

A Ciência e os caminhos do desenvolvimento

**Análise da complexidade regulatória e evolutiva de genes de soja
(*Glycine max*) a partir de redes de co-expressão gênica**

*Fabício Almeida-Silva, Fabricio Brum Machado, Rajesh Kumar Gazara, Kanhu Charan
Moharana, Thiago Motta Venancio*

A soja (*Glycine max* (L.) Merr.) é a oleaginosa mais produzida no mundo e desempenha um papel fundamental para a agroindústria brasileira, compondo uma fração significativa do PIB nacional. Redes de co-expressão gênica têm sido amplamente utilizadas em estudos transcriptômicos a fim de elucidar a complexidade regulatória e evolução de genes e suas funções. No presente trabalho, nós construímos redes de co-expressão gênica utilizando diversos tecidos de soja. Nós utilizamos transcritos de 1298 amostras de 12 tecidos distintos, disponíveis publicamente no banco de dados SRA (*Sequence Read Archive*) do NCBI. Os dados brutos de sequenciamento (i.e. “reads”) foram mapeados contra o genoma de referência (a2.v1) usando o programa STAR. Os níveis de transcrição foram estimados e normalizados (em TPM) usando o programa StringTie e o pacote zFPKM, em ambiente R. As redes foram construídas e exploradas utilizando WGCNA e *igraph*, também em ambiente R. Exploramos as propriedades gerais das redes construídas e identificamos módulos regulatórios de grande importância, em virtude de suas associações com tecidos específicos e estágios do desenvolvimento da planta. Esses módulos regulatórios foram submetidos a análises de enriquecimento de processos biológicos e vias metabólicas. As análises revelaram genes e vias que podem ter contribuições fundamentais para o desenvolvimento da planta. Também mapeamos a presença de fatores de transcrição nestes módulos, visando a identificação de reguladores importantes da dinâmica regulatória global. Os resultados sugerem um número de fatores de transcrição que podem regular as variações fisiológicas importantes que ocorrem em cada tecido da soja. Os fatores de transcrição apontados podem ser utilizados em futuras validações experimentais. Por fim, utilizamos a rede de co-expressão para explorar as funções de genes duplicados e seus destinos durante a evolução da soja, em particular aqueles derivados de eventos de poliploidização.

Palavras-chave: Regulação da transcrição, Redes de co-expressão, Bioinformática.

Instituição de fomento: CAPES, CNPq, FAPERJ.