Genetic diversity in the *Helicobacter pylori* cag pathogenicity island and effect on expression of anti-CagA serum antibody in UK patients with dyspepsia

T M Peters, R J Owen, E Slater, R Varea, E L Teare, S Saverymuttu

Abstract

Aims—To investigate variation within the cag pathogenicity island (PAI) of *Helicobacter pylori* isolated from patients with dyspepsia in mid-Essex, and to evaluate the effect on expression of anti-CagA antibody.

Methods—Sixty two isolates of *H pylori* cultured from gastric biopsies were screened by specific PCR assays for the presence of cagA and other gene markers (cagD and cagE, and virD4) in the cag PAI. An enzyme linked immunosorbent assay (ELISA) kit (Viva Diagnostica helicobacter p120) was used to test for anti-CagA IgG antibody in matching sera. Isolates were also genotyped by vaculating cytotoxin polymerase chain reaction (PCR) analysis, and tested for absence of the complete cag PAI (empty site PCR assay).

Results—Forty one of the *H pylori* isolates had a cag PAI containing cagA. One strain had no cagA but other cag PAI loci were present, whereas the remaining 20 strains had no detectable cag PAI markers. Anti-CagA IgG antibody was detected in 34 sera by the ELISA assay, and when compared with the cag PAI genotype of the infecting strain, accuracy, sensitivity, and specificity were 92%, 87%, and 100%, respectively. The seven discrepant or borderline strains in the ELISA were all vacA s1 but differed in other genotypic markers.

Conclusions—The cag PAI was widely distributed in *H pylori* from patients with dyspepsia in mid-Essex who had different gastric pathologies. Infection with a strain having an uninterrupted cag PAI was associated with the presence of anti-CagA antibody in most patients. Discrepant ELISA results, mostly for elderly patients with duodenal ulcers, were attributed to cagA associated variation, particularly to the presence of mixed cagA+/cagA− cell variants in the infecting strain population. Tests for anti-CagA serum antibody were unreliable for predicting severity of clinical disease associated with *H pylori* infection in this series of patients.

Keywords: *Helicobacter pylori*, cagA; cag pathogenicity island; CagA antibody assay

Seroepidemiological surveys show that *Helicobacter pylori* has a worldwide distribution, and that up to 40% of the UK population is infected. About 80% of *H pylori* infections do not result in clinically relevant disease, even though such infections are often persistent and can be lifelong in many individuals. It has been proposed that certain strains of *H pylori* may be more virulent, and a variety of strain characteristics associated with disease severity have been identified. One key marker is CagA, a high (120–140 kDa) molecular weight immunodominant protein encoded by the cagA gene, which is part of the cag pathogenicity island (PAI). The cagA gene is an important marker for the cag PAI and there is evidence that cagA is present in at least 70% of *H pylori* organisms infecting the European population; its presence is closely correlated with the production of CagA protein, a serological marker of potentially virulent strains. The presence of *H pylori* anti-CagA antibodies is widely associated with increased risk for the development of peptic ulcer disease, atrophic gastritis, and gastric cancer. Nevertheless, the reliability, accuracy, and clinical usefulness of CagA serological and cagA genetic tests for predicting the risk of chronic gastritis, gastric cancer, and peptic ulcer disease remains controversial.

The aim of our study was to investigate the effect that diversity within the *H pylori* cag PAI might have on CagA production, and the consequent reliability of the enzyme linked immunosorbent assay (ELISA) for the detection of host serum anti-CagA IgG antibody, and hence for use as a predictive marker for more severe *H pylori* related disease. We examined strains of *H pylori* isolated from gastric biopsies of 62 patients with dyspepsia to determine the frequency of cagA and other key genes in the cag PAI, and tested for anti-CagA antibody expression in matching sera.

Materials and methods

BACTERIAL STRAINS AND GROWTH CONDITIONS

The 62 isolates of *H pylori*, for which matching patient sera were available, were obtained between 1996 and 1998 from gastric biopsies of patients (39 men, age range 29.8–75 years; and 23 women, age range 27.5–77.8 years; mean 52.3 years) in mid-Essex undergoing routine endoscopy for a range of gastroduodenal symptoms. Patients were excluded from the
Table 1 Primers used to amplify and identify markers in the cag pathogenicity island of H. pylori

<table>
<thead>
<tr>
<th>Region amplified</th>
<th>Primer designation</th>
<th>Primer sequence (5’→3’)</th>
<th>Amplicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>cagA</td>
<td>F1</td>
<td>GATAAACGGCAGGCTTTTGGAG</td>
<td>349 bp</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>CTGCAAAAAGATTGTTTGGGCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D008</td>
<td>TAATGCTAAATTAGAATTCTAGGCA</td>
<td>298 bp</td>
</tr>
<tr>
<td></td>
<td>R008</td>
<td>TAGATAATCTAAACACATACGCGGAT</td>
<td></td>
</tr>
<tr>
<td>cagD/E</td>
<td>picf</td>
<td>GAGGTTTGGCTTTGCTTGGC</td>
<td>2091 bp</td>
</tr>
<tr>
<td></td>
<td>psr</td>
<td>CAACAAAGCCTGCTTTCTAT</td>
<td>2091 bp</td>
</tr>
<tr>
<td>vlrD4</td>
<td>524r</td>
<td>AACCTTTGGCTTTTGGTTGC</td>
<td>1993 bp</td>
</tr>
<tr>
<td></td>
<td>524f</td>
<td>TTTACCTCTGTATTTGGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>AACTTTCGGCTTTTGTTGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ESr</td>
<td>TCAATGCGGAGGGGCGAGTG</td>
<td>360 bp</td>
</tr>
</tbody>
</table>

ES, empty site; f, forward primer; r, reverse primer.

study if they had previously received acid suppression treatment, *H. pylori* eradication treatment, or were current users of non-steroidal anti-inflammatory drugs. Patient groups based on endoscopic diagnosis before treatment were broadly as follows: the peptic ulcer group (13 patients); and the non-ulcer dyspeptic group with diagnoses of gastritis and/or duodenitis (16 patients), oesophagitis (19 patients), gastric neoplasia (two patients), and endoscopically normal mucosa including normal oesophagus, stomach, and duodenum (nine patients).

The biopsy specimens were homogenised in Griffith's tubes, inoculated on to 10% (vol/vol) Columbia blood and selective (DENT) agars (Oxoid, Basingstoke, UK) and incubated at 37°C under microaerophilic conditions (5% O₂, 5% CO₂, 2% H₂, 88% N₂) for five days. After primary isolation, all isolates of *H. pylori* were cultured for two to three days on 10% (vol/vol) Columbia blood agar (Oxoid) at 37°C as above. Stock cultures were preserved on glass beads in nutrient broth (Oxoid) containing 10% (vol/vol) glycerol over liquid nitrogen or at −80°C.

Reference cultures of *H. pylori* (in lyophilised form) as controls for the various polymerase chain reaction (PCR) assays were obtained from the National Collection of Type Cultures (NCTC, London), namely: NCTC 11637 (the type strain, cagA+), NCTC 12455 (strain 26695, cagA+), NCTC 12908 (G50, cagA−), and NCTC 13081 (strain Tx30a, cagA−).

ELISA FOR DETECTION OF SERUM IgG ANTIBODIES TO CAGA PROTEIN
Serum samples were obtained from blood, taken at the time of endoscopy, from the 62 patients and were stored at −80°C until serological testing was performed. An ELISA using a highly purified p120 antigen of *H. pylori* (Viva Diagnostica Helicobacter p120 ELISA; supplied by Axis-Shield, Dundee, UK) was performed according to the manufacturer’s instructions to detect anti-CagA serum IgG. The cut off values for the ELISA as recommended by the manufacturer were: < 5 units, negative; 5–7.5 units, borderline; > 7.5 units, positive. Control sera used with the kit were used for semiquantitative analysis.

EXTRACTION, PRIMERS FOR CAG PAI LOC1, AND PCR CONDITIONS
Genomic DNA was extracted from strains of *H. pylori* grown on 10% Columbia blood agar and purified by means of the Isoquick kit (Orca Research, Bothell, Washington, USA) or the cetrimethammonium bromide (CTAB) method. Diluted DNA (100 ng) was added to each PCR reaction containing: 10 mM Tris/HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.05 mM each deoxynucleotide, and 0.4 mM each oligonucleotide primer, which were synthesised by MWG-BIOTECH (Milton Keynes, UK). One unit Taq polymerase (Life Technologies, Paisley, UK) was added and PCR was performed in a Sprint thermal cycler (Hybaid Ltd, Ashford, UK).

Details of the primers and PCR conditions for the various assays have been described elsewhere. Briefly, these were as follows. PCR conditions for the two assays for cagA (a marker for the right hand end of the cag PAI and for the cagI region) were as described previously for the F1/B1 primers and the D008/F008 primers. The assay for cagD and cagE (also referred to as part of the picAB operon, where picA comprises cagC and cagD, and picB is cagE), which is located within the cagI region upstream of cagA, was used as a second marker for that region. The primers were designed to amplify both cagD and cagE and to generate a 2091 bp product based on the published sequences for NCTC 12455 (table 1). The PCR amplification comprised one cycle of 94°C for 110 seconds followed by 25 cycles of 94°C for 50 seconds, 54°C for 50 seconds, 72°C for 110 seconds, and finally 72°C for three minutes. The gene locus identified as a virD4 homologue (*Hp 524*) was selected as a marker for the cagH region, and two primers were designed to give a 1993 bp product based on the NCTC 12455 sequence (table 1). The PCR amplification conditions used comprised one cycle at 94°C for two minutes followed by 30 cycles at 94°C for one minute, 56°C for one minute, and finally 72°C for two minutes. The empty site (ES) assay was used to test for absence of the complete cag PAI. A 360 bp fragment was amplified with primers 2 and 25 specific for the genome regions flanking the left and right ends of the cag PAI insert location for NCTC 12455 (table 1). The PCR amplification conditions used were 15 cycles at 94°C for one minute, 57°C for one minute, and 72°C for one minute.

VACCULATING CYTOTOXIN GENOTYPING
The strains of *H. pylori* that were negative or borderline in the CagA ELISA were genotyped based on the signal (s1 and s2) and mid (m1 and m2) regions of the vacuolating cytotoxin gene (vacA). Genotyping was performed using previously described PCR conditions and the following primers: VA1-F and VA1-R for the s region, giving fragments of 259 bp for s1 and 289 bp for s2; and VAG-F and VAG-R for the m region, giving fragments of 570 bp for m1 and 645 bp for m2.
Table 2  Serology results with the CagA ELISA for serum from Helicobacter pylori infected individuals

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Positive</th>
<th>Negative</th>
<th>Borderline</th>
</tr>
</thead>
<tbody>
<tr>
<td>cagA+ (cagDE/virD4 not tested)</td>
<td>34</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>cagA−/cagDE+ (virD4 not tested)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

ELISA, enzyme linked immunosorbent assay; PAI, pathogenicity island.

**Results**

**Cag PAI Genotypes**

The 62 isolates of *H. pylori*, each from a different patient, were tested for the presence of the cag PAI using the panel of PCR assays, and the results are shown in table 2. A hierarchical scheme of assays was used to minimise the number of individual PCR tests, because it was established previously that a positive cagA result was a reliable predictor of positivity for cagD and cagE (picAB). Overall, 34 strains were shown to have cagA with one or both of the two cagA assays. Seventeen of those strains were negative with the F1/B1 primers but were positive with the D008/F008 primers. Strains of *H. pylori* that were confirmed negative for cagA in the above assays were tested also for the presence of cagD and cagE, which are key components of the cagI region of the cag PAI. No strains with the cagA−/cagDE+ genotype were identified.

Strains of *H. pylori* that were negative in the above PCR assays were then tested using the virD4 assay, which provided a marker of the cagII region. One positive strain appeared to have a diverged cag PAI, in which the two markers (cagA and cagDE) for the cagI region had been deleted, although the cagII region indicated by virD4 was conserved. Because the remaining 20 strains of *H. pylori* were negative in the above PCR assays, it may be that they have neither a complete nor a partial cag PAI. However, owing to the large size (about 40 kb) of the cag PAI, the absence of these specific markers did not necessarily indicate its complete absence. Such cagA−/cagDE−/virD4− strains were the second most common group and a feature of 20 of 62 of all strains in our series of patients. The distribution of the *H. pylori* cagA+ and cagA− strains according to the patient disease categories was as follows: 12 strains, peptic ulcer group; 19 strains, oesophagitis group; 16 strains, gastritis group, 14 strains endoscopically normal mucosa; and one strain, gastric cancer. The relative numbers of strains with the cagA positive genotype were similar in each disease group.

**Detection of Anti-CagA Antibody and Associations with Cag PAI Genotype**

The 62 human sera that matched the strains of *H. pylori* isolated from gastric biopsies before treatment were investigated. When tested for the presence of anti-CagA antibody by ELISA, a total of 34 patients showed CagA positivity, with two additional sera giving borderline results. Twenty one sera gave a negative ELISA result. Figure 1 shows the distribution of semiquantitative values obtained in the ELISA assay.

Table 2 lists the serology results with the CagA ELISA for serum from *Helicobacter pylori* infected individuals.

**Figure 1** The range of semiquantitative values for *Helicobacter pylori* CagA enzyme linked immunosorbent assay. Each point represents one strain. A, negative zone; B, borderline zone; C, positive zone; and D, strong positive zone.

**Table 3** Evaluation of ELISA for detecting anti-CagA antibody in sera of *Helicobacter pylori* infected patients

<table>
<thead>
<tr>
<th>Genotype</th>
<th>cagA+</th>
<th>cagA−</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA test +</td>
<td>36</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>ELISA test −</td>
<td>5</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>21</td>
<td>62</td>
</tr>
</tbody>
</table>

Borderline results are included as positive results. Sensitivity is 36/36, or 100%; specificity is 21/21, or 100%; predictive value of a positive test is 36/36, or 100%; predictive value of a negative test is 21/21, or 80.8%; accuracy is 36 + 21/62, or 91.9%.

**ATYPICAL ELISA RESULTS AND GENOTYPES OF THE INFECTING *H. PYLORI* STRAINS**

Atypical results were obtained for seven sera and paired cultures of *H. pylori*. Sera from five patients gave a negative CagA ELISA result although the infecting strain of *H. pylori* in each case was cagA positive (table 4). These five strains were cagA positive with the D008/D009 primers but were cagA negative with the F1/B1 primers. Two other atypical strains (H6 associated with duodenal ulcer, and H339 from a 68 year old patient with gastric neoplasia) were
also cagA+ but were borderline in the CagA ELISA assay. Additional genotyping (vacA sm and ES assay) of these seven strains showed that they were genomically diverse. The ES assay results were unusual because two strains gave a positive ES amplicon even though they contained cagA, and the other five strains gave a non-specific product in the ES assay. These atypical CagA ELISA results were most often associated with cultures and sera from patients aged over 55 with more severe disease such as duodenal ulcer (five patients) or gastric neoplasia (one patient).

### Discussion

Serological detection of infection with a cagA containing strain of *H pylori* by means of an anti-CagA ELISA is the only non-invasive diagnostic test at present available for assessing strain virulence potential and possible disease risk. It is important, therefore, to establish the reliability of CagA serology as a predictive test for determining the cagA genotype of the infecting strain because various serological assays are now available. Our study was designed to determine the link between the composition of the strain cag PAI and the host response to exported CagA protein, as measured by anti-CagA IgG serum production. One possible problem in assessing such a serological test is the reliability of the cagA PCR assay because the correct design of primers is crucially important as a result of strain genomic diversity. Our results indicated that the F1/B1 primer pair resulted in a 50% underestimation of the number of strains containing cagA. In a previous study of *H pylori* from Italian patients, the problem of different sets of cagA primers giving different results was highlighted and attributed to divergence in the primer target sequences. It is important, therefore, to investigate other genes within the cag PAI to establish alternative markers of the degree of conservation.

For most of the 62 patients in our study, we found a direct correlation between the presence of the cagA gene in the infecting strain of *H pylori* and anti-CagA antibody seropositivity of the host, giving an accuracy of 91.9% for the CagA ELISA assay. The sensitivity of the assay was 87.8% because five patients were serologically CagA negative, even though the cagA gene and the rest of the cag PAI were conserved in the genomes of each of the infecting strains.

These apparently “false negative” results in the ELISA assay have several possible explanations. First, the cagA gene could have been functional but mutations might have arisen in other loci within the cag PAI that modified the ability of the cell to export active CagA protein, so causing a lack of seropositivity. However, this is unlikely because most investigations show a close correlation between the presence of the cagA gene and expressed CagA protein. Second, sequencing of the cagA gene in *H pylori* shows a region of internal duplication present in some strains, which may be responsible for the heterogeneity seen in CagA size. This variation has been reported to be located in the 3′ region, with differences in the number and size of repeats. Our negative PCR results for the F1/B1 primers indicated variation within cagA of four of these “atypical” strains, and also for another strain (H339), which gave a borderline CagA ELISA result. Thus, genetic variation within cagA might generate antigenic diversity different from and not reflected in the CagA antigen coating the ELISA plate well, thereby resulting in failure to detect CagA antibody. Because details of the strain(s) used as source of antigen in the kit were not available, it was not possible to assess this further.

Another possible explanation of the discrepancy between the genetic and phenotypic assays was that the cultures of the infecting *H pylori* were a cagA+/cagA− mixture of cell types. Such mixtures of *H pylori* have been documented previously. In addition, the results of the ES PCR assay for most of these strains indicated that cagA− variants were present as well as cagA+ forms. If the proportion of cagA+ to cagA− cell types in the infecting *H pylori* population was low, then there might have been insufficient CagA protein exported in vivo to induce a host response, even though there was enough DNA from cagA− cells on subculture to give a positive result in the ES PCR assay. Host factors might also have contributed to the observed discrepancies because it has been suggested that host responses to *H pylori* may vary in different populations and in different age ranges. These particular patients were mostly over 55 years of age and had a severe disease (gastric ulcer or neoplasia), which might have contributed to a diminished immunological response to the *H pylori* infection. However, in older patients the incidence of positive antibodies initially increases with age, such that at 55 to 60 years, 50–60% of the general population will show raised antibody values, irrespective of whether they have an active infection, whereas antibody responses in the very elderly may be weak or non-existent because of a general decline in the immune system.

In summary, we found that the cag PAI was common in *H pylori* infected patients attending an open access endoscopy clinic in mid-Essex, and there was no reliable evidence that the cag PAI containing strains were more often associated with peptic ulcers in this series of patients. In general, the results using the Viva Diagnostics Helicobacter p120 ELISA kit for detecting
anti-CagA antibody in patient sera closely matched the infecting strain cag PAI genotype. False negative results (10% of sera) were best interpreted in the context of mixed infections, in particular cagA+/cagA− culture mixtures, as well as diversity within cagA, although limitations of the kit cannot be excluded because of some difficulties in interpreting borderline results. We conclude that the increasing availability of diagnostic tests, such as the CagA ELISA, provides the potential to make improvements in the management of \( H \) pylori infection in primary care. However, the results of such tests should be interpreted with caution in view of the fact that some negative ELISA results may lead to incorrect assumptions about the cagA status of the infecting \( H \) pylori strain. The ELISA appeared to be least reliable in older patients with more severe disease. Consequently, the value of predicting the clinical outcome of an \( H \) pylori infection based on cagA status alone must remain uncertain until the role of other possible determinants, such as host immunological and geographical status factors, as well as other \( H \) pylori strain specific virulence factors, have been clarified.

This work was supported in part by a three year award from the PHS Central R & D Fund (1995–98).

