

CRISPR-like elements and MGEs in *Arabidopsis thaliana* nuclear genome

Gorbenko I.V.^{1*}, Petrushin I.S.^{1,2}, Konstantinov Yu.M.^{1,2}

¹ Siberian Institute of Plant Physiology and Biochemistry, SB RAS, Irkutsk, Russia

² Irkutsk State University, Irkutsk, Russia

* gorbenko@sifibr.irk.ru

Key words: *Arabidopsis*, CRISPR, transposon

Previously we reported on the discovery of CRISPR-like elements in *Arabidopsis thaliana* genome. In the ecotype Col-0 we found 1372 arrays containing direct repeats and possessing 1 to 8 spacers with a length ranging from 61 to 743 bp. Most arrays possess 1–2 spacers and thus resemble mini-CRISPR from archaeal viruses. The length distribution of the detected arrays is discrete, which speaks in favor of the canonical CRISPR structure of the detected elements. The fact that fairly aligning arrays are often found across different ecotypes suggests their ancient origin. According to present data, CRISPR can distribute across genomes residing in special transposons – casposons. We suggest that transposons in general could be the mechanism of CRISPR emerging in plant genomes. More than 30 % of CRISPR-like elements found in *A. thaliana* Col-0 were located within mobile genetic elements (MGEs), mainly belonged to the DNA/MuDR, LTR/Gypsy, and RC/Helitron superfamilies. It should be noted that CRISPR were found inside VANDAL-type MGEs, that according to recent data can remain active due to their ability to demethylate DNA. CRISPR-containing MGEs are basically located in the pericentromeric regions, probably because the TEs there retain their structural integrity more. Analyzing the diversity of CRISPR residing in TEs, we found that arrays with a high degree of homology are usually located within MGEs that are of the same superfamily, which speaks in favor of their possible emergence as a component of MGEs. It is possible that CRISPR arrays spread throughout the plant genome as a part of transposons, but then 70 % of MGE-hosts were destroyed. Given this origin, plant CRISPR-like systems may be trans-acting and utilize the molecular machinery of transposons as Cas proteins. Considering the abundance of CRISPR-like elements, we suggest that they could play a role as a component of defense mechanism against pathogenic DNA and foreign MGEs.