Data in Brief

Genome assembly of *Chryseobacterium* sp. strain IHBB 10212 from glacier top-surface soil in the Indian trans-Himalayas with potential for hydrolytic enzymes

Mohinder Pal, Mohit Kumar Swarnkar, Hena Dhar, Sanjay Chhibber, Arvind Gulati

---

**ARTICLE INFO**

**Keywords:**

*Chryseobacterium* sp.  
Complete genome sequence  
Esterases  
Hydrolytic enzymes

**ABSTRACT**

The cold-active esterases are gaining importance due to their catalytic activities finding applications in chemical industry, food processes and detergent industry as additives, and organic synthesis of unstable compounds as catalysts. In the present study, the complete genome sequence of 4,843,645 bp with an average 34.08% G + C content and 4260 protein-coding genes are reported for the low temperature-active esterase-producing novel strain of *Chryseobacterium* isolated from the top-surface soil of a glacier in the cold deserts of the Indian trans-Himalayas. The genome contained two plasmids of 16,553 and 11,450 bp with 40.54 and 40.37% G + C contents, respectively. Several genes encoding the hydrolysis of ester linkages of triglycerides into fatty acids and glycerol were predicted in the genome. The annotation also predicted the genes encoding proteases, lipases, amylases, β-glucosidases, endoglucanases and xylanases involved in biotechnological processes. The complete genome sequence of *Chryseobacterium* sp. strain IHBB 10212 and two plasmids have been deposited vide accession numbers CP015199, CP015200 and CP015201 at DDBJ/EMBL/GenBank.

---

1. Direct link to deposited data

https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA313490

2. Introduction

The microorganisms inhabiting the cold environments have been explored for the industrially important enzymes [1–3]. The genus *Chryseobacterium* has been reported for the production of antifungal protease by *C. aquaticum* [4], mannanase by *C. indologenes* [5], α-amylase by *C. taeanense* [6], metallo-protease by *C. gleum* [7,8], and proteases by *C. polytrichastri* [9]. The genomic analyses have also supported the hydrolytic enzyme activities by *Chryseobacterium* sp. strain P1–3 [10], *Chryseobacterium* sp. strain 5RB126 [21], *C. gallinarum* strain DSM 27622 [11], and *C. polytrichastri* strain EMR1:04 [9]. The strain IHBB 10212 exhibited the hydrolysis of p-nitrophenyl butyrate (pNPB), p-nitrophenyl decanote (pNPD) and p-nitrophenyl caprylate (pNPC) substrates of esterases in quantitative assay [12]. The present study reports the genes encoding hydrolytic enzymes by the functional annotation of complete genome sequence of *Chryseobacterium* sp. strain IHBB 10212.
3. Experimental design, materials and methods

The strain IHBB 10212 was grown in tryptone soya broth (Himedia, India) at 28 ± 0.1 °C for 48 h and 200 rpm. Genomic DNA was isolated using phenol-chloroform-isooamyl alcohol extraction by Marmur method, assessed for quality by NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and quantified using Qubit 2.0 fluorometer (Invitrogen, USA). The DNA was sheared using Covaris g-TUBE, assessed for quality using Bioanalyzer DNA 12000 chip (Agilent Technologies, USA) and employed for preparation of genomic library using PacBio SMRTbell library preparation kit v1.0 (PacBio Technologies, USA). The library was quantified and sequenced on PacBio RS II system (Pacific Biosciences, USA) on one SMRT cell employing P5 polymerase-C3 sequencing chemistry with 180-min movie [13]. The subreads were assembled de novo using RS hierarchical genome assembly process (HGAP) protocol version 2.0 in SMRT analysis version 2.2.0 (Pacific Biosciences, USA). The functional annotations were performed using Rapid Annotation Subsystem Technology (RAST) server [14]. The circular genome map was generated using GView online server [15]. The cluster of orthologous groups (COG) functional categories were determined by using WebMGA online server [16]. The genomic islands (GIs) were predicted using IslandViewer 3 web-based interface by integrating IslandPath-DIMOB and SIGI-HMM prediction methods [17].

4. Genomic analysis

The genome sequencing produced 551,857,867 nucleotide bases generated through 74,645 reads (N50 size 13,222 bp and mean read length 7393 bp). The genome assembly resulted in a gapless and complete circular genome with total size of 4,871,648 bases with 90-fold coverage. The circular chromosome was 4,843,645 bp in size with an average G + C content 40.54% (Fig. 1). Two plasmids of 16,553 and 11,450 bp were also identified with the average G + C content 40.54 and 40.37%. The genomic features included 4260 protein-coding sequences and 21 rRNAs and 75 tRNAs (Table 1). Among the protein-coding genes, 2927 genes (68.7%) were classified into 21 categories of COG families (Table 2). Nineteen putative GIs were predicted in the chromosome. The smallest GI was 4018 bp in size having 2 genes while the largest GI was 44,550 bp in size with 30 genes. Seven GIs contained the mobile genetic elements like transposases and integrases (Fig. 2).

The functional annotations also predicted the gene clusters coding for esterases (26 genes), proteases (23 genes), lipases (1 gene), lipases/esterases (6 genes) amylases (4 genes), endoglucanases (1 gene), β-glucosidases (7 genes) and xylanase (5 genes). Among the gene clusters coding for esterases, 1 gene each for sialic acid-specific 9-O-acetylbacter, 2′,3′-cyclic-nucleotide 2′-phosphodiesterase (EC 3.1.4.16), 3′,5′-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), acyl carrier protein phosphodiesterase (EC 3.1.4.14), alkaline phosphatase (EC 3.1.3.1), arylesterase-like protein slr1119, CheB methyltransferase (EC:3.1.1.61), esterase/lipase, esterase/lipase/cholesterase family protein, esterase/lipase-like protein, FIG006285: ICC-like protein.

![Fig. 1. Circular genome map of Chryseobacterium sp. strain IHBB 10212 generated using GView server. From outside to the center: genes on forward strand, genes on reverse strand, GC content (%), GC skew, colors of cluster of orthologous groups (COG) categories and scale in kbp.](image-url)
phosphoesterase, glycerylphosphoryl diester phosphodiesterase (EC 3.1.4.46), lipase/esterase, pectinesterase (EC 3.1.1.11), putative carboxyesterase, putative phosphoesterase, serine esterase, thioesterase family protein, and thioesterase superfamily, 2 genes each for alkaline phosphodiesterase I (EC 3.1.4.1)/nucleotide pyrophosphatase (EC 3.6.1.9), thioesterase, phospholipase/carboxylesterase, putative esterase, and 5 genes for rhamnogalacturonan acetylesterase were predicted. A gene encoding for lipase (EC 3.1.1.3) was also predicted. Several genes encoding for proteases were predicted including 1 gene each for alkaline serine protease, ATP-dependent Clp protease ATP-binding subunit ClpX, ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92), ATP-dependent protease La (EC 3.4.21.53) type I, CAAX amino terminal protease family protein, carboxyl-terminal protease, HtrA protease/chaperone protein, membrane protease family protein BA0301, membrane-associated zinc metalloprotease, metalloprotease MEP2, protease II (EC 3.4.21.83), protease IV, protease precursor, protease synthase and sporation negative regulatory protein PAI 2, putative stomatin/prohibitin-family membrane protease subunit PA4582, putative stomatin/prohibitin-family membrane protease subunit YbbK, serine alkaline protease (subtilisin E), subtilisin-like serine proteases and zinc metalloprotease superfamily, and 4 genes for zinc protease. Additionally genome sequence also showed coding sequences for amylasases including 1 gene for α-amylase (neopolulpanase) SusA (EC 3.2.1.135) and 4 genes for cytoplasmic α-amylase (EC 3.2.1.1). Genes related to β-glucosidases were also predicted which included 1 gene each for glucan endo-1,6-β-glucosidase (EC 3.2.1.75) and periplasmic β-glucosidase (EC 3.2.1.21), and 4 genes for β-glucosidase (EC 3.2.1.21). The functional annotations also predicted genes for xylanases including 1 gene each for endo-1,4-β-xylanase B precursor, xylanase and 3 genes for endo-1,4-β-xylanase A precursor (EC 3.2.1.8).

The analysis also predicted in Chryseobacterium sp. strain IHBB 10212 genome the presence of 4 CpsA and 1 CspG genes responsible for survival during the cold shock. The upregulation of cold-shock proteins have been reported for survival against the cold shock by the microbial organisms inhabiting the cold environments [18]. The genome contained genes for the nucleotide excision repair system including genes for excinuclease ABC (UvrABC), transcription-repair coupling factor and UvrD helicases which are reported for repairing DNA damage by UV radiations [19,20]. The genomic analysis has provided an understanding on production of hydrolytic enzymes and genes related to cold adaptation and repairing DNA damage by UV radiations in the cold deserts.

5. Nucleotide sequence accession numbers

Complete genome sequence of Chryseobacterium sp. strain IHBB 10212 has been deposited vide the accession number CP015199 to DDBJ/EMBL/GenBank. The two plasmids have been assigned the accession numbers CP015200 and CP015201.

Conflict of interest

There is no conflict of interest among the authors.

Acknowledgements

We acknowledge the Director, CSIR-Institute of Himalayan Bioresource Technology for providing the laboratory facilities. This work was funded by the Council of Scientific and Industrial Research, Govt. of India under the Network Project “Plant-Microbe and Soil Interactions” (BSC 0117). The grant of the Rajiv Gandhi National Fellowship to Mohinder Pal by the University Grants Commission (RGNF-SC-HIM-13671), Govt. of India, is also acknowledged.

References

[1] H. Dhar, M.K. Swarnkar, A. Gulati, A.K. Singh, R.C. Kasana, Draft genome sequence of a cellulase-producing psychrotrophic Pseudomonas strain, BIBR 3415, isolated from the cold environment of the western Himalayas, India, Genome Announc. 3 (2015) e01581-14.
[2] A.N. Yadav, S.G. Sachan, P. Verma, R. Kaushik, A.K. Saxena, Cold active hydrolytic enzymes production by psychrotrophic bacilli isolated from three sub-glacial lakes of NW Indian Himalayas, J. Basic Microbiol. (2016) 294–307.
[3] J. Charlesworth, B.P. Burns, Extremophilic adaptations and biotechnological applications in diverse environments, AMS Microbiol. (2016) 251–261.
[4] M. Gandhi Pragash, K.B. Narayanan, P.R. Naik, N. Sakthivel, Characterization of Chryseobacterium aquamarinum strain PUPC1 producing a novel anti-fungal protease from rice rhizosphere soil, J. Microbiol. Biotechnol. 19 (2009) 99–107.
[5] S. Rattanawisuk, M.K. Cairns, Chryseobacterium indolgenes, novel mannanase-producing bacteria, Songklanakarin J. Sci. Technol. 31 (2009) 1–5.
[6] S.L. Wang, Y.C. Liang, T.W. Liang, Purification and characterization of a novel alkali-stable α-amylase from Chryseobacterium thermosolvens TKU001, and application in antifungal and probiotic, Process Biochem. 46 (2011) 745–750.
[7] P.N. Chaudhari, B.L. Chaudhari, S.B. Chincholkar, Iron containing keratinolytic metallo-protease produced by Chryseobacterium gleum, Process Biochem. 48 (2013) 141–151.
[8] A. Riffel, F. Lucas, P. Heeb, A. Brandelli, Characterization of a new keratinolytic bacterium that completely degrades native feather keratin, Arch. Microbiol. 179 (2003) 258–265.
[9] R. Kumar, D. Singh, M.K. Swarnkar, A.K. Singh, S. Kumar, Genome assembly of Chryseobacterium polytrichastri ERMR1-04, a psychrotolerant bacterium with cold active proteases, isolated from East Rathong Glacier in India, Genome Announc. 3 (2015) e01305-15.
[10] G.S. Park, S.J. Hong, C.H. Lee, A.R. Khan, I. Ullah, B.K. Jung, J.B. Choi, Y. Kwak, C.G. Bock, H.Y. Jung, J.H. Shin, Draft genome sequence of Chryseobacterium sp. strain P1–3, a keratinolytic bacterium isolated from poultry waste, Genome Announc. 2 (2014) e01237-14.
[11] G.S. Park, S.J. Hong, B.K. Jung, A.R. Khan, Y.J. Park, C.F. Park, A. Lee, Y. Kwak, Y.J. Lee, D.W. Lee, C. Lee, Complete genome sequence of a keratin-degrading bacterium Chryseobacterium gallinarum strain DSM 27622 T isolated from chicken, J. Biotechnol. 211 (2015) 66–67.
[12] M. Hasanuzzaman, K.M. Umahay-Brones, S.M. Ziaor, N. Morita, Y. Nodasaka, I. Yamoto, H. Okuyama, Isolation, identification, and characterization of a novel, oil-degrading bacterium, Pseudomonas aeruginosa TI, Curr. Microbiol. 49 (2004) 108–114.
[13] C.S. Chin, D.H. Alexander, P. Marks, A.A. Klammer, J. Drake, C. Hein, A. Chun, A. Copeland, J. Huddleston, E.E. Eichler, S.W. Turner, J. Korlach, Nonhybrid, complete genome sequencing of human brain tumor cells using the Ion Torrent PGM, Genome Announc. 2 (2014) 151–154.
[14] S. Endo, M. Taniyama, A. Hamada, K. Miyazaki-Miyamoto, M. Koga, K. Tomita, M. Tani, K. Ogi, R. Fujisawa, F. Ito, T. Saito, R. Kasai, K. Katawaki, Y. Miki, S. Kato, Prokaryotic genome sequences of Chryseobacterium indologenes TKU001, and application in antifungal and probiotic, Process Biochem. 46 (2011) 745–750.

Fig. 2. Genomic islands of Chryseobacterium sp. strain IHBB 10212 predicted by using IslandViewer 3 online tool. Orange color represents genomic islands predicted based on SIRI-HMM, blue color represents genomic islands predicted based on IslandPath-DIMOR, and red color represents genomic islands predicted based on integrated results of the two methods. Black plot represents G + C content (%).
[16] S. Wu, Z. Zhu, L. Fu, B. Niu, W. Li, WebMGA: a customizable web server for fast metagenomic sequence analysis, BMC Genomics 12 (2011) 444.

[17] B.K. Dhillon, M.R. Laird, J.A. Shay, G.I. Winsor, R. Lo, F. Nizam, S.K. Pereira, N. Waglechner, A.G. McArthur, M.G.I. Langille, F.S.L. Brinkman, IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis, Nucleic Acids Res. 43 (2015) W104–W108.

[18] A.K. Singh, K. Sad, S.K. Singh, S. Shivaji, Regulation of gene expression at low temperature: role of cold-inducible promoters, Microbiology 160 (2014) 1291–1296.

[19] C. Kisker, J. Kuper, B. Van Houten, Prokaryotic nucleotide excision repair, Cold Spring Harb. Perspect. Biol. 5 (2013) a012591.

[20] A. Sancar, J.T. Reardon, Nucleotide excision repair in E. coli and man, Adv. Protein Chem. 69 (2004) 43–71.

[21] T. Morohoshi, W.Z. Wang, N. Someya, T. Ikeda, Complete genome sequence of Chryseobacterium sp. strain StRB126, an N-acylhomoserine lactone-degrading bacterium isolated from potato root, Genome Announc 2 (2014) e00952-14.