Helicobacter pylori: Clonal Population Structure and Restricted Transmission within Families Revealed by Molecular Typing

SHAN-RUI HAN,1 HANS-CHRISTOPH E. ZSCHAUSCH,1 HEINZ-GEORG W. MEYER,1 THOMAS SCHNEIDER,2 MICHAEL LOOS,1 SUCHARIT BHAKDI,1 AND MARKUS J. MAEGERER1*

Department of Medical Microbiology1 and Department of Paediatrics,2 Johannes Gutenberg University, D-55101 Mainz, Germany

Received 3 April 2000/Returned for modification 11 May 2000/Accepted 31 July 2000

Helicobacter pylori infects up to 50% of the human population worldwide. The infection occurs predominantly in childhood and persists for decades or a lifetime. H. pylori is believed to be transmitted from person to person. However, tremendous genetic diversity has been reported for these bacteria. In order to gain insight into the epidemiological basis of this phenomenon, we performed molecular typing of H. pylori isolates from different families. Fifty-nine H. pylori isolates from 27 members of nine families were characterized by using restriction fragment length polymorphism analysis of five PCR-amplified genes, by pulsed-field gel electrophoresis (PFGE) of chromosomal DNA, and by vacA and cagA genotyping. The 16S rRNA gene exhibited little allelic variation, as expected for a unique bacterial species. In contrast, the vacA, flaA, ureAB, lspa-glmM genes were highly polymorphic, with a mean genetic diversity of 0.83, which exceeds the levels recorded for all other bacterial species. In conjunction with PFGE, 59 H. pylori isolates could be differentiated into 21 clonal types. Each individual harbored only one clone, occasionally with a clonal variant. Identical strains were always found either between siblings or between a mother and her children. Statistical analysis revealed clonality of population structure in all isolates. The results of this study suggest the possible coexistence of a large array of clonal lineages that are evolving in each individual in isolation from one another. Transmission appears to occur primarily from mother to child and perhaps between siblings.

Helicobacter pylori-associated gastritis is today recognized as the major cause of duodenal and gastric ulcers, gastric adenocarcinoma and mucosa-associated lymphoid tissue lymphoma (9). H. pylori infects up to 50% of the human population worldwide. The infection occurs predominantly in childhood and persists for decades or a lifetime. H. pylori is believed to be transmitted from person to person. However, tremendous genetic diversity has been reported for these bacteria. In order to gain insight into the epidemiological basis of this phenomenon, we performed molecular typing of H. pylori isolates from different families. Fifty-nine H. pylori isolates from 27 members of nine families were characterized by using restriction fragment length polymorphism analysis of five PCR-amplified genes, by pulsed-field gel electrophoresis (PFGE) of chromosomal DNA, and by vacA and cagA genotyping. The 16S rRNA gene exhibited little allelic variation, as expected for a unique bacterial species. In contrast, the vacA, flaA, ureAB, lspa-glmM genes were highly polymorphic, with a mean genetic diversity of 0.83, which exceeds the levels recorded for all other bacterial species. In conjunction with PFGE, 59 H. pylori isolates could be differentiated into 21 clonal types. Each individual harbored only one clone, occasionally with a clonal variant. Identical strains were always found either between siblings or between a mother and her children. Statistical analysis revealed clonality of population structure in all isolates. The results of this study suggest the possible coexistence of a large array of clonal lineages that are evolving in each individual in isolation from one another. Transmission appears to occur primarily from mother to child and perhaps between siblings.

MATERIALS AND METHODS

Subjects. Nine index children from different families, designated families A through I, were referred to the Department of Pediatrics, Johannes Gutenberg University, Mainz, Germany, due to recurrent abdominal pain. One or more gastric (antral and corpus) and duodenal biopsy specimens were obtained from each of these patients and from at least one additional family member (see Fig. 2). In total, 27 members from nine families were investigated, represented by 2 to 6 members in each family. These patients were 23 children (11 boys and 12 girls with an age range of 2 to 18 years) and 4 adults from three families (2 women and 2 men from 29 to 44 years of age). The other parents were not sampled. They lived either within or in close proximity to Mainz.

Culture and PCR. Biopsy specimens were cultured on Columbia agar with 7% human erythrocytes and an H. pylori-selective supplement at 37°C under microaerobic conditions. H. pylori was identified and bacterial genomic DNA was prepared as previously described (23). PCR amplification of the vacA, flaA, ureAB, lspa-glmM (formerly ureCD), or cagA gene fragments was performed using the primer pairs described previously (23). The cagA gene status was determined by the presence or absence of cagA amplicons. The vacA genotypes, including signal(s) sequences and midregion (m) types, were characterized by a one-step PCR method (22). For ribotyping, a 1.5-kb fragment was amplified with broad-specificity 16S ribosomal DNA (rDNA) primer pairs pA4 (5'-AGAGTTTGATCCTGCGCTAG-3') and pH-b (5'-AAGGAGGTGACCTGGAGCAG-3') (15), using an initial denaturation step at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C, annealing at 55°C, and extension at 72°C (each for 1 min), and a final extension at 72°C for 10 min. In order to eliminate contaminating DNA within PCR reagents, UV light and 8-methoxypsoralen (Sigma,
patterns of 16S rDNA was defined as a ribotype. The restriction fragment length polymorphism (PCR-RFLP) type, which defined the A combination of these 10 patterns was designated a PCR-based locus-specific and the DNA digests were analyzed on ethidium bromide-stained 2% agarose mended by the supplier (New England Biolabs, Schwalbach/Taunus, Germany), prepared in 0.5

Biolabs) served as a size standard. For pulsed-field gel electrophoresis (PFGE), Laboratories, Hercules, Calif.). A lambda ladder PFG marker (New England

bromide and photographed. A combination of three restriction patterns in an subjected to electrophoresis for 20.3 h at 200 V and a 0.5- to 60.4-s switch time

f o r1ha troom temperature, and then washed (18). Approximately 1 ml of cell suspension with an optical density at 600 nm of 0.9 was needed for the prepa-

tion of two agarose plugs. The required volume of cell suspension was centri-

fuged, and bacteria were resuspended in 10 mM Tris-HCl–10 mM EDTA (pH

4°C for 3 months.

4°C. The plugs could be stored in 10 mM Tris-HCl–50 mM EDTA (pH

8.0) at 4°C. The index of association between loci (hij) is the number of isolates in the sample, and n(n−1) is a correction for bias in small samples (38). Genotypic diversity was calculated by the same formula, except that the frequency of the ith RFLP type and

is the frequency of the RFLP type and

is the proportion of loci at which dissimilar alleles occur, i.e., the proportion of mismatches (38). The dendrogram was constructed from PCR-RFLP typing data by the unweighted pair group cluster method with arithmetic means in the computer program MEGA (28).

The methods previously described by Taylor et al. (41) were modified. Fresh 2-day cultures of H. pylori were harvested and washed three times using normal saline. Alternatively, in order to inhibit DNase activity, 0.9-ml portions of normal saline. Alternatively, in order to inhibit DNase activity, 0.9-ml portions of

buffer was replaced with 100

of fresh enzyme buffer containing 25 U of NotI, Ncol, or Smal (New England Biolabs) and incubated overnight at 37°C (25°C for Smal). The plugs were chilled on ice for 15 min, and the enzyme buffer was replaced with 10 mM Tris-HCl–10 mM EDTA (pH 8.0), followed by loading of the plugs into the running gel (Pulsed-field certified agarose; Bio-Rad Laboratories, Hercules, Calif.). A lambda ladder PFGE marker (New England Biolabs) served as a size standard. For pulsed-field gel electrophoresis (PFGE), CHEF-DR III PFGE systems (Bio-Rad Laboratories) were used. An agarose gel prepared in 0.5× Tris-borate-EDTA running buffer (Bio-Rad Laboratories) was subjected to electrophoresis for 20.3 h at 200 V and a 0.5- to 60.4-s switch time at a constant temperature of 14°C. Agarose gels were stained with ethidium bromide and photographed. A combination of three restriction patterns in an isolate was designated a PFGE type.

Statistical analysis. For the RFLP data, different patterns of each of the analyzed genes were treated as different alleles. A distinct combination of 10 individual patterns in an H. pylori isolate was designated a RFLP type (clonal type) (14). All combinations of six alleles (6RFLP patterns) were calculated. Genetic diversity (h) is calculated as h = 1 − \sum_i p_i^2 (n(i−1)), where \(p_i\) is the frequency of the ith allele at the locus, n is the number of isolates in the sample, and n(n−1) is a correction for bias in small samples (38). Genotypic diversity was calculated by the same formula, except that the frequency of the ith RFLP type and

is the frequency of the ith RFLP type and

is the proportion of loci at which dissimilar alleles occur, i.e., the proportion of mismatches (38). The dendrogram was constructed from PCR-RFLP typing data by the unweighted pair group cluster method with arithmetic means in the computer program MEGA (28).

RESULTS

Figure 1 illustrates the RFLP and PFGE patterns obtained with nine isolates from the mother, the father, and two children of family A. These results are representative. Whenever possible, analyses were performed on isolates obtained from the antrum (a), corpus (c), and duodenum (d). Most patients harbored only a single H. pylori isolate. If a second isolate was detected, it differed only in a single restriction pattern, except for one isolate (H4a). An example is the isolate from the antrum of the mother (A1a), which varied from the corpus isolate A1c exclusively in the flaA-HhaI patterns. We considered isolate A1a to be a mutational variant of the parent isolate A1c.

Conspicuously, apparently identical isolates were very frequently found within families, where they were identified between siblings or between mother and child. The collective results of molecular typing are summarized in Fig. 2. For 59 H. pylori isolates, analysis of the 16S rRNA gene demonstrated an average of 2.5 RFLP patterns, a mean genetic diversity of 0.08, and a mean genotypic diversity of 0.14. The sparse allelic variation in the 16S RNA gene was as expected for a unique bacterial species (14). In contrast, the vacA, flaA, ureAB, and lspaA-glmM genes were highly polymorphic, with an average of 11.25 RFLP patterns per gene, a mean genetic diversity of 0.83, and a mean genotypic diversity of 0.91 for all of 59 H. pylori isolates. These values exceed the level of diversity recorded for all other bacterial species (19). The results of PFGE typing were concordant with the PCR-RFLP analysis in 46 of 52 cases. Three variants (A1a, D2a, and I2c) were distinct from their parent strains by PCR-RFLP analysis but showed PFGE patterns identical to those of the parent strains. Vice versa, three variants (B1c, H4a, and H4c) were distinguishable from their parent strains by PFGE but not by PCR-RFLP analysis. Hence, detection of strain variants is sometimes possible only by implementing both PCR-RFLP and PFGE analysis.

The genetic relationship among the 59 H. pylori isolates exhibited a tree-like phylogenetic structure with a deep but limited branching pattern (Fig. 2). Twenty-one clonal types clustered into 10 divisions (designated I to X). Isolates from one family tended to cluster into the same division. Isolates from different families were sometimes also found to be related, as evident in divisions II and IV.

Siblings tended to harbor identical isolates (families A, C, D, G, and H), but distinct H. pylori isolates could also be found between siblings (families D, E, F, and I). We were able to analyze H. pylori isolates from only one pair of parents (family A). The mother harbored the same isolate as her children, whereas the allelic composition of the parent strain was unrelated to the isolates. Intriguingly, the mother in family B also had the same isolate as her son. In contrast, the H. pylori strain of the father in family F was distinct from the isolates recovered in either of his children. These results indicate that H. pylori can be transmitted within families and possibly from parent (mother) to child or between siblings. However, a common source of exposure for H. pylori infection could not be completely ruled out.

The complete set of H. pylori isolates and subsets of the population were analyzed for multilocus linkage disequilibrium (Table 1). The 59 isolates differed on average at 6.80 of 10 loci examined, and the Iw was 0.43. The Monte Carlo procedure indicated that the Iw was significantly different from zero. Moreover, calculation of \(h_{ij}\) for the 21 PCR-RFLP types resulted in a value of 1.42, which was significantly different from zero (\(P < 0.001\)) based on the Monte Carlo procedure. These results, analyzed at the level of all isolates or RFLP types, demonstrated a significant degree of linkage disequilibrium among the H. pylori isolates from different families, consistent with the clonal structure described by Maynard Smith et al. (30).

DISCUSSION

Previous studies have suggested that H. pylori is panmictic and that the immense genetic diversity among H. pylori strains
FIG. 1. Representative molecular typing results for *H. pylori* isolates from family A. PCR-RFLP analyses of 16S rDNA fragments digested with *Hae*III (A), 16S rDNA fragments digested with *Hin*fi (B), *flaA* fragments digested with *Hha*I (C), *flaA* fragments digested with *Sau*3AI (D), *ureAB* fragments digested with *Hae*III (E), *lspA-glmA* fragments digested with *Hha*I (F), and *vacA* fragments digested with *Hph*I (G) and PFGE of genomic DNA digested with *Not*I (H) are shown. Lanes: M, size markers; A1a to A4c, *H. pylori* isolates, designated according to the origins of the isolates (A, family A; 1 to 4, family members; a, c, and d, antrum, corpus, and duodenum, respectively).
FIG. 2. Genetic relationships among 59 H. pylori isolates from nine different families. The dendrogram was constructed from PCR-RFLP typing data by the unweighted pair group cluster method with arithmetic means. Each distinct combination of 10 individual gene and enzyme patterns was designated a PCR-RFLP type, corresponding to a clonal type. Ten major lineages, termed I through X, separating at a genetic distance of 0.3 are indicated to the left. The columns to the right indicate the designation of H. pylori isolates, the origin of the isolates, and the molecular typing results. n.d., not determined; n.t., not typeable.
TABLE 1. Analysis of association between loci in 59 *H. pylori* isolates from family members

<table>
<thead>
<tr>
<th>Isolate group</th>
<th>Isolates</th>
<th>PCR-RFLP types only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>X</td>
</tr>
<tr>
<td>vacA s1 type</td>
<td>44</td>
<td>6.83</td>
</tr>
<tr>
<td>vacA s2 type</td>
<td>10</td>
<td>4.02</td>
</tr>
<tr>
<td>vacA m1 type</td>
<td>34</td>
<td>6.05</td>
</tr>
<tr>
<td>vacA m2 type</td>
<td>25</td>
<td>6.51</td>
</tr>
<tr>
<td>cagA positive</td>
<td>34</td>
<td>6.78</td>
</tr>
<tr>
<td>cagA negative</td>
<td>25</td>
<td>5.88</td>
</tr>
<tr>
<td>From antrum</td>
<td>27</td>
<td>7.13</td>
</tr>
<tr>
<td>From corpus</td>
<td>26</td>
<td>6.90</td>
</tr>
<tr>
<td>All</td>
<td>59</td>
<td>6.80</td>
</tr>
</tbody>
</table>

* X, mean number of allelic mismatches between pairs of *H. pylori* isolates or between PCR-RFLP types; *I₄*, index of association between loci; *P*, probability based on the Monte Carlo procedure. Values of *I₄* that differ significantly from zero indicate a clonal structure in bacterial populations, with rare or absent recombination. In contrast, values of *I₄* that do not differ significantly from zero indicate a panmictic structure, with frequent recombination.

*H. pylori* transmission occurs predominantly from parent (mother) to child, perhaps combined with the horizontal transmission between siblings. However, Rowland (36) has commented on the problems surrounding the interpretation of those data. The hypothesis that *H. pylori* transmission occurs predominantly from parents to child, and not vice versa, is supported by the observations that (i) *H. pylori* acquisition occurs predominantly in childhood and (ii) the prevalence of *H. pylori* infection in parents is much higher than that in children. Additionally, interspousal transmission of *H. pylori* occurs rarely (29, 40). Our molecular analysis of *H. pylori* suggests the vertical spread from parent (mother) to child, perhaps combined with the horizontal transmission between siblings, as the pathway for *H. pylori* transmission. If this proves to be correct, *H. pylori* would represent the first human pathogen recognized to display this remarkably restricted mode of transmission.

ACKNOWLEDGMENTS

We are indebted to J. G. Lörén, Departament de Microbiologia i Parasitologia Sanitàries, Facultat de Farmàcia, Divisió de Ciències, de la Salut, Universitat de Barcelona, Barcelona, Spain, for providing computer programs of linkage disequilibrium and helpful discussions. We acknowledge G. Rippin, Department of Medical Statistics and Documentation, University of Mainz, for helpful discussions on statistical methods. We thank D. E. Taylor and Q. Jiang, Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, Alberta, Canada, and B. Geilhausen, Department of Medical Microbiology and Hygiene, University of Cologne, Cologne, Germany, for helpful communications concerning the PFGE method.

REFERENCES


