POSTER - SYSTEMS BIOLOGY AND MODELING

THE “SWEET THIRSTY SPIDER”: BUILDING A DROUGHT-RELATED SUGARCANE GENE NETWORK

Pedro Cristovão Carvalho (carvalhopc@usp.br)

Danyel Contiliani (danyelfc1@gmail.com)

Renato Gustavo Hoffmann Bombardelli (rghbombardelli@usp.br)

Silvana A Creste Dias De Souza (silvanacreste@gmail.com)

Claudia Barros Monteiro Vitorello (cbmontei@usp.br)

Antonio Figueira (figueira@cena.usp.br)

The development of new crop cultivars represents an important step for food security under a scenario of climate change and exponential population growth. In this context, plant breeding requires a refined selection of candidate genes to ensure that subsequent years of development will produce significant results using biotechnological approaches. However, for sugarcane (Saccharum spp.), the selection of candidate genes is hindered by the lack of genome annotation, generally established by orthology, and might not reflect the actual information regarding the gene functions and related pathways. System biology may help to overcome the lack of annotation, allowing the integration of diverse omics data to develop more robust biological networks with solid information about candidate genes and their effects in related pathways. Here, we describe the development of a drought-related co-expression network based on 16 publicly available transcriptome data from sugarcane upon contrasting water regimes.

All data was standardized against the same assembled transcriptome, providing patterns of differential expression of sugarcane transcripts under water deficit, which were then analyzed for the presence of eigengenes and expression hubs by weighted correlation network analysis (WGCNA), establishing the edges between sugarcane nodes. The initial network comprises 36,265 nodes, representing the group of transcripts previously described within five clusters and 1,108,458 edges, including possible interactions among those transcripts. All the nodes from the initial network were further annotated by BLAST, EggNOG, and Interpro, aggregating all the available information for function, related pathways, orthologs, and domains for each node. Further, results from BLAST, when cross-referenced with the STRING database, revealed protein-protein interactions, which can reinforce the original edges from the WGCNA analysis while providing novel information regarding the direct interaction between nodes, together with the metabolic interaction information from KEGG. Even though the network was developed based on specific transcriptomes, it is being future-proofed by enrichment with all the corresponding gene and transcript identifications from other available genomes and transcriptomes, allowing analysis with any current or future dataset.

Palavras-chave: climate change; saccharum; sugarcane; systems biology; water deficit.