Supplementary Material

*Azohydromonas caseinilytica* sp. nov., a nitrogen fixing bacterium isolated from forest soil by using optimized culture method

Ram Hari Dahal1,2, Dhiraj Kumar Chaudhary1,3, Dong-UK Kim4 and Jaisoo Kim1\*

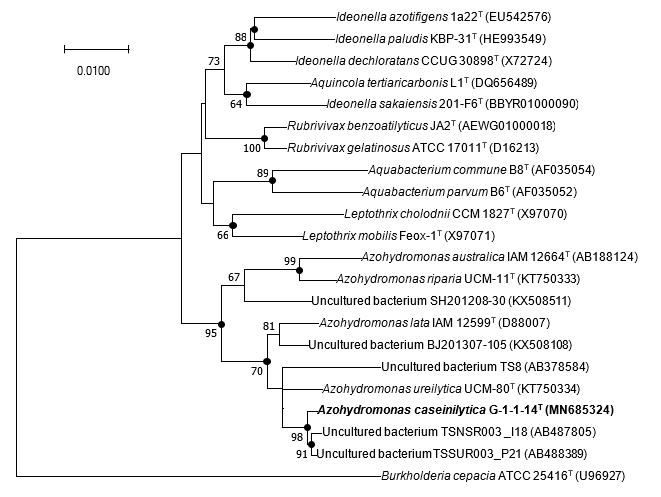
1Department of Life Science, College of Natural Sciences, Kyonggi University, Suwon, Kyonggi-Do 16227, Republic of Korea

2Department of Microbiology, School of Medicine, Kyungpook National University, Daegu 41944, Republic of Korea

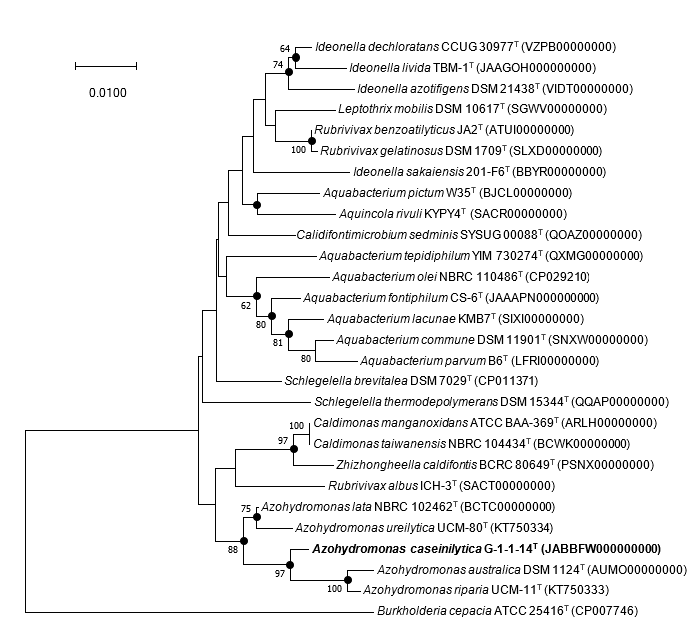
3Department of Environmental Engineering, Korea University Sejong Campus, Sejong City 30019, Republic of Korea

4Department of Biological Science, College of Science and Engineering, Sangji University, Wonju, Republic of Korea

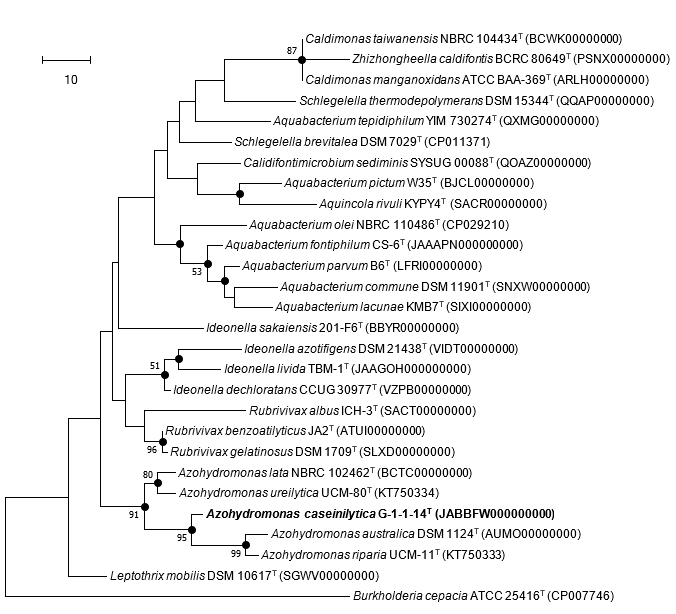
**\***Correspondence: jkimtamu@kgu.ac.kr; Tel.: +82-31-249-9648



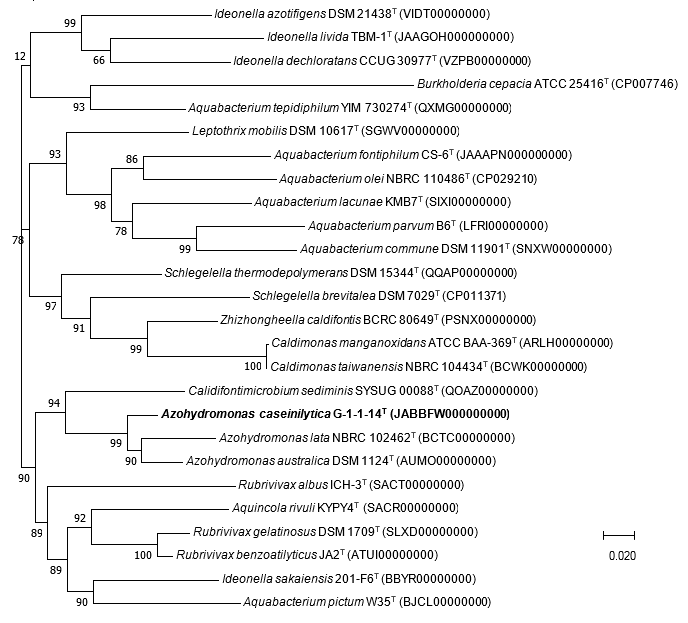
**Supplementary Figure 1.** Neighbour-joining tree based on 16S rRNA gene sequences showing the phylogenetic position of strain G-1-1-14T among closely related members of the order *Burkholderiales* along with previously uncultured bacterial clones. Filled circles indicate nodes recovered by treeing methods (neighbour–joining, maximum–likelihood and maximum–parsimony). The numbers at the nodes indicate the percentage of 1000 bootstrap replicates yielding this topology; only values >50% are shown. *Burkholderia cepacia* ATCC 25416T was used as an out-group. GenBank accession numbers are given in parentheses. Bar, 0.0100 substitutions per nucleotide position.



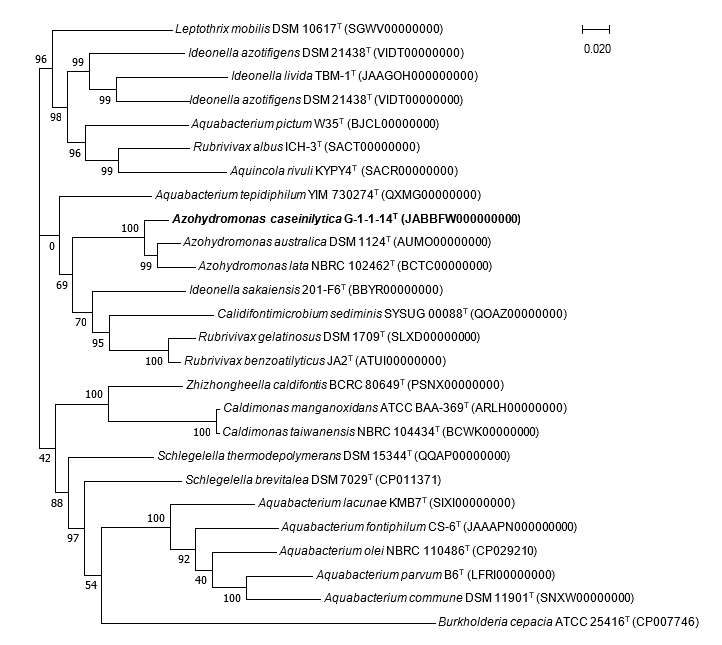
**Supplementary Figure 2.** Neighbour–joining tree based on 16S rRNA gene sequences showing the phylogenetic position of strain G-1-1-14T among closely related members of the order *Burkholderiales*. Filled circles indicate nodes recovered by all three treeing methods (neighbour–joining, maximum–likelihood and maximum–parsimony). The numbers at the nodes indicate the percentage of 1000 bootstrap replicates yielding this topology; only values >50% are shown. *Burkholderia cepacia* ATCC 25416T was used as an out-group. GenBank accession numbers are given in parentheses. Bar, 0.0100 substitutions per nucleotide position.



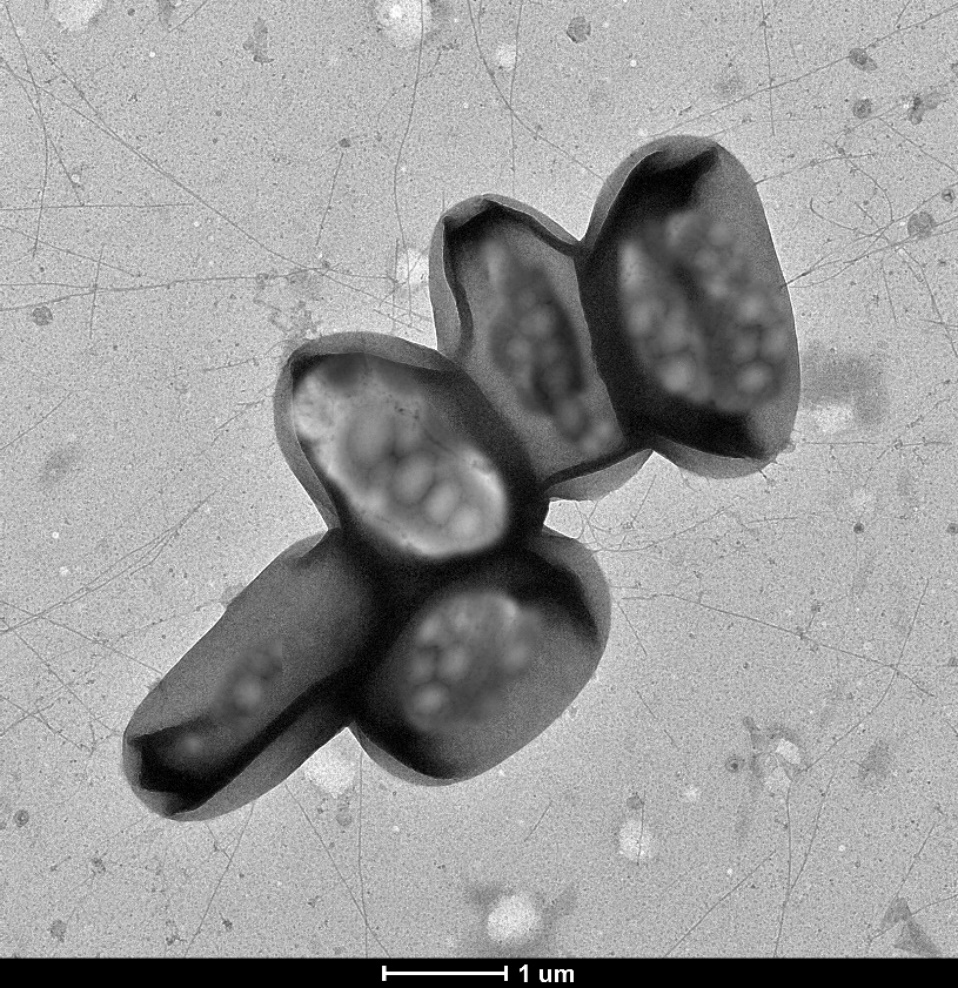
**Supplementary Figure 3.** Maximum–parsimony tree based on 16S rRNA gene sequences showing the phylogenetic position of strain G-1-1-14T among closely related members of the order *Burkholderiales*. Filled circles indicate nodes recovered by all three treeing methods (neighbour–joining, maximum–likelihood and maximum–parsimony). The numbers at the nodes indicate the percentage of 1000 bootstrap replicates yielding this topology; only values >50% are shown. *Burkholderia cepacia* ATCC 25416T was used as an out-group. GenBank accession numbers are given in parentheses. Bar, 10 substitutions per nucleotide position.



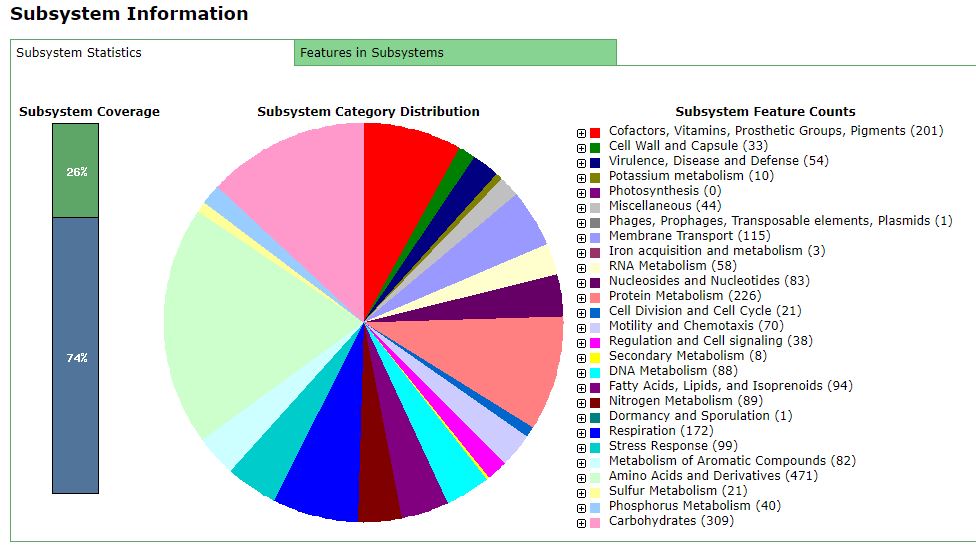
**Supplementary Figure 4.** MLSA (Multilocus sequence analyses) tree based on *recA* genes showing the phylogenetic position of strain G-1-1-14T among the closest phylogenetic members. Bootstrap values are presented at the branch nodes. Bar, 0.020 substitutions per nucleotide position.



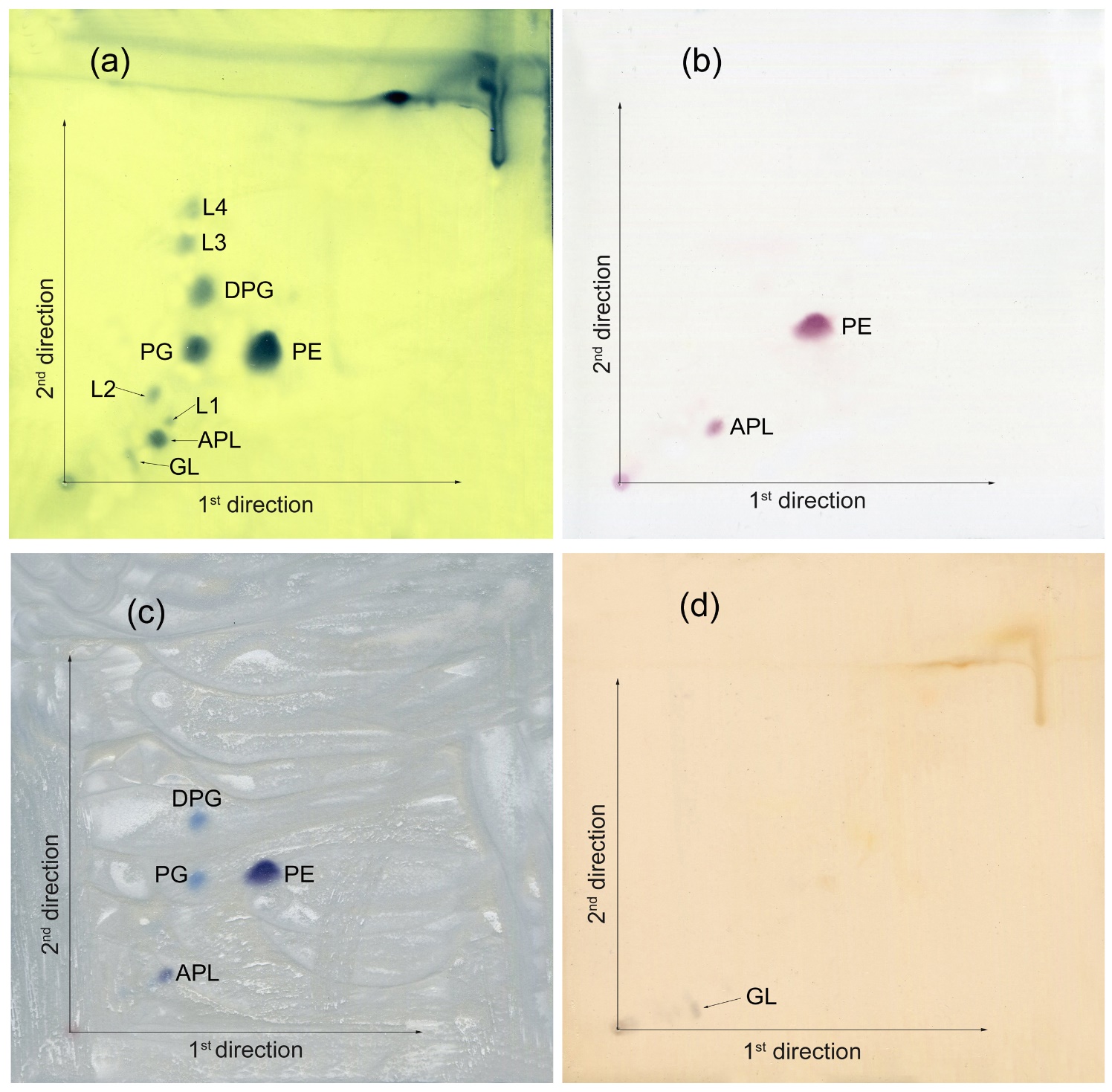
**Supplementary Figure 5.** MLSA (Multilocus sequence analyses) tree based on *rpoB* genes showing the phylogenetic position of strain G-1-1-14T among the closest phylogenetic members. Bootstrap values are presented at the branch nodes. Bar, 0.020 substitutions per nucleotide position.



**Supplementary Figure 6.** Transmission electron microscopic image of strain G-1-1-14T grown on R2A agar at 28 °C for 7 days. Bar, 1 µm



**Supplementary Figure 7.** Subsystem feature counts for strain G-1-1-14T analysed by RAST (Rapid Annotation using Subsystem technology).



**Supplementary Figure 8.** Thin-layer chromatograms of the polar lipids from strain G-1-1-14T determined on TLC (thin layer chromatography). The components were seen by staining with 5% molybdophosphoric acid in ethanol and heating them to 180 °C for 15 min (a); ninhydrin at 110 °C for 15 min (b); molybdenum at room temperature (c); and *α*-napthol-sulphuric acid at 110 °C for 15 min (d). Abbreviations: PE, phosphatidylethanolamine; DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; GL, unidentified glycolipid; APL, unidentified aminophospholipid; L1–L4, unidentified polar lipids.

**Supplementary Table 1.** Average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) of strain G-1-1-14T with reference strains.

|  |  |  |  |
| --- | --- | --- | --- |
| Strains | G-1-1-14T | | |
| OrthoANI% |  | dDDH% |
| *Azohydromonas lata* NBRC 102462T | 84.6 |  | 28.6 |
| *Azohydromonas australica* DSM 1124T | 86.5 |  | 31.5 |
| *‘Azohydromonas aeria’* t3-1-3 | 85.3 |  | 29.7 |

**Supplementary Table 2.** The genes and associated functional roles in strain G-1-1-14T genome revealed by RAST (Rapid Annotation using Subsystem technology) analyses.

|  |  |
| --- | --- |
| Genes | Functional Role |
| *nifA* | Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA |
| *nifB* | Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB |
| *frdN* | 4Fe-4S ferredoxin, nitrogenase-associated |
| *nifX* | Nitrogenase FeMo-cofactor carrier protein NifX |
| *nifX2* | NifX-associated protein |
| *nifE* | Nitrogenase FeMo-cofactor scaffold and assembly protein NifE |
| *nifN* | Nitrogenase FeMo-cofactor scaffold and assembly protein NifN |
| *nifQ* | Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein NifQ |
| *nifV* | Homocitrate synthase (EC 2.3.3.14) |
| *nifW* | Nitrogenase stabilizing/protective protein NifW |
| *nifH* | Nitrogenase (molybdenum-iron) reductase and maturation protein NifH |
| *nifD* | Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1) |
| *nifK* | Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1) |
| *nifZ* | NifZ protein |
| *nifT* | NifT protein |
| *nifO* | Nitrogenase-associated protein NifO |
| Avin2460 | LRV (FeS)4 cluster domain protein clustered with nitrogenase cofactor synthesis |

**Supplementary Table 3.** N2-fixation regulatory proteins in G-1-1-14T genome.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Accession** | **Contig** | **Locus** | **From** | **To** | **Protein product** | **Length** | **Protein Name** |
| [NZ\_JABBFW010000002.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000002.1) | 2 | *ntrC* | 495336 | 496817 | [WP\_169159272.1](https://www.ncbi.nlm.nih.gov/protein/WP_169159272.1) | 493 | Nitrogen regulation protein NR(I) |
| [NZ\_JABBFW010000001.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000001.1) | 1 | - | 470052 | 470390 | [WP\_028997364.1](https://www.ncbi.nlm.nih.gov/protein/WP_028997364.1) | 112 | MULTISPECIES: P-II family nitrogen regulator |
| [NZ\_JABBFW010000071.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000071.1) | 71 | - | 2456 | 2794 | [WP\_169164069.1](https://www.ncbi.nlm.nih.gov/protein/WP_169164069.1) | 112 | P-II family nitrogen regulator |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | - | 236951 | 237358 | [WP\_169161004.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161004.1) | 135 | Nitrogen fixation protein NifQ |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | - | 237984 | 238457 | [WP\_169161007.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161007.1) | 157 | NifX-associated nitrogen fixation protein |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | *nifX* | 238454 | 238867 | [WP\_169161008.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161008.1) | 137 | Nitrogen fixation protein NifX |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | *nifT* | 263908 | 264120 | [WP\_169161034.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161034.1) | 70 | Putative nitrogen fixation protein NifT |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | - | 265384 | 265650 | [WP\_169161036.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161036.1) | 88 | Nitrogen fixation protein NifZ |
| [NZ\_JABBFW010000008.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000008.1) | 8 | - | 281902 | 282237 | [WP\_169161053.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161053.1) | 111 | Nitrogen fixation protein NifZ |
| [NZ\_JABBFW010000009.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000009.1) | 9 | - | 54376 | 55188 | [WP\_169161139.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161139.1) | 270 | Carbon-nitrogen hydrolase family protein |
| [NZ\_JABBFW010000009.1](https://www.ncbi.nlm.nih.gov/nuccore/NZ_JABBFW010000009.1) | 9 | *ptsN* | 145280 | 145750 | [WP\_169161211.1](https://www.ncbi.nlm.nih.gov/protein/WP_169161211.1) | 156 | PTS IIA-like nitrogen regulatory protein PtsN |

**Supplementary Table 4.** The distribution of biosynthetic gene clusters in *Azohydromnonas caseinilytica* G-1-1-14T revealed by antiSMASH 5.0. APE Vf, APE Vf biosynthetic gene cluster from *Aliivibrio fischeri* ES114; NRP, non-ribosomal peptide; NRPS, non-ribosomal synthesized peptide; NRPS-like, NRPS -like fragment; NAGGN, *N*-acetylglutaminylglutamine acid; RiPP-like, other unspecified ribosomally synthesised and post-translationally modified peptide product (RIPP) cluster; RRE-containing, RiPP recognition element (RRE) - element containing cluster.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cluster** | **Contig** | **Type** | **From** | **To** | **Most similar known cluster** | **Similarity** | **Core biosynthetic genes** | **Additional biosynthetic genes** |
| 1 | 1 | Hserlactone | 398,138 | 418,842 | - | - | 1 | 1 |
| 2 | 2 | Terpene | 298,872 | 320,568 | - | - | 2 | 5 |
| 3 | 6 | RiPP-like | 85,304 | 96,164 | - | - | 1 | 1 |
| 4 | 11 | RRE-containing | 116,078 | 136,347 | Burkholderic acid; NRP + Polyketide:Modular type I | 17% | 1 | 1 |
| 5 | 13 | Terpene | 54,074 | 77,764 | [Carotenoid; terpene](https://mibig.secondarymetabolites.org/go/BGC0000640/1) | 66% | 2 | 2 |
| 6 | 13 | RRE-containing, RiPP-like | 108,265 | 130,418 | - | - | 3 | 1 |
| 7 | 15 | RiPP-like | 74,921 | 85,853 | - | - | 1 | 1 |
| 8 | 25 | NRPS | 38,162 | 82,487 | - | - | 2 | 8 |
| 9 | 28 | NRPS-like | 21,719 | 64,559 | Burkholderic acid; NRP + Polyketide:Modular type I | 6% | 1 | 9 |
| 10 | 32 | Redox-cofactor | 37,760 | 59,950 | - | - | 3 | 3 |
| 11 | 34 | Arylpolyene | 1 | 36,282 | [APE Vf; Other](https://mibig.secondarymetabolites.org/go/BGC0000837/1) | 10% | 1 | 11 |
| 12 | 38 | Lanthipeptide | 24,646 | 42,830 | [Pyoverdin; NRP](https://mibig.secondarymetabolites.org/go/BGC0000413/1) | 1% | 1 | 3 |
| 13 | 41 | NAGGN | 10,213 | 25,010 | - | - | 3 | 3 |

**Supplementary Table 5.** Cellular fatty acid profiles (% of totals) of strain G-1-1-14T and other type strains of the genus *Azohydromonas*.

Strains: 1, G-1-1-14T; 2, *Azohydromonas ureilytica* UCM-80T; 3, *Azohydromonas lata* KACC 15149T; 4, *Azohydromonas riparia* UCM-11T; 5, *Azohydromonas australica* KACC 15148T; 6, ‘*Azohydromonas areia*’ t3-1-3. All data were obtained from this study unless otherwise indicated. –, not detected.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Fatty acid** | **1** | **2** | **3** | **4** | **5** | **6†** |
| Saturated |  |  |  |  |  |  |
| C10:0 | 0.3 | 0.3 | 0.2 | – | 0.2 | – |
| C12:0 | 1.9 | 1.1 | 1.6 | 0.7 | 1.9 | 4.6 |
| C13:0 | – | – | – | – | 0.2 | – |
| C14:0 | 4.1 | 4.4 | 3.9 | 4.0 | 4.2 | 3.2 |
| C16:0 | 26.1 | 28.1 | 26.2 | 23.5 | 20.2 | 17.3 |
| C17:0 | 0.5 | 0.6 | 1.3 | – | 0.7 | – |
| C18:0 | 0.7 | – | 0.7 | 1.1 | 0.6 | 1.1 |
| Unsaturated |  |  |  |  |  |  |
| C14:1 *ω*5c | 0.1 | 0.1 | 0.2 | – | – | – |
| C15:1 *ω*6c | 0.2 | 0.3 | 0.6 | – | 0.4 | – |
| C16:1 *ω*5c | – | 0.3 | – | – | – | – |
| C17:1 *ω*6c | – | – | 1.1 | – | – | – |
| C18:3 *ω*6c | – | – | – | – | – | 0.8 |
| Hydroxy |  |  |  |  |  |  |
| C12:0 2-OH | 0.9 | 1.5 | 1.1 | 1.9 | 2.0 | – |
| C14:0 2-OH | 0.3 | – | – | 0.6 | – | – |
| C8:0 3-OH | 0.1 | – | – | – | – | – |
| C10:0 3-OH | 4.6 | 2.9 | 2.3 | 3.2 | 2.5 | 4.1 |
| C12:0 3-OH | 0.4 | – | – | – | 2.1 | – |
| C15:0 3-OH | 0.6 | – | – | – | – | – |
| C16:0 3-OH | 0.2 | 0.2 | – | – | – | – |
| iso-C15:0 3-OH | 0.2 | – | – | – | – | – |
| Branched saturated |  |  |  |  |  |  |
| iso-C15:0 | – | 0.3 | – | – | – | – |
| iso-C16:0 | 0.7 | 0.7 | 0.3 | 0.9 | 0.5 | – |
| iso-C19:0 | 0.3 | – | – | – | – | – |
| cyclo-C17:0 | 12.7 | 7.3 | 1.1 | 5.9 | 3.7 | 0.9 |
| anteiso-C15:0 | – | – | – | 1.6 | – | – |
| Summed features\* |  |  |  |  |  |  |
| 3 | 30.6 | 37.3 | 44.5 | 38.3 | 40.0 | 36.8 |
| 8 | 14.5 | 14.6 | 14.9 | 18.3 | 20.8 | 29.7 |

\*Summed features represent groups of two or three fatty acids that are treated together for the purpose of evaluation in the MIDI system and include both peaks with discrete ECLs as well as those where the ECLs are not reported separately. Summed feature 3 listed as iso-C15:0 2-OH and/or C16:1*ω*7*c* and summed feature 8 listed as C18:1*ω*7*c* and/or C18:1*ω*6*c*.

†Data from Xue *et al.*(Xue et al. 2020).

**References**

Xue H, Piao C gen, Bian D ran, et al (2020) Azohydromonas aeria sp. nov., isolated from air. J Microbiol 58:543–549