REVIEW

Helicobacter pylori virulence and cancer pathogenesis

Yoshio Yamaoka*[,1,2] & David Y Graham[2]

ABSTRACT: *Helicobacter pylori* is human gastric pathogen that causes chronic and progressive gastric mucosal inflammation and is responsible for the gastric inflammation-associated diseases, gastric cancer and peptic ulcer disease. Specific outcomes reflect the interplay between host-, environmental- and bacterial-specific factors. Progress in understanding putative virulence factors in disease pathogenesis has been limited and many false leads have consumed scarce resources. Few *in vitro–in vivo* correlations or translational applications have proved clinically relevant. Reported virulence factor-related outcomes reflect differences in relative risk of disease rather than specificity for any specific outcome. Studies of individual virulence factor associations have provided conflicting results. Since virulence factors are linked, studies of groups of putative virulence factors are needed to provide clinically useful information. Here, the authors discuss the progress made in understanding the role of *H. pylori* virulence factors CagA, vacuolating cytotoxin, OipA and DupA in disease pathogenesis and provide suggestions for future studies.

Helicobacter pylori causes gastric inflammation and is etiologically related to gastric adenocarcinoma and primary B-cell lymphoma of mucosa-associated lymphoid tissue. Gastric cancer is a model for inflammation-induced cancer [1,2]. The infection is typically life-long and is associated with a decades-long acute and chronic inflammatory response that results in progressive mucosal damage. This has resulted in the highly regulated acid secretory and digestive enzyme producing mucosa being transformed through a series of different types of metaplastic and dysplastic epithelia to eventually result in gastric adenocarcinoma [2]. There is also a strong environmental component involved in the eventual outcome of an *H. pylori* infection. The powerful effect of the environment is reflected in rapid changes in population risk as seen in western countries in the late 19th and early 20th centuries where gastric cancer incidence fell and duodenal ulcers became the prevalent manifestation of the infection. Later, the incidence of both gastric cancer and duodenal ulcer falling was the result of the rapid decline in the prevalence of *H. pylori.* Another example is in Japan where the incidence of gastric cancer fell by approximately 60% between 1965 and 1995 despite no change in virulence or prevalence of the most common infecting strains [3,4].

The result of the interactions between host, environment and bacterial factors is reflected in the most common clinical outcome in a specific population (e.g., atrophic gastritis and gastric carcinoma vs duodenal ulcer disease) [5]. Host factors are reflected in polymorphisms in host genes that govern the intensity of the inflammatory response that also influence the risk of a specific clinical outcome [6].

[1]Department of Environmental & Preventive Medicine, Oita University Faculty of Medicine, 1-1 Idaigaoka, Hasama-machi, Yufu-City, Oita 879-5593, Japan
[2]Department of Medicine–Gastroenterology, Michael E DeBakey Veterans Affairs Medical Center & Baylor College of Medicine, Houston, TX 77030, USA
*Author for correspondence: Tel.: +81 97 586 5740; Fax: +81 97 586 5749; yyamaoka@oita-u.ac.jp

Overall, the clinical presentation of *H. pylori* infections reflects the pattern and severity of gastritis. For example, *H. pylori*-infected individuals living in areas where diets are seasonal with long periods without fresh fruits and vegetables and where food preservation is largely dependent on the use of salt and smoking, have a strong tendency to develop progressive atrophy, which is linked to gastric ulcers and gastric cancer [4]. By contrast, in environments such as Africa, south India or south Asia where fresh fruits and vegetable are available all year round, the mucosal damage tends to remain nonatrophic, the incidence of gastric cancer is low and duodenal ulcer and its complications are the predominant clinical manifestations [7].

However, even in low cancer incidence areas, the presence of polymorphisms in host proinflammatory genes can result in early development of atrophic gastritis and an increased risk of gastric cancer especially if the infected strain also contains virulence factors associated with an enhanced inflammatory response. Therefore while *H. pylori*–host interactions play an important role in disease pathogenesis, bacterial virulence factors are often key factors determining the outcome. In this review, we discuss the role of direct *H. pylori*–host interactions in disease pathogenesis with particular attention to the role of bacterial virulence factors. Between 5 and 10% of *H. pylori*'s 1600 genes are thought to be *H. pylori*-specific, including the putative virulence factors CagA, vacuolating cytotoxin (VacA), OipA and DupA. There are many other putative virulence factors possibly related to clinical outcomes such NapA, HepA and outer membrane proteins including BabA, SabA, HopQ and HomB [8–12]. Their role in pathogenesis remains unlcear and others proved to be false leads (e.g., IceA). For example, IceA turned out to be a restriction enzyme and proof that it was a false lead took years and consumed tremendous resources.

Here, the authors focus on CagA, VacA, OipA and DupA. These were chosen as there is considerable evidence regarding their role in disease pathogenesis, as well as each has a biologic basis for a potential role in disease causation.

CagA

CagA, is a highly immunogenic protein encoded at one end of the *cag* pathogenicity island (PAI). The *cag* PAI is an approximately 40-kbp insertion sequence thought to have been incorporated into *H. pylori* by horizontal transfer. The *cag* PAI encodes a type IV secretion system (T4SS; i.e., a molecular syringe), which injects CagA and other proteins into host cells [13]. The *H. pylori* CagL protein is thought to be a specialized adhesin targeted to the pilus surface, which binds to and activates the integrin α5β1 receptor on gastric epithelial cells. This interaction triggers CagA delivery into target cells [14–16]. Following entry CagA interacts with host cell molecules including the cytoplasmic SHP-2, which has oncogenic activity [17]. It is consensus that *in vivo* CagA-expressing *H. pylori* are associated with an enhanced host inflammatory response and with an increased risk of a clinical outcome, such as peptic ulcer or gastric cancer. However, these same clinical diseases are also caused by infections with CagA-negative *H. pylori*, consistent with the hypothesis that any host or bacterial factor that increases the inflammatory reaction to the infection should also be expected to increase the risk of an important clinical outcome.

CagA has been the subject of innumerable *in vitro* studies and has also been used in animal studies of *H. pylori*-related disease pathogenesis. The many discrepancies in the *in vivo* and *in vitro* data are as yet unexplained. For example, it has been reported that Mongolian gerbils developed gastric cancer with wild-type CagA-expressing *H. pylori* and not with isogenic *cagA* mutants [18,19]. However, although gastric cancer is reported to develop rapidly in this model, the majority of laboratories have been unsuccessful or have at great difficulty in producing gastric cancer with *H. pylori* in Mongolian gerbils irrespective of the strain used [20,21]. CagA-containing transgenic mice have developed gastric cancer and other neoplasms suggesting that CagA can be an oncoprotein when expressed widely [22]. Whether studies with transgenic animals or when *cagA* is transfected into cells actually reflects what occurs naturally when it is injected by bacteria attached to the surface of intact and highly polarized gastric cells remains unclear. Studies are needed to identify how much of the extensive *in vitro* data are relevant to the host–bacterial interactions in life.

The prevalence of CagA among *H. pylori* in different regions varies greatly (e.g., from almost 100% in east Asia to 50% or less in some western countries) [23]. Considerable resources have been expended investigation whether the difference in disease incidence (e.g., gastric cancer) between

east Asian and western countries is related to CagA, CagA structure or some other factors. As noted above, in first part of the 20th century and, currently in the mountainous regions of central and south America, gastric cancer is extremely common; however, the CagA structure is western and not east Asian type. Even in Japan where the east Asian type of CagA is the rule, the incidence of gastric cancer in Japan has fallen rapidly without a change in *H. pylori*, which is consistent with environmental–host–bacteria factors playing a key role in outcome. Nonetheless, differences in CagA structure have proven useful for population epidemiology and for exploring the effects of the infection on changes in signaling pathways in host cells.

cagA-positive strains can be divided into east Asian and western types based on differences in the sequences in the 3′ region of *cagA* [24], which contain a different number of copies of the EPIYA tyrosine phosphorylation site motif (Glu–Pro–Ile–Tyr–Ala). The first-repeat region consists of EPIYA-A/EPIYA-B segments. The segments in the second-repeat region in western and east Asian strains are named EPIYA-C and EPIYA-D, respectively (Figure 1) [17]. CagA sequences are typed based on the order of EPIYA segments, such as type ABC, ABCC, ABCCC and so on for western-type CagA or for east Asian-type CagA as ABD, among others).

EPIYA-D CagA segments exhibits greater *in vitro* binding ability for SHP-2 than do EPIYA-C segments [17]. There are data suggesting that within the same geographic area infection with east Asian-type *cagA* strains is associated with a higher risk of peptic ulcer or gastric cancer than infection with western-type *cagA* strains [25–27]. For example, data from Okinawa suggested that the pattern of gastritis in those

with east Asian type CagA was that of a rapid progression to atrophic gastritis compared with infections with western types of CagA [28]. However, the individuals with these different types of CagA may also differ in terms of other factors unrelated to CagA type, such as when they or their families migrated to Okinawa and diet among others. These host-associated factors are probably important. For example, prior studies have shown that those migrating from an area with a high incidence of gastric cancer generally retain their original risk after migration, whereas their descendents risk tends to become similar to those of the resident population.

The prevalence of second-repeat regions is also higher among patients with gastric cancer in regions where gastric cancer is common. For example, in Columbia, a high cancer incident country, the age-standardized incidence rates per 100,000 population is 17.4 and 57% of isolates examined had 2% EPIYA-C, which contrasted to only 4% of isolates from the low gastric cancer incidence in the USA [23]. However, it is thought that the number of repeats may be an acquired trait. For example, strains with more than one second repeat are rare among children, but become increasingly common in patients with atrophic gastritis [29]. As *H. pylori* infections progress, the intragastric conditions change requiring continued evolution of the infecting *H. pylori* strain, which experience innumerable generations over the lifespan of the host. Strains with multiple second repeats are less able to survive in acidic environments and, thus, their appearance is likely suppressed until hypochlorhydria occurs [30]. The fact that they then appear and become dominant suggests that the presence of multiple second repeats is associated with a survival advantage in stomachs with
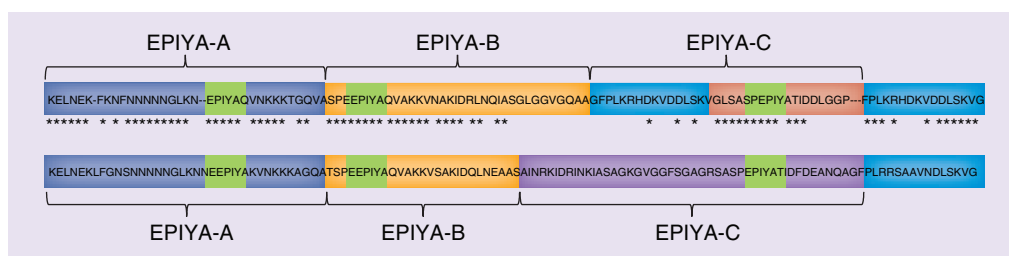

Figure 1. Structural polymorphism in CagA. Western-type CagA contain the EPIYA-A, EPIYA-B and EPIYA-C segments, whereas east Asian-type CagA contain the EPIYA-A, EPIYA-B and EPIYA-D segments. EPIYA-motif in each segment represent the tyrosine phosphorylation sites of CagA. The asterix means that the nucleotide sequences are the same between western-type CagA and east Asian-type CagA.

hypochlorhydria. Hypochlorhydria is also associated with increased growth of non-*H. pylori* bacteria in the stomach and thus increased competition for resources. While, the nature of any survival advantage is as yet unknown, one possibility is to produce a more robust inflammatory response thus allowing for increased availability of nutrients to the organism.

There are *in vitro* data showing that Csk, an important molecule involved in intracellular signaling, prefers to bind EPIYA-A and EPIYA-B motifs [31] and that c-Src only phosphorylates EPIYA-C and -D motifs whereas c-Abl phosphorylates EPIYA-A, -B, -C and -D motifs [32]. Studies are needed that compare *H. pylori* that are thought to differ in virulence in terms of growth rates under different conditions, their ability to stimulate inflammation, to survive intracellularly, to increase intracellular signaling or activate additional signaling pathways, and so on, Studies are also needed regarding the effect of an increase in second repeat number on bacteria–host interactions *in vivo*.

The type of CagA has been studied in relation to gastric cancer incidence. Although both the Vietnam and South Korean strain typically possess the east Asian type CagA, the incidence of gastric cancer in Vietnam is half of that in South Korea [33]. Most Vietnam and South Korea strains have one EPIYA-A, EPIYA-B, and EPIYA-D segment (i.e., ABD) [34]. Vietnamese strains also have a unique 18-bp deletion located upstream of the EPIYA-A segment. The availability of a wide range of CagA subtypes provides a wealth of reagents to examine if and how they may be involved in the pathogenesis of gastric cancer.

Despite the innumerable *in vitro* studies attempting to identify a precise role for CagA in the pathogenesis of gastric cancer, its role *in vivo* remains unknown. Most, but not all, *in vitro* studies have used nonpolarized gastric cancer cells or polarized nongastric cells [35,36]. In one interesting study, polarized cells were grown as monolayers, thus mirroring the *in vivo* condition in that *H. pylori* were able only to contact the apical surface [37]. In this simulated natural condition, changes in cell shape and mobility were not possible due to their being in a polarized monolayer and the *H. pylori*–cell interaction only resulted in alterations of the apical surface, which become a site especially suitable for bacterial replication. In addition, when nutrients were removed from the fluids overlying the cells, only CagA-positive organisms were able to obtain all their necessary nutrients and iron directly from the cells and survive. By contrast, CagA-negative strains were not able to survive. *In vivo*, both can survive on the cell surface and both tend to attach at the intracellular junctions, a site where nutrients can potentially transit from the host to the external environment.

Paracellular permeability to small molecules such as sucrose is increased in *H. pylori* infection [38]. This increase correlates best to the presence of the inflammatory infiltrate and is seen in both *H. pylori* and in non-*H. pylori* lymphocytic gastritis. Both polymorphonuclear cells and intraepithelial lymphocytes can transit through the intracellular tight junctions without damaging them. During transit, intracellular permeability is transiently increased and the transmembrane potential difference slightly reduced but overall is maintained. *H. pylori* can also be found beneath the tight junctions suggesting that the trafficking may be bidirectional. Although there are a number of studies showing that *H. pylori* can affect tight junctions *in vitro*, there are little *in vivo* data to support these observations [39]. Interestingly, recent studies using semimonolayer gastric cells, MKN28 cells showed that barrier dysfunction requires functional urease activity and was independent of the *cag* pathogenicity island (including CagA) or VacA [36]. Clearly, more *in vivo* and relevant *in vitro* experiments are needed to clarify the role of CagA and the number of EPIYA motifs.

VacA

Almost all *H. pylori* contain the *vacA* gene that encodes a vacuolating cytotoxin. It remains unclear what role VacA plays in disease pathogenesis as most observation are based on *in vitro* experiments of unclear clinical relevance. VacA was recognized because of vacuolation following *in vitro* exposure of cell lines. *In vitro* studies have documented multiple activities, such as membrane channel formation, cytochrome C release from mitochondria leading to apoptosis, binding to cell membrane receptors resulting the initiation of a proinflammatory response and inhibition of T-cell activation and proliferation [23]. It remains unclear which, if any, of these vitro observations have *in vivo* counterparts. Early studies showed *H. pylori* antigens in the submucosa and more detailed analysis showed that *H. pylori* were also present within the mucosal cells [40]. More recent studies have

even been able to culture *H. pylori* from perigastric lymph nodes removed at surgery [40]. Clearly, a small proportion of *H. pylori* are able to invade and live within the mucosa. The question about how has been possibly answered by studies suggesting that VacA has a major and possibly primary role related to autophagy where it assists in producing the vacuole where *H. pylori* can survive intracellularly [41]. It is possible that this is its most important role in disease pathogenesis.

Like CagA, a number of association studies have been carried out, relating differences in *vacA* gene structure and clinical outcomes. Differences have been described in the signal (s) region (s1 and s2) and the middle (m) regions (m1 and m2) that contribute to variations in the *in vitro* vacuolating activity [42]. s1/m1 strains appear to be the most cytotoxic *in vitro* and s2/m2 strains have no *in vitro* cytotoxic activity; s1/m2 strains have reduced activity compared with s1/m1. It has been suggested that. compared with those with s2 or m2 strains, patients infected with *vacA* s1 or m1 strains may be at increased risk for developing peptic ulcers and/ or gastric cancer [42–44]. However, the risk of peptic ulcer or gastric cancer is best correlated with the extent and severity of inflammation and few studies have controlled for the presence of CagA or other variables known to be associated with severity of inflammation. The majority of *H. pylori* strains in east Asia are *vacA* s1 genotype and in east Asia s region genotypes have proven to be independent of clinical outcomes, thus not confirming the reported significance of the VacA genotype correlations reported from western countries [23].

The m1 genotype is common in northern Asian countries where gastric cancer is frequent whereas m2 genotype is predominant in southeast Asia (e.g., Taiwan and Vietnam) [23]. Generally, the incidence of gastric cancer is higher in northern regions than in the southern regions of east Asia and this north–south difference is also evident in Vietnam where the m1 genotype and gastric cancer also show a north–south gradient [33]. Even in east Asia, in areas where non-s1/m1 strains are common, *vacA* m genotypes correlate roughly with risk of different *H pylori*-related diseases.

vacA s1 and m1 genotype have also be subdivided into subtypes (i.e., s1a, s1b and s1c [42], and m1a, m1b and m1c, respectively) [45]. *vacA* genotypes vary geographically, for example, *vacA* s1c and m1b are common in east Asia, s1a and m1c

are common in south Asia [45,46], the predominant genotype in central Asia is *vacA* m1c [46], m1a is common in Africa and ethnic Europeans [46], s1a and s1b genotypes are common in ethnic Europeans whereas s1b genotypes are more prevalent in the Iberian Peninsula, Latin America [47] and Africa [47]. However, it remains unclear whether subtypes of s1 and m1 genotypes relate to clinical outcomes or to other factors governing differences in the extent, severity or distribution of gastritis.

The region of *vacA* between the s region and the m region (the intermediate or i region) has been separated into i1, i2 and i3 subtypes [48]. Attempts have also been make to link i subtypes to pathogenicity and risk of disease. When i subtypes were first reported, it was hypothesized that the *vacA* i genotype would be more helpful in assessing risk of gastric cancer than was typing using the s and m regions. This was followed by studies linking the *vacA* i genotype to peptic ulcer [49,50] and polymorphisms at amino acid position 196 of *vacA* (in the i region) and severe outcomes in South Korea [25]. Hopes for a predictive value of *vacA* i genotyping were dashed when no associations were found between the i region and clinical outcomes east and southeast Asia [51] and a long-term follow-up study from Portugal reported that *vacA* i genotyping did not improve the prediction of progression in relation to other *vacA* loci [52]. Like the phoenix, *vacA* i genotyping continues to rise from the dead, with the description of another, putatively disease-related region, between the i region and the m region called the deletion or d region [53]. The d region consisted of two groups (i.e., those with no deletion, d1 and d2 genotype with a 69–81-bp deletion). The d1 genotype was associated with mucosal atrophy in western strains, but in east Asia, d genotyping proved useless as all strains were classified as s1/i1/d1. Follow-up studies are needed. For example, a study from Italy looked at *vacA* genotypes over 15 years (from January–June 1989 to January–June 2005) [54]. They reported that although while the prevalence of s1m1 (thought to be more virulent genotype) increased, there was a reduction of peptic ulcer disease in the same area.

In summary, despite innumerable papers attempting to relate *vacA* genotypes to outcome or disease pathogenesis, no truly consistent associations or demonstrable biologic basis for the putative associations as appeared. *In vivo*, the only consistent association with risk of a

clinically important outcome has been the extent and severity of inflammation. One would expect that important observations would be universally applicable or the reason why not would be clearly explained. It is possible that differences in the prevalence of *vacA* genotypes primarily represents differences related to human migration and that putative disease associations are likely to have occurred by chance rather than as a cause and effect relationship. In addition, the available studies have tended not to combine the possible effects of other putative virulence factors. Finally, it would be interesting to examine the effects of genotyping on autophagy or other possible important biologic functions. Much has been done with genotyping, but little useful has been accomplished to date.

OipA

OipA, an outer membrane protein is an adhesin [55]. OipA's functional status is regulated by slipped-strand mispairing based on the number of CT dinucleotide repeats in the 5′ region of the genes that can be either switched 'on' (functional) or switched 'off' (nonfunctional) [55]. OipA was originally identified as it is also a proinflammatory response-inducing protein involved in the induction of IL-8 from gastric epithelial cells [55]. Transcription of *IL-8* genes is both OipA- and *cag* PAI-dependent through interactions with different binding sites are involved, such as transcription factors within the IL-8 promoter for NFκB, AP-1 and ISRE-like element [56–58]. Other *in vitro* studies have identified a role for OipA in inflammation and actin dynamics through the phosphorylation of signaling pathways that interact with *cag* PAI (CagA)-related pathways **(Figure 2)** [56,59–63]. OipA and *cag* PAI and/or CagA interactions are postulated to be involved in β-catenin signaling that affects cell–cell junctions and proliferation **(Figure 2)** [19].

Both OipA and *cag* PAI/CagA-related pathways and OipA-specific pathways have been described. For example, although both OipA and the *cag* PAI activate the IL-8 promoter, the pathways diverge upstream of IRF-1 with only OipA being involved in the STAT-1→IRF-1→ISRE-like pathway **(Figure 2)** [56]. AP-1 activity is stimulated by a complex network of signaling pathways that involve MAPKs (JNK, ERK and p38) which activate various transcription factors resulting in the induction of transcription of *fos* and *jun*, thus increasing the number of AP-1 complexes available to activate AP-1

target genes, such as *IL-8*. Activated expression of c-Fos and c-Jun proteins, as well as *c-fos* and *c-jun* mRNA, are OipA- and *cag* PAI-dependent [57–58,60–61]. Overall, the data are consistent with the notion that both OipA and the *cag* PAI are involved in JNK and ERK phosphorylation [60,64] with only OipA being involved in p38 phosphorylation [61,64–65].

OipA is also involved in phosphorylation of FAK tyrosine 397, 576, 577, 861 and 925, stress fiber formation and cell morphology **(Figure 2)** [62], as well as in phosphorylation of AKT serine 473, paxillin phosphorylation at tyrosine 31 and 118 and actin stress fiber formation [66], With regard to stress fiber formation and changes in cell morphology, *cag* PAI involvement appears limited to phosphorylation of AKT threonine 308 and phosphorylation of tyrosine 118 of paxillin **(Figure 2)** [63].

FoxO1 and FoxO3a are nuclear substrates of *H. pylori*-induced PI3K/Akt cell survival signaling involved in IL-8 production [67]. OipA regulates IL-8 release through PI3K/Akt, which, in turn, requires FoxO1/3a inactivation **(Figure 2)**. By contrast, *cag* PAI-mediated IL-8 production is FoxO1/3a-independent, such that both OipA and *cag* PAI (CagA) regulate intracellular signaling pathways together.

Wild-type *H. pylori* 7.13 infects gerbils and has been associated with gastric cancer in gerbils; gerbils infected with its isogenic *oipA* mutant failed to developed gastric cancer [19], suggesting a possible role for OipA in the development of gastric cancer, at least in Mongolian gerbils.

There is other support for OipA being involved in disease pathogenesis. For example, challenge of human volunteers with an *oipA* 'on'/whole *cag* PAI-negative clinical isolate (Baylor strain 100 or ATCC BAA-945) caused severe inflammation consistent with its role as a proinflammatory molecule [68]. Examination of presence of multiple putative virulence factors (e.g., the *cag* PAI, *vacA, iceA, oipA* and *babA*) in relation to clinical outcomes using multiple logistic regression analysis showed that only the oipA 'on' status was an independent determinant predictor of duodenal ulcer (adjusted odds ratio [aOR]: 5.0; 95% CI: 2.1–11.9) [69]. This observation was confirmed in another study using immunoblot analysis for 4 adhesins: OipA, BabA, BabB and SabA [70]. Multiple logistic regression analysis showed that only the OipA-positive status was an independent determinant predictor of gastric cancer versus gastritis (aOR: 4.8; 95% CI:
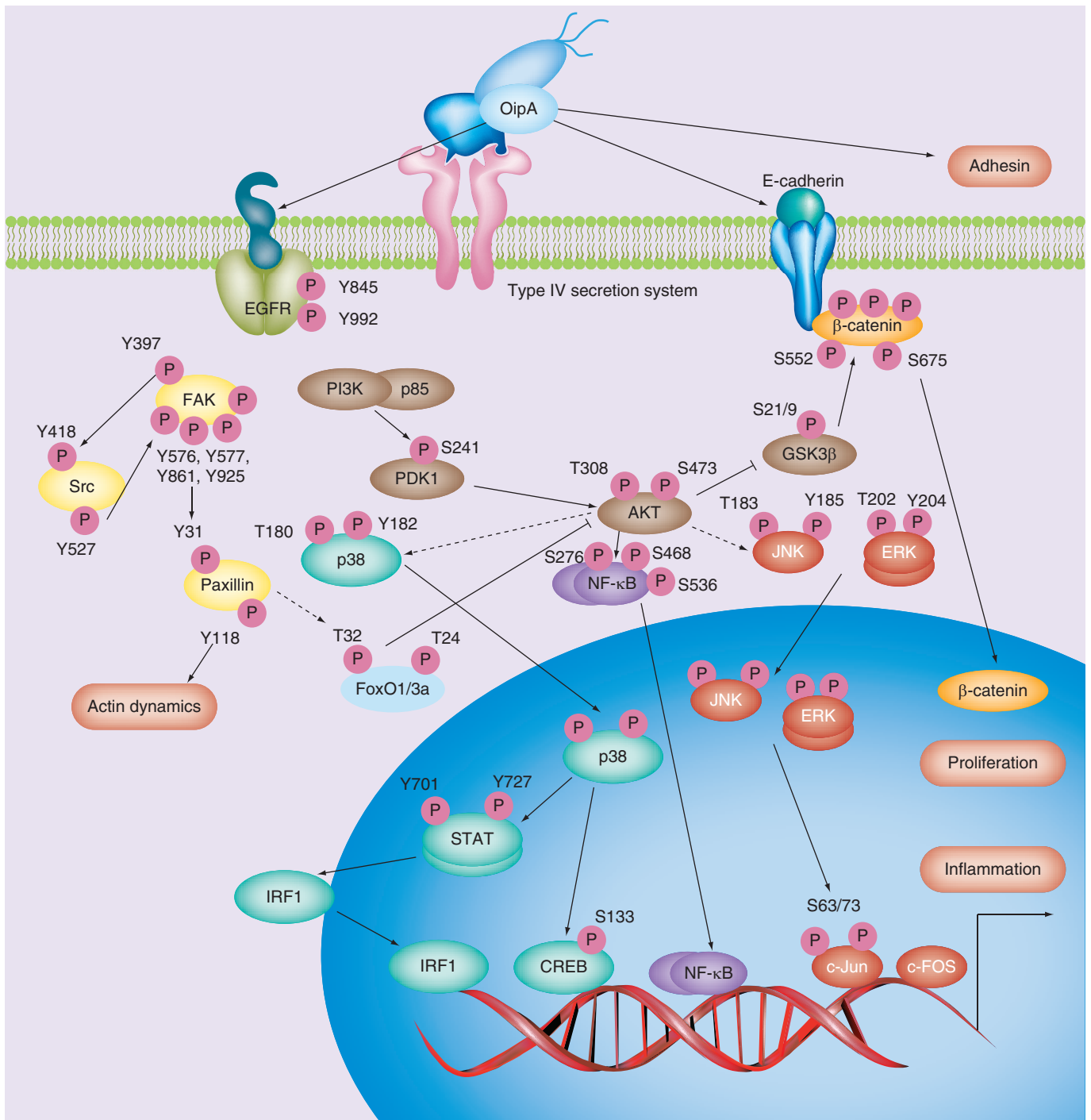
Figure 2. Pathogenesis of OipA-related signaling. OipA is reported to regulate various epithelial cell signaling pathways. The dashed lines show that the pathways are still unclear.
EGFR: EGF receptor; S: Serine; T: Threonine.

1.4–16.8) and duodenal ulcer versus gastritis (aOR: 4.0; 95% CI: 1.6–10.2).

It is important to acknowledge the association between the *cag* PAI and OipA as clinical *cag* PAI-positive isolates are typically also have *oipA*

'on' status [69–73]. In addition, *oipA* status is also linked to the *vacA* s region type, which in turn is closely linked to the presence of the *babA* gene, with encodes a blood group binding adhesin [74]. Such data support the notion, compared with

studies of isolated factors or studies searching for a 'most virulent' factor, putative virulence factors probably interact synergistically. We believe that studies of groups of putative virulence factors are needed to provide clinically useful information [23]. It is also important to reconfirm that the presence of CagA, VacA and OipA are linked such that typically *H. pylori* either produce all of these proteins or none of them and that clinical outcomes, such as peptic ulcer and gastric cancer, are associated with strains with and without these virulence factors. However, strains with recognized virulence factors tend to produce more severe inflammation and are associated with higher risk of these important clinical outcomes.

DupA

DupA (encoded by the duodenal ulcer promoting gene A) was identified based on its positive relationship to duodenal ulcer and inverse relationship to gastric cancer [75]. The presence of DupA was also associated with elevated IL-8 production in the antrum, (i.e., antrum-predominant gastritis – a feature of duodenal ulcer disease) and has been reported to induce IL-12 production from monocytes [76]. The initial study described *dupA* in 42% of patients with duodenal ulcer versus only 9% in patients with gastric cancer, irrespective of nationality [75]. Subsequent studies [23,77], and a meta-analysis, have confirmed that *dupA*-positive *H. pylori* increased the duodenal ulcer risk by approximately 40% (aOR: 1.41; 95% CI: 1.12–1.76), but only in Asian countries (aOR: 1.57; 95% CI: 1.19–2.06) [78]. These differences in results are partially explained; however, recent studies showing that many strains have frame shift mutations in *dupA* and are unable to produce intact DupA proteins. Thus, studies based on finding the gene tend to overestimate its prevalence and emphasizing the importance of confirming the presence of the protein.

dupA has also been divided into *dupA1* with a 1884-bp allele and *dupA2* a truncated version with mutations [76]. Full-sequenced data of *H. pylori* showed that the length of *dupA* is strain dependent with strains Shi470 and G27 having an approximately 600-bp longer open reading frame (approximately 2500 bp) than strain J99 due an additional N-terminal region of *dupA*. These data suggest *dupA* genotypes can also be defined as long type and short types in relation to location of the signal sequence of the N-terminal region. In Okinawa, only the long-type *dupA* was significantly associated with severe gastroduodenal diseases [79]. It remains unclear whether the short-type *dupA* produces a functional protein, or even whether long-type *dupA* without frame shift mutations are functional. Most studies have not shown a relationship between the presence of *dupA* and CagA, VacA or OipA [78].

Overall, studies with DupA have been very instructive and emphasize the importance of multinational studies to confirm the universality of putative associations, as well as the need to confirm studies based only on molecular tools with those confirming expression of the proteins less we continue to be misled.

DupA as a type IV secretion system

Putative virulence factors should be linked to a biologic mechanism responsible for the observed actions. DupA has been suggested to form a T4SS with *vir* genes that make up the *dupA* gene cluster. Three gene clusters that code for T4SS have been recognized: a protein translocation system encoded by the *cag* PAI, a DNA-uptake system encoded by the ComB cluster and an unknown cluster in the plasticity zone [80]. *virB4*, one of the constituents of T4SS, is highly homologous to *dupA* and *dupA* and the adjacent six *vir* gene homologs (*virB8*, *B9*, *B10*, *B11*, *virD4* and *D2*) in the plasticity zone have been suggested to be the basis for a third T4SS named dupA cluster or *tfs3a* **(Figure 3)** [23]. A putative T4SS having *virB4* sequence, but without *dupA* has been named *tfs3b* [81]. Many strains, such as strain G27, contain complete *dupA* cluster, but are unlikely to produce a functional T4SS because *dupA* has a frame shift mutation. Another example is strain J99, which has a mutation in *dupA* and contains only part of complete *dupA* cluster. Only strains with a functional *dupA* and a complete *dupA* cluster would be expected to possess fully functional T4SS. One study compared the prevalence of *dupA* and *vir* gene homologs and the associations between the status of *dupA* clusters and clinical outcomes reported that a complete *dupA* cluster was associated with an increased risk of duodenal ulcer disease compared strains predicted to not produce a T4SS and results were independent of *cag* PAI status (aOR: 2.13; 95% CI: 1.13–4.03) [80]. Further progress will require the availability of reliable immunologic reagents for DupA. However, the genes in the plasticity zone remain an attractive area for future investigations.
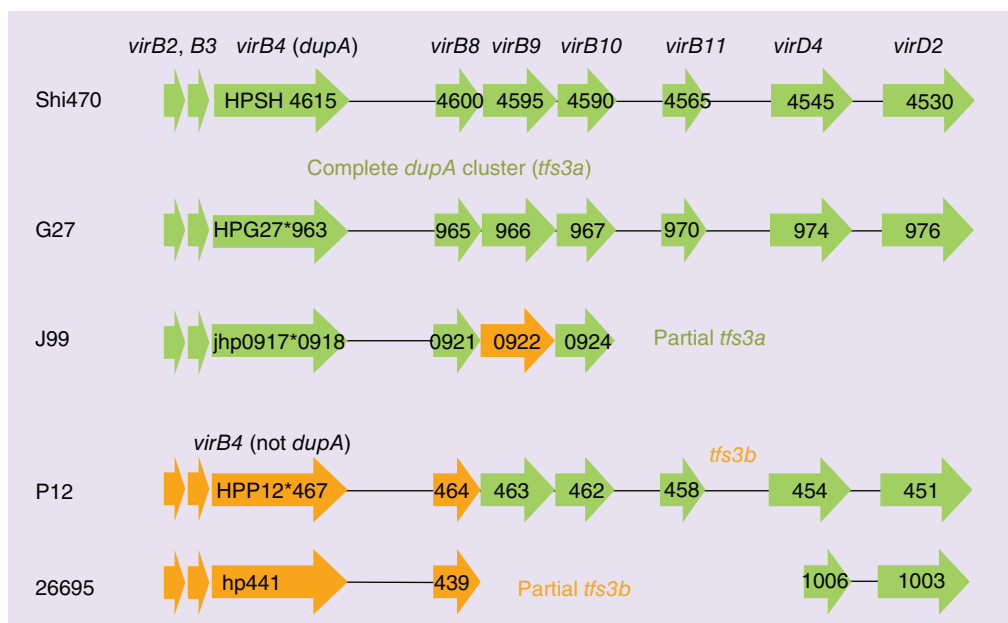
Figure 3. Type IV secretion system in plasticity zone. Gene arrangement in plasticity region of strain Shi470, which has complete *dupA* cluster with functional *dupA* gene sequence, and comparison with corresponding regions in other genome sequences are presented. Genes encoding type IV secretion system components are indicated as green arrows, and frameshift mutations are indicated by asterisks. Genes encoding proteins with 90–95% sequence similarity to the Shi470 proteins are shown in green color, genes encoding proteins with 50–75% sequence similarity are shown in orange color.

Sequencing to detect putative virulence factors: the future

Sequencing technology provides rapid methods of identifying novel putative virulence factors [82–84] by comparing the genome sequences of isolates from patient with or without a specific disease either in one region or across geographic areas [83,85] and by comparing the genome sequences of isolates from one patient during the progression of the disease to identify genes gained or lost [86]. This approach should assist greatly in understanding the biology of the organism–host interaction over time. At the present time it seems unlikely that cancer-specific genes will be found. Clinically, duodenal ulcer and gastric cancer are at different ends of the spectrum of disease; however, in retrospect, that observation is based on temporal associations and differences in the patterns of gastritis associated with the different diseases. Duodenal ulcer can only occur during the phase when acid secretion is maintained as progressive inflammation will eventually lead to a fall in acid secretion below the point where the ulcer disease can be maintained and it will burn out [87]. Thus, in regions where atrophic gastritis is common, there may be only a short window

when duodenal ulcers can occur and is consistent with the finding of duodenal ulcer scars in a proportion of patients presenting with atrophic gastritis and gastric cancer. Thus, comparisons between strains from duodenal ulcer and gastric cancer or between gastric ulcer and gastric cancer are unlikely to identify 'cancer' genes, especially with gastric ulcer, which is linked to corpus gastritis and risk for gastric cancer. It is more likely that comparison of gastric cancer patients across geographic regions would be a better strategy. However, data obtained with the massively parallel sequencing technology should provide considerable data useful to understanding the interactions between the host and the bacteria during different phases of the disease and different gastric histologies, but at the risk of provided many false leads.

Conclusion

Eradication of *H. pylori* before the development of significant and irreversible gastric damage can prevent cancer. This should not be surprising given that gastric cancer is known to be an inflammation-associated malignancy. However, the fact that *H. pylori* eradiation in patients with

severe atrophy and irreversible changes reduced development of metachronous cancers was welcome but somewhat surprising [88]. This result suggested the presence of direct *H. pylori*–host interactions resulting in genetic instability of the host genome.

Future perspective

H. pylori contains approximately 1600 genes and, to date, it is likely that only a fraction of the potential virulence genes have been identified. These techniques and sequential sampling of stomach during the course of the infection provide an opportunity to link changes in the bacteria with environmental factors, such as diet (e.g., salt intake), which are critical determinants of outcome. We expect that epidemiologic studies associating the virulence factor with a clinical outcome will continue to be published, but hopefully they will include representative populations from different regions and continents. Associations from small convenience samples, without a postulated biologic mechanism, and without considering the presence of other virulence and environmental factors have to date resulted in many false leads and wasting of resources. Future studies should examine putative virulence factors as groups or as part of a virulence complex rather than individually. This approach is likely to provide a more complete and richer understanding of the mechanisms underlying how *H. pylori* induces gastric inflammation and leads to severe gastroduodenal diseases by combining bacterial factors with other factors, such as environmental factors and host factors.

EXECUTIVE SUMMARY

Background

- *Helicobacter pylori* causes gastric inflammation and is etiologically responsible for gastric adenocarcinoma and primary B-cell lymphoma of mucosa-associated lymphoid tissue as well as gastric and duodenal ulcer disease.

- As with other chronic infectious diseases associated with a long latent period (e.g., syphilis and tuberculosis), only a proportion of infected individuals develop a clinically important outcome and the specific outcomes reflect the interplay between host-, environmental- and bacterial-specific factors.

- While it is important to recognized that *H. pylori* provides an outstanding model of bacterial-induced chronic mucosal inflammation and inflammation-associated malignancy, all *H. pylori*-associated diseases have been associated with strains possessing few if any of the currently recognized putative virulence factors.

CagA & VacA genotyping

- Despite innumerable *in vitro* studies attempting to identify a precise role for CagA and *vacA* genotyping in relation to clinical outcome or disease pathogenesis, no truly consistent associations or demonstrable biologic basis has yet appeared.

CagA, VacA & OipA

- The presence of CagA, VacA and OipA are typically linked such that *H. pylori* either produce all of these proteins or none of them. Importantly, clinical outcomes, such as peptic ulcer and gastric cancer, are associated with strains with and without these virulence factors. However, strains these virulence factors tend to produce more severe inflammation and are associated with a higher risk of important clinical outcomes.

DupA as a type IV secretion system

- Studies with DupA have emphasized the importance of multinational studies to confirm the universality of putative associations found in one region, as well as the requirement to confirm protein expression as results based only on genotypic analyses have provided misleading conclusion. Further progress will require the availability of reliable immunologic reagents for DupA. However, the genes in the plasticity zone including DupA remain an attractive area for future investigations.

Sequencing to detect putative virulence factors

- Data obtained with the massively parallel sequencing technology should provide considerable data useful to understanding the interactions between the host and the bacteria during different phases of the disease and different gastric histologies, but at the risk of provided many false leads.

Gastric cancer is an inflammation-induced genetic disease. There are at least five known pathways where direct *H. pylori*–host cell interactions result in host cell genetic instability (AID, double-stranded breaks, methylation and mismatch repair) and aberrant miRNA expression [1,89–93]. Each of these are also reduced or eliminated following *H. pylori* eradication. The role of *H. pylori* infection in inducing genetic instability appears to be a fruitful as well as what role putative virulence factors might play.

Disclaimer

Financial & competing interests disclosure

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1   Chiba T, Marusawa H, Ushijima T. Inflammation-associated cancer development in digestive organs: mechanisms and roles for genetic and epigenetic modulation. *Gastroenterology* 143(3), 550–563 (2012).

••   **Demonstrates that chronic inflammation is a strong risk factor for the development of cancer including *Helicobacter pylori*-related gastric cancer. AID, in particular, plays a key role in inflammation-associated carcinogenesis.**

2   Correa P. Gastric cancer: overview. *Gastroenterol. Clin. North Am.* 42(2), 211–217 (2013).

3   Atherton JC, Blaser MJ. Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. *J. Clin. Invest.* 119(9), 2475–2487 (2009).

4   Graham DY, Lu H, Yamaoka Y. African, Asian or Indian enigma, the East Asian *Helicobacter pylori*: facts or medical myths. *J. Dig. Dis.* 10(2), 77–84 (2009).

5   Lee YY, Mahendra RS, Graham DY. *Helicobacter pylori* infection – a boon or a bane: lessons from studies in a low-prevalence population. *Helicobacter* 18(5), 338–346 (2013).

6   Amieva MR, El-Omar EM. Host–bacterial interactions in *Helicobacter pylori* infection. *Gastroenterology* 134(1), 306–323 (2008).

7   Graham DY. *Helicobacter pylori* infection in the pathogenesis of duodenal ulcer and gastric cancer: a model. *Gastroenterology* 113(6), 1983–1991 (1997).

8   Belogolova E, Bauer B, Pompaiah M *et al.* *Helicobacter pylori* outer membrane protein HopQ identified as a novel T4SS-associated virulence factor. *Cell Microbiol.* 15(11), 1896–1912 (2013).

9   Ilver D, Arnqvist A, Ogren J *et al.* *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 279(5349), 373–377 (1998).

10   Mahdavi J, Sonden B, Hurtig M *et al.* *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science* 297(5581), 573–578 (2002).

11   Oleastro M, Cordeiro R, Yamaoka Y *et al.* Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. *Gut Pathog.* 1(1), 12 (2009).

12   Shiota S, Watada M, Matsunari O *et al.* *Helicobacter pylori iceA*, clinical outcomes, and correlation with *cagA*: a meta-analysis. *PLoS ONE* 7(1), e30354 (2012).

13   Tegtmeyer N, Wessler S, Backert S. Role of the *cag*-pathogenicity island encoded type IV secretion system in *Helicobacter pylori* pathogenesis. *FEBS J.* 278(8), 1190–1202 (2011).

14   Backert S, Clyne M, Tegtmeyer N. Molecular mechanisms of gastric epithelial cell adhesion and injection of CagA by *Helicobacter pylori*. *Cell Commun. Signal.* 9, 28 (2011).

15   Barden S, Lange S, Tegtmeyer N *et al.* A helical RGD motif promoting cell adhesion: crystal structures of the *Helicobacter pylori* type IV secretion system pilus protein CagL. *Structure* 21(11), 1931–1941 (2013).

16   Kwok T, Zabler D, Urman S *et al.* *Helicobacter* exploits integrin for type IV secretion and kinase activation. *Nature* 449(7164), 862–866 (2007).

17   Hatakeyama M. Anthropological and clinical implications for the structural diversity of the *Helicobacter pylori* CagA oncoprotein. *Cancer Sci.* 102(1), 36–43 (2011).

18   Franco AT, Israel DA, Washington MK *et al.* Activation of beta-catenin by carcinogenic *Helicobacter pylori*. *Proc. Natl Acad. Sci. USA* 102(30), 10646–10651 (2005).

19   Franco AT, Johnston E, Krishna U *et al.* Regulation of gastric carcinogenesis by *Helicobacter pylori* virulence factors. *Cancer Res.* 68(2), 379–387 (2008).

20   Sugimoto M, Ohno T, Yamaoka Y. Expression of angiotensin II type 1 and type 2 receptor mRNAs in the gastric mucosa of *Helicobacter pylori*-infected Mongolian gerbils. *J. Gastroenterol.* 46(10), 1177–1186 (2011).

21 Sugimoto M, Ohno T, Graham DY, Yamaoka Y. *Helicobacter pylori* outer membrane proteins on gastric mucosal interleukin 6 and 11 expression in Mongolian gerbils. *J. Gastroenterol. Hepatol.* 26(11), 1677–1684 (2011).

22 Ohnishi N, Yuasa H, Tanaka S *et al.* Transgenic expression of *Helicobacter pylori* CagA induces gastrointestinal and hematopoietic neoplasms in mouse. *Proc. Natl Acad. Sci. USA* 105(3), 1003–1008 (2008).

23 Yamaoka Y. Mechanisms of disease: *Helicobacter pylori* virulence factors. *Nat. Rev. Gastroenterol. Hepatol.* 7(11), 629–641 (2010).

24 Yamaoka Y, Kodama T, Kashima K, Graham DY, Sepulveda AR. 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(8), 2258–2263 (1998).

•• There are two main genotypes in *CagA* (east Asian and western-type). This paper also emphasize that the repeat region of CagA is epidemiologically related to the development of gastric cancer.

25 Jones KR, Jang S, Chang JY *et al.* Polymorphisms in the intermediate region of VacA impact *Helicobacter pylori*-induced disease development. *J. Clin. Microbiol.* 49(1), 101–110 (2011).

26 Vilaichone RK, Mahachai V, Tumwasorn S *et al.* Molecular epidemiology and outcome of *Helicobacter pylori* infection in Thailand: a cultural cross roads. *Helicobacter.* 9(5), 453–459 (2004).

27 Hirai I, Sasaki T, Kimoto A *et al.* Infection of less virulent *Helicobacter pylori* strains in asymptomatic healthy individuals in Thailand as a potential contributing factor to the Asian enigma. *Microbes Infect.* 12(3), 227–230 (2010).

28 Matsunari O, Shiota S, Suzuki R *et al.* Association between *Helicobacter pylori* virulence factors and gastroduodenal diseases in Okinawa, Japan. *J. Clin. Microbiol.* 50(3), 876–883 (2012).

29 Yamaoka Y, Reddy R, Graham DY. *Helicobacter pylori* virulence factor genotypes in children in the United States: clues about genotype and outcome relationships. *J. Clin. Microbiol.* 48(7), 2550–2551 (2010).

30 Yamaoka Y, El-Zimaity HM, Gutierrez O *et al.* Relationship between the *cagA* 3' repeat region of *Helicobacter pylori,* gastric histology, and susceptibility to low pH. *Gastroenterology* 117(2), 342–349 (1999).

31 Selbach M, Paul FE, Brandt S *et al.* Host cell interactome of tyrosine-phosphorylated bacterial proteins. *Cell Host Microbe* 5(4), 397–403 (2009).

32 Mueller D, Tegtmeyer N, Brandt S *et al.* c-Src and c-Abl kinases control hierarchic phosphorylation and function of the CagA effector protein in western and east Asian *Helicobacter pylori* strains. *J. Clin. Invest.* 122(4), 1553–1566 (2012).

33 Uchida T, Nguyen LT, Takayama A *et al.* Analysis of virulence factors of *Helicobacter pylori* isolated from a Vietnamese population. *BMC Microbiol.* 9, 175 (2009).

34 Xia Y, Yamaoka Y, Zhu Q, Matha I, Gao X. A comprehensive sequence and disease correlation analyses for the C-terminal region of CagA protein of *Helicobacter pylori*. *PLoS ONE* 4(11), e7736 (2009).

35 Fiorentino M, Ding H, Blanchard TG *et al.* *Helicobacter pylori*-induced disruption of monolayer permeability and proinflammatory cytokine secretion in polarized human gastric epithelial cells. *Infect. Immun.* 81(3), 876–883 (2013).

36 Wroblewski LE, Shen L, Ogden S *et al.* *Helicobacter pylori* dysregulation of gastric epithelial tight junctions by urease-mediated myosin II activation. *Gastroenterology* 136(1), 236–246 (2009).

37 Tan S, Tompkins LS, Amieva MR. *Helicobacter pylori* usurps cell polarity to turn the cell surface into a replicative niche. *PLoS Pathog.* 5(5), e1000407 (2009).

38 Borch K, Sjostedt C, Hannestad U *et al.* Asymptomatic *Helicobacter pylori* gastritis is associated with increased sucrose permeability. *Dig. Dis. Sci.* 43(4), 749–753 (1998).

39 Wex T, Monkemuller K, Kuester D *et al.* Zonulin is not increased in the cardiac and esophageal mucosa of patients with gastroesophageal reflux disease. *Peptides* 30(6), 1082–1087 (2009).

40 Ito T, Kobayashi D, Uchida K *et al.* *Helicobacter pylori* invades the gastric mucosa and translocates to the gastric lymph nodes. *Lab. Invest.* 88(6), 664–681 (2008).

41 Raju D, Hussey S, Ang M *et al.* Vacuolating cytotoxin and variants in Atg16L1 that disrupt autophagy promote *Helicobacter pylori* infection in humans. *Gastroenterology* 142(5), 1160–1171 (2012).

42 Atherton JC, Cao P, Peek RM Jr *et al.* Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. *J. Biol. Chem.* 270(30), 17771–17777 (1995).

• First paper describing the presence of mosaic genotypes in *vacA* gene in the signal sequences and middle region. This also shows that s1m1 genotype is the most toxic whereas s2m2 genotype is nontoxic.

43 Sugimoto M, Zali MR, Yamaoka Y. The association of *vacA* genotypes and *Helicobacter pylori*-related gastroduodenal diseases in the Middle East. *Eur. J. Clin. Microbiol. Infect. Dis.* 28(10), 1227–1236 (2009).

44 Sugimoto M, Yamaoka Y. The association of *vacA* genotype and *Helicobacter pylori*-related disease in Latin American and African populations. *Clin. Microbiol. Infect.* 15(9), 835–842 (2009).

45 Mukhopadhyay AK, Kersulyte D, Jeong JY *et al.* Distinctiveness of genotypes of *Helicobacter pylori* in Calcutta, India. *J. Bacteriol.* 182(11), 3219–3227 (2000).

46 Yamaoka Y, Orito E, Mizokami M *et al.* *Helicobacter pylori* in North and South America before Columbus. *FEBS Lett.* 517(1–3), 180–184 (2002).

47 Sugimoto M, Yamaoka Y. The association of *vacA* genotype and *Helicobacter pylori*-related disease in Latin American and African populations. *Clin. Microbiol. Infect.* 15(9), 835–842 (2009).

48 Rhead JL, Letley DP, Mohammadi M *et al.* A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. *Gastroenterology* 133(3), 926–936 (2007).

49 Hussein NR, Mohammadi M, Talebkhan Y *et al.* Differences in virulence markers between *Helicobacter pylori* strains from Iraq and those from Iran: potential importance of regional differences in *H. pylori*-associated disease. *J. Clin. Microbiol.* 46(5), 1774–1779 (2008).

50 Basso D, Zambon CF, Letley DP *et al.* Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. *Gastroenterology* 135(1), 91–99 (2008).

51 Ogiwara H, Graham DY, Yamaoka Y. *vacA* i-region subtyping. *Gastroenterology* 134(4), 1267–1267 (2008).

52 Ferreira RM, Figueiredo C, Bonet C *et al.* *Helicobacter pylori vacA* intermediate region genotyping and progression of gastric preneoplastic lesions. *Am. J. Gastroenterol.* 107(1), 145–146 (2012).

53 Ogiwara H, Sugimoto M, Ohno T *et al.* Role of deletion located between the intermediate and middle regions of the *Helicobacter pylori vacA* gene in cases of gastroduodenal diseases. *J. Clin. Microbiol.* 47(11), 3493–3500 (2009).

54 De Francesco, V, Margiotta M, Zullo A *et al. Helicobacter pylori vacA* arrangement and related diseases: a retrospective study over a period of 15 years. *Dig. Dis. Sci.* 54(1), 97–102 (2009).

55 Yamaoka Y, Kwon DH, Graham DY. A M(r) 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. *Proc. Natl Acad. Sci. USA* 97(13), 7533–7538 (2000).

•• **Identifies the role of one outer membrane protein of *H. pylori*, which was denoted as OipA.**

56 Yamaoka Y, Kudo T, Lu H *et al.* Role of interferon-stimulated responsive element-like element in interleukin-8 promoter in *Helicobacter pylori* infection. *Gastroenterology* 126(4), 1030–1043 (2004).

57 Choi IJ, Fujimoto S, Yamauchi K, Graham DY, Yamaoka Y. *Helicobacter pylori* environmental interactions: effect of acidic conditions on *H. pylori*-induced gastric mucosal interleukin-8 production. *Cell Microbiol.* 9, 2457–2469 (2007).

58 Lu H, Wu JY, Beswick EJ *et al.* Functional and intracellular signaling differences associated with the *Helicobacter pylori* AlpAB adhesin from western and east Asian strains. *J. Biol. Chem.* 282(9), 6242–6254 (2007).

59 Kudo T, Lu H, Wu JY *et al.* Pattern of transcription factor activation in *Helicobacter pylori*-infected Mongolian gerbils. *Gastroenterology* 132(3), 1024–1038 (2007).

60 Wu JY, Lu H, Sun Y *et al.* Balance between polyoma enhancing activator 3 and activator protein 1 regulates *Helicobacter pylori*-stimulated matrix metalloproteinase 1 expression. *Cancer Res.* 66(10), 5111–5120 (2006).

61 Lu H, Wu JY, Kudo T *et al.* Regulation of interleukin-6 promoter activation in gastric epithelial cells infected with *Helicobacter pylori*. *Mol. Biol. Cell* 16(10), 4954–4966 (2005).

62 Tabassam FH, Graham DY, Yamaoka Y. OipA plays a role in *Helicobacter pylori*-induced focal adhesion kinase activation and cytoskeletal re-organization. *Cell Microbiol.* 10(4), 1008–1020 (2008).

63 Tabassam FH, Graham DY, Yamaoka Y. *Helicobacter pylori* activate epidermal growth factor receptor- and phosphatidylinositol 3-OH kinase-dependent Akt and glycogen synthase kinase 3beta phosphorylation. *Cell Microbiol.* 11(1), 70–82 (2009).

64 Choi KD, Kim N, Lee DH *et al.* Analysis of the 3′ variable region of the *cagA* gene of *Helicobacter pylori* isolated in Koreans. *Dig. Dis. Sci.* 52(4), 960–966 (2007).

65 Kudo T, Lu H, Wu JY *et al.* Regulation of RANTES promoter activation in gastric epithelial cells infected with *Helicobacter pylori*. *Infect. Immun.* 73(11), 7602–7612 (2005).

66 Tabassam FH, Graham DY, Yamaoka Y. Paxillin is a novel cellular target for converging *Helicobacter pylori*-induced cellular signaling. *Am. J. Physiol Gastrointest. Liver Physiol.* 301(4), G601–G611 (2011).

67 Tabassam FH, Graham DY, Yamaoka Y. *Helicobacter pylori*-associated regulation of forkhead transcription factors FoxO1/3a in human gastric cells. *Helicobacter* 17(3), 193–202 (2012).

68 Graham DY, Opekun AR, Osato MS *et al.* Challenge model for *Helicobacter pylori* infection in human volunteers. *Gut* 53(9), 1235–1243 (2004).

69 Yamaoka Y, Kikuchi S, El-Zimaity HM *et al.* Importance of *Helicobacter pylori oipA* in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. *Gastroenterology* 123(2), 414–424 (2002).

70 Yamaoka Y, Ojo O, Fujimoto S *et al. Helicobacter pylori* outer membrane proteins and gastroduodenal disease. *Gut* 55(6), 775–781 (2006).

71 Ando T, Peek RM, Pride D *et al.* Polymorphisms of *Helicobacter pylori* HP0638 reflect geographic origin and correlate with cagA status. *J. Clin. Microbiol.* 40(1), 239–246 (2002).

72 Dossumbekova A, Prinz C, Mages J *et al. Helicobacter pylori* HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of *hopH* gene polymorphisms. *J. Infect. Dis.* 194(10), 1346–1355 (2006).

73 Lehours P, Menard A, Dupouy S *et al.* Evaluation of the association of nine *Helicobacter pylori* virulence factors with strains involved in low-grade gastric mucosa-associated lymphoid tissue lymphoma. *Infect. Immun.* 72(2), 880–888 (2004).

74 Fujimoto S, Olaniyi OO, Arnqvist A *et al. Helicobacter pylori* BabA expression, gastric mucosal injury, and clinical outcome. *Clin. Gastroenterol. Hepatol.* 5(1), 49–58 (2007).

75 Lu H, Hsu PI, Graham DY, Yamaoka Y. Duodenal ulcer promoting gene of *Helicobacter pylori*. *Gastroenterology* 128(4), 833–848 (2005).

•• **Identifies the role of one gene located in the plasticity zone of *H. pylori*, which was related to the development of duodenal ulcer and was denoted as DupA.**

76 Hussein NR, Argent RH, Marx CK *et al. Helicobacter pylori dupA* is polymorphic, and its active form induces proinflammatory cytokine secretion by mononuclear cells. *J. Infect. Dis.* 202(2), 261–269 (2010).

77 Yamaoka Y. Roles of the plasticity regions of *Helicobacter pylori* in gastroduodenal pathogenesis. *J. Med. Microbiol.* 57(Pt 5), 545–553 (2008).

78 Shiota S, Matsunari O, Watada M, Hanada K, Yamaoka Y. Systematic review and meta-analysis: the relationship between the *Helicobacter pylori dupA* gene and clinical outcomes. *Gut Pathog.* 2(1), 13 (2010).

79 Takahashi A, Shiota S, Matsunari O *et al.* Intact long-type *dupA* as a marker for gastroduodenal diseases in Okinawan subpopulation, Japan. *Helicobacter* 18(1), 66–72 (2012).

80 Jung SW, Sugimoto M, Shiota S, Graham DY, Yamaoka Y. The intact *dupA* cluster is a more reliable *Helicobacter pylori* virulence marker than *dupA* alone. *Infect. Immun.* 80(1), 381–387 (2012).

81 Kersulyte D, Lee W, Subramaniam D *et al. Helicobacter pylori*'s plasticity zones are novel transposable elements. *PLoS ONE* 4(9), e6859 (2009).

82 Fischer W, Windhager L, Rohrer S *et al.* Strain-specific genes of *Helicobacter pylori*: genome evolution driven by a novel type IV secretion system and genomic island transfer. *Nucleic Acids Res.* 38(18), 6089–6101 (2010).

83 Kawai M, Furuta Y, Yahara K *et al.* Evolution in an oncogenic bacterial species with extreme genome plasticity: *Helicobacter pylori* east Asian genomes. *BMC Microbiol.* 11, 104 (2011).

84 McClain MS, Shaffer CL, Israel DA, Peek RM Jr, Cover TL. Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer. *BMC Genomics* 10, 3 (2009).

85 McNamara D. *Helicobacter pylori* infection and the pathogenesis of gastric cancer: a paradigm for host–bacterial interactions 40(7), 504–509 (2008).

86 Kennemann L, Didelot X, Aebischer T *et al. Helicobacter pylori* genome evolution during human infection. *Proc. Natl Acad. Sci. USA* 108(12), 5033–5038 (2011).

87 Ubukata H, Nagata H, Tabuchi T *et al.* Why is the coexistence of gastric cancer and duodenal ulcer rare? Examination of factors related to both gastric cancer and duodenal ulcer. *Gastric Cancer* 14(1), 4–12 (2011).

88 Fukase K, Kato M, Kikuchi S *et al.* Effect of eradication of *Helicobacter pylori* on incidence of metachronous gastric carcinoma after

endoscopic resection of early gastric cancer: an open-label, randomised controlled trial. *Lancet* 372(9636), 392–397 (2008).

89 Kim JJ, Tao H, Carloni E *et al. Helicobacter pylori* impairs DNA mismatch repair in gastric epithelial cells. *Gastroenterology* 123(2), 542–553 (2002).

90 Matsumoto Y, Marusawa H, Kinoshita K *et al. Helicobacter pylori* infection triggers

aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.* 13(4), 470–476 (2007).

91 Toller IM, Neelsen KJ, Steger M *et al.* Carcinogenic bacterial pathogen *Helicobacter pylori* triggers DNA double-strand breaks and a DNA damage response in its host cells. *Proc. Natl Acad. Sci. USA* 108(36), 14944–14949 (2011).

92 Ushijima T, Hattori N. Molecular pathways: involvement of *Helicobacter pylori*-triggered inflammation in the formation of an epigenetic field defect, and its usefulness as cancer risk and exposure markers. *Clin. Cancer Res.* 18(4), 923–929 (2012).

93 Salama NR, Hartung ML, Muller A. Life in the human stomach: persistence strategies of the bacterial pathogen *Helicobacter pylori*. *Nat. Rev. Microbiol.* 11(6), 385–399 (2013).