On evolutionary analysis of gene networks by the Orthoscape software

Z.S. Mustafin[1]*, D.A. Afonnikov[1, 2], Yu.G. Matushkin[1, 2], S.A. Lashin[1, 2]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
* e-mail: mustafinzs@bionet.nsc.ru

Motivation and Aim: There is a huge amount of networks of different types in biology. There is also a number of software intended to visualization and analysis of such networks. One of the key software frameworks to work with biological networks is Cytoscape (http://cytoscape.org/). Recently we had presented Orthoscape, a Cytoscape application for evolutionary analysis and visualization of gene networks and gene sets [1]. Such an analysis, which includes phylostratigraphic analysis and Darwinian selection analysis, may shed light on origin and evolution of complex traits determined by gene networks

Methods and Algorithms: We used KEGG (http://www.kegg.jp/) as the source of biological data. We used CyKEGGParser plugin [2] to get networks from the KEGG Pathway and original KEGG API to get lists or homologs with identity and SW Score values, protein domains, nucleotide and amino acid sequences and taxonomic information. K_a/K_s was calculated by PAML [3] using pairwise sequence comparisons for the taxa under analysis. Sequences were aligned using the Needleman–Wunsch algorithm realization adapted from (http://zhanglab.ccmb.med.umich.edu/NW-align/). The gene divergence value based on every gene-ortholog pair. K_a/K_s result, is allows us to discriminate the diversifying and the stabilizing selection.

Results: 81 human diseases networks from the KEGG Pathway were analyzed. The most evolutionary young genes influenced into immune system group of diseases. They have also shown the highest evolutionary rate (the highest K_a/K_s value) in comparison with the Primates. The oldest genes have been found within the substance dependencies group. There are small amount of genes at all and the most part of them influenced in basic biological processes like breath regulation.

Availability: http://apps.cytoscape.org/apps/orthoscape; https://github.com/ZakharM/Orthoscape

References
1. Mustafin Z.S., Lashin S.A., Matushkin Y.G., Gunbin K.V., Afonnikov D.A. (2017) Orthoscape:a cytoscape application for grouping and visualization KEGG based gene networks by taxonomy and homology principles. BMC Bioinformatics. 18(S1):1-9.
2. Nersisyan L., Samsonyan R., Arakelyan A. (2014) CyKEGGParser: tailoring KEGG pathways to fit into systems biology analysis workflows. F1000Research. 145:1-11.
3. Yang Z. (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24(8):1586-1591.