

DISTRIBUTION AND EVOLUTION OF *CAGA*, *CAGE*, AND *VACA* GENES AMONG *HELICOBACTER PYLORI* ISOLATES IN BULGARIA

Liliya Boyanova[✉], Lyudmila Boyanova^{id}, Daniel Yordanov^{id},
Raina Gergova^{id}, Rumyana Markovska^{id}

Received on March 20, 2026

Presented by B. Petrunov, Member of BAS, on May 26, 2026

Abstract

Helicobacter pylori is a species with a highly plastic genome encoding numerous virulence factors, such as the pathogenicity island (*cagPAI*) and vacuolating cytotoxin A (*VacA*). *cagPAI* consists of more than 26 genes, with *cagA* and especially *cagE* being essential predictors of *cagPAI* intactness. *vacA* gene has numerous allelic variations, among which *slm1l1* is the most virulent. The aim of our study was to assess *cagA/cagE* gene frequency and *vacA* allele status among clinical *H. pylori* isolates from Bulgaria for the period 2017–2024 and to compare the data with that from our previous study (2004–2010). In the current study, the rates of *cagA*, *cagE*, and the combination *cagA(+)/cagE(+)* were significantly lower compared to the previous study, as were the rates of the combinations *vacA slm1l1* and *cagA+/cagE+/slm1l1*. In conclusion, *H. pylori* evolution showed a statistically significant decrease in the prevalence of *cagA+/cagE+/vacA slm1l1* virulence genotype over 14 years, observed in the groups of total patients, as well as in the subgroups of adults and non-ulcer patients.

Key words: *Helicobacter pylori*, *cagPAI*, *cagA*, *cagE*, *vacA*, evolution, virulence

Introduction. Among its numerous virulence determinants, vacuolating cytotoxin A gene (*vacA* gene) and *cag* pathogenicity island (*cagPAI*) are most intensively studied.

CagA, encoded by *cagPAI*, is an oncoprotein that interferes with numerous signalling molecules in the host cell [1].

CagE, part of the inner membrane complex of *cagPAI*, is an ATP-ase that provides the energy necessary for the translocation of CagA [2, 3]. This is an important factor in virulence, as it is also associated with malignancy risk [4, 5]. Lack of *cagE* alters the integrity of T4SS; therefore, the presence of *cagE* is a better predictor of *cagPAI* intactness than *cagA* [4, 6].

VacA is a protoxin with a structure similar to A/B toxin of other bacteria. Its functions include pore formation, formation of intracellular vacuoles, induction of apoptosis, and interference with the host immune response [7]. There are three main regions: 's' signal region, 'i' intermediate region, and 'm' middle region. The *vacA* gene shows genetic diversity with different allelic variations: s1a-d, s2, m1a-c, m2a-b, m3, and i1-3 [8]. Genotype s1m1i1 is considered to be the most virulent due to the secretion of large levels of toxin, which causes a strong vacuolisation effect, and its association with peptic ulcer and adenocarcinoma. In contrast, s2 lacks cytotoxic activity, and m2 does not cause vacuolisation in HeLa cells [8]. In the combination s1m2, the presence of i1 determines the vacuolating activity [7].

The aim of the study was to examine the frequency of *cagA* and *cagE* genes and the rate of the main allelic variations of the *vacA* gene (s1a, s1b, s2, m1, m2, i1 and i2), as well as their combinations in *H. pylori* clinical isolates obtained in the period 2017–2024. We analysed the frequency of the genes and gene combinations in different population groups. Moreover, we compared the current results with data from Bulgaria for the period 2004–2010 published by BOYANOVA et al. [9], in order to evaluate the evolution of *H. pylori* virulence.

Materials and methods. Clinical isolates. We collected 127 single-genotype isolates during our routine diagnostics in the period 2017–2024. Isolation and identification were done as previously described [9]. Adults consisted of 113 patients with an age range (AR) of 22–93 years and a mean age (MA) of 48.7 years. The number of children was 14 (AR: 5–17, MA: 11.9). The number of males and females was 53 (AR: 10–93, MA: 45) and 74 (AR: 5–77, MA: 46.4), respectively. The frequency of clinical diagnoses was as follows: chronic gastritis – 71.7% (91/127), gastroesophageal reflux disease (GERD) – 8.7% (11/127), ulcerative disease – 7.9% (10/127), both gastritis and GERD – 2.4% (3/127), and 12 patients with other conditions, such as acute gastritis, autoimmune gastritis, or without clinical diagnosis. Informed consent was obtained from all participants included in the study, or from the parents of the children.

We used the data on *cagA*, *cagE* genes from our previous publication [9] and added data on *vacA* alleles.

DNA extraction. Bacterial DNA was obtained through thermoextraction for 10 min at 100 °C in a thermoblock and subsequently stored at –70 °C.

Molecular-genetic tests. The *cagA*, *cagE* genes, and the allelic variants of *vacA* were detected using conventional PCR with the Taq PCR kit (New England

Biolabs, Massachusetts, USA). We used the same primers and reaction conditions as previously described [9,10]. The results were read by gel-electrophoresis with ethidium bromide.

Statistical analysis. Differences in the prevalence of *H. pylori* genotypes between groups were assessed using the chi-squared test and Fisher's exact test. A p -value < 0.05 was considered significant.

Results. In 70.1% of the isolates, *cagA* was positive, while *cagE* was positive in 42.5% of the isolates. The rate of *cagA*(+)/*cagE*(+) was the same, since all *cagE*(+) were also *cagA*(+). The frequency of *cagA*(+)/*cagE*(-) was 27.6%, while that of *cagA*(-)/*cagE*(-) was 29.9%. None of the isolates were *cagA*(-)/*cagE*(+), (Table 1).

Regarding the virulent alleles of the *vacA* gene, 79.5% were positive for s1, 26.0% for m1, and 48.8% for i1. The highly virulent combination s1m1i1 was detected in 26.0% of the tested isolates. The combination *cagA*+/*cagE*+/s1m1i1 was found in 15.7%.

Compared with our previous results [9], a sharp decline in virulence determinants was observed. In recent years, the rate of *cagA* has been significantly lower in both the total number of patients and in the subgroups of non-ulcer patients, males, and adults. Compared to the previous study, the frequency of both *cagE* and the combination *cagA*(+)/*cagE*(+) was significantly lower in the total number of patients and in the subgroups of non-ulcer patients, males, females, and adults. In contrast, in the present study, the rate of *cagA*(+)/*cagE*(-) was significantly higher in the overall number of patients, males, and adults. Similarly, the *cagA*(-)/*cagE*(-) rate was higher in the total number of patients, non-ulcer patients, males, and adults (Table 1).

For the period 2017–2024, the rates of *vacA* m1 and i1 alleles also decreased significantly (Table 2). Consequently, the allelic combinations s1m1, s1m1i1, and s1m2i1 were much lower in the overall population than in 2004–2010. The combination s1m1i1 declined in adults and males. Additionally, a decrease in the *cagA*+/*cagE*+/s1m1i1 combination was observed in the total number of patients, non-ulcer patients, males, and adults (Table 1).

Discussion. The rate of *cagA*(+) strains in our study was 70.1%. It was higher than that reported by Polish authors (60.0%), but lower than in Asian countries like India (80%), Korea, and Japan, where the reported *cagA* rate as high as 90.0% [1,11].

The frequency of *cagE* was 42.5%, which is lower than that reported in Morocco (51.4%), Iran (61.0%), England (62.0%), Turkey (59.3%), Malaysia (59.0%), Brazil (89.3%), and India (85.4%) [3,12].

In the current study, the most common allelic combination of *vacA* gene was s1m2i2 (29.1%), while in the previous period the dominant combination was s1m1i1 (37.9%). The frequency of the s1m1i1 genotype in the 2017–2024 period was 26.0%. The proportion of the *cagA*+/*cagE*+/s1m1i1 combination in the

Table 1

Evolution of *H. pylori* associated with *cagA* and *cagE* and s,m,i genotypes in strains from 2017–2024 compared to strains from 2004–2010

Patient groups Genotypes	Study years	All patients (<i>n</i> = 219*/127)		Ulcer patients (<i>n</i> = 64*/10)		Non-ulcer patients (<i>n</i> = 155*/117)		Men (<i>n</i> = 135*/53)		Women (<i>n</i> = 84*/74)		Children (<i>n</i> = 12*/14)		Adults (<i>n</i> = 207*/113)	
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
<i>cagA</i> ⁺	2004–2010*	186	84.9	60	93.8	126	81.3	114	84.4	72	85.7	10	83.3	176	85.0
	2017–2024	89	70.1	9	90.0	80	68.4	34	64.2	55	74.3	11	78.6	78	69.0
	<i>p</i> -value	0.001		0.527		0.016		0.005		0.107		1.000		0.001	
<i>cagE</i> ⁺	2004–2010*	150	68.5	54	84.4	96	61.9	99	73.3	51	60.7	5	41.7	145	70.0
	2017–2024	54	42.5	6	60.0	48	41.0	22	41.5	32	43.2	5	35.7	49	43.4
	<i>p</i> -value	< 0.0001		0.087		0.0009		< 0.0001		0.038		1.000		< 0.001	
<i>cagA</i> ⁺ / <i>cagE</i> ⁺	2004–2010*	149	68.0	54	84.4	95	61.3	98	72.6	51	60.7	5	41.7	145	70.0
	2017–2024	54	42.5	6	60.0	48	41.0	22	41.5	32	43.2	5	35.7	49	43.4
	<i>p</i> -value	< 0.0001		0.087		0.001		< 0.0001		0.038		1.000		< 0.001	
<i>cagA</i> ⁺ / <i>cagE</i> [−]	2004–2010*	37	16.9	6	9.4	31	20.0	16	11.9	21	25.0	5	41.7	32	15.5
	2017–2024	35	27.6	3	30.0	32	27.4	13	24.5	22	29.7	6	42.9	29	25.7
	<i>p</i> -value	0.020		0.097		0.191		0.042		0.592		1.000		0.037	
<i>cagA</i> [−] / <i>cagE</i> ⁺	2004–2010*	1	0.5	0	0.0	1	0.6	1	0.7	0	0.0	0	0.0	1	0.5
	2017–2024	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	<i>p</i> -value	1.000		1.000		1.000		1.000		1.000		1.000		1.000	
<i>cagA</i> [−] / <i>cagE</i> [−]	2004–2010*	32	14.6	4	6.2	28	18.1	20	14.8	12	14.3	2	16.7	30	14.5
	2017–2024	38	29.9	1	10.0	37	31.6	18	34.0	20	27.0	3	21.4	35	31.0
	<i>p</i> -value	0.0008		0.527		0.014		0.005		0.051		1.000		0.0007	
slmli1	2004–2010	83	37.9	28	43.8	55	35.5	57	42.2	26	35.6	5	41.7	78	37.7
	2017–2024	33	26.0	3	30.0	30	25.6	12	22.6	21	28.4	3	21.4	30	26.5
	<i>p</i> -value	0.025		0.505		0.088		0.012		0.731		0.401		0.048	
<i>cagA</i> ⁺ / <i>cagE</i> ⁺ / slmli1	2004–2010	71	32.4	27	42.2	44	28.4	51	37.8	20	23.8	3	25	68	32.9
	2017–2024	20	15.7	1	10.0	19	16.2	7	13.2	13	17.6	1	7.1	19	16.8
	<i>p</i> -value	0.0006		0.079		0.021		0.0008		0.433		0.306		0.002	

*strains from 2004–2010

a *p*-value < 0.05 is considered statistically significant

T a b l e 2

Evolution of *H. pylori* associated with *vacA* s,m,i alleles in strains from 2017–2024 compared to strains from 2004–2010

Genotype	2004–2010* <i>n</i> , %	2017–2024** <i>n</i> , %	<i>p</i> -value***
s1a	185 (84.5%)	99 (78.0%)	<i>p</i> = 0.146
s1b	2 (0.9%)	2 (1.6%)	<i>p</i> = 0.626
s1 (a+b)	187 (85.4%)	101 (79.5%)	<i>p</i> = 0.180
s2	30 (13.7%)	25 (19.7%)	<i>p</i> = 0.170
m1	90 (41.1%)	33 (26.0%)	<i>p</i> = 0.005
m2	127 (58.0%)	80 (63.0%)	<i>p</i> = 0.366
i1	138 (63.0%)	62 (48.8%)	<i>p</i> = 0.013
i2	79 (36.1%)	60 (47.2%)	<i>p</i> = 0.053
s1m1	87 (39.7%)	34 (26.8%)	<i>p</i> = 0.019
s1m2	98 (44.7%)	55 (43.3%)	<i>p</i> = 0.823
s2m2	28 (12.8%)	25 (19.7%)	<i>p</i> = 0.091
s1m1i1	83 (37.9%)	33 (26.0%)	<i>p</i> = 0.025
s1m2i1	48 (21.9%)	16 (12.6%)	<i>p</i> = 0.032
s1m2i2	51 (23.3%)	37 (29.1%)	<i>p</i> = 0.250
s2m2i2	24 (11.0%)	22 (17.3%)	<i>p</i> = 0.102
<i>cagA/cagE/s1m1i1</i>	71 (32.4%)	20 (15.7%)	<i>p</i> = 0.0009

* strains from 2004–2010, *n* = 219

** strains from 2017–2024, *n* = 127

*** a *p*-value < 0.05 is considered statistically significant

recent period was twofold lower (15.7%) than in the previous period (32.4%). The frequency of the leading allelic combination varies across different geographic regions. In a study among Jordan patients, the most common combination was s2m2 (21.5%), followed by m2s1 (15.2%), while m1s1 was the least frequent (5.1%) [13]. In a Moroccan study the most prevalent combination was s2m2i2 (39.3%), followed by s1m1i1 (20.9%) [4]. Peruvian researchers reported a rate of 79.1% for *cagA*+/s1m1 [14].

H. pylori virulence appear to decrease over the years. The more virulent Eastern-type strains tend to evolve towards the less virulent Western-type *H. pylori*, which is described as a global phenomenon [15]. Taiwanese authors reported a decrease in *H. pylori* rates, gastric cancer, and duodenal ulcer [16].

In our study, a decrease in the virulence of *H. pylori* is observed in terms of both *cagPAI* and *vacA* genes/gene combinations. The rates of *cagA*, *cagE*, and the combination *cagA*(+)/*cagE*(+) have significantly decreased compared with the period 2004–2010. Meanwhile, the less virulent combinations *cagA*(+)/*cagE*(-) and *cagA*(-)/*cagE*(-) have significantly increased. The findings for the *vacA* gene

were consistent with this trend. The most virulent *vacA* combination, s1m1l1, has significantly declined in total number of patients, as well as in males, and adults. Additionally, the combination of *cagA/cagE/s1m1l1* decreased in both the aforementioned groups and in non-ulcer patients (Table 1). Such a trend was not observed in ulcer patients, which is in concordance with literature data indicating that virulence determinants are linked to ulcer diseases [2,5,11]. The combination *cagA(+)/cagE(+)* was associated with a 1.97-fold higher risk for duodenal ulcer in Morocco [4]. No evolution of virulence was observed within the group of children. A possible explanation is that *H. pylori* infection is often asymptomatic during childhood, which may delay diagnosis.

There may be several reasons for the decrease in *H. pylori* virulence determinants. According to a meta-analysis [17], *cagA(+)* strains are more likely to be eradicated. This was confirmed by ZHANG et al. [18], who explained that a more severe inflammatory reaction leads to increased blood supply to gastric mucosa, which enhances the action of antibiotics. Another possible reason is the higher fitness of *cagPAI(-)* strains, which provides a selective advantage under certain conditions [5]. A ten-year follow-up study [19] showed that long-term colonization in a single individual often leads to mutations, reflecting the highly adaptive properties of the pathogen. In this study, 5 out of 90 patients' isolates changed from *cagA(+)* to *cagA(-)* over the 10-year period [19].

Conclusion. In the current study (2017–2024), no significant differences were observed in the *cagA/cagE/vacA(s,m,i)* genotypes among Bulgarian patients with respect to gender, age, or diagnosis. Compared with the period 2004–2010, *H. pylori* strains showed a significant decrease in virulence genes in the groups of the total number of patients, adults, and non-ulcer patients, but not in children and ulcer patients. The decrease in virulence may be due to the higher eradication success of more virulent *cagA(+)* strains compared with the remaining strains and/or the greater fitness of *cagPAI(-)* strains compared with those with functional *cagPAI*. Future studies are needed to follow up on and analyse the reported trend.

REFERENCES

- [1] KLESIEWICZ K., M. ORCZYKOWSKA-KOTYNA, I. SKIBA-KUREK et al. (2025) Prevalence and antimicrobial resistance of highly virulent *cagA*-positive *Helicobacter pylori* strains in Southern Poland, Eur. J. Clin. Microbiol. Infect. Dis., **44**(2), 405–416.
- [2] GOU L., X. YANG, J. YUN et al. (2024) Roles of the components of the *cag*-pathogenicity island encoded type IV secretion system in *Helicobacter pylori*, Future Microbiol., **19**(14), 1253–1267.

- [3] BARHOINE M., F. MOUSTAOU, O. HAMMANI et al. (2025) 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*, **14**(4), 279.
- [4] EL KHADIR M., S. A. BOUKHRIS, S. O. ZAHIR et al. (2021) *CagE*, *cagA* and *cagA* 3' region polymorphism of *Helicobacter pylori* and their association with the intra-gastric diseases in Moroccan population, *Diagn. Microbiol. Infect. Dis.*, **100**(3), 115372.
- [5] TRAN S. C., K. N. BRYANT, T. L. COVER (2024) The *Helicobacter pylori* *cag* pathogenicity island as a determinant of gastric cancer risk, *Gut Microbes*, **16**(1), 2314201.
- [6] MARKOVSKA R., L. BOYANOVA, D. YORDANOV et al. (2018) Status of *Helicobacter pylori* *cag* pathogenicity island (*cagPAI*) integrity and significance of its individual genes, *Infect. Genet. Evol.*, **59**, 167–171.
- [7] JUNAID M., A. K. LINN, M. B. JAVADI et al. (2016) Vacuolating cytotoxin A (*VacA*) – A multi-talented pore-forming toxin from *Helicobacter pylori*, *Toxicon*, **118**, 27–35.
- [8] JARZAB M., J. SKORKO-GLONEK (2024) There are no insurmountable barriers: Passage of the *Helicobacter pylori* *VacA* toxin from bacterial cytoplasm to eukaryotic cell organelle, *Membranes*, **14**(1), 11.
- [9] BOYANOVA L., D. YORDANOV, G. GERGOVA et al. (2011) Benefits of *Helicobacter pylori* *cagE* genotyping in addition to *cagA* genotyping: a Bulgarian study, *Antonie van Leeuwenhoek*, **100**(4), 529–535.
- [10] BOYANOVA L., L. BOYANOVA, P. HADZHIYSKI et al. (2023) Mixed (multiple-genotype) *Helicobacter pylori* infections in Bulgarian patients, *Diagn. Microbiol. Infect. Dis.*, **107**(4), 116073.
- [11] SHETTY V., R. LINGADAKAI, G. C. PAI et al. (2021) Profile of *Helicobacter pylori* *cagA* and *vacA* genotypes and its association with the spectrum of gastroduodenal disease, *Indian J. Med. Microbiol.*, **39**(4), 495–499.
- [12] GHOLIZADE TOBNAGH S., S. Z. BAKHTI, S. LATIFI NAVID et al. (2017) Role of plasticity region genes and *cagE* gene of *cagPAI* of *Helicobacter pylori* in development of gastrointestinal (GI) diseases, *Asian Pac. J. Cancer Prev.*, **18**(1), 43–49.
- [13] AL-HYASSAT M. M., H. I. AL-DAGHISTANI, L. F. ABU-NIAAJ et al. (2025) Allelic variation of *Helicobacter pylori* *vacA* gene and its association with gastric pathologies in clinical samples collected in Jordan, *Microorganisms*, **13**(8), 1841.
- [14] GUZMÁN J., D. CASTILLO, A. D. GONZÁLEZ-SICCHA et al. (2024) *Helicobacter pylori* *cagA*, *vacA*, *iceA* and *babA* genotypes from Peruvian patients with gastric intestinal metaplasia, *Cancers*, **16**(8), 1476.
- [15] RETNAKUMAR R. J., P. CHETTRI, S. C. LAMTHA et al. (2025) Genome-wide accumulations of non-random adaptive point mutations drive westward evolution of *Helicobacter pylori*, *BMC Microbiol.*, **25**(1), 229.
- [16] CHEN M. J., M. J. BAIR, P. Y. CHEN et al. (2022) Declining trends of prevalence of *Helicobacter pylori* infection and incidence of gastric cancer in Taiwan: An updated cross-sectional survey and meta-analysis, *Helicobacter*, **27**(5), e12914.
- [17] WANG D., Q. LI, Y. GONG et al. (2017) The association between *vacA* or *cagA* status and eradication outcome of *Helicobacter pylori* infection: A meta-analysis, *PLoS One*, **12**, e0177455.
- [18] ZHANG S. H., X. ZHU, B. M. LI et al. (2018) The effect of virulence genotypes

- of *Helicobacter pylori* on eradication therapy in children, Saudi J. Gastroenterol., **24**(4), 249–254.
- [19] GHADERSOLTANI P., S. SHORAKA, A. SADJADI et al. (2024) Long-term assessment of *Helicobacter pylori* *cagA* EPIYA motif changes and pathology outcomes in gastric biopsies of dyspeptic patients: 10-year follow-up, BMC Gastroenterol., **24**(1), 466.

*Department of Medical Microbiology, Medical University of Sofia,
2 Zdrave St, 1431 Sofia, Bulgaria*

e-mails: liliya.boianova@medfac.mu-sofia.bg,

<https://orcid.org/0000-0001-8926-2741>

l.boyanova@hotmail.com, <https://orcid.org/0000-0001-9622-0873>

dr.daniel.yordanov.md@gmail.com, <https://orcid.org/0000-0002-3983-7778>

rtgergova@gmail.com, <https://orcid.org/0000-0003-3605-6447>

markovska73@abv.bg, <https://orcid.org/0000-0003-1634-8768>