Variation in the number of EPIYA-C repeats in CagA protein from Colombian Helicobacter pylori strains and its ability to induce hummingbird phenotype in gastric epithelial cells

Andrés Javier Quiroga1,2, Antonio Huertas2, Alba Lucía Cómbita1,2, María Mercedes Bravo1,2

1 Grupo de Investigación en Agentes Infecciosos y Cáncer, Instituto Nacional de Cancerología
2 Grupo de Investigación en Biología del Cáncer, Instituto Nacional de Cancerología, Bogotá D.C., Colombia.

Introduction. Studies using Western Helicobacter pylori strains have shown that a risk factor for gastric cancer is the number of EPIYA-C motifs in the cytotoxin-associated A protein. CagA is delivered into epithelial cells, where it becomes tyrosine phosphorylated in their EPIYA repeats and induces cytoskeleton rearrangements.

Objectives. The objective of this study was to evaluate H. pylori cagA positive strains isolated from Colombian patients with gastroduodenal diseases for the number of EPIYA-C repeats in cagA and their ability to induce cytoskeleton rearrangements in epithelial cells.

Materials and methods. We analyzed the 3' EPIYA repeats region of cagA by PCR in 93 H. pylori cagA positive strains from 49 patients with gastritis, 17 with gastric cancer, and 24 with duodenal ulcer. AGS cells exposed to the various H. pylori isolates were evaluated for rearrangements in their cytoskeleton.

Results. Strains with one EPIYA-C were the most frequent in gastritis and duodenal ulcer patients. Strains with three EPIYA-C were mainly found in gastric cancer. We found a significantly higher risk of gastric cancer for individuals infected with strains harboring three EPIYA-C motifs (OR=12.4, CI95%: 2.32-66.3). Strains from gastric cancer showed significantly higher percentages of induction of cytoskeleton rearrangements in comparison with those from gastritis (p Mann-Whitney<0.005).

Conclusions. H. pylori strains with three EPIYA-C repeats can confer an increased risk of cancer to infected individuals.

Key words: Helicobacter pylori; polymorphism, genetic; gastritis, stomach neoplasm, duodenal ulcer.

Variación en el número de repeticiones EPIYA-C en la proteína CagA de aislamientos colombianos de Helicobacter pylori y su capacidad para inducir fenotipo colibrí en células epiteliales gástricas

Introducción. En los aislamientos de Helicobacter pylori del hemisferio occidental, se ha observado que el número de repeticiones EPIYA-C en la proteína CagA es un factor de riesgo para cáncer gástrico. La proteína CagA es introducida en la célula epitelial y, posteriormente, es fosforilada en las tirosinas presentes en los motivos EPIYA e induce rearreglos en el citoesqueleto.

Objetivos. Nuestro propósito fue evaluar el número de repeticiones EPIYA-C y la habilidad para inducir rearreglos en el citoesqueleto en los aislamientos de H. pylori positivos para cagA, provenientes de pacientes colombianos con enfermedad gastroduodenal.

Materiales y métodos. Mediante PCR, se analizó la región 3´ que contiene las repeticiones EPIYA-C, en 93 aislamientos de H. pylori positivos para cagA provenientes de 49 pacientes con gastritis, 17 con cáncer gástrico y 24 con úlcera duodenal. Los rearreglos del citoesqueleto se evaluaron mediante cultivos simultáneos de células AGS con las cepas de H. pylori.

Resultados. En gastritis y úlcera duodenal se observó la mayor frecuencia de aislamientos con EPIYA C; los aislamientos con tres repeticiones EPIYA-C se encontraron con mayor frecuencia en cáncer gástrico. Encontramos un riesgo de cáncer gástrico significativamente mayor para individuos infectados con cepas con tres repeticiones EPIYA-C (OR=12.4; IC95% 2.32-66.3). Los aislamientos provenientes de cáncer gástrico mostraron mayores porcentajes de inducción de rearreglos en el citoesqueleto que los observados con aislamientos provenientes de gastritis (prueba de Mann-Whitney menor de 0,005).

Conclusiones. La infección con cepas de H. pylori con tres repeticiones EPIYA-C puede conferir un mayor riesgo de desarrollar cáncer gástrico.

Palabras clave: Helicobacter pylori, proteína CagA, polimorfismo, gastritis, neoplasias del estómago, úlcera duodenal.
Helicobacter pylori is a human gastric pathogen which chronically infects more than half the world’s population (1). It colonizes the human stomach and persists for decades, causing chronic gastritis. In a small fraction of individuals, the infection progresses to peptic ulcer disease, gastric adenocarcinoma, or mucosa-associated lymphoid tissue (MALT) lymphoma (2-7). The causes for the different outcomes of H. pylori infections may include bacterial virulence, host susceptibility, and environmental factors (8-11).

Individuals infected with strains harbouring the cag pathogenicity island (cag-PAI) are at a higher risk of developing duodenal ulcer, atrophic gastritis, and gastric cancer (8,12-15). The cag-pathogenicity island (PAI) is a 40 kb DNA segment which contains genes encoding for a type IV secretion system (16,17), which forms a syringe-like structure that penetrates epithelial cells and allows the translocation of CagA, the product of cagA gene (16,18). Once introduced into the cell, CagA is phosphorylated by src family kinases on specific EPIYA sequences repeats (18-20) and then binds and activates phosphatase SHP-2 to induce an elongated cell shape known as hummingbird phenotype (21).

CagA proteins show size variation due to the presence of different types and/or numbers of repeat sequences containing the EPIYA motifs within the C-terminal variable region (7). Four types of EPIYA segments have been described: A, B, C, and D (22). East Asian H. pylori strains commonly have an ABD pattern, while Western H. pylori strains almost always contain EPIYA-A and EPIYA-B segments, followed by one to three EPIYA-C segments. It has been shown that CagA with more EPIYA-C motifs induced more cytoskeleton rearrangements (hummingbird phenotype) (22) and Western H. pylori strains with more than one EPIYA-C motifs have been significantly associated with gastric cancer (6,23-25).

In this study we analyzed the number of CagA EPIYA-C segments and the ability of the CagA protein expressed by H. pylori strains isolated from Colombian patients with gastroduodenal diseases to induce a hummingbird phenotype.

Materials and methods

Bacterial strains

A total of 93 H. pylori cagA positive strains originally isolated from patients living in Bogotá, Colombia between 1998 and 2002 (12,15) were obtained from the H. pylori stock collection of the Instituto Nacional de Cancerología, in Bogotá, Colombia. These strains were isolated from antral gastric biopsy specimens from 52 patients with gastritis (29 male and 23 female; mean age 52 years; age range 31-78 years), 17 patients with gastric cancer (12 male, 5 female; mean age 56, age range 42-67 years), and 24 patients with duodenal ulcer (11 male, 13 female; mean age 37, range 19-57). Two reference strains were included: NCTC 11638 and NCTC 11637. These strains 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 under microaerophilic conditions. Chromosomal DNA was prepared by DNazol (Invitrogen, Carlsbad, CA) extraction method, according to the manufacturer’s instructions. Ethics committee of the Instituto Nacional de Cancerología approved the protocol for this study. All the strains were isolated from patients that provided informed consent.

Amplification of the 3'-region of cagA by PCR

The primers CAGTF 5’-ACCCTAGTCGGTAATGGG-3’ and CAGTR 5’-GCTTTAGCTTCTGAYACYGC, described by Yamaoka et al. (24), were used to amplify the entire 3’ repeat region of cagA. PCR was performed in a volume of 25 µL containing 10 mM Tris, 50 mM KCl, 2.5 mM MgCl₂, 200 µM dNTPs, 25 pmol of the primers, 100 ng H. pylori genomic DNA, and 1U Taq polymerase (Promega, Madison, WI). The PCR conditions included an initial denaturation step at 94°C for 5 min, followed up by 35 cycles consisting of 1 min at 94°C, 1 min at 52°C, and 1 min at 72°C, and then an extension step at 72°C for 7 min. PCR products were analyzed by 1.5% agarose gel electrophoresis with ethidium bromide staining.

Sequence analysis

To sequence the 3’ repeat region of cagA, PCR amplified products were excised from a 1.5% agarose gel and purified by using a GFX PCR...
DNA purification kit (Amersham, Buckinghamshire, UK), and cloned into pCR 4 TOPO (Invitrogen, Paisley, UK). DNA sequencing was performed with a 7-Deaza-dGTP Cy5/5.5 Dye Primer Cycle DNA sequencing kit (Visible Genetics, Toronto, Canada) by using M13R universal primer labelled with the Cy5.5 fluorescent dye. M13 fluorescent primers targeting the margins of the cloning sites were used for DNA sequencing of both strands. Nucleotide sequences were aligned and analyzed with BLAST tools. The previous published cagA sequences of strains NCTC 11638 (gene bank access AF282853) and NCTC 11637 (gene bank access AF202973) were also included in the analysis. The deduced peptide sequences containing the EPIYA motifs were aligned by CLUSTALW (http://www.ebi.ac.uk/Tools/clustalw2).

**AGS cell hummingbird phenotype formation**

AGS cell hummingbird phenotype formation was evaluated according to Argent et al. with minor modifications (6). Briefly, AGS cells were grown in six-well plates at a density of 2x10⁵ cells/well in RPMI 1640 (Sigma Chemical Co., St. Louis, MO), supplemented with 10% heat-inactivated fetal bovine serum (Invitrogen/GIBCO BRL, Grand Island, NY) and incubated for 24 hours at 37°C in a 5% CO₂-air humidified atmosphere to obtain monolayers with a 70% to 80% confluence. *H. pylori* strains were harvested from 48 to 72 hour growth plates, suspended in RPMI 1640 medium, adjusted to an optical density at 600 nm of 0.1 (1 OD₆₀₀ₙₐₜ = 2.4 x 10⁸ CFU/mL) and added to the AGS cells at a multiplicity of infection (MOI) of 100. *H. pylori* strains were cocultured in triplicate with AGS cells for 24 hours at 37 °C in a 5% CO₂-air humidified atmosphere, and then the cells were examined for hummingbird phenotype formation by microscopy (Axiovert 135 microscope, Zeiss, Germany) reading three randomly chosen fields in each of three wells. The results were reported as mean percentages of cells exhibiting the cellular elongation effect in the three wells. A hummingbird cell was defined as a cell with a needle-like protrusion >2 µm (6).

**Statistical analysis**

The statistical analysis of data was made by using univariate logistical regression, X² test, and non-parametric Mann-Whitney test in SPSS for Windows version 16.0 software. A p <0.05 value was considered statistically significant.
seven isolates were noted to harbor one and two EPIYA-C motifs and one isolate harbored one and three EPIYA-C motifs.

**Association of the predicted number of EPIYA-C motifs with gastric cancer**

We analyzed the association between the number of EPIYA-C motifs at the 3’-end of \( cagA \) and the clinical origin of the strains (Table 1). Clinical isolates involving more than one strain were classified according to the highest observed number of EPIYA-C motifs. Strains with one EPIYA-C motif were the most frequent in gastritis and duodenal ulcer and were the least frequent in cancer patients. Strains with two EPIYA-C motifs showed frequencies between 34.6% and 45.8%, with gastritis showing

### Table 1. Association between the *Helicobacter pylori* \( cagA \) EPIYA-C number and the clinical origin of strains.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>No more than 1 EPIYA-C n (%)</th>
<th>2 EPIYA-C n (%)</th>
<th>2 EPIYA-C OR a</th>
<th>CI 95%</th>
<th>3 EPIYA-C n (%)</th>
<th>3 EPIYA-C OR a</th>
<th>CI 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastritis</td>
<td>31 (59.6)</td>
<td>18 (34.6)</td>
<td>1</td>
<td>--</td>
<td>3 (5.8)</td>
<td>1</td>
<td>--</td>
</tr>
<tr>
<td>Duodenal ulcer</td>
<td>12 (50)</td>
<td>11 (45.8)</td>
<td>1.6</td>
<td>0.6-4.3</td>
<td>1 (4.2)</td>
<td>0.9</td>
<td>0.08-9.1</td>
</tr>
<tr>
<td>Gastric cancer</td>
<td>5 (29.4)</td>
<td>6 (35.3)</td>
<td>2.1</td>
<td>0.5-7.7</td>
<td>6 (35.3)</td>
<td>12.4^b</td>
<td>2.32-66.3</td>
</tr>
</tbody>
</table>

a 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 at least one EPIYA-C fragment.

b \( p<0.003 \).

---

Figure 2. Deduced amino acid sequences of six \( cagA \) variable regions, amplified by PCR with primers CAGTF5’-ACCCAGTGGTAATGGG-3’ and CAGTR 5’-GTCTTAGTCTGAGACYGC-3’. The type of EPIYA motif: A, B or C, is shown on top.

...
the smallest frequency, and duodenal ulcer had the highest. In gastric carcinoma patients, isolates harboring three EPIYA-C motifs represented 35.3% of cases, a higher percentage than the proportions observed in gastritis (5.8%) and duodenal ulcer (4.2%). We observed a significant higher risk of gastric cancer for individuals infected with strains harboring three EPIYA-C motifs in comparison with individuals infected with strains harboring one EPIYA-C motif (OR = 12.4, CI 95%: 2.32-66.3).

**Higher induction of cytoskeleton rearrangements in AGS cells cocultured with H. pylori isolates from gastric cancer**

In order to assess, for clinical strains, whether the number of EPIYA-C motifs in the CagA protein affected the extent of induction of hummingbird phenotype, we cocultured AGS cells with 69 H. pylori strains (40 from gastritis, 15 from gastric cancer and 14 from duodenal ulcer). In spite of the high variability observed in the percentages of hummingbird cells in the cocultures, the mean percentages of hummingbird cells increased when the number of EPIYA-C motifs in the strains was high (one EPIYA-C: 38.2 %, two EPIYA-C: 41.3%, three EPIYA-C: 45.81%). These differences were not statistically significant (Figures 3 and 4A). When we analyzed the extent of humming bird cells induction in relation to the type of gastroduodenal disease, strains from gastric cancer showed significantly higher percentages of induction of cytoskeleton rearrangements, in comparison with gastritis (gastric cancer: 48.86% vs. gastritis: 36.4%, p< 0.005) (figure 4b).

---

Figure 3. Hummingbird cell formation after coculture of AGS cells with *Helicobacter pylori* reference strain 11638 (A) and with strain 1039 expressing CagA with three EPIYA-C motifs, from a patient with gastric cancer (B). Magnification, X20; bar, 20 µm.

Figure 4. Percentage of cells with hummingbird phenotype according to number of EPIYA-c motifs (A) and according to gastroduodenal disease (B).
Discussion

Infection with cagA positive H. pylori strains is correlated with a significantly higher risk of atrophic gastritis, gastric cancer, and duodenal ulcer (8,14). However, most persons colonized by these strains never develop severe gastroduodenal disease and remain asymptomatic. In Colombia, 72% to 90% of H. pylori strains possess the cagA gene (15,27), and the incidence of gastric cancer is high (28). It has been shown that the number of EPIYA-C segments on CagA influences the degree of virulence as well as the oncogenic potential of individual cagA positive H. pylori strains (25,29). Further, among Western strains, the most important factor for cancer risk is the number of CagA EPIYA-C motifs (25). In order to evaluate, in a high gastric cancer incidence population, the clinical relevance of the variations of the 3'-region of the cagA gene, we analyzed the number of EPIYA-C repeats and the ability to induce a hummingbird phenotype of H. pylori cagA positive strains isolated from Colombian patients with gastroduodenal diseases. Most of the strains possessed a cagA gene with one EPIYA-C motif in the 3'-region; its frequency was 49%, similar to the one reported by Yamaoka for Colombian isolates (55%) (30), but lower to the reported for isolates from South Africa (73%) (6), USA (80%) (30), France (68%) (30), and Greece (67.3%) (31). The frequency of strains harboring a cagA gene with two or three EPIYA-C motifs was 41%, similar to the frequency reported by Yamaoka for Colombian strains (44 %) (30) and higher than that reported for strains from USA (17 %) (30), France (32%) (30), South Africa (27%) (6), and Greece (21%) (31). Geographic variations in the frequency of H. pylori strains with one, two, or three EPIYA-C repeats may be explained by different acidic conditions in the gastric mucosa, which could be related to differences among populations in the frequencies of cytokine gene polymorphisms that attenuate gastric acid secretions (32). However, Argent et al. found no association between the number of EPIYA-C motifs and acid secretion (33).

We found 8.6% of isolates displaying a double band with sizes corresponding to different numbers of EPIYA-C repeats, all of them in benign diseases. This goes in accordance with the reports of Basso et al. (25) and Panayotopoulou et al. (31), who reported a similar frequency of isolates with this characteristic. Probably some of these strains reflect the adaptation of H. pylori to changes in the environment of the stomach by microevolution generating clones within an individual with different types of CagA proteins and also different degrees of virulence (33-35), and some could correspond to mixed infections.

As observed in previous studies (25,36), strains with cagA with higher number of EPIYA-C motifs were associated with a higher gastric cancer risk than strains with only one EPIYA-C motif. The risk of gastric cancer was 12 times higher in patients infected with cagA positive strains with three EPIYA-C motifs as compared with patients infected with cagA positive strains with one EPIYA-C motif. The high frequency of strains with more than one EPIYA-C motif could explain in part the high incidence of gastric cancer in Colombia.

Higashi et al (22) demonstrated, by generating EPIYA mutants from the strain NTCC 11637, that CagA with more EPIYA-C motifs induced more cytoskeleton rearrangements (hummingbird phenotype). Argent et al. also reported an association between the number of EPIYA-C motifs and the extent of AGS cell elongation in strains from South Africa (6). In this study, we did not observe this association; instead, we found a high variability in the extent of hummingbird phenotype induction for strains with one and two EPIYA-C (3% to 60%), and for strains with three EPIYA-C (35% to 61%). Although the mean percentages of cytoskeleton rearrangement induction in AGS cells tended to increase in the direction one EPIYA-C → two EPIYA-C → three EPIYA-C, the differences were not significant. These results suggest that the induction of the hummingbird phenotype must be influenced by other mechanisms besides translocation and phosphorylation of CagA on EPIYA-C motifs. Interestingly, strains from gastric cancer showed a significantly higher ability to induce elongation than strains from gastritis. Probably, the ability to induce elongation is associated with the virulence of cagA positive strains.

In conclusion, our study documents a high frequency of H. pylori cagA positive strains with more than one EPIYA-C repeats in a high gastric cancer incidence country like Colombia. We do not confirm a correlation between the number of EPIYA-C repeats and the extent of the induction of morphological changes in epithelial cells. We confirmed that, among CagA positive strains, those with three EPIYA-C motifs are associated with a higher gastric cancer risk. The number of EPIYA-C motifs could be used as a risk marker for individuals with nonatrophic gastritis infected with
cagA positive strains; this marker could be useful in a country like Colombia, with a high prevalence of infections with cagA positive H. pylori strains.

Conflicts of interest
We did not identify any conflict of interest.

Financing
This research was supported by a grant from the Instituto Nacional de Cancerología, Bogotá D.C., Colombia.

References


