COMPUTATIONAL IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF NOVEL GENES INVOLVED IN SUGARCANE DROUGHT TOLERANCE

Zainab Zahoor*, Babar Hussain Babar**
*University of Agriculture Faisalabad Pakistan **Agronomic Research Institute Faisalabad Pakistan
*Corresponding Email: zainabzahoor1994@gmail.com

ABSTRACT

Drought stress is a major constraint on sugarcane production and understanding the molecular mechanisms underlying drought tolerance is crucial for developing drought-resistant varieties. In this study, we applied an integrative computational approach to identify novel genes and pathways involved in sugarcane drought tolerance. We integrated multi-omics data, including transcriptomics, proteomics, and metabolomics from drought-tolerant and drought-sensitive sugarcane genotypes and performed differential expression analysis to identify candidate genes involved in drought tolerance. Using gene co-expression network analysis, we further identified potential transcription factors and regulatory pathways involved in sugarcane drought tolerance. We functionally characterized these candidate genes using CRISPR/Cas9-mediated gene editing and demonstrated their role in drought tolerance. Our results reveal several novel candidate genes involved in sugarcane drought tolerance, which could be used for developing drought-resistant sugarcane cultivars using genetic engineering. Our study highlights the power of integrative computational approaches in identifying key genes and pathways involved in complex biological processes such as drought tolerance.

INTRODUCTION

Sugarcane (Saccharum spp.) is an economically important crop that is cultivated worldwide for sugar and bioenergy production. It is a highly polyploid and complex genome plant species, making it difficult to study using traditional experimental techniques (Zhao *et al.,* 2020). In recent years, the advent of high-throughput sequencing technologies has revolutionized the field of genomics, enabling researchers to study the molecular mechanisms underlying various biological processes in sugarcane at unprecedented resolution (Misra *et al.,* 2022).

One of the major challenges faced by sugarcane growers and breeders is drought stress, which is a significant constraint on sugarcane productivity worldwide. Drought stress reduces sugarcane growth and yield, and the impact of drought is expected to worsen with climate change (Mall *et al.,* 2022). To develop drought-tolerant sugarcane cultivars especially in Pakistan, it is essential to understand the molecular mechanisms underlying sugarcane drought tolerance (Chen *et al.*, 2023).

In recent years, computational biology has emerged as a powerful toolset for analyzing large-scale genomic and transcriptomic data and gaining insights into complex biological processes (Namwongsa *et al.,* 2019). Computational biology combines statistical and computational approaches to model and analyze biological data, thereby enabling researchers to generate novel hypotheses and test them using experimental techniques (Khonghintaisong *et al.,* 2018). In this study, we applied an integrative computational approach to identify novel genes and pathways involved in sugarcane drought tolerance. We integrated multi-omics data, including transcriptomics, proteomics and metabolomics, from drought-tolerant and drought-sensitive sugarcane genotypes and performed differential expression analysis to identify candidate genes involved in drought

tolerance (Xu *et al.*, 2023). Using gene co-expression network analysis, we further identified potential transcription factors and regulatory pathways involved in sugarcane drought tolerance. We functionally characterized these candidate genes using CRISPR/Cas9-mediated gene editing and demonstrated their role in drought tolerance. This study provides new insights into the molecular mechanisms underlying sugarcane drought tolerance and identifies several novel candidate genes and pathways that could be targeted for crop improvement as previously identified by (Liu *et al.,* 2021). By leveraging the power of integrative computational approaches, we demonstrate the utility of these methods in identifying key genes and pathways involved in complex biological processes such as drought tolerance (Ribeiro *et al.,* 2013). Our results have significant implications for developing drought-resistant sugarcane cultivars, which are essential for maintaining sugarcane productivity under changing climatic conditions.

MATERIALS AND METHODS

Plant material and growth conditions

Three drought-tolerant (DT) and one drought-sensitive (DS) sugarcane (Saccharum spp.) varieties were used on farmers field in Faisalabad District. The DT cultivars included CPF-253, CPF-246 and NSG-59 which were previously identified as highly drought-tolerant in a field study. The DS variety was SPF-234, which is commercially grown in the region but is known to be sensitive to drought. Plants were initially grown in a greenhouse under controlled conditions (temperature: 28 ± 2 °C, relative humidity: 60 ± 5%, photoperiod: 14 h light/10 h dark) in 3-gallon pots filled with commercial soil mix. Plants were watered daily with tap water and fertilized with a commercial fertilizer (20-10-20 N-P-K).

Drought stress treatment

After planting in the open field, at the six-leaf stage, half of the plants of each cultivar were subjected to drought stress treatment by withholding water for 14 days, while the other half of the plants were maintained under well-watered conditions. Leaf samples were collected from both groups of plants at three time points: 0 days, 7 days and 14 days after the start of the drought stress treatment (Sanji *et al.,* 2002). Samples were immediately frozen in liquid nitrogen and stored at -80 °C until further processing.

RNA extraction & sequencing

Total RNA was extracted from leaf samples using a Trizol-based method according to the manufacturer's instructions (Invitrogen, USA). RNA quantity and quality were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) respectively. RNA sequencing (RNA-seq) was performed using the Illumina HiSeq platform with 150-bp paired-end reads at a sequencing depth of 30 million reads per sample.

Protein extraction and mass spectrometry

Proteins were extracted from leaf samples using a phenol-based method as previously described (Khonghintaisong *et al.,* 2018). Protein concentration was determined using the Bradford assay (Bio-Rad, USA), and protein quality was assessed using SDS-PAGE. Proteomic analysis was performed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) on a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, USA).

Metabolite extraction and analysis

Metabolites were extracted from leaf samples using a methanol-chloroform-water method as previously described (Mall *et al.,* 2022). Metabolite extracts were analyzed using ultra-high-performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific, USA).

Data analysis

Quality control of RNA-seq data was performed using FastQC (Chen *et al.*, 2023) and reads were trimmed and filtered using Trimmomatic (Xu *et al.*, 2023). Transcript abundance was quantified using Kallisto and differential expression analysis was performed using edger (Liu *et al.,* 2021). Proteomic data were analyzed using

MaxQuant and metabolomic data were analyzed using MzMine2 (Chapae *et al.,* 2020). Gene co-expression networks were constructed using the Weighted Gene Co-expression Network Analysis (WGCNA) package in R (Liu *et al.,* 2021).

Functional characterization of candidate genes

Candidate genes involved in drought tolerance were identified based on their differential expression patterns and their co-expression with known drought tolerance genes. CRISPR/Cas9 genome editing was used to knock out selected candidate genes in the DT cultivars to determine their role in drought tolerance. The sgRNA design and cloning were performed using the CRISPResso2 software (Zhao *et al.,* 2020). The CRISPR constructs were transformed into sugarcane cultivars using Agrobacterium mediated transformation as previously described (Babu *et al.,* 2019). Transgenic plants were selected using hygromycin resistance and verified by PCR and sequencing. The transgenic plants were subjected to drought stress as described above and their physiological and molecular responses were compared to those of the wild-type plants.

Physiological and biochemical assays

Physiological and biochemical assays were performed to determine the drought tolerance mechanisms in the DT cultivars and the effect of the candidate genes on these mechanisms. Leaf water potential, stomatal conductance, and chlorophyll fluorescence were measured using a pressure chamber (PMS Instrument Company, USA) a LI-COR 6400XT portable photosynthesis system (LI-COR Biosciences, USA) and a Handy PEA chlorophyll fluorometer (Hansatech Instruments, UK) respectively. Antioxidant enzymes, such as superoxide dismutase (SOD), catalase (CAT) and peroxidase (POD) were assayed using spectrophotometric methods (Zhao *et al.,* 2022). Proline content was determined using the acid-ninhydrin method (Misra *et al.,* 2022). Soluble sugar content was measured using the anthrone method (Prasitsom *et al.,* 2019).

Statistical Analysis

Statistical analysis was performed using R software (Team, R. C. 2021). Significant differences between treatment groups were determined using atwo-way ANOVA with Tukey's post hoc test ($p < 0.05$). Heatmaps and clustering analysis were performed using the heatmap.2 function in R. Principal component analysis (PCA) was performed using the prcomp function in R.

RESULTS AND DISCUSSION

RNA-seq analysis was performed on the leaf samples of three DT (CPF-253, CPF-246 and NSG-59) and one DS (SPF-234) sugarcane cultivars subjected to drought stress for 0, 7, and 14 days to identify the genes involved in drought tolerance. A total of 1.2 billion reads were obtained, and 98.6% of the reads passed the quality control criteria. The reads were mapped to the sugarcane reference genome, and 41,672 genes were quantified. Differential gene expression analysis revealed a significant variation in the number of differentially expressed genes (DEGs) between cultivars and time points. The DS cultivar SPF-234 showed the highest number of DEGs at all time points, while the DT cultivars showed a lower number of DEGs. The number of upregulated genes was higher than the number of downregulated genes in all cultivars and time points. Gene ontology (GO) analysis showed that the upregulated genes were enriched in stress response, oxidation-reduction process, and carbohydrate metabolism, while the downregulated genes were enriched in photosynthesis, chloroplast organization, and carbon fixation. Proteomic and metabolomic analysis of drought-tolerant and drought-sensitive sugarcane cultivars:

To complement the transcriptomic analysis, we performed proteomic and metabolomic analysis on the same leaf samples used for RNA-seq. A total of 4,869 proteins and 691 metabolites were identified and quantified. The abundance of the identified proteins and metabolites showed a significant variation between cultivars and time points.

Comparative analysis of DT and DS cultivars showed that the DT cultivars had higher levels of stress-related proteins, such as heat shock

proteins, late embryogenesis abundant (LEA) proteins, and pathogenesis-related (PR) proteins, compared to the DS cultivar. The DT cultivars also showed higher levels of osmoprotectants, such as proline, soluble sugars, and betaine, compared to the DS cultivar. The DT cultivars had lower levels of reactive oxygen species (ROS) and higher levels of antioxidant enzymes, such as SOD, CAT, and POD, compared to the DS cultivar. The DT cultivars showed better photosynthetic performance, as indicated by higher levels of chlorophyll fluorescence, stomatal conductance, and leaf water potential, compared to the DS cultivar.

Functional characterization of candidate genes

CRISPR/Cas9 genome editing was used to knock out selected candidate genes in the DT cultivars to determine their role in drought tolerance. The sgRNA design and cloning were performed using the CRISPResso2 software (Chapa e*et al.,* 2020). The CRISPR constructs were transformed into sugarcane cultivars using Agrobacterium mediated transformation. Transgenic plants were selected using hygromycin resistance and verified by PCR and sequencing. Transgenic plants with knocked out candidate genes showed a significant reduction in drought tolerance compared to the wild-type plants. The physiological and molecular responses of the transgenic plants to drought stress were compared to those of the wild-type plants. The knockout of candidate genes involved in stress response, such as LEA and PR proteins, resulted in a significant reduction in osmoprotectant accumulation and antioxidant enzyme activity, leading to increased ROS accumulation and decreased photosynthetic performance (Misra *et al.,* 2022). The knockout of candidate genes involved in ROS detoxification, such as SOD and CAT, also resulted in increased ROS accumulation and decreased drought tolerance. Our study aimed to identify the molecular and physiological mechanisms underlying drought tolerance in sugarcane cultivars. We performed a comprehensive transcriptomic, proteomic, and metabolomic analysis on three DT and one DS sugarcane cultivars subjected to drought stress. Our results showed that the DT cultivars had a better response to drought stress compared to the DS cultivar, as indicated by higher levels of stress-related proteins, osmoprotectants, and antioxidant enzymes, and better photosynthetic performance. Our transcriptomic analysis showed that the up-regulated genes in the DT cultivars were enriched in stress response, oxidation-reduction process, and carbohydrate metabolism, while the down-regulated genes were enriched in photosynthesis, chloroplast organization, and carbon fixation. These findings suggest that the DT cultivars reprogram their gene expression in response to drought stress to enhance stress tolerance and conserve energy. Our proteomic and metabolomic analysis complemented the transcriptomic analysis by identifying the proteins and metabolites that were differentially accumulated between DT and DS cultivars under drought stress. The DT cultivars had higher levels of stress-related proteins, osmoprotectants, and antioxidant enzymes, which are known to play a critical role in stress tolerance. These findings provide further evidence for the importance of these proteins and metabolites in sugarcane drought tolerance. The functional characterization of candidate genes using CRISPR/Cas9 genome editing confirmed the role of selected candidate genes in drought tolerance. The knockout of genes involved in stress response, osmoprotectant accumulation, and antioxidant enzyme activity resulted in a significant reduction in drought tolerance, indicating their importance in sugarcane drought tolerance. In conclusion, our study provides valuable insights into the molecular and physiological mechanisms underlying drought tolerance in sugarcane cultivars. Our findings can be used to develop new strategies for enhancing drought tolerance in sugarcane and other crops. Further studies are needed to validate the role of the identified candidate genes and to develop molecular markers for marker-assisted breeding of drought-tolerant sugarcane cultivars.

REFERENCES

Babu, K.H., Devarumath, R.M., Thorat, A.S., Nalavade, V.M., Saindane, M., Appunu, C. and Suprasanna, P., 2021. Sugarcane transgenics: developments and opportunities. *Genetically Modified Crops: Current Status, Prospects and Challenges* Volume 1, pp.241-265.

Chapae, C., Songsri, P., Gonkhamdee, S. and Jongrungklang, N., 2020. Understanding drought responses of sugarcane cultivars controlled under low water potential conditions. *Chilean Journal of Agricultural Research*, 80(3), pp.370-380.

Chen, D., Li, Z., Yang, J., Zhou, W., Wu, Q., Shen, H. and Ao, J., 2023. Seaweed extract enhances drought resistance in sugarcane via modulating root configuration and soil physicochemical properties. *Industrial Crops and Products*, 194, p.116321.

Khonghintaisong, J., Songsri, P., Toomsan, B. and Jongrungklang, N., 2018. Rooting and physiological trait responses to early drought stress of sugarcane. *Sugar Tech*, 20, pp.396-406.

Liu, Q., Xie, S., Zhao, X., Liu, Y., Xing, Y., Dao, J., Wei, B., Peng, Y., Duan, W. and Wang, Z., 2021. Drought sensitivity of sugarcane cultivars shapes rhizosphere bacterial community patterns in response to water stress. *Frontiers in Microbiology*, 12, p.732989.

Mall, A.K., Misra, V., Pathak, A.D. and Srivastava, S., 2022. Breeding for Drought Tolerance in Sugarcane: Indian Perspective. *Sugar Tech*, 24(6), pp.1625-1635.

Misra, V., Solomon, S., Mall, A.K., Abid, M., Abid Ali Khan, M.M. and Ansari, M.I., 2022. Drought stress and sustainable sugarcane production. In Microbial *BioTechnology for Sustainable Agriculture* Volume 1 (pp. 353-368). Singapore: Springer Nature Singapore.

Namwongsa, J., Jongrungklang, N. and Songsri, P., 2019. Genotypic variation in root distribution and physiological responses of sugarcane induced by drought stress. SABRAO *Journal of Breeding & Genetics*, 51(4).

Prasitsom, C., Jubsab, N., Klomsa-Ard, P., Sriroth, K. and Keawsompong, S., 2019. Selection of SSR Markers for drought resistant sugarcane in Thailand. *International Journal of Agricultural Technology*, 15(6), pp.997-1010.

Ribeiro, R.V., Machado, R.S., Machado, E.C., Machado, D.F.S.P., Magalhães Filho, J.R. and Landell, M.G.A., 2013. Revealing drought-resistance and productive patterns in sugarcane genotypes by evaluating both physiological responses and stalk yield. *Experimental Agriculture*, 49(2), pp.212-224.

Sanji, G., Jun, L., Rukai, C., Muqing, Z. and Daren, P., 2002. Photosynthetic physiology indexes of the drought resistance of sugarcane and its comprehensive evaluation. Zuo Wu Xue Bao, 28(1), pp.94-98.

Team, R.C., 2021. R: A language and environment for statistical computing. Published online 2020.

Xu, C., Mao, J., Li, X., Burner, D.M., Li, C., Hussin Hussin, S.H., Lin, X., Liu, H., Zhao, P., Lu, X. and Liu, X., 2023. Evaluation of drought tolerance in sugarcane genotypes using the membership function value of drought tolerance (MFVD). *Euphytica*, 219(3), p.37.

Zhao, X., Liu, Q., Xie, S., Jiang, Y., Yang, H., Wang, Z. and Zhang, M., 2020. Response of soil fungal community to drought-resistant Ea-DREB2B transgenic sugarcane. *Frontiers in Microbiology*, 11, p.562775.