SUGARCANE SIGNALING AND REGULATORY NETWORKS

Glaucia Mendes Souza
Chemistry Institute / University of São Paulo (IQ/USP)
FAPESP Process 2008/52146-0 | Term: Nov 2008 to Oct 2015 | Thematic Project

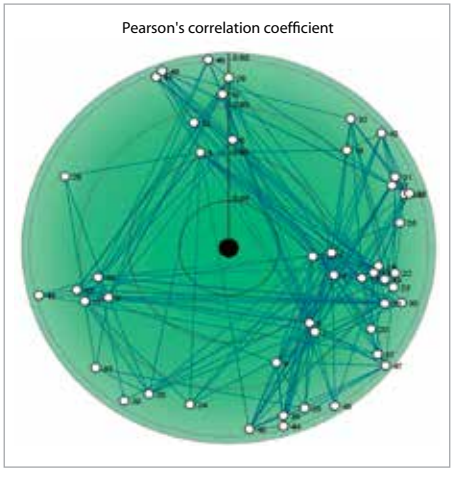

Figure 1. Gene co-expression network with 4CL maize gene. The central dot represents the 4CL gene and the concentric circles show the Person correlation coefficient limits until 0.8 cutoff. Genes with mutual correlation above 0.8 are connected by blue lines

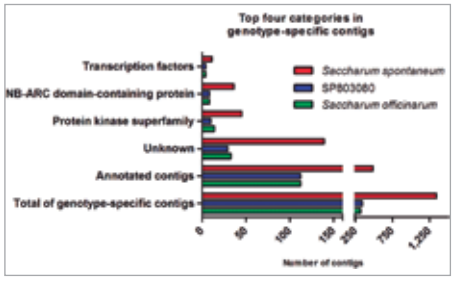

Figure 2. Sugarcane ORFeome top four categories of genotype-specific contigs based on Phytozome annotations. Only contigs from leaf samples in each genotype were considered

We aim to study signaling and regulatory networks in sugarcane and to develop tools for a systems biology in this grass. As a starting point we intend to characterize three agronomical traits of interest: drought, brix and lignin content. We will study gene categories with a well-known regulatory role (Transcription Factors, Protein Kinases and Phosphatases), continue studies on the Transcriptome, produce transgenics, develop a database and computational tools to integrate the several levels of information and we will initiate the whole genome sequencing of a Brazilian sugarcane cultivar. In parallel, we intend to implement the ChIP-HTS technology in sugarcane, to identify TF targets and gene promoters. The results will have multiple direct consequences on breeding programs that frequently select for CREs and TF changes in search for genotypes better adapted to the environment and with increased agronomical performance. PKs activate signaling cascades in response to environmental stimuli and our studies point to a predominant role of PKs in the regulation of sucrose content and drought responses. To identify new genes associated to brix, drought and lignin content we will characterize the transcriptome of genotypes and cultivars that contrast for these traits using oligonucleotide arrays. Genes of interest will be functionally evaluated by generating transgenics altered for their expression. To integrate the immense amount of public data and that generated by this project a robust computational infrastructure and database will be developed. The SUCEST-FUN database will integrate the SUCEST sequences, promoters, CREs, expression data, agronomical, physiological and biochemical characterization of sugarcane cultivars. We will also participate in the development of the GRASSIUS database to establish sugarcane, rice, maize and sorghum regulatory networks.

Aim 1 – Identification of genes associated to sucrose, lignin content and drought tolerance by transcriptome analyses using microarray technique

Aim 2 – Evaluation of gene function through the generation and analysis of transgenic sugarcane plants

Aim 3 – Sugarcane gene promoter sequence and regulatory motif identification by sequencing R570 BAC clones and SP80-3280 genome using 454 platform.

Aim 4 – Development of the SUCEST-FUN Database and computational tools to support all activities of the BIOEN Program.

SUMMARY OF RESULTS TO DATE AND PERSPECTIVES

In this Thematic Project we have developed a custom oligoarray platform to study the expression of sugarcane genes. The oligoarray contains 21.092 probes, of which 14.554 detect expression from sense transcripts (SS) and 4.137 detect the expression of antisense transcripts (AS). Transcripts with altered expression were identified in samples contrasting for sugar content (2.761 SS, 126 AS), lignin content (420 SS, 17 AS) and drought tolerance (3.379 SS, 243 AS). We also verified that sugarcane has a higher proportion of transcripts with circadian rhythms than other species.

The sugarcane ORFeome was sequenced from two ancestral and one hybrid genotype using a full-length enriched library. We found 38,195 sugarcane-specific transcripts and observed that less than 1.6% of all transcripts were ancestor-specific. Transgenic sugarcane plants were produced overexpressing and/or repressing three candidate genes selected from the transcriptome results. The expression modification of downstream elements was observed with the use of the custom oligoarray. Sugarcane putative promoters from candidate genes were identified by high-throughput sequencing using the Roche 454 platform and by ChIPSeq experiments using an antibody against the RNA polymerase II.

We started to study the sugarcane development and maturation in the field in two different seasons using a System Biology approach. Physiological, morphological and biochemical measurements were collected in four different developmental stages. Samples from four different tissues in four different developmental stages were collected for transcriptome, proteome and metabolome analysis which are being conducted.

During the project, the SUCEST-FUN Regulatory Network Database (CaneRegNet), which contains data and tools of interest for sugarcane functional genomicists and molecular breeders, was developed. The CaneRegNet database assembles different sugarcane databases such as the Sugarcane Expressed Sequence Tags Genome Project, the SUCAST and the SUCAMET Catalogues, which include expression data and the GRASSIUS database. The tools available in the database are: Cane Genome, Cane Transcriptome, Cane Gene Expression, Cane Physiology, Functional Annotation, Cane Transgenic Plant, Generic Tools and Publication. We have started the sugarcane genome sequencing in this project and we have already sequenced sugarcane BACs and shotgun reads using Roche 454 sequencing platform, pair-end and mate-pair libraries using Illumina HiSeq and long reads using Moleculo Illumina.

MAIN PUBLICATIONS

Nishiyama MY, Ferreira SS, Tang PZ, Becker S, Pörtner-Taliana A, Souza GM. 2014 Full-length enriched cDNA libraries and ORFeome analysis of sugarcane hybrid and ancestor genotypes. *Plos One*. **9**: e107351.

Ferreira SS, Nishiyama M, Paterson AH, Souza GM. 2013 Biofuel and energy crops: high-yield Saccharinae take center stage in the post-genomics era. *GenomeBiology.com (London. Print)*. **14**: 210.

Hotta CT, Nishiyama MY, Souza GM. 2013 Circadian rhythms of sense and antisense transcription in sugarcane, a highly polyploid crop. *Plos One*. **8**: e71847.

Lembke CG, Nishiyama MY, Sato PM, SOUZA GM. 2012. Identification of sense and antisense transcripts regulated by drought in sugarcane. *Plant Molecular Biology*. **79(4-5)**: 461-477.

Souza GM, Berges H, Bocs S, Casu R, D Hont A, Ferreira JE, Henry R, Ming R, Potier B, Sluys M, Vincentz M, Paterson AH. 2011. The sugarcane genome challenge: strategies for sequencing a highly complex genome. *Tropical Plant Biology*. 1-12.

Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann H, Garcia AAF, Souza GM. 2011. Sugarcane improvement: how far can we go?. *Current Opinion in Biotechnology*, v. x.

Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM. 2010. Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. *Plant Biotechnology Journal*. **8**: 263-276.

Yilmaz A, Nishiyama-Junior MY, Garcia-Fuentes B, Souza GM, Janies D, Gray J, Grotewold E. 2008. GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiology (Bethesda)*. 10.110, p. 10.1104/pp.108.

Glaucia Mendes Souza

Instituto de Química
Universidade de São Paulo (USP)
Departamento de Bioquímica
Av. Prof. Lineu Prestes 748, sala 954
CEP 05508-900 – São Paulo, SP – Brasil

+55-11-3091-8511
glmsouza@iq.usp.br