

## THE HAPLOTYPE-PHASED GENOME ASSEMBLY FOR *FICUS CARICA* L.: AN ANCIENT CROP WITH PROMISING PERSPECTIVES

CASTELLACCI M.\*, USAI G.\*, VANGELISTI A.\*, VENTIMIGLIA M.\*, SIMONI S.\*, NATALI L.\*, CAVALLINI A.\*, MASCAGNI F.\*, GIORDANI T.\*

\*) Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80 56124, Pisa, Italy

*fig tree, fig genome, genome assembly, genetic variability*

The availability of genome sequences facilitates the application of modern crop breeding procedures. In this context, haplotype-resolved genomes represent crucial resources to drive the understanding of the genotype-phenotype relationship and assess the presence of allele-specific expression and regulation patterns. This is especially true for fruit trees, maintained via clonal propagation to control their heterozygosity levels. Among these, the fig tree (*Ficus carica* L.) has the potential to increase its economic relevance due to its fruit nutritional and nutraceutical features, paired with the plant ability to adapt to adverse environmental conditions, i.e., marginal soils and drought. Despite these promising features, high-quality genomic resources were not delivered until recently.

Here we describe the fig haplotype-phased genome assembly that was obtained by combining single-molecule real-time technology with a method of chromosome conformation capture. In total, 538 homologous sequences were obtained representing the two pseudo-haplotypes, accounting for 98% of the fig genome (estimated size: 356 Mb). Four hundred sequences were assigned to a fig chromosome, delivering a high-quality assembly (fig genome v2.0). Protein-coding genes were derived by integrating transcriptomics data analysis, protein alignment and *de novo* gene calling, obtaining a total of 33,954 and 33,379 protein-coding genes per pseudo-haplotype, respectively, 82% of which were annotated with a known function. The repetitive component was analysed with similar efficiency.

Based on such reference genome, the genetic variability of several fig

varieties (from Spanish, Tunisian and Turkish fig collections) will be assessed with a genotyping-by-sequencing approach in the FIGGEN project of the PRIMA action. These genomics data will allow determining the linkage of molecular markers with traits of interest, such as fruit quality and environmental adaptation, favouring the species genetic improvement.