MAJOR ARTICLE

Natural Variation in Populations of Persistently Colonizing Bacteria Affect Human Host Cell Phenotype

Rahul A. Aras,[1,a] Yongchan Lee,[1,a] Sung-Kook Kim,[3] Dawn Israel,[2] Richard M. Peek, Jr.,[2] and Martin J. Blaser[1]

[1]Departments of Medicine and Microbiology, New York University School of Medicine and VA Medical Center, New York; [2]Division of Gastroenterology, Vanderbilt University Medical Center, Nashville, Tennessee; [3]Kyungpook National University Hospital, South Korea

The highly diverse bacterium *Helicobacter pylori,* which persistently colonizes the human stomach, provides models to study the role of genome plasticity in host adaptation. Within *H. pylori* populations from 2 colonized individuals, intragenomic recombination between *cagA* DNA repeat sequences leads to deletion or duplication of tyrosine phosphorylation sites in the CagA protein, which is injected by a type IV secretion system into host cells. Experimental coculture of gastric epithelial cells with the strains containing these naturally occurring CagA phosphorylation site variants induced markedly divergent host cell morphologic responses. Mutants were constructed in which a phosphorylation site was either added or deleted in the expressed CagA protein; coculture studies confirmed that the naturally occurring differences in CagA phosphorylation are responsible for the observed phenotypic variation. These findings indicate that within an individual host, intragenomic recombination between *H. pylori* repetitive DNA produces strain variants differing in their signals to host cells.

Prokaryotic organisms, even when closely related, have varied genomes [1, 2], reflecting phenotypic differences that facilitate adaptation to alternative environments [3]. For bacteria that are subject to changing constraints, genomic plasticity leads to a pool of individuals with varied adaptations [2]. Such genetic diversity is generated through multiple mechanisms, including spontaneous point mutations, recombination with other bacterial cells, and intragenomic rearrangements involving mobile gene elements [4] or repetitive DNA sequences [5]; this variation may be spontaneous or programmed [2, 4]. Host selection of variants that are well adapted to particular environmental constraints is one mechanism for regulating host-microbial interactions [6].

Received 26 September 2002; accepted 21 March 2003; electronically published 31 July 2003.

Financial support: National Institutes of Health (grants R01 GM63270, R01 DK58587, and 5T32 AI07180-21, and the National Cancer Institute Cancer Core); Medical Research Service of the Department of Veterans Affairs.

[a] R.A.A. and Y.L. contributed equally to this work.

Reprints or correspondence: Dr. Yongchan Lee, Dept. of Internal Medicine, Div. of Gastroenterology, Yonsei University College of Medicine, Seodaemunku Shinchondong 134, Seoul, Korea 120-752 (leeyc@yumc.yonsei.ac.kr).

The Journal of Infectious Diseases 2003;188:486–96

0022-1899/2003/18804-0002$15.00

Helicobacter pylori, a gram-negative curved bacteria whose presence is associated with gastric cancer and peptic ulcer disease [7], are highly diverse, compared with other studied organisms in the human biosphere [8]. The identification of genetically divergent subclones within individual hosts [5, 9–12] indicates that *H. pylori* diversification continues during its decades-long host colonization and is consistent with the organism existing as a quasi species [11]. A genetically diverse *H. pylori* population provides the host with a repertoire of varied phenotypes from which a subpopulation with optimal fitness may be selected; this strategy allows *H. pylori* to adapt to changing environments and creates a dynamic equilibrium between the microbes and their host [13]. For *H. pylori,* neither spontaneous point mutation nor recombination with other bacterial cells is sufficient to explain the magnitude of observed intrahost genetic variation [9, 10]. Computational analysis of the fully sequenced *H. pylori* strain 26695 identified a large number of direct DNA repeats [14]. Recombination between paired repeats permits deletion or duplication of the intervening DNA segment plus one copy of the repeat [15] and may be a general mechanism through which *H. pylori* and other organisms regulate gene content [5, 9].

Approximately 60% of US *H. pylori* strains contain a ~40-kb pathogenicity island (termed "*cag* island"), which is associated with an increased risk of peptic ulcer disease or distal gastric cancer in hosts carrying the strains [7]. The *cag* island encodes a type IV secretion system [16–19] that injects the *cagA* product into host gastric epithelial cells [20–24], where it undergoes tyrosine phosphorylation by Src-like kinases [25, 26]. The translocated and phosphorylated CagA protein interacts with signal transduction pathway molecules (i.e., SHP-2) [25–28], resulting in host cell morphologic changes [21, 23]. The 3′ portion of *cagA*, which encodes the tyrosine phosphorylation site or sites, is diverse in both size and sequence, in part because of differences in the number of DNA repeats [29, 30]. The presence of repetitive DNA in *cagA* provides a model system to examine the role of intragenomic prokaryotic recombination in promoting diversity in loci critical for specific host cell responses.

To test the hypothesis that intrahost *cagA* diversity exists and leads to injection of divergent CagA proteins into gastric epithelial cells, which results in differential host responses, we examined multiple *H. pylori* isolates from individual hosts. Analyses of 2 pairs of subclones with divergent *cagA* alleles indicates that *H. pylori* can use recombination between direct DNA repeats to delete or duplicate CagA tyrosine phosphorylation sites and that the host cell responses reflect the number of such sites present.

MATERIALS AND METHODS

Bacterial strains. A total of 75 *H. pylori* isolates (62 from 24 people and 13 from a monkey) (table 1) were obtained from the *H. pylori* stock collection from the New York University (NYU) *Helicobacter/Campylobacter* strain reference collection stored at −70°C. These included 34 single-colony *H. pylori* isolates (antrum, 11; corpus, 13; and cardia, 10) obtained via biopsy from 15 patients during upper endoscopy. Sixteen paired isolates were obtained from antral biopsy specimens at both initial and follow-up (7–10 years later) endoscopies of 8 patients from The Netherlands [11]. After 48 h of growth on trypticase soy agar plates, the entire population of bacterial cells recovered from each biopsy was harvested and stored at −70°C. Nashville patient B41 (from whom strain J99 was isolated and then subjected to genomic sequence analysis [17]) underwent a repeat endoscopy 6 years after his initial endoscopy; 12 single-colony isolates were examined, and randomly amplified polymorphic DNA (RAPD) and sequence analyses identified all isolates as being closely related [10]. We also examined strain J166, which had been used for gastric challenge of 4 male rhesus monkeys, and 13 isolates obtained from 1 monkey 10 months after experimental challenge [31]. All single-colony isolates were obtained directly from the biopsy samples and propagated as pools for <5 passages before freezing at −70°C, as described above. All bacterial cells were grown on trypticase soy agar with 5% sheep's blood (BBL) for 1 passage at 37°C in a microaerobic environment for 48 h, and chromosomal DNA was prepared by means of a phenol extraction method [32].

DNA techniques. Polymerase chain reaction (PCR) of the *cagA* 3′ region was performed with primers cag3 (5′-GGAACCCTAGRCGGTAATG) and cag4 (5′-CGATAGACAAGCTCAAAGAT). All PCRs were run for 35 cycles with reaction mixtures containing 100 ng of template DNA, 200 ng of each primer, and 0.5 U of *Taq* polymerase (Qiagen) in a 50-μL volume. Products were electrophoresed on a 2% agarose gel containing 5 μg of ethidium bromide and visualized under UV light. The amplified products were purified with the QIAquick gel extraction kit (Qiagen); the DNA sequence was determined on both strands with an automated Applied Biosystems sequencer in the NYU Cancer Center Core Laboratory and was analyzed with Sequencer 3.1.1 (Gene Code). Examination of restriction fragment–length polymorphism was performed, as described elsewhere [33], on 10-μL aliquots of PCR products that were digested with 10 U of *Alu*I, *Mbo*I, or *Mse*I (New England Biolabs) for 3 h at 37°C, as recommended by the manufacturer. RAPD PCR analysis was performed with primers 1247, 1254, 1283, or D14307, as described elsewhere [34].

AGS cell culture. AGS human gastric epithelial cells (ATCC CRL 1739) were grown in RPMI 1640 medium (Life Technologies), supplemented with 10% fetal bovine serum, in an atmosphere of 5% CO_2 at 37°C. For coculture experiments, *H. pylori* was grown in brucella broth with 5% fetal bovine serum overnight, and cells were suspended in PBS (pH 7.0). *H. pylori* was added to AGS cells at a bacteria-to-cell ratio of 100:1 for all experiments. The morphological changes associated with the scatter factor–like ("hummingbird") phenotype characterized by spreading and elongated growth of the cell, as described elsewhere [23], were enumerated in 3 different fields of duplicate coculture plates by 2 independent observers blinded to the experimental conditions. Statistical analyses was performed by 2-tailed Student's *t* test in Microsoft Excel.

Immunoblot analysis. AGS cells (5×10^5) were cocultured for 4 h with *H. pylori* cells at a bacteria-to–AGS cell multiplicity of 100. Nonadherent cells were removed by washing 5 times with PBS. Cells were harvested in 1 mL of ice-cold PBS* (PBS, 1 mmol/L EDTA, 1 mmol/L orthovanadate, 1 mmol/L phenylmethylsulfonyl fluoride, 1 μmol/L leupeptin, and 1 μmol/L pepstatin) by cell scrapers, collected by centrifugation, and resuspended in 50 μL of PBS*. For SDS-PAGE, an equal volume of 2× sample buffer (125 mmol/L Tris [pH 6.8], 2% β-mercaptoethanol, 4% SDS, 10% glycerol, and 0.006% bromophenol blue) was added to the sample before boiling for 5 min. Proteins were separated on a 6% polyacrylamide gel, electrotransfered to a polyvinylidene diflouride membrane, and examined for

Table 1. Characteristics of *Helicobacter pylori* strains studied for variation in *cagA*.

Designation of host source	No. of related isolates examined	Strain origin[a]	Isolation of single colony or entire population	Interval between samplings, years	Reference
147	2	Antrum, corpus	S	NA	This study
258	2	Antrum, corpus	S	NA	This study
263	2	Antrum, cardia	S	NA	This study
265	2	Corpus, cardia	S	NA	This study
266	2	Antrum, corpus	S	NA	This study
267	2	Antrum, cardia	S	NA	This study
269	2	Antrum, corpus	S	NA	This study
270	3	Antrum, corpus, cardia	S	NA	This study
284	2	Antrum, corpus	S	NA	This study
289	3	Antrum, corpus, cardia	S	NA	This study
291	3	Antrum, corpus, cardia	S	NA	This study
292	2	Corpus, cardia	S	NA	This study
293	2	Antrum, corpus	S	NA	This study
294	3	Antrum, corpus, cardia	S	NA	This study
295	2	Corpus, cardia	S	NA	This study
J99	14	Initial = antrum	EP	6	[10]
		Follow-up = 5 antrum, 5 corpus, 1 cardia, 2 duodenum	S		
4qs	2	Initial = antrum	EP	8.7	[11]
		Follow-up = antrum	EP		
6qs	2	Initial = antrum	EP	8.1	[11]
		Follow-up = antrum	EP		
7qs	2	Initial = antrum	EP	7	[11]
		Follow-up = antrum	EP		
8qs	2	Initial = antrum	EP	9.8	[11]
		Follow-up = antrum	EP		
10qs	2	Initial = antrum	EP	10.2	[11]
		Follow-up = antrum	EP		
11qs	2	Initial = antrum	EP	7.8	[11]
		Follow-up = antrum	EP		
12qs	2	Initial = antrum	EP	9	[11]
		Follow-up = antrum	EP		
13qs	2	Initial = antrum	EP	7.4	[11]
		Follow-up = antrum	EP		
Rhesus	14	Initial = J166 prechallenge strain	EP	0.8	[31]
		Follow-up = 3 antrum	S		

NOTE. EP, sweep of entire *H. pylori* population cultured from biopsy sample; NA, not applicable because the sample was obtained during a single endoscopic procedure; S, single-colony isolates of *H. pylori* cultured from biopsy sample.

[a] Origin of biopsy sample from which strain was isolated.

the presence of CagA by polyclonal rabbit α-CagA IgG [35] or α-phosphotyrosine (PY20; Transduction Laboratories). The horseradish peroxidase–conjugated antibody was visualized by enhanced chemiluminescence (Amersham Pharmacia Biotech), as recommended by the manufacturer.

Assessment of AGS cell apoptosis and interleukin (IL)–8 induction. AGS cell apoptosis and IL-8 induction were measured as described elsewhere [36]. In short, AGS cells (5×10^5) were cocultured for 4 h with *H. pylori* cells at a bacteria-to–AGS cell multiplicity of 100. DNA fragmentation, reflecting apoptosis, was quantified with a commercially available ELISA (Roche Molecular Biochemicals). For quantification of IL-8, AGS cell monolayers in 6-well plates were cocultured with or without *H. pylori* for 24 h and supernatants removed from the wells were centri-

fuged at 15,000 *g*. IL-8 protein was measured by ELISA (R&D Systems) on supernatants.

Complementation of the* H. pylori *strain 147A*cagA *tyrosine phosphorylation site and deletion of the strain 147C CagA tyrosine phosphorylation site. To reintroduce the deleted 102-bp *cagA* region into *H. pylori* strain 147A, the right-junction of the *cag* island from strain 147C (containing an intact phosphorylation site) was amplified by primers cag3 and glr (5′-CAAGCTAAAGAATGGCTCAAATTG) and cloned into pGEM-T-EZ (Promega) to create p147C. Sequence analysis of the 2-kb insert identified a unique *Cla*I recognition site (1158 bp downstream of the region deleted in strain 147A) that was used to introduce a *Cla*I-digested kanamycin resistant *aphA* (kanamycin-resistance) cassette [37], thereby creating p147CKO. To confirm correct insertion of the *aphA* cassette, p147CKO was digested with *Cla*I, and products were electrophoresed on a 1% agarose gel and visualized under ultraviolet light. Subsequently, *H. pylori* strain 147AP was created by p147CKO to transform *H. pylori* strain 147A, as described elsewhere [38]; transformants were selected on brucella broth plates containing 10% newborn calf serum and 25 μg/mL kanamycin (BBNK). Chromosomal DNA was prepared from all transformants, and PCR with primers cag3 and cag4 was performed to confirm restoration of the 102-bp DNA fragment into strain 147A.

We next sought to create strain 147CP, a 147C derivative, in which the repeat containing its single putative CagA tyrosine phosphorylation site was experimentally deleted. To accomplish this, DNA from a transformant of strain 147AP that acquired an *aphA* cassette but did not have restoration of the 102-bp fragment (strain 147AP-*aphA*) was used as template for PCR with primers cag3 and glr. The resulting 5-kb PCR product was used to transform strain 147C, and transformants selected on BBNK plates. Chromosomal DNA was prepared from all transformants and PCR with primers cag3 and cag4 performed to confirm deletion of the 102-bp DNA fragment containing the CagA tyrosine phosphorylation site.

RESULTS

Identification of* H. pylori *subclones divergent in* cagA *size. To examine intrahost *cagA* diversity, PCR with the cag3/cag4 primers that flank the *cagA* 3′ repeat region was performed on 75 isolates obtained from 25 individual hosts (24 human and 1 rhesus; table 1). From initial and follow-up gastric biopsy samples in 10 hosts, a total of 41 isolates were obtained 10 months to 10 years apart. From the other 15 hosts, a total of 34 isolates were obtained from the gastric antrum, corpus, and/or cardia during a single endoscopy. From 2 individuals, paired strains yielded *cagA* PCR products that differed in size. Strains 147A (antrum) and 147C (corpus), obtained during a single endoscopy, yielded products of ~500 and 600 bp (figure 1*A*), respectively. RAPD and restriction fragment–length polymorphism analyses yielded identical profiles for the 2 isolates, indicating their origin from a common ancestor (figure 1*B*). From strains 7aqs and 7bqs, obtained 7 years apart from the same individual [11], products of ~750 and 700 bp were amplified, respectively (figure 2*A*). RAPD, amplified fragment–length polymorphisms, and sequence analysis indicated that these isolates also are clonal variants [11]. These results thus indicate that, within an individual host, *H. pylori* subpopulations may vary in *cagA* alleles, distinguishable by their size differences.

Sequence analyses of* cagA *from paired subclones. To more closely examine the divergence between the 2 pairs of subclones, sequences of the *cagA* PCR products from strains 147A and 147C and strains 7aqs and 7bqs were determined. Comparison of the 147A (498 bp) and 147C (600 bp) PCR products revealed that the latter contained a 102-bp insertion, which included a 33-bp segment identical to the sequence immediately downstream (figure 1*C*). Thus, *cagA* from strain 147C contained a 69-bp region that was flanked by 33-bp direct repeats, and this 102-bp segment (69 + 33) was absent from 147A. Translation of the 147C product identified a single putative (SVSPEPIYA) tyrosine phosphorylation site [25, 26] that was absent in strain 147A (figure 1*D*). Sequence comparison of the 7aqs (756 bp) and 7bqs (714 bp) PCR products showed that the latter contains a 84-bp duplication, preceded by 25-bp direct DNA repeats (with 22 bp of identity), followed by a 126-bp deletion flanked by 24-bp direct DNA repeats (figure 2*B*). Analyses of the *cagA*-translated products indicate that 7aqs contains 3 putative phosphorylation sites, whereas, in 7bqs, one of the predicted phosphorylation sites has been deleted; in both genes, the open-reading frame proceeds to the same terminus (figure 2*C*). In each pair of strains, the *cagA* sequences are otherwise 99.9% identical within a pair; the high degree of identity supports the hypothesis that the observed indels result from intragenomic rearrangements and not from horizontal acquisition from another *H. pylori* strain [5, 33], which would differ in nucleotide sequence by ~6% on average [17].

Intragenomic recombination results in deletion of CagA phosphorylation sites. To determine whether the DNA rearrangements identified in the naturally occurring paired subclones (figure 1 and figure 2) affect intracellular CagA phosphorylation, AGS gastric epithelial cells were cocultured with these *H. pylori* strains (or with controls), and CagA translocation and phosphorylation were examined by immunoblot that used α-CagA and α-pTyr antibodies. As expected, no products consistent with intracellular CagA presence were identified in immunoblots performed on AGS cell lysates cocultured with *cag*-negative strain 88-22 or without coculture (figure 3*A*), thus confirming the specificity of the antibodies used. For the AGS cells cocultured with the paired *H. pylori* strains, immunoblots that used α-CagA antibodies identified ~130-kDa products in-

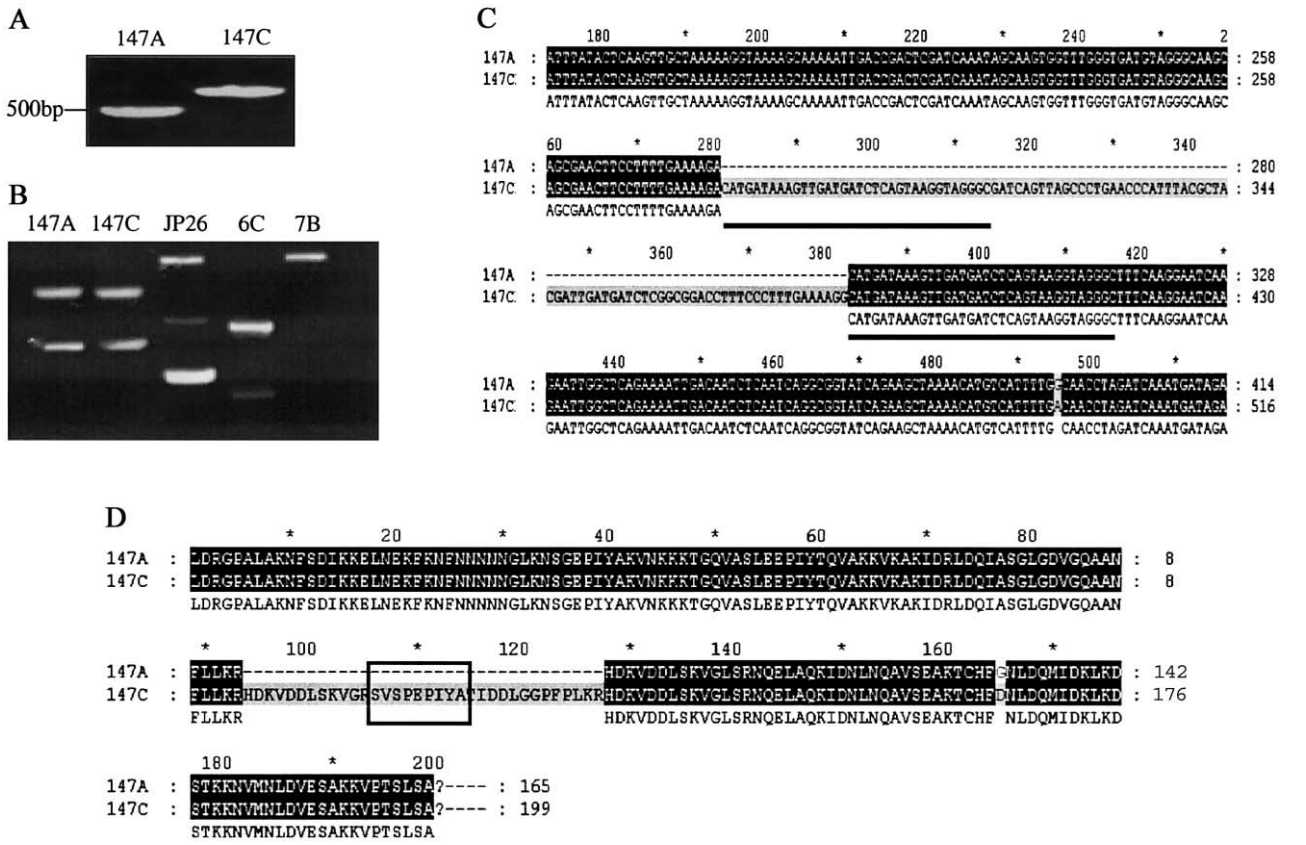

Figure 1. Paired *Helicobacter pylori* subclones isolated from an individual host differ in presence of a 102-bp *cagA* region encoding a putative tyrosine phosphorylation site. *A,* Polymerase chain reaction (PCR) with primers cag3 and cag4 identified 2 subclones, obtained from an individual host (147A [from the antrum] and 147C [from the corpus]), that differed in *cagA* allele size. *B,* Randomly amplified polymorphic DNA analysis with primer 1247 showed subclones 147C and 147A to be identical and easily distinguishable from 3 unrelated strains. Similar results were observed with 3 other primers (data not shown). *C,* Sequence analysis of the 2 PCR products identified a 102-bp deletion flanked by 33-bp identical repeats *(underlined)* in isolate 147A, compared with 147C. *D,* Translation of nucleotide sequences of 147C and 147A identified a predicted phosphorylation site *(boxed)* [21] in isolate 147C that was deleted in 147A.

dicating the cell association of the CagA protein produced by each of the 4 strains, as expected [19–24]; the size differences were also expected on the basis of the sequence variation.

Immunoblots of the AGS cells cocultured with strain 7aqs or 7bqs using α-pTyr antibodies indicate that, for both strains, the translocated CagA is phosphorylated, which is consistent with these proteins containing 3 and 2 phosphorylation sites, respectively. For the AGS cells cocultured with strain 147C or 147A, the α-pTyr antibodies recognized the CagA protein from 147C, but not 147A, indicating phosphorylation of the former, but not the latter, protein (figure 3*B*), which is consistent with the *cagA* sequence data from these 2 subclones (figure 1*C*). These results indicate that natural variation within a host resulting from intragenomic recombination in the *cagA* 3′ region affects intracellular CagA phosphorylation status.

Differential host cell responses to coculture with* H. pylori *subclones differing in CagA structure. Translocation of CagA into AGS cells leads to marked changes in host cell physiology, including apoptosis [36], cytokine production [16], and morphology [21–23]. Histopathological examination of patient B147, a 58-year-old white man from whom strains 147A and 147C were isolated and in whom gastric erythema was visualized during upper gastrointestinal tract endoscopy, revealed an average of 7 and 48 *H. pylori*/high-power field were present in tissue sections from the gastric antrum (from which 147A was isolated) and corpus (from which 147C was isolated), respectively. The presence of 3^+–4^+ acute and chronic inflammation accompanied by 3^+ glandular atrophy in the gastric antrum, and 3^+ acute and chronic inflammation and 2^+ glandular atrophy in the gastric corpus are consistent with differential host responses to colonization by the 147A (antrum) and 147C (corpus) isolates.

To determine which host cell responses are dependent on CagA phosphorylation, we examined apoptosis, IL-8 production, and morphology of AGS cells cocultured with *H. pylori* strains 7aqs, 7bqs, 147A, or 147C. For paired strains 7aqs and

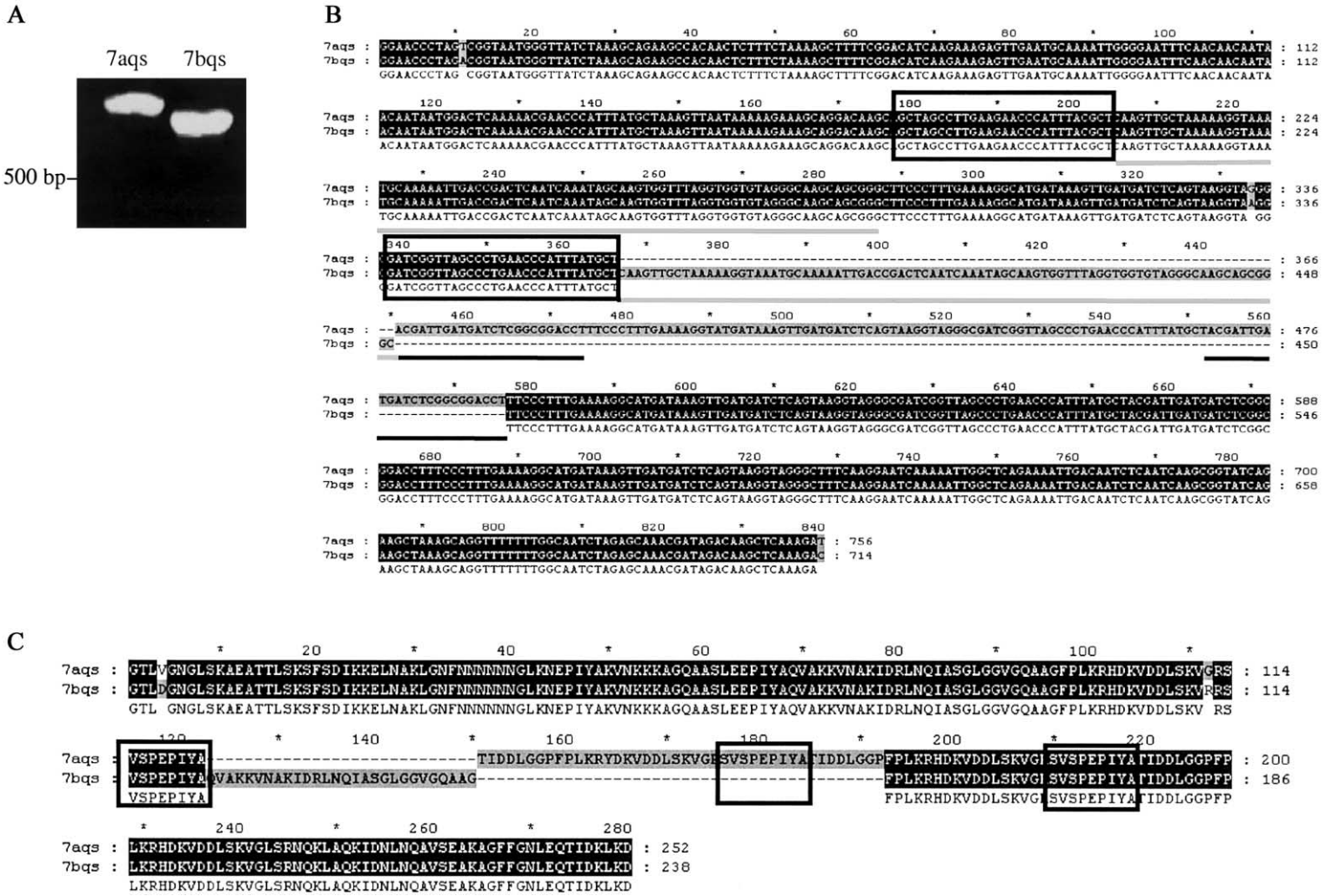

Figure 2. Deletion of a putative *cagA* tyrosine phosphorylation site in strains isolated 7 years apart from the same host. *A,* Polymerase chain reaction (PCR) with primers cag3 and cag4 identified 2 subclones, 7aqs and 7bqs, that differed in *cagA* allele size. *B,* Sequence analysis of the 2 PCR products identified an 84-bp duplication *(gray underline),* followed by a 126-bp deletion in strain 7bqs. The duplicated sequences are preceded by nearly identical 25-bp repeats *(black boxes),* and the deletion is flanked by identical 24-bp repeats *(black underline). C,* Translation of nucleotide sequences of 7aqs and 7bqs identified 3 putative phosphorylation sites *(boxed)* in isolate 7aqs, 1 of which was deleted in isolate 7bqs.

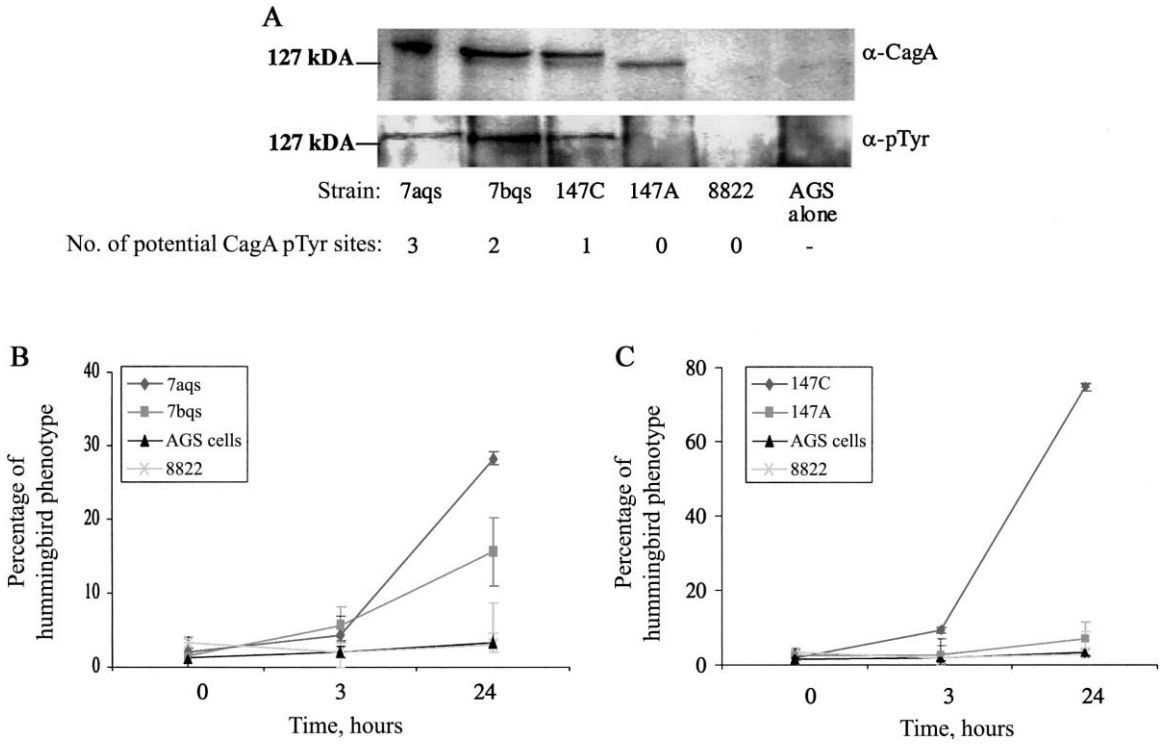

Figure 3. Differences in CagA phosphorylation status affect host cell morphologic responses. *A,* Immunoblots of lysates of AGS cells cocultured with *Helicobacter pylori* cells with α-CagA antibodies. For strains 147C, 147A, 7aqs, and 7bqs, CagA is translocated into gastric epithelial cells. As expected, 88-22 (*cag*-negative) and uninfected AGS cells yield no translocated CagA product, confirming the antibody specificity. Immunoblots performed with anti-pTyr antibodies show that the CagA fragments from 7aqs and 7bqs undergo tyrosine phosphorylation, which is consistent with their having 3 and 2 predicted phosphorylation sites, respectively. Strain 147C, but not 147A, is phosphorylated, which is consistent with deletion of the one predicted phosphorylation site in isolate 147A. As expected, AGS cells alone or those incubated with 88-22 yield no product. *B,* AGS cells cocultured with 7aqs or 7bqs (and controls). The presence of 3 phosphorylation sites (7aqs) is associated with significantly ($P = .02$) increased induction of the host hummingbird response, compared with 2 phosphorylation sites (7bqs). *C,* Strain 147C, containing 1 predicted phosphorylation site, induced a strong hummingbird phenotype at 24 h, whereas there was essentially none for 147A (0 sites); a difference that was significantly different ($P = .007$).

7bqs, although coculture with AGS cells led to significantly greater apoptosis and IL-8 induction than observed in AGS cells cultured alone, there were no significant differences between the 2 strains (figure 4*A*). Parallel results were obtained for paired strains 147A and 147C (figure 4*B*). These findings indicate that CagA phosphorylation status does not affect apoptosis and are consistent with previous studies [39], which indicates that IL-8 induction is CagA independent. That there is no significant difference in IL-8 induction levels between paired strains suggests that all strains are producing a functional type IV secretion system [39].

Host cell morphologic changes are characterized by spreading and elongated growth of the cell and by the presence of lamellipodia and filopodia, a response termed the "hummingbird" phenotype [23]. Next, we examined whether the observed differences in CagA phosphorylation status between the paired subclones affect induction of the hummingbird phenotype. For each pair of strains, after 24 h of coculture, the subclone containing the greater number of predicted phosphorylation sites induced a more pronounced hummingbird phenotype (28.3% for 7aqs vs. 15.7% for 7bqs [$P = .02$] and 74.7% for 147C vs. 7.0% for 147A [$P = .007$]) (figure 3*B* and 3*C*). As expected, control *cag*-negative strain 88-22 induced minimal (3.0%) morphologic changes. That isolate 147A, which contains no predicted CagA tyrosine phosphorylation sites (figure 1), induced a nearly identical level of hummingbird phenotype, compared with strain 88-22, confirms that host cell morphologic changes are dependent on CagA phosphorylation [21]. These results indicate that, within a host carrying *H. pylori,* subclones expressing CagA proteins that differ in phosphorylation status can elicit differential host cell responses.

CagA phosphorylation is necessary to induce host cell morphologic changes. To determine whether the CagA phosphorylation differences between 147C and 147A are sufficient to explain the divergence in observed host cell morphologic changes, by complementation, we restored the deleted tyrosine phosphorylation site encoded by *cagA* from strain 147A to create strain 147AP. Phosphorylation status and hummingbird phenotype induction were examined after coculture of this mutant with AGS cells. Immunoblots of lysates of AGS cells that

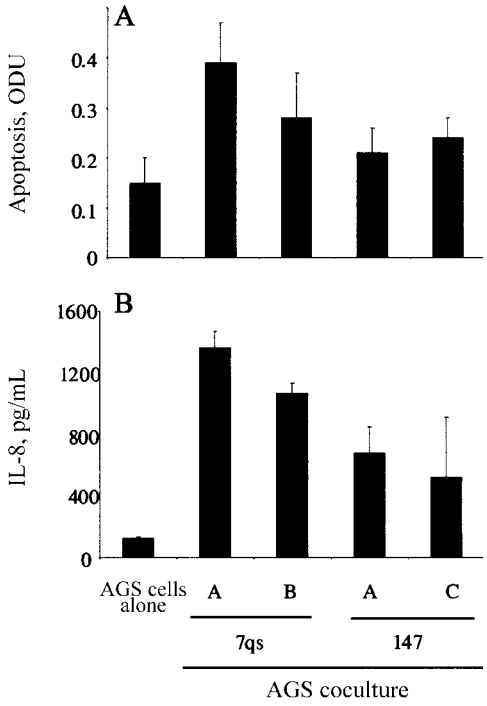

Figure 4. Differences in CagA phosphorylation status do not affect host cell apoptosis or interleukin (IL)–8 production. *A,* To determine whether CagA phosphorylation status of *Helicobacter pylori* cells in coculture with AGS cells affect AGS cell apoptosis, we measured apoptosis induction by ELISA. *H. pylori* strains 7aqs, 7bqs, 147A, and 147C induced greater levels of apoptosis than AGS cells cultured alone; however, there were no significant differences between paired strains 7aqs and 7bqs or 147A and 147C. ODU, optical density units. *B,* To determine whether CagA phosphorylation status of *H. pylori* cells in cocultured with AGS cells affects cytokine production, IL-8 levels were measured by ELISA. *H. pylori* strains 7aqs, 7bqs, 147A, and 147C induced significantly greater levels of IL-8 than did AGS cells cultured alone ($P < .01$); however, there were no significant differences between paired strains 7aqs and 7bqs, or 147A and 147C.

had been cocultured with 147AP using α-CagA or α-pTyr antibodies indicated that the CagA from strain 147AP was translocated into AGS cells and subsequently phosphorylated (figure 5*A*). This result confirms that the 102-bp sequence deleted in strain 147A contained the unique and obligatory tyrosine phosphorylation site. That AGS cells cocultured with complemented strain 147AP developed morphologic changes at a frequency comparable to that of strain 147C but significantly ($P = .007$) higher frequency than that of strain 147A (figure 5*B*) indicates that CagA tyrosine phosphorylation is both necessary and sufficient for hummingbird phenotype induction. For further confirmation, we created an isogenic mutant of strain 147C, termed 147CP, in which the DNA segment encoding the unique tyrosine phosphorylation site was deleted. Coculturing AGS cells with 147CP confirmed the absence of tyrosine phosphorylation (figure 5*A*) and induced the hummingbird morphology (figure 5*A*) at a significantly ($P = .008$) lower frequency than for strain 147C (figure 5*B*). These results confirm that the specific CagA sequences are both necessary and sufficient to induce tyrosine phosphorylation and the resulting host cell morphologic changes [21, 28].

DISCUSSION

H. pylori shows extensive intrahost genetic diversity [5, 9–12], but the loci involved, the mechanisms through which diversity is generated, and the effects on host-microbial interactions are not well studied. Here, we show for the first time that, within an individual host, recombination between *H. pylori* direct DNA repeats resulted in deletion (or duplication) of phosphorylation sites in its CagA protein. We further show that variation in the number of these phosphorylation sites strongly influences the morphologic changes observed after CagA translocation into gastric epithelial cells [23]. These findings are consistent with the hypothesis that *H. pylori* populations exist as a quasi species in a host [11] and that this genetic diversity gives rise to bacterial phenotypes that substantially vary their interactions with their colonized host [13]. That isolates divergent in *cagA* size were isolated from only 2 (8%) of the 25 hosts studied suggests that this variation may be occurring in a minor proportion of hosts carrying *H. pylori cagA*-positive strains; however, because only 2–14 isolates (median, 2 isolates) were examined from each host, it is difficult to determine whether these numbers accurately represent *cagA* allelic variation in vivo. We hypothesize that, if more isolates were examined from each host, we would document a higher rate of carrying strains with differing *cagA* alleles.

Previous reports [40–42] suggest that *H. pylori* is able to invade epithelial cells in the gastric mucosa; however, the role of invasion in *H. pylori* pathogenesis remains unclear. Although electron microscopy, which shows *H. pylori* invasion of epithelial cells, was not performed in this study, our observation that a *cagA*-phosphorylation mutant induced significantly fewer morphologic changes during coculture with AGS cells than did its wild-type counterpart indicates that *cagA* allelic differences, rather than *H. pylori* invasion, are a major determinant of host cell responses.

Although persons carrying *cag*-positive *H. pylori* strains have a higher risk of developing peptic ulceration [43] and of distal gastric cancer [7], compared with persons carrying *cag*-negative strains or no *H. pylori* at all, the biological function of CagA remains unknown. We have hypothesized that coevolution of *H. pylori* and humans has selected for *H. pylori* cells that, when subject to environmental stimuli (e.g., low pH), are able to send signals to the host that are transduced, which then lead to amelioration of the noxious stress [44–46]. The signaling pathway involved with CagA appears to fulfill these conditions.

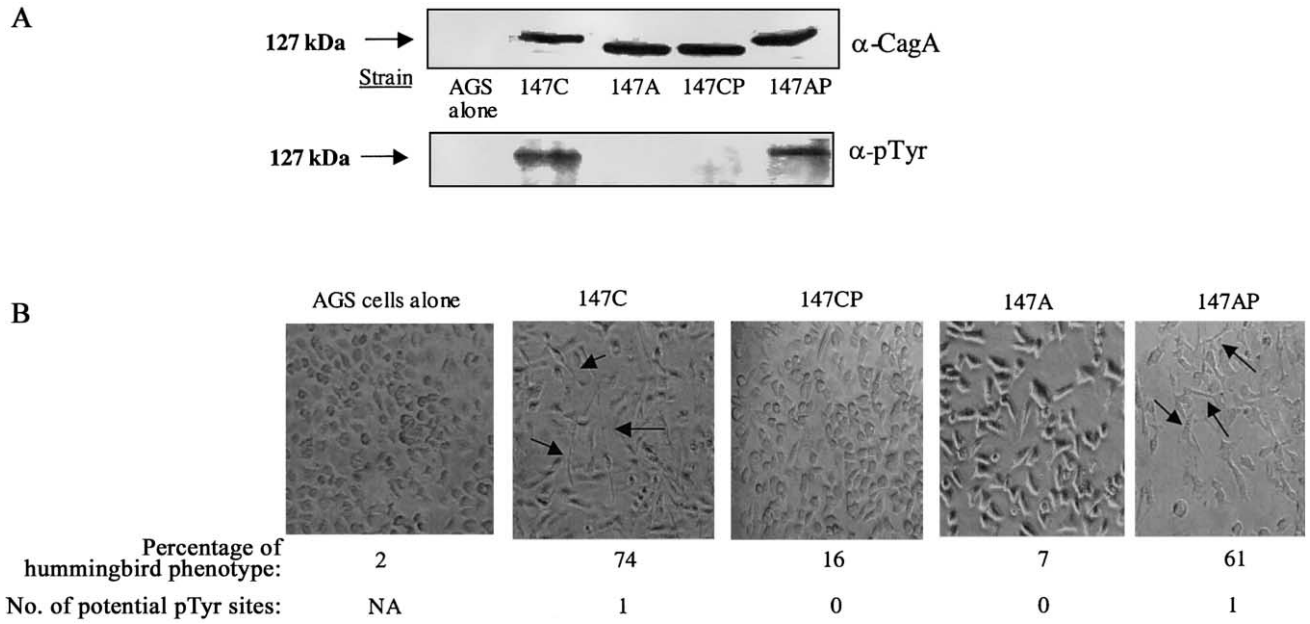

Figure 5. Effects of complementation of the 102-bp deleted region in isolate 147A and deletion of 102-bp region in 147C. *A,* To confirm that the deleted portion in isolate 147A contained a functional phosphorylation site, we created 147AP, a 147A in which the predicted phosphorylation site was restored through complementation, and 147CP, in which the predicted phosphorylation site was deleted. AGS cells were infected with 147AP or 147CP, and immunoblots performed with anti-CagA and anti-pTyr were performed to confirm CagA phosphorylation status. As expected, 147C CagA was phosphorylated after translocation into host cells, but 147A CagA was not. That 147AP CagA was translocated and phosphorylated in host epithelial cells and 147CP was translocated but not phosphorylated confirms that the 102-bp portion deleted in isolate 147A contains a functional phosphorylation site. *B,* Examination of host cell morphology changes after CagA translocation, indicate that 147AP induces the hummingbird phenotype *(arrows)* at a frequency comparable to 147C, but significantly greater than 147A, indicating that CagA phosphorylation is necessary and sufficient for the hummingbird phenotype response.

At low environmental pH, *cagA* transcription is up-regulated [47, 48]. This results in more substrate for type IV system–mediated injection into the gastric epithelium, which would perturb these cells to a greater extent, whether or not they are dependent on tyrosine phosphorylation. The enhanced intraepithelial cell signaling leads to greater induction of IL-1β [49], which is a potent inhibitor of gastric acidity [50]. Because reduced acidity diminishes the stimulus for *cagA* transcription, the proposed relationship is homeostatic through the negative feedback loop described elsewhere [44].

However, that CagA translocation and phosphorylation affect principal host cell signaling pathways [25–28] and induce morphologic changes that may affect cell cycle [50] and death [21, 23] indicate that the overall effects of the interactions are complex and may be host specific. Previous PCR studies have demonstrated that *H. pylori* single-colony isolates obtained from a single biopsy specimen can differ in *cagA* presence [51]. An *H. pylori* population that contains subclones varying in *cagA* alleles with differing potential for phosphorylation provides a repertoire of phenotypically diverse cells subject to host-specific selective pressures. In quasi species, the generation of subclones varying in genotype (and phenotype) permits the host to select the microbial populations most "fit" for colonization [52]. In changing environments, the fitness criteria are dynamic [53], and the selective pressures may differ in distinct host macro- and microniches.

In the patient carrying isolates 147A and 147C, histopathological examination performed on the biopsy samples revealed that more-severe chronic inflammation and atrophic gastritis were present in the antrum than in the corpus. That strain 147C (from corpus), with CagA that is phosphorylated when translocated into epithelial cells, was isolated from a relatively more acidic environment [54] than for strain 147A (isolated from antrum) suggests the hypothesis that CagA phosphorylation may play a role in limiting acid stress. The development of atrophic gastritis is detrimental to *H. pylori* survival in its host [55]; we hypothesize that, within an atrophic (low acid) niche, selection for subclones lacking tyrosine phosphorylation sites might result in less-intense host responses that minimize atrophy progression. Responding to environment differences in ways that increase host tolerance may improve the likelihood for prolonged *H. pylori* colonization [13]. That strain 7aqs, in which CagA contains 3 predicted tyrosine phosphorylation sites, elicited a significantly lower level of host cell hummingbird morphologic changes than did strain 147C, in which the CagA protein contains 1 predicted tyrosine phosphorylation

site, suggests the presence of other influences on hummingbird induction such as CagA expression levels or type IV secretion system efficiency.

Genomic analysis of 2 *H. pylori* strains has identified few of the known effector mechanisms for environmental adaptation (e.g., stringent response and 2-component regulatory systems [17]), which suggests that *H. pylori* differs from other organisms in the regulation of gene expression. Although *H. pylori* cells develop point mutations [56], most nucleotide divergence in *H. pylori* is due to synonymous substitutions [57], which does not influence phenotypic variation. The natural competence of *H. pylori* [58] allows for horizontal acquisition of DNA but requires the presence of multiple *H. pylori* strains [33] and is limited by restriction barriers [59]. Thus, neither point mutation nor recombination with exogenous DNA appears to be sufficient to explain the high level of intrahost diversity observed [10, 11]. However, computational analysis of the strain 26695 and J99 genomic sequences identified large numbers of direct DNA repeats [14] that are widely and nonrandomly distributed throughout the chromosome (R.A.A., J. Kang, A. I. Tschumi, Y. Harasaki, M.J.B., unpublished data). Recombination between direct DNA repeats allows for the deletion or duplication of the intervening sequence plus one copy of the repeat [15] and is used by *H. pylori* to regulate gene content [5, 9]. Identification of paired DNA repeats in known hypervariable regions in other bacterial species [60, 61] suggests that recombination between repetitive DNA may be a general mechanism used by prokaryotes to promote programmed genomic plasticity.

In conclusion, the identification of *H. pylori* subclones within individual hosts that differ in their ability to elicit specific host responses supports the hypothesis that *H. pylori* exists in a dynamic equilibrium with its host [45, 46] and that, by programmed genome plasticity through recombination between DNA repeats, a large number of phenotypically diverse cells are available for host selection. Furthermore, this work supports the suggested hypothesis that, through modification of *cag* island genes, an *H. pylori* strain may generate an array of clones within an individual differing in their virulence [62]. The identification of apparently identical naturally occurring *H. pylori* cells that differ only in CagA phosphorylation status provides a model system for future in vivo studies examining the biological consequences of the particular host-microbial interactions.

References

1. Boucher Y, Nesbo CL, Doolittle WF. Microbial genomes: dealing with diversity. Curr Opin Microbiol **2001**; 4:285–9.
2. Dobrindt U, Hacker J. Whole genome plasticity in pathogenic bacteria. Curr Opin Microbiol **2001**; 4:550–7.
3. Ochman H, Moran NA. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science **2001**; 292:1096–9.
4. Woolhouse ME, Taylor LH, Hayden DT. Population biology of multihost pathogens. Science **2001**; 292:1109–12.
5. Aras RA, Takata T, Ando T, van der Ende A, Blaser MJ. Regulation of the *Hpy*II restriction-modification system of *Helicobacter pylori* by gene deletion and horizontal reconstitution. Mol Microbiol **2001**; 42:369–82.
6. Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, Gordon JI. Molecular analysis of commensal host-microbial relationships in the intestine. Science **2001**; 291:881–4.
7. Peek RM Jr, Blaser MJ. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. Nat Rev Cancer **2002**; 2:28–37.
8. Suerbaum S, Smith JM, Bapumia K, et al. Free recombination within *Helicobacter pylori.* Proc Natl Acad Sci USA **1998**; 95:12619–24.
9. Bjorkholm B, Lundin A, Sillen A, et al. Comparison of genetic divergence and fitness between two subclones of *Helicobacter pylori.* Infect Immun **2001**; 69:7832–88.
10. Israel DA, Salama N, Krishna U, et al. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. Proc Natl Acad Sci USA **2001**; 98:14625–30.
11. Kuipers EJ, Israel DA, Kusters JG, et al. Quasispecies development of *Helicobacter pylori* observed in paired isolates obtained years apart from the same host. J Infect Dis **2000**; 181:273–82.
12. van der Ende A, Rauws EA, Feller M, Mulder CJ, Tytgat GN, Dankert J. Heterogeneous *Helicobacter pylori* isolates from members of a family with a history of peptic ulcer disease. Gastroenterology **1996**; 111:638–47.
13. Blaser MJ, Berg DE. *Helicobacter pylori* genetic diversity and risk of human disease. J Clin Invest **2001**; 107:767–73.
14. Rocha EPC, Danchin A, Viari A. Analysis of long repeats in bacterial genomes reveals alternative evolutionary mechanisms in *Bacillus subtilis* and other competent prokaryotes. Mol Biol Evol **1999**; 16:1219–30.
15. Chedin F, Dervyn E, Dervyn R, Ehrlich SD, Noirot P. Frequency of deletion formation decreases exponentially with distance between short direct repeats. Mol Microbiol **1994**; 12:561–9.
16. Akopyants NS, Clifton SW, Kersulyte D, et al. Analyses of the *cag* pathogenicity island of *Helicobacter pylori.* Mol Microbiol **1998**; 28:37–53.
17. Alm RA, Ling LS, Moir DT, et al. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori.* Nature **1999**; 397:176–80.
18. Censini S, Lange C, Xiang Z, et al. *cag,* a pathogenicity island of *Helicobacter pylori,* encodes type I–specific and disease-associated virulence factors. Proc Natl Acad Sci USA **1996**; 93:14648–53.
19. Tummuru MK, Sharma SA, Blaser MJ. *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 **1995**; 18:867–76.
20. Asahi M, Azuma T, Ito S, et al. *Helicobacter pylori* CagA protein can be tyrosine phosphorylated in gastric epithelial cells. J Exp Med **2000**; 191:593–602.
21. Backert S, Moese S, Selbach M, Brinkmann V, Meyer TF. Phosphorylation of tyrosine 972 of the *Helicobacter pylori* CagA protein is essential for induction of a scattering phenotype in gastric epithelial cells. Mol Microbiol **2001**; 42:631–44.
22. Odenbreit S, Puls J, Sedlmaier B, Gerland E, Fisher W, Haas R. Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. Science **2000**; 287:1497–500.
23. Segal ED, Cha J, Lo J, Falkow S, Tompkins LS. Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori.* Proc Natl Acad Sci USA **1999**; 96:14559–64.
24. Stein M, Rappuoli R, Covacci A. Tyrosine phosphorylation of the *Helicobacter pylori* CagA antigen after *cag*-driven host cell translocation. Proc Natl Acad Sci USA **2000**; 97:1263–8.
25. Selbach M, Moses S, Hauck CR, Meyer TF, Backert S. Src is the kinase of the *Helicobacter pylori* CagA protein in vitro and in vivo. J Biol Chem **2002**; 277:6775–8.
26. Stein M, Bagnoli F, Halenbeck R, Rappuoli R, Fantl WJ, Covacci A. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. Mol Microbiol **2002**; 43:971–81.
27. Higashi H, Tsutsumi R, Muto S, et al. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **2002**; 295:683–6.

28. Puls J, Fischer W, Haas R. Activation of *Helicobacter pylori* CagA by tyrosine phosphorylation is essential for dephosphorylation of host cell proteins in gastric epithelial cells. Mol Microbiol **2002**; 43:961–70.
29. Evans DL Jr, Evans DG. Direct repeat sequences in the *cagA* gene of *Helicobacter pylori:* a ghost of a chance encounter? Mol Microbiol **1997**; 23:409–10.
30. Covacci A, Censini S, Bugnoli M, et al. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci USA **1993**; 90:5791–6.
31. Dubois A, Berg DE, Incecik ET, et al. Host specificity of *Helicobacter pylori* strains and host responses in experimentally challenged nonhuman primates. Gastroenterology **1999**; 116:90–6.
32. Wilson K. Preparation of genomic DNA from bacteria. In: Ausubel FA, Brent R, Kingston RE, et al., eds. Current protocols in molecular biology. New York: Wiley Interscience, **1987**:2.4.1–2.4.2.
33. Kersulyte D, Chalkauskas H, Berg DE. Emergence of recombinant strains of *Helicobacter pylori* during human infection. Mol Microbiol **1999**; 31:31–43.
34. Akopyanz N, Bukanov NO, Westblom TU, Berg DE. PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori.* Nucleic Acids Res **1992**; 20:6221–5.
35. Ando T, Peek RM Jr, Lee YC, Krishna U, Kusugami K, Blaser MJ. Host cell responses to genotypically similar *Helicobacter pylori* isolates from United States and Japan. Clin Diagn Lab Immunol **2002**; 9:167–75.
36. Israel DA, Salama N, Arnold CA, et al. *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. J Clin Invest **2001**; 107:611–20.
37. Cover TL, Tummuru MKR, Cao P, Thompson SA, Blaser MJ. Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains. J Biol Chem **1994**; 269:10566–73.
38. Donahue JP, Israel DA, Peek RM, Blaser MJ, Miller GG. Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori.* Mol Microbiol **2000**; 37:1066–74.
39. Fischer W, Puls J, Buhrdorf R, Gebert B, Odenbreit S, Haas R. Systematic mutagenesis of the *Helicobacter pylori cag* pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. Mol Microbiol **2001**; 42:1337–48.
40. Bjorkholm B, Zhukhovitsky V, Lofman C, et al. *Helicobacter pylori* entry into human gastric epithelial cells: a potential determinant of virulence, persistence, and treatment failures. Helicobacter **2000**; 5:148–55.
41. Amieva MR, Salama NR, Tompkins LS, Falkow S. *Helicobacter pylori* enter and survive within multivesicular vacuoles of epithelial cells. Cell Microbiol **2002**; 4:677–90.
42. Kwok T, Backert S, Schwarz H, Berger J, Meyer TF. Specific entry of *Helicobacter pylori* into cultured gastric epithelial cells via a zipper-like mechanism. Infect Immun **2002**; 70:2108–20.
43. Nomura AM, Perez-Perez GI, Lee J, Stemmermann G, Blaser MJ. Relation between *Helicobacter pylori cagA* status and risk of peptic ulcer disease. Am J Epidemiol **2002**; 155:1054–9.
44. Blaser MJ. Hypotheses on the pathogenesis and natural history of *Helicobacter pylori*–induced inflammation. Gastroenterology **1992**; 102: 720–7.
45. Blaser MJ, Kirschner D. Dynamics of *Helicobacter pylori* colonization in relation to the host response. Proc Natl Acad Sci USA **1999**; 96:8359–64.
46. Kirschner DE, Blaser MJ. The dynamics of *Helicobacter pylori* infection of the human stomach. J Theor Biol **1995**; 176:281–90.
47. Karita M, Tummuru MK, Wirth HP, Blaser MJ. Effect of growth phase and acid shock on *Helicobacter pylori cagA* expression. Infect Immun **1996**; 64:4501–7.
48. Allan E, Clayton CL, McLaren A, Wallace DM, Wren BW. Characterization of the low-pH response of *Helicobacter pylori* using genomic DNA arrays. Microbiology **2001**; 147:2285–92.
49. El-Omar EM. The importance of IL-1β in *Helicobacter pylori* associated disease. Gut **2001**; 48:743–7.
50. Peek RM Jr, Blaser MJ, Mays DJ, et al. *Helicobacter pylori* strain-specific genotypes and modulation of the gastric epithelial cell cycle. Cancer Res **1999**; 59:6124–31.
51. Wirth HP, Yang M, Peek RM Jr, Hook-Nikanne J, Fried M, Blaser MJ. Phenotypic diversity in Lewis expression of *Helicobacter pylori* isolates from the same host. J Lab Clin Med **1999**; 133:488–500.
52. Domingo E. Genetic variation and quasi-species. Curr Opin Genet Dev **1992**; 2:61–3.
53. Webb GF, Blaser MJ. Dynamics of bacterial phenotype selection in a colonized host. Proc Natl Acad Sci USA **2002**; 99:3135–40.
54. Ruiz B, Correa P, Fontham ET, Ramakrishnan T. Antral atrophy, *Helicobacter pylori* colonization, and gastric pH. Am J Clin Pathol **1996**; 105:96–101.
55. Karnes WE Jr, Samloff IM, Siurala M, et al. Positive serum antibody and negative tissue staining for *Helicobacter pylori* in subjects with atrophic body gastritis. Gastroenterology **1991**; 101:167–74.
56. Bjorkholm B, Sjolund M, Falk PG, Berg OG, Engstrand L, Andersson DI. Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori.* Proc Natl Acad Sci USA **2001**; 98:14607–12.
57. Wang G, Humayun MZ, Taylor DE. Mutation as an origin of genetic variability in *Helicobacter pylori.* Trends Microbiol **1999**; 7:488–93.
58. Tsuda M, Karita M, Nakazawa T. Genetic transformation in *Helicobacter pylori.* Microbiol Immunol **1993**; 37:85–9.
59. Ando T, Xu Q, Torres M, Kusugami K, Israel DA, Blaser MJ. Restriction-modification system differences in *Helicobacter pylori* are a barrier to interstrain plasmid transfer. Mol Microbiol **2000**; 37:1052–65.
60. Singh KK, Zhang X, Patibandla AS, Chien P Jr, Laal S. Antigens of *Mycobacterium tuberculosis* expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. Infect Immun **2001**; 69:4185–91.
61. van Soolingen D, de Haas PE, Hermans PW, Groenen PM, van Embden JD. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis.* J Clin Microbiol **1993**; 31:1987–95.
62. Covacci A, Falkow S, Berg DE, Rappuoli R. Did the inheritance of a pathogenicity island modify the virulence of *Helicobacter pylori?* Trends Microbiol **1997**; 5:205–8.