

GENOME SEQUENCING

The Complete Genome Sequence of *Amorphophallus titanum*, the Corpse Flower

Linda Frisse¹, Marc A Martinez², Stacy Pirro

¹ National Center of Biotechnology Information, National Institutes of Health, ² Carnivero

Keywords: genome, viridiplantae, corpse flower

<https://doi.org/10.56179/001c.37841>

Biodiversity Genomes

The Corpse Flower, or Titan Arum (*Amorphophallus titanum*) is a flowering plant in the family Araceae. endemic to a limited range in the rainforests of Sumatra, Indonesia. It is notable for two reasons: it produces the world's largest known unbranched flower, and it produces a strong odor of rotting meat to attract pollinators. We present the whole genome sequence of this species. A total of 335,712,220 paired-end Illumina reads consisting of 100.7G bases were obtained by Illumina sequencing the leaf tissue of a single individual. The reads were assembled by a *de novo* method followed by contig extension using related species as references. The raw and assembled data is publicly available via Genbank: Sequence Read Archive (SRR11565159) and genome assembly (GCA_024336825).

Introduction

Amorphophallus titanum is notable both for the large size and strong smell of its flower. The species does not have an annual flowering cycle, but instead produces a bloom every several years, although some individuals produce a bloom on a shorter cycle. The flower can reach over 10 ft (3m) in height and consists of a fragrant spadix of flowers wrapped by a spathe, which looks like a large petal. The spathe is a deep green on the outside and dark burgundy red on the inside, with a deeply furrowed texture.

When in bloom, the flower has an extremely strong odor of rotting meat. The smell attracts insects that normally feed on dead animals and these act as pollinators. The species has male and female forms, and is also capable of self-pollination. Self-pollination has occurred spontaneously in botanical garden settings when there has only been a single specimen present, and also manually by horticultural staff.

The species is endemic to a small region in Sumatra, Indonesia. Because of shrinking habitat, the species is considered endangered in the wild.

Methods

A single live cultivated plant was used for this study. The specimen was obtained from Carnivero (Austin, Texas, USA).

DNA extraction was performed using the Qiagen DNAeasy genomic extraction kit using the standard process. A paired-end sequencing library was constructed using the Illumina TruSeq kit, according to the manufacturer's instructions. The library was sequenced on an Illumina Hi-Seq platform in paired-end, 2 × 150bp format. The resulting fastq files were trimmed of

adapter/primer sequence and low-quality regions with Trimmomatic v0.33 (Bolger, Lohse, and Usadel 2014). The trimmed sequence was assembled by SPAdes v2.5 (Bankevich et al. 2012) followed by a finishing step using Zanfona v1.0 (Kieras 2021) to make additional contig joins based on conserved regions in related species.

Results

The genome assembly yielded a total sequence length of 942,822,506 bp over 9,603 scaffolds with an N50 of 15.37 MB.

Data availability

Raw and assembled data is publicly available via GenBank:

RAW GENOME DATA

<https://trace.ncbi.nlm.nih.gov/Traces/sra/?run=SRR11565159>

ASSEMBLED GENOME

https://www.ncbi.nlm.nih.gov/assembly/GCA_024336825

Funding

Funding was provided by Iridian Genomes, grant# **IRGEN_RG_2021-1345** Genomic Studies of Eukaryotic Taxa

This work was supported in part by the National Center for Biotechnology Information of the National Library of Medicine (NLM), National Institutes of Health.

Conflict of Interest Statement

The authors declare they have no conflict of interests.

Submitted: July 13, 2022 EDT, Accepted: August 12, 2022 EDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for more information.

REFERENCES

- Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. 2012. “SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.” *Journal of Computational Biology* 19 (5): 455–77. <https://doi.org/10.1089/cmb.2012.0021>.
- Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- Kieras, M. 2021. *Zanfona, a genome finishing process for use with paired-end short reads*. <https://github.com/zanfona734/zanfona>.