Journal of Medical Microbiology (2010), **59**, 89–95 DOI 10.1099/jmm.0.011684-0

Helicobacter pylori oipA, *vacA* and *dupA* genetic diversity in individual hosts

Mario José Matteo,[1] Rita Inés Armitano,[1] Gabriela Granados,[1] Andrés Dario Wonaga,[2] Christian Sánches,[2] Martín Olmos[3] and Mariana Catalano[1]

[1]Departamento de Microbiología, Parasitología e Inmunología, Facultad de Medicina, Universidad de Buenos Aires, Buenos Aires, Argentina

[2]Servicio de Gastroenterología, Hospital Escuela 'Don José de San Martín' Facultad de Medicina, Buenos Aires, Argentina

[3]Servicio de Endoscopía, Hospital General de Agudos Juan A. Fernández, Buenos Aires, Argentina

Correspondence
Mariana Catalano
catalano@fmed.uba.ar

Received 7 April 2009
Accepted 26 July 2009

Helicobacter pylori putative virulence factors can undergo a continuously evolving mechanism as an approach to bacterial adaptation to the host changing environment during chronic infection. *oipA*, *vacA* and *dupA* genetic diversity among isolates from multiple biopsies (niches) from the antrum and corpus of 40 patients was investigated. A set of 229 isolates was examined. Direct DNA sequence analysis of amplified fragments was used to study *oipA* 'on/off' expression status as well as the presence of C or T insertion in *jhp0917* that originates a continuous (*jhp0917*–*jhp0918*) *dupA* gene. *vacA* alleles were identified by multiplex PCR. Different inter-niches *oipA* CT repeat patterns were observed in nine patients; in six of these, 'on' and 'off' mixed patterns were found. In three of these nine patients, different *vacA* alleles were also observed in a single host. Inter-niche *dupA* differences involved the absence and presence of *jhp0917* and/or *jhp0918* or mutations in *dupA*, including those that may originate a non-functional gene, and they were also present in two patients with mixed *oipA* CT patterns and in another seven patients. Evidence of mixed infection was observed in two patients only. In conclusion, *oipA* and *dupA* genes showed similar inter-niche variability, occurring in approximately 1/4 patients. Conversely, *vacA* allele microevolution seemed to be a less common event, occurring in approximately 1/10 patients, probably due to the mechanism that this gene evolves '*in vivo*'.

INTRODUCTION

Helicobacter pylori is characterized by a great genetic diversity evidenced by an apparently unlimited number of unique strains that differ in genome size, gene order, genetic content and allelic profile (Cooke *et al.*, 2005). Since *H. pylori* exhibits high levels of genetic diversity between clinical isolates (Go *et al.*, 1996), the presence of certain genotypes of virulence factors might be considered as an expression of a selective advantage conferred to strains and, therefore, a potential useful tool for predicting the clinical outcome of gastric mucosa colonization. Putative virulence genes are generally classified into three categories: (i) strain-specific genes such as *cag* pathogenicity island (*cag*PAI) genes and genes located in the plasticity region (e.g. *jhp0947* and *dupA* genes); (ii) phase-variable genes (e.g. *oipA*, *sabA*, *sabB*, *babB*, *babC* and *hopZ*); and (iii) genes with variable structures/genotypes (e.g. *vacA*) (Yamaoka, 2008). However, reports on the clinical predictive value of putative virulence factor status and disease outcomes are controversial, at least in all ethnic groups (Chiarini *et al.*, 2008; Dossumbekova *et al.*, 2006; Douraghi *et al.*, 2008; Erzin *et al.*, 2006; Gomes *et al.*, 2008; Höcker & Hohenberger, 2003; Lu *et al.*, 2005a). In addition, according to the microevolution model, different putative virulence factors might undergo a continuously evolving mechanism as an approach to bacterial adaptation to the changing host environment during infection (Marshall *et al.*, 1998; Alvi *et al.*, 2007). It appears that *H. pylori* continuously evolves and spreads within the niches of the stomach as a fluid population of phenotypic variants able to rapidly adapt to changing conditions within the host gastric environment (Carroll *et al.*, 2004; Falush *et al.*, 2001; Israel *et al.*, 2001; Kuipers *et al.*, 2000; Prouzet-Mauléon *et al.*, 2005; Salama *et al.*, 2007).

However, *cag*PAI genotype, *vacA* alleles, *oipA* status and *dupA* presence have been used as stable entities to link the severity of clinical outcomes (Lu *et al.*, 2005b).

Abbreviation: RAPD, randomly amplified polymorphic DNA.

The GenBank/EMBL/DDBJ accession numbers for the new sequence data for *oipA* and *dupA* are FJ789571–FJ789588 and FJ763588–FJ76364, respectively.

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

In the present study, we investigated the microevolution of: (i) *oipA*, representing an example of a contingency locus; (ii) *vacA*, which could be indicative of horizontal gene exchange and associated recombination during colonization; and (iii) *dupA* (*jhp0917–jhp0918*), a *virB* homologue, present in the plasticity region and associated with an increased risk of duodenal ulcer.

METHODS

Isolates. The same set of isolates previously analysed for *cag*PAI genotypes obtained from multiple biopsies from 40 patients belonging to a study approved by a local Ethics Committee was used (Matteo *et al.*, 2007). The study included all patients with positive *H. pylori* cultures from 118 consecutive patients referred to the Division of Gastroenterology of The Clinical University Hospital Don José de San Martín and to the Endoscopy Service of Hospital Juan A. Fernández for gastroscopy on clinical grounds.

From two of these 40 patients, isolates were also obtained 1 year after eradication therapy failure. Twenty-six patients had superficial gastritis, nine had gastric erosion and five had gastric ulceration. Six biopsy specimens were obtained from each patient during the same endoscopy process: a1, from the mid greater curvature of the antrum; a2, from the lesser curvature facing the incisura angularis; a3, from the antral lesser curvature within 2 cm of the pylorus; c1, from the middle portion of the greater curvature of the corpus; c2, from the greater curvature within 3 cm proximal to the antral–corpus border; and c3, from the lesser curvature within 3 cm of the Z line. All biopsy specimens were cultured separately as previously described (Matteo *et al.*, 2007).

DNA extraction and strain delineation. DNA was extracted from confluent cultures with fewer than three '*in vitro*' passages by standard protocols, using SDS, lysozyme and proteinase K, followed by phenol/chloroform extraction and ethanol precipitation. Confluent culture for DNA extraction was obtained using a pool of colonies from the isolation plate of each niche. When results suggested intra-niche diversity, single colonies of bacteria were grown by subculturing from the isolation plate of each biopsy conserved at −80 °C. *lspA–glmM* RFLP and randomly amplified polymorphic DNA (RAPD)-PCR were used for strain delineation in an attempt to identify mixed (multi-strain) infection (Matteo *et al.*, 2007).

***oipA*, *vacA* and *dupA* gene analysis.** *oipA* status was investigated using the primers OIPF (CAAGCGCTTAACAGATAGGC) and OIPR (AAGGCGTTTTCTGCTGAAGC), which amplified at around 430 bp (Yamaoka *et al.*, 2000). PCR products were purified using Wizard-PCR Preps (Promega) according to the manufacturer's instructions. Sequencing was performed on both strands using an ABI 373 DNA sequencer (Applied Biosystems).

vacA alleles s1, s2, m1 and m2 were analysed by multiplex PCR as previously described (Leanza *et al.*, 2004).

The criterion for the presence of the *dupA* gene was the insertion of C or T after position 1385 of *jhp0917* (GenBank accession no. AB196363) resulting in a frameshift leading to a continuous *jhp0917–jhp0918* gene. In order to screen for the presence of *jhp0917* and *jhp0918*, two sets of primers were employed for partial amplification of each gene [JHP0917F (TGGTTTCTACTGACAGAGCGC) and JHP0917R (AACACGCTGACAGGACAATCTCCC); and JHP0918F (CCTATACGCTAACGCGCGCTC) and JHP0918R (AAGCTGAAGCGTTTGTAACG), respectively] (Lu *et al.*, 2005a). When PCRs indicated the absence of *jhp0917* or *jhp0918*, amplification reaction results were confirmed using a combination of JHP0917F and JHP0918R or JHP0917R and JHP0918F primers for the investigation of a possible mismatch with the clinical isolate gene sequences. Looking for a continuous *dupA* gene, the set A and set B primers proposed by Arachchi *et al.* (2007) [DUPF (ATAGCGATAACCAACAAGAT) and DUPR (AAGCTGAAGCGTTTGTAACG) (set A) and DUPF (ATTCACGCCTAAGACCTCA) and DUPR (AAGCTGAAGCGTTTGTAACG) (set B)] were used and the 632 bp PCR products obtained with set A were purified and sequenced on both DNA strands as mentioned above. The nucleotide sequences were analysed by BioEdit version 5.0.9 and BLAST software at the National Center for Biotechnology Information. The amino acid sequences were analysed by GeneDoc version 2.6.002.

RESULTS AND DISCUSSION

A total of 229 isolates were obtained from 40 patients, including those isolates recovered from the six biopsies from the two patients 1 year after eradication therapy failure. Among the 229 isolates, a total of 18 *oipA* CT repeat patterns were identified (Table 1). The number of CT repeats ranged from 6 to 13, and 136/229 isolates (59.4 %) had an in-frame nucleotide pair ('on'). The pattern containing 6 CT repeats was the most frequently associated with the 'on' status (76/136, 55.8 %), and the pattern with 7 CT repeats was the most common genetic basis for a non-functional ('off') *oipA* gene (20/93, 21.5 %). A single specific *oipA* CT repeat pattern was observed among isolates recovered from all antrum and corpus biopsies in 31/40 patients. However, in two of these 31 patients, mixed infection was previously evidenced by *lspA–glmM* RFLP and RAPD DNA patterns with different spatial distributions of strains in the stomach (Matteo *et al.*, 2007). In addition to DNA fingerprinting diversity, these strains also showed different *cag*PAI genotypes. Inter-niche *oipA* CT repeat differences were observed in nine patients (22.5 %) and mixed 'on' and 'off' patterns were observed in six of them (Table 2, patients 24, 58, 60, 230, 400 and 420). The remaining three patients showed mixed patterns in 'on' or in 'off' (Table 2, patients 63, 260 and 520). Moreover, patient number 60 also showed inter-niche *oipA* CT pattern variation since isolates with 7 CT and 8 CT 'off' repeats were identified by growth of single colonies. *lspA–glmM* RFLP and RAPD DNA fingerprinting did not show evidence of mixed infection in these nine patients (Matteo *et al.*, 2007). Therefore, isolates with different *oipA* CT repeat patterns in a single host seemed to be variants or diversification of single strains. Isolates recovered from the six niches of the two patients with eradication therapy failure exhibited the original *lspA–glmM* RFLP and RAPD DNA patterns as well as the original *oipA* CT repeat patterns: 6 CT 'on' and 9 (2 + 7) 'off', respectively.

Three of the nine patients with inter-niche *oipA* CT repeat pattern variation were the only individuals among the 40 patients who also showed inter-niche mixed *vacA* alleles (Table 2, patients 60, 63 and 400). The analysis of multiple single colonies from the six niches from these three patients confirmed these results and showed that isolates within each patient exhibited indistinguishable *lspA–glmM* RFLP

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

Table 1. Frequency of the *oipA* CT repeat pattern

Sequence of signal peptide encoding region of *oipA*	CT no.	No. of isolates
'On' status		
ATGAAAAAAGCCCTCTTACTAA**CTCTCTCTCTCT**CGTTTTGGCTC	6	76
M K K A L L L T L S L S F W L		
ATGAAAAAAGCCCTCTTACTAA**CTCTCTCTCTCTCTCTCT**CGTTCTGGCTC	9	13
M K K A L L L T L S L S L S F W L		
ATGAAAAAAGCTCTCTTACTAA**CTTTCTCTCTCT**TGTTTTGGCTC	1+4	9
M K K A L L L T F S L L F W L		
ATGAAAAAAACCCTTTTA**CTCTTTCTCTCTTTCT**CGTTCTGGCTC	2+3+1	6
M K K T L L L F L S F S F W L		
ATGAAAAAAGCCCTCTTA**CTCTCTCTCTTTCTCT**CGTTCTGGCTC	5+2	19
M K K A L L L S L F L S F W L		
ATGAAAAAAACCCTTTTACTAA**CCCTCTCTCTCT**CGTTCTGGCTC	5 (6−CC)	4
M K K T L L L T L S L S F W L		
ATGAAAAAAGCCCTCTTA**CTCTCTCTCTCTCTCT**CGTTTTGGCTC	8 (2+6)	9
M K K A L L L S L S L S F W L		
'Off' status		
ATGAAAAAAGCTCTCTTACTAA**CTCTCTCTCT**CGTTTTGGCTC	5	15
M K K A L L L T L S L V L A		
ATGAAAAAAGCCCTCTTACTAA**CTCTCTCTCTCTCT**CGTTTTGGCTC	7	20
M K K A L L L T L S L S R F G		
ATGAAAAAAGCCCTCTTACTAA**CTCTCTCTCTCTCTCT**CGTTTTGGCTC	8	1
M K K A L L L T L S L S L V L A		
ATGAAAAAAGCTCTCTTACTAA**CTCTCTCTCTCTCTCTCTCT**CGTTTTGGCTC	10	8
M K K A L L L T L S L S L S R F G		
ATGAAAAAAGCTCTCTTACTAA**CTCTCTCTCTCTCTCTCTCTCT**CGTTCTGGCTC	11	6
M K K A L L L T L S L S L S L V L A		
ATGAAAAAAGCTCTCTTACTAA**CTCTCTCTCTCTCTCTCTCTCTCTCT**CGTTCTGGCT	13	2
M K K A L L L T L S L S L S L S R S G		
ATGAAAAAAACTCTCTTA**CTCTCTCTCTCT**CGTTTTGGCTC	6 (2+4)	6
M K K T L L L S L S R F G		
ATGAAAAAAGCTCTTTTA**CTCTCTCTCTCTCT**CGTTCTGGCTC	7 (2+5)	12
M K K A L L L S L S L V L A		
ATGAAAAAAGCTCTCTTA**CTCTCTCTCTCTCTCTCT**CGTTCTGGCTC	9 (2+7)	16
M K K A L L L S L S L S R S G		
ATGAAAAAAGCTCTCTTA**CTCTCTCTCTCTCTCTCT**CGTTTTGGCTC	10 (2+8)	2
M K K A L L L S L S L S L V L A		
ATGAAAAAAGCCCTCTTACTAA**CTCTCTCTTTCTCT**CGTTTTGGCTC	4+2	6
M K K A L L L T L S F S R F G		

and RAPD DNA patterns. Since *vacA* can be a marker of horizontal gene exchange (Aviles-Jimenez *et al.*, 2004; Marshall *et al.*, 1998; Prouzet-Mauléon *et al.*, 2005), the existence of co-colonization or superinfection with another unidentified strain(s) is the likely explanation for the emerging recombinant variants with different *vacA* alleles (Aviles-Jimenez *et al.*, 2004). In patients without inter-niche *vacA* allele differences, either *vacA*-s1m1 or *vacA*-s2m2 was identified in all isolates from 18 (45 %) and 19 (47.5 %) patients, respectively. The prevalence of the *vacA*-s2m2 allele is higher than that reported in other Latin American populations (Faundez *et al.*, 2002; Mattar *et al.*, 2005). Moreover, in a previous study that included 88 patients, we found a *vacA*-s2m2 prevalence of 22 % (Leanza *et al.*, 2004). However, in this previous study, almost 1/4 patients had ulcer disease while in the present study, only 1/8 patients exhibited ulcer disease.

jhp0917 and *jhp0918* genes were absent in isolates recovered from all niches of 17/40 (42.5 %) patients. Therefore, *dupA* gene PCRs were negative in these isolates. Evidence of mixed infection was suggested in only one of these patients. In another seven patients (17.5 %), isolates from all niches exhibited the presence of *jhp0917*, *jhp0918* and *dupA* genes. However, both *jhp0917* and *jhp0918* or *dupA* positive PCRs did not always suggest a functional gene. In an additional patient with these three positive PCRs, sequencing of *dupA* partially amplified from isolates from the six niches indicated the presence of an earlier stop codon (Fig. 1, 12a1–c3). *jhp0917* and *jhp0918* were present in all the

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

Table 2. Patients with inter-niche differences in *oipA* CT repeat patterns, *vacA* gene alleles, and the presence of *jhp0917*, *jhp0918* or *dupA* genes

Identical: the same genotype, gene status or DNA patterns in all the designated niches. +, Presence of the gene; −, absence of the gene.

Patient no. and stomach niches*	*oipA* CT repeat patterns	*vacA* alleles	*jhp0917*	*jhp0918*	*dupA*	*lspA–glmM* RFLP	RAPD
24; a1, c1, c2, c3	On 6	Identical	Identical	Identical	Identical	Identical	Identical
24; a2, a3	Off 7	Identical	Identical	Identical	Identical	Identical	Identical
45; a1 and a3	Identical	Identical	+	+	+ (truncated)	Identical	Identical
45; a2	Identical	Identical	+	+	+	Identical	Identical
45; c1–c3	Identical	Identical	−	−	−	Identical	Identical
53; a2	Identical	Identical	+	+	+	Identical	Identical
53; c2 and c3	Identical	Identical	−	−	−	Identical	Identical
58; a1	On 6	Identical	−	−	−	Identical	Identical
58; a2, a3	Off 7	Identical	−	−	−	Identical	Identical
58; c1–c3	On 6	Identical	+	+	+	Identical	Identical
60; a1	On 6	s1m1	Identical	Identical	Identical	Identical	Identical
60; c1, c3	Off 7	s2m2	Identical	Identical	Identical	Identical	Identical
60; c2	Off 7/off 8†	s2m2	Identical	Identical	Identical	Identical	Identical
63; a1, a3	On 6	s2m2	Identical	Identical	Identical	Identical	Identical
63; a2, c1	On (1+4)	s2m2	Identical	Identical	Identical	Identical	Identical
63; c2	On 5 (6−CC)	s1m1	Identical	Identical	Identical	Identical	Identical
63; c3	On (5+2)	s2m2	Identical	Identical	Identical	Identical	Identical
69; a1–a3	Identical	Identical	−	−	−	Identical	Identical
80; a1, a2, c1	Identical	Identical	−	−	−	Identical	Identical
80; a3	Identical	Identical	+	+	+	Identical	Identical
120; a1, a3, c2, c3	Identical	Identical	+	+	−‡	Identical	Identical
120; a2 and c1	Identical	Identical	+	−	−	Identical	Identical
230; a1, a3, c1, c2, c3	Off 10	Identical	Identical	Identical	Identical	Identical	Identical
230; a2	On 9	Identical	Identical	Identical	Identical	Identical	Identical
260; a1, a2, c2, c3	Off 9 (2+7)	Identical	Identical	Identical	Identical	Identical	Identical
260; a3, c1	Off 10 (2+8)	Identical	Identical	Identical	Identical	Identical	Identical
350; a1–a3, c1	Identical	Identical	+	+	+	Identical	Identical
350; c2, c3	Identical	Identical	−	−	−	Identical	Identical
400; a1, c1, c2	Off 11	s1m1	Identical	Identical	Identical	Identical	Identical
400; a2	On 9	s2m2	Identical	Identical	Identical	Identical	Identical
400; a3, c3	On 13	s1m1	Identical	Identical	Identical	Identical	Identical
420; a1, a2, a3, c3	On 6	Identical	Identical	Identical	Identical	Identical	Identical
420; c1, c2	Off 7	Identical	Identical	Identical	Identical	Identical	Identical
470; a1–a3	Identical	Identical	−	−	−	Identical	Identical
470; c2	Identical	Identical	+	+	+	Identical	Identical
480; a2	Identical	Identical	+	−	−	Identical	Identical
480; c1 and c2	Identical	Identical	+	+	+ (truncated)	Identical	Identical
520; a1	Off 10 (2+8)	Identical	+	+	+	Identical	Identical
520; c1, c2	Off 10 (2+8)	Identical	+	−	−	Identical	Identical
520; a2, a3, c3	Off 11	Identical	+	−	−	Identical	Identical

*Stomach niches: a1, from the mid greater curvature of the antrum; a2, from the greater curvature facing the incisura angularis; a3, from the antral lesser curvature within 2 cm of the pylorus; c1, from the middle portion of the greater curvature of the corpus; c2, from the greater curvature within 3 cm proximal to the antral–corpus border; and c3, from the lesser curvature within 3 cm of the Z line. *H. pylori*-positive niches are indicated in each patient.

†Inter-niche mixed *oipA* CT repeat patterns. Isolates with 7 CT repeats and isolates with 8 CT repeats in the same niche.

‡The use of a multiple set of primers could allow the identification of the presence of the *dupA* gene.

isolates from one more patient. However, *dupA* seemed to be absent using both sets of specific primers (2.5 %). These results did not ensure real absence of *dupA*, which would be demonstrated if multiple sets of primers were used to avoid probable sequence mismatches. A similar *dupA* PCR result with positive *jhp*0917 and *jhp*0918 amplification was observed in another patient but with isolates from one niche only (c3). In the remaining niches (a1–c2), *jhp*0917

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

```
                     360        370        380        390        400        410        420        430
                  |....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....
AB196363          QDNKAS QSIDKVIRDT YNYMGEQKNQ INLLEIAKNL GSSEQDFNEI LKSQGEKIYF KNFQDCLDFS KSPLSVINMD AFAN
350a1-a3, c1      ...... .......... .....D.... .......... .......... .......... .......... .......... ....
58c1-c3           ...... .......... .......... .......... .......... .......... .......... .........Y ....
90a1-a3, c1, c2   ...... .......... .....D.... .......... .......... .......... .......... .......... ....
29a1-c3           ...... .......... .......... .......... .......... .......... .......... .......... ....
45c1              ...... .......... .....D.... .......... .......... .......V.S .K........ .......... ....
45a1              ...... .......... .....D.... .......... .......... .......V.S .K........ .......... ....
45a2              ...... .......... .....D.... .......... .......... .......V.S .K........ .......... ....
60a1-c3           ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
480c2, c3         ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
58c1-c3           ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
8a1-a3, c2, c3    ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
7a1-c3            ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
390a1, c1, c3     ...... .......... .......... .......... .......... .......V.. .K........ .......... V...
22a1              ...... .......... .......... .......... .......... .......... .......... .......... ....
22a2-c3           ...... .......... .......... .......... .......... .......V.. .K........ .......... ....
71a1-c3           ...... .......... ...I.D.... .......... .......... .......V.. .......... .......... ....
53a2              ...... .......... ...I.D.... .......... .......... .......V.. .......... .......... ....
12a1              ...... .......... .......... .......... .......... .......... .......... .......... ....
12a2              ...... .......... .......... .......... .......... .......... .......... .......... ....
12a3              ...... .......... .......... .......... .......... .......... .......... .......... ....
12c1              ...... .......... .......... .......... .......... .......... .......... .......... ....
12c2              ...... .......... .......... .......... .......... .......... .......... .......... ....
12c3              ...... .......... .......... .......... .......... .......... .......... .......... ....

                     440        450        460        470        480        490        500        510
                  |....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....
AB196363          DKKLMG LIAMYLFHKL FFEAKEHNKP FFLFIDETKD YIMHPIMFAY IANALAQARK INGTLCMAFQ KISQVKELGI DKAK
350a1-a3, c1      ...... .......... .......... .......... .......... .T........ .......... .......... ....
58c1-c3           ...... .......... .......... .......... .......... .......... .......... .......... ....
90a1-a3, c1, c2   ...... .......... .......... .......... .......... .T........ .......... .......... ....
29a1-c3           ...... .......... .......... .......... .......... .......... .......... .......... ....
45c1              ...... .......... .......... .......... ........T. .......... .......... .......... ....
45a1              ...... .......... .......... .......... ........T. .......... .......... .......... ....
45a2              ...... .......... .......... .......... ........T. .......... .......... .......... ....
60a1-c3           ...... .......... .......... .......... ........T. .......... .......... .......... ....
480c2, c3         ...... .......... .......... .......... ........T. .......... .......... .......... ....
58c1-c3           ...... .......... .......... .......... ........T. .......... .......... .......... ....
8a1-a3, c2, c3    ...... .......... .......... .......... ........T. .......... .......... .......... ....
7a1-c3            ...... .......... .......... .......... ........T. .......... .......... .......... ....
390a1, c1, c3     ...... .......... .......... .......... ........T. .......... .......... .......... ....
22a1              ...... .......... .......... .......... .......... .......... .......... .........H ....
22a2-c3           ...... .......... .......... .......... ........T. .......... .......... .......... ....
71a1-c3           ...... .......... .......... .......... .......... .......... .......... .......... ....
53a2              ...... .......... .......... .......... .......... .......... .......... .......... ....
12a1              ..N.W. *
12a2              ..N.W. *
12a3              ..N.W. *
12c1              ..N.W. *
12c2              ..N.W. *
12c3              ..N.W. *

                     520        530        540        550        560        570        580
                  |....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
AB196363          SLIGNL SQVIIYPTKD TDELIECGVP LSDSEINFLH NTDMRVRQVL VKNIVTNASA FIEIDLKKIC
350a1-a3, c1      ...... .......... .......... .......... .....A.... ....------ ----------
58c1-c3           ...... .......... .......... .......... .....A.... ....------ ----------
90a1-a3, c1, c2   ...... .......... ....M..... .......... .....A.... ....------ ----------
29a1-c3           ...... .......... .......... .......... .....A.... ....------ ----------
45c1              ...... .......... .......... ....*
45a1              ...... .......... .......... ....*
45a2              ...... .......... .......... .......... .....A.... ....------ ----------
60a1-c3           ...... .......... .......... .......... .....A.... ....------ ----------
480c2, c3         ...... .......... .......... ....*
58c1-c3           ...... .......... .......... .......... .....A.... ....------ ----------
8a1-a3, c2, c3    ...... .......... .......... .......... .....A.... ....------ ----------
7a1-c3            ...... .......... .......... .......... .....A.... ....------ ----------
390a1, c1, c3     ...... .......... .......... .......... .....A.... ....------ ----------
22a1              ...... .......... .......... .......... .....A.... ....------ ----------
22a2-c3           ...... .......... .......... .......... .....A.... ....------ ----------
71a1-c3           ...... .......... .......... .......... .....A.... ....------ ----------
53a2              ...... .......... .......... .......... .....A.... ....------ ----------
```

Fig. 1. Amino acid sequence of 632 bp of the *dupA* gene. Patient numbers and niches from antrum (a) and corpus (c) from where isolates were recovered are indicated. GenBank accession numbers are FJ763588–FJ76364.

and *jhp*0918 and *dupA* PCRs were positive. However, in this patient, two different strains were evident by *lspA–glmM* RFLP and RAPD DNA fingerprinting: one located in a1–c2 and the other one in c3.

In three different patients (7.5 %), isolates from all niches showed the presence of *jhp0917* only, and in an additional patient (2.5 %), all isolates exhibited the presence of *jhp0918* only.

Intra-niche differences in the presence or absence of *jhp0917*, *jhp0918* or *dupA* were evident in nine patients (22.5 %) (Table 2). The most remarkable results suggesting *dupA* gene microevolution were those observed in patients 45 and 480 (Table 2), where in isolates from different niches of the single host there was not only loss or acquisition of *jhp0917* and *jhp*0918 but also mutations in *dupA* that originated a non-functional gene. Moreover, in

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

one of the two patients with eradication therapy failure, the presence of *jhp0917*, *jhp0918* and *dupA* was different in the isolates recovered pre- and post-therapy, even though all isolates seemed to belong to the same strain. In the biopsies obtained before therapy, isolates showed the presence of *jhp0917* and/or *jhp0918* (Table 2, patient 120). However, in those obtained after therapy from the same stomach sites, absence of *jhp0917* and *jhp0918* in all niches was observed and the analysis of single colonies confirmed this result.

A total of 65 *dupA* PCR products obtained with the set A primers were sequenced and all exhibited the C or T insertion after position 1385 (65 % C and 35 % T). Fig. 1 shows the translated proteins observed. Mutations that gave amino acid substitutions were observed in several patients and these substitutions seemed to characterize isolates from each host. Isolates from patient 45 were the only ones with serine at position 410, suggesting the same origin in these three isolates. However, two of the three *dupA* gene positive isolates showed an early stop codon at position 545 of the protein. This stop codon seems to originate by the substitution of G/T at position 1633 (GenBank accession no. AB196363) absent in the remaining isolate with the non-truncated *dupA* gene (Fig. 1). The same early stop codon was also observed in the six isolates from patient 60, while the remaining amino acid sequences differentiated isolates from both patients (Fig. 1). The exception was patient 22, where isolates from niches a2–c3 showed different amino acid substitutions from those from a1 (Fig. 1). All isolates from patient 12 showed a deletion at position 1311 of the *dupA* DNA sequence that originated a frameshift mutation and a truncated gene; however, they still have a T insertion after position 1385. Adenine insertion after position 1426 of the *dupA* DNA sequence, which creates a premature stop codon (GenBank accession no. DQ834307), frequently found in isolates from Brazil (Gomes *et al.*, 2008), was not identified in our isolates.

Prior to the investigation of the relevance between the status of a putative virulence determinant and the colonization outcomes, it will be important to ascertain whether microevolution of this virulence factor is a common phenomenon. In this study, microevolution was observed for all three putative virulence determinants, but with different frequencies. *oipA* and *dupA* genes showed similar inter-niche variability: around 1/4 patients. Inter-niche *oipA* variability was also reported by Kim *et al.* (2009). These authors observed different *oipA* genotyping statuses in a single host with a slightly higher frequency: 1/3 patients. Single host variability of *oipA* and *dupA* status can be expected since the former is a phase-variable gene and the latter is located in the plasticity region where gene content is highly variable. In this study, in all patients except one, inter-niche differences in both genes seemed to represent the diversification of a single strain.

In contrast to *oipA* and *dupA* genes, *vacA* inter-niche differences seem to be less common events, occurring in around 1/10 patients, probably due to the mechanism this gene evolves '*in vivo*' and because mixed infections could not be evidenced at a higher rate in the study population. This point is illustrated by the Kim *et al.* (2009) study, which showed inter-niche *vacA*-m and *vacA*-s allele differences in around 1/8 and 3/8 patients, respectively, but in a population with 67.5 % of infection due to multiple strains.

ACKNOWLEDGEMENTS

This study was supported by grants BID 1728 OC-AR PICT 2006-0458 provided by Agencia Nacional de Promoción Científica y Tecnológica and by M007 from Universidad de Buenos Aires (UBACyT 2008-2010).

REFERENCES

Alvi, A., Devi, S. M., Ahmed, I., Hussain, M. A., Rizwan, M., Lamouliatte, H., Mégraud, F. & Ahmed, N. (2007). Microevolution of *Helicobacter pylori* type IV secretion systems in an ulcer disease patient over a ten-year period. *J Clin Microbiol* **45**, 4039–4043.

Arachchi, H. S., Kalra, V., Lal, B., Bhatia, V., Baba, C. S., Chakravarthy, S., Rohatgi, S., Sarma, P., Mishra, M. V. & other authors (2007). Prevalence of duodenal ulcer-promoting gene (*dupA*) of *Helicobacter pylori* in patients with duodenal ulcer in North Indian population. *Helicobacter* **12**, 591–597.

Aviles-Jimenez, F., Letley, D. P., Gonzalez-Valencia, G., Salama, N., Torres, J. & Atherton, J. C. (2004). Evolution of the *Helicobacter pylori* vacuolating cytotoxin in a human stomach. *J Bacteriol* **186**, 5182–5185.

Carroll, I. M., Ahmed, N., Beesley, S. M., Khan, A. A., Ghousunnissa, S., Moráin, C. A., Habibullah, C. M. & Smyth, C. J. (2004). Microevolution between paired antral and paired antrum and corpus *Helicobacter pylori* isolates recovered from individual patients. *J Med Microbiol* **53**, 669–677.

Chiarini, A., Calà, C., Bonura, C., Gullo, A., Giuliana, G., Peralta, S., D'Arpa, F. & Giammanco, A. (2008). Prevalence of virulence-associated genotypes of *Helicobacter pylori* and correlation with severity of gastric pathology in patients from western Sicily, Italy. *Eur J Clin Microbiol Infect Dis* **28**, 437–446.

Cooke, C. L., Huff, J. L. & Solnick, J. V. (2005). The role of genome diversity and immune evasion in persistent infection with *Helicobacter pylori*. *FEMS Immunol Med Microbiol* **45**, 11–23.

Dossumbekova, A., Prinz, C., Mages, J., Lang, R., Kusters, J. G., Van Vliet, A. H., Reindl, W., Backert, S., Saur, D. & other authors (2006). *Helicobacter pylori* HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of *hopH* gene polymorphisms. *J Infect Dis* **194**, 1346–1355.

Douraghi, M., Mohammadi, M., Oghalaie, A., Abdirad, A., Mohagheghi, M. A., Hosseini, M. E., Zeraati, H., Ghasemi, A., Esmaieli, M. & Mohajerani, N. (2008). *dupA* as a risk determinant in *Helicobacter pylori* infection. *J Med Microbiol* **57**, 554–562.

Erzin, Y., Koksal, V., Altun, S., Dobrucali, A., Aslan, M., Erdamar, S., Dirican, A. & Kocazeybek, B. (2006). Prevalence of *Helicobacter pylori vacA*, *cagA*, *cagE*, *iceA*, *babA2* genotypes and correlation with clinical outcome in Turkish patients with dyspepsia. *Helicobacter* **11**, 574–580.

Falush, D., Kraft, C., Taylor, N. S., Correa, P., Fox, J. G., Achtman, M. & Suerbaum, S. (2001). Recombination and mutation during long-

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06

term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. *Proc Natl Acad Sci U S A* **98**, 15056–15061.

Faundez, G., Troncoso, M. & Figueroa, G. (2002). *cagA* and *vacA* in strains of *Helicobacter pylori* from ulcer and non-ulcerative dyspepsia patients. *BMC Gastroenterol* **2**, 20–24.

Go, M. F., Kapur, V., Graham, D. Y. & Musser, J. M. (1996). Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. *J Bacteriol* **178**, 3934–3938.

Gomes, L. I., Rocha, G. A., Rocha, A. M., Soares, T. F., Oliveira, C. A., Bittencourt, P. F. & Queiroz, D. M. (2008). Lack of association between *Helicobacter pylori* infection with *dup*A-positive strains and gastroduodenal diseases in Brazilian patients. *Int J Med Microbiol* **298**, 223–230.

Höcker, M. & Hohenberger, P. (2003). *Helicobacter pylori* virulence factors – one part of a big picture. *Lancet* **362**, 1231–1233.

Israel, D. A., Salama, N. R., Krishna, U., Rieger, M., Atherton, J. C., Falkow, S. & Peek, R. M., Jr (2001). *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. *Proc Natl Acad Sci U S A* **98**, 14625–14630.

Kim, Y. S., Kim, N., Kim, J. M., Kim, M. S., Park, J. H., Lee, M. K., Lee, D. H., Kim, J. S., Jung, H. C. & Song, I. S. (2009). *Helicobacter pylori* genotyping findings from multiple cultured isolates and mucosal biopsy specimens: strain diversities of *Helicobacter pylori* isolates in individual hosts. *Eur J Gastroenterol Hepatol* **21**, 522–528.

Kuipers, E. J., Israel, D. A., Kusters, J. G., Gerrits, M. M., Weel, J., van der Ende, A., van der Hulst, R. W., Wirth, H. P., Hook-Nikanne, J. & other authors (2000). Quasispecies development of *Helicobacter pylori* observed in paired isolates obtained years apart from the same host. *J Infect Dis* **181**, 273–282.

Leanza, A. G., Matteo, M. J., Crespo, O., Antelo, P., Olmos, J. & Catalano, M. (2004). Genetic characterisation of *Helicobacter pylori* isolates from an Argentinean adult population based on cag pathogenicity island right-end motifs, *lspA-glmM* polymorphism and *iceA* and *vacA* genotypes. *Clin Microbiol Infect* **10**, 811–819.

Lu, H., Graham, P. I., Yang, D. & Yamaoka, Y. (2005a). Duodenal ulcer promoting gene of *Helicobacter pylori*. *Gastroenterology* **128**, 833–848.

Lu, H., Yamaoka, Y. & Graham, D. Y. (2005b). *Helicobacter pylori* virulence factors: facts and fantasies. *Curr Opin Gastroenterol* **21**, 653–659.

Marshall, D. G., Dundon, W. S., Beesley, S. M. & Smyth, C. J. (1998). *Helicobacter pylori* – a conundrum of genetic diversity. *Microbiology* **144**, 2925–2939.

Mattar, R., dos Santos, A. F., Eisig, J. N., Rodrigues, T. N., Silva, F. M., Lupinacci, R. M., Iriya, K. & Carrilho, F. J. (2005). No correlation of *bab*A2 with *vac*A and *cag*A genotypes of *Helicobacter pylori* and grading of gastritis from peptic ulcer disease patients in Brazil. *Helicobacter* **10**, 601–608.

Matteo, M. J., Granados, G., Pérez, C. V., Olmos, M., Sanchez, C. & Catalano, M. (2007). *Helicobacter pylori* cag pathogenicity island genotype diversity within the gastric niche of a single host. *J Med Microbiol* **56**, 664–669.

Prouzet-Mauléon, V., Hussain, M. A., Lamouliatte, H., Kauser, F., Megraud, F. & Ahmed, N. (2005). Pathogen evolution in vivo: genome dynamics of two isolates obtained 9 years apart from a duodenal ulcer patient infected with a single *Helicobacter pylori* strain. *J Clin Microbiol* **43**, 4237–4241.

Salama, N. R., Gonzalez-Valencia, G., Deatherage, G. B., Aviles-Jimenez, F., Atherton, J. C., Graham, D. Y. & Torres, J. (2007). Genetic analysis of *Helicobacter pylori* strain populations colonizing the stomach at different times postinfection. *J Bacteriol* **189**, 3834–3845.

Yamaoka, Y. (2008). Roles of the plasticity regions of *Helicobacter pylori* in gastroduodenal pathogenesis. *J Med Microbiol* **57**, 545–553.

Yamaoka, Y., Kwon, D. H. & Graham, D. Y. (2000). A M_r 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. *Proc Natl Acad Sci U S A* **97**, 7533–7538.

Downloaded from www.microbiologyresearch.org by
IP: 93.91.26.109
On: Sat, 28 Nov 2015 13:02:06