Determination of *Helicobacter pylori* Virulence by Simple Gene Analysis of the *cag* Pathogenicity Island

TSUNEO IKENOUE,* SHIN MAEDA, KEIJI OGURA, MASAO AKANUMA, YUZO MITSUNO, YASUO IMAI, HARUHIKO YOSHIDA, YASUSHI SHIRATORI, AND MASAO OMATA

Department of Gastroenterology, Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

Received 10 May 2000/Returned for modification 4 August 2000/Accepted 20 September 2000

*N. pylori* is a gram-negative, spiral-shaped, microaerophilic bacterium that infects human gastric mucosa and is recognized as a major cause of chronic active gastritis, peptic ulcer disease, gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue lymphoma (10, 17, 18, 20, 23, 33). Although the pathogenesis of *N. pylori* infection is not well understood, there are several putative virulence factors that may contribute to mucosal damage by *N. pylori* infection.

The cytotoxin-associated-gene (*cag*) pathogenicity island (PAI) is an approximately 40-kb cluster of genes in the *N. pylori* chromosome (4, 29) and is divided into two regions, *cagI* and *cagII*. There are at least 14 and 16 open reading frames (ORFs) in *cagI* and *cagII*, respectively. Some of the ORFs in the *cag* PAI are believed to encode proteins which have similarities to other bacterial secretion systems, such as the *Bordetella pertussis* toxin secretion system (4).

The *cag* PAI is considered to be one of the major virulence factors of *H. pylori* (4). Extensive studies of the *cag* gene, located in the most downstream portion of the *cag* PAI, have indicated that the *cagA* gene is associated with peptic ulcer disease, gastric cancer, and mucosa-associated lymphoid tissue lymphoma in the stomach (3, 5, 11, 13, 14, 22, 24, 25, 27, 30, 32). Blaser et al. (3) revealed that *CagA* antibodies were more frequently detected in *H. pylori*-infected patients with gastric cancer than in those without gastric cancer (odds ratio, 1.9). Furthermore, Parsonnet et al. (24) showed that subjects infected with *H. pylori* who had *CagA* antibodies were more likely to develop gastric cancer as compared with uninfected subjects (odds ratio, 5.8), while *H. pylori*-infected subjects without *CagA* antibodies were at only slightly and not significantly increased risk for cancer (odds ratio, 2.2). Thus, the *cagA* gene is conventionally used as a marker of pathogenic strains. However, several studies suggest that the *cagA* gene cannot be used as a suitable marker for *cag* PAI-associated virulence for the following reasons: (i) although *cag* PAI-intact *H. pylori* strains are shown to induce interleukin-8 secretion from gastric epithelial cells (1, 4, 6, 7, 15, 26), an inactivation of some *cag* PAI genes such as *cagE* but not *cagA* causes a marked reduction in the ability of *H. pylori* to induce interleukin-8 induction (1, 4, 8, 21, 31); (ii) we have previously shown that some Japanese strains obtained from patients with nonulcer dyspepsia lack most of the *cag* PAI genes, including the promoter region of the *cagA* gene, despite the presence of the *cagA* gene itself, indicating that the presence of the *cagA* gene does not always signify the presence of an intact *cag* PAI and an ability to produce *CagA* protein (15). Although recent studies have revealed that the *CagA* protein is translocated into the host cells and tyrosine phosphorylated, the precise role of the *CagA* protein in *H. pylori* pathogenesis is still unknown (2, 19, 28). These findings may suggest that a gene other than *cagA* can be used as a marker for *cag* PAI-associated virulence.

By using our recombinant *CagA* protein and antibodies, we

### Table 1. Sequences and locations of oligonucleotide primers

| Primer  | Primer sequence | Location |
|---------|-----------------|----------|
| cagA-F1 | 5’-AACAGGACAAGTACTACGAC-3’ | 157–161† |
| cagA-F2 | 5’-GATAAACAGGCAAAATTTTGTC-3’ | 157–161† |
| cagA-R1 | 5’-TTATATCCGTGGTGGTGGC-3’ | 157–161† |
| cagA-R2 | 5’-TGAGAAAAAGATTTTGTCGACA-3’ | 157–161† |
| cagE-F1 | 5’-CCGGATTTGTGTTTGTGGTC-3’ | 157–161† |
| cagE-R1 | 5’-GAAGTGGTTAAAAAATCAATGCC-3’ | 157–161† |
| cagT-F1 | 5’-CAATCTTATATGCTAGTGC-3’ | 157–161† |
| cagT-R1 | 5’-CATCACACACACCTTGTAGT-3’ | 157–161† |
| cagAP-F1 | 5’-CTGGAATAAAATGTGAATCG-3’ | 18,738–18,757† |
| cagAP-F2 | 5’-CTCCTTCTCCCAACATT-3’ | 18,738–18,757† |
| LEC-F1 | 5’-ACATCTTGTCCATATATACGCT-3’ | 3,920–3,942‡ |
| LEC-F2 | 5’-ATAAGCTGTCTTGGTGCATAGA-3’ | 3,920–3,942‡ |
| LEC-R1 | 5’-CTTTCGATTGCACATTGCT-3’ | 4,303–4,323‡ |
| LEC-R2 | 5’-ATCATTAGCTCTGTTAG-3’ | 4,303–4,323‡ |

* Corresponding author. Mailing address: Department of Gastroenterology, Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. Phone: 81-3-3815-5411 ext. 3070. Fax: 81-3-3814-0021. E-mail: ikenoue-2im@h.u-tokyo.ac.jp.

* GenBank accession no.: *, AF001357; †, U60176; ‡, AC000108.
previously showed a high prevalence of CagA-producing *H. pylori* strains in Japan (13, 14). Furthermore, by using Southern blot hybridization with DNA probes obtained by cloning 15 different ORFs in *cag* PAI, we clarified the details of DNA structure in the entire *cag* PAI (15) and found that Japanese strains deficient in the *cagA* gene lacked most of the *cag* PAI genes, including the *cagE* gene and the promoter region of the *cagA* gene. However, it is troublesome to check all *cag* PAI genes by Southern blotting.

Thus, in the present study, we attempted to establish a simple and practical method for determining the structure of *cag* PAI and consequently discriminating between Japanese *H. pylori* strains causing different rates of disease progression.

A total of 204 *H. pylori* isolates was obtained from *H. pylori*-infected adults who had undergone upper gastrointestinal endoscopy at Tokyo University Hospital. The patients consisted of 145 men and 59 women with a mean age of 58.5 years (ranging from 22 to 85 years). Patient endoscopic findings were as follows: gastric cancer in 55 patients, gastric ulcer in 22, duodenal ulcer in 17, both gastric and duodenal ulcers in 14, and chronic gastritis in 96.

FIG. 1. Structure of *cag* pathogenicity island and locations of PCR primers. (A) Locations of PCR primers for *cagE*, *cagT*, and the LEC. Two sets of primers were used to detect the *cagA* gene: primer set A1 (cagA-F1 and cagA-R1) and set A2 (cagA-F2 and cagA-R2). For detection of the promoter region of *cagA*, primer sets AP1 (cagAP-F1 and cagA-R2) and AP2 (cagAP-F2 and cagA-R2) were used, and for the left end of *cagII*, primer sets LEC1 (LEC-F1 and LEC-R1) and LEC2 (LEC-F2 and LEC-R2) were used. The expected lengths of PCR products using each primer set are shown. The sequence and location of each primer are shown in Table 1. (B) Locations of PCR primers for the *cagA* gene and its promoter region. The expected lengths of PCR products using each primer set are also shown.
–80°C in brucella broth with 5% (vol/vol) fetal bovine serum containing 16% (vol/vol) glycerol. DNA was prepared as described previously (15).

Five different loci allowing for structure screening of cag PAI were selected on the basis of our previous Southern blot analysis (15). The previous study revealed that when the ORFs of cag PAI were deleted, the deletions started from the region between cagA and the cagA promoter region through cagQ (cagI) and continued from cagS through cag-13 or cag-8 (cagII) (15). Thus, cagA, the cagA promoter region, and cagE were selected to represent cagI, and cagT and the left end of cagII (LEC) were selected to represent cagII. Therefore, overall five loci were selected.

Pairs of oligonucleotide primers were used to detect the presence of the cag PAI genes cagA, the cagA promoter region, cagE, cagT, and the LEC, containing both inside and outside genes of cag PAI, and these primer pairs were designed on the basis of published sequences reported by Censini et al. (GenBank accession number, U60176), Akopyants et al. (GenBank accession number, AC000108), Tomb et al. (GenBank accession number, AE000511), and ourselves (GenBank accession number, AF001357) (Table 1; Fig. 1). As shown in Fig. 1A and B, two sets of primers were used to detect the cagA gene (sets A1 and A2), the cagA promoter region (sets AP1 and AP2), and the LEC (sets LEC1 and LEC2). To detect cagE and cagT, one set of primers was used for each gene, set E1 and set T1, respectively. H. pylori strains ATCC 43526 and 43579, which have been determined to have the entire cag PAI (15), were used as positive controls for each PCR. Since eight cagA gene-positive strains from Western countries, including Tx30a, kindly provided by J. C. Atherton (Nottingham University, United Kingdom), were determined to lack the entire cag PAI (15), these were used as negative controls. The genomic DNAs from other bacterial species—Escherichia coli, Pseudomonas aeruginosa, Serratia marcescens, Haemophilus influenzae, Streptococcus pneumoniae, Campylobacter fetus, Campylobacter jejuni, Klebsiella pneumoniae, Citrobacter freundii, and Enterobacter aerogenes—were tested using each primer set to assess the specificity of each PCR.

For histological analysis, biopsy specimens from corpus and antrum were embedded in paraffin, stained with hematoxylin and eosin, and examined by two pathologists blinded to the patient’s clinical diagnosis or characteristics of the H. pylori strain. The presence of chronic active gastritis was determined by scoring the following parameters on the basis of the updated Sydney System (9): density of inflammatory infiltration (0 to 3)

### TABLE 2. Relationship between presence of cag PAI genes and clinical diagnosis

| Diagnosis            | Patients with cag type infection: |  |
|----------------------|-----------------------------------|--|
|                      | Intact PAI | Partially deleted PAI | Totally deleted PAI |
| Peptic ulcer disease | 53         | 0                      | 0                   |
| Gastric cancer       | 55         | 0                      | 0                   |
| Chronic gastritis    | 89         | 5                      | 2                   |
| Total                | 197        | 5                      | 2                   |

FIG. 2. Results of cag PAI gene PCR in 204 Japanese H. pylori isolates. (A) Positivity of cag PAI gene PCR by each primer set: *, seven cagA promoter PCR-negative isolates also negative for both cagE and cagT PCR; **, two isolates negative for LEC PCR and also negative for cagA PCR. (B) Types of cag PAI structures determined in this study.
and density of neutrophil infiltration (0 to 3). For each parameter, 0 is none, 1 is mild, 2 is moderate, and 3 is severe.

PCR amplification specificity for \( cagA \), the \( cagA \) promoter region, \( cagE \), \( cagT \), and the LEC was assessed by testing \( H. pylori \) strains ATCC 43526 and 43579 and eight \( cagA \) gene-negative strains from Western countries, as well as 10 other bacterial species. Only \( H. pylori \) strains ATCC 43526 and 43579 were positive for PCR amplification of all five loci. Eight \( cagA \) gene-negative Western strains and the other bacterial species tested were all negative for PCR of all five loci. Thus, the specificity of PCR for each PAI locus was 100%.

To assess the sensitivity of PCR for each locus, excluding the \( cagA \) promoter region, PCR was performed with 30 \( H. pylori \) isolates from Japanese patients whose \( cag \) PAI gene status was determined by Southern blot analysis in our previous study (15). PCR results for \( cagA \) (primer sets A1 and A2), \( cagE \) (primer set E1), \( cagT \) (primer set E1), and the LEC (primer sets LEC1 and LEC2) were completely consistent with those of previous Southern blot analyses (15). Thus, if at least two sets of primers were used, the specificity of PCR for each PAI locus was 100%.

As shown in Fig. 2 and Table 2, 202 out of 204 (99.0%) isolates were positive for \( cagA \) and LEC, and 197 out of 204 (96.6%) isolates were positive for the \( cagA \) promoter region, \( cagE \), and \( cagT \). Since two \( cagA \)-negative strains were also negative for all other genes tested and the remaining five out of seven \( cagA \) promoter-negative strains were negative for \( cagE \) and \( cagT \), the \( cag \) PAI genes present in Japanese \( H. pylori \) isolates were divided into three types; intact-PAI, partially deleted-PAI, and totally deleted-PAI genes (Fig. 2).

Recently, Jenk et al. reported that the presence of the entire \( cag \) PAI is highly related to duodenal ulcers but that the clinical outcome of \( H. pylori \) infection is not reliably predicted by analyzing several genes of the \( cag \) PAI, including \( cagA \), \( cagE \), and \( cagT \) (12). In their study, the presence of \( cagE \) was completely consistent with that of \( cagA \) but not \( cagT \). In contrast, our study revealed consistency in the presence of \( cagE \) with \( cagT \) but not \( cagA \), indicating that the strain diversity may exist.

FIG. 3. Relationship between gastric inflammation and the presence of \( cag \) PAI genes of infected \( H. pylori \) strains. A total of 64 patients (59 infected with strains with intact type, 4 with partially deleted type, and 1 with totally deleted type) were assessed for (A) inflammatory infiltration and (B) neutrophil infiltration in corpus and antral mucosa. Each valuable was scored on a four-point histological scale (0, none; 1, mild; 2, moderate; and 3, severe). *, statistically significant by Mann-Whitney U test.
in relation to cag PAI genes among Western countries and Japan.

Strains with partially or totally deleted cag PAIs, which lack both cagE and cagT, were more frequently found in more patients with chronic gastritis only (7 out of 96 patients [7.3%]) than with peptic ulcer disease (0 out of 53; \( P = 0.042 \)) or with gastric cancer (0 out of 55; \( P = 0.039 \)) (Table 2). Furthermore, by assessing inflammation activity in the gastric mucosa of 64 patients (59 infected with intact-PAI-type strains, 4 with partially deleted-PAI-type strains, and 1 with a totally deleted-PAI-type strain), we found no significant differences in inflammatory infiltration of corpus between patients with intact type strains and those with partially or totally deleted type strains. However, inflammatory infiltration in antrum (Fig. 3A) and neutrophil infiltration in corpus and antrum (Fig. 3B) were significantly milder in patients with partially or totally deleted type strains than in those with intact type strains. These findings suggest that the strains with partially or totally deleted PAI may have weaker ability to cause disease progression than those with intact PAI.

In the present study, the partially or totally deleted type strains in Japan lacked cagE, cagT, and the cagA gene promoter region, regardless of the presence of the cagA gene itself. Therefore, they could be discriminated from intact type strains by detection of cagE, cagT, or the cagA gene promoter region but not by detection of the cagA gene itself. Although the cagA gene is conventionally used as a marker for virulence, especially with the PCR amplification method, our results indicate that not the cagA gene itself but the promoter region of the cagA gene could be a better marker. However, due to the diversity of the cagA gene promoter region sequences, designating specific primers to detect this region may be difficult. Since the cagE gene is located near the cagA gene promoter region and retained consistently within this region, it seems valid to choose the cagE gene as a substitute for the cagA gene promoter region. Since the primer sets designed for cagE PCR in this study were extremely specific and sensitive, at least for Japanese strains, we conclude that cagE PCR can be used as a practical method for screening the status of the cagA gene structure, which may be related to disease progression, for a large number of samples in order to test clinical significance or to conduct an epidemiological survey.

In conclusion, the results of the present study indicate that cagE is more accurate, as a marker of an intact cag PAI, than the cagA gene and that it seems to be more useful in discriminating between \( H. pylori \) strains with different rates of disease progression in Japan. Detection of the cagE gene by PCR amplification with specific primers can be used as a simple and practical method for their discrimination.

REFERENCES

1. Akopyants, N. S., S. W. Clifton, D. Kersulyte, J. E. Crabtree, B. E. Youree, C. A. Reece, N. O. Bukhanov, E. S. Drzezka, B. A. Roe, and D. E. Berg. 1998. Analyses of the cag pathogenicity island of \( H. pylori \). Mol. Microbiol. 28:57–53.

2. Asahi, M., T. Azuma, S. Ho, H. Suto, Y. Nagai, M. Tsubokawa, Y. Tohyama, S. Maeda, M. Omata, T. Suzuki, and C. Sasakawa. 2000. \( H. pylori \) CagA protein can be tyrosine phosphorylated in gastric epithelial cells. J. Exp. Med. 191:593–602.

3. Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura. 1995. Infection with \( H. pylori \) strains possessing cagA associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res. 55:2111–2115.

4. Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiaia, M. Borodovsky, R. Rappuoli, and A. Covacci. 1996. cag, a pathogenicity island of \( H. pylori \), encodes type-specific and disease associated virulence factors. Proc. Natl. Acad. Sci. USA 93:14648–14653.
Interleukin-8 response of gastric epithelial cell lines to Helicobacter pylori stimulation in vitro. Infect. Immun. 63:1681–1687.

27. Shimoyama, T., S. Fukuda, M. Tanaka, T. Minami, Y. Saito, and A. Munakata. 1997. High prevalence of the CagA-positive Helicobacter pylori strains in Japanese asymptomatic patients and gastric cancer patients. Scand. J. Gastroenterol. 32:465–468.

28. Stein, M., R. Rappuoli, and A. Coracci. 2000. Tyrosine phosphorylation of the Helicobacter pylori CagA antigen after cag-driven host cell translocation. Proc. Natl. Acad. Sci. USA 97:1263–1268.

29. Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, R. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenny, L. M. Fitzegerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter. 1997. The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature 338:539–547.

30. Tummuru, M. K. R., T. L. Cover, and M. J. Blaser. 1993. Cloning and expression of a high-molecular-mass major antigen of Helicobacter pylori: evidence of linkage to cytotoxin production. Infect. Immun. 61:1799–1809.

31. Tummuru, M. K. R., S. A. Sharma, and M. J. Blaser. 1995. Helicobacter pylori p6B, a homologue of the Bordetella pertussis toxin secretion protein, is required for induction of IL-8 in gastric epithelial cells. Mol. Microbiol. 18:867–876.

32. Weel, J. E. L., R. W. M. van der Hulst, Y. Gerrits, P. Roorda, M. Feller, J. Dankert, G. N. J. Tytgat, and A. van der Ende. 1996. The interrelationship between cytotoxin-associated gene A, vacuolating cytotoxin, and Helicobacter pylori-related diseases. J. Infect. Dis. 173:1171–1175.

33. Wotherspoon, A. C., C. Doglioni, T. C. Diss, L. Pan, A. Moschini, M. de Boni, and P. G. Isaacson. 1993. Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of Helicobacter pylori. Lancet 342:575–577.