol and treated with DNase I (Ambion) to remove DNA contamination. The absence of DNA contamination was checked by PCR with primer set UreBF/UreBR (Table 1) and quantified by measuring the absorbance at 260 nm. Total RNA (2 μg) was reverse transcribed to cDNA with a RevertAid First Strand cDNA Synthesis kit (Fermentas). The cDNA was then amplified with two different primer sets, Set5F/Set5R and Jhp0917F/Jhp0917R (Table 1). All cDNA samples were amplified with primers UreBF and UreBR (Table 1) to check the integrity of the cDNA. Real-time PCR was carried out in a final volume of 12 μl with StepOnePlus Real-Time PCR System (Applied Biosystems) with 75 ng cDNA, 10 pmol each primer (DupAsetIF/DupAsetIR, RpsTF/RpsTR; Table 1) and Power SYBR Green master mix (Applied Biosystems). Cycling conditions were 10 min at 95 °C for 5 min, followed by 40 cycles of 95 °C for 15 s, 55 °C for 30 s and 62 °C for 30 s. The threshold cycle number (Ct) of triplicate reactions was determined using StepOne software version 2.1, and the mean Ct of triplicate reactions was determined. The internal control gene rpsT was amplified simultaneously in separate reaction tubes under the same conditions. The levels of expression of the dupA gene in relation to clinical outcome. For univariate analysis, a χ² test was used. A value of P<0.05 was considered statistically significant.

**Statistical analysis.** A univariate analysis was performed to determine the risk of having the dupA gene in relation to clinical outcome. For univariate analysis, a χ² test was used. A value of P<0.05 was considered statistically significant.

**Table 1.** Primers used in this study for amplification and sequencing

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>Sequence (5'→3')</th>
<th>Amplicon (bp)</th>
<th>Annealing temp. (°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jhp0917F/Jhp0917R</td>
<td>TGGTTTCTACTGACAGAGGCG/AAACCGGCTACAGGAAATCCTCCC</td>
<td>307</td>
<td>55</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td>Jhp0918F/Jhp0918R</td>
<td>CCTATATGCTAAGGCGCCTC/AGCCTGAAAGCTTGTTAAGG</td>
<td>276</td>
<td>55</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td>UreBF/UreBR</td>
<td>CGTCCCGGAATAGCTGCCATAGT/GTATGCTGCTGCAAGGCTTAA</td>
<td>480</td>
<td>50</td>
<td>This study</td>
</tr>
<tr>
<td>Set5F/Set5R</td>
<td>CTAGCGGACAGAATTTTAATGAT/GCTAATTTTGTGACGATATT</td>
<td>350</td>
<td>56</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAsetIR</td>
<td>CGTGATCACAATTTAGTGTCTT/CCAGAACTGCTGATCTTCTTTCT</td>
<td>214</td>
<td>54</td>
<td>Gomes et al. (2008)</td>
</tr>
<tr>
<td>RpsTF/RpsTR</td>
<td>GGCAAATCATGATCCGGAGA/GGTTCAATCAGAGGTTTTTCT</td>
<td>217</td>
<td>55</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAsetIR</td>
<td>GGGGATTGTGCGCTGACGTG/GCAAGGTTGCTGGTGAGT</td>
<td>805</td>
<td>58</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/Jhp0918R</td>
<td>GGGGATTGTGCGCTGACGTG/GCAAGGTTGCTGGTGAGT</td>
<td>1022</td>
<td>55</td>
<td>This study</td>
</tr>
<tr>
<td>DupAsetIF/DupAR</td>
<td>CGTGATCACAATTTAGTGTCTT/CCAGAACTGCTGATCTTCTTTCT</td>
<td>563</td>
<td>50</td>
<td>This study</td>
</tr>
<tr>
<td>CagA5cF/CagA3cR</td>
<td>GTGTAAGAGGCTGTGCTGCCG/GGGTTGATGATATTTCCATAA</td>
<td>350</td>
<td>55</td>
<td>Chattopadhyay et al. (2004)</td>
</tr>
</tbody>
</table>
RESULTS
A total of 140 H. pylori strains were isolated from the enrolled 221 subjects, who underwent endoscopy. Subjects with abdominal discomfort, acidity and loss of appetite but no frank ulceration were considered to be NUD patients, whilst those with duodenal ulceration that was visible endoscopically were considered to be DU patients. The strains were isolated from 83 DU patients and 57 NUD patients. In the 83 DU cases comprising 51 males and 32 females, the mean age was 46 ± 10.72 and 43.7 ± 9.36 years, respectively, whilst among the 57 NUD patients comprising 36 males and 21 females, the mean age was 32.4 ± 7.22 and 33.13 ± 6.84 years, respectively. The genomic DNA from these 140 strains was used for further PCR-based analysis.

Distribution of jhp0917, jhp0918, cagA and vacA
We first studied the presence of the ORFs jhp0917 and jhp0918 in the 140 strains from south-east India (83 from DU patients and 57 from NUD patients) using PCR and dot-blot hybridization. The ORF jhp0917 was targeted with the specific primers jhp0917F and jhp0917R and yielded a 307 bp amplicon (Table 1 and Fig. 1). Similarly, jhp0918 was amplified with the gene-specific primers Jhp0918F and Jhp0918R producing an amplicon of 276 bp (Table 1 and Fig. 1). All strains that were positive for both jhp0917 and jhp0918 by PCR were also positive in the dot-blot hybridization. In addition, five PCR-negative strains were found to be positive by dot-blot hybridization, with strong binding of the probe (Fig. 2). The hybridization results indicated that interpretation of jhp0917 and jhp0918 positivity should not be considered based only on the PCR result. The PCR and dot-blot hybridization data indicated the absence of both jhp0917 and jhp0918 in 68.6 % (96/140) of strains isolated from south-east India. Among the positive strains, three strains were positive only for jhp0917 and another three strains were positive only for jhp0918. Thus, the PCR and dot-blot hybridization results showed that 38 strains had both the jhp0917 and jhp0918 ORFs.

The cagA and vacA status was determined using primers and protocols described previously (Chattopadhyay et al., 2004; Mukhopadhyay et al., 2000). The cagA gene was present in 92.1 % (129/140) of the tested strains from this region, whilst 70.0 % (98/140) of the strains had the vacA s1m1 allele (Table 2). The other two alleles of vacA, s1m2 and s2m2, were present in 17.1 % (24/140) and 12.9 % (18/140) of the strains, respectively (data not shown). The status of the cagA and vacA genes in the 38 jhp0917+ jhp0918-positive strains isolated from the south-east Indian population was 34/38 (89.5 %) and 22/38 (57.9 %), respectively. Four strains isolated from the DU patients were positive for jhp0917+ jhp0918 but negative for cagA and had the s2m2 allele of vacA.

Fig. 1. Schematic representation of the dupA gene of clinical isolates with reference to strain J99 with ORFs jhp0917 and jhp0918. The locations of the various sets of primers are indicated, with the resulting PCR products indicated below.

Fig. 2. PCR (a) and dot-blot hybridization (b) results. Lane 1, strain J99 (positive control) with an amplicon size of 307 bp; lanes 2–8, clinical isolates. In the dot-blot hybridization, lanes 2, 5 and 6 show examples of a positive result despite a negative result by PCR.
Table 2. Prevalence of the dupA, cagA and vacA genes among the studied strains in India

<table>
<thead>
<tr>
<th>Gene</th>
<th>Total (n=140)</th>
<th>DU (n=83)</th>
<th>NUD (n=57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dupA</td>
<td>28 (20.0 %)</td>
<td>23 (27.7 %)</td>
<td>5 (8.8 %)</td>
</tr>
<tr>
<td>cagA</td>
<td>129 (92.1 %)</td>
<td>76 (91.6 %)</td>
<td>53 (93.0 %)</td>
</tr>
<tr>
<td>vacA s1m1</td>
<td>98 (70.0 %)</td>
<td>59 (71.1 %)</td>
<td>39 (68.4 %)</td>
</tr>
</tbody>
</table>

Sequence analysis

Four different primer sets were used for sequencing a small fragment of jhp0917 to search for the 1 nt insertion that characterizes dupA, as a single PCR set was unable to yield PCR fragments from all 38 strains. Using these four sets of primers, the junction region of the 38 jhp0917 + jhp0918-positive strains was sequenced and compared with the published sequence of strain J99 (GenBank accession no. AE001439). We observed insertion of a C after nt 1385 in the 3’ region of jhp0917 in 35 of the 38 clinical isolates. The remaining three isolates had the same sequence as strain J99 where no insertion of either T or C was found, but these strains were still PCR and dot-blot positive with the primers described earlier, indicating these were not the true dupA (Fig. 3).

Expression of dupA

We performed RT-PCR of the 35 dupA-positive strains and found that the dupA transcript was present in 28 strains and absent in the remaining seven, indicating that the dupA gene was not expressed in all strains (Fig. 4a). Real-time PCR was carried out with randomly selected dupA-positive strains and it was found that the level of dupA transcripts varied in different strains (Fig. 4b).

dupA status in the south-east Indian population

In our study, 38 clinical strains (31 DU and seven NUD) were positive for both jhp0917 and jhp0918 by PCR and dot-blot hybridization. We found by sequencing that 35 strains had insertion of a C nucleotide, whilst three strains had no insertion of C or T, similar to strain J99. Analysis by RT-PCR showed that 28 strains of these 35 were positive for the dupA transcript. On the basis of sequencing and RT-PCR findings, we confirmed that 28 strains (23 DU and five NUD) were dupA positive and seven (five DU and two NUD) strains were dupA negative. Our study showed that the prevalence of dupA in DU patients (23/83, 27.7 %) was significantly higher than in NUD patients (5/57, 8.7 %). (P=0.001, odds ratio=6.49, 95 % confidence interval=1.71–28.94).

DISCUSSION

Recent studies have suggested the possibility of using genetic markers in the plasticity zone as indicators of pathogenicity for H. pylori infection, despite a lack of credible knowledge regarding the functions of the putatively encoded proteins in this cluster. It seems that these determinants may play a key role in determining the virulence capacity of H. pylori strains, either directly or by encoding factors that may lead to varying clinical outcomes. An association between some of the ORFs in the plasticity zone and various disease categories has been reported previously. For instance, Occhialini et al. (2000) found that two single ORFs (jhp0940 and jhp0947) were more prevalent in strains isolated from patients with gastric adenocarcinoma in Costa Rica, and Santos et al. (2003) showed an association between jhp0947 and DU as well as GC in Brazilian patients. This was confirmed for the jhp0947 and jhp0949 ORFs in DU patients from the Netherlands (De Jonge et al., 2004).

Our study in a south-east Indian population demonstrated that the dupA gene was 6.5 times more prevalent in DU patients than in NUD patients, indicating that the dupA gene was significantly associated with DU (P=0.001). Associations between the presence of dupA and H. pylori diseases varies around the world (Argent et al., 2007; Gomes et al., 2008; Hussein, 2010; Nguyen et al., 2010). However, a number of factors, from geographical variations to study procedures, have to be considered. In some studies, only one set of primer pairs for jhp0917 and jhp0918 was used (Douraghi et al., 2008; Lu et al., 2005; Pacheco et al., 2008; Zhang et al., 2008). Our study showed that 13.2 % (5/38) of dot-blot-positive strains for jhp0917 and jhp0918 failed to provide an amplicon using an initial

![Fig. 3. Nucleotide sequence alignment of partial sequences of dupA from nt 1355 to nt 1406 for various H. pylori strains. The start codon of the dupA gene was taken as nt 1. Asterisks indicate matching nucleotides between all the above strains and hyphens represent deletions. Sequence alignment was done using the CLUSTAL version 2.0.12 multiple sequence alignment tool. Nucleotides that differed between strains are highlighted by shading.]
set of primers, but different sets of primers yielded amplicons from the same strains. Hence, the use of multiple primer pairs is recommended for detection of the dupA gene in future studies. In addition, sequence analysis showed that 7.9% (3/38) of the jhp0917+ jhp0918-positive strains did not have an insertion of C or T after nt 1385 in the 3′ region of jhp0917, indicating that they did not form the dupA gene. This report is inconsistent with previous reports in other populations, which indicated that all clinical isolates possessed a continuous dupA gene (Douraghi et al., 2008; Gomes et al., 2008; Schmidt et al., 2009). Moreover, in our study, no strain was detected with insertion of a T after nt 1385 of jhp0917. RT-PCR analysis showed that in 20.0% (7/35) of the dupA-positive strains, a dupA transcript could not be detected. This contradicts the findings of Nguyen et al. (2010) in which dupA was always expressed. Real-time PCR analysis showed that the expression level of the dupA transcripts varied from strain to strain. A recent systematic review demonstrated the importance of the presence of the dupA gene for DU, especially in Asian countries (Shiota et al., 2010). Arachchi et al. (2007) showed that dupA was present in 37.5 and 22.8% of DU and functional dyspepsia patients, respectively, from north India, but in our study, the dupA gene was present in 27.7 and 8.7% of DU and NUD patients, respectively. The reason for this variation in dupA prevalence in India may be due to the fact that their study did not include sequencing of the intergenic region of jhp0917–jhp0918 to check for the insertion of a nucleotide after position 1385 or the RNA expression profile of the dupA gene, or might be related to the geographical genomic variation of H. pylori, as India is a large country with much diversity. Some studies have reported a dupA gene with a single-nucleotide polymorphism that created a premature stop codon and may have considerable effects on protein expression or function (Gomes et al., 2008; Hussein, 2010; Moura et al., 2012; Queiroz et al., 2011). Moreover, Douraghi et al. (2008) reported that dupA was inversely associated with the histological feature of dysplasia, an important pre-malignant and pre-cancerous lesion associated with increased incidence of cancer in the Iranian population. As a result, the dupA gene may be applicable as a protective marker against GC development. However, we were unable to study this hypothesis as we did not have samples from GC patients. A recent study showed that the presence of a complete dupA cluster (type IV secretory system with vir genes around the dupA gene) appeared to be important in DU development (Jung et al., 2012).

In conclusion, infection with dupA-positive H. pylori increases the overall risk for DU and this evidence was significant in our Indian study. The dupA gene can be considered an important biomarker for DU in the Indian population. However, further studies are required to determine the function of dupA and its relationship with disease. The discrepancy of dupA association with disease outcomes could be related to the limitation of PCR techniques for detecting the intact dupA gene or may be a consequence of the plasticity of H. pylori, which contributes to its genetic diversity, and requires additional studies for a firm conclusion.

ACKNOWLEDGEMENTS

J.A. thanks the Indian Council of Medical Research for a Senior Research Fellowship [no. 3/1/RF/36/MDP/2007 (22588)]. This work was supported in part by the ICMR, Government of India, Program...
of Founding Research Center for Emerging and Reemerging Infectious Diseases, Ministry of Education, Culture, Sports, Science and Technology of Japan and Department of Biotechnology (no. BT/PR10407/BRB/10/604/2008).

REFERENCES


Of Founding Research Center for Emerging and Reemerging Infectious Diseases, Ministry of Education, Culture, Sports, Science and Technology of Japan and Department of Biotechnology (no. BT/PR10407/BRB/10/604/2008).


