

University of São Paulo
“Luiz de Queiroz” College of Agriculture

Differential gene expression in roots of sugarcane hybrids provides insights into
drought stress tolerance

Ana Letycia Basso Garcia

Thesis presented to obtain the degree of Doctor of
Science. Area: Genetics and Plant Breeding

Piracicaba
2023

Ana Letycia Basso Garcia
Agronomist

Differential gene expression in roots of sugarcane hybrids provides insights into drought stress tolerance

Advisor:
Prof. Dr. **GABRIEL RODRIGUES ALVES MARGARIDO**

Thesis presented to obtain degree of Doctor in Science.
Area: Genetics and Plant Breeding.

Piracicaba
2023

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Garcia, Ana Letycia Basso

Differential gene expression in roots of sugarcane hybrids provides insights into drought stress tolerance / Ana Letycia Basso Garcia - - Piracicaba, 2023.

63 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. RNA-Seq 2. *Saccharum* 3. Estresse abiótico 4. Perfil de expressão gênica
I. Título

DEDICATION

To my beloved dad (*in memoriam*), who always supported my dreams financially and emotionally and inspired me to follow on the scientific journey since I was a kid. He was an example of ideonality and passion for the work. He has told me in my dreams that science is the right and best way to go, despite there is to much behind the nature and the God laws.

To my mother, who has always trusted in my potential and gives me the strength to continue, even after a major failure. She never lets me give up, she wipes away my tears, and she laughs with me at every victory.

To all my ancestors who, by chance, made the best combination to share their genes, granting me the opportunity to exist to explore such brillant heterosis.

ACKNOWLEDGEMENTS

I thank God for giving me the light and wisdom to investigate and understand nature and to contribute to human development through science.

I express my deepest gratitude to my thesis advisor, Mr. Gabriel Rodrigues Alves Margarido, for his dedicated support, patience, and guidance. I would like to extend my sincere thanks to the technical support of Professor Monalisa Sampaio Carneiro from the Federal University of São Carlos (UFSCar) and her team, and Professor Anete Pereira de Souza and her group from the University of Campinas (Unicamp). I am also thankful to Dr Philippa Borrill from the John Innes Centre (United Kingdom), for her valuable discussions and remote collaboration during the COVID-19 pandemic.

I could not have undertaken this journey without the support of my lab mates for deep technical discussions and feedback, specially to Victor Hugo Mello Pessoa, for gene discovery and color pallets choices; Guilherme Kenichi Hosaka, for writing style discussions and coding support; Fernando Henrique Correr, for coding and bioinformatics support; Amanda Avelar de Oliveira and Lorena for strong moral support and scientific insights. I also would like to thank colleagues, peers, and friends at the Department of Genetics.

I would be remiss in not mentioning my family and friends, especially my parents. Their belief in me has kept my spirits and motivation high during this process. I would also like to thank my current employer Veracel Celulose S.A., for being comprehensive with my dedication to the thesis and stimulating its conclusion through the individual development plan.

Lastly, I thank The São Paulo Research Foundation (FAPESP) for their financial support in developing this project. Additionally, this endeavor would not have been possible without the scholarship support from CAPES and CNPq.

CONTENTS

RESUMO.....	6
ABSTRACT	7
1. GENERAL INTRODUCTION	8
2. CONCLUSIONS	11
References	13

RESUMO

Expressão gênica diferencial em raízes de híbridos de cana-de-açúcar fornece informações sobre a tolerância ao estresse por déficit hídrico

A cana-de-açúcar é uma das principais culturas que contribuem para a economia brasileira. Os programas de melhoramento genético de cana-de-açúcar têm sido desafiados a desenvolver cultivares cada vez mais produtivas e que possam tolerar condições ambientais adversas, como o estresse hídrico. A análise da expressão gênica é uma forma eficaz de identificar genes envolvidos na resposta à seca e entender como esses genes regulam as respostas fisiológicas da planta. O objetivo deste trabalho foi identificar genes em raízes de cana-de-açúcar que regulam a resposta ao estresse por déficit hídrico no solo e identificar padrões de expressão gênica associados à tolerância à seca. Para isso, avaliamos a expressão gênica em seis híbridos comerciais de cana-de-açúcar (RB867515, RB855113, RB855536, RB92579, SP79-1011 e SP80-3280). Foram coletadas raízes de um experimento em esquema fatorial 2 x 2 instalado em casa de vegetação, em que os fatores corresponderam aos efeitos de disponibilidade de água no solo (tratamento controle e tratamento com privação de água por cinco dias) e ao efeito de genótipos, agrupados em suscetíveis ou tolerantes à seca. Nós coletamos raízes e realizamos RNA-seq para três réplicas biológicas de cada tratamento. Essas sequências foram utilizadas para montar um transcriptoma *de novo*, que foi anotado funcionalmente e utilizado como referência para estimar os níveis de expressão gênica a partir de uma estratégia de *quasi*-mapping. Em seguida, fizemos uma análise de expressão gênica diferencial para identificar os principais genes responsivos aos efeitos principais de estresse por seca e de grupo de cultivar. Foram encontrados 1.913 genes com expressão aumentada e 1.550 genes suprimidos para efeito principal de estresse por seca. A grande maioria desses genes são relacionados a respostas para estresses abióticos, e incluem precursores de quinases, fatores de transcrição e proteínas que conferem homeostase celular. Complementarmente, identificamos 12.939 genes diferencialmente expressos (DEGs) para o efeito principal de cultivares, sugerindo divergência significativa de expressão gênica entre cultivares tolerantes e suscetíveis. Esses genes estão, em sua maioria, relacionados ao estresse biótico e à resposta geral à estresses. Por fim, identificamos 412 DEGs que possuem interação entre os efeitos de tratamento e de grupo de cultivar, dos quais mais de 86% foram induzidos em genótipos tolerantes sob estresse por seca. Muitos desses genes enriqueceram ontologias gênicas relacionadas a processos biológicos definidos como: via de sinalização ativada por etileno, regulação negativa de citocinina e sinalização celular. Concluímos que a tolerância à seca está possivelmente associada a características gênicas inatas presentes nos genótipos tolerantes. Alguns desses genes estão envolvidos em vias de sinalização hormonal e são induzidos em raízes de cultivares tolerantes sob estresse por seca. Sugerimos, portanto, um estudo aprofundado para os genes de tolerância à seca que identificamos neste trabalho, para buscar compreender os mecanismos envolvidos na regulação desses genes. Esses resultados são valiosos na compreensão de mecanismos genéticos de tolerância a seca, que contribuirá com a melhor eficiência no desenvolvimento de novas cultivares cana-de-açúcar tolerantes ao estresse por seca

Palavras-chave: RNA-Seq, *Saccharum*, Estresse abiótico, Perfil de expressão gênica

ABSTRACT

Differential gene expression in roots of sugarcane hybrids provides insights into drought stress tolerance

Sugarcane is one of the main crops that contributes to the Brazilian economy. A major challenge for sugarcane breeding programs has been developing cultivars that are more productive and tolerant to adverse environmental conditions, such as drought stresses. Gene expression analysis is an effective way to identify genes involved in drought response and understand how these genes regulate plant physiology under stress. The aim of this study was to identify genes in sugarcane roots that regulate the response to soil water deficit stress and to identify gene expression patterns associated with tolerance to drought. For this, we evaluated gene expression in six commercial sugarcane hybrids (RB867515, RB855113, RB855536, RB92579, SP79-1011, and SP80-3280). Roots were collected from a 2 x 2 factorial experiment in a greenhouse. One of the main factors tested was the effect of water availability in the soil (control treatment and treatment with water deprivation for five days) and the other factor was the effect of genotypes, grouped as susceptible or tolerant to drought. We performed RNA-seq for three biological replicates of roots of each treatment. These sequences were used to assemble a *de novo* transcriptome, which was functionally annotated and used as a reference to estimate gene expression levels in each sample from a *quasi*-mapping strategy. We then performed a differential gene expression analysis to identify genes responsive to the main effects of drought stress and cultivar group. Finally, we explored the interaction between these two main effects. We found 1,913 upregulated genes and 1,550 downregulated genes for the main effect of drought stress, showing common responses to stress between the two groups of cultivars. The majority of these genes were related to abiotic stress responses and include kinase precursors, transcription factors, and proteins that confer cellular homeostasis. In addition, we identified 12,939 differentially expressed genes (DEGs) for the main effect of cultivars, suggesting a significant divergence in gene expression between tolerant and susceptible cultivars. These genes are mostly related to biotic stress and the general response to stress. Finally, we identified 412 DEGs that have an interaction between the effects of the treatment and cultivar group, in which more than 86% were induced in tolerant genotypes under drought stress. Many of these genes enriched gene ontologies related to defined biological processes such as: ethylene-activated signaling pathway, negative cytokinin regulation, and cell signaling. Therefore, we conclude that drought tolerance is possibly associated with innate gene characteristics present in tolerant genotypes. Some of these genes are involved in hormonal signaling pathways and are induced in the roots of tolerant cultivars under drought stress. Thus, we suggest an in-depth investigation of the drought tolerance genes that we identified in this work, to better understand the role of each of these genes in regulating drought stress tolerance. These results are valuable in understanding the genetic mechanisms of drought tolerance, which will contribute to better efficiency in the development of new drought-tolerant sugarcane cultivars.

Keywords: RNA-Seq, *Saccharum*, Abiotic stress, Expression profile

1. GENERAL INTRODUCTION

Sugarcane (*Saccharum spp.*) is the main sugar crop and an important feedstock to produce ethanol fuel. Given the need to reduce greenhouse gas emissions, it has emerged as a valuable source of fiber and various forms of bioenergy to reduce the dependence on petroleum (VANDENBERGHE et al., 2022). Brazil leads the global sugarcane production in 2022/23 with 610.8 Mt, followed by India (490.5 Mt) (CONAB, 2023; PRESS INFORMATION BUREAU, 2023). Brazil is expected to produce 36% of the world's sugarcane demand by 2030. Its sugarcane will account for 20% of global sugar production and 84% of world sugarcane-based ethanol production. Most of the ethanol fuel produced in Brazil supplies the growing domestic market, supporting the potential growth of the mixing of biofuels with gasoline and diesel (OECD-FAO, 2021).

Due to the increasing demand for sugarcane-based materials, Brazil has invested in innovative technologies, such as the development of new commercial sugarcane hybrids with higher yield potential (CURSI et al., 2021). The purpose has been to increase the potential to breakdown the production into sugar and ethanol or to only produce bioenergy, according to market demands. Based on that, sugarcane breeding programs aim to develop genotypes with high yield potential and that are able to tolerate adverse conditions, regarding pest and disease resistance and tolerance to abiotic stresses (CURSI et al., 2021; MORAIS et al., 2015). Sugarcane has high yield potential in tropical weather but requires high water availability in the soil in the early stages of development, especially during the stages of intense growth and stalk elongation (ENDRES et al., 2018). Furthermore, due to climate changes and the expansion of fields to marginal areas, water availability is sometimes deficient. Thus, currently available genotypes cannot reach their maximum productive potential (DAROS; OLIVEIRA; BARBOSA, 2015).

Soil water availability during plant development is decisive for sugarcane yield, so drought can lead to losses greater than 60% in sugarcane production (BASNAYAKE et al., 2015; GENTILE et al., 2015). During plant development, the water deficit affects cell expansion and reduces plant growth, in addition to membrane integrity, osmotic adjustment, and plant-water relations (BUDZINSKI et al., 2019). The response to drought stress involves a complex set of genetic and physiological mechanisms. Drought stress results in changes to gene expression by activation of specific pathways due to the interaction of signaling molecules that orchestrate transcription regulators, such as transcription factors (KIDO et al., 2012). Briefly, signaling molecules present in the plasma membrane of plant roots, such as receptor-like kinases (RLKs), detect the water deficit in the soil and start signal transduction, which increases the level of reactive

oxygen species (ROS) and activates a cascade of signaling molecules that transduce the signal to nuclear content. So, the nuclear machinery is designed to work to manage genetic imbalance and avoid cellular damage (TAIZ et al., 2017).

Gene expression analysis is valuable for understanding gene profiles and discovering differentially expressed patterns in response to biotic or abiotic stress. RNA-Seq is an efficient and widely used technique for cDNA-mediated transcriptome sequencing. The technique has been used to uncover the expression profiles of genes related to abiotic stress in sugarcane, such as nitrogen deficiency (YANG et al., 2019) and drought tolerance (BELESINI et al., 2017; CONTILIANI et al., 2023; FERREIRA et al., 2012; GENTILE et al., 2013; TELLES et al., 2019; VANTINI et al., 2015). Nevertheless, little is known about how genetic response is modulated in sugarcane roots, as most published papers regarding drought stress are focused on genetic mechanisms of response and physiological reactions in leaves. Considering that plant tissues are specialized, gene expression patterns change between organs under the same environmental conditions.

Commercial sugarcane cultivars are interspecific hybrids between *S. officinarum* and *S. spontaneum*. Due to its complex genome, which is allopolyploid and aneuploid, breeding for specific traits is always a challenge (THIRUGNANASAMBANDAM; HOANG; HENRY, 2018). Moreover, not much is known about its genome architecture and genes related to quantitative traits. However, genomic and molecular tools have been developed to help breeding programs and increase annual genetic gains for this crop, which is currently stable (CURSI et al., 2021; DAL-BIANCO et al., 2012). A Brazilian project sequenced the genome of the commercial hybrid SP80-3280, in which more than 373.000 putative genes and potential regulatory regions were identified (SOUZA et al., 2019). In addition to that, a monoploid genome of the cultivar R570 (GARSMEUR et al., 2018) and an allele-defined genome of *S. spontaneum* (ZHANG et al., 2018) are currently published and available. Even with the high contribution of these works, the genome is still incomplete. Because of that, transcriptome studies and differential expression analysis are a relevant source to identify key genes and pathways involved in response to abiotic stresses and help reveal genome behavior under harsh environmental conditions, such as drought.

In this study, we investigated the genetic responses to drought stress in the roots of drought-tolerant and susceptible hybrid sugarcane cultivars. Using RNA-Seq analysis, we identified differentially expressed genes that respond to drought and evaluated the gene expression profile of key genes in drought-stimulated pathways in both the tolerant and susceptible groups of cultivars. We also explore the marginal effects of drought stress and the type of tolerance to drought. Although most studies focus on above-ground plant organs, this study explored changes in gene

expression in roots. This research aimed to reveal and explore drought-responsive genes and cultivar-specific genes that play a role in tolerance to drought stress. Our results contribute to the understanding of the genetic mechanisms underlying drought tolerance in sugarcane hybrids and the molecular role that roots play in supporting drought-tolerant plants in stressful conditions.

2. CONCLUSIONS

Our research revealed consistent gene expression patterns in all cultivars under drought stress in response to abiotic stress. We found that signaling molecules, regulation of the antioxidant system, and alterations in hormone biosynthesis, mainly ABA and ethylene, play an important role in the response to drought stress that occurs in sugarcane roots after being subjected to five days of drought. Genes that regulate the regulation of cell homeostasis were induced under drought stress to prevent cellular dehydration and ROS detoxification. They are a key mechanism that the plant activates to maintain its functions under drought stress.

The differences between the drought-tolerant and drought-susceptible cultivar groups were evidenced by the differential expression of genes that act on the activation of the immune system, including the presence of signaling molecules and disease-resistant proteins. The cultivars within each group showed similar expression patterns for these genes, but an opposite expression pattern from the cultivars of the other group. This suggests that, in addition to the selection process for breeding being biased to keep more disease-tolerant cultivars, molecular intersections between biotic and abiotic stress responses can occur.

When we studied the interaction between the effects of treatment and the cultivar group, we found key genes that act on signal transduction and hormone biosynthesis in the ABA and ethylene pathways. The ethylene pathway plays a role in drought tolerant plants. This was reinforced by the expression patterns of members of the AP2/ERF transcription factor family. Members of this family contribute to regulation of gene expression in roots under abiotic stress and mediate response with plant shoots, contributing with mechanisms to avoid loss of water during low moisture in the soil in the above ground parts. Furthermore, tolerant plants revealed genes that cause reduced plant growth and development, aiming to save energy. We identified key genes in the innate immune system, acting mainly in signal transduction or hormone biosynthesis in the ABA and ethylene pathways, regulating drought stress tolerance in sugarcane roots and plant development. Our results reveal that the innate genetic differences between tolerant and susceptible cultivars may indicate tolerance to drought stress.

Although there are some genes that are differentially expressed under drought stress, regardless of cultivar tolerance to drought, there are other genes that the expression pattern is dependent on the cultivar ability to support drought. As a general response to drought stress, we found that roots exposed to drought stress induced genes associated with cellular homeostasis and hormone signaling molecules pathways. On the other hand, we discovered that the biggest difference between drought-tolerant and drought-susceptible cultivars in this study regards to the

immune system, with special attention to disease-resistant proteins and protein kinases. When exploring the marginal effects of both factors, we found that most differentially expressed genes were induced in drought-tolerant plants. This was useful in inferring putative genes that can be used in sugarcane breeding programs to develop new drought tolerant commercial sugarcane cultivars. Due to the complexity of the sugarcane genome, our findings must be explored in more depth at the levels of genome and transcript isoforms. Moreover, including other cultivars and considering different time points of drought stress may produce a more accurate result on specific genes. This study provides important insights about the genetic responses of sugarcane to drought stress, which may help breeders develop cultivars that are more resistant to it in the future. Understanding these genetic responses will help sugarcane breeding programs focus on creating cultivars that are more tolerant to adverse environmental conditions such as drought and gain or improve yield in a dynamic climate change scenario.

REFERENCES

- BASNAYAKE, J. et al. Sugarcane for water-limited environments. Variation in stomatal conductance and its genetic correlation with crop productivity. **Journal of Experimental Botany**, v. 66, n. 13, p. 3945–3958, jul. 2015.
- BELESINI, A. A. A. et al. De novo transcriptome assembly of sugarcane leaves submitted to prolonged water-deficit stress. **Genetics and Molecular Research Mol. Res**, v. 16, n. 162, p. 16028845–16028845, 2017.
- BUDZINSKI, I. G. F. et al. Network analyses and data integration of proteomics and metabolomics from leaves of two contrasting varieties of sugarcane in response to drought. **Frontiers in Plant Science**, v. 10, p. 1524, 28 nov. 2019.
- CONAB. **Acompanhamento da safra brasileira**. [s.l.: s.n.]. v. 11
- CONTILIANI, D. F. et al. Drought-triggered leaf transcriptional responses disclose key molecular pathways underlying leaf water use efficiency in sugarcane (*Saccharum* spp.). **Frontiers in Plant Science**, v. 14, p. 1182461, 8 maio 2023.
- CURSI, D. E. et al. History and Current Status of Sugarcane Breeding, Germplasm Development and Molecular Genetics in Brazil. **Sugar Tech**, p. 1–22, 2021.
- DAL-BIANCO, M. et al. Sugarcane improvement: How far can we go? **Current Opinion in Biotechnology**, 2012.
- DAROS, E.; OLIVEIRA, R. A.; BARBOSA, G. V. S. **45 anos de variedades RB de cana de açúcar**. 1. ed. Curitiba: Graciosa, 2015.
- ENDRES, L. et al. Morphological changes recorded in different phenophases of sugarcane plants subjected to water stress in tropical field conditions. **AJCS**, v. 12, n. 07, p. 1835–2707, 2018.
- FERREIRA, T. H. et al. microRNAs Associated with Drought Response in the Bioenergy Crop Sugarcane (*Saccharum* spp.). **PLoS ONE**, v. 7, n. 10, p. e46703, 11 out. 2012.
- GARSMEUR, O. et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications**, v. 9, n. 1, p. 2638, 6 dez. 2018.
- GENTILE, A. et al. Effects of drought on the microtranscriptome of field-grown sugarcane plants. **Planta**, v. 237, n. 3, p. 783–798, 2013.
- GENTILE, A. et al. MicroRNAs and drought responses in sugarcane. **Frontiers in Plant Science**, v. 6, n. February, p. 1–13, 2015.
- KIDO, É. A. et al. New insights in the sugarcane transcriptome responding to drought stress as revealed by supersage. **The Scientific World Journal**, v. 2012, p. 1–14, 2 maio 2012.

MORAIS, L. K. et al. **Melhoramento Genético da Cana-de-Açúcar**. Aracaju: Embrapa Tabuleiros Costeiros, 2015.

OECD-FAO. OECD-FAO Agricultural Outlook 2021-2030. Em: **Agricultural Outlook**. OECD-FAO Agricultural Outlook. [s.l.] OECD, 2021.

PRESS INFORMATION BUREAU. **Final estimates of production of major crops released for the year 2022-23**. [s.l.: s.n.].

SOUZA, G. M. et al. Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. **GigaScience**, v. 8, n. 12, 1 dez. 2019.

TELLES, B. R. et al. Prolonged Water Deficit Reveals New Profile of Sugarcane Gene Expression and Metabolic Pathway Related to Tolerance. **Sugar Tech**, v. 21, n. 3, p. 451–461, 2019.

THIRUGNANASAMBANDAM, P. P.; HOANG, N. V.; HENRY, R. J. The Challenge of Analyzing the Sugarcane Genome. **Frontiers in plant science**, v. 9, 14 maio 2018.

VANDENBERGHE, L. P. S. et al. Beyond sugar and ethanol: The future of sugarcane biorefineries in Brazil. **Renewable and Sustainable Energy Reviews**, v. 167, p. 112721, 1 out. 2022.

VANTINI, J. S. et al. Differential gene expression in drought-tolerant sugarcane roots. **Genetics and Molecular Research**, v. 14, n. 2, p. 7196–7207, 2015.

YANG, Y. et al. Transcripts and low nitrogen tolerance: Regulatory and metabolic pathways in sugarcane under low nitrogen stress. **Environmental and Experimental Botany**, v. 163, p. 97–111, 2019.

ZHANG, J. et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. **Nature Genetics**, v. 50, n. 11, p. 1565–1573, 8 nov. 2018.