

# V<sup>th</sup> International Symposium

on Pomegranate & Minor Mediterranean Fruits

Towards a Resilient and Prosperous Future



VIRTUAL SYMPOSIUM | BOOK OF ABSTRACTS & PROGRAMME

15-17 February 2022



V<sup>th</sup> International Symposium  
on Pomegranate & Minor Mediterranean Fruits  
15-17 FEB 2022



## Haplotype-phased genome assembly for *Ficus carica* breeding

### **TOMMASO GIORDANI**

Via del Borghetto 80, 56124 Pisa, Italy

### **GABRIELE USAI**

Via del Borghetto 80, Pisa, Italy

### **ALBERTO VANGELISTI**

Via del Borghetto 80, Pisa, Italy

### **FLAVIA MASCAGNI**

Via del Borghetto 80, Pisa, Italy

### **MARIA VENTIMIGLIA**

Via del Borghetto 80, Pisa, Italy

### **SAMUEL SIMONI**

Via del Borghetto 80, Pisa, Italy

### **LUCIA NATALI**

Via del Borghetto 80, Pisa, Italy

### **ANDREA CAVALLINI**

Via del Borghetto 80, Pisa, Italy

The availability of the genome sequence is a key prerequisite to applying modern breeding procedures to crops, and it is increasingly important to decipher the sequence of the two haplotypes that make up the genome, representing a pivotal resource to study the allele-specific expression and regulation patterns. This is all the more true for fruit trees, whose condition of heterozygosity is maintained by clonal propagation. The fig tree (*Ficus carica* L.) has a great potential for commercial expansion thanks to its esteemed nutritional and nutraceutical characteristics, combined with its ability to adapt well to marginal soils and difficult environmental conditions. However, high-quality genomic resources have been released only in recent years. Here, we report on assembling a haplotype-phased genome, achieved by combining single-molecule, real-time sequencing technology with a methodology of chromosome conformation capture. Two pseudo-haplotypes of 538 sequences were obtained, representing ~98% of the estimated 356 Mb fig genome. A total of 400 out of 538 sequences were associated with the 13 fig chromosomes generating the fig genome v2.0. *Ad hoc* approaches including RNA-seq data, protein alignment and *de novo* prediction allowed us to predict 34,288, and 33,698 protein-coding genes per pseudo-haplotype, respectively, and ~82% of the total protein-coding genes were functionally annotated. The same level of accuracy was achieved to analyze the repetitive DNA. The sequence of the fig genome is being used to evaluate genetic variability of fig varieties on available Spanish, Tunisian and Turkish fig collections using a genotyping by sequencing approach in the frame of a PRIMA project, FIGGEN. This is the prerequisite for carrying out genome-wide association studies (GWAS) to identify genes or molecular markers linked to several traits linked to fruit quality and the adaptation to difficult environmental conditions, leading to the genetic improvement of this species.