

Draft Genome Sequence of the Bioelectricity-Generating and Dye-Decolorizing Bacterium *Proteus hauseri* Strain ZMd44

Nan Wang,^a I-Son Ng,^{a,b} Po Ting Chen,^c Yuzhe Li,^a Yi-Chung Chen,^c Bor-Yann Chen,^d Yinghua Lu^{a,b}

Department of Chemical and Biochemical Engineering, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, China^a; Key Laboratory for Chemical Biology of Fujian Province, Xiamen University, Xiamen, China^b; Department of Biotechnology, Southern Taiwan University of Science and Technology, Yung Kang, Tainan City, Taiwan^c; Department of Chemical and Materials Engineering, National I-Lan University, I-Lan, Taiwan^d

***Proteus hauseri* ZMd44 (CGMCC 6746), as a crucial biodecolorizing, bioelectricity-generating, and copper-resistant bacterium, is distinguished from the urinary pathogens *Proteus penneri* and *Proteus mirabilis*. To further investigate the genetic functions of this strain, the genome sequence and annotation of its open reading frames, which consist of 3,875,927 bp (G+C content, 38.12%), are presented here.**

Received 13 November 2013 Accepted 7 December 2013 Published 16 January 2014

Citation Wang N, Ng I-S, Chen PT, Li Y, Chen Y-C, Chen B-Y, Lu Y. 2014. Draft genome sequence of the bioelectricity-generating and dye-decolorizing bacterium *Proteus hauseri* strain ZMd44. *Genome Announc*. 2(1):e00992-13. doi:10.1128/genomeA.00992-13.

Copyright © 2014 Wang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to I-Son Ng, yswu@xmu.edu.cn, or Po Ting Chen, ptchen@mail.stust.edu.tw.

Proteus hauseri ZMd44 is a new classification of the *Proteus* genus (1), which was found in Lanyang Plain, northeast Taiwan. It has been used for application in microbe fuel cells (MFCs) with wastewater using azo dye Reactive Blue 160 (RBu160) (2). Moreover, the interactive characteristics of bioelectricity generation and dye decolorization in MFCs have been established (3). Recently, *P. hauseri* ZMd44 has been found to express McoA-laccase by induction with copper sulfate (4). Moreover, *P. hauseri* cells were significantly suppressed and swarming motility was inhibited with supplementation of Cu(II). This implies that its copper-related genes potentially contain vital therapeutic functions, since *P. hauseri* lost capabilities to cause swarming motility and urinary tract infections in the presence of Cu(II). This is why the genome sequencing for this new species is of great importance and significance.

High-throughput DNA sequencing of *P. hauseri* ZMd44 was conducted on the Illumina GA IIx 90-bp paired-end platform, with an average insert size of 250 bp, at Yourgene Bioscience (Taiwan). In total, 14,455,522 reads were achieved, resulting in 3.73-fold genomic coverage. The reads were filtered to remove adapter sequences and achieve the quality trim of at least 35 bp. A draft genome was generated by *de novo* assembly, using a high-volume-read-accommodating algorithm by storing data in de Bruijn graphs (5). This genome is a single circular chromosome of 3,875,927 bp with a mean G+C content of 38.12%. Fifty-one contigs with an N₅₀ value of 197,041 and a largest contig of 1,426,865 bp were constructed. Gene prediction using GeneMark (6) revealed a total of 3,147 open reading frames (ORFs). The predicted proteins were annotated by implementing a BLASTp search against the NCBI NR database.

A comparative genome analysis of *P. hauseri* ZMd44 with the closely related *Proteus mirabilis* HI4320 (7) revealed that the strains are similar, with 2,274 known protein-coding genes, 490 hypothetical protein-coding genes, and 383 no-hits genes. The known coding genes for proteins and enzymes are involved in the

three catalogues of biological process (44.9%), cellular component (23.5%), and molecular function (31.6%). On the other hand, only four ORFs encode proteins that are predicted to be involved in the copper resistance properties of *P. hauseri* ZMd44. The proteins are copper homeostasis protein CutC, copper resistance protein Crp, multicopper oxidase Mco, and repressor protein FtsI. Among these, the highest sequence identity of Crp to other pathogens microbes is only 33.3% with that of the PcoD from *Escherichia coli* strain K-12 substrain W3110 (accession no. YP_490102), followed by 12.8% with that of CopD from *Shewanella putrefaciens* CN-32 (accession no. YP_001181786). Thus, the lower similarity of the copper-resistant protein has suggested that Crp is the key protein of *P. hauseri* ZMd44 for copper resistance. This draft genome sequence helps accelerate the understanding of the characteristics of this strain not only in application of MFC and bioremediation processes, but also by potentially using the copper-related genes in therapeutic therapy of urinary tract infection caused by *Proteus* species.

Nucleotide sequence accession number. The whole-genome shotgun project for *P. hauseri* ZMd44 has been deposited in GenBank under the accession no. [AWXP000000000](https://www.ncbi.nlm.nih.gov/nuccore/AWXP000000000).

ACKNOWLEDGMENTS

We are grateful for the financial support by the National Natural Science Foundation of China (no. 21206141), the Fundamental Research Funds for the Central Universities (no. 2011121017), and the Fund from Fujian Provincial Department of Science & Technology (no. 2012I0009).

We also sincerely appreciate the academic connection program between Xiamen University (China) and National I-Lan University (Taiwan), and we thank Yourgene Bioscience (Taiwan) for assistance with the genomic data analysis in this study.

REFERENCES

1. O'Hara CM, Brenner FW, Steigerwalt AG, Hill BC, Holmes B, Grimont PA, Hawkey PM, Penner JL, Miller JM, Brenner DJ. 2000. Classification of *Proteus vulgaris* biogroup 3 with recognition of *Proteus hauseri* sp. nov., nom.

- rev. and unnamed *Proteus* genomospecies 4, 5 and 6. *Int. J. Syst. Evol. Microbiol.* 50:1869–1875. <http://ijs.sgmjournals.org/content/50/5/1869.short>.
2. Chen BY, Zhang MM, Chang CT, Ding Y, Lin KL, Chiou CS, Hsueh CC, Xu H. 2010. Assessment upon azo dye decolorization and bioelectricity generation by *Proteus hauseri*. *Bioresour. Technol.* 101:4737–4741. <http://dx.doi.org/10.1016/j.biortech.2010.01.133>.
3. Chen BY, Wang YM, Ng IS. 2011. Understanding interactive characteristics of bioelectricity generation and reductive decolorization using *Proteus hauseri*. *Bioresour. Technol.* 102:1159–1165. <http://dx.doi.org/10.1016/j.biortech.2010.09.040>.
4. Zheng X, Ng IS, Ye C, Chen BY, Lu Y. 2013. Copper ion-stimulated McoA-laccase production and enzyme characterization in *Proteus hauseri* ZMd44. *J. Biosci. Bioeng.* 115:388–393. <http://dx.doi.org/10.1016/j.jbiosc.2012.10.012>.
5. Besemer J, Borodovsky M. 1999. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res.* 27:3911–3920. <http://dx.doi.org/10.1093/nar/27.19.3911>.
6. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
7. Pearson MM, Sebaihia M, Churcher C, Quail MA, Seshasayee AS, Luscombe NM, Abdellah Z, Arrosmith C, Atkin B, Chillingworth T, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabinowitsch E, Walker D, Whithead S, Thomson NR, Rather PN, Parkhill J, Mobley HL. 2008. Complete genome sequence of uropathogenic *Proteus mirabilis*, a master of both adherence and motility. *J. Bacteriol.* 190:4027–4037. <http://dx.doi.org/10.1128/JB.01981-07>.