



Transcriptome analysis of muskmelon (*Cucumis melo*) for moisture stress tolerance

P SUPRIYA^{1*} and K V BHAT²

ICAR-National Academy of Agricultural Research Management, Hyderabad, Telangana 500 030, India

Received: 24 June 2020; Accepted: 31 October 2022

Keywords: Differential expression, Gene ontology, Pathway analysis, RNA-Seq

The major goal of crop improvement is to increase yield and quality with minimum inputs. Climate change and the occurrence of extreme stress events such as drought, salinity, high temperature are major challenges for agriculture. Abiotic constraints, in particular moisture stress, have a major impact on crop productivity. The interaction between moisture stress and changes in plant gene expression have been intensively studied in many species including *Arabidopsis* (Kreps *et al.* 2002, Oono *et al.* 2003, Liu *et al.* 2008, Matsui *et al.* 2008), rice (Zhao *et al.* 2007, Degenkolbe *et al.* 2009, Lenka *et al.* 2011), maize (Hayano-Kanashiro *et al.* 2009, Luo *et al.* 2010, Zheng *et al.* 2010), wheat (Aprile *et al.* 2009) and sorghum (Buchanan *et al.* 2005). The reported studies have provided evidence of the signaling pathways related to moisture stress (Chaves *et al.* 2003). ABA triggers a signaling cascade that alters the transcriptome and up-regulates genes that encode several proteins and enzymes associated in response to drought (Schroeder *et al.* 2001). At the transcriptome level, the variations in response to stress should be modulated both immediately and with relevance to the particular stress experienced. In the present study, we performed transcriptome analysis of muskmelon (*Cucumis melo* L.) to identify genes responsible for moisture stress tolerance.

An analysis of transcript level changes can be used to detect new signaling proteins and enzymatic reactions that are essential for ensuring plant stress tolerance. An experiment was conducted at ICAR- National Bureau of Plant Genetic Resources, New Delhi during 2019–20 and the data was received in the form of assembled transcripts (Supplementary file). In the present study, the plants were grown in pots with regular irrigation and moisture stress condition, in which plants were exposed to moisture stress 30 days after sowing (DAS) without receiving irrigation under the stress condition. Fresh leaves were collected from

plants 60 days after treatment and total RNA was extracted from leaves using TRIZOL RNA isolation protocol.

Functional annotation of the assembled transcripts: The transcripts were used for Blastx (E-value $\leq 1e-05$) analysis against the publicly accessible protein databases such as NCBI non redundant protein (Nr) database and UniProt. Blast2Go platform was used for assigning Gene Ontology (GO) terms into biological processes, molecular functions and cellular components to those sequences which have a significant match with the sequences available in the public domain. BLAST2GO program was used to get GO annotations of unique assembled transcripts.

RNA-Seq mapping: A total of 12859 and 13448 transcripts of muskmelon control and stress samples respectively were used in the current study. Both the sets of transcripts were mapped to the 26,307 contigs by taking combined assembly as reference for mapping. The parameters used in the RNA-Seq mapping pipeline were minimum read length fraction to 0.9, minimum similarity to 0.95 and Reads Per Kilobase of Million mapped reads (RPKM) value as expression value.

Differential gene expression analysis: Differential gene expression analysis was carried out using CLC genomics work bench. The aim of study was to identify and characterize the genes involved in moisture stress tolerance in control and stress samples. The P-value limit of 0.05 was used to filter statistically significant results.

Metabolic pathway analysis: Kyoto Encyclopedia of Genes and Genomes (KEGG) mapping was used to identify the metabolic pathways. Such annotations of the names and description of genes or proteins, gene ontology terms, conserved domains and possible metabolic pathways would provide a valuable tool for investigating functions, processes and pathways associated in stress tolerance of muskmelon.

Next-generation sequencing (NGS) based RNA sequencing (RNAseq) enabled the simultaneous acquisition of gene discovery sequences as well as the detection of transcripts involved in complex biological processes.

Functional annotation and classification: In total, 90% of all transcripts had a significant match with one of

¹ICAR-National Academy of Agricultural Research Management, Hyderabad, Telangana; ²ICAR-National Bureau of Plant Genetic Resources, New Delhi. *Corresponding author email: puramsupriya@gmail.com

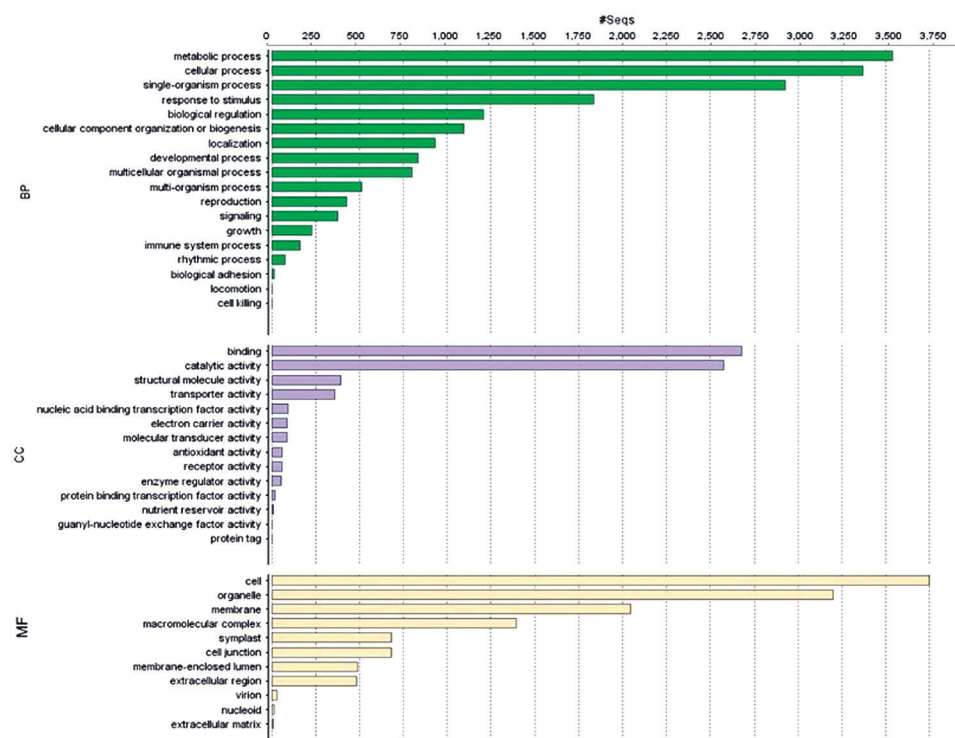


Fig 1 Distribution of GO terms of transcripts.

the databases such as NCBI non-redundant protein (Nr) and UniProt. The majority of the annotated sequences were matching with *Cucumis sativus* which belongs to Cucurbitaceae family, implying that the sequences of the muskmelon transcripts obtained in the present study were assembled and annotated properly. BLAST2GO was used to assign GO terms for the assembled transcripts and a total of 6263 and 6430 transcripts were assigned at least one GO term from control and stress samples, respectively. GO terms were majorly grouped into 25 functional categories including biological processes (BP), cellular component (CC) and molecular function (MF) in control and stress samples (Fig 1).

Metabolic process (30.2%) was the most prevalent in the biological process group followed by cellular process (28.2%) and response to stimulus (11.6%). In the category molecular function, the most represented GO terms were of binding activity (46%) and catalytic activity (39.8%). The GO terms of cellular component category showed significant representation of cell (38%), organelle membrane (36%) and macro molecular complex (6%).

Expression analysis: RNA-Seq mapping of control and stress transcripts was done by mapping high quality reads on assembled transcripts and RPKM values were calculated. Mortazavi *et al.* (2008) demonstrated that RPKM value makes reliable comparisons of gene expression levels. It was found that 1015 and 1087 non-redundant transcripts were found in control and stress samples, respectively. After extracting unique gene list, these genes were searched in annotated transcripts and were further categorized (Table 1).

Differential gene expression: The aim of a differential

expression analysis is to illustrate genes that have significantly increased in abundance through experimental conditions. After normalization of RPKM values, genes with expression changes no less than 2 during stress conditions and FDRs less than 0.05 were identified as differentially expressed genes. Normalization allows reliable comparisons of expression levels between as well as within samples (Loven *et al.* 2012). The methods of normalization vary between and within sample comparisons. Analysis of differential expression showed that 122 stress-responsive genes were widely expressed in both the samples and 72 genes in the stress sample were highly

upregulated. Primarily, DREB genes, MYB transcription factors, kinases, heat shock proteins, zinc finger and AP2/ERF domain containing transcription factors were among the upregulated genes during stress condition. Transcription factors and regulatory proteins are essential gene groups which can control downstream expression of drought-responsive genes (Singh and Laxmi 2015). The genes induced during water stress conditions work not only in protecting cells by generating essential metabolic proteins but also in controlling genes for signal transduction in response to stress.

Metabolic pathway analysis: The annotated sequences were mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) using KOBAS server to recreate the metabolic pathways involved during stress conditions. Among these, 4 genes are involved in pathways such as plant-hormone signal transduction, protein processing in endoplasmic reticulum, Inositol phosphate metabolism, and Sulfur metabolism (Table 2).

Table 1 Distribution of stress responsive genes

Description	No. of transcript
Ethylene responsive genes	6
Calcium-binding/dependent	39
Heat shock protein	19
Dehydration responsive genes	28
Other metabolite responsive genes	5
Stress responsive genes	18
Other responsive gene	69
Total	184

Table 2 Genes involved in different pathways

Transcript id	Gene name	Pathway
trans_s_3189	AP2/ERF domain-containing transcription factor (ERF33)	Plant-hormone signal transduction
trans_s_12451	Abscisic acid responsive elements binding protein 2 (ABF2-1)	Plant-hormone signal transduction
trans_s_6519	Heat shock protein (Hsp45.9)	Protein processing in endoplasmic reticulum
trans_s_12142	SAL1-like protein	Inositol phosphate metabolism and sulfur metabolism

In the present study, transcriptome analysis of muskmelon was carried out to investigate the modifications occurring in gene expression in response to moisture stress. Comparison of gene expression levels between control and stress samples revealed that 72 genes were highly upregulated in stress sample. The primary enriched biological process associated with these genes was hormone-mediated signaling pathway. These genes will not only promote understanding of the genetic basis of the response to moisture stress but will also expedite genetic development in muskmelon. Further, the functional characterization of the identified genes and pathways described here may lead to new targets for plant stress tolerance enhancement, which will be particularly relevant in the face of climate change and the growing prevalence of different forms of abiotic stress.

SUMMARY

Transcriptome is the whole set of RNA molecules transcribed in a cell at a particular time under particular environmental conditions. Assessing the transcriptome and estimating the degree of expression of all genes under different conditions is a crucial step in understanding the dynamic processes that take place during development. In the present study, genes that play a major role in moisture stress were identified using high-throughput transcriptome sequencing analysis. High quality assembled transcripts of *Cucumis melo* var. agrestis control and stress samples were compared using BlastX with the protein databases available in the public domain. Gene Ontology (GO) analysis revealed that a total of 6263 and 6430 transcripts were assigned at least one GO term from control and stress samples, respectively. Changes of expression level in these transcripts for moisture stress tolerance were detected using RNA-seq mapping. Among stress responsive genes, a total of 122 genes were commonly expressed in both control and stress samples and 72 genes were highly upregulated in stress sample when compared to control sample.

REFERENCES

- Aprile A, Mastrangelo A, De Leonardis A, Galiba G, Roncaglia E, Ferrari F, De Bellis L, Turchi L, Giuliano G and Cattivelli L. 2009. Transcriptional profiling in response to terminal drought stress reveals differential responses along the wheat genome. *BMC Genomics* **10**: 279.
- Buchanan C D, Lim S, Salzman R A, Kagiampakis I, Morishige D T, Weers B D and Mullet J E. 2005. Sorghum bicolor's transcriptome response to dehydration, high salinity and ABA. *Plant Molecular Biology* **58**(5): 699–720.
- Chaves M M, Maroco J P and Pereira J S. 2003. Understanding plant responses to drought from genes to the whole plant. *Functional Plant Biology* **30**(3): 239–64.
- Degenkolbe T, Do P T, Zuther E, Repsilber D, Walther D, Hinch D K and Kohl K I. 2009. Expression profiling of rice cultivars differing in their tolerance to long term drought stress. *Plant Molecular Biology* **69**(1–2): 133–53.
- Hayano-Kanashiro C, Calderon-Vazquez C, Ibarra-Laclette E, Herrera-Estrella L and Simpson J. 2009. Analysis of gene expression and physiological responses in three Mexican maize landraces under drought stress and recovery irrigation. *PLoS One* **4**(10): 7531.
- Kreps J A, Wu Y, Chang H S, Zhu T, Wang X and Harper J F. 2002. Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant physiology* **130**(4): 2129–41.
- Lenka S K, Katiyar A, Chinnusamy V and Bansal K C. 2011. Comparative analysis of drought responsive transcriptome in Indica rice genotypes with contrasting drought tolerance. *Plant Biotechnology Journal* **9**: 315–27.
- Liu H H, Tian X, Li Y J, Wu C A and Zheng C C. 2008. Microarray-based analysis of stress regulated microRNAs in *Arabidopsis thaliana*. *RNA-a publication of the RNA society* **14**(5): 836–43.
- Loven J, Orlando D A, Sigova A A, Lin C Y, Rahl P B, Burge C B, Levens D L, Lee T I and Young R A. 2012. Revisiting global gene expression analysis. *Cell* **151**(3): 476–82.
- Luo M, Liu J, Lee R D, Scully B T and Guo B. 2010. Monitoring the expression of maize genes in developing kernels under drought stress using oligomicroarray. *Journal of Integrative Plant Biology* **52**(12): 1059–74.
- Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo T A, Okamoto M, Nambara E, Nakajima M and Kawashima M. 2008. Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant and Cell physiology* **49**(8): 1135–49.
- Mortazavi A, Williams B A, McCue K, Schaeffer L and Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**: 621–28.
- Oono Y, Seki M, Nanjo T, Narusaka M, Fujita M, Satoh R, Satou M, Sakurai T, Ishida J and Akiyama K. 2003. Monitoring expression profiles of Arabidopsis gene expression during rehydration process after dehydration using ca. 7000 full-length cDNA microarray. *Plant Journal* **34**(6): 868–87.
- Schroeder J I, Allen G J, Hugouvieux V, Kwak J M and Waner D. 2001. Guard cell signal transduction. *Annual Review of Plant Physiology and Plant Molecular Biology* **52**: 627–58.
- Singh D and Laxmi A. 2015. Transcriptional regulation of drought response: a tortuous network of transcriptional factors. *Frontiers in plant science* **6**: 895.
- Zhao B, Liang R, Ge L, Li W, Xiao H, Lin H, Ruan K and Jin Y. 2007. Identification of drought-induced microRNAs in rice. *Biochemical and Biophysical Research Communications* **354**(2): 585–90.
- Zheng J, Fu J, Gou M, Huai J, Liu Y, Jian M, Huang Q, Guo X, Dong Z and Wang H. 2010. Genome-wide transcriptome analysis of two maize inbred lines under drought stress. *Plant Molecular Biology* **72**(4–5): 407–21.