

CHARACTERIZATION OF THE REPETITIVE COMPONENT OF THE FIG (*FICUS CARICA*) GENOME

USAI G.*, MASCAGNI F.*, ZUCCOLO A.***, GIORDANI T.*, NATALI L.*, CAVALLINI A.*

*) Dept. of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa (Italy)

**) Institute of Life Sciences, Scuola Superiore Sant'Anna, Pisa (Italy)

Ficus carica, genome structure, repetitive component, retrotransposons

The fig tree (*Ficus carica* L.) is an ancient crop with promising perspectives for Italian Mediterranean agriculture. Fig is widely grown throughout the temperate world, both for its fruit and as an ornamental plant. In recent years, large interest arose on the nutraceutical properties of fig fruit, especially dried.

Considering the low level of genetic improvement of the present fig cultivars, a genomic approach can be useful to speed up the development of new cultivars. Despite its economic, cultural and ecological importance in many areas of the world, fig is however still poorly characterized at genetic and genomic level compared to other fruit tree crops.

Ficus is one of the thirty-seven genera of the Moraceae family. *Ficus carica* L. is a diploid species, with a small/medium genome size. The aim of this work was to gain an insight into fig genome structure.

After DNA isolation, Illumina sequencing was carried out using both HiSeq2000 and MiSeq platforms. Different assembling and clustering approaches were performed for characterizing the repetitive component of the genome. The fig genome resulted composed for around 60% of repeated sequences, of which none was especially redundant. All types of repeats already described in plants were found, however the majority of assembled repeats were not identified. Among identified repeats, ribosomal DNA constituted around 5% of the genome. The most represented repeats were LTR-retrotransposons (20% of the genome), with *Gypsy* elements much more frequent than *Copia*. LTR-retrotransposons were also characterized with regard to their age of insertion in the genome.

A comparative analysis of *Ficus carica*, *Morus notabilis* and *Malus domestica* was carried out using RepeatExplorer, performing a hybrid clustering of Illumina sequence reads. These species all belong to the *Rosales* order: *F. carica* and *M. notabilis* to the *Moraceae* family while *M. domestica* to the *Rosaceae* family. Results highlighted a large diversification among the repetitive component of the genomes of these related species, even within the *Moraceae* family. In fact, excluding ribosomal DNA related clusters, only 15 out of 227 hybrid repeat clusters were found across all the species, while fig and *M. notabilis* shared only 96/227 clusters. Finally, 29 clusters were specific to the fig, 24 to *M. notabilis* and 45 to *M. domestica*.

This comparison was complemented by means of phylogenetic analysis. Paralogs of the reverse transcriptase coding domain were identified in the three aforementioned species and used to build Neighbor-Joining trees for both Ty1-*Copia* and Ty3-*Gypsy* elements.

Studies are in progress to extend the characterization of fig genome to its gene component, leading to the identification of important genes involved in agronomic and productive traits as fruit

production and quality, biotic and abiotic stress resistance, and synthesis and accumulation of metabolites. These data will offer a first start for new opportunities of fig breeding.