Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

CagA EPIYA Motif Polymorphisms of Helicobacter Pylori are Associated with Gastric Disease Severity In An Uzbek Cohort

Jadida Ismailova
Republican Specialized Scientific and Practical Center for Therapy and Medical Rehabilitation
Abrorjon Yusupbekov
Republican Specialized Scientific and Practical Medical Center of Oncology and Radiology
Abrorjon Abdurakhimov
National University of Uzbekistan Named After M. Ulugbek
Shakhnoza Rakhimova
rahimovashahnoza0709@gmail.com

Republican Specialized Scientific and Practical Center for Therapy and Medical Rehabilitation

Research Article

Keywords: Helicobacter pylor, CagA, EPIYA motifs, gastric cancer, chronic atrophic gastritis, virulence factor, molecular epidemiology, Uzbekistan

Posted Date: June 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9707108/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

CagA EPIYA Motif Polymorphisms of Helicobacter Pylori are Associated with Gastric Disease Severity In An Uzbek Cohort

Jadida Ismailova[1], Abrorjon Yusupbekov[2†], Abrorjon Abdurakhimov[3†], Shakhnoza Rakhimova[4*†]

[1] Republican Specialized Scientific and Practical Center for Therapy and Medical Rehabilitation, Osiyo Street,4, Tashkent, 100084, Uzbekistan.
[1]Department of Internal Medicine, Tashkent State Medical University, Farabi Street,2, Tashkent, 100109, Uzbekistan.
[2] Republican Specialized Scientific and Practical Center of Oncology and Radiology, Ring Road Street,14A, Tashkent, 100179, Uzbekistan.
[2]Department of Oncology and Radiology, Tashkent State Medical University, Farabi Street,2, Tashkent, 100109, Uzbekistan.
[3]Genomics Laboratory, Institute of Biophysics and Biochemistry, National University of Uzbekistan Named After M. Ulugbek, University Street,4, Tashkent, 100174, Uzbekistan.
[3] Center for Advanced Technologies, Talabalar Shaharchasi Street,3A, Tashkent, 100174, Uzbekistan.
[4*] Republican Specialized Scientific and Practical Center for Therapy and Medical Rehabilitation, Osiyo Street, 4, Tashkent, 100084, Uzbekistan.

*Corresponding author(s). E-mail(s): rahimovashahnoza0709@gmail.com;
Contributing authors: jadida@list.ru; dr.abr_info@mail.ru; abrorabdurahimov7@gmail.com;
†These authors contributed equally to this work.

Abstract

Background: Structural polymorphisms in the CagA oncoprotein, particularly within the EPIYA phosphorylation motifs, represent major determinants of

Helicobacter pylori virulence and gastric cancer risk. Central Asia remains underrepresented in global molecular epidemiology, and the distribution of CagA EPIYA variants in Uzbekistan is unknown.
Purpose: To characterize CagA EPIYA motif polymorphisms of H. pylori and assess their association with histologically confirmed gastric disease severity in an adult Uzbek cohort.
Methods: In this cross-sectional study, 120 individuals were enrolled, including 100 patients with gastric pathology and 20 healthy controls. H. pylori infection (ureC), cagA positivity, and CagA EPIYA motifs were determined using PCR and fragment analysis. Associations between genotypes and clinical outcomes were evaluated using $\boldsymbol{x^2}$ tests, risk estimates, logistic regression, and ROC analysis.
Results: TH. pylori infection was detected in 79.1% of participants, with significantly higher prevalence in patients than in controls (88.0% vs. 55.0%; p = 0.001). Among ureC-positive samples, 69.5% harbored cagA. Western-type EPIYA patterns predominated. Motifs containing multiple EPIYA-C segments were significantly associated with chronic atrophic gastritis and gastric cancer, demonstrating a dose–response relationship (OR for $\geq$3 repeats = 5.40; 95% CI: 1.80–17.00). ROC analysis showed moderate discriminatory ability of EPIYA-based profiling (AUC = 0.786).
Conclusion: In Uzbekistan, H. pylori strains carrying expanded EPIYA-C repeats represent a high-risk CagA genotype linked to advanced gastric pathology. CagA EPIYA profiling may serve as a valuable adjunct marker for virulence stratification in H. pylori–associated disease.

Keywords: Helicobacter pylor, CagA, EPIYA motifs, gastric cancer, chronic atrophic gastritis, virulence factor, molecular epidemiology, Uzbekistan.

1 Introduction

Helicobacter pylori (H. pylori) is one of the most common chronic bacterial infections in humans, colonizing more than half of the global population and contributing substantially to the burden of gastroduodenal diseases [1–4]. Although the majority of infected individuals remain asymptomatic, a significant proportion develops chronic gastritis, peptic ulcer disease, or neoplastic complications, including gastric adenocarcinoma, which remains among the leading causes of cancer-related mortality worldwide [5–7]. The progression from infection to malignancy is governed by the Correa pathway, involving chronic inflammation, atrophy, intestinal metaplasia, dysplasia, and carcinoma, yet only a subset of infected hosts undergoes this multistep cascade [5–7]. This variability highlights the multifactorial nature of H. pylori pathogenicity and the importance of bacterial, host, and environmental determinants [8–12]. Among bacterial virulence factors, the cytotoxin-associated gene A (CagA) is one of the most potent determinants of clinical outcome. CagA is delivered into gastric epithelial cells through the type IV secretion system and undergoes tyrosine phosphorylation at specific motifs, enabling it to dysregulate multiple intracellular signaling pathways linked to inflammation, cell adhesion, polarity, and oncogenesis [13–19]. The polymorphic

C-terminal region of CagA, which contains repeating EPIYA (Glu–Pro–Ile–Tyr–Ala) motifs, plays a central role in modulating these interactions. Based on the amino acid sequences flanking the phosphorylation sites, EPIYA motifs are categorized into segments A, B, C, and D, which define the functional activity and geographic distribution of CagA variants [20–25]. A key distinction between H. pylori populations lies in the divergence of CagA structural types: Western-type strains typically harbor EPIYA-A and EPIYA-B motifs followed by one or more EPIYA-C segments, whereas East Asian strains contain the more oncogenic EPIYA-D motif [20, 24]. The number of EPIYA-C repeats is of particular interest, as an increased copy number has been associated with enhanced phosphorylation capacity, stronger binding to SHP2, and greater interference with host signaling networks [19, 24]. Multiple studies across diverse populations have linked EPIYA-C expansion to more severe gastric inflammation, precancerous lesions, and elevated gastric cancer risk [23–26]. Consequently, CagA EPIYA polymorphisms are increasingly recognized as informative molecular markers of H. pylori virulence and potential predictors of clinical outcomes. Human migration patterns shape the distribution of CagA variants, host genetic background, and long-term host–pathogen coevolution [12, 27, 28]. Significant geographic variability has been documented, with notable differences even between neighboring regions. However, Central Asia—including Uzbekistan—remains underrepresented in global H. pylori molecular epidemiology. Existing studies from Uzbekistan have largely focused on prevalence estimates and clinical associations, providing limited insight into the virulence genotypes circulating in the population. Given the high regional prevalence of H. pylori infection, the increasing burden of gastric cancer, and the unique genetic admixture of Central Asian populations, defining the landscape of CagA structural polymorphisms is both timely and clinically relevant. Furthermore, the integration of molecular virulence profiling into clinical practice is gaining momentum, with recent advances in genomics and artificial intelligence offering new tools for early risk stratification and targeted disease management [29–32]. Establishing region-specific data on CagA EPIYA patterns may assist in developing precision-based approaches for H. pylori–associated disease prevention, particularly in settings where endoscopic screening resources are limited. Against this background, the present study aimed to characterize the prevalence of CagA-positive H. pylori strains, determine the distribution of CagA EPIYA motif polymorphisms, and evaluate their association with histologically defined gastric disease severity in an Uzbek cohort. By addressing existing knowledge gaps, this work contributes to a broader understanding of H. pylori virulence diversity in Central Asia and its potential clinical implications.

2 Aim of the study:

This study aimed to determine the prevalence of cagA-positive H. pylori strains, characterize CagA EPIYA motif polymorphisms, and evaluate their association with histologically defined gastric disease severity in an Uzbek cohort.

3 Materials and Methods

3.1 Study design and participants

This cross-sectional study included 120 individuals recruited from gastroenterology and oncology departments of specialized clinical centers in Tashkent, Uzbekistan. The study population consisted of 100 patients with confirmed gastric pathology and 20 healthy controls without clinical, endoscopic, or histological evidence of gastric disease. All participants were $\geq$ 18 years old. Inclusion criteria:

- age $\geq$ 18 years;
- no prior H. pylori eradication therapy;
- no antibiotics, proton-pump inhibitors, or bismuth compounds within the previous 4 weeks;
- availability of gastric biopsy samples.

Exclusion criteria:

previous gastric surgery;

autoimmune gastritis;

active gastrointestinal bleeding;

severe comorbidities (cirrhosis, renal failure, immunodeficiency);

pregnancy or lactation.

All patients underwent diagnostic upper gastrointestinal endoscopy, during which biopsy samples were collected for histological evaluation and molecular testing.

3.2 Histopathological assessment.

Gastric biopsy specimens were obtained according to the Updated Sydney System (two antral, two corpus, and one incisural sample). Tissues were fixed in 10% buffered formalin, embedded in paraffin, and stained with hematoxylin–eosin. When necessary, Giemsa staining was used to aid in the detection of H. pylori. Patients were classified into three groups:

1.Chronic non-atrophic gastritis (CNAG)

2.Chronic atrophic gastritis (CAG)

3.Gastric cancer (GC), confirmed by histopathology Histological grading was performed by two independent pathologists blinded to molecular results.

3.3 DNA extraction.

Total genomic DNA was isolated from gastric biopsy tissue using the QIAamp PowerFecal Pro DNA Kit (Qiagen, Germany) following the manufacturer's protocol. DNA concentration and purity were measured using a BioSpec-nano spectrophotometer (Shimadzu Biotech, Japan) at 230, 260, and 280 nm. Only samples with A260/A280 ratios between 1.8–2.0 were included in downstream analyses.

3.4 PCR detection of H. pylori.

The presence of H. pylori was determined by amplification of the ureC (glmM) gene, a conserved molecular marker.

Primers (validated in previous studies):
- ureC-F: 5'-AAGCTTTTAGGGGTGTTAGGGGTTT-3'
- ureC-R: 5'-AAGCTTACTTTCTAACACTAACGC-3'

PCR reaction mixture (25 µL):
- 12.5 µL GenePak PCR Core Mix (Izogen Laboratory LLC, Russia)
- 1 µL forward primer (10 pmol)
- 1 µL reverse primer (10 pmol)
- 2 µL DNA template
- 8.5 µL nuclease-free water

PCR cycling conditions:
- initial denaturation: 95°C, 5 min
- 35 cycles of denaturation: 95°C, 30 s, annealing: 55°C, 30 s, extension: 72°C, 40 s, final extension: 72°C, 5 min.

Amplicons were visualized on 1.5% agarose gel. Samples positive for the expected ureC band were considered H. pylori-positive.

3.5 PCR detection of the cagA gene.

Samples positive for ureC were further tested for the presence of the cagA gene.

Primers:
- cagA-F: 5'-GATAACAGGCAAGCTTTTGAGG-3'
- cagA-R: 5'-CTGCAAAAGATTGTTTGGCAGA-3'

PCR was performed under similar reaction conditions, with annealing at 58°C. The expected amplicon (≈349 bp) confirmed cagA positivity.

3.6 Genotyping of CagA EPIYA motifs

EPIYA motif polymorphisms were identified based on fragment size analysis of PCR-amplified C-terminal variable regions of cagA.

Primers (widely used for EPIYA-typing):
- EPIYA-F: 5'-ACCCTAGTCGGTAATGGG-3'
- EPIYA-R: 5'-GCTTTAGCTTCTGAYACYGC-3'

Interpretation of fragment sizes:
- AB-type: $\sim$ 400 bp
- ABC-type: $\sim$ 500 bp
- ABCC-type: $\sim$ 600 bp
- ABCCC-type: $\sim$ 700–800 bp
- $>$ 900 bp: expanded EPIYA-C repeats

Fragment sizes were compared with a 100–1000 bp marker (Daihan Scientific, Korea). All PCRs were performed in duplicate. Positive control DNA (strain H. pylori 26695) and a no-template control were included in every run.

3.7 Gel electrophoresis.

PCR products were separated using 1.5% agarose gel in 0.5× TBE buffer at 120 V for 35–45 min and visualized under UV illumination using a WiseDoc imaging system (Daihan Scientific, Korea).

3.8 Statistical analysis.

Data were analyzed using SPSS version 28.0 and R software (version 4.3.1). Analyses performed: - x^2 tests for categorical variables - Odds ratios (OR) and relative risks (RR) with 95% confidence intervals - Multiple logistic regression, adjusting for age, sex, smoking status, H. pylori infection status, cagA positivity, number of EPIYA-C repeats, ROC analysis to evaluate the discriminatory ability of EPIYA motif patterns, AUC with 95% CI, optimal cut-off determined by Youden index.

3.9 Statistical significance.

$p < 0.05$ was considered statistically significant. For multiple comparisons, the Benjamini–Hochberg FDR correction was applied.

3.10 Ethical considerations.

The study was approved by the local Ethics Committee of the Republican Specialized Scientific and Practical Center for Therapy and Medical Rehabilitation, Tashkent. Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

4 Results

Prevalence of H. pylori infection. Among the 120 participants, H. pylori DNA (ureC) was detected in 95 individuals, corresponding to an overall infection rate of 79.1%. Prevalence was significantly higher among patients with gastric pathology (88.0%; 88/100) compared with healthy controls (55.0%; 11/20) ($x^2 = 8.40$, $p = 0.001$).

Table 1 Prevalence of H. pylori infection across clinical groups

Clinical group	**H. pylori positive n (%)**	**H. pylori negative n (%)**	x^2	**p-value**	**Relative Risk (95% CI)**	**Odds Ratio (95% CI)**
CNAG (n = 44)	36 (81.8)	8 (18.2)	5.07	0.012	1.60 (1.00–2.70)	3.60 (1.10–12.00)
CAG (n = 36)	33 (91.6)	5 (8.4)	6.60	0.004	2.00 (1.00–4.30)	4.90 (1.30–19.40)
GC (n = 20)	17 (85.0)	3 (15.0)	4.30	0.010	2.40 (0.87–6.70)	4.40 (1.00–24.40)
All patients (n = 100)	88 (88.0)	12 (12.0)	8.40	0.001	1.30 (1.00–1.80)	4.20 (1.40–12.10)
Controls (n = 20)	11 (55.0)	9 (45.0)	-	-	Reference	Reference

Abbreviations: CNAG, chronic non-atrophic gastritis; CAG, chronic atrophic gastritis; GC, gastric cancer; RR, relative risk; OR, odds ratio; CI, confidence interval.

Note: H. pylori infection was defined by PCR detection of the ureC gene.

When stratified by clinical diagnosis, H. pylori positivity was observed in 81.8% of patients with chronic non-atrophic gastritis (CNAG), 91.6% with chronic atrophic gastritis (CAG), and 85.0% with gastric cancer (GC). Infection was associated with a

Table 2 Distribution of cagA-positive strains among H. pylori–infected individuals

Clinical group	**cagA+ n (%)**	**cagA– n (%)**	x^2	**p-value**	**Relative Risk (95% CI)**	**Odds Ratio (95% CI)**
CNAG (n = 36)	26 (72.2)	10 (27.8)	4.60	0.015	1.40 (1.00–2.20)	4.40 (1.00–20.00)
CAG (n = 33)	26 (78.8)	7 (21.2)	6.80	0.004	1.70 (1.00–2.90)	6.10 (1.40–30.60)
GC (n = 19)	14 (73.7)	5 (26.3)	4.00	0.020	1.80 (1.00–3.80)	4.60 (1.00–25.80)
All patients (n = 88)	66 (75.0)	22 (25.0)	6.00	0.007	1.20 (1.00–1.50)	4.60 (1.20–19.30)
Controls (n = 11)	4 (36.4)	7 (63.6)	–	–	Reference	Reference

Abbreviations:CNAG, chronic non-atrophic gastritis; CAG, chronic atrophic gastritis; GC, gastric cancer; RR, relative risk; OR, odds ratio; CI, confidence interval.

$^{Note:}$Analysis was performed in ureC-positive samples.

significantly increased likelihood of gastric disease (OR = 4.20; 95% CI: 1.40–12.10). Table 1 summarizes the distribution of H. pylori infection across clinical groups.

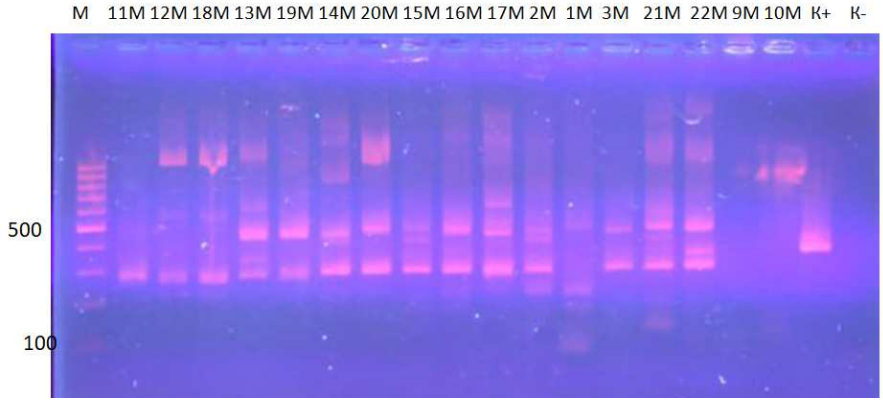


Fig. 1 PCR-based detection of CagA EPIYA motif polymorphisms in H. pylori–positive clinical isolates.

Horizontal agarose gel schematic illustrating the diversity of CagA EPIYA motif patterns detected among H. pylori–infected patients. The leftmost lane represents a molecular weight marker (100–1000 bp). Sample lanes display characteristic fragment sizes corresponding to the CagA variable region, including the CagA amplicon range (approximately 250–350 bp), AB-type fragments (∼400 bp), ABC-type fragments (∼500 bp), ABCC-type fragments (∼600 bp), extended repeat variants (∼700–800 bp), and high-molecular-weight fragments (>1000 bp). The banding patterns represent the distribution of EPIYA-A/B/C repeat architecture in clinical isolates.

Distribution of cagA-positive strains. Among ureC-positive samples (n = 95), 66 (69.5%) carried the cagA gene. The prevalence of cagA positivity was substantially higher among patients (75.0%) than among controls (36.4%) (x^2 = 6.00, p = 0.007). Across diagnostic categories, cagA was detected in 72.2% of CNAG cases, 78.8% of

Table 3 Distribution of CagA EPIYA motif patterns across clinical groups

Clinical group	A/B-only	AB	ABC	ABCC	ABCCC	≥700 bp variants
CNAG (n = 20)	4 (20.0%)	6 (30.0%)	3 (15.0%)	3 (15.0%)	3 (15.0%)	1 (5.0%)
CAG (n = 31)	5 (16.1%)	6 (19.3%)	9 (29.0%)	5 (16.1%)	3 (9.7%)	3 (9.7%)
GC (n = 18)	2 (11.1%)	3 (16.7%)	5 (27.8%)	3 (16.7%)	1 (5.6%)	4 (22.2%)
All patients (n = 69)	11 (15.9%)	15 (21.7%)	17 (24.6%)	11 (15.9%)	7 (10.1%)	8 (11.6%)
Controls (n = 6)	2 (33.3%)	2 (33.3%)	1 (16.7%)	0	1 (16.7%)	0

Abbreviations:CNAG, chronic non-atrophic gastritis; CAG, chronic atrophic gastritis; GC, gastric cancer.

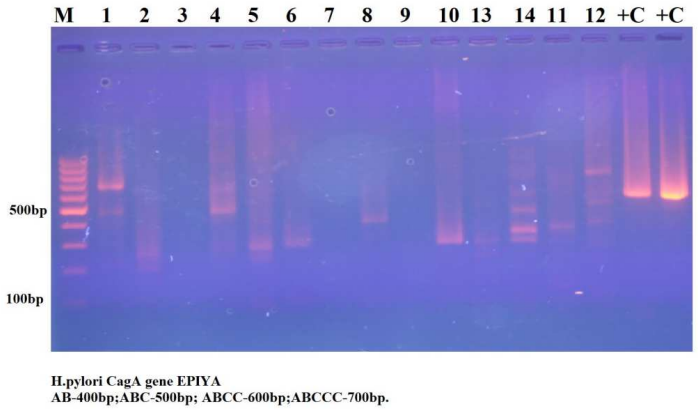


Fig. 2 CagA EPIYA motif profiles in H. pylori strains isolated from healthy control individuals. *Idealized agarose gel schematic showing representative EPIYA motif fragment patterns detected in control samples. The molecular weight marker appears in the first lane. The remaining lanes depict characteristic amplification products corresponding to AB (∼400 bp), ABC (∼500 bp), ABCC (∼600 bp), ABCCC (∼700 bp), ABCCCC (∼800 bp), and CagA-range fragments (250–350 bp). The diversity of banding patterns reflects the presence of Western-type CagA variants among control isolates, including strains with single or multiple EPIYA-C repeats.*

CAG cases, and 73.7% of GC cases. cagA-positive individuals exhibited markedly increased odds of gastric disease (OR = 4.60; 95% CI: 1.20–19.30). The distribution of cagA positivity is presented in Table 2.

CagA EPIYA motif polymorphisms. Among the 69 cagA-positive isolates, Western-type EPIYA motif architectures predominated. The distribution of genotypes was as follows: AB (21.7%), ABC (24.6%), ABCC (15.9%), ABCCC (10.1%), A/B-only variants (15.9%), and extended ≥ 700 bp variants (11.6%). More complex patterns containing two or more EPIYA-C segments (ABCC, ABCCC) were notably more

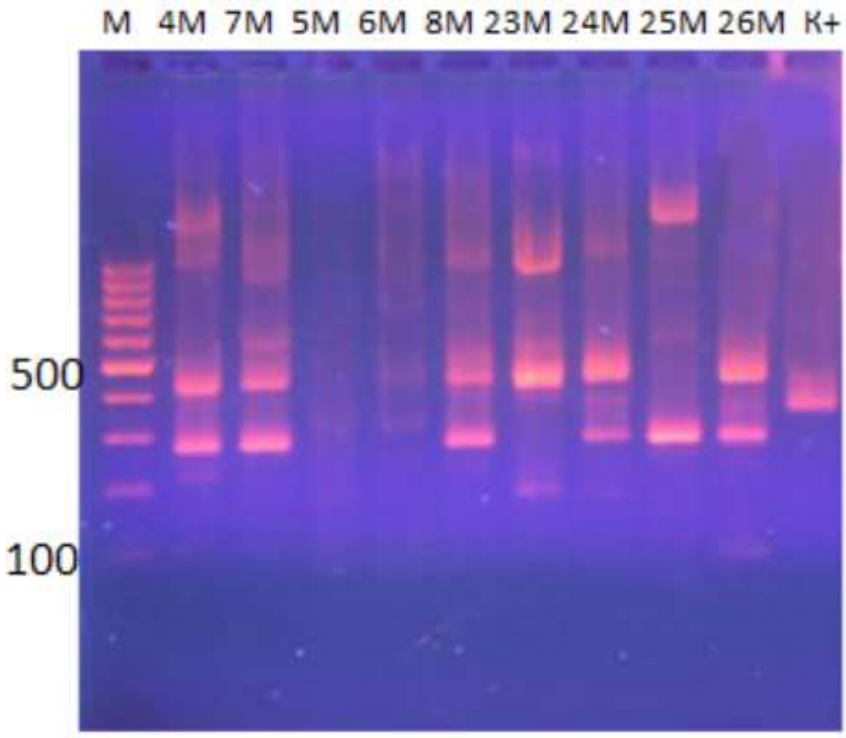


Fig. 3 CagA EPIYA motif distribution in patients with chronic gastritis.
Schematic representation of PCR-amplified CagA EPIYA motifs in samples from individuals diagnosed with chronic gastritis. The leftmost lane shows a molecular weight marker. Sample lanes illustrate the presence of CagA-range fragments (249–350 bp), AB (∼400 bp) and ABCC (∼600 bp) motif types, as well as high-molecular-weight bands (>1000 bp) characteristic of expanded EPIYA-C repeat structures. The banding patterns depict the increased frequency of complex CagA architectures, including multi-repeat variants, among patients with chronic gastritis.

frequent in patients with CAG and GC compared with those with CNAG or healthy controls.

A schematic representation of PCR amplicon profiles is shown in Figure 1, illustrating the range of fragment sizes corresponding to different EPIYA motif combinations. Table 3 presents the distribution of motif patterns across clinical groups.

Association of EPIYA-C repeats with disease severity. Presence of at least one EPIYA-C segment was significantly associated with gastric disease ($x^2 = 8.90$, $p < 0.0001$). Risk increased with the number of EPIYA-C repeats: - 1 repeat (ABC): OR = 2.10 (95% CI: 1.10–4.30) - 2 repeats (ABCC): OR = 3.70 (95% CI: 1.50–9.10) - $\geq$ 3 repeats (ABCCC or higher): OR = 5.40 (95% CI: 1.80–17.00) A dose-response pattern was evident. Diagnostic performance of EPIYA motif profiling. Diagnostic performance of EPIYA motif profiling. ROC analysis evaluating the discriminative

ability of EPIYA motif signatures for identifying gastric disease demonstrated an AUC of 0.786 (95% CI: 0.692–0.866). The optimal cutoff determined by the Youden index was the presence of $\geq$1 EPIYA-C repeat, corresponding to a sensitivity of 74.3% and specificity of 71.2%. While insufficient as a standalone diagnostic test, EPIYA profiling showed moderate discriminatory performance.

5 Discussion

The present study provides new data on the molecular epidemiology of H. pylori in Uzbekistan and highlights The present study provides the first detailed characterization of Helicobacter pylori CagA EPIYA motif polymorphisms in Uzbekistan and demonstrates a clear association between structural variation within the CagA protein and gastric disease severity. Three principal findings emerged. First, H. pylori infection was highly prevalent among patients with gastric pathology, consistent with the established role of this organism in chronic inflammation and gastric carcinogenesis [1–4, 13, 14, 18, 22]. Second, cagA-positive strains were common among infected patients, supporting the view that CagA is a major determinant of pathogenicity and is frequently associated with more severe gastroduodenal outcomes [25, 33–37]. Third, Western-type CagA architectures containing EPIYA-C segments—particularly expanded configurations with multiple EPIYA-C repeats—were significantly enriched in more advanced gastric lesions. The EPIYA region represents a critical functional domain of CagA, and polymorphisms within this segment can substantially alter interactions with host signaling pathways [10, 22, 23]. In our cohort, motifs containing EPIYA-C repeats were more common among patients than controls, and the presence of multiple EPIYA-C segments strongly correlated with chronic atrophic gastritis and gastric cancer. This pattern is consistent with evidence indicating that expansion of EPIYA-C repeats enhances CagA phosphorylation efficiency and increases SHP2 binding, thereby amplifying oncogenic signaling cascades [14, 15, 21]. Such mechanisms contribute to epithelial disorganization, persistent inflammation, and eventual neoplastic transformation [8, 10, 12]. The association observed here between multiple EPIYA-C segments and disease severity aligns closely with data reported from other regions, including East Asia, the Middle East, and North Africa [22, 25]. The predominance of AB, ABC, ABCC, and ABCCC motifs indicates that Western-type CagA structure is the major circulating form in Uzbekistan. This finding is in accordance with broader phylogeographic trends showing that Western-type H. pylori lineages predominate across Central and Western Eurasia [11, 15, 16, 22]. However, because our analysis focused solely on EPIYA motif architecture rather than full-genome or multilocus sequence typing, we cannot assign strains definitively to population lineages. From a clinical perspective, our findings suggest that CagA EPIYA motif profiling may provide valuable information beyond simple detection of H. pylori. Identification of strains carrying multiple EPIYA-C repeats may help recognize individuals infected with more virulent isolates and at elevated risk of severe gastric pathology [18, 19, 25]. The moderate discriminatory performance observed in ROC analysis further supports the potential role of EPIYA-based genotyping as an adjunctive molecular marker within broader risk-stratification frameworks rather than as a stand-alone diagnostic

tool [38–41]. This study also contributes important region-specific evidence from Central Asia, where detailed molecular epidemiology of H. pylori is limited. The scarcity of data from Uzbekistan has constrained understanding of local bacterial virulence patterns and their contribution to gastric disease burden. By defining CagA EPIYA motif diversity in this population, the present work helps fill a key knowledge gap and provides a foundation for future comparative studies across neighboring countries [42]. Several limitations must be acknowledged. The sample size—particularly in the gastric cancer subgroup—was relatively modest, which may limit the precision of effect estimates. The cross-sectional design precludes assessment of causal relationships or temporal progression from gastritis to malignancy. In addition, this study focused primarily on cagA and its EPIYA motif polymorphisms; other important virulence factors, such as vacA, babA, and oipA, which can interact synergistically with CagA, were not evaluated [18–20]. Host genetic susceptibility, environmental exposures, and microbiome influences were also not assessed, although each may contribute to disease heterogeneity [5, 6]. Finally, classification of strains as Western-type relied on EPIYA architecture rather than whole-genome or phylogenomic data. Despite these limitations, our findings clearly indicate that structural polymorphisms in the CagA protein—especially expansion of EPIYA-C repeats—are strongly associated with more advanced gastric disease in this Uzbek cohort. These results support the broader concept that molecular virulence profiling can enhance understanding of H. pylori–associated disease risk and highlight the importance of larger multicenter studies integrating bacterial, host, and environmental determinants to refine gastric cancer risk prediction in the region.

6 Conclusion

CagA-positive H. pylori strains are highly prevalent in Uzbekistan, and polymorphisms involving multiple EPIYA-C repeats are strongly associated with advanced gastric pathology, including chronic atrophic gastritis and gastric cancer. These results highlight the potential value of EPIYA motif profiling as an adjunct risk-stratification tool in populations with a high burden of H. pylori infection. Integrating molecular virulence characterization into clinical algorithms may improve early identification of individuals at elevated risk of gastric cancer, particularly in resource-limited settings.

7 Limitations

This study has several limitations. First, the overall sample size—particularly within the gastric cancer subgroup—was modest, which may have reduced statistical power and affected the precision of effect estimates. Second, the cross-sectional design precludes assessment of temporal disease progression and does not allow causal inference. Third, the molecular analysis was restricted to cagA and its EPIYA motif polymorphisms; other key H. pylori virulence determinants, such as vacA, babA, or oipA, were not evaluated and may modify clinical outcomes. Fourth, host-related and environmental contributors—including genetic susceptibility, dietary patterns, smoking, medication exposure, and socioeconomic conditions—were not incorporated into the analysis. Finally, classification of strains as Western-type was based solely on

EPIYA motif architecture rather than on whole-genome sequencing or phylogenetic characterization, which may underestimate the true diversity of circulating lineages.

8 Highlights

- H. pylori infection was significantly more common in patients with gastric pathology than in healthy controls in this Uzbek cohort. - cagA-positive strains predominated among infected patients and were associated with more severe gastric lesions. - Western-type CagA EPIYA motifs (AB, ABC, ABCC, ABCCC) represented the dominant structural patterns. - Multiple EPIYA-C repeats were strongly associated with chronic atrophic gastritis and gastric cancer. - CagA EPIYA motif profiling shows promise as a molecular marker for detecting high-risk H. pylori strains.

9 Future work

Future research should expand the molecular epidemiological framework for H. pylori in Central Asia by incorporating larger, multicenter cohorts and analyzing a broader panel of virulence genes, including vacA, babA, oipA, and iceA. Integrative approaches combining bacterial genomics, host genetic markers, epigenetics, microbiome composition, and environmental exposures are needed to more fully elucidate the mechanisms driving differential disease progression. Additionally, the application of emerging technologies—such as omics-based diagnostics, machine learning, and AI-assisted endoscopic evaluation—may enhance early detection of high-risk infections and refine risk-stratification algorithms. Longitudinal studies are particularly important to clarify causal pathways linking specific CagA EPIYA variants to the Correa cascade and gastric carcinogenesis.

10 Practical implications

The findings of this study have several practical implications. Molecular characterization of H. pylori virulence factors—particularly cagA status and EPIYA motif structure—may help identify patients infected with highly pathogenic strains. Detection of multiple EPIYA-C repeats could support risk stratification in individuals with chronic gastritis or premalignant gastric lesions, especially in high-prevalence regions such as Central Asia. Incorporating virulence profiling into clinical and epidemiological assessments may aid in designing targeted surveillance strategies and guiding individualized management. Although EPIYA analysis is not yet sufficient as a stand-alone diagnostic tool, it represents a promising adjunct marker that, with further validation, could contribute to precision-oriented approaches to gastric cancer prevention.

11 Author Contributions

Jadida Ismailova and Abrorjon Yusupbekov were the initiators of the proposed concept and made significant contributions to the final manuscript. Abrorjon Abdurakhimov and Jadida Ismailova were responsible for the development of the theoretical framework and provided support in data collection. Shakhnoza Rakhimova conducted a

thorough review of the final manuscript, played a role in the conceptual development, and made valuable additions. All authors made substantial contributions to the article and have given their approval for the submitted version.

12 Acknowledgements

We would like to express our gratitude to all individuals and institutions who contributed to this work. We also acknowledge the valuable feedback and support from colleagues and reviewers. For details regarding the inclusion of individuals in the author list, please refer to the Author Disclosure Form.

13 Funding

This research received no external funding.

14 Ethics Approval and Consent to Participate

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Republican Specialized Scientific and Practical Medical Center for Therapy and Medical Rehabilitation (Tashkent, Uzbekistan). Written informed consent was obtained from all participants prior to inclusion in the study.

15 Consent for Publication

All authors approved the final version of the manuscript and consented to its publication.

16 Competing Interests / Conflict of Interest

The authors declare no competing interests.

17 Data Availability

Data availability statement: No sequencing data were generated in this study. The study employed PCR-based genotyping using primers derived from previously published literature. Therefore, no datasets have been deposited in a public repository. All relevant data supporting the findings of this study are presented within the manuscript.

References

[1] Chen Y, Malfertheiner P, Yu H, Kuo C, Chang Y, Meng F, et al. Global prevalence of Helicobacter pylori infection and incidence of gastric cancer between 1980 and 2022. Gastroenterology. 2024;166(4):605–619. doi:10.1053/j.gastro.2023.12.022.

[2] Zamani M, Ebrahimtabar F, Zamani V, Miller WH, Alizadeh-Navaei R, Shokri-Shirvani J, Derakhshan MH. Systematic review with meta-analysis: the worldwide prevalence of Helicobacter pylori infection. Aliment Pharmacol Ther. 2018;47(7):868–876. doi:10.1111/apt.14561.

[3] Hooi JK, Lai WY, Ng WK, Suen MM, Underwood FE, Tanyingoh D, et al. Global prevalence of Helicobacter pylori infection: systematic review and meta-analysis. Gastroenterology. 2017;153(2):420–429. doi:10.1053/j.gastro.2017.04.022.

[4] Murray CJL. The Global Burden of Disease Study at 30 years. Nat Med. 2022;28(10):2019–2026. doi:10.1038/s41591-022-01990-1.

[5] Oh HJ, Kim CH, Park H, Park B. Quantifying the effects of the Correa pathway from Helicobacter pylori infection to gastric cancer: causal inference found in 6.8 million Koreans. BMC Cancer. 2025;26(1):174. doi:10.1186/s12885-025-15507-9.

[6] Song H, Ekheden IG, Zheng Z, Ericsson J, Nyrén O, Ye W. Incidence of gastric cancer among patients with gastric precancerous lesions: observational cohort study in a low risk Western population. BMJ. 2015;351:h3867. doi:10.1136/bmj.h3867.

[7] Hoang T, Lee J, Kim J. Comorbidity risk score in association with cancer incidence: results from a cancer screenee cohort. Cancers (Basel). 2020;12(7):1834. doi:10.3390/cancers12071834.

[8] Morgan E, Arnold M, Camargo MC, Gini A, Kunzmann AT, Matsuda T, et al. The current and future incidence and mortality of gastric cancer in 185 countries, 2020–40: a population-based modelling study. EClinicalMedicine. 2022;47:101404. doi:10.1016/j.eclinm.2022.101404.

[9] Savitri CMA, Matsumoto T, Fauzia KA, Alfaray RI, Waskito LA, Rezkitha YAA, et al. Helicobacter pylori pathogenic factors and their interactions with the gastric microbiome. Helicobacter. 2025;30(5):e70072. doi:10.1111/hel.70072.

[10] Rakesh, Munjal SD, Sengar R, Sharma N, Rautela S, Singh M. Molecular mechanisms of Helicobacter pylori infection: a review. J Pure Appl Microbiol. 2025;19(3):1755–1770. doi:10.22207/jpam.19.3.61.

[11] Baj J, Korona-Głowniak I, Forma A, Maani A, Sitarz E, Rahnama-Hezavah M, et al. Mechanisms of the epithelial–mesenchymal transition and tumor microenvironment in Helicobacter pylori-induced gastric cancer. Cells. 2020;9(4):1055. doi:10.3390/cells9041055.

[12] Bridge DR, Merrell DS. Polymorphism in the Helicobacter pylori CagA and VacA toxins and disease. Gut Microbes. 2013;4(2):101–117. doi:10.4161/gmic.23797.

[13] Díaz P, Valderrama MV, Bravo J, Quest AFG. Helicobacter pylori and gastric cancer: adaptive cellular mechanisms involved in disease progression. Front Microbiol. 2018;9:5. doi:10.3389/fmicb.2018.00005.

[14] Oghalaie A, Saberi S, Esmaeili M, Ebrahimzadeh F, Barkhordari F, Ghamarian A, et al. Helicobacter pylori peptidyl prolyl isomerase expression is associated with the severity of gastritis. J Gastrointest Cancer. 2016;47(4):375–380. doi:10.1007/s12029-016-9849-x.

[15] He B, Hu Y, Wu Y, Wang C, Gao L, Gong C, et al. Helicobacter pylori CagA elevates FTO to induce gastric cancer progression via a "hit-and-run" paradigm. Cancer Commun (Lond). 2025;45(5):608–631. doi:10.1002/cac2.70004.

[16] Sharafutdinov I, Tegtmeyer N, Linz B, Rohde M, Vieth M, Tay AC, et al. A single-nucleotide polymorphism in Helicobacter pylori promotes gastric cancer development. Cell Host Microbe. 2023;31(8):1345–1358.e6. doi:10.1016/j.chom.2023.06.016.

[17] Palrasu M, Zaika E, El-Rifai W, Garcia-Buitrago M, Piazuelo MB, Wilson KT, et al. Bacterial CagA protein compromises tumor suppressor mechanisms in gastric epithelial cells. J Clin Invest. 2020;130(5):2422–2434.

[18] Sharndama HC, Mba IE. Helicobacter pylori: an up-to-date overview on the virulence and pathogenesis mechanisms. Braz J Microbiol. 2022;53(1):33–50. doi:10.1007/s42770-021-00675-0.

[19] Tegtmeyer N, Neddermann M, Asche CI, Backert S. Subversion of host kinases: a key network in cellular signaling hijacked by Helicobacter pylori CagA. Mol Microbiol. 2017;105(3):358–372. doi:10.1111/mmi.13707.

[20] Altanbayar O, Amgalanbaatar A, Battogtokh C, Bayarjargal N, Belick D, Vasconcelos MK, et al. Characterization of the cagA-gene in Helicobacter pylori in Mongolia and detection of two EPIYA-A enriched CagA types. Int J Med Microbiol. 2022;312(3):151552. doi:10.1016/j.ijmm.2022.151552.

[21] Ansari S, Akada J, Matsuo Y, Shiota S, Kudo Y, Okimoto T, et al. Epitope peptides of Helicobacter pylori CagA antibodies from sera by whole-peptide mapping. J Gastroenterol. 2019;54(12):1039–1051. doi:10.1007/s00535-019-01584-8.

[22] Salvatori S, Marafini I, Laudisi F, Monteleone G, Stolfi C. Helicobacter pylori and gastric cancer: pathogenetic mechanisms. Int J Mol Sci. 2023;24(3):2895. doi:10.3390/ijms24032895.

[23] Li Q, Liu J, Gong Y, Yuan Y. Association of CagA EPIYA-D or EPIYA-C phosphorylation sites with peptic ulcer and gastric cancer risks. Medicine (Baltimore). 2017;96(17):e6620. doi:10.1097/MD.0000000000006620.

[24] Abdi E, Latifi-Navid S, Zahri S, Yazdanbod A, Safaralizadeh R. Helicobacter pylori genotypes determine risk of non-cardia gastric cancer and intestinal- or diffuse-type GC in Ardabil: a very high-risk area in Northwestern Iran. Microb Pathog. 2017;107:287–292. doi:10.1016/j.micpath.2017.04.007.

[25] Maciel DN, De Oliveira Santos-Dutra HC, Rocha VL, Rasmussen LT, Barbosa MS. Impact of Helicobacter pylori virulence genotypes cagA, vacA, oipA, and babA2 on severity of gastropathies in Brazilian patients. Int J Mol Sci. 2025;26(19):9471. doi:10.3390/ijms26199471.

[26] Jouimyi MR, Bounder G, Boura H, Essaidi I, Bendahmane A, Benomar H, et al. The EPIYA-ABCC motif of Helicobacter pylori cagA gene and gastric carcinogenesis in Casablanca population. Afr Health Sci. 2022;22(1):573–580. doi:10.4314/ahs.v22i1.66.

[27] Retnakumar RJ, Chettri P, Lamtha SC, Sivakumar KC, Dutta P, Sen P, et al. Genome-wide accumulations of non-random adaptive point mutations drive westward evolution of Helicobacter pylori. BMC Microbiol. 2025;25(1):229. doi:10.1186/s12866-025-03944-2.

[28] Hanafiah A, Lopes BS. Genetic diversity and virulence characteristics of Helicobacter pylori isolates in different human ethnic groups. Infect Genet Evol. 2019;78:104135. doi:10.1016/j.meegid.2019.104135.

[29] Yu X, Li L, He Q. Rapid diagnosis of Helicobacter pylori infection status based on endoscopic features and deep learning algorithms. Front Public Health. 2026;14:1765267. doi:10.3389/fpubh.2026.1765267.

[30] Alwahaibi N. Transforming gastric biopsy diagnostics: integrating omics technologies and artificial intelligence. Biomedicines. 2026;14(2):407. doi:10.3390/biomedicines14020407.

[31] Zhang Y, Shang Y, Li Z, Kuang Z, Huang S, Ye Q, et al. Integrative profiling of Helicobacter pylori clinical isolates: virulence genes, antimicrobial susceptibility and genetic diversity in gastric cancer risk stratification. Microb Genom. 2026;12(3). doi:10.1099/mgen.0.001599.

[32] Barhoine M, Moustaoui F, Hammani O, Aghrouch M, Lemkhente Z, Belhabib Z, et al. The effect of Helicobacter pylori gene combinations of cagA, cagE, virB11, vacA, and babA on the outcome of gastric disease in a Southern Moroccan population. Pathogens. 2025;14(3):279. doi:10.3390/pathogens14030279.

[33] Othman KI, Balaky STJ. Investigating the potential association of Helicobacter pylori cagA, vacA s1/s2, iceA1, iceA2, babA2, sabA, and oipA genotypes with gastric disease severity. Cell Mol Biol (Noisy-le-grand). 2025;71(1):88–95. doi:10.14715/cmb/2025.70.1.9.

[34] Kazemi E, Kahrizi D. Lack of association between gastric cancer and hopQ alleles in Helicobacter pylori. Genetika. 2016;48(3):893–902. doi:10.2298/GENSR1603893K.

[35] Lee D, Ha J, Shin J, Kim K, Choi J, Park S, et al. Increased risk of severe gastric symptoms by virulence factors vacAs1c, alpA, babA2, and hopZ in Helicobacter pylori infection. J Microbiol Biotechnol. 2021;31(3):368–379. doi:10.4014/jmb.2101.01023.

[36] Tran SC, Bryant KN, Cover TL. The Helicobacter pylori cag pathogenicity island as a determinant of gastric cancer risk. Gut Microbes. 2024;16(1):2314201. doi:10.1080/19490976.2024.2314201.

[37] Ofori EG, Adinortey CA, Bockarie AS, Kyei F, Tagoe EA, Adinortey MB. Helicobacter pylori infection, virulence genes' distribution and accompanying clinical outcomes: the West Africa situation. Biomed Res Int. 2019;2019:7312908. doi:10.1155/2019/7312908.

[38] Gawron AJ, Shah SC, Altayar O, Davitkov P, Morgan D, Turner K, Mustafa RA. AGA technical review on gastric intestinal metaplasia—natural history and clinical outcomes. Gastroenterology. 2020;158(3):705–731.e5. doi:10.1053/j.gastro.2019.12.001.

[39] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7–30. doi:10.3322/caac.21387.

[40] Polk DB, Peek RM. Helicobacter pylori: gastric cancer and beyond. Nat Rev Cancer. 2010;10(6):403–414. doi:10.1038/nrc2857.

[41] Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–209. doi:10.1038/nature13480.

[42] Ismailova J, Yusupbekov A, Abdurakhimov A, Nazarov K. Clarithromycin resistance and treatment outcomes in Helicobacter pylori infections: molecular genetic study from Uzbekistan. Gastroenterol Res Pract. 2026;2026(1). doi:10.1155/grp/9614183.