Journal of Clinical Microbiology

The Geographic Origin of Helicobacter pylori Influences the Association of the *homB* Gene with Gastric Cancer

Jieun Kang, Kathleen R. Jones, Sungil Jang, Cara H. Olsen, Yun-Jung Yoo, D. Scott Merrell and Jeong-Heon Cha

J. Clin. Microbiol. **2012, 50(3):1082. DOI: 10.1128/JCM.06293-11.**
Published Ahead of Print 28 December 2011.

JCM
Journals.ASM.org

The Geographic Origin of *Helicobacter pylori* Influences the Association of the *homB* Gene with Gastric Cancer

Jieun Kang,[a] Kathleen R. Jones,[b] Sungil Jang,[a] Cara H. Olsen,[c] Yun-Jung Yoo,[a] D. Scott Merrell,[b] and Jeong-Heon Cha[a,d]

Department of Oral Biology, Oral Science Research Center, BK21 Project, Research Center for Orofacial Hard Tissue Regeneration, Yonsei University College of Dentistry, Seoul, South Korea[a]; Department of Microbiology and Immunology, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA[b]; Department of Preventive Medicine and Biometrics, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA[c]; and Oral Cancer Research Institute, Yonsei University College of Dentistry, Seoul, South Korea[d]

We found that South Korean *Helicobacter pylori* isolates predominantly carry *homB* at locus B and that there is no association between the *homB* allele and the *cagA* allele or the development of gastric cancer within this population. Uniquely, several East Asian strains carried multiple copies of the *hom* genes.

Helicobacter pylori colonizes the gastric mucosas of over 50% of the world's population (6, 13) and is the etiological agent of gastritis, duodenal ulcers, gastric ulcers, gastric adenocarcinoma, and mucosa-associated lymphoid tissue lymphoma (4, 5, 19, 20, 22). Due to this bacterium's association with gastric cancer, which is the second most common cause of cancer-associated death (14), the World Health Organization has classified *H. pylori* as a class I carcinogen (8). Gastric cancer mortality rates vary geographically; the highest rates are in East Asian countries like China, Japan, and South Korea, which also display high rates of *H. pylori* infection (7, 8, 21, 23). Clearly, gastric diseases are due, at least in part, to infection by *H. pylori*, and the ultimate disease developed appears to be affected by variability in *H. pylori* virulence factors. Recently, we presented detailed epidemiological studies of *cagA* and *vacA* from a collection of 260 isolates from South Korea (9, 10, 11). Our studies showed that there is a significant association between infection with *H. pylori* strains carrying the EPIYA-ABD *cagA* genotype and the development of gastric cancer (11). Moreover, the majority of *H. pylori* isolates encoded the most virulent CagA (EPIYA-ABD) and VacA (s1/i1/m1) proteins (9, 11). The polymorphisms in *cagA* and *vacA*, alone and in concert, impact the progression to severe gastric disease, but the impact of these two virulence factors alone is not sufficient to explain the vast discrepancy in gastric cancer rates between East Asian and Western populations. Thus, it is important to examine the impacts of different virulence factors among both Western and East Asian populations (9, 10, 11, 16).

In vitro, *Helicobacter* outer membrane B (HomB) promotes the secretion of the proinflammatory cytokine interleukin-8 (IL-8) and increases *H. pylori*'s ability to adhere to host cells (15). More importantly, *homB* presence is significantly associated with development of peptic ulcer disease in Portuguese children and young adults (15, 18) and with gastric cancer development and the presence of *cagA* in U.S. and Colombian populations (12). These findings suggest that the outer membrane protein HomB is a novel virulence factor. Thus, it and other members of the small paralogous family of *hom* adhesion molecules are currently being investigated. The two best-studied *hom* genes, *homA* and *homB*, are 90% identical at the nucleotide level (2, 3, 18). These *homA* and *homB* genes can be present at two different loci within the *H. pylori* genome: locus A and locus B. Strains can carry a single copy of one of the *hom* genes, a double copy of a single gene, a single copy of each gene, or neither gene (15, 16). Previous studies suggest geographic differences, either in distribution, location, or copy number, of the *hom* genes in the genome and suggest that these differences influence any association with disease outcome (15, 16, 17).

In the present study, our collection of 260 South Korean isolates was assessed for any associations between the distribution of the *homA* or *homB* genes and disease state, as well as any associations between the *hom* genes and the different *cagA* and *vacA* alleles. The South Korean isolates include 115 isolates from patients diagnosed with gastritis, 60 isolates from patients diagnosed with gastric ulcers, 55 isolates from patients diagnosed with duodenal ulcers, and 30 isolates from patients diagnosed with gastric cancer (9, 10, 11). A complete description of all strains can be found in Table S1 in the supplemental material.

To analyze the *hom* genotype of the South Korean *H. pylori* strains, the presence of the *homA* and/or *homB* gene(s) was identified by a single PCR with the *hom* primers hf and hr (Table 1 and Fig. 1). We successfully genotyped 225 samples for which we had complete epidemiological data for the *hom* genes (Fig. 2; see also Table S2 in the supplemental material). Of note, two strains showed an amplicon with a length that was intermediate (approximately 146 bp) of what was expected for either the *homA* PCR product (~128 bp) or the *homB* PCR product (~161 bp). This intermediate-length *hom* or *ihom* genotype has been previously described and shown to be due to random deletions and/or insertions within the *hom* genes (12).

Once the strains were genotyped for the presence of the two *hom* genes, we next sought to define the copy numbers and locations of the genes. This was accomplished through two additional PCRs (Fig. 1). The distribution of *homA/homB* is shown in Fig. 2. Within this pop-

Received 4 November 2011 **Accepted** 19 December 2011
Published ahead of print 28 December 2011
Address correspondence to D. Scott Merrell, dmerrell@usuhs.mil, or Jeong-Heon Cha, Jcha@yuhs.ac.
J. Kang and K. R. Jones contributed equally to this article.
Supplemental material for this article may be found at http://jcm.asm.org/.

doi:10.1128/JCM.06293-11

TABLE 1 Primer sequences

Primer name	Primer sequence	Reference
hf (F1-jhp0870/jhp0649)	AGAGGGTGTTTGAAACGCTCAATA	18
hr (R1-jhp0870/jhp0649)	GGTGAATTCTTCTGCGGTTTG	18
Af (F1-jhp0648/HP0709)	TAATTTCGCGCAAAAACATC	18
Ar (R1-jhp0650/HP0711)	ATTCCAGCGCCTAATGGAC	18
Bf (F1-jhp0869/HP0935)	AAGAGGATTGCGTGGTGGAGTTG	18
Br (R1-jhp0871/HP0936)	GGGTTGCCTTTGGGCTTGGA	18
K-Af	TGGAATATTGATATAAAGAAGTG	This study
K-Ar	GGGTTTAATAGGATGAGCCGC	This study
K-Bf	GATTTTCCCCACTCTTTTTATGG	This study
K-Br	GGTTTTTGTCCATGAACATGC	This study

ulation, 212 isolates carried a single *hom* gene at locus B (35, −/*homA*; 175, −/*homB*; and 2, −/*ihom*). This is in contrast to Western strains that carry a single *hom* gene at locus A 100% of the time (15, 16). Also, this distribution is different from what has been reported for Western strains, which show a much more evenly distributed population of isolates carrying *homA* or *homB* (15, 16). In our population, three isolates were indeterminate for the presence of the *homB* gene at locus A but were positive for occupation at locus B ($homB^{+/-}$/*homB*). These 3 $homB^{+/-}$/*homB* isolates were included as *homB*-positive strains for the statistical analysis but were eliminated from the data set when assessing the impact of multiple copies of the *hom* genes. Six strains carried multiple *hom* genes (4 *homA/homA*, 1 *homB/homB*, and 1 *homA/homB*), which is again in contrast to previous studies that suggested that Western strains carry multiple copies of *homA* and/or *homB* but East Asian strains do not (15, 16). Our finding is perhaps not surprising since our collection of East Asian isolates is much larger than the collection previously examined.

Finally, four strains failed to amplify either *hom* gene at either locus A or locus B and therefore were considered *hom* negative.

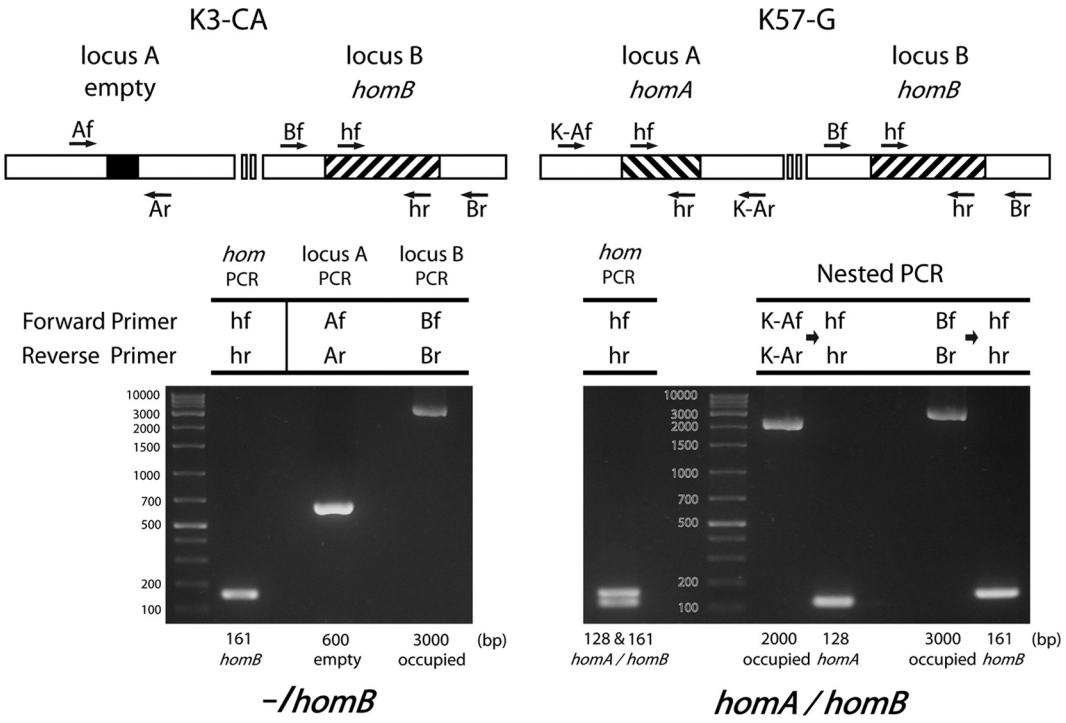

FIG 1 Genotyping of the *hom* genes at the respective loci. (Top) Schematic representation of the two loci where the *hom* genes are traditionally found, locus A and locus B. The annealing positions (arrows) and names of the primers used in this study are shown. The presence of a *hom* gene in a particular locus is depicted by the presence of a dashed box. (Bottom) The strains were genotyped for the *hom* gene by a single PCR with the hf and hr primers. A PCR amplicon of 128 bp indicates the presence of the *homA* gene, and an amplicon of 161 bp denotes the presence of the *homB* gene. In order to determine the location (locus A or B) of the *hom* gene, two additional PCRs were performed. To amplify locus A, primers Af and Ar were used, and to amplify locus B, the Bf and Br primers were used, as previously described (18). If an indeterminate result was obtained from the PCR using the Af and Ar primers or Bf and Br primers, another PCR with the K-Af and K-Ar primers or K-Bf and K-Br primers (Table 1) was performed. These K-Af/K-Ar and K-Bf/K-Br primers were designed according to the genome sequences of the Korean *H. pylori* strains HP51 and HP52 (GenBank accession numbers CP000012 and CP001680, respectively). For locus A, a resulting amplicon of 300 to 900 bp indicates that locus A is empty, whereas the presence of a 2,000- to 2,500-bp amplicon confirms that locus A is occupied by a *hom* gene. In the case of locus B, a 1,300- to 1,800-bp amplicon denotes that B is empty and the presence of a 2,500- to 4,000-bp amplicon indicates that locus B is occupied by a *hom* gene. (Left) For the K3-CA strain, the PCR with the hf and hr primers yielded a single amplicon of 161 bp (second lane), denoting the presence of the *homB* gene. The PCR with the Af and Ar primers (amplifying locus A) yielded a 600-bp product, indicating that locus A is empty (third lane), whereas a 3,000-bp amplicon produced from the PCR using the Bf and Br primers indicates that locus B is occupied (fourth lane). These results indicate that K3-CA has a genotype of −/*homB*. (Right) For K57-G and any strains that carried both *homA* and *homB*, an additional set of nested PCRs was also performed. First, PCRs with the hf and hr primers yielded two differently sized amplicons, a 128-bp amplicon, indicating the presence of a *homA* allele, and a 161-bp amplicon, indicating the presence of a *homB* allele (first lane). Next, a PCR using the Af and Ar primers or the K-Af and K-Ar primers yielded a 2,000-bp amplicon, which denotes an occupied locus. In the case of K57-G (right, third lane), the K-Af and K-Ar primers were used to amplify locus A because it showed no band for PCR using the Af/Ar primers. This PCR product was then purified and used as the template in a PCR using the hf and hr primers (fourth lane). This PCR yielded a 128-bp amplicon, which indicates that *homA* is located at locus A. Next, a PCR was performed using the Bf and Br primers, which yielded a 3,000-bp product (fifth lane). This again indicates an occupied locus B, and this PCR product was purified and used as the template in a PCR using the hf and hr primers. This PCR yielded a 161-bp amplicon (right, sixth lane), which indicates that *homB* is located at locus B. Thereby, these results indicate that K57-G has a genotype of *homA/homB*.

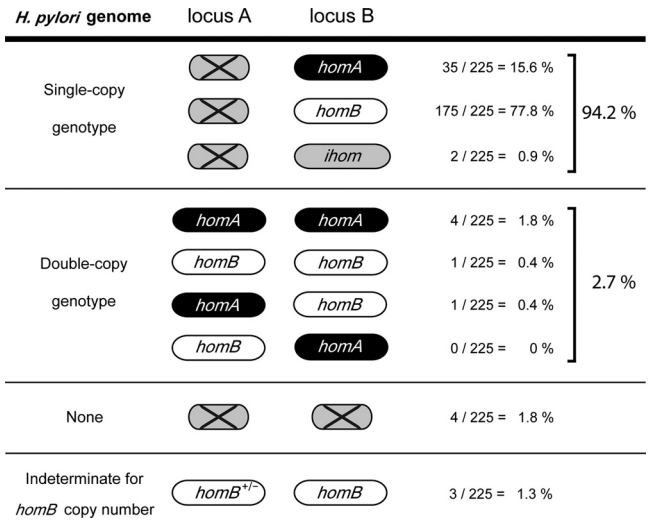

FIG 2 A schematic of the distribution of *hom* genes at the respective loci within this South Korean population is shown. *homB*$^{+/-}$ indicates that the amplification of locus A was unsuccessful. Therefore, the strains are *homB* positive and there is at least a single copy of *homB* found at locus B, but it cannot be determined whether or not they have two copies of *homB*. These strains were included as *homB*-positive strains for the statistical analysis but were eliminated from the data set when assessing the impact of multiple copies of the *hom* genes. To the right of the schematic, the percentage of the overall population of each individual genotype is indicated.

For these four −/− strains, PCR amplification of locus A and locus B yielded products that were indicative of an empty locus. However, each of the four −/− strains indicated the presence of a *hom* gene through the *hom* PCR amplification: two had *homA*, one had *ihom*, and one had *homA* and *homB*. This suggests that for these strains, *homA* or *homB* is presumably located at an alternate unknown location within the genome. It is interesting to speculate that perhaps *homA* or *homB* may be carried in the normal locus for the virtually unstudied *homC* or *homD* gene. Further study is clearly necessary to elucidate whether the location of the *homA/homB* genes corresponds to a functional difference and whether *homA/homB* can be located at other locations besides locus A and locus B. Of note, no strains carrying *homA*/−, *homB*/−, or *homB/homA* alleles were found.

We next assessed whether there was an association between the distribution of the individual *hom* genes and the disease state. A complete breakdown of the relationship between the *hom* alleles and the disease state is provided in Tables S1 and S2 in the supplemental material. The Fisher exact test was used to analyze two-way associations with SAS version 9.1 software (SAS Institute Inc., Cary, NC). There was no association between the distribution of the *hom* gene and the disease state among this population of East Asian strains ($P = 0.9978$), a result which is in direct contrast to the association found in Western populations (12, 15). In fact, there was no association regardless of which disease states were assessed (Table 2).

Since there is a statistical association between the presence of *homB* and gastric cancer in Western populations, which carry the single *hom* gene at locus A, and not in East Asian populations, which carry the single *hom* gene at locus B, these data perhaps suggest that the location of the *hom* gene within the genome is important. Genes carried at one particular locus could be expressed at greater levels; the promoters of *homA* and *homB* may differ enough to influence transcriptional levels of each gene, or the different loci within the genome may provide the ability to bind and influence overall levels of the *hom* transcript for different enhancers/inhibitors of each *hom* gene.

Recently, it has become clear that individual virulence factors interact in order to impact *H. pylori* pathogenesis (1, 9, 10, 24). Since *homB* is associated with *cagA* within Western populations (12, 15), we next assessed the distribution of *hom* genes in combination with the *cagA* alleles, the *vacA* alleles, and the disease state. A complete breakdown of the strains based on these factors is provided in Table S3 in the supplemental material. We first assessed if there was any association between the distribution of the *hom* alleles and the different *cagA* alleles (a canonical EPIYA-ABD versus all other EPIYA motifs). We found that there was no association between the distribution of the *hom* alleles and the different *cagA* alleles ($P = 0.0872$). Furthermore, when each gene was considered separately, there was no association between the distribution of *homA* ($P = 0.6139$) or that of *homB* ($P = 0.2217$) across the different *cagA* alleles. We next analyzed the association between the distribution of the *hom* alleles and the different *vacA* alleles (see Table S3 in the supplemental material) and found that the distribution of *vacA* alleles among the two *hom* genes was statistically significant ($P = 0.0142$). This association was dependent only on the *homB* allele ($P = 0.0275$), since there was no association between the distribution of the *homA* allele and that of the *vacA* allele ($P = 0.3955$). The overall association between the *vacA* alleles and the *hom* alleles was also influenced by the distribution of *cagA* alleles; the association was present in the non-EPIYA-ABD population ($P = 0.0319$) but not in the EPIYA-ABD population ($P = 0.1014$). Due to this difference, we used log-linear modeling to determine if there was a three-way association between the *cagA*, *vacA*, and *hom* alleles. However, no association between these three virulence factors was identified ($P = 0.681$). Another interesting aspect of this *vacA/hom* association was that it appeared to be influenced by the age of the patient; the association only became evident in the population above 60 years of age ($P = 0.0076$). A higher-order association between the *cagA*, *vacA*, and *hom* alleles and disease states was also assessed, but no significant associations existed between the virulence factors and the disease state.

TABLE 2 Significance of the associations between the distributions of *homA/homB* and the disease states

Disease comparison	*P* value for distribution of:		
	hom[a]	*homA*	*homB*
Across all diseases	0.9978	0.9802	0.9335
Gastritis vs all other diseases	0.8953	0.7215	0.7311
Duodenal ulcers vs all other diseases	1.0000	1.0000	1.0000
Gastric ulcers vs all other diseases	0.9219	1.0000	1.0000
Gastric cancer vs all other gastric diseases	0.8805	0.7931	0.6121
Peptic ulcers (both duodenal and gastric ulcers) vs gastritis and gastric cancer	1.0000	1.0000	1.0000
More severe diseases (gastric ulcer and gastric cancer) vs less severe diseases (gastritis and duodenal ulcer)	0.9379	0.8481	0.7146
Gastritis vs peptic ulcers vs gastric cancer	0.9803	0.9084	0.7894

[a] *hom* represents the distribution of the *homA*, *homB*, and *hom*-negative strains for the different disease states listed.

In conclusion, this is the first study to assess the association between the presence of the *homB* gene and gastric cancer in a population of predominantly East Asian strains; we found that the impact of the *homB* allele on disease is geographically dependent. In Western strains, there is a more even distribution of the *homA* and *homB* genes, while in East Asian strains, *homB* is more common (Fig. 2) (15, 16). Moreover, Western strains carry a single *hom* gene at locus A, whereas East Asian strains carry a single *hom* gene at locus B (Fig. 2) (15, 16). This study was the first to identify the presence of any East Asian isolates that carry multiple copies of the *hom* genes. Interestingly, in this population, no association between the presence of *homB* and the progression to gastric cancer was found (Table 2), suggesting that a hierarchy of virulence factors exists and that virulence factors have different impacts on disease based on the presence of other virulence factor polymorphisms. Within East Asian strains, EPIYA-ABD CagA appears to be the "master" virulence factor. *En masse*, these data exemplify the need for information about the presence and function of different virulence factors within different populations and the need to develop geographically tailored treatment regimens.

ACKNOWLEDGMENTS

We thank Jeannette Whitmire for critical reading of the manuscript.

This work was supported by the Basic Science Research Program through the National Research Foundation of South Korea (NRF), funded by the Ministry of Education, Science, and Technology (R13-2003-013-04002-0) (to J.-H.C.).

The contents of this article are the sole responsibility of the authors and do not necessarily represent the official views of the Department of Defense or the funding agency.

REFERENCES

1. **Akada JK, et al.** 2010. *Helicobacter pylori* CagA inhibits endocytosis of cytotoxin VacA in host cells. Dis. Model Mech. **3**:605–617.
2. **Alm RA, et al.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397**:176–180.
3. **Alm RA, Trust TJ.** 1999. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. J. Mol. Med. **77**:834–846.
4. **Blaser MJ.** 1998. *Helicobacter pylori* and gastric diseases. BMJ **316**:1507–1510.
5. **Ernst PB, Gold BD.** 2000. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu. Rev. Microbiol. **54**:615–640.
6. **Eurogast Study Group.** 1993. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. Gut **34**:1672–1676.
7. **Gwack J, et al.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br. J. Cancer **95**:639–641.
8. **International Agency for Research on Cancer.** 1994. Infection with *Helicobacter pylori*. IARC Monogr. Eval. Carcinog. Risks Hum. **61**:177–240.
9. **Jang S, et al.** 2010. Epidemiological link between gastric disease and polymorphisms in VacA and CagA. J. Clin. Microbiol. **48**:559–567.
10. **Jones KR, et al.** 2011. Polymorphisms in the intermediate region of VacA impact *Helicobacter pylori*-induced disease development. J. Clin. Microbiol. **49**:101–110.
11. **Jones KR, et al.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. J. Clin. Microbiol. **47**:959–968.
12. **Jung SW, Sugimoto M, Graham DY, Yamaoka Y.** 2009. *homB* status of *Helicobacter pylori* as a novel marker to distinguish gastric cancer from duodenal ulcer. J. Clin. Microbiol. **47**:3241–3245.
13. **Matysiak-Budnik T, Megraud F.** 1997. Epidemiology of *Helicobacter pylori* infection with special reference to professional risk. J. Physiol. Pharmacol. **48**(Suppl 4):3–17.
14. **Neugut AI, Hayek M, Howe G.** 1996. Epidemiology of gastric cancer. Semin. Oncol. **23**:281–291.
15. **Oleastro M, et al.** 2008. Evaluation of the clinical significance of *homB*, a novel candidate marker of *Helicobacter pylori* strains associated with peptic ulcer disease. J. Infect. Dis. **198**:1379–1387.
16. **Oleastro M, et al.** 2009. Allelic diversity and phylogeny of *homB*, a novel co-virulence marker of *Helicobacter pylori*. BMC Microbiol. **9**:248.
17. **Oleastro M, et al.** 2009. Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. Gut Pathog. **1**:12.
18. **Oleastro M, Monteiro L, Lehours P, Megraud F, Menard A.** 2006. Identification of markers for *Helicobacter pylori* strains isolated from children with peptic ulcer disease by suppressive subtractive hybridization. Infect. Immun. **74**:4064–4074.
19. **Parsonnet J, et al.** 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N. Engl. J. Med. **325**:1127–1131.
20. **Parsonnet J, et al.** 1994. *Helicobacter pylori* infection and gastric lymphoma. N. Engl. J. Med. **330**:1267–1271.
21. **Shin A, et al.** 2005. A nested case-control study of the association of *Helicobacter pylori* infection with gastric adenocarcinoma in Korea. Br. J. Cancer **92**:1273–1275.
22. **Talley NJ, et al.** 1991. Gastric adenocarcinoma and *Helicobacter pylori* infection. J. Natl. Cancer Inst. **83**:1734–1739.
23. **Tokudome S, et al.** 2007. Are there any real *Helicobacter pylori* infection-negative gastric cancers in Asia? Asian Pac. J. Cancer Prev. **8**:462–463.
24. **Yamaoka Y, et al.** 1998. Relationship of *vacA* genotypes of *Helicobacter pylori* to *cagA* status, cytotoxin production, and clinical outcome. Helicobacter **3**:241–253.