

A NEW REFERENCE GENOME SEQUENCE FOR CULTIVATED OLIVE TREE

CARBONE F.*, SCALABRIN S.***, BAGNARESI P.***, TACCONI G.***, SALIMONTI A.*, ZELASCO S.*, FORGIONE I.*, SIRANGELO T. M.*, DESIDERIO F.***, CATTIVELLI L.***, MORGANTE M.****, *****

*) Centro di ricerca Olivicoltura, Frutticoltura e Agrumicoltura, CREA, C.da Li Rocchi Vermicelli 87036 Rende (CS)

**) IGA Technology Services, Via J.Linussio 51 Z.I.U., 33100 Udine

***) Centro di ricerca Genomica e Bioinformatica, CREA, Via S. Protaso 302, 29017 Fiorenzuola d'Arda (PC)

****) Dipartimento di Scienze agroalimentari, ambientali e animali, Università di Udine, Via delle Scienze 208, 33100 Udine

*****) Istituto di Genomica Applicata, Via J.Linussio 51 Z.I.U., 33100 Udine

olive, reference genome, de-novo assembly, Hi-C

Olive (*Olea europaea* L. subsp. *europaea* var. *europaea*) is the oldest tree crop in the Mediterranean basin, highly heterozygous, diploid and is characterized by phenotypic plasticity and wide genetic variability allowing it to survive under different and complex agroecological conditions.

The Italian heterozygous cultivar 'Leccino' was sequenced and 79% of the estimated 1.4-gigabase haploid genome was assembled. Previous studies highlighted that the olive genome structure is peculiar among plant genomes, with a very large percentage of satellite DNA, related to a few tandem repeat families (about 31% of the whole genome), probably resulting in a large portion of the genome being unassembled. The genome has been sequenced using a combination of NGS (Illumina) and long read (ONT) technologies and a combination of assembly approaches producing contig and scaffold N50 sizes of 981 Kb and 40.70 Mb, respectively. We produced both a diploid haplotype-phased as well as a haploid assembly and we anchored 1.1 Gb of sequences of the haploid assembly to 23 pseudochromosomes by Hi-C and two genetic maps. A total of 88,708 protein-coding genes were predicted in the diploid assembly. The high-quality assembly of the Italian cultivar 'Leccino', together with the recently released genome of the Spanish cultivar 'Arbequina', greatly increased the accuracy of the olive genome and set 'Leccino' as reference variety for genomics and biological studies in olive tree providing a platform for isolation and functional characterization of genes and then for genetic improvement of the species.

A Genome Browser has been set up to visualize and to explore genome sequence and annotation data, and it is currently available at the reserved area of web portal <http://olgenome.crea.gov.it>.

Work supported by the Italian Ministry of agricultural food and forestry policies Project OLGENOME DM 13939 – 24/04/2018