2024 open GPB
Open Conference on Grapevine Physiology and Biotechnology
LOGROÑO - LA RIOJA
7-11 July 2024

A comprehensive and accurate annotation for the grapevine T2T genome

Antonio Santiago[1,2], David Navarro-Payá[1], Pascual Villalba-Bermell[1], Gustavo G. Gomez[1], Iñigo De Martín Agirre[1], Amandine Velt[3], Marco Moretto[4], Hua Xiao [5*], Yongfeng Zhou [5*], Camille Rustenholz[3*], José Tomás Matus[1*]

[1] *Institute for Integrative Systems Biology (I2SysBio), Universitat de València-CSIC, Paterna, 46980, Valencia, Spain*
[2] *Universitat Politècnica de València (UPV), 46022 Valencia, Spain*
[3] *SVQV, INRAE-University of Strasbourg, Colmar 68000, France*
[4] *Fondazione Edmund Mach Via E. Mach 1, Research and Innovation Centre, Via E. Mach 1, 38010 San Michele all'Adige, Italy*
[5] *Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China*

[*] *Corresponding author: tomas.matus@uv.es, camille.rustenholz@inrae.fr, xiaohua01@caas.cn, zhouyongfeng@caas.cn**

Addressing the opportunities and challenges of genomics methods in grapevine (*Vitis vinifera* L.) requires the development of a comprehensive and accurate reference genome and annotation. We aimed to create a new gene annotation for the PN40024 grapevine reference genome by integrating the highly accurate and complete T2T assembly and the manually curated PN40024.v4 annotation. Here, we present a novel workflow to enhance the annotation of the T2T genome by incorporating past community input found in PN40024.v4. The pipeline's containerization will improve the workflow's reproducibility and flexibility, facilitating its inclusion as a shared workflow on the Grapedia portal, the grapevine genomics encyclopedia. The pipeline includes both structural and functional annotation of genes, including lncRNAs and miRNAs. Moreover, we provide conversions between different annotation versions, allowing the comparison and integration of various grapevine datasets. To ensure transparency and traceability, we have modified the structure of the gene IDs, retaining the Vitvi prefix but also referencing the genome version and annotation. This will allow us to handle any gene model issues between different annotation versions and to easily distinguish the version from the annotation in publications. The annotation workflow will soon be available on the Grapedia portal (https://grapedia.org/), where it can potentially be applied to other cultivars. This annotation version will also serve as the basis for the new grape gene reference catalogue, which will provide a comprehensive and updated resource for grapevine genomics.

Keywords: Bioinformatics, genome annotation, grapevine, Grapedia, PN40024.