JB Journals.ASM.org

GENOME ANNOUNCEMENT

Draft Genome Sequences of *Helicobacter pylori* Strains 17874 and P79

Ceara D. Clancy,[a] Brian M. Forde,[a] Stanley A. Moore,[b] and Paul W. O'Toole[a,c]

Department of Microbiology, University College Cork, Cork, Ireland[a]; Department of Biochemistry, University of Saskatchewan, Saskatoon, Saskatchewan, Canada[b]; and Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland[c]

***Helicobacter pylori* is a human pathogen that colonizes the human gastric mucosa, causing gastritis, duodenal and gastric ulcers, and gastric carcinoma. Here we announce the draft genomes of *H. pylori* strain 17874, commonly used for studying motility, and P79, a strain for which plasmid vectors have been developed.**

Helicobacter pylori genomes sequenced to date exhibit significant variation. *H. pylori* CCUG 17874 was originally isolated from the gastric antrum of a patient in Perth, Australia, and is the type strain for the species (4) and is often used for flagellum biogenesis studies. P79 is a derivative of strain P1, transformed with 17874 chromosomal DNA to generate a streptomycin-resistant mutant (3). This readily transformable strain facilitates *in vivo* studies of *H. pylori*. The genomes of these strains were sequenced to provide a clearer genomic platform for investigation of *H. pylori* motility.

The *H. pylori* 17874 and P79 genomes were sequenced at the Beijing Genomics Institute (BGI) on the Illumina HiSeq platform, generating a paired-end library containing 20,154,284 and 13,298,804 reads of 90 bp, respectively. In a reference-guided assembly strategy using MIRA (version 3.2.1), reads for both genomes were mapped to the genomes of *H. pylori* 26695 (GenBank accession no. NC_000915) (5) and J99 (NC_000921.1) (1). A *de novo* assembly using Velvet was also performed and aligned to the MIRA assembly to close gaps. Strain 17874 and P79 contigs were assembled into 80 and 48 scaffolds, respectively. Protein coding regions were identified using the NCBI Prokaryotic Genome Automated Annotation Pipeline (PGAAP) and manually curated, with particular interest in flagellum-related genes. Predicted coding regions were identified with a minimum cutoff size of 30 amino acids.

H. pylori 17874 and P79 have genome sizes of 1,615,763 bp and 1,641,495 bp, respectively, and GC contents of 38.97% and 38.86%, respectively. Both strains are positive for *cagA* and *vacA*, well-described virulence factors (2). Strain-unique genes were identified using a pairwise bidirectional BLASTP comparison, where the query sequence has no detectable homologues. The 17874 genome contains 1,639 open reading frames, with 35, 45, and 24 unique genes that are absent in 26695, J99, and P79, respectively. Sixteen genes from 26695 and 6 genes from J99 are absent in strain 17874. *H. pylori* P79 contains 1,699 open reading frames, with 40, 52, and 36 unique genes that are absent in 26695, J99, and 17874, respectively. Twelve genes from 26695 and 6 genes from J99 are absent in P79. Twenty-one genes are unique to the 17874 and P79 genomes compared across these four strains.

The majority of strain-unique genes identified encode hypothetical protein products. Of note, strain 17874 possesses a unique type II restriction enzyme, and P79 possesses a unique hypothetical membrane protein that is absent in 26695 and J99. Strains 17874 and P79 lack metal-binding proteins present in both 26695 and J99 but possess Cag island protein B. All major flagellar and outer membrane proteins are present and intact in both 17874 and P79 compared to 26695 and J99. A hypothetical protein with predicted involvement in ATPase activity during flagellum biogenesis is absent in P79.

Nucleotide sequence accession numbers. The draft genome sequence of *H. pylori* 17874 has been deposited in GenBank, available through BioProject accession no. PRJNA76569 and project identification (ID) no. 76569. Similarly, the draft sequence of P79 is available in GenBank through BioProject accession no. PRJNA76567 and project ID no. 76567.

ACKNOWLEDGMENTS

This project was supported by a Research Frontiers Programme award (09_RFP_GEN2443) from Science Foundation Ireland to P.W.O. and by an Embark scholarship from IRCSET to C.D.C.

REFERENCES

1. **Alm RA, et al.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397**:176–180.
2. **Basso D, Plebani M, Kusters JG.** 2010. Pathogenesis of *Helicobacter pylori* infection. Helicobacter **15**:14–20.
3. **Heuermann D, Haas R.** 1998. A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation. Mol. Gen. Genet. **257**:519–528.
4. **Marshall BJ, Royce H, Annear D.** 1984. Original isolation of *Campylobacter pyloris* from human gastric mucosa. Microbios Lett. **25**:83–88.
5. **Tomb J-F, et al.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388**:539–547.

Received 14 February 2012 Accepted 20 February 2012

Address correspondence to Paul W. O'Toole, pwotoole@ucc.ie.

doi:10.1128/JB.00230-12