

GENOME EVOLUTION IN GRAPEVINE: THE ROLE OF TRANSPOSABLE ELEMENTS AND WHOLE GENOME DUPLICATIONS

MORGANTE M.***, THE FRENCH-ITALIAN PUBLIC CONSORTIUM FOR GRAPEVINE GENOME CHARACTERIZATION

*) Istituto di Genomica Applicata, Via Linussio 51, 33100 Udine (Italy)

***) Dipartimento di Scienze Agrarie e Ambientali, Università di Udine, Via delle Scienze 208, 33100 Udine (Italy)

grape, whole genome duplication, paleopolyploidy, LINE, intron

The grapevine (*Vitis vinifera*) is economically the principal fruit plant in the world. The grapevine genome represented the first fruit tree whose sequence was made publicly available. Now, more than 90% of a 12X genome sequence assembly is anchored in less than 200 pieces on the 19 chromosomes.

Examination of the reference genomic sequence obtained from the highly-homozygous genotype PN40024 reveals that 41.4 % of the genome is composed of repetitive sequences and/or TEs. The repetitive portion is slightly higher than that in rice, which however has a somewhat smaller sized genome. All known classes and superfamilies of TEs are represented in the grapevine genome, with a large prevalence of class I elements. In contrast to Arabidopsis and rice, in *Vitis vinifera* we observe large regions along the chromosomes that alternate between high and low gene density. Among class I elements, LTR-retrotransposons prevail with a larger abundance of the Copia that of the Gypsy superfamily, but LINES (long interspersed nuclear elements) are also present in high copy number and are more abundant than in any previously sequenced plant genome. An analysis of the genomic distribution of the repetitive elements shows that, unlike in other plants, grapevine introns are quite rich in repeats and TEs, which considerably increase their size. In particular, when compared to other plant genomes, mainly LINE retrotransposons seem to have contributed specifically to the larger intron size observed in grapevine.

The re-sequencing of a second genotype, the cultivar Tocai friulano, allowed the identification of structural variations (SVs) between the two genotypes. We frequently observed size differences between Tocai and the reference sequence that could be attributed to recent insertions into PN40024 of TE elements that are missing in Tocai. Many of these insertions, that were also experimentally validated, are attributable to LINE elements and have occurred in the introns. Transcriptional activity of these elements can be detected in specific tissues. The finding of SVs between two grapevine varieties confirms the dynamic nature of the genomes, revealing that TE activity is an important source for the generation of genetic diversity.

In contrast to the young age and considerable transcriptional and transpositional activity that can be detected in the repetitive fraction of the grapevine genome, a remarkable evolutionary stability in genome and chromosome structure, compared to other sequenced plants, permitted to reveal that a vast fraction of the genome can be found as triplicated segments. Such an arrangement can be explained by hypothesizing a paleohexaploidization event: the comparative analysis of the Arabidopsis, poplar and papaya genomes confirms that the hexaploidization event predates the divergence among all these species and is therefore to be placed at the basis of eudicot evolution.

Grapevine, like papaya, did not undergo any further round of genome duplication while poplar underwent one and Arabidopsis two. Thanks to small number of genomic rearrangements that grape underwent after the hexaploidization event, it was possible to reconstruct the ancestral genome that gave rise to the hexaploid ancestor and that was formed by seven chromosomes.