SUGARCANE OMICS: STATUS OF RESEARCH IN CROP IMPROVEMENT

Muhammad Ehsan Khan
Sugarcane Research and Development Board
Corresponding Author email: ehsansrdb@gmail.com

ABSTRACT
Sugarcane, a vital member of the Poaceae family, plays a significant role in global sucrose production, contributing approximately 80% of the world's total sucrose output. Furthermore, sugarcane serves as a crucial raw material for bioethanol production, offering an alternative source of renewable energy. In the realm of sugarcane genomics, substantial progress has been achieved in transferring genes from various plant and non-plant sources to sugarcane, facilitated by efficient transformation systems, selectable marker genes, and genetic engineering tools. Genetic engineering techniques enable the cloning and characterization of beneficial genes, as well as the enhancement of commercially important traits in elite sugarcane varieties, leading to the development of an ideal cultivar. Sugarcane is a complex polyploid crop, making it challenging to identify a single technique that best confirms polygenic and phenotypic characteristics. To gain a deeper understanding of how basic omics approaches can be applied to sugarcane, particularly in terms of agronomic attributes, industrial quality traits, and responses to various biotic and abiotic stresses. Genetic improvements in sugarcane face obstacles such as its complex genome, low fertility ratio, lengthy production cycle, and susceptibility to numerous biotic and abiotic stresses. Biotechnological interventions are expected to offer solutions to these challenges and enhance sugarcane crop production. Thus, this review article presents up-to-date insights into how advanced omics data, including genomics, proteomics can be harnessed to advance sugarcane cultivation.

Keywords: Studies, status, Sugarcane, Omics

INTRODUCTION

Sugarcane holds the primary position as the world's leading sugar source, contributing to a substantial 80% of global sugar production. Modern sugarcane varieties present a complex genome characterized by high polyploidy and aneuploidy, stemming from a genomic composition of roughly 70%–80% derived from *Saccharum officinarum* and 10%–20% from *S. spontaneum*. This combination leverages the high sugar content of *S. officinarum* with the resilience, disease resistance, and ratooning ability of *S. spontaneum*. Conventional breeding for sugarcane improvement faces considerable challenges due to the crop's limited genetic diversity and intricate genome structure. Recent research efforts in sugarcane have delved into molecular biology, encompassing cytogenetic analysis and various omics investigations, including genomics, transcriptomics and proteomics. These endeavors aim to enhance yields, elevate sucrose content, bolster resistance to biotic and abiotic stresses, and unravel the genetic regulation and underlying mechanisms governing sugarcane's traits and characteristics. Omics approaches rely heavily on understanding the intricate relationships among genetic composition, genes and proteins. However, their effectiveness is greatly dependent on analytical methods like bioinformatics, computational analysis, and various other biological disciplines. These approaches have yielded a wealth of new information concerning the molecular mechanisms underlying sugarcane's resistance and tolerance to herbicides, cold, drought, salinity stress, and plant development. Furthermore, proteomic methods such as two-dimensional difference gel

electrophoresis (2D-DIGE) and isobaric tags for relative and absolute quantitation (iTRAQ) have revealed various differentially expressed proteins (DEPs) and their functions in signal transduction pathways in response to both biotic and abiotic stresses. In more recent times, metabolite analysis has deepened our comprehension of complex regulatory processes, focusing on potential metabolites such as saccharides and derivatives. High-throughput technologies enable the determination of metabolic phenotypes, shedding light on resistance mechanisms.

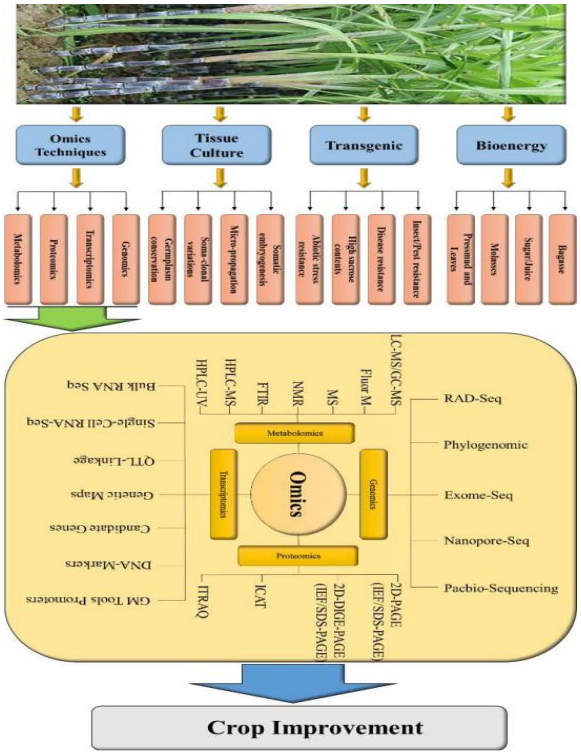

Figure-1 Graphical strategy showing the role of biotechnological interventions for the development of sugarcane crop

Genomics Studies in Sugarcane

Until now, the assembly of the modern sugarcane cultivar genome has remained a formidable challenge due to its intricate nature, characterized by interspecific hybridization, polyploidy, and aneuploidy. The monoploid genome size of *Saccharum officinarum* measures 930 Mb, while that of *Saccharum spontaneum* is 750 Mb. A recent breakthrough by Zhang *et al.*, (2018) involved the assembly of the autopolyploid sugarcane *S. spontaneum* genome using haploid S. spontaneum (AP85-441). This effort resulted in the creation of 32 pseudo-chromosomes, each holding eight homologous groups consisting of four members, encompassing 35,525 genes with well-defined alleles. This remarkable achievement revealed that the reduction in the number of essential chromosomes from *S. spontaneum*, from ten to eight, was the outcome of the fission of two ancestral chromosomes followed by their translocation to four chromosomes. Notably, 80% of nucleotide-binding site encoding genes associated with disease resistance.

Geneticists are actively exploring the connections between the genomes of complex sugarcane and other closely related crops and plants. The genomic complexity of plants in the Poaceae family varies from diploid to triploid, with gene function preservation and origin closely tied to genomic homology. The expansion of genomes in grasses is significantly influenced by TEs, particularly transposons and retro-transposons. Among these, the most abundant reverse transcription element is the LTR (long terminal repeat) retro-transposon, facilitated by transposase proteins in

the insertion-deletion mechanism. The active site of transcription regulates the retro-transposon's movement, reinserting it into the genome after each breeding cycle to increase its copy number.

In the context of sugarcane, TEs can be activated and studied using functional transcriptomic approaches (Mustafa *et al.,* 2018; Shingote *et al.,* 2019). Grasses, especially sugarcane, exhibit a significant degree of TEs, which can be systematically activated and analyzed through functional transcription techniques. This approach helps unveil the complexity of sugarcane traits, such as sucrose accumulation, fiber content, and research on pathogen-resistance proteins. Notably, recent studies have highlighted mutant-like transposases as prominent transposon transcripts in the sugarcane transcriptome (Chuong *at al.,* 2017).

Transcriptomics Studies in Sugarcane

Transcriptome analysis furnishes crucial gene-related data through various in silico techniques, including probe hybridization arrays, expressed sequenced tags (ESTs), or known genes from closely related crops. Brazil boasts one of the largest sugarcane EST databases, housing approximately 238,000 ESTs sourced from 26 different cDNA libraries constructed using various tissues from a diverse range of Brazilian sugarcane varieties (Vettore *at el.,* 2003; Cardoso-Silva *et al.,* 2014). These ESTs have been organized into 43,141 putative unique transcripts, encompassing 26,803 contigs and 16,338 singletons, collectively referred to as sugarcane assembled sequences. Nevertheless, the absence of an accurate and comprehensive reference genome for sugarcane poses a challenge in terms of gene function prediction and harnessing the transcriptome dataset. As a workaround, researchers commonly turn to the reference genome of Sorghum bicolor, which exhibits a high degree of homology (95%) in genic regions between sugarcane and sorghum genomes. Notably, approximately 47% of the unigenes in the sugarcane transcriptome data exhibited top hits in BLASTx searches matching Sorghum bicolor proteins. While high-throughput RNA-Seq has found widespread use in eukaryotic transcriptome analyses, it has limitations, such as generating short reads that necessitate extensive computational assemblies and cannot span full-length transcripts. To overcome these limitations, single-molecule long-read sequencing technologies, such as Pacific Biosciences long-read isoform sequencing (Iso-Seq), have been developed and widely adopted in transcriptome sequencing. This approach offers a superior alternative for sequencing more complete transcriptomes and enhances the accuracy of gene model prediction and validation.

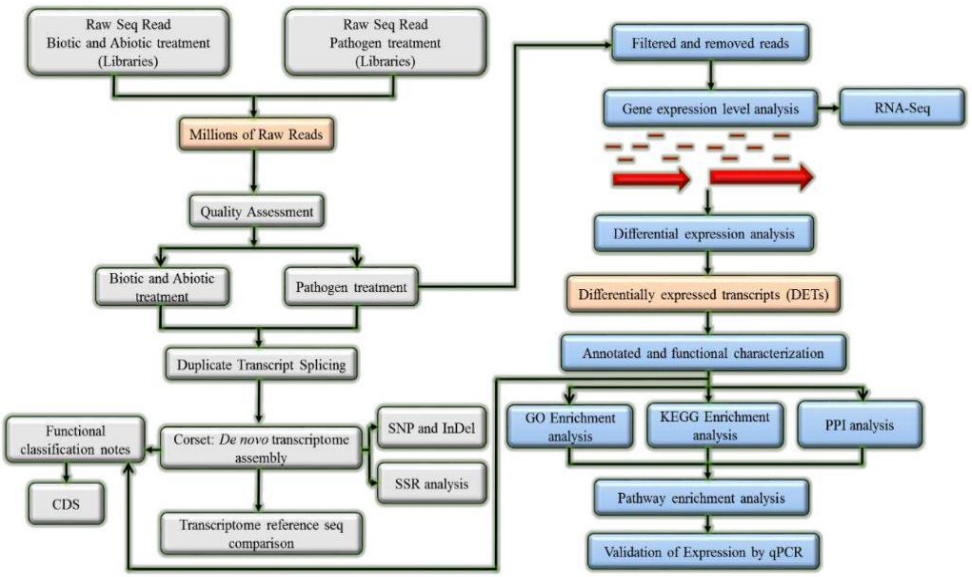

Figure-2 A workflow sketch of sugarcane transcriptome analysis including construction of reference transcriptome from the de novo assembly and annotation and

functional characterization of differentially expressed transcripts (DETs)

Proteomics studies in Sugarcane

In addition to transcriptomics, proteomics approaches also yield fresh insights into intricate biological processes. Therefore, understanding the mechanisms for quantifying proteins and their post-translational modifications is fundamental for comprehending biological systems. While the genome remains static, an individual's proteome dynamically responds to environmental stimuli and intracellular metabolite levels.

To assess differential and comparative protein expression levels in sugarcane under various biotic and abiotic stresses, researchers employ diverse protein isolation and quantitation tools. These tools include techniques like two-dimensional electrophoresis (2-DE), mass spectrometry (MS), and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF-MS), as reviewed by Barnabas *et al*. in 2015. Additionally, iTRAQ represents one of the primary quantitation methods utilized in differential plant proteomic research. In the realm of sugarcane, numerous current proteomic studies have been conducted using both gel-based and gel-free tools under different biotic and abiotic stress conditions.

Proteomics approaches have been employed to study the regulation of the proteome during the development processes of sugarcane, as well as to profile the proteomes of various sugarcane organs. To date, numerous research reports have outlined the characteristics of sugarcane proteomes, encompassing topics such as somatic embryogenesis in both embryogenic and non-embryogenic callus induced by putrescine and exposure to different red and blue light conditions, the cell wall proteomes of different sugarcane organs at various developmental stages, and the proteome features of sugarcane stems, cell wall remodeling in suspension cells, and lignin composition during stem development.

Fonseca and collaborators characterized the cell wall proteomes of both young and mature leaves and stems of sugarcane. Their findings revealed that sugarcane leaves and young stems exhibited the highest rate of lipid metabolism (LM) compared to other species. They identified 277 cell wall proteins (CWPs) in sugarcane. Another study by Calderan-Rodrigues and colleagues focused on characterizing the cell wall proteome of young sugarcane culms, with the aim of identifying proteins involved in cell wall biogenesis. Their work identified 84 distinct cell wall proteins related to lipid metabolism and oxido-reductase activity.

Challenges in Cane Breeding

More recently, sugarcane breeding has primarily emphasized the augmentation of its sucrose content, as sucrose serves as the primary substrate for sugar and ethanol production. Consequently, there has been no distinct categorization of sugarcane genotypes, as both first-generation technology and raw sugar production rely on the quantity of sucrose accumulated in the stalk, a characteristic greatly beneficial to the sugarcane industry. However, the advent of second-generation technology, which relies on cellulose, necessitates a return to genetic diversity within sugarcane breeding projects.

The use of omics techniques and modern bioinformatics tools has simplified the annotation of sequences and the elucidation of regulatory mechanisms in sugarcane, thereby enhancing our understanding of its genome, genetics, physiology, and molecular biology. Regulatory genes associated with sucrose synthesis, their allelic variations, copy numbers, and expression patterns have been functionally characterized in contemporary sugarcane cultivars.

A significant challenge lies in deciphering the complete genome sequence of sugarcane due to its intricate ploidy and aneuploidy characteristics. Breeding programs should leverage these tools and incorporate them into their selection

processes to generate superior new cultivars that can meet the current and future requirements of the industry and the expectations of society at large.

CONCLUSION

Sugarcane possesses all the attributes necessary to serve as a primary raw material for energy, biofuels, and electricity production, making it a valuable resource in the bioprocessing and bio-refinery industries. Among GM crops, more than 90% are insect-resistant or herbicide-resistant, resulting in a substantial reduction, around 37%, in the use of chemical pesticides, a 22% increase in yields, and a 68% rise in growers' profits. The discovery of genes through various "omics" methods is crucial for advancing sugarcane improvement programs and unraveling the mechanisms governing plant adaptation and responses to biotic and abiotic environments. EST-SSRs, and development of new markers and their integration into genetic maps is set to expedite breeding and improvement initiatives. The present era has witnessed the progression of gene silencing and overexpression techniques, enabling the exploration of their vital functions and the generation of novel phenotypes that would otherwise be unattainable through conventional methods. Moreover, genomics studies related to genes of interest should place greater emphasis on the interaction of multiple genes and metabolic pathways with the environment, with a preference for using model crop plants. Furthermore, CRISPR and epigenetic molecular events, which are highly relevant to plant adaptation in changing environments, deserve heightened attention. Finally, a crucial step toward crop improvement is fostering transparent communication between molecular biologists and plant physiologists on one side and farmers, breeding companies, and the public on the other, to collaboratively address economic, sociological, legal, and ethical challenges.

REFERENCES

Waclawovsky, A.J.; Sato, P.M.; Lembke, C.G.; Moore, P.H.; Souza, G.M. Sugarcane for bioenergy production: An assessment of yield and regulation of sucrose content. Plant Biotechnol. J. 2010, 8, 263–276.

Ali, A.; Pan, Y.-B.; Wang, Q.-N.; Wang, J.-D.; Chen, J.-L.; Gao, S.-J. Genetic diversity and population structure analysis of Saccharum and Erianthus genera using microsatellite (SSR) markers. Sci. Rep. 2019, 9, 395.

D'Hont, A.; Grivet, L.; Feldmann, P.; Glaszmann, J.; Rao, S.; Berding, N. Characterisation of the double genome structure of modern sugarcane cultivars (Saccharum spp.) by molecular cytogenetics. Mol. Gen. Genet. MGG 1996, 250, 405–413.

Zhang, J.; Zhang, X.; Tang, H.; Zhang, Q.; Hua, X.; Ma, X.; Zhu, F.; Jones, T.; Zhu, X.; Bowers, J. Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. Nat. Genet. 2018, 50, 1565.

Garsmeur, O.; Droc, G.; Antonise, R.; Grimwood, J.; Potier, B.; Aitken, K.; Jenkins, J.; Martin, G.; Charron, C.; Hervouet, C. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat. Commun. 2018, 9, 2638.

Mustafa, G.; Joyia, F.A.; Anwar, S.; Parvaiz, A.; Khan, M.S. Biotechnological interventions for the improvement of sugarcane crop and sugar production. In Sugarcane-Technology and Research; IntechOpen: London, UK, 2018; pp. 113–138.

Barnabas, L.; Ramadass, A.; Amalraj, R.S.; Palaniyandi, M.; Rasappa, V. Sugarcane proteomics: An update on current status, challenges, and future prospects. Proteomics 2015, 15, 1658–1670.

Kaur, L.; Dharshini, S.; Ram, B.; Appunu, C. Sugarcane Genomics and Transcriptomics. In Sugarcane Biotechnology: Challenges and Prospects; Springer: Cham, Switzerland, 2017; pp.13-32.

Sanghera, G.S.; Singh, R.P.; Tyagi, V.; Thind, K. Recent genomic approaches for sugarcane improvement: Opportunities and challenges. In Quality and Quantum Improvement in Field Crops; Agrobios: Jodhpur, India, 2017; pp. 109–152.

Khueychai, S.; Jangpromma, N.; Daduang, S.; Jaisil, P.; Lomthaisong, K.; Dhiravisit, A.; Klaynongsruang, S. Comparative proteomic analysis of leaves, leaf sheaths, and roots of drought-contrasting sugarcane cultivars in response to drought stress. Acta Physiol. Plant. 2015, 37, 88.

Augustine, S.M.; Narayan, J.A.; Syamaladevi, D.P.; Appunu, C.; Chakravarthi, M.; Ravichandran, V.; Tuteja, N.; Subramonian, N. Overexpression of EaDREB2 and pyramiding of EaDREB2 with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (Saccharum spp. hybrid). Plant Cell Rep. 2015, 34, 247–263.

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

Su, Y.; Xu, L.; Wang, Z.; Peng, Q.; Yang, Y.; Chen, Y.; Que, Y. Comparative proteomics reveals that central metabolism changes are associated with resistance against Sporisorium scitamineum in sugarcane. BMC Genom. 2016, 17, 800.

Allwood, J.W.; Ellis, D.I.; Goodacre, R. Metabolomic technologies and their application to the study of plants and plant–host interactions. Physiol. Plant. 2008, 132, 117–135.