Sains Malaysiana 53(3)(2024): 605-621
http://doi.org/10.17576/jsm-2024-5303-10

Functional Analysis of *cag* Pathogenicity Island Diversity in *Helicobacter pylori* Isolated from Multi-Ethnic Patients in Malaysia

(Analisis Fungsian Kepelbagaian Pulau Kepatogenan *cag* dalam *Helicobacter pylori* Dipencilkan daripada Pesakit Pelbagai Etnik di Malaysia)

Shaza Azlin Razak[1], Alfizah Hanafiah[1,*], Asif Sukri[2], Hui-Min Neoh[3] & Noraziah Mohamad Zin[4]

[1]*Department of Medical Microbiology and Immunology, Faculty of Medicine, Universiti Kebangsaan Malaysia, 56000 Cheras, Kuala Lumpur, Malaysia*
[2]*Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[3]*UKM Medical Molecular Biology Institute (UMBI), Universiti Kebangsaan Malaysia, 56000 Cheras, Kuala Lumpur, Malaysia*
[4]*Programme of Biomedical Science, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia*

Received: 13 October 2023/Accepted: 9 February 2024

ABSTRACT

*cag*PAI is an important pathogenic marker contributing to disease severity caused by *H. pylori* infection, encoding proteins for type four secretion systems (T4SS) and the pro-inflammatory cytokines (IL-8) induction in infected host. The heterogeneity of *cag*PAI genes results in different clinical outcomes in *H. pylori* infection. The aim of this study was to investigate the functional diversity of *cag*PAI genes in clinical *H. pylori* strains isolated from patients with gastroduodenal diseases. Among the isolates investigated, 6 out of 27 (22.2%) harbor intact *cag*PAI, 3 strains (11.1%) have complete absence of *cag*PAI, and 18 strains (66.7%) contain a partially deleted island. IL-8 concentration is observed significantly higher among *cag*PAI-positive isolates ($P = 0.045$). Of note, IL-8 secretion in the gastric cells infected with *H. pylori* strains isolated from Malay patients was significantly higher than the gastric cells infected with the strains from non-Malay patients ($P = 0.046$). The secretion of IL-8 appears in similarly high among intact and partially deleted *cag*PAI isolates. East Asian *cagA* isolates induce a notably higher level of IL-8 than that of Western *cagA*. Inflammatory activities are observed and IL-8 concentration in the milder stage was found in significantly higher quantities than in the severe stage ($P = 0.027$). The 'hummingbird' phenotype was observed to cause severe effects in infected cells with *cag*PAI-intact *H. pylori*. The genetic variability of *cag*PAI isolates from multi-ethnicity patients may have unique implications for disease outcomes with respect to proinflammatory secretion.

Keywords: CagA; *cag*PAI; *Helicobacter pylori*; 'hummingbird' phenotype; interleukin-8

ABSTRAK

*cag*PAI merupakan penanda patogen yang penting dalam mempengaruhi keterukan penyakit akibat jangkitan *Helicobacter pylori*, yang mana ia mengekod sistem perembesan jenis empat (T4SS) serta rembesan pro-radang sitokin (IL-8) dalam hos. Perbezaan dalam hasil klinikal kerana jangkitan *H. pylori* adalah akibat daripada kepelbagaian gen *cag*PAI. Tujuan penyelidikan ini adalah untuk mengkaji variasi kefungsian gen *cag*PAI bagi strain klinikal *H. pylori* daripada pesakit gastrousus. Keutuhan *cag*PAI yang lengkap dikesan terhadap 6 daripada 27 strain (22%) kajian, manakala 3 strain (11.1%) tidak mempunyai *cag*PAI dan 18 strain (66.7%) mengandungi *cag*PAI separa lengkap. Kepekatan IL-8 adalah lebih tinggi setelah dijangkiti pencilan positif-*cag*PAI ($P = 0.045$). Tambahan lagi rembesan IL-8 akibat jangkitan isolat *H. pylori* daripada pesakit Melayu adalah secara signifikannya lebih tinggi berbanding infeksi terhadap sel gastrik dalam kalangan pesakit bukan Melayu ($P = 0.046$). Rembesan IL-8 pencilan yang utuh

*cag*PAI dengan *cag*PAI tidak lengkap adalah pada aras yang tinggi. Kepekatan IL-8 akibat jangkitan isolat *cagA* Timur Asia dilihat lebih tinggi berbanding isolat *cagA* Barat. Pemerhatian terhadap aktiviti keradangan mendapati kepekatan IL-8 pada peringkat yang lebih ringan adalah nyata sekali lebih banyak berbanding peringkat yang teruk ($P = 0.027$). Jangkitan *H. pylori* yang utuh *cag*PAI dilihat menyebabkan fenotip 'hummingbird' yang lebih teruk terhadap sel. Kepelbagaian genetik isolat *cag*PAI daripada pesakit berbilang etnik berkemungkinan memberi implikasi unik berkenaan dengan rembesan proinflamasi terhadap hasil penyakit.

Kata kunci: CagA; *cag*PAI; fenotip 'hummingbird'; *Helicobacter pylori*; interleukin-8

Introduction

Helicobacter pylori is a pathogen known to infect the human stomach and causing gastrointestinal diseases ranging from gastritis to gastric cancer (Hatakeyama 2009; Peeks Jr. & Crabtree 2006). The prevalence of *H. pylori* varies depending on the geographical and socio-economic region. *H. pylori* prevalence was reported to be higher in developing countries (85% - 95%) compared with developed countries (30% - 50%) (Burucoa & Axon 2017; Hunt et al. 2011). According to Goh (2018), the prevalence of *H. pylori* infections in Malaysia varied for different ethnicities: Malays (25.9%), Chinese (48.5%), and Indians (61.8%). *H. pylori* strains which harbor virulence factors, including *vacA*, *babA*, *iceA*, and *cagA* are associated with higher disease severity in infected patients than strains without such virulence factors. One of the important virulent determinants to influence clinical outcomes is the cytotoxin-associated gene pathogenicity island (*cag*PAI). The *cag*PAI is a chromosomal region with a size estimation of 40 kbp, containing ~ 30 genes (Terry et al. 2005) which are responsible for the translocation of oncoprotein CagA into gastric epithelial cells via a type four secretion system (T4SS) (Tegtmeyer, Wessler & Backert 2011) and interleukin-8 (IL-8) induction (Cover, Lacy & Ohi 2020).

A few studies have shown *cag*PAI diversity across different geographical populations (Olbermann et al. 2010; Yakoob et al. 2009) and showed that patients infected with *H. pylori* harboring intact *cag*PAI are at a higher risk for contracting severe gastric diseases, including gastric cancer. *H. pylori* strains harboring intact *cag*PAI induce more severe inflammation in gastric mucosa and are usually regarded as more pathogenic than strains with partial or complete *cag*PAI deletion (Nilsson et al. 2003; Patra et al. 2011; Rizzato et al. 2020).

One of the crucial outcomes in *H. pylori* infection is the production of proinflammatory cytokines, namely IL-8 (Allison et al. 2009; Yamaoka 2010). Production of proinflammatory cytokines, particularly IL-8, is usually the hallmark of *H. pylori* infection. The secretion of IL-8 is often involved with the NF-κB signalling pathway, which is crucial to epithelial cell integrity and possible metastasis. Reports from previous studies have shown that in the presence of *cag*PAI-positive strains, infected gastric epithelial cells produced high levels of proinflammatory cytokines including IL-8 (Guillemin et al. 2002; Mueller et al. 2004). The level of IL-8 induced by *H. pylori* strains may have a correlation with *cag*PAI region diversity, hence playing crucial role with respect to the extent of gastric mucosal damage and disease outcomes.

CagA is divided into Western and East Asian types, depending on the amino acid sequences at the C-terminal known as EPIYA motifs. A few studies have reported severe clinical outcomes due to infection with *H. pylori* harboring Western-type CagA with higher number of EPIYA-C repeats (Argent et al. 2004; Khaledi et al. 2020). In contrast to Western-type CagA, East Asian CagA contains only a single EPIYA-D at the end of the motif, yet it impacts the degree of severity in infected gastric tissues (Xia et al. 2009). A few studies (Argent et al. 2004; Keikha & Karbalaei 2021) have primarily suggested that the presence of multiple repeats of EPIYA-C in Western-type CagA is associated with worse severe sequelae than the presence of single EPIYA-C repeat. Unlike Western-type CagA, distinguishing clinical manifestations in patients infected with *H. pylori* harboring East Asian type EPIYA is complicated since it contains only a single EPIYA-D, and most cases change from simple gastritis to gastric cancer (Xia et al. 2009). It has also been suggested by Crew and Neugut (2006) that East Asian CagA might have a higher potential to cause worse severity in gastric diseases, including gastric cancer than Western-type CagA. CagA is also important for the secretion of IL-8 through the phosphorylation of CagA protein, although the mechanism of IL-8 secretion can also occur through

the non-phosphorylation mechanism of CagA (Brandt et al. 2005; Wang et al. 2023). Consequently, the presence of the *cagA* gene has been associated with higher grades of inflammation, which may lead to the development of the most severe gastrointestinal diseases. According to a couple of studies by Censini et al. (1996) and Markovska et al. (2018), the *cagA* gene is closely related to intact *cag*PAI, yet the existence of the *cagA* gene does not guarantee the intactness of *cag*PAI.

The 'hummingbird' effect or phenotype is characterized by the elongation and scattering of gastric epithelial cells upon infection with *H. pylori*. This phenomenon occurs when phosphorylated EPIYA motifs interact with cells' SH2-containing phosphatase (SHP2), resulting in focal adhesion kinase (FAK) inactivation (Suzuki et al. 2005; Takahashi-Kanemitsu, Knight & Hatakeyama 2020; Tsutsumi et al. 2003), and hence leading towards the elongation and scattering of infected cells. The quantity and type of CagA EPIYA motifs play important roles in impacting the degree of morphological change in infected gastric cells (Alfizah & Ramelah 2012; Chang et al. 2016).

The variation of *cag*PAI, as well as its association with clinical outcomes that are responsible for IL-8 induction, has not been extensively studied with respect to the multi-ethnic population in Malaysia. The aim of this study was to evaluate the functional capabilities of *cag*PAI genes involved in IL-8 induction and 'hummingbird' phenotype formation among Malaysian *H. pylori* isolates.

Materials and Methods

Cultivation of *H. pylori* on culture media

H. pylori isolates from frozen stock were cultivated onto Columbia blood agar supplemented with sheep blood (7%) under microaerophilic conditions at 37 °C for 3 - 5 days. Culture of *H. pylori* was confirmed by morphology observation and biochemical tests. Successfully grown strains were used for genomic DNA extraction (Hanafiah et al. 2020), IL-8 assay, and 'hummingbird' phenotype examination.

*cag*PAI gene and *cagA* EPIYA motif characterization

The detection of *cag*PAI genes in *H. pylori* strains was conducted using polymerase chain reaction (PCR) using primers listed and protocol as described in study by Hanafiah et al. (2020). Prior to individual *cag*PAI gene detection, the presence of *cag*PAI was determined by existence of flanking genes (Olbermann et al. 2010). *cag*PAI-positive strains underwent subsequent PCRs to further identify all *cag*PAI genes with primers as described in previous research (Olbermann et al. 2010; Ta et al. 2012). Upon confirmation, the products were run on agarose gel and visualized with gel documentation. All materials and instruments used were mentioned in the previous study (Hanafiah et al. 2020).

Molecular amplification of *cagA* EPIYA motifs and categorization were conducted using selected primer sets as previously published article (Jones et al. 2009), presented in Table A.1. Determination of the number of nucleotide sequences that encoded the EPIYA motifs as well as the types of motifs were carried out using the forward primer cag2 (Rudi et al. 1998) and the reverse primers cagA-P1C, cagA-P2CG, cagA-P2TA, cagA-P3E (Azuma et al. 2002) and cagA-pD(R) (Jones et al. 2009). Each PCR reaction mixture contained mastermix (Lucigen, Wisconsin, USA), 10 µml of each primer, 10 ng/µL of extracted DNA and RNAse free water. Each mixture was incubated at 95 °C for 3 min, followed by 30 cycles at 95 °C for 30 s, 57 °C for 60 s, and 72 °C for 30 s and a final extension at 72 °C for 5 min (Argent, Zhang & Atherton 2005). PCR products were electrophoresed on a 1.5% (wt/vol) agarose gel.

According to previous studies (Argent, Zhang & Atherton 2005; Jones et al. 2009), EPIYA-A motif is identified by amplification of primers cag2 with cagA-P1C, whereas amplification of primers cag2 with a 1:1 mixture of cagA-P2TA and cagA-P2CG represents EPIYA-B motif. Amplification with primers cag2 and cagA-P3E identifies as either EPIYA-C or EPIYA-D motif, hence additional cag2 with cagA-pD(R) primers to confirm the presence of EPIYA-D motif. Identification of EPIYA-A, -B and -C motifs combination signifies *H. pylori* strain containing Western CagA, whereas East-Asian CagA contains a combination of EPIYA-A, -B and -D motifs. CagA without the existence of either EPIYA-C or -D were characterized as undefined CagA. The identifications were further confirmed by gene sequencing, in which the DNA sequences were analysed and blasted against NCBI database. Instruments used were similar to study by Hanafiah et al. (2020).

Gastroadenocarcinoma cell culture

Human gastric adenocarcinoma cell line (AGS) (RRID: CVCL_0139) was grown in RPMI 1640 media (Sigma Aldrich, Massachusetts, USA) supplemented with FBS (10%) (Gibco, New York, USA) and antibiotics (Gibco,

New York, USA). Incubation was conducted at 37 °C with 5% CO_2 until the culture reached confluence of 90%. Cultivation was carried out in T-25 flask prior to experimentation. Cell viability was measured and standardized prior to experimentation.

H. pylori INFECTION OF AGS CELLS

AGS cells were grown until 90% confluence was achieved. The cell culture was washed twice with pre-warmed PBS and detached from the flask using pre-warmed TrypLE Express (Gibco, New York, USA). The enzymatic reaction was deactivated by introducing pre-warmed media into the mixture. Concentration of viable AGS cells was calculated using hemocytometer.

The cell suspension then transferred to 6-well plate at starting cell concentration of 6.0×10^5/mL in each well, and was cultured in media under incubation conditions until 90% confluence was achieved. Twenty-four hours prior to bacterial infection, the media were replenished without FBS or antibiotics. *H. pylori* cells at logarithmic phase were suspended in serum- and antibiotic-free media and used for AGS cell infection at multiplicity of infection (MOI) of 100 for 24 h at 37 °C.

IL-8 DETECTION AND QUANTIFICATION

One mL of media suspension from the flask was centrifuged at 1000 rpm, for 20 min at 4 °C. The supernatant was then transferred into new sterile tube and kept at -80 °C until analysis. Detection and quantification of IL-8 concentration for each lysate was conducted using enzyme-linked immunosorbent assay (ELISA) kit (Fine Test, Beijing, China) according to the manufacturer's protocol. Each sample was run in triplicates and AGS cells without infection served as negative control.

'HUMMINGBIRD' PHENOTYPE OBSERVATION

Morphological changes of AGS cells infected with *H. pylori* strains were observed after 24 h of infection. Elongation of cells was identified and calculated as percentage, wherein number of elongated AGS cells upon total number of AGS cells was measured against 100 percent, and numbers were observed under inverted microscopy at magnification of 400×. Comparison was made between non-infected and infected AGS cells of *H. pylori* with different *cag*PAI statuses. Each experiment was conducted in duplicate.

HISTOPATHOLOGICAL EXAMINATION

The biopsies were fixed in formalin and paraffin-embedded, and sections were cut and stained with hematoxylin-eosin. The sections were also stained with Warthin-Starry for superior *H. pylori* visualization, when necessary. All gastric biopsies were scored based on the updated Sydney System (0, none; 1, mild; 2, moderate; 3, marked) (Dixon et al. 1996).

STATISTICAL ANALYSIS

Statistical analysis was carried out using SPSS software version 23 (SPSS Inc, Illinois, USA). Differences in IL-8 concentration between samples were analyzed using Student's t-test. Data were reported as mean ± standard error measurement (SEM). Chi-square test or Fischer's exact test was used to analyze data comparing *cag*PAI intactness, CagA types, patient's ethnicity and disease severity. One-way ANOVA was performed to analyse association of IL-8 secretion with the *cag*PAI gene's existence, CagA categories and histopathological examination. Kruskal-Wallis test with multiple pairwise comparisons via Mann-Whitney U test were used to analyse the association of 'hummingbird' phenotype induction with *cag*PAI integrity status, CagA types and histopathological examination. *P* value less than 0.05 were considered statistically significant.

RESULTS AND DISCUSSION

H. pylori SAMPLES AND *cag*PAI GENOTYPING

A total of 27 *H. pylori* strains were used in this study. The strains were isolated from patients with different ethnicity: Indians (n = 12, 44.4%), Chinese (n =10, 37%), Malays (n = 4, 14.8%) and other ethnicities (n = 1, 3.8%). Most of the patients were diagnosed with gastritis (n = 13, 48.1%), chronic gastritis (n = 12, 44.4%), and peptic ulcers (n = 2, 7.4%).

Regarding *cag*PAI integrity, six isolates (22.2%) were found to harbor intact *cag*PAI, three (11.1%) of the *H. pylori* isolates had a complete absence of *cag*PAI, while the remaining 18 isolates (66.7%) contained a partially deleted island. Most gene deletion occurred in the strains were *cag2* (n = 15, 55.6%) and *cag14* (n = 15, 55.6%), followed by deletion in *cag16* (n = 9, 33.3%). Identification of individual *cag*PAI genes were illustrated previously (Hanafiah et al. 2020) via gel electrophoresis, presented in Figure A.1 (Appendices). The *cag*PAI-

positive isolates were further characterized based on CagA EPIYA motifs. Eight of these strains (29.6%) were identified as Western type, ten strains (37%) were of East Asian type, and six (22.2%) were of undefined type due to the lack of a third EPIYA motif for classification. Among the undefined strains, five presented with the AB motif and only one strain had the ABB EPIYA motif. A representative result of EPIYA motifs PCR amplification is shown in Figure A.2 (Appendices). A description on the CagA types among isolates with different *cag*PAI statuses is shown in Table 1.

IL-8 QUANTIFICATION UPON *H. pylori* INFECTION TO AGS CELLS

IL-8 concentration measurement was conducted upon *H. pylori* strain infections in AGS cells. The standard curve using IL-8 standard dilutions was determined based on the manufacturer's instructions. The amount of IL-8 induced upon infection in this study ranged from 160.60 pg/mL to 656.60 pg/mL. The positive control used for the experiment is the supernatant of infected AGS cells by standard *H. pylori* strain (ATCC 43526), whereas the negative control is the supernatant of uninfected or unstimulated AGS cells. The measurement of negative control in average is 24.71 pg/mL.

Figure 1 shows the mean IL-8 concentration secreted upon infection with *H. pylori* isolates with different *cag*PAI integrity statuses. The mean concentration of IL-8 secreted due to infection with strains of intact *cag*PAI was 432.00 (± 81.43) pg/mL whereas for infection with partially deleted *cag*PAI was 434.91 (± 34.36) pg/mL, and mean IL-8 production by negative-*cag*PAI strains was 239.20 (± 32.27) pg/mL. Analysis upon infection of *cag*PAI-positive *H. pylori* in AGS cells showed that they appeared to be producing a significantly higher concentration of IL-8 compared to those from infection with *cag*PAI-negative isolates (P = 0.045). Range differences in IL-8 concentrations secreted between the infection of intact or partially deleted *cag*PAI were minimal, in with the latter secreting a slightly higher concentration.

TABLE 1. Detection of CagA EPIYA type in *H. pylori* isolates of various *cag*PAI integrity statuses

*cag*PAI status	CagA type, n (%)			
	Western	East Asian	Undefined	Negative
Complete	2 (25)	1 (10)	3 (50)	-
Partially deleted	6 (75)	9 (90)	3 (50)	-
Negative	-	-	-	3 (100)

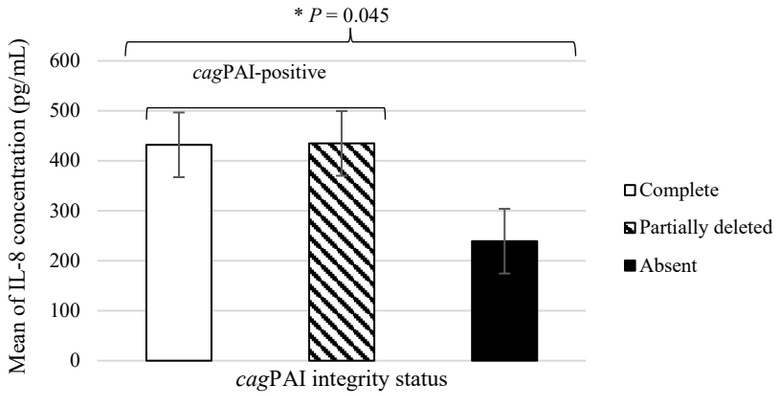

FIGURE 1. IL-8 concentration secreted by AGS cells infected with *H. pylori* strains with different *cag*PAI statuses. IL-8 concentration was measured using ELISA kit. Mean IL-8 concentration comparison of *H. pylori* with different *cag*PAI integrity status groups; *cag*PAI-positive group consisted of both complete and partially deleted *cag*PAI. The mean ± SEM of triplicate results are presented. A Student's t-test was performed for statistical evaluation between *cag*PAI-positive and -negative strains upon IL-8 secretion. * There was a significant difference in IL-8 production between infection with *cag*PAI-positive and -negative *H. pylori*

The concentrations of IL-8 were also measured and analysed according to existence of the *cag*PAI's individual genes upon infection against AGS cells, with the observations displayed in Table A.2 (Appendices). The levels of IL-8 were significantly higher with the existence of *cag1*/*cagζ*, *cag5*/*cagβ*, *cag6*/*cagZ*, *cag8*/*cagX,* and *cag21*/*cagG* ($P = 0.045$). Furthermore, infection caused by strains with positive *cag17*/*cagN*, *cag18*/*cagL,* and *cag22*/*cagF* has resulted to an elevation of IL-8 compared to strains that lack of the genes ($P = 0.029$). Interestingly, *H. pylori* isolates which were harboring *cag7*/*cagY* and *cag19*/*cagI* exhibited an increased secretion of IL-8 upon infection, with their significance values at $P = 0.025$ and $P = 0.009$, respectively. Despite no significant differences were observed, *cag*PAI genes that were prone towards deletions including cag2/*cagε*, *cag14*/*cagQ* and *cag16*/*cagM* showed lack of impact towards IL-8 secretion upon AGS cells.

The differences in IL-8 concentrations produced by AGS cells upon infection of *H. pylori* strains that varied with *cagA* are illustrated in Figure 2. Infection with East Asian *cagA* induced a higher IL-8 concentration compared with the Western *cagA*, while *cagA* that lacked a third EPIYA exhibited a lower secretion of IL-8 among infected cells. In addition, there was no statistical correlation ($P = 0.131$) between *cag*PAI status or *cagA* type and disease severity with regard to IL-8 secretion.

An evaluation of IL-8 concentrations induced upon infection with *H. pylori* isolated from patients of different ethnicities was conducted, and this analysis is shown in Table 2. There was a significant difference in the IL-8 concentration induced by *H. pylori* isolated from Malay patients compared with that non-Malay patients ($P = 0.046$). IL-8 was secreted in greater quantities among patients with acute infections (446.8 pg/mL) compared with those with chronic infections (392.62 pg/mL) or peptic ulcers (312.50 pg/mL), although there was no statistical correlation observed ($P = 0.480$). The majority of gastritis cases (84.6%, n = 11) were infected with *H. pylori* strains harboring incomplete *cag*PAI, whereas almost half (41.7%, n = 5) of the chronic gastritis patients were infected by intact *cag*PAI, and the peptic ulcer cases were infected by partially deleted strains.

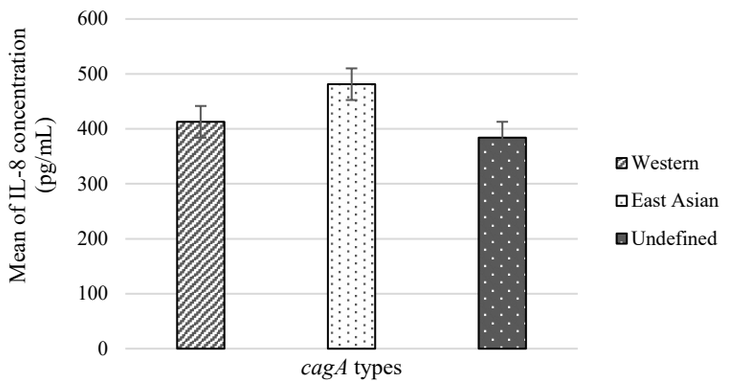

FIGURE 2. Quantification of IL-8 concentration induced by AGS cells infected with *H. pylori* of variouses *cag*A classifications. Characteristics of *cagA* were based on the third EPIYA motif detected; results are presented as mean ± SEM of triplicate. A one-way ANOVA was performed to compare the impact of *cagA* variations on IL-8 secretion. No statistical differences were observed between IL-8 concentration and *cagA* types ($P = 0.131$)

TABLE 2. Differences in *cag*PAI integrity statuses and IL-8 concentrations upon infection with *H. pylori* strains from different patients' ethnicities. A Student's t-test was carried out to compare IL-8 concentrations among the Malay and non-Malay infected patients. * There was a significant difference in IL-8 concentrations induced by *H. pylori* strains from the Malay and non-Malay patients ($P = 0.046$)

Ethnicity, n (%)	*cag*PAI status			IL-8 mean concentration (pg/mL) ± SE
	Complete	Partially deleted	Negative	
Malay	2 (50)	2 (50)	0	558.83 ± 47.06*
Non-Malay	4 (17)	16 (70)	3 (13)	387.07 ± 32.69

The association between IL-8 induction to mononuclear infiltration and neutrophil activity through the histopathological scores of patients' gastric mucosa was analysed, and the results are displayed in Table 3. Similar trends of lower IL-8 secretion were seen with higher or severe scorings in both histopathological examinations, whereas higher IL-8 amounts were detected with lower scorings. Based on the data, gastric mucosa with a mild score for mononuclear infiltration showed elevated concentrations of IL-8 compared with those with a severe score ($P = 0.027$). Further, gastric mucosa with a mild score for mononuclear cell infiltration showed elevated IL-8 concentrations compared with those with a higher/severe score ($P = 0.027$). No significant relationship was observed between the histological scores and disease severity.

'HUMMINGBIRD' PHENOTYPE FORMATION

Morphological changes upon *H. pylori* infection in AGS cells were analyzed through the observation of the hummingbird phenotype. An illustration of *cag*PAI diversity and the elongation of the infected AGS cells is shown in Figure 3. The infection of strains harboring *cag*PAI appeared to cause more hummingbird formation among infected AGS cells compared with *cag*PAI-negative strains ($P = 0.005$). Interestingly, *H. pylori* strains which contained intact *cag*PAI yielded more severe effects compared with strains with partially deleted *cag*PAI ($P = 0.014$) as shown in Figure 4. *cag*PAI-positive strains with many gene deletions were observed to have the lowest percentage of cell elongation.

TABLE 3. Distribution of *H. pylori cag*PAI integrity according to histopathological scoring of infected gastric mucosa, as well as IL-8 secretion and 'hummingbird' phenotype formation. A one-way ANOVA was executed to analyze differences in IL-8 concentrations between scorings in histopathological evaluations. The scoring was done according to the updated Sydney System (Dixon et al. 1996), in which the higher the score the worse the inflammation in response to infection. *Statistical difference was observed in IL-8 secretion between mild and severe scorings of mononuclear cell infiltration ($P = 0.027$). **Statistical difference was observed in 'hummingbird' phenotype between mild and severe scorings of mononuclear infiltration

Histopathological examination	Score	*cag*PAI, n (%)			IL-8 mean concentration (pg/mL) ± SE	'hummingbird' phenotype (%) ± SE
		Intact	Partial	Absent		
Mononuclear cell infiltration	0	0	1 (5.6)	0	459.9 ± 0	13.0 ± 0
	1	2 (33.3)	3 (16.7)	0	543.26 ± 50.64 *	27.9 ± 6.8 **
	2	3 (50)	11 (61.1)	3 (100)	392.89 ± 40.3	19.1 ± 3.1
	3	1 (16.7)	3 (16.7)	0	320.73 ± 62.6	35.3 ± 2.8
Total score	Mean ± SE	1.83 ± 0.31	1.89 ± 0.18	2.0 ± 0		
Neutrophil activity	0	1 (16.7)	2 (11.1)	1 (33.3)	491.63 ± 114.7	21.6 ± 9.1
	1	3 (50)	12 (66.7)	1 (33.3)	415.33 ± 43.15	22.0 ± 2.9
	2	1 (16.7)	4 (22.2)	1 (33.3)	368.43 ± 59.16	23.3 ± 6.8
	3	1 (16.7)	0	0	347.4 ± 0	40.5 ± 0
Total score	Mean ± SE	1.33 ± 0.42	1.11 ± 0.14	1.0 ± 0.58		

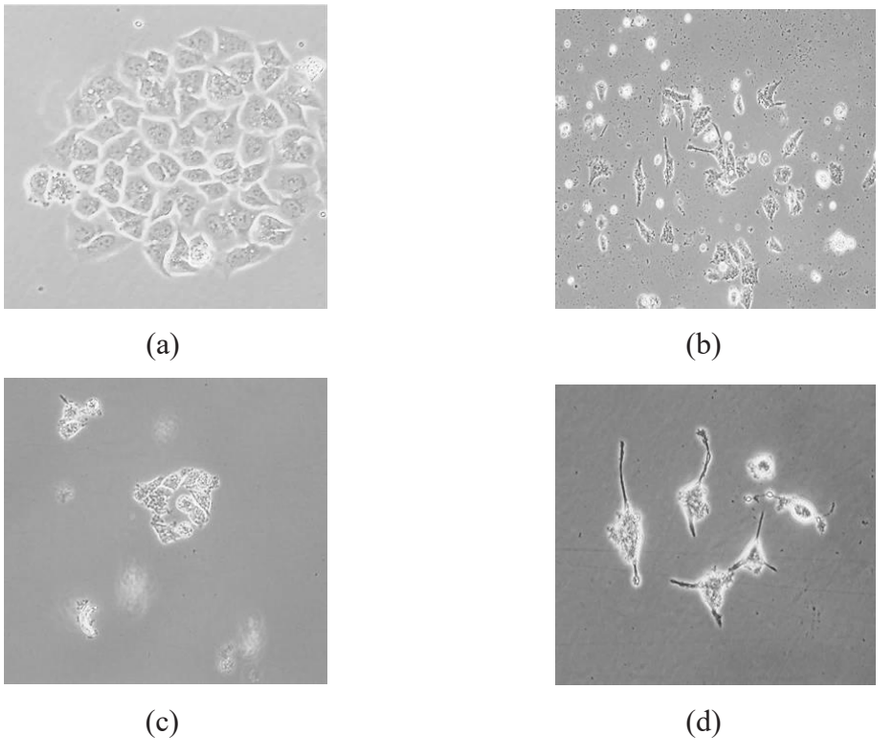

FIGURE 3. Microscopic images of 'hummingbird' phenotype of AGS cells upon *H. pylori* infection. The images are presented at 200x magnification except image D. (a) Uninfected AGS cells. (b) 'Hummingbird' formation in AGS cells after infection with *cag*PAI-positive *H. pylori* strains. (c) No 'hummingbird' effect on AGS morphology upon infection with *cag*PAI-negative *H. pylori*. (d) 'hummingbird' phenotype in AGS cells at 400x magnification. A Mann-Whitney U test was conducted for statistical analysis. A significant difference in 'hummingbird' formations was observed between *cag*PAI-positive and -negative *H. pylori* strains ($P = 0.005$)

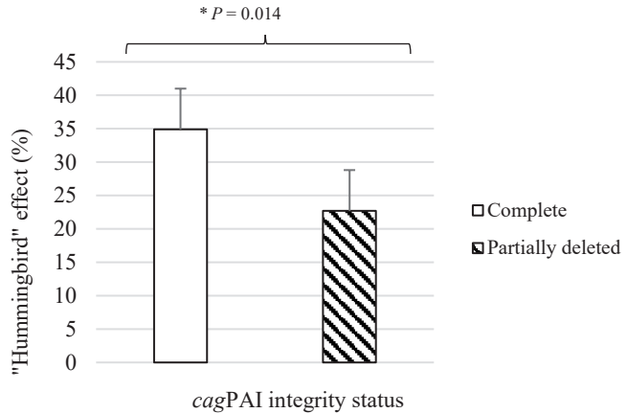

FIGURE 4. Percentage of 'hummingbird' phenotype observed in AGS cells upon *cag*PAI-positive *H. pylori* infection with different integrity status. Mean 'hummingbird' effect comparison between *H. pylori* with complete and partially deleted *cag*PAI. A Mann-Whitney U test was performed to compare 'hummingbird' phenotype between complete- and incomplete-*cag*PAI infection. *Infection by complete *cag*PAI *H. pylori* induced significantly greater 'hummingbird' effect compared to partially deleted *cag*PAI ($P = 0.014$)

Evaluations of 'hummingbird' phenotype according to *H. pylori* isolate's individual *cag*PAI genes were carried out as well, with the comparisons displayed in Table A.2 (Appendices). The existence of *cag16/cagM* and *cag24/cagD* had shown impact towards the 'hummingbird' formation in infected AGS cells with their significant values of $P = 0.025$ and $P = 0.014$, respectively. There were no significant differences observed on the existence of other *cag*PAI genes upon formation of 'hummingbird' effect.

With regard to CagA, there was no significant difference in 'hummingbird' formation between strains with EPIYA motifs of Western and East Asian types ($P = 0.894$). The correlation between the 'hummingbird' effect and the histopathological scores is displayed in Table 3. The histopathological scores of mononuclear infiltrations displayed a significant impact on cell elongation with severe scoring ($p = 0.030$), but no statistical difference upon other analyses. In addition, no significant relationship was observed between the 'hummingbird effect' and disease severity ($P = 0.181$).

The *cag*PAI is an important virulent factor that contributes to the severity outcomes of infected patients. The integrity of *cag*PAI implies the existence of genes in the pathogenicity island wherein intact *cag*PAI indicates the existence of all *cag* genes, while incomplete or rearranged *cag*PAI indicates the deletion of at least one *cag* gene and a *cag*PAI-negative signifies the absence of *cag*PAI entirely.

The integrity of *cag*PAI has been associated with an increased risk of developing gastric cancer (Khatoon et al. 2017; Waskito et al. 2018). Malaysia is a multi-cultural country that consists of three major ethnicities, namely Malays, Malaysian Indians and Malaysian Chinese. *H. pylori* infection differs among major ethnicities in Malaysia; the inf ection is the highest in Malaysian Indians and followed by Malaysian Chinese, while the lowest infection rates are among the Malays. Although *H. pylori* infection is the highest in Malaysian Indians, the incidence of gastric cancer is low in this population, and this phenomenon has been described as the 'Indian enigma' (Goh et al. 2007). The presence of *cag*PAI in *H. pylori* has been associated with a more severe degree of stomach disease than *H. pylori* with an absence of *cag*PAI (Markovska et al. 2018). In our previous study, we demonstrated the significant differences in composition of *cag*PAI genes among *H. pylori* isolates from different ethnicities in Malaysia as well as the contribution of *cag*PAI intactness to the severity of gastric diseases (Hanafiah et al. 2020). We extended our previous study to determine the impact of *cag*PAI intactness in *H. pylori* isolates from multi-ethnic populations in Malaysia on IL-8 secretion and 'hummingbird' phenotype formation in AGS cells. There is lack of studies conducted in Malaysia on the functional properties of *H. pylori cag*PAI despite the disproportionate incidence of gastric cancer among the three major ethnicities in the country.

The intactness and rearrangement of *cag*PAI genes depend on the geographical region from which *H. pylori* is isolated. Most *cag*PAI-positive *H. pylori* strains in this study possessed a rearranged or partially deleted *cag*PAI, which contrasted with most studies that have emphasized the conservation of intact *cag*PAI. The majority of the strains in our study showed partially deleted *cag*PAI genotype (61.9%), a similar to those of previous studies (Ahmadzadeh et al. 2015; Alfizah et al. 2013; Hanafiah et al. 2020). As explained in a previous study (Hanafiah et al. 2020), the lack of intactness in the *cag*PAI within these strains was due to different gene detection method having been used in the other study (Fazeli et al. 2017). As in other studies, *cag*PAI might not be conserved as deletions within islands can occur, and socio-demographic and geographical factors also play essential roles in the variation of *cag*PAI diversity in *H. pylori* (Olbermann et al. 2010; Ta et al. 2012). In Malaysia's neighbouring island country, Indonesia, complete *cag*PAI was observed in 55.8% of *H. pylori* isolates (Waskito et al. 2018), while in Vietnam 94% of *H. pylori* isolates harbored complete *cag*PAI (Phuc et al. 2021). As in this study, *H. pylori* isolates from the study conducted in Indonesia (Waskito et al. 2018) were collected from different ethnicities, while the study conducted in Vietnam by Phuc et al. (2021), *H. pylori* isolates were collected from only one ethnicity. Thus, ethnicity plays an essential role in determining the heterogeneity of *cag*PAI in *H. pylori* isolates.

The possibility of IL-8 induction caused by *H. pylori* strain infection is attributed to the genomic variability that *cag*PAI possesses. Several studies have reported a correlation between *cag*PAI integrity and IL-8 production, wherein intact *cag*PAI induced higher secretion while partially deleted island caused a reduction in IL-8 secretion (Nguyen et al. 2010; Nilsson et al. 2003; Schmidt et al. 2010). In the present study, concentrations of IL-8 were observed to be higher among *cag*PAI-positive strains than among *cag*PAI-negative strains. However, there was no significant difference in

IL-8 induction with the infection of *H. pylori* between complete and incomplete/rearranged *cag*PAI. As in a report by Salih et al. (2014), the level of IL-8 was similarly high in both intact and partially deleted *cag*PAI infections. In this study, most of the partially deleted *cag*PAI strains were harboring East Asian *cagA*, suggesting that East Asian *cagA* might contribute to an elevated IL-8 level in infected cells. This observation could indicate that deletions within *cag*PAI in the strains may not have a severe impact on the epithelial cell response towards infection. In addition, there is also the possibility of other factors influencing bacterial virulence. Aside from CagA translocation into gastric epithelial cells, the production of IL-8 by *H. pylori*-infected gastric epithelial cells can also occur through the translocation of peptidoglycan into the gastric cells (Lamb & Chen 2010) and the expression of outer inflammatory protein A (OipA) that mediates CagA translocation into the cells (Horridge et al. 2017). In addition, the heterogeneity of the T4SS components of *cag*PAI is also pertinent to the stimulation of IL-8 in gastric epithelial cells (Choi et al. 2021). Further studies examining these matters should be conducted.

cagA is an important virulence marker within *cag*PAI that is attributed to disease pathogenesis. The EPIYA-repeat region of CagA in Western *H. pylori* isolates is an arrangement of EPIYA-A, EPIYA-B, and EPIYA-C segments (A–B–C-type CagA), whereas CagA from East Asian *H. pylori* isolates also possesses EPIYA-A and EPIYA-B segments as well as having a distinct EPIYA-D segment instead of EPIYA-C. Upon the translocation of the effector protein into the host cell, phosphorylation occurs in the CagA EPIYA motifs, which leads to the induction of proinflammatory cytokines, including IL-8 (Brandt et al. 2005; Wang et al. 2023). Previous studies have found that infection with *H. pylori* strains containing EPIYA-D induced a higher level of IL-8 compared with Western-type strains (Argent et al. 2008; Hatakeyama 2004), hence leading to severe gastric inflammation and disease (Higashi et al. 2002). *H. pylori* harboring East Asian EPIYA motifs are mostly isolated to areas with a high incidence of gastric cancer, such as China (Xue et al. 2021). Despite no significant difference being observed, the results of this study demonstrated the possible severity of East Asian-type strains since the IL-8 concentration in these strains was higher than that of Western *cagA*. This result supports the importance of EPIYA motif types in impacting disease severity.

Interestingly, one of the peptic ulcer cases infected by *H. pylori* with East Asian CagA presented with a higher level of IL-8 than the undefined CagA (EPIYA AB) of the same disease. Compared with EPIYA A and EPIYA B, EPIYA C and EPIYA D have a greater impact on tyrosine phosphorylation, thus inducing higher IL-8 secretion in infected gastric mucosal cells (Higashi et al. 2002; Schneider et al. 2009). Although we ascertained the importance of EPIYA in IL-8 induction, other factors might also play a significant role in pathogenesis of *H. pylori* infection.

The translocation of oncogenic CagA protein into gastric epithelial cells results in the phosphorylation of CagA with intracellular signaling proteins, disturbing the cellular signaling process and eventually leading to epithelial-mesenchymal transition-like cells with increased migration and invasiveness (Baj et al. 2020). This phenomenon is known as the 'hummingbird' phenotype. *H. pylori* expresses 'needle-like' T4SS machinery to translocate CagA into the cells (Cover, Lacy & Ohi 2020). We observed that the percentage of 'hummingbird' phenotype formations in cells infected with *H. pylori* possessing intact *cag*PAI was significantly higher than that in cells infected with *H. pylori* harboring the partial deletion of *cag*PAI. Our results indicate that *cag*PAI with complete genes are essential to inducing more dramatic cell scattering than that of *cag*PAI with gene deletion. Consistent with previous findings by Sukri et al. (2022), no significant difference was observed in 'hummingbird' phenotype formation between Western and East Asian *H. pylori* strains.

Histopathological examinations showed prominent neutrophilic and mononuclear cell infiltration to infected gastric mucosa. The activation and migration of neutrophils from blood vessels towards the sites of infections during the innate response (Koller et al. 2009), eventually led to mononuclear immune cell activity. Previous studies have shown that neutrophilic infiltration in gastric mucosa is attributed to *H. pylori* infection, and the severity of mucosal damage is related to the degree or scoring of neutrophil activity (Blaser 1992; Wallace 1991). The observed data in this study showed that the lower activity or infiltration of neutrophils induced elevated IL-8 concentration despite no significant differences between high and low scoring activity. By contrast, the level of IL-8 was significantly higher with mild scoring for mononuclear infiltration, which contributed to chronic gastritis in the infected patients. Most strains that contributed to higher amounts of IL-8 harbored partially deleted *cag*PAI, and the severity of the disease was only associated at the level of inflammation. As mentioned by Hanafiah et al. (2020), the lack of genes in *cag*PAI might influence the impact of pathogenesis in

H. pylori infection, in turn leading to lesser damage to the gastric mucosa, which is also reflected on IL-8 induction. Further investigation is highly encouraged to understand the complexity of *H. pylori* pathogenicity.

Although several studies have explored IL-8 induction by epithelium upon infection of *cag*PAI-intact *H. pylori*, our data focus on impact of *cag*PAI diversity from multi-ethnic populations upon IL-8 secretion and 'hummingbird' phenotype in AGS cell line. Furthermore, we examined the complete genes of *cag*PAI in this study (total of 27 genes), instead of only several *cag*PAI genes in other studies (Alfizah et al. 2013; Hamat et al. 2013; Schmidt et al. 2010). Moreover, there is a lack of information regarding *cag*PAI integrity of Malaysian *H. pylori* and its association towards *in vitro* studies. Hence, this study has shown novelty in its findings despite the method is rather conventional. Limitation of this study is the small sample size due to difficulties in growing *H. pylori* strains plus restricted time and resources.

Conclusions

This study documented the variability of *cag*PAI genes in *H. pylori* strains isolated from multi-ethnic patients. Interestingly, it was the non-Malay patients that has the highest frequency for harboring intact *cag*PAI. The integrity of *cag*PAI was acknowledged as an important factor influencing the virulence impact of *H. pylori* harboring an island. IL-8 secretion induced by *H. pylori* with various levels of *cag*PAI content appeared to vary in different isolates. Deletion among genes appeared to have a lesser impact on IL-8 production in infected cells, which was reflected by the disease outcomes of patients. Future study directions include the assessment of heterogeneity in the T4SS of *H. pylori* and a proteomics study of *cag*PAI expression.

Acknowledgements

The authors would like to express gratitude to Universiti Kebangsaan Malaysia for providing both permission and the facilities to conduct and publish this research, as well as to technical staff of the Department of Medical Microbiology and Immunology, Faculty of Medicine, Universiti Kebangsaan Malaysia for their technical assistance and guidance. Part of the data have been exhibited in poster presentation at the 30th Annual Meeting of the Korean College of *Helicobacter* and Upper Gastrointestinal Research & International Symposium on *Helicobacter* and Upper Gastrointestinal Diseases (2022). This research was supported by grant of Faculty of Medicine, Universiti Kebangsaan Malaysia [FF-2019-472]. Studentship was also provided to S.A.R. under MyBrain15 program from the Ministry of Higher Education, Malaysia. The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the Medical Research Ethic Committee of University (UKM1.5.3.5/244/JEP-2016-095, date of approval: 17 March 2016). Patient consent was waived due to reason that the current study used *H. pylori* stock cultures as samples. However, these isolates were previously obtained from patients in former studies (research no. ETP-2013-042 and GUP-2011-307) wherein written informed consent has been obtained from the patients. The authors declare no conflict of interest.

References

Ahmadzadeh, A., Ghalehnoei, H., Farzi, N., Yadegar, A., Alebouyeh, M. & Aghdaei, H.A. 2015. Association of *cag*PAI integrity with severeness of *Helicobacter pylori* infection in patients with gastritis. *Pathol. Biologie.* 63: 252-257.

Alfizah, H. & Ramelah, M. 2012. Variant of *Helicobacter pylori* CagA proteins induce different magnitude of morphological changes in gastric epithelial cells. *Malays. J. Pathol.* 34(1): 29-34.

Alfizah, H., Rukman, A.H., Norazah, A., Hamizah, R. & Ramelah, M. 2013. Ethnicity association of *Helicobacter pylori* virulence genotype and metronidazole susceptibility. *World J. Gasteroenterol*. 19: 1283-1291.

Allison, C.C., Kufer, T.A., Kremmer, E., Kaparakis, M. & Ferrero, R.L. 2009. *Helicobacter pylori* induces MAPK phosphorylation and AP-1 activation via a NOD1-dependent mechanism. *J. Immunol*. 183: 8099-8109.

Argent, R.H., Zhang, Y. & Atherton, J.C. 2005. Simple method for determination of the number of *Helicobacter pylori* CagA variable-region EPIYA tyrosine phosphorylation motifs by PCR. *J. Clin. Microbiol*. 43: 791-795.

Argent, R.H., Hale, J.L., El-Omar, E.M. & Atherton, J.C. 2008. Differences in *Helicobacter pylori* CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. *J. Med. Microbiol*. 57: 1062-1067.

Argent, R.H., Kidd, M., Owen, R.J., Thomas, R.J., Limb, M.C. & Atherton, J.C. 2004. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of *Helicobacter pylori*. *Gastroenterol*. 127: 514-523.

Azuma, T., Yamakawa, A., Yamazaki, S., Fukuta, K., Ohtani, M., Ito, Y., Dojo, M., Yamazaki, Y. & Kuriyama, M. 2002. Correlation between variation of the 3' region of the cagA gene in *Helicobacter pylori* and disease outcome in Japan. *J. Infect. Dis*. 186: 1621-1630.

Baj, J., Forma, A., Sitarz, M., Portincasa, P., Garruti, G., Krasowska, D. & Maciejewski, R. 2020. *Helicobacter pylori* virulence factors-mechanisms of bacterial pathogenicity in the gastric microenvironment. *Cells* 10: 27.

Blaser, M.J. 1992. Hypotheses on the pathogenesis and natural history of *Helicobacter pylori*-induced inflammation. *Gastroenterol*. 102(2): 720-727.

Brandt, S., Kwok, T., Hartig, R., König, W. & Backert, S. 2005. NF-κB activation and potentiation of proinflammatory responses by the *Helicobacter pylori* CagA protein. *Proc. Natl. Acad. Sci*. USA. 102: 9300-9305.

Burucoa, C. & Axon, A. 2017. Epidemiology of *Helicobacter pylori* infection. *Helicobacter* 22: S1.

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. USA*. 93: 14648-14653.

Chang, C.C., Kuo, W.S., Chen, Y.C., Perng, C.L., Lin, H.J. & Ou, Y.H. 2016. Fragmentation of CagA reduces hummingbird phenotype induction by *Helicobacter pylori*. *PLoS ONE* 11(3): e0150061.

Choi, Y.H., Lai, J., Kim, M.A., Kim, A., Kim, J., Su, H., Ge, L. & Cha, J.H. 2021. CagL polymorphisms between East Asian and Western *Helicobacter pylori* are associated with different abilities to induce IL-8 secretion. *J. Microbiol*. 59: 763-770.

Cover, T.L., Lacy, D.B. & Ohi, M.D. 2020. The *Helicobacter pylori* cag type IV secretion system. *Trends Microbiol*. 28: 682-695.

Crew, K.D. & Neugut, A.I. 2006. Epidemiology of gastric cancer. *World J. Gastroenterol*. 12: 354-362.

Dixon, M.F., Genta, R.M., Yardley, J.H. & Correa, P. 1996. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. *Am. J. Surg. Pathol*. 20(10): 1161-1181.

Fazeli, Z., Alebouyeh, M., Tavirani, M.R., Yadegar, A., Farzi, N. & Zali, M.R. 2017. Functional cytotoxin associated gene A in *Helicobacter pylori* strains and its association with integrity of *cag*-pathogenicity island and histopathological changes of gastric tissue. *Arch. Clin. Infect. Dis.* 12(3): e62955.

Goh, K.L. 2018. Lessons learnt from the epidemiology of *Helicobacter pylori* infection in Malaysia: JGHF Marshall and Warren Lecture 2017. *J. Gastroenterol. Hepatol.* 33(6): 1177-1184.

Goh, K.L, Cheah, P.L., Md, N., Quek, K.F. & Parasakthi, N. 2007. Ethnicity and *H. pylori* as risk factors for gastric cancer in Malaysia: A prospective case control study. *Am. J. Gastroenterol*. 102: 40-45.

Guillemin, K., Salama, N.R., Tompkins, L.S. & Falkow, S. 2002. *cag* pathogenicity island-specific responses of gastric epithelial cells to *Helicobacter pylori* infection. *Proc. Natl. Acad. Sci. USA*. 99: 15136-15141.

Hamat, R.A., Nor Amalina, E., Malina, O., Zamberi, S., Alfizah, H., Rizal, A.M., Adnan, A. & Mohamed, R. 2013. Ethnic differences in the prevalence, clinical outcome and *cag* pathogenicity island (*cag*PAI) virulence gene profiles of *Helicobacter pylori* strains from Malaysia. *Pertanika J. Trop. Agric. Sci.* 36(4): 289-298.

Hanafiah, A., Razak, S.A., Neoh, H.M., Zin, N.M. & Lopes, B.S. 2020. The heterogeneic distribution of *Helicobacter pylori cag* pathogenicity island reflects different pathologies in multiracial Malaysian population. *Braz. J. Infect. Dis*. 24: 545-551.

Hatakeyama, M. 2009. *Helicobacter pylori* and gastric carcinogenesis. *J. Gastroenterol*. 44: 239-248.

Hatakeyama, M. 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. *Nat. Rev. Cancer*. 4: 688-694.

Higashi, H., Tsutsumi, R., Fujita, A., Yamazaki, S., Asaka, M., Azuma, T. & Hatakeyama, M. 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. *Proc. Natl. Acad. Sci. USA*. 99: 14428-14433.

Horridge, D.N., Begley, A.A., Kim, J., Aravindan, N., Fan, K. & Forsyth, M.H. 2017. Outer inflammatory protein a (OipA) of *Helicobacter pylori* is regulated by host cell contact and mediates CagA translocation and interleukin-8 response only in the presence of a functional *cag* pathogenicity island type IV secretion system. *Pathog. Dis.* 75: ftx113.

Hunt, R.H., Xiao, S.D., Megraud, F., Leon-Barua, R., Bazzoli, F., van der Merwe, S., Vaz Coelho, L.G., Fock, M., Fedail, S., Cohen, H., Malfertheiner, P., Vakil, N., Hamid, S., Goh, K.L., Wong, B.C., Krabshuis, J. & Le Mair, A. 2011. World Gastroenterology Organization: *Helicobacter pylori* in developing countries. World Gastroenterology Organisation Global Guideline. *J. Gastrointestin. Liver. Dis*. 20(3): 299-304.

Jones, K.R., Joo, Y.M., Jang, S., Yoo, Y., Lee, H.S., Chung, I., Olsen, C.H., Whitmire, J.M., Merrel, D.S. & Cha, J.H. 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. *J. Clin. Microb*. 47(4): 959-968.

Keikha, M. & Karbalaei, M. 2021. EPIYA motifs of *Helicobacter pylori cagA* genotypes and gastrointestinal diseases in the Iranian population: A systematic review and meta-analysis. *New Microbes and New Infections* 41: 100865.

Khaledi, M., Bagheri, N., Validi, M., Zamanzad, B., Afkhami, H., Fathi, J., Rahimian, G. & Gholipour, A. 2020. Determination of *CagA* EPIYA motif in *Helicobacter pylori* strains isolated from patients with digestive disorder. *Heliyon*. 6(9): e04971.

Khatoon, J., Prasad, K.N., Prakash Rai, R., Ghoshal, U.C. & Krishnani, N. 2017. Association of heterogenicity of *Helicobacter pylori* cag pathogenicity island with peptic ulcer diseases and gastric cancer. *Br. J. Biomed. Sci.* 74: 121-126.

Koller, B., Bals, R., Roos, D., Korting, H.C., Griese, M. & Hartl, D. 2009. Innate immune receptors on neutrophils and their role in chronic lung disease. *Eur. J. Clin. Invest*. 39: 535-547.

Lamb, A. & Chen, L.F. 2010. The many roads travelled by *Helicobacter pylori* to NFκB activation. *Gut Microbes*. 1: 109-113.

Markovska, R., Boyanova, L., Yordanov, D., Stankova, P., Gergova, G. & Mitov, I. 2018. Status of *Helicobacter pylori cag* pathogenicity island (*cag*PAI) integrity and significance of its individual genes. *Infect. Genet. Evol*. 59: 167-171.

Mueller, A., Merrell, D.S., Grimm, J. & Falkow, S. 2004. Profiling of microdissected gastric epithelial cells reveals a cell type-specific response to *Helicobacter pylori* infection. *Gastroenterol*. 127(5): 1446-1462.

Nguyen, L.T., Uchida, T., Tsukamoto, Y., Trinh, T.D., Ta, L., Mai, H.B., Le, H.S., Ho, D.Q.D., Hoang, H.H., Matsuhisa, T., Okimoto, T., Kodama, M., Murakami, K., Fujioka, Y., Yamaoka, Y. & Moriyama, M. 2010. Clinical relevance of *cag*PAI intactness in *Helicobacter pylori* isolates from Vietnam. *Eur. Soc. Clin. Microb*. 29(6): 651-660.

Nilsson, C., Sillén, A., Eriksson, L., Strand, M., Enroth, H., Normark, S., Falk, P. & Engstrand, L. 2003. Correlation between *cag* pathogenicity island composition and *Helicobacter pylori*-associated gastroduodenal disease. *Infect. Immun*. 71(11): 6573-6581.

Olbermann, P., Josenhans, C., Moodley, Y., Uhr, M., Stamer, C., Vauterin, M., Suerbaum, S., Achtman, M. & Linz, B. 2010. A global overview of the genetic and functional diversity in the *Helicobacter pylori cag* pathogenicity island. *PLoS Genet*. 6: e1001069.

Patra, R., Chattopadhyay, S., De, R., Datta, S., Chowdhury, A., Ramamurthy, T., Balakrish Nair, G., Berg, D.E. & Mukhopadphyay, A.K. 2011. Intact *cag* pathogenicity island of *Helicobacter pylori* without disease association in Kolkata, India. *Int. J. Med. Microbiol.* 301: 293-302.

Peek Jr., R.M. & Crabtree, J.E. 2006. *Helicobacter* infection and gastric neoplasia. *J. Pathol*. 208(2): 233-248.

Phuc, B.H., Tuan, V.P., Dung, H., Binh, T.T., Tung, P.H., Tri, T.D., Thuan, N.P.M., Khien, V.V., Trang, T.T.H., Akada, J., Matsumoto, T. & Yamaoka, Y. 2021. *Helicobacter pylori* type 4 secretion systems as gastroduodenal disease markers. *Sci. Rep*. 11(1): 4584.

Rizzato, C., Torres, J., Obazee, O., Camorlinga-Ponce, M., Trujillo, E., Stein, A., Mendez-Tenorio, A., Bravo, M.M., Canzian, F. & Kato, I. 2020. Variations in *cag* pathogenicity island genes of *Helicobacter pylori* from Latin American groups may influence neoplastic progression to gastric cancer. *Sci. Rep*. 10(1): 6570.

Rudi, J., Kolb, C., Maiwald, M., Kuck, D., Sieg, A., Galle, P.R. & Stremmel, W. 1998. Diversity of *Helicobacter pylori* vacA and cagA genes and relationship to VacA and CagA protein expression, cytotoxin production, and associated diseases. *J. Clin. Microb*. 36: 944-948.

Salih, B.A., Guner, A., Karademir, A., Uslu, M., Ovali, M.A., Yazici, D., Bolek, B.K. & Arikan, S. 2014. Evaluation of the effect of *cag*PAI genes of *Helicobacter pylori* on AGS epithelial cell morphology and IL-8 secretion. *Antonie van Leeuwenhoek* 105: 179-189.

Schmidt, H.M.A., Andres, S., Nilsson, C., Kovach, Z., Kaakoush, N.O., Engstrand, L., Fock, K.M., Forman, D. & Mitchell, H. 2010. The *cag* PAI is intact and functional but HP0521 varies significantly in *Helicobacter pylori* isolates from Malaysia and Singapore. *Eur. J. Clin. Microbiol. Infect. Dis.* 29: 439-451.

Schneider, N., Krishna, U., Romero-Gallo, J., Israel, D.A., Piazuelo, M.B., Camargo, M.C., Sicinschi, L.A., Schneider, B.G., Correa, P. & Peek Jr., R.M. 2009. Role of *Helicobacter pylori* CagA molecular variations in induction of host phenotypes with carcinogenic potential. *J. Infect. Dis*. 199: 1218-1221.

Sukri, A., Hanafiah, A., Kosai, N.R., Mohammed Taher, M. & Mohamed, R. 2022. New insight on the role of *Helicobacter pylori* cagA in the expression of cell surface antigens with important biological functions in gastric carcinogenesis. *Helicobacter* 27(5): e12913.

Suzuki, M., Mimuro, H., Suzuki, T., Park, M., Yamamoto, T. & Sasakawa, C. 2005. Interaction of CagA with Crk plays an important role in *Helicobacter pylori*-induced loss of gastric epithelial cell adhesion. *J. Exp. Med*. 202: 1235-1247.

Ta, L.H., Hansen, L.M., Sause, W.E., Shiva, O., Millstein, A., Ottemann, K.M., Castillo, A.R. & Solnick, J.V. 2012. Conserved transcriptional unit organization of the *cag* pathogenicity island among *H. pylori* strains. *Front. Cellular. Infect. Microbiol*. 2: 46.

Takahashi-Kanemitsu, A., Knight, C.T. & Hatakeyama, M. 2020. Molecular anatomy and pathogenic actions of *Helicobacter pylori* CagA that underpin gastric carcinogenesis. *Cell. Mol. Immunol.* 17: 50-63.

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

Terry, C.E., McGinnis, L.M., Madigan, K.C., Cao, P., Cover, T.L., Liechti, G.W., Peek Jr., R.M. & Forsyth, M.H. 2005. Genomic comparison of *cag* pathogenicity island (PAI)-positive and -negative *Helicobacter pylori* strains: Identification of novel markers for *cag* PAI-positive strains. *Infect. Immun*. 73(6): 3794-3798.

Tsutsumi, R., Higashi, H., Higuchi, M., Okada, M. & Hatakeyama, M. 2003. Attenuation of *Helicobacter pylori* CagA x SHP-2 signaling by interaction between CagA and C-terminal Src kinase. *J. Biol. Chem.* 278(6): 3664-3670.

Wallace, J.L. 1991. Possible mechanisms and mediators of gastritis associated with *Helicobacter pylori* infection. *Scand. J. Gastroenterol*. *Suppl.* 187: 65-70.

Wang, H., Zhao, M., Shi, F., Zheng, S., Xiong, L. & Zheng, L. 2023. A review of signal pathway induced by virulent protein CagA of *Helicobacter pylori*. *Front. Cell. Infect. Microbiol.* 13: 1062803.

Waskito, L.A., Miftahussurur, M., Lusida, M.I., Syam, A.F., Suzuki, R., Subsomwong, P., Uchida, T., Hamdan, M., Nasronudin & Yamaoka, Y. 2018. Distribution and clinical associations of integrating conjugative elements and *cag* pathogenicity islands of *Helicobacter pylori* in Indonesia. *Sci. Rep*. 8: 6073.

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

Xue, Z., You, Y., He, L., Gong, Y., Sun, L., Han, X., Fan, R., Zhai, K., Yang, Y., Zhang, M., Yan, X. & Zhang, J. 2021. Diversity of 3' variable region of cagA gene in *Helicobacter pylori* strains isolated from Chinese population. *Gut. Pathog*. 13: 23.

Yakoob, J., Abid, S., Abbas, Z., Jafri, W., Ahmad, Z., Ahmed, R. & Islam, M. 2009. Distribution of *Helicobacter pylori* virulence markers in patients with gastroduodenal diseases in Pakistan. *BMC. Gastroenterol*. 9: 87.

Yamaoka, Y. 2010. Mechanisms of disease: *Helicobacter pylori* virulence factors. *Nat. Rev. Gastroenterol. Hepatol*. 7: 629-641.

*Corresponding author; email: alfizah@ppukm.ukm.edu.my

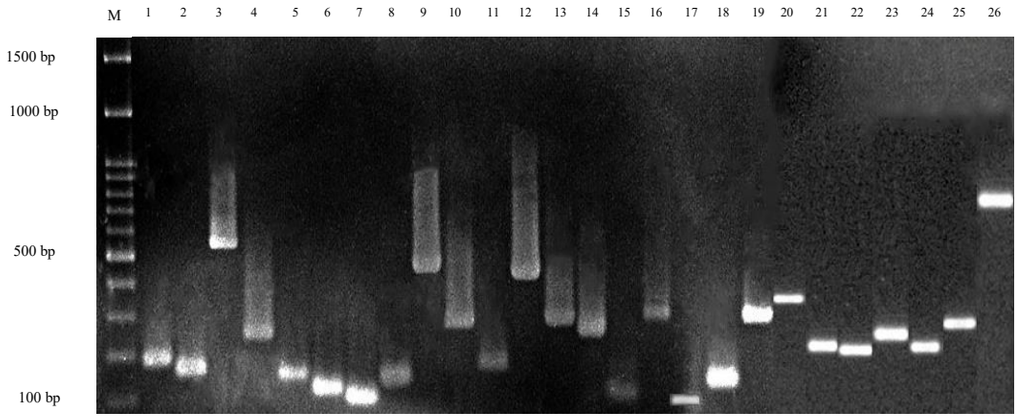

FIGURE A.1. Gel electrophoresis of *cag*PAI genes in *H. pylori* representative strain. Identification of individual genes were conducted and reported previously by Hanafiah et al. (2020). *H. pylori* strain ATCC 43526 was used as positive control. Lane M: molecular weight marker 100 bp; Lane 1; *cag1*; Lane 2; *cag2*; Lane 3; *cag3/cagδ* ; Lane 4: *cag4/cagγ*; Lane 5: *cag5/cagβ*; Lane 6: *cag6/cagZ*; Lane 7: *cag7/cagY*; Lane 8: *cag8/cagX*; Lane 9: *cag9/cagV*; Lane 10: *cag10/cagW*; Lane 11: *cag11/cagU*; Lane 12: *cag12/cagT*; Lane 13: *cag13/cagS*; Lane 14: *cag14/cagQ*; Lane 15: *cag15/cagP*; Lane 16: *cag16/cagM*; Lane 17: *cag17/cagN*; Lane 18: *cag18/cagL*; Lane 19: *cag19/cagI*; Lane 20: *cag20/cagH*; Lane 21: *cag21/cagG*; Lane 22: *cag22/cagF*; Lane 23: *cag23/cagE*; Lane 24: *cag24/cagD*; Lane 25: *cag25/cagC*; Lane 26: *cagα*

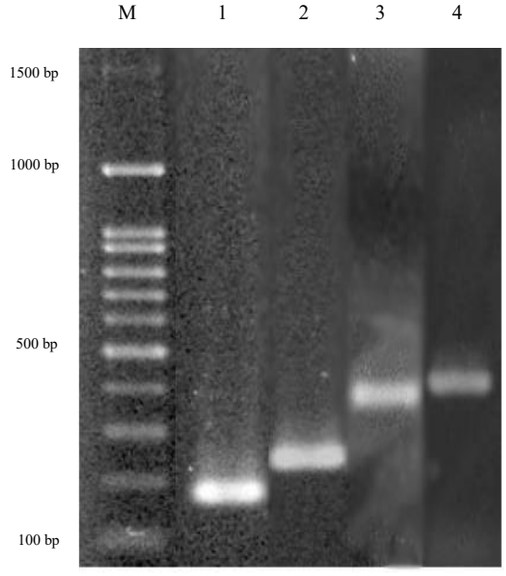

FIGURE A.2. PCR amplification of *cag*A EPIYA motifs in *H. pylori* strain. Lane M: molecular weight marker 100 bp; Lane 1: EPIYA-A; Lane 2: EPIYA-B; Lane 3: EPIYA-C; Lane 4: EPIYA-D

TABLE A.1. Details of primer sequences used for amplification of *H. pylori cagA* EPIYA motifs

EPIYA motif	Forward primer	Sequence (5' – 3')	Reference	Reverse primer	Sequence (5' – 3')	Reference
EPIYA-A	cag2	GGAACCCTAGTCGGTAATG	Rudi et al. (1998)	cagA-P1C	GTCCTGCTTTCTTTTTATTAACTTKAGC	Azuma et al. (2002)
EPIYA-B				cagA-P2TA	TTTAGCAACTTGAGTATAAATGGG	
				cagA-P2CG	TTTAGCAACTTGAGCGTAAATGGG	
EPIYA-C				cagA-P3E	ATCAATTGTAGCGTAAATGGG	
EPIYA-D				cagA-pD(R)	TTGATTTGCCTCATCAAAATC	Jones et al. (2009)

TABLE A.2. Mean concentration of IL-8 and 'hummingbird' phenotype according to existence of individual *cag*PAI genes (Part 1)

*cag*PAI gene	Gene's existence	IL-8 concentration [(ρg/mL) ± SD]	'Hummingbird' effect (% ± SD)
cag1 (cagζ)	Positive (n = 24)	434.2 ± 156.1 [a]	25.8 ± 11.3
	Negative (n = 3)	239.2 ± 55.9	0
cag2 (cagε)	Positive (n = 12)	445 ± 166.4	29.4 ± 10.2
	Negative (n = 15)	386.5 ± 156	17.7 ± 13.7
cag3 (*cagΔ*)	Positive (n = 23)	433.9 ± 159.6	26.8 ± 10.3
	Negative (n = 4)	289.8 ± 110.9	0.5 ± 1
cag4 (cagγ)	Positive (n = 22)	424.3 ± 156.4	26.3 ± 10.3
	Negative (n = 5)	360.6 ± 185.3	7.8 ± 16.3
cag5 (*cagβ*)	Positive (n = 24)	434.2 ± 156.1 [b]	25.8 ± 11.3
	Negative (n = 3)	239.2 ± 55.9	0
cagα	Positive (n = 21)	433.5 ± 167.1	26.3 ± 10.4
	Negative (n = 6)	338.9 ± 116.1	11.1 ± 16.9
cag6 (*cagZ*)	Positive (n = 24)	434.2 ± 156.1 [c]	25.8 ± 11.3
	Negative (n = 3)	239.2 ± 55.9	0
cag7 (*cagY*)	Positive (n = 22)	443.6 ± 156.1 [d]	26.8 ± 10.5
	Negative (n = 5)	275.6 ± 101.1	5.6 ± 11.4
cag8 (cagX)	Positive (n = 24)	434.2 ± 156.1 [e]	25.8 ± 11.3
	Negative (n = 3)	239.2 ± 55.9	0
cag9 (cagW)	Positive (n = 23)	433.9 ± 159.6	26.8 ± 10.3
	Negative (n = 4)	289.8 ± 110.9	0.5 ± 1
cag10 (cagV)	Positive (n = 23)	433.9 ± 159.6	26.8 ± 10.3
	Negative (n = 4)	289.8 ± 110.9	0.5 ± 1
cag11 (cagU)	Positive (n = 21)	422.3 ± 161.1	27 ± 11.2
	Negative (n = 6)	378.1 ± 167	8.3 ± 10.5
cag12 (*cagT*)	Positive (n = 23)	433.9 ± 159.6	26.8 ± 10.3
	Negative (n = 4)	289.8 ± 110.9	0.5 ± 1
cag13 (*cagS*)	Positive (n = 22)	435 ± 163.2	26.8 ± 10.5
	Negative (n = 5)	313.8 ± 110.1	5.6 ± 11.4

[a b c d]There is a significant difference upon IL-8 concentration according to the existence of *cag1*/*cagζ*, *cag5*/*cagβ*, *cag6*/*cagZ* dan *cag8*/*cagX* ($P = 0.045$)

[d]There is a significant impact towards IL-8 induction according to the existence of *cag7*/*cagY* ($P = 0.025$)

TABLE A.2. Mean concentration of IL-8 and 'hummingbird' phenotype according to existence of individual *cag*PAI genes (Part 2)

*cag*PAI gene	Gene's existence	IL-8 concentration [(ρg/mL) ± SD]	'Hummingbird' effect (% ± SD)
cag14 (*cagQ*)	Positive (n = 12)	451.9 ± 180.6	32.5 ± 6.5
	Negative (n = 15)	381 ± 140.3	15.2 ± 12.6
cag15 (*cagP*)	Positive (n = 22)	437.1 ± 153	26.3 ± 11.4
	Negative (n = 5)	304.2 ± 161.3	7.9 ± 11.7
cag16 (cagM)	Positive (n = 18)	439.6 ± 163.8	28.9 ± 10.2 [I]
	Negative (n = 9)	353.3 ± 146.6	10.9 ± 11.1
cag17 (*cagN*)	Positive (n = 22)	443.6 ± 156.2 [f]	26.8 ± 10.5
	Negative (n = 5)	275.8 ± 101	5.7 ± 11.6
cag18 (cagL)	Positive (n = 22)	443.6 ± 156.2 [g]	26.8 ± 10.5
	Negative (n = 5)	275.8 ± 101	5.7 ± 11.6
cag19 (*cagI*)	Positive (n = 21)	454.3 ± 151.6 [h]	26.8 ± 10.8
	Negative (n = 6)	266.4 ± 93.2	9.1 ± 13.3
cag20 (*cagH*)	Positive (n = 24)	432.9 ± 163.2	26.6 ± 10.5
	Negative (n = 3)	323.1 ± 121.6	6.4 ± 13.2
cag21 (*cagG*)	Positive (n = 24)	434.2 ± 156.1 [i]	25.8 ± 11.3
	Negative (n = 3)	239.2 ± 55.9	0
cag22 (*cagF*)	Positive (n = 22)	443.6 ± 156.2 [j]	26.8 ± 10.5
	Negative (n = 5)	275.8 ± 101	5.7 ± 11.6
cag23 (*cagE*)	Positive (n = 21)	436.9 ± 156.8	27.5 ± 10.3
	Negative (n = 6)	327.1 ± 154.7	6.9 ± 10.8
cag24 (*cagD*)	Positive (n = 19)	425 ± 168.3	28.5 ± 10.7 [II]
	Negative (n = 8)	383 ± 146	9.6 ± 9.5
cag25 (*cagC*)	Positive (n = 20)	444.5 ± 164	27.5 ± 10.6
	Negative (n = 7)	321.3 ± 114.2	9.6 ± 12.2

[f g j]There is a significant difference upon IL-8 concentration according to the existence of *cag17/cagN, cag18/cagL* and *cag22/cagF* ($P = 0.029$)
[h]There is a significant difference upon IL-8 concentration according to the existence of *cag19/cagI* ($P = 0.009$)
[i]There is a significant difference upon IL-8 concentration according to the existence of *cag21/cagG* ($P = 0.045$)
[I]There is a significant difference in 'hummingbird' formation according to existence of *cag16/cagM* ($P = 0.025$)
[II]There is a significant difference in 'hummingbird' formation according to existence of *cag24/cagD* ($P = 0.014$)