JOURNAL OF CLINICAL MICROBIOLOGY, Nov. 2004, p. 5302–5308 Vol. 42, No. 11
0095-1137/04/$08.00+0 DOI: 10.1128/JCM.42.11.5302–5308.2004

The *cag* Pathogenicity Island of *Helicobacter pylori* Is Disrupted in the Majority of Patient Isolates from Different Human Populations

Farhana Kauser,[1]† Aleem A. Khan,[2]† M. Abid Hussain,[1]† Ian M. Carroll,[3] Naheed Ahmad,[4] Santosh Tiwari,[2] Yogesh Shouche,[5] Bimal Das,[6] Mahfooz Alam,[1] S. Mahaboob Ali,[2] C. M. Habibullah,[2] Rafaela Sierra,[7] Francis Megraud,[8] Leonardo A. Sechi,[9] and Niyaz Ahmed[1]*

Pathogen Evolution Group, Centre for DNA Fingerprinting and Diagnostics,[1] and Centre for Liver Research and Diagnostics, Deccan College of Medical Sciences and Allied Hospitals,[2] Hyderabad, Department of Botany, Patna University, Patna,[4] National Centre for Cell Science, Pune,[5] and Department of Microbiology, All India Institute of Medical Sciences, New Delhi,[6] India; The Moyne Institute of Preventive Medicine, Trinity College, Dublin, Ireland[3]; Instituto de Investigaciones en Salud, Universidad de Costa Rica, San José, Costa Rica[7]; Laboratoire de Bacteriologie–C.H.U. Pellegrin, Bordeaux, France[8]; and Department of Biomedical Sciences, University of Sassari, Sassari, Italy[9]

Received 7 May 2004/Returned for modification 1 June 2004/Accepted 25 June 2004

The *cag* pathogenicity island (*cag*-PAI) is one of the major virulence determinants of *Helicobacter pylori*. The chromosomal integrity of this island or the lack thereof is speculated to play an important role in the progress of the gastroduodenal pathology caused by *H. pylori*. We determined the integrity of the *cag*-PAI by using specific flanking and internally anchored PCR primers to know the biogeographical distribution of strains carrying fully integral *cag*-PAI with proinflammatory behavior in vivo. Genotypes based on eight selected loci were studied in 335 isolates obtained from eight different geographic regions. The *cag*-PAI appeared to be disrupted in the majority of patient isolates throughout the world. Conservation of *cag*-PAI was highest in Japanese isolates (57.1%). However, only 18.6% of the Peruvian and 12% of the Indian isolates carried an intact *cag*-PAI. The integrity of *cag*-PAI in European and African strains was minimal. All 10 strains from Costa Rica had rearrangements. Overall, a majority of the strains of East Asian ancestry were found to have intact *cag*-PAI compared to strains of other descent. We also found that the *cagE* and *cagT* genes were less often rearranged (18%) than the *cagA* gene (27%). We attempted to relate *cag*-PAI rearrangement patterns to disease outcome. Deletion frequencies of *cagA*, *cagE*, and *cagT* genes were higher in benign cases than in isolates from severe ulcers and gastric cancer. Conversely, the *cagA* promoter and the left end of the *cag*-PAI were frequently rearranged or deleted in isolates linked to severe pathology. Analysis of the *cag*-PAI genotypes with a different biogeoclimatic history will contribute to our understanding of the pathogen-host interaction in health and disease.

Infection with *Helicobacter pylori* has plausible associations with a variety of clinical outcomes. This includes chronic gastritis, duodenal ulcer, gastric cancer, lymphomas and seemingly paradoxical relationships with gastroesophageal reflux disease and adenocarcinoma of the gastric cardia (12, 13, 15, 29, 38). *H. pylori* inhabits at least 50% of the world's population, with the highest incidence recorded in industrially underdeveloped areas, including Asia (70 to 80%) and Africa (70 to 90%), and a waning occurrence in developed areas such as North America (30 to 40%), South America (80 to 90%), and Europe (30 to 70%) (33). Prevalence rates vary between populations and between groups within the same population (49).

Variation in the clinical outcome of *H. pylori* induced pathology is multifactorial, involving a complex interplay between the host immune responses, pathogen virulence factors, and niche characteristics. Many putative virulence factors have been identified in *H. pylori* that contribute to its pathogenesis. Most important among them are the 128-kDa cytotoxin-associated gene-encoded antigen gene (*cagA*) (18, 45) and the vacuolating cytotoxin antigen gene (*vacA*) (4, 43). The *cagA* gene is present downstream of an ~40-kb cluster of virulence genes known as the *cag* pathogenicity island (*cag*-PAI) located at the 3′ end of the glutamate racemase gene. In some strains, the presence of an insertion element IS*605* may disrupt an otherwise-uninterrupted unit into two regions, *cag* I and *cag* II, consisting of at least 14 and 16 open reading frames (ORFs), respectively (11). These genes encode a type IV secretion system that forms a syringe-like structure to translocate the immunodominant *cagA* protein into the gastric epithelial cells. *cag*-PAI has also been implicated in the induction of interleukin-8 (IL-8) in cultured gastric cells. As shown by the systematic mutagenesis studies of individual genes encoding the *cag*-PAI, there are at least 17 of 27 genes that are found to be essential for translocation of *cagA* into host cells (syringe-like function), and 14 were necessary for *H. pylori* to fully induce

* Corresponding author. Mailing address: Pathogen Evolution Group, CDFD, ECIL Rd., Nacharam, Hyderabad 500 076, India. Phone: 91 40 27150008. Fax: 91 40 27155610. E-mail: niyaz@cdfd.org.in.

† F.K., A.A.K., and M.A.H. contributed equally to this study.

transcription of IL-8 (17). This property contributes to the proinflammatory power of the strains and thus also to its virulence capability. Nevertheless, the induction of IL-8 is not exclusively linked to *cag*-PAI. Exposure of gastric epithelial cells to *cag*-PAI-positive *H. pylori* can activate the proto-oncogenes c-*fos* and c-*jun*, a crucial step in the development of *H. pylori*-related neoplasia (30). An intact *cag*-PAI has therefore been thought to contribute toward full proinflammatory power of *H. pylori*. However, the island has proved to be prone to disruption due to various genetic rearrangements occurring within and outside the constituent genes. Further still, the impact of intactness or rearrangement of the island or the constituent genes on the progression of gastroduodenal pathology has been debated. It has been argued that abolition of *cagE* gene results in a considerable reduction of IL-8 production (2, 11, 14, 36, 38, 51), whereas strains capable of eliciting an IL-8 response irrespective of an intact *cag*-PAI have been described (23, 28). Some strains causing nonulcer dyspepsia have regions of the promoter of *cagA* gene deleted (28), although this gene has long been regarded as a marker for the functionality of the *cag*-PAI. However, due to rearrangements often inhibiting *cagA*, it does not seem to be a reliable indicator of the virulence spectrum of strains (7). Other members of the *cag*-PAI have also been scrutinized for their involvement in virulence characteristics. The *cagE* gene served as a better marker of an intact *cag*-PAI in Japanese strains (22, 28) and French isolates (5).

The *cagG* gene was shown to be a better indicator for the presence of an intact *cag*-PAI than the *cagE* gene in French isolates (41). Deletion of *cagE*, *cagT*, *cagA*, *cagG*, and *cagM* genes was reported in all cases of chronic gastritis, gastric ulcer and gastric cancer, indicating that the pathogenicity of *H. pylori* may not be determined by *cag*-PAI genes alone in such cases (24). In a series of patients from Taiwan, the presence of *cagA*, *cagC*, *cagE*, *cagF*, *cagN*, and *cagT* genes in the PAI showed no relationship to the type of disease and/or the histological features present in the patients (46). The involvement of these genes and others in eliciting a strong immune response has also been contested (45). Rearrangements in *cag*-PAI in *H. pylori* isolates in different countries have been reported in isolation (5, 22, 23, 26, 28, 37, 40, 47). However, there are no comprehensive data available on the abundance of intact versus rearranged *cag*-PAI in strains inhabiting geographically and culturally distinct patient populations on a global scale. Such data on geographical distribution of *cag*-PAI rearrangement patterns may provide a window into the trinity of bacterial virulence, host genetic predisposition, and niche characteristics in a better way. These concerns, as well as our continued efforts (1, 9, 10, 21) to study the complexities associated with the infection dynamics and evolutionary genetics of this interesting pathogen with regard to host-pathogen interactions at all levels, led to the inception of the present study. We have used here simple and nested PCR-based genotyping approaches to demonstrate for the first time with a large number of strains that the *cag*-PAI is not perpetually intact in majority of *H. pylori* in the world.

MATERIALS AND METHODS

Patients and bacterial isolates. *cag*-PAI analyses were performed on a total of 335 *H. pylori* DNAs obtained from patient isolates from eight different countries: Spain, Peru, South Africa, England, Ireland, Japan, Costa Rica, and India. These isolates were recovered from unrelated patients undergoing upper gastrointestinal endoscopy after informed consent was obtained. Bacterial isolates were cultured from antral biopsies as described previously (1). Genomic DNA samples were prepared, quantified, and preserved according to the methods described elsewhere (1). Most of the strains we used were described in other studies, and their microbiological and clinical characteristics have been elaborated elsewhere (1, 9, 25, 35, 48). To remove biases, isolates were picked randomly based merely on geographic representation. Classification and tabulation based on clinical corroboration was attempted only after the experiments were completed. Isolates including genomic DNA preparations belonging to patients from Peru, Spain, Japan, England, and South Africa were a gift from Asish Mukhopadhyay and Douglas E. Berg (Washington University, St. Louis, Mo.). Genomic DNAs from Irish *H. pylori* isolates were obtained from single antral gastric biopsies from patients experiencing peptic ulcer disease at the Meath-Adelaide and St. James's Hospitals in Dublin, Ireland.

***cag*-PAI PCR analysis.** PCR analyses were carried out with eight oligonucleotide pairs (Table 1) as described by Ikenoue et al. (22) to amplify five different loci spread over the *cag* I and *cag* II regions. The presence or absence of the gene *cagA*, the promoter regions of *cagA*, *cagE*, and *cagT*, and the left end of *cag*-PAI (LEC) was detected. All PCRs were performed in a total volume of 10 μl containing 0.25 μl of genomic DNA (50 ng/μl), 10× PCR buffer (Applied Biosystems), 2 mM $MgCl_2$ (MBI), 80 μM deoxynucleoside triphosphates (Amersham Biosciences), 10 pM concentrations of each forward and reverse primer, and 1 U of *Taq* polymerase (Applied Biosystems). The total volume was made up with autoclaved Milli-Q water. *H. pylori* 26695 DNA was taken as a positive control. A negative control lacking either DNA or primers was also included. The absence of *cag*-PAI was confirmed in isolates reported to be negative for the same by using previously described primers flanking the *cag*-PAI (11, 22, 25). Appropriate measures were taken to avoid contamination. PCR was performed in a 9700 thermal cycler (Applied Biosystems). Amplification conditions were optimized as follows: initial denaturation for 5 min at 94°C was followed by 40 cycles of denaturation at 90°C for 30 s, annealing at 52°C for 30 s, and extension at 70°C for 1 min. After a final extension at 70°C for 10 min, the amplified products were resolved on a 1% agarose gel in 1× TAE buffer (40 mM Tris-acetate–1 mM EDTA [pH 8.0]) containing ethidium bromide (0.5 μg/ml) and then visualized under UV light.

All of the *cag*-PAI rearrangement profiles generated in the present study were archived in the form of a searchable database (http://www.cdfd.org.in/amplibase/HP).

RESULTS

PCR amplification of *cag*-PAI regions. We determined intactness or otherwise of the *cag*-PAI genes by using specific flanking and internally anchored PCR primers. This was essentially to know the geographic prevalence of strains with different *cag*-PAI rearrangement patterns and whether these patterns could be associated with the clinical outcome of the infection. The reasons for spanning the entire *cag*-PAI region with eight specific PCR primers have been already explained elsewhere (28) based on Southern blot analysis and, more recently, by Ikenoue et al. (22). In our experiments, all of the PCR results were 100% reproducible on all occasions. All of the samples were independently tested at the Centre for DNA Fingerprinting and Diagnostics (CDFD) and the Deccan College of Medical Sciences. Our results, upon crosschecking, were found to be the same within and between the two labs. PCR amplification protocols were portable and easy to perform. The PCR amplification procedure was tested on different positive and negative control samples and also with paired isolates and serial isolates obtained from the same patients months apart. Overall, the PCR technique that we used was robust, reproducible, portable, easy to perform, and highly specific for the corresponding regions of the *cag*-PAI (Fig. 2).

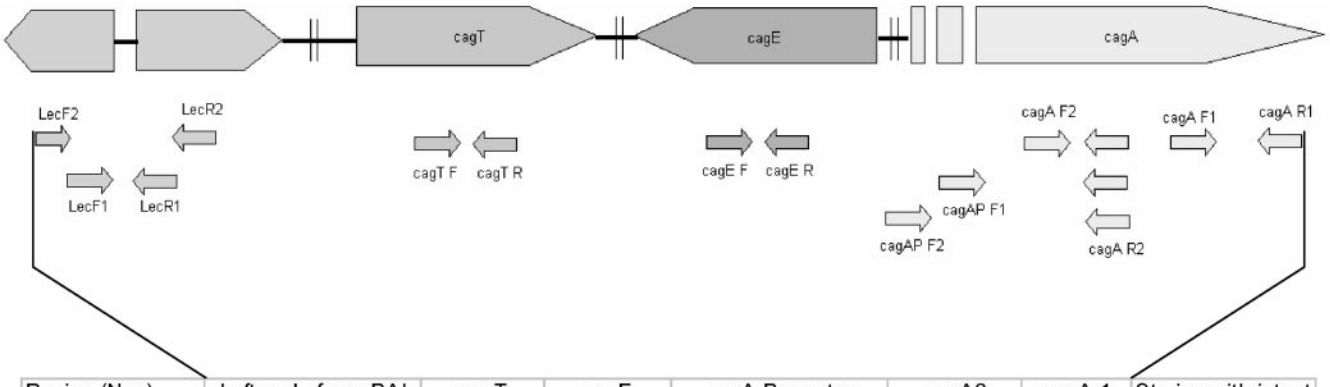
TABLE 1. Details of the PCR primers used for rearrangement analysis within *cag*-PAI

Primer pairs	Sequence	Amplicon size (bp)	Locus name	Coordinates in HP26695 genome
cagAF1 cagAR1	AACAGGACAAGTAGCTAGCC TATTAATGCGTGTGTGGCTG	349	HP0547/JHP0495	582638–583137
cagAF2 cagAR2	GATAACAGGCAAGCTTTTGA TCTGCCAAACAATCTTTTGCAG	701	HP0547/JHP0495	580092–580440
cagAP-F1 cagAR2	GTGGGTAAAAATGTGAATCG TCTGCCAAACAATCTTTTGCAG	730	Promoter region of HP0547/JHP0495	579697–580440
cagAP-F2 cagAR2	CTACTTGTCCCAACCATTTT TCTGCCAAACAATCTTTTGCAG	1181	Promoter region of HP0547/JHP0495	579231–580440
cagEF1 cagER1	GCGATTGTTATTGTGCTTGTAG GAAGTGGTTAAAAAATCAATGCCCC	329	HP0544/JHP0492	577492–579697
cagTF1 cagTR1	CCATGTTTATACGCCTGTGT CCATGTTTATACGCCTGTGT	301	HP0532/JHP0522	565171–565471
LecF1 LecR1	ACATTTTGGCTAAATAAACGCTG TCTCCATGTTGCCATTATGCT	384	JHP0521/JHP0522	547222–547242
LecF2 LecR2	ATAGCGTTTTGTGCATAGAA ATCTTTAGTCTCTTTAGCTT	877	JHP0520/JHP521	–[a]

[a] –, Ikenoue et al. (22).

For the strains with an incomplete or rearranged *cag*. PAI, *PCRs we*re repeated for all of the regions to reconfirm the results.

The percentage of PCRs positive for the *cagE* and *cagT* genes was always more (82%) than any other *cag*-PAI region studied (Fig. 1 to 3). These genes were more frequently intact compared to the *cagA* gene (72.8%). Deletions were frequently observed in the three ORFs located in the extreme left of the



Region (Nos)	Left end of cag-PAI		cagT	cagE	cagA Promoter		cagA2	cagA 1	Strains with intact
	LEC1	LEC2			P1	P2			cag-PAI
Japan (28)	25	22	25	28	28	27	27	26	16 (57.1%)
India (63)	36	29	49	51	26	15	44	31	8 (12.0%)
Peru (43)	41	28	37	41	27	21	41	29	8 (18.6%)
Spain (30)	23	19	20	30	26	10	27	21	4 (13.3%)
England (66)	50	41	55	47	40	32	41	35	2 (3.03%)
Ireland (76)	71	50	62	52	62	32	74	52	12 (15.8%)
Africa (19)	14	14	18	18	17	15	19	8	2 (11.1%)
Costa Rica (10)	2	3	9	8	3	1	8	5	0 (0%)

FIG. 1. Schematic representation of the loci used for analyzing integrity of the *cag*-PAI in strains from eight countries. Primer sets used amplified eight regions, namely, LEC1 (primers LecF1/LecR1), LEC2 (primers LecF2 and LecR2), cagT (primers cagTF and cagTR), cagE (primers cagEF and cagER), cagA P1 (primers cagAP-F1 and cagAR2), cagA P2 (primers cagAP-F2 and cagAR2), cagA2 (primers cagAF2 and cagAR2), and cagA1 (primers cagAF1 and cagAR1).

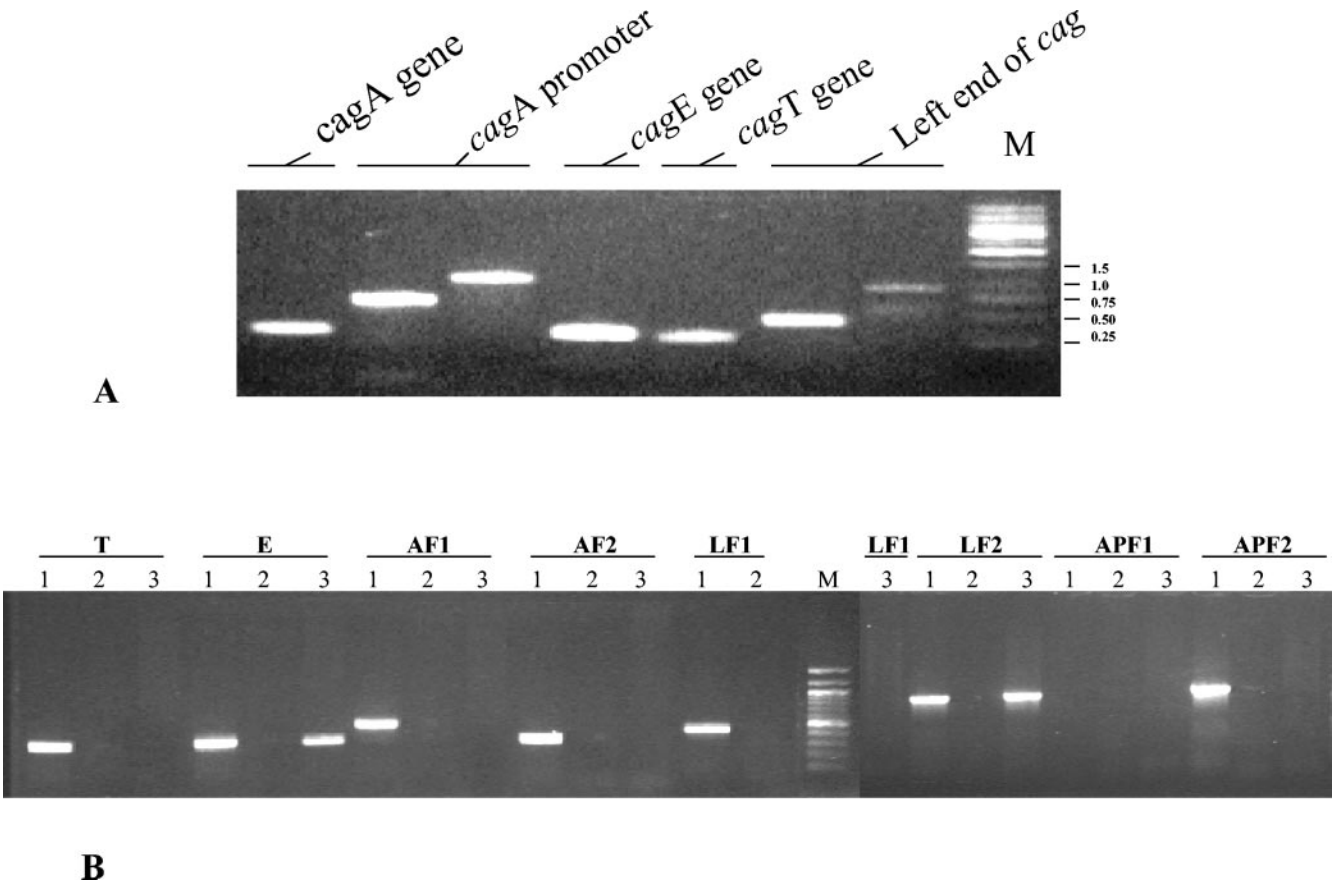


FIG. 2. (A) Agarose gel electrophoresis image showing PCR amplification with strain L22 (from India) for the *cag* PAI regions analyzed. This strain was intact for all of the regions studied within the *cag* PAI. M, 1-kb molecular size marker with corresponding base pair sizes indicated on the extreme right end in kilobases. (B) PCR amplification of *cagT*, *cagE*, and *cagA* (AF1 and AF2), the left end of *cag* (LF1 and LF2), and the promoter region of *cagA* (AP-F1 and AP-F2) with three randomly selected strains that either had a rearranged *cag*-PAI (CR 7 from Costa Rica [lanes 1] and P190 from northern India [lanes 3]) or completely deleted *cag* PAI (D185B from India [lanes 2]). Strain CR7 showed deletions in AP-F1, whereas strain P190 had only the *cagE* and LF2 regions intact. M, 100-bp molecular size marker with bands ranging from 100 to 1,000 bp. The corresponding PCR product sizes for *cag* PAI genes are summarized in Table 1.

cag-PAI, namely, JHP520, -521, and -522 (annotated in the J99 genome), coding for the predicted ribose 5-phosphate isomerase and a membrane protein. In addition, the promoter region of the *cagA* gene was also disrupted in the majority of the isolates. Within the *cagA* gene, deletions were more frequently observed involving the 3′ region of the gene corresponding to *H. pylori* J99 ORF JHP0495 than at the 5′ end of this ORF. A very high amount of nucleotide diversity has already been reported in this region (25). All of the results corresponding to rearrangement patterns of *cag*-PAI and the constituent genes are shown in Fig. 1 and 3.

Biogeographic distribution of *cag*-PAI rearrangements. Our analyses revealed that the highest proportion of the strains from Japan (57.1%) were harboring an intact *cag*-PAI. A good fraction of Peruvian strains (18.6%), carried an intact *cag*-PAI. Only 12% of the 63 isolates recovered from northern, southern, and western parts of India had an intact *cag*-PAI. Fifty percent of the strains from West India (Pune) were devoid of the entire island of genes analyzed. Three of the ten (30%) strains from far-northern India (Ladakh) had an intact *cag*-PAI. Frequency of rearrangements in the *cagA* gene and the LEC region was higher for Indian isolates from Pune and Delhi than for isolates from Hyderabad and Ladakh. Conservation of *cag*-PAI from strains in Europe and Africa was minimum. Majority of strains from Ireland were found to be rearranged except for the 15.8% strains that carried an intact *cag*-PAI. Only 2 of the 66 English strains analyzed had an intact *cag*-PAI. None of the strains from Costa Rica had intact *cag*-PAI. Strains with intermediate genotype (deletions within the *cag*-PAI) were more common compared to strains with intact *cag*-PAI in all of the geographic regions studied. *cagA* gene was present in 72.8% of strains world over while both *cagE* and *cagT* genes were present in 82% of these isolates, indicating that the presence of *cagA* gene alone is not a marker for an intact *cag*-PAI. Conservation of *cagE* region was 100% in Japanese and Spanish strains and lowest in Ireland. Conservation of *cagA* promoter region was also highest among Japanese strains and was lowest in Costa Rican and Indian strains. The LEC region was found to be more frequently rearranged in isolates from Costa Rica than in other isolates in the study. The extreme right end of the *cagA* gene was found to be frequently rearranged in all of the African isolates studied compared to all other isolates. However, the proximal portion of this gene was present in 100% of the African isolates, the

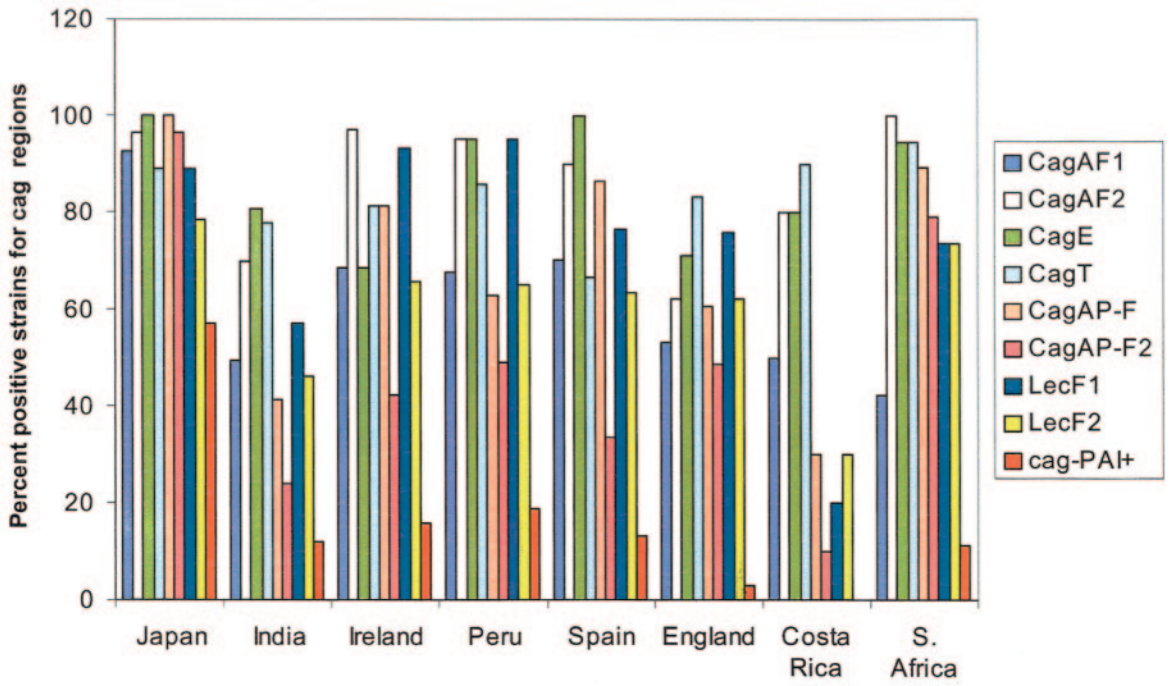


FIG. 3. Deletion (or rearrangement) analysis of the *cag*-PAI in strains belonging to eight different geographic regions. Values on *y* axes denote the percentage of strains with intactness in the genomic region analyzed.

highest level among all of the isolates analyzed from different regions. In all, a total of 33.33% of the strains representing Asian gene pool (Japan, native Peruvian, and Ladakh) were intact with respect to all of the genomic landmarks analyzed. Conversely, only 10.25% of the strains representative of all other gene pools (Indo-Europe, African, and American) were found to be harboring an intact PAI.

Corroboration with clinical outcome. Although no association could be discerned between the presence of *cagA* gene and the clinical outcome, the deletion frequency of both *cagE* and *cagT* genes was higher in nonulcer dyspepsia than in cases of ulcer and cancer. *cagA*, *cagE*, and *cagT* were found to be more frequently rearranged in isolates linked to rather benign infections than in isolates recovered from patients suffering from severe ulcers and gastric cancers worldwide. The *cagA* gene was intact in 60% of the isolates from benign cases, whereas it was intact in >82% of the isolates recovered from patients with ulcers and cancers. The *cagT* gene was conserved in 100 and 80% of the isolates analyzed from cases of severe gastroduodenal pathology and gastritis, respectively. Conversely, the *cagA* promoter and LEC regions were rearranged more frequently in isolates linked to severe pathology than isolates responsible for a more benign outcome worldwide (Fig. 4). Many of the Japanese and Peruvian strains linked to nonulcer gastritis were also found to be intact with respect to all of the regions analyzed. Surprisingly, none of the five cancer strains from Costa Rica were found to be intact.

DISCUSSION

Pathogenic *H. pylori* may have evolved from its benign counterparts by acquiring large amounts of genetic information in the form of PAIs that would have conferred upon it an increased fitness for colonizing new host niches (19). The ancient circumstances associated with such acquisition and the impact of the same on human gastric disease epidemiology is a subject of intense argument. Our study sought to perform an extensive survey of the genetic rearrangements within the *cag*-PAI by using a large number of *H. pylori* strains from across the globe. It appears that rearrangement in *cag*-PAI is quite a prevalent phenomenon, and its constituent genes are under selection pressure more than any other regions in the chromosome. A recent study, in which paired antral and corpus isolates differed with respect to polymorphism of *cagA* locus, supports this idea (9). However, these paired isolates were identical with respect

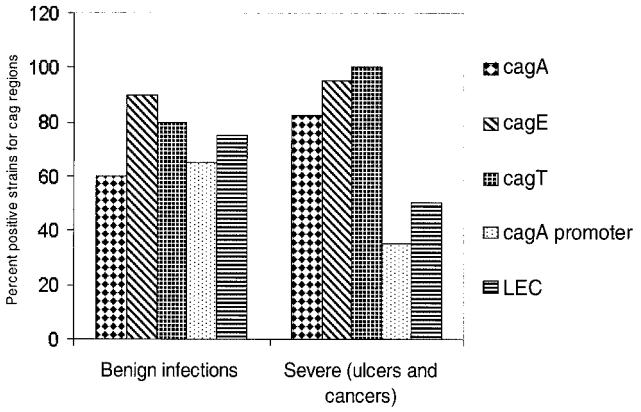


FIG. 4. Deletion of the *cag*-PAI regions with respect to disease condition. Values on the *y* axis denote the percentages of strains with intactness in the genomic region analyzed.

to other important loci, such as the *vacA*, *ribAP*, *vacAP*, and *cagA-glr* motifs, ~50 fluorescent amplified fragment length polymorphism loci, and many RAPD [random(ly) amplified polymorphic DNA] loci (9). Mechanisms underlying such rearrangements are unknown. Insertion elements have been reported to cause rearrangements in the *H. pylori* genome (3, 50). The core sequence of the 31-bp direct repeat flanking the *cag*-PAI has sequence homology with the flanking ends of IS*605* element (11). Insertion of the latter in functionally silent regions of PAIs has been reported (20, 42). Strains lacking the *cag*-PAI contained the IS*605* (5, 23, 28), although its location within the *cag*-PAI could not be demonstrated. In contrast, certain studies have failed to identify IS*605* elements within the *cag*-PAI and their association to the strains with rearranged genes (23, 28, 37, 47). Apart from this, one can also guess the involvement of certain dedicated gene products (such as DNA helicases), encoded by chromosomal regions flanking the *cag*-PAI, in propagating the island (16). However, in our study the rearrangement profiles obtained for regions within the *cag*-PAI were independent of the insertion deletion and substitution activities (in-del-sub) involving putative helicase (*hel*), IS*605* and IS*606* at the extreme right end of the *cag*-PAI (data not shown). These in-del-sub profiles have been previously reported for a majority of isolates from our collection (10, 21, 25).

Genotype-phenotype correlation in case of *cag*-PAI genotypes is a difficult exercise, although there are reports describing the association between the presence of the *cag*-PAI and disease type (5, 6, 34). The pathogenic role of the *cag*-PAI as a whole or in part, in disease development, is not yet completely understood (8, 12, 32, 39). *H. pylori* strains differing in virulence potentials can colonize in a mixed infection and emergence of recombinant strains, with deletion of *cag*-PAI, cannot be ruled out. It is therefore possible to isolate strains with rearranged *cag*-PAI from gastric ulcer and cancer patients. Isolates linked to severe pathology (including gastric cancers) in our study did not reveal an intact *cag*-PAI. In case of chronic infections, It is possible that strains with intact *cag*-PAI may subsequently rearrange the island long after the initial damage (leading to severe pathology after many years). On the basis of our analyses, the *cag*-PAI was found to be most intact in Japanese strains. This is in agreement with reports of an increased occurrence of gastric cancer and augmented severity of gastric ulcers in this part of the world (27, 31, 44). However, it will be premature to link disease outcome to rearrangement patterns within *cag*-PAI unless a large number of strains with precisely defined clinico-epidemiological and pathological features are analyzed. Also, the *cag*-PAI may not be the principal virulence factor, as suggested by the absence or sporadic distribution of the *cag*-PAI genes among strains from varied clinical outcome. This may be due to the fact that the development of ulcer disease is a complex process that also involves factors other than the *cag*-PAI, such as the constituents of the outer membrane, *vacA*, and the nuclear activating protein, etc. Nonetheless, the PCR-based rearrangement profiling described here could provide the clinician with an educated guess on the present proinflammatory power of the infecting strain. This will be immensely helpful in case of patients colonized only with a single strain and not undergoing eradication therapy. The genotypic signatures thus developed may also help track the routes of infection in communities and may be used as reference material in the form of strain databases. Finally, the present study of *H. pylori cag*-PAI genotypes from different human populations adds to our understanding of bacterium-host interactions and the ecology of the bacterium in mixed infections. It will also be helpful to confirm the importance of *H. pylori* geographical genomics in studying gastroduodenal pathology in humans.

ACKNOWLEDGMENTS

We thank Seyed E. Hasnain for his patronage, encouragement, guidance, and support. We thank Douglas E. Berg (Washington University, St. Louis, Mo.) for sharing some genomic DNA of *H. pylori* isolates from his collection of strains. We thank Robert H. Gilman and John Atherton for DNA samples from Peruvian and English strains, respectively. We thank Martin Blaser, Richard Novick, and Mark Achtman for allowing us to use Ladakh isolates included in this study and Cyril J. Smyth for his patronage. We also thank G. Balakrish Nair (ICDDRB, Bangladesh) for guidance.

This study was funded in part by a grant from the Department of Biotechnology, Government of India, to N.A., C.M.H., B.D., and Y.S. I.M.C. was a visiting scientist at DCMS and CDFD and was supported by the Irish Health Research Board. N.A. is a staff scientist and currently leads the pathogen evolution program at CDFD.

REFERENCES

1. **Ahmed, N., A. A. Khan, A. Alvi, S. Tiwari, C. S. Jyothirmayee, F. Kauser, M. Ali, and C. M. Habibullah.** 2003. Genomic analysis of *Helicobacter pylori* from Andhra Pradesh, South India: molecular evidence for three major genetic clusters. Curr. Sci. **85:**101–108.
2. **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.
3. **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. de Jonge, G. Carmel, P. J. Tummino, A. Caruso, M. U. 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.
4. **Atherton, J. C., T. L. Cover, R. J. Twells, M. R. Morales, C. J. Hawkey, and M. J. Blaser.** 1999. Simple and accurate PCR-based system for typing vacuolating cytotoxin alleles of *Helicobacter pylori.* J. Clin. Microbiol. **37:**2979–2982.
5. **Audibert, C., C. Burucoa, B. Janvier, and J. L. Fauchère.** 2001. Implication of the structure of the *Helicobacter pylori cag* pathogenicity island in induction of interleukin-8 secretion. Infect. Immun. **69:**1625–1629.
6. **Backert, S., T. Schwarz, S. Miehlke, C. Kirsch, C. Sommer, T. Kwok, M. Gerhard, U. B. Goebel, N. Lehn, W. Koenig, and T. F. Meyer.** 2004. Functional analysis of the *cag* pathogenicity island in *Helicobacter pylori* isolates from patients with gastritis, peptic ulcer, and gastric cancer. Infect. Immun. **72:**1043–1056.
7. **Blaser, M. J.** 1999. Allelic variation in *Helicobacter pylori*: progress but no panacea. Gut **45:**477–482.
8. **Blaser, M. J., and D. E. Berg.** 2001. *Helicobacter pylori* genetic diversity and risk of human disease. J. Clin. Investig. **107:**767–773.
9. **Carroll, I. M., N. Ahmed, S. M. Beesley, A. A. Khan, S. Ghousunnissa, C. A. O'Morain, C. M. Habibullah, and C. J. Smyth.** 2004. Microevolution between paired antral and paired antrum and corpus *Helicobacter pylori* isolates recovered from individual patients. J. Med. Microbiol. **53:**1–9.
10. **Carroll, I. M., N. Ahmed, S. M. Beesley, A. A. Khan, S. Ghousunnissa, C. A. O'Morain, and C. J. Smyth.** 2003. Fine-structure molecular typing of Irish *Helicobacter pylori* isolates and their genetic relatedness to strains from four different continents. J. Clin. Microbiol. **41:**5755–5759.
11. **Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiara, M. Borodovsky, R. Rappuoli, and A. Covacci.** 1996. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type-specific and disease associated virulence factors. Proc. Natl. Acad. Sci. USA **93:**14648–14653.
12. **Covacci, A., J. L. Telford, G. D. Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328–1333.
13. **Cover, T. L., and M. J. Blaser.** 1999. *Helicobacter pylori* factors associated with disease. Gastroenterology **117:**257–261.
14. **Crabtree, J. E., Z. Xiang, I. J. Lindley, D. S. Tompkins, R. Rappuoli, and A. Covacci.** 1995. Induction of interleukin-8 secretion from gastric epithelial

cells by a *cagA*-negative isogenic mutant of *Helicobacter pylori.* J. Clin. Pathol. **48:**967–969.
15. **El-Omar, E. M., M. Carrington, W. H. Chow, K. E. McColl, J. H. Bream, H. A. Young, J. Herrera, J. Lissowska, C. C. Yuan, N. Rothman, G. Lanyon, M. Martin, J. F. Fraumeni, and C. S. Rabkin.** 2000. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature **404:**398–402.
16. **Firth, N., K. I. Ihler, and R. A. Skurray.** 1996. Structure and function of the F factor and mechanism of conjugation, p. 2377–2401. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
17. **Fischer, W., P. L. Jürgen, R. Buhrdorf, B. Gebert, S. Odenbreit, and R. Haas.** 2001. Systematic mutagenesis of the *Helicobacter pylori cag* pathogenicity island: essential genes for *cagA* translocation in host cells and induction of interleukin-8. Mol. Microbiol. **42:**1337–1348.
18. **Graham, D. Y., and Y. Yamaoka.** 1998. *H. pylori* and *cag*A: relationships with gastric cancer, duodenal ulcer, and reflux esophagitis and its complications. Helicobacter **3:**145–151.
19. **Groisman, E. A., and H. Ochman.** 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87:**791–794.
20. **Hacker, J., and J. B. Kaper.** 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. **54:**641–679.
21. **Hussain, M. A., F. Kauser, A. A. Khan, S. Tiwari, C. M. Habibullah, and N. Ahmed.** 2004. Implications of molecular genotyping of *Helicobacter pylori* isolates from different human populations by genomic fingerprinting of enterobacterial repetitive intergenic consensus regions for strain identification and geographic evolution. J. Clin. Microbiol. **42:**2372–2378.
22. **Ikenoue, T., S. Maeda, K. O. Gura, M. Akanuma, Y. Mitsuno, Y. Imai, H. Yoshida, Y. Shiratori, and M. Omata.** 2001. Determination of *Helicobacter pylori* virulence by simple gene analysis of the *cag* pathogenicity island. Clin. Diagn. Lab. Immunol. **8:**181–186.
23. **Jenks, P., F. Megraud, and A. Labigne.** 1998. Clinical outcome after infection with *Helicobacter pylori* does not appear to be reliably predicted by the presence of any of the genes of the *cag* pathogenicity island. Gut **43:**752–758.
24. **Kawamura, O., M. Murakami, O. Araki, T. Yamada, S. Tomizawa, Y. Shimoyama, K. Minashi, M. Maeda, M. Kusano, and M. Mori.** 2003. Relationship between gastric disease and deletion of *cag* pathogenicity island genes of *Helicobacter pylori* in gastric juice. Dig. Dis. Sci. **48:**47–53.
25. **Kersulyte, D., A. K. Mukhopadhyay, B. Velapatiño, W. W. Su, Z. J. Pan, C. Garcia, V. Hernandez, Y. Valdez, R. S. Mistry, R. H. Gilman, Y. Yuan, H. Gao, T. Alarcón, M. López-Brea, G. B. Nair, A. Chowdhury, S. Datta, M. Shirai, T. Nakazawa, R. Ally, I. Segal, B. C. Y. Wong, S. K. Lam, F. Olfat, T. Borén, L. Engstrand, O. Torres, R. Schneider, J. E. Thomas, S. Czinn, and D. E. Berg.** 2000. Differences in genotypes of *Helicobacter pylori* from different human populations. J. Bacteriol. **182:**3210–3218.
26. **Kidd, M., A. J. Lastovica, J. C. Atherton, and J. A. Louw.** 2001. Conservation of the *cag* pathogenicity island is associated with *vacA* alleles and gastroduodenal disease in South African *Helicobacter pylori* isolates. Gut **49:**11–17.
27. **Machida, M. A., S. Sasazuki, M. Inoue, S. Natsukawa, K. Shaura, Y. Koizumi, Y. Kasuga, T. Hanaoka, and S. Tsugane.** 2004. Association of *Helicobacter pylori* infection and environmental factors in non-cardia gastric cancer in Japan. Gastric Cancer **7:**46–53.
28. **Maeda, S., H. Yoshida, T. Ikenoue, K. Ogura, F. Kanai, N. Kato, Y. Shiratori, and M. Omata.** 1999. Structure of *cag* pathogenicity island in Japanese *Helicobacter pylori* isolates. Gut **44:**336–341.
29. **Marshall, B. J., and J. R. Warren.** 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet **i:**1311–1315.
30. **Meyer-Ter-Vehn, T., A. Covacci, M. Kist, and H. L. Pahl.** 2000. *Helicobacter pylori* activates MAP kinase cascades and induces expression of the proto-oncogenes c-*fos* and c-*jun*. J. Biol. Chem. **275:**16064–16072.
31. **Miwa, H., M. F. Go, and N. Sato.** 2002. *Helicobacter pylori* and gastric cancer: the Asian enigma. Am. J. Gastroenterol. **97:**1106–1112.
32. **Montecucco, C., and R. Rappuoli.** 2001. Living dangerously: how *Helicobacter pylori* survives in the human stomach. Nat. Rev. Mol. Cell. Biol. **2:**457–466.
33. **Nancy, A. L.** 2002. *Helicobacter pylori* and ulcers: a paradigm revisited. [Online.] http://www.faseb.org/opa/pylori/pylori.html.
34. **Nilsson, C., A. Sille′n, L. Eriksson, M.-L. Strand, H. Enroth, S. Normark, P. Falk, and L. Engstrand.** 2003. Correlation between *cag* pathogenicity island composition and *Helicobacter pylori*-associated gastroduodenal disease. Infect. Immun. **71:**6573–6581.
35. **Occhialini, A., A. Marais, R. Alm, F. Garcia, R. Sierra, and F. Mégraud.** 2000. Distribution of open reading frames of plasticity region of strain J99 in *Helicobacter pylori* strains isolated from gastric carcinoma and gastritis patients in Costa Rica. Infect. Immun. **68:**6240–6249.
36. **Ogura, K., M. Takahashi, S. Maeda, T. Ikenoue, F. Kanai, H. Yoshida, Y. Shiratori, K. Mori, K. Mafune, and M. Omata.** 1998. Interleukin-8 production in primary cultures of human gastric epithelial cells induced by *Helicobacter pylori.* Dig. Dis. Sci. **43:**2738–2743.
37. **Owen, R. J., T. M. Peters, R. Varea, E. L. Teare, and S. Saverymuttu.** 2001. Molecular epidemiology of *Helicobacter pylori* in England: prevalence of *cag* pathogenicity island markers and IS*605* presence in relation to patient age and severity of gastric disease. FEMS. Immunol. Med. Microbiol. **30:**65–71.
38. **Peek, R. M., G. Miller, K. Tham, G. Perez-Perez, X. Zhao, J. Atherton, and M. J. Blaser.** 1995. Heightened inflammatory response and cytokine expression in vivo to *cagA1 Helicobacter pylori* strains. Lab. Investig. **71:**760–770.
39. **Peek, R. M., and M. J. Blaser.** 2002. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. Nat. Rev. Cancer **2:**28–37.
40. **Peters, T. M., R. J. Owen, E. Slater, R. Varea, E. L. Teare, and S. Saverymuttu.** 2001. Genetic diversity in the *Helicobacter pylori cag* pathogenicity island and effect on expression of anti-*cag*A serum antibody in UK patients with dyspepsia. J. Clin. Pathol. **54:**219–223.
41. **Ping, H. I., I. Hwang, D. Cittelly, K.-H. Lai, H. M. T. El-Zimaity, O. Gutierrez, J. G. Kim, M. S. Osato, D. Y. Graham, and Y. Yamaoka.** 2002. Clinical presentation in relation to diversity within the *Helicobacter pylori cag* pathogenicity island. Am. J. Gastroenterol. **97:**2231–2238.
42. **Portnoy, D. A., and S. Falkow.** 1981. Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia pestis.* J. Bacteriol. **148:**877–883.
43. **Reyrat, J. M., P. Vladimir, P. Emanuele, M. Cesare Rino, Rappuoli, and J. L. Telford.** 1999. Towards deciphering the *Helicobacter pylori* cytotoxin. Mol. Microbiol. **34:**197–204.
44. **Rozen, P.** 2004. Cancer of the gastrointestinal tract: early detection or early prevention? Eur. J. Cancer Prevent. **13:**71–75.
45. **Segal, E. D., J. Cha, J. Lo, S. Falkow, and L. S. Tompkins.** 1999. Altered states: involvement of phosphorylated *cag*A in the induction of host cellular growth changes by *Helicobacter pylori.* Proc. Natl. Acad. Sci. USA **96:**14559–14564.
46. **Sheu, S. M., B. S. Sheu, H. B. C. Yang Li, T. C. Chu, and J. J. Wu.** 2002. Presence of *iceA1* but not *cagA*, *cagC*, *cagE*, *cagF*, *cagN*, *cagT*, or *orf13* genes of *Helicobacter pylori* is associated with more severe gastric inflammation in Taiwanese. J. Formos. Med. Assoc. **101:**18–23.
47. **Slater, E., R. J. Owen, M. Williams, and R. E. Pounder.** 1999. Conservation of the *cag* pathogenicity island of *Helicobacter pylori*: associations with vacuolating cytotoxin allele and IS*605* diversity. Gastroenterology **117:**1308–1315.
48. **Suerbaum, S., J. M. Smith, K. Bapumia, G. Morelli, N. H. Smith, E. Kunstmann, I. Dyrek, and M. Achtman.** 1998. Free recombination within *Helicobacter pylori.* Proc. Natl. Acad. Sci. USA **95:**12619–12624.
49. **Suerbaum, S., and P. Michetti.** 2002. *Helicobacter pylori* infection. N. Engl. J. Med. **347:**1175–1186.
50. **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, and B. A. Dougherty.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori.* Nature **388:**539–547.
51. **Tummuru, M., S. Sharma, and M. J. Blaser.** 1995. *Helicobacter pylori picB*, a homologue of the *Bordetella pertussis* toxin secretion protein, is required for induction of IL-8 in gastric epithelial cells. Mol. Microbiol. **18:**867–876.