



# Draft Genome of *Kangiella* sp. Strain TOML190, Isolated from the Surface of the Striped Shore Crab, *Pachygrapsus crassipes*

Abdullah Adham Zulmajdi,<sup>a</sup> Tatsuru Okazaki,<sup>a</sup>  Tetsushi Mori<sup>a</sup>

<sup>a</sup>Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

**ABSTRACT** *Kangiella* sp. strain TOML190 is a strain from the *Kangiella* genus that was isolated from the surface of a crustacean. Genetic background analysis of this strain shows that it harbors unique features possibly related to its symbiotic adaptation to its residing host.

*Kangiella* are unique for their ability to produce iso-branched fatty acids (1–3) and have been used to understand bacterial genome reduction (4). Thus far, 11 strains have been reported (3), and a majority of the strains were isolated from marine environments, apart from *Kangiella spongicola* (5). Here, we present the draft genome sequence of *Kangiella* sp. strain TOML190, a strain showing 96.9% 16S rRNA gene similarity to its closest relative, *Kangiella* sp. strain HZ709 (GenBank number [MN330026.1](#)) (Fig. 1). This bacterium is unique in this genus, as it was isolated from the surface of the striped shore crab, *Pachygrapsus crassipes*.

Upon collection from Ohama Beach, Kanagawa, Japan, in 2020, *P. crassipes* was immediately rinsed with filter-sterilized seawater to remove surface contaminants. A dental water flosser was used to vigorously detach microbes from the crab's surface, and the collected supernatant was plated on marine broth (MB) agar for 3 days at 30°C. Out of 24 randomly selected colonies from the MB agar plates, 8 were identified as strain TOML190 via 16S rRNA gene analysis and were cryopreserved. A single colony was revived in liquid MB and cultivated with agitation at 30°C, and genomic DNA was extracted using a DNeasy UltraClean microbial kit (Qiagen). A genomic library was prepared using the rapid barcoding kit (Oxford Nanopore Technologies [ONT]), and sequencing was performed on a MinION Mk1B device with a FLO-MIN106D flow cell. The raw sequencing data were base called with Guppy v4.3.4 (ONT) using the high-accuracy mode, and the length and quality score of the sequences were obtained using NanoPlot v1.36.1 (6). The sequenced reads were filtered (length, >1,000 bp; Q-score, >12) using NanoFilt v2.5.0 (6). In total, 144,684 reads (average length, 3,405 bp) were obtained.

*De novo* assembly was performed using Flye v2.8.3 with –iterations 3 and –plasmids as options (7), and polishing was performed using Medaka v1.3.2 (ONT). The assembled genome was a single linear contig of 2,454,686 bp with GC content of 43.4% and 50× genome coverage. No plasmids were identified. BUSCO assessment (v5.2.2) was conducted against the *Oceanospirillales* lineage data set (8) to validate the quality of the genome. The BUSCO score (92.6%) and the genome length showing similarity to other *Kangiella* strains suggest that a near-complete draft genome sequence was attained. Gene annotation was performed using the RAST server (9), resulting in 2,332 coding sequences. Finally, annotated genes were analyzed using KofamKOALA (10), and metabolic pathways were identified using KEGG Mapper (11).

Features of strain TOML190 were highlighted by comparing its genome with those of other *Kangiella* species from the KEGG database. Primary pathway comparison showed that it lacks the sulfate permease and harbors a biofunctional sulfate adenyltransferase-adenylsulfate kinase. It also harbors genes (*cyoABCDE*) encoding a cytochrome *o*-ubiquinol

**Editor** Frank J. Stewart, Montana State University

**Copyright** © 2022 Zulmajdi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

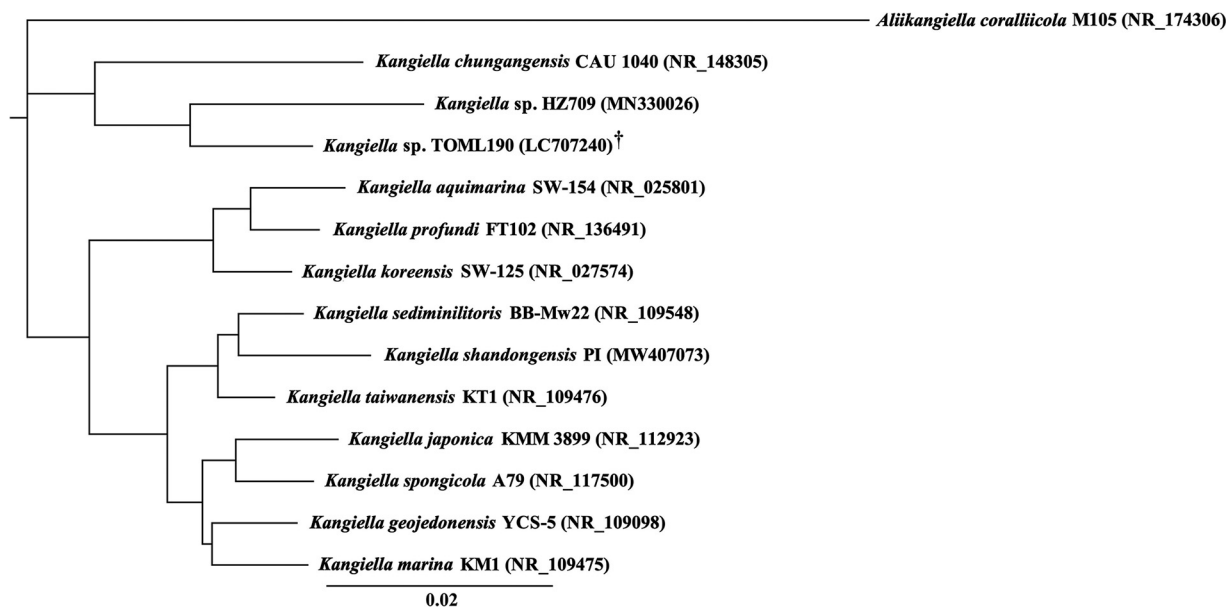
Address correspondence to Tetsushi Mori, moritets@go.tuat.ac.jp.

The authors declare no conflict of interest.

**Received** 17 May 2022

**Accepted** 21 July 2022

**Published** 4 August 2022



**FIG 1** 16S rRNA gene similarity of *Kangiella* sp. TOML190 (†) with currently reported *Kangiella* strains. The 16S rRNA gene of TOML190 was extracted from the draft genome presented in this work. The phylogenetic tree was generated using Geneious Prime v2020.2.5 based on the neighbor-joining tree and Tamura-Nei genetic distance model. The scale bar represents a 2% nucleotide sequence difference. *Kangiella* sp. HZ709 has not been published but was included in the tree generation due to its highest similarity to strain TOML190. *Aliikangiella coralliicola* was used as the outgroup. Accession numbers are derived from recently curated versions of the gene sequences.

oxidase that are expressed at high oxygen levels (12). These genes were not identified in currently reported *Kangiella* strains. The elucidated features of strain TOML190 suggest that it showed possible signs of adaptation to its residing host. We believe that *Kangiella* sp. TOML190 may serve as an important candidate to further understand the role of *Kangiella* in symbiotic settings.

**Data availability.** This whole-genome sequencing project and the 16S rRNA gene sequence of *Kangiella* sp. TOML190 have been deposited at DDBJ/EMBL/GenBank under the accession numbers [BQYL01000001](https://doi.org/10.1093/bioinformatics/bty149) and [LC707240](https://doi.org/10.1093/bioinformatics/bty149), respectively. The ONT base-called FASTQ files are available in the NCBI Sequence Read Archive (SRA) under the accession number [DRA014096](https://doi.org/10.1093/bioinformatics/bty149).

## ACKNOWLEDGMENTS

This work was financially supported by the Japan Science and Technology (JST) Agency, under the Fusion Oriented Research for Disruptive Science and Technology Program and the Grant in Aid for Scientific Research (B) (JSPS KAKENHI grant number JP21H02280), both awarded to T.M.

## REFERENCES

- Jean WD, Huang SP, Chen JS, Shieh WY. 2012. *Kangiella taiwanensis* sp. nov. and *Kangiella marina* sp. nov., marine bacteria isolated from shallow coastal water. *Int J Syst Evol Microbiol* 62:2229–2234. <https://doi.org/10.1099/ijs.0.037010-0>.
- Kim JH, Ward AC, Kim W. 2015. *Kangiella chungangensis* sp. nov. isolated from a marine sand. *Antonie Van Leeuwenhoek* 107:1291–1298. <https://doi.org/10.1007/s10482-015-0423-5>.
- Peng L-Y, Yin R, Gao S-K, Jiang H-N, Liu X-X, Ma Y, Zhou Y-X. 2021. *Kangiella shandongensis* sp. nov., a novel species isolated from saltern in Yantai, China. *Antonie Van Leeuwenhoek* <https://doi.org/10.1007/s10482-021-01682-9>.
- Wang J, Lu Y, Nawaz MZ, Xu J. 2018. Comparative genomics reveals evidence of genome reduction and high extracellular protein degradation potential in *Kangiella*. *Front Microbiol* 9:1224. <https://doi.org/10.3389/fmicb.2018.01224>.
- Ahn J, Park JW, McConnell JA, Ahn YB, Häggblom MM. 2011. *Kangiella spongicola* sp. nov., a halophilic marine bacterium isolated from the sponge *Chondrilla nucula*. *Int J Syst Evol Microbiol* 61:961–964. <https://doi.org/10.1099/ijs.0.021733-0>.
- 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>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Manni M, Berkeley MR, Seppely M, Simao FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654. <https://doi.org/10.1093/molbev/msab199>.

9. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
10. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36:2251–2252. <https://doi.org/10.1093/bioinformatics/btz859>.
11. Kanehisa M, Sato Y, Kawashima M. 2021. KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci* 31:47–53. <https://doi.org/10.1002/pro.4172>.
12. Tseng CP, Albrecht J, Gunsalus RP. 1996. Effect of microaerophilic cell growth conditions on expression of the aerobic (*cyoABCDE* and *cydAB*) and anaerobic (*narGHJI*, *frdABCD*, and *dmsABC*) respiratory pathway genes in *Escherichia coli*. *J Bacteriol* 178:1094–1098. <https://doi.org/10.1128/jb.178.4.1094-1098.1996>.