Heterogeneity in the *Helicobacter pylori* vacA and cagA genes: association with gastroduodenal disease in South Africa?

M Kidd, A J Lastovica, J C Atherton, J A Louw

Abstract

**Background**—*Helicobacter pylori* infection is universally associated with gastritis, but only sometimes with clinically significant disease. Candidate virulence markers seem to be useful in identifying the pathogenic infections in some populations.

**Aims**—To investigate the association between putative virulence markers and disease in an African population.

**Methods**—Fifty-nine *H pylori* strains isolated from dyspeptic patients (11 with peptic ulceration, eight with gastric adenocarcinoma, and 28 with no pathology other than gastritis) were studied for differences in the genes vacA and cagA.

**Results**—Forty-seven (80%) of 59 strains had the vacA signal sequence genotype s1 (cagA s1a, s1b, or s1c) and 12 (20%) had subtype s2. vacA mid-region analysis revealed that 40 (68%) strains were vacA m1 and 19 (32%) were m2. All 14 strains from patients with peptic ulceration were vacA s1, in contrast to 23 (66%) of 35 strains from patients with gastritis alone (p<0.01). vacA s2 was found exclusively in patients with gastritis alone (p<0.01). All strains isolated from patients with gastric adenocarcinoma were s1bm1 (p<0.05 versus gastritis alone). cagA was detectable in 56 (95%) of 59 isolates. Strains from patients with peptic ulceration (12/13 versus 19/30 with gastritis alone, p=0.05) had the shortest fragment length in the 3’ region of cagA, while 4/10 strains from patients with gastric cancer had the longest fragment length in this region (p<0.02 versus gastritis alone).

**Conclusion**—In this study, the vacA s1 genotype, and fragment length of the 3’ region of cagA identified isolates associated with significant clinical disease. The vacAs1bml genotype seems to be strongly associated with gastric cancer.

(Gut 1999;45:499–502)

Keywords: adenocarcinoma; cagA; *Helicobacter pylori*; peptic ulceration; South Africa; vacuolating cytotoxin

*Helicobacter pylori*, a spiral shaped gastric organism, is the cause of chronic gastritis, and is important in the pathogenesis of peptic ulceration, mucosa associated lymphoid tissue lymphoma, and gastric adenocarcinoma.1–3 Although infection is universally associated with gastritis, the development of clinically significant disease seems to depend on a number of factors, including the virulence of the infecting strain, the susceptibility of the host, and environmental cofactors.4 Focus on the organism has resulted in the identification of a number of non-conserved candidate virulence markers. These include in vitro production of vacuolating cytotoxin activity, the presence of certain types of the vacuolating cytotoxin gene *(vacA)*, possession of the gene *(cagA)* which is a marker for the *cag* pathogenicity island, and the ability to activate neutrophils directly.4,5 *vacA* alleles are mosaics of any combination of signal sequence type (s1a, s1b, or s2) and mid-region (m1 or m2).6 The link between the ability of a strain to induce epithelial vacuolation in vitro and peptic ulceration in vivo is significant but not invariable.4 Infection with strains with the *vacA* s1a genotype in the USA, has been linked to gastric inflammation, duodenal ulceration, in vitro cytotoxin activity, and the presence of *cagA*, while *vacA* s2 strains seem to have the least pathogenic potential.7 In other studies, *vacA* s1 strains are also associated with peptic ulcer disease.8,9

The gene *cagA* is present in 60–70% of *H pylori* strains and encodes a high molecular weight antigenic protein (120–140 kDa).10 11 CagA producing strains have been detected in patients with peptic ulceration more frequently than in patients with chronic gastritis alone.12 13 In addition, recent evidence has shown an association between infection with these strains and the development both of atrophic gastritis and of adenocarcinoma of the stomach.14 15 Recent genotypic analysis has revealed the presence of substantial variability in the 3’ region of *cagA*, which has been postulated to alter the immunogenicity of the protein. These alterations in the gene and its protein seem to correlate with clinical outcome in vivo.16

Based on historic observations, it has been postulated that clinically significant gastroduodenal disease is not common in the African setting, although contrasting views, based on objective parameters (prospective endoscopic studies) exist.17 18 This has led some to state that there is no established correlation between the organism and any pathology in Africa apart from gastritis,17 and, even more contentiously, it has been suggested that *H pylori* infection may be protective in the African setting, particularly against gastric cancer.19 We have attempted to assess whether associations...
between markers of *H pylori* virulence and disease in Western populations applied in the African setting.

In this report, we document the genotypic profile of *H pylori* isolates from the Western Cape in South Africa, as well as the relation between putative virulence markers and gastrointestinal disease in these individuals.

**Material and methods**

**PATIENTS, BIOPSY SAMPLING, AND CULTIVATION OF *H pylori* STRAINS**

*H pylori* isolates from 47 infected patients (15 black, seven white, 25 coloured) with *H pylori* infection undergoing upper gastrointestinal endoscopy were used in this study. Endoscopic and histological diagnoses were recorded for all patients. In 28 of the patients, no endoscopic pathology was evident, but histology confirmed the presence of chronic active gastritis; 11 patients had duodenal ulcers and eight had gastric adenocarcinoma. Gastric biopsy specimens obtained from the antrum of all 47 patients and additionally from the fundus of the eight patients with gastric adenocarcinoma. Gastric biopsy specimens obtained from the antrum of 47 patients and additionally from the fundus of five patients were cultured as described previously. Briefly, specimens were transported in a jar under microaerobic conditions (Oxoid Gas Generating BR38, Basingstoke, Hampshire, UK). The specimen was cultured on tryptose blood agar (CM 233; Oxoid Ltd) containing lysed horse blood (10% vol/vol) at 37°C in a microaerobic atmosphere (12% CO₂, 88% air (6% CO₂), 95% humidity) for a minimum of seven days. *H pylori* was identified by colony morphology and positive urease reaction (Christensen’s urea slope). The strains were numbered, and all analyses were performed without prior knowledge of the clinical diagnosis. Reference strains NCTC11639 and CCUG17135 were used in this study.

**PREPARATION OF SAMPLES FOR PCR AMPLIFICATION**

Genomic DNA was isolated from each colony using the High PCR (polymerase chain reaction) purification kit from Boehringer Mannheim (Johannesburg, South Africa). Briefly, bacterial cells were resuspended in phosphate buffered saline (PBS), and lysozyme (10 mg/ml) was added. Thereafter, cells were further digested by proteinase K and the DNA was isolated/purified by chaotropic salt/glass fibre methodology. Approximately 10 µg DNA was obtained from 10⁶ bacterial cells.

**PCR AMPLIFICATION AND DETECTION OF AMPLIFIED DNA PRODUCTS**

For *vacA* strains and the two reference strains, all 59 strains yielded the 286 bp amplicon (s2). None of the strains yielded a PCR product of any other size. Further analysis of the 47 strains with the s1 genotype revealed that 46 (98%) were *vacA* s1b and 1 (2%) was *vacA* s1a. All 59 *H pylori* strains and the two reference strains contained DNA which was amplifiable by VAGF and VAGR, resulting in fragments of 570 bp (m1) in 40 (68%) (and in CCUG17135) and 645 bp (m2) in 19 (32%) (and in NCTC11639) respectively. None had DNA amplified by both sets of primers or gave products of sizes other than predicted. Among the 59 isolates studied, five of the six possible combinations of *vacA* signal and mid-regions were identified. The s1a/m1 combination was

**STATISTICAL METHODS**

Data from non-identical strains (n=59) were analysed using the χ² test or Fisher’s exact test as appropriate. Probability levels less than 0.05 were considered statistically significant.

**RESULTS**

**STRAIN COLLECTION**

Single colonies (n=34) were obtained from 34 patients (21 with gastritis alone, eight with peptic ulceration, and five with gastric adenocarcinoma). Two or more morphologically distinct (either light/dark or large/small paired) colonies (n=26) were isolated from three patients with peptic ulceration, three with gastric adenocarcinoma, and seven with gastritis alone. Random amplified polymorphic DNA (RAPD) analysis, *vacA* and *cagA* typing showed that the morphologically distinct colonies were identical in only one patient (gastric adenocarcinoma, data not shown). Twenty five of the 26 strains were therefore available for analysis from this group. In total, 59 of the 60 strains isolated were available for analysis: 35 strains from 28 patients with gastritis only, 14 from 11 patients with peptic ulceration, and 11 from the eight patients with gastric adenocarcinoma.

**DETERMINATION OF *vacA* GENOTYPES**

The gene *vacA* was detectable in all *H pylori* isolates. By using the primers VA1F and VA1R to amplify the *vacA* signal sequences, the predicted PCR product of 259 bp for genotype s1 was obtained from 47 (80%) of 59 strains as well as the two reference strains. Twelve (20%) of the strains yielded the 286 bp amplicon (s2). None of the strains yielded a PCR product of any other size. Further analysis of the 47 strains with the s1 genotype revealed that 46 (98%) were *vacA* s1b and 1 (2%) was *vacA* s1a.

All 59 *H pylori* strains and the two reference strains contained DNA which was amplifiable by VAGF and VAGR, resulting in fragments of 570 bp (m1) in 40 (68%) (and in CCUG17135) and 645 bp (m2) in 19 (32%) (and in NCTC11639) respectively. None had DNA amplified by both sets of primers or gave products of sizes other than predicted. Among the 59 isolates studied, five of the six possible combinations of *vacA* signal and mid-regions were identified. The s1a/m1 combination was
**Table 1** Distribution of vacA signal sequence and mid-region types in strains isolated from patients with different gastric diseases

<table>
<thead>
<tr>
<th>vacA signal sequence type</th>
<th>vacA mid-region type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s1a</td>
</tr>
<tr>
<td>No of strains</td>
<td></td>
</tr>
<tr>
<td>Gastritis alone (n=35)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Peptic ulceration (n=13)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Adenocarcinoma (n=10)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>p Value</td>
<td>NS</td>
</tr>
</tbody>
</table>

*Compared with gastritis alone.
†Compared with peptic ulceration.
‡Compared with gastritis alone.

<table>
<thead>
<tr>
<th>Fragment length of cagA 3' amplicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of strains</td>
</tr>
<tr>
<td>Gastritis alone (n=30)</td>
</tr>
<tr>
<td>Peptic ulceration (n=13)</td>
</tr>
<tr>
<td>Adenocarcinoma (n=10)</td>
</tr>
<tr>
<td>p Value</td>
</tr>
</tbody>
</table>

*Compared with gastritis alone.
†Compared with peptic ulceration.
‡Compared with adenocarcinoma.

**Discussion**

The *Helicobacter pylori* vacuolating cytotoxin has been implicated in the pathogenesis of peptic ulceration. A US study showed that strains with vacA signal type s1a are often associated with peptic ulceration, while vacA s2 strains are usually found in patients with no ulcers. vacA s1b strains seemed to be immediately ulcerogenic and vacA alleles with this signal sequence were found to be less cytotoxic in vitro than vacA s1a strains. In a preliminary study, in a small number of South African individuals, we noted the lack of the cytotoxic vacA s1a allele in this study group.

In this expanded study, our results show a high prevalence of the vacA s1b allele and confirm the low prevalence of the vacA s1a allele in this population. Analysis of disease groups showed that the vacA s1 allele was found consistently (100%) in strains from patients with peptic ulceration. In contrast to strains from the USA and Asia, no vacA s1a alleles were present in any of these cases. South African strains with the vacA s1b allele produce low levels or no in vitro vacuolating activity on cultured epithelial (HeLa) cells. It is tempting to speculate that widespread prevalence of weakly or non-cytotoxic strains may be reflected in the apparent low frequency of *H pylori* associated diseases noted in Africa. As has been noted, however, the assumption on which the “low prevalence” of disease in Africa is based can be questioned.

In addition, a high prevalence of vacA s1b is not limited to South Africa; vacA s1b alleles are also highly prevalent in Brazil and Portugal, where reported prevalences of peptic ulceration and gastric adenocarcinoma are higher. Further studies are needed on the relative contributions of strain virulence, host susceptibility, and environmental cofactors to disease in all these populations. In this study, the vacA s2 allele was found exclusively in the group of individuals with no endoscopic pathology, in keeping with observations suggesting that vacA s2 strains are minimally or non-ulcerogenic. As in other studies, the vacA middle region types were not independently associated with the occurrence of peptic ulceration.

Investigation of strains isolated from patients with gastric cancer have focused on the

---

Table 2 Distribution of cagA 3' variability in 56 cagA strains

Fifty six (95%) of 59 strains and the two reference strains were cagA+, utilising primers F1/B1 to identify a 349 bp conserved amplicon from the S region of the gene. The gene cagA was present in 13 (93%) of 14 strains isolated from patients with peptic ulceration and 33 (94%) of 35 strains from patients with gastritis alone. All 10 isolates from patients with gastric adenocarcinoma were cagA+.

Fifty three of the 56 cagA+ strains had identifiable PCR products with primer set cag2/cag4 differing in fragment length from 400 to 757 bp. Amplicons smaller than 600 bp were found more frequently in *H pylori* strains isolated from patients with peptic ulceration (12/13, 92%) than in those with gastritis alone (19/30, 63%; p=0.05; table 2). Strains from patients with gastritis alone were also found more frequently to express amplicons of 600–700 bp (10/30, 30%; p=0.07). PCR fragments of >700 bp were more often found in patients with gastric adenocarcinoma (4/10, 40%) than in either patients with peptic ulceration (0/13; p=0.02) or patients with gastritis alone (1/30, 3%; p<0.01).

**PREVALENCE AND SIZE VARIATION OF THE GENE cagA IN H PYLORI ISOLATES**

Fifty six (95%) of 59 strains and the two reference strains were cagA+, utilising primers F1/B1 to identify a 349 bp conserved amplicon from the S region of the gene. The gene cagA was present in 13 (93%) of 14 strains isolated from patients with peptic ulceration and 33 (94%) of 35 strains from patients with gastritis alone. All 10 isolates from patients with gastric adenocarcinoma were cagA+.
presence or absence of the gene cagA. In this study we found the great majority of strains to be cagA+. Data from the rest of Africa are scarce, but suggest that cagA is commonly expressed. It seems unlikely that a low population prevalence of cagA+ strains can be invoked to explain the apparent low prevalence of gastric adenocarcinoma in Africa. As expected cagA+ strains were predominant in our gastric cancer patients (100%). The association between vacA genotypes and gastric adenocarcinoma has not previously been studied in the African setting, and here our findings are striking. All strains in this study were vacA genotype s1b/m1. A US study showed that strains with vacA mid-region type m1 were more strongly associated with epithelial injury (epithelial degeneration, mucus depletion, and microscopic erosions) than mid-region type m2. We hypothesise that infection with vacA m1 strains, and the associated damage to the antral and corpus mucosa may play a role in the aetiology of gastric adenocarcinoma in our study population. The association between vacA m1 and gastric adenocarcinoma as well as the notable absence of the vacA s2 genotype needs to be followed up in other studies in Africa and elsewhere.

The nucleotide sequence of the gene cagA contains internal duplications of a 102 base pair fragment, encoding a proline rich region in the 3’ region. Differences in this region (possibly by a mechanism to generate antigenic diversity) have been suggested to generate proteins with different sizes and immunogeneticities. We evaluated whether there was a biological significance to the variability in fragment length of the 3’ region of cagA in a population with a high prevalence (>95%) of isolates containing this gene. As in other studies, we found the length of this region to be shorter in strains isolated from patients with peptic ulceration than in patients with gastritis alone. In contrast, the fragment length was longer in strains isolated from patients with gastric cancer than in patients with peptic ulceration or gastritis alone. These results are similar to the preliminary results from a Japanese study which showed that a structural subtype of cagA, encoding a protein with a higher molecular weight, and detectable by its larger size, was associated with gastric carcinoma and atrophy. Thus, our data support the view that cagA variants provide an additional marker to distinguish disease associated strains of H pylori.

In summary, we have shown the high prevalence of vacA s1 alleles in strains isolated from South African individuals with peptic ulceration and gastric cancer, and noted the consistent association of vacA s1b/m1 strains with patients with gastric cancer. While attention has been focused on the positive correlation with disease, the segregation of vacA s2 alleles with “non-pathogenic” strains may be an excellent negative marker. The presence of the gene cagA in the majority of isolates, irrespective of associated pathology, suggests a limited use of this as a virulence factor. However, variability noted in the fragment length of the 3’ region of the gene correlates with disease in this population and may prove to be a useful virulence marker, if confirmed.