**Helicobacter pylori** virulence and the diversity of gastric cancer in Asia

Lam Tung Nguyen, Tomohisa Uchida, Kazunari Murakami, Toshio Fujioka and Masatsugu Moriyama

1Department of Molecular Pathology, Faculty of Medicine, Oita University, Oita 879-5593, Japan  
2Department of Gastroenterology, Faculty of Medicine, Oita University, Oita 879-5593, Japan  
3Department of Human Environmental and Social Medicine, Faculty of Medicine, Oita University, Oita 879-5593, Japan

Infection with cagPAI positive strains of Helicobacter pylori is recognized as being associated with an increased risk of gastric cancer. This article reviews the current knowledge on the structures and pathological functions of cagPAI and the CagA protein, focusing especially on the molecular mechanism through which CagA may be involved in gastric carcinogenesis. The possible link between the geographical distribution of cagPAI and cagA variations and gastric cancer diversity in Asia is also discussed.

**Introduction**

Despite a worldwide trend of decline over past decades, gastric cancer remains a significant health problem in Asia (Parkin, 2001; Parkin et al., 2005). The association between gastric cancer and Helicobacter pylori has been well-established (Forman et al., 1991; Uemura et al., 2001) and an estimated 77% of the world’s non-cardia cancer is attributable to *H. pylori* infection (Parkin, 2006). In spite of a general consensus that this relationship exists, there is still controversy as to how *H. pylori* might contribute to the large geographical variations in the gastric cancer incidence rate within Asia, which has long been known as the ‘Asian enigma’ or ‘Asian paradox’ (Miwa et al., 2002; Singh & Ghoshal, 2006).

From multiple epidemiological research studies, it is now clear that infection with CagA-producing strains of *H. pylori* further increases the risk of gastric cancer (Brenner et al., 2004; Huang et al., 2003). Recent studies have shed new light on the mechanism through which CagA intrudes into the host cell cytosol, dysregulates various cellular signal transduction pathways and thus may play an important role in gastric carcinogenesis (Backert et al., 2000; Hatakeyama, 2004, 2006; Odenbreit et al., 2000). Meanwhile, molecular epidemiological data have gradually revealed more details of the geographical distribution of *H. pylori* genetic variations (Falush et al., 2003; Ghose et al., 2002; Kersulyte et al., 2000; Yamaoka et al., 2002). In this review, we summarize new insights into the pathological activities of CagA and its possible role in gastric oncogenesis. The hypothetical relationship between cagPAI and CagA polymorphisms and gastric cancer diversity in Asia will also be discussed in the light of these new discoveries.

**cag pathogenicity island**

cagPAI is an approximately 40 kb cluster of genes that has been acquired through horizontal transmission from an unknown extraneous source and integrated into the *H. pylori* chromosome relatively recently in evolution (Censini et al., 1996; Covacci et al., 1999). cagPAI is divided into two segments, the upstream cag II and the downstream cag I, consisting of about 14 and 16 genes, respectively (Censini et al., 1996; Tomb et al., 1997). In many strains, an insertion sequence called IS605 is interposed between the two segments. Detailed analysis has revealed that cagPAI harbours seven genes homologous with virB4, virB7, virB8, virB9, virB10, virB11 and virD4, which encode component proteins of the type IV secretion system (TFSS) in *Agrobacterium tumefaciens* (Backert & Meyer, 2006; Censini et al., 1996) (Table 1).

It is well-known that cagPAI has at least three important pathological functions. First, it contains the cagA, a gene responsible for producing CagA protein, which is believed to have oncogenic potential (Hatakeyama, 2004, 2006). Second, cagPAI encodes TFSS, a syringe-like structure specialized in the transfer of bacterial components such as CagA protein and peptidoglycan into host cells (Backert et al., 2000; Odenbreit et al., 2000; Viala et al., 2004). Third, cagPAI is implicated in the release of various inflammatory cytokines from the host cells including interleukin 8 (IL-8) (Censini et al., 1996; Yamaoka et al., 1997). By systematic mutagenesis, Fischer et al. (2001) demonstrated that 17 and 14 out of 27 examined genes in cagPAI are indispensable for CagA translocation and full induction of IL-8, respectively (Table 1). These findings indicate the importance of an intact cagPAI for the assembly of a TFSS that is fully functional for CagA delivery and stimulation of IL-8 secretion. cagPAI is usually subject to internal disruptions.
designated EPIYA motifs, located at the C terminus of the protein (Hatakeyama, 2004, 2006; Higashi et al., 2002a). Functionally, EPIYA motifs are the binding targets of many host cell proteins, especially for Src homology 2 (SH2) domain-containing tyrosine phosphatase (SHP-2) (Higashi et al., 2002b; Higashi et al., 2005; Selbach et al., 2002; Stein et al., 2000, 2002; Tsutsumi et al., 2003). Four different EPIYA motifs (EPIYA-A, EPIYA-B, EPIYA-C and EPIYA-D) have been defined based on the distinctive amino acid sequences flanking each of them. From the alignment of these EPIYA motifs, two major types of CagA protein have been identified. The Western type, which represents the CagA of \( H. pylori \) strains prevalent in Europe, America, Australia and Africa, contains EPIYA-A and EPIYA-B, followed by up to five repeated sequences of EPIYA-C. The Eastern type, which represents the CagA of \( H. pylori \) strains circulating in Japan, Korea and China, also possesses EPIYA-A and EPIYA-B, but the third motif is EPIYA-D instead of EPIYA-C (Hatakeyama, 2004, 2006; Higashi et al., 2002a). On rare occasions, various types of untypical combinations among the four kinds of EPIYA motif may be observed including: AB, AAB, ABB, AABC, AABCC, AAABC, AAC, ABABC, ABBC, ABCCBC, ABABC, AC, ACC, ACCC, AABD, AAABD, AAD, ABD, ABDAB, ABDBD, ABABD, AD, BD, BBD, BABD etc. (Fig. 1).

A recent study has uncovered another important domain of CagA termed the CagA multimerization (CM) motif

### Table 1. Contribution of the constituent genes of cagPAI to CagA translocation and IL-8 induction

<table>
<thead>
<tr>
<th>Name of ( cag ) gene according to:</th>
<th>Involved in:*</th>
<th>CagA translocation</th>
<th>IL-8 induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomb et al. (1997)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP547 cagA</td>
<td>++</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>ND cagB</td>
<td>++</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>HP546 cagC</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>HP545 cagD</td>
<td>++</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP544 cagE (virB4)</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>HP543 cagF</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HP542 cagG</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP541 cagH</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP540 cagI</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP539 cagL</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP538 cagN</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP537 cagM</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ND cagO</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>HP536 cagP</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HP535 cagQ</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ND cagR</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>HP534 cagS</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HP532 cagT (virB7)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP531 cagU</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP530 cagV</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP529 cagW (virB8)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP528 cagX (virB9)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP527 cagY (virB10)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP526 cagZ</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP525 cagZ (virB11)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP524 cagβ (virD4)</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HP523 cagβ</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP522 cagδ</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HP521 cagγ</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HP520 cagζ</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*Fischer et al. (2001).

(Kauser et al., 2004), and even partial deletions within this locus may have an impact on bacterial virulence, and thus also on clinical outcome.

With regard to the role of \( cagPAI \) in \( H. pylori \)-host cell interaction, microarray transcriptional analysis has revealed that \( cagPAI \) is essential for altered expression of 92% of the genes in epithelial cells infected with \( H. pylori \) (El-Etr et al., 2004). In contrast, the \( cagPAI \) negative strain evokes very little transcriptional response in comparison with uninfected cells (Guillemin et al., 2002). These findings once again emphasize the importance of \( cagPAI \) in the pathogenesis of \( H. pylori \)-associated diseases.

### CagA protein

CagA is structurally characterized by the presence of repeated 5 amino acid sequences (Glu-Pro-Ile-Tyr-Ala), designated EPIYA motifs, located at the C terminus of the protein (Hatakeyama, 2004, 2006; Higashi et al., 2002a). Functionally, EPIYA motifs are the binding targets of many host cell proteins, especially for Src homology 2 (SH2) domain-containing tyrosine phosphatase (SHP-2) (Higashi et al., 2002b; Higashi et al., 2005; Selbach et al., 2002; Stein et al., 2000, 2002; Tsutsumi et al., 2003). Four different EPIYA motifs (EPIYA-A, EPIYA-B, EPIYA-C and EPIYA-D) have been defined based on the distinctive amino acid sequences flanking each of them. From the alignment of these EPIYA motifs, two major types of CagA protein have been identified. The Western type, which represents the CagA of \( H. pylori \) strains prevalent in Europe, America, Australia and Africa, contains EPIYA-A and EPIYA-B, followed by up to five repeated sequences of EPIYA-C. The Eastern type, which represents the CagA of \( H. pylori \) strains circulating in Japan, Korea and China, also possesses EPIYA-A and EPIYA-B, but the third motif is EPIYA-D instead of EPIYA-C (Hatakeyama, 2004, 2006; Higashi et al., 2002a). On rare occasions, various types of untypical combinations among the four kinds of EPIYA motif may be observed including: AB, AAB, ABB, AABC, AABCC, AAABC, AAC, ABABC, ABBC, ABCCBC, ABABC, AC, ACC, ACCC, AABD, AAABD, AAD, ABD, ABDAB, ABDBD, ABABD, AD, BD, BBD, BABD etc. (Fig. 1).

A recent study has uncovered another important domain of CagA termed the CagA multimerization (CM) motif

![Western CagA](Fig. 1. Classification of CagA protein. CagA is categorized based on the alignment of different EPIYA motifs. Western CagA has EPIYA-A and EPIYA-B, followed by 1–5 EPIYA-C. Eastern CagA contains EPIYA-A, EPIYA-B and 1–2 EPIYA-D. Rarely, various types of untypical arrangements among the EPIYA motifs can be observed.)

![Eastern CagA](Fig. 1. Classification of CagA protein. CagA is categorized based on the alignment of different EPIYA motifs. Western CagA has EPIYA-A and EPIYA-B, followed by 1–5 EPIYA-C. Eastern CagA contains EPIYA-A, EPIYA-B and 1–2 EPIYA-D. Rarely, various types of untypical arrangements among the EPIYA motifs can be observed.)

![Other variations](Fig. 1. Classification of CagA protein. CagA is categorized based on the alignment of different EPIYA motifs. Western CagA has EPIYA-A and EPIYA-B, followed by 1–5 EPIYA-C. Eastern CagA contains EPIYA-A, EPIYA-B and 1–2 EPIYA-D. Rarely, various types of untypical arrangements among the EPIYA motifs can be observed.)
which is a prerequisite for subsequent formation of the CagA-SHP-2 complex, and dysregulation of the β-catenin signal (Kurashima et al., 2008; Ren et al., 2006).

**Biological activity of CagA protein inside the host cell**

**Membrane-associated activities of CagA**

Upon delivery into epithelial cells via the TFSS, CagA anchors to the inner surface of cell membrane (Higashi et al., 2005). At these sites, it recruits and modifies the distribution of zona occludens-1 (ZO-1) and junctional adhesion molecule (JAM), which are normally located at cell junctions. By interaction with ZO-1 and JAM, CagA alters the structure and function of the apical-junctional complex (Amieva et al., 2003; Saadat et al., 2007). Apical-junctional dysfunction may result in loss of control over cytoskeletal architecture, cell polarity, proliferation and differentiation, which is characteristic of oncogenic transformation (Bagnoli et al., 2005; Jamora & Fuchs, 2002).

CagA is then phosphorylated by several Src family kinases (SKFs) on the tyrosine residues of all the EPIYA motifs albeit with much weaker affectivity at EPIYA-A and EPIYA-B than at EPIYA-C and EPIYA-D (Higashi et al., 2002a; Naito et al., 2006; Selbach et al., 2002; Stein et al., 2002). This is a very important process since nearly 80% of host cell transcriptional changes elicited by CagA are phosphorylation dependent (El-Etr et al., 2004). Moreover, the number of EPIYA motifs is proportional to the phosphorylation level and biological activity of CagA protein, and accordingly is associated with gastric cancer (Argent et al., 2004; Naito et al., 2006).

**Phosphorylation-dependent activities of CagA**

After phosphorylation, the phosphorylated CagA binds to SHP-2 via the EPIYA-C or EPIYA-D motif (Higashi et al., 2002a, b). Both the N-SH2 and C-SH2 domains of SHP-2 are required for stable complex formation with CagA, signifying the importance of the preceding CagA multimerization mediated by the CM motif (Higashi et al., 2002a; Ren et al., 2006). SHP-2–CagA interaction correlates not only with the number of phosphorylated sites but also with CagA types. In this regard, Eastern CagA has strong affinity for SHP-2 due to a 6 amino acid sequence overlying EPIYA-D (Y-A-T-I-D-F), which is identical to the consensus ligand binding motif for the SH-2 domain of SHP-2 (De Souza et al., 2002). In contrast, the corresponding sequence surrounding EPIYA-C in Western CagA (Y-A-T-I-D-D) differs from the consensus motif in 1 terminal amino acid (D versus F/W) and thus exhibits a weaker binding ability (Higashi et al., 2002a).

Activated by phosphorylated CagA, SHP-2 inhibits focal adhesion kinase (FAK), an enzyme that modulates cell adhesion, migration and survival (Parsons, 2003; Tsutsumi et al., 2006). A decline in FAK activity triggers a cytoskeletal rearrangement phenomenon called the ‘hummingbird phenotype’, which is characterized by the elongation and spreading of the host cells (Backert et al., 2001; Segal et al., 1999; Tsutsumi et al., 2006). CagA-activated SHP-2 also stimulates extracellular signal-regulated kinase (Erk) through both Ras-dependent and -independent pathways (Higashi et al., 2004; Neel et al., 2003). Corresponding to its stronger binding affinity for SHP-2, Eastern CagA is reported to exhibit a greater ability to induce the hummingbird phenotype and to have more influence on Erk activity as well as cell growth than the Western type (Fu et al., 2007; Higashi et al., 2002a). Therefore, Eastern CagA is considered to be pathologically more virulent than its Western counterpart (Hatakeyama, 2004, 2006).

Although the majority of phosphorylated CagA binds to SHP-2, a smaller proportion interacts with the C-terminal Src kinase (CSK) (Tsutsumi et al., 2003). Activation of CSK by interaction with CagA in turn inhibits SFKs. Therefore, together with direct suppression of SFKs by phosphorylated CagA (Selbach et al., 2003), the indirect inhibitory pathway via CSK forms dual feedback loops that down-regulate SFK phosphorylation activity. As a result, these two negative feedback loops may mitigate the phosphorylation-related cytotoxicity of CagA. Additionally, CagA-mediated inhibition of SFKs favours the dephosphorylation of other host cell proteins such as vinculin, ezrin and cortactin, resulting in impaired adhesion of host cells to extracellular matrix and disrupted cellular actin organization, which partially contributes to the hummingbird phenotype (Moese et al., 2007; Selbach et al., 2003, 2004) (Fig. 2).

**Phosphorylation-independent activities of CagA**

Cag A is able to dysregulate several other signalling cascades in a phosphorylation-independent manner. CagA interacts with hepatocyte growth factor (HGF) receptor c-Met, thereby stimulating cell growth, motility and invasiveness (Churin et al., 2003). CagA binds to growth factor receptor-bound protein 2 (Grb-2) and activates the Ras/Erk pathway via son of sevenless (SOS) leading to cell scattering and proliferation. The interaction between CagA and Grb-2 relies strictly on the presence of EPIYA motifs but does not require tyrosine phosphorylation (Mimuro et al., 2002). Notably, both c-Met and Grb-2 have been implicated in malignant transformation (Cheng et al., 1998; Peruzzi & Bottaro, 2006).

Recently, Murata-Kamiya et al. (2007) found that by interacting with E-cadherin, CagA disturbs the E-cadherin/β-catenin complex, leading to cytoplasmic and nuclear accumulation of β-catenin. This is consistent with a previous study in which β-catenin nuclear translocation in gastric mucosa infected with cagPAI positive H. pylori was shown to increase in a CagA-dependent manner.
Deregulated β-catenin signalling is believed to drive transcriptional upregulation of genes implicated in oncogenic transformation (Fodde & Brabletz, 2007; Hoppler & Kavanagh, 2007). Furthermore, strong nuclear and cytoplasmic expression of β-catenin in various types of gastric neoplasm has also been reported (Tsukashita et al., 2003). These findings raise the possibility that CagA causes aberrant expression of β-catenin, which in turn may be involved in gastric tumorigenesis.

Formerly, CagA was thought to have no role in the induction of pro-inflammatory responses such as IL-8 production based on observations that cagA-knockout strains still retain a considerable ability to induce IL-8 production in gastric epithelial cells (Censini et al., 1996; Fischer et al., 2001). However, this general concept was challenged by Brandt et al. (2005) who demonstrated that CagA indeed induced IL-8 release via the Ras/Raf/Mek/Erk/NF-κB signalling pathway independently of SHP-2 and c-Met. These findings prompted a not-yet-proven assumption that CagA–Grb-2 interaction, which was reported to activate the Ras/Erk pathway, might be an upstream event of CagA-induced NF-κB stimulation. The explicit involvement of CagA in the NF-κB/IL-8-activating pathway signifies its direct role in the chronic mucosal inflammation that precedes, and is thought to be linked with, cancer (Balkwill & Mantovani, 2001; Coussens & Werb, 2002; Karin, 2006).

Recently, Matsumoto et al. (2007) discovered that activation-induced cytidine deaminase (AID) was upregulated via the NF-κB-dependent pathway in gastric epithelial cells infected with cagPAI positive H. pylori. Aberrant expression of AID, a DNA- and RNA-editing enzyme, induces cumulative mutations of somatic genes such as p53 inside host cells. Although not directly implicated, CagA

(Franco et al., 2005). Deregulated β-catenin signalling is believed to drive transcriptional upregulation of genes implicated in oncogenic transformation (Fodde & Brabletz, 2007; Hoppler & Kavanagh, 2007). Furthermore, strong nuclear and cytoplasmic expression of β-catenin in various types of gastric neoplasm has also been reported (Tsukashita et al., 2003). These findings raise the possibility that CagA causes aberrant expression of β-catenin, which in turn may be involved in gastric tumorigenesis.

Formerly, CagA was thought to have no role in the induction of pro-inflammatory responses such as IL-8 production based on observations that cagA-knockout strains still retain a considerable ability to induce IL-8 production in gastric epithelial cells (Censini et al., 1996; Fischer et al., 2001). However, this general concept was challenged by Brandt et al. (2005) who demonstrated that CagA indeed induced IL-8 release via the Ras/Raf/Mek/Erk/NF-κB signalling pathway independently of SHP-2 and c-Met. These findings prompted a not-yet-proven assumption that CagA–Grb-2 interaction, which was reported to activate the Ras/Erk pathway, might be an upstream event of CagA-induced NF-κB stimulation. The explicit involvement of CagA in the NF-κB/IL-8-activating pathway signifies its direct role in the chronic mucosal inflammation that precedes, and is thought to be linked with, cancer (Balkwill & Mantovani, 2001; Coussens & Werb, 2002; Karin, 2006).

Recently, Matsumoto et al. (2007) discovered that activation-induced cytidine deaminase (AID) was upregulated via the NF-κB-dependent pathway in gastric epithelial cells infected with cagPAI positive H. pylori. Aberrant expression of AID, a DNA- and RNA-editing enzyme, induces cumulative mutations of somatic genes such as p53 inside host cells. Although not directly implicated, CagA
may have some relationship to that mechanism for a number of reasons. Firstly, CagA is involved in the majority of cagPAI-dependent activities (El-Etr et al., 2004). Secondly, CagA itself is able to activate NF-kB (Brandt et al., 2005). Finally, a cagE knockout mutation strain, which failed to translocate CagA into epithelial cells, reportedly had no effect on p53 mutation, as was observed in the cagPAI negative strain (Matsumoto et al., 2007).

**cagPAI, CagA and gastric cancer diversity in Asia**

Gastric cancer is still one of the most common malignant diseases. Overall, the estimated worldwide incidence of gastric cancer in 2002 was 934,000 cases, ranking this cancer fourth behind cancers of the lung, breast and colorectum. Gastric cancer mortality accounted for about 10.4% of all cancer deaths, with 700,000 cases annually, the second highest rate after only lung cancer (Parkin et al., 2005). About 56% of newly diagnosed gastric cancers arise in East Asia, of which 42% are reported from China and 12% from Japan [GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide, version 2.0, CancerBase no 5 (http://www-dep.iarc.fr/); International Agency for Research on Cancer]. In Asia, gastric cancer distribution is geographically characterized by a wide variation, the difference in the incidence rate between regions of high and low prevalence reaching up to 40-fold (GLOBOCAN 2002). Based on the age-standardized incidence rate of gastric cancer, Asian countries can be categorized as high risk (Japan, Korea, China), intermediate risk (Vietnam, Singapore) or low risk (Thailand, India, Bangladesh etc.) (Fig. 3).

One intriguing phenomenon is the ‘Asian paradox’, whereby the *H. pylori* infection rate among the general population is very high in areas with little risk of gastric cancer such as South-Central Asia (India, Pakistan, Bangladesh) and South-Eastern Asia (Thailand, Indonesia), but lower in Eastern Asia, where the gastric cancer rate is the highest in the world (Japan, Korea) (Parkin, 2004, 2006). Thus, despite the undeniable role of *H. pylori* in gastric cancer, this paradox suggests that any explanation should not focus only on bacterial prevalence. As a multifactorial model of gastric malignant transformation is currently accepted, other aspects such as bacterial virulence, host genetic susceptibility and diet should also be taken into consideration. In this article we analyse the possible links between cagPAI disruption, the CagA genotype and the diversity of gastric cancer in Asia.

CagA intactness is thought to be an important determinant of *H. pylori* virulence and clinical outcomes. Nilsson et al. (2003) argued that infection with *H. pylori* containing a complete cagPAI conferred a fivefold increased risk of developing severe gastroduodenal diseases. An intact cagPAI has been detected in 86% of strains isolated from gastric cancer versus 7% of those from non-ulcer dyspepsia (Ali et al., 2005). Meanwhile, precancerous lesions have been frequently found in the stomachs colonized by strains with the full cagPAI but rarely in those harbouring strains that lack parts of cagPAI (Ali et al., 2005). A study in Mongolian gerbils has shown that intact cagPAI facilitates *H. pylori* colonization in the body (corpus) of the stomach, resulting in atrophic corpus-predominant gastritis, a well-known precancerous condition (Rieder et al., 2005). Consequently, enhanced carcinoma risk has been observed in 71% of animals inoculated with an intact cagPAI strain compared to 9% of those infected with a partially deleted cagPAI strain (Rieder et al., 2005). In humans, intact cagPAI has been detected in 95% of isolates from gastric ulcers and gastric cancer, both characterized by corpus-predominant gastritis, versus only 6.9% of those from duodenal ulcer, which is characterized by antrum-predominant gastritis and reduced risk of gastric cancer. The authors concluded that in comparison with partially deleted cagPAI strains, those with intact cagPAI would...
increase the risk of gastric carcinoma 10-fold in infected subjects (Ali et al., 2005). In Japan, where the gastric cancer incidence rate is high, cagPAI is very well-conserved with nearly all _H. pylori_ strains carrying an intact island (Ikenoue et al., 2001). Conversely, in India, where the gastric cancer incidence rate is about 10 times lower, 63–87 % of strains have a cagPAI that is partially deleted (Ali et al., 2005; Kauser et al., 2004). Surprisingly, the deletions usually occur in the promoter region of cagA, which would lead to abolition of gene transcription as reported by Maeda et al. (1999). Thus, although Indian _H. pylori_ strains often accommodate cagA (Ali et al., 2005; Chattopadhay et al., 2002; Chaudhuri et al., 2003; Datta et al., 2003), many of them might not produce CagA protein and could be regarded as functionally cagA negative. Similarly, disrupted cagPAI and deletion of the promoter region of cagA are also very common in _H. pylori_ isolates from Turkey (Salih et al., 2007), possibly corresponding to the low gastric cancer rate there. Taken together, it can be speculated that the differences in the prevalence of intact cagPAI strains contribute partly to gastric cancer diversity in Asia.

As discussed above, Eastern CagA is considered more virulent and thus associated with a more severe clinical outcome than its Western counterpart. Azuma et al. (2004) demonstrated that in both the antrum and body of the stomach, the grades of inflammation, activity and mucosal atrophy were significantly higher in chronic gastritis patients infected with Eastern cagA positive strains than in those infected with Western cagA positive strains. As expected, the relationship between Eastern CagA, and gastric cancer was also found to be stronger. In Okinawa, where both genotypes of cagA are present, 85 % of _H. pylori_ isolates from gastric cancer patients carried Eastern cagA whereas only 15 % carried the Western genotype (Satomi et al., 2006). The same authors noted that Eastern cagA strains were considerably more prevalent in a region with a high gastric cancer rate (Fukui prefecture) than in a region with a lower rate (Okinawa prefecture) (Azuma et al., 2004; Satomi et al., 2006). A similar result was also obtained by Vilaichone et al. (2004) who reported that 85 % of _H. pylori_ isolates from ethnic Chinese people living in Thailand harboured Eastern cagA, whereas only 18 % of those isolates from indigenous Thais carried such a gene. Accordingly, gastric cancer was found to be common among Chinese Thais but rare among the indigenous Thais (Vilaichone et al., 2004). From these studies, it is tempting to speculate that the prevalence of Eastern cagA strains may partly correlate with gastric cancer risk. A growing body of molecular epidemiological data shows that in Korea, China and Japan (except Okinawa) almost all _H. pylori_ strains have Eastern cagA (Azuma et al., 2004; Choi et al., 2007; Kanada et al., 2008; Uchida et al., 2007; Yamaoka et al., 2002; Zhou et al., 2004a, b), corresponding well to the high incidence rate of gastric cancer in this region. In contrast, the fact that nearly all _H. pylori_ isolates from India, Pakistan and Turkey bear Western cagA is likely to reflect the reduced risk of gastric cancer in these populations (Mukhopadhyay et al., 2000; Saribasak et al., 2004; Yamaoka et al., 2002). Therefore, the striking difference in the geographical distribution of cagA genotypes may partially contribute to the wide variation of gastric cancer among the countries, and to a larger extent, among Asian countries.

It should be noted that other virulence factors of _H. pylori_ may also account for some of the differences in gastric cancer incidence. For example, variations in the signal region(s) (which may be type s1 or s2), and the middle region (m) (which may be type m1 or m2) of the vaculating cytotoxin-encoding gene (vacA) have been shown to correlate with _H. pylori_ cytotoxin activity (Atherton et al., 1995, 1997). In this respect, vacA s1/m1 is more toxic than vacA s1/m2 and vacA s2/m2, and has been found to be strongly associated with gastric cancer (Kidd et al., 1999; Miehlke et al., 2000). In East Asia (Japan, Korea), nearly all _H. pylori_ strains carry vacA s1/m1 while in countries with a low incidence of gastric cancer, such as Thailand and Turkey, this virulence genotype is carried in less than 60 % of strains (Chomvarin et al., 2008; Erzin et al., 2006; Yamaoka et al., 2002). Recently, a new _H. pylori_ vaculating cytotoxin determinant has been described and termed the intermediate region (i). Two i regions have been identified, i1 and i2, and the i1 region has been shown to be associated with gastric cancer (Basso et al., 2008; Rhead et al., 2007). Interestingly, the great majority of _H. pylori_ isolates in East Asia have vacA i1 (Ogiwara et al., 2008). Therefore, the combination of many virulence factors including an intact cagPAI, Eastern cagA and vacA s1/m1/i1 in most of the _H. pylori_ strains isolated from East Asia might be an important contributor to the high incidence of gastric cancer in this region.

**Concluding remarks**

New data suggest that previous studies using a single serological method to detect _H. pylori_ infection might have considerably underestimated the relationship between this bacterium and gastric cancer. Ekstrom et al. (2001) demonstrated that the adjusted odds ratio (OR) for non-cardia gastric cancer among _H. pylori_-positive subjects increased from 2.2 to 21 when immunoblot analysis with anti-CagA antibody was added to the conventional ELISA method. Another population-based study, in which both ELISA and Western blotting were used, has shown that the OR for non-cardia gastric cancer rose from 3.7 to 18.3 for any _H. pylori_ infection and from 5.7 to 28.4 for cagA positive _H. pylori_ infection (Brenner et al., 2004). These results suggest that the correlation between _H. pylori_ infection and gastric cancer might be changed if additional factors were to be integrated into a large population-based study. Accordingly, the _H. pylori_-gastric cancer relationship should be reinvestigated in the context of cagPAI rearrangement and cagA polymorphisms because serological positivity for _H. pylori_ alone does not fully reflect the
true risk of gastric cancer. With more comprehensive epidemiological data on the trinity of host genetic susceptibility, lifestyle and bacterial virulence factors, the Asian paradox may prove to be not as enigmatic as it seems.

References


