

GENOME SEQUENCING

The complete genome sequence of date palm (*Phoenix dactylifera* L.) variety 'Mejhouli'

Karim Rabeh¹, Fatima Gaboun², Stacy Pirro³, Adil Essarioui⁴

¹ Oasis System Research Unit, Regional Center of Agricultural Research of Errachidia, National Institute of Agricultural Research, ² Biotechnology Unit, National Institute for Agronomic Research (INRA), ³ Iridian Genomes, ⁴ Oasis System Research Unit, Regional Center of Agricultural Research of Errachidia, National Institute of Agricultural Research

<https://doi.org/10.56179/001c.156371>

Biodiversity Genomes

We present a high-quality genome sequence of Moroccan date palm (*Phoenix dactylifera* L.), variety 'Mejhouli', an economically and ecologically important tree to arid oasis regions. The genome was sequenced using Illumina Hi-Seq platform in paired-end sequencing and assembled with SPAdes v3.15.4. We present the complete genome sequence of *Phoenix dactylifera* L. Genomic DNA from the cultivar 'Mejhouli' was collected in Errachidia, Morocco, in July 2025 and sequenced using the Illumina platform. The reads were assembled using a de novo method followed by a finishing step. The raw and assembled data are publicly available via GenBank.

Introduction

Date palm (*Phoenix dactylifera* L.) is a keystone perennial crop of arid and semi-arid regions, supporting oasis agroecosystems and rural livelihoods under extreme environmental conditions. The Moroccan cultivar 'Mejhouli' (also known as 'Medjool') is globally prized for its large, tender, fleshy, elongated fruits, high sugar content, and high yield potential, as well as for its tolerance to drought and moderate salinity. Despite its agronomic and economic importance, the genomic basis of its key traits remains poorly characterized, limiting functional studies and crop improvement. Here, we present the first whole-genome sequence of 'Mejhouli', providing a high-quality reference genome for this variety. This resource establishes a foundation for trait-driven functional genomics, comparative analyses across cultivars, and molecular breeding strategies aimed at enhancing fruit quality, stress resilience, and productivity in climate-constrained environments.

Methods

Leaf tissue from a single live cultivated individual was used for genome sequencing in this study. DNA extraction was performed using the Qiagen DNEasy genomic extraction kit using the standard process. Paired-end sequencing libraries were constructed using the Illumina TruSeq kit according to the manufacturer's instructions. The libraries were sequenced on an Illumina Hi-Seq platform in paired-end, 2 × 150bp format. The resulting fastq files were trimmed of adapter/primer sequences and low-quality regions

with Trimmomatic v0.33 (Bolger et al. 2014). The trimmed sequence was assembled by SPAdes v3.15.4 (Bankevich et al. 2012) followed by a finishing step using Zanfona (Kieras et al. 2021).

Results and Data Availability

The genome assembly size was 486 MB with 44,654 scaffolds/contigs, and 34.9% GC.

taxname

Phoenix dactylifera

accession

JBTXPJ0000000000

.....

Funding

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

Submitted: February 06, 2026 EST. Accepted: February 10, 2026 EST. Published: February 10, 2026 EST.

REFERENCES

- Bankevich, Anton, Sergey Nurk, Dmitry Antipov, 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., K. O'Neill, and S. Pirro. 2021. "Zanfona, a Genome Assembly Finishing Tool for Paired-End Illumina Reads." <https://github.com/zanfona734/zanfona>.