Original article

CagL amino acid sequence polymorphism of *Helicobacter pylori* in Vietnamese patients with gastroduodenal diseases

Tran Ha Duy Khang[1], Nguyen Thi Mai Ngan[2], Thai Thi Hong Nhung[3], Pham Van Linh[4*]

[1] Medical student at Faculty of Medicine, Hong Bang International University, Ho Chi Minh city, Vietnam

[2] Department of Medical Genetics, University of Medicine and Pharmacy, Hue University, Hue, Vietnam

[3]Department of Internal Medicine, Faculty of Medicine Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam

[4]Faculty of Medicine, Hong Bang International University, Ho Chi Minh city, Vietnam

***Corresponding author:** Phạm Văn Lình; email: linhpv@hiu.vn

Received: 05/02/2026; **Accepted:** 01/04/2026; **Published:** 30/04/2026

DOI: 10.34071/jmp.2026.2.953

Abstract

Background: *H. pylori* CagA protein plays the most critical pathogenesis role of gastroduodenal diseases. CagA translocation is dependent on the CagL protein encoded by the *cagL* gene. **Objectives:** This study aimed to investigate CagL amino acid sequence polymorphisms of *H. pylori* and its associations with gastroduodenal diseases.

Materials and methods: The *cagL* partial nucleotide sequences of 72 *H. pylori* strains from Vietnamese patients with gastroduodenal disorders were investigated via Sanger sequencing, then were translated into amino acid sequences.

Results: The findings showed nine variants of amino acid polymorphism within the CagL hypervariable motif at residues 58 - 62, in which the DEIGK variant was the most prevalent, accounting for 40.28%, the DKIGK variants represented 30.56%, both DKMGE and DTTGE variants accounted for 8.33%. The five remaining variants, including YEIGK, NEIGQ, NKIGQ, DEIGQ and DTIGK, had low frequencies (1.39 - 4.17%). Notably, we observed a novel amino acid sequence polymorphism pattern at residues 140-144 with four variants as EAELQ, EGKLK, LGKLK, and EAKLQ, which accounted for 61.11%, 30.56%, 5.56%, and 2.78%, respectively. Stratifying gastroduodenal diseases, the EAELQ variant rates in the gastric cancer and precancerous lesions groups were 86.67% and 70%, respectively, whereas those in the non-atrophic gastritis and peptic ulcers groups were only 47.37% and 44.44%, respectively. The difference in these rates was statistically significant (p = 0.038). Multivariable logistic regression analysis demonstrated that the EAELQ variant was independently associated with premalignant and malignant gastric lesions, after adjustment for age and gender, OR=3.82 (1.40 - 11.20).

Conclusion: This preliminary study highlighted the genetic diversity of the *H. pylori cagL* gene, revealing distinct polymorphism patterns unique to Vietnam. The EAELQ variant at CagL residues 140–144 may serve as an indicator of gastric cancer risk.

Keywords: gastric cancer, *cagL* gene, *Helicobacter pylori*, amino acid polymorphism, gastroduodenal diseases**.**

1. INTRODUCTION

Helicobacter pylori (H. pylori), a microaerophillic Gram-negative bacterium, infects more than half of the population in the world [1]. This bacterium infection causes chronic gastritis (CG) and can subsequently lead to other gastroduodenal disorders, such as peptic ulcers (PU), gastric cancer (GC), and gastric mucosa-associated lymphoid tissue lymphoma [2]. In 1994, the International Agency for Research on Cancer identified this bacterium as a Group I definite carcinogen that causes GC [3]. Nevertheless, not all H. pylori infections result in severe clinical outcomes. Several factors influence the progression of H. pylori-induced gastroduodenal diseases, including host susceptibility, environmental factors and, most importantly, bacterial toxins [4]. H. pylori CagA protein encoded by cytotoxin-associated gene A (cagA gene) within the cag pathogenicity island (cagPAI) plays the most critical pathogenesis role [5]. A number of studies have shown that cagA-positive H. pylori strains are identified as risk factors for GC [6,7]. H. pylori CagA protein is translocated into host gastric epithelial cells via the Type IV secretion system (T4SS).

The CagL protein, a component of T4SS encoded by the cagL gene, plays an important role in the CagA translocation [8]. The CagL arginine-glycine-aspartate (RGD) motif, at residues 76–78, is essential for the interaction of T4SS with integrin receptors during translocation. Although the RGD motif

is conservative, its adjacent regions are highly polymorphic, particularly the CagL hypervariable motif (CagLHM) at residues 58–62. More than 30 variations of CagLHM have been identified globally, and geographical diversity has been associated with clinical outcomes, especially GC [9–11]. Furthermore, CagL can trigger the intracellular signal pathway, causing an inflammatory response independent of CagA translocation [11]. Therefore, CagL amino acid polymorphisms should be investigated for their association with gastroduodenal diseases.

Vietnam has a high frequency of H. pylori infection (about 60%) [12]. H. pylori eradication treatment is becoming increasingly challenging due to growing antibiotic resistance [13,14]. In 2022, the reported number of new cases of GC in Vietnam was 16,277 [15]. The prevalence of cagA-positive H. pylori in East Asian nations, including Vietnam, is extremely high [16–18]. Notably, cagA-positive H. pylori strains were frequently observed not only in the GC group but also in benign groups in East Asian nations [19–22]. As a result, cagA positivity cannot be used as a useful biomarker for GC risk in East Asia [19]. In addition to cagA, some H. pylori virulent genes have been deeply investigated in Vietnam, such as cagE, vacA, iceA, babA, oipA, homB, etc [22–25]. To date, there has been no research on the H. pylori cagL gene in Vietnam.

In that circumstance, investigating polymorphisms of H. pylori virulent cagL gene is essential not only for better understanding H. pylori genetic diversity and pathogenesis but also for identifying H. pylori strains with high toxicity. This will serve as a foundation for developing more effective and targeted strategies to prevent H. pylori-related gastroduodenal disorders, including GC. Therefore, this study was carried out to investigate CagL amino acid sequence polymorphisms of H. pylori and its associations with gastroduodenal diseases.

2. MATERIALS AND METHODS

2.1. Study population

2.1.1. Inclusion criteria

- Patients who underwent upper gastrointestinal endoscopy having gastroduodenal mucosa lesions, such as gastritis, peptic ulcers, and visible neoplastic suspicious lesions.

+ nAG or PU was diagnosed via gastroduodenal mucosa lesions recorded from upper gastrointestinal endoscopy.

+ PCL or GC was diagnosed by histopathological examination on gastric biopsy specimens: Atrophic gastritis and intestinal metaplasia were graded using the updated Sydney System [26], whereas dysplasia and gastric cancer were recognized using the 2019 WHO Classification of Tumours [27].

- *H. pylori* infection was identified by rapid urease test (RUT) on antrum and corpus biopsy specimens and was confirmed by PCR with *ureC*-specific primers for extracted DNA samples [28].

- *H. pylori* with *cagA*-positivity was detected using PCR as previously described [22].

- *H. pylori* with *cagL*-positivity was detected using PCR as described in the method section.

2.1.2. Exclusion criteria

- Patients who were received therapy with antibiotics and/or proton pump inhibitors and/or an H2 blocker within 4 weeks prior to endoscopy.

- Patients who had undergone gastrectomy.

2.2. Methods

2.2.1. Sample size

This study was conducted on patients who underwent upper gastrointestinal endoscopy and were obtained gastric mucosa biopsy specimens at the Hospital of Can Tho University of Medicine and Pharmacy (CTUMP) or the Hue University of Medicine and Pharmacy Hospital (HueUMP Hospital) from June 2019 to December 2024. DNA samples extracted from *H. pylori* and *cagA*-positive biopsy specimens or isolates were stored at -20°C at the Department of Medical Genetics, Hue University of Medicine and Pharmacy. The Ethics Committee of the Hue University of Medicine and Pharmacy approved the research on the *H. pylori cagL* gene on October 5, 2023, under the code H2023/474. Therefore, *H. pylori cagL* positivity was identified from October, 2023 by PCR assays.

Finally, we collected 72 DNA samples from patients with gastroduodenal diseases infected with *cagL*-positive *H. pylori*, including 19 patients with non-atrophic gastritis (nAG), 18 patients with PU, 20 patients with chronic gastritis with precancerous lesions (PCL), and 15 patients with gastric cancer (GC).

2.2.2. Amplifying H. pylori cagL *fragment*

- The *cagL* fragment was amplified by PCR assay using primers cagL-F-B4-new and cagL-R-B5-new primers that were modified from Yadegar's original cagL-F-B4 and cagL-R-B5 primers [11].

+ Forward primer (cagL-F-B4-new): 5'-GCAGAAGATATAACAAGCGGCTTAAAG-3'

+ Reverve primer (cagL-R-B5-new): 5'-ATTAGAAATCATAGCCTATCGTCTCAG-3'

- Each PCR reaction was performed in a total

volume of 50 µL containing 25 µL of GoTaq Green MasterMix 2´ (Promega Corp., Madison, WI, USA), 2.5 µL (10 pmol/µL) of forward and reverse primers, 2 µL (100 ng/µL) of genomic DNA, and 18 µL of nuclease-free sterile water.

- Temperature conditions included initial denaturation at 95°C for 5 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 50 sec, extension at 72°C for 1 min, and a final extension at 72°C for 8 min. PCR reaction was carried out in the thermal cycler SureCycler 8800 (Agilent Technologies, Malaysia).

- The PCR products were electrophoresed on a 1% agarose gel with SafeViewTM Classic (Applied Biological Materials Inc. (abm), Richmond, BC, Canada) at 80 V for 1 h. The product size was 689 bp.

- PCR technique was performed at Department of Medical Genetics, Hue University of Medicine and Pharmacy.

2.2.3. Genotyping *H. pylori cagL*

- Amplicons were subjected to Sanger direct sequencing at 1st BASE Laboratories (Malaysia) using an ABI PRISM 3730xl Genetic Analyzer (Applied Biosystems, USA). Results were received as .ab1 and .seq files.

- We used BioEdit software version 7.2.5 (Informer Technologies Inc.) to read DNA sequencing traces and translate them in-frame into amino acid sequences. The amino acid sequences of our *H. pylori* strains were aligned with those of two reference strains, F32 (AP011943.1) and 26695 (NC_000915.1), which represent East Asian and Western-type strains, respectively. The F32 and 26695 strain sequences are publicly available on GenBank (https://www.ncbi.nlm.nih.gov/genbank/). The *cagL* partial nucleotide sequences were deposited in GenBank database under accession numbers from PQ878555 to PQ878603, from PV035783 to PV035801, from PV126523 to PV126526, which will be publicly accessible from February 2026.

2.2.4. Phylogenetic analysis

A phylogenetic tree of 72 our clinical *H. pylori* samples and two reference strains was constructed based on *cagL* partial sequences using Molecular Evolutionary Genetics Analysis (MEGA11) software (Pennsylvania, USA). Particularly, sequences of this study were aligned against reference sequences using ClustalW multiple alignment tool. A Neighbor-joining tree was constructed with bootstrap method at 1000 replications. The phylogenetic tree was into two major clade (clade A and clade B) based on topology-defined bifurcation points.

2.2.5. Statistical analysis

Statistical analysis was carried out on the VassarStats website (http://vassarstats.net) (Richard Lowry, Vassar College, Poughkeepsie, NY). ANOVA and Tukey HSD tests were performed to assess significant differences in mean age between groups with gastroduodenal disease. A chi-square test was also used to determine significant relationships between *cagL* amino acid polymorphisms and gastroduodenal diseases; if more than 20% of the expected cell frequencies were less than 5, Fisher's exact tests were used. Multivariable logistic regression analysis was performed using R package version 4.5.3 to evaluate the independent association between the *H. pylori* CagL amino acid polymorphism and premalignant and malignant gastric lesions, with adjustment for potential confounders such as age and gender. Odds ratios (OR) and 95% confidence intervals (CI) were calculated. A p-value of less than 0.05 was statistically significant.

3. RESULTS

3.1. Patient characteristics

Table 1. Characteristics of patients with gastroduodenal diseases infected with *cagL*-positive *H. pylori*

Characteristics	nAG	PU	PCL	GC	p-value
Age ($\overline{X} \pm$ SD)	41.74 ± 15.08	44.56 ± 14.83	37.00 ± 12.63	64.87 ± 13.97	**< 0.0001***
Gender (%)					
Male (n = 48)	11 (57.89)	14 (77.78)	12 (60.00)	11 (73.33)	0.524**
Female (n = 24)	8 (42.11)	4 (22.22)	8 (40.00)	4 (26.67)	
Total (N = 72)	**19**	**18**	**20**	**15**	

Notes: **p*-value was calculated by ANOVA test. Tukey HSD test showed the significant differences in the mean age between nAG group, or PU group, or PCL group and GC group, with $p < 0.01$.

**p-value was calculated by Fisher's exact test.

The GC group's mean age was 64.87 ± 13.97, which was greater than that of the other groups. The distribution of genders within the groups of gastroduodenal diseases did not differ.

3.2. Amino acid polymorphism of CagL in groups of gastroduodenal diseases

Table 2. Distribution of amino acid polymorphisms of CagL classical motifs in groups of gastroduodenal diseases

Motif	Poly-morphism	Number (%)	nAG (%)	PU (%)	PCL (%)	GC (%)	p-value
CagLHM	DEIGK	29 (40.28)	9 (47.37)	3 (16.67)	10 (50)	7 (46.67)	
	DKIGK	22 (30.56)	6 (31.58)	7 (38.89)	3 (15	6 (40)	
	DKMGE	6 (8.33)	2 (10.53)	2 (11.11)	2 (10)	0 (0)	
	DTTGE	6 (8.33)	0 (0.00)	2 (11.11)	2 (10)	2 (13.33)	
	YEIGK	3 (4.17)	1 (5.26)	0 (0.00)	2 (10)	0 (0)	> 0.05
	NEIGQ	2 (2.78)	0 (0.00)	1 (5.56)	1 (5)	0 (0)	
	NKIGQ	2 (2.78)	1 (5.26)	1 (5.56)	0 (0)	0 (0)	
	DEIGQ	1 (1.39)	0 (0.00)	1 (5.56)	0 (0)	0 (0)	
	DTIGK	1 (1.39)	0 (0.00)	1 (5.56)	0 (0)	0 (0)	
RGD	RGD	70 (97.22)	18 (94.74)	18 (100)	20 (100)	14 (93.33)	> 0.05
	KGD	2 (2.78)	1 (5.26)	0 (0)	0 (0)	1 (6.67)	
LXXL	LALL	72 (100)	19 (100)	18 (100)	20 (100)	15 (100)	> 0.05
	Other	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
FEANE	FEANE	61 (84.72)	16 (84.21)	16 (88.89)	17 (85)	12 (72.73)	> 0.05
	FETNE	11 (15.28)	3 (15.79)	2 (11.11)	3 (15)	3 (27.27)	
TASLI	TASLI	67 (93.06)	17 (89.47)	17 (94.44)	18 (90)	15 (100)	
	TVSLI	2 (2.78)	2 (10.53)	0 (0)	0 (0)	0 (0)	> 0.05
	TTSLI	3 (4.17)	0 (0)	1 (5.56)	2 (10)	0 (0)	
SKIIVK	SKIIVK	69 (95.83)	18 (94.74)	17 (94.44)	20 (100)	14 (93.33)	> 0.05
	SKVIVK	3 (4.17)	1 (5.26)	1 (5.56)	0 (0)	1 (6.67)	
Total		**72**	**19**	**18**	**20**	**15**	

Among the nine variants of CagLHM, the DEIGK variant was the most prevalent, accounting for 40.28%. The DKIGK variants represented 30.56%, both DKMGE and DTTGE variants accounted for 8.33%. The five remaining variants had low frequencies: YEIGK (4.17%), NEIGQ (2.78%), NKIGQ (2.78%), DEIGQ (1.39%) and DTIGK (1.39%). The other motifs, including RGD, LXXL, FEANE, TASLI, and SKIIVK, were observed to be mostly conserved. The distribution of CagL amino acid polymorphism did not differ between groups of gastroduodenal diseases.

Table 3. Distribution of novel amino acid polymorphisms of CagL at residues 140 - 144 in groups of gastroduodenal diseases

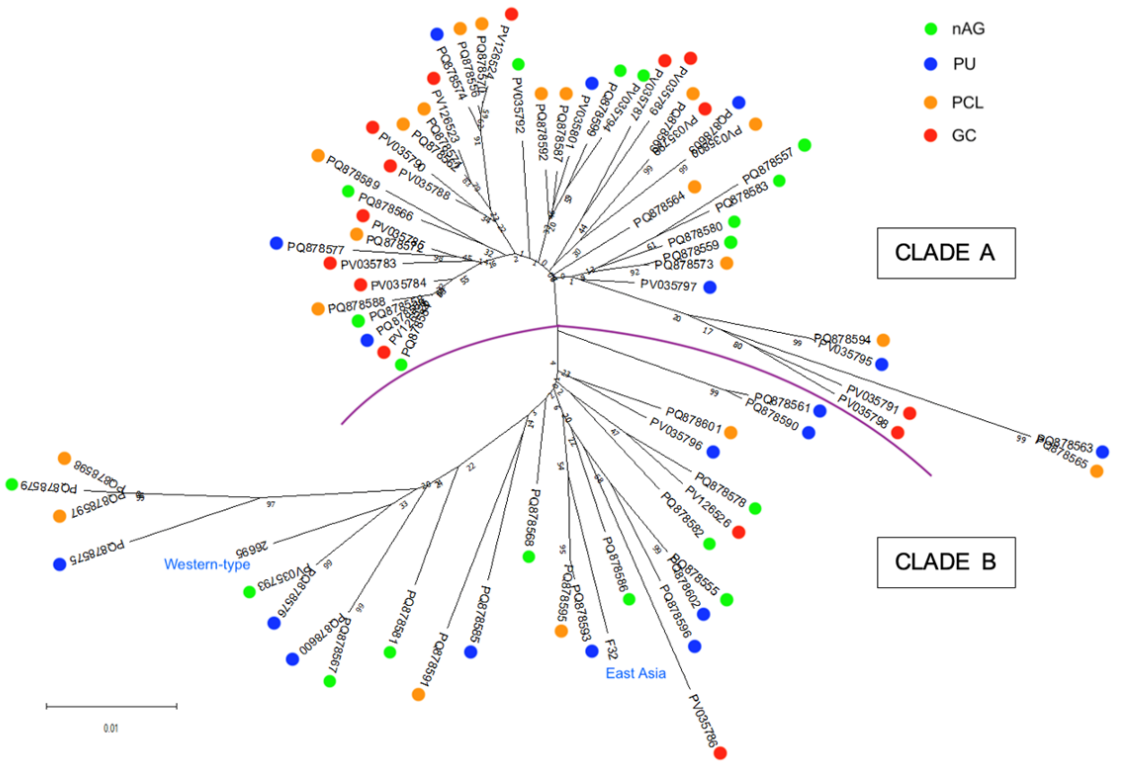
Residues 140 - 144	Number (%)	nAG (%)	PU (%)	PCL (%)	GC (%)	p-value
EAELQ	44 (61.11)	9 (47.37)	8 (44.44)	14 (70.00)	13 (86.67)	**0.038**
EGKLK	22 (30.56)	8 (42.11)	8 (44.44)	4 (20.00)	2 (13.33)	0.120
LGKLK	4 (5.56)	1 (5.26)	1 (5.56)	2 (10.00)	0 (0)	0.900
EAKLQ	2 (2.78)	1 (5.26)	1 (5.56)	0 (0)	0 (0)	0.710
Total	**72**	**19**	**18**	**20**	**15**	

We detected a novel CagL amino acid polymorphic pattern at residues 140 - 144, including four variants as EAELQ, EGKLK, LGKLK, and EAKLQ, which accounted for 61.11%, 30.56%, 5.56%, and 2.78%, respectively. Stratifying gastroduodenal diseases, the EAELQ variant rates in the GC and PCL groups were 86.67% and 70%, respectively, whereas those in the nAG and PU groups were only 47.37% and 44.44%, respectively. The difference between these rates was statistically significant (p = 0.038).

Table 4. Multivariable logistic regression analysis of factors associated with premalignant and malignant gastric lesions (PCL and GC) versus benign conditions (nAG and PU)

Variable	Unit	OR (95% CI)	p-value
Age	per 1-year increase	1.02 (0.99 - 1.05)	0.182
Gender	male vs female	0.76 (0.25 - 2.27)	0.620
EAELQ variant	present vs absent	**3.82 (1.40 - 11.20)**	**0.011**

The EAELQ variant was independently associated with premalignant and malignant gastric lesions, whereas age and gender were not significantly associated.



3.3. *CagL* phylogenetic tree

Topology-based phylogenetic clustering revealed a significant enrichment of GC and PCL within Clade A, whereas Clade B was predominantly associated with nAG and PU disease. This association was statistically significant, particularly nAG, PU, PCL and GC in Clade A were 10, 8, 15 and 13 patients, respectively; in Clade B were 9, 10, 5 and 2 patients, respectively (Fisher's exact test, p = 0.038). Notably, both reference strains F32 and 26695 were located within Clade B, suggesting that the Clade A enriched for GC and PCL represents a distinct *cagL* sequence lineage.

4. DISCUSSION

The genome of *H. pylori* exhibits considerable polymorphism, characterized by notable geographical variations, which explains the differences in pathogenicity observed between East Asian and Western-type strains. Although the CagA protein is a virulence factor that plays a crucial role in *H. pylori* pathogenesis, it cannot be used as a useful biomarker in East Asian strains [19]. In our study, all patients were infected with *cagA*-positive *H. pylori*. In term of *H. pylori* pathogenesis, the translocation of CagA protein into host gastric epithelial cells is necessary to initiate intracellular signal pathways. CagL protein, a T4SS component, is critical for interacting with integrin receptors, such as $\alpha_5\beta_1$, $\alpha_V\beta_5$, $\alpha_V\beta_3$, and $\alpha_V\beta_6$ on epithelial cell surfaces, allowing CagA to translocate [9,29-31]. Moreover, CagL can induce an inflammatory response in cells that have not undergone CagA translocation [11]. The primary site of binding for the interaction between CagL and integrin receptors is the RGD motif located at residues 76–78 [10]. The subsequent motif is the

LXXL motif, with the amino acid leucine at residues 79 and 82, which enhances the adherence of CagL to various cell lines through the integrin $\alpha_v\beta6$, playing a crucial role in this process [32]. Furthermore, an auxiliary sequence consisting of phenylalanine-glutamic-alanine-asparagine-glutamic (at residues 86–90), referred to as the FEANE motif, triggers the interaction between the RGD motif and integrins [31]. It is proposed that a motif consisting of the amino acids threonine, alanine, serine, leucine, and isoleucine (TASLI) at residues 170–174 has a role in integrin binding, separate from the RGD motif. The TASLI motif also has a role in maintaining the conformation of RGD due to its position opposite RGD [33]. Additionally, the CagL C-terminal region spanning residues 232–237, referred to as serine-lysine-isoleucine-isoleucine-valine-lysine (SKIIVK), facilitates interleukin-8 secretion [11,30]. In our investigation, all five of the aforementioned motifs exhibited high conservative feature and were not shown to be related to gastroduodenal diseases.

Notably, CagL contains a hypervariable motif (CagLHM) located at residues 58 - 62, which has been observed to have over 30 variants worldwide [9]. The current study identified nine variants, which include four globally prevalent variants as DKMGE, NEIGQ, NKIGQ, and DKIGK. Nevertheless, the NEIGQ and NKIGQ variants exhibited extremely low prevalence in our study. On the other hand, the DEIGK variant was unexpectedly identified as the most prevalent, accounting for 40.28% of our samples, although it was rather uncommon worldwide. Gorrell et al.[9] conducted a global analysis of CagLHM geographical diversity among 506 *H. pylori* strains, and the DEIGK variant was only found in six strains in Malaysia (three isolates), New Zealand (two isolates), and Taiwan (one isolate). The prevalence of the DEIGK variant, together with the minority presence of NEIGQ and NKIGQ variants, in the current study emphasized a unique characteristic of *H. pylori* strains in Vietnam. A number of publications have indicated a significant relationship between CagLHM and gastroduodenal diseases, particularly the Y58E59 polymorphism, which results in an increased risk of GC [29,34]. Additionally, the NKMGK variant has been found to be related to PUD [11]. In contrast, Caliskan et al (2022) demonstrated that new patterns of CagL polymorphism, specifically excluding CagLHM, were correlated with clinical outcomes al [35]. Our findings did not reveal an association between CagLHM polymorphism and gastroduodenal diseases.

Surprisingly, we found a sequence of five amino acids at residues 140–144 of the CagL that exhibited high polymorphism, with four types of variants: EAELQ, EAKLQ, EGKLK, and LGKLK. The predominant variants were EAELQ (61.11%) and EGKLK (30.56%). This study is the first to examine amino acid polymorphism at residues 140 - 144. Searching publications that showed the CagL sequence, we found that, among 309 analyzed *H. pylori* strains from other countries, EGKLK variant was the most frequent (297/309) and only three strains possessed EAELQ variant [10,11,35-38]. As a result, the region of residues 140 - 144 is highly conservative among *H. pylori* isolates worldwide, with the exception of Vietnam, which accounts for the lack of prior studies on this segment of CagL. Apart from us, only Ogawa et al. conducted an analysis of five distinct *H. pylori* strains obtained from five Vietnamese patients with GC [36]. The five strains possessed the AE variant at residues 141–142, whereas the other strains from various Southeast Asian countries carried the GK variant. Nevertheless, these investigators did not carry out research on *H. pylori* isolates from Vietnamese patients with other gastroduodenal disorders, and hence they could not draw a comprehensive conclusion.

The current study found a high polymorphism at residues 140 - 144, which is associated with gastroduodenal diseases. In particular, the EAELQ variant increased the risk of gastric carcinogenesis. Multivariable logistic regression analysis demonstrated that the EAELQ variant was independently associated with premalignant and malignant gastric lesions. Individuals infected with *H. pylori* carrying this variant had approximately 3.8-fold higher odds of disease compared with those without the variant, even after adjustment for age and gender. This finding suggests that the EAELQ variant at residues 140 - 144 of *H. pylori* CagL may play a role in the early stages of gastric carcinogenesis.

The amino acid sequence at residues 140 - 144 is part of CagL a_5 helix [33]. Amino acids, such as alanine (A), glutamic (E), leucine (L), glutamine (Q), and lysine (K), are amino acids with high helix-forming propensities. Notably, alanine has the highest helix-forming propensity, whereas glycine (G) has the lowest propensity [39, 40]. Therefore, the E**A**ELQ variant resulted in a more stable of the a_5 helix compared to the E**G**KLK variant. As a result of the increased CagA translocation capacity, *H. pylori* strains with the EAELQ variant were more hazardous than those with the EGKLK variant. Further research, particularly *in vivo* experiments, is required to support this hypothesis.

Additionally, a *cagL*-based phylogenetic analysis was performed to explore the relationship between *H. pylori* genetic variation and the spectrum of gastroduodenal diseases. We observed a non-random distribution of clinical phenotypes across the phylogenetic tree. The *cagL* sequences derived from patients with GC and patients with PCL were significantly enriched within Clade A, whereas sequences associated with nAG and PU were predominantly distributed within Clade B. This finding suggests that *cagL* sequence variants associated with advanced gastric lesions may form a distinct sequence lineage that is not well represented by commonly used reference strains.

This study presents an initial report on the genetic diversity of *cagL* gene, the association between the CagL amino acid polymorphism at residues 140–144, as well as genetic variation, and gastric carcinogenesis. Nonetheless, it has certain limitations, such as the small sample size. Additional research on a greater number of *H. pylori* isolates is necessary to ascertain whether the EAELQ variant may serve as a marker for the risk of gastric carcinogenesis. Future studies with larger and more diverse populations, including multicenter designs, are warranted to improve the generalizability of the findings. In addition, prospective cohort studies would be valuable to further evaluate the role of the EAELQ variant in disease progression. Despite these limitations, our findings provide preliminary evidence supporting the potential role of the *H. pylori* CagL EAELQ variant in gastric carcinogenesis.

5. CONCLUSION

This preliminary study highlighted the genetic diversity of the *H. pylori cagL* gene, revealing distinct polymorphism patterns unique to Vietnam. A novel amino acid polymorphism was identified at residues 140–144, which is associated with gastric carcinogenesis, with the EAELQ variant potentially serving as a novel indicator of gastric cancer risk.

ACKNOWLEDGEMENTS

This study was partially funded by Hong Bang International University, Vietnam, code SVTC17.01.

REFERENCES

1. Kao CY, Sheu BS, Wu JJ. Helicobacter pylori infection: An overview of bacterial virulence factors and pathogenesis. Biomed J. 2016;39(1):14-23.

2. Hooi JKY, Lai WY, Ng WK, Suen MMY, Underwood FE, Tanyingoh D, et al. Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. Gastroenterology. 2017;153(2):420-429.

3. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol 60.; 1994.

4. Ogorodnik E, Raffaniello RD. Analysis of the 3'-variable region of the cagA gene from Helicobacter pylori strains infecting patients at New York City hospitals. Microb Pathog. 2013;56:29-34.

5. Cover TL, Lacy DB, Ohi MD. The Helicobacter pylori Cag Type IV Secretion System. Trends Microbiol. 2020;28(8):682-695.

6. Plummer M, van Doorn LJ, Franceschi S, Kleter B, Canzian F, Vivas J, et al. Helicobacter pylori cytotoxin-associated genotype and gastric precancerous lesions. J Natl Cancer Inst. 2007;99(17):1328-1334.

7. Sheikh AF, Yadyad MJ, Goodarzi H, Hashemi SJ, Aslani S, Assarehzadegan MA, et al. CagA and vacA allelic combination of Helicobacter pylori in gastroduodenal disorders. Microb Pathog. 2018;122:144-150.

8. Hatakeyama M. Structure and function of Helicobacter pylori cagA, the first-identified bacterial protein involved in human cancer. Proc Jpn Acad Ser B Phys Biol Sci. 2017;93(4):196-219.

9. Gorrell RJ, Zwickel N, Reynolds J, Bulach D, Kwok T. Helicobacter pylori CagL Hypervariable Motif: a Global Analysis of Geographical Diversity and Association with Gastric Cancer Downloaded from. J Infect Dis. 2016;213(12):1927-1931.

10. Román-Román A, Martínez-Santos VI, Castañón-Sánchez CA, Albañil-Muñoz AJ, González-Mendoza P, Soto-Flores DG, et al. CagL polymorphisms D58/K59 are predominant in Helicobacter pylori strains isolated from Mexican patients with chronic gastritis. Gut Pathog. 2019;11(1).

11. Yadegar A, Mohabati Mobarez A, Zali MR. Genetic diversity and amino acid sequence polymorphism in Helicobacter pylori CagL hypervariable motif and its association with virulence markers and gastroduodenal diseases. Cancer Med. 2019;8(4):1619-1632.

12. Quach DT, Vilaichone RK, Van Vu K, Yamaoka Y, Sugano K, Mahachai V. Helicobacter pylori infection and related gastrointestinal diseases in southeast Asian countries: An expert opinion survey. Asian Pacific Journal of Cancer Prevention. 2018;19(12):3565-3569.

13. Quach DT, Mai BH, Tran MK, Dao LV, Tran HV, Vu KT, et al. Vietnam Association of Gastroenterology (VNAGE) consensus on the management of Helicobacter pylori infection. Front Med (Lausanne). 2023;9.

14. Tran VH, Nguyen TMN, Le PTQ, Nguyen THT, Nguyen TCL, Ha TMT. Current status of Helicobacter pylori resistance to clarithromycin and levofloxacin in Vietnam: Results from molecular analysis of gastric biopsy specimens. J Glob Antimicrob Resist. 2023;36:76-82.

15. Le NT, Van Nguyen T, Le LT, Nguyen LC. Dietary protein intake and stomach cancer, insights from a case-control study. Sci Rep. 2025;15:6909.

16. Jones KR, Whitmire JM, Merrell DS. A tale of two toxins: Helicobacter pylori CagA and VacA modulate host pathways that impact disease. Front Microbiol. 2010;1(NOV).

17. Yamaoka Y. Mechanisms of disease: Helicobacter

pylori virulence factors. Nat Rev Gastroenterol Hepatol. 2010;7(11):629-641.

18. Zhu X, Zhao Y, Zhu C, Wang Y, Liu Y, Su J. Rapid detection of cagA-positive Helicobacter pylori based on duplex recombinase aided amplification combined with lateral flow dipstick assay. Diagn Microbiol Infect Dis. 2022;103(1):115661.

19. Maeda S, Yoshida H, Kanai F, Ikenoue T, Kato N, Shiratori Y, et al. Major virulence factors, VacA and CagA, are commonly positive in Helicobacter pylori isolates in Japan. Gut. 1998;42:338-343.

20. Yamaoka Y, Kodama T, Gutierrez O, Kim JG, Kashima K, Graham DY. Relationship between Helicobacter priori iceA, cagA, and vacA status and clinical outcome: Studies in four different countries. J Clin Microbiol. 1999;37(7):2274-2279.

21. Chomvarin C, Namwat W, Chaicumpar K, Mairiang P, Sangchan A, Sripa B, et al. Prevalence of Helicobacter pylori vacA, cagA, cagE, iceA and babA2 genotypes in Thai dyspeptic patients. International Journal of Infectious Diseases. 2008;12(1):30-36.

22. Nguyen TMN, Tran VH, Ha TMT. Helicobacter pylori cagA, vacA, and iceA genotypes and clinical outcomes: a cross-sectional study in central Vietnam. Brazilian Journal of Microbiology. 2024.

23. Nguyen TH, Ho TTM, Nguyen-Hoang TP, et al. The endemic Helicobacter pylori population in Southern Vietnam has both South-East Asian and European origins. Gut Pathog. 2021;13(1).

24. Thai THN, Nguyen HP, Nguyen THY, Nguyen TBH, Nguyen TH, Nguyen TMN, et al. Genetic diversity of the oipA gene among Helicobacter pylori isolates and clinical outcome in Vietnam. Infection, Genetics and Evolution. 2023;112.

25. Nguyen TL, Uchida T, Tsukamoto Y, Trinh DT, Ta L, Mai BH, et al. Helicobacter pylori infection and gastroduodenal diseases in Vietnam: A cross-sectional, hospital-based study. BMC Gastroenterol. 2010;10:no pagination.

26. Dixon MF, Genta RM, Yardley JH, Correa P. Classification and grading of gastritis: the updated Sydney system. Am J Surg Pathol. 1996;20(10):1161-1181.

27. Kushima R, Lauwers GY, Rugge M. Gastric dysplasia. In: WHO Classification of Tumours Editorial Board, ed. WHO Classification of Tumours: Digestive Systemic Tumours. Lyon: International Agency for Research on Cancer; 2019:71-75.

28. Peek RM, Miller GG, Tham KT, Pé GI, Rez RP, Cover TL, et al. Detection of Helicobacter pylori Gene Expression in Human Gastric Mucosa. J Clin Microbiol. 1995;33(1):28-32.

29. Bickley J, Owen RJ, Fraser AG, Pounder RE. Evaluation of the polymerase chain reaction for detecting the urease C gene of Helicobacter pylori in gastric biopsy samples and dental plaque. J Med Microbiol. 1993;39(5):338-344.

30. Tran VH, Ha TMT, Le PTQ, Nguyen VN, Phan TN, Paglietti B. Helicobacter pylori 23S rRNA gene mutations associated with clarithromycin resistance in chronic gastritis in Vietnam. The Journal of Infection in Developing Countries. 2018;12(7):526-532.

31. Jones KR, Joo YM, Jang S, Yoo YJ, Lee HS, Chung IS, et al. Polymorphism in the cagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol. 2009;47(4):959-968.

32. Yamaoka Y, Malaty HM, Michael SO, Graham DY. Conservation of Helicobacter pylori Genotypes in Different Ethnic Groups in Houston, Texas. J Infect Dis. 2000;181:2083-2086.

33. Phan TN, Santona A, Tran VH, Tran TNH, Le VA, Cappuccinelli P, et al. Genotyping of Helicobacter pylori shows high diversity of strains circulating in central Vietnam. Infection, Genetics and Evolution. 2017;52:19-25.

34. Truong BX, Mai VTC, Tanaka H, Ly LT, Thong TM, Hai HH, et al. Diverse characteristics of the cagA gene of Helicobacter pylori strains collected from patients from Southern Vietnam with gastric cancer and peptic ulcer. J Clin Microbiol. 2009;47(12):4021-4028.

35. Yeh YC, Cheng HC, Yang HB, Chang WL, Sheu BS. H. pylori CagL-Y58/E59 Prime Higher Integrin α5β1 in Adverse pH Condition to Enhance Hypochlorhydria Vicious Cycle for Gastric Carcinogenesis. PLoS One. 2013;8(8).

36. Wiedemann T, Hofbaur S, Loell E, Rieder G. A C-terminal coiled-coil region of CagL is responsible for Helicobacter pylori-induced IL-8 expression. Eur J Microbiol Immunol (Bp). 2016;6(3):186-196.

37. Conradi J, Tegtmeyer N, Wozna M, Wissbrock M, Michalek C, Gagell C, et al. An RGD helper sequence in CagL of Helicobacter pylori assists in interactions with integrins and injection of CagA. Front Cell Infect Microbiol. 2012;2(70):1-11.

38. Barden S, Niemann HH. Adhesion of several cell lines to helicobacter pylori CagL Is mediated by integrin αvβ6 via an RGDLXXL motif. J Mol Biol. 2015;427(6):1304-1315.

39. Bönig T, Olbermann P, Bats SH, Fischer W, Josenhans C. Systematic site-directed mutagenesis of the Helicobacter pylori CagL protein of the Cag type IV secretion system identifies novel functional domains. Sci Rep. 2016;6.

40. Yeh YC, Chang WL, Yang HB, Cheng HC, Wu JJ, Sheu BS. H. pylori CagL amino acid sequence polymorphism Y58E59 induces a corpus shift of gastric integrin α5β1 related with gastric carcinogenesis. Mol Carcinog. 2011;50(10):751-759.

41. Caliskan R, Polat Sari S, Ercan B, Peker, KD, Omac-Sonmez M, Akgul O, et al. New CagL Amino Acid Polymorphism Patterns of Helicobacter pylori in Peptic Ulcer and Non-Ulcer Dyspepsia. Medicina (Lithuania). 2022;58(12).

42. Ogawa H, Iwamoto A, Tanahashi T, Okada R, Yamamoto K, Nishiumi S, et al. Genetic variants of Helicobacter pylori type IV secretion system components CagL and CagI and their association with clinical outcomes. Gut Pathog. 2017;9(1).

43. Choi YH, Lai J, Kim MA, Kim A, Kim J, Su H, et al. CagL polymorphisms between East Asian and Western Helicobacter pylori are associated with different abilities to induce IL-8 secretion. Journal of Microbiology.

2021;59(8):763-770.

44. Yeh YC, Kuo HY, Chang WL, Yang HB, Lu CC, Cheng HC, et al. H. pylori isolates with amino acid sequence polymorphisms as presence of both HtrA-L171 & CagL-Y58/E59 increase the risk of gastric cancer. J Biomed Sci. 2019;26(1).

45. Glatzová D, Mavila H, Saija MC, Chum T, Cwiklik L, Brdička T, et al. The role of prolines and glycine in the transmembrane domain of LAT. FEBS Journal. 2021;288(13):4039-4052.

46. Pace CN, Scholtz JM. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. Biophys J. 1998;75:422-427.