Genome Sequences of Three *Helicobacter pylori* Strains Isolated from Atrophic Gastritis and Gastric Ulcer Patients in China

Yuanhai You,^a Lin Liu,^b Maojun Zhang,^a Xifang Han,^b Lihua He,^a Yuanfang Zhu,^b Peixiang Ni,^b and Jianzhong Zhang^a

State Key Laboratory for Infectious Disease Prevention and Control and National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, People's Republic of China,^a and BGI-Shenzhen, Shenzhen, People's Republic of China^b

***Helicobacter pylori* is a bacterial pathogen which can lead to several human gastric diseases. Here we describe the genome sequences of three strains isolated from atrophic gastritis and gastric ulcers patients in China. The data will permit genomic characterization of traits that may contribute to various gastric diseases.**

Helicobacter pylori is widely known to be the primary cause of several severe diseases of the digestive system, such as gastritis, peptic ulceration, gastric adenocarcinoma, and MALT (mucus-associated lymphoid tissue) lymphoma (11, 16). It is also presumed to have coevolved with its human host, and it therefore reflects ancient human migrations. A number of phylogenetic studies have subdivided *H. pylori* into several distinct biogeographic populations, including hpAfrica1, hpAfrica2, hpEastAsia, and hpEurope. East Asian type (hpEastAsia) is further classified into at least three subtypes: East Asian (hspEAsia), Pacific (hspMaori), and native American (hspAmerind). The East Asia subtype (hspEAsia) may be related to the high incidence of gastric cancer in East Asia. Recent studies illustrate the East Asian group differ greatly from the European group in electron transfer and redox reactions, which implies a model of adaptive evolution and selection through proteome diversification and modulation of translational fidelity (5, 9, 10, 13).

Genome sizes among clinical isolates of *H. pylori* vary considerably, with some showing differences up to 25%. To date, there are over 30 completed genomes available on GenBank. Among these strains, there are in total six strains isolated from Japan and Korea, belonging to the hspEAsia group (1–4, 6, 7, 9, 12, 14, 15). Despite the fact that China is one of the countries with highest incidence of gastrointestinal diseases, few Chinese isolates have been sequenced so far (8).

In this study, we sequenced draft genomes of three strains isolated from Heilongjiang province, including isolates from two patients with atrophic gastritis (HLJ193 and HLJ256) and one patient with gastric ulcer (HLJ271). These strains were shown have a high genomic diversity compared to isolates involved in chronic superficial gastritis and gastric cancer in our previous studies (17). The strains were cultured on Columbia agar base supplemented with 5% sheep blood, and DNAs were extracted as previously described. For each strain, whole-genome sequencing was performed using Illumina HiSeq 2000 by generating paired-end libraries (500 bp and 2 kb) following the manufacturer's instructions. The read lengths were 90 bp and 50 bp for each library, from which more than 100 Mb of high-quality data were generated. Next, the paired-end reads from the two libraries were *de novo* assembled into scaffolds. Gene prediction was performed using Glimmer. The tRNA genes were searched for with tRNAscan-SE2. rRNA genes were searched for with RNAmmer3. Protein BLAST4 was run using the translated coding sequences as a query against the reference sequence (*H. pylori* strain 51). The best result for

each BLAST search was imported as the gene annotation. Finally, we obtained 36 contigs, with a total length of 1,576,324 bp, and predicted 1,568 CDSs within the draft genome of strain HLJHP256. For strain HLJHP193, the contig number was 46, with a total length of 1,569,611 bp, and 1,596 CDSs were predicted. For strain HLJHP271, the contig number was 41, with a total length of 1,588,141 bp, and 1,633 CDSs were predicted. The average G+C content of the genome for each strain was 38.5%.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession numbers [ALJI000000000](#) (HLJ193), [ALKA000000000](#) (HLJ256), and [ALKB000000000](#) (HLJ271). The versions described in this paper are the first versions, ALJI000000000.1, ALKA000000000.1, and ALKB000000000.1.

ACKNOWLEDGMENT

This research was supported by a grant from the National Technology R&D Program in the 12th Five-Year Plan of China (no. 2012BAI06B02).

REFERENCES

1. Alm R, et al. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397: 176–180.
2. Avasthi T, et al. 2011. Genomes of the two chronological isolates (*Helicobacter pylori* 2017 and 2018) of the West African *Helicobacter pylori* strain 908, obtained from a single patient. *J. Bacteriol.* 193:3385–3386.
3. Baltrus D, et al. 2009. The complete genome sequence of *Helicobacter pylori* strain G27. *J. Bacteriol.* 191:447–448.
4. Devi S, et al. 2010. Genome of *Helicobacter pylori* strain 908. *J. Bacteriol.* 192:6488–6489.
5. Falush D, et al. 2003. Traces of human migrations in *Helicobacter pylori* populations. *Science* 299:1582–1585.
6. Farnbacher M, et al. 2010. Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC Genomics* 11:335.
7. Furuta Y, et al. 2011. Birth and death of genes linked to chromosomal inversion. *Proc. Natl. Acad. Sci. U. S. A.* 108:1501–1506.

Received 3 August 2012 Accepted 31 August 2012

Address correspondence to Jianzhong Zhang, zhangjianzhong@icdc.cn, or Peixiang Ni, nipx@genomics.cn.

Y.Y., L.L., M.Z., X.H., L.H., and Y.Z. contributed equally to this work.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.01399-12

8. Guo Y, et al. 2012. Genome of *Helicobacter pylori* strain XZ274, an isolate from a tibetan patient with gastric cancer in China. *J. Bacteriol.* **194**:4146–4147.
9. Kawai M, et al. 2011. Evolution in an oncogenic bacterial species with extreme genome plasticity: *Helicobacter pylori* East Asian genomes. *BMC Microbiol.* **11**:104.
10. Linz B, et al. 2007. An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* **445**:915–918.
11. Marshall B. 1994. *Helicobacter pylori*. *Am. J. Gastroenterol.* **89**:S116–S128.
12. McClain MS, et al. 2009. Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer. *BMC Genomics* **10**:3.
13. Moodley Y, et al. 2009. The peopling of the Pacific from a bacterial perspective. *Science* **323**:527–530.
14. Oh J, et al. 2006. The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. *Proc. Natl. Acad. Sci. U. S. A.* **103**:9999–10004.
15. Tomb J, et al. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**:539–547.
16. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, et al. 2001. *Helicobacter pylori* infection and the development of gastric cancer. *N. Engl. J. Med.* **345**:784–789.
17. You Y, et al. 2012. Comparative genomics of *Helicobacter pylori* strains of China associated with different clinical outcome. *PLoS One* **7**:e38528. doi: 10.1371/journal.pone.0038528.