genomeA
Journals.ASM.org

Draft Genome Sequence of the Endosymbiont "*Candidatus* Ruthia magnifica" UCD-CM (Phylum *Proteobacteria*)

Ruth D. Lee,[a] Guillaume Jospin,[a] David A. Coil,[a] Jonathan A. Eisen[a,b]

Genome Center, University of California Davis, Davis, California, USA[a]; Department of Evolution and Ecology, Department of Medical Microbiology and Immunology, University of California Davis, Davis, California, USA[b]

Here, we present the draft genome of the endosymbiont "*Candidatus* Ruthia magnifica" UCD-CM, a member of the phylum *Proteobacteria*, found from the gills of a deep-sea giant clam, *Calyptogena magnifica*. The assembly consists of 1,160,249 bp contained in 18 contigs.

Received 23 June 2014 **Accepted** 26 June 2014 **Published** 17 July 2014

Citation Lee RD, Jospin G, Coil DA, Eisen JA. 2014. Draft genome sequence of the endosymbiont "*Candidatus* Ruthia magnifica" UCD-CM (phylum *Proteobacteria*). Genome Announc. 2(4):e00717-14. doi:10.1128/genomeA.00717-14.

Address correspondence to Jonathan A. Eisen, jaeisen@ucdavis.edu.

The gammaproteobacterial endosymbiont "*Candidatus* Ruthia magnifica" was previously found to be an obligate, intracellular autotroph in one species of giant clam, *Calyptogena magnifica* (1–3). "*Candidatus* Ruthia magnifica" possesses the ability to fix carbon for its host, although the specific biochemical mechanisms of this ability remain elusive (4, 5).

Calyptogena magnifica was collected from a 28 May 2002 deep-sea exploration of a hydrothermal vent located in the Galápagos Rift, via the submersible, DSV *Alvin*, dive 3790 (6). Gill tissue was dissected and frozen in liquid nitrogen. Genomic DNA was extracted as previously described for environmental samples (7). Illumina paired-end libraries were made using a modified version of the Nextera kit by Illumina but with homegrown transposase.

A total of 3,587,578 paired-end reads were generated on an Illumina MiSeq, at a read length of 160 bp. Quality trimming and error correction of the reads resulted in 3,500,962 high-quality reads. All sequence processing and assembly was performed using the A5 assembly pipeline (8). This pipeline automates the processes of error correction, data cleaning, scaffolding, contig assembly, and quality control. The resulting assembly produced 17,632 contigs, with an N_{50} of 591. Screening the contigs using NCBI BLASTx against the NCBI's nonredundant GenBank database showed a preponderance of non "*Candidatus* Ruthia magnifica" (human, *Escherichia coli*, or other) hits. A consequent BLAST filter against a "*Candidatus* Ruthia magnifica" reference database, however, identified 18 of these contigs as "*Candidatus* Ruthia magnifica" and increased the genome N_{50} to 105,440. The resulting genome consisted of 1,160,249 bp, with a GC content of 34% and an overall coverage estimate of 19×. Scaffolds were verified by mapping error-corrected reads to the assembly using the Burrows-Wheeler Aligner (BWA) (9). Completeness of the genome was assessed using PhyloSift software (10), which searches for a list of 37 highly conserved, single-copy marker genes (11), of which all 37 were found in this assembly.

Automated annotation was performed using the RAST server (12). "*Candidatus*" sp. strain UCD-CM contains 1,215 predicted protein-coding genes and 40 predicted noncoding RNAs.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession number JARW00000000. The version described in this paper is the first version, JARW01000000.

ACKNOWLEDGMENTS

Illumina sequencing was performed at the DNA Technologies Core facility in the Genome Center at the University of California Davis.

We thank John Zhang and Aaron Darling for technical assistance and library preparation. We would also like to express our appreciation to Timothy Shank and Dan Fornari, the principal investigators of the cruise of the R/V *Atlantis* and *Alvin* during which this sample was collected. We also thank the *Atlantis* and *Alvin* crews for all their work.

This work was funded from startup funds to J.A.E.

REFERENCES

1. **Roeselers G, Newton ILG, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, Fisher MC, Fontanez KM, Lau E, Stewart FJ, Richardson PM, Barry KW, Saunders E, Detter JC, Wu D, Eisen JA, Cavanaugh CM.** 2010. Complete genome sequence of *Candidatus* Ruthia magnifica. Stand. Genomic Sci. **3:**163–173. http://dx.doi.org/10.4056/sigs.1103048.
2. **Cavanaugh CM.** 1983. Symbiotic chemoautotrophic bacteria in marine invertebrates from sulfide-rich nabitats. Nature **302:**58–61. http://dx.doi.org/10.1038/302058a0.
3. **Felbeck H, Somero GN.** 1982. Primary production in deep-sea hydrothermal vent organisms: roles of sulfide-oxidizing bacteria. Trends Biochem. Sci. **7:**201–204. http://dx.doi.org/10.1016/0968-0004(82)90088-3.
4. **Fisher CR, Childress JJ, Arp AJ, Brooks JM, Distel DL, Dugan JA, Felbeck H, Fritz LW, Hessler RR, Johnson KS, Kennicutt MC, Lutz RA, Macko SA, Newton A, Powell MA, Somero GN, Soto T.** 1988. Variation in the hydrothermal vent clam, *Calyptogen magnifica*, at the Rose Garden vent on the Galapagos spreading center. Deep Sea Res. A Oceanogr. Res. Pap. **35:**1811–1831. http://dx.doi.org/10.1016/0198-0149(88)90051-9.
5. **Newton IL, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, Fisher MC, Fontanez KM, Lau E, Stewart FJ, Richardson PM, Barry KW, Saunders E, Detter JC, Wu D, Eisen JA, Cavanaugh CM.** 2007. The *Calyptogena magnifica* chemoautotrophic symbiont genome. Science **315:**998–1000. http://dx.doi.org/10.1126/science.1138438.
6. **Shank T, Fornari D, Yoerger S, Humphris S, Bradley A, Hammond S, Lupton J, Schierer D, Collier R, Reysenbach A-L, Ding K, Seyfried W, Butterfield D, Olson E, Lilley M.** 2003. Deep submergence synergy: Alvin

and ABE explore the Galápagos rift at 86°W. Eos **84:**432–433. http://dx.doi.org/10.1029/2003EO410001.

7. **Kuske CR, Banton KL, Adorada DL, Stark PC, Hill KK, Jackson PJ.** 1998. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl. Environ. Microbiol. **64:**2463–2472.
8. **Tritt A, Eisen JA, Facciotti MT, Darling AE.** 2012. An integrated pipeline for de novo assembly of microbial genomes. PLoS One **7:**e42304. http://dx.doi.org/10.1371/journal.pone.0042304.
9. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:**1754–1760. http://dx.doi.org/10.1093/bioinformatics/btp324.
10. **Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM, Eisen JA.** 2014. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ **2:**e243. http://dx.doi.org/10.7717/peerj.243.
11. **Wu D, Jospin G, Eisen JA.** 2013. Systematic identification of gene families for use as "markers" for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. PLoS One **8:**e77033. http://dx.doi.org/10.1371/journal.pone.0077033.
12. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, 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. http://dx.doi.org/10.1186/1471-2164-9-75.