RESEARCH ARTICLE

# The complete genome sequence and annotation of a psychrophilic *Cryobacterium* species GCJ02 isolated from cryomorphic soil of a virgin forest

**Dong-Sheng Xue**[1], **Shu-Yang Zhang**[1], **Chong-Ju Wang**[1], **Chun-Jie Gong**[1,*]
[1]Key Laboratory of Fermentation Engineering (Ministry of Education), Hubei Provincial Cooperative Innovation Center of Industrial Fermentation, Hubei Key Laboratory of Industrial Microbiology, National "111" Center for Cellular Regulation and Molecular Pharmaceutics, Hubei University of Technology, Wuhan 430068, China.

**Abstract:** A novel psychrotroph *Cryobacterium*sp. GCJ02 was isolated and characterized, which showed growth well at 4 ºC. The assembled whole genome of strain GCJ02 is 4.39 Mb, including 4,139 protein coding genes with G+C content of 68.41 mol%. In this study, we report the complete genome sequence of a novel strain of the genus *Cryobacterium*, affording feasibility to elucidate the molecular mechanism of cold adaptation, and facilitate genetic manipulation of this bacterium.
**Keywords:** *Cryobacterium*, genome, psychrophilic

***Correspondence to:** Chun-Jie Gong, Key Laboratory of Fermentation Engineering (Ministry of Education), Hubei Provincial Cooperative Innovation Center of Industrial Fermentation, Hubei Key Laboratory of Industrial Microbiology, National "111" Center for Cellular Regulation and Molecular Pharmaceutics, Hubei University of Technology, Wuhan 430068, P.R. China; E-mail: 20141083@hbut.edu.cn

**Received:** February 09, 2019; **Accepted:** February 24, 2019; **Published Online:** April 08, 2019
**Citation:** Dong-Sheng Xue, Shu-Yang Zhang, Chong-Ju Wang, Chun-Jie Gong, 2019. The complete genome sequence and annotation of a psychrophilic *Cryobacterium* species GCJ02 isolated from cryomorphic soil of a virgin forest. *Applied Environmental Biotechnology*, 4(2): 19-23
http://doi.org/10.26789/AEB.2019.01.006.


## 1 Introduction

The genus *Cryobacterium* is a member of the family *Microbacteriaceae* in the phylum *Actinobacteria*. The genus *Cryobacterium* was first discovered (Suzuki et al., 1997), and contained only one species *Cryobacterium psychrophilum*, which was firstly named "*Curtobacterium psychrophilum*". *C. psychrophilum* was renamed to *Cryobacterium psychrophilum* in 1997. Up to now, the genus *Cryobacterium* comprises eight others reported species, namely *Cryobacterium psychrotolerans* (Zhang et al., 2007), *Cryobacterium arcticum* (Bajerski et al., 2011), *Cryobacterium mesophilium* (Dastager et al., 2008), *Cryobacterium roopkundense* (Reddy et al., 2010), *Cryobacterium flavum, Cryobacterium luteum* (Liu et al., 2009), *Cryobacterium levicorallium* (Liu et al., 2013) and *Cryobacterium aureum* (Liu et al., 2018). Of the known species, eight are psychrophilic and only one is mesophilic.

Species of this genus were isolated from geographical diverse ecosystems. The distribution of *Cryobacterium* has been found in various environments, such as glacier, soil, and sedimentary rock (Lee et al., 2016). Much attention was paid to *Cryobacterium* in the investigations of diversity, ecological distribution, and natrual products. Among its important functions, there are the production of amylase, protease and lipase. These productions were reported with the *Cryobacterium* sp. MLB-32 in 2014 (Singh, 2014). These characteristics were confirmed by the complete sequencing of the type species *C. arcticum* DSM 22823$^T$, *C. roopkundense* DSM 21065$^T$, and *C. luteum* CGMCC 1.11210$^T$.

Despite the potential importance of *Cryobacterium* in biotechnology and synthesis of natural products, few genomic analysis have been carried out to facilitate the applications of biotechnology and searching for natural products. In this study, we provided a high-quality complete sequence of *Cryobacterium* sp. GCJ02, isolated from cryomorphic soil of a virgin forest. This bacterium was determined to perform a whole genomic sequencing due to its excellent cold-adaptation at low temperature. Additionally, the genomic analyse will contribute to investigation of the microbial diversity in virgin forest (undisturbed area), and perform comparative genomic analysis of *Cryobacterium* speciess in databases of bioinformatics.

## 2 Materials and Methods

### 2.1 Bacterial Isolate and DNA Extraction

*Cryobacterium* sp. GCJ02 (=JCM 32391) was isolated from the cryomorphic soil of a virgin forest without any human activity in China. Humans are prohibited to enter into the area for conservation of biological diverisity. The

strain was characterized as a novel species of the genus *Cryobacterium* based on taxonomy study. GCJ02 showed excellent ability to adapt to low temperature. According to the physiological tests, GCJ02 was characterized as Gram-positive, rod-shaped, aerobic, and non-motile (Dongsheng Xue et al., unpublished data). The whole genome sequencing was carried out to perform the genetic manipulation and obtain the informations of cold adaptation of GCJ02.

The genomic DNA of GCJ02 was extracted using GenElute$^{TM}$ Bacterial Genomic DNA Isolation Kit (Sigma, USA) followed the manufacture's instructions. The extracted DNA was qualified using NanoDrop 2000 Microvolume UV-Vis spectrophotometer (Thermo Scientific).

## 2.2 Phylogenetic and Phylogenomic Analyses

Based on the description of the putative novel *Cryobacterium* species, the phylogenetic and taxonomic position of *Cryobacetrium* sp. GCJ02 was determined with 16S rRNA gene sequence (1,485 bp) from 30 strains' sequence within the *Cryobacterium* genus and related genus downloaded from GenBank database (http://www.ncbi.nlm.nih.gov) and EzTaxon-e server (http://eztaxon-e.ezbiocloud.net) (Yoon et al., 2016). Phylogenetic analysis was performed using the program MEGA 7.0 (Kumar et al., 2016), distance matrices were calculated according to the Kimura two-parameter model. Phylogenetic trees were inferred using the maximum-likelihood methods (Tamura et al., 1993). Bootstrap values were determined based on 1,000 replications. DNA-DNA hybridization (DDH) estimate values between strain GCJ02$^T$ and two close type strains (*C. arcticum* SK-1$^T$ and *C. psychrotolerans* CGMCC 1.5382$^T$) were analyzed using the genome-to-genome distance calculator (GGDC 2.0) with the alignment method of BLAST+ (Meierkolthoff et al., 2013; Auch et al., 2010; Auch et al., 2010), respectively.

## 2.3 Genome Sequencing and Assembly

The genome DNA of strain GCJ02 was broken into pieces around 10 kb, and SMRTbell library was prepared using the PacBio DNA Template Prep Kit for the strain. The PacBio RS library with targeted insert size of 10 kb was constructed. Next-generation sequencing was performed on PacBio RS platforms (Pacific Biosciences, Menlo Park, CA, USA) in one SMRT cell. Raw reads were filtered to retrieve high quality subreads. The clean data of genome was assembled using HGAP (Chin et al., 2013), aligned against the reference genome by BLAST, and unmapped regions were extracted and analyzed. Naming and classification of insertion sequences (ISs) in this study were according to Wolk et al. (2010).

## 2.4 Accession Number

The whole genome sequence of *Cryobacterium* sp. GCJ02 was submitted to the GenBnak database with the Accession number CP030033. The strain GCJ02 was deposied in the Japan Collection of Microorganisms under the accession number JCM 32391.

## 2.5 Genome Annotation

Based on the whole genome of GCJ02, the protein coding sequences (CDS)s were predicted by Glimmer (Version 3.02) (Delcher, 1999). The functions of predicted genes were annotated using NR, Swiss-prot, KEGG, and COG databases. Functional rRNA genes were detected using RNAmmer (Version 1.2) (Lagesen et al., 2007). The sequences and structure of tRNA were determined by tRNAscan-SE (Version 1.3.1) (Lowe and Eddy, 1997). The tandem repeats were picked up with RepeatMasker (Version 4.04) (Benson, 1999), and the ncRNA were sought by Rfam 12.0 (Griffiths-Jones et al., 2005). A circular genome map was generated using the Circos visualization tool created by Krzywinski et al. (2009).

# 3 Results and Discussion

The phylogenetic position of *Cryobacterium* sp. GCJ02 was determined using the complete 16S rRNA gene sequences from 30 strains in the genus *Cryobacterium* and others downloaded from GenBank database. The resulting phylogeny suggests that GCJ02 is closely ralated to *Cryobacterium arcticum* SK-1T, isolated from a soil sample in Northeast Greenland (Table 1). The pairwise alignment indicated that the percent identity between the two sequences is 99.49. Furthermore, the phylogenetic tree also indicated that the *Cryobacerium* genus held a monophyletic clade. The Average

**Table 1.** Top matches of the 16S rDNA sequence of strain GCJ02 against known bacterial sequences from the Genbank database (Ezbiocloud)

| Hit taxon name | Hit strain name | Accession | Similarity (%) |
|---|---|---|---|
| *Cryobacterium arcticum* | SK-1(T) | GQ406814 | 99.49 |
| *Cryobacterium psychrotolerans* | CGMCC 1.5382(T) | jgi.1076200 | 97.71 |
| *Cryobacterium roopkundense* | RuGl7(T) | EF467640 | 96.92 |
| *Cryobacterium flavum* | CGMCC 1.11215(T) | jgi.1076268 | 96.88 |
| *Cryobacterium luteum* | CGMCC 1.11210(T) | jgi.1076264 | 96.74 |
| *Cryobacterium levicorallinum* | Hh34(T) | JF267312 | 96.7 |
| *Cryobacterium psychrophilum* | DSM 4854(T) | AJ544063 | 96.46 |
| *Cryobacterium mesophilum* | MSL-15(T) | EF466127 | 96.33 |

Nucleotide Identity (ANI) and the DNA-DNA hybridization estimate values between strain GCJ02 and the closest type strain *C. arcticum* SK-1T were 83.99 and 28.0%, respectively. Thus, based on the 16S rRNA gene phylogeny, ANI and DDH results, the strain GCJ02 is a novel species.

## 3.1 General Genomic Features

The GCJ02 genome size is 4,385,465 bp, which is distributed on 1 contig. A circlar map of chromosome and the features of the genome is shown in Fig. 1 and Table 2, respectively. The contig N50 is 4,395,815 bp, and the G+C contents are 68.41 mol%, consistent with the reported datas of *Cryobacterium* species, ranging from 64.3% to 68.6%. The distribution of basepair is shown in Table 2 and the sequence length is

4,385,465 after circlization. Sequencing of the library generated 97,291 subreads with an average length of 12,474 bp with one SMRT cell of PacBio RS sequencing. Raw reads were filtered to retrieve high quality subreads. The clean data was assembled using hierarchical genome-assembly process (HGAP). The genome sequence of strain GCJ02 was submitted to GenBank of National Centre for Biotechnology Information (NCBI) with Accession Number CP030033. The summarized information on the genes of GCJ02 is exhibited in Table 2. Based on the prediction, there are 4,298 genes, including 4,139 coding sequences (CDS), and total 159 RNAs (10 rRNAs, 50 tRNAs and 99 ncRNAs). Among the CDS, a total of 1,885 gene functions were annotated by aligning to KEGG database (Kanehisa et al., 2010). The predicted CDS were related to a various of metabolic pathways which associated with amino acid, nucleotide, carbohydrate, and energy metabolism; synthesis of glycan, lipid, cofactor, vitamin, and other secondary metabolites; cellular process; membrane transport; signal transduction; transcription; translation; environmental adaptation etc. As genomic elements, the predicted CDS reflect the global functions of the cells.

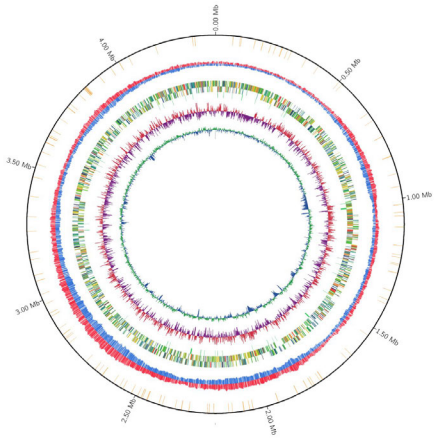

**Figure 1.** Circular map of the Cryobacterium sp. GCJ02 chromosome From the center to outside: GC content, GC skew, tRNA and rRNA, genes in backward strand, colored according to COG function categories, genes in forward strand, colored according to COG function categories, m4C and m6A sites in genome, Distribution of genes belonging to the restriction modification system. The map was generated using CGView

**Table 2.** *Cryobacterium* sp. GCJ02 genome statistics

| Attribute | Value |
|---|---|
| Genome size (bp) | 4,385,465 |
| G+C content | 68.41 |
| Total genes | 4298 |
| rRNAs | 10 |
| tRNAs | 50 |
| ncRNAs | 99 |
| Protein coding genes (CDS) | 4139 |

## 3.2 Insights from the Genome Sequence

### 3.2.1 Butirosin and Neomycin Biosynthesis

Butirosin and neomycin were aminoglycoside antibiotics, and identified in 2000 and in 2005, respectively. Lots of biosynthetic gene clusters of butirosin and neomycin encoding for related enzymes have been characterized in the last decade. The butirosin and neomycin are associated with bacteria, such as *Bacillus circulans*(Kudo et al., 1999a) and *E. coli* (Kudo et al., 1999b). The genes are present in the genome of *Cryobacterium* sp. GCJ02.

Important genes related to biosynthesis of butirosin and neomycin were detected, for example, CDSs encoding for glucokinase, glycosyl hydrolase, glucose dehydrogenase, and N-acetylmannosamine-6-phosphate epimerase. The gene clusters coding for enzymes involved in butirosin and neomycin biosynthesis were not reported in genus *Cryobacterium*. We propose that some *Cryobacterium* species have the potential to produce antibiotics to resist external pathogens.

### 3.2.2 Identification of Benzoate Degradation Genes

In the recent years,the natural environment has been polluted by harmful organic pollutants, including aromatics, halogen and polyaromatic compounds. Microorganisms play an important role in degradation of toxic chemicals and provide an efficient measure for environmental protection. In the genome of GCJ02, six genes are putative for involvement in benzoate degradation, specific genes encoding for glyoxalase, acetyltransferase, and hydroxyacyl-CoA dehydrogenase, indicating the ability to metabolize benzoate. Five proteins encoded by genes involed in aminobenzoate degradation were found in the genome including amidase, acylphosphatase and enoyl-CoA hydratase.

# 4 Conclusions

Genomic analysis of a putative novel *Cryobacterium*. sp GCJ02 reveals a high degree of consistency between genotypes, especially in butirosin and neomycin biosynthesis and benzoic acid degradation. The genome sequence of GCJ02 also reveals the genes involved in biosynthesis of amino acid, metabolism, and other cellular processes. Functional analysis of these genes will be reported in the future. The genomic analysis of this putative novel species facilitates to elucidate the molecular mechanism of cold-adaptation and perform the genetic manipulation of the bacterium.

# Acknowledgement

This work was financially supported by The Foundation of Hubei Provincial Department of Education (B2018040).

# References


Auch, A.F., Klenk, H.P. and GoKer, M., 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Standards in Genomic Sciences, 2(1), 142-148. http://www.doi.org/10.4056/sigs.541628

Auch, A.F., Von Jan, M., Klenk, H.P. and GoKer, M., 2010. Digital dna-dna hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Standards in Genomic Sciences, 2(1), 117-134. http://www.doi.org/10.4056/sigs.531120

Bajerski, F., Ganzert, L., Mangelsdorf, K., André Lipski and Wagner, D., 2011. *Cryobacterium* arcticum sp. nov. a psychrotolerant bacterium from an arctic soil. International Journal of Systematic and Evolutionary Microbiology, 61(Pt 8), 1849-1853. http://www.doi.org/10.1099/ijs.0.027128-0

Benson, G. 1999. Tandem repeats finder: a program to analyze dna sequences. Nucleic Acids Research, 27(2), 573-580.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., and Heiner, C., 2013. Nonhybrid, finished microbial genome assemblies from long-read smrt sequencing data. Nature Methods, 10(6), 563. http://www.doi.org/ 10.1038/nmeth.2474

Dastager, S.G., Lee, J.C., Ju, Y.J., Park, D.J., Kim, C.J., 2008. *Cryobacterium* mesophilum sp. nov., a novel mesophilic bacterium. International Journal of Systematic and Evolutionary Microbiology, 58(10), 1241-1244. http://www.doi.org/10.1099/ijs.0.65650-0

Delcher, A., 1999. Improved microbial gene identification with glimmer. Nucleic Acids Research, 27(23), 4636-4641.

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., and Bateman, A., 2005. Rfam: annotating non-coding rnas in complete genomes. Nucleic Acids Research, 33 (Database issue), D121-4. http://www.doi.org/10.1093/nar/gki081

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., and Horsman, D., 2009. Circos: an information aesthetic for comparative genomics. Genome Research, 19(9), 1639-1645. http://www.doi.org/10.1101/gr.092759.109

Kudo, F., Hosomi, Y., Tamegai, H., and Kakinuma, K. 1999a. Purification and characterization of 2-deoxy-scyllo-inosose synthase derived from bacillus circulans. a crucial carbocyclization enzyme in the biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. The Journal of Antibiotics, 52(2), 81-88.

Kudo, F., Tamegai, H., Fujiwara, T., Tagami, U., Hirayama, K., and Kakinuma, K., 1999b. Molecular cloning of the gene for the key carbocycle-forming enzyme in the biosynthesis of 2-deoxystreptamine-containing aminocyclitol antibiotics and its comparison with dehydroquinate synthase. Journal of Antibiotics, 52(6), 559-71.

Kumar, S., Stecher, G., and Tamura, K., 2016. Mega 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, msw054. http://www.doi.org/10.1093/molbev/msw054

Lagesen, K., Hallin, P., Einar, A.R., Staerfeldt, H.H., and Ussery, D.W., 2007. Rnammer: consistent and rapid annotation of ribosomal rna genes. nucleic acids res. Nucleic Acids Research, 35(9), 3100-3108. http://www.doi.org/10.1093/nar/gkm160

Lee, J., Cho, A., Yang, J.Y., Woo, J., Lee, H.K. and Hong, S.G., 2016. Complete genome sequence of *cryobacterium arcticum*, strain pamc 27867, isolated from a sedimentary rock sample in northern victoria land, antarctica. Genome Announcements, 4(5), e00885-16. http://www.doi.org/10.1128/genomeA.00885-16

Liu, Q., Liu, H., Wen, Y., Zhou, Y. and Xin, Y., 2009. *Cryobacterium* flavum sp. nov. and *Cryobacterium* luteum sp. nov., isolated from glacier ice. International Journal of Systematic and Evolutionary Microbiology, 59,2738-2742. http://www.doi.org/10.1099/ijs.0.033738-0

Liu, Q., Liu, H., Wen, Y., Zhou, Y. and Xin, Y., 2013. *Cryobacterium* levicorallinum sp. nov., a psychrophilic bacterium isolated from glacier ice. International Journal of Systematic and Evolutionary Microbiology, 63,2819-2822. http://www.doi.org/10.1099/ijs.0.046896-0

Liu, Q., Xin, Y. H., Chen, X. L., Liu, H. C. and Chen, W. X., 2018. *Cryobacterium* aureum sp. nov. a psychrophilic bacterium isolated from glacier ice collected from the ice tongue surface. International Journal of Systematic and Evolutionary Microbiology, 68(4),1173-1176. http://www.doi.org/10.1099/ijsem.0.002647

Lowe, T.M. and Eddy, S.R., 1997. Trnascan-se: a program for improved detection of transfer rna genes in genomic sequence. Nucleic Acids Research, 25(5), 955-964.

Meierkolthoff, J.P., Auch, A.F., Klenk, H.P. and Markus Göker., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics, 14(1), 60. http://www.doi.org/10.1186/1471-2105-14-60.

Reddy, G.S.N., Pradhan, S., Manorama, R. and Shivaji, S., 2010. *Cryobacterium* roopkundense sp. nov. a psychrophilic bacterium isolated from glacial soil. International Journal of Systematic and Evolutionary Microbiology, 60(4), 866-870. http://www.doi.org/10.1099/ijs.0.011775-0

Singh, P., Singh, S. M. and Dhakephalkar, P., 2014. Diversity, cold active enzymes and adaptation strategies of bacteria inhabiting glacier cryoconite holes of high arctic. Extremophiles, 18(2), 229-242. http://www.doi.org/10.1099/ijs.0.027128-0

Stothard, P. and Wishart, D.S., 2005. Circular genome visualization and exploration using cgview. Bioinformatics, 21(4), 537-539. http://www.doi.org/10.1093/bioinformatics/bti054

Suzuki, K., Sasaki, J., Uramoto, M., Nakase, T. and Komagata, K., 1997. *Cryobacterium psychrophilum* gen. nov. sp. nov. nom. rev. comb. nov. an obligately psychrophilic actinomycete to accommodate "*curtobacterium psychrophilum*" inoue and komagata 1976. International Journal of Systematic Bacteriology, 47(2), 474. http://www.doi.org/10.1038/nrmicro2186

Tamura, K. and Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial dna in humans and chimpanzees. Molecular Biology and Evolution, 10(3), 512-526. http://www.doi.org/10.1093/oxfordjournals.molbev.a040023

Wolk, C.P., Lechno-Yossef, S., and Jager, K.M., 2010. The insertion sequences of anabaena sp. strain pcc 7120 and their

effects on its open reading frames. Journal of Bacteriology, 192(20), 5289-5303. http://www.doi.org/10.1128/JB.00460-10

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J. and Chun, J., 2016. Introducing ezbiocloud: a taxonomically united database of 16s rrna and whole genome assemblies. International Journal of Systematic and Evolutionary Microbiology, 67(5), 1613-1617. http://www.doi.org/10.1099/ijsem.0.001755

Zhang, D.C., Wang, H.X., Cui, H.L., Yang, Y., Liu, H.C. and Dong, X.Z., 2007. *Cryobacterium psychrotolerans* sp. nov. a novel psychrotolerant bacterium isolated from the china no. 1 glacier. International Journal of Systematic and Evolutionary Microbiology, 57(4), 866-869. http://www.doi.org/10.1099/ijs.0.64750-0