JB Journals.ASM.org

GENOME ANNOUNCEMENT

# Complete Genome Sequence of *Rahnella* sp. Strain Y9602, a Gammaproteobacterium Isolate from Metal- and Radionuclide-Contaminated Soil

**Robert J. Martinez,[a] David Bruce,[b] Chris Detter,[b] Lynne A. Goodwin,[b] James Han,[c] Cliff S. Han,[b] Brittany Held,[b] Miriam L. Land,[d] Natalia Mikhailova,[c] Matt Nolan,[c] Len Pennacchio,[c] Sam Pitluck,[c] Roxanne Tapia,[b] Tanja Woyke,[c] and Patricia A. Sobecky[a]**

Department of Biological Sciences, University of Alabama, Tuscaloosa, Alabama, USA[a]; Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA[b]; US DOE Joint Genome Institute, Walnut Creek, California, USA[c]; and Oak Ridge National Laboratory, Bioscience Division, Oak Ridge, Tennessee, USA[d]

***Rahnella* sp. strain Y9602 is a gammaproteobacterium isolated from contaminated subsurface soils that is capable of promoting uranium phosphate mineralization as a result of constitutive phosphatase activity. Here we report the first complete genome sequence of an isolate belonging to the genus *Rahnella*.**

*Rahnella* sp. strain Y9602 is a gammaproteobacterium isolated from a mixed-waste-contaminated subsurface (i.e., low pH and high concentrations of nitrate, heavy metals, and radionuclides) at the U.S. Department of Energy (DOE) Oak Ridge Reservation in Oak Ridge, TN (13). Phenotypic characterization of Y9602 indicated resistance to heavy metals, uranium tolerance, and constitutive phosphatase activity. Subsequent studies indicated that the strain precipitated 95% of soluble uranium [U(VI)] as an insoluble autunite mineral during oxic and anoxic growth conditions when supplied with glycerol-3-phosphate as the sole carbon and phosphorus source (1, 2, 12). Additionally, recent microbial ecology studies of metal-, metalloid-, organic-, and radionuclide-contaminated soils have identified enrichments of *Rahnella* species that show promise for remediation strategies (6, 10, 13, 15). Whole-genome sequencing was thus conducted on *Rahnella* sp. Y9602 to provide insights into its physiology, which promotes metal and radionuclide sequestration.

The complete genome of *Rahnella* sp. Y9602 was generated at the DOE Joint Genome Institute (JGI) using a combination of Illumina (3) and 454 (11) technologies. For this genome, an Illumina GAII shotgun library (38,026,764 reads totaling 2,890.0 Mb), a 454 Titanium standard library (260,211 reads), and a paired-end 454 library (269,624 reads totaling 194.8 Mb) were generated and sequenced. All general aspects of library construction and sequencing can be found on the JGI website (http://www.jgi.doe.gov/). The initial draft assembly contained 54 contigs in 3 scaffolds. The 454 Titanium standard data and the 454 paired end data were assembled together with Newbler, version 2.3. Illumina sequencing data were assembled with Velvet, version 0.7.63 (16). We integrated the 454 Newbler consensus shreds, the Illumina Velvet consensus shreds, and the read pairs in the 454 paired-end library using parallel Phrap, version SPS-4.24 (High Performance Software, LLC). The software Consed (4, 5, 7) was used in the following finishing process. Illumina data were used to correct potential base errors and increase consensus quality by using the software Polisher, developed at JGI (A. Lapidus, unpublished data). Possible misassemblies were corrected by using GapResolution (C. Han, unpublished data) or Dupfinisher (8) or by sequencing subcloned PCR products. Gaps between contigs were closed by editing in Consed, by PCR, and by Bubble PCR (J.-F. Cheng, unpublished data) primer walks. Protein-encoding genes were determined by Prodigal (9) and GenePrimp (14). The total size of the genome is 5,614,252 bp (chromosome and plasmids), and the final assembly is based on 191.2 Mb of 454 draft data (35× coverage) and 3,166.1 Mb of Illumina draft data (576× coverage).

The *Rahnella* sp. Y9602 chromosome has a GC content of 52.4%, is 4,864,217 bp in size, and contains two plasmids, pRAHAQ01 and pRAHAQ02, consisting of 616,549 bp (52.1% GC content) and 133,486 bp (48.3% GC content), respectively. The genome contains 5,184 protein-encoding genes, 73 pseudogenes, 76 tRNA genes, and 7 rRNA gene operons.

**Nucleotide sequence accession numbers.** The complete genome sequence of *Rahnella* sp. Y9602 is available in DDBJ/EMBL/GenBank under the accession number CP002505, in the IMG database under accession number 649633088, and in the Genomes OnLine Database under accession number Gc01605. The GenBank accession numbers for plasmids pRAHAQ01 and pRAHAQ02 are CP002506 and CP002507, respectively.

## ACKNOWLEDGMENTS

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231 and U.S. Department of Energy grant no. DE-FG02-04ER63906.

## REFERENCES

1. **Beazley MJ, et al.** 2009. Nonreductive biomineralization of uranium(VI) phosphate via microbial phosphatase activity in anaerobic conditions. Geomicrobiol. J. **26**:431–441.
2. **Beazley MJ, et al.** 2007. Uranium biomineralization as a result of bacterial phosphatase activity: insights from bacterial isolates from a contaminated subsurface. Environ. Sci. Technol. **41**:5701–5707.
3. **Bennett S.** 2004. Solexa Ltd. Pharmacogenomics **5**:433–438.
4. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. Genome Res. **8**:186–194.
5. **Ewing B, et al.** 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res. **8**:175–185.
6. **Geissler A, et al.** 2009. Biogeochemical changes induced in uranium mining waste pile samples by uranyl nitrate treatments under anaerobic conditions. Geobiology **7**:282–294.

**Received** 20 January 2012 **Accepted** 2 February 2012

Address correspondence to Robert J. Martinez, rmartinez@ua.edu.



doi:10.1128/JB.00095-12

7. **Gordon D, et al.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8**:195–202.
8. **Han CS, Chain P.** 2006. Finishing repetitive regions automatically with Dupfinisher, p 142–147. *In* Arabnia HR, Valafar H (ed), Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology. CSREA Press, Las Vegas, NV.
9. **Hyatt D, et al.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11**:119.
10. **Kourtev PS, et al.** 2006. Responses of the anaerobic bacterial community to addition of organic C in chromium(VI)- and iron(III)-amended microcosms. Appl. Environ. Microbiol. **72**:628–637.
11. **Margulies M, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
12. **Martinez RJ, et al.** 2007. Aerobic uranium(VI) bioprecipitation by metal-resistant bacteria isolated from radionuclide- and metal-contaminated subsurface soils. Environ. Microbiol. **9**:3122–3133.
13. **Martinez RJ, et al.** 2006. Horizontal gene transfer of $P_{IB}$-type ATPases among bacteria isolated from radionuclide- and metal-contaminated subsurface soils. Appl. Environ. Microbiol. **72**:3111–3118.
14. **Pati A, et al.** 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat. Methods **7**:455–457.
15. **Valverde A, et al.** 2011. Diversity and community structure of culturable arsenic-resistant bacteria across a soil arsenic gradient at an abandoned tungsten-tin mining area. Chemosphere **85**:129–134.
16. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.