



# Draft Genome Sequence of the Polychlorinated Biphenyl Degradar *Comamonas testosteroni* Strain YAZ2, Isolated from a Natural Landscape in the Tohoku Region of Japan

Tomijiro Hara,<sup>a</sup> Yumiko Takatsuka,<sup>a</sup> Yuh Shiwa,<sup>b</sup> Kenji Yokota<sup>c</sup>

<sup>a</sup>Environmental Microbiology Research Section, Laboratory for Complex Energy Processes, Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto, Japan

<sup>b</sup>Department of Molecular Microbiology, Tokyo University of Agriculture, Setagaya, Tokyo, Japan

<sup>c</sup>Department of Agricultural Chemistry, Tokyo University of Agriculture, Setagaya, Tokyo, Japan

**ABSTRACT** We report a draft genome sequence of *Comamonas testosteroni* strain YAZ2, a polychlorinated biphenyl (PCB) degrader that was isolated from a PCB-unpolluted environment. The assembled genome contains a single 5.4-Mb chromosome and an 87-kb plasmid. The *bph* gene cluster, which is involved in PCB degradation, was found on the chromosome.

*Comamonas testosteroni* strain YAZ2 is a polychlorinated biphenyl (PCB) degrader that was isolated from PCB-uncontaminated soil, collected in Yonezawa (37°54' 3.76"N, 140°6' 13.01"E) in the Tohoku region of Japan, by enrichment culture using biphenyl as the sole carbon source (1). W-minimal medium was used as the isolation medium, as described by Kimbara et al. (2). This strain was shown to degrade 5 mg/L PCBs with 50% 3-chlorine substitution by 83.1 ± 1.5% (1).

Genomic DNA was extracted from YAZ2 cells freshly cultured in W-medium supplied with biphenyl using a bacterial DNA preparation-solution kit (Jena Bioscience, Jena, Germany). The extracted DNA was confirmed to be of sufficient quality for next-generation sequencing (NGS) using a Thermo Fisher Scientific NanoDrop spectrophotometer (purity), Invitrogen Qubit fluorometer (double-stranded DNA [dsDNA]), and Agilent 2200 TapeStation system (electrophoresis) (Hokkaido System Science Co., Ltd., Hokkaido, Japan). The genome was sequenced using (i) an Illumina HiSeq platform with 350-bp paired-end and 6- to 10-kb mate-pair libraries (Hokkaido System Science Co., Ltd.) and (ii) a Roche 454 GS Junior sequencer (Roche Diagnostics K.K., Tokyo, Japan). The Illumina paired-end and mate-pair libraries yielded 21,847,630 and 10,944,900 paired-end 100-bp reads, respectively, and the Roche 454 library yielded 202,587 reads, with an average read length of 444 bp. Sequence assembly of the YAZ2 genome was performed cooperatively by Bits Co., Ltd. (Tokyo, Japan). Default parameters were used for all software unless otherwise specified. The raw reads from Illumina sequencing were trimmed using Cutadapt v1.1 (3) and Trimmomatic v0.32 (4) and subsequently assembled using Velvet v1.2.10 (5), resulting in 55 contigs/scaffolds, amounting overall to 5,661,919 bp, with an  $N_{50}$  value of 4,645,585 bp. The remaining gaps were filled by manual recursive mapping of Illumina paired-end reads against the contigs/scaffolds (minimum number of reads, 10) (6). Further scaffolding by manual mapping of Illumina mate-pair reads linked 3 of the 55 contigs/scaffolds into 1 scaffold, and the remaining unlinked contigs were removed from the assembly. Because the resulting assembly did not contain any plasmid sequences, Roche 454 sequence reads were assembled using the NewblerGS *de novo* assembler v3.0, yielding 118 contigs. Among these contigs, 1 contig (87 kb), with the start and end overlapping with 1 Illumina contig, showed the highest similarity, based on BLASTN analysis, to the plasmid sequence of *Acidovorax* sp. strain JS42 (GenBank accession number NC\_008765), belonging to the same family as *C. testosteroni*, i.e., *Comamonadaceae*. Thus, this contig

**Editor** J. Cameron Thrash, University of Southern California

**Copyright** © 2022 Hara 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 Tomijiro Hara, t-hara@iae.kyoto-u.ac.jp.

The authors declare no conflict of interest.

**Received** 24 August 2021

**Accepted** 16 December 2021

**Published** 13 January 2022

was determined to be a plasmid sequence. Annotation analysis was performed to identify the proteins and estimate their functions using the DFAST v1.4.0 pipeline with default settings, excluding the Prodigal and tRNAscan-SE options (<https://dfast.ddbj.nig.ac.jp>) (7).

Overall, the draft genome sequence of YAZ2 contained a single chromosome in one scaffold (5,448,456 bp, with a gap ratio of 0.38%, a G+C content of 61.6%, and coverage of 607.9×) generated from the Illumina data and one plasmid sequence (87,186 bp, with a G+C content of 62.4% and coverage of 1,032.5×) generated from Roche 454 sequence data. It also contained 5,001 putative protein-encoding genes, 4 rRNAs, and 79 tRNAs; the *bph* operon (*bphEGFA1A2A3BCDA4*), the gene cluster involved in PCB degradation, was found on the chromosome.

**Data availability.** This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers [AP024703](https://accession.ddbj.ac.jp/entry/CP024703) (chromosome) and [AP024827](https://accession.ddbj.ac.jp/entry/CP024827) (plasmid). The raw sequence data have been deposited under accession numbers [DRA012593](https://accession.ddbj.ac.jp/entry/DRA012593) and [DRA012754](https://accession.ddbj.ac.jp/entry/DRA012754).

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI grant numbers JP26550066 and JP18K19881.

## REFERENCES

1. Hara T, Takatsuka Y. Aerobic polychlorinated biphenyl-degrading bacteria isolated from the Tohoku region of Japan are not regionally endemic. *Can J Microbiol*, in press. <https://doi.org/10.1139/cjm-2021-0056>.
2. Kimbara K, Hashimoto T, Fukuda M, Koana T, Takagi M, Oishi M, Yano K. 1988. Isolation and characterisation of a mixed culture that degrades polychlorinated biphenyls. *Agric Biol Chem* 52:2885–2891. <https://doi.org/10.1080/00021369.1988.10869161>.
3. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>.
4. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
5. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
6. Hara T, Takatsuka Y, Shiwa Y, Yokota K, Shin-i T. 2021. Manual recursive mapping method for the draft genome sequence of *Comamonas testosteroni* strain YAZ2. *figshare* <https://doi.org/10.6084/m9.figshare.17132309.v7>.
7. 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>.