

Global co-expression network analysis unveils important aspects of evolution and transcriptional regulation in soybean (*Glycine max*)

Fabrício de Almeida Silva, Fabricio Brum Machado, Kanhu Charan Moharana,
Rajesh Kumar Gazara, Thiago Venancio

Universidade Estadual do Norte Fluminense Darcy Ribeiro

Abstract

Soybean (*Glycine max* (L.) Merr.) is one of the most important crops worldwide, representing a significant fraction of Brazilian GNP. Gene co-expression networks (GCN) have been largely used to elucidate regulatory complexity and evolution of genes and their functions. Here, we have reconstructed a GCN using 1298 publicly available samples from 12 distinct tissues. Sequencing reads were mapped against the soybean genome (Wm82.a2.v1) and relative transcript abundance estimated in Transcripts per million mapped reads (TPM). The network was reconstructed and visualized with the R packages WGCNA and igraph, respectively. The top 10% most highly connected genes with the highest module membership were considered intramodular hubs. We explored the network properties and found critically important modules that are up-regulated in specific tissues. Enrichment analyses of these modules revealed biological processes and pathways that are essential to some particular tissues and may have elementary contributions to the plant development. We also identified transcription factors (TFs) among intramodular hubs, which may be important regulators, shaping the transcriptional landscape in particular tissues. The top hubs for each module were identified and we found that they tend to encode proteins with critical roles, such as succinate dehydrogenase, peroxidases, and RNA polymerase subunits. Further, we analyzed the distribution of soybean paralogous genes across the network modules to better comprehend the fate of duplicate genes in polyploid organisms. Most of the duplicate gene pairs were present in different modules, supporting their subfunctionalization in different tissues and providing insights on the evolutionary importance of polyploidization in soybean genome complexity and evolution.

Funding: CAPES, CNPq, FAPERJ