Data Article

# Draft genome sequence data of the starch-utilizing yeast *Saccharomycopsis fibuligera* MBY1320 isolated from Nuruk

Eun-Hee Park[a], Myoung-Dong Kim[a,b,*]

[a] *Department of Food Science and Biotechnology, Kangwon National University, Chuncheon 24341, Korea*
[b] *Institute of Fermentation and Brewing, Kangwon National University, Chuncheon 24341, Korea*

## ARTICLE INFO

## ABSTRACT

*Saccharomycopsis fibuligera* MBY1320 was isolated from Nuruk, a traditional Korean starter for makgeolli (rice wine) production. This isolate has previously been reported to exhibit thermotolerance and starch-degrading activities. In this data description, we present the draft genome sequence of *S. fibuligera* MBY1320. The genome contained 19,138,941 bp in 16 contigs, and the internal transcribed spacer region of *S. fibuligera* MBY1320 rRNA gene showed the highest similarity with that of *S. fibuligera* NRRL Y-2388. The BioProject sequence has been deposited at DDBJ/EMBL/GenBank and Mendeley database. The GenBank accession numbers are PRJNA598085 for the BioProject data, SAMN13698230 for the BioSample data, and GCA_012062855 for the GenBank data.

---

* Corresponding author.
  *E-mail address:* mdkim@kangwon.ac.kr (M.-D. Kim).

## Specifications Table

| | |
|---|---|
| Subject | Biology |
| Specific subject area | Microbiology, Genomics, Biotechnology |
| Type of data | Table |
| How data were acquired | The draft genome sequence was determined using a Pac-Bio RSII instrument |
| Data format | Raw, analyzed and deposited |
| Parameters for data collection | Isolation of *S. fibuligera* MBY1320 from Nuruk, a traditional Korean starter for makgeolli (rice wine) fermentation. Genomic DNA extraction and sequencing procedures were used. |
| Description of data collection | Genomic DNA was isolated from a pure culture of the *S. fibuligera* MBY1320 strain |
| Data source location | Institution: Kangwon National University<br>City/Town/Region: Chuncheon, Kangwon-do, Republic of Korea<br>Latitude and longitude: 37°87′ N, 127°74′ E. |
| Data accessibility | The draft genome sequence of *S. fibuligera* MBY1320 is available at DDBJ/EMBL/GenBank under the accession no. GCA_012062855. The BioSample, BioProject, and assembly/WGS accession numbers are SAMN13698230, PRJNA598085(https://www.ncbi.nlm.nih.gov/bioproject/PRJNA598085), and WXEZ01, respectively.<br>All sequence files can be accessed at Mendeley Data (https://data.mendeley.com/datasets/vmr97×5mr7). |
| Related research article | D. H. Choi, E. H. Park, M. D. Kim, Characterization of starch-utilizing yeast *Saccharomycopsis fibuligera* isolated from nuruk, Microbiol. Biotechnol. Lett. 42(2014), 407–412. DOI: 10.4014/kjmb.1409.09006 [3]. |

## Value of the Data

- The draft genome of *S. fibuligera* MBY1320, isolated from Nuruk, will facilitate the analysis of the strain genetic information. Thus, this data could improve our understanding of starch-utilizing yeast.
- Data from the *S. fibuligera* MBY1320 strain will be useful for further study of starch metabolism in yeast.
- Data on the genome sequence of MBY1320 could be used to search and characterize biotechnologically-relevant enzymes and gene clusters in the future.

## 1. Data Description

The starch-degrading yeast *S. fibuligera* is used in food fermentation and uses rice and cassava [1] as substrates. *S. fibuligera* produces both $\alpha$-amylase and glucoamylase for starch-degradation [2,5]. Glucoamylase is also an essential enzyme in the food and fermentation industries. The *S. fibuligera* MBY1320 strain was isolated from Nuruk, a traditional Korean rice wine fermentation starter, showed high glucoamylase activity and thermotolerance [3].

Fig. 1 shows a phylogenetic tree comparing the internal transcribed spacer(ITS) region for *S. fibuligera* MBY1320 and other yeasts. The ITS of *S. fibuligera* MBY1320 (GenBank accession number MT944985) showed the highest similarity with that of *S. fibuligera* NRRL Y-2388.

Table 1 summarizes the main features for the draft genome sequence of *S. fibuligera* MBY1320. The assembled draft genome contained 19,138,941 bp distributed across 16 contigs, with a GC content of 38.2%. The longest contig had a length of 3974,907 bp, while the N50 contig was 2645,300 bp long. MAKER [4] predicted 2844 protein-coding genes that contained 524 hypothetical protein candidates. The predicted genes were annotated at 97% based on BLASTp similarity searches against an nr database selection. The data includes auxotrophic markers, such as the phosphoribosyl anthranilate isomerase gene (*TRP1*) [6] as well as the $\beta$-isopropyl malate dehydrogenase gene (*LEU2*) [7]. The data can be used to construct a host-vector tool for
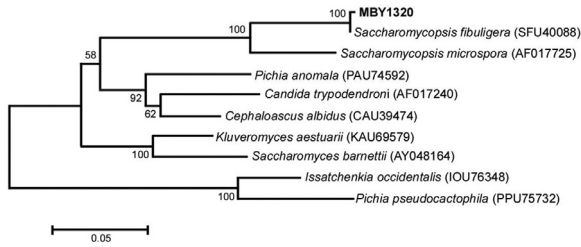
**Fig. 1.** Phylogenetic tree showing the *Saccharomycopsis fibuligera* ITS region compared with that of other yeasts. The tree was constructed using the neighbor-joining method. The numbers on the nodes correspond to the percentages, with which clusters appeared during the 1000 pseudoreplicate bootstrap tests. The bars denote the relative branch lengths. The ITS region was identified by GenBank accession numbers (shown in parentheses).

**Table 1**
Features of the draft genome of *S. fibuligera* MBY1320.

| Attribute | *S. fibuligera* MBY1320 Value |
| --- | --- |
| Total length (bp) | 19,138,941 |
| Number of contigs | 16 |
| N50 | 2645,300 |
| Number of predicted genes | 2844 |
| GC content (%) | 38.2 |

establishing an auxotrophic transformation system using the *S. fibuligera* MBY1320 strain. This draft genome sequence of *S. fibuligera* MBY1320 represents a new genomic information source that can provide valuable information for starch-related applications.

## 2. Experimental Design, Materials and Methods

Strain *S. fibuligera* MBY1320 was isolated from Nuruk, and was grown in YEPD liquid medium for 24 h at 30 °C. According to the manufacturer's instructions, genetic DNA was extracted using G-DEXTMIIc Genomic DNA Extraction Kit (iNtRON, Korea).

Sequences obtained by BLAST searching against the GenBank database were manually aligned with the ITS sequence of *S. fibuligera* MBY1320 using CLUSTAL_W. The phylogenetic tree was constructed with the neighbor-joining method in MEGA7.0.26.

The genome of *S. fibuligera* MBY1320 was sequenced using PacBio RSII (Pacific Biosciences, USA). The reads were assembled *de novo* into 16 contigs using HGAP2. The open reading frames were predicted using MAKER(http://www.yandell-lab.org/software/maker) [4] and subsequently annotated by SNAP(http://korflab.ucdavis.edu/software). Search of putative genes were conducted with BLASTP in NR database.

## Ethics Statement

Not applicable.

## CRediT Author Statement

**Eun-Hee Park:** Writing – Original draft; Writing – Revier & Edting; Visualization; Data curation; **Myoung-Dong Kim**: Concptualization, Supervision, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## Acknowledgments

## References

[1] Z. Chi, Z. Chi, G. Liu, F. Wang, L. Ju, T. Zhang, *Saccharomycopsis fibuligera* and its applications in biotechnology, Biotechnol. Adv. 27 (2009) 423–431.
[2] Z. Chi, J.M. Wang, Z.M. Chi, F. Ye, Trehalose accumulation from corn starch by *Saccharomycopsis fibuligera* A11 during 2-l fermentation and trehalose purification, J. Ind. Microbiol. Biotechnol. 37 (2010) 19–25.
[3] D.H. Choi, E.H. Park, M.D. Kim, Characterization of starch-utilizing yeast *Saccharomycopsis fibuligera* isolated from nuruk. Microbiol, Biotechnol. Lett. 42 (2014) 407–412.
[4] C. Holt, M. Yandell, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects, BMC Bioinformatics 12 (2011) 491.
[5] A.M. Knox, J.C. Preez, S.G. Kilian, Starch fermentation characteristics of *Saccharomyces cerevisiae* strains transformed with amylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*, Enzyme Microb. Technol. 34 (2004) 453–460.
[6] E.H. Park, M.D. Kim, Isolation of the phosphoribosyl anthranilate isomerase gene (*TRP1*) from starch-utilizing yeast *Saccharomycopsis fibuligera*, J. Microbiol. Biotechnol. 25 (2015) 1324–1327.
[7] E.H. Park, S.H. Yeo, M.D. Kim, Cloning of the *LEU2* gene from the amylolytic yeast *Saccharomycopsis fibuligera*, Food Sci. Biotechnol. 24 (2015) 2151–2154.