Fine-Structure Molecular Typing of Irish Helicobacter pylori Isolates and Their Genetic Relatedness to Strains from Four Different Continents

Ian M. Carroll, 1* Niyaz Ahmed, 2 Sarah M. Beesley, 1 Aleem A. Khan, 3 Sheikh Ghausunnissa, 2 Colm A. O Moráin, 4 and Cyril J. Smyth 1

Department of Microbiology, Moyne Institute of Preventive Medicine, Trinity College, Dublin 2, 1 and Department of Gastroenterology, The Adelaide and Meath Hospital, Tallaght, Dublin 24, 4 Ireland, and Centre for DNA Fingerprinting and Diagnostics (CDFD), Nacharam, 2 and Deccan College of Medical Sciences and Allied Hospitals, 5 Hyderabad, India

Received 4 March 2003/Returned for modification 29 May 2003/Accepted 1 August 2003

Genotyping of 74 Irish Helicobacter pylori isolates was performed at four different loci (vacA signal sequence and mid-region, insertion-deletion polymorphisms at the 3′ end of the cag pathogenicity island, and cagA). The predominant vacA alleles and insertion-deletion motifs suggest an ancestral relationship between Irish isolates and either specific East Asian or Northern European strains. In addition, fluorescent amplified fragment length polymorphism-PCR genotyping and phylogenetic analysis of 32 representative Irish H. pylori isolates and 22 isolates from four different continents demonstrated that the Irish H. pylori isolates examined were weakly clonal and showed some association with both European and Asian isolates. These three genotyping techniques show that Irish H. pylori isolates have distinctive features that may have evolved in this insular European population.

Helicobacter pylori displays genetic diversity between strains at a very high level (17). This diversity has been shown to occur at specific loci within H. pylori isolates (3, 4, 21, 23). The vacA gene, which encodes the vacuolating cytotoxin, is a polymorphic locus present in all H. pylori strains. While this gene contains conserved regions, sequence variations occur in the signal sequence and the mid-region of the gene as a mosaic structure (3). The signal sequence can be one of four allelic subtypes (s1a, s1b, s1c, and s2) (25) and the mid-region can be one of two allelic subtypes (m1 and m2) (3). There have been numerous reports correlating these vacA allelic types with particular gastric diseases (5, 20), e.g., the vacA s1 m1 genotype has been associated with gastric carcinoma in Germany (18).

The cagA gene (cytotoxin-associated gene) is another polymorphic locus that is thought to be a marker for virulence and is associated with increased severity of disease in some geographic regions (6). The 3′ end of the cag pathogenicity island (cag PAI) (the region between the 3′ end of the cagA gene and the 3′ end of the glutamate racemase gene [gtr]) contains different motifs that have previously been used for studying relationships between H. pylori isolates from different geographic regions (14). To demonstrate geographic partitioning of H. pylori isolates, other studies have investigated sequences of putative virulence and housekeeping genes (1).

Although H. pylori is a panmictic organism (10), a study encompassing isolates from different regions of the world identified weakly clonal groupings for this bacterium (1). There have also been indications of nonrandom geographic distributions of H. pylori genotypes that may aid in the understanding of the evolution of H. pylori and its clinical consequences in different parts of the world (14, 24). The predominant vacA alleles have been shown to distinguish H. pylori isolates in different geographic regions (13, 15, 16, 19, 24).

Fluorescent amplified fragment length polymorphism (FAFLP) analysis is a genotyping technique that has been used to investigate outbreaks of infection by Streptococcus pyogenes, Staphylococcus aureus, and Neisseria meningitidis (7, 11, 12). FAFLP analysis generates a specific profile for each strain under examination. These profiles can be grouped together, resulting in generalized patterns. Recently, FAFLP was used to examine the geographic partitioning of H. pylori isolates [N. Ahmed, A. A. Khan, D. E. Berg, and C. M. Habibullah, abstract from the Int. Workshop on Campylobacter, Helicobacter and Related Organisms (CHRO), 2001, Freiburg, Germany, Int. J. Med. Microbiol. 291(Suppl. 31):D-11, 2001].

For the present study, Irish H. pylori isolates were characterized by FAFLP and insertion-deletion analyses, along with vacA genotyping and cag gene status, and compared to isolates from other geographic regions, namely India, United Kingdom, Peru, Spain, Japan, and Africa.

Seventy-four clinical H. pylori isolates were examined that had been obtained from single antral gastric biopsies from patients experiencing peptic ulcer disease at the Meath-Adelaide and St. James’s Hospitals in Dublin. Culture and storage of isolates were performed as described previously (8). Twenty-two H. pylori strains from six different geographic regions were also analyzed (six from India, five from Peru, five from the United Kingdom, four from Spain, one from Africa, and one from Japan). Genomic DNA preparations were provided by Douglas E. Berg and Asish Mukhopadhyay (Washington University Medical School, St. Louis, Mo.) from 16 non-Irish isolates (3, 4, 21, 23).

* Corresponding author. Mailing address: Department of Microbiology, Moyne Institute of Preventive Medicine, Trinity College, Dublin 2, Ireland. Phone: 353 1 6081728. Fax: 353 1 6799294. E-mail: carrolli @tcd.ie.
and non-Indian H. pylori strains originally obtained by Robert H. Gilman, Teresa Alarcon, Manuel Lopez Brea, and John Atherton. The Indian preparations were from the Centre for DNA Fingerprinting and Diagnostics (Hyderabad, India).

Genomic DNA was isolated as described previously (8). The sequences of the oligonucleotides used as PCR primers and their corresponding PCR product sizes are listed in Table 1. The signal sequences and mid-regions of the vacA gene, the cagA gene, and the insertion-deletion motifs at the 3' end of the cag PAI were amplified in separate reactions as reported previously (3, 14, 25). Negative controls for PCR included sterile distilled water and Escherichia coli DNA from strain XL1-Blue. The nucleotide sequences of representative vacA and cagA PCR products were determined to confirm their identities.

Of the 74 isolates examined, 69% were of vacA type s1a m2, 27% were of vacA type s1a m1, and 4% were of vacA type s2 m2 (Table 2). The most common signal sequence and mid-region alleles were s1a and m2, respectively, which is in accordance with previous studies of Irish H. pylori isolates (22). The frequency of the predominant vacA allele found in this study is similar to those found in Taiwan (82%) (16) and China (85%) (19). However, vacA alleles have been shown to differ in frequency and type among East Asian isolates (13, 16, 19), e.g., s1c is the predominant signal sequence allele among East Asian isolates (24). Thus, Irish H. pylori isolates share vacA similarities with isolates from specific countries within Eastern Asia. This may reflect particular adaptations of H. pylori to specific host populations (9, 24) rather than indicating geographic relationships.

All of the 74 Irish H. pylori isolates were found to contain the cagA gene. This does not agree with a previous Irish study (8), but as no asymptomatic patients were included in the present study, this may have influenced the present finding. Moreover, the cagA status of this collection was not known before the study commenced. Characterization of the 3' end of the cag PAI has shown that eight different insertion-deletion motifs can be present (Ia, Ib, Ic, II, IIIa, IIIb, IV, and V) (14). These insertion-deletion motifs have been shown to be well suited to population-level surveys of H. pylori genotypes (14). Type I motifs predominate in isolates from Spanish and Peruvian populations but are less common in strains from Northern European populations. Type II motifs are mostly found among Japanese and Chinese strains, but are also common in Northern European strains. Type III motifs predominate among isolates from Indian populations. When insertion-deletion analysis was performed on the isolates, 63 (85%) yielded useable PCR products, 4 (5%) produced multiple PCR products (which made it impossible to determine the different motifs at

### Table 1. Oligonucleotide sequences of primers and PCR product sizes

<table>
<thead>
<tr>
<th>Amplified gene or DNA region</th>
<th>Primer</th>
<th>Primer sequence (5'-3')</th>
<th>PCR product size (bp [type])</th>
</tr>
</thead>
<tbody>
<tr>
<td>cagA</td>
<td>CAGAF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>GATAACAGGCAAGCTTTTGGAG</td>
<td>349</td>
</tr>
<tr>
<td>Signal sequence of vacA</td>
<td>VAI-F&lt;sup&gt;b&lt;/sup&gt;</td>
<td>ATGGAAATACACAAACACAC</td>
<td>259 (s1)</td>
</tr>
<tr>
<td>Signal sequence sla of vacA</td>
<td>SS1-F&lt;sup&gt;c&lt;/sup&gt;</td>
<td>GTCAGCATCACCGCAAC</td>
<td>190</td>
</tr>
<tr>
<td>Mid-region of vacA</td>
<td>VAI-R&lt;sup&gt;b&lt;/sup&gt;</td>
<td>CTGGTTGAATGCGCAAC</td>
<td>286 (s2)</td>
</tr>
<tr>
<td>Extreme right end of the cag PAI</td>
<td>cagR5280&lt;sup&gt;c&lt;/sup&gt;</td>
<td>GTTATTACAAAAGGTGTTCAGAC</td>
<td>1,200 (IV)</td>
</tr>
</tbody>
</table>

**Type IV insertion-deletion motif specific**
- Xins.R<sup>c</sup> | CGCTCTCTAAATGTTCTAGGA | 520 |
- cagF4584<sup>c</sup> | Same as above |
- cagR5280<sup>c</sup> | Same as above |

**Type II insertion-deletion motif specific**
- IS606-1692<sup>c</sup> | CTAACAAATTGGGATGTGTC | 400 |
- cagR5280<sup>c</sup> | Same as above |

**Type III insertion-deletion motif specific**
- fcn unk<sup>c</sup> | TGGATTAAATCTTAATTGAATTAC | 320 |
- cagR5280<sup>c</sup> | Same as above |

**Type I and IV insertion-deletion motif specific**
- cagF4856<sup>c</sup> | GCGTCAAGGAGAATATCTTTAGCG | 350 (I) |
- cagR5280<sup>c</sup> | Same as above |

### Table 2. vacA typing and cagA insertion-deletion motif analysis of 74 Irish H. pylori isolates<sup>a</sup>

<table>
<thead>
<tr>
<th>vacA or insertion-deletion type</th>
<th>No (%) of strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>vacA types</td>
<td></td>
</tr>
<tr>
<td>s1a..................................</td>
<td>71 (95.9)</td>
</tr>
<tr>
<td>s2..................................</td>
<td>3 (4.1)</td>
</tr>
<tr>
<td>m1..................................</td>
<td>20 (27)</td>
</tr>
<tr>
<td>m2..................................</td>
<td>20 (27)</td>
</tr>
<tr>
<td>s1a m2................................</td>
<td>51 (68.9)</td>
</tr>
<tr>
<td>s1a m1................................</td>
<td>20 (27)</td>
</tr>
<tr>
<td>s2 m2................................</td>
<td>3 (4.1)</td>
</tr>
</tbody>
</table>

**Insertion-deletion types**
- Ia.................................. | 6 (9.5)           |
- II.................................. | 31 (49.2)         |
- IIIa................................ | 3 (4.8)           |
- IIIb................................ | 1 (1.6)           |
- IV.................................. | 22 (34.9)         |

<sup>a</sup> For insertion-deletion motif analysis, only 63 isolates were typeable.
this locus), and 7 (10%) did not yield PCR products under the test conditions used, despite repeated attempts. Five insertion-deletion motifs were detected. Of the 63 characterizable isolates, 31 (49%) were found to have type II motifs, 22 (35%) had type IV motifs, 6 (9%) had type IIIa motifs, and 1 (2%) had a type IIIb motif (Table 2). These findings suggest that Irish \textit{H. pylori} isolates are similar to Northern European strains, as type II motifs are common in this region, but not to East Asian isolates, as the frequencies of occurrence are much higher (95%) (14). However, although type II motifs are common in Irish and Northern European isolates, type I motifs also have a high prevalence among Northern European strains, whereas Irish isolates show a high prevalence of type IV motifs. Interestingly, the only isolate yielding a type IV motif in a study of 500 strains was from England (14). A more extensive study of English strains by insertion-deletion analysis is required to clarify the frequency of occurrence of type IV motifs among isolates from different ethnic groups, including the Irish diaspora. Thus, this collection of Irish \textit{H. pylori} isolates displays a rare insertion-deletion genotype that may be representative of an insular European population.

FAFLP-PCR was carried out as previously detailed (2). The fingerprints for each Irish isolate were generated by using three selective primers (\textit{Eco}RI plus A, G, and C), and the fingerprints for each non-Irish isolate were generated by using one selective primer (\textit{Eco}RI plus G). Amplified products were sized in base pairs within the user-defined categories of marker sizes. The presence or absence of amplicons within the categories was scored by a user-defined Genotyper macro. Allele scores (presence or absence of amplicons) were converted into a binary format (1 or 0, respectively). This binary format was converted to a nucleotide sequence (1 = G and 0 = A); therefore, fingerprint profiles could be aligned and a neighbor-joining tree could be constructed based on these profiles using Genotyper (Fig. 1) or ClustalX and Treeview (Fig. 2).

As \textit{vacA} and insertion-deletion genotyping suggested potential geographic relationships between Irish \textit{H. pylori} isolates and isolates from specific countries in East Asia and from Northern Europe, respectively, FAFLP-PCR analysis was applied. This technique encompasses the whole genome rather than specific loci and has been previously used to determine the geographic relationship of \textit{H. pylori} isolates [Ahmed et al., Int. J. Med. Microbiol. \textbf{291}(Suppl. 31):D-11, 2001]. FAFLP-PCR analysis was used to examine 45 of the 74 Irish \textit{H. pylori} isolates. Of these 45 isolates, 32 (71%) yielded informative profiles. The remaining 13 isolates either did not yield profiles with all three selective primers or did not yield profiles at all. However, the 32 characterizable isolates were still genotypically representative of the larger set of isolates. Twenty-five of these isolates (78%) yielded insertion-deletion motifs (type II, 13 [52%]; type IV, 9 [36%]; type Ia, 2 [8%]; type IIIb, 1 [4%]), while the remaining isolates were not typeable by insertion-deletion analysis (7 [22%] of the 32 isolates; cf. with 11 of 74 [15%]). The distribution of \textit{vacA} alleles in this selected group of isolates was s1a m2 (19 [60%]), s1a m1 (10 [31%]), and s2 m2 (3 [9%]). The amplification reactions yielded unique profiles for each isolate, and a neighbor-joining phylogenetic tree was constructed based on these profiles (Fig. 1). Analysis of this tree showed that the majority (28 [87.5%]) of isolates in this group differ in percentage heterogeneity by approximately 7.8% with respect to their FAFLP profiles. This supports earlier observations by which there appears to be a relative genetic restriction of \textit{vacA} alleles and insertion-deletion motifs, indicating that Irish \textit{H. pylori} isolates are weakly clonal.

Seven FAFLP-PCR profiles randomly chosen from among those obtained for Irish isolates plus 22 profiles previously generated in the same way at the Centre for DNA Fingerprint-
ing and Diagnostics for strains encompassing six different geographic regions [Ahmed et al., Int. J. Med. Microbiol. 291 (Suppl. 31):D-11, 2001] were phylogenetically analyzed. The binary data obtained from FAFLP-PCR analysis were used to construct a neighbor-joining phylogenetic tree (Fig. 2). This tree shows two main clusters (1 and 2) and six subgroups (A to F). The fingerprints used to construct this tree were previously generated by using only one selective primer, whereas fingerprints generated by three selective primers were used to construct Fig. 1. Using only one selective primer reduces the sensitivity of the technique. This explains why isolate FR165 shows about 9% heterogeneity with respect to isolates J563 and MI571 in Fig. 1, while these isolates show little difference in Fig. 2. However, each fingerprint is still unique to each isolate and the geographic separation of isolates can clearly be seen in Fig. 2. Subgroups A, B, and C are in cluster 1 and subgroups D, E, and F are in cluster 2. Each subgroup contained either isolates from one region or a combination of isolates from different geographic locations (A, mainly India; B, mainly United Kingdom; C, Ireland; D, mainly United Kingdom and Spain; E, mainly Spain and Peru; and F, Spain, Peru, and Africa). The fact that Irish \emph{H. pylori} isolates do not cluster with isolates from any other region suggests that they may have a unique genotype and are distinct from those from other regions. However, the Irish isolates do reside in cluster 1 with Asian and United Kingdom isolates, demonstrating that they are more like isolates from these regions than like Spanish, Peruvian, and African isolates. These findings are in accordance with a recent study that demonstrated that Indian and European \emph{H. pylori} isolates grouped in the same subpopulation and that East Asian and a subset of European isolates share an ancestral relationship and diverged from each other recently (9).

As indicated by the present genotyping study and by other studies (22), Irish \emph{H. pylori} isolates appear to exhibit restricted diversity. The clustering of the Irish isolates into one subgroup (C) by FAFLP-PCR confirms that Irish isolates are weakly clonal. This weakly clonal behavior was not observed within strains from other geographic regions, as some isolates from specific regions clustered in two or three different subgroups (Indian isolates clustered in subgroups A, B, and E, United Kingdom isolates clustered in subgroups B and D, Peruvian isolates clustered in subgroups D, E, and F, and Spanish isolates clustered in subgroups A, D, E, and F). The clustering of Peruvian and Spanish isolates (Fig. 2) is in accordance with previous studies (14).

In summary, the results obtained for this study suggest that Irish \emph{H. pylori} isolates have associations with Northern European isolates, as revealed by insertion-deletion analysis. Yet the \textit{vacA} allelic predominance among these Irish isolates displays similarities with isolates from China and Taiwan. Further, FAFLP-PCR genotyping confirmed that Irish isolates are

![Phylogenetic tree of various \emph{H. pylori} isolates from different geographic regions. The neighbor-joining network was developed from the binary data obtained by the Genotyper macro. Clustering of isolates is based on sharing of FAFLP fragments. Clustering resulted in two main clusters (1 and 2) and six distinct subgroups (A to F).](http://jcm.asm.org/ on October 30, 2017 by guest)
weakly clonal, display more homogeneous fingerprint profiles than do isolates from other geographic locations, and have some associations with both European and Asian isolates. These three genotyping techniques show that Irish H. pylori isolates have distinctive features that may have evolved in this insular European population.

Ian Carroll was the recipient of a postgraduate studentship from the Health Research Board (HRB) of Ireland, whose support is gratefully acknowledged. Niyaz Ahmed acknowledges with thanks a research grant from the Department of Biotechnology, Government of India.

We thank Douglas E. Berg and Ashis Mukhopadhyay for the supply of genomic DNA from H. pylori isolates from non-Irish and non-Indian origins, originally obtained by Robert H. Gilman, Teresa Alarcon, Manuel Lopez Brea, and John Atherton, and Seyed E. Hasnain for his suggestions and encouragement. We also thank Ross Fitzgerald for his helpful suggestions.

REFERENCES


