

Genome Assembly of *Chryseobacterium polytrichastri* ERM1:04, a Psychrotolerant Bacterium with Cold Active Proteases, Isolated from East Rathong Glacier in India

Rakshak Kumar, Dharam Singh, Mohit Kumar Swarnkar, Anil Kumar Singh, Sanjay Kumar

Department of Biotechnology, CSIR, Institute of Himalayan Bioresource Technology, Palampur, Himachal Pradesh, India

R.K. and D.S. contributed equally to this work.

This represents CSIR-IHBT communication number 3927.

We report here the genome assembly of a psychrotolerant bacterium, *Chryseobacterium polytrichastri* ERM1:04, which secretes cold-active proteases. The bacterium was isolated from a pristine location, the East Rathong Glacier in the Sikkim Himalaya. The 5.53-Mb genome provides insight into the cold-active industrial enzyme and adaptation in the cold environment.

Received 19 September 2015 Accepted 25 September 2015 Published 5 November 2015

Citation Kumar R, Singh D, Swarnkar MK, Singh AK, Kumar S. 2015. Genome assembly of *Chryseobacterium polytrichastri* ERM1:04, a psychrotolerant bacterium with cold active proteases, isolated from East Rathong Glacier in India. *Genome Announc* 3(6):e01305-15. doi:10.1128/genomeA.01305-15.

Copyright © 2015 Kumar 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/4.0/).

Address correspondence to Rakshak Kumar, rakshak@ihbt.res.in.

The East Rathong Glacier in the Sikkim Himalaya has been explored for changes in its glacier volume and geologic importance (1–3), but reports on microbes there are lacking. The cold environment of glaciers may harbor a diverse bacterial community with a potential source of industrially important cold-active enzymes (4). Realizing this importance, we isolated cold-adapted bacteria at 5°C from an East Rathong Glacier sample and found few potentially prospective protease-secreting bacteria after screening, using a plate assay, as described earlier (5). Amongst them, ERM1:04 is a yellow-pigmented, Gram-negative, aerobic, and nonmotile bacterium that was isolated from a soil sample collected from the moraines (27°33′049″N 88°07′533″E; 4,613 meters above sea level [masl]). Using its 16S rRNA gene sequence (1,506 bp), phylogenetically, it was found to branch with the type strain of *Chryseobacterium polytrichastri* YG4-6 with 98.92% sequence similarity, confirming that the strain belongs to species *C. polytrichastri* (6). The type strain (YG4-6) was reported from a glacier in China; however, its genome data are not available (7). Therefore, to elucidate genomic insights into industrial enzymes and its cold adaptation, the strain ERM1:04 was chosen for whole-genome sequencing.

Genomic DNA from strain ERM1:04 was isolated using a standard genomic DNA isolation method. The quality and quantity of extracted genomic DNA were determined by a Qubit 2.0 fluorometer (Invitrogen, USA). Genomic DNA was sheared using Covaris g-Tubes to construct a library with a 10-kb DNA insert size using the PacBio SMRTbell template preparation kit version 1.0, as described earlier (8). The prepared SMRTbell library was sequenced using the PacBio RS II system on two SMRT cells using P6 polymerase-C4 sequencing chemistry with a 240-min movie. Two SMRT cells produced 1,832,764,456 nucleotide bases generated from 118,266 reads (N_{50} size, 27,831 bp; mean read length, 15,496 bp). The filtered subreads were assembled *de novo* using CLC Genomics Workbench version 8.0.3 (Qiagen Aarhus A/S).

The strain ERM1:04 genome was assembled in 10 contigs, for a total of 5,532,908 bp in size with 34.07% G+C content (longest contig length, 1,673,155; N_{50} contig length, 1,207,894; mean contig length, 553,291).

The protein-coding sequences (CDSs) were predicted by the Prokaryotic Genome Annotation Pipeline (PGAP) version 2.10 (NCBI, 2013), which predicted 4,938 genes, 4,524 CDSs, 24 rRNAs (5S, 16S, 23S), 75 tRNAs, 1 noncoding RNA (ncRNA), and 314 pseudogenes. Functional annotation was performed using the Rapid Annotations using Subsystems Technology (RAST) server (9), which predicted 5,176 coding sequences and 106 RNAs. These predicted genes were functionally assigned to 357 RAST subsystem categories. The genomic analysis identified genes responsible for industrially important enzymes, such as protease, catalase, permease, esterase/lipase, chitinase, alcohol dehydrogenase, glycosyltransferase, and alkaline phosphatase. Our experimental data of cold-active protease is well corroborated with the presence of various genes encoding important proteases, such as Clp protease, HtrA protease, and metalloproteases in the genome. Cold-shock proteins (CspA and CspG), which aid in the survival of bacteria in harsh environments, were also found.

This is the first report of genomic data of the species *C. polytrichastri*, and it will aid in comparative genomic studies for understanding the evolutionary mechanism of cold adaptation and cold-active industrial enzymes.

Nucleotide sequence accession numbers. This whole-genome shotgun project of *C. polytrichastri* ERM1:04 has been deposited at DDBJ/EMBL/GenBank under the accession no. [LIRF00000000](https://www.ncbi.nlm.nih.gov/nuclink/LIRF00000000). The version described in this paper is version LIRF01000000.

ACKNOWLEDGMENTS

R.K. thanks DST, Govt. of India, for financial assistance under the DST-INSPIRE Faculty Scheme (DST/INSPIRE/04/2014/001280), and the Sikkim State Council of Science and Technology, DST & CC, Govt. of Sikkim

for their support in sample collection. We acknowledge support from CSIR Network projects SIMPLE (BSC0109) and CeHAB (BSC0209).

REFERENCES

1. Agrawal A, Tayal S. 2013. Assessment of volume change in East Rathong Glacier, eastern Himalaya. *Int J Geoinfo* 9:73–82.
2. Agrawal A, Tayal S. 27 July 2015. Mass balance reconstruction since 1963 and mass balance model for East Rathong Glacier, eastern Himalaya, using remote sensing methods. *Geogr Ann A* <http://dx.doi.org/10.1111/geoa.12109>.
3. Sharma BM, Tayal S, Chakraborty P, Bharat GK. 2015. Chemical characterization of meltwater from East Rathong Glacier vis-à-vis western Himalayan glaciers, p 181–190. *In* Joshi R, Kumar K, Palni LMS (ed), *Dynamics of climate change and water resources of northwestern Himalaya*. Springer International Publishing, Switzerland.
4. Elleuche S, Schröder C, Sahm K, Antranikian G. 2014. Extremozymes—biocatalysts with unique properties from extremophilic microorganisms. *Curr Opin Biotechnol* 29:116–123. <http://dx.doi.org/10.1016/j.copbio.2014.04.003>.
5. Sacherer P, Défago G, Haas D. 1994. Extracellular protease and phospholipase C are controlled by the global regulatory gene *gacA* in the biocontrol strain *Pseudomonas fluorescens* CHA0. *FEMS Microbiol Lett* 116:155–160. <http://dx.doi.org/10.1111/j.1574-6968.1994.tb06694.x>.
6. Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, Park S-C, Jeon YS, Lee J-H, 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>.
7. Chen XY, Zhao R, Chen ZL, Liu L, Li XD, Li YH. 2015. *Chryseobacterium polytrichastri* sp. nov., isolated from a moss (*Polytrichastrum formosum*), and emended description of the genus *Chryseobacterium*. *Antonie van Leeuwenhoek* G107:403–410.
8. Kiran S, Swarnkar MK, Pal M, Thakur R, Tewari R, Singh AK, Gulati A. 2015. Complete genome sequencing of protease-producing novel *Arthrobacter* sp. strain IHBB 11108 using PacBio single-molecule real-time sequencing technology. *Genome Announc* 3(2):e00346-15. <http://dx.doi.org/10.1128/genomeA.00346-15>.
9. 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>.