

## **BSA-SEQ ANALYSIS REVEALS GENOMIC REGIONS CARRYING CANDIDATE GENES FOR MALE STERILITY PHENOTYPE IN GLOBE ARTICHOKE**

ACQUADRO A.\*, ZAYAS A.\*\*\*, MARTINA M.\*, POLLI M. F.\*, DI NARDO G.\*\*\*, COMINO C.\*, GILARDI G.\*\*\*, MARTIN E., PORTIS E.\*

\*) DISAFA, Plant Genetics and Breeding, University of Torino, Italy

\*\*) IICAR-CONICET (Instituto de Investigaciones en Ciencias Agrarias de Rosario), Zavalla (SF), Argentina

\*\*\*) DBIOS, Department of Life Sciences and Systems Biology, University of Torino, Italy

*BSA-seq, genomics, globe artichoke, male sterility*

Globe artichoke (*Cynara cardunculus* var. *scolymus* L.;  $2n=2x=34$ ) is a highly heterozygous species traditionally vegetatively propagated. It is mainly cultivated in the Mediterranean region with Italy leading world producer, but, over time, it has spread to the Americas and China as well. By exploiting male sterile (MS) genotypes, a number of F1 hybrids have been successfully introduced in cultivation. However, the genetic bases of this trait have been to date poorly explored.

An F2 population of 250 offsprings, obtained by selfing an F1 hybrid derived from a cross between a MS globe artichoke and a male fertile cultivated cardoon, was genotyped and phenotyped for pollen viability. Two main classes were identified in the F2 progeny, including individuals producing viable or not viable pollen, in a ratio 195:55. This segregation fits with a monogenic Mendelian segregation model (3:1), and suggests that one gene (ms) might affect male sterility when in homozygous recessive state.

We applied BSA sequencing (BSA-seq) with the goal of identifying genomic regions/genes affecting male sterility, and associated markers exploitable in globe artichoke breeding programs. We performed Illumina sequencing of two bulks (60X each), including 15 MS and 15 male fertile plants, as well as the sequencing of the two parents (40X each). Clean reads were aligned to the globe artichoke reference genome and SNPs calling was performed.

Data were analysed with QTLseqr (<https://github.com/bmansfeld/QTLseqr>), an R package for NGS data deriving from BSA.

The analysis identified four QTL regions likely involved in male sterility on chrs. 4, 5, 12, and 14. The highest values were recorded in chr. 14, with a peak observed near the 14.5 Mb for Gprime (24.8) and near 13.2Mb for DeltaSNP (0.31). By analysing the sequence and polymorphisms present in that region in both the parental lines, up to 60 candidate genes were spotted. A SnpEff analysis (<http://pcingola.github.io/SnpEff>) on those genes, followed by PROVEAN (Protein Variation Effect Analyzer; <http://provean.jcvi.org>) survey, permitted to identify one of them, a cytochrome P450 (*CYP703A2*), carrying a deleterious substitution (R/G), only present in the male sterile parent. Mutations in *CYP703A2* have been reported to be strictly associated with pollen sterility in both *Arabidopsis* and rice. A 3D model of *CYP703A2* from globe artichoke was generated by homology modelling and predicted that the R/G substitution dramatically affects protein folding.

Finally, a single dCAP marker was developed based on *CYP703A2* genomic region in the two parents. A set of primers was designed amplifying a small region (~100bp) around the SNP of interest and the restriction enzyme *HpaI* was selected after in silico PCR reaction for its ability of allowing a clear discrimination between MS and fertile genotypes in our population.