Review

Correspondence
Yoshio Yamaoka
yyamaoka@bcm.tmc.edu

Roles of the plasticity regions of *Helicobacter pylori* in gastroduodenal pathogenesis

Yoshio Yamaoka

Department of Medicine – Gastroenterology, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX 77030, USA

Putative virulence genes of *Helicobacter pylori* are generally classified into three categories: strain-specific genes, phase-variable genes and genes with variable structures/genotypes. Among these, there has recently been considerable interest in strain-specific genes found outside of the *cag* pathogenicity island, especially genes in the plasticity regions. Nearly half of the strain-specific genes of *H. pylori* are located in the plasticity regions in strains 26695 and J99. Strain HPAG1, however, seems to lack a typical plasticity region; instead it has 43 HPAG1-specific genes which are either undetectable or incompletely represented in the genomes of strains 26695 and J99. Recent studies showed that certain genes or combination of genes in this region may play important roles in the pathogenesis of *H. pylori*-associated gastroduodenal diseases. Most previous studies have focused on the plasticity region in strain J99 (*jhp0914–jhp0961*) and the *jhp0947* gene and the duodenal ulcer promoting (*dupA*) gene are good candidate markers for gastroduodenal diseases although there are some paradoxical findings. The *jhp0947* gene is reported to be associated with an increased risk of both duodenal ulcers and gastric cancers, whereas the *dupA* gene, which encompasses *jhp0917* and *jhp0918*, is reported to be associated with an increased risk of duodenal ulcers and protection against gastric cancers. In addition, recent studies showed that approximately 10–30% of clinical isolates possess a 16.3 kb type IV secretion apparatus (*tfs3*) in the plasticity region. Studies on the plasticity region have only just begun, and further investigation is necessary to elucidate the roles of genes in this region in gastroduodenal pathogenesis.

Background

Helicobacter pylori is a well-recognized pathogen that chronically infects more than 50% of the world's population. *H. pylori* plays an important role in the development of peptic ulcers, gastric adenocarcinoma and gastric mucosa-associated lymphoid tissue lymphoma (MALToma). The infection remains latent in the majority of infected patients, with only approximately 20% of infected individuals developing severe disease. It is unclear what determines the outcome of an infection; however, it is thought to involve an interplay between the virulence of the infecting strain, host genetics and environmental factors. Experience with other bacterial pathogens suggests that *H. pylori*-specific factors may exist that influence the pathogenicity of *H. pylori*.

Many putative virulence genes of *H. pylori* have been reported to determine clinical outcome, and these are generally classified into three categories (Table 1). The first category contains strain-specific genes, which are present in only some *H. pylori* strains. Among this group, the best studied is the *cag* pathogenicity island (PAI), which encodes a bacterial type IV secretory apparatus (Censini *et al.*, 1996). The *cag* PAI contains approximately 30 genes,

including the *cagA* gene in the 3' end of the island, which injects CagA and possibly other bacterial proteins into host cells (Asahi *et al.*, 2000; Backert *et al.*, 2000; Odenbreit *et al.*, 2000; Segal *et al.*, 1999; Stein *et al.*, 2000; Viala *et al.*, 2004). Strains that possess the *cag* PAI/*cagA* are statistically more likely to be associated with peptic ulcer and gastric cancer than are strains lacking the PAI/*cagA* (Blaser *et al.*, 1995; van Doorn *et al.*, 1998). To date, the complete genomes of three *H. pylori* strains (26695, J99 and HPAG1) have been sequenced (Alm *et al.*, 1999; Oh *et al.*, 2006; Tomb *et al.*, 1997). These are all *cag* PAI-positive strains; however, many other strain-specific genes lie outside of the *cag* PAI (Table 2) and nearly half of the strain-specific genes of *H. pylori* are located in the plasticity region. Genes in this region are also good candidates for *H. pylori* virulence factors, as described in detail in this review.

The second group is the phase-variable genes such that the gene status can be changed during growth or in different conditions. Based on the comparison of the three completed genomes of *H. pylori*, six genes encoding outer-membrane proteins (*oipA*, *sabA*, *sabB*, *babB*, *babC* and *hopZ*) are thought to undergo phase variation (Alm *et al.*, 1999; Oh *et al.*, 2006; Tomb *et al.*, 1997). The

Table 1. Virulence factors of *H. pylori*

Category	DNA/protein status	Major virulence factors
Strain-specific genes	Gene positive or negative	<i>cag</i> pathogenicity island (PAI) Plasticity regions
Phase-variable genes	Gene positive, but some produce functional proteins and others not	<i>oipA</i> , <i>sabA</i> , <i>babA</i> (regulated by slipped strand mispairing)
Genes with different structures/genotypes	Gene positive, but function and/or production levels of proteins different between strains	<i>vacA</i> genotypes (combination of s1/s2 and m1/m2) <i>cagA</i> repeat region (East Asian type and Western type) <i>alpAB</i> (East Asian type and Western type)

functional status is regulated by a slipped strand mispairing mechanism and is mediated by the number of CT dinucleotide repeats in the 5' region of the genes. Among these, *OipA* (outer inflammatory protein) and *SabA* (sialic acid binding adhesin) are reported to be associated with peptic ulcers and gastric cancer (Yamaoka *et al.*, 2002, 2006). Recently, another putative virulence factor, the outer-membrane protein *BabA* (blood group antigen binding adhesin), was shown to be regulated by the slipped strand mispairing mechanism in some clinical isolates (Colbeck *et al.*, 2006; Hennig *et al.*, 2006) and has been reported to be associated with peptic ulcers and gastric cancer (Fujimoto *et al.*, 2007; Gerhard *et al.*, 1999; Yamaoka *et al.*, 2006).

The last group is the genes with variable structures/genotypes depending on the strain. For example, specific *vacA* genotypes containing different mosaic combinations of signal regions and middle region allelic types have been associated with different clinical outcomes (Atherton *et al.*, 1995). In addition, the structure of many genes differs between Western strains and East Asian strains, and the structural differences in some genes (e.g. 3' repeat region of the *cagA*, *alpAB* genes) are reported to influence virulence (Lu *et al.*, 2007; Yamaoka *et al.*, 1998b, 1999a).

Among these virulence factors, there has recently been considerable interest in strain-specific genes outside of the *cag* PAI: the plasticity regions. This review focuses on the plasticity regions and describes the current knowledge of the regions in relation to clinical outcomes.

Table 2. Strain-specific genes

Category	Strain		
	26695	J99	HPAG1
R-M system	6	6	5
Transposases	14	1	0
<i>cag</i> genes	0	0	3
Other	22	15	7
Hypothetical	90	37	28
Total	132	59	43

Plasticity regions of *H. pylori*

Comparison of the complete genome sequences of two *H. pylori* strains (26695 and J99) revealed several regions whose G+C content was lower than that of the rest of the *H. pylori* genome (35 % compared with 39 %), suggestive of horizontal DNA transfer from other species (Alm *et al.*, 1999; Tomb *et al.*, 1997). One such region is the *cag* PAI and the other has been termed the plasticity region based on the variability in gene content between different isolates. Nearly half of the strain-specific genes of *H. pylori* are located in the plasticity region (Alm & Trust, 1999).

In *H. pylori* strain J99, the plasticity region has been reported to range from *jhp0914* to *jhp0961* (Alm *et al.*, 1999). This region is continuous in strain J99; however, only 6 and 10 of the 48 open reading frames (ORFs) present in the plasticity region of strain J99 were found in strains 26695 and HPAG1, respectively. Two DNA microarray studies have analysed the gene content of 15 (Salama *et al.*, 2000) and 56 (Gressmann *et al.*, 2005) isolates of *H. pylori*. Table 3 shows the prevalence of *jhp0914*–*jhp0961* in 72 strains, including strains 26695, J99 and HPAG1, and confirms that the ORFs in the plasticity region really do display diversity, except for *jhp0915*, which was present in all strains studied.

Although the ORFs of the major part of the plasticity region encode putative proteins with unknown functions, some have been found to share similarity with genes encoding functional proteins. For example, in the plasticity region of strain J99, *jhp0917* and *jhp0918* are homologous to *vir* factors; *jhp0919*, *jhp0920* and *jhp0931* are homologous to DNA topoisomerase I (*topA*) involved in DNA replication; *jhp0921*, *jhp0922* and *jhp0923* are homologous to DNA transformation competence ComB8 (*jhp0921* and *jhp0922*) or ComB9 (*jhp0923*) homologues; *jhp0928* is homologous to a methylase gene; *jhp0935* is homologous to partitioning protein A; and *jhp0941* and *jhp0951* are homologous to integrase/recombinase genes (*xerCD* family) involved in DNA restriction, modification, recombination and repair systems.

Strain HPAG1 seems to lack a typical plasticity region; instead it has 43 HPAG1-specific genes which are either undetectable or incompletely represented in the genomes of strains 26695 and J99 and are scattered throughout the

Table 3. Prevalence of strain-specific genes in the plasticity region of strain J99: analyses for 72 strains

Gene number	Strains hybridized/positive (%)
<i>jhp0914</i>	17
<i>jhp0915</i>	100
<i>jhp0916</i>	43
<i>jhp0917</i>	39
<i>jhp0918</i>	36
<i>jhp0919</i>	36
<i>jhp0920</i>	38
<i>jhp0921</i>	36
<i>jhp0922</i>	42
<i>jhp0923</i>	40
<i>jhp0924</i>	39
<i>jhp0925</i>	19
<i>jhp0926</i>	19
<i>jhp0927</i>	61
<i>jhp0928</i>	58
<i>jhp0929</i>	50
<i>jhp0930</i>	24
<i>jhp0931</i>	47
<i>jhp0932</i>	47
<i>jhp0933</i>	51
<i>jhp0934</i>	47
<i>jhp0935</i>	49
<i>jhp0936</i>	46
<i>jhp0937</i>	22
<i>jhp0938</i>	85
<i>jhp0939</i>	85
<i>jhp0940</i>	22
<i>jhp0941</i>	89
<i>jhp0942</i>	86
<i>jhp0943</i>	42
<i>jhp0944</i>	21
<i>jhp0945</i>	51
<i>jhp0946</i>	49
<i>jhp0947</i>	26
<i>jhp0948</i>	31
<i>jhp0949</i>	24
<i>jhp0950</i>	60
<i>jhp0951</i>	44
<i>jhp0952</i>	88
<i>jhp0953</i>	39
<i>jhp0954</i>	63
<i>jhp0955</i>	94
<i>jhp0956</i>	86
<i>jhp0957</i>	94
<i>jhp0958</i>	72
<i>jhp0959</i>	13
<i>jhp0960</i>	38
<i>jhp0961</i>	44

genome evenly without indication of islands (Table 2). Accordingly, other *H. pylori* strains would be expected to contain their own set of strain-specific genes, and in fact novel candidates for virulence factors have been found experimentally in the plasticity regions of additional *H.*

pylori strains (for example, a novel 16.3 kb segment, *tfs3*; see below) (Kersulyte *et al.*, 2003).

Plasticity regions and clinical outcomes

The strain-specific genes are likely to play a similar role to the *cag* PAI in the pathogenesis of *H. pylori* infection. The highly variable gene content in the plasticity regions may play an important role in this regard, and some loci within the plasticity regions have previously been proposed to serve as markers of clinical outcomes. Most previous studies have focused on the plasticity region in strain J99, probably due to the fact that strain J99 was isolated from a duodenal ulcer patient whereas the other fully sequenced strains, HPAG1 and 26695, were from gastritis patients (Alm *et al.*, 1999; Oh *et al.*, 2006; Tomb *et al.*, 1997).

In an initial systematic study, Occhialini *et al.* (2000) examined the prevalence of 21 ORFs inside the plasticity region of strain J99 (*jhp0914*–*jhp0961*) in a small group of 17 gastric cancer and 26 gastritis patients from Costa Rica and found an increased prevalence of the *jhp0940* and *jhp0947* genes in *H. pylori* from gastric cancer patients. A follow-up study to further examine this relationship was performed with 200 strains of *H. pylori* from Brazil, including strains from duodenal ulcer, gastric cancer and *H. pylori* gastritis patients (Santos *et al.*, 2003). The original hypothesis was only partially confirmed as the new study demonstrated an association between *jhp0947* (but not *jhp0940*) and both duodenal ulcer and gastric cancer. de Jonge *et al.* (2004) also reported that disruption of the *jhp0945*–*0947*–*0949* locus in *H. pylori* strain 1061 significantly decreased its ability to induce interleukin (IL)-12 production in monocyte THP-1 cells, and the requirement for the *jhp0945*–*0947*–*0949* locus for IL-12 induction was subsequently confirmed using four wild-type *H. pylori* strains. These authors also reported that the presence of the *jhp0947* and *jhp0949* genes, but not *jhp0945*, was significantly associated with duodenal ulcer when compared with gastritis in a small group of 26 duodenal ulcer and 19 gastritis patients from the Dutch population. Recently, Rizwan *et al.* (2008) reported that recombinant JHP0940 protein elicited strong and significant levels of TNF- α and IL-8 in human macrophages and enhanced translocation of nuclear factor (NF)- κ B in cultured macrophages. However, they could not find a relationship between the presence of *jhp0940* and clinical outcomes or *cag* PAI status in seven countries studied (India, South Africa, Japan, Costa Rica, Peru, France and Spain). In India, South Africa and France, most strains possessed the *jhp0940* gene, whereas in Spain fewer than 10% of strains possessed the gene irrespective of disease status. As a preliminary study, Salih *et al.* (2007) recently examined the prevalence of *jhp0926*, *jhp0931*, *jhp0933*, *jhp0944* and *jhp0945* in a small group containing 21 gastritis and 14 peptic ulcer patients from Turkey and found that the *jhp0931* gene was more likely to be found in *H. pylori* from patients with peptic ulcer disease than in those with gastritis (14.2% vs 0%).

However, the *jhp0931* gene was not related to clinical outcomes in the original study in the Costa Rican population (Occhialini *et al.*, 2000).

The overall conclusion from the series of studies examining ORFs within the plasticity region of strain J99 is that the *jhp0947* gene currently appears to be the best candidate for a disease marker. The *jhp0947* gene is homologous to *jhp0938* (*hp0990*) and *jhp253* (*hp1333*); however, its function is unknown. The 5' region of *jhp0947* is also homologous to that of *jhp0477* (*hp0528*), which is part of the *cag* PAI (*virB9* homologue) and has been identified as an important structural component of the type IV secretion system (Tanaka *et al.*, 2003). Interestingly, the presence of *jhp0947* was associated with the presence of *cag* PAI in Brazilian strains (Santos *et al.*, 2003), whereas this relationship was not observed in Dutch strains (de Jonge *et al.*, 2004). There are no studies on the *jhp0947* gene in East Asian countries, where the strains are considerably different from those in Western countries. Further studies are necessary to confirm the roles of the *jhp0947* gene product in gastroduodenal pathogenesis.

Lehours *et al.* (2004) reported that the presence of *jhp0950*, which was not included in the 21 ORFs selected by Occhialini *et al.* (2000), is a candidate marker for gastric extranodal marginal zone B cell lymphoma of the mucosa-associated lymphoid tissue (MALT)-type (MZBL) in the French population. Importantly, *jhp0950* seems to be a disease-specific marker for MZBL since the prevalence of *jhp0950* (74 %) was significantly higher in MZBL than in all other diseases examined (49 % for gastritis, 49 % for duodenal ulcer and 39 % for gastric adenocarcinoma). The function of *jhp0950* is unknown; however, a recent study suggests that *jhp0942*–*jhp0944*–*jhp0945*–*jhp0947*–*jhp0949* may be expressed as an operon, since these genes are consecutive and oriented in the same direction (Occhialini *et al.*, 2000). The *jhp0950* gene is also arranged in the same orientation as the *jhp0949* gene and may also form part of the operon that includes the *jhp0949* gene.

Novel type IV secretion system in the plasticity region

Type IV secretion systems are widely distributed in prokaryotes and are structurally complex molecular machines, typically composed of a cell envelope-spanning translocation channel, cytoplasmic ATPases and a pilus. The most studied type IV secretion systems of *H. pylori* are the secretion system for effector proteins such as CagA (the *cag* PAI system), and the DNA import system via natural transformation (the ComB system), which is found in all *H. pylori* strains (Hofreuter *et al.*, 1998, 2001; Smeets & Kusters, 2002; Censini *et al.*, 1996). The *Agrobacterium tumefaciens* VirB/D4 type IV secretion system serves as a prototype system for which detailed structural and functional data are available. The VirB/D4 system consists of 11 VirB proteins (VirB1–VirB11) and the substrate recognition factor VirD4, which are assembled into the

transport apparatus necessary for the delivery of proteins or nucleoprotein complexes into target cells. In *H. pylori*, the *cag* PAI contains seven homologues of the VirB/D4 type IV secretion system, *virB4*, *virB7*, *virB8*, *virB9*, *virB10*, *virB11* and *virD4* (Censini *et al.*, 1996; Selbach *et al.*, 2002; Covacci *et al.*, 1999; Buhrdorf *et al.*, 2003), and the ComB system contains homologues of all *vir* genes, except *virB1*, *virB5*, *virB11* and *virD4* (Hofreuter *et al.*, 2001). Mutation studies suggest that VirB homologues in the *cag* PAI, but not the ComB, system are involved in the induction of proinflammatory cytokine IL-8 and activation of transcription factors such as NF- κ B in gastric epithelial cells (Censini *et al.*, 1996; Fischer *et al.*, 2001; Selbach *et al.*, 2002; Glocker *et al.*, 1998).

Kersulyte *et al.* (2003) identified a novel 16.3 kb type IV secretion apparatus (*tfs3*) in the plasticity region of a strain from a Peruvian patient with gastric cancer (PeCan18B), seven of the 16 ORFs of which were homologues of the *virB/D* operon of *A. tumefaciens*, the third putative type IV secretion system present in *H. pylori*. The clustering of DNA with a lower G+C content in *tfs3* is suggestive of horizontal DNA transfer from other species. Four of these genes are transmembrane pore genes (*virB7*, *virB8*, *virB9* and *virB10*), and three encode cytoplasmic membrane-associated ATPases that move their cognate macromolecule substrates VirB4, VirB11 and VirD4 to and through the pore. Full-length and partial *tfs3* elements were each found in approximately one-fifth of clinical *H. pylori* strains from Spain, Peru, India and Japan. Our subsequent study confirmed that 8 % (Japan), 13 % (Korea) and 33 % (Colombia) of clinical isolates possess complete *tfs3* (Lu *et al.*, 2005a). Strain 26695 contains a truncated *tfs3* element that corresponds to 7369 bp of the 5' region of full-length *tfs3* from strain PeCan18B (18 144 bp). The function of the *tfs3* elements is still unknown, and there is no evidence to support a critical role in transformability, general bacterial conjugation, intracellular entry, survival or mouse colonization (Kersulyte *et al.*, 2003). In addition, there was no correlation between the presence of complete or partial *tfs3* and clinical outcomes (Kersulyte *et al.*, 2003), and the presence of complete *tfs3* was not linked to other *vir* homologue genes (Lu *et al.*, 2005a). Although other type IV secretion systems of *H. pylori* show homology to components of the *tfs3* system, it is not known whether they work independently and are unable to complement each other functionally or not, as shown between the *cag* PAI and ComB system (Hofreuter *et al.*, 2001). Further studies are necessary to elucidate the function of this third type IV secretion system in *H. pylori* and to determine the interplay among the three type IV secretion systems.

Identification of the duodenal ulcer promoting gene (*dupA*) as another *virB4* homologue in the plasticity region

The plasticity region of strain J99 contains two *virB4* homologues (*jhp0917* and *jhp0918*). The *jhp0917* gene

encodes a protein of 475 amino acids, but lacks a region homologous to the C-terminus of VirB4, while the *jhp0918* gene encodes a product of 140 amino acids that is homologous to the missing VirB4 region. Lu *et al.* (2005a) recently found that all of the clinical isolates that they studied (eight each from Japan, Korea and Colombia) contained a 1 bp insertion (C or T) in the 3' region of the *jhp0917* gene (after position 1385 in strain C142; GenBank accession no. AB196363), resulting in a frameshift leading to a continuous gene. Subsequent studies in Brazilian strains reported that 86 out of 89 (97 %) isolates contained a 1 bp insertion after position 1385 (Gomes *et al.*, 2008), confirming that *jhp0917–0918* usually form one continuous gene homologous to the intact *virB4* gene, with strain J99 being a rare exception.

More importantly, Lu *et al.* (2005a) reported that the *jhp0917–0918* gene is a marker for development of duodenal ulcer disease and for protection against gastric adenocarcinoma. They also found that the gene is a marker for a protective effect against atrophy and intestinal metaplasia. Consequently, they have designated the *jhp0917–jhp0918* gene the duodenal ulcer promoting (*dupA*) gene. They examined 500 *H. pylori* isolates, 160 from Japan, 175 from Korea and 165 from Colombia, and found that the prevalence of the *dupA* gene was significantly greater among strains isolated from duodenal ulcer patients (42 %) than from patients with *H. pylori* gastritis (21 %), with gastric ulcer (27 %) or with gastric cancer (9 %) ($P < 0.001$ for duodenal ulcer vs gastritis or cancer). Importantly, these associations were consistently observed in isolates from both Asian and Western countries. Although a number of putative *H. pylori* virulence genes have been associated with increased risks of a clinical outcome such as peptic ulcer or gastric cancer, none have clearly been linked to one specific *H. pylori*-related disease such as duodenal ulcer, thus *dupA* is the first disease-specific virulence marker in *H. pylori*.

As expected, the prevalence of the *jhp0917* and *jhp0918* genes was closely linked: only 10 of 500 (2 %) strains were *jhp0917*-negative/*jhp0918*-positive while the status of the two genes in the remaining strains was either both positive or both negative (Lu *et al.*, 2005a). Subsequent studies confirmed that most isolates were both positive or both negative: for example, 474 of 482 (98 %) in Brazilian strains (Gomes *et al.*, 2008) and 149 of 157 (95 %) in Iranian strains (Douraghi *et al.*, 2008). This relationship was also observed in North Indian strains; however, in this case only 75 % (122/166) of the strains were either both positive or both negative, with 15 (9 %) *jhp0917*-positive/*jhp0918*-negative and 27 (16 %) *jhp0917*-negative/*jhp0918*-positive (Arachchi *et al.*, 2007). The absence of the *jhp0917* and/or *jhp0918* genes (especially where one is present and the other is absent) is probably due to primer mismatches rather than to gene truncation or genetic decay. We could categorize strains as intact *dupA*-positive only if positive for both *jhp0917* and *jhp0918* genes with a 1 bp insertion after position 1385 corresponding to strain C142.

Possible functions of DupA

The function of DupA is still not fully understood. The *dupA* gene encodes homologues of VirB4 ATPase, which is thought to be involved in DNA uptake/DNA transfer and protein transfer. Bioinformatic analyses show that the N-terminus of DupA (encoded by *jhp0917*; position 3–201) has homology to members of the FtsK/SpoIIIE family, whereas the central region (encoded by *jhp0917* 3'–*jhp0918*; position 203–610) shares homology with the TraG/TraD family (Lu *et al.*, 2005a). The FtsK/SpoIIIE domain contains a putative ATP-binding P-loop motif and is involved in cell division and peptidoglycan synthesis or modification, and has been implicated in intercellular chromosomal DNA transfer. Members of the TraG/TraD family are potential NTP hydrolases that are essential for DNA transfer in bacterial conjugation and are thought to mediate interactions between DNA processing and mating pair formation systems.

In vitro experiments using *dupA*-deleted and -complemented mutants showed that the absence of the *dupA* gene was associated with increased susceptibility to low pH (Lu *et al.*, 2005a). In addition, the presence of the *dupA* gene was associated with increased IL-8 production from the antral gastric mucosa *in vivo* as well as from the gastric epithelial cells *in vitro* (Lu *et al.*, 2005a). The authors further confirmed that the *dupA* gene is involved in the activation of transcription factors that bind to the IL-8 promoter, such as NF- κ B and AP-1. These findings are in agreement with *in vivo* results showing higher neutrophil infiltration and IL-8 production in the antral gastric mucosa of duodenal ulcer patients compared with that in other diseases (Yamaoka *et al.*, 1998a, 1999b). It is possible that the *dupA* gene acts in combination with other *vir* homologues in the plasticity region to form a type IV secretion system similar to the *cag* PAI, which is also involved in IL-8 secretion; however, a search for genes in the plasticity region of strain J99 did not find potential partner genes for *dupA* (*jhp0917–jhp0918*) that were components of a type IV secretion system. Recent preliminary data from my laboratory showed that IL-8 induction from gastric epithelial cells was decreased in *dupA* mutants of some strains compared with their parental strains, but was not reduced in *dupA* mutants of other strains (Y. Yamaoka, unpublished data). Therefore, it is possible that some strains form a novel type IV secretion system involving the *dupA* gene that may have a role in IL-8 induction, whereas others do not. The *dupA* gene might be an ancestral remnant of a type IV secretion system in some strains. In other words, only strains that are both *dupA*-positive and form a novel type IV secretion system might be involved in gastroduodenal diseases. Further studies are required to elucidate the roles of genes surrounding the *dupA* gene.

DupA and clinical outcomes: discrepancies in recent studies

In late 2007 and early 2008, four studies evaluating the relationship between the presence of *dupA* and clinical

outcomes in different geographical areas were published (Arachchi *et al.*, 2007; Argent *et al.*, 2007; Douraghi *et al.*, 2008; Gomes *et al.*, 2008). Detailed distribution of the *dupA* gene in relation to different disease states is summarized well by Douraghi *et al.* (2008). One study examining North Indian strains supported the original findings and reported that the *dupA* gene was more common in strains from patients with duodenal ulcer (38 %; 36/96) than in those from patients with functional dyspepsia (23 %; 16/70) ($P < 0.05$); this study did not examine isolates from patients with gastric cancer (Arachchi *et al.*, 2007). In contrast, a study on Iranian strains reported that the presence of the *dupA* gene was independent of the clinical outcome, including duodenal ulcer, gastric ulcer and gastric cancer; however, histological analyses confirmed that the presence of the *dupA* gene was inversely associated with the presence of the precancerous lesion gastric dyspepsia and with the presence of lymphoid follicles, which represent a relatively common histological feature of chronic gastritis (Douraghi *et al.*, 2008). Therefore, this study provided further support for *dupA* as a protective marker against gastric cancer, in agreement with the original study.

In a study examining strains from Brazilian children and adults, the prevalence of the *dupA* gene was extremely high (92 %; 445/482) irrespective of the nature of the gastro-duodenal diseases, gastritis, duodenal ulcer and gastric cancer (Gomes *et al.*, 2008). Interestingly, the frequency of the *dupA* gene was significantly higher in strains from children than in those from adults ($P = 0.01$). *H. pylori* infection is typically acquired in childhood and persists throughout life unless treated with a combination of anti-acid and antimicrobial therapy, so it is speculated that the *dupA* gene might be lost during long-term infections in which the gastric mucosa gradually develop chronic atrophic gastritis and gastric cancer. Therefore, their results might partially support the original hypothesis that *dupA* is a marker associated with reduced gastric damage that leads to the development of gastric cancer. Familial studies or longitudinal studies will be necessary to confirm the possible loss of *dupA* over time.

Another study examining strains from four countries (Belgium, South Africa, China and the United States) reported no significant association between the presence of the *dupA* gene and peptic ulcer or gastric cancer in any individual population, although the prevalence of these diseases in each population may have been too small to reveal a true effect; for example, there was only one case of gastric cancer in the Chinese population studied (Argent *et al.*, 2007). When the authors combined Belgian and South African populations, the presence of the *dupA* gene was significantly associated with the presence of gastric cancer [70 % (26/37) in gastric cancer vs 44 % (40/92) in gastritis] ($P < 0.01$). There was also a non-significant trend towards an association between the *dupA* gene and duodenal ulcer [60 % (32/53)] ($P = 0.057$) in the combined Belgian and South African population. However, it is not relevant to combine random populations since it is unclear

whether the South African strains were taken from patients of European descent and recent studies confirmed that the genomic structures of some South African strains (e.g. HpAfrica2 type) were relatively different from those of strains from the European population (i.e. HpEurope type) (Falush *et al.*, 2003; Linz *et al.*, 2007).

The relationship between the presence of the *dupA* gene and the *cag* PAI/*cagA* gene is also somewhat controversial. One study involving Indian patients with gastritis or duodenal ulcer reported that the presence of the *dupA* gene was significantly associated with the *cagA*-positive genotype (Arachchi *et al.*, 2007). Interestingly, two studies reported that the presence of *dupA* was associated with the *cagA*-positive genotype in strains isolated from adults with duodenal ulcer, but not in those from patients with other diseases or from children (Douraghi *et al.*, 2008; Gomes *et al.*, 2008); however, these relationships were not observed in the remaining two studies (Argent *et al.*, 2007; Lu *et al.*, 2005a).

Overall, there are distinct geographical variations in the prevalence of the *dupA* gene, and there appears to be an association between *dupA* and duodenal ulcers in some populations but not in others. As Argent *et al.* (2007) reported, the association of *dupA* with duodenal ulcer in only some populations could reflect differences in the definition or diagnosis of ulcers or in the use of drugs that either cause or heal ulcers in these populations. In addition, the discrepancy could be related to the limitation of PCR techniques for detecting the intact *dupA* gene. In some studies, only one set of primer pairs for *jhp0917* and *jhp0918* was used (Douraghi *et al.*, 2008; Lu *et al.*, 2005a): use of multiple primer pairs is recommended for detection of the *dupA* gene in future studies. None of the previous reports considered the 1 bp insertion after position 1385 as a criterion for the presence of the *dupA* gene, although the insertion is very common (97–100 %) (Gomes *et al.*, 2008; Lu *et al.*, 2005a). More importantly, Gomes *et al.* (2008) reported frameshift mutations in 14/86 (16 %) *dupA*-positive sequenced samples; a single adenine insertion after position 1426 of *dupA* or at position 2998 of the *jhp0917*–*jhp0918* gene of the J99 strain that created a premature stop codon and may have considerable effects on protein expression or function. In their study, they counted the truncated samples as *dupA*-positive; however, it is clear that these mutated sequences would not produce intact DupA protein. It is definitely unacceptable to ignore such mutations that occur at an unexpectedly high frequency; however, it is impossible to detect them using a simple PCR method. It might be better to detect intact *dupA* by measuring intact DupA protein using immunoblotting techniques, which has not been reported previously. In addition, as mentioned above, only strains that are *dupA*-positive and form a novel type IV secretion system might be involved in gastroduodenal diseases. If this is true, examining the presence of DupA/*dupA* alone might not be sufficient. Studies on DupA and the role of the plasticity region are still in their early stages, and great progress is expected in the near future.

Conclusions

There are few studies investigating the roles of genes in the plasticity regions in *H. pylori*; however, as described in this review, there are several good candidate markers for gastroduodenal diseases, including the *jhp0947* gene and the *dupA* gene. It has been proposed that a *H. pylori* virulence factor should (1) have a disease or other *in vivo* correlation, (2) be epidemiologically consistent across populations and regions and (3) be biologically plausible and the effect should be reduced or eliminated by gene deletion and be restored by complementation (Lu *et al.*, 2005b). Only criterion (1) is cleared for the *jhp0947* gene, and the studies have been limited to strains from non-Asian countries. Further studies in the Asian population will be necessary to confirm the importance of the *jhp0947* gene for gastroduodenal diseases. Although there are some paradoxical findings, studies to date suggest that DupA clears all three criteria and might be the first disease-specific marker for increased risk of duodenal ulcers and protection from gastric malignancies. In contrast, although a number of putative virulence factors of *H. pylori* have been reported, including the *jhp0947* gene, their presence has typically been associated with an increased risk of both gastric cancer and peptic ulcers. However, there are currently no established methods to detect the intact *dupA* gene or DupA protein and further studies are necessary, first to establish such detection procedures and subsequently to elucidate the role of DupA, the interaction of the *dupA* gene with surrounding genes, and possible formation of a novel type IV secretion system. In addition, most previous studies have focused on the plasticity region in strain J99; however, there are many strain-specific genes in strains 26695 and HPAG1, not only in the plasticity regions but also outside of these regions, that may be involved in gastritis. Novel strain-specific genes are expected to be present in other clinical isolates, and we expect that identification of the roles of such strain-specific genes will give further insight into the pathogenesis of *H. pylori*-induced gastroduodenal diseases. Recently, the high parallel Genome Sequencer 20 System (GS-20) (454 Life Sciences) was introduced and successfully used to sequence *H. pylori* genomes (Oh *et al.*, 2006). Such instrumentation provides an opportunity to rapidly sequence multiple individual *H. pylori* isolates, and will allow more efficient identification of novel strain-specific genes that are related to specific diseases.

Acknowledgements

This material is based upon work supported in part by National Institutes of Health (NIH) grants R01 DK62813.

References

Alm, R. A. & Trust, T. J. (1999). Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. *J Mol Med* 77, 834–846.

- Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C. & other authors (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397, 176–180.
- Arachchi, H. S. J., Kalra, V., Lal, B., Bhatia, V., Baba, C. S., Chakravarthy, S., Rohatgi, S., Sarma, P. M., Mishra, V. & other authors (2007). Prevalence of duodenal ulcer-promoting gene (*dupA*) of *Helicobacter pylori* in patients with duodenal ulcer in North Indian population. *Helicobacter* 12, 591–597.
- Argent, R. H., Burette, A., Miendje Deyi, V. Y. & Atherton, J. C. (2007). 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* 45, 1204–1206.
- Asahi, M., Azuma, T., Ito, S., Ito, Y., Suto, H., Nagai, Y., Tsubokawa, M., Tohyama, Y., Maeda, S. & other authors (2000). *Helicobacter pylori* CagA protein can be tyrosine phosphorylated in gastric epithelial cells. *J Exp Med* 191, 593–602.
- Atherton, J. C., Cao, P., Peek, R. M., Jr, Tummuru, M. K., Blaser, M. J. & Cover, T. L. (1995). Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. *J Biol Chem* 270, 17771–17777.
- Backert, S., Ziska, E., Brinkmann, V., Zimny-Arndt, U., Fauconier, A., Jungblut, P. R., Naumann, M. & Meyer, T. F. (2000). Translocation of the *Helicobacter pylori* CagA protein in gastric epithelial cells by a type IV secretion apparatus. *Cell Microbiol* 2, 155–164.
- Blaser, M. J., Perez-Perez, G. I., Kleanthous, H., Cover, T. L., Peek, R. M., Chyou, P. H., Stemmermann, G. N. & Nomura, A. (1995). Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. *Cancer Res* 55, 2111–2115.
- Buhrdorf, R., Forster, C., Haas, R. & Fischer, W. (2003). Topological analysis of a putative *virB8* homologue essential for the *cag* type IV secretion system in *Helicobacter pylori*. *Int J Med Microbiol* 293, 213–217.
- Censini, S., Lange, C., Xiang, Z., Crabtree, J. E., Ghiara, P., Borodovsky, M., Rappuoli, R. & Covacci, A. (1996). *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci U S A* 93, 14648–14653.
- Colbeck, J. C., Hansen, L. M., Fong, J. M. & Solnick, J. V. (2006). Genotypic profile of the outer membrane proteins BabA and BabB in clinical isolates of *Helicobacter pylori*. *Infect Immun* 74, 4375–4378.
- Covacci, A., Telford, J. L., Del, G. G., Parsonnet, J. & Rappuoli, R. (1999). *Helicobacter pylori* virulence and genetic geography. *Science* 284, 1328–1333.
- de Jonge, R., Kuipers, E. J., Langeveld, S. C., Loffeld, R. J., Stoof, J., van Vliet, A. H. & Kusters, J. G. (2004). The *Helicobacter pylori* plasticity region locus *jhp0947-jhp0949* is associated with duodenal ulcer disease and interleukin-12 production in monocyte cells. *FEMS Immunol Med Microbiol* 41, 161–167.
- Douraghi, M., Mohammadi, M., Oghalaie, A., Abdirad, A., Mohagheghi, M. A., Eshagh Hosseini, M., Zeraati, H., Ghasemi, A., Esmaili, M. & Mohajerani, N. (2008). *dupA* as a risk determinant in *Helicobacter pylori* infection. *J Med Microbiol* 57, 554–562.
- Falush, D., Wirth, T., Linz, B., Pritchard, J. K., Stephens, M., Kidd, M., Blaser, M. J., Graham, D. Y., Vacher, S. & other authors (2003). Traces of human migrations in *Helicobacter pylori* populations. *Science* 299, 1582–1585.
- Fischer, W., Puls, J., Buhrdorf, R., Gebert, B., Odenbreit, S. & Haas, R. (2001). Systematic mutagenesis of the *Helicobacter pylori* *cag* pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Mol Microbiol* 42, 1337–1348.

- Fujimoto, S., Olaniyi, O. O., Arnqvist, A., Wu, J. Y., Odenbreit, S., Haas, R., Graham, D. Y. & Yamaoka, Y. (2007). *Helicobacter pylori* BabA expression, gastric mucosal injury, and clinical outcome. *Clin Gastroenterol Hepatol* 5, 49–58.
- Gerhard, M., Lehn, N., Neumayer, N., Boren, T., Rad, R., Schepp, W., Miehke, S., Classen, M. & Prinz, C. (1999). Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. *Proc Natl Acad Sci U S A* 96, 12778–12783.
- Glocker, E., Lange, C., Covacci, A., Bereswill, S., Kist, M. & Pahl, H. L. (1998). Proteins encoded by the *cag* pathogenicity island of *Helicobacter pylori* are required for NF-kappaB activation. *Infect Immun* 66, 2346–2348.
- Gomes, L. I., Rocha, G. A., Rocha, A. M., Soares, T. F., Oliveira, C. A., Bittencourt, P. F. & Queiroz, D. M. (2008). Lack of association between *Helicobacter pylori* infection with *dupA*-positive strains and gastroduodenal diseases in Brazilian patients. *Int J Med Microbiol* (in press).
- Gressmann, H., Linz, B., Ghai, R., Pleissner, K. P., Schlapbach, R., Yamaoka, Y., Kraft, C., Suerbaum, S., Meyer, T. F. & Achtman, M. (2005). Gain and loss of multiple genes during the evolution of *Helicobacter pylori*. *PLoS Genet* 1, e43.
- Hennig, E. E., Allen, J. M. & Cover, T. L. (2006). Multiple chromosomal loci for the *babA* gene in *Helicobacter pylori*. *Infect Immun* 74, 3046–3051.
- Hofreuter, D., Odenbreit, S., Henke, G. & Haas, R. (1998). Natural competence for DNA transformation in *Helicobacter pylori*: identification and genetic characterization of the *comB* locus. *Mol Microbiol* 28, 1027–1038.
- Hofreuter, D., Odenbreit, S. & Haas, R. (2001). Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Mol Microbiol* 41, 379–391.
- Kersulyte, D., Velapatino, B., Mukhopadhyay, A. K., Cahuayme, L., Bussalleu, A., Combe, J., Gilman, R. H. & Berg, D. E. (2003). Cluster of type IV secretion genes in *Helicobacter pylori*'s plasticity zone. *J Bacteriol* 185, 3764–3772.
- Lehours, P., Dupouy, S., Bergey, B., Ruskoné-Foumestraux, A., Delchier, J. C., Rad, R., Richy, F., Tankovic, J., Zerbib, F. & other authors (2004). Identification of a genetic marker of *Helicobacter pylori* strains involved in gastric extranodal marginal zone B cell lymphoma of the MALT-type. *Gut* 53, 931–937.
- Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F. & other authors (2007). An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* 445, 915–918.
- Lu, H., Hsu, P. I., Graham, D. Y. & Yamaoka, Y. (2005a). Duodenal ulcer promoting gene of *Helicobacter pylori*. *Gastroenterology* 128, 833–848.
- Lu, H., Yamaoka, Y. & Graham, D. Y. (2005b). *Helicobacter pylori* virulence factors: facts and fantasies. *Curr Opin Gastroenterol* 21, 653–659.
- Lu, H., Wu, J. Y., Beswick, E. J., Ohno, T., Odenbreit, S., Haas, R., Reyes, V. E., Kita, M., Graham, D. Y. & Yamaoka, Y. (2007). Functional and intracellular signaling differences associated with the *Helicobacter pylori* AlpAB adhesin from Western and East Asian strains. *J Biol Chem* 282, 6242–6254.
- Occhialini, A., Marais, A., Alm, R., Garcia, F., Sierra, R. & Megraud, F. (2000). Distribution of open reading frames of plasticity region of strain J99 in *Helicobacter pylori* strains isolated from gastric carcinoma and gastritis patients in Costa Rica. *Infect Immun* 68, 6240–6249.
- Odenbreit, S., Puls, J., Sedlmaier, B., Gerland, E., Fischer, W. & Haas, R. (2000). Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science* 287, 1497–1500.
- Oh, J. D., Kling-Backhed, H., Giannakis, M., Xu, J., Fulton, R. S., Fulton, L. A., Cordum, H. S., Wang, C., Elliott, G. & other authors (2006). The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. *Proc Natl Acad Sci U S A* 103, 9999–10004.
- Rizwan, M., Alvi, A. & Ahmed, N. (2008). Novel protein antigen (JHP940) from the genomic plasticity region of *Helicobacter pylori* induces TNF-alpha and interleukin-8 secretion by human macrophages. *J Bacteriol* 190, 1146–1151.
- Salama, N. R., Guillemin, K., McDaniel, T. K., Sherlock, G., Tompkins, L. & Falkow, S. (2000). A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc Natl Acad Sci U S A* 97, 14668–14673.
- Salih, B. A., Abasiyanik, M. F. & Ahmed, N. (2007). A preliminary study on the genetic profile of *cag* pathogenicity-island and other virulent gene loci of *Helicobacter pylori* strains from Turkey. *Infect Genet Evol* 7, 509–512.
- Santos, A., Queiroz, D. M., Menard, A., Marais, A., Rocha, G. A., Oliveira, C. A., Nogueira, A. M., Uzeda, M. & Megraud, F. (2003). New pathogenicity marker found in the plasticity region of the *Helicobacter pylori* genome. *J Clin Microbiol* 41, 1651–1655.
- Segal, E. D., Cha, J., Lo, J., Falkow, S. & Tompkins, L. S. (1999). Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*. *Proc Natl Acad Sci U S A* 96, 14559–14564.
- Selbach, M., Moese, S., Meyer, T. F. & Backert, S. (2002). Functional analysis of the *Helicobacter pylori* *cag* pathogenicity island reveals both VirD4-CagA-dependent and VirD4-CagA-independent mechanisms. *Infect Immun* 70, 665–671.
- Smeets, L. C. & Kusters, J. G. (2002). Natural transformation in *Helicobacter pylori*: DNA transport in an unexpected way. *Trends Microbiol* 10, 159–162.
- Stein, M., Rappuoli, R. & Covacci, A. (2000). Tyrosine phosphorylation of the *Helicobacter pylori* CagA antigen after *cag*-driven host cell translocation. *Proc Natl Acad Sci U S A* 97, 1263–1268.
- Tanaka, J., Suzuki, T., Mimuro, H. & Sasakawa, C. (2003). Structural definition on the surface of *Helicobacter pylori* type IV secretion apparatus. *Cell Microbiol* 5, 395–404.
- Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S. & other authors (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388, 539–547.
- van Doorn, L. J., Figueiredo, C., Sanna, R., Plaisier, A., Schneeberger, P., de Boer, W. & Quint, W. (1998). Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. *Gastroenterology* 115, 58–66.
- Viala, J., Chaput, C., Boneca, I. G., Cardona, A., Girardin, S. E., Moran, A. P., Athman, R., Mémet, S., Huerre, M. R. & other authors (2004). Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* *cag* pathogenicity island. *Nat Immunol* 5, 1166–1174.
- Yamaoka, Y., Kita, M., Kodama, T., Sawai, N., Tanahashi, T., Kashima, K. & Imanishi, J. (1998a). Chemokines in the gastric mucosa in *Helicobacter pylori* infection. *Gut* 42, 609–617.
- Yamaoka, Y., Kodama, T., Kashima, K., Graham, D. Y. & Sepulveda, A. R. (1998b). Variants of the 3' region of the *cagA* gene in *Helicobacter pylori* isolates from patients with different *H. pylori*-associated diseases. *J Clin Microbiol* 36, 2258–2263.
- Yamaoka, Y., El-Zimaity, H. M., Gutierrez, O., Figura, N., Kim, J. G., Kodama, T., Kashima, K. & Graham, D. Y. (1999a). Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. *Gastroenterology* 117, 342–349.

Yamaoka, Y., Kodama, T., Kita, M., Imanishi, J., Kashima, K. & Graham, D. Y. (1999b). Relation between clinical presentation, *Helicobacter pylori* density, interleukin 1beta and 8 production, and *cagA* status. *Gut* 45, 804–811.

Yamaoka, Y., Kikuchi, S., El-Zimaity, H. M., Gutierrez, O., Osato, M. S. & Graham, D. Y. (2002). Importance of *Helicobacter pylori oipA* in

clinical presentation, gastric inflammation, and mucosal interleukin 8 production. *Gastroenterology* 123, 414–424.

Yamaoka, Y., Ojo, O., Fujimoto, S., Odenbreit, S., Haas, R., Gutierrez, O., El-Zimaity, H. M., Reddy, R., Arnqvist, A. & Graham, D. Y. (2006). *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. *Gut* 55, 775–781.