

Genome Sequences of Three *Helicobacter pylori* Strains from Patients with Gastric Mucosa-Associated Lymphoid Tissue Lymphoma

Hsuan-Chen Wang,^a Feng-Chi Cheng,^b Ming-Shiang Wu,^c Hung-Yu Shu,^d H. Sunny Sun,^e Yu-Chun Wang,^a Ih-Jen Su,^a Chi-Jung Wu^{a,f}

National Institute of Infectious Diseases and Vaccinology^a and Institute of Population Health Sciences,^b National Health Research Institutes, Miaoli, Taiwan; Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan^c; Department of Bioscience Technology, Chang Jung Christian University, Tainan, Taiwan^d; Institute of Molecular Medicine, College of Medicine, National Cheng Kung University, Tainan, Taiwan^e; Department of Internal Medicine, National Cheng Kung University Hospital, Tainan, Taiwan^f

Most of the published complete genome sequences of *Helicobacter pylori* strains are limited to clinical isolates associated with gastritis, peptic ulcers, or gastric cancer. The genome sequences of three *H. pylori* strains isolated from patients with gastric mucosa-associated lymphoid tissue (MALT) lymphoma are presented here to facilitate studies of *H. pylori*-associated MALT lymphomagenesis.

Received 3 March 2015 Accepted 17 March 2015 Published 23 April 2015

Citation Wang H-C, Cheng F-C, Wu M-S, Shu H-Y, Sun HS, Wang Y-C, Su I-J, Wu C-J. 2015. Genome sequences of three *Helicobacter pylori* strains from patients with gastric mucosa-associated lymphoid tissue lymphoma. *Genome Announc* 3(2):e00229-15. doi:10.1128/genomeA.00229-15.

Copyright © 2015 Wang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Ih-Jen Su, suihjen@nhri.org.tw, or Chi-Jung Wu, wucj@nhri.org.tw.

Helicobacter pylori, a Gram-negative microaerophilic bacterium, has been implicated in a variety of human gastrointestinal diseases, including chronic gastritis, peptic ulcers, and gastric cancer (1). *H. pylori* also plays an oncogenic role in the development of mucosa-associated lymphoid tissue (MALT) lymphoma. Chronic immunological stimulation by *H. pylori* leads to the formation of acquired MALT in the gastric mucosa. These infiltrating B cells actively proliferate and occasionally undergo malignant transformation (2). While several virulence factors of *H. pylori* that are involved in the pathogenic processes of peptic ulcers and gastric cancer have been described, such as CagA and VacA, the virulence determinants associated with MALT lymphomagenesis remain undetermined (3, 4). Furthermore, with the exception of *H. pylori* B38, a gastric MALT lymphoma strain, most of the published complete genome sequences of *H. pylori* strains are limited to clinical isolates associated with gastritis, peptic ulcers, gastric cancer, or asymptomatic colonization (4). In the present report, the genome sequences of three *H. pylori* strains, HPML01, HPML02, and HPML03, which were isolated from three patients with gastric MALT lymphoma, were analyzed to facilitate studies on *H. pylori*-associated MALT lymphomagenesis.

Whole-genome sequencing of HPML01, HPML02, and

HPML03 was performed using 454 sequencing technology. Genomic shotgun and 8-kb paired-end libraries were constructed and sequenced with a 454-GS Junior instrument (Roche Diagnostics, Indianapolis, IN). The sequence reads of the three strains were assembled into 33, 55, and 45 contigs and 4, 2, and 4 scaffolds, respectively, using 454 Newbler (version 2.7; 454 Life Sciences, Branford, CT). All gaps between the contigs were filled by PCR and Sanger sequencing, and the complete genome sequences were assembled and analyzed using OMIGA 2.0 and the CG-Pipeline software (5). The final draft genomes consisted of circular chromosomes of 1,629,815 bp, 1,562,125 bp, and 1,629,114 bp in length for HPML01, HPML02, and HPML03, respectively, as well as a circular plasmid of 6,220 bp for HPML03. The pertinent statistics of the genomes are summarized in Table 1.

The genome sequences were annotated using the NCBI Basic Local Alignment Search Tool. HPLM01, HPLM02, and HPLM03 harbored genes encoding previously known virulence factors associated with peptic ulcers and gastric cancer, such as *cagA*, *cagE*, *vacA*, *iceA*, *babA*, *hopQ*, *oipA*, *sabA*, *hopZ*, *babB*, *babC*, *sabB*, and *homB*, while the last four genes were not present in strain B38 (3, 4). A whole-genome comparison using the OrthoMCL software (6) revealed that there are 9 genes shared by HPLM01, HPLM02,

TABLE 1 Summary of statistics for three sequenced gastric MALT lymphoma *H. pylori* strains

Strain	Genome content	Accession no.	Genome length (bp) ^a	Coverage depth (fold) ^b	G+C content (%) ^a	No. of predicted genes ^c	No. of coding sequences ^c
HPML01	Circular chromosome	AP014710	1,629,815	67	38.69	1,707	1,671
HPML02	Circular chromosome	AP014711	1,562,125	45	38.92	1,751	1,715
HPML03	Circular chromosome and circular plasmid	AP014712 (chromosome) and AP014713 (plasmid)	1,629,114 and 6,220	60	38.67	1,744	1,708

^a Determined by OMIGA.

^b Determined by 454 Newbler.

^c Determined by CG-Pipeline software.

and HPLM03 that are absent in the five published *H. pylori* strains (namely, 26695 [recovered from a patient with gastritis], HP J99 [duodenal ulcer], HPAG1 [atrophic gastritis], G27 [unspecified], and Shi470 [chronic gastritis]) (7–11), eight of which are hypothetical proteins. HPLM01, HPLM01, and HPLM03 also were found to share many gene substitution, deletion, and insertion sites that are not present in the five published strains. Although the biological significance of this genetic variability needs to be functionally validated, knowledge of the genome sequences of MALT lymphoma *H. pylori* strains opens new avenues for the further genomics-based exploration of virulence determinants contributing to MALT lymphomagenesis.

Nucleotide sequence accession numbers. The genome sequences of HPLM01, HPLM02, and HPLM03 were deposited into the whole-genome sequencing (WGS) database of DDBJ/EMBL/GenBank under the accession numbers listed in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work was supported by an intramural grant from the National Health Research Institutes, Taiwan (IV-101-SP-13).

We thank the Center for Genomic Medicine of the National Cheng Kung University for support of this project.

REFERENCES

1. Suerbaum S, Michetti P. 2002. *Helicobacter pylori* infection. *N Engl J Med* 347:1175–1186. <http://dx.doi.org/10.1056/NEJMra020542>.
2. Isaacson PG, Du MQ. 2004. MALT lymphoma: from morphology to molecules. *Nat Rev Cancer* 4:644–653. <http://dx.doi.org/10.1038/nrc1409>.
3. Lehours P, Ménard A, Dupouy S, Bergey B, Richy F, Zerbib F, Ruskoné-Fourmestreaux A, Delchier JC, Mégraud F. 2004. Evaluation of the association of nine *Helicobacter pylori* virulence factors with strains involved in low-grade gastric mucosa-associated lymphoid tissue lymphoma. *Infect Immun* 72:880–888. <http://dx.doi.org/10.1128/IAI.72.2.880-888.2004>.
4. Thiberge JM, Boursaux-Eude C, Lehours P, Dillies MA, Creno S, Coppée JY, Rouy Z, Lajus A, Ma L, Burucoa C, Ruskoné-Fourmestreaux A, Courillon-Mallet A, De Reuse H, Boneca IG, Lamarque D, Mégraud F, Delchier JC, Médigue C, Bouchier C, Labigne A, Raymond J. 2010. From array-based hybridization of *Helicobacter pylori* isolates to the complete genome sequence of an isolate associated with MALT lymphoma. *BMC Genomics* 11:368. <http://dx.doi.org/10.1186/1471-2164-11-368>.
5. Kislyuk AO, Katz LS, Agrawal S, Hagen MS, Conley AB, Jayaraman P, Nelakuditi V, Humphrey JC, Sammons SA, Govil D, Mair RD, Tatti KM, Tondella ML, Harcourt BH, Mayer LW, Jordan IK. 2010. A computational genomics pipeline for prokaryotic sequencing projects. *Bioinformatics* 26:1819–1826. <http://dx.doi.org/10.1093/bioinformatics/btq284>.
6. Li L, Stoeckert CJ, Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189. <http://dx.doi.org/10.1101/gr.1224503>.
7. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388:539–547. <http://dx.doi.org/10.1038/41483>.
8. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397:176–180. <http://dx.doi.org/10.1038/16495>.
9. Oh JD, Kling-Bäckhed H, Giannakis M, Xu J, Fulton RS, Fulton LA, Cordum HS, Wang C, Elliott G, Edwards J, Mardis ER, Engstrand LG, Gordon JI. 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. <http://dx.doi.org/10.1073/pnas.0603784103>.
10. Baltus DA, Amieva MR, Covacci A, Lowe TM, Merrell DS, Ottemann KM, Stein M, Salama NR, Guillemin K. 2009. The complete genome sequence of *Helicobacter pylori* strain G27. *J Bacteriol* 191:447–448. <http://dx.doi.org/10.1128/JB.01416-08>.
11. Kersulyte D, Kalia A, Gilman RH, Mendez M, Herrera P, Cabrera L, Velapatiño B, Balqui J, Paredes Puente de la Vega F, Rodriguez Ulloa CA, Cok J, Hooper CC, Dailide G, Tamma S, Berg DE. 2010. *Helicobacter pylori* from Peruvian Amerindians: traces of human migrations in strains from remote Amazon, and genome sequence of an Amerind strain. *PLoS One* 5:e15076. <http://dx.doi.org/10.1371/journal.pone.0015076>.