Development of a Scheme for Genotyping *Helicobacter pylori* Based on Allelic Variation in Urease Subunit Genes

ROBERT J. OWEN, 1, *ELEANOR R. SLATER, 1 JACQUELINE XERRY, 1 TANSY M. PETERS, 2 E. LOUISE TEARE, 2 AND ANDREW GRANT 3

Laboratory of Enteric Pathogens, Central Public Health Laboratory, 1 and PHLS Statistics Unit, 3 London, and Public Health Laboratory, Chelmsford, 2 United Kingdom

Received 5 May 1998/Returned for modification 6 July 1998/Accepted 23 September 1998

*Helicobacter pylori* urease subunit genes in 383 isolates from 10 countries were investigated by PCR-restriction fragment length polymorphism (*Hae*III) analysis. Eighty-two different *ureAB* profiles were documented by reference to known sequences. Variation among 51% of strains was accounted for by 10 predominant patterns, which provided a unique framework for categorizing isolates with geographically diverse origins.

*Helicobacter pylori* is the causative agent worldwide of chronic gastritis and is a significant risk factor in the development of peptic ulceration and gastric cancer (5). Intense urease activity is a key pathogenicity factor (12), and in urease, the first genomic region to be sequenced, genes are highly conserved (2, 13, 28). Because DNA restriction digest analysis of isolates from different patients indicates a high level of intragenic diversity, the approach has been applied widely for genotyping (1, 2, 6, 7, 10, 15, 17, 18, 23–27). In view of the importance of urease as a candidate vaccine (14) and as a target for PCR specific-detection assays (15, 19), we have developed a scheme for indexing *ureAB* variation to facilitate identification of clinical isolates.

The 383 isolates of *H. pylori* were collected between 1992 and 1997 from gastric or duodenal biopsy tissue taken during routine endoscopy of patients presenting with symptoms of dyspepsia (10, 11, 18) or from the gastric juice of healthy volunteers (21). Three reference cultures (NCTC 11637, NCTC 11638, and NCTC 12455 strain 26695) were included. Isolates were mostly from United Kingdom (n = 323), but strains from nine other countries (Australia [3 isolates], Canada [8 isolates], China [1 isolate], Italy [4 isolates], Nigeria [7 isolates], Peru [2 isolates], South Africa [4 isolates], Turkey [26 isolates] and the United States [5 isolates]) were also included in the study to determine geographical diversity. Strains were cultured on 10% Columbia blood agar at 37°C in a variable atmosphere incubator (Don Whitley Scientific Ltd., Shipley, United Kingdom) under microaerobic conditions (5% O₂, 5% CO₂, 2% H₂, and 88% N₂) and preserved at −196°C. Genomic DNA was extracted, the *ureAB* region was amplified, and restriction fragment length polymorphism (RFLP) analysis was performed as described previously (11, 18, 23). DNA fragment sizes were estimated from migration distances by using polynomial curve-fitting functions (16). Profiles were recorded in Microsoft Excel as derived fragment sizes (in base pairs) and analyzed as numerical strings according to their positions in the amplicon (Table 1). A novel Visual Basic macro (details are available from the corresponding author) was written to calculate genotype frequencies and to search for matches and partial matches with existing database profiles (9).

The 2.41-kb *ureAB* amplicon was obtained from all 383 isolates of *H. pylori*. A total of 82 *ureAB* *Hae*III RFLP patterns were defined and designated by a *ureAB* (UAB) profile number (UAB1 to UAB82). Examples of typical profiles are shown

<table>
<thead>
<tr>
<th>Profile</th>
<th>Band sizes (bp)</th>
<th>Frequency</th>
<th>Reference strain</th>
<th>Other designations</th>
</tr>
</thead>
<tbody>
<tr>
<td>uAB1</td>
<td>315, 457, 444, 678, 515</td>
<td>25</td>
<td>NCTC 13085</td>
<td>60190, ATCC 49503</td>
</tr>
<tr>
<td>uAB2</td>
<td>315, 457, 444, 383, 295, 515</td>
<td>18</td>
<td>NCTC 11637*</td>
<td>RPH 13487, CCUG 17874, ATCC 43504</td>
</tr>
<tr>
<td>uAB3</td>
<td>315, 457, 1,122, 515</td>
<td>23</td>
<td>NCTC 12455</td>
<td>26695*</td>
</tr>
<tr>
<td>uAB4</td>
<td>772, 1,059, [63], 515</td>
<td>14</td>
<td>NCTC 12385</td>
<td></td>
</tr>
<tr>
<td>uAB5</td>
<td>772, 827, 232, [63], 515</td>
<td>12</td>
<td>NCTC 13082</td>
<td></td>
</tr>
<tr>
<td>uAB6</td>
<td>772, 827, 295, 515</td>
<td>7</td>
<td>NCTC 13089</td>
<td></td>
</tr>
<tr>
<td>uAB10</td>
<td>315, 457, 827, 295, 515</td>
<td>11</td>
<td>NCTC 11638</td>
<td>RPH 13491, CCUG 17875</td>
</tr>
<tr>
<td>uAB11</td>
<td>315, 901, 615, [63], 515</td>
<td>9</td>
<td>NCTC 13090</td>
<td></td>
</tr>
<tr>
<td>uAB19</td>
<td>315, 457, 444, 165, 450, [63], 515</td>
<td>7</td>
<td>NCTC 13088</td>
<td></td>
</tr>
<tr>
<td>uAB32</td>
<td>772, 444, 678, 515</td>
<td>6</td>
<td>NCTC 13094</td>
<td></td>
</tr>
</tbody>
</table>

* Corresponding author. Mailing address: Helicobacter Reference Unit, Laboratory of Enteric Pathogens, Central Public Health Laboratory, 61 Colindale Ave., London NW9 5HT, United Kingdom. Phone: (44) 0181 200 4400 ext. 3740. Fax: (44) 0181 905 9929. E-mail: rowen@phls.co.uk

* Bands are ordered from 5′ to 3′ ends according to predicted locations of *Hae*III sites within known *ureA* and *ureB* gene sequences (3, 13, 28). The 63-bp fragment (in brackets) was not detected in the RFLP analysis.

* Based on the numbers of infected patients and excluding multiple isolates from individual patients.

* Type strain of species.

* Genotype sequence has been published (28).

©1998, American Society for Microbiology. All Rights Reserved.
studies showing considerable diversity in the region associated with cell wall synthesis and not to urease activity (4). Used for genotyping, its main role is now recognized to be restriction fragment length polymorphism (25). Although ureC (operon) was selected as the core pattern, as it was the most frequent and gave rise to the other common patterns by a series of single or double mutations. The 315-bp region at the 5' end of ureA and the 515-bp region at the 3' end of ureB were highly conserved, whereas most variation occurred in the middle region of ureB. An analysis of strains representing the 10 common types indicated that ureAB profiling was more discriminatory than ureC (HhaI) and ureC (MboI) profiling but marginally less discriminatory than combined ureC (HhaI-MboI) profiling (25). Although ureC (revised designation, glmM) is used for genotyping, its main role is now recognized to be related to cell wall synthesis and not to urease activity (4).

Our results confirm and extend the findings of previous studies showing considerable diversity in the ureAB region of H. pylori isolated from different individuals (2, 18). All strains produced functionally highly active urease, and mutations within the subunit genes were synonymous for the three sequenced strains. The effect of such variations on urease antigenicity is unknown but may be of practical significance when considering the effectiveness of urease-based vaccines in diverse and heterogeneous human populations.

A novel finding was that about half the strains were represented by just 10 ureAB RFLP profiles, which indicated a higher degree of conservation than expected. Furthermore, H. pylori with the same urease profile infected unrelated individuals, and in some cases, isolates were from widely separated parts of the world. Although it is generally believed that a large and very genetically diverse population of H. pylori circulates in the community, certain genomic markers, such as cagA and vacAs1, are widely distributed and partially conserved (8, 22), and it would appear that this is also true for certain ureAB polymorphisms. Further evidence of this conservation was apparent from published data on strains isolated in the United States showing profiles closely matching uAB3 and uAB5 (6). The availability of H. pylori urease gene sequences enabled precise sizes to be assigned to each RFLP profile and an exact genomic location to be identified. This novel development of RFLP analysis has considerable potential for studying allelic variation in other regions of the H. pylori genome, such as the highly polymorphic vacA locus (22).

In conclusion, we found that HaeIII RFLP analysis was an easy-to-perform, reproducible, and discriminatory means of documenting diversity within the H. pylori ureAB operon. Furthermore, differences were independent of variation in other loci associated with pathogenicity, such as cagA, vacA, 48-kDa HiiA, and 26-kDa antigen genes, although analysis of such loci with multiple restriction enzymes could be applied to further discriminate strains when necessary (2, 20). The scheme was developed mainly on U.K. isolates of H. pylori, but it appears to be equally applicable to isolates from other parts of the world and thus provides a practical general-purpose genotyping framework to facilitate interlaboratory comparability.

We thank the various collaborators for providing isolates.

### References


9. Grant, A. Unpublished data.
25. Slater, E. Unpublished data.