Torres, Lino E.; González, Lidice; Melián, Karelia; Alonso, Jordis; Moreno, Arenis; Hernández, Mayrín; Reyes, Orlando; Bermúdez, Ludisleydis; Campos, Javier; Pérez-Pérez, Guillermo; Rodríguez, Boris L.

EPIYA motif patterns among Cuban Helicobacter pylori CagA positive strains


Instituto Nacional de Salud

Bogotá, Colombia

Available in: http://www.redalyc.org/articulo.oa?id=84322454003
PUBLICACIÓN ANTICIPADA EN LINEA

El Comité Editorial de Biomédica ya aprobó para publicación este manuscrito, teniendo en cuenta los conceptos de los pares académicos que lo evaluaron. Se publica anticipadamente en versión pdf en forma provisional con base en la última versión electrónica del manuscrito pero sin que aún haya sido diagramado ni se le haya hecho la corrección de estilo.

Siéntase libre de descargar, usar, distribuir y citar esta versión preliminar tal y como lo indicamos pero, por favor, recuerde que la versión impresa final y en formato pdf pueden ser diferentes.

Citación provisional:


Recibido: 26-05-11
Aceptado: 14-09-11
Publicación en línea: 14-09-11
EPIYA motif patterns among Cuban *Helicobacter pylori* CagA positive strains

**Patrón de los motivos EPIYA de cepas cubanas de *Helicobacter pylori* CagA-positivas**

**EPIYA pattern in Cuban *H. pylori* CagA+ isolates**

Lino E. Torres¹, Lidice González¹, Karelia Melián², Jordis Alonso², Arlenis Moreno¹, Mayrán Hernández¹, Orlando Reyes¹, Ludisleydis Bermúdez¹, Javier Campos³, Guillermo Pérez-Pérez⁴, Boris L. Rodríguez¹

¹ Department of Research and Development, Division of Production and Development of Equipments and Diagnostics, National Centre for Scientific Research, Ciudad de La Habana, Cuba.

² Department of Gastroenterology, Medical and Surgery Research Centre (CIMEQ), Ciudad de La Habana, Cuba.

³ Department of Molecular Biology, Biotechnology Division, National Centre for Scientific Research, Ciudad de La Habana, Cuba.

⁴ Departments of Medicine and Microbiology, New York University, School of Medicine, USA.

**Corresponding author:**

Boris Luis Rodríguez, Development and Research Department, Division of Equipments and Diagnostics, National Centre for Scientific Research, Ave. 25 and 158, Cubanacán, Playa, AP 6412, Ciudad de La Habana, Cuba.

Tel: (53) 7 2080959; fax: (53) 7 2087538.

boris.rodriguez@cnic.edu.cu
Author contributions: Torres LE and Rodríguez BL contributed equally to this work; Torres LE, González L, Campos J and Rodríguez BL designed the research; González L, Torres LE, Reyes O, Moreno A, Hernández M, Bermúdez L collected and processed the samples and performed the research; Melián K, Alonso J performed the endoscopy and managed the patient data; Torres LE, González L, Pérez-Pérez G and Rodríguez BL analyzed the data and wrote the paper.
Introduction. It is known that polymorphisms in C-terminal region of CagA influence gastric disease development on *Helicobacter pylori* infection. Additionally, the geographic distribution of these polymorphisms has been associated with the appearance of more severe gastroduodenal pathologies.

Objective. To determine the CagA phosphorylation motifs pattern (EPIYA pattern) in Cuban *H. pylori* isolates, and to study its association with patient's pathologies.

Materials and methods. DNAs from 95 *H. pylori cagA*-positive strains were used to amplify the 3’ variable region of cagA gene by PCR using two different strategies. Additionally, new primers were designed to identify either Western or Eastern CagA EPIYA motif type by PCR. To confirm the PCR results, PCR products from 14 representative isolates were purified and sequenced

Results. The distribution of the EPIYA motif found was: 2 AB (2.1 %), 1 AC (1.1 %), 1 BC (1.1 %), 70 ABC (73.6 %), 19 ABCC (20 %), and 2 ABCCC (2.1 %). Sequencing analysis confirmed the PCR classification in the 14 studied strains and showed three strains with unusual nucleotide sequences, not reported before. Distribution of the EPIYA-ABC pattern was equivalent in all pathologies (78.9 % in gastric ulcer, 72.5 % in duodenal ulcer and 72.2 % in non-ulcer dyspepsia).

Conclusion. The PCR results using the new primers confirmed that all studied strains carried the Western CagA type. No specific EPIYA motif was associated with peptic ulcer. This is the first report that shows EPIYA motif distribution in *H. pylori* isolates from the Caribbean region.

Key words: *Helicobacter pylori*, gastric neoplasms, virulence factors, Cuba.
Introducción. Se conoce que el polimorfismo en la región C-terminal de CagA influye en el desarrollo de la enfermedad gástrica durante la infección por *Helicobacter pylori*. 

Objetivo. Determinar el número y tipo de patrones de fosforilación de CagA (patrón EPIYA) en aislamientos cubanos de *H. pylori*, y estudiar su asociación con las patologías gástricas.

Materiales y métodos. Los ADNs de 95 cepas de *H. pylori* cagA-positivas, fueron empleados para amplificar la región 3' variable del gen cagA por PCR mediante el empleo de diferentes estrategias. Se diseñaron además nuevos cebadores para clasificar los aislados según el tipo occidental o del este asiático de CagA por PCR. Los productos de PCR obtenidos de 14 aislados representativos fueron purificados y secuenciados para confirmar los resultados del PCR.

Resultados. La distribución de los patrones EPIYA encontrada fue: 2 AB (2,1%), 1 AC (1,1%), 1 BC (1,1%), 70 ABC (73,6%), 19 ABCC (20%), y 2 ABCCC (2,1%). El análisis de la secuenciación confirmó las clasificaciones realizadas por PCR en las 14 cepas estudiadas y mostró 3 cepas con secuencias nucleotídicas únicas, no reportadas anteriormente. La distribución del patrón EPIYA-ABC fue equivalente en todas la patologías encontradas (78,9 % en úlcera gástrica, 72,5 % en úlcera duodenal y 72,2 % en dispepsia no ulcerada).

Conclusión. La mayoría de los aislados cubanos presentaron las combinaciones de motivos EPIYA menos virulentas (ABC). Los resultados del empleo de los nuevos cebadores y el análisis de la secuenciación confirmó que todas las cepas estudiadas portaban el CagA tipo occidental. Ninguno de los patrones específicos de EPIYA presentaron asociación con la úlcera péptica. Este es el primer reporte que muestra la distribución de los motivos EPIYA en los aislados de *H. pylori* de la región del Caribe.

Palabras clave: *Helicobacter pylori*, neoplasias gástricas, factores de virulencia, Cuba.
*H. pylori* strains producing the cytotoxin associated gene A protein (CagA) have been related with the development of severe gastroduodenal diseases, atrophic gastritis and gastric adenocarcinoma (1,2). Gastric carcinoma (GC) is one of the leading causes of cancer-related deaths in many parts of the world (3). The CagA protein is delivered into the epithelial cells for a Type IV Secretion System (4) and it is phosphorylated by host kinases on the tyrosine phosphorylation motifs (TPMs) called EPIYA (Glu-Pro-Ile-Tyr-Ala), located in the carboxyl-terminal region (5,6). Upon phosphorylation, CagA interacts with the phosphatase SHP-2 (7), activating the Erk MAP kinase cascade eliciting a cell-morphological transformation, which is characterized by a dramatic cytoskeletal rearrangements of epithelial cells (1,8). Additionally, phosphorylated CagA can stimulate multiple motility-inducing signaling cascades such as the activation of cAbl (9) and Crkl cascades(10). Therefore, deregulation of SHP-2 by CagA plays an important role in gastric carcinoma development (1). On the other hand, CagA also disturbs cell functions in a tyrosine phosphorylation-independent manner, activating a nuclear factor that induces transcription of several inflammatory genes (11). Furthermore, CagA is also capable of activating NF-κB, which in turn induces interleukin-8 expression(12) and more recently, it has been reported that this bacterial protein deregulate the β-catenin signaling (13).

CagA TPMs are classified as type A, B, C and D based on specific amino acid sequences surrounding the EPIYA motif (7). The *H. pylori* strains isolated from Western countries mainly consist of EPIYA-A, B, and a single or multiple repeat of C segment, while East Asian strains (Eastern strains) mainly have a combination of EPIYA-A, B and D motifs (7). SHP-2 binds both EPIYA-C and D motifs, but EPIYA-D exhibits greater SHP-2 binding activity and therefore a stronger ability to induce severe changes in epithelial cells (7,14). Consequently, the predominant presence of strains carrying an EPIYA-ABD type has been associated with the higher GC rate in East Asia (1,15,16). Furthermore, Western strains harboring multiple EPIYA-C motifs have been more frequently isolated from patients with gastric adenocarcinoma (17,18). The above observations support the hypothesis that
specific TPM patterns could determine the carcinogenic potential of *H. pylori* isolates, and partly explaining the geographic differences in GC incidence (1). In fact, very recently CagA have been recognized as the first bacterial oncoprotein which acts in mammalian cells (19) and its carcinogenic activity is modulated by specific TPM patterns (20). Therefore, accurate detection of the EPIYA pattern may become a useful prognostic tool for gastric malignancy due to *H. pylori* infection. Despite the relevance of TPM patterns in CagA protein as virulence markers, there are not reports about the EPIYA types among Cuban *H. pylori* isolates, or from strains isolated in the Caribbean region, despite that gastric cancer death rate varies significantly between countries in this region (21). In this disease, the highest mortality rates are exhibited by the Haitian males and Jamaican females with 25.1/100 000 and 11.6/100 000, respectively, while 9.7 and 5.6/100 000 are the lowest rates in the area, reported for Cuban males and females, respectively (21, http://www.sld.cu/servicios/estadisticas/). Therefore, the aim of this work was to study the EPIYA motif pattern of CagA gene in *H. pylori* isolates from a Caribbean population.

**MATERIALS AND METHODS**

*Patients and H. pylori isolates*

A total of 95 cagA positive *H. pylori* strains were isolated from antrum biopsies in a previous study involving 130 consecutive patients (77 male and 53 female) with a mean age of 49.1 (range 18 to 88) who underwent routine endoscopy due to dyspeptic complaints at CIMEQ Hospital, Havana, Cuba (22). Patients were classified in three groups (40 with duodenal ulcer, 19 with gastric ulcer, and 36 with non-ulcer dyspepsia) according to endoscopic and histological observations. *H. pylori* strains J99 and CCUG-17874 were used as reference strains, and kindly provided by Professor Francis Megraud from Pellegrin Hospital, Bordeaux, France and Professor Ann-Mary Svennerholm, Gothenburg University, Sweden, respectively. All the strains were cultured as previously described (22).

**DNA extraction and PCR amplification**
Genomic DNAs of the strains were extracted by CTAB method with phenol/chloroform and isopropanol precipitation as previously described (22). Purified DNAs were stored at -20ºC until use. PCR amplifications were carried out in a 25 μl reaction mixture containing 2.5 μl 10X PCR buffer, 0.2 mM of each deoxynucleotide triphosphate, 0.6 mM sense and antisense primers, 4 mM magnesium chloride, 1.25 U Taq DNA polymerase (CIGB, Cuba) and 100 ng genomic DNA. The PCR had an initial step at 94ºC for 1 min, followed by 40 cycles at 94ºC for 1 min, 60ºC for 1 min and 72ºC for 1 min, and a final extension at 72ºC for 5 min, using a Master Cycler apparatus (Eppendorf, Germany). Several primers were used for subtyping cagA variable region (table 1). PCR products were separated on 1.6 % agarose gel electrophoresis containing ethidium bromide. Images were digitalized using the Gene Genius system (Syngene, England).

**Sequence analysis of the 3' region of cagA gene**

DNAs amplified with primers cag2 and cag4 (Table 1) from 14 representative strains of different EPIYA patterns were isolated by using the ADN-GFX™ purification kit (Amersham-Pharmacia, Germany) and sent for sequence analysis to Macrogen Inc., Korea. Sequencing reactions were performed for both DNA strands, and the ClustalW2 multiple-sequence alignment software was used to analyse sequencing results.

**Specific subtyping of cagA gene into EPIYA-C or D**

Two new reverse primers were synthesized, cagA-EpiC and cagA-EpiD (table 1), to differentiate EPIYA motifs C and D by PCR. These reverse primers were used with forward primer cag2 (Table 1) and were designed using cagA gene sequences from Western and East Asian isolates published in Genbank Database and employing the Genrunner Software. Professor Guillermo Pérez-Pérez, from New York University, USA, kindly provided us with four DNAs from Eastern CagA isolates which were used as controls in our experiments.

**Accession numbers**
The nucleotide sequences of the cagA variable region from *H. pylori* strains in this study have been published in the GenBank database under accession numbers EU443785 (strain Hpcnic-35) and EU444054 to EU444066 (strains Hpcnic-22, Hpcnic-47, Hpcnic-74, Hpcnic-8, Hpcnic-10, Hpcnic-14, Hpcnic-15, Hpcnic-16, Hpcnic-17, Hpcnic-27, Hpcnic-19, Hpcnic-28 and Hpcnic-3).

**Statistical analysis**

Differences between groups were tested by χ² test. P values < 0.05 were considered to be significant. The Statistic software, version 8 for Windows, was used for statistical analysis.

**Results**

**Determination of EPIYA pattern by PCR**

Detection of the EPIYA pattern was made first by PCR amplification of 3′ variable region of cagA gene, by using primers originally designed to detect cagA gene (primers cag2 and cag4, table 1). The strategy renders an amplicon of varying sizes according to the number of EPIYA motifs (24). These PCR reactions generated DNA fragments of 400 to 750 bp that were visualized on agarose gel electrophoresis (figure 1). DNAs from the reference strains J99 and CCUG-17874 exhibited their expected sizes of the cagA gene fragment (figure 1). The appearance of a double band in the electrophoresis of PCR products may reveal the presence of a co-infection or microevolution of a single strain in a given patient. There were only three strains in which the presence of strains with more than one EPIYA pattern was confirmed (figure 1), and all of them carried a mixed BC-ABC EPIYA pattern. Although, the above PCR methods allow estimating the EPIYA pattern of *H. pylori* strains, the exact prediction of the EPIYA motifs of a given strain cannot be established. Therefore, the elegant PCR method from Argent et al, (25) was carried out to determine the exact EPIYA type and number in our isolates (data not shown). This methodology amplifies each EPIYA sequence in three different PCRs, using the forward primer cag2 with primers cagA-P1C, cagA-P2CG and cagA-P3 to amplify EPIYA motifs A, B and C, respectively (table 1).
The EPIYA motifs in 95 Cuban *H. pylori* strains shown that most of isolates (73.7%) carries three TPMs with a single copy of the EPIYA-C motif (table 2). Additionally, three of our strains had an EPIYA-C that differs from the normal size of this motif, and one of them also lacks most of its EPIYA-B segment (table 2). The nucleotide sequence of these three strains (Hpcnic-27, Hpcnic-35 and Hpcnic-47) were studied in detail by sequencing analysis (figure 2). Furthermore, nineteen (20 %) Cuban strains had the EPIYA-ABCC type and only two strains (2.1%) had the ABCCC pattern (table 2).

**Subtyping of cagA gene by PCR**

Despite of the capacity of the above mentioned PCRs to determine the EPIYA pattern of *H. pylori* strains, none of them can distinguish between EPIYA-C and D motifs because designed primers to amplify the C motif are unspecific, and also amplify EPIYA-D. Therefore, we designed two new reverse primers to be used with the forward primer cag2 (table 1), for the specific detection of the EPIYA-C and D motifs. The new specific primer for TPM-C anneals to nucleotide positions 2928 to 2945, in *cagA* gene (HP0547) from the reference strain 26695. The specific primer for EPIYA-D anneals to nucleotide positions 2890 to 2910 of *cagA* gene (AB090073) from the Eastern strain F-13. The generated PCR-products by the new primers for EPIYA-C and D were visualized as bands of 397 and 399 bp on agarose gel electrophoresis, respectively (figure 3, Panels A and B, respectively). In the case of Western strains containing multiple C-segments (figure 3, Panel C), there might be a 102-bp increment over the first C motif (397 bp band), resulting in a new 499-bp amplicon for CC type strains (figure 3, Panel C, line 2) and a third 602-bp amplicon for CCC type strains (figure 3, Panel C, line 3). From the 95 Cuban strains analyzed with primer cagAEpi-C (table 1), 93 carried the EPIYA-C motif and none of the 95 isolates rendered an amplicon when the cagAEpi-D primer was used. In contrast, the four DNAs from Eastern *cagA* isolates were positive with cagAEpi-D primer and negative with cagAEpi-C primer (figure 3). The two Cuban strains that lack the TPM-C were also C negative by the new
specific primer to the EPIYA-C motif (data not shown) and one of them (Hpcnic-74) was confirmed by sequence analysis (figure 2).

**Sequencing analysis of the cagA gene variable region**

Nucleotide sequencing reactions were performed in the 3’ variable region of cagA gene from 14 representative strains using the cag2 and cag4 primers (table 1) to amplify this region. The PCR predictions perfectly matched the EPIYA pattern in the fourteen strains. In figure 2, we showed the variability found in the deduced peptide sequences of 9 studied strains. Three strains showing an EPIYA-C of different sizes in PCR analyses (table 2) were sequenced twice, in order to be sure of the results. The strain Hpcnic-47 had a deletion of 15 aa just after the EPIYA-C motif (figure 2) and the strain Hpcnic-27 had an insertion of 22 aa between the second and third EPIYA repeats. Furthermore, the strain Hpcnic-35 showed almost full deletion of the EPIYA-B sequence and a 21 aa deletion of EPIYA-C motif, coding for an unusually short cagA variable region. These specific mutations in the EPIYA-C fragment had not been reported before in Database, and reflect the high intragenic recombination in the cagA variable region.

The 95 H. pylori strains were isolated from 19 cases of gastric ulcer, 40 from duodenal ulcer and 36 from non-ulcer dyspepsia, and the EPIYA-ABC motifs were detected in 78.9 %, 72.5 % and 72.2 %, respectively (Table 2). Thus, there was not a specific EPIYA motifs associated with peptic ulcer. Furthermore, no correlation was observed between the type and number of the EPIYA motifs and any other gastroduodenal diseases (table 2).

**Discussion**

Characterization of the EPIYA motif types seems to be a better virulence marker than detection of cagA gene prevalence alone to predict the pathogenic potential of H. pylori (1). The prevalence of CagA protein and its corresponding gene in Cuban H. pylori strains has already been detected (22,26). However, the EPIYA motif patterns of Cuban H. pylori isolates has never been investigated before.
PCR methodology is simpler and cheaper than sequencing analysis to determine EPIYA combinations in *H. pylori* strains. There are several set of primers that hybridized into cagA gene around the 3´ variable region and also generates a single band (27). Thus PCR methodologies, the reported previously (27) and the developed with primers cag2 and cag4 (25) have the same usefulness. Also, both PCR strategies could detect the presence of strains with more than one EPIYA pattern in the same patient.

Most of analyzed Cuban *H. pylori* isolates (73.7%) carries a cagA gene with the common EPIYA- ABC pattern (table 2), that resembles the CagA type reported in various Western populations from Europe, Africa and America (17,18,28,29). The percentage of ABC type strains in our case is slightly higher than those above mentioned, suggesting a possible lesser degree of virulence among Cuban CagA positive strains. Further studies, including *H. pylori* strains from representative areas of our country are required to confirm this tendency.

The prevalence of multiple EPIYA-C motif in Cuban strains, was very low. This isolates have been reported to induce *in vitro* higher levels of tyrosine phosphorylation, SHP-2 binding activity and morphological damage to cells. Thus, CagA protein with a greater number of EPIYA-C repeats is considered to be pathophysiologically more virulent and carcinogenic (17,18,29). There are only a few studies in Latin-America that have reported EPIYA patterns of cagA-positive *H. pylori* isolates. Most of the strains from Latin-America also had the EPIYA-ABC pattern reported in this study. On the other hand, differences in the prevalence of strains with multiple EPIYA-C have been reported in Costa Rica 10/33 (30.3 %) (30), Peru 3/26 (11.5 %) (28) and Mexico 6/30 (20 %) (31). While, three different studies from Colombia reported more than 31 % multiple C strains (32-34). It is important to mentioned that Costa Rica and Colombia have higher prevalence of gastric cancer than the other countries.

The newly designed primers (cagAEpi-C and cagAEpi-C) allow determining by PCR the EPIYA status of cagA positive *H. pylori* isolates. The results obtained let all Cuban strains to be classified according to the Western CagA status and provide evidence that is possible to
classify *H. pylori* strains in Western or Eastern CagA types by a simple PCR method, without performing sequencing analysis. However, both primers should be validated using further Eastern and Western *H. pylori* isolates. As far as we know, there are only two similar methods published to discriminate the EPIYA-C and D repeats using PCR (35,36), but the designed primers and methodology here are different from those already published. Our PCR strategy could be used in combination with the previous ones, since it has been reported the usefulness of different sets of primers in the amplification of polymorphic genes in *H. pylori*, especially in studies from different populations (36,37).

The efficacy of PCR methods from this study was confirmed by sequencing analysis; we found a great correlation in the fourteen analyzed isolates between sequencing and PCR results. Three isolates showed variation in the EPIYA-C motif, reinforcing that this region at *cagA* gene is highly polymorphic. Previous studies have reported differences in the response induced in gastric epithelial cells, depending on the number and type of EPIYA motif, being EPIYA-C the most relevant TPM (7). Moreover, it has been reported that heterogeneity in this variable region, such as lack of EPIYA-B motif (31) or changes in the multimerization sequence (10) could explain the diversity in the response induced in gastric epithelial cells by infection with *H. pylori* strains. Taking these observations into account, it would be very interesting to study how the polymorphism found in our isolates influences the biological activity of CagA.

In our study there was no significant difference between the number of EPIYA motifs in *H. pylori* from patients with different disease. These findings are in agreement with several reports that only found association to gastric adenocarcinoma and not to ulcer diseases (17,18). Unfortunately, we could not study a relationship of EPIYA-C repeats to precancerous lesions or GC, since we have no patients having those pathologies. This fact, must be influenced by the low gastric cancer death rate (6.9 / 100 000 in 2009, http://www.sld.cu/servicios/estadisticas/) in Cuban population.
We hypothesize that the absence and relatively low percentage of *H. pylori* isolates carrying the more carcinogenic D or multiple C EPIYA motifs could be the main factor that influences the low rate of GC in our country. The absence of East-Asian CagA type in our study was an expected result, since this type is almost restricted to East Asian patients (38-41). EPIYA-D motif represents the CagA genotype with the highest association to premalignant lesions and GC development in infected persons (15,16). The small percentage of multiple EPIYA-C strains found here also indicates a lower risk for carcinogenicity among our isolates. In fact, studies from Costa Rica and Colombia showed higher percentages of multiple C isolates (30,32-34) as mentioned above, particularly in populations having a gastric cancer death rate over 20 / 100 000 (42); whereas Mexico, with a similar multiple C percentages as Cuba (31), have GC rates below 10 / 100 000 (42). However, two other molecular-epidemiological studies have found that the number of EPIYA-C motif was essentially equal in areas showing divergent incidence of GC. In the first study, strains from Nashville, EUA (low GC zone) and Colombia (high GC zone), no differences in proportion of multiple C motifs was found, indicating that other factors may influence differences in gastric cancer rates in these regions (33). Similar results were reported by Sicinschi *et al.*, who found that the distribution of strains with one, two or three EPIYA-C motifs was not significantly different in two areas from Colombia with different gastric cancer levels (34). Due to the small number of strains characterized in the previous reports and that both reports came from the same research group, new molecular-epidemiological studies are needed to determine whether geographic differences in premalignant lesions and GC levels are influenced by the number of EPIYA-C motifs carried by infecting *H. pylori* strains.

In conclusion, the EPIYA motif pattern of Cuban *H. pylori* isolates is reported for the first time. All Cuban isolates characterized harbor the Western *cagA* gene type, with a predominant pattern of EPIYA-ABC and a relatively low incidence of multiple C motifs among the strains. These results assume that Cuban *H. pylori* circulating strains could have low carcinogenic potential, a finding that requires further studies. The new specific primers
designed in this study are capable of classifying *H. pylori* cagA positive strains into Western or Eastern type. Additionally, mutations not reported before were found in the EPIYA-C motif of the cagA gene.

**Conflicts of interest**

The authors have declared no conflicts of interest.

**Financing**

This work was supported by a grant from the National Center for Scientific Research, Havana, Cuba.

**References**


<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>AT † (°C)</th>
<th>Product Size (bp) §</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>cag2</td>
<td>GGAACCCTAGTCGGAATG</td>
<td>50</td>
<td>-</td>
<td>(23)</td>
</tr>
<tr>
<td>cag4</td>
<td>ATCTTGAGCTTGATC</td>
<td>50</td>
<td>Variable</td>
<td>(23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(400-750)</td>
<td></td>
</tr>
<tr>
<td>cagA-P1C</td>
<td>GTCCTGCTTTTCTTTATTAACCTTTAGC</td>
<td>60</td>
<td>174</td>
<td>(24)</td>
</tr>
<tr>
<td>cagA-P2 CG</td>
<td>TTTAGCAACTTGAGCGTAAATGGG</td>
<td>60</td>
<td>222</td>
<td>(24)</td>
</tr>
<tr>
<td>cagA-P3E</td>
<td>ATCAATTGTAGCGTAAATGGG</td>
<td>60</td>
<td>378</td>
<td>(24)</td>
</tr>
<tr>
<td>cagEpi-C</td>
<td>AAAGGTCCGCGAGATCA</td>
<td>60</td>
<td>397</td>
<td>This study</td>
</tr>
<tr>
<td>cagEpi-D</td>
<td>GCCTGCTTTGATTTGCCTCATC</td>
<td>60</td>
<td>399</td>
<td>This study</td>
</tr>
</tbody>
</table>

† Annealing temperature in centigrades degree.

§ The size corresponds to the cagA sequence for reference strain 26695, which exhibits the EPIYA-ABC pattern, and only the size showed with primer cagEpi-D was taken from the Eastern strain F-13.
Table 2. Prevalence of the EPIYA pattern among 95 Cuban *H. pylori* strains and relationship to clinical outcome.

<table>
<thead>
<tr>
<th>EPIYA type</th>
<th># of strains (%)</th>
<th>Non-ulcer dyspepsia</th>
<th>Duodenal ulcer</th>
<th>Gastric ulcer</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>2 (2.1)</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>AC†</td>
<td>1 (1.1)</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>BC</td>
<td>1 (1.1)</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ABC§</td>
<td>70 (73.6)</td>
<td>26</td>
<td>29</td>
<td>15</td>
</tr>
<tr>
<td>ABCC</td>
<td>19 (20)</td>
<td>7</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>ABCCC</td>
<td>2 (2.1)</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>95</strong></td>
<td><strong>36</strong></td>
<td><strong>40</strong></td>
<td><strong>19</strong></td>
</tr>
</tbody>
</table>

† Strain Hpcnic-35 carrying a deletion of 21 aa in the EPIYA-C motif of *cagA* gene.

§ Two of the ABC strains showed specific mutations, strain Hpcnic-47 (from gastric ulcer patient) had a deletion of 15 aa just after the EPIYA-C motif and Hpcnic-27 (from non ulcer dyspepsia patient) had an insertion of 22 aa after the EPIYA-B motif.
Figure 1. Electrophoretic analysis of PCR products of cagA variable region amplifications from Cuban *H. pylori* isolates. PCR were developed using primers cag2 and cag4. The shown results are the representative band sizes found among studied Cuban *H. pylori* strains. Lane 1: CCUG 17874 (449 bp, AB), Lane 2: J99 (500 bp, BC), Lane 3: Hpcnic-30 (500 and 554 bp, representative of the three BC-ABC detected patterns), Lane 4: Hpcnic-14 (554 bp, ABC), Lane 5: Hpcnic-27 (600 bp, ABC with an insertion of 22 aas), Lane 6: Hpcnic-19 (654 bp, ABCC), Lane 7: Hpcnic-22 (754 bp, ABCCC), M: 100 bp DNA ladder (Promega, USA).
Figure 2. Alignment of deduced amino acid sequences of the 3’ cagA variable region from nine Cuban *H. pylori* strains. These sequences are representative of all the different detected TPM patterns. Normal EPIYA-AB (Hpcnic-74), EPIYA-AC with a deletion of 21 aa in the EPIYA-C motif of cagA gene (Hpcnic-35), Normal EPIYA-BC (Hpcnic-3), normal EPIYA-ABC (Hpcnic-14 and Hpcnic-17), EPIYA-ABC with a deletion of 15 aa just after the EPIYA-C motif (Hpcnic-47), EPIYA-ABC with an insertion of 22 aa between the second and third EPIYA repeats (Hpcnic-27), normal EPIYA-ABCC (Hpcnic-19) and normal EPIYA-ABCCC (Hpcnic-22).

| Hpcnic-74       | NFDI KELNKL-GFNNNNNNGLKN--------EPIYA KVNKKKAGQASLEPIYA QV |
| Hpcnic-35      | NFDI KELNKE-FKNFNNNN-GLKNST-----EPIYA KVNKKKTVQASPE----- |
| Hpcnic-3       | NFDI KELNLKFGFNNNNNNGLKNST--------EPIYTQVQ |
| Hpcnic-47      | NFDI KELNKE-FKNFNNNNNGLKNST-----EPIYA KVNKKKAGQASPEPEPIYA QV |
| Hpcnic-14      | NFDI KELNKE-FKNFNNNN-GLKNST-----EPIYA KVNKKKAGQASPEPEPIYA QV |
| Hpcnic-17      | NFDI KELNKE-FKNFNNNS-GLKNKDDKQPFEPIYA QVKNKKTVQASPEPEPIYA QV |
| Hpcnic-27      | NFDI KELNKE-FKNFNNNN-GLKNKGDGPEPIYA QVKNKKTVQASPEPEPIYA QV |
| Hpcnic-19      | NFDI KELNLKFGFNNNN-GLKNST--------EPIYA KVNKKKAGQASPEPEPIYA QV |
| Hpcnic-22      | NFDI KELNKLGIFNNNNNGLKNST--------EPIYA KVNKKKAGQASPEPEPIYA QV |

********** * ****

| Hpcnic-74       | AKKVKDNMLNQIAS--------------------------GLGVQGQA-------------------------- |
| Hpcnic-35      | AKKVKDNMLNQIAS--------------------------GLGVQGQA-------------------------- |
| Hpcnic-3       | AKKVKDNMLNQIAS--------------------------GLGVQGQA-------------------------- |
| Hpcnic-47      | AKKVTQKIDQLNQAS--------------------------GFPGVQA-GFPLRKHDVDDL |
| Hpcnic-14      | AKKVTQKIDQLNQAS--------------------------GFPGVQA-GFPLRKHDVDDL |
| Hpcnic-17      | AKKVTQKIDQLNQAT--------------------------SGFLGVQQA-GFPLRKHDVDDL |
| Hpcnic-27      | ARKVTQKIDQLNQASGFGVVQAGFPLRKHDVDDLSVGVQQA-GFPLRKHDVDDL |
| Hpcnic-19      | AKKVKDNMLNQAS--------------------------GLGVQGQA-GFPLRKHDVDDL |
| Hpcnic-22      | AKKVKDNMLNQAS--------------------------GLGVQGQA-GFPLRKHDVDDL |

Hpcnic-74       ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-35      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-3       ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-47      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-14      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-17      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-27      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-19      ------------------------------------------EPIYATIDDLGG------------------------------------------ |
Hpcnic-22      ------------------------------------------EPIYATIDDLGG------------------------------------------ |

Hpcnic-74       ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-35      ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-3       ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-47      ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-14      ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-17      ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-27      ------------------------------------------FPLKRHDVDDLSVGVQQAELAQKIDNLD |
Hpcnic-19      ------------------------------------------FPLKRHDVDSLKVGRSVSPEPIYATIDDLGG |
Hpcnic-22      ------------------------------------------FPLKRHDVDSLKVGRSVSPEPIYATIDDLGG |

. **:*:*:* *:*:*:*
**Figure 3.** PCR method used to classify *H. pylori* cagA gene into Western or Eastern type. Panel A: Electrophoretic analysis of PCR products amplified with primers cag2 and the new primer cagAEpi-C (specific to amplify EPIYA-C motif). Lane 1: J99 strain. Lanes 2-5: Cuban cagA Western strains (Hpcnic-8, Hpcnic-10, Hpcnic-14, Hpcnic-17). Lanes 6-9: The four Eastern cagA isolates. Panel B: PCR products obtained with primers cag2 and the new primer cagAEpi-D (specific to amplify EPIYA-D motif). Lanes 1-4: The four Eastern strains. Lanes 5-9: Cuban cagA Western strains. Panel C: PCR products obtained with primers cag2 and the new primer cagAEpi-C. Lane 1: 26695 strain, Lane 2: Hpcnic-19, Lane 3: Hpcnic-22, Lane 4: Eastern cagA isolate (C-negative). M: [Panel A and B, 1 kb DNA ladder (Promega, USA) and Panel C, 100 bp DNA ladder (Roche, Germany)].