genomeA
Journals.ASM.org

Genome Sequence and Annotation of *Helicobacter pylori* Strain Hp238, Isolated from a Taiwanese Patient with Mucosa-Associated Lymphoid Tissue Lymphoma

Cheng-Yen Kao,[a] Jenn-Wei Chen,[b] Yi-Ting Huang,[a] Shew-Meei Sheu,[c] Bor-Shyang Sheu,[c] Jiunn-Jong Wu[a,b]

Department of Medical Laboratory Science and Biotechnology, National Cheng Kung University, Tainan, Taiwan[a]; Center of Infectious Disease and Signaling Research, National Cheng Kung University, Tainan, Taiwan[b]; Department of Internal Medicine, College of Medicine, National Cheng Kung University, Tainan, Taiwan[c]

We present the complete genome sequence of *Helicobacter pylori* strain Hp238, isolated from a Taiwanese patient with gastric mucosa-associated lymphoid tissue lymphoma. Importantly, *H. pylori* strain Hp238 can multiply in THP-1 cells after internalization through the induction of autophagosome formation. These genome data will help to identify genes associated with *H. pylori* intracellular multiplication and pathogenesis.

Received 4 January 2015 **Accepted** 13 January 2015 **Published** 19 February 2015

Citation Kao C-Y, Chen J-W, Huang Y-T, Sheu S-M, Sheu B-S, Wu J-J. 2015. Genome sequence and annotation of *Helicobacter pylori* strain Hp238, isolated from a Taiwanese patient with mucosa-associated lymphoid tissue lymphoma. Genome Announc 3(1):e00006-15. doi:10.1128/genomeA.00006-15.

Address correspondence to Jiunn-Jong Wu, jjwu@mail.ncku.edu.tw.

Helicobacter pylori is a Gram-negative spiral-shaped microaerophilic bacterium that infects 50% of the population worldwide (1). Persistent infection with *H. pylori* increases the risk of developing gastroduodenal diseases, including peptic ulcer, duodenal ulcer, and gastric adenocarcinoma (2, 3). The most common histopathological features of gastric malignancies are adenocarcinoma and lymphoma of mucosa-associated lymphoid tissue (MALT). It has been calculated that the risk of gastric adenocarcinoma and MALT lymphoma in *H. pylori*-infected individuals is 3- to 6-fold higher than in those who are uninfected (4). Genome diversity, especially involving disease-associated genes, is highly associated with different gastroduodenal diseases, since *H. pylori* has coevolved with human populations (5). The study of whole-genome sequences from strains isolated from patients with MALT lymphoma is necessary to better understand the pathogenesis of the disease and to identify virulence genes useful in the identification of patients with higher risk.

We report the genome sequence of *H. pylori* strain Hp238, isolated from a 57-year-old Taiwanese man diagnosed with MALT lymphoma. In a previous study, we determined that Hp238 can multiply in THP-1 cells after internalization through the induction of autophagosomes formation (6). Although CagA and VacA participate in Hp238 multiplication in THP-1 cells, whether other virulence factors are involved in this multiplication remained unclear. This genome sequence will provide valuable information for a better understanding of bacterial pathogenesis. Particularly in Taiwan, this is the first report of a genome sequence of *H. pylori*.

The Hp238 strain was sequenced using the 454 junior platform (Roche, Germany), generating a library containing 252,098 single reads with an average length of 405 bp and a 33-fold average coverage. The genome sequence was generated by a *de novo* assembly using the GS Assembler version 2.6 software (Roche). This strategy provided 80 large contigs. As a reference, we employed the genome sequence of *H. pylori* strain B38 (NC_012973.1). The contigs were then remapped according to a previous study (7), and the results showed two scaffolds with 67 gaps. Finally, gaps were filled in through PCR and direct sequencing. The Hp238 genome has a size of 1,586,473 bp and a GC content of 38.7%.

Sequences were annotated at NCBI (http://www.ncbi.nlm.nih.gov/genome/annotation_prok; strain Hp238) or with the RAST prokaryotic genome annotation server (http://www.nmpdr.org/FIG/wiki/view.cgi/Main/RAST). RAST annotation results showed that the genome contained 92.3% coding regions and 1,663 genes, including 39 RNA genes. The average length for protein-coding genes was found to be 897 bp. The results showed the presence of major virulence markers, including *cagA* (typing of the repeating EPIYA motifs revealed the ABD type), *vacA* (typing of s and m regions, s1c/m2), and *babA*. Moreover, genome and gene comparison analysis by RAST with other fully and partially sequenced strains demonstrates that this genome is most similar to the *H. pylori* 51 strain isolated in eastern Asia (Korea), followed by *H. pylori* strains Shi470 and 35A.

Nucleotide sequence accession number. The complete genome sequence of strain Hp238 has been deposited in GenBank under the accession number CP010013.

ACKNOWLEDGMENTS

We thank Robert Jonas for helpful comments on this manuscript. We also thank Chien-Yi Tung and Jia-Ping Pan for helping on comparative genomic analysis.

This study was supported by grants NSC101-2320-B-006-020-MY3 and NSC101-2320-B- 006-029-MY3 from the National Science Council, Taiwan.

REFERENCES

1. **Calvet X, Ramirez Lázaro MJ, Lehours P, Mégraud F.** 2013. Diagnosis and epidemiology of *Helicobacter pylori* infection. Helicobacter **18**(Suppl 1):5–11. http://dx.doi.org/10.1111/hel.12071.
2. **Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH,**

Orentreich N, Sibley RK. 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med **325:**1127–1131. http://dx.doi.org/10.1056/NEJM199110173251603.
3. **Rauws EA, Tytgat GN.** 1990. Cure of duodenal ulcer associated with eradication of *Helicobacter pylori*. Lancet **335:**1233–1235. http://dx.doi.org/10.1016/0140-6736(90)91301-P.
4. **Kim SS, Ruiz VE, Carroll JD, Moss SF.** 2011. *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. Cancer Lett **305:** 228–238. http://dx.doi.org/10.1016/j.canlet.2010.07.014.
5. **Yamaoka Y, Graham DY.** 2014. *Helicobacter pylori* virulence and cancer pathogenesis. Future Oncol **10:**1487–1500. http://dx.doi.org/10.2217/fon.14.29.
6. **Wang YH, Wu JJ, Lei HY.** 2009. The autophagic induction in *Helicobacter pylori*-infected macrophage. Exp Biol Med (Maywood) **234:**171–180. http://dx.doi.org/10.3181/0808-RM-252.
7. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and psi-blast: a new generation of protein database search programs. Nucleic Acids Res **25:**3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.