

PRELIMINARY DATA ON PISTACHIO (*PISTACIA VERA* L.) GENOME SEQUENCING AND ASSEMBLY

MARCHESE A.*, CARBONELL-BEJERANO² P.** , MARTÍNEZ E.***, MORALES A. M.***, LEE C.****, CATTIVELLI L.***** , BAGNARESI P.***** , ARMADORO S.***, GIOVINO A.***** , MARRA F. P.***** , MONROE G. J.****

*) Department of Agricultural, Food and Forest Sciences, University of Palermo, Viale delle Scienze - Ed. 4, 90128 Palermo, Italy

**) Institute for Grape and Wine Sciences (ICVV, CSIC - Gobierno de La Rioja - UR), 26007 Logroño, Spain

***) Regional Institute of Agri-Food and Forestry Research and Development of Castilla-La Mancha (IRIAF), IVICAM. Ctra. Toledo-Albacete s/n, 13700 Tomelloso, Spain

****) Department of Plant Sciences, University of California Davis, Davis, CA, 95616, USA

*****) CREA Research Centre for Genomics and bioinformatics, Fiorenzuola d'Arda, 29017, Italy

*****) Council for Agricultural Research and Economics (CREA), Research Centre for Plant Protection and Certification (CREA-DC), 90011 Bagheria, Italy

*****) Department of Architecture (DARCH), University of Palermo, Viale delle Scienze–Ed. 8, 90128 Palermo, Italy

Pistacia vera, PacBio long read sequencing, transcriptomics, pan genome, climate change

The pistachio (*P. vera* L., $2n = 30$) is a xerophytic, deciduous tree that belongs to the Anacardiaceae family, originated from Central and West Asia arid regions, including Iran and Afghanistan. It is largely cultivated for its nuts, having high economic, nutritional, and medicinal values. This dioecious species shows peculiar resilience to abiotic stresses, making it ideal for marginal, arid and salinized areas. However, the current global warming scenario may result in complications for its cultivation, since it requires hours of winter cold for correct vernalization of the buds and flowering synchronization of male and female genotypes. Many pistachio

varieties and male genotypes were selected under different climate conditions than these currently present in growing areas, placing pistachio fruit set at risk. In order to reduce this risk, an accurate selection supported by genomic technology is mandatory. Today, the rapid progress of genome sequencing combined with transcriptomic analysis has facilitated gene discovery in many species, however very few genomics studies are available in pistachio. In the present work, early flowering female and male genotypes (Mateur and T41, respectively), late flowering female genotypes (Kerman, Sirora and Napoletana) and a late flowering male genotype (Chaparrillo) were selected for genome sequencing with PacBio HiFi technology and assembly using the Hifiasm bioinformatics tool. The initial contig draft for the reference genome assembly of 'Kerman' cultivar comprises 602 Mb, with contig N50 = 28.4 Mb and 98.0% of BUSCO completeness. Highly continuous genome assemblies were indeed obtained for all genotypes ($6 \leq LG50 \leq 8$ for the primary contigs of all six cultivars, which is close to chromosome level). In addition, in all cases the two haplotypes could be assembled in at least 95% of their totality, according to a BUSCO analysis. The comparison of the different assembled haplotypes identifies a high heterozygosity and inter-varietal diversity. For annotation, transcriptome Iso-seq was performed in different tissues and phenological stages for all six genotypes; mRNA sequencing data were aligned to Kerman assembly, and putative protein-coding gene models were predicted. As the conventional breeding methods are not highly efficient to quickly develop varieties to respond to climate change, the identification of gene variants associated with agronomically important phenotypes could be exploited in breeding programs for appropriate male genotype and varietal choice to ensure the sustainability of pistachio crop in future climatic conditions and cultivation scenarios.