Elucidation of the draft genome sequence of *Ideonella azotifigens* DSMZ21438 a novel aerobic diazotroph of the Betaproteobacteria isolated from grass rhizosphere soil

Richard Allen White III,1,5-6 Jeffrey S. Norman,3 Emily E. Mclachlan,¹ Joseph P. Dunham,4 Aaron Garoutte,3 Maren L. Friesen1-3

¹Department of Plant Pathology, Washington State University, Pullman, WA USA

²Department of Crop and Soil Sciences, Washington State University, Pullman, WA USA

³Department of Plant Biology, Michigan State University, East Lansing, MI, USA

⁴SeqOnce Biosciences Inc. Pasadena, CA, USA

⁵RAW Molecular Systems (RMS) LLC, Spokane, WA USA

⁶Australian Centre for Astrobiology, University of New South Wales, Sydney, NSW, Australia

Correspondence: Maren L. Friesen, Department of Plant Pathology, Department of Crop and Soil Sciences, Washington State University. Email: m.friesen@wsu.edu

**Abstract:** Here, we report the draft genome sequence of the type strain of *Ideonella azotifigens* DSMZ21438ᵀ (formally 1a22ᵀ = JCM15503ᵀ). *Ideonella azotifigens* DSMZ21438ᵀ is the first member of the genus isolated from rhizosphere soil, providing a framework for further study into non-alphaproteobacterial nitrogen fixation and synthetic biology applications.

The 891,561 paired-end shotgun reads were quality filtered and decontaminated with the ATLAS pipeline, then assembled with Unicycler. The genome size is 6,257,981 bp, an N50 size of 7,849 bp, with a G+C content of 66.71%, and with 5,882 predicted protein-coding genes. *I. azotifigens* DSMZ21438ᵀ represents the first member of the genus isolated from rhizosphere soil, providing a framework for further study into non-alphaproteobacterial nitrogen fixation and synthetic biology applications.

**Data Set:** OSF repo [https://osf.io/r9y3g/](https://osf.io/r9y3g/).

**Data Set License:** CC-By Attribution 4.0 International

**Keywords:** Betaproteobacterial diazotroph, free-living nitrogen fixer, synthetic biology
1. Summary

Rhizosphere soils represent one of the most microbially diverse ecosystems on the planet (1-3). Understanding microbes that fix atmospheric nitrogen (i.e., diazotrophs) within the rhizosphere interface is critical for soil health and crop production. Many of these diazotrophic microbes are symbiotic forming nodules within plant roots (e.g., Rhizobium) but some non-symbiotic free-living diazotrophs such as *Azotobacter* and *Ideonella* (4-7). As much as 20% of the global biological nitrogen fixation comes from free-living non-symbiotic diazotrophs, which could act as biofertilizers (4-7).

*Ideonella azotifigens* DSMZ21438\(^T\) (formally 1a22\(^T\) = JCM15503\(^T\)) was isolated from grass rhizosphere from a 30-year old fallow agricultural field (7). *Ideonella* genus officially proposed by the isolation of *Ideonella dechloratans* (8). The genus *Ideonella* recently diverged within *Rubrivivax–Roseateles–Leptothrix–Azohydromonas–Aquincola–Ideonella* branch within the order *Burkholderiales* (9). *Ideonella* has been isolated from diverse environments, including activated sludge (8), recycling plant (10-11), and a freshwater marsh (12).

The resulting *de novo* assembly was contained on 1219 contigs, with a genome size of 6,257,981 bp, an N50 size of 7,849 bp, with a G+C content of 66.71%. The genome is 95.66% complete with 1.02% contamination. Prokka annotation predicts 64 tRNAs, 1 tmRNAs, 1 copy of the 5S-16S-23S operon, 0 CRISPRs, 22 misc RNA (or non-coding RNAs), 1 repeat region, and 5,882 predicted protein-coding genes.

*I. azotifigens* DSMZ21438\(^T\) represents the first member of the genus isolated from rhizosphere soil (7). Here we provide the draft genome sequence of *I. azotifigens* DSMZ21438\(^T\), which will provide a genomic blueprint of designing diazotrophic synthetic rhizospheres (13).
2. Data description

This Whole Genome Shotgun project and the version described in this paper has been deposited at DDBJ/ENA/GenBank under the accession VIDT00000000. All code can be found at www.github.com/friesenlab/ideonella-azotifigens_DSMZ21438. Fourteen contigs were removed for being <200 bp for GenBank submission, including the version discussed here. However, all assemblies and annotations are available on the Open Science Framework (OSF) repo for this genome (https://osf.io/r9y3g/).

3. Methods

A single colony of I. azotifigens DSMZ21438T was inoculated in Burk’s broth with nitrogen (0.25 g L⁻¹ NH₄Cl) for DNA extraction in an ambient atmosphere for six days. DNA was extracted and purified using the MasterPure DNA Extraction Kit (Epicentre, Madison WI, USA), following manufacturer’s guidelines. DNA was quantified using Qubit Fluorometer 2.0 (Invitrogen, Carlsbad, CA, USA), then quality checked using a Nanodrop-1000 (Thermo Fisher, Waltham, MA, USA). SeqOnce RhinoSeq kit was used for Illumina library preparation following the manufacturer’s protocols (https://seqonce.com/rhinoseq/). Michigan State University Research Technology Support Facility (RTSF) sequencing core completed DNA sequencing, library quantification, and sequenced on HiSeq 4000 in 150 bp paired-end read format.

Paired-end shotgun reads were quality filtered, assembled, and decontaminated with the ATLAS pipeline (version 1.0) (14). The 891,561 paired-end (150 bp format) quality controlled were then assembled with Unicycler (version 0.4.7) using default Illumina assembly parameters (15). CheckM (version 1.0.12) was used to estimate completeness and
contamination (16). Annotation was completed using Prokka (version 1.13.3) with -rfam flag to obtain rRNAs and tmRNAs (17).

Acknowledgments

We acknowledge support from NSF IOS 1331218 to MLF and NSF DEB 1547024 to MLF and JSN. We acknowledge with special thanks to Dr. Daniel H. Buckley of Cornell for original characterization and isolation of strain DSMZ21438T as well as excellent discussions about this unique diazotroph.

References


