Proceedings of the Joint Congress SIBV-SIGA
Pisa, Italy – 19/22 September, 2017
ISBN **978-88-904570-7-4**

Poster Communication Abstract – 6.29

EXPLORING LONG NONCODING RNAs IN WHEAT WILD RELATIVES

PIERI A., PÈ M.E., BERTOLINI E.

Institute of Life Sciences, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127 Pisa (Italy)

lncRNAs, noncoding genome, Triticum urartu, Aegilops tauschii, wheat

The understanding of the complex bread wheat genome (*T. aestivum* AABBDD) can be facilitated by the study of its wild relatives. In this work we chose the donors of the A and D genomes, *T. urartu* and *Ae. tauschii*, to investigate the presence and function of long noncoding RNAs (lncRNAs). To date, the major efforts have been mostly focused on the characterization of the coding fraction of the genome, neglecting noncoding RNAs loci that have been shown to be important regulators of gene expressions in a broad range of species, including plants, where they can play key roles also in response to biotic and abiotic stresses.

To perform a comprehensive annotation of the lncRNAs in *T. urartu* and *Ae. tauschii* a total of 68 public RNA-seq libraries were retrieved, comprising different organs: root, shoot, leaf, seedling and spike were available for both species whereas seed, pistil, sheath, stem and stamen only for *Ae. tauschii*. A two steps pipeline was applied: i) *Transcriptome reconstruction*: short-reads are aligned two times to the respective reference genome to benefit from the splice junction information retrieved from the first mapping iteration, successively the entire transcriptome is independently *de novo* reconstructed using a genome-guided approach. ii) *Chasing lncRNAs:* according to the main features of lncRNA a stringent pipeline is applied to filter out *bona-fide* lncRNAs from the whole set of transcripts *de-novo* reconstructed.

We predicted 14,515 *T. urartu* and 20,908 *Ae. tauschii bona-fide* lncRNAs, showing features similar to those of other plant and animal counterparts, such as a reduced transcript length, with a median around 460 nucleotides, and number of exons, most of them are monoexonic. Thousands lncRNAs were significantly modulated in different organs and exhibited organ specific expression with a predominant accumulation in the spikes, sustaining the hypothesis of their critical biological role in reproductive organs. Interestingly, most of the organ-specific lncRNAs were found to be associated with transposable elements (TEs), suggesting a possible role of TEs in lncRNAs origin and differentiation. Although the majority of *T. urartu* and *Ae. tauschii* lncRNAs appears to be species-specific, we found 38 lncRNAs perfectly conserved between *T. urartu* and *Ae. tauschii*. In addition, we identified thousands of lncRNAs promoter sequences and lncRNAs transcripts conserved in the two progenitors and also in the A and D genome of bread wheat.

Our work provides the first comprehensive atlas of wheat wild relatives lncRNAs and shed new light on their characteristics and conservation across different species.