Helicobacter pylori Diversity and Gastric Cancer Risk

Timothy L. Cover

Department of Medicine and Department of Pathology, Microbiology and Immunology; Vanderbilt University School of Medicine, Nashville, Tennessee, USA; Veterans Affairs Tennessee Valley Healthcare System, Nashville, Tennessee, USA

ABSTRACT

Gastric cancer is a leading cause of cancer-related death worldwide. Helicobacter pylori infection is the strongest known risk factor for this malignancy. An important goal is to identify H. pylori-infected persons at high risk for gastric cancer, so that these individuals can be targeted for therapeutic intervention. H. pylori exhibits a high level of intraspecies genetic diversity, and over the past two decades, many studies have endeavored to identify strain-specific features of H. pylori that are linked to development of gastric cancer. One of the most prominent differences among H. pylori strains is the presence or absence of a 40-kb chromosomal region known as the cag pathogenicity island (PAI). Current evidence suggests that the risk of gastric cancer is very low among persons harboring H. pylori strains that lack the cag PAI. Among persons harboring strains that contain the cag PAI, the risk of gastric cancer is shaped by a complex interplay among multiple strain-specific bacterial factors as well as host factors. This review discusses the strain-specific properties of H. pylori that correlate with increased gastric cancer risk, focusing in particular on secreted proteins and surface-exposed proteins, and describes evidence from cell culture and animal models linking these factors to gastric cancer pathogenesis. Strain-specific features of H. pylori that may account for geographic variation in gastric cancer incidence are also discussed.

About 2 million new cancer cases each year worldwide are attributable to infections (1). Hepatitis viruses, papillomavirus, and Helicobacter pylori are responsible for most of these malignancies (arising in the liver, cervix, and stomach, respectively). Since H. pylori is the only bacterium known to be a common cause of cancer in humans, the relationship between H. pylori and gastric cancer is of particular interest.

A large body of evidence links H. pylori to two types of stomach cancer—gastric adenocarcinoma and gastric lymphoma. This review focuses on gastric adenocarcinoma, the most common type of stomach cancer. Epidemiological studies have shown that the risk of gastric cancer is higher in H. pylori-infected persons than in H. pylori-negative persons and that H. pylori infection precedes the development of gastric cancer (2–4). H. pylori is associated with adenocarcinoma of the distal (noncardia) stomach but not cancer of the proximal stomach. Experimental orogastric infection of Mongolian gerbils with H. pylori can result in the development of gastric cancer (5), which provides further evidence of a causative role. Consequently, the International Agency for Research on Cancer (World Health Organization) classifies H. pylori as a group I carcinogen (4), a category that includes well-known carcinogens such as tobacco smoke and asbestos.

H. pylori colonizes the stomach and elicits a gastric mucosal inflammatory response termed “gastritis” in both humans and experimentally infected animals. Once established in the human stomach, H. pylori and gastric inflammation can persist for many decades in the absence of antimicrobial treatment. Longitudinal studies indicate that gastritis is one of the first detectable changes in a stepwise pathway of histologic abnormalities that can ultimately culminate in gastric cancer: inflammation, gastric atrophy (loss of specialized cell types such as parietal cells and chief cells), intestinal metaplasia (presence of intestinal-type epithelium in the stomach), and dysplasia (6, 7). The development of gastric cancer in the setting of H. pylori infection is thought to be a long-term consequence of many alterations, including chronic inflammation (which contributes to the pathogenesis of many types of malignancy) (8), DNA damage, activation of gastric stem cells, changes in cell proliferation and apoptosis, changes in epithelial differentiation and polarity, degradation of tumor suppressors, and impaired gastric acidification, leading to bacterial overgrowth with species not found in the normal acidic stomach (6, 7).

EPIDEMIOLOGY OF GASTRIC CANCER

The incidence of gastric cancer varies markedly throughout the world, and it occurs about twice as commonly in males than females (3). The highest incidence rates are currently observed in East Asia (about 60 cases per 100,000 males in Japan and Korea) (3). In all parts of the world, H. pylori is the strongest known risk factor for gastric cancer (3, 4). Regions of the world with a low prevalence of H. pylori infection tend to have a relatively low incidence of gastric cancer, but geographic variation in gastric cancer rates cannot be explained entirely by variations in H. pylori prevalence. For example, populations in many parts of Africa and India have a high prevalence of H. pylori infection but a relatively low incidence of gastric cancer (3).

Although H. pylori is the strongest known risk factor for gastric cancer, most H. pylori-infected persons tolerate the presence of this organism over an entire lifetime without any adverse effects, and some persons may even derive health benefits from H. pylori (9). An important goal is to define the factors that determine whether gastric cancer will develop, so that the subset of H. pylori-infected persons with the highest risk of gastric cancer can be identified and targeted for therapeutic interventions.

Published 26 January 2016

Citation


Copyright

© 2016 Cover. This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license, which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to timothy.l.cover@vanderbilt.edu.
TABLE 1 Strain-specific H. pylori features that correlate with gastric cancer risk

<table>
<thead>
<tr>
<th>Gene or region</th>
<th>Encoded protein(s)</th>
<th>Feature of gene in H. pylori strains*</th>
<th>Higher gastric cancer risk</th>
<th>Lower gastric cancer risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>cag PAI</td>
<td>CagA and T4SS</td>
<td>Present</td>
<td>Absent</td>
<td></td>
</tr>
<tr>
<td>cagA</td>
<td>Effector protein</td>
<td>More EPIYA motifs*</td>
<td>Fewer EPIYA motifs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EPIYA-D motif*</td>
<td>Lack of EPIYA-D motif</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EPIYA-B motif*</td>
<td>EPIYT-B motif</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High levels of CagA production*</td>
<td>Lower levels of CagA production</td>
<td></td>
</tr>
<tr>
<td>vacA</td>
<td>Secreted toxin</td>
<td>s1, l, m1 forms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>babA</td>
<td>OMP</td>
<td>In frame</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subE</td>
<td>OMP</td>
<td>Present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>homB</td>
<td>OMP</td>
<td>In frame</td>
<td></td>
<td></td>
</tr>
<tr>
<td>oipA (hopH)</td>
<td>OMP</td>
<td>Present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hopQ</td>
<td>OMP</td>
<td>Type I form</td>
<td>Absence of type I form</td>
<td></td>
</tr>
<tr>
<td>dupA</td>
<td>VirB4 homolog</td>
<td>Absent</td>
<td>Present</td>
<td></td>
</tr>
</tbody>
</table>

* cagA-positive strains are associated with a higher risk of gastric cancer than are cagA-negative strains. Among cagA-positive strains, those producing CagA with the features indicated with an asterisk are associated with a higher risk of gastric cancer.

H. PYLORI STRAIN-SPECIFIC PROPERTIES ASSOCIATED WITH GASTRIC CANCER

H. pylori is characterized by a high level of intraspecies genetic diversity (10, 11). Diversity in the nucleotide sequences of individual genes is attributable to a high mutation rate, as well as a high rate of intraspecies recombination (12, 13). Strains from unrelated persons not only differ in the sequences of individual genes but also exhibit differences in gene content and chromosomal organization. The core genome of H. pylori consists of about 1,100 genes (present in all H. pylori strains), and each strain typically contains several hundred additional genes that are not universally present.

Over the past two decades, numerous studies have analyzed H. pylori isolates from persons with different disease states in an effort to identify strain-specific features that correlate with the presence of gastric cancer or premalignant histologic lesions. This review focuses on strain-specific variations in secreted proteins or surface-exposed proteins that correlate with increased gastric cancer risk (Table 1) and describes the actions of these factors in cell culture and animal models that link these proteins to gastric cancer pathogenesis.

The cag pathogenicity island. One of the most striking variations among H. pylori strains from unrelated persons is the presence or absence of a chromosomal region known as the cag pathogenicity island (PAI). Individual strains may contain an intact cag PAI (about 40 kb), a cag PAI that has undergone chromosomal rearrangements, or an incomplete cag PAI that lacks one or more genes (14). The cag PAI encodes an antigenic effector protein (CagA) and contains about 18 genes required for the entry of CagA into host cells through a type IV secretion system (T4SS)-mediated process (15–17). In addition to its role in translocating CagA into host cells, the type IV secretion system is required for H. pylori-induced upregulation of proinflammatory cytokine secretion by gastric epithelial cells (18).

Upon entry into host cells, CagA undergoes tyrosine phosphorylation by Src and Abl family kinases at sites known as EPIYA motifs (16, 17, 19). CagA in either its phosphorylated form or its nonphosphorylated form can interact with at least 10 host cell components, resulting in a complex assortment of cellular alterations (16, 17). These include changes in epithelial cell shape and polarity (20, 21), disruption of apical epithelial junctional complexes (22), changes in epithelial differentiation (resulting in an invasive phenotype, resembling an epithelial to mesenchymal transition) (23), enhanced proliferation and inhibited apoptosis (24, 25), activation of gastric stem cells (26), and degradation of the p53 tumor suppressor (27). When experimentally introduced into the stomachs of Mongolian gerbils, H. pylori strains containing cagA and an intact cag PAI promote the development of premalignant changes and gastric cancer, whereas cagA mutant strains or mutant strains with a defective cag T4SS fail to cause gastric cancer (28–31). Similarly, strains containing cagA and an intact cag PAI cause more severe gastric inflammation in gerbils than do cagA mutant strains or mutant strains with a defective cag T4SS (28–31). Transgenic mice expressing CagA spontaneously develop gastric epithelial hyperplasia, adenocarcinoma of the stomach and small intestine, and hematologic malignancies (32). Therefore, CagA has been designated a bacterial oncoprotein (17).

A large body of evidence indicates that the risk of gastric cancer or premalignant lesions is higher in persons infected with cagA-positive H. pylori strains than in persons infected with cagA-negative strains (33–35). The increased risk of gastric cancer observed with cagA-positive strains (which often contain the entire cag PAI) is attributed to the cellular effects of CagA (described above), combined with an enhanced gastric mucosal inflammatory response (36).

Most studies linking CagA-producing strains to increased gastric risk have analyzed H. pylori isolates to determine whether the cagA gene is present or absent or have assessed the presence of anti-CagA serum antibody responses as an indication of infection with CagA-producing strains. The risk of gastric cancer among persons infected with CagA-producing strains can be further stratified by analyzing CagA amino acid sequence variations that influence the protein’s activity. For example, the level of CagA tyrosine phosphorylation in host cells is influenced by the number of EPIYA motifs within the protein that can undergo phosphorylation (Fig. 1). CagA proteins containing higher numbers of EPIYA motifs exhibit enhanced binding to intracellular targets and enhanced activity within host cells (37–39), and strains harboring higher numbers of EPIYA motifs (Fig. 1) are associated with greater gastric cancer risk (40, 41). The intracellular activity of CagA is also influenced by amino acid sequence variations within the tyrosine phosphorylation motifs. For example, strains...
containing EPIYT-B motifs are less frequently associated with gastric cancer than are strains containing EPIYA-B motifs (42).

Among *H. pylori* strains that produce CagA, there is variation in the levels of CagA production. This variation is attributable to nucleotide sequence variation downstream of the *cagA* transcriptional start site, within the 5' untranslated region of the *cagA* transcript (Fig. 1) (43, 44). Two recent studies showed that strains producing high levels of CagA are linked to an increased risk of premalignant lesions compared to strains producing lower levels of CagA (43, 44).

Vacuolating toxin. *H. pylori* secretes a protein known as VacA through an autotransporter or type V secretion pathway (45–47). VacA was originally identified based on its capacity to cause vacuolation of epithelial cells (48) but is now known to have a much broader range of activities (45–47). The amino acid sequence, structure, and cellular effects of VacA are unrelated to those of any other known bacterial toxins (49). Most VacA-induced cellular alterations are attributable to its capacity for pore formation in cell membranes (45–47).

All *H. pylori* strains contain a *vacA* gene, and nearly all secrete a VacA protein, but there is considerable variation among strains in VacA-mediated effects on host cells. This is due to differences among strains in the levels of VacA produced or secreted (50), as well as amino acid sequence variation among VacA proteins. Several families of vacA alleles have been described based on sequence differences in regions of *vacA* designated the s, i, and m regions (Fig. 1) (51–53). Type s1/i1/m1 forms of VacA are more active in assays of cell-vacuolating activity than are forms of VacA designated s2, i2, or m2, and type s2 forms of VacA lack detectable activity in this assay (51, 52, 54, 55). The differences in activity of type s1 and s2 forms of VacA are attributable to impaired channel-forming properties of the type s2 protein (54).

Thus far, most studies of VacA cellular effects have been conducted with type s1/i1/m1 forms of VacA. Studies with this form of the protein indicate that it enters host cells through a clathrin-independent process and can ultimately localize in endosomes as well as mitochondria (56–58). Consequences of VacA intoxication in epithelial cells include altered endosomal function (59), changes in mitochondrial membrane permeability (56), stimulation of autophagy (60), reactive oxygen species accumulation (61), and sometimes cell death (45–47). VacA inhibits the acid-secretory capacity of parietal cells (62) and causes functional alterations in a variety of immune cells. Inhibitory effects of VacA on T cells have been studied in the most detail (63, 64), but B cells, eosinophils, mast cells, and dendritic cells are also targeted (45–47).

Strains containing *vacA* alleles classified as s1, i1, or m1 (encoding the more active forms of VacA) are associated with a higher risk of gastric cancer or premalignant lesions (such as intestinal metaplasia) than strains classified as s2, i2, or m2 (34, 41, 52, 65). In a mouse model, *H. pylori* strains producing s1/i1 forms of VacA...
induced more severe and extensive intestinal metaplasia and inflammation in the stomach than strains producing s2/i2 toxins (65). The increased risk of gastric cancer associated with strains producing more active forms of VacA may be a consequence of several actions, including the capacity of VacA to stimulate gastric epithelial cell injury, alter parietal cell function and gastric acidification, and interfere with immune cell function.

**Outer membrane proteins and DupA.** The *H. pylori* genome contains about 60 genes that are predicted to encode outer membrane proteins (OMPs). While many OMP-encoding genes are conserved among all *H. pylori* strains, others may be present or absent in individual strains, and production of OMPs can be regulated by slipped-strand mispairing within polynucleotide repeat regions. Two of the most extensively studied OMPs, BabA and SabA, function as adhesins that mediate *H. pylori* binding to gastric epithelial cells. BabA binds to the fucosylated Lewis b histo-blood group antigen on host cells, and SabA binds to the sialyl dimeric Lewis x glycosphingolipid (66, 67). In addition to its role in mediating *H. pylori* adherence to epithelial cells, SabA can function as a sialic acid-dependent hemagglutinin and has a role in nonopsonic activation of neutrophils (68).

Several studies have reported that infection with *H. pylori* strains containing in-frame babA or sabA genes is associated with an increased risk of gastric cancer, premalignant changes, or enhanced gastric inflammation, compared to infection with strains that lack these genes or that harbor out-of-frame genes (69–72). When challenged with a BabA-producing *H. pylori* strain, transgenic mice expressing the BabA receptor developed more severe gastritis, atrophy, and anti-parietal cell antibodies than infected wild-type animals (73). These results indicate that BabA-mediated adherence modulates the outcome of infection.

Three other OMPs (HomB, HopQ and HopH [OipA]) have been linked to gastric cancer. HomA and HomB are two closely related *H. pylori* OMPs; individual strains may contain one or both of the corresponding genes (74). Similarly, two forms of HopQ (designated type I and type II) have been described; individual strains may contain one or both of the corresponding genes (75). Strains containing homB or type I hopQ or in-frame hopH (oipA) alleles have been associated with an increased risk of gastric cancer, compared to strains lacking these features (72, 76–78). The presence of an in-frame hopH (oipA) allele contributes to development of gastric cancer in Mongolian gerbils and has been linked to increased gastric inflammation in mice (29, 79). In comparison to a wild-type strain, a homB mutant strain exhibited diminished adherence to gastric epithelial cells and reduced capacity to stimulate production of the proinflammatory cytokine interleukin 8 (IL-8) (74). HopQ is required for maximal activity of the *cag* T4SS (80). Cell culture experiments suggest that HopH (OipA) can contribute to *H. pylori* adherence (81) and also may have a role in regulating expression of proinflammatory cytokines and other processes in gastric epithelial cells (82). HopH (OipA) does not contribute to stimulation of cytokine expression unless the *cag* PAI is also present (81, 83).

Finally, a gene designated *dupA* (for “duodenal-ulcer-promoting gene”), located within a nonconserved region of the *H. pylori* chromosome known as the plasticity region, is reported to be a marker of gastric cancer risk. Strains containing this gene have been associated with a reduced risk of gastric atrophy and gastric cancer compared to strains lacking this gene (84, 85). There is variation among strains in the length of the *dupA* ORF, and this may also influence the risk of gastric cancer (86, 87). The *dupA* gene exhibits weak sequence similarity to *virB* components of type IV secretion systems, but at present it is not known whether *dupA* has a similar activity in *H. pylori*. DupA is reported to have a role in stimulating IL-8 production in gastric epithelial cells and promoting *H. pylori* survival at low pH (84), and strains containing active forms of the *dupA* gene induce proinflammatory cytokine production in mononuclear cells (86).

Associations of specific OMPs or *dupA* with gastric cancer have been detected less consistently than associations between the *cag* PAI or *vacA* and gastric cancer. Nevertheless, a substantial body of experimental evidence indicates that several OMPs and *dupA* influence *H. pylori*-host interactions, and as discussed below, there is an association between the presence of several of these genes and the presence of the *cag* PAI and type I *vacA*. These OMPs and DupA may contribute to gastric cancer pathogenesis by augmenting *H. pylori* adherence, enhancing the activity of the *cag* T4SS, influencing *H. pylori*-induced signaling in host cells, or stimulating proinflammatory immune responses.

**RELATIONSHIPS AMONG STRAIN-SPECIFIC DETERMINANTS OF GASTRIC CANCER RISK**

Intraspecies genetic recombination occurs commonly in *H. pylori*, and the species is considered to have a recombinational population structure (12, 13). Therefore, most polymorphisms are distributed randomly among individual strains (12). Interestingly, the strain-specific features associated with gastric cancer (Table 1) tend to be distributed nonrandomly, even though these genes are localized at unlinked sites in the *H. pylori* chromosome. For example, the majority of *cag* PAI-positive strains contain type I *vacA*, *babA*, *homB*, type I *hopQ*, and in-frame *oipA* alleles, and the major portion of *cag* PAI-negative strains contain type II *vacA*, type II *hopQ*, and out-of-frame *oipA* alleles and lack *babA* and *homB* (51, 69, 74, 75, 81, 83, 88, 89).

The nonrandom distribution of strain-specific features associated with gastric cancer (Table 1) is presumably attributable to selective forces that favor certain combinations. Recent studies in cell culture models have revealed functional interactions among these proteins that may account for the observed associations. For example, both CagA and VacA are used by *H. pylori* to acquire nutrients such as iron from host cells (90), and this process may require a balanced activity of the two proteins. In addition, the effects of CagA on gastric epithelial cells are attenuated by the presence of VacA, and the actions of VacA are attenuated by the presence of CagA (91–94). VacA stimulates degradation of CagA through autophagy (61), which may account for some of the observed antagonism. Functional interactions between OMPs and the T4SS also have been reported (80, 95). For example, BabA and HopQ can potentiate *cag* T4SS-dependent phenotypes (80, 95). In addition, *cag*-positive strains stimulate expression of a host cell gene required for synthesis of sialyl Lewis x (the receptor for the SabA adhesin) (96). By inducing higher levels of gastric inflammation, *cag*-positive strains can indirectly stimulate increased SabA-mediated attachment of the bacteria to epithelial cells.

Several studies have analyzed the correlation between multiple strain-specific features and the development of gastric cancer. These studies indicate that the risk of gastric cancer is highest in persons infected with strains harboring multiple constituents listed in Table 1 (e.g., *cagA*, type I *vacA* and *babA*) (34, 69). Strains harboring few or none of these features are less frequently
associated with gastric cancer. Collectively, these studies suggest that there is a spectrum of strains, ranging from those associated with a very high incidence of gastric cancer to those associated with a very low risk of gastric cancer.

**GEOGRAPHIC VARIATION IN GASTRIC CANCER INCIDENCE**

The incidence of gastric cancer varies markedly throughout the world, for reasons that are not well understood. Since the incidence of gastric cancer is particularly high in East Asia (3), efforts have been directed toward comparing *H. pylori* isolates from East Asia with isolates from other parts of the world. One of the most striking observations is that several of the strain-specific features linked to high gastric cancer risk (including the *cagA PAI*, type s1 *vacA* and *babA*) are present in nearly all East Asian *H. pylori* isolates (88, 97, 98). Conversely, *cagA*-negative strains containing type s2 *vacA* alleles and lacking *babA* are commonly found in the United States and Western Europe but are rarely isolated in East Asia (51, 88, 89). The predominance of strains harboring *cagA*, type s1 *vacA*, and other strain-specific markers linked to gastric cancer (Table 1) may be one of the factors contributing to a high rate of gastric cancer in East Asia.

Most *H. pylori* isolates from East Asia constitute a distinct group based on multilocus sequence typing of housekeeping genes (99), and genes under positive selection (including *cagA PAI*, type s1 *vacA* and *babA*) are highly divergent in East Asian strains compared to strains isolated elsewhere in the world (100). The distinctive properties of *CagA* in East Asian strains have been studied in the most detail. Specifically, a CagA tyrosine phosphorylation motif (EPIYA-D) found exclusively in East Asian strains is associated with a higher level of CagA tyrosine phosphorylation within cells and greater cellular effects than are seen with non-Asian forms of CagA (37, 101, 102). Correspondingly, strains producing CagA proteins with EPIYA-D motifs have been associated with a higher risk of gastric cancer than strains producing other forms of CagA (103). There has been relatively little effort to analyze the functions of other proteins besides CagA in East Asian strains compared to non-Asian strains, but East Asian forms of AlpA and AlpB outer membrane proteins are reported to have different effects on signaling in gastric epithelial cells than forms of AlpA/AlpB found elsewhere in the world (104). Thus, specialized properties of CagA and other constituents in East Asian *H. pylori* strains may contribute to the high incidence of gastric cancer in that part of the world. Variations in host genetics and environmental factors are also likely to contribute to geographic differences in gastric cancer incidence.

**RELATIONSHIP BETWEEN BACTERIAL AND HOST RISK FACTORS FOR GASTRIC CANCER**

Multiple host-related factors are known to be determinants of gastric cancer risk (105). In the context of *H. pylori* infection, polymorphisms in genes involved in cytokine production have been studied in the most detail (6, 105, 106). An association between gastric cancer risk and polymorphisms linked to IL-1β production is relevant because this cytokine not only contributes to gastric inflammation but also regulates gastric acid secretion. Several studies have shown that the contribution of specific bacterial factors to gastric cancer risk is augmented in persons who have specific genetic risk factors. For example, the risk of gastric cancer is particularly high in persons harboring certain polymorphisms in genes encoding cytokines (IL-1β, tumor necrosis factor [TNF], and IL-10) who are infected with *H. pylori* strains containing type s1 *vacA* and *cagA* (34, 106, 107). These studies indicate that the risk of gastric cancer is determined by both bacterial and host factors.

The composition of the human diet is another factor that influences gastric cancer risk. For example, a high-salt diet and a diet low in fruits and vegetables have been associated with increased gastric cancer risk (3, 108). A relationship between composition of the diet and strain-specific *H. pylori* risk factors for gastric cancer also has been observed in experiments with animal models of gastric cancer. Specifically, Mongolian gerbils infected with a *cagA PAI*-positive strain and fed high-salt or low-iron diets had an increased incidence of gastric cancer compared to infected animals fed a regular diet, whereas the high-salt and low-iron diets did not confer an increased risk of gastric cancer in animals infected with a *cagA* mutant strain or uninfected animals (30, 31). The mechanisms by which dietary composition influences gastric cancer risk are not yet well understood. Changes in *H. pylori* gene transcription in response to the composition of the diet may be one mechanism. For example, *H. pylori* produces increased levels of CagA when exposed to high-salt conditions (109), and low-iron conditions stimulate enhanced activity of the *cag* T4SS (30). These results in animal models suggest that certain dietary risk factors for gastric cancer are relevant mainly in persons who are infected with *cagPAI*-positive strains.

Multilocus sequence typing of housekeeping genes in *H. pylori* isolates from human populations throughout the world has allowed the identification of groups of strains with distinct geographic distributions (99). The observed patterns of geographic diversity suggest that *H. pylori* has been present in humans for at least 100,000 years (110), that *H. pylori* accompanied humans out of Africa in multiple waves of migration beginning about 60,000 years ago (111), and that *H. pylori* strains subsequently diversified in relative isolation in various parts of the world (111). Evolutionary theory posits that prolonged association of pathogenic organisms with hosts should lead to a progressive loss of virulence. Since *H. pylori* has been associated with humans for at least 100,000 years, one might anticipate a gradual reduction in the capacity of these bacteria to cause disease.

The development of a commensal or symbiotic relationship between *H. pylori* and humans is presumably dependent on co-evolution of the bacteria and human hosts over a prolonged time period. In many parts of Africa, a high proportion of humans carry *H. pylori*, but the rate of gastric cancer is relatively low (a phenomenon known as the African enigma) (3). Similarly, a recent study in the country of Colombia revealed that humans of African origin, when infected with *H. pylori* strains of African origin, had relatively benign gastric pathology with little evidence of progression to gastric cancer (112). The low incidence of disease in these populations illustrates a relationship between *H. pylori* and humans reflecting coevolution of both species over a prolonged period of time. Introduction of *H. pylori* into a noncoevolved human population may result in a less favorable outcome. In support of this hypothesis, mismatch between the geographic origin of *H. pylori* strains and the geographic ancestry of human hosts has been associated with more severe gastric pathology and development of premalignant gastric lesions (112). These findings bolster the view that gastric cancer risk is influenced by both bacterial and host factors and suggest that disruptions in coevolved bacterial–human relationships may contribute to elevated gastric cancer risk.
Comparison of H. pylori strains isolated from patients with gastric cancer to isolates from patients with non-malignant gastric histologies has led to the identification of multiple strain-specific constituents that contribute to gastric cancer pathogenesis (Table 1). Most previous studies have analyzed H. pylori strains isolated from single gastric biopsy specimens obtained at the time when gastric cancer was diagnosed. In future studies, more robust sampling approaches will allow analysis of H. pylori strains at earlier time points (prior to the development of gastric cancer) and will allow the detection of multiple strains of H. pylori within individual stomachs. Inhibition of T4SS-mediated phenomena has been observed in vitro when cell lines are coinfected with two different H. pylori strains (113), and similarly, diminished severity of disease has been observed in human subjects coinfected with cagA-positive and cagA-negative strains (114). Therefore, further studies of the functional consequences of coinfection with multiple strains of H. pylori are warranted.

Thus far, only a few candidate strain-specific H. pylori genes have been evaluated to detect possible links to gastric cancer, and often the analysis has been limited to a determination of whether a gene is present or absent. In future studies, it will be important to use a more comprehensive approach for analyzing H. pylori genetic variation to permit analysis of a larger number of strain-specific bacterial factors. In addition to determining whether genes are present or absent, important insights will be gained by investigating whether the encoded proteins are produced, analyzing levels of gene transcription or protein production, and analyzing variations in gene sequences that are linked to variations in protein function. It will also be important to investigate further the functions of the strain-specific H. pylori features listed in Table 1 as well as newly identified strain-specific constituents in cell culture and animal models and thereby elucidate the mechanisms by which these factors contribute to gastric cancer pathogenesis.

Recent studies have revealed functional interactions among several bacterial factors linked to gastric cancer pathogenesis (91–94). Developing an improved understanding of these functional interactions and the effects of such interactions on gastric cancer pathogenesis is also an important goal. For example, since VacA can inhibit the actions of CagA in cell culture models, it will be important to determine whether imbalances in VacA and CagA production can influence gastric cancer risk.

Further investigation of the geographic variation in gastric cancer incidence is also warranted. In particular, there has been relatively little effort thus far to evaluate whether the variations in amino acid sequences of H. pylori proteins in strains from different geographic regions are accompanied by alterations in protein function and whether this sequence variation influences the development of gastric cancer. It will also be important to investigate further the coevolved relationships between H. pylori and humans that minimize the risk of gastric cancer, as well as the molecular basis by which disruption in these relationships leads to increased cancer risk.

In summary, a large body of evidence indicates that there is a correlation between the risk of gastric cancer and strain-specific features of H. pylori strains, and studies in cell culture systems and animal models provide mechanistic support for the observed correlations. In future studies, it should be possible to define more clearly the role of strain-specific H. pylori constituents in gastric carcinogenesis. A better understanding of this topic may lead to the development of improved methods to identify H. pylori-infected persons at high risk for development of gastric cancer, so they can be targeted for therapeutic interventions.

ACKNOWLEDGMENTS
This work was supported by NIH AI039657, CA116087, and the Department of Veterans Affairs (2I01BX000627).

FUNDING INFORMATION
HHS | National Institutes of Health (NIH) provided funding to Timothy L Cover under grant numbers AI039657 and CA116087. U.S. Department of Veterans Affairs (VA) provided funding to Timothy L Cover under grant number 2I01BX000627.

REFERENCES


