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

Michiko Takahashi,a Yuki Matsumoto,a Takako Ujihara,b Hiromichi Maeda,b Kazuhiro Hanazaki,c Keizo Nagasaki,a,d Hiroaki Takeuchi,e Shigenobu Matsuzaki,f

aFaculty of Science and Technology, Kochi University, Nankoku, Kochi, Japan
bScience Research Center, Kochi University, Nankoku, Kochi, Japan
cDepartment of Surgery, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan
dFaculty of Agriculture and Marine Science, Kochi University, Nankoku, Kochi, Japan
eDepartment of Medical Laboratory Sciences, Health and Science, International University of Health and Welfare Graduate School, Chiba, Japan
fDepartment 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 μg/ml vancomycin) agar plates under microaerobic conditions (10% CO2) 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 XS 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: ≥Q30 is qualified, ≥20 bp limited), and the Nanopore reads were trimmed using NanoFilt v2.7.1 (8) (parameters: ≥Q10 is qualified, ≥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×. 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...
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, ureI, 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. The raw data are available in the DDBJ Sequence Read Archive (DRA) under the accession number DRA011836.

**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, Koki 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**


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

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