Genome Sequences of *Marichromatium gracile* HOL-1 and Its Purple Photosynthetic Coisolate, *Afifella* sp. H1R

J. A. Kyndt,ᵃ S. Dubey,ᵃ N. Frazier,ᵃ T. E. Meyerᵇ†

ᵃCollege of Science and Technology, Bellevue University, Bellevue, Nebraska, USA
ᵇDepartment of Chemistry and Biochemistry, the University of Arizona, Tucson, Arizona, USA

ABSTRACT We sequenced the genomes of both the purple sulfur gammaproteobacterium *Marichromatium gracile* HOL-1 and another purple photosynthetic organism, strain H1R, that was originally isolated as an unidentified contaminant. Through genome sequencing, we have now identified organism H1R as a species of *Afifella*. A whole-genome-based phylogenetic analysis of both species is provided.

Strain H1R was isolated as a contaminant of *Marichromatium gracile* HOL-1, and although it was known to be a photosynthetic purple nonsulfur bacterium, its true nature has been elusive for decades. The species was simply referred to as "organism H1R" (red isolate from HOL-1), a purple photosynthetic bacterium of unknown taxonomy (1–3). It was found to contain two soluble cytochromes (c_5 and c_3); however, the closest homologue of the c_5 cytochrome appeared to be from *Shewanella* (2), which is not a purple photosynthetic bacterium, indicating that there may have been another contaminant. We have now sequenced the genomes of both *Marichromatium gracile* HOL-1 and the isolated H1R species.

Marichromatium gracile HOL-1 was originally isolated in 1980 from marine waters from the Southern California coast and was cultivated on Pfennig's medium I (DSMZ medium 28), in an effort to study novel cytochrome *c* proteins and photosynthetic mechanisms of purple sulfur bacteria. Strain H1R strain had been separated from it by one of us (T.E.M.) decades ago, by repetitive plating on *Rhodospirillaceae* medium (DSMZ medium 27) supplemented with 3% NaCl. This separation occurred before strain HOL-1 was deposited as a pure culture into the DSMZ culture collection (DSM 1712). Genomic DNA was isolated from glycerol freezer stocks of both species. DNA was purified using the GeneJET DNA purification kit (Thermo Scientific). Qubit and NanoDrop DNA analyses showed 260/280 nm absorbance ratios of 1.81 for strain HOL-1 and 1.60 for strain H1R. The sequencing libraries were prepared using the Illumina Nextera DNA Flex library prep kit. The genomes were sequenced with an Illumina MiniSeq using 500 μL of a 1.8 pM library. Paired-end (2 × 150 bp) sequencing was performed, and the number of obtained reads is given in Table 1. Quality control of the reads was performed using FASTQC within Basespace (Illumina; version 1.0.0), using a k-mer size of 5 and contamination filtering. We assembled the genomes *de novo* using Unicycler (version 0.4.8) through PATRIC (4, 5). The results of this assembly approach are summarized in Table 1. The HOL-1 genome was 3,749,065 bp in length, with a GC content of 68.49%, while the H1R genome was 4,233,774 bp in length, with a GC content of 63.72%. The genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (6–8). An EvalG genome quality analysis, using the checkM algorithm (9), showed an estimated 100% completeness and 0% contamination for both genomes (Table 1). Default parameters were used for all software unless otherwise noted.

TABLE 1 Overview of the properties of the *Marichromatium gracile* HOL-1 and *Afifella* sp. H1R whole-genome sequences

Organism	No. of reads	Genome length (bp)	No. of contigs	N_{50} (bp)	GC content (%)	No. of genes	No. of tRNAs	Coverage (×)	Completeness (%)	GenBank accession no.
Marichromatium gracile HOL-1	1,872,258	3,749,065	68	144,724	68.49	3,330	48	75	100	JAKEDQ010000000
Afifella sp. H1R	2,828,530	4,233,774	10	2,471,581	63.72	3,973	50	101	100	JAKEDY010000000

Both a whole-genome phylogenetic analysis using RAxML (10, 11) (Fig. 1A) and a JSpecies comparison (12) of average nucleotide identity (ANI) showed that H1R belongs to the genus *Afifella* (formerly *Rhodobium*). The closest relative, with an ANI of 95.5%, is *Afifella aestuarii* JA968 (13). This is close to the arbitrary species cutoff of 95% (12), and further physiological studies will be needed to determine a possible species differentiation. The *Marichromatium gracile* HOL-1 whole-genome comparison (Fig. 1B) showed *Marichromatium gracile* DSM 203^T (14) as its closest relative, with an ANI of 97.4%, confirming its species placement. Besides expanding the relative small number of genomes available for these genera (6 and 7 in NCBI GenBank), these new genome sequences and the whole-genome-derived phylogenetic placement help elucidate the previous cytochrome *c* analysis of *Marichromatium gracile* HOL-1 and its, up until now, unidentified coisolate (1–3).

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers JAKEDQ000000000 (*Marichromatium gracile* HOL-1) and JAKEDY000000000 (*Afifella* sp. H1R). The versions described in this paper are versions JAKEDQ010000000 (HOL-1) and JAKEDY010000000 (H1R). The raw sequencing reads have been submitted to SRA, and the corresponding accession numbers are SRR17553093 (HOL-1) and SRR17553527 (H1R).
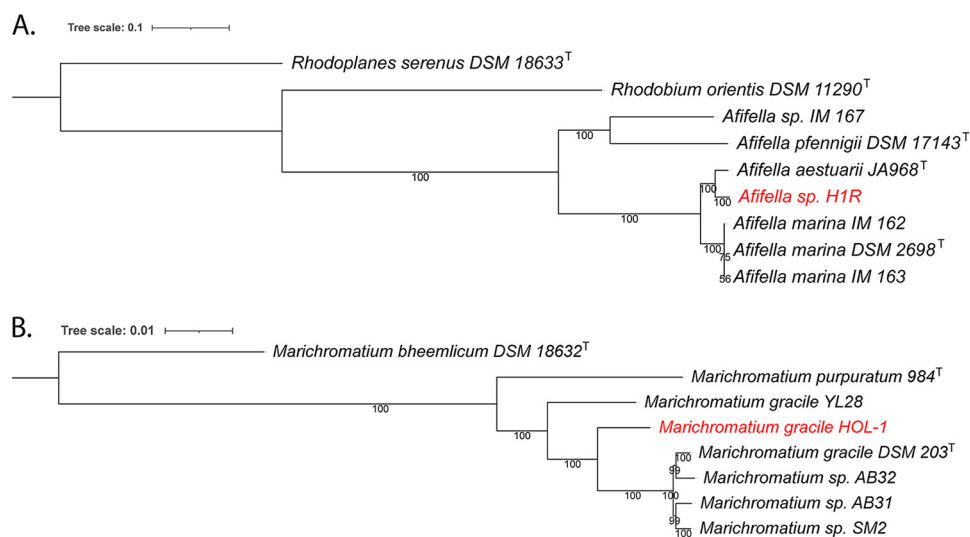
FIG 1 (A) Whole-genome-based phylogenetic tree of *Afifella* sp. H1R and its closest relatives; (B) whole-genome-based phylogenetic tree of *Marichromatium* related species. The phylogenetic trees were generated using the CodonTree method within PATRIC (5), which used PGFams as homology groups. Five hundred (A) and 1,000 (B) PGFams were used from these selected genomes using the CodonTree analysis, and the aligned proteins and coding DNA from single-copy genes were used for RAxML analysis (10, 11). The support values for the phylogenetic tree were generated using 100 rounds of the rapid-bootstrapping option of RAxML (5). *Rhodoplanes serenus* DSM 18633^T and *Marichromatium bheemlicum* DSM 18632^T were used as an outgroups. Interactive Tree Of Life (iTOL) was used for the tree visualization (15).

REFERENCES

1. Ambler RP. 1991. Sequence variability in bacterial cytochrome c. Biochim Biophys Acta 1058:42–47. https://doi.org/10.1016/s0005-2728(05)80266-x.
2. Gordon EH, Pike AD, Hill AE, Cuthbertson PM, Chapman SK, Reid GA. 2000. Identification and characterization of a novel cytochrome c(3) from *Shewanella frigidimarina* that is involved in Fe(III) respiration. Biochem J 349:153–158. https://doi.org/10.1042/0264-6021:3490153.
3. Bartsch RG. 1991. The distribution of soluble metallo-redox proteins in purple phototrophic bacteria. Biochim Biophys Acta 1058:28–30. https://doi.org/10.1016/s0005-2728(05)80262-2.
4. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
5. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535–D542. https://doi.org/10.1093/nar/gkw1017.
6. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105.
7. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.
8. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
9. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
10. Stamatakis A, Hoover P, Rougemont JJS. 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758–771. https://doi.org/10.1080/10635150802429642.
11. Stamatakis AJB. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
12. Richter M, Rosselló-Móra R, Glöckner FO, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.
13. Buddhi S, Suresh G, Deepshikha G, Sasikala C, Ramana CV. 2020. *Afifella aestuarii* sp. nov., a phototrophic bacterium. Int J Syst Evol Microbiol 70:327–333. https://doi.org/10.1099/ijsem.0.003756.
14. Imhoff JF, Rahn T, Künzel S, Keller A, Neulinger SC. 2020. Osmotic adaptation and compatible solute biosynthesis of phototrophic bacteria as revealed from genome analyses. Microorganisms 9:46. https://doi.org/10.3390/microorganisms9010046.
15. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:256–259. https://doi.org/10.1093/nar/gkz239.