JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2010, p. 2885–2891 Vol. 48, No. 8
0095-1137/10/$12.00 doi:10.1128/JCM.00401-10

Clinical Relevance and Diversity of Two Homologous Genes Encoding Glycosyltransferases in *Helicobacter pylori*[▽]

Mónica Oleastro,[1,2] Andrea Santos,[1] Rita Cordeiro,[1] Baltazar Nunes,[3] Francis Mégraud,[2,4]* and Armelle Ménard[2,4]

Departamento de Doenças Infecciosas, Instituto Nacional Saúde Dr. Ricardo Jorge, Lisbon, Portugal[1]; INSERM U853, 33076 Bordeaux, France[2]; Departamento de Epidemiologia, Instituto Nacional Saúde Dr. Ricardo Jorge, Lisbon, Portugal[3]; and Université Victor Segalen Bordeaux 2, Laboratoire de Bactériologie, 33076 Bordeaux, France[4]

Received 26 February 2010/Returned for modification 15 April 2010/Accepted 8 June 2010

Helicobacter pylori **is known to be a major cause of peptic ulceration. The *jhp0562* gene, encoding a glycosyltransferase involved in the synthesis of the lipopolysaccharide, was associated with peptic ulcer disease (PUD) in children. The β-(1,3)-galactosyltransferase [β-(1,3)GalT] gene (*jhp0563*), involved in Lewis (Le) antigen expression, is highly similar to *jhp0562*. The clinical significance and diversity of both genes were examined by PCR and sequencing of clinical strains ($n = 117$) isolated from children with PUD ($n = 57$) and nonulcer dyspepsia (NUD; $n = 60$). The prevalence of the *jhp0562* gene was significantly higher in strains with a more-virulent profile (strains positive for the *cag* pathogenicity island [PAI], *vacA* sl allele, *babA*, *homB*, phase-variable gene *oipA* "on" [i.e., functional], and *hopQ* I allele). The distribution of genotypes according to clinical outcome showed that the presence of *jhp0562* represented one of the greatest risks for the development of PUD. Moreover, the triple-positive genotype for the *cag* PAI, *jhp0562*, and *homB* provided the best discriminatory model for distinguishing PUD and NUD outcomes in children. Sequence and *in vitro* expression analyses of *jhp0562* showed the presence of a complete open reading frame, while the β-(1,3)GalT gene was shown to be a phase-variable gene. The regular presence of *jhp0562* in strains with a truncated β-(1,3)GalT gene suggests that *jhp0562* may also be implicated in the regulation of Le antigen expression. Overall, the results of this study suggest that the *jhp0562* gene is of great clinical relevance, being a useful comarker for severe *H. pylori*-related disease and contributing to host adaptation.**

Helicobacter pylori colonizes the human gastric mucus layer and is the major causative agent of chronic active gastritis, peptic ulcer disease (PUD), gastric carcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma (8, 16, 19, 28). Progression to diseases of greater severity occurs only in certain infected individuals and seems to depend on a number of factors, including host genetic susceptibility, environmental factors, and differences in *H. pylori* strain virulence.

Comparison of *H. pylori* strains isolated from patients with symptomatic gastric disease, such as gastric cancer or PUD, with strains isolated from patients with no such history is useful in identifying bacterial genetic markers associated with gastric disease. Recently, the *jhp0562* gene was isolated from a duodenal ulcer (DU)-associated strain, by means of PCR-based subtractive hybridization, in which the genomic content of two *H. pylori* strains isolated from two young children, one presenting with DU and the other with recurrent abdominal pain and gastritis, was analyzed (27). The *jhp0562* gene was found to be highly associated with PUD in children (15 with PUD and 30 with nonulcer dyspepsia [NUD] with gastritis) (27). This gene encodes a glycosyltransferase involved in the synthesis of the chemical structure of the lipopolysaccharide (LPS). It is present in the genome of the J99 strain but absent in the 26695 strain and located immediately upstream of the β-(1,3)-galactosyltransferase [β-(1,3)GalT] gene (*jhp0563* or *HP0619* according to the annotation of these strains), which in turn encodes a β-(1,3)GalT involved in type I Lewis (Le) antigen (Le^a and Le^b) synthesis of the LPS (3, 31). *jhp0562* and the β-(1,3)GalT gene are highly similar (>80%), especially at their 5′ and 3′ ends (2, 9).

The present study aimed to evaluate (i) the clinical relevance of the *jhp0562* gene by examining its association with *H. pylori* virulence-associated genes and with clinical outcome as well as (ii) the *jhp0562* gene polymorphisms among clinical isolates, including *in vitro* gene expression. Due to the high-level similarity between *jhp0562* and the β-(1,3)GalT gene, the genetic polymorphism of the latter was also studied. *H. pylori* strains from a pediatric population were chosen, since children are by definition recently colonized, therefore constituting a privileged natural model with less interference of environmental and age-related factors.

MATERIALS AND METHODS

Patients and *H. pylori* bacterial strains. A total of 117 *H. pylori* strains isolated from children attending the three pediatric gastroenterology units in the Lisbon area (Portugal) for upper gastrointestinal symptoms were examined in this study. These strains belong to the collection of *H. pylori* strains in the Department of Infectious Diseases of the Instituto Nacional Saúde Dr. Ricardo Jorge, in Lisbon, Portugal. Fifty-seven children presented with PUD (53 with DU and 4 with gastric ulcers), and 60 presented with abdominal pain and gastritis only, categorized as NUD. The characteristics of the study population are described in Table 1.

* Corresponding author. Mailing address: INSERM U853, Université *Victor Segalen* Bordeaux 2, Laboratoire de Bactériologie, Bat. 2B RDC Zone Nord, 33076 Bordeaux Cedex, France. Phone: 33 (0)5 56 79 59 10. Fax: 33 (0)5 56 79 60 18. E-mail: francis.megraud@chu-bordeaux.fr.

▽ Published ahead of print on 16 June 2010.

TABLE 1. Distribution of the 117 *H. pylori* strains in the study by clinical outcome, peptic ulcer disease or nonulcer dyspepsia, patient gender, and patient age

Parameter	Value by clinical outcome		*P* value
	PUD	NUD	
No. of strains	57[a]	60	
% of male patients	55.0	68.4	0.183[b]
Median age ± SD (yr)	12.02 ± 3.6	9.27 ± 3.15	0.001[c]

[a] Of the 57 cases of peptic ulcer disease, 53 were duodenal ulcers and the remaining 4 were gastric ulcers.
[b] Determined by the χ^2 test.
[c] Determined by the Mann-Whitney test.

Two *H. pylori* strains with sequenced genomes, 26695 (ATCC 700392), negative for *jhp0562* and positive for the β-(1,3)GalT gene, and J99 (ATCC 700824), positive for both genes (Table 2), were used as control strains (2, 35).

All of the strains were cultured from antral biopsy specimens as previously reported (20). In the case of possible mixed infections, all of the colonies were recovered from the plate and used for the study. Genomic DNA was extracted from a 48-h culture by using the QIAamp DNA minikit (Qiagen GmbH, Hilden, Germany) and stored at −20°C until required for analysis.

Genotyping of *jhp0562* and *H. pylori* virulence factors. The *jhp0562* gene was detected by PCR using a set of primers previously described (27), F1-*jhp0562*/*jhp0563* (5′-TGAAAAGCCCTTTTGATTTTG-3′), and R1-*jhp0562*/*jhp0563* (5′-GCTGTAGTGGCCACATACACG-3′). This PCR generates two PCR products with 301 and 602 bp in the J99 strain, corresponding to the *jhp0562* and β-(1,3)GalT (*jhp0563*) genes, respectively, and only one PCR product with 558 bp in the 26695 strain, corresponding to the β-(1,3)GalT gene (HP0619) (27). PCR fragments from clinical strains presenting different sizes were identified by sequencing.

The characterization of the two major *H. pylori* virulence genes, the *cag* pathogenicity island (PAI) (for which both *cagA* and *cag* empty site genotypes were determined) and the *vacA* s region, was performed by PCR (1, 5, 36). The status of *H. pylori babA*, *homB*, *oipA*, *hopQ*, *sabA*, and *hopZ* outer membrane protein-encoding genes was determined by PCR and sequencing in some cases as previously described (6, 12, 17, 27, 29, 40).

Diversity in *jhp0562* and β-(1,3)GalT genes. To study the diversity at the *jhp0562*–β-(1,3)GalT gene locus, as well as polymorphisms in the sequences of these genes, PCR primers designed in the conserved genomic flanking open reading frame (ORF) (*jhp0561* and *jhp0564*, according to J99 strain annotation) were used for the amplification of this locus (F1-*jhp0561*, 5′-CATAACCGCATGCGGGTGTTT-3′; R1-*jhp0564*, 5′-CGCACAAAGAATCAGCGTCAT-3′). Sequencing of the PCR products was achieved on both strands by using the initial set of PCR primers and highly specific internal primers F2-*jhp0562* (5′-CTACCCATAAAGTCATCCACTA-3′), R2-*jhp0563* (5′-AGGCAAACATCAAAGAGAC-3′), and R3-*jhp0563* (5′-TAACTTCATGATTTTCAATAGG-3′).

Reverse transcription. RNA was extracted from a 48-h culture by using the High Pure RNA isolation kit (Roche Diagnostics GmbH, Mannheim, Germany) according to the manufacturer's recommendations. Also, as recommended by the manufacturer, a DNase I digestion step was performed to eliminate residual genomic DNA. Reverse transcription reactions were carried out with 500 ng of RNA by using the Transcriptor First Strand cDNA synthesis kit (Roche Diagnostics GmbH) according to the manufacturer's instructions. The cDNA obtained was then used as the template in PCRs using the F1-*jhp0562*/*jhp0563*–R1-*jhp0562*/*jhp0563* set of primers. In order to check the quality and integrity of the RNA, a PCR targeting the 23S rRNA gene was performed in parallel (21). The absence of DNA contamination in RNA preparations was controlled by performing PCR without the reverse transcription step.

Statistical analysis. Statistical analysis was performed with the statistical package SPSS for Windows 17.0 (2008). Association between the genotypes was first evaluated by the χ^2 and Fisher's exact tests. Then, a principal-component analysis for categorical data was performed in order to measure the global association of the genotypes studied. To confirm the results of the former analysis, a cluster analysis was used to draw a dendrogram based on the same variables by using the average aggregation method and Euclidian distance.

To identify possible associations between the *H. pylori* genotypes and clinical outcome, a univariate analysis, using Fisher's exact test, was performed. The results were confirmed by unconditional binary logistic regression analysis. Lastly, in order to evaluate which combinations of genotypes presented the most discriminatory capacity to distinguish the clinical outcome, a binary logistic

TABLE 2. Distribution of the *jhp0562* and β-(1,3)GalT genes among the sequenced *Helicobacter pylori* strains

Gene characteristic	Value for *Helicobacter pylori* strain (GenBank accession no.)										
	J99 (AE001439)	26695 (AE000511)	HPAG1 (CP000241)	Shi470 (CP001072)	B38 (FM991728)	G27 (CP001173)	P12 (CP001217)	51 (CP000012)	52 (CP001680)	B128 (NZ_ABSY01000005)
jhp0562-like gene Name	*jhp0562*		*HPAG1_0600*	*HPSH_03775*				*KHP_0701*	*HPKB_0725*	*HPB128_187g39*
% identity with *jhp0562*[a] (nt overlap)	100	NA[f]	94.9 (999)	90.9 (999)	NA	NA	NA	92.1 (999)	91.2 (999)	94.1 (999)
jhp0563-like gene Name	*jhp0563*[d]	*HP0619*	*HPAG1_0601-HPAG1_0602*	*HPSH_03770-HPSH_03765*	*HELPY_0753-HELPY_0752*	*HPG27_579-HPG27_580*	*HPP12_0628*	*KHP_0700*	*HPKB_0724-HPKB_0723*	*HPB128_187g38-HPB128_187g37*
% identity with *jhp0563*[a] (nt overlap)	100	78.5 (1,352)	91.3 (1,339)	87.4 (1,356)	83.8 (1,328)	89.5 (1,300)	86.0 (1,290)	86.3 (1,320)	88.8 (552)	90.0 (1,354)
No. of tandem-repeat sequences[b] (21 nt)	2[e]	1	6	6	4	4	3	7	3	2[e]
Status[c]	Off	Off	Off	Off	Off	Off	On	On	Off	Off

[a] The percent identity between the nucleotide sequences was determined using the LALIGN software (15), which identifies multiple matching subsegments in two sequences (http://www.ch.embnet.org/software/LALIGN_form.html). The *jhp0562* and *jhp0563* sequences from the J99 strain (NC_000921) were used as the queries.
[b] Tandem-repeat sequences were determined according to the method of Salaun et al. (32, 33).
[c] *jhp0563* is a phase-variable gene, in contrast to *jhp0562*. The gene status always refers to the entire sequence, even in cases for which the *jhp563*-like ORF is fragmented into two ORFs. "On" indicates a functional status, and "off" indicates a nonfunctional status. nt, nucleotides.
[d] In the J99 strain, the *jhp0563* annotation was revised and is available on the Magnifying Genomes (MaGe) website for the Microbial Genome Annotation System (https://www.genoscope.cns.fr/agc/mage/wwwpkgdb/Login/log.php?pid=19). The complete *jhp0563* ORF ranges from 625,240 to 626,575 bp (AE001439).
[e] The J99 and B128 strains presented five and four repeats of 21-nt-length sequences, respectively. Only two of these repetitions were directly adjacent to each other for the B128 strain, thus truly corresponding to tandem-repeat sequences. In addition, J99 presented a two-tandem-21-nt-repeat sequence, which was repeated twice.
[f] NA, not applicable.

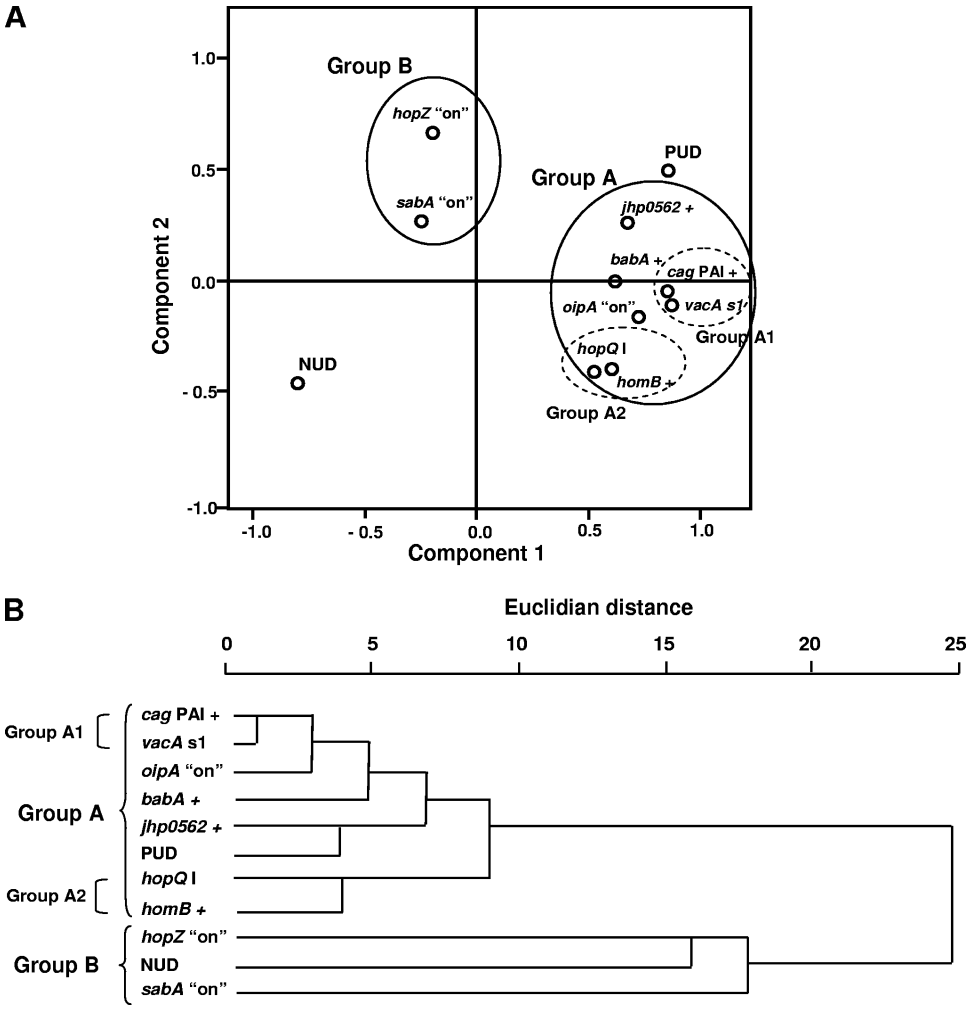

FIG. 1. Analysis of the association of *Helicobacter pylori* virulence factors between themselves and according to clinical outcome. Complete data (for the *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA* and *sabA* status, and *hopQ* and *hopZ* alleles) were available for 83 *H. pylori* strains isolated from children with peptic ulcer disease ($n = 41$) and nonulcer dyspepsia ($n = 42$). (A) Principal-component analysis plot. (B) Dendrogram constructed using the average aggregation method and Euclidian distance.

regression analysis was performed, using the backward stepwise method based on the likelihood ratio test. All of the tests performed were two-tailed, and the level of significance was always set at 5%.

Nucleotide sequence accession numbers. Full sequences of the *jhp0562* and β-(1,3)GalT genes from the 34 *H. pylori* clinical strains are available at GenBank with the accession numbers EF648399, EF648401 to EF648403, and EF648416 to EF648445.

RESULTS

Association of *jhp0562* with other virulence genes. The PCR using the F1-*jhp0562*/*jhp0563*–R1-*jhp0562*/*jhp0563* primer set yielded PCR amplicons for all of the 117 *H. pylori* clinical strains tested. Thus, the relationships between the *jhp0562* genes and the other virulence markers were assessed for all of the strains. The *jhp0562* gene was strongly associated with positive genotypes for the *cag* PAI ($P < 0.001$), *vacA* s1 allele ($P < 0.001$), *babA* ($P < 0.001$), *homB* ($P = 0.007$), phase-variable gene *oipA* "on" (i.e., functional) ($P < 0.001$), and the *hopQ* I allele ($P = 0.014$).

Global analysis of the *H. pylori* virulence factors. In order to evaluate how the different genotypes were associated, a principal-component analysis for categorical data was performed, including all of the genotypes and also the clinical outcome (Fig. 1A). The analysis was performed on 83 strains for which complete data were available for all of the variables tested. Indeed, 34 strains presented a multiple genotype regarding *hopQ* alleles (presence of both *hopQ* I and II alleles), reflecting results from other studies (6, 7), and were excluded. The first two components accounted for 55% of the total information (43% for the horizontal axis [component 1] and 12% for the vertical axis [component 2]). Two main clusters were identified on each side of the horizontal axis (Fig. 1). The first included strains positive for *jhp0562*, the *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA* "on," and the *hopQ* I allele (group A). The second cluster included strains positive for *sabA* and *hopZ* "on" status (group B), although these two genotypes were more weakly associated than the genotypes in group A. Two subclusters, one including *cag* PAI and *vacA* s1 allele (A1) and the other com-

TABLE 3. Statistical analysis of the association between virulence-associated genes and clinical outcome in 117 *H. pylori* strains

Genotype	No. (%) of strains associated with:		*P* value[b]	OR (95% CI)		
	PUD (n = 57)	NUD (n = 60)		Crude	Adjusted[c]	Adjusted for final model[d]
jhp0562	48 (84.2)	21 (35.0)	<0.001	9.14 (3.74–22.32)	6.55 (1.54–27.89)	5.82 (1.64–20.63)
cag PAI	43 (75.4)	10 (16.7)	<0.001	15.36 (6.19–38.08)	6.68 (1.27–35.28)	5.58 (1.64–19.04)
vacA s1	40 (70.2)	11 (18.3)	<0.001	11.14 (4.65–26.69)	0.94 (0.12–7.48)	NI
babA	27 (47.4)	7 (11.7)	<0.001	6.81 (2.65–17.52)	3.51 (0.63–19.71)	NI
homB	41 (71.9)	21 (35.0)	<0.001	4.52 (2.10–9.93)	4.68 (0.96–22.77)	3.43 (1.02–11.54)
oipA "on"	40 (70.2)	18 (30.0)	<0.001	5.69 (2.55–12.69)	0.44 (0.07–2.75)	NI
hopQ I[a]	29 (70.7)	17 (40.5)	0.008	3.55 (1.43–8.85)	0.88 (0.16–4.96)	NI
sabA "on"	25 (43.9)	34 (56.7)	0.197	0.60 (0.29–1.24)	0.79 (0.21–2.92)	NI
hopZ "on"	36 (63.2)	39 (65.0)	0.85	0.92 (0.43–1.97)	2.3 (0.49–11.07)	NI

[a] n = 83; strains with multiple *hopQ* genotypes were excluded.
[b] *P* values were determined using Fisher's exact test.
[c] Adjusted by logistic regression.
[d] Adjusted by binary logistic regression (final model determined by the backward stepwise method). NI, genotype not included in the final model.

prised of the *hopQ* I allele and *homB* (A2), were also observed. In the context of the correlation between genotypes and clinical outcome, PUD was closely related to the genotypes included in group A and particularly close to *jhp0562*. By contrast, the clinical outcome NUD was not tightly associated with any genotype (Fig. 1).

The dendrogram drawn from the results of the cluster analysis confirmed the presence of the formally identified groups as well as the close relatedness of *jhp0562* to PUD (Fig. 1B).

Distribution of *jhp0562* and other virulence genotypes according to clinical outcome. The 117 *H. pylori* clinical strains isolated from patients with PUD or NUD were examined for the presence of the *jhp0562* gene, as well as the *H. pylori* virulence-associated *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA*, *hopQ*, *sabA*, and *hopZ* genes, in order to assess which genotypes were associated with disease. The results of the binary logistic regression analysis are shown in Table 3. The univariate analysis showed that PUD was strongly associated with the presence of *jhp0562* ($P < 0.001$; odds ratio [OR], 9.14), the *cag* PAI ($P < 0.001$; OR, 15.36), *vacA* s1 allele ($P < 0.001$; OR, 11.14), *babA* ($P < 0.001$; OR, 6.81), *homB* ($P < 0.001$; OR, 4.52), *oipA* gene with "on" status ($P < 0.001$; OR, 5.69), and the *hopQ* type I allele ($P = 0.008$; OR, 3.55). On the other hand, functional *sabA* and *hopZ* genes were more prevalent among NUD strains, although the difference was not statistically significant. After we adjusted for all of the genotypes, only the *cag* PAI and *jhp0562* remained associated with PUD, although *homB* was at the limit of statistical significance (Table 3).

In order to evaluate which genotype combinations presented the most discriminatory capacity to distinguish the clinical outcome, a binary logistic regression analysis was performed. We used the backward stepwise method, which selects a subset among all of the genotypes that best predicts the clinical outcome with a small but statistically significant number of variables. The resulting parsimonious model that discriminates between PUD and NUD was the triple-positive genotype for the *cag* PAI, *jhp0562*, and *homB* (Table 3). In fact, this combination was present in 49.1% of the PUD-associated strains and in only 11.9% of the NUD-associated strains ($P < 0.001$; OR, 7.17; 95% confidence interval [CI], 2.79 to 18.45).

Diversity in *jhp0562* and β-(1,3)GalT genes. With regard to the currently available *H. pylori* complete genomes (n = 10; Table 2), a BLAST search revealed that the *jhp0563*-like ORF is always present and only at the expected locus, and it seems to be regulated by phase variation and rarely expressed *in vitro*. Moreover, the tandem-repeat sequences (21 nucleotides [nt]) differ by up to seven repeats in those strains. The *jhp0562*-like ORF, on the contrary, is not consistently present either at the expected locus or at another locus.

Using the F1-*jhp0562*/*jhp0563*–R1-*jhp0562*/*jhp0563* primer set, PCR fragments corresponding to *jhp0562*, the β-(1,3)GalT gene, or both genes were detected in all of the *H. pylori* clinical strains tested, indicating the constant presence of at least one of these two genes. The fragment corresponding to *jhp0562* presented a constant size of 301 bp, whereas the length of the PCR product corresponding to the β-(1,3)GalT gene ranged from 535 to 675 bp, as confirmed by sequencing. Four different combinations of these two genes were observed (Fig. 2): only one PCR fragment, corresponding either to the *jhp0562* gene (profile 1) or to the β-(1,3)GalT gene (profile 2), or two PCR fragments, corresponding to each of the two genes (profile 3) or corresponding to two β-(1,3)GalT gene fragments of different lengths (profile 4). Overall, the presence of both genes (profile 3) was observed in 41.3% of the strains, the presence of the β-(1,3)GalT gene only (profile 2) was observed in 37.6% of the isolates, and the presence of *jhp0562* only (profile 1) accounted for 17.5% of the strains, while the presence of two β-(1,3)GalT gene fragments of different lengths (profile 4) was rare (3.6% of the strains). The distributions of the four profiles differed markedly between the two disease groups (Table 4). The presence of both genes was most prevalent in PUD-associated strains (29/57, 50.9%), while the majority of NUD-associated strains harbored only the β-(1,3)GalT gene (37/60, 61.7%). Furthermore, the presence of the *jhp0562* gene alone was observed in 19 of the 57 PUD-associated strains (33.3%) but in only 1 NUD-associated strain (1.7%) ($P < 0.001$). Inversely, the presence of the β-(1,3)GalT gene alone was significantly more frequent in NUD-associated strains (37/60, 61.7%) than in PUD-associated strains (7/57, 12.3%) ($P < 0.001$).

In order to check the status of the *jhp0562* and β-(1,3)GalT

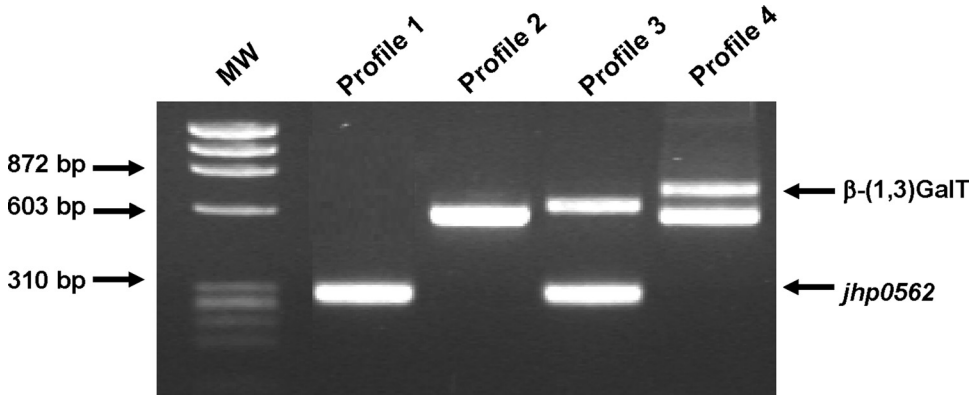

FIG. 2. *jhp0562* and β-(1,3)GalT gene profiles observed in *Helicobacter pylori* strains. The gene profiles were determined by PCR performed on 117 *Helicobacter pylori* strains and amplicons were separated on a 1.5% agarose gel. MW, molecular weight marker X174 RF DNA/HaeIII fragments (72 to 1,353 bp).

genes in *H. pylori* clinical strains, the *jhp0562-jhp0563* locus (according to the J99 strain annotation) was amplified and sequenced. The β-(1,3)GalT gene is a phase-variable gene and is out of frame in most of the sequenced *H. pylori* strains, while this kind of regulation is not observed for *jhp0562*, which, when present, displays a complete coding sequence (Table 2). A total of 34 clinical strains were analyzed at this locus: 8 strains with *jhp0562* only, 13 strains with the β-(1,3)GalT gene only, and 13 strains carrying both genes. Amplification and sequencing results of these clinical strains showed that these genes occupy the same region in the chromosome as in reference strains and display similar genomic organizations, i.e., when the two genes are present, the *jhp0562* gene is always upstream of the β-(1,3)GalT gene. Among the strains analyzed, the *jhp0562* gene had a constant length of 999 bp, except for two cases with lengths of 993 and 1,002 bp, and the corresponding ORF was always in frame. The β-(1,3)GalT gene presented a variable length (from 1,230 to 1,409 bp) due to a variation in the number of 21-nt tandem-repeat previously reported sequences (31–33), and 16 of the 26 (61.5%) sequences analyzed were out of frame, due to the presence of homopolymeric tract of C repeats (ranging from 8 to 15) in the middle region of the gene, resulting in a nonsense frameshift mutation (31). Interestingly, in the 13 strains in which the β-(1,3)GalT gene was present alone, 9 ORFs (69.2%) were in frame, whereas among the 13 isolates positive for both genes, 12 (92.3%) carried an out of frame β-(1,3)GalT gene.

TABLE 4. Distribution of the *jhp0562* and β-(1,3)GalT genes by clinical outcome among the 117 *H. pylori* strains studied

Copy no. of the genes according to the size of PCR fragments (profile designation)[a]	No. (%) of strains associated with:		*P* value[b]
	PUD (n = 57)	NUD (n = 60)	
Single copy of *jhp0562* gene (1)	19 (33.3)	1 (1.7)	<0.001
Single copy of the β-(1,3)GalT gene (2)	7 (12.3)	37 (61.7)	<0.001
jhp0562 and β-(1,3)GalT genes (3)	29 (50.9)	20 (33.3)	NS
Two copies of the β-(1,3)GalT gene (4)	2 (3.5)	2 (3.3)	NS

[a] According to lanes 1 to 4 of Fig. 1.
[b] *P* values were determined using Fisher's exact test. NS, not statistically significant.

Study of the *in vitro* transcripts of the *jhp0562* and β-(1,3)GalT genes. The total RNAs from the sequenced 26695 and J99 strains as well as from the 34 clinical strains mentioned above were reverse transcribed with random primers and then used in a control PCR targeting the 23S rRNA gene and in a PCR discriminating between *jhp0562* and β-(1,3)GalT genes. First, the 23S rRNA gene transcript was observed for all of the strains tested (not shown). With regard to the second PCR, only the fragment corresponding to the *jhp0562* gene was observed for the J99 strain, and the expected PCR fragment corresponding to the β-(1,3)GalT gene was lacking. No reverse transcription-PCR product corresponding to the β-(1,3)GalT gene was detected for strain 26695 (data not shown), whose β-(1,3)GalT gene is out of frame. With regard to clinical strains, the *jhp0562* gene was found to be expressed in all of the 21 strains positive for this gene, and the β-(1,3)GalT gene was expressed differently according to the frame status of the sequence, i.e., an absence of amplification for the 16 out-of-frame sequences and an amplification for the 10 in-frame sequences.

DISCUSSION

Research over the past few years has led to the identification of several *H. pylori* markers for severe gastric disease. This study reports on the clinical relevance and diversity of the *H. pylori jhp0562* gene, encoding a protein from the cell envelope, involved in the LPS biosynthesis, and recently associated with PUD in children (27).

According to several statistical approaches, a positive genotype for *jhp0562* was significantly more prevalent in strains with a more-virulent profile, more specifically in strains carrying a positive genotype for the *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA* "on," and the *hopQ* I allele, than in the less-virulent strains. These results suggest that the *jhp0562* gene could be a good marker for *H. pylori* virulence. Moreover, when evaluating the distribution of *H. pylori* genotypes according to clinical outcome, an increased risk for the development of PUD was associated with strains carrying *jhp0562*, confirming previous

results obtained with a smaller group of strains (from 15 patients with PUD and 30 patients with NUD) (27). The *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA* "on," and *hopQ* I allele genotypes were also independently associated with PUD. However, the multivariate analysis showed that only the presence of the *cag* PAI and *jhp0562* are significantly associated with PUD. Finally, the triple-positive genotype for *cag* PAI, *jhp0562*, and *homB* provided the best discriminatory model for separating outcomes of PUD and NUD. Although these results were obtained with a small and select group of strains which will require further validation, overall they suggest that the *jhp0562* gene can be a new comarker of PUD. It is important to emphasize that the clinical strains used in the present study were isolated from children, thus leading to a more reliable model of the early natural conditions of the infection and one that is less influenced by environmental determinants. Several studies report that pediatric *H. pylori* strains display more distinct virulence profiles than those isolated from adults (13, 18, 25, 34). Accordingly, the severity of *H. pylori*-associated disease in younger subjects may be closely related to the virulence of the strain, independently of the host factors, which may not be the case among adults. The fact that the results of this study are in contrast with those from previous studies of adults in the same geographic area, for which only *cagA* was associated with PUD, is added proof (25, 26).

In the present study, *jhp0562*, the *cag* PAI, *vacA* s1 allele, *babA*, *homB*, *oipA* "on," and *hopQ* I were strongly associated with each other. Similar findings have been previously reported in other studies, excluding such results for *jhp0562*, whose frequency was not determined (6, 10, 12, 25, 37, 40). Thus, the coexpression of several genes most likely results from a shared selective pressure, contributing to the fitness of the strains *in vivo*. Therefore, the strong linkage of *jhp0562* with PUD may be due to its association with other virulence genes. One can also speculate that all of these factors act synergistically in causing damage to the host.

The *jhp0562* gene is a strain-specific gene, located upstream of the 80%-similar β-(1,3)GalT gene (2, 24, 35). This genomic organization was confirmed among *H. pylori* clinical isolates presenting both of these genes. It was, however, observed that the *jhp0562*–β-(1,3)GalT gene locus could be occupied only by the *jhp0562* gene [the β-(1,3)GalT gene being absent], and this profile was significantly more prevalent among PUD-associated strains than among NUD-associated isolates (33.3% versus 1.7%, respectively). Interestingly, the occurrence of the β-(1,3)GalT gene locus alone was the most prevalent genotype among NUD-associated strains, and this prevalence was significantly higher than that observed among PUD-associated strains (61% versus 12.3%, respectively). These observations confirm the previously described association of *jhp0562* with PUD in children (27).

H. pylori expresses Le antigens in its LPS, both type I (Le^a and Le^b) and type II (Le^x and Le^y), and they are structurally related to the human blood group antigens also expressed in gastric epithelial cells (22, 23). The exact role of these Le antigenic structures during *H. pylori* infection is still unclear, although their importance for bacterial colonization, adherence, or evasion of host immune response has been hypothesized (4, 11, 14, 39). A recent study reported that mutagenesis of *jhp0562* resulted in the loss of expression of all Le types, suggesting that the product of this gene is truly an essential glycosyltransferase for Le expression (30).

The β-(1,3)GalT gene encodes a β-(1,3)-galactosyltransferase involved in type I Le antigen synthesis (3, 31). Expression of this gene is regulated by a slipped-strand repair mechanism, while *jhp0562* is a non-phase-variable homologue of the β-(1,3)GalT gene (2, 24, 31). A variation in the gene status was confirmed for the β-(1,3)GalT gene among the clinical strains examined. Furthermore, studies of the *in vitro* expression of the β-(1,3)GalT gene showed that this gene was expressed differently according to the frame status of the sequence, confirming its regulation by a slipped-strand repair mechanism. The possibility of a translational control locus of Le gene expression was also suggested for the β-(1,3)GalT gene (31).

Interestingly, β-(1,3)GalT functionality seemed to be related to the presence or absence of the homologue gene *jhp0562*. In fact, for 92.3% of the isolates harboring both genes, the β-(1,3)GalT gene was nonfunctional, while a complete corresponding *jhp0562* ORF was present. In contrast, in the absence of *jhp0562*, only 33.3% of the isolates presented a truncated β-(1,3)GalT ORF. These results offer further support that the products of the two genes have the same cell function and that one product may functionally replace the other.

In vivo regulation of the β-(1,3)GalT gene was initially observed in the 21-nt tandem-repeat sequences which showed length variation during colonization (32), and it was proposed that this gene uses instability within simple sequence repeats as a gene switching mechanism (33). Pohl et al. also showed that the variation in the β-(1,3)GalT gene homopolymeric tract length affected Le^b expression (31). More importantly, using a Le^b-expressing transgenic mouse model, they provided evidence that *H. pylori* can change the Le phenotype to match the Le phenotype of its host, supporting previous studies in humans and rhesus monkeys (31, 38, 39). The change to a type I Le phenotype was found to be linked to a variation in the β-(1,3)GalT gene homopolymeric tract length, suggesting that this locus is under host selection pressure (31). Moreover, the possibility of recombination between the β-(1,3)GalT gene and its upstream homologue, *jhp0562*, was suggested (30). Therefore, based on this and based on the results of the present study, one can speculate that the presence of the *jhp0562* gene in strains with a truncated β-(1,3)GalT may constitute an alternative regulation mechanism for Le antigen expression involved in host adaptation.

In conclusion, the strong correlation of *jhp0562* with PUD and with the main *H. pylori*-associated virulence markers suggests that this gene is a good candidate as a virulence marker of *H. pylori*. Moreover, the product of *jhp0562*, a glycosyltransferase, may be essential for the survival of the bacteria *in vivo*, as it is most likely implicated in host adaptation, contributing to the persistence of the infection.

ACKNOWLEDGMENTS

We thank the Sociedade Portuguesa de Gastrenterologia and the Programme de Cooperation Scientifique et Technique Franco-Portugais of the French Embassy in Portugal for supporting the project.

We thank the Fundação para a Ciência e Tecnologia, project PPCDT/SAL-IMI/57297/2004, for financial support.

REFERENCES

1. **Akopyants, N. S., S. W. Clifton, D. Kersulyte, J. E. Crabtree, B. E. Youree, C. A. Reece, N. O. Bukanov, E. S. Drazek, B. A. Roe, and D. E. Berg.** 1998. Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. Mol. Microbiol. **28:**37–53.
2. **Alm, R. A., L. S. L. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176–180.
3. **Appelmelk, B. J., M. C. Martino, E. Veenhof, M. A. Monteiro, J. J. Maaskant, R. Negrini, F. Lindh, M. Perry, G. Del Giudice, and C. M. J. E. Vandenbroucke-Grauls.** 2000. Phase variation in H type I and Lewis a epitopes of *Helicobacter pylori* lipopolysaccharide. Infect. Immun. **68:**5928–5932.
4. **Appelmelk, B. J., and C. M. J. E. Vandenbroucke-Grauls.** 2000. *H. pylori* and Lewis antigens. Gut **47:**10–11.
5. **Atherton, J. C., P. Cao, R. M. Peek, M. K. R. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*—association of specific *vacA* types with cytotoxin production and peptic ulceration. J. Biol. Chem. **270:**17771–17777.
6. **Cao, P., and T. L. Cover.** 2002. Two different families of *hopQ* alleles in *Helicobacter pylori*. J. Clin. Microbiol. **40:**4504–4511.
7. **Cao, P., K. J. Lee, M. J. Blaser, and T. L. Cover.** 2005. Analysis of *hopQ* alleles in East Asian and Western strains of *Helicobacter pylori*. FEMS Microbiol. Lett. **251:**37–43.
8. **Correa, P., J. Fox, E. Fontham, B. Ruiz, Y. Lin, D. Zavala, N. Taylor, D. Mackinley, E. De Lima, H. Portilla, and G. Zarama.** 1990. *Helicobacter pylori* and gastric carcinoma. Cancer **66:**2569–2574.
9. **Doig, P., B. L. de Jonge, R. A. Alm, E. D. Brown, M. Uria-Nickelsen, B. Noonan, S. D. Mills, P. Tummino, G. Carmel, B. C. Guild, D. T. Moir, G. F. Vovis, and T. J. Trust.** 1999. *Helicobacter pylori* physiology predicted from genomic comparison of two strains. Microbiol. Mol. Biol. Rev. **63:**675–707.
10. **Dossumbekova, A., C. Prinz, J. Mages, R. Lang, J. G. Kusters, A. H. Van Vliet, W. Reindl, S. Backert, D. Saur, R. M. Schmid, and R. Rad.** 2006. *Helicobacter pylori* HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of *hopH* gene polymorphisms. J. Infect. Dis. **194:**1346–1355.
11. **Edwards, N. J., M. A. Monteiro, G. Faller, E. J. Walsh, A. P. Moran, I. S. Roberts, and N. J. High.** 2000. Lewis X structures in the O antigen side-chain promote adhesion of *Helicobacter pylori* to the gastric epithelium. Mol. Microbiol. **35:**1530–1539.
12. **Gerhard, M., N. Lehn, N. Neumayer, T. Boren, R. Rad, W. Schepp, S. Miehlke, M. Classen, and C. Prinz.** 1999. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proc. Natl. Acad. Sci. U. S. A. **96:**12778–12783.
13. **González-Valencia, G., J. C. Atherton, O. Muñoz, M. Dehesa, A. Madrazo de la Garza, and J. Torres.** 2000. *Helicobacter pylori vacA* and *cagA* genotypes in Mexican adults and children. J. Infect. Dis. **182:**1450–1454.
14. **Heneghan, M. A., C. F. McCarthy, and A. P. Moran.** 2000. Relationship of blood group determinants on *Helicobacter pylori* lipopolysaccharide with host Lewis phenotype and inflammatory response. Infect. Immun. **68:**937–941.
15. **Huang, X., and M. Miller.** 1991. A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. **12:**337–357.
16. **Kuipers, E. J., A. M. Uyterlinde, A. S. Pena, R. Roosendaal, G. Pals, G. F. Nelis, H. P. M. Festen, and S. G. M. Meuwissen.** 1995. Long-term sequelae of *Helicobacter pylori* gastritis. Lancet **345:**1525–1528.
17. **Lehours, P., A. Ménard, S. Dupouy, B. Bergey, F. Richy, F. Zerbib, A. Ruskone-Fourmestraux, J. C. Delchier, and F. Mégraud.** 2004. Evaluation of the association of nine *Helicobacter pylori* virulence factors with strains involved in low-grade gastric mucosa-associated lymphoid tissue lymphoma. Infect. Immun. **72:**880–888.
18. **Li, J., Z. Ou, F. Wang, Y. Guo, R. Zhang, J. Zhang, P. Li, W. Xu, and Y. He.** 2009. Distinctiveness of the *cagA* genotype in children and adults with peptic symptoms in South China. Helicobacter **14:**248–255.
19. **Mégraud, F., and H. Lamouliatte.** 1992. *Helicobacter pylori* and duodenal ulcer. Evidence suggesting causation. Dig. Dis. Sci. **37:**769–772.
20. **Mégraud, F., N. Lehn, T. Lind, E. Bayerdorffer, C. O'Morain, R. Spiller, P. Unge, S. V. Van Zanten, M. Wrangstadh, and C. F. Burman.** 1999. Antimicrobial susceptibility testing of *Helicobacter pylori* in a large multicenter trial: the MACH 2 study. Antimicrob. Agents Chemother. **43:**2747–2752.
21. **Ménard, A., A. Santos, F. Mégraud, and M. Oleastro.** 2002. PCR-restriction fragment length polymorphism can also detect point mutation A2142C in the 23S rRNA gene, associated with *Helicobacter pylori* resistance to clarithromycin. Antimicrob. Agents Chemother. **46:**1156–1157.
22. **Monteiro, M. A., B. J. Appelmelk, D. A. Rasko, A. P. Moran, S. O. Hynes, L. L. MacLean, K. H. Chan, F. St. Michael, S. M. Logan, J. O'Rourke, A. Lee, D. E. Taylor, and M. B. Perry.** 2000. Lipopolysaccharide structures of *Helicobacter pylori* genomic strains 26695 and J99, mouse model *H. pylori* Sydney strain, *H. pylori* P466 carrying sialyl Lewis X, and *H. pylori* UA915 expressing Lewis B—classification of *H. pylori* lipopolysaccharides into glycotype families. Eur. J. Biochem. **267:**305–320.
23. **Monteiro, M. A., K. H. N. Chan, D. A. Rasko, D. E. Taylor, P. Y. Zheng, B. J. Appelmelk, H. P. Wirth, M. Q. Yang, M. J. Blaser, S. O. Hynes, A. P. Moran, and M. B. Perry.** 1998. Simultaneous expression of type 1 and type 2 Lewis blood group antigens by *Helicobacter pylori* lipopolysaccharides. J. Biol. Chem. **273:**11533–11543.
24. **Oh, J. D., H. Kling-Backhed, M. Giannakis, J. Xu, R. S. Fulton, L. A. Fulton, H. S. Cordum, C. Wang, G. Elliott, J. Edwards, E. R. Mardis, L. G. Engstrand, and J. I. Gordon.** 2006. The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. Proc. Natl. Acad. Sci. U. S. A. **103:**9999–10004.
25. **Oleastro, M., R. Cordeiro, J. Ferrand, B. Nunes, P. Lehours, I. Carvalho-Oliveira, A. I. Mendes, D. Penque, L. Monteiro, F. Mégraud, and A. Ménard.** 2008. Evaluation of the clinical significance of *homB*, a novel candidate marker of *Helicobacter pylori* strains associated with peptic ulcer disease. J. Infect. Dis. **198:**1379–1387.
26. **Oleastro, M., R. Cordeiro, Y. Yamaoka, D. Queiroz, F. Mégraud, L. Monteiro, and A. Ménard.** 2009. Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. Gut Pathog. **1:**12.
27. **Oleastro, M., L. Monteiro, P. Lehours, F. Mégraud, and A. Ménard.** 2006. Identification of markers for *Helicobacter pylori* strains isolated from children with peptic ulcer disease by suppressive subtractive hybridization. Infect. Immun. **74:**4064–4074.
28. **Parsonnet, J., S. Hansen, L. Rodriguez, A. B. Gelb, R. A. Warnke, E. Jellum, N. Orentreich, J. H. Vogelman, and G. D. Friedman.** 1994. *Helicobacter pylori* infection and gastric lymphoma. N. Engl. J. Med. **330:**1267–1271.
29. **Peck, B., M. Ortkamp, K. D. Diehl, E. Hundt, and B. Knapp.** 1999. Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*. Nucleic Acids Res. **27:**3325–3333.
30. **Pohl, M., and M. J. Blaser.** 2009. A role for the *beta-(1,3)galT* upstream homolog, *jhp0562*, in *Helicobacter pylori* Lewis antigen synthesis, abstr. P-113, p. 111. 15th Int. Workshop Campylobacter Helicobacter Related Organisms. International Workshop on Campylobacter, Helicobacter, and Related Organisms, Niigata, Japan.
31. **Pohl, M., J. Romero-Gallo, J. Guruge, D. Tse, J. Gordon, and M. Blaser.** 2009. Host-dependent Lewis (Le) antigen expression in *Helicobacter pylori* cells recovered from Leb-transgenic mice. J. Exp. Med. **206:**3061–3072.
32. **Salaün, L., S. Ayraud, and N. J. Saunders.** 2005. Phase variation mediated niche adaptation during prolonged experimental murine infection with *Helicobacter pylori*. Microbiology **151:**917–923.
33. **Salaün, L., B. Linz, S. Suerbaum, and N. J. Saunders.** 2004. The diversity within an expanded and redefined repertoire of phase-variable genes in *Helicobacter pylori*. Microbiology **150**(Pt. 4)**:**817–830.
34. **Talarico, S., B. Gold, J. Fero, D. Thompson, J. Guarner, S. Czinn, and N. Salama.** 2009. Pediatric *Helicobacter pylori* isolates display distinct gene coding capacities and virulence gene marker profiles. J. Clin. Microbiol. **47:**1680–1688.
35. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539–547.
36. **Tummuru, M. K. R., T. L. Cover, and M. J. Blaser.** 1993. Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. Infect. Immun. **61:**1799–1809.
37. **van Doorn, L., C. Figueiredo, R. Sanna, A. Plaisier, P. Schneeberger, W. De Boer, and W. Quint.** 1998. Clinical relevance of the *cag*A, *vacA*, and *iceA* status of *Helicobacter pylori*. Gastroenterology **115:**58–66.
38. **Wirth, H., M. Yang, E. Sanabria-Valentín, D. Berg, A. Dubois, and M. Blaser.** 2006. Host Lewis phenotype-dependent *Helicobacter pylori* Lewis antigen expression in rhesus monkeys. FASEB J. **20:**1534–1536.
39. **Wirth, H. P., M. Q. Yang, R. M. Peek, K. T. Tham, and M. J. Blaser.** 1997. *Helicobacter pylori* Lewis expression is related to the host Lewis phenotype. Gastroenterology **113:**1091–1098.
40. **Yamaoka, Y., S. Kikuchi, H. M. T. El-Zimaity, O. Gutierrez, M. S. Osato, and D. Y. Graham.** 2002. Importance of *Helicobacter pylori oipA* in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. Gastroenterology **123:**414–424.