Madras Agric. J., **2018**, *105 (7-9)(2)*: 396-400; doi: 10.29321/MAJ.2018.000169

Differential Transcript Expression Profiling of Sugarcane (co 99004) for Elevated Temperature Stress

R. Gomathi[1*], S. Kohila[1] and Lakshmi Kasirajan[2]
[1]Division of Crop Production, ICAR - Sugarcane Breeding Institute, Coimbatore.
[2]Division of Crop Improvement, ICAR - Sugarcane Breeding Institute, Coimbatore.

Sugarcane is one of the world's major C_4 crops and mainly grown in tropical and sub-tropical regions for sugar and bioenergy. Climatological extremes including very high temperatures are predicted to have a catastrophic loss of crop productivity. Based on physiological and biochemical assays (data Unpublished), Co 99004 was selected as heat stress tolerant sugarcane variety. To investigate the differential expression of transcripts between control and heat stressed meristematic tissues of Co 99004, an RNA-Seq experiment was conducted. Meristematic tissues were collected from both control and treated samples at formative phase of the crop (150 days). Totally 2,308 transcripts were differentially expressed in the meristematic tissues of sugarcane in response to heat stress. About 458 transcripts and 305 transcripts were up and down regulated respectively in response to heat stress in sugarcane. There were common differential expressed transcripts detected between up and down regulated transcripts. Among the 458 up regulated transcripts, 328 transcripts were found to be commonly upregulated in both control and heat stress. Among the down regulated transcripts, 292 transcripts were down regulated in both treated and control. There were 130 transcripts exclusively upregulated during heat stress, while no transcripts were exclusively down regulated. The distribution of transcript expression level varied in different ranges of fold change. Transcripts coding for dehydration responsive element binding 1A like, grain softness protein, Malate synthase1, lipoxygenase and avenin were exclusively expressed during heat stress in sugarcane. These findings reveal novel target genes for subsequent research on the regulation of elevated temperature stress tolerance.

Key words: Sugarcane, Climate change, Elevated temperature, Transcripts, RNA-Seq.

Weather and climate associated events (*i.e.,* CO_2, temperature, precipitation, and other extreme weather) are the key factors for sugarcane production worldwide, particularly in many developing countries (Zhao and Li, 2015). High temperature is a most important factor that affects the rate of plant growth (Hatfield and Prueger, 2015). It has detrimental effects on plants by affecting growth, development and metabolism leading to major loss of yield potential reported by Kaushal *et al*. (2016). Hatfield *et al*. (2011) said that heat stress reduced yield from 2.5% to 10% in many agronomic species. Kohila and Gomathi (2018) have reported that, heat stress (45°C) reduced chlorophyll content and stability, F_v/F_m ratio, leaf gas exchange, relative water content (RWC), nitrate reductase (NR) and sucrose-metabolizing enzymes (SPS, SS, AI, NI) in all the sugarcane genotype and species clones. In contrast, it increased primary and secondary metabolites, Superoxide dismutase and Peroxidase activities, lipid peroxidation (LP), membrane injury index (MII) and soluble sugar in sugarcane. Temperature stress reduces the germination and early seedling growth in sugarcane. Porter (2005) said that it accelerates the senescence and reduce crop yield and leads to plant death (Sharma *et al*., 2005). Wahid (2007) described that, synthesis of primary and secondary metabolites had close relationship to heat resistance in sugarcane. Qin *et al*. (2008) stated that, identification of novel genes and study of their expression patterns in response to temperature stress may offer a molecular basis for improving heat tolerance in crops. De Setta *et al*. (2014) and Cardoso-Silva *et al*. (2014) also reported that more than five thousand sugarcane genes remain undiscovered. New sequencing efforts of sugarcane transcriptomes need to increase the primary set of sequence information, and next generation sequencing (NGS) will allow more efficient trancriptomic studies Manners *et al*. (2011). Illumina technology and digital gene-expression (RNA-Seq) is a novel approach for expression study and has reported many genes involved in different stress response and plant growth (Hao *et al*., 2011; Li *et al*., 2016).

To our knowledge, heat stress in sugarcane has received much less attentiveness than the other abiotic stresses. Sugarcane varietal evolution in the future requires yield stability even under harsh climates. Understanding the metabolic and molecular signal transcription processes and the interaction to high temperatures is necessary. Therefore, this research will provide a valuable resource for understanding the molecular mechanism of heat stress tolerance in sugarcane.

*Corresponding author's email: gomathi_sbi@yahoo.co.in

Material and Methods

Genetic material used for the study

Seven sugarcane genotypes were selected based on their performance from the previous studies conducted at Plant Physiology Section, Division of Crop Production, ICAR - Sugarcane Breeding Institute, Coimbatore. It consists of five commercial canes (Co 06022, Co 0315, Co 8021, Co 86032 and Co 99004) and two wild spp. (*Spontaneum* Spp. *viz.,*Taiwan -96 and SES -150) genotypes.

Experimental design and heat stress treatment

All the seven genotypes were planted in 6 inches pots filled with the mixture of red soil, sand and farmyard manure which was previously homogenized with enough water to bring the soil for field capacity. The experiment was laid out in a Completely Randomized Block Design (CRD) with three replications. Normal agronomic practices were followed for all of these plants. In order to impose heat stress, a set of three pots/genotypes was exposed to a temperature of 4°C to 5°C above its optimum temperature range, an increase which corresponds tightly to climate change model predictions. Control plants were grown under optimal conditions at 37 / 28 ± 2°C day /night with a 12-hour photoperiod. Heat stressed plants were grown at 45 / 32 ± 2°C during the day/ night with a 12-hour photoperiod and for a total of 15 days, with 60-70% relative humidity, and light intensity 395–410 µmol m^{-2} s^{-1}.

RNA extraction and mRNA library construction

Based on physiological and biochemical assays, Co 99004 was selected as heat stress tolerant sugarcane variety (Kohila and Gomathi, 2018). Meristematic tissues were collected from three pots (one pot/replicate) from both control and treated samples at formative phase of the crop (150 days). Those samples were considered as one biological replicates, immediately frozen in liquid nitrogen and then stored at -80°C for total RNA isolation. Transcriptome sequencing on the Illumina HiSeq 2500 platform works had done in Sci Genome Labs Private Limited, Cochin, Kerala.

cDNA library construction and Illumina sequencing

The total RNA was isolated from the meristematic tissues of sugarcane genotype Co 99004 (both control and treated (high temperature) using Trizol reagent (Ambion, Life technologies, North America.) according to the manufacturer's instructions. Construction of cDNA library and transcriptome sequencing on the Illumina HiSeq 2500 platform. Equal amount of total RNA (4µg) was used for library construction and transcriptome sequencing. Qbit was used for measuring the quantity of the library before sequencing. After constructing libraries, they were sequenced on Illumina HiSeq 2500 instrument.

Data pre-processing

After sequencing, raw reads were processed to remove adapter sequences, bases of low quality and trimming low quality reads. High quality RNA-seq reads were processed by Cutadapt software (version 1.8) (http://github.com/marcelm/cutadapt). Bowtie2 (version 2.2.6) (http://Bowtie2-bio.sourceforge.net/index.shtml) software removed contaminations in RNA-seq very fast and sensitively. Finally, clean reads were produced from raw reads.

De novo transcriptome assembly

Clean reads were used for normalization and were assembled into expressed transcripts as full-length sequences, find a solution to other similar sequences and duplicated genes by using Trinity (version 2.3.2) software (http://trinityrnaseq.sourceforge.net). In-house script tool was used for a sequence of this RNA pool again, filter the contigs with FPKM>=1 for every fragment were sequenced and capture the unique contigs from multiple samples.

Differential gene expression analysis and identification of plant metabolic networks

The transcripts with different level fold changes of FPKM in response to high temperature stress were compared to FPKM value of the control sample. More than 2 fold changes and adjusted p value $p<0.05$ of were identified as differentially expressed genes by DeSeq (version1.16.0) database (http://www.bioconductor.org/ packages/ release/ bioc/ html/ DeSeq.html).The plant metabolic pathway analysis was performed based on PMN database (http://www.plantcyc.org/).

Functional annotation of transcripts

The assembled transcripts were annotated using in-house pipeline CANoPI (Contig annotator Pipeline). The transcripts were blasted against –NR NCBI database and then functionally annotated through gene ontology annotation, pathway annotation, gene and protein annotation, organism annotation. The assembled mRNA were queried to non-redundant protein database from NCBI and checked for similarity (under the parameters of E-value <=1e-5, Similarity >= 40%) to existing known protein from various plant genomes. Organisms' similarity distribution and top hits for organism was found using BLASTX. By using CANoPI, significant BLASTX hit and UniProt hit provided the protein status, sub cellular location of protein and EC numbers for sequences. The transcripts were compared with GO database to classify unigenes based on biological processes, molecular function and cellular components. A cluster of orthologous genes: eggNOG database was used to annotate any sequences present and it's corresponding orthologous groups.

Validation of transcripts using RT-qPCR

Total RNA was extracted from shoot apical meristematic tissues and it was treated with DNAse free kit to remove genomic DNA contaminations. This was performed using one step Real-Time PCR System (Applied Biosystem).The cycle threshold (Ct) values of the candidate genes were used to evaluate

their expression stability by using NormFinder applet for MS Excel. Using a reference gene, the fold change of target gene expression levels comparing the meristematic tissues of both control and treated samples were determined using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Results and Discussion

To identify the molecular mechanism involved in response to high temperature stress of sugarcane variety Co 99004, two gene expression profiles were constructed with the help of Illumina Sequencing-Paired End libraries from control and treated meristematic tissues RNA Integrity Number (RIN) value of this sample was 9.7. It is greater than or equal to 8 is considered good quality for that. The HiSeq 2500 platform was used to generate 5,52,40,382 and 3,16,72,900 of raw reads.

De novo transcriptome assembly

The raw reads were pre processed after the quality checking. The low-quality sequences (phred score < 20 bp) of the paired end reads and vague raw reads were trimmed using silva database (http://www.arb-silva.de/) for removing the rRNA from raw reads. Transcriptome assembly of 1,42,859 contigs with a mean GC of 49.64% was generated.

Table 1. Overview of the sequencing and assemble of *Saccharum officinarum* L

Items	Reads		Transcripts
	Control	Treated	
Total number of raw reads	5,52,40,382	3,16,72,900	-
Total number raw reads removed	1,59,59,754 (28.9%)	1,82,35,606 (57.6%)	-
Total number of clean reads	3,92,80,628	1,34,37,294	-
Average Q30	90.04	90.01	-
Number of clean bases (Mb)	3,755.6	1,279.59	-
Maximum sequence length (bp)	-		21,435
Average length (bp)	-		780
Total number of assembled transcripts	-		1,42,859
Assembled Transcripts after cd-hit-est	-		1,18,017

Approximately 59.6 % of the assembled transcripts were ~200 to >5000 bp in length. The longest sequence in the assembled transcriptome was 21435 bp, which had a similarity of (99 %) with a protein of Maturase K (Intron maturase) from *Saccharum officinarum*.

Gene expression analysis

The assembled transcripts (length ≥ 200bp) was aligned by Bowtie2 program. The fragments per kb per million fragments (FPKM) was used to estimate the gene expression level. Transcripts with FPKM values ≥ 1.0 was considered the for further analysis.

Table 2. Read alignment and expression summary of *Saccharum officinarum*

Items	Reads	
	Control	Treated
Number of reads	1,96,40,314	67,18,647
Number of reads aligned	1,77,97,516	60,10,842
Alignment Percentage	90.62	89.47
Number of transcripts with FPKM ≥ 1.0	60,733	72,763

The number of transcripts with FPKM (≥ 1.0) was 60,733 and 72,763 for control, treated samples respectively, it is referred to as the high temperature response transcripts (HTRTs) in the present study in the later sections.

Differentially expressed transcripts

Changes in the gene expression levels in response to heat stress was compared to control. Differentially expressed transcripts were identified with help of DeSeq program. The numbers of transcripts were considered for DeSeq is 118,017. The FPKM value of each gene from the treatment vs. control was calculated according to the following criteria: *p*-value < 0.05 and log2fold change (FC) > 2. Totally 2308 transcripts were differentially expressed in the meristematic tissues of sugarcane exposed to heat stressed, which is shown in the volcano plot (Up regulated in red; Down regulated as blue) Fig.1. Among the heat responsive DE transcripts, 1137 transcripts were up regulated and 1171 transcripts were down regulated in heat stress tolerant sugarcane cultivar Co 99004 in comparison to control. Based on the values of log two fold change of the up and down regulated transcripts transcripts were distributed in different ranges (Fig. 2).

DE transcripts categorized into functional groups

Differentially expressed transcripts of heat stress tolerant sugarcane variety were annotated using

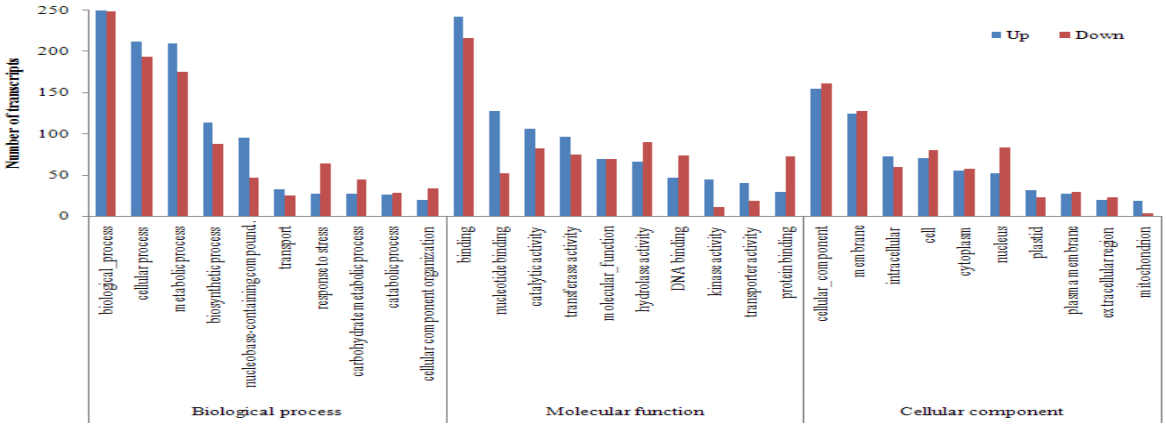

Fig. 1. GO term distribution of differentially expressed Up and Down regulated transcripts

UniPort database to draw the GO tree to classify the up and down regulated transcripts into putative functional groups. On the basis of Gene Ontology (GO) analysis, we analyzed 1137 up regulated transcripts were categorized into 82 GO functional subcategories, which were summarized into three main categories: “Biological process” “metabolic process” and “cellular process” were most abundantly represented with 1219 (42.62 %), 955 (33.39 %) and 686 (23.90 %) transcripts, respectively. GO term distribution of the differentially expressed 1171 down regulated transcripts were functional categorized into 76 subcategories, which were summarized into three main categories: “Biological process” “metabolic process” and “cellular process” were most abundantly represented with 1264 (44.19 %), 841 (29.40 %) and 719 (25.13 %) transcripts, respectively.

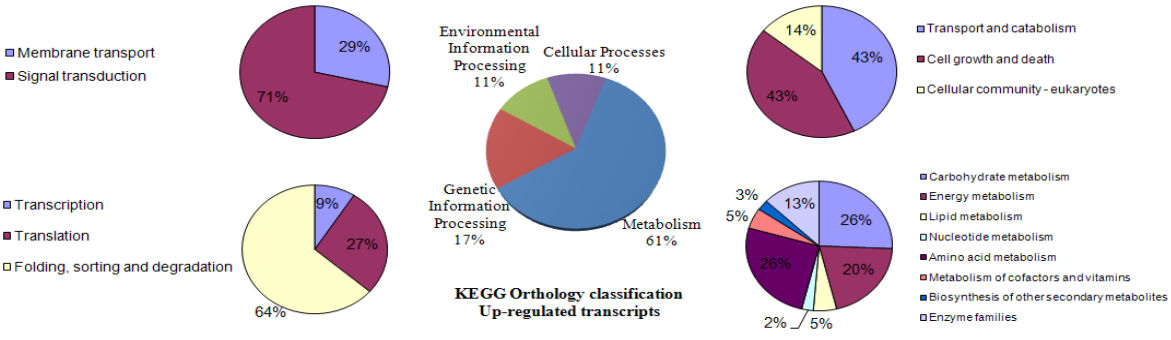

Fig. 2. KEGG classification differentially expressed up regulated transcripts

KEGG orthology and pathway classification of the DE transcripts

The orientations of the differentially expressed up-regulated transcripts in metabolic pathways were analyzed by querying the KEGG database. In total, 1137 up-regulated transcripts were assigned to five specific KEGG database pathways, including metabolism (30.19 % transcripts), cellular processes (9 % transcripts), environmental information processing (11 % transcripts), genetic information processing (17 % transcripts) and hypothetical/unknown. The largest category contained 69.8 % of up regulated transcripts sequences with hypothetical/ unknown/ uncharacterized function (69.8 % transcripts) (Fig 3).

In total, 1171 Down-regulated transcripts were assigned into five specific pathways; the largest category contained 54.99 % of Down-regulated transcripts with hypothetical/unknown/ uncharacterized function. The pathways with the most representation by the Down-regulated transcripts were “metabolic” pathways (40 % transcripts) “genetic information processing” (30 % transcripts) “environmental-information processing” (15 % transcripts) and “Cellular Processes” (13 % transcripts)

Venn diagram analysis

As shown in Venn diagram, there was common differential expressed transcripts detected between up and down regulated transcripts in both treated and control (Fig.4). Venn analysis indicated that six transcripts were regulated in both treated and control. Totally, 458 up regulated transcripts and 305 down regulated transcripts were detected both treated and control. Of the 458 shared DE up regulated transcripts, 328 up regulated transcripts had similar response patterns to both treated and control and 130 up regulated transcripts were detected only in the treated. Among the down regulated transcripts, 292 common transcripts were down regulated in both treated and control and 13 transcripts were degraded only in control.

In the current study, we employed next generation sequencing method RNA-Seq to analyze the transcripts of sugarcane and characterized candidate genes related to the heat stress in sugarcane during formative phase. A total of 1,42,859 transcripts were assembled the longest sequence in the assembled transcripts was 21,435 bp. We used the assorted protein databases, *viz.,* NCBI plant non-redundant (nr), UniProt, GO, KEGG pathway; KOG, COG and eggNOG using the BLASTX algorithm for annotated the assembled transcripts by mapping them. Based on the assembled transcripts, we identified several transcripts differentially expressed in treated compare to control, which annotated to involve in the response

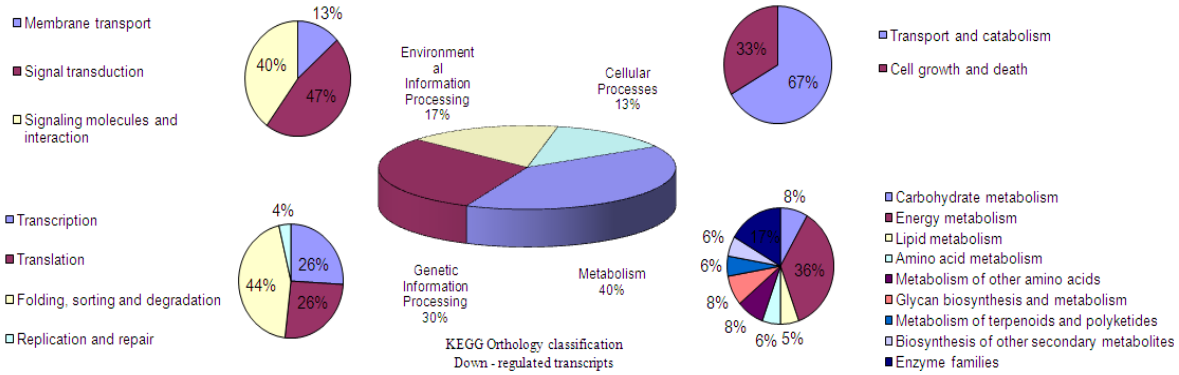

Fig. 3. KEGG classification differentially expressed down regulated transcripts

oxidative stress, stress, heat shock protein and response to heat. There were 130 transcripts exclusively up regulated during heat stress, while no transcripts were exclusively down regulated. The distribution of transcript expression level varied in different ranges of fold change. Transcripts coding for dehydration responsive element binding 1A like, grain softness protein, Malate synthase1, lipoxygenase and avenin were exclusively expressed during heat stress in sugarcane. These findings reveal novel target genes for subsequent research on the regulation of elevated temperature stress tolerance.

Acknowledgment

The authors thank the DST-SERB, New Delhi for financial support and to the Director, ICAR-Sugarcane Breeding Institute, Coimbatore for providing necessary facility and support to carry out the research work successfully.

References

Cardoso-Silva, C.B., Costa, E.A., Mancini, M.C., Balsalobre, T.W., Canesin, L.E. and Pinto, L.R. 2014. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. *PLoS One,* **9(2):** 88462.

De Setta, N., Monteiro-Vitorello, C.B., Metcalfe, C.J., Cruz, G.M.Q., Del Bem, L.E. and Vicentini, R. 2014. Building the sugarcane genome for biotechnology and identifying evolutionary trends. *BMC Genomics*, **15:** 540.

Hao, Q.N., Shou, X.A., Sha, A.H., Wang, C., Zhou, R., and Chen, S.L. 2011. Identification of genes associated with nitrogen use efficiency by genome wide transcriptional analysis of two soybean genotypes. *BMC Genomics,* **15:** 525.

Hatfield, J.L., Boote, K.J., Kimball, B.A., Ziska, L.H., Izaurralde, R.C., Ort, D., Thomson, A., and Wolfe, D. 2011. Climate impacts on agriculture: Implications for crop production. *Agronomy Journal,* **103:** 351-370.

Hatfield, J.L. and Prueger, J.H. 2015. Temperature extremes: Effect on plant growth and development. *Weather Clim Extremes*, **10:** 4-10.

Kaushal, N., Bhandari, K., Siddique, K.H.M. and Nayyar, H. 2016. Food crops face rising temperatures: An overview of responses, adaptive mechanisms, and approaches to improve heat tolerance. **2:** 1-42.

Kohila, S. and Gomathi, R. 2018. Adaptive physiological and biochemical response of sugarcane genotypes to high-temperature stress. *Ind. J Plant Physiol.,* doi. org/10.1007/s40502-018-0363-y.

Li, M., Liang, Z., Zeng, Y., Jing, Y., Wu, K., Liang, J., He, S., Wang, G., Mo, Z., Tan, F., Li, S. and Wang, L. 2016. De novo analysis of transcriptome reveals genes associated with leaf abscission in sugarcane (*Saccharum officinarum* L.). *BMC Genomics,* **17:** 195.

Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2[-Delta Delta C (T)] Method. *Methods*, **25:** 402-408.

Manners, J. and Casu, R. 2011. Transcriptome Analysis and Functional Genomics of Sugarcane. *Trop. Plant Biol.,* **4:** 9-21.

Porter, J.R. 2005. Rising temperatures are likely to reduce crop yields. *Nature,* **436:** 174.

Qin, D., Wu, H., Peng, H., Yao, Y., Ni, Z., Li, Z., Zhou, C. and Sunm, Q. 2008. Heat stress responsive transcriptome analysis in heat susceptible and tolerant wheat by using Wheat Genome Array. *BMC Genomics,* **9:** 432.

Sharma, P., Sharma, N. and Deswal, R. 2005. The molecular biology of the low-temperature response in plants. *BioEssays*, **27:** 1048-1059.

Wahid, A., Farooq, M., Rasheed, R., Gelani, S. and Rasul, E. 2007. Sugarcane under thermal stress: some biotechnological considerations. In: Kumar Ashwani (ed) Plant genetic transformation and molecular markers. Pointer Publishers Jaipur, pp 109-123.

Zhao, D. and Li, Y.R. 2015. Climate Change and Sugarcane Production: Potential Impact and Mitigation Strategies. *International Journal of Agronomy,* 1-10.