Genome Sequences of *Allochromatium palmeri* and *Allochromatium humboldtianum* Expand the *Allochromatium* Family Tree of Purple Sulfur Photosynthetic Bacteria within the *Gammaproteobacteria* and Further Refine the Genus

@John A. Kyndt, a Terry E. Meyerb

a College of Science and Technology, Bellevue University, Bellevue, Nebraska, USA
b Department of Chemistry and Biochemistry, The University of Arizona, Tucson, Arizona, USA

**ABSTRACT** New genomes of two *Allochromatium* strains were sequenced. Whole-genome and average nucleotide identity based on BLAST (ANIb) comparisons show that *Allochromatium humboldtianum* is the nearest relative of *Allochromatium vinosum* (ANIb, 91.5%), while both *Allochromatium palmeri* and *Thermochromatium tepidum* are more distantly related (ANIb, <87%). These new sequences firmly establish the position of *Allochromatium* on the family tree.

*Chromatium vinosum* (now *Allochromatium vinosum*) is the prototypic purple sulfur bacterium, and it is the only species in the genus to have had a genome sequence determined (1). Moreover, there are several genera that are fairly closely related to *Allochromatium*, including *Thiocystis*, *Thermochromatium*, *Chromatium*, and *Thiorhodococcus* (2), although the relationships are not clear despite single-gene comparisons (3); therefore, a whole-genome comparison including multiple *Allochromatium* species is needed.

*Allochromatium palmeri* DSM 15591T was originally isolated from a cave system in the Bahamas (4), while *Allochromatium humboldtianum* DSM 21881T was isolated from marine sediments in Peru (5). Cultures were grown and genomic DNA was prepared by the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). DNA analysis showed A260/A280 ratios of 1.60 for *A. palmeri* and 1.96 for *A. humboldtianum*. The sequencing libraries were prepared using the Illumina Nextera DNA Flex library preparation kit and were sequenced by an Illumina MiniSeq sequencer using 500 µl of a 1.8 pM library. Paired-end (2 × 150-bp) sequencing generated 2,433,982 reads and 192 Mbp for *A. palmeri* and 3,349,346 reads and 252.2 Mbp for *A. humboldtianum*. Quality control of the reads was performed using FastQC within BaseSpace (version 1.0.0; Illumina), using a k-mer size of 5 and contamination filtering. We assembled the genome de novo through PATRIC (6) using SPAdes (version 3.10.0) (7) for *A. palmeri* and Unicycler for *A. humboldtianum*. The assembly yielded 196 contigs (>300 bp) and an N50 value of 74,142 bp for *A. palmeri* (45× coverage), while *A. humboldtianum* was assembled into 86 contigs with an N50 value of 305,111 bp (55× coverage). The *A. palmeri* genome had a GC content of 62.5% and a length of 4,272,782 bp, whereas the *A. humboldtianum* genome had a GC content of 63.9% and a length of 4,584,820 bp. The genomes were annotated using the RAST tool kit (RASTtk) (8) within PATRIC (6). This annotation showed *A. palmeri* to have 4,134 coding sequences and 45 tRNAs and *A. humboldtianum* to contain 4,391 coding sequences and 47 tRNAs. Default parameters were used for all software applications unless otherwise noted.

A JSpeciesWS comparison (9) of average nucleotide identity based on BLAST (ANIb)
showed 86.6% identity between *A. palmeri* and *A. humboldtianum* (Table 1). *A. humboldtianum* is closer to *Allochromatium vinosum* with 91.5% ANIb, while *A. palmeri* showed 86.6% ANIb. All of these ANIb values are clearly below the proposed 95% cutoff value for genome definition of a species (9). *Thermochromatium tepidum* is about equidistant from all three of the *Allochromatium* species; however, *Allochromatium warmingii* appears to be more distant from all of them.

Whole-genome-based phylogenetic analysis was performed with RAxML within PATRIC (10, 11) using all of the *Allochromatium* and related genomes (1, 12–16). This analysis grouped all of the *Allochromatium* species (Fig. 1); however, it also placed *Thermochromatium tepidum* within this group. Consistent with the ANIb analysis, *A. warmingii* is more distant from the other *Allochromatium* species. Further genetic and physiological studies may be needed to determine whether a nomenclature change of the latter species is warranted. The addition of these new *Allochromatium* genomes has substantially strengthened the phylogenetic tree of this genus.

**Data availability.** These whole-genome shotgun projects have been deposited in DDBJ/ENA/GenBank under the accession numbers WNKT00000000 for *Allochromatium palmeri* and JABZEO000000000 for *Allochromatium humboldtianum*. The versions described in this paper are versions WNKT010000000 and JABZEO010000000. The raw sequencing reads have been submitted to SRA, and the accession numbers are SRR12110462 for *Allochromatium palmeri* and SRR12110432 for *Allochromatium humboldtianum*.

**TABLE 1** ANIb comparisons

| Strain                  | ANIb (%) with strain: |
|-------------------------|------------------------|
|                         | *A. vinosum* DSM 180<sup>T</sup> | *A. humboldtianum* DSM 21881<sup>T</sup> | *A. palmeri* DSM 15591<sup>T</sup> | *T. tepidum* ATCC 43061<sup>T</sup> |
| *A. humboldtianum* DSM 21881<sup>T</sup> | 91.5                   |                     |                     |                     |
| *A. palmeri* DSM 15591<sup>T</sup>     | 86.6                   | 86.8               |                     |                     |
| *T. tepidum* ATCC 43061<sup>T</sup>    | 84.3                   | 84.9               | 82.2               |                     |
| *A. warmingii* DSM 173<sup>T</sup>     | 76.6                   | 76.5               | 76.4               | 74.7               |

**FIG 1** Whole-genome-based phylogenetic tree of all sequenced *Allochromatium* and related species. The phylogenetic tree was generated using the Codon Tree method within PATRIC (6), which used PATRIC global protein families (PGFams) as homology groups; 467 PGFams were found among these selected genomes using the Codon Tree 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 are shown on the tree branches and were generated using 100 rounds of the rapid bootstrapping option of RAxML. *Thiorhodovibrio* was used as an outgroup. Interactive Tree Of Life (iTOL) was used for the tree visualization (17).
ACKNOWLEDGMENT

This work was sponsored by the Wilson Enhancement Fund for Applied Research in Science at Bellevue University.

REFERENCES

1. Weissgerber T, Ziggan R, Bruce D, Chang YJ, Detter JC, Han C, Hauser L, Jeffries CD, Land M, Muk AC, Tapia R, Dahl C. 2011. Complete genome sequence of Allochromatium vinosum DSM 180T. Stand Genomic Sci 5:311–330. https://doi.org/10.4056/sigs.2335270.

2. Imhoff JF, Suling J, Petri R. 1998. Phylogenetic relationships among the Chromatiaceae, their taxonomic reclassification and description of the new genera Allochromatium, Isochromatium, Marichromatium, Thiococcus, Thiohalocapsa and Thromochromatium. Int J Syst Bacteriol 48:1129–1143. https://doi.org/10.1099/00207713-48-4-1129.

3. Imhoff JF, Rahn T, Künzel S, Neulinger SC. 2017. Photosynthesis is widely distributed among proteobacteria as demonstrated by the phylogeny of PhuLM reaction center proteins. Front Microbiol 8:2679. https://doi.org/10.3389/fmicb.2017.02679.

4. Herbert RA, Ranchou-Peyruse A, Duran R, Guyoneaud R, Schwabe S. 2005. Characterization of purple sulfur bacteria from the South Andros Black Hole cave system: highlights taxonomic problems for ecological studies among the genera Allochromatium and Thiocapsa. Environ Microbiol 7:1260–1268. https://doi.org/10.1111/j.1462-2920.2005.00815.x.

5. Serrano W, Schrübbers J, Amann R, Fischer U. 2015. Allochromatium humboldtianum sp. nov., isolated from soft coastal sediments. Int J Syst Evol Microbiol 65:2980–2985. https://doi.org/10.1099/ijsem.0.000364.

6. 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.

7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

8. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsmo K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Overbeek RA, McNeil NLK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

9. Richter M, Rosselló-Móra R, Glöckner FO, Peoples 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.

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. Sattley WM, Swingley WD, Burchell BM, Gurbani SA, Kujawa CM, Nuccio DA, Schladweiler J, Shaffer KN, Stokes LM, Touchman JW, Blankenship RE, Madigan MT. 2019. The complete genome of the thermophilic, anoxicogenic phototrophic gammaproteobacterium Thermochromatium tepidum. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP039268 (accession number NZ_CP039268).

13. Varghese N, Submissions S. 2016. Whole genome shotgun sequence of Allochromatium warmingii strain DSM 173. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NZ_FNOW00000000 (accession no. NZ_FNOW00000000).

14. Aviles FA, Meyer TE, Kynadt JA. 2020. Draft genome sequences of Thiorhodococcus mannitoliphagus and Thiorhodococcus minor, purple sulfur photosynthetic bacteria in the gammaproteobacterial family Chromatiaceae. Microbiol Resour Announc 9:e00193-20. https://doi.org/10.1128/MRA.00193-20.

15. Luedin SM, Liechti N, Storelli N, Danza F, Wittwer M, Potherf JS, Tonnall MA. 2018. The complete genome sequence of Chromatium okenii LaCa, a purple sulfur bacterium with a turbulent life. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NZ_PPHG00000000 (accession no. NZ_PPHG00000000).

16. Saint MK, Hanada S, Tank M. 2020. Caldichromatium gen. nov., sp. nov., a thermophilic purple sulfur bacterium member of the family Chromatiaceae isolated from Nakabusa hot spring, Japan. GenBank https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP048029 (accession no. NZ_CP048029).

17. 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.