AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Whole-Genome Sequence of *Monascus purpureus* GB-01, an Industrial Strain for Food Colorant Production

Toshitaka Kumagai,[a] Masatoshi Tsukahara,[b] Naoya Katayama,[c] Katsuro Yaoi,[d] Sachiyo Aburatani,[e,f] Kohji Ohdan,[c] Kazuhiro E. Fujimori[d]

[a]Fermlab, Inc., Tokyo, Japan
[b]Biojet Company Ltd., Okinawa, Japan
[c]Biochemical Research Laboratory of Ezaki Glico Co. Ltd., Osaka, Japan
[d]Bioproduction Research Institute (BPRI), National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan
[e]Biotechnology Research Institute for Drug Discovery (BRD), National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan
[f]Computational Bio Big Data Open Innovation Lab (CBBD-OIL), National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan

ABSTRACT We report the draft genome sequence of *Monascus purpureus* GB-01, an industrial strain used as a food colorant. *De novo* assembly of long reads resulted in 121 chromosomal contigs and 1 mitochondrial contig, and sequencing errors were corrected by paired-end short reads. This genome sequence will provide useful information for azaphilone pigments and mycotoxin citrinin biosynthesis.

Utilization of the filamentous fungus *Monascus purpureus* (phylum Ascomycota) began more than 2,000 years ago for the production of red mold rice, a fermented food in Southeast Asian countries (1). Currently, it is widely used for production of yellow and red azaphilone pigments as food colorants. Recently, several research groups began conducting genetic and genomic studies of *Monascus ruber* (2, 3), *Monascus aurantiacus* (4), and *Monascus purpureus* (5, 6) using modern approaches to elucidate the biosynthesis of pigments and the mycotoxin citrinin. In addition, the first genome sequence of *M. purpureus* YY-1 was reported recently (7).

In this study, we obtained the *M. purpureus* strain GB-01 from Ezaki Glico Co., where the strain was collected half a century ago for the purpose of red pigment production. We selected GB-01 as a representative strain with high pigment production from among the stocks available at this company. *M. purpureus* GB-01 cells were grown in PSD100 medium (100 g/liter D-glucose, 38 g/liter polypeptone, 1.0 g/liter $MgSO_4 \cdot 7\ H_2O$, and 2.0 g/liter $NaNO_3$ [pH 5.5]). at 33.5°C for 4 days. The genomic DNA of *M. purpureus* GB-01 was isolated using an Isoplant II kit for short reads and a NucleoSpin plant II kit for long reads according to the manufacturer's instructions. Paired-end short reads were generated on the Illumina MiSeq platform using the Nextera DNA library preparation kit to generate libraries with different insert lengths and the MiSeq reagent kit 3 for sequencing runs. The numbers of reads totaled ~4.00 million and 4.67 million, and their mean insert lengths were 449 and 784 bases, respectively. Long reads of *M. purpureus* GB-01 were generated on the PacBio RS II platform using the PacBio SMRTbell template prep kit 1.0 and PacBio DNA/polymerase binding kit P6. The total number of subreads was ~1.23 million, the total number of bases was ~2,862 million, the mean subread length was ~2,318, bases and the N_{50} subread length was ~2,536 bases. To extract the mitochondrial reads, all long reads were mapped to the *Aspergillus nidulans* FGSC A4 complete mitochondrial genome sequence (GenBank accession number JQ435097) using Minialign 0.5.3 (8). The reads that did not map to the mitochondrial genome were processed as chromosomal reads and assembled *de novo* using Canu 1.7 (9). The assembly errors were corrected with the PacBio genomic

Citation Kumagai T, Tsukahara M, Katayama N, Yaoi K, Aburatani S, Ohdan K, Fujimori KE. 2019. Whole-genome sequence of *Monascus purpureus* GB-01, an industrial strain for food colorant production. Microbiol Resour Announc 8:e00196-19. https://doi.org/10.1128/MRA.00196-19.

Editor Jason E. Stajich, University of California, Riverside

Address correspondence to Kazuhiro E. Fujimori, k-fujimori@aist.go.jp.

Received 20 February 2019
Accepted 21 May 2019
Published 13 June 2019

consensus tool using the Arrow algorithm (Pacific Biosciences). In addition, to eliminate small indels, a final polish of the assembly was performed using Pilon 1.22 with nonredundant short paired-end reads of approximately 208.7× total coverage (10, 11). In the final assembly, we obtained 121 chromosomal contigs with 24.3 million total bases with an N_{50} value of 327,944 bases and 1 circular mitochondrial contig with 27,264 bases.

Obtaining multiple genomic sequences from *Monascus* spp. will help establish the molecular machineries for pigment and fungal toxin biosynthesis and may lead to the development of engineered strains with improved pigment productivity and lower mycotoxin levels.

Data availability. The draft genome sequence of GB-01 was deposited in DDBJ/GenBank under accession numbers BIYA00000000 for chromosomes and AP019407 for mitochondria, SRA accession number DRA007939, and BioProject number PRJDB7887.

ACKNOWLEDGMENTS

We thank N. Aoyagi and M. Hamajima for technical assistance with sample preparation and processing.

This study was supported by the New Energy and Industrial Technology Development Organization (NEDO).

K.Y., S.A., K.O., N.K., and K.E.F. participated in project conception and coordination; M.T. and K.E.F. participated in study design and carried out the experiments; T.K. and S.A. participated in data analyses; and K.E.F. wrote the whole manuscript. All the authors read and approved the final manuscript.

REFERENCES

1. Chen W, He Y, Zhou Y, Shao Y, Feng Y, Li M, Chen F. 2015. Edible filamentous fungi from the species *Monascus*: early traditional fermentations, modern molecular biology, and future genomics. Compr Rev Food Sci Food Saf 14:555–567. https://doi.org/10.1111/1541-4337.12145.
2. Liu Q, Xie N, He Y, Wang L, Shao Y, Zhao H, Chen F. 2014. *MpigE*, a gene involved in pigment biosynthesis in *Monascus ruber* M7. Appl Microbiol Biotechnol 98:285–296. https://doi.org/10.1007/s00253-013-5289-8.
3. Huang T, Tan H, Lu F, Chen G, Wu Z. 2017. Changing oxidoreduction potential to improve water-soluble yellow pigment production with *Monascus ruber* CGMCC 10910. Microb Cell Fact 16:208. https://doi.org/10.1186/s12934-017-0828-0.
4. Ning ZQ, Cui H, Xu Y, Huang ZB, Tu Z, Li YP. 2017. Deleting the citrinin biosynthesis-related gene, *ctnE*, to greatly reduce citrinin production in *Monascus aurantiacus* Li AS3.4384. Int J Food Microbiol 241:325–330. https://doi.org/10.1016/j.ijfoodmicro.2016.11.004.
5. Balakrishnan B, Chen C-C, Pan T-M, Kwon H-J. 2014. Mpp7 controls regioselective Knoevenagel condensation during the biosynthesis of *Monascus* azaphilone pigments. Tetrahedron Lett 55:1640–1643. https://doi.org/10.1016/j.tetlet.2014.01.090.
6. Liang B, Du XJ, Li P, Sun CC, Wang S. 2018. Investigation of citrinin and pigment biosynthesis mechanisms in *Monascus purpureus* by transcriptomic analysis. Front Microbiol 9:1374. https://doi.org/10.3389/fmicb.2018.01374.
7. Yang Y, Liu B, Du X, Li P, Liang B, Cheng X, Du L, Huang D, Wang L, Wang S. 2015. Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, *Monascus purpureus* YY-1. Sci Rep 5:8331. https://doi.org/10.1038/srep08331.
8. Suzuki H. Minialign. https://github.com/ocxtal/minialign/.
9. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
10. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
11. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. Bioinformatics 31:2032–2034. https://doi.org/10.1093/bioinformatics/btv098.