Draft Genome Sequence of a Papaverine-Degrading, Gram-positive *Arthrobacter* sp., Isolated from Soil Near Hohenheim, Germany

Ondrej Reznicek, Sandra J. Facey, Bernhard Hauer

Institute of Technical Biochemistry, University of Stuttgart, Stuttgart, Germany

We present the 4.8-Mb draft genome of a soil bacterium identified as *Arthrobacter* sp. This Gram-positive soil bacterium is able to use the aromatic compound papaverine as sole carbon source and will be examined for novel oxygenases.

Address correspondence to Bernhard Hauer, bernhard.hauer@itb.uni-stuttgart.de.

Using papaverine as a substrate, a soil bacterium was isolated in Hohenheim, Germany, in the late 1970s [1]. Morphological and physiological tests identified this microorganism as *Nocardia* sp. During cultivation on papaverine as the sole source of carbon and nitrogen, several substrate intermediates were found to be excreted into the medium. A partial degradation pathway for papaverine involving initial ring hydroxylation was proposed [1, 2]. Here, we present a draft genome sequence by shotgun sequencing to explore its enzymatic content for novel oxygenases capable of converting bulky aromatic substrates such as papaverine.

The draft genome of this microorganism was obtained using Illumina shotgun and mate-pair sequencing, which resulted in paired-end reads of 250 bp. Raw data, consisting of 9,583,484 reads, were clipped and trimmed with CASAVA Illumina software. Reads with more than one N or a final length of <20 bases were removed. Quality trimmed reads were error-corrected using Musket version 1.0.6, with a 21-bp k-mer size for correction. The error-corrected reads were digitally normalized by normalize_by_median.py from the "khmer" package version 0.3, with a coverage cutoff from 80, and reads with <21 bases were discarded, yielding a data set of 2,365,326 reads. Allpaths LG release 47547 software was used to assemble the digitally normalized reads into scaffolds. Gap closure and refinement of the scaffolds was done with SOAP GapCloser version 1.12 and SEQuel version 1.0.2, respectively. All reads were aligned against the assembled scaffolds with Bowtie2 version 2.1.0. Assembly of the digitally normalized reads resulted in three genomic scaffolds in total, with a GC content of 62%. Among the three, scaffold_1 was the largest (2.83 Mb), followed by scaffold_2 (1.1 Mb) and scaffold_3 (917 kb).

For taxonomic classification, phylogenetic analysis based on the 16S rRNA sequence within scaffold_1 was performed using the eubacteria primers 8F (5′ AGAGTTTGATCCTGGCTCAG 3′) and 1492R (5′ CGGTTACCTTGTTACGACTT 3′). An NCBI BLAST search showed an absolute sequence identity of the formerly identified *Nocardia* sp. as now related to the genus *Arthrobacter*. This new classification was additionally confirmed with a BLAST search within the EzTaxon-e databank [3].

The Glimmer program [4] was used to search for open reading frames (ORFs) coding for oxygenases within the scaffold sequences. The ORFs were blasted against the NCBI nonredundant, Gen3D, SMART, Pfam, TIGRFam, SUPERFAMILY, HAMP, PIRSF, COILS, and InterProScan databases. Out of a total of 4,599 proteins, the global annotation revealed 11 monooxygenases and 4 α-subunits of Rieske nonheme iron aromatic ring-hydroxylating oxygenases (RHOs). Such RHOs are prevalently responsible for microbial degradation of aromatic compounds [5, 6] and have been proposed to be involved in this bacterium's initial conversion of papaverine. Being one of the first isolated soil bacteria capable to degrade papaverine as a complex aromatic substrate for energy supply, further investigation will be done to characterize these oxygenases in more detail for application in biotechnology [7] and organic synthesis.

Nucleotide sequence accession numbers. The scaffolds of this whole-genome shotgun project has been deposited in the European Nucleotide Archive (ENA) under the accession numbers CVLG01000001 to CVLG01000003.

ACKNOWLEDGMENTS

REFERENCES

1. **Haase-Aschoff K, Lingens F**. 1979. Mikrobieller Abbau von Papaverin. Hoppe-Seyler's Z Physiol Chem **360:**621–632. http://dx.doi.org/10.1515/bchm2.1979.360.1.621.
2. **Hauer B, Haase-Aschoff K, Lingens F**. 1982. Papaverine degradation with papaverine mutants of a *Nocardia* sp. Hoppe-Seyler's Z Physiol Chem **363:**499–506. http://dx.doi.org/10.1515/bchm2.1982.363.1.499.
3. **Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H, Park SC, Jeon YS, Lee JH, Yi H, Won S, Chun J**. 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol **62:**716–721. http://dx.doi.org/10.1099/ijs.0.038075-0.
4. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL**. 1999. Improved

microbial gene identification with GLIMMER. Nucleic Acids Res **27:** 4636–4641. http://dx.doi.org/10.1093/nar/27.23.4636.

5. **Peng R-H, Xiong A-S, Xue Y, Fu X-Y, Gao F, Zhao W, Tian Y-S, Yao Q-H.** 2008. Microbial biodegradation of polyaromatic hydrocarbons. FEMS Microbiol Rev **32:**927–955. http://dx.doi.org/10.1111/j.1574 -6976.2008.00127.x.

6. **Ferraro DJ, Gakhar L, Ramaswamy S**. 2005. Rieske business: structure–function of Rieske non-heme oxygenases. Biochem Biophys Res Commun **338:**175–179. http://dx.doi.org/10.1016/j.bbrc.2005.08.222.

7. **Furukawa K**. 2000. Engineering dioxygenases for efficient degradation of environmental pollutants. Curr Opin Biotechnol **11:**244–249. http:// dx.doi.org/10.1016/S0958-1669(00)00091-4.