

Short Research Communication

# Draft Genome Sequence of *Roseovarius* sp. A-2, an Iodide-Oxidizing Bacterium Isolated from Natural Gas Brine Water, Chiba, Japan

Tri Yuliana<sup>1,2</sup>, Nobuyoshi Nakajima<sup>3</sup>, Shigeki Yamamura<sup>3,4</sup>, Masaru Tomita<sup>5,6</sup>, Haruo Suzuki<sup>5,6</sup> and Seigo Amachi<sup>1</sup>✉

1. Graduate School of Horticulture, Chiba University, Matsudo City, Chiba, Japan;
2. Food Technology Department, Padjadjaran University, Bandung, Indonesia;
3. Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan;
4. Center for Regional Environmental Research, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan;
5. Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan;
6. Systems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa, Kanagawa, Japan.

✉ Corresponding author: Seigo Amachi. Graduate School of Horticulture, Chiba University, 648 Matsudo, Matsudo City, Chiba 271-8510, Japan. Tel: +81-47-308-8867; Fax: +81-47-308-8867; Email: amachi@faculty.chiba-u.jp

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Published: 2017.04.06

## Abstract

*Roseovarius* sp. A-2 is a heterotrophic iodide (I<sup>-</sup>)-oxidizing bacterium isolated from iodide-rich natural gas brine water in Chiba, Japan. This strain oxidizes iodide to molecular iodine (I<sub>2</sub>) by means of an extracellular multicopper oxidase. Here we report the draft genome sequence of strain A-2. The draft genome contained 46 tRNA genes, 1 copy of a 16S-23S-5S rRNA operon, and 4,514 protein coding DNA sequences, of which 1,207 (27%) were hypothetical proteins. The genome contained a gene encoding *loxA*, a multicopper oxidase previously found to catalyze the oxidation of iodide in *Iodidimonas* sp. Q-1. This draft genome provides detailed insights into the metabolism and potential application of *Roseovarius* sp. A-2.

Key words: *Roseovarius* sp. A-2, draft genome, iodide oxidation, *loxA*.

## Introduction

*Roseovarius* sp. A-2 is a heterotrophic iodide (I<sup>-</sup>)-oxidizing bacterium, and was isolated from natural gas brine water in Chiba, Japan, which contained very high concentration (0.7 mM) of iodide [1]. Based on 16S rRNA gene sequence analysis, iodide-oxidizing bacteria are divided into two distinct groups within the class Alphaproteobacteria. One of the groups is most closely related to *Roseovarius tolerans* and *Roseovarius mucosus* with sequence similarities of 94 to 98%. The other group is closely related to the genera *Rhodothalassium* and *Kordiimonas*, and the newly proposed *Iodidimonas* gen. nov [2]. Shiroyama *et al.* [3] recently characterized an iodide-oxidizing enzyme of *Roseovarius* sp. A-2. It was

an extracellular protein, and had significant oxidizing activities not only for iodide but also for *p*-phenylenediamine and hydroquinone. Tandem mass spectrometric analysis of this enzyme revealed that it is homologous to *loxA*, a multicopper oxidase previously found as a component of iodide-oxidizing enzyme of *Iodidimonas* sp. Q-1 [4]. *loxA* is a novel type of bacterial multicopper oxidase, and expected to be used as an enzyme-based antimicrobial system due to its strong molecular iodine (I<sub>2</sub>)-producing activity [5]. We have generated a draft genome sequence of *Roseovarius* sp. A-2 to better understand the metabolism of this strain and to provide insights into future practical applications of iodide-oxidizing

bacteria, in general.

*Roseovarius* sp. A-2 was grown in Marine Broth 2216 (Becton Dickinson, Sparks, MD) and DNA was extracted using a DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany). Whole-genome sequencing was performed using paired-end sequencing on an Illumina MiSeq. The sequencer produced 300-bp paired-end reads that were obtained from 550-bp inserts. The quality of the reads was checked using FastQC [6]. PhiX contaminations [7] were removed using bbdduk [8], and the PhiX free reads were trimmed using Trimmomatic [9]. The reads were assembled using SPAdes version 3.9.0 [10]. After the removal of low-coverage contigs, the resulting assembly contains 126 contigs consisting of 4,584,578 bp, with a G+C content of 62.8 %. The genome size and G+C content for strain A-2 were within the range of those for other *Roseovarius* spp. including *R. tolerans* (3.7 Mb and 63.9%), *R. mucosus* (4.2 Mb and 61.9%), and *R. indicus* (5.5 Mb and 64.8%) [11-13]. Genome annotation was performed using Prokka v1.11, which is a pipeline comprising several bioinformatic tools [14]. Briefly, Aragorn [15] detected 46 tRNA genes, Barrnap predicted 1 copy of a 16S-23S-5S rRNA operon, and Prodigal [16] identified 4,514 protein coding DNA sequences (CDS), of which 400 contained signal peptides identified using SignalP [17]. Of the 4,514 proteins, 1,207 were hypothetical proteins of unknown function, 2,491 were annotated by UniProtKB [18], 621 by Pfam [19], 255 by NCBI's CDD (Conserved Domain Database) [20], and 21 by HAMAP [21].

The genome contained four multicopper oxidase genes. Among these, one was closely related to *ioxA*, which was previously found to be involved in iodide oxidation in *Iodidimonas* sp. Q-1 [4]. A key gene encoding phosphofructokinase (*pfk*) for glycolysis was not detected, but a complete set of genes for the TCA cycle and pentose phosphate pathway were predicted. The inability to grow on glucose has been observed commonly in *Roseovarius* spp., and this was probably due to lack of *pfk* [22, 23]. Genes involved in starch, sucrose, and galactose metabolism were also absent. The genome contained nitrate reductase (*nar*), but not nitrite reductase (*nir*) or nitrogenase (*nif*) genes. The genome did not contain the photosynthetic gene cluster (*puf*, *bch*, and *crt* genes), which has been found in *Iodidimonas* sp. Q-1 [24] and other *Roseovarius* spp. [11, 12]. There were no autotrophic CO<sub>2</sub> fixation pathway genes such as *rbc* genes. The genome contained at least 39 genes coding for cytochrome *c*. Nearly complete sets of genes for the flagellar system (*flg*, *fli*, and *mot*), type II secretion system (*gsp*), type IV secretion system (*virB*), general secretion system (*sec*),

and twin-arginine translocation pathway (*tat*) were identified. A wide variety of ABC transporters, including those involved in sulfate, tungstate, molybdate, zinc, iron(III), phosphate, phosphonate, taurine, spermidine, putrescine, glycine betaine, sorbitol/mannitol, glycerol, *sn*-glycerol 3-phosphate, and urea transport were identified. Various proteins involved in aerobic metabolism, such as NADH dehydrogenase (*nuo*), succinate dehydrogenases (*sdh* and *frd*), cytochrome-*c* oxidase (*cbb<sub>3</sub>*-type and quinol oxidase), catalase-peroxidase (*katG*), and superoxide dismutase (*sodB*) were identified. These results suggest that *Roseovarius* sp. A-2 is an aerobic heterotrophic iodide (I<sup>-</sup>)-oxidizing bacterium, and that it is well adapted to brine water environments.

## Nucleotide sequence accession numbers

This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number BDIY00000000. The version described in this paper is the first version, BDIY01000000, which consists of sequences BDIY01000001 to BDIY01000126.

**Table 1.** Genome features of *Roseovarius* sp. A-2.

Genome size	4.58 Mb
GC content	62.8%
Number of contigs	126
Total contig size	4,584,578 bp
Largest contig	670,842 bp
<i>N</i> <sub>50</sub>	160,967
<i>L</i> <sub>50</sub>	8
Protein encoding genes	4,514
tRNAs	46
rRNA	1

## Acknowledgments

This work was supported financially by Directorate-General of Higher Education, the Ministry of Research, Technology and Higher Education, Indonesia awarded to Tri Yuliana. This work was supported in part by research funding from Yamagata Prefecture and Tsuruoka City.

## Competing Interests

The authors have declared that no competing interest exists.

## References

- Amachi S, Muramatsu Y, Akiyama Y, Miyazaki K, Yoshiki S, Hanada S, Kamagata Y, Ban-nai T, Shinoyama H, Fujii T. Isolation of

- iodide-oxidizing bacteria from iodide-rich natural gas brines and seawaters. *Microbial Ecol.* 2005; 49: 547–557.
2. Iino T, Ohkuma M, Kamagata Y, Amachi S. *Iodidimonas muriae* gen. nov., sp. nov., an aerobic iodide-oxidizing bacterium isolated from brine of a natural gas and iodine recovery facility, and proposals of *Iodidimonadaceae* fam. nov., *Iodidimonadales* ord. nov., *Emcibacteraceae* fam. nov. and *Emcibacterales* ord. nov., *Int J Syst Evol Microbiol.* 2016; 66: 5016–5022.
  3. Shiroyama K, Kawasaki Y, Unno Y, Amachi S. A putative multicopper oxidase, IoxA, is involved in iodide oxidation by *Roseovarius* sp. strain A-2. *Biosci Biotechnol Biochem.* 2015; 79: 1898–1905.
  4. Suzuki M, Eda Y, Ohsawa S, Kanesaki Y, Yoshikawa H, Tanaka K, Muramatsu Y, Yoshikawa J, Sato I, Fujii T, Amachi S. Iodide oxidation by a novel multicopper oxidase from the Alphaproteobacterium strain Q-1. *Appl Environ Microbiol.* 2012; 78: 3941–3949.
  5. Yuliana T, Ebihara K, Suzuki M, Shimonaka C, Amachi S. A novel enzyme-based antimicrobial system comprising iodide and a multicopper oxidase isolated from Alphaproteobacterium strain Q-1. *Appl Microbiol Biotechnol.* 2015; 99: 10011–10018.
  6. [Internet] FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
  7. Huntemann M, Ivanova N, Kyrpides NC, Pati A. Large-scale contamination of microbial isolate genomes by Illumina PhiX control. *Stand Genomic Sci.* 2015; 10: 18.
  8. [Internet] BBDMap short read aligner, and other bioinformatics tools. <https://sourceforge.net/projects/bbmap/>
  9. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30: 2114–2120.
  10. 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. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012; 19: 455–477.
  11. Voget S, Bruns H, Wagner-Döbler I, Schulz S, Daniel R. Draft genome sequence of *Roseovarius tolerans* EL-164, a producer of *N*-acylated alanine methyl esters and *N*-acylhomoserine lactones. *Genome Announcements.* 2015; 3: e01096-15.
  12. Riedel T, Spring S, Fiebig A, Scheuner C, Petersen J, Göker M, Klenk H-P. Genome sequence of the *Roseovarius mucosus* type strain (DSM 17069<sup>T</sup>), a bacteriochlorophyll *a*-containing representative of the marine *Roseobacter* group isolated from the dinoflagellate *Alexandrium ostenfeldii*. *Stand Genomic Sci.* 2015; 10: 17.
  13. Rosana ARR, Orata FD, Xu Y, Simkus DN, Bramucci AR, Boucher Y, Case RJ. Draft genome sequences of seven strains isolated from a polymicrobial culture of coccolith-bearing (C-type) *Emiliania huxleyi* M217. *Genome Announcements.* 2016; 4: e00673-16.
  14. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014; 30: 2068–2069.
  15. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 2004; 32:11–16.
  16. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* 2010; 11: 119.
  17. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011; 8: 785–786.
  18. UniProt Consortium. Activities at the universal protein resource (UniProt). *Nucleic Acids Res.* 2014; 42: D191–D198.
  19. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 2016; 44: D279–D285.
  20. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 2015; 43: D222–D226.
  21. Lima T, Auchincloss AH, Coudert E, Keller G, Michoud K, Rivoire C, Bulliard V, de Castro E, Lachaize C, Baratin D, Phan I, Bougueleret L, Bairoch A. Hamap: a database of completely sequenced microbial proteome sets and manually curated microbial protein families in UniProtKB/Swiss-Prot. *Nucleic Acids Res.* 2009; 37: D471–D478.
  22. Labrenz M, Collins MD, Lawson PA, Tindall BJ, Schumann P, Hirsch P. *Roseovarius tolerans* gen. nov., sp. nov., a budding bacterium with variable bacteriochlorophyll *a* production from hypersaline Ekho Lake. *Int J System Bacteriol.* 1999; 49: 137–147.
  23. Biebl H, Allgaier M, Lünsdorf H, Pukall R, Tindall BJ, Wagner-Döbler I. *Roseovarius mucosus* sp. nov., a member of the *Roseobacter* clade with trace amounts of bacteriochlorophyll *a*. *Int J System Evol Microbiol.* 2005; 55: 2377–2383.
  24. Ehara A, Suzuki H, Kanesaki Y, Yoshikawa H, Amachi H. Draft genome sequence of strain Q-1, an iodide-oxidizing Alphaproteobacterium isolated from natural gas brine water. *Genome Announcements.* 2014; 2: e00659-14.