

## Heat shock response elements in the promoters of heat stress activated LTR retrotransposons in *Macrostomum lignano*

V. Vavilova<sup>1\*</sup>, E. Berezikov<sup>1,2</sup>, K. Ustyantsev<sup>1</sup>

<sup>1</sup> *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

<sup>2</sup> *European Research Institute for the Biology of Ageing, Groningen, The Netherlands*

\* e-mail: valeriya-vavilova@bionet.nsc.ru

**Key words:** *Macrostomum lignano*, heat shock, LTR retrotransposons

**Motivation and Aim:** Retrotransposon activity is generally repressed by host cell due to their impact on the genome stability, although some retrotransposons can become active again under the influence of various external stressors. However, the molecular mechanisms underlying the heat shock (HS) induced activation of retrotransposons in animals are poorly understood. *Macrostomum lignano* is a free-living flatworm that is increasingly used as model organism to study stem cell biology and regeneration. We have observed that LTR retrotransposons (LTR-RTs)-encoded transcripts were among the first top ten genes with increased expression after HS in this animal, opening up possibilities for the detailed investigation of the underlying molecular mechanism.

**Methods and Algorithms:** The HS elevated transcripts of the *M. lignano* LTR-RTs were mapped back to the genome assembly (<http://gb.macgenome.org/>). The corresponding full-length DNA sequences spanning both long terminal repeats were extracted and their internal structure was identified as in [1]. The core promoter sequences were mined inside long terminal repeats ([http://www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html)) and HS elements (HSEs) [2] were identified using regular expressions. HSEs were also screened in promoters of *M. lignano* HS protein genes and those of the known LTR-RTs available at GyDB (<http://gydb.org>). The phylogenetic analysis was performed using IQ-tree webserver (<http://www.iqtree.org/>).

**Results:** *M. lignano* HS-activated LTR-RTs belong to two phylogenetically distinct clusters of *Ty3/gypsy* group: *Mag* and *CsRNI*, with copy numbers of 49 and 10 in the *Mlig\_3\_7* genome assembly. Elements of both clusters have multiple canonical HSEs (nGAAnnTTCnnGAAn) upstream of the core promoter regions, which are also present in the promoters of *M. lignano* HS protein genes. Interestingly, previously the canonical HSEs were found only for the *Ty1/copia* LTR-RTs of *Arabidopsis thaliana*, which is also activated upon HS [3]. HSEs were absent in long terminal repeats of other known LTR-RTs available at GyDB, for which there is no data for the HS activation.

**Conclusion:** *M. lignano* LTR-RTs acquired multiple canonical HSEs upstream of their core promoters similar to the ONSEN element in *A. thaliana*, suggesting similar activation mechanisms and, thereby, possible convergent evolution of these elements. This is the first example of HSEs in animal LTR-RTs.

**Acknowledgements:** The project was supported by the RFBR (Grant No. 18-34-00288).

### References

1. Ustyantsev K. et al. (2017) Convergence of retrotransposons in oomycetes and plants. *Mobile DNA*. 8(4).
2. Mittal D. et al. (2011) Binding affinities and interactions among different heat shock element types and heat shock factors in rice (*Oryza sativa* L.). *FEBS Journal*. 278(17):3076-3085.
3. Cavrak V.V. (2014) How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genetics*. 10(1):e1004115.