Genetic variation in Helicobacter pylori
Pan, Z.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
cagA-positive *Helicobacter pylori* populations in China and the Netherlands are distinct

*Infection and Immunity* 1998; 66: 1822-1826
cagA-Positive Helicobacter pylori Populations in China and The Netherlands Are Distinct

Arié van der Ende, 1,2* Zhi-Jun Pan, 1,2* Albert Bart, 1 René W. M. van der Helst, 3 Monique Feller, 3 Shu-Dong Xiao, 2 Guido N. J. Ttygat, 3 and Jacob Dankert 3

Departments of Medical Microbiology 1 and Gastroenterology, 3 Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, and Shanghai Institute of Digestive Disease, Shanghai Second Medical University, Shanghai, People's Republic of China

Received 26 November 1997/Returned for modification 23 January 1998/Accepted 11 February 1998

The aim of this research was to study whether and to what extent Chinese cagA-positive Helicobacter pylori isolates differ from those in The Netherlands. Analysis of random amplified polymorphic DNA (RAPD)-PCR-assessed DNA fingerprints of chromosomal DNA of 24 cagA-positive H. pylori isolates from Dutch (n = 12) and Chinese (n = 10) patients yielded the absence of clustering. Based on comparison of the sequence of a 243-nucleotide part of cagA, the Dutch (group I) and Chinese (group II) H. pylori isolates formed two separate branches with high confidence limits in the phylogenetic tree. These two clusters were not observed when the sequence of a 248-bp part of glmM was used in the comparison. The number of nonsynonymous substitutions was much higher in cagA than in glmM, indicating positive selection. The average levels of divergence at the nucleotide and protein levels between group I and II isolates were found to be high, 13.3 and 17.9%, respectively. Possibly, the pathogenicity island (PAI) that has been integrated into the chromosome of the ancestor of H. pylori now circulating in China contained a different cagA than the PAI that has been integrated into the chromosome of the ancestor of H. pylori now circulating in The Netherlands. We conclude that in China and The Netherlands, two distinct cagA-positive H. pylori populations are circulating.

Helicobacter pylori infection in humans is one of the most widespread infections today, and its cure prevents peptic ulcer recurrence (26, 35). Besides symptomatic gastritis and peptic ulcer disease (PUD), H. pylori infection is strongly associated with gastric cancer, gastric mucosa-associated lymphoid tissue (MALT), and atrophic gastritis (3, 9, 24).

The heterogeneity of the clinical outcome of H. pylori infection may be related to differences among the host or to differences in virulence among H. pylori strains. The latter assumption is supported by the finding that the presence of cag toxin-associated gene A (cagA) has been found to be associated with PUD (7). PUD patients are virtually all infected with cagA-positive H. pylori and have serum antibodies as well as antibodies at the mucosal level against a 120- to 128-kDa protein encoded by this gene (7, 38). In contrast, only 60% of patients with functional dyspepsia (FD) are positive for this protein. The presence of cagA-positive H. pylori is also related to an increased risk to develop atrophic gastritis, intestinal metaplasia (16, 35), or gastric cancer (25).

Recently, the complete genomic sequence of H. pylori has become available (30). A 40-kb region of the H. pylori chromosome containing cagA was sequenced earlier by Censini et al. (4). This locus, comprising at least 40 genes, has a GC content different from that of the rest of the chromosome, forms a so-called pathogenicity island (PAI), and is assumed to have been integrated into the H. pylori chromosome only recently (4, 6). The proteins encoded by the PAI genes possess features similar to those of bacterial type II, type III, and most notably type IV secretion systems. It was hypothesized that such proteins may function to export macromolecules that may be involved in the H. pylori-host cell interaction (6).

China is one of the countries with a high prevalence of H. pylori infection and a high incidence of gastroduodenal diseases (39). The prevalence of H. pylori infection increases with age to about 70% of the people over 30 years old (22, 33, 39). The prevalence of cagA-positive H. pylori populations in Chinese patients with PUD and FD is almost universally high (21). Data obtained from this recent study further suggested that H. pylori genotypes distinct from those present in Western Europe may circulate in China.

The aim of this study is to investigate this hypothesis by comparison of the random amplified polymorphic DNA (RAPD)-PCR-assessed genotype of 24 randomly collected cagA-positive H. pylori isolates from 12 Dutch (14 isolates) and 10 Chinese patients. We used four different primers in each of four amplifications of H. pylori genomic DNA. In addition, part of cagA and glmM of the H. pylori isolates was sequenced. Sequences were analyzed for similarity by a computer-based program by using the neighbor-joining algorithm of Saitou and Nei (27).

**MATERIALS AND METHODS**

Patients and H. pylori isolates. In this study, 24 cagA-positive H. pylori isolates, 14 from 12 Dutch patients (6 with PUD and 7 with FD) and 10 from 10 Chinese patients (5 with PUD and 5 with FD), were used. Isolates were randomly collected from the collection present in the Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands. From two Dutch patients, five H. pylori isolates were analyzed. These isolates were cultured from biopsy specimens taken with 6-year (isolates 79A and 79B) and 4-year isolates (161A and 161B) from intervals, respectively. Culture of the H. pylori isolates and assessment for the presence of cagA by PCR and Western blotting were recently described (21, 38).

Preparation of genomic DNA for PCR. The chromosomal DNA of H. pylori was prepared as previously described (32). Briefly, stored bacterial suspensions were thawed, sonicated on horse blood agar plates, and cultured at 37°C for 3 days in a microaerobic environment. Bacteria were harvested, and genomic DNA...
was extracted by phenol-chloroform-isooctyl alcohol extraction and ethanol precipitation (32).

Genome typing by RAPD-PCR. PCR-based RAPD fingerprinting was performed by the method of Akopyants et al. (1), with minor modifications (32). Briefly, 20 µL of chromosomal DNA and 5 µL of one of the primer sets (Perkin-Elmer, Elmer Nederland BV, Gouda, The Netherlands) 1254 (5'-CCGCGGAGGCTAGA'), 1281 (5'-AACGCGCAACGCGAGGACTAGC'), 1283 (5'-GCGATCCCCAGCGATAGC'), and 1247 (5'-AAATCCGCACTCGTGATAGCT') were used in PCR as previously described (32). The PCR fragments were sequenced by robotic protocol using a fluorescent dye-labeled PCR primer (Perkin-Elmer Biosystems). Sequences were analyzed with Sequencher software version 3.1 (Applied Maths, Kortrijk, Belgium). Patterns were normalized to RAPD patterns from Neisseria meningitidis ET present every five lanes on each gel. The patterns generated by each of the four RAPD primers were combined and compared by using unweighted pair group method for arithmetic averages (UPGMA) clustering with Dice coefficient applied.

Fluorescence-based DNA sequencing and analysis. PCR products obtained with primer cagA5 (5'-GGAAATTTGCTGTGTGCTGCGG-3'), positions 1395 to 1519, according to Covacci et al. (5) and primer cagA6 (5'-GCGTCTTATCAGGCCCAAG-3'), positions 1289 to 1314, according to Labigne et al. (14), were used in PCR as previously described (15). The PCR fragments were sequenced by robotic protocol, using a fluorescent dye-labeled PCR primer (Perkin-Elmer Biosystems). Sequences were analyzed with Sequencher software version 3.1 (Applied Maths, Kortrijk, Belgium). Patterns were normalized to RAPD patterns from Neisseria meningitidis ET present every five lanes on each gel. The patterns generated by each of the four RAPD primers were combined and compared by using unweighted pair group method for arithmetic averages (UPGMA) clustering with Dice coefficient applied.

RESULTS

RAPD-PCR of H. pylori isolates from Dutch and Chinese patients. Assessment by RAPD-PCR of chromosomal DNA of 22 cagA-positive H. pylori isolates, 12 from 12 Dutch patients and 10 from 10 Chinese patients, showed that each isolate had a unique RAPD pattern. The initial isolate 79A and isolate 79J cultured from sequential biopsy specimens taken from the same patient were identical. Likewise, the initial isolate 161A was identical to isolate 161L. Clustering analysis did not reveal any clusters of isolates on the basis of either clinical manifestations or origin of geographic area.

Comparison of cagA sequences of H. pylori isolates from Dutch and Chinese patients. Comparison of a 243-bp part of the cagA region between nucleotides 1537 and 1780 (notation according to Covacci et al. (5)) from the 24 clinical H. pylori isolates showed 21 alleles, with mutations at 67 possible positions (Fig. 1). Both sequentially recovered H. pylori isolates from two Dutch patients (strains 161A and 161L; notation according to Covacci et al. (5)) showed that each isolate had a unique RAPD pattern. The initial isolate 79A and isolate 79J cultured from sequential biopsy specimens taken from the same patient were identical. Likewise, the initial isolate 161A was identical to isolate 161L. Clustering analysis did not reveal any clusters of isolates on the basis of either clinical manifestations or origin of geographic area.
level. The levels of average divergence of the cagA sequence among the group II strains were similar, 4.8 and 5.8% at the nucleotide and amino acid levels, respectively. Evidently, the difference in the cagA sequence was more extensive (two to three times larger) when the strains of the two groups were compared with each other (Table 2).

Comparison of glmM sequences of H. pylori isolates from Dutch and Chinese patients. To compare sequence heterogeneity of cagA, located on the PAI, with that of a gene outside the PAI, part of glmM (formerly called ureC [18]) was sequenced. Of the 24 H. pylori isolates, the same 240-bp part of glmM was sequenced as described by Kansau et al. (14). Twenty-two alleles with mutations at 32 possible positions were found (Fig. 3). The two sequentially recovered H. pylori isolates from each of the two Dutch patients (strains 161A and 161L; strains 79A and 79J) were identical. The total number of 32 nucleotide substitutions resulted in only 3 possible amino acid substitutions. The $d_{	ext{SS}}/d_{	ext{SN}}$ ratio ($d_{	ext{SS}}/d_{	ext{SN}} = 0.1103/0.0068 = 16.2$) was much higher in glmM than in cagA. In contrast to the cagA sequence, clustering analysis of glmM did not result in any robust cluster formation.

**TABLE 1.** Proportion (Jukes-Cantor corrected) of synonymous and nonsynonymous substitutions per site among the 243-nucleotide sequenced part of cagA between nucleotides 1573 and 1780 (notation according to Covacci et al. [5]) of 25 H. pylori isolates

<table>
<thead>
<tr>
<th>Source of isolates</th>
<th>No. of isolates</th>
<th>Proportion of substitutions (mean ± SE)</th>
<th>$d_{	ext{SN}}$</th>
<th>$d_{	ext{SS}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch patients</td>
<td>12*</td>
<td>0.102 ± 0.071</td>
<td>0.027 ± 0.021</td>
<td>3.8</td>
</tr>
<tr>
<td>Chinese patients</td>
<td>10</td>
<td>0.140 ± 0.053</td>
<td>0.025 ± 0.010</td>
<td>5.6</td>
</tr>
</tbody>
</table>

* The cagA sequences of H. pylori isolate 79J (identical to 79A but isolated from the same patient 4 years later) and 161L (identical to 161A but isolated from the same patient 6 years later) were not taken into account.

**DISCUSSION**

Data obtained from a recent report suggested that H. pylori genotypes circulating in China are distinct from those in Western Europe due to allelic variation in cagA (21). The aim of our study was to provide evidence that Chinese patients and Dutch patients are colonized with distinct cagA-positive H. pylori strains.

RAPD-PCR analysis of 14 H. pylori isolates from 12 Dutch patients and 10 from Chinese patients demonstrated a high level of genetic diversity among the 24 strains. In previous studies using this technique, it was shown that H. pylori comprises a genetically highly heterogeneous group, with patient-to-patient variation (1). In addition, patients can harbor a heterogeneous H. pylori population (12, 31, 33, 37). On the basis of the RAPD-PCR patterns, the 24 H. pylori strains could be clustered according to neither the various clinical entities nor the geographic origin of the patient. Results obtained with multilocus enzyme electrophoresis suggested clustering of 23 H. pylori isolates into four clusters (11). The authors concluded that the genetic diversity in H. pylori may be sufficient to classify H. pylori strains into four or more cryptic species.

**TABLE 2.** Sequence diversity among a part of 243 nucleotides of the cagA region between 1573 and 1780 (notation according to Covacci et al. [5]) of H. pylori isolates from 12 Dutch and 10 Chinese patients

<table>
<thead>
<tr>
<th>Source of isolates</th>
<th>No. of isolates</th>
<th>% Differences (mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Nucleotide</td>
</tr>
</tbody>
</table>

* Group I, Dutch patients; group II, Chinese patients.

* The cagA sequences of H. pylori isolate 79J (identical to 79A but isolated from the same patient 4 years later) and 161L (identical to 161A but isolated from the same patient 6 years later) were not taken into account.
substitutions were observed (not shown). Such a bias toward positions 1249 and 1519 (according to the notation of Covacci et al.) may have reached its balance. We hypothesize that most cagA variation of H. pylori occurs during the acute phase of infection, in which the incoming H. pylori has to adapt to harsh conditions present in the human stomach, resulting in new H. pylori variants. These so-called sequential bottlenecks might also give an explanation of the finding that patients can carry heterogeneous populations of H. pylori in one patient (12, 32, 34, 37). It may be that the different variants grow out at different sites in the stomach. Recombination within the chromosome of the bacteria and/or between different variants may further increase heterogeneity (2, 10). However, evidence for recombination within the cagA sequences was not found in the set of 25 H. pylori strains analyzed by a computer program using the algorithm of Maynard Smith (19). The many non-synonymous mutations could also imply that CagA of H. pylori from patients from different geographic areas are antigenically different, especially of H. pylori isolates from Dutch and Chinese patients.

In summary, we conclude that two distinct cagA-positive H. pylori populations are circulating in China and the Netherlands. Most likely, the PAI that has been integrated into the chromosome of the ancestor of H. pylori now circulating in China contained a different cagA than the PAI that has been integrated into the chromosome of the ancestor of H. pylori now circulating in the Netherlands.

ACKNOWLEDGMENTS

This work was supported by grants from the Dutch Ministry of Education and Science, the Royal Dutch Academy of Science, and the Chinese Ministry of Public Health (1994).

REFERENCES

tance of a pathogenicity island modify the virulence of Helicobacter pylori? Trends Microbiol. 8:205-209.


17. Kumar, S., K. Tamura, and M. Nei. 1993. MEGA2, Molecular Evolutionary Genetics Analysis v. 1.01. The Pennsylvania State University, University Park. (Distributed by the authors.)


