

Discordant evolution of YUCCA family proteins demonstrated along sequence

I. Turnaev^{1*}, K. Gunbin¹, V. Suslov¹, D. Afonnikov^{1,2}

¹*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

²*Novosibirsk State University, Novosibirsk, Russia*

* e-mail: turn@bionet.nsc.ru

Key words: auxin biosynthesis pathway, horizontal gene transfer, discordant evolution

Motivation and Aim: The origin of the tryptophan dependent pathway of auxin biosynthesis in terrestrial plants is a subject for discussion. The answer to this question related to the evolution of two pathway's enzymes, TAA and YUCCA. Here we perform phylogenetic analysis of the evolution of the YUCCA protein family in plants.

Methods and Algorithms: Proteins of YUCCA family and their closest homologues were extracted from NCBI database using protein BLAST with a threshold e-value $< 1e-5$. Homologs of protein YUC2 (AT4G13260) of *Arabidopsis thaliana* were the BLAST queries; only full-size proteins were taking into account. Promals used for alignment reconstruction. The phylogenetic tree was built using different methods: Maximum Likelihood, Bayesian, Approximated Maximum Likelihood (AML). To assess the degree of anomaly of evolution (the degree of qualitative differences between the topologies of phylogenetic trees of various alignment fragments and known species tree) along the alignment we reconstructed phylogenetic trees for the alignment fragments within 60 aa sliding windows. We group obtained phylogenetic trees by their sliding window position from 1 to 50, 25 to 75, 50–100 etc. For each of these groups we build a phylogenetic network with the SplitsTree program. We compared network for these sequence segments with the species tree topology and identified the deviation from the species tree.

Results: As result, we identified two regions with discordant evolution incompatible with the species tree topology: aa 95–185 and 210–280. Phylogeny for those regions of the YUCCA protein is significantly different from the known species phylogeny. We propose that the cause of this discordancy may be the horizontal transfer events of fragments of the coding part of genes from non-plant taxa (perhaps symbionts of ferns or gymnosperms plants) to the ancestral proteins of the subfamily YUC10-11 terrestrial plants in the above-named regions of protein.

Acknowledgements: The work was supported by ICG SB RAS budget project (No. 0324-2018-0017).