

SURVEY OF TRANSPOSABLE ELEMENTS EXAPTATION EVENTS IN SUNFLOWER (*HELIANTHUS ANNUUS L.*) GENOME

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

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

**) Department of Biotechnology, University of Verona, Cà Vignal 1, Strada Le Grazie 15, 37134 Verona (Italy)

***) Center for Desert Agriculture, Biological and Environmental Sciences & Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900 (Saudi Arabia); Institute of Life Sciences, Scuola Superiore Sant'Anna,

Helianthus annuus, transposable elements, exaptation, genome evolution, bioinformatic tools

The ability to move within the genome makes the activity of transposable elements (TEs) a source of variability, contributing to the evolution of their host. The coding regions of TEs that during evolution lost their self-replicative capability, can acquire new functions and start evolving under phenotypic selective pressure. These “new” genes derived from the molecular domestication of TEs are defined as exapted transposable element genes (ETEs). Despite their biological relevance, there are currently few protocols for ETEs systematic discovery. The sunflower (*Helianthus annuus* L.) has an estimated genome size of 3.6 Gbp and its repetitive component, mostly represented by TEs, amounts to about 80% of it, providing a significant source for exaptation. A library of long terminal repeat retroelements (LTR-REs) and terminal inverted repeat (TIR) elements identified in sunflower was used to build TEs specific hidden Markov model (HMM) profiles. HMM profiles were then used to search all the available gene products of sunflower. The 20,016 putative ETEs identified with this approach were further investigated for the characteristics that distinguish TEs from genes, namely repetitiveness, expression, siRNA coverage and similarity with already known TEs; through this process 3,530 sunflower ETEs were validated. The functional characterization of these genes by Gene Ontology (GO) showed a significant involvement in disparate cellular functions, suggesting that ETEs affected several biological processes during sunflower evolution. Furthermore, distribution of functional domains in ETEs, identified by Pfam, revealed that a relevant fraction of TIR-derived genes belong to different transcription factor (TF) families; this result is in agreement with previous studies indicating that ETEs derived by DNA-TEs are mostly TFs. In conclusion, the identification and characterization of ETEs in sunflower represents a useful genomic resource, but also highlights that the phenomenon of exaptation plays a crucial role in the creation of sequences with novel functions, thus driving species evolution.