Proceedings of the LXIV SIGA Annual Congress
*Online, 14/16 September, 2021*
ISBN: **978-88-944843-2-8**

**Poster Communication Abstract – 7.11**

# GENOME-WIDE IDENTIFICATION OF MICROSATELLITE MARKERS IN ANEMONE CORONARIA L. AND DEVELOPMENT OF A USER-FRIENDLY WEB RESOURCE

MARTINA M.*, ACQUADRO A.*, BARCHI L.*, GULINO D.*, BRUSCO F.**, RABAGLIO M.**, PORTIS F.***, PORTIS E.*, LANTERI S.*

*) DISAFA, Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco (TO), Italy.
**) Biancheri Creazioni, 18033 Camporosso (IM), Italy
***) Yebokey, Turin, Italy.

*microsatellites, ornamentals, breeding, Anemone*

*Anemone coronaria* L. (family *Ranunculaceae*, *2n=2x*=16, estimated genome size = 9,08Gb) is a perennial allogamous plant grown as a cut flower or in gardens. Commercial cultivars have been developed through mass selection within segregating progenies obtained by crossing highly heterozygous genotypes, and maintain a certain level of genetic variability.

Due to its high genome size, the development of molecular markers in *A. coronaria* is, up to now, very limited. However, it is crucial for cultivar fingerprinting, detection of genetic diversity, assessment of population structure as well as for mapping genes of interest and selection of desirable traits. Microsatellites or simple sequence repeat (SSR) are one of the most informative markers, due to their codominant inheritance and a multi-allelic nature, versatility and easy of detection.

Here we obtained a first *A. coronaria* draft genome by sequencing a haploid plant obtained through *in vitro* androgenesis of the commercial line 'MISTRAL® Magenta', by short-reads Illumina sequencing (PE150). The assembly, comprising of 2x109contigs, and covering about 65% of the estimated genome size, was generated through the MEGAHIT pipeline. The genome sequence was used to catalogue the genome's content of microsatellite loci. Perfect, imperfect and compound SSRs were in-silico mined using the SciRoKo SSR-search module, which identified 401.822 perfect SSR motifs, equivalent to an overall density across the genome of 65,8 SSRs/Mbp. The perfect SSRs were categorized for the numbers of repeats present and their overall length; furthermore, 188.987 imperfect SSRs were also identified.

The gene annotation of the assembled draft genome was obtained through a *de novo* approach based on RepeatMasker v4.1.0 and the MAKER pipeline. The obtained gene annotation made it possible to investigate the distribution of genic SSRs. Overall, 3223 perfect SSRs (0.80% of the total) and 1261 imperfect SSRs (0.67%) were associated with as many genes, representing the 0.23% of the gene space. These microsatellites were estimated to cover a total of 134Kb, value which translate to a density across the gene space of 57.48 and 22.52 SSRs/Mbp for perfect and imperfect motifs, respectively.

Finally, we developed a user-friendly "*Anemone coronaria* Microsatellite DataBase" (AnCorDB), which

allows the identification of SSR markers in terms of type of repeat (perfect vs. imperfect), sequence, repeat number, and motif type. The incorporation of the Primer3 script inside the database website makes it possible to design couples of primers for downstream application of the identified markers. Eight genotypes belonging to eight commercial lines were used to validate 150 of the identified SSRs.

Our work represents a novel step in A. coronaria breeding and provides new molecular tools for population structure studies, genetic diversity analysis, cultivar fingerprinting, genetic mapping and marker assisted selection.