

## Molecular evolution analysis of genetic network components related to plant trichome development

A.V. Doroshkov<sup>1,2\*</sup>, D.K. Konstantinov<sup>1</sup>, K.V. Gunbin<sup>1,2</sup>, D.A. Afonnikov<sup>1,2</sup>

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

<sup>2</sup> *Novosibirsk State University, Novosibirsk, Russia*

\* *e-mail: AD@bionet.nsc.ru*

**Key words:** molecular evolution, trichome, MBW complex, gene regulatory network

*Motivation and Aim:* Trichomes are involved in many significant functions such as the transpiration, thermoregulation and protection from insect attacks. On the other hand specialized cell formation is an fruitful model system for analyzing the molecular mechanisms of plant cell differentiation, including cell fate choices, cell cycle control, and cell morphogenesis. However, the evolution of specialized epidermal cell formation genetic network remains unclear.

*Methods and Algorithms:* We reconstructed the network of interactions between known leaf pubescence genes using *A. thaliana* as a model organism using Cytoscape and Pathway Studio software. For each network node (gene) the extraction of sets of homological protein sequences was carried out using the DELTA-BLAST. Then multiple sequence alignment was conducted with MAFFT algorithm. The PhyML and IQTree maximum likelihood algorithms was used to reconstruct the phylogenetic trees of protein families and bootstrap resampling technique was used for testing the topology. In order to find phylogenetic tree branches at which accelerated evolution occurred, we reconstruct protein ancestral sequence state spectra in each inner tree nodes using Bayesian approach implemented in PhyloBayes. After that for each protein family we reconstructed relative rate matrix of amino acid substitutions by modeestimator software and compared it with the inferred frequencies of amino acid substitutions on each tree branch that estimated by the ancestral sequence state spectra comparison on these branches. Additionally, for each residue of each ancestral proteins we predicted its structural properties using RaptorX Property pipeline. As a result of these procedures, for each internal branch of protein family phylogenetic tree we derived several statistical values describing acceleration of sequence and structure evolution.

*Results:* The main stages of gene network evolution has been traced down to the evolutionary time of appearance of its components. It was found that the LCA of dicotyledonus plants in a comparison with *A. thaliana* have a reduced set of R2R3-MYB factors (7 versus 13), and at the LCA of all flowering plants also have a reduced set of R3-MYB inhibitors (1 versus 7). The reduction of the network component sets resulted in at least 4 protein families: *ttg1*, *egl3*, *glabra*, *cpc*. The above described analysis of protein evolution in each inner branches of that protein families demonstrated that in contrast to *glabra* protein family accelerated protein evolution events associated with *cpc*, *egl* and *ttg1* concentrated on orthologs formation events related with divergence between flower plants and gymnosperms.

*Conclusion:* Thus, we found that the main adaptive evolution events of the gene network involved in plant trichome development related with divergence between flower plants and gymnosperms.