Published in final edited form as:

Helicobacter. 2013 February ; 18(1): 66–72. doi:10.1111/j.1523-5378.2012.00994.x.

Intact long-type *dupA* as a marker for gastroduodenal diseases in Okinawan subpopulation, Japan

Ayaka Takahashi¹, Seiji Shiota¹, Osamu Matsunari¹, Masahide Watada¹, Rumiko Suzuki¹, Saori Nakachi¹, Nagisa Kinjo², Fukunori Kinjo², and Yoshio Yamaoka^{1,3,*}

¹Department of Environmental and Preventive Medicine, Oita University Faculty of Medicine, Yufu, 879-5569, Japan

²Department of Endoscopy, University Hospital, University of the Ryukyus, 207 Uehara, Nishihara, Okinawa 903-0215, Japan

³Department of Medicine-Gastroenterology, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, Texas, 77030 USA

Abstract

Background—*Helicobacter pylori dupA* can be divided into two types according to the presence or absence of the mutation. In addition, full-sequenced data revealed that *dupA* has two types with different lengths depend on the presence of approximately 600 bp in the putative 5' region (presence; long-type and absence; short-type), which has not been taken into account in previous studies.

Methods—A total of 319 strains isolated from Okinawa, the south islands of Japan, were included. The status of *dupA* and *cagA* was determined by polymerase chain reaction. The presence of mutations in long-type *dupA* was determined by DNA sequencing.

Results—The prevalence of long-type *dupA* was 26.3% (84/319). Sequence analysis showed that there were only 6 cases (7.1%) with point mutations lead to stop codon among 84 long-type *dupA* strains studied. Interestingly, intact long-type *dupA* without frameshift mutation, but not short-type *dupA* was significantly associated with gastric ulcer and gastric cancer than gastritis ($P = 0.001$ and $P = 0.019$, respectively). After adjustment by age, gender and *cagA*, the presence of intact long-type *dupA* was significantly associated with gastric ulcer and gastric cancer compared with gastritis (odds ratio [OR] = 3.35, 95% confidence interval [CI] = 1.55–7.24 and OR = 4.14, 95% CI = 1.23–13.94, respectively).

Conclusions—Intact long-type *dupA* is a real virulence marker for severe outcomes in Okinawa, Japan. The previous information gained from PCR-based methods without taking long-type *dupA* into account must be interpreted with caution.

Keywords

Helicobacter pylori; intact long-type *dupA*; gastric ulcer; duodenal ulcer; gastric cancer

Corresponding author: Yoshio YamaokaMD, PhD, Department of Environmental and Preventive Medicine, Oita University Faculty of Medicine, 1-1 Idaigaoka, Hasama-machi, Yufu-City, Oita 879-5593, Japan, Tel: +81-97-586-5740; Fax: +81-97-586-5749, yyamaoka@oita-u.ac.jp.

disclosures

Statements of Interests

All of the authors declare no competing interests.

Introduction

Helicobacter pylori infection is now accepted to be linked to severe gastritis-associated diseases, including peptic ulcer and gastric cancer (GC) (1). The infection remains latent in the majority of infected patients, only a minority of individuals with *H. pylori* infection ever develop it (2). In addition to host, environmental, and dietary factors, another possible reason for the varying outcomes of *H. pylori* infection relates to differences in the virulence of *H. pylori* strains.

Several *H. pylori* virulence factors associated with peptic ulcer and GC have been reported, including *cagA*, *vacA*, *babA* and *oipA* (1, 3–6). However, no such factors have been exclusively linked to a specific *H. pylori*-related disease (e.g., GC). In 2005, our group (7) described a novel virulence factor; duodenal ulcer promoting (*dupA*) gene, which located in the plasticity region of the *H. pylori* genome. *dupA* is homologous to *virB4*, a gene encoding a component protein of the type IV secretion system (T4SS). We reported that infections with *dupA*-positive strains increased the risk for duodenal ulcer (DU) but were protective against GC both in Asian and Western countries (7). However, roles of *dupA* were still controversial since several studies were unable to reproduce the observation in other populations (8–12).

Recent full-sequenced data of *H. pylori* revealed that the length of the *dupA* open reading frame (ORF) depends on the strains; that of strains Shi470 and G27 have approximately 600 bp longer ORF (approximately 2,500 bp) than that of strain J99 due to the additional 5' region of *dupA* (13, 14). This suggests that *dupA* has two genotypes according to the sequence of the putative 5' region (presence; long-type and absence; short-type); however all previous studies did not take the additional 5' region into account. In addition, previous studies indicated that some strains have a presence of mutations in *dupA*, which created a premature stop codon (10, 15, 16). Hussein *et al.* recently classified a *dupA* allele with 1,884 bp (short-type) as *dupA1* and a truncated version by the mutations as *dupA2* (15). They also mentioned the presence of the long-type in “Note added in proof” (15). We hypothesized that only intact *dupA* (i.e., long-type without frameshift mutation) is a real virulence marker for severe outcomes. We therefore examined the *dupA* genotypes, presence of mutations, and the association with clinical outcomes in Okinawan subpopulation in Japan.

Methods

Patients

H. pylori strains were obtained from the gastric mucosa of *H. pylori*-infected patients who underwent endoscopy at University of the Ryukyus (Okinawa, Japan) between February 1993 and March 2005. We included patients with *H. pylori* infection confirmed by culture. Gastritis, DU, gastric ulcer (GU), and GC were identified by endoscopy, and GC was further confirmed by histopathology. Patients with a history of gastric resection were excluded. Written informed consent was obtained from the all participants, and the protocol was approved by the Ethics Committee of University of the Ryukyus.

Isolation and genotyping of *H. pylori*

Antral biopsy specimens were obtained for isolation of *H. pylori* using standard culture methods (17). *H. pylori* DNA was extracted from confluent plate cultures using a commercially available kit (QIAGEN, Valencia, CA). The *dupA* status and the presence of *cagA* were determined by polymerase chain reaction (PCR). The presence of long-type *dupA* was determined with the primer sets (*dupA*-F2 and *dupA*-R2) according to the *dupA* sequence of strain Shi470 deposited in Genbank (Table 1 and Figure 1). *dupA*-F2 primer was located on the start codon which include the approximately 600 bp 5'-region of *dupA* of

strain Shi470, which were absent in strain J99. *dupA*-R2 primer was located on the end of *dupA* of strain Shi470. We confirmed these regions were well conserved by BLAST search. These primers generated PCR products with 2,499 bp in strain Shi470. Amplification of *H. pylori* genomic DNA sequences was performed in a total volume of 50 µL containing 5 µL of PCR buffer, 2.5mM MgCl₂, 2.5 mM of each deoxynucleotide, 2.5 U of Takara LA Taq® (Takara Biotechnology Co., LTD, Japan), 0.2 µM of each primer, and more than 10 ng of *H. pylori* DNA. The PCR conditions were initial denaturation for 5 min at 95°C, 35 amplification steps (95°C for 30 s, 56°C for 30 s, and 72°C for 30 s), and a final extension cycle of 7 min at 72°C.

The status of *cagA* was determined by PCR for conserved region of *cagA* and for direct sequencing using primer pair *cagTF*; 5'-ACC CTA GTC GGT AAT GGG-3' and *cagTR*; 5'-GCT TTA GCT TCT GAY ACY GC-3' (Y = C or T) designed in the 3' repeat region of *cagA*, as described previously (18). We also constructed a new primer pair, *cagOMF*; 5'-AGC AAA AAG CGA CCT TGA AA-3' and *cagOMR*; 5'-ATT CAC GAG CTT GAG CCA CT- 3', to exclude the false negative cases (19). The PCR conditions were initial denaturation for 5 min at 95°C, 35 amplification steps (95°C for 30 s, 56°C for 30 s, and 72°C for 30 s), and a final extension cycle of 7 min at 72°C, using Blend Taq® DNA polymerase (TOYOBO, Japan). These primers generated PCR products with approximately 500 to 950 bp depend on the number of the repeat sequences. The absence of *cagA* was confirmed by the presence of *cagA* empty site using primer pair (Luni1 and R5280), as previously described (20).

Nucleotide sequence of *dupA*

To examine the presence of mutations in *dupA*, we performed DNA sequencing. First, we used primer pairs (*dupA*-F0, *dupA*-F1 and *dupA*-R1) to generate the PCR product including the long-type *dupA*. *dupA*-F1, *dupA*-F0 and *dupA*-R1 primers were located in HPSH_0425 and HPSH_0410 of strain Shi470, respectively. These primer pairs (*dupA*-F0 and *dupA*-R1, *dupA*-F1 and *dupA*-R1) generated PCR products with 3,096 and 2,918 bp in strain Shi470, respectively. The PCR conditions of these primers were the same as the case of primer pair *dupA*-F2 and *dupA*-R2. Six primers (*dupA*-F2, *dupA*-F3, *dupA*-F4, *dupA*-F5, *dupA*-F6, *dupA*-F7) were used for DNA sequencing by using PCR product of *dupA*-F0 and *dupA*-R1, *dupA*-F1 *dupA*-R1 as shown in Table 2. In case primer pairs (*dupA*-F0 and *dupA*-R1, *dupA*-F1 *dupA*-R1) did not show positive band, we used PCR product using *dupA*-F2 and *dupA*-R2. In such case, we used reverse primer F3R, which reverse sequence of *dupA*-F3 primer to confirm the sequence around the start codon. In case primers (*dupA*-F2 and *dupA*-R2) did not show the positive bands, the presence of short-type *dupA* was determined with the primer sets (*dupA*-F5 and *dupA*-R3) according to the *dupA* sequence deposited in Genbank (Figure 1). *dupA*-F5 were designed by Lu *et al.* (7). This primer pair generated PCR products with 468 bp in strain Shi470.

PCR products were purified with QIAquick Purification Kit (QIAGEN Inc., Santa Clarita, CA) and DNA direct sequencing was performed using AB 3130 Genetic Analyzer (Applied Biosystems by Life Technologies, Carlsbad, CA) according to the manufacturer's instruction.

Statistical analysis

The univariate association was quantified by the chi-square test. A multivariate logistic regression model was used to calculate the odds ratios (OR) and 95% confidence interval (CI) of the clinical outcomes by including age, gender, and *H. pylori* genotypes. Spearman rank coefficients (*r*) were also determined to evaluate the association between *cagA* and intact long-type *dupA*. A P value of less than 0.05 was accepted as statistically significant.

The SPSS statistical software package version 18.0 (SPSS, Inc., Chicago, IL) was used for all statistical analyses.

Results

The full-length of *dupA* in the full-sequenced *H. pylori* strains

By November 2011, full-sequenced data of 31 *H. pylori* strains were deposited in Genbank. Among them, 12 strains have *dupA* and their *dupA* structure was shown in Figure 1. The length of *dupA* (2,499 bp) in strain Shi470, Cuz20, SouthAfrica7, F57, Puno135 and SNT49 have approximately 600 bp longer than that of strain J99 due to further extension in the 5' region of *dupA*. Strain G27 also has this additional sequence in the 5' region of *dupA*; however this strain contains stop codon at position of 1,184 bp in *dupA*. Strain Gambia94/24 has similar length of *dupA* with J99. Strain 908, and its chronological sub-clones (strains 2017 and 2018) also has similar length of *dupA* with J99; however, this gene was reported to be located out of ORF.

The prevalence of intact long-type *dupA* and *cagA*

A total of 319 patients (196 males and 123 females, mean age 55.5 ± 13.9 years, 88 with gastritis, 114 with DU, 95 with GU, and 22 with GC) were included in this study. To examine the presence of long-type *dupA*, we used primer sets (*dupA*-F2 and *dupA*-R2) according to the *dupA* sequence of strain Shi470 (Figure 1). The prevalence of long-type *dupA* was 26.3% (84/319) (Table 3). On the other hand, the prevalence of short-type *dupA* was 6.6% (21/319).

Previous reports showed that approximately 16–29% of *dupA* was truncated due to insertion/deletion in the gene (10, 16). It suggests that even though the strain has long-type *dupA*, if *dupA* is truncated by mutations, it will be non-functional. We therefore checked the prevalence of frameshift mutations among strains with long-type *dupA*. Full-length sequence of *dupA* was completed in all 84 strains with long-type *dupA*. Among them, frameshift mutation was found in only six strains (7.1%), which were isolated from patients with 2 gastritis, 1 DU and 3 GU. These strains can be considered as non-intact long-type *dupA*. As a result, the prevalence of intact long-type *dupA* was 24.5% (78/319).

The prevalence of *cagA* was overall 87.5% (279/319). There was a significant positive correlation between the presence of *cagA* and intact long-type *dupA* ($r = 0.12$, $P = 0.02$).

The association between intact long-type *dupA* and *cagA* with clinical outcomes

The prevalence of intact long-type *dupA* was significantly higher in strains from patients with GU and GC than that from gastritis (34.7, 36.4, and 13.6%, $P = 0.001$, $P = 0.019$, respectively) (Table 3). That of DU was also higher than that from gastritis (21.9 vs. 13.6%); however the differences did not reach statistical significance ($P = 0.13$). The prevalence of short-type *dupA* was not significantly different among different clinical outcomes (gastritis 5.7%, DU 5.3%, GU 9.5%, and GC 4.5%, respectively). The prevalence of *cagA* was significantly higher in strains from patients with DU than that from gastritis (90.4 vs. 80.7%, $P = 0.04$), whereas the prevalence of *cagA* was not significantly different between the strains from patients with GU, GC and gastritis ($P = 0.14$ and 0.08 , respectively).

After adjustment by age, gender and *cagA* in multivariate analysis, the presence of intact long-type *dupA* was significantly associated with GU compared with gastritis (OR = 3.35, 95% CI = 1.55–7.24) (Table 4). Similarly, intact long-type *dupA* was independently associated with GC (OR = 4.14, 95% CI = 1.23–13.94). These data suggests that intact long-type *dupA* is superior marker for the risk of developing GU and GC to the presence of *cagA*.

Nucleotide sequence accession numbers

Nucleotide sequences of long-type *dupA* are available under the DDBJ accession numbers xxxx to xxxx (under submission).

Discussion

In 2005, we found the first disease-specific *H. pylori* virulence factor that induced DU and had a suppressive action on GC, which we named as *dupA* (7). Although recent meta-analysis showed that *dupA* was associated with DU, the association of *dupA* on clinical outcomes is different in different populations (21). Interestingly, Gomes *et al.* reported that there were frameshift mutations in 14/86 (16%) of *dupA* (10) which created a premature stop codon and may have considerable effects on protein expression or function. It is clear that strains with these mutated sequences would not produce intact DupA protein. Intriguingly, the presence of *dupA* without stop codon was more frequently found in strains from patients with DU than gastritis and GC in their following study (16). However, importantly, there were few point mutations in *dupA* in our strains studied, suggesting that frameshift mutations do not an important issue to be considered for evaluating the intact status of *dupA* in Japanese population, at least in Okinawa.

In our initial study in 2005, primer pair for *dupA* was constructed based on strain J99 (7). However, full-sequenced data of *H. pylori* revealed that the length of the *dupA* depends on the strains; *dupA* of strain Shi470 is approximately 600 bp longer than that of strain J99 due to the extension of the 5' region of *dupA* (13, 14). Strains Gambia94/24 and J99 are chimera in comparison with other strains. The recombination point is located at the position -67 bp from the start codon of *jhp0917* of J99. It is not clear whether the evolution of *dupA* are related with this chimera or not. In all previous studies, the 5' region has not been taken into account in investigating the importance of *dupA*. Importantly, we found that long-type *dupA*, but not short-type *dupA* was significantly associated with severe gastroduodenal diseases. In addition, we found that frameshift mutation was found in only six strains (7.1%) in long-type *dupA*, which inconsistent with previous reports from other countries (10, 16). Importantly, the presence of intact long-type *dupA* was superior distinguishing marker than *cagA* for severe gastroduodenal diseases. Our data suggest that intact long-type *dupA* might produce a functional DupA protein. In contrast, non-intact long-type *dupA* may not produce functional DupA protein. This might be the reason why the importance of *dupA* on clinical outcomes was conflicting in different studies. The previous information gained from PCR-based methods without taking long-type *dupA* into account must be interpreted with caution.

In our previous study, *in vitro* experiments showed that the presence of *dupA* was associated with increased susceptibility to low pH (7). Recent study showed that the gastric acid output was significantly higher in the *dupA*-positive than the *dupA*-negative group (22). These findings assume the positive association between the presence of *dupA* and DU. However, intact long-type *dupA* was significantly associated with GU and GC, but not with DU in the present study. Although our meta-analysis showed no association between *dupA* and GC, several studies showed the positive trend for *dupA* and GC (21). Initial experiments showed that the presence of *dupA* was associated with increased IL-8 production from both *in vitro* and *vivo* (7). Furthermore, one study showed that the presence of *dupA* is associated with increased IL-8 production from gastric mucosa *in vivo* (15). Importantly, *in vitro* experiments using *dupA* mutants showed that *dupA* substantially increased *H. pylori*-induced IL-12 production by CD14⁺ mononuclear cells (15). The production of other Th1-associated cytokines and IL-8 were also modestly induced by the cells. These data suggest that *dupA*-positive *H. pylori* can also be associated with GU and GC by inducing inflammatory cytokines but not increased gastric acid. It is necessary to examine whether intact long-type *dupA* strains can induce the inflammation or not in the future study.

In addition, the *vir* genes exist before and after the region of the *dupA* locus (23). In strain Shi470, for example, *virB2*, *virB3*, *virB4* (*dupA*), *virB8*, *virB9*, *virB10*, *virB11*, *virD4*, and *virD2* were detected (13). These are structurally similar to the T4SS called *cag* PAI and ComB and thought to be the third T4SS. These observations suggest that only strains that are intact *dupA*-positive and form a novel T4SS might be involved in gastroduodenal diseases. Intriguingly, we recent reported that complete *dupA* cluster (possessing *dupA* and all adjacent *vir* genes) was significantly associated with DU in the United States (24). Nevertheless, the presence of intact long-type *dupA* alone was significantly associated with severe outcomes in this study. At this point, we have not examined the presence of complete *dupA* cluster in Okinawan samples. Complete *dupA* cluster with intact long-type *dupA* might be true virulence factors. Further study from other geographic area will help to elucidate of the importance of intact long-type *dupA* and complete *dupA* cluster.

However, our study includes several limitations. The *cagA*-negative strains in this study were relatively prevalent than that from other part of Japan (17). This finding was consistent with previous study from Okinawa (25). Okinawa consists of islands in southwestern Japan. Though the prevalence of *H. pylori* in Okinawa is not different from other part of Japan (26, 27), the incidence of GC (6.3 deaths/100,000 population) in Okinawa is the lowest in Japan (mean mortality rate of Japan; 11.8 deaths/100,000 population) in 2009 (This is available in the Center for Cancer Control and Information Services, National Cancer Center, Japan, <http://www.ncc.go.jp/>). Recent our study showed that *H. pylori* strains in Okinawa were different from typical Japanese strains (19). Therefore, we cannot apply as representative of Japan. Further study from other part of Japan is necessary to confirm the importance of intact long-type *dupA* in Japan. Second, our study was based on the results of PCR. Among 103 *dupA*-positive strains, 21 strains (20.0%) showed positive with primer sets only for short-type *dupA* but not long-type one. It remains unclear whether short-type *dupA* without mutation produce a functional DupA protein as well as intact long-type *dupA*. Measuring intact DupA protein using immunoblotting techniques will be necessary in the future study.

In conclusion, intact long-type *dupA* was significantly associated with GU and GC from gastritis in Okinawa, Japan. Interestingly, intact long-type *dupA* was superior marker than *cagA*. These observations suggest that only intact long-type *dupA*-positive strains might be involved in gastroduodenal diseases.

Acknowledgments

We thank Ms. Miyuki Matsuda for excellent technical assistance.

This report is based on work supported in part by grants from the National Institutes of Health (DK62813) (YY), and grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (22390085 and 22659087) (YY), (23790798) (SS) and Special Coordination Funds for Promoting Science and Technology from the MEXT of Japan.

References

1. Suerbaum S, Michetti P. *Helicobacter pylori* infection. N Engl J Med. 2002 Oct; 347(15):1175–1186. [PubMed: 12374879]
2. Kusters JG, van Vliet AH, Kuipers EJ. Pathogenesis of *Helicobacter pylori* infection. Clin Microbiol Rev. 2006 Jul; 19(3):449–490. [PubMed: 16847081]
3. Atherton J, Cao P, Peek RJ, Tummuru M, Blaser M, Cover T. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem. 1995 Jul; 270(30):17771–17777. [PubMed: 7629077]
4. Basso D, Zambon C, Letley D, Stranges A, Marchet A, Rhead J, et al. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology. 2008 Jul; 135(1):91–99. [PubMed: 18474244]

5. Gerhard M, Lehn N, Neumayer N, Borén T, Rad R, Schepp W, et al. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. *Proc Natl Acad Sci U S A*. 1999 Oct; 96(22):12778–12783. [PubMed: 10535999]
6. Yamaoka Y, Kikuchi S, el-Zimaity H, Gutierrez O, Osato M, Graham D. Importance of *Helicobacter pylori oipA* in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. *Gastroenterology*. 2002 Aug; 123(2):414–424. [PubMed: 12145793]
7. Lu H, Hsu P, Graham D, Yamaoka Y. Duodenal ulcer promoting gene of *Helicobacter pylori*. *Gastroenterology*. 2005 Apr; 128(4):833–848. [PubMed: 15825067]
8. Argent R, Burette A, Miendje Deyi V, Atherton J. The presence of *dupA* in *Helicobacter pylori* is not significantly associated with duodenal ulceration in Belgium, South Africa, China, or North America. *Clin Infect Dis*. 2007 Nov; 45(9):1204–1206. [PubMed: 17918084]
9. Douraghi M, Mohammadi M, Oghalaie A, Abdirad A, Mohagheghi M, Hosseini M, et al. *dupA* as a risk determinant in *Helicobacter pylori* infection. *J Med Microbiol*. 2008 May; 57(Pt 5):554–562. [PubMed: 18436587]
10. Gomes L, Rocha G, Rocha A, Soares T, Oliveira C, Bittencourt P, et al. Lack of association between *Helicobacter pylori* infection with *dupA* -positive strains and gastroduodenal diseases in Brazilian patients. *Int J Med Microbiol*. 2008 Apr; 298(3-4):223–230. [PubMed: 17897881]
11. Schmidt H, Andres S, Kaakoush N, Engstrand L, Eriksson L, Goh K, et al. The prevalence of the duodenal ulcer promoting gene (*dupA*) in *Helicobacter pylori* isolates varies by ethnic group and is not universally associated with disease development: a case-control study. *Gut Pathog*. 2009; 1(1): 5. [PubMed: 19338650]
12. Nguyen L, Uchida T, Tsukamoto Y, Kuroda A, Okimoto T, Kodama M, et al. *Helicobacter pylori dupA* gene is not associated with clinical outcomes in the Japanese population. *Clin Microbiol Infect*. 2010 Aug; 16(8):1264–1269. [PubMed: 19832706]
13. Kersulyte D, Lee W, Subramaniam D, Anant S, Herrera P, Cabrera L, et al. *Helicobacter Pylori*'s plasticity zones are novel transposable elements. *PLoS One*. 2009; 4(9):e6859. [PubMed: 19727398]
14. Kersulyte D, Kalia A, Gilman RH, Mendez M, Herrera P, Cabrera L, et al. *Helicobacter pylori* from Peruvian Amerindians: Traces of Human Migrations in Strains from Remote Amazon, and Genome Sequence of an Amerind Strain. *PLoS One*. 2010; 5(11):e15076. [PubMed: 21124785]
15. Hussein N, Argent R, Marx C, Patel S, Robinson K, Atherton J. *Helicobacter pylori dupA* is polymorphic, and its active form induces proinflammatory cytokine secretion by mononuclear cells. *J Infect Dis*. 2010 Jul; 202(2):261–269. [PubMed: 20533870]
16. Queiroz DM, Rocha GA, Rocha AM, Moura SB, Saraiva IE, Gomes LI, et al. *dupA* polymorphisms and risk of *Helicobacter pylori* -associated diseases. *Int J Med Microbiol*. 2011 Mar; 301(3):225–228. [PubMed: 21050811]
17. Yamaoka Y, Kodama T, Kita M, Imanishi J, Kashima K, Graham D. Relationship of *vacA* genotypes of *Helicobacter pylori* to *cagA* status, cytotoxin production, and clinical outcome. *Helicobacter*. 1998 Dec; 3(4):241–253. [PubMed: 9844065]
18. Yamaoka Y, Osato M, Sepulveda A, Gutierrez O, Figura N, Kim J, et al. Molecular epidemiology of *Helicobacter pylori*: separation of *H. pylori* from East Asian and non-Asian countries. *Epidemiol Infect*. 2000 Feb; 124(1):91–96. [PubMed: 10722135]
19. Matsunari O, Shiota S, Suzuki R, Watada M, Kinjo N, Murakami K, et al. Association between *Helicobacter pylori* virulence factors and gastroduodenal diseases in Okinawa, Japan. *J Clin Microbiol*. 2012 Mar; 50(3):876–883. [PubMed: 22189111]
20. Mukhopadhyay AK, Kersulyte D, Jeong JY, Datta S, Ito Y, Chowdhury A, et al. Distinctiveness of genotypes of *Helicobacter pylori* in Calcutta, India. *J Bacteriol*. 2000 Jun; 182(11):3219–3227. [PubMed: 10809703]
21. Shiota S, Matsunari O, Watada M, Hanada K, Yamaoka Y. Systematic review and meta-analysis: the relationship between the *Helicobacter pylori dupA* gene and clinical outcomes. *Gut Pathog*. 2010; 2(1):13. [PubMed: 21040520]
22. Imagawa S, Ito M, Yoshihara M, Eguchi H, Tanaka S, Chayama K. *Helicobacter pylori dupA* and gastric acid secretion are negatively associated with gastric cancer development. *J Med Microbiol*. 2010 Dec; 59(Pt 12):1484–1489. [PubMed: 20829397]

23. Yamaoka Y. Roles of the plasticity regions of *Helicobacter pylori* in gastroduodenal pathogenesis. J Med Microbiol. 2008 May; 57(Pt 5):545–553. [PubMed: 18436586]
24. Jung SW, Sugimoto M, Shiota S, Graham DY, Yamaoka Y. The intact *dupA* cluster is a more reliable *Helicobacter pylori* virulence marker than *dupA* alone. Infect Immun. 2012 Jan; 80(1): 381–387. [PubMed: 22038914]
25. Azuma T, Yamazaki S, Yamakawa A, Ohtani M, Muramatsu A, Suto H, et al. Association between diversity in the Src homology 2 domain--containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis. 2004 Mar; 189(5): 820–827. [PubMed: 14976598]
26. Nobuta A, Asaka M, Sugiyama T, Kato M, Hige S, Takeda H, et al. *Helicobacter pylori* infection in two areas in Japan with different risks for gastric cancer. Aliment Pharmacol Ther. 2004 Jul; 20(Suppl 1):1–6. [PubMed: 15298598]
27. Ito S, Azuma T, Murakita H, Hirai M, Miyaji H, Ito Y, et al. Profile of *Helicobacter pylori* cytotoxin derived from two areas of Japan with different prevalence of atrophic gastritis. Gut. 1996 Dec; 39(6):800–806. [PubMed: 9038660]

Figure 1. Schematic representation of the ORFs of the two *dupA* allele types

Strains Shi470, Cuz20, SouthAfrica7, F57, Puno135 and SNT49 have approximately 600 bp longer than that of other strains (e.g., strain J99 or Gambia94/24) in the 5' region of *dupA*. Strain G27 also has this additional sequence in the 5' region of the *dupA*; however this strain contains stop codon at position of 1184 bp in *dupA* (black circle). Strain Gambia94/24 has a *dupA* ORF of approximately 1800 bp. Strain J99 also has short type *dupA*; however this strain contains stop codon in *dupA* (black triangle). Strains 908 and chronological sub-clone 2017, 2018 have the same *dupA* pattern as Gambia94/24; however, this gene was reported to be located out of open reading frame and therefore is not shown in this figure. Other 19 strains do not possess *dupA*. Each arrow shows the annealing positions of the forward primers and reverse primers.

Table 1

Oligonucleotide primers used for PCR

Primer	Sequence (5'-3')	PCR condition	Product size (bp)
<i>dupA</i> -F2	ATG TTT CTT GGT TTA GAG GG		
<i>dupA</i> -R2	TTA TAC ATA TTG AAT ATT CTC GC	95°C 30 s, 56°C 30 s, and 72°C 30 s	2499
<i>dupA</i> -F5	GGT TTC TAC TGA CAG AGC GC		
<i>dupA</i> -R3	CGT ATT TAG TCA GTA AGT TGG CG	95°C 30 s, 56°C 30 s, and 72°C 30 s	468
<i>cagTF</i>	ACC CTA GTC GGT AAT GGG		
<i>cagTR</i>	GCT TTA GCT TCT GAY ACY GC (Y=C or T)	95°C 30 s, 56°C 30 s, and 72°C 30 s	500
<i>cagOMF</i>	AGC AAA AAG CGA CCT TGA AA		
<i>cagOMR</i>	ATT CAC GAG CTT GAG CCA CT	95°C 30 s, 56°C 30 s, and 72°C 30 s	950
Lini1	ACA TTT TGG CTA AAT AAA CGC TG		
R5280	GGT TGC ACG CAT TTT CCC TTA ATC	95°C 30 s, 56°C 30 s, and 72°C 30 s	550

Table 2Oligonucleotide primers used for DNA sequencing of *Helicobacter pylori* *dupA*

Primer	Sequence (5'-3')	Corresponding DNA sequence
<i>dupA</i> -F0	TGG CGT GTG GCA GTC TAA TGC	+497 – +476
<i>dupA</i> -F1	GCA AAC GCT CAA ACT ATT GCC	+311 – +296
<i>dupA</i> -F2	ATG TTT CTT GGT TTA GAG GG	1 – 20
<i>dupA</i> -F3	CAG AAC ACA AGC TTT AAA TGA ATT G	219 – 243
<i>dupA</i> -F4	ATG AGT TCT ATA CTA ACA GAC TTT GAG CC	616 – 646
<i>dupA</i> -F5	GGT TTC TAC TGA CAG AGC AC	1017 – 1036
<i>dupA</i> -F6	GAT AAT TGG TAG CAC AGG AAG CGG	1411 – 1434
<i>dupA</i> -F7	GGC TCT AGC GAA CAA GAT TTT AAT GAG	1787 – 1812
<i>dupA</i> -R1	GCT CAA CAA AAT GCC CAC CAG TCG C	2577 – 2601

Table 3

Prevalence of *dupA* and *cagA* among each disease

	Total	AG	DU	GU	GC
N	319	88	114	95	22
Age	55.5±13.9	55.0±13.1	54.5±13.2	54.8±15.2	65.4±11.7
Male	196 (61.4%)	41 (46.6%)	69 (60.5%)*	70 (73.7%)*	16 (72.7%)*
<i>cagA</i>	279 (87.5%)	71 (80.7%)	103 (90.4%)*	84 (88.4%)	21 (95.5%)
long type <i>dupA</i>	84 (26.3%)	14 (15.9%)	26 (22.8%)	36 (37.9%)*	8 (36.4%)*
intact long <i>dupA</i>	78 (24.5%)	12 (13.6%)	25 (21.9%)	33 (34.7%)*	8 (36.4%)*
short type <i>dupA</i>	21 (6.6%)	5 (5.7%)	6 (5.3%)	9 (9.5%)	1 (4.5%)

DU: duodenal ulcer, GU: gastric ulcer, GC: gastric cancer

* P < 0.05, compared with that of gastritis

Table 4

Multivariate analyses of the risk for gastric ulcer (a), and gastric cancer (b) by age, gender, *cagA*, and long-type *dupA*

(a)			
	Adjusted OR	95% CI	P
Age (per one year)	0.99	0.97–1.01	0.71
Gender (male)	3.20	1.68–6.06	< 0.001
<i>cagA</i>	1.18	0.49–2.84	0.7
intact long-type <i>dupA</i>	3.35	1.55–7.24	0.002
(b)			
	Adjusted OR	95% CI	P
Age (per one year)	1.09	1.03–1.15	0.001
Gender (male)	2.22	0.69–7.06	0.17
<i>cagA</i>	4.85	0.56–41.71	0.15
intact long-type <i>dupA</i>	4.14	1.23–13.94	0.02

OR: Odds ratio, CI: confidence interval