A Comparison of 16S Ribosomal DNA Sequences from Five Isolates of Helicobacter pylori

BRUCE W. ECKLOFF, RAYMOND P. PODZORSKI, BRUCE C. KLINE,* and FRANKLIN R. COCKERILL III

Department of Biochemistry and Molecular Biology, and Department of Laboratory Medicine, Section of Clinical Microbiology, Mayo Clinic and Foundation, Rochester, Minnesota 55905

Other workers have found that clinical isolates of Helicobacter pylori exhibit very extensive DNA sequence polymorphisms when they are examined by ribotyping or some other genomic sequence characterization technique. In fact, it is rare to find similar clones, much less identical ones, among isolates. We found that the levels of divergence between the 16S ribosomal DNA sequences of individual organisms and the consensus sequence of the five isolates which we examined ranged from 0.2 to 0.5%. In contrast, other workers have shown that levels of divergence between the 16S ribosomal DNA sequence of H. pylori and the 16S ribosomal DNA sequences of four other Helicobacter species range from 2.7 to 8.0%. Our results show that the H. pylori 16S ribosomal DNA is not very polymorphic and support the conclusion that H. pylori is a unique species.

Strains of Helicobacter pylori, the suspected agent of type B gastritis and gastric and duodenal ulcers, appear to exhibit high levels of genetic variability. The possibility that such variability occurs was suggested by the results of clonal analyses of 134 isolates by molecular techniques (1, 3, 7, 8, 10). In the study of Taylor et al. (10), genomic DNAs from 20 isolates were examined for restriction fragment length polymorphism patterns and only two patterns matched. Perhaps more remarkably, the sizes of only a minority of the fragments matched. Similarly, Linton et al. (7) used ribotyping to show that none of the 12 strains that they examined were identical. These authors observed from 50 to 70% dissimilarity in fragment sizes. The results of these two studies are consistent with the results of Akopyanz et al. (1), who used random amplification of a single small primer to map the sequence sites in H. pylori genomes. In this study, 64 isolates were examined, and no 2 isolates gave the same pattern of amplicons. Again, the patterns of amplicon sizes were highly variable.

Beyond academic interests, several important, medically related questions are raised by these findings. One question concerns the utility of sequence information for identification. A second question is whether organisms clinically classified as H. pylori are actually H. pylori are correctly identified since the number of phenotypic parameters used in routine diagnostic examinations is limited. For example, any microaerophilic, gram-negative, curved organism that produces urease, oxidase, and catalase and is isolated from a human gut is classified as H. pylori. Might there be several species of human pathogens that fit this profile? We know that there are several Helicobacter species whose phenotypes fit this profile but are distinguished as separate species by other factors, and one of the major characteristics is 16S ribosomal DNA sequence data (4).

Bacterial phylogeny and identification are coming to rest squarely on ribosomal sequence information. Thus, the third important question is whether the extensive polymorphism observed in genomic sequences and ribotypes of H. pylori extends to 16S ribosomal DNA sequences and, if so, to what degree. To begin to answer to these questions, we amplified and sequenced the 16S ribosomal genes of five strains of H. pylori.

MATERIALS AND METHODS

Bacteria. Five isolates of H. pylori were examined. One isolate was strain ATCC 43504, and the other four isolates were clinical isolates recovered from endoscopic biopsies of patients evaluated at the Mayo Clinic in Rochester, Minn. The latter four isolates were designated MC123, MC268, MC903, and MC937. The nonidentity of the isolates was verified in our laboratories by Susan Wilkinson, who used pulsed-field gel electrophoresis performed as described by Taylor et al. (10) and genomic DNAs digested with NruI (data not shown).

DNA sequencing. Our sequencing strategy was to make 16S amplicons from purified H. pylori DNA by using universal primers 24 and pH (Table 1) and the PCR. DNA was purified from approximately 5 × 10^6 cells in the exponential phase of growth by using phenol-chloroform extraction (9). Each 50-μM PCR mixture contained 50 mM KCl, 10 mM Tris (pH 8.3), 1.5 mM MgCl2, 200 μM dATP, 200 μM dCTP, 200 μM dGTP, 200 μM dTTP, 50 pmol of each primer, 2.5 U of Taq polymerase (Perkin-Elmer Cetus), and 1 μg of extracted H. pylori DNA. The cycling conditions used were denaturation for 1 min at 94°C, annealing for 2 min at 50°C, and extension for 3 min at 72°C for 35 cycles, followed by a final extension for 10 min at 72°C. An identical mixture lacking H. pylori target DNA was used as a negative control in each amplification experiment. The 16S amplicons were purified by using PCR Magic Prep columns (Promega) according to the manufacturer’s instructions. A 5-μl portion of each purified amplicon was sequenced by using a dye terminator reagent kit, including Taq polymerase, and the protocol recommended by the manufacturer (Applied Biosystems, Inc., Foster City, Calif.); 3.2 pmol of sequencing primer was used per reaction mixture, and the cycling conditions used were denaturation at 96°C for 30 s, annealing at 50°C for 15 s, and extension for 4 min at 60°C for 25 cycles. The sequencing reaction products were separated from dye-labelled precursors by using Centri-Sep columns (Princeton Separations, Inc., Adelphia, N.J.) according to the manufacturer’s instructions. The sequencing reaction products were then analyzed with a model 373A automated DNA sequencer (Applied Biosystems, Inc.).

Nucleotide sequence accession numbers. The nucleotide
sequences of the isolates which we used have been deposited in the GenBank data base under the following accession numbers: strain MC123, U01328; strain MC238, U01329; strain ATCC 43504, U01330; strain MC937, U01331; and strain MC903, U01332. The 16s ribosomal DNA sequence of H. pylori ATCC 43504 sequence determined by Eaton et al. (4) was deposited in the GenBank data base under accession number M55305, and the sequences of the seven strains studied by Höök-Nikanne et al. (6) were deposited in the GenBank data base under accession numbers M55303 to M55309.

**RESULTS AND DISCUSSION**

H. pylori has been reported to contain three copies of the 16s ribosomal gene (10); we did not determine the numbers of copies in the isolates which we studied. To sequence the 16s ribosomal genes, we used universal PCR primers and genomic DNA to produce amplicons. The amplicons were then partially purified and sequenced by using dye-labelled dideoxy terminators, the Sanger technique, and an automated sequencer. Within the sensitivity of this method, all of the copies that were amplified appeared to be identical because we did not detect peaks that were obvious mixtures of bases. A possible exception was at position 1434, as discussed below.

Table 1 shows the consensus sequence determined for the five isolates. The sequence of each individual isolate can be deduced by altering the consensus sequence at the positions indicated in Table 2. As Table 2 shows, deviations from the consensus sequence were found at 13 separate positions. The number of deviations from the consensus sequence were two (strains ATCC 43504 and MC930), three (strains MC123 and MC937), and five (strain MC268). We found that strains MC123 and MC43504 share a common variation, a C → T transition at position 1105, while strains MC903 and MC937 also share a common variation, an A → G transversion, and the change at position 92 is a single-base deletion.

The 16s ribosomal DNA sequence of H. pylori ATCC 43504 is 97.4% identical to the sequence of Helicobacter acinonyx;
that is, the two sequences differ at almost 30 positions (4). It is interesting that because of mutations at positions 91, 94, and 95 strain MC268 has the same sequence as *H. acinonyx* (that is, AGAAGTGGA). Also, the mutation at position 1419 in strain MC268 changes a G to an A. A is also found at position 1419 in *H. acinonyx*. Thus, four of the five differences between the strain MC268 sequence and the consensus sequence are found in *H. acinonyx*. However, *H. acinonyx* still differs from strain MC268 at approximately 25 other positions.

Our results are largely in agreement with those of Eaton et al., whose unpublished *H. pylori* ATCC 43504 sequence has been deposited in the GenBank data base. These authors used rRNA and reverse transcriptase to make single-stranded DNA for sequencing. Our sequence differs from the sequence of Eaton et al. at two locations; C is missing from the sequence of Eaton et al. at our position 184, and we found a G at our position 727, whereas Eaton et al. found a T. The results of a reexamination by Eaton et al. of their raw data suggest that our results are correct (3a). In various experiments we observed an A or C at our position 1434, but Dewhirst and colleagues (3b) have shown that there is always a C at this position in *Helicobacter* species.

Previously, Höök-Nikanne et al. (6) sequenced 397 bp of 16S DNA of seven *H. pylori* strains isolated on three continents. The region which these authors sequenced corresponds to our coordinates 71 to 468 (Fig. 1). We examined the data of Höök-Nikanne et al. (which have been deposited in the GenBank data base) and compared their results with to our results. The data agreed completely with the following exceptions: (i) like Eaton et al., Höök-Nikanne et al. did not find C in position 184; (ii) at position 71 we found only G, and Höök-Nikanne et al. found only C; and (iii) Höök-Nikanne et al. reported three single-base-pair differences from the consensus sequence in two strains, at positions 91 (T → G), 129 (C → T), and 463 (A → T). We also found a change at position 91 in one strain; however, the change which we found was a deletion (Table 2).

In addition to the differences between our results and those of other workers, we were able to identify the unknown bases in GenBank sequence M88157 at our positions 913, 928 to 931, 30 bp at the 3' end of the 16S DNA sequence (Fig. 1, boldface type) beyond the bases reported in GenBank sequence M88157 even though we examined an additional 50 bp; we did not count the last 20 bp because this region represented primer H, which was used to make the amplon.

The results of a large number of studies have established the utility and validity of microbial phylogeny based on 16S DNA sequence data (see reference 11 for a review). For the most part, 16S DNA sequence data represent sequence information obtained from a single isolate of a species. Indeed, this type of data base should be useful for clinical laboratories as well. However, the following immediate questions occur: to what extent do clinical specimens deviate from the data base entry, and how much deviation is acceptable? Our results address the first question but not the second.

Deviation can be assessed by comparison with a type strain, by pairwise comparisons (Table 2), or by comparison with a consensus sequence (Table 2). While each of these approaches has its merits, comparison with a consensus sequence minimizes disparity and avoids distortion due to selection of an arbitrary standard. Presumably, the consensus sequence is closer to the *H. pylori* ancestral sequence than the sequence of any given isolate(s). Since the sample size in this study was small (five isolates), we doubt that our *H. pylori* consensus sequence is the final one. In fact, partial sequencing results for an additional five isolates indicated that A and G occur equally at our position 91. Thus, the consensus sequence seems to have a purine at this position. Our results clearly establish that 16S ribosomal DNA sequences in *H. pylori* are relatively stable and that the high level of genomic variability ascribed to this
organism does not extend to the 16S ribosomal genes. Indeed, the sequences of the five H. pylori strains differ from the consensus sequence by no less than 0.2% and no more than 0.5%; in contrast, five Helicobacter species, including H. pylori, H. acinonyx, Helicobacter felis, Helicobacter mustelae, and Helicobacter muridarum, differ in their 16S ribosomal DNA sequences by 2.7 to 8.0% (4). Thus, our data support the notion H. pylori is a separate species. Finally, we raised the possibility above that the morphological and physiological parameters used clinically to isolate and identify H. pylori might not be as discerning as molecular data. Thus far, our results do not support this possibility.

ACKNOWLEDGMENTS

We are deeply indebted to Floyd Dewhirst for his helpful discussions and generous sharing of data; his many kindnesses are highly appreciated. We also thank Susan Wilkinson for characterizing the Mayo Clinic H. pylori genomic restriction polymorphisms by pulsed-field gel electrophoresis.

REFERENCES

4a. Dewhirst, F. E. Personal communication.
4b. Dewhirst, F. E., K. A. Eaton, and S. Paster. Personal communication.