**Helicobacter pylori** CagA protein polymorphisms and their lack of association with pathogenesis

Nicole Acosta, Andrés Quiroga, Pilar Delgado, María Mercedes Bravo, Carlos Jaramillo

**AIM:** To investigate *Helicobacter pylori* (*H. pylori*) CagA diversity and to evaluate the association between protein polymorphisms and the occurrence of gastric pathologies.

**METHODS:** One hundred and twenty-two clinical isolates of *H. pylori* cultured from gastric biopsies obtained from Colombian patients with dyspepsia were included as study material. DNA extracted from isolates was used to determine *cagA* status, amplifying the C-terminal *cagA* gene region by polymerase chain reaction. One hundred and six strains with a single amplicon were sequenced and results were used to characterize the 3' variable region of the *cagA* gene. To establish the number and type of tyrosine phosphorylation motifs Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine (EPIYA) bioinformatic analysis using Amino Acid Sequence Analyzer-Amino Acid Sequence Analyzer software was conducted. Analysis of the association between the number of EPIYA motifs and the gastric pathology was performed using χ² test and analysis of the presence of EPIYA-C motifs in relation to the pathology was made by logistic regression odds ratios. Comparisons among EPIYA types found and those reported in GenBank were performed using a proportion test in Statistix Analytical Software version 8.0.

**RESULTS:** After amplification of the 3' of the *cagA* gene, 106 from 122 isolates presented a single amplicon and 16 showed multiple amplicons. As expected, diversity in the size of the *cagA* unique fragments among isolates was observed. The 106 strains that presented a single amplicon after 3' *cagA* amplification came from patients with gastritis (19 patients), atrophic gastritis (21), intestinal metaplasia (26), duodenal ulcer (22) and gastric cancer. DNA sequence analysis showed that the differences in size of 3' *cagA* unique fragments was attributable to the number of EPIYA motifs: 1.9% had two EPIYA motifs, 62.3% had three, 33.0% had four and 2.8% had five motifs. The majority of tested clinical strains (62.3%) were found to harbor the ABC combination of EPIYA motifs and a significant statistical difference was observed between the frequencies of ABCC tyrosine phosphorylation motifs and Western strains sequences deposited in GenBank.

**CONCLUSION:** The present report describes a lack of association between *H. pylori* CagA-protein polymorphisms and pathogenesis. ABCC high frequency variations compared with Western-strains sequences deposited in GenBank require more investigation.

© 2010 Baishideng. All rights reserved.

**Key words:** *Helicobacter pylori* CagA-protein polymorphisms; Molecular characterization; Bioinformatic analysis; Pathogenesis; Cancer

Acosta N, Quiroga A, Delgado P, Bravo MM, Jaramillo C. *Helicobacter pylori* CagA protein polymorphisms and their lack...
of association with pathogenesis. World J Gastroenterol 2010; 16(31): 3936-3943 Available from: URL: http://www.wjgnet.com/1007-9327/full/v16/i31/3936.htm DOI: http://dx.doi.org/10.3748/wjg.v16.i31.3936

INTRODUCTION

Helicobacter pylori (H. pylori) is a spiral gram-negative microaerophilic bacillus[1]. This bacteria is one of the most common worldwide human pathogens[2], it is present in at least 50% of the world’s population, with the highest incidence recorded in industrially underdeveloped areas, including Asia, Africa and South America[3]. H. pylori colonizes the human stomach and persists for several decades, causing chronic gastritis and peptic ulcer diseases[4]. Studies have suggested that chronic infection by H. pylori is an important risk factor for the development of gastric carcinoma[5,6]. For this reason, H. pylori was defined as a type I carcinogen by The International Agency for Research on Cancer - IARC[7].

The cytotoxin-associated antigen A, CagA, was identified in 1989[8]. It is encoded in the cag pathogenicity island (cag PAI), a segment of 40 kb that codifies components required to ensemble a type IV secretion system (TFSS)[9]. More than 90% of isolated strains from East Asia including Korea, Japan, and China are known to harbor cagA, while 50%-60% of isolated strains from Western countries are positive for it[10]. This gene shows variation which is explained by adaptive evolution, where a genetically diverse H. pylori population provides the host with a repertoire of varied phenotypes from which a subpopulation with optimal fitness may be selected[11]. This evolution would operate through recombination between H. pylori direct DNA repeats that result in deletion (or duplication) of phosphorylation sites in cagA gene[12,13] or the entire genomic cag PAI[14].

The cagA gene product, CagA, is directly translocated from H. pylori into the gastric epithelia cells the bacteria are attached to via TFSS[15,16] and upon localizing in the inner surface of the plasma membrane, CagA undergoes tyrosine phosphorylation by Ab1 and Src family kinases on specific tyrosine residues within a Glutamine acid-Proline-tyrosine-tyrosine-phosphorylation by Ab1 and Src family kinases on specific tyrosine residues within a Glutamine acid-Proline-tyrosine-Alanine (EPIYA) motif[17-20]. Once CagA is phosphorylated, it interacts with Src homology phosphatase 2 (SHP-2) which stimulates downstream signaling cascades involved in the reorganization of the cytoskeleton, resulting in cellular morphological changes such as the “hummingbird” phenotype[21]. Among the various CagA activities that disturb cellular functions, deregulation of SHP-2 by CagA is of potential importance in gastric carcinogenesis because mutations in PTPN11, the gene encoding human SHP-2, have been identified in human malignancies[22,23].

CagA protein varies in size according to the strain[24,25]. The structure of the gene reveals a 5’ highly conserved region and CagA size variation is due to the presence of different types and/or numbers of repeat sequences containing the EPIYA motifs within the C-terminal variable region[26]. Four types of EPIYA segments have been described: A, B, C, and D, each of which contains a single EPIYA motif[27]. Moreover, Panayotopoulou et al[28] and Kanada et al[29] described the pattern around the EPIYA motif to determine the type to which it corresponds as follows: EPIYA-A, EPIYAKNNKK_K/A/T/V/S/GQ; EPIYA-B, E(S/P)IY(A/T)_(Q/K)VAKKVNAKI; EPIYA-C, EPIYATIDDLG and EPIYA-D, EPIYATIDFEANQAG. Earlier studies have shown that CagA protein nearly always contains EPIYA-A and EPIYA-B sites, followed by one to three EPIYA-C repeats in Western-type H. pylori isolates[30] or by one EPIYA-D motif in East Asian-type isolates[31]. Src kinase in gastric epithelial cells phosphorylates CagA on EPIYA-C tyrosine residues[31-33]. Consequently, among Western CagA strains, the number of EPIYA-C sites is directly associated with the level of tyrosine phosphorylation. Thus, Western CagA proteins with a greater number of EPIYA-C sites are pathophysiologically more virulent and probably more carcinogenic[84]. The tyrosine phosphorylation status of CagA is important for the pathogenicity of H. pylori[20] and this variable number of EPIYA could be of clinical relevance in gastroduodenal diseases. For the above stated reasons it is important to analyze the genetic variability of cagA gene from clinical strains in relation to the associated pathologies. In this study, we used a polymerase chain reaction (PCR)-sequencing-bioinformatics strategy to characterize the CagA variable region of H. pylori from Colombian isolates. Additionally, the association between CagA diversity and the severity of gastroduodenal disease was analyzed.

MATERIALS AND METHODS

Clinical strains and culture conditions
A total of 122 H. pylori strains obtained from the stock collection at the Instituto Nacional de Cancerología, in Bogotá, Colombia, were grown on blood agar plates, supplemented with 7% horse serum (Invitrogen, Grand Island, NY), 1% Vitox (Oxoid, Basingstoke, UK), and Campylobacter selective supplement (Oxoid, Basingstoke, UK), at 37°C for 3 days in microaerophilic conditions. Three cagA positive control strains NCTC 11637, NCTC 11638 and ATCC 43579, were also included. Isolates belonged to patients with different types of gastric pathologies including benign, mild and severe conditions associated with H. pylori infection. Histopathology diagnosis was recorded for all voluntary participants.

DNA extraction and PCR assay
Genomic DNA was extracted using AquaPure Genomic DNA isolation kit, BIO-RAD, according to manufacturer’s instructions and obtained DNA was stored at -20°C until PCR amplification. In order to amplify the variable cagA region, a PCR assay was carried out in a volume of 50 µL containing 50 mmol/L KCl, 20 mmol/L Tris-HCl, pH (8.4), MgCl2 1.75 mmol/L, 0.2 mmol/L of each dNTP, 1 pmol/µL of each primer (CAGTF 5-3: ACCCTAGTC GAYACYGC), previously reported by Yamaoka et al[34], 1.25 units of Taq DNA Polymerase [Invitrogen, Carlsbad
Sequencing of 3' variable region of the cagA region

Sequencing was performed using an ABI PRISM® 310 Genetic Analyzer [Applied Biosystems, Foster City (California), USA], BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA). The reactions were done in a volume of 10 µL containing 0.5 x of Premix, 0.5 x of buffer, 0.16 µmol/L of each primer and 2 µL of purified DNA. The conditions were: one cycle at 96°C for 1 min, followed by 25 cycles of 10 s at 96°C, 5 s at 50°C and 4 min at 60°C. All of the 106 sequences obtained were deposited into GenBank/EMBL/DDBJ database with the following accession numbers: FJ755476 and FJ915841 to FJ915945.

Bioinformatics analysis

The sequences (forward and reverse) were edited and assembled using CLC DNA Workbench [CLC Bio A/S, (Aarhus C), Denmark]. For the characterization and quantification of the EPIYA motifs located in the C-terminal of CagA protein, a software called Amino Acid Sequence Analyzer (AASA) was designed for the study, in order to look for the type and number of EPIYA motifs, using the 6 open reading frames of each sequence. The characterization of tyrosine phosphorylation motifs, which contains EPIYA sequences, was done as previously described by Higashi et al[30]. Clustal W program[38] was used to generate a multiple alignment from the amino acid sequences of each strain.

In order to test the capability of the software to establish phosphorylation motifs in an accurate manner and the facility of its use, all controls were processed as a first phase. Sequence data results was as expected according to GenBank databases, so based on these results evaluation of clinical strain was done.

Statistical analysis

Analysis of association between the number of EPIYA motifs and the gastric pathology described was performed using χ² test. Analysis of the presence of one or more than one EPIYA-C motifs in relation to the pathology was made by logistic regression odds ratios (OR); a 95% confidence interval (CI) was calculated using SPSS statistical software package version 16.0 (SPSS Inc., Chicago, IL, USA). Comparisons among EPIYA types found and those reported in GenBank were performed using a Proportion test in Statistical Analysis Software, version 8.0 (Software, 1985-2003).

RESULTS

PCR assay

After amplification of the 3' of the cagA gene, 106 clinical isolates from 122 isolates presented a single amplicon (a unique band), and 16 showed multiple amplicons (two or more fragments) (Figure 1, lane 4). As expected, diversity in the size of the cagA fragment among isolates was observed, and PCR products ranged from 343 to 811 bp. The 106 strains that presented a single amplicon after 3'-cagA amplification came from patients with superficial gastritis (19 patients), atrophic gastritis (21), intestinal metaplasia (26), duodenal ulcer (22) and gastric cancer (18).

Bioinformatics analysis of peptide sequences

After sequencing of cagA 3’ region gene, corresponding peptide sequences from all of the strains with single amplicons were deduced. The combination of the different EPIYA motifs was determined using the ClustalW and AASA software, based on the classification defined by Higashi et al[30] and Panayotopoulou and collaborators[32]. The pattern EPIYA(K/Q)VNAKK(A/T)GQ that corresponds to EPIYA-A; the pattern E(P/S)IY(A/T)(Q/R)K)VAKKV(N/T)(A/Q)KI, to EPIYA-B; and the pattern EPIYATIDDL(G/R) to EPIYA-C were found (Figure 2). All the EPIYA motifs were Western type (Table 1), we did not found strains harboring EPIYA-D, a characteristic pattern of Eastern type.

DNA sequence analysis revealed that 2 of 106 strains (1.9%) had two EPIYA motifs, 66 strains (62.3%) three EPIYA motifs, 35 strains (33.0%) four EPIYA motifs and 3 strains (2.8%) five EPIYA motifs. In 66 out of the 106 (62.3%) strains, the pattern of the EPIYA motifs was ABC type (Table 1). Moreover, 49 of 106 strains had a modified EPIYA-B motif (EPIYT) instead of EPIYA (Figure 2) and there was no association between this type of tyrosine phosphorylation motif (EPIYT) containing a threonine residue instead of an alanine residue and the studied pathologies (P = 0.51) (data not shown).

Studies have shown that CagA proteins with more EPIYA-C motifs are expected to be more active biologically than those with a small number of EPIYA-C motifs because they interact more effectively with SHP-2 phosphatases and therefore it could perturb SHP-2-dependent signaling pathways, inducing greater morphological changes or probably contributing to generation of gastric cancer[30,38]. For this reason, in order to determine if there was any relation between the severity of the gastroduodenal...
Statistical analyses were applied. Statistical tests

A $\chi^2$ test performed showed that in this group of Colombian isolates, the variation of all EPIYA motifs in CagA was not directly associated to the outcome of the disease caused by $H. \text{pylori}$ ($P = 0.52$). Then, atrophic gastritis, intestinal metaplasia, duodenal ulcer and gastric cancer risks were estimated with respect to a reference group made up of patients with gastritis, infected by $cagA$ positive strains with one EPIYA-C motif and those with more than one EPIYA-C motif. No association between the number of EPIYA-C motifs and the pathology was found using logistical regression (Table 2). Proportion test frequencies of the EPIYA genotypes obtained in this study were compared with those of $CagA$ sequences including the C-terminal variable region reported by Argent et al. and sequences collected from Colombia deposited in GenBank. Statistical differences between the frequencies of ABCC pattern in Western strains previously reported by Argent et al. and our results were established, $P < 0.002$ (Table 1 and Figure 3).

**DISCUSSION**

The present work considered amplification and sequencing of the 3' variable $cagA$-gene region and a final bioinformatics analysis of the corresponding C-terminal of CagA protein using AASA software as a single and rapid method for $H. \text{pylori}$ CagA characterization. Determination of number and type of $H. \text{pylori}$ CagA phosphorylation motifs has been suggested by some researchers as a way to predict clinical outcome of $H. \text{pylori}$ associated pathologies and as a prognosis tool by others. So, in order to examine the viability of mentioned approach, 122 Co-

**Table 1** Prevalence of different types of CagA tyrosine phosphorylation motifs compared with data obtained from Argent et al. and Colombian strains reports in GenBank (%)

| Type(s) of EPIYA motif | BC | AABC | ABC | ABCC | ABCCC | AABCC | ABABC | AC | D | Total |
|------------------------|----|------|-----|------|-------|-------|-------|----|---|-------|
| Gastritis              | 1  | 0    | 10  | 6    | 0     | 1     | 0     | 1  | 0 | 19    |
| Atrophic gastritis     | 0  | 0    | 15  | 5    | 0     | 0     | 1     | 0  | 0 | 21    |
| Intestinal metaplasia  | 0  | 0    | 16  | 9    | 0     | 0     | 0     | 0  | 0 | 26    |
| Duodenal ulcer         | 0  | 0    | 14  | 8    | 0     | 0     | 0     | 0  | 0 | 22    |
| Gastric cancer         | 1  | 1    | 11  | 5    | 1     | 0     | 0     | 0  | 0 | 18    |
| This study total       | 1 (0.9) | 2 (1.9) | 66 (62.3) | 33 (31.1) | 1 (0.9) | 1 (0.9) | 1 (0.9) | 1 (0.9) | 0 (0.0) | 106 |
| Colombian$^1$          | 2 (2.2) | 2 (2.2) | 49 (53.3) | 29 (31.5) | 1 (1.1) | 2 (2.2) | 1 (1.1) | 1 (1.1) | 0 (0.0) | 87 |
| Argent et al.$^{2,3}$  | 3 (0.7) | 3 (0.7) | 262 (63.3) | 81 (19.6) | 19 (4.6) | 1 (0.2) | 1 (0.2) | 4 (1.0) | 1 (0.1) | 374 |

$^1$CagA sequences of Colombian strains recovered on GenBank; $^2$Data for Western CagA EPIYA types (-A, -B, -C); $^3$The EPIYA –D was not included in the analysis of this table because it was absent in the Colombian strains. EPIYA: Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine.

**Figure 2** Multiple alignment of the amino acid sequence of the 3' region from $cagA$. First Square (A) shows the Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine (EPIYA) motifs which had a modified EPIYA-B motif (EPIYT) instead of EPIYA. Second Square (B) shows the CagA multimerization (CM) motif which contains the amino acid sequence FPLXRXXXVXDLSKVG. The shaded squares represent the localization of EPIYA motif in the alignment. Amino acids are designated by standard one-letter codes. (∙) same residue; (-) gap. GenBank accession numbers and EPIYA motif of each isolate in this figure: CagA 1-1093 (FJ915841) (ABC), CagA 1-1095 (FJ915842) (ABC), CagA 1-2007 (FJ915843) (ABC), CagA 04-A033 (FJ915904) (ABC), CagA 1-3125 (FJ915893) (ABC), CagA 5-22023 (FJ915921) (ABC), CagA 5-22366 (FJ915913) (ABC), CagA 5-22025 (FJ915878) (BC), CagA 1-1099 (FJ915912) (ABC) and CagA01-3004 (FJ915971) (ABCC).
Table 2  Association between cagA Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine-C motif number and gastric pathology (a)

| Pathology               | No more than 1 EPIYA-C motif | At least 2 or more EPIYA-C motif | OR  | 95% CI |
|-------------------------|------------------------------|---------------------------------|-----|--------|
| Gastritis               | 12 (16.9)                    | 7 (20)                          | 1.00| -      |
| Atrophic gastritis      | 16 (22.5)                    | 5 (14.2)                        | 1.867| 0.47-7.34 |
| Intestinal metaplasia   | 17 (23.9)                    | 9 (25.7)                        | 1.102| 0.32-3.78 |
| Duodenal ulcer          | 14 (19.7)                    | 8 (22.8)                        | 1.021| 0.30-4.51 |
| Gastric cancer          | 12 (16.9)                    | 6 (17.1)                        | 1.167| 0.28-3.65 |
| Total                   | 71                           | 35                              |     |        |

EPIYA: Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine.

Figure 3  Comparison between the prevalence of different types of CagA tyrosine phosphorylation motifs from this study compared with data obtained from Argent et al[39] and Colombian strains reported in gene bank. There is a statistical difference between the frequency of ABC pattern in Western strains reported in GenBank (19.6%) and our results (31.1%). EPIYA: Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine.

lombian clinical isolates, all of them derived from adults, were analyzed. From 122 initial strains included, 106 isolates that presented a single PCR product were considered for latter analysis. As expected, diversity in the size of the cagA fragment among isolates was observed, and PCR products ranged from 343 to 811 bp. This is explained by the fact that cagA varies in size depending on the number of EPIYA motifs it has, and because each type of EPIYA adds a certain number of amino acids, since each type is surrounded by a specific amino acid sequence. EPIYA-A represents an additional 32 amino acids in the protein, EPIYA-B 40 more amino acids, and EPIYA-C 34 more amino acids for Western CagA type[40][41]. The remaining 16 strains that showed two or more PCR products were excluded. However, they were considered for further analysis due to: (1) they could represent subclones attributable to a clone microevolution process demonstrating the high recombination rate of H. pylori plastic genome; or (2) they could represent a coinfection process with multiple H. pylori strains coexisting in the same host. A microevolution process at cagA gene has been described previously in the same individual and also in different family members[32][42][43]. Otherwise strains with a polymorphism in the cagA variable region have been also observed by Panayotopoulou et al[22], Aras et al[12] and Reyes-Leon et al[43]. The EPIYA patterns found in this study match with those reported by Panayotopoulou et al[22], but there were differences for EPIYA-A in which the amino acid Lysine (K) in the 6th position (being the amino acid E of the EPIYA motif the 1st position) changed to K/Q as previously reported[19]. EPIYA-B in which the pattern in positions 12 and 13 changed from being only Asparagine and Alanine (N, A) to (N/T) and (A/Q) respectively and finally EPIYA-C in which the amino acid in the 11th position changed from Glycine (G) to (G/R) (data not shown, accession number: FJ915913), which had been previously reported by Occhialini[34]. There was no association between the phosphorylation motif, EPIYT, which has a threonine residue instead of an alanine residue and the studied pathologies (P = 0.51). However, it has been reported that isolates that harbor the ABCC genotype and have a modification in the 5th residue (EPIYT) in the B type may induce lower levels of cellular elongation and interleukin-8 secretion than isolates with the normal ABCC pattern[42].

EPIYA results from this study were compared with those reported by Argent et al[39]. The analyses showed that in our samples the ABC pattern of EPIYA motifs is the most common (62.3%) (Table 1 and Figure 3), which is in agreement with Western sequences reported before (63.3%)[39]. Interestingly, we found a statistical difference between the frequency of ABCC patterns in Western strains reported in GenBank (19.6%) and our results (31.1%) (Table 1) (P = 0.002). A search in the GenBank data base was performed in order to look for sequences reported in Colombia[40][46] that included the C-terminal variation region of CagA. A comparison between the frequencies of the different EPIYA patterns with our results showed that there is not a statistically significant difference (Table 1) (P = 0.98). Current research showed that the variation of the EPIYA motifs in CagA protein is not directly associated with the outcome of the disease caused by H. pylori and that there is no association between the number of EPIYA motifs and the pathology (Table 2). Similar results had been reported before in Iranian and Iraqi populations[47], where it seems there is no explicit positive correlation between the number of EPIYA motifs and various gastroduodenal diseases associated to H. pylori infection. In contrast, Yamaoka et al[47] showed that 7 out of 8 H. pylori Colombian strains (87.5%) with more than 3 EPIYA-C motifs were from gastric cancer patients. More recently Sicinschi et al[45] evaluated the 3’ cagA region in 66 isolates from Colombian patients with gastric precancerous lesions, from areas with low (31 strains) and high risk of gastric cancer (35 strains) in the south of the country. The proportion of strains bearing one EPIYA-C (62.2%), two EPIYA-C (34.3%) and three EPIYA-C (1.5%) motifs were similar to the ones observed in our population of strains. In contrast with our results, they observed a significant association between the presence of two or three EPIYA-C motifs and more severe lesions. Particularly they found a very high prevalence of strains with more than two EPIYA-C motifs in individuals with intestinal metaplasia (16/27, 59%). Sicinschi et al[45] included in the study only isolates
from men aged between 39 and 69 years in order to have high prevalence of *H. pylori* infection and preneoplastic lesions; 40% of the analyzed strains were from intestinal metaplasia. Our study included isolates from the Andes Mountains in the central part of the country, a high risk gastric cancer area; we included strains from men and women aged between 19 and 80 years, furthermore we included strains from gastric cancer patients. These differences in the design of the studies and also variations in geographic areas could explain the different results obtained. A limitation of our study is the small size of the analyzed population of strains. Although we could not find differences between EPIYA-C motifs and clinical outcomes, high prevalence of strains with more than one EPIYA-C motif might explain in part the high incidence of gastric cancer in Colombia.

It also must be considered that the absence of association between the CagA polymorphisms and the pathogenesis could be due to other factors such as genetic features of the host and cellular and extracellular environmental variables, which could influence the development of gastric cancer in combination with the type and number of EPIYA motifs in CagA protein. Moreover, another possibility as stated by Püls *et al*[^54^], using site-specific mutagenesis, suggests that tyrosine phosphorylation at EPIYA-C is sufficient, but not exclusive, to activate translocated CagA, suggesting that other motifs besides EPIYA-C are used for phosphorylation of CagA proteins as well. For this reason it is not only important to study the tyrosine phosphorylation motif, EPIYA, but also other new repetitions that have been discovered recently, such as the 7-AA sequence KIDQLNQ, which occurs in and near the CagA variable region[^28^]. Another important amino acid sequence is FPLXXXVXDLSKVG (Figure 2, second square), which surrounds the EPIYA-C motif and EPIYA-D in Western and East Asian CagA species, this was called CagA multimerization (CM) motif[^54^]. This CM motif was characterized in all of our strains, and interestingly in the strain 5-22019 (GenBank access number: FJ915938) the mentioned motif does not only surround the EPIYA-C motif but also the EPIYA-A. For this reason, it is important to study the molecular interaction between CagA proteins through the CM motif, because CagA multimerization is critically involved in the formation of the CagA-SHP2 signaling complex, during cagA-positive *H. pylori* infection[^54^]. Thus, it will be important to study the biological activity of Colombian CagA variants that harbor a variation in the C-terminal of the CagA protein, in order to comprehend how the interaction between the EPIYA-C motif and the SRC-2 protein in the epithelial cells works.

At the structural level it is important to determine how the EPIYA motifs are located in the CagA protein and if it influences CagA-SRC interaction, and consequently contribute to elucidate the biological activity of the CagA protein in the development of carcinogenesis in epithelial cells through this interaction. It is also important for future studies to check the evolution of the cagA gene in patients with atrophic gastritis, because this pathology has been reported as an early step in the development of gastric cancer[^51^] and this combined with genetic variability on the cagA gene generated by multiple rearrangements such as insertions, deletions, and substitution events on a 102nt region encoding the EPIYA-C motif could increase the probability of developing a gastric cancer pathology[^52^].

This kind of study has already been developed through hybridization assays with whole-genome DNA microarrays studies from Colombian strains[^53^] and monitoring the acquisition of additional EPIYA-C motifs throughout adulthood to determine if they contribute to *H. pylori* biological activity at a later age[^25^].

Our findings suggest that CagA polymorphisms EPIYA motifs, in Colombian patients are not clearly associated with the outcome of the disease. For this reason, it could be important to evaluate if the EPIYA motifs variation can cause an effect in the CagA protein structure and if so, if it can be correlated with its interaction inside the gastric epithelial cells. CagA has a striking functional similarity with Gab, Grb2 associated binder adaptor, in terms of being a protein that recruits SHP-2 to the plasma membrane and activates SHP-2 phosphatase by forming a physical complex[^54^]. Accordingly, it could be important to understand the role of CagA variants inside the gastric epithelial cells at protein-protein interaction level. On the other hand, microevolution in cagA gene has important implications for *H. pylori* pathogenesis because changing CagA type may provide a partial explanation of why studies examining the relationship between bacterial virulence and disease do not show tighter associations[^43^]. Finally, the CagA variable region characterization was successfully achieved and it gave an important register of the CagA protein polymorphisms in Colombian patients.

**ACKNOWLEDGMENTS**

We would like to thank colleagues who developed the bioinformatics software (AASA): Eric Gerard Rothstein Morris, Miguel Andrés Yáñez Barreto and Sonia Vivas from the Engineering Faculty, Los Andes University, Bogotá, Colombia.

**COMMENTS**

**Background**

*Helicobacter pylori* (*H. pylori*) is one of the most common worldwide human pathogens and it is present in at least 50% of the world’s population. Studies have suggested that chronic infection by *H. pylori* is an important risk factor for the development of gastric carcinoma, due to the presence of an important virulence factor, CagA protein.

**Research frontiers**

CagA protein is a potential factor in the development of gastric carcinogenesis in human gastric epithelial cells. However, the relation between the polymorphisms on CagA and the pathogenesis of *H. pylori* in gastric cells has not been addressed in terms of number and type of Glutamine acid-Proline-Isoleucine-Tyrosine-Alanine (EPIFY) motif located in the C-terminal of CagA protein.

**Innovations and breakthroughs**

Some studies argue that there is an association between the EPIYA motifs in CagA and the pathology presented by the patients. In these studies, the authors showed that this association is not present in Colombian patients. Furthermore, their results suggest that this association is not equal in different regions of the world because there are others factors such as structural conformation and

[^28^]: Püls et al. (2004)  
[^43^]: Püls et al. (2004)  
[^51^]: Püls et al. (2004)  
[^52^]: Püls et al. (2004)  
[^53^]: Püls et al. (2004)  
[^54^]: Püls et al. (2004)
protein-protein interactions that could affect the final outcome of the transloca-
tion of CagA in the gastric epithelial cells.

Applications
Because there is no association in Colombian samples between the motifs in
the C-terminal of CagA and the pathology, it may be important to understand
how these motifs interact with their target on gastric cells (Src kinase and Src
homology phosphatase 2 proteins) in order to elucidate the function of CagA
in the development of cancer gastric.

Terminology
CagA is one of the major virulence factors of H. pylori. This protein is involved
in the development of gastric cancer because when this protein is translocated
into gastric cells it causes changes in the signal transduction pathway which are
involved in the proliferation of the cells.

Peer review
The paper by Acosta et al assesses EPIYA in Colombian patients with dyspepsia
and finds no association with pathology.

REFERENCES
1 Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet 1985; 1: 1273-1275
2 Everhart JE. Recent developments in the epidemiology of Helicobacter pylori. Gastroenterol Clin North Am 2000; 29: 559-578
3 Kauser F, Khan AA, Hussain MA, Carroll IM, Ahmad N, Ti-
wardi S, Shouche Y, Das B, Alam M, Ali SM, Habibullah CM, Sierra S, Megraud F, Sechi LA, Ahmed N. The cag pathoge-
nicity island of Helicobacter pylori is disrupted in the majority
of patient isolates from different human populations. J Clin Microbiol 2004; 42: 5302-5308
4 Dooley CP, Cohen I, Fitzgibbons PL, Bauer M, Appleman MD, Perez-Perez GI, Blaser MJ. Prevalence of Helicobacter pylori
protein infection and histologic gastritis in asymptomatic per-
sons. N Engl J Med 1989; 321: 1562-1566
5 Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yama-
guchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. Helicobacter pylori infection and the development of gastric
cancer. N Engl J Med 2001; 345: 784-789
6 Wong BC, Lam SK, Wong WM, Chen JS, Zheng TT, Feng
RE, Lai KC, Hu WH, Yuen ST, Leung SY, Fong DY, Ho J, Ch-
ing CK, Chen JS. Helicobacter pylori eradication to prevent
cancer gastric in a high-risk region of China: a randomized
controlled trial. JAMA 2004; 291: 187-194
7 IARC. IARC monographs of the evaluation of carcinogenic
risks to humans. IARC 1994; 61: 177-241
8 Cover TL, Dooley CP, Blaser MJ. Characterization of and
human response to proteins in Helicobacter pylori
broth culture supernatants with vacuolizing cytotoxin activ-
ity. Infect Immun 1990; 58: 603-610
9 Han da O, Naito Y, Yoshikawa T. CagA protein of Helico-
bacter pylori: a hijacker of gastric epithelial cell signaling. Biochem Pharmacol 2007; 73: 1697-1702
10 Kim JM, Kim JS, Jung HC, Song JS, Kim CY. Virulence fac-
tors of Helicobacter pylori in Korean isolates do not influence
proinflammatory cytokine gene expression and apoptosis in human gastric epithelial cells, nor do these factors influence the
clinical outcome. J Gastroenterol 2000; 35: 898-906
11 Blaser MJ, Berg DE. Helicobacter pylori genetic diversity
and risk of human disease. J Clin Invest 2001; 107: 767-773
12 Aras RA, Lee Y, Kim SK, Israel D, Peek RM Jr, Blaser MJ.
Natural variation in populations of persistently colonizing bacteria affect human host cell phenotype. J Infect Dis 2003;
188: 486-496
13 Evans DJ Jr, Evans DG. Helicobacter pylori CagA: analysis
of sequence diversity in relation to phosphorylation motifs
and implications for the role of CagA as a virulence factor.
Helicobacter 2001; 6: 187-198
14 Kersulyte D, Chalkauskas H, Berg DE. Emergence of recom-
binant strains of Helicobacter pylori during human infection.
Mol Microbiol 1999; 31: 31-45
15 Backert S, Ziska E, Brinkmann V, Zimmy-Arntz U, Faucon-
nier A, Jungblut FR, Naumann M, Meyer TF. Translocation of the Helicobacter pylori CagA protein in gastric epithelial
cells by a type IV secretion apparatus. Cell Microbiol 2000; 2: 155-164
16 Segal ED, Cha J, Lo J, Falkov S, Tompkins LS. Altered
states: involvement of phosphorylated CagA in the induc-
tion of host cellular growth changes by Helicobacter pylori.
Proc Natl Acad Sci USA 1999; 96: 14559-14564
17 Selbach M, Moese S, Hauck CR, Meyer TF, Backert S. Src is
the kinase of the Helicobacter pylori CagA protein in vitro
and in vivo. J Biol Chem 2002; 277: 6775-6778
18 Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S,
Misselwitz R, Berger J, Seward N, König W, Backert S. Helicobacter exploits integrin for type IV secretion and kinase
activation. Nature 2007; 449: 862-866
19 Argent RH, Zhang Y, Atherton JC. Simple method for deter-
mation of the number of Helicobacter pylori CagA variable-region EPIYA tyrosine phosphorylation motifs by PCR.
J Clin Microbiol 2005; 43: 791-795
20 Asahi M, Azuma T, Ito S, Ito Y, Sato H, Nagai Y, Tsubokawa
M, Tohyma Y, Maeda S, Orama M, Suzuki T, Sasakawa C.
Helicobacter pylori CagA protein can be tyrosine phosphor-
ylated in gastric epithelial cells. J Exp Med 2000; 191: 593-602
21 Odenbreit S, Püls J, Sedlmaier B, Gerland E, Fischer W,
Haas R. Translocation of Helicobacter pylori CagA into
gastric epithelial cells by type IV secretion. Science 2000; 287: 1497-1500
22 Panayotopoulou EG, Sgouras DN, Papadakos K, Kalliaro-
polous A, Papanoteodoris G, Mentis AF, Archimandritis AJ.
Strategy to characterize the number and type of repeating
EPIYA phosphorylation motifs in the carboxyl terminus of CagA protein in Helicobacter pylori clinical isolates. J Clin Microbiol 2007; 45: 488-495
23 Selbach M, Moese S, Hauck CR, Meyer TF, Backert S. Src is
the kinase of the Helicobacter pylori CagA protein in vitro
and in vivo. J Biol Chem 2002; 277: 6775-6778
24 Stein M, Rappuoli R, Covacci A. Tyrosine phosphoryla-
tion of the Helicobacter pylori CagA antigen after cag-driven host
cell translocation. Proc Natl Acad Sci USA 2000; 97: 1263-1268
25 Tammer I, Brandt S, Hartig R, König W, Backert S. Activa-
tion of Abl by Helicobacter pylori: a novel kinase for CagA and crucial mediator of host cell scattering. Gastroenterology 2007;
132: 1309-1319
26 Bentires-Alj M, Paoz JC, David FS, Keilhack H, Halmos B,
Naoki K, Maris JM, Richardson A, Bardelli A, Sugarbaker
DJ, Richards WG, Du J, Giraud L, Minna JD, Loh ML, Fisher
DE, Velculescu VE, Vogelstein B, Meyerson M, Sellers WR,
Neel BG. Activating mutations of the noonan syndrome
gene SH2P2/PTPN11 gene in human solid tumors and
adult acute myelogenous leukaemia. Cancer Res 2004; 64:
8816-8820
27 Tartaglia M, Niemeyer CM, Fragale A, Song X, Buechner
J, Jung A, Hählen K, Hasle L, Hich J, Gellb BD. Somatic
mutations in PTPN11 in juvenile myelomonocytic leukaemia,
myelodysplastic syndromes and acute myeloid leukaemia.
Nat Genet 2003; 34: 148-150
28 Covacci A, Censini S, Bugnoli M, Petracca R, Burroni D,
Macchia G, Massone A, Papini E, Xiang Z, Figura N. Mole-
cular characterization of the 128-kDa immunodominant anti-
gen of Helicobacter pylori associated with cytotoxicity and
duodenal ulcer. Proc Natl Acad Sci USA 1993; 90: 5791-5795
29 Yamaoka Y, Osato MS, Sepulveda AR, Gutierrez O, Figura
N, Kim JS, Kodama T, Shima K, Graham DY. Molecular
epidemiology of Helicobacter pylori: separation of H. pylori
from East Asian and non-Asian countries. Epidemiol Infect 2000;
124: 91-96
30 Higashi H, Tsutsumi R, Fujita A, Yamazaki S, Asaka M,
Azuma T, Hatakeyama M. Biological activity of the Helico-
bacter pylori virulence factor CagA is determined by varia-
tion in the tyrosine phosphorylation sites. Proc Natl Acad Sci
Acosta N et al. CagA protein polymorphisms and pathogenesis

USA 2002; 99: 14428-14433

31 Kanada R, Uchida T, Tsukamoto Y, Nguyen LT, Hijiya N, Matsuura K, Kodama M, Okimoto T, Murakami K, Fujioka T, Yanagisawa S, Moriyama M. Genotyping of the cagA gene of Helicobacter pylori on immunohistochemistry with East Asian CagA-specific antibody. Pathol Int 2008; 58: 218-225

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

33 Azuma T, Yamazaki S, Yamakawa A, Ohtani M, Muramatsu A, Sato H, Ito Y, Dogo M, Yamazaki Y, Kuriyama M, Keida Y, Higashi H, Hatakeyama M. Association between diversity in the Src homology 2 domain-containing tyrosine phosphatase binding site of Helicobacter pylori CagA protein and gastric atrophy and cancer. J Infect Dis 2004; 189: 820-827

34 Backert S, Moese S, Sellaich M, Brinkmann V, Meyer TF. Phosphorylation of tyrosine 972 of the Helicobacter pylori CagA protein is essential for induction of a scattering phenotype in gastric epithelial cells. Mol Microbiol 2001; 42: 631-644

35 Higashi H, Tsutsui M, Ruto S, Sugiyama T, Azuma T, Asaka M, Hatakeyama M. SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein. Science 2002; 295: 683-686

36 Argent RH, Kidd M, Owen RJ, Thomas RJ, Limb MC, Atherton JC. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of Helicobacter pylori. Gastroenterology 2004; 127: 514-523

37 Yamaoka Y, El-Zimaity HM, Gutierrez O, Figura N, Kim JG, Kodama T, Kashima K, Graham DY. Relationship between the cagA 3′ repeat region of Helicobacter pylori gastric histology, and susceptibility to low pH. Gastroenterology 1999; 117: 342-349

38 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680

39 Argent RH, Hale JL, El-Omar EM, Atherton JC. Differences in Helicobacter pylori CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. J Med Microbiol 2008; 57: 1062-1067

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

41 Higashi H, Yokoyama K, Fujii Y, Ren S, Yuasa H, Saadat I, Murata-Kamiya N, Azuma T, Hatakeyama M. EPIYA motif is a membrane-targeting signal of Helicobacter pylori virulence factor CagA in mammalian cells. J Biol Chem 2005; 280: 23130-23137

42 Reyes-Leon A, Atherton JC, Argent RH, Puente JL, Torres J. Heterogeneity in the activity of Mexican Helicobacter pylori strains in gastric epithelial cells and its association with diversity in the cagA gene. Infect Immun 2007; 75: 3445-3454

43 Argent RH, Thomas RJ, Aviles-Jimenez F, Letey DP, Limb MC, El-Omar EM, Atherton JC. Toxigenic Helicobacter pylori infection precedes gastric hypochlorhydria in cancer relatives, and H. pylori virulence evolves in these families. Clin Cancer Res 2008; 14: 2227-2235

44 Occhialini A, Marais A, Urduci M, Sienna R, Muñoz N, Covacci A, Mérraud F. Composition and gene expression of the cag pathogenicity island in Helicobacter pylori strains isolated from gastric carcinoma and gastritis patients in Costa Rica. Infect Immun 2001; 69: 1902-1908

45 Sicinski LA, Correa P, Peek RM, Camargo MC, Piazuelo MB, Romero-Gallo J, Hobbs SS, Krishna U, Delgado A, Mera R, Bravo LE, Schneider BG. CagA C-terminal variations in Helicobacter pylori strains from Colombian patients with gastric precancerous lesions. Clin Microbiol Infect 2010; 16: 369-378

46 Yamaoka Y, Orito E, Mizokami M, Gutierrez O, Saitou N, Kodama T, Osato MS, Kim JG, Ramirez FC, Mahachai V, Graham DY. Helicobacter pylori in North and South America before Columbus. FEBS Lett 2002; 517: 180-184

47 Shokrzadeh L, Baghaei K, Yamaoka Y, Dabiri H, Jafari F, Sahabehtiani N, Tahami A, Sugimoto M, Zojaji H, Zali MR. Analysis of 3′-end variable region of the cagA gene in Helicobacter pylori isolates from Iranian population. J Gastroenterol Hepatol 2010; 25: 172-177

48 Püls J, Fischer W, Haas R. Activation of Helicobacter pylori CagA by tyrosine phosphorylation is essential for dephosphorylation of host cell proteins in gastric epithelial cells. Mol Microbiol 2002; 43: 961-969

49 Evans DJ Jr, Evans DG. Direct repeat sequences in the cagA gene of Helicobacter pylori: a ghost of a chance encounter? Mol Microbiol 1997; 23: 409-410

50 Ren S, Higashi H, Lu H, Azuma T, Hatakeyama M. Structural basis and functional consequence of Helicobacter pylori CagA multimerization in cells. J Biol Chem 2006; 281: 32344-32352

51 Correa P. A human model of gastric carcinogenesis. Cancer Res 1988; 48: 3554-3560

52 Alvi A, Devi SM, Ahmed I, Hussain MA, Rizwan M, Lamouliatte H, Mérraud F, Ahmed N. Microevolution of Helicobacter pylori type IV secretion systems in an ulcer disease patient over a ten-year period. J Clin Microbiol 2007; 45: 4039-4043

53 Kraft C, Stack A, Josenhans C, Niehus E, Dietrich G, Correa P, Fox JG, Falush D, Suerbaum S. Genomic changes during chronic Helicobacter pylori infection. J Bacteriol 2006; 188: 249-254

54 Hatakeyama M. Helicobacter pylori CagA--a potential bacterial oncoprotein that functionally mimics the mammary Gab family of adaptor proteins. Microbes Infect 2003; 5: 143-150

S- Editor Wang YR  L- Editor O'Neill M  E- Editor Lin YP