




# Complete Genome Sequence of *Helicobacter pylori* Strain 3401, a Suitable Host for Bacteriophages KHP30 and KHP40

Michiko Takahashi,<sup>a</sup> Yuki Matsumoto,<sup>a\*</sup> Takako Ujihara,<sup>b</sup> Hiromichi Maeda,<sup>c</sup> Kazuhiro Hanazaki,<sup>c</sup> Keizo Nagasaki,<sup>a,d</sup> Hiroaki Takeuchi,<sup>e</sup>  Shigenobu Matsuzaki<sup>c,f</sup>

<sup>a</sup>Faculty of Science and Technology, Kochi University, Nankoku, Kochi, Japan

<sup>b</sup>Science Research Center, Kochi University, Nankoku, Kochi, Japan

<sup>c</sup>Department of Surgery, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan

<sup>d</sup>Faculty of Agriculture and Marine Science, Kochi University, Nankoku, Kochi, Japan

<sup>e</sup>Department of Medical Laboratory Sciences, Health and Science, International University of Health and Welfare Graduate School, Chiba, Japan

<sup>f</sup>Department of Medical Laboratory Science, Faculty of Health Sciences, Kochi Gakuen University, Kochi, Kochi, Japan

**ABSTRACT** *Helicobacter pylori* 3401, isolated from a patient with duodenal ulcers in Japan, is susceptible to the bacteriophages KHP30 and KHP40. In this study, we report the complete genome sequence of *H. pylori* 3401. This study may lead to the establishment of phage therapy against *H. pylori* infection.

*Helicobacter pylori* is a Gram-negative spiral bacterium that colonizes the human gastric mucosa and is estimated to infect half of the world's population (1). With the emergence of antibiotic-resistant *H. pylori* strains, the development of bacteriophage therapy is desirable (2, 3). In 2012, two *H. pylori* bacteriophages (KHP30 and KHP40) that have a lytic effect on several strains of *H. pylori* were discovered (4, 5). *H. pylori* 3401 is used as a suitable host for these bacteriophages (4, 5). Here, we report the complete genome sequence of *H. pylori* 3401. This study has no ethical issues regarding research using humans or animals.

*H. pylori* 3401 was isolated from duodenal ulcer biopsy samples from a Japanese patient at Yamaguchi University Hospital, Yamaguchi, Japan, in 1991 (6). The isolated strain was cultured on BEV (*Brucella* supplemented with 10% [vol/vol] equine serum and 10  $\mu$ g/ml vancomycin) agar plates under microaerobic conditions (10% CO<sub>2</sub>) at 37°C (5). A single colony was subcultured in BEV broth, and genomic DNA was extracted using the phenol-chloroform method.

DNA sequencing was conducted using an Illumina MiSeq platform (San Diego, CA, USA) with v2 chemistry (300-bp paired-end reads) and a Nanopore GridION X5 platform with a FLO-MIN106 R9.41 flow cell (Oxford Nanopore Technologies, Oxford, UK), after the sequencing libraries were prepared using a Nextera DNA Flex library prep kit (Illumina) and a ligation sequencing kit (SQK-LSK109, Oxford Nanopore Technologies), respectively. The Illumina reads were trimmed using fastp v0.20.1 (7) (parameters:  $\geq$ Q30 is qualified,  $\geq$ 20 bp limited), and the Nanopore reads were trimmed using NanoFilt v2.7.1 (8) (parameters:  $\geq$ Q10 is qualified,  $\geq$ 1,000 bp limited). The statistics of the raw/trimmed reads are shown in Table 1. Both sets of trimmed reads were hybrid-assembled using Unicycler v0.4.8 (9); then, the assembled contig was checked for other plasmids and contaminants using metaplasmid SPAdes v3.13.1 (10) and BlobTools v1.0 (11) with default parameters. The circular structure of this genome was also confirmed using Bandage v0.8.1 (12). The resulting circular contig was polished using Pilon v1.23 (13) with default parameters. The coverage of the genome was 47 $\times$ . DFAST v1.2.4 (14) was used for auto-annotation.

The *H. pylori* 3401 assembly consisted of a circular chromosome of 1,591,853 bp (GC content, 39%). It contained 1,540 coding sequences (CDSs), four rRNAs, and 36

**Citation** Takahashi M, Matsumoto Y, Ujihara T, Maeda H, Hanazaki K, Nagasaki K, Takeuchi H, Matsuzaki S. 2021. Complete genome sequence of *Helicobacter pylori* strain 3401, a suitable host for bacteriophages KHP30 and KHP40. *Microbiol Resour Announc* 10:e00647-21. <https://doi.org/10.1128/MRA.00647-21>.

**Editor** Catherine Putonti, Loyola University Chicago

**Copyright** © 2021 Takahashi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Shigenobu Matsuzaki, [smatsuzaki@kochi-gu.ac.jp](mailto:smatsuzaki@kochi-gu.ac.jp).

\*Present address: Yuki Matsumoto, Department of Bacteriology, Osaka City University, Graduate School of Medicine, Abeno, Osaka, Japan.

**Received** 7 July 2021

**Accepted** 3 October 2021

**Published** 21 October 2021

**TABLE 1** Statistics of raw reads and quality controlled reads

| Type of reads    | No. of sequences | Sum of length (bp) | Minimum length (bp) | Avg length (bp) | Maximum length (bp) | Nanopore $N_{50}$ (bp) |
|------------------|------------------|--------------------|---------------------|-----------------|---------------------|------------------------|
| Illumina_raw     | 795,261          | 122,023,362        | 35                  | 153.4           | 156                 |                        |
| Illumina_trimmed | 780,932          | 116,710,771        | 31                  | 149.5           | 152                 |                        |
| Nanopore_raw     | 52,944           | 116,384,428        | 1                   | 2,198.3         | 28,359              | 6,084                  |
| Nanopore_trimmed | 21,908           | 59,325,730         | 1,000               | 2,707.9         | 27,925              | 6,673                  |

tRNA genes. Multilocus sequence typing (MLST) in PubMLST (<https://pubmlst.org/>) identified known allele numbers only for the *efp* gene as 2220, and all four sequence types containing allele 2220 in the *efp* gene (ST2789, ST2798, ST2824, and ST2845) were isolated in Japan, while previously undescribed alleles were detected for the genes *atpA*, *mutY*, *ppa*, *trpC*, *urel*, and *yphC*. These results will contribute to elucidating the infection mechanism of bacteriophages KHP30 and KHP40 against *H. pylori* and help to develop therapeutic strategies.

**Data availability.** The complete genome sequence of *H. pylori* 3401 has been deposited in GenBank under the accession number [AP024599](https://www.ncbi.nlm.nih.gov/nuccore/AP024599). The raw data are available in the DDBJ Sequence Read Archive (DRA) under the accession number [DRA011836](https://www.dra-ncbi.nlm.nih.gov/DRAS/DRAS011836).

## ACKNOWLEDGMENTS

We are grateful to Makoto Taniguchi, Oral Microbiome Center, Taniguchi Dental Clinic, Takamatsu, Kagawa, Japan, for technical support.

We thank the Research Instrument and Radio-isotope Research, Division of Biological Research, Science Research Center, Kochi University, for the use of the research instruments. This study was partially supported by grants-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (grant number 20K15582 to M.T.).

## REFERENCES

- Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. 2001. *Helicobacter pylori* infection and the development of gastric cancer. *N Engl J Med* 345: 784–789. <https://doi.org/10.1056/NEJMoa001999>.
- Vitor JMB, Vale FF. 2011. Alternative therapies for *Helicobacter pylori*: probiotics and phytomedicine. *FEMS Immunol Med Microbiol* 63:153–164. <https://doi.org/10.1111/j.1574-695X.2011.00865.x>.
- Lin DM, Koskella B, Lin HC. 2017. Phage therapy: an alternative to antibiotics in the age of multi-drug resistance. *World J Gastrointest Pharmacol Ther* 8:162–173. <https://doi.org/10.4292/wjgpt.v8.i3.162>.
- Uchiyama J, Takeuchi H, Kato S-I, Takemura-Uchiyama I, Ujihara T, Daibata M, Matsuzaki S. 2012. Complete genome sequences of two *Helicobacter pylori* bacteriophages isolated from Japanese patients. *J Virol* 86: 11400–11401. <https://doi.org/10.1128/JVI.01767-12>.
- Uchiyama J, Takeuchi H, Kato S-I, Gamoh K, Takemura-Uchiyama I, Ujihara T, Daibata M, Matsuzaki S. 2013. Characterization of *Helicobacter pylori* bacteriophage KHP30. *Appl Environ Microbiol* 79:3176–3184. <https://doi.org/10.1128/AEM.03530-12>.
- Karita M, Kouchiyama T, Okita K, Nakazawa T. 1991. New small animal model for human gastric *Helicobacter pylori* infection: success in both nude and euthymic mice. *Am J Gastroenterol* 86:1596–1603.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Laetsch DR, Blaxter ML. 2017. BlobTools: interrogation of genome assemblies. *F1000Res* 6:1287. <https://doi.org/10.12688/f1000research.12232.1>.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualisation of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.