

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CAMPUS DE JABOTICABAL**

**DINÂMICA DO PERFIL TRANSCRICIONAL DE DUAS
CULTIVARES DE CANA-DE-AÇÚCAR CONTRASTANTES À
SECA E SUBMETIDAS AO DÉFICIT HÍDRICO PROLONGADO**

Daniela Konrad

Bióloga

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CAMPUS DE JABOTICABAL**

**DINÂMICA DO PERFIL TRANSCRICIONAL DE DUAS
CULTIVARES DE CANA-DE-AÇÚCAR CONTRASTANTES À
SECA E SUBMETIDAS AO DÉFICIT HÍDRICO PROLONGADO**

Daniela Konrad

Orientador: Profa. Dra. Maria Inês Tiraboschi Ferro

Coorientadora: Dra. Juliana da Silva Vantini

Coorientador: Prof. Dr. Daniel Guariz Pinheiro

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias - Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia - Genética e Melhoramento de Plantas

K82d

Konrad, Daniela

Dinâmica do perfil transcricional de duas cultivares de cana-de-açúcar contrastantes à seca e submetidas ao déficit hídrico prolongado / Daniela Konrad. -- Jaboticabal, 2019

94 p. : tabs., fotos

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),
Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientadora: Maria Inês Tiraboschi Ferro

Coorientadora: Juliana da Silva Vantini

1. Cana-de-açúcar. 2. Déficit hídrico prolongado. 3. Expressão diferencial. 4. RNA-Seq. 5. RT-qPCR. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

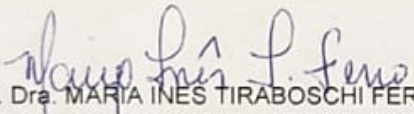
Essa ficha não pode ser modificada.


CERTIFICADO DE APROVAÇÃO

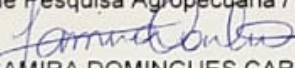
TÍTULO DA DISSERTAÇÃO: DINÂMICA DO PERFIL TRANSCRICIONAL DE DUAS CULTIVARES DE CANA-DE-AÇÚCAR CONTRASTANTES À SECA E SUBMETIDAS AO DÉFICIT HÍDRICO PROLONGADO

AUTORA: DANIELA KONRAD
ORIENTADORA: MARIA INES TIRABOSCHI FERRO
COORIENTADOR: DANIEL GUARIZ PINHEIRO
COORIENTADORA: JULIANA DA SILVA VANTINI

Aprovada como parte das exigências para obtenção do Título de Mestra em AGRONOMIA (GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:


Profa. Dra. MARIA INES TIRABOSCHI FERRO
Departamento de Tecnologia / FCAV - UNESP


Pesquisadora POLIANA FERNANDA GIACHETTO
Empresa Brasileira de Pesquisa Agropecuária / EMBRAPA - Campinas/SP


Pesquisadora Dra. SAMIRA DOMINGUES CARLIN CAVALLARI
Centro de Cana-IAC / Ribeirão Preto/SP

Jaboticabal, 18 de outubro de 2019

DADOS CURRICULARES DO AUTOR

Daniela Konrad – nascida em 22 de novembro de 1994, no município de Palotina, Paraná, filha de Luís Arthur Konrad e Sueli Genero Konrad. Em 2012, ingressou no curso de Bacharelado em Ciências Biológicas na Universidade Federal do Paraná – UFPR – Palotina, vindo a graduar-se em fevereiro de 2017. Em agosto de 2017 iniciou o curso de mestrado junto ao Programa de Pós-Graduação em Agronomia (Genética e Melhoramento de Plantas) na Faculdade de Ciências Agrárias e Veterinárias (UNESP/Jaboticabal), sob a orientação da Profa. Dra. Maria Inês Tiraboschi Ferro, o que resultou no presente trabalho.

*“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito.
Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.” (Martin
Luther King)*

DEDICATÓRIA

À Deus, por sua infinita bondade, amparo e consolo nos momentos mais difíceis, por conceder-me esta vida maravilhosa;

Aos meus pais, Luis Arthur Konrad e Sueli Genero Konrad, que sempre deram o melhor de si para propiciar as condições para minha formação profissional e do meu irmão Mateus Konrad, a quem também dedico este trabalho, servindo-me de carinho, incentivo, confiança e compreensão em todos os momentos da minha vida. Se não fossem vocês eu não teria superado cada etapa da minha vida e todas as dificuldades que passei até chegar aqui. Obrigado por torceram tanto por mim e me ajudarem nos momentos difíceis. Amo muito vocês!

Ao meu noivo Gustavo Mastria, por sempre acreditar no meu potencial, até quando nem eu mesma acreditei. Por nunca permitir que eu desistisse frente às dificuldades e me incentivar a todo momento a atingir meus objetivos. Agradeço pela compreensão, carinho e amor que dedicou durante essa etapa da minha vida, sem se importar com a distância que ela colocou entre nós. Agradeço por acrescentar tanto na minha vida, você é um presente de Deus. Te amo muito!

AGRADECIMENTOS

Infinitamente a Deus, por sempre me guiar nos melhores caminhos e me dar muita força para lutar pelos meus objetivos, mesmo quando a saudade da família apertou no peito. Obrigada Senhor Deus por esta etapa concluída e por tudo que providenciaste em minha vida. E igualmente a minha família e meu noivo, que sempre me deram todo suporte e amor do mundo, e mesmo longe nunca estiveram ausentes;

À Profa. Dra. Maria Inês Tiraboschi Ferro, pela orientação, ensinamentos, correções, incentivos, conselhos nos momentos difíceis e por toda a sua experiência. Obrigada por tudo e pela oportunidade de realizar este trabalho;

Ao mestrando Rafael Correia da Silva, pelos ensinamentos, parceria e análises de bioinformática. Agradeço muito a paciência, as sugestões e a disponibilidade em participar deste trabalho. Muito obrigada, este trabalho é nosso;

À Dra. Juliana da Silva Vantini, minha coorientadora, pela experiência, amizade, muitas risadas, correções e sugestões. Obrigada por todo o aprendizado que me proporcionou e por me incentivar tanto durante a jornada deste trabalho;

Ao Prof. Dr. Daniel Guariz Pinheiro, meu coorientador, pelas sugestões, ensinamentos, e análises de bioinformática;

Ao Prof. Dr. Jesus Aparecido Ferro, pelas contribuições valiosas para o desenvolvimento deste projeto e pela disponibilidade do laboratório;

Aos pesquisadores Dra. Poliana Fernanda Giachetto e Dr. Antônio Nhani Júnior, pelos ensinamentos e pelas análises de bioinformática. Suas correções e sugestões foram de fundamental importância e contribuíram muito para a realização deste trabalho. Muito obrigada;

Aos membros da banca de qualificação: Profa. Dra. Maria Inês Tiraboschi Ferro, Dra. Luciana Rossini Pinto e Dr. Antonio Nhani Junior, pela disponibilidade, sugestões e correções;

Aos companheiros (as) de trabalho do Laboratório de Bioquímica e Biologia Molecular e do Laboratório de Bioinformática, em especial a Maria Oliza, Rafa, Jú Vantini, Dani Jovino, Helen, Francini, Flávia, Michelli, pela amizade, auxílios, boa convivência e por todos os momentos agradáveis que me proporcionaram;

À todos os meus familiares, e aos familiares do meu noivo, por todos os momentos de descontração e muita alegria vividos durante este período, pelas conversas e pelo apoio nos momentos difíceis;

Às amigas de Palotina, Ana, Vanessa, Barbara, Esther, Marlova, Nina, Cíntia, Pri, Aryane, Dara, Milena, Marília, por apoiarem este sonho desde sempre, e acreditarem no meu potencial, sempre me incentivando a seguir meu caminho, e por representarem

o sentido real da palavra AMIGO na minha vida;

À querida amiga Lígia de Oliveira Amaral, que esteve compartilhando um lar comigo nos últimos dois anos, por toda amizade, carinho, e incentivo. Sempre pronta a ouvir, com ótimos conselhos, e me dando todo apoio necessário. Muito obrigada por ser essa pessoa iluminada que você é, e por todos os momentos de descontração. Agradeço demais pela nossa amizade;

À todas as novas amizades feitas em Jaboticabal, da UNESP ou da cidade, que são aquelas pessoas especiais que tornam a vida mais leve, trazendo momentos de descontração e alegria, sou muito grata por ter conhecido cada um;

À Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista “Júlio de Mesquita Filho” (FCAV/UNESP) pela oportunidade de realização do mestrado e pelos cursos e eventos oferecidos, e a todos os professores que contribuíram para esta etapa, proporcionando uma ótima formação;

À todos os alunos, funcionários e professores do Departamento de Tecnologia que de alguma maneira contribuíram para a realização deste trabalho;

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001;

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de estudos (Processo FAPESP n 2018/01326-0);

Agradeço a todos, imensamente.

SUMÁRIO

	Página
1 Introdução	1
2 Revisão de literatura	5
2.1 A cultura da cana-de-açúcar: origem, aspectos gerais e econômicos . .	5
2.2 Déficit hídrico e mecanismos de resposta das plantas	8
2.3 Transcriptoma e Tecnologias de Sequenciamento de Nova Geração (NGS)	12
2.4 RNA-Seq e Montagem <i>de novo</i>	15
2.5 PCR quantitativa em tempo real (RT-qPCR)	18
2.6 Estudos genéticos e moleculares em cana de açúcar na resposta ao déficit hídrico	20
3 Material e Métodos	23
3.1 Experimento em casa de vegetação e material vegetal	23
3.2 Extração, quantificação e análise de integridade do RNA total	24
3.3 Construção das bibliotecas de DNA complementar (cDNA) e sequencia- mento Illumina	25
3.4 Análise dos dados do sequenciamento.	25
3.4.1 Pré-processamento das leituras: remoção de adaptadores e filtros de qualidade.	25
3.4.2 Nova montagem <i>de novo</i> , avaliação do transcriptoma de referência e mapeamento das leituras nos transcritos	25
3.4.3 Análise de expressão diferencial por cultivar, déficit hídrico e tempo	27
3.4.4 Anotação dos transcritos	28
3.4.5 Diferença entre proporções de termos de ontologia e ortólogos .	28
3.5 Seleção dos transcritos-alvo a validação por PCR quantitativo em tempo real (qRT-PCR)	29
3.5.1 Obtenção dos RNAs das folhas de cana-de-açúcar	29
3.5.2 Síntese da Primeira Fita de cDNA	29
3.5.3 Desenho dos Oligonucleotídeos Iniciadores dos transcritos alvos para o RT- qPCR	30
3.5.4 Titulação dos Oligonucleotídeos Iniciadores	31
3.5.5 Curva Padrão e Eficiência de Amplificação da PCR	32
3.5.6 Quantificação Relativa da Expressão Gênica pelo Método do CT Comparativo ($\Delta\Delta$ Ct)	32
4 Resultados e discussão	35

4.1	Qualidade do material, sequenciamento das bibliotecas de cDNA via RNA-Seq	35
4.2	Análise do transcriptoma das duas cultivares, aos 30, 60 e 90 dias de déficit hídrico	35
4.2.1	Processamento das leituras brutas	35
4.2.2	Nova montagem <i>de novo</i> e avaliação do transcriptoma	35
4.2.3	Identificação dos transcritos diferencialmente expressos (DEs) por cultivar, déficit hídrico e tempo de estresse	37
4.2.4	Anotação dos transcritos	40
4.2.5	Categorização dos transcritos DEs das cultivares a partir de ortólogos (KEGG) e ontologia (GO)	41
4.2.6	Enzimas diferencialmente expressas no metabolismo da sacarose envolvidas na resposta ao déficit hídrico prolongado	50
4.2.7	Relação dos transcritos DEs na resposta ao déficit hídrico prolongado	53
4.3	Validação dos resultados do RNA-Seq pela técnica de PCR quantitativa em tempo real (RT-qPCR)	57
5	Conclusões	65
6	Referências bibliográficas	67

DINÂMICA DO PERFIL TRANSCRICIONAL DE DUAS CULTIVARES DE CANA-DE-AÇÚCAR CONTRASTANTES À SECA E SUBMETIDAS AO DÉFICIT HÍDRICO PROLONGADO

RESUMO – Períodos prolongados de seca têm se tornado mais frequentes em algumas regiões do Brasil. Além disso, a expansão da cultura de cana-de-açúcar para regiões com deficiência hídrica prolongada torna a produção sucroenergética limitada nestes locais. A melhor forma de contornar esse problema é utilizar cultivares tolerantes a este estresse. Neste trabalho, o perfil de expressão gênica da cana-de-açúcar foi avaliado, a partir de duas cultivares com respostas contrastantes ao déficit hídrico: uma delas com comportamento considerado tolerante (SP81-3250), e a outra altamente exigente em água (RB855453), considerada sensível. Ambas foram submetidas a três potenciais hídricos do solo (controle (sem estresse hídrico), déficit hídrico moderado e déficit hídrico severo) a partir de 60 dias após o plantio. Essas plantas foram avaliadas molecular e fisiologicamente em três épocas distintas: 30, 60 e 90 dias após a aplicação dos tratamentos, sendo este um dos poucos estudos realizados até o momento sobre a resposta de plantas de cana-de-açúcar sob déficit hídrico prolongado, estresse esse que foi realizado no período conhecido como fase de formação da cana-de-açúcar, compreendendo o período mais crítico por demanda de água. A análise global da expressão gênica através da tecnologia de RNA-Seq mostrou alterações significativas em resposta ao déficit hídrico entre as duas cultivares. Os transcritos do genótipo tolerante apresentaram uma maior capacidade de reação das plantas frente ao déficit hídrico prolongado, enquanto que no genótipo sensível, diversos mecanismos de sobrevivência da planta foram reprimidos. A cultivar tolerante apresentou indução de transcritos associados ao transporte e metabolismo de carboidratos, transcrição e produção de energia. No estresse mais severo, a cultivar tolerante apresentou, em contraste à sensível, a indução do acúmulo de sacarose por diversos caminhos, mostrando uma conservação desse açúcar apesar do estresse. Além disso, na análise de expressão diferencial das enzimas do metabolismo da glutathione, cujo papel importante em condições de estresse abiótico tem sido amplamente discutido nas mais diferentes plantas, observou-se uma maior indução das enzimas de produção da glutathione na cultivar tolerante em comparação à sensível. Foram detectados alguns transcritos diferencialmente expressos na condição de seca que não receberam anotação, os quais podem ser estudados para análise do seu papel na tolerância ao déficit hídrico. Os resultados encontrados trazem novas informações para a compreensão e melhoramento direcionado para o aumento da tolerância ao déficit hídrico prolongado em cana-de-açúcar.

Palavras-chave: Cana-de-açúcar, déficit hídrico prolongado, expressão diferencial, montagem *de novo*, RNA-Seq, RT-qPCR

TRANSCRIPTIONAL PROFILE DYNAMICS OF TWO CONTRASTING SUGARCANE CULTIVARS IN RESPONSE TO DROUGHT STRESS SUBMITTED TO PROLONGED WATER DEFICIT

ABSTRACT – Extended periods of drought have become more frequent in some regions of Brazil. In addition, the expansion of sugarcane cultivation to regions with prolonged water deficiency makes sugarcane production limited at these locations. The best way around this problem is to use stress-tolerant cultivars. In this work, the sugarcane gene expression profile was evaluated from two cultivars with contrasting responses to water deficit: one of them with tolerant behavior (SP81-3250), and the other highly demanding in water (RB855453), considered sensitive. Both were submitted to three soil water potentials (control (without water stress), moderate water deficit and severe water deficit) from 60 days after planting. These plants were evaluated molecularly and physiologically at three different times: 30, 60 and 90 days after the application of the treatments. This is one of the few studies carried out so far on the response of sugarcane plants under prolonged water deficit. This stress was realized during the period known as sugarcane formation phase, comprising the most critical period by water demand. The overall analysis of gene expression through RNA-Seq technology showed significant changes in response to water deficit between the two cultivars. The transcripts of the tolerant genotype showed a higher reaction capacity of the plants to prolonged water deficit, while in the sensitive genotype, several plant survival mechanisms were repressed. The tolerant cultivar presented induction of transcripts associated with carbohydrate transport and metabolism, transcription and energy production. In the most severe stress, the tolerant cultivar showed, in contrast to the sensitive one, the induction of sucrose accumulation by several ways, showing a conservation of this sugar despite the stress. Moreover, in the differential expression analysis of glutathione metabolism enzymes, whose important role in abiotic stress conditions has been widely discussed in the most different plants, there was a greater induction of glutathione production enzymes in tolerant cultivar compared to sensitive. Some unannotated differentially expressed transcripts were detected and could be studied to analyze their role in tolerance to water deficit. The results bring new information for understanding and improvement aimed at increasing tolerance to prolonged water deficit in sugarcane.

Keywords: Sugarcane, prolonged water deficit, differential expression, *de novo* assembly, RNA-Seq, RT-qPCR

LISTA DE TABELAS

		Página
1	Descritivos agronômicos das cultivares utilizadas no estudo. Adaptado de Marin (2019).	24
2	Sequência dos oligonucleotídeos iniciadores utilizadas na RT-qPCR.	31
3	Parâmetros e estatísticas da montagem <i>de novo</i> do transcriptoma de folhas de cana-de-açúcar.	36
4	Principais métricas de avaliação das montagens <i>de novo</i> do transcriptoma.	36
5	Anotações dos transcritos por referência.	40
6	Perfil de expressão ao longo do tempo dos transcritos DEs (p-valor ajustado < 0,1) em ambas as cultivares e níveis de déficit hídrico, que receberam termos do KEGG Pathways da rota do metabolismo da glutatona (ko00480).	45
7	Número de transcritos DEs na cultivar tolerante sob déficit hídrico moderado e severo, em relação ao controle, que receberam anotação nos mecanismos de resposta à seca do DroughtDB (Alter et al., 2015) (p-valor < 0,05).	54
8	Número de transcritos DEs na cultivar sensível sob déficit hídrico moderado e severo, em relação ao controle, que receberam anotação nos mecanismos de resposta à seca do DroughtDB (Alter et al., 2015) (p-valor < 0,05).	55
9	Resultados obtidos através da leitura das amostras de cDNA no equipamento NanoDrop.	58
10	Concentrações dos oligonucleotídeos dos transcritos-alvo e genes de referência usados para validação dos resultados do RNA-Seq por meio de RT-qPCR.	58

LISTA DE FIGURAS

		Página
1	Fases do desenvolvimento da cana-de-açúcar. Fonte: Gascho e Shih (1983).	6
2	Produção de cana-de-açúcar, etanol e açúcar por estado no Brasil (2019).	8
3	Via genérica de transdução de sinal em resposta à sinalização de estresse abiótico.	10
4	Funções dos genes induzidos pelo estresse hídrico na resposta e tolerância ao estresse. Fonte: (Shinozaki, Yamaguchi-Shinozaki, 2007), citado em Ferro (2008).	12
5	Fluxograma básico para obtenção de um transcriptoma utilizando a plataforma NGS. Adaptado de Pinto et al. (2011).	17
6	Características das curvas de amplificação da qRT-PCR, que traçam o sinal de fluorescência versus o número do ciclo. Fonte: Bustin e Mueller (2005).	19
7	Esquema utilizado para o delineamento experimental, considerando todas as avaliações realizadas.	23
8	Agrupamento (análise de componentes principais) dos transcritos por amostra, com base na contagem de expressão normalizada com o programa RSEM.	26
9	Diagrama de Venn do número de transcritos diferencialmente expressos na cultivar de cana-de-açúcar tolerante sob déficit hídrico moderado (A) e severo (B), aos 30, 60 e 90 dias após o início dos tratamentos (p-valor ajustado < 0.1).	38
10	Diagrama de Venn do número de transcritos diferencialmente expressos na cultivar sensível sob déficit hídrico moderado (A) e severo (B), aos 30, 60 e 90 dias após o início dos tratamentos (p-valor ajustado < 0.1).	38
11	Mapa de calor para termos de ontologia gênica (GO) relativos a processos biológicos, função molecular e componente celular, atribuídos aos transcritos de folhas de cana-de-açúcar diferencialmente expressos sob déficit hídrico moderado em comparação à ausência de estresse, em ao menos um tempo (p-valor ajustado < 0,05).	41
12	Mapa de calor para termos de ontologia gênica (GO) relativos a processos biológicos, atribuídos aos transcritos de folhas de cana-de-açúcar diferencialmente expressos sob déficit hídrico severo, em comparação à ausência de estresse, em ao menos um tempo (p-valor ajustado < 0,05).	43

13	Transcritos diferencialmente expressos na cultivar tolerante, entre as plantas controle e sob estresse severo, agrupados por rotas metabólicas do KEGG (p -valor ajustado $<0,1$).	44
14	Transcritos que apresentaram expressão diferencial na cana-de-açúcar sob déficit hídrico na rota do metabolismo da Glutaciona (ko00480), marcados em vermelho. Somente os significativos (p -valor ajustado $<0,1$). Fonte: KEGG Pathways.	47
15	Comparação dos metabolismos entre as cultivares através dos transcritos DEs exclusivos da cultivar tolerante (somente aqueles com p -valor ajustado $<0,05$).	49
16	Enzimas (número EC.) DEs nas cultivares tolerante e sensível sob déficit hídrico moderado e severo, mapeadas na via de metabolismo da sacarose (ko00500), adaptadas do KEGG Pathways (p -valor $<0,05$).	51
17	Eletoferograma e valor de RIN dos "pools" de RNA total das folhas de cana-de-açúcar da cultivar tolerante, em todos os tempos e níveis de déficit hídrico, geradas pelo Bioanalyzer (Agilent 2100 Bioanalyzer).	57
18	Curva de "melting" dos genes 14-3-3, Actina, β -Tubulina, Peroxidase extracelular, Glutaciona S-transferase, Poligalacturonase e Receptor do tipo quinase (GbRLK).	59
19	Curva de diluição dos genes endógenos.	60
20	Curva de diluição dos transcritos-alvo.	61
21	Comparação dos valores de expressão, obtidos por RNA-Seq e RT-qPCR em dois transcritos diferencialmente expressos.	62
22	Comparação dos valores de expressão diferencial do transcrito Receptor do tipo quinase, obtidos por RNA-Seq e RT-qPCR.	62
23	Comparação dos valores de expressão diferencial do transcrito Glutaciona S-transferase, obtidos por RNA-Seq e RT-qPCR.	63

Introdução

Estresses ambientais como a seca limitam o crescimento das plantas, e estão entre os principais fatores que causam perdas na produção de diversas culturas no mundo todo (Mahajan, Tuteja, 2005; Ferreira et al., 2017), sendo que a produtividade total chega a redução de mais de 50% (Rodríguez et al., 2006). Essas perdas têm se tornado grande preocupação para todos os países, frente à crescente necessidade alimentar com o aumento alarmante da população mundial (Mahajan, Tuteja, 2005).

A cana-de-açúcar pertence ao gênero *Saccharum*, e sua base genética é composta de híbridos derivados de *S. officinarum*, *S. sinense*, *S. barberi* e *S. spontaneum* (Menossi et al., 2008; Li et al., 2016). Ela é uma importante cultura de áreas tropicais e subtropicais em todo o mundo, e, como toda cultura agrícola, está vulnerável a quaisquer condições ambientais adversas, sendo a seca é um dos principais fatores abióticos responsável pela diminuição de sua produtividade (Basnayake et al., 2012; Marchiori, 2014). Assim, torna-se cada vez mais importante a obtenção de cultivares capazes de tolerar/resistir à seca, ou utilizar a água mais eficientemente em períodos de estresse, visando o máximo de desempenho (Li et al., 2016).

A tolerância das plantas à seca trata-se de uma característica quantitativa complexa, regulada por muitos genes (Hassan et al., 2015) que, por diversos mecanismos, trabalham isoladamente ou em conjunto, para gerar uma resposta fisiológica complexa de defesa contra o déficit hídrico. Em síntese, uma combinação de eventos bioquímicos e moleculares, ativados pela percepção do sinal de estresse, geram essa resposta (Nepomuceno et al., 2001; Zhu, 2001; Rodríguez et al., 2006).

Em resposta à limitação hídrica e ambientes extremos, as plantas desenvolveram diferentes estratégias adaptativas que as permitiram manter o seu metabolismo (Mitra, 2001; Cia et al., 2012; Li et al., 2016; Ferreira et al., 2017). As estratégias iniciais incluem modificações morfológicas e em processos metabólicos, celulares e fisiológicos (Li et al., 2016). Além disso, sob condições de estresse, as plantas podem aumentar ou diminuir a expressão de determinados genes e proteínas relacionados à tolerância.

Sistemas convencionais de melhoramento de culturas são limitados para a cana-de-açúcar, em função da elevada complexidade genética devido à sua natureza poliploide

e aneuploide (Menossi et al., 2008; Li et al., 2016) e complexidade das características de tolerância ao estresse (Rodriguez et al., 2005). Isolar e identificar genes diferencialmente expressos em condições de déficit hídrico e as vias metabólicas relacionadas, são essenciais para o desenvolvimento de cultivares tolerantes. Assim, o estudo da dinâmica do perfil de expressão gênica se tornou uma ferramenta importante para investigar, em nível transcricional, as respostas de um organismo frente às mudanças ambientais (Molina et al., 2008; Hassan et al., 2015; Li et al., 2016).

Um dos recursos mais poderosos para descrever e compreender as respostas genéticas das plantas frente ao estresse ambiental, é o estudo do transcriptoma (Lestari et al., 2006; Ambrosone et al., 2013). Entre as técnicas moleculares utilizadas para estes estudos, está a técnica de RNA-Seq, cuja característica mais atraente é o fato de permitir o mapeamento e quantificação de transcritos diferencialmente expressos, independente de conhecimento do genoma completo da cultura (Rodrigues, 2011; Belesini, 2015). O sequenciamento de RNA (RNA-Seq) tem outras vantagens, incluindo a detecção mais sensível dos níveis de transcritos quando comparado à outros métodos; também por ser quantitativo, permitindo a obtenção do perfil global dos genes expressos; o potencial para detectar leituras contendo modificações pós-transcricionais ou sequências rearranjadas que não podem ser mapeadas diretamente no genoma, entre outras (Marguerat, Bähler, 2010; Cardoso-Silva et al., 2014).

Entre as técnicas utilizadas para a análise quantitativa da expressão gênica, para a validação de resultados obtidos em experimentos com RNA-Seq, está a técnica de RT-qPCR (PCR quantitativo em tempo real). Ela é bastante precisa para quantificar níveis de expressão de um determinado gene, bem como oferece resultados rápidos e dados reprodutíveis (Stolf, 2007). Esta técnica tem sido utilizada com sucesso na validação da expressão gênica em cana-de-açúcar tolerante ao déficit hídrico (Vantini, 2013).

Na literatura, vários trabalhos são encontrados envolvendo déficit hídrico severo aplicado na cana-de-açúcar por um curto período de tempo para a identificação de genes relacionados à tolerância ao estresse (Gupta et al., 2010; Vargas et al., 2014; Vantini et al., 2015; Li et al., 2016). No entanto, esses experimentos podem levar a resultados pouco conclusivos em relação à sensibilidade ou tolerância à seca das plantas em condições reais de campo, onde a restrição de água geralmente dura vários meses.

Estudos anteriores realizados pelo nosso grupo de pesquisa foram alguns dos pioneiros a investigar déficit hídrico prolongado na cana-de-açúcar. Primeiro, foram investigados os parâmetros fisiológicos de cada cultivar em resposta aos tratamentos - tempo e nível de estresse hídrico, encontrando diferenças significativas principalmente

na associação de CO₂ e alterações na fisiologia das folhas. A seguir, foi feita a avaliação e detecção dos perfis globais de transcrição (Belesini et al., 2017; Telles et al., 2018), isto é, quais aspectos gerais foram alterados entre as cultivares.

Diante disso, este trabalho tem por objetivo dar continuidade à investigação do déficit hídrico prolongado na cana-de-açúcar, através da mineração dos dados obtidos pela tecnologia de RNA-Seq a partir de folhas de duas cultivares com respostas contrastantes ao déficit hídrico, descritas como tolerante (SP81-3250) e sensível (RB855453). As cultivares foram submetidas à um déficit hídrico prolongado, sob dois níveis de tensão do solo (estresse moderado e estresse severo) e as avaliações foram realizadas aos 30, 60 e 90 dias após a aplicação do estresse, a fim de identificar transcritos diferencialmente expressos (DEs) envolvidos no caráter de tolerância e/ou sensibilidade ao déficit hídrico, revelando vias metabólicas e processos biológicos associados.

Nossas hipóteses são que as cultivares respondem de maneira diferente frente ao déficit hídrico ao longo do tempo, possuindo perfis únicos de resposta. Entre as perguntas a serem respondidas estão: quais são os processos biológicos mais evidentes nas respostas das duas cultivares? Quais rotas metabólicas estão favorecidas? O que está modificado no metabolismo de carboidratos, especialmente da sacarose? Como se distribuem todos os transcritos DEs com anotação no DroughtDB? Para a validação dos resultados de expressão obtidos pelo RNA-Seq, foi utilizada a técnica de RT-qPCR.

Revisão de literatura

2.1 A cultura da cana-de-açúcar: origem, aspectos gerais e econômicos

A cana-de-açúcar (*Saccharum* spp.) é uma monocotiledônea que pertence à divisão Magnoliophyta, classe Liliopsida, ordem Cyperales, família Poaceae, bem como o arroz, milho e sorgo (Tzvelev, 1989; Palhares, 2010; EMBRAPA, 2019). O centro de origem desta cultura fica na região da Nova Guiné, de onde espalhou-se em várias linhas ao sul do Oceano Pacífico, aparecendo como planta produtora de açúcar na Índia (Delgado, Cesar, 1977; Silva et al., 2010). Entretanto, foram as terras da América que ofereceram à cana-de-açúcar condições favoráveis para o seu desenvolvimento, sendo introduzida no Brasil já nos anos seguintes à época do descobrimento (Mozambani et al., 2006 apud Telles, 2016). Desde então, o Brasil se destaca entre os maiores produtores de cana-de-açúcar do mundo, e também dos subprodutos açúcar e etanol.

Esta cultura, que pertence ao gênero *Saccharum*, tem sua base genética composta por híbridos derivados de *S. officinarum* (clones nobres), *S. sinense* (clones chineses), *S. barberi* (clones do norte da Índia) e *S. spontaneum* (Menossi et al., 2008; Li et al., 2016). Os cruzamentos interespecíficos realizados na Índia, buscando melhoramento para uma cana com maior capacidade de armazenamento de sacarose, e também com características de rusticidade, vigor, resistência e tolerância a doenças e a fatores abióticos, resultaram em uma composição genética bastante complexa das espécies dessa cultura (Ferro, 2008). Plantas do gênero *Saccharum* apresentam altos níveis de ploidia e algumas apresentam número atípicos de cromossomos (aneuploidia), sendo que a espécie *S. officinarum*, espécie produtora de açúcar, é octaplóide ($2n = 80$) e apresenta $x=10$ cromossomos básicos (D'Hont et al., 1995; D'Hont et al., 1996), enquanto *S. spontaneum*, espécie silvestre que possui grande variabilidade, apresenta $2n$ de 40 à 128, e possui $x=8$ cromossomos básicos (D'Hont et al., 1996). As cultivares modernas são híbridos derivados do cruzamento dessas duas espécies.

Em relação ao crescimento da planta, a cana-de-açúcar apresenta quatro estágios fenológicos (Figura 1). O primeiro deles, brotação e emergência, depende da qualidade da muda, ambiente, época e manejo do plantio. Neste estágio ocorre o enraizamento

inicial e o aparecimento das primeiras folhas. Na segunda fase, de perfilhamento, ocorre a emissão de colmos por uma mesma planta, que recebem a denominação de perfilhos. Através do desenvolvimento das raízes de cada perfilho ocorre a formação do sistema radicular de touceira da cana-de-açúcar. O crescimento dos colmos é caracterizado pelo crescimento dos entrenós e desenvolvimento dos mesmos, ganhando altura e começando a acumular açúcar na base. É estimulado por luz, umidade e calor. Além disso, o crescimento radicular torna-se mais intenso. Por fim, o último estágio denominado maturação, caracteriza-se pelo intenso armazenamento de açúcar nos colmos (EMBRAPA, 2019).

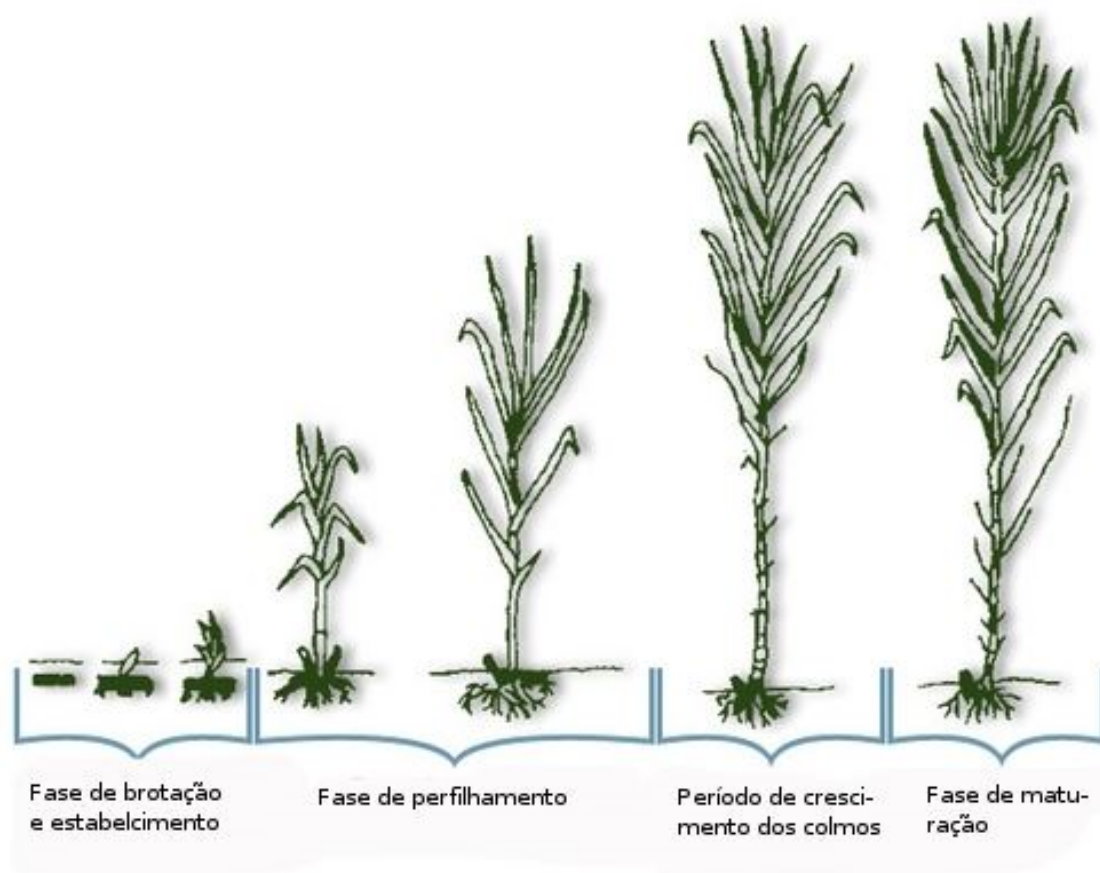


Figura 1: Fases do desenvolvimento da cana-de-açúcar. Fonte: Gascho e Shih (1983).

A cana-de-açúcar possui alto poder de absorção e fixação de gás carbônico, por ter comportamento fisiológico do tipo C4, caracterizado por possuir máxima eficiência em converter energia radiante em energia química (Paula et al., 2010). Entretanto, a eficiência fotossintética da planta é influenciada pelas características das cultivares, além das variações climáticas durante o seu desenvolvimento (Rodrigues, 1995).

As formas cultivadas desta gramínea semi-perene produzem colmos de alguns metros de altura. O plantio da cana-de-açúcar pode ser realizado em diferentes épocas

do ano, e por isso seu ciclo evolutivo pode ser de 12 e 18 meses, cana de ano e cana de ano e meio, respectivamente (Freitas, 2011; Scarpate, 2011), proporcionando de cinco a seis cortes a cada plantio, sendo a primeira colheita proveniente do plantio a partir das mudas (cana-planta) e as outras colheitas feitas a partir da rebrota da cana-planta (cana-soca) (Freitas, 2011).

O Brasil, além de ser o maior produtor mundial de cana, é historicamente o maior produtor de açúcar no mundo, responsável por 20% da produção global e 45% da exportação mundial (UNICA, 2019a). Segundo o último levantamento publicado em abril pela Conab (2019), a área colhida de cana-de-açúcar no Brasil foi de aproximadamente 8,5 milhões de hectares na safra 2018/2019, apresentando uma diminuição de 1,6% se comparada a 2017/2018, sendo que a produção foi de 620 milhões de toneladas, apresentando redução de 2% em relação à safra anterior, mas ainda assim, garantindo ao país a posição de maior produtor mundial desta cultura.

A produção se concentra principalmente na região Centro-Sul do país, sendo que o estado de São Paulo lidera a produção, seguido pelo estado de Goiás. A Figura 2 mostra os valores de produção de cana-de-açúcar, açúcar e etanol da safra de 2018/2019 por estado no Brasil, segundo dados oficiais reunidos pela UNICA (UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR)(UNICA, 2019b).

Além da importância econômica e social da cana-de-açúcar para o Brasil pela produção do açúcar, fabricação de fármacos, ração animal e fertilizantes orgânicos, ela apresenta-se como uma alternativa ao uso do petróleo, tanto pela produção de etanol combustível, como pela co-geração de energia elétrica limpa, a partir da biomassa (bagaço e outros resíduos) gerada nos processos de obtenção dos subprodutos da cana (Vantini, 2013; Marin, 2019; CONAB, 2019). De acordo com os dados apresentados, nota-se a grande importância do setor sucroalcooleiro na sociedade brasileira, pela produção e exportações de açúcar e etanol, que contribuem com o equilíbrio da balança comercial, e também pelo potencial na geração de empregos diretos e indiretos (Marques et al., 2006).

Sendo assim, para que os altos níveis de produção sejam mantidos, a cultura da cana-de-açúcar tem se expandido no Brasil e em outros países do mundo, para regiões com condições edafoclimáticas não muito favoráveis ao seu desenvolvimento e produção (Ferro, 2008).

Considerando que, a restrição hídrica é um dos principais fatores que limitam a produção dessa cultura (Basnayake et al., 2012), nas últimas décadas os melhoristas vem tentando compreender o processo de tolerância através do estudo de genes diferencialmente expressos durante condições de estresse, afim de identificar genes e vias relacionadas à respostas ao estresse, através de técnicas moleculares modernas, para

que sejam criadas novas variedades que consigam usar a água mais eficientemente ou que sejam capazes de tolerar/resistir à seca (Li et al., 2016).

Estados	Cana-de-açúcar (mil toneladas)	Açúcar (mil toneladas)	Etanol Anidro (mil m ³)	Etanol Hidratado (mil m ³)	Etanol Total (mil m ³)
Acre	0	0	0	0	0
Alagoas	16.202	1.195	180	310	490
Amazonas	237	13	0	5	5
Bahia	3.743	108	80	165	246
Ceará	0	0	0	0	0
Espírito Santo	3.174	147	116	21	137
Goiás	70.001	1.670	988	3.904	4.892
Maranhão	1.964	22	123	25	148
Mato Grosso	17.349	370	657	1.148	1.804
Mato Grosso do Sul	49.505	944	793	2.483	3.276
Minas Gerais	63.063	3.063	902	2.322	3.224
Pará	999	47	36	10	46
Paraíba	5.589	118	153	229	382
Paraná	35.641	2.122	484	1.131	1.615
Pernambuco	11.425	733	80	352	432
Piauí	1.167	78	18	19	37
Rio de Janeiro	1.058	21	0	58	58
Rio Grande do Norte	2.429	118	17	88	105
Rio Grande do Sul	39	0	0	2	2
Rondônia	31	0	0	2	2
Santa Catarina	0	0	0	0	0
São Paulo	333.294	18.172	5.202	10.743	15.944
Sergipe	1.896	99	19	82	101
Tocantins	2.024	0	70	86	155
Região Centro-Sul	573.125	26.510	9.141	21.812	30.953
Região Norte-Nordeste	47.707	2.530	775	1.374	2.150
Brasil	620.832	29.040	9.917	23.186	33.103

Figura 2: Produção de cana-de-açúcar, etanol e açúcar por estado no Brasil. Fonte: ÚNICA, ALCOPAR, BIOSUL, SIAMIG, SINDALCOOL, SIFAEG, SINDAAF, SUDES e MAPA (2019).

2.2 Déficit hídrico e mecanismos de resposta das plantas

No ambiente onde são cultivadas, as plantas estão frequentemente submetidas a condições externas desfavoráveis, gerando estresses que afetam negativamente o seu crescimento, desenvolvimento e/ou produtividade (Bianchi et al., 2016).

Estresse é a condição fisiológica alterada, provocada por fatores que tendem a causar desequilíbrios nas plantas (SANTANA, 2011). Estes podem ser de origem biótica ou abiótica, podendo ocorrer concomitantemente. Entre os estresses abióticos (ambientais), estão: frio, elevadas temperaturas, salinidade, seca (déficit hídrico), inundações, deficiência ou excesso de nutrientes minerais, metais pesados e outros (Wang et al., 2003; Mahajan, Tuteja, 2005; Souza et al., 2012).

Segundo Rodríguez et al. (2006), o estresse abiótico é a principal causa de perda de colheitas no mundo, reduzindo a produtividade em mais de 50% para as principais culturas. Essas perdas têm se tornado grande preocupação para todos os países frente ao aumento alarmante da população mundial, e conseqüentemente à crescente necessidade alimentar (Mahajan, Tuteja, 2005).

Dentre os estresses abióticos, o déficit hídrico é um dos principais responsáveis pela diminuição da produtividade nas lavouras (Basnayake et al., 2012), incluindo a cana-de-açúcar (Inman-Bamber, Smith, 2005). Taiz e Zeiger (2004) definiram déficit hídrico como todo conteúdo de água de uma célula ou tecido que está abaixo do conteúdo de água presente no estado de maior hidratação.

Quando as mudanças ocasionadas pelo estresse ocorrem na planta, o sistema radicular é quem faz inicialmente a detecção dessas mudanças e sinaliza seus efeitos às demais células, tecidos e órgãos (Wang et al., 2003; Taiz, Zeiger, 2004; Graça, 2009). O sistema radicular percebe a falta de água no solo e desencadeia a emissão de sinais químicos que promovem uma série de eventos de ordem molecular, como modificações no comportamento estomático durante o déficit hídrico, com várias respostas fisiológicas, metabólicas e de desenvolvimento (Bray, 1997; Wilkinson, Davies, 2002; Graça, 2009).

Os efeitos gerais da seca, principalmente os morfológicos, no crescimento das plantas, estão bem descritos na literatura. Contudo, os efeitos primários do déficit hídrico a nível bioquímico e molecular nas plantas ainda não foram completamente entendidos, uma vez que interações complexas resultantes de mais de um estresse podem afetar a planta em um mesmo período (Mittler et al., 2001; Chaves et al., 2003).

As plantas possuem uma complexa via de respostas ao déficit hídrico. As respostas a nível celular iniciam com a interação entre componentes extracelulares, denominados ligantes ou elicitores, e proteínas da membrana plasmática, denominadas receptoras, que fazem ligação e interagem com a molécula extracelular (Mahajan, Tuteja, 2005). Neste caso, o estresse hídrico serve como elicitor para as células da planta. Basicamente, no momento em que a perda de água da célula é percebida, uma via de transdução de sinal celular é desencadeada para ativar genes específicos, convertendo o estresse físico em uma resposta bioquímica (Bray, 1997).

O sinal celular emitido resulta na geração de mensageiros secundários como o cálcio, espécies reativas de oxigênio (ERO) e fosfatos de inositol. Esses mensageiros modulam o nível de cálcio intracelular, e essa variação é percebida por proteínas de ligação ao cálcio (sensores de Ca^{2+}), as quais interagem com outras proteínas, possivelmente iniciando uma cascata de fosforilação que resulta na ativação de genes responsivos ao estresse hídrico, ou de fatores de transcrição que controlam/ativam

esses genes, levando à respostas fisiológicas (Figura 3). Os produtos gênicos levam a planta à se adaptar e sobreviver superando as condições desfavoráveis (Mahajan, Tuteja, 2005).

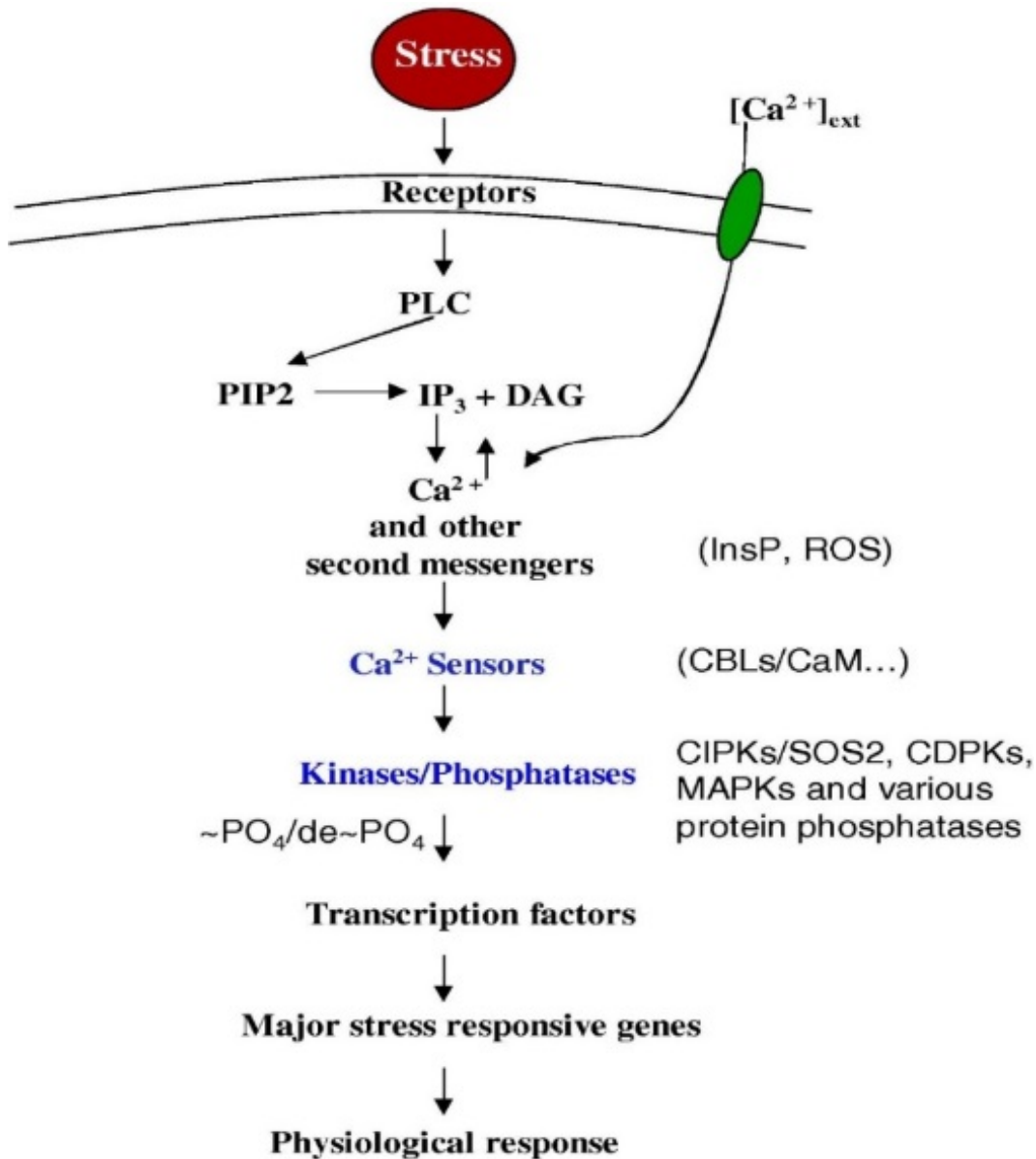


Figura 3: Via genérica de transdução de sinal em resposta à sinalização de estresse abiótico. Representa a visão geral da via de sinalização sob condições de estresse. Adaptado de Mahajan e Tuteja (2005).

Acredita-se que existam vários mecanismos de sinalização diferentes, uma vez que nem todos os genes são ativados pelo estresse (Bray, 1997). Além de ser espécie-específicas e genótipo-dependentes (Hall et al., 1982; Lopes et al., 2011), as respostas das plantas à seca estão relacionadas com a duração e severidade da exposição ao

estresse (Pinheiro, Chaves, 2010), idade e estágio fenológico no período de seca, bem como do órgão e tipo celular (Chaves et al., 2003; Gao et al., 2013).

Segundo Mahajan e Tuteja (2005), os genes responsivos ao estresse podem ser categorizados como genes induzidos precocemente e genes tardios. O genes precoces normalmente se expressam de forma transitória, minutos após a percepção do sinal de estresse, e incluem sensores de cálcio e vários fatores de transcrição, pois a indução dos mesmos não requer a síntese de novas proteínas. Já os genes induzidos tardiamente são aqueles que são ativados pelo estresse horas depois, e a expressão destes é frequentemente mantida. Estes genes incluem os principais genes responsivos ao estresse.

Entre os principais mecanismos de sinalização que operam durante o déficit hídrico, está a produção do ácido abscísico (ABA), fitohormônio que regula a expressão de vários genes relacionados ao estresse e leva ao fechamento dos estômatos. Porém, nem todos os genes induzidos pelo déficit hídrico são regulados pelo ABA (Bray, 1997; Mahajan, Tuteja, 2005), classificando os sinais de percepção e transmissão do sinal de estresse que controlam a expressão de genes induzidos pela seca em ABA-dependentes e ABA-independentes (Yamaguchi-Shinozaki, Shinozaki, 2005).

Segundo Shinozaki et al. (2003), os produtos dos genes induzidos por seca, identificados em *Arabidopsis thaliana*, podem ser classificados em dois grupos: proteínas funcionais e proteínas reguladoras (Figura 4).

O primeiro grupo inclui proteínas que provavelmente funcionam na tolerância ao estresse (proteínas funcionais), tais como chaperonas, osmotina, proteínas LEA (“Late Embryogenesis Abundant proteins”), proteínas de canais de água, enzimas detoxificadoras como as antioxidantes, proteases, enzimas chave na biossíntese de vários osmoprotetores como a prolina, transportadores de açúcar, entre outras.

O segundo grupo é constituído pelas proteínas reguladoras, que atuam na transdução de sinais e na regulação da expressão de genes responsivos ao estresse. Este grupo inclui vários fatores de transcrição, proteínas quinases, proteínas fosfatases, enzimas envolvidas no metabolismo de fosfolípidos e outras moléculas sinalizadoras tais como proteínas ligantes de calmodulina (Shinozaki et al., 2003; Shinozaki, Yamaguchi-Shinozaki, 2007; Ferro, 2008).

Os efeitos do estresse hídrico sobre as plantas incluem a redução nas taxas de assimilação de CO₂, crescimento, transpiração e abertura estomática, além da diminuição do tamanho das células foliares e do potencial de água na planta (Taiz, Zeiger, 2004; Shinozaki, Yamaguchi-Shinozaki, 2007).

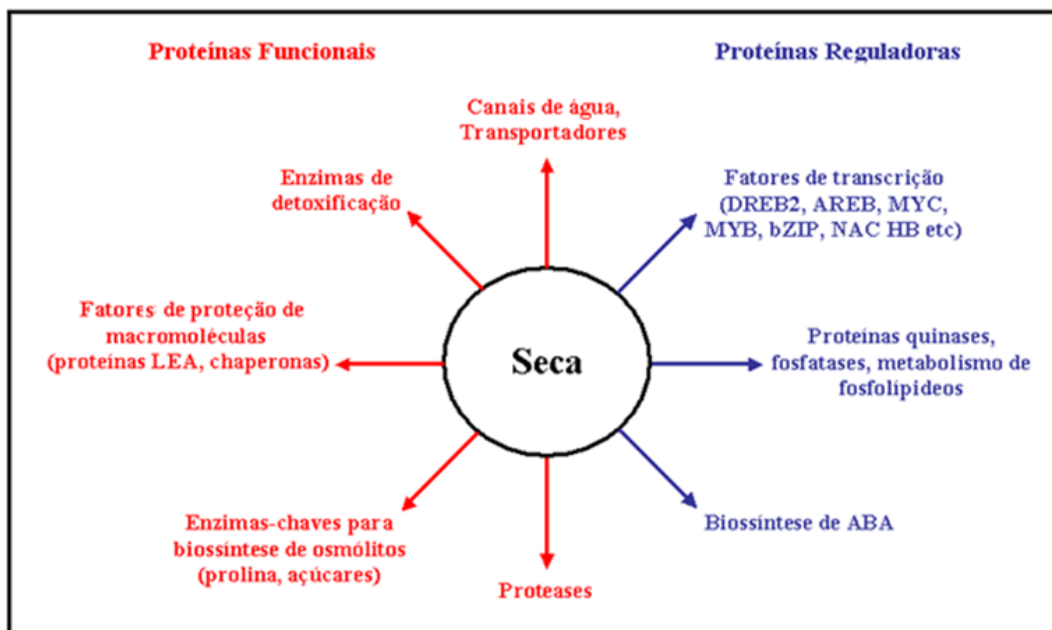


Figura 4: Funções dos genes induzidos pelo estresse hídrico na resposta e tolerância ao estresse. Produtos gênicos classificados em dois grupos: proteínas funcionais e proteínas reguladoras. Fonte: (Shinozaki, Yamaguchi-Shinozaki, 2007), citado em Ferro (2008).

Dessa maneira, o estresse hídrico afeta a produtividade final, pois influencia diretamente no crescimento dos perfilhos e na altura final dos colmos da cana-de-açúcar (Silva, Costa, 2004; Inman-Bamber, Smith, 2005; Ferro, 2008). Sendo assim, muitos esforços tem sido realizados para melhorar a produtividade dos cultivos agrícolas em condições de seca (Cattivelli et al., 2008).

2.3 Transcriptoma e Tecnologias de Sequenciamento de Nova Geração (NGS)

A descoberta da estrutura dupla hélice do DNA por Watson et al. (1953), que permitiu o avanço no entendimento dos fenômenos que ocorrem nos organismos vivos, abriu a possibilidade de uma nova estratégia para os estudos genéticos: a genômica (Watson et al., 1953; Collins et al., 2003). Esta ciência, que estuda os genomas, está dividida em genômica estrutural e funcional (Faleiro et al., 2011).

Considerando que o genoma é um elemento virtual estático, ou seja, em um mesmo indivíduo, todas as células, salvo exceções, tem a mesma estrutura do genoma, o estudo da estrutura deles é feito pela genômica estrutural. Para que a mensagem contida no genoma (DNA) seja convertida em uma função gênica, é necessário a transcrição dos genes em um outro tipo de molécula, o RNA, que será traduzido em proteína (Nelson et al., 2004; Furlan et al., 2007; Costa, 2011).

Desta maneira, a genômica funcional tem contribuído para a compreensão do papel dos genes e das proteínas no funcionamento dos organismos, pois neste tipo de estudo

obtem-se as informações do que está sendo expresso nas células e tecidos (RNAs e proteínas) (Furlan et al., 2007; Costa, 2011).

O RNA é considerado uma “ponte” de transferência de informações biológicas entre o DNA e as proteínas (Costa et al., 2010). Segundo Biazuzo (2013), o processo pelo qual a informação biológica contida em um gene é processada em um produto gênico funcional, é denominado expressão gênica. A transformação de DNA em RNA (transcrição) é o início da expressão.

O termo transcriptoma representa o conjunto de todos os RNAs – transcritos – de um determinado organismo, codificados pelos genes em uma dada célula, que respondem fenotipicamente à condição que levou a transcrição, e que são avaliados qualitativa e quantitativamente em determinados estágios específicos de desenvolvimento, ou durante uma condição fisiológica específica que estão submetidos (Wang et al., 2009; Kogenaru et al., 2012).

Assim, um grande número de transcriptomas podem ser gerados em um mesmo organismo, considerando as diferentes condições (fisiológicas ou de desenvolvimento) em que ele se encontra (Faleiro et al., 2011).

Entre várias outras aplicações, o estudo do transcriptoma determina quando e onde os genes são ativados ou reprimidos, em vários tipos de células e tecidos (Adams, 2008). Desta maneira, trata-se de um importante recurso para analisar as respostas das plantas frente a uma situação de estresse, pois, através dele, pode-se quantificar número de transcritos e compreender a dinâmica de expressão de cada um deles ao longo do desenvolvimento e sob diferentes condições (Lestari et al., 2006; Wang et al., 2009).

Segundo Giachetto e Higa (2014), para a agricultura, o monitoramento e a identificação das variações nos níveis de expressão gênica permitem identificar vias biológicas alteradas em situações onde diferentes perturbações externas são inseridas no sistema, como por exemplo o estresse causado por uma condição de restrição hídrica.

Diversos grupos de pesquisa têm se dedicado a esclarecer os mecanismos utilizados pelas plantas em resposta ao déficit hídrico. Para estes estudos, diferentes tecnologias foram utilizadas para quantificar a expressão gênica, entre elas: sequenciamento de regiões específicas de fragmentos de cDNA; macro e microarranjos, baseados na hibridização de sondas e análise de fragmentos de cDNA amplificados via PCR (Wang et al., 2009).

Entre as várias tecnologias de análise dos transcriptomas, as técnicas de RNA-seq e de microarranjo de DNA destacam-se como as mais utilizadas na quantificação da expressão de genes em larga escala (Wang et al., 2009; Kogenaru et al., 2012).

Nas abordagens utilizando o sequenciamento de regiões específicas dos fragmentos

de cDNA, há a necessidade de uma grande quantidade de mRNA e, ainda assim, apenas uma parte dos transcritos são analisados, e as diferentes isoformas não são distinguíveis umas das outras, pois uma porção significativa das “tags” curtas geradas nesse método não podem ser exclusivamente mapeadas para o genoma de referência (Wang et al., 2009).

O método baseado em hibridação (microarranjo) é de alto rendimento e relativamente barato. Por anos esta técnica tem sido aplicada com sucesso na expressão simultânea de milhares de genes e na descoberta de genes em grande escala, bem como no rastreamento de polimorfismos e no mapeamento de clones de DNA genômico.

Entretanto, essas tecnologias tem várias limitações, que dificultam a análise da expressão gênica, entre elas a dependência do conhecimento existente sobre a sequência do genoma a ser avaliada; a dificuldade na análise de sequências altamente relacionadas devido à ocorrência de hibridação cruzada (não específica) e a natureza analógica do sinal torna difícil detectar e quantificar com segurança espécies com abundância baixa. Ainda, há grandes desafios no que diz respeito à reprodutibilidade dos resultados entre diferentes laboratórios e entre plataformas (Xiang, Chen, 2000; Shendure, 2008; Wang et al., 2009).

Desta maneira, apesar de inicialmente os microarranjos terem sido fundamentais nas análises dos transcriptomas, o desenvolvimento recente de novos métodos de sequenciamento de DNA de alta produtividade forneceu um novo método de análise dos transcriptomas, o qual supera efetivamente as limitações da abordagem de microarranjos, tendo claras vantagens sobre as abordagens existentes (Shendure, 2008; Wang et al., 2009).

Este método, denominado RNA-Seq, revolucionou os estudos da área da transcriptômica, pois o conhecimento de sequência prévia não é necessário; a quantificação é “digital” e não analógica, fornecendo uma medida mais precisa do nível de transcritos, assim como suas isoformas. A técnica permite ainda uma quantificação acurada da expressão diferencial, possibilitando a geração de novos conhecimentos sobre mecanismos moleculares de uma dada característica de interesse (Shendure, 2008; Giachetto, Higa, 2014).

As tecnologias que são usadas para o sequenciamento de todo transcriptoma, através de bibliotecas de cDNA, RNA-seq, promovem o sequenciamento de DNA em plataformas capazes de gerar informação sobre milhões de fragmentos em uma única corrida, são denominadas tecnologias de nova geração (NGS - do inglês, Next-generation Sequencing) (Jain, 2011; Haas et al., 2012).

Dentre as plataformas NGS disponíveis, as principais são: sequenciamento por ligação da Life Technologies/SOLiD e sequenciamento por síntese da Illumina/Solexa

(Wang et al., 2009; Vliet, 2010; Jain, 2011). As plataformas têm seus prós e contras e, na maioria dos casos, o acesso ou as instalações locais influenciarão a escolha final da tecnologia de sequenciamento (Vliet, 2010).

A plataforma Roche 454 resulta em leituras mais longas (média de 400 pb), que propiciam mais facilmente o mapeamento único em comparação as leituras curtas das tecnologias Illumina e SOLiD; porém, o número de leituras geradas é muito menor. As tecnologias Illumina e SOLiD, por gerarem maior número de leituras, além de tornar mais econômicas, oferecem a vantagem de maior alcance dinâmico de detecção de expressão de transcritos e medição de expressão gênica mais precisa. Melhorias recentes nessas duas tecnologias superaram a desvantagem de menor comprimento de leitura em comparação a plataforma Roche 454, e por isso, o uso desta tecnologia foi descontinuado. Embora as vantagens e desvantagens das plataformas Illumina e SOLiD são semelhantes, o uso da tecnologia Illumina superou. Ela foi comparada ao microarranjo, e foi superior na detecção de genes pouco expressos, novos transcritos e variações gênicas – "splicing alternativo"(Marioni et al., 2008; Jain, 2011). Essa tecnologia tem sido muito utilizada para avaliar transcriptoma em larga escala, que é o caso do RNA-Seq.

2.4 RNA-Seq e Montagem *de novo*

Nas últimas décadas houve um progresso considerável no desenvolvimento de métodos para mapeamento e quantificação dos níveis de expressão gênica em todo o nível do transcriptoma (Wang et al., 2009). Dentre as diferentes técnicas utilizadas para avaliar o transcriptoma, podemos destacar as tecnologias de sequenciamento de nova geração (NGS), em escala muito grande, como é o caso do RNA-Seq.

O termo RNA-Seq tem sido utilizado para representar o transcriptoma revelado por sequenciamento de cDNA através de NGS (Pinto et al., 2011). Esta técnica permite o mapeamento e quantificação de transcriptomas (Rodrigues et al., 2011), e, segundo Ambrosone et al. (2013), essa ferramenta tem sido utilizada para descrever as respostas genéticas das plantas aos estímulos ambientais.

Este método permite o sequenciamento e quantificação de transcritos a uma grande resolução e a um alcance dinâmico que independe do tamanho do transcrito. Entre as principais vantagens deste método está a ausência quase total de ruídos, alta cobertura e grande variedade na detecção, sendo capaz de identificar de uma a numerosas cópias de RNA por célula (Pinto et al., 2011).

Além disso, a detecção dos transcritos não fica restrita somente àqueles correspondentes a uma sequência genômica pré-existente, o que torna o método atrativo para a pesquisa em organismos cujos genomas ainda não foram determinados. Como não

existem etapas de clonagem, o RNA-Seq requer uma quantidade muito menor de amostras de RNA, e uma vez que é um método quantitativo, pode ser usado para determinar os níveis de expressão gênica de maneira mais acurada do que os microarranjos. A técnica é capaz de capturar a dinâmica do transcriptoma através de diferentes tecidos e condições, sem necessidade de que normalizações sofisticadas sejam aplicadas sobre os conjuntos de dados (Wang et al., 2009; Oshlack et al., 2010; Giachetto, Higa, 2014).

Segundo Giachetto e Higa (2014), a tecnologia RNA-Seq também permite a identificação sistemática de SNPs em regiões transcritas, os quais podem ser utilizados na detecção de expressão alelo-específica e também como marcadores moleculares. O método utiliza-se da clonagem *in vitro* e de sistemas de suporte sólido para as unidades de sequenciamento, o que permite que milhares de leituras possam ser determinadas de uma só vez com a plataforma (Carvalho et al., 2010; Rodrigues et al., 2011).

As etapas do RNA-Seq, para qualquer organismo, são basicamente: preparação da amostra, sequenciamento, captura da imagem e análise dos dados (Metzker, 2010), como podemos observar na Figura 5.

Em síntese, o RNA total é isolado e o mRNA é enriquecido e transformado em bibliotecas de cDNA, que será amplificada na plataforma NGS (Wang et al., 2009). Após o sequenciamento, pelo qual as leituras são obtidas, e o número de leituras depende da plataforma NGS utilizada, se houver um genoma de referência as leituras são mapeadas contra este genoma (Pinto et al., 2011). Caso não haja um genoma de referência determinado para a espécie em estudo, é feita a montagem *de novo* das leituras.

A análise dos dados (última etapa) é a mais crítica do processo, uma vez que é produzida uma enorme quantidade de dados, requerendo uma infraestrutura computacional adequada para execução das análises, bem como locais para o armazenamento físico desses dados. Além disso, o grande número de sequências a serem processadas e analisadas torna esta etapa a mais demorada da tecnologia RNA-Seq (Wilhelm et al., 2010). Entretanto, o mais difícil e crucial é transformar os dados obtidos em informação biológica.

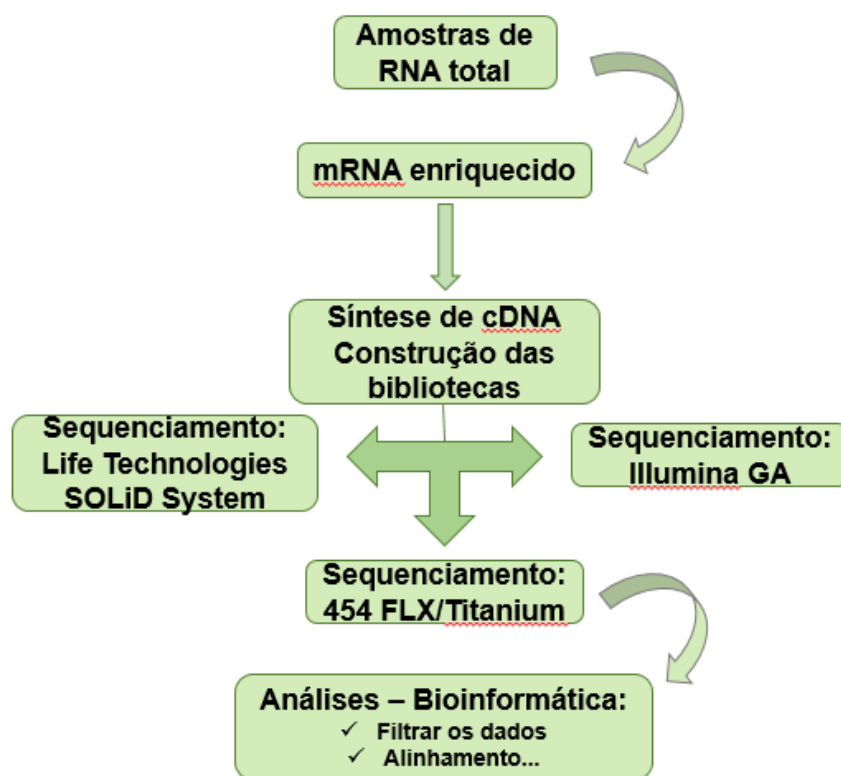


Figura 5: Fluxograma básico para obtenção de um transcriptoma utilizando a plataforma NGS. Adaptado de Pinto et al. (2011).

Além de capacidade computacional, para que essas análises sejam possíveis também é necessário e fundamental a montagem de um transcriptoma referência e criação de um índice gênico bem anotado. A primeira iniciativa de criação desse índice gênico para cana-de-açúcar (*S. officinarum*) surgiu no projeto SUCEST, que gerou 238.000 sequências expressas (ESTs) a partir de diversos tecidos de variedades brasileiras, a qual resultou em 43.141 transcritos putativos únicos de 33.000 genes (Vettore et al., 2001; Vettore et al., 2003).

Posteriormente, surgiram outras iniciativas como o Sugarcane Gene Index (SGI) ou *S. officinarum* Gene Index (SoGI), desenvolvido pelo The Institute for Genomic Research (TIGR) e Dana-Farber Cancer Institute (DFCI), a partir de um conjunto de 282.683 ESTs montadas em 42.377 sequências consensos (transcritos únicos) (Devarumath et al., 2013).

Com o aumento na capacidade de rendimento dos sequenciadores, a quantidade de dados de fragmentos de sequências de transcritos gênicos aumentou de forma exponencial para diversas espécies, inclusive para a cana-de-açúcar, e essas bases de dados deixaram de ser atualizadas. Sendo assim, apesar de iniciativas como essas terem surgido e terem sido importantes referências no passado, atualmente tornaram-se defasadas para utilização como referência nas análises de transcriptomas de cana-de-

açúcar, em especial nos estudos de expressão gênica. Além disso, surgiram novas estratégias para montagem de transcriptomas com base em dados de RNA-Seq, as quais levam em consideração as características das plantas não-modelos, tais como a cana-de-açúcar.

Entre as estratégias de montagem, está a montagem *de novo* dos dados do sequenciamento. O termo montagem *de novo* vem do latim, e significa “desde o princípio”. A realização por diferentes métodos, da montagem *de novo*, é utilizada para determinar a sequência de DNA quando, para a espécie em estudo, não há genoma disponível para o uso como referência (Martins, 2013).

Diante disto, o sequenciamento (RNA-Seq) de nova geração é uma técnica promissora na identificação de genes nos mais diversos processos biológicos e pode ser utilizado para o estudo do transcriptoma em cana-de-açúcar, possibilitando a identificação de genes envolvidos no caráter de tolerância à seca, o que pode auxiliar na seleção de cultivares tolerantes a este tipo de estresse, ou até mesmo na identificação de genes relacionados ao caráter de sensibilidade ao déficit hídrico, neste caso visando o possível aconselhamento da não utilização em regiões que certamente passarão por períodos de seca. Em ambos os casos, a identificação desses genes pode gerar melhor conhecimento dos mecanismos envolvidos com tolerância à seca, o que é útil para a canavicultura e, conseqüentemente, para o setor sucroalcooleiro.

2.5 PCR quantitativa em tempo real (RT-qPCR)

Uma técnica muito utilizada para análises de expressão gênica, quantificação dos níveis de transcrição, e também para a validação de resultados obtidos em experimentos com RNA-Seq, é a RT-qPCR (reação em cadeia da polimerase da transcrição reversa em tempo real - “real-time reverse transcription polymerase chain reaction”).

Esta técnica envolve: transcrição reversa (RT), reação em cadeia da polimerase (PCR) e a análise da PCR assim que os seus produtos se formam (q: quantitativa). Ela é bastante precisa para quantificar níveis de expressão de um determinado gene, bem como oferece resultados rápidos e dados reprodutíveis (Bustin, Mueller, 2005; Stolf, 2007).

Entre as principais diferenças da PCR em tempo real em relação à PCR convencional, é que é possível observar a medição do produto de PCR amplificado a cada ciclo durante a reação de PCR, e não apenas no final (Figura 6). Isto é possível devido a presença de uma câmera de vídeo no equipamento, que registra a luz emitida por um fluoróforo utilizado na reação, o qual se incorpora ao produto recém-sintetizado, permitindo que a quantidade de material seja determinada com precisão e em tempo real (Gachon et al., 2004). Como o resultado é visualizado em tempo real durante a

amplificação da sequência de interesse, a técnica permite gerar resultados quantitativos, diferentemente da PCR convencional, que reproduz resultados qualitativos.

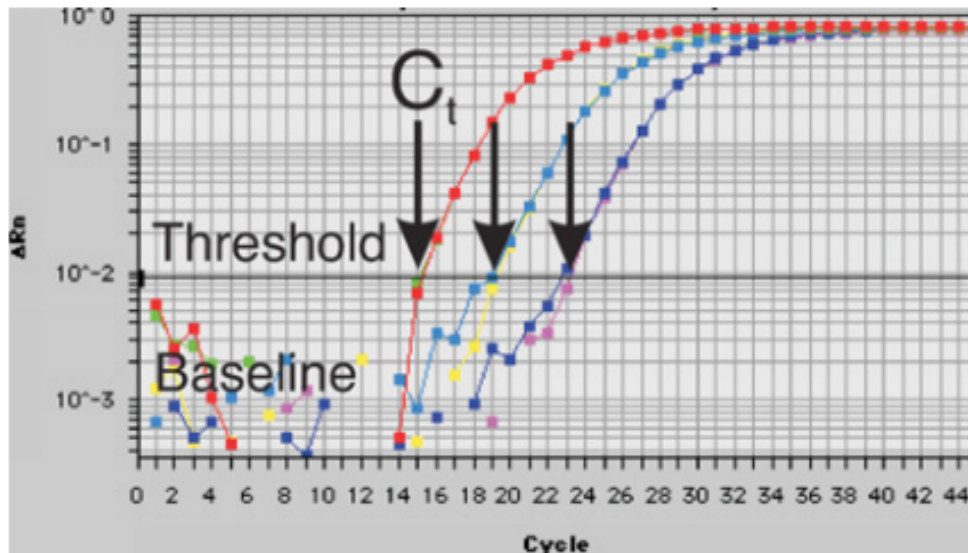


Figura 6: Características das curvas de amplificação da qRT-PCR, que traçam o sinal de fluorescência versus o número do ciclo. O gráfico mostra as curvas de amplificação de três amostras, executadas em duplicatas. Baseline: O sinal de fluorescência durante os ciclos iniciais da PCR que estão abaixo do limiar de detecção do instrumento definem a linha de base (baseline) para o gráfico de amplificação. Valores de C_t indicados pelas setas: representam as frações do ciclo em que o instrumento pode detectar com segurança pela primeira vez a fluorescência derivada da reação de amplificação. Treshold: Um aumento na fluorescência acima do limite (treshold) indica a detecção do produto de PCR acumulado. Fonte: Bustin e Mueller (2005).

Gachon et al. (2004) apresentaram um resumo dos princípios da PCR quantitativa em tempo real. Esta técnica é utilizada para medir com precisão as diferentes quantidades de um mesmo produto gênico alvo, em amostras independentes. No caso da qRT-PCR em tempo real, as amostras de RNA devem ser transcritas reversamente através da enzima transcriptase reversa, a qual converte RNA em cDNA, que é muito mais estável.

As amostras são então submetidas à amplificação por PCR juntamente com um fluoróforo presente no mix de reação, que emite luz quando o produto específico é reconhecido ou sintetizado. Um limiar de fluorescência é definido na parte linear da curva, que corresponde à fase com a melhor eficiência de amplificação. A partir disso um ' C_t ' é definido como o número de ciclos necessários para atingir esse limite de fluorescência.

É desenhada então uma curva de calibração a partir de corridas paralelas, usando quantidades iniciais conhecidas do alvo específico. Esta curva permite o cálculo da quantidade do produto alvo presente em uma amostra de interesse.

O detalhamento de todas as etapas nos permite observar as principais características do método da PCR quantitativa em tempo real: especificidade conseguida pela utilização de dois oligonucleotídeos específicos da sequência alvo; alta sensibilidade em detectar as amostras de DNA ou RNA, rapidez em fornecer dados confiáveis (20 minutos à 2 horas), além do resultado ser visualizado em tempo durante a amplificação da sequência de interesse, com capacidade de gerar resultados quantitativos com alta precisão.

Esta técnica tem sido utilizada com sucesso na validação da expressão gênica em cana-de-açúcar tolerante ao déficit hídrico (Vantini et al., 2015; Li et al., 2016).

2.6 Estudos genéticos e moleculares em cana de açúcar na resposta ao déficit hídrico

As cultivares de cana-de-açúcar em todo o mundo são híbridos interespecíficos resultantes de melhoramento clássico, conhecidas como *Saccharum* spp. A complexidade do genoma da cana, estimado em torno de 10 Mbp, juntamente com altas ploidia, aneuploidia e polimorfismos (D'Hont et al., 1994; D'Hont et al., 1996; Grivet, Arruda, 2002; D'Hont, 2005; Aitken et al., 2014), criam desafios para os programas de melhoramento convencionais e moleculares.

Diversos grupos de pesquisa têm se dedicado a esclarecer os mecanismos utilizados pela cana-de-açúcar para se defender do déficit hídrico, o que resultou em vários trabalhos sobre a análise do transcriptoma das plantas submetidas a esse estresse (Rodrigues et al., 2011; Ferreira et al., 2012; Gentile et al., 2013; Vargas et al., 2014; Cardoso-Silva et al., 2014; Vantini et al., 2015; Cheng et al., 2015; Li et al., 2016; Belesini et al., 2017; Telles et al., 2018).

A grande maioria dos estudos com cana-de-açúcar em casa de vegetação envolvem estresses hídricos (déficit) drásticos e aplicados em um curto período de tempo para identificar e propor cultivares tolerantes (Rodrigues et al., 2009; Gupta et al., 2010).

Assim, foi proposto um projeto à FAPESP (aprovado em 2013, Proc. n 2013/11.617-9), para avaliar o comportamento fisiológico e molecular de duas cultivares de cana-de-açúcar com diferentes padrões de tolerância ao déficit hídrico, ao longo de um período de 90 dias de estresse, no período conhecido como fase de formação da cana-de-açúcar, compreendendo o período mais crítico por demanda de água.

A cultivar SP81-3250 foi utilizada como padrão de tolerância ao déficit hídrico e a cultivar RB855453 foi utilizada como sensível à deficiência hídrica (Pincelli, Silva, 2012). Estas duas cultivares foram submetidas a três potenciais hídricos do solo (controle (sem déficit hídrico), déficit hídrico moderado e déficit hídrico severo) a partir de 60 dias após o plantio e foram amostradas aos 30, 60 e 90 dias após a aplicação dos

tratamentos e avaliadas em relação a parâmetros fisiológicos (número de perfilhos, área foliar, potencial hídrico foliar, índice de cor verde da folha, fotossíntese e concentração interna de CO₂) bem como em relação à expressão gênica (RNA-Seq).

Cabe aqui salientar que todas as variáveis fisiológicas (número de perfilhos, área foliar, potencial hídrico foliar, índice de cor verde da folha, fotossíntese e concentração interna de CO₂) analisadas foram eficientes em distinguir as cultivares SP81-3250 e RB855453 como tolerante e sensível à deficiência hídrica, respectivamente, uma vez que a cultivar tolerante apresentou menores reduções dos parâmetros avaliados que a cultivar sensível (Telles et al., 2018).

Esses resultados fisiológicos nos permitiram avançar nas outras etapas do trabalho, possibilitando a investigação, a nível molecular, de quais genes estão envolvidos na resposta das plantas ao déficit hídrico prolongado (Belesini et al., 2017; Telles et al., 2018).

Compreender as respostas da cana-de-açúcar em diferentes níveis de déficit hídrico em plantas com idades superiores às utilizadas pela maioria dos trabalhos publicados até o momento é de grande importância para prever com maior segurança o comportamento a campo das cultivares estudadas.

Os transcriptomas das cultivares SP81-3250 e RB855453 em diferentes condições de déficit hídrico e em diferentes épocas geraram uma grande quantidade de dados (2.367.679.551 sequências), sendo assim, há muito a ser explorado neste acervo de transcritos. Diante disso, o presente trabalho objetivou dar continuidade a análise do transcriptoma gerado, de maneira a compreender a dinâmica do perfil transcricional ao longo do período do estresse (30, 60 e 90 dias após o déficit hídrico) e dentro de cada tensão do solo (estresse moderado e estresse severo) nas duas cultivares de cana-de-açúcar e, conseqüentemente, identificar transcritos diferencialmente expressos (DEs) e/ou vias metabólicas (ou processos biológicos) envolvidos no caráter de tolerância e/ou sensibilidade da cana-de-açúcar submetida ao déficit hídrico prolongado.

Material e Métodos

3.1 Experimento em casa de vegetação e material vegetal

O experimento para obtenção das amostras foi instalado e conduzido em casa de vegetação, em ambiente controlado. Todos os detalhes sobre a obtenção das mudas, instalação e avaliação do experimento, bem como as etapas de coleta das folhas e extração de RNA total, já foram descritas detalhadamente em estudos anteriores do nosso grupo de pesquisa (Belesini et al., 2017; Telles et al., 2018), as quais aqui descrevemos brevemente.

O delineamento experimental utilizado foi em blocos casualizados, com esquema fatorial 2 por 3 (duas cultivares X três níveis de potencial hídrico do solo), em três épocas amostrais (30, 60 e 90 dias após tratamento), perfazendo um total de 18 unidades experimentais e três repetições, totalizando 54 parcelas, demonstrado na Figura 7.

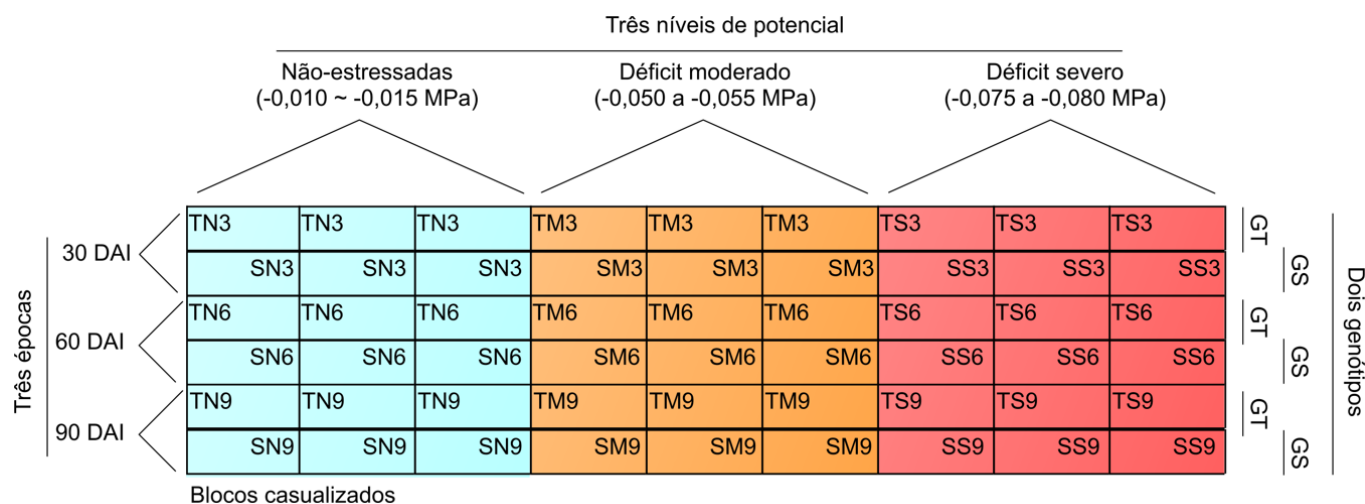


Figura 7: Delineamento experimental utilizado na montagem do experimento, considerando todas as avaliações realizadas. GS: genótipo sensível. GT: genótipo tolerante. N: plantas não estressadas (controle). M: plantas sob déficit hídrico moderado. S: plantas sob déficit hídrico severo. 3, 6 e 9: 30, 60 e 90 dias após aplicação dos tratamentos (DAT).

Foram utilizadas duas cultivares comerciais de cana-de-açúcar consideradas potenciais em produtividade: SP81-3250 utilizada como padrão de tolerância, e RB855453, sensível à deficiência hídrica (Pincelli, Silva, 2012; Telles et al., 2018). As cultivares também apresentam outras diferenças agronômicas (Marin, 2019), expostas na Tabela 1.

Tabela 1: Descritivos agronômicos das cultivares utilizadas no estudo. Adaptado de Marin (2019).

Descritivo	SP81-3250	RB855453
Destaque	Rica e produtiva	Rica e ereta
Regime hídrico	Tolerante ao estresse	Demandante em água
Exigência de solos	ND	Exigente
Maturação	Média	Precoce
Rendimento de transporte	Bom	Bom
Colheita mecânica	Boa	Boa
Brotação	Sem restrição	ND
Brotação de soca com palha	Excelente	Excelente
Fechamento de entrelinhas	Bom	ND
Nematoides	Resposta instável	ND
Florescimento	Regularmente	Todos os anos
Maturadores	Resposta instável	Resposta excelente

ND: Não descrito

As mudas foram cultivadas sem déficit hídrico até completarem 60 dias de idade. Após isso, iniciou-se a aplicação dos tratamentos com três distintos potenciais hídricos do solo: sem déficit hídrico (controle, -0,010 a -0,015 MPa); potencial hídrico moderado (-0,050 a -0,055 MPa) e potencial hídrico severo (-0,075 a -0,080 MPa). As coletas foram realizadas aos 30, 60 e 90 dias após o início dos tratamentos (DAT), totalizando, respectivamente, 90, 120 e 150 dias após o plantio (DAP).

3.2 Extração, quantificação e análise de integridade do RNA total

As coletas para as avaliações moleculares foram realizadas aos 30, 60 e 90 dias após o início dos tratamentos (DAT). As folhas +1 (primeira folha totalmente expandida e com lígula aparente) foram coletadas de cada planta e imediatamente armazenadas em ultrafreezer (-80C).

Os RNAs totais de todas as 54 amostras foram extraídos utilizando-se o Kit comercial Purelink RNA Mini Kit da Ambion®. Os mesmos foram avaliados quanto à qualidade, quantidade e integridade, utilizando os equipamentos Nanodrop ND-1000 (Thermo Fisher Scientific Inc.), Qubit® 2.0 (Invitrogen) e Bioanalyzer 2100 (Agilent Technologies, USA), respectivamente (Belesini, 2015; Telles, 2016).

3.3 Construção das bibliotecas de DNA complementar (cDNA) e sequenciamento Illumina

O RNA mensageiro utilizado na construção das bibliotecas foi isolado e purificado utilizando-se o kit de preparação de amostras TruSeq RNA Sample Preparation v2 (Illumina), a partir de 1,5 ug do RNA total extraído das folhas. As bibliotecas de cDNA foram construídas de acordo com a recomendação do fabricante, conforme descrito em Belesini (2015) e Telles (2016).

O sequenciamento dos cDNAs por meio de RNA-Seq foi realizado nas plataformas HiScanSQ System e HiSeq 2500 da Illumina, seguindo as instruções do fabricante para sequenciamento "paired-end" de fragmentos de 50 pares de bases e de 100 pares de bases (Belesini, 2015; Telles et al., 2018).

3.4 Análise dos dados do sequenciamento.

Nessa seção, descrevemos como foram realizadas as etapas de análise dos dados do sequenciamento até a obtenção dos sinais biológicos em resposta ao estresse hídrico.

3.4.1 Pré-processamento das leituras: remoção de adaptadores e filtros de qualidade.

As leituras brutas ou "reads" no formato FASTq, geradas no sequenciamento, foram pré-processadas com o programa Trimmomatic (v. 0.36) (Bolger et al., 2014), que faz a remoção dos adaptadores, sequências menores que um tamanho arbitrário pré-determinado, sequências de baixa qualidade por um algoritmo de janela de qualidade ou "sliding window" e sequências ambíguas, que apresentam uma alta taxa de nucleotídeos "N".

Os parâmetros utilizados nesse programa foram: janela de qualidade - 4:5; poda abrupta da ponta 5' ("leading"): 5 bases; poda abrupta da ponta 3' ("trailing"): 5 e tamanho mínimo de sequências ("minlen") de 25.

3.4.2 Nova montagem *de novo*, avaliação do transcriptoma de referência e mapeamento das leituras nos transcritos

Em trabalhos anteriores realizados em nosso grupo de pesquisa com estes dados de transcriptoma de folhas de cana-de-açúcar sob déficit hídrico, uma montagem *de novo* já havia sido realizada a partir das leituras processadas provenientes do sequenciamento dos cDNAs via RNA-Seq (Telles et al., 2018).

Contudo, com base nas métricas da montagem do transcriptoma realizada, em especial a fragmentação dos transcritos e a exclusão de uma amostra potencialmente contaminada, concluímos que uma nova montagem deveria ser feita para testar qual das duas montagens seria mais adequada para prosseguir com a mineração dos dados.

As bibliotecas filtradas, sem adaptadores e incluindo apenas as regiões de alta qualidade foram utilizadas na nova montagem *de novo*, realizada com o programa Trinity (v.2.6.6) (Grabherr et al., 2011). Uma única montagem reuniu o transcriptoma de ambas as cultivares (SP81-3250/tolerante e RB855453/sensível).

Os seguintes parâmetros foram alterados no Trinity: cobertura mínima de k-mers para que o contig seja montado pela etapa "Inchworm": "-min_kmer_cov" alterado para 3; "-min_glue", que determina o número mínimo de leituras necessárias para união de duas ou mais sequências contíguas foi aumentado para 3, e o tamanho mínimo de sequências contíguas foi de 300. Os demais parâmetros de configuração do Trinity foram mantidos para a realização da montagem *de novo*.

Nesta nova montagem, foi excluída a amostra SM6_rep3, que, em um teste de agrupamento, foi separada dos demais transcritos, o que levantou a suspeita de contaminação desta amostra e subsequente exclusão das leituras obtidas a partir da mesma, conforme pode ser observado na Figura 8.

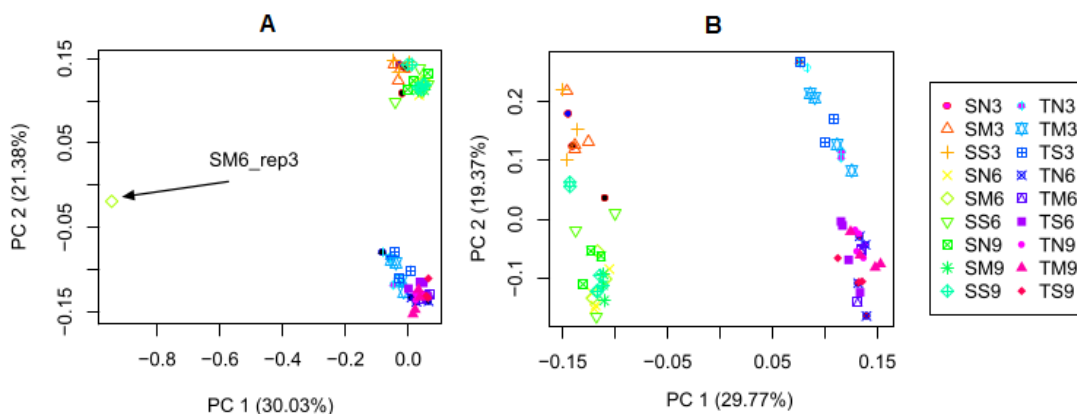


Figura 8: Agrupamento (análise de componentes principais) dos transcritos por amostra, com base na contagem de expressão normalizada com o programa RSEM, antes (A) e depois (B) da remoção da amostra SM6_rep3 (indicada pela seta), exibindo o perfil divergente da amostra e a separação dela nos agrupamentos.

A seguir, a nova montagem foi comparada com aquela realizada anteriormente, a partir de dois avaliadores de montagem: o RSEM-EVAL (Li, Dewey, 2011) e o TransRate (Smith-Unna et al., 2016). Para ambas as avaliações, foi feito o alinhamento das leituras pré-processadas como entrada ou "query", usando como referência ou "subject" o transcriptoma de referência, através do programa Bowtie2 (Langmead,

Salzberg, 2012), configurado com o parâmetro “verysensitive”, que, apesar de tornar a etapa mais lenta, aumenta a sensibilidade na busca por alinhamentos válidos.

No caso do RSEM, é necessário fornecer uma distribuição esperada do tamanho dos transcritos pra aquele organismo – neste caso, foram utilizados os genes anotados do genoma do sorgo, obtidos na forma de um arquivo no formato “General Feature File” ou “gff”, retirado do banco de dados “Sbicolor_79” da base de dados online PlantGDB (Duvick et al., 2007), para a geração deste perfil com o comando “rsem-eval-estimate-transcript-length-distribution”.

Para o avaliador TransRate, foi necessário fornecer o conjunto de proteínas de sorgo anotadas como base de comparação para analisar a qualidade da montagem, proteínas que também foram extraídas do PlantGDB (Duvick et al., 2007).

3.4.3 Análise de expressão diferencial por cultivar, déficit hídrico e tempo

Os transcritos foram quantificados a partir do alinhamento das leituras contra o transcriptoma referência com o programa Kallisto (v. 0.44.0) (Bray et al., 2016), que gerou contagens brutas na unidade de transcritos por milhão (TPM). Essas contagens foram usadas como entrada no programa DESeq2 (Love et al., 2014), com o pacote txImport (Soneson et al., 2015), para o cálculo dos transcritos diferencialmente expressos (DEs).

Na análise utilizando o programa DESeq2 (v.1.14.1), foram realizadas todas as comparações possíveis, considerando as três variantes (cultivar, intensidade do déficit hídrico e data de avaliação) em relação ao seu equivalente controle (sem déficit hídrico), obtendo assim o resultado da diferença de expressão para cada par de fatores na unidade de “log2FoldChange”(L2FC). Foram considerados diferencialmente expressos apenas aqueles transcritos cuja razão de módulo de L2FC foi maior que 2, e de significância estatística (p-valor) menor que 0,05 (usado para as análises exploratórias) e p-valor ajustado pelo teste de falsos positivos de Benjamini-Hochberg, abaixo de 0,1 (usado para as análises quantitativas).

A análise de expressão diferencial no DESeq2 é baseada em modelos estatísticos de modelos lineares generalizados (GLMs) para dados de contagem, baseados na distribuição binomial β -negativa típica de RNA-Seq, para então realizar o cálculo da expressão diferencial e normalizações baseadas no tamanho de bibliotecas.

As análises de expressão diferencial foram feitas usando a suite do R v.3.3.3 (R Core Team, 2013). Os Diagramas de Venn das contagens de transcritos DEs (p-valor ajustado <0,1) entre cultivares, níveis de déficit hídrico e datas de avaliação foram feitos com o InteractiVenn (Heberle et al., 2015).

3.4.4 Anotação dos transcritos

Para a caracterização, os transcritos de tamanho mínimo de 150 nucleotídeos foram anotados utilizando o programa UBLAST do USEARCH11 contra os seguintes genomas obtidos do PGSB (Spannagl et al., 2015): *Arabidopsis thaliana*, *Sorghum bicolor*, *Zea mays*, *Aegilops tauschii*, *Hordeum vulgare* e *Oriza sativa*, além dos genoma monoploide mosaico de cana-de-açúcar do CIRAD (Garsmeur et al., 2018) da cultivar R570.

Para anotação da montagem de referência, primeiro os nucleotídeos foram alinhados com BlastX com o programa Blast+ (Camacho et al., 2009), e depois passaram por uma busca de predição de domínios proteicos usando HMMScan (Finn et al., 2011). A partir desses resultados, foi realizada a predição de proteínas com o programa Transdecoder (Haas et al., 2013), usando como critérios de seleção para as proteínas aquelas com ao menos 900 nucleotídeos com o parâmetro “–retain_long_orfs 900”, ou com alinhamentos no Blast+ “–retain_blastp_hits” ou HMMER “–retain_pfam_hits” conforme descrito acima, retornando apenas a melhor proteína por transcrito, com o parâmetro “–single_best_orf”.

Em seguida, as proteínas foram anotadas utilizando o pacote de programas ou “pipeline” do InterProScan (Finn et al., 2016) para inferência de sinais proteicos por diversos métodos. Também foi realizada uma anotação baseada em ortólogos com o programa Egnog-mapper (v. 20 nov. 2017) (Huerta-Cepas et al., 2017), utilizando como parâmetros para a anotação aqueles baseados em Xiong (2006), que são: identidade mínima de 50% e o e-value mínimo de 0,0001, para transcritos traduzidos pelo TransDecoder cujas proteínas tem tamanho mínimo de 150 aminoácidos, considerado um valor mínimo seguro para atribuição de ortólogos.

A anotação final foi feita contra o banco de dados DroughtDB (Alter et al., 2015), uma base de referência para estresse hídrico. Essa anotação foi feita com o programa Diamond (Buchfink et al., 2015), que realiza alinhamentos de maneira muito mais rápida que o BlastX, sendo que os parâmetros de alinhamento utilizados foram: cobertura mínima da sequência de entrada (“query”) de 50%, e identidade mínima de 50%.

3.4.5 Diferença entre proporções de termos de ontologia e ortólogos

As análises InterProScan e Egnog geraram termos KEGG, COG e GO para cada transcrito. Esses termos foram vinculados a uma matriz de contagens (TPM) normalizada pelo DESeq2.

Essa matriz de contagem foi importada no pacote estatístico STAMP (v.2.1.3) (Parks et al., 2014) para cálculo da diferença entre médias de contagens atribuídas a esses termos, usando o teste *t* de Welch com correção de falsos-positivos (FDR)

de Benjamini-Hochberg, sendo aceitos como significativos os resultados com p-valor ajustado inferior a 0,1.

3.5 Seleção dos transcritos-alvo a validação por PCR quantitativo em tempo real (qRT-PCR)

Finalizadas todas as análises de bioinformática do transcriptoma dessas cultivares de cana-de-açúcar submetidas ao déficit hídrico prolongado, foi possível selecionar manualmente transcritos candidatos (alvos) de tolerância à seca para a validação por PCR quantitativo em tempo real (RT- qPCR).

A seleção foi realizada aleatoriamente a partir dos critérios: (a) se está diferencialmente expresso de acordo com a análise do DESeq2 (considerando p-valor<0,05); (b) se possui anotação e pertence a alguma via já descrita como participante à resposta da planta no estresse hídrico; (c) se esse transcrito foi selecionado no teste do TransRate (não-fragmentado); (d) se é possível anotar as sequências usando repositórios de acesso público, como o BLASTN e BLASTX contra o banco NR do NCBI, o InterProScan versão online, e o AMIGO BLAST (Carbon et al., 2008), que permite obter os termos de ontologia (GO) para uma ou algumas poucas sequências.

Assim, foram selecionados quatro transcritos diferencialmente expressos exclusivamente na cultivar tolerante (p-valor<0,05): Peroxidase Extracelular, Receptor do tipo quinase (GbRLK), Glutathione S-transferase e Poligalacturonase. Além disso, foram utilizados três genes de referência para a normalização dos dados: Proteína 14-3-3, actina e β -tubulina. Para a validação dos transcritos por RT-qPCR, foram realizadas as etapas abaixo descritas.

3.5.1 Obtenção dos RNAs das folhas de cana-de-açúcar

Os RNAs de folhas de cana-de-açúcar utilizados para a síntese da primeira fita de DNA complementar (cDNA) foram os mesmos utilizados para construção das bibliotecas de RNA-Seq e que se encontravam armazenados a -80°C.

3.5.2 Síntese da Primeira Fita de cDNA

A síntese da primeira fita do DNA complementar (cDNA) foi feita utilizando valores calculados a partir das quantificações obtidas anteriormente no fluorômetro Qubit® 2.0 (Invitrogen), em ng/ μ L das amostras de RNA total iniciais (Belesini, 2015; Telles, 2016), obtendo-se um “pool” de RNA total (três repetições) de cada época de amostragem, tanto das plantas controle quanto das plantas sob os diferentes níveis de déficit hídrico analisadas, da cultivar tolerante, totalizando nove amostras.

O procedimento da síntese do cDNA foi feito com o Kit “SuperScript III First-Strand Synthesis SuperMix for RT- qPCR” (Invitrogen, Carlsbad, CA, EUA). As reações foram feitas separadamente para cada “pool” de amostras de RNA de cada condição experimental.

As soluções seguiram o seguinte protocolo: 1 μg de RNA total, 2 μL de RT “enzime mix” contendo “SuperScript III RT” e “RNaseOut”, 10 μL de 2X RT “Reaction Mix”, contendo “primer oligo(dT)20” (2,5 μM), “random hexamers” (2,5ng/ μL), 10 mM MgCl₂ e dNTPs e água ultra pura para completar o volume para 20 μL . As amostras foram incubadas a 25°C por 10 minutos, 50°C por 30 minutos, 85°C por 5 minutos e transferidas para um recipiente com gelo por 1 minuto. Em seguida, foi adicionado 1 μL da enzima E. coli “RNase H” em cada amostra e estas foram incubadas a 37°C por 20 minutos.

A quantificação dos cDNAs foi determinada no espectrofotômetro NanoDrop ND-1000 (“Thermo Fisher Scientific Inc.”), através da leitura de absorbância no comprimento de onda 260 nm, 280 nm e 230 nm. Todas as amostras apresentaram boa quantidade e qualidade do material. O cDNA fita simples foi estocado a -20°C até o uso nas próximas etapas da RT- qPCR.

3.5.3 Desenho dos Oligonucleotídeos Iniciadores dos transcritos alvos para o RT- qPCR

Os oligonucleotídeos utilizados para a amplificação dos transcritos alvos foram desenhados a partir das sequências FASTA disponibilizadas. Já as sequências dos oligonucleotídeos dos genes de referência (endógenos), foram obtidas do trabalho de Vantini et al. (2015), que analisou a expressão gênica diferencial em raízes de cana-de-açúcar tolerantes à seca.

O desenho destes oligonucleotídeos (transcritos alvos) foi realizado utilizando-se o programa “Primer Express” versão 3.0 (“Applied Biosystems”, Foster, CA, USA), seguindo as regras, como tamanho entre 20 e 25 pares de bases, temperatura de anelamento entre 55°C e 60°C e concentração de GC superior a 50%.

Os melhores iniciadores foram selecionados e, em seguida, submetidos a uma análise de bioinformática que alinhou os iniciadores contra a montagem *de novo* realizada, a fim de identificar regiões da sequência que poderiam apresentar alto grau de homologia com outras partes do transcriptoma e, conseqüentemente, poderiam amplificar outras regiões de sequências semelhantes que não a dos transcritos de interesse.

A Tabela 2 mostra os oligonucleotídeos iniciadores dos genes de referência ou endógenos, codificantes de: Proteína 14-3-3 (clone SCCSD2C03A12.g - CA297824);

actina (clone SCMCLV1032G08.g - CA299974) e β -tubulina (clone SCEZFL4046D04.g -CA222437) (Vantini et al., 2015) e dos transcritos alvos que codificam Peroxidase extracelular, Receptor do tipo quinase (GbRLK), Glutathione S-transferase e Poligalacturonase.

Tabela 2: Sequência dos oligonucleotídeos iniciadores dos genes de referência e dos transcritos alvo utilizadas na RT-qPCR.

Proteína	Oligonucleotídeo	Sequência
*PROTEÍNA 14-3-3	14-3-3 – F	5' GAGCCAGATCAGCAAGAGCAAT 3'
	14-3-3 – R	5' GCGGAGAGCACCATGAATG 3'
*ACTINA	ACT - F	5' ATGGAGGCTGCTGGAATCC 3'
	ACT - R	5' ATCCACGTCGCACTTCATGA 3'
* β -TUBULINA	β -TUB - F	5' GGAGGAGTACCCTGACAGAATGA 3'
	β -TUB - R	5' CAGTATCGGAAACCTTTGGTG 3'
Peroxidase extracelular	PER - F	5' CACTTGCCCTCAGCTGTACTAT 3'
	PER - R	5' TCAGCTCTCATGGCAGAGAA 3'
GbRLK	GbRLK - F	5' GCGATCGTTCTGGCTACAAT 3'
	GbRLK – R	5' CGACCCTCCTCAACGACTTG 3'
Glutathione S-transferase	Glut - F	5' GCGACATGATGCTGGTGAAG 3'
	Glut - R	5' GTGGACGTTTGGCTGGAAAT 3'
Poligalacturonase	Poly - F	5' CCTCACAGGGTACTCCTACTCTTC 3'
	Poly - R	5' GCACGAGGAGAGGACTGACT 3'

*: utilizados como normalizadores

3.5.4 Titulação dos Oligonucleotídeos Iniciadores

Para a boa funcionalidade (alta eficiência e ausência de formação de dímeros) nos experimentos de RT- qPCR com o fluoróforo GoTaq® qPCR, foi realizada uma série de padronizações. Na titulação (primeira padronização), fixamos uma concentração de cDNA em 100ng através de um “pool” de cDNA das nove amostras, e variamos a concentração dos iniciadores a 100/100 nM, 300/300 nM e 600/600 nM (F/R, respectivamente), o que possibilitou a escolha da concentração mais adequada para cada um deles.

Todas as reações de RT- qPCR do presente trabalho foram realizadas em equipamento “7500 Real-Time PCR” (Applied Biosystems, Foster, CA, USA, 2009), utilizando o kit GoTaq® qPCR Master Mix (Promega, Madison, Wisconsin, EUA). A reação consistiu em 6,5 μ L de GoTaq® qPCR Master Mix; 2,5 μ L do mix de oligonucleotídeos (Forward + Reverse); 2 μ L de cDNA a 50 ng/ μ L e água ultra pura para um volume final de 12,5 μ L. Os parâmetros utilizados foram: 50°C/2 min, 95°C/10 min e 40 ciclos de 95°C/15 seg e 60°C/1 min.

3.5.5 Curva Padrão e Eficiência de Amplificação da PCR

Após verificar a melhor concentração dos iniciadores a ser utilizada no experimento, foi construída uma curva padrão através de diluições seriadas de cDNA (“pool” de amostras) para verificação da eficiência de amplificação da reação.

A eficiência de amplificação de cada gene foi calculada por meio da construção de uma curva de diluição em série do “pool” de cDNA de 2:1, a partir de 100ng de concentração, com 7 pontos (100 ng, 50 ng, 25 ng, 12,5 ng, 6,25 ng, 3,12 ng e 1,56 ng). Foram determinados 7 pontos de diluição para que houvesse uma margem de segurança, sendo que, segundo Vantini (2013), a eficiência de amplificação pode ser calculada com o mínimo de 5 pontos.

Os resultados de todas as curvas foram analisados no mesmo programa “Applied Biosystems” 7500 real-time PCR e foram consideradas eficientes aquelas reações que apresentaram valores de R^2 igual a 0,99 (pipetagem precisa) e % de EFF: 90 a 110% (eficiência aceitável) (Invitrogen, 2018).

Depois de verificadas as melhores eficiências para cada transcrito alvo e gene de referência, foram definidas as melhores concentrações de cDNA e de oligonucleotídeos.

3.5.6 Quantificação Relativa da Expressão Gênica pelo Método do CT Comparativo ($\Delta\Delta Ct$)

Nas próximas etapas da RT- qPCR, para as análises do ΔCt , todas as reações foram feitas em triplicata e os cDNAs das amostras foram quantificadas em relação aos transcritos alvo e de referência (três datas de avaliação – em três níveis de estresse hídrico (controle, moderado e severo)). Em cada placa foram pipetados também três réplicas dos controles negativos (água no lugar de cDNA) para cada um dos genes alvo e referência.

As reações de RT-qPCR seguiram o seguinte protocolo: 6,25 μL de (GoTaq® qPCR Master Mix) (Promega, Madison, Wisconsin, EUA), 2,5 μL do mix de oligonucleotídeos (Forward + Reverse; a concentração usada foi de acordo com o melhor resultado da titulação dos mesmos), 2 μL de cDNA (50 ng/ L) e água ultra pura para o volume final de 12,5 μL . As reações foram conduzidas em aparelho Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems) e submetidas a 50°C por 2 min, 10 min a 95°C e 40 ciclos de 15s a 95°C e 1min a 60°C.

Os resultados obtidos foram analisados no programa Expression Suite Software v1.01 (“Applied Biosystems”, Foster, CA, USA). Este programa disponibiliza valores de “score” para cada amostra, que são medidas de estabilidade da expressão de um candidato ou controle endógeno em relação a outros controles endógenos ou candidatos

selecionados, através do qual é feita a normalização dos controles endógenos. São considerados satisfatórios valores inferiores a 1,5 (Vandesompele et al., 2002).

Os dados obtidos ao final da reação de RT- qPCR foram gerados na forma de valores de Ct ("Cycle threshold"), que representam o ciclo de início da detecção do produto amplificado, de cada gene alvo e endógeno. A determinação dos níveis de expressão dos genes alvos foi realizada pela quantificação relativa utilizando a expressão $2^{-\Delta\Delta Ct}$ (Livak, Schmittgen, 2001).

Segundo Vantini (2013), este método de análise apresenta os resultados em relação ao ΔCt de um gene normalizador. Para cada tratamento foi detectado o valor de Ct ("Cycle threshold"), tanto para o transcrito alvo quanto para o de referência (ou endógeno). O valor do Ct do gene alvo foi subtraído do valor da média do Ct do melhor gene de referência definido após as etapas iniciais, para ajustar a reação, obtendo-se então o valor de ΔCt . O valor do ΔCt da amostra controle (calibrador = controles em cada um dos três tempos avaliados) foi subtraído do valor do ΔCt do tratamentos, tendo-se o valor de $\Delta\Delta Ct$. Este valor foi utilizado na fórmula do nível de expressão, onde o número 2 representa a somatória da eficiência do gene alvo e do gene de referência, considerando que ambos os genes possuem 100% de eficiência.

Resultados e discussão

4.1 Qualidade do material, sequenciamento das bibliotecas de cDNA via RNA-Seq

As amostras de RNAs obtidas a partir das folhas +1 foram isoladas e tiveram sua pureza e integridade avaliadas conforme procedimentos descritos em estudos anteriores realizados pelo nosso grupo de pesquisa (Belesini et al., 2017; Telles et al., 2018).

O sequenciamento das bibliotecas na plataforma Illumina HiScanSQ System (2x50 pb) e HiSeq2500 (2x100 pb) gerou um total de 2.367.679.551 leituras brutas (1.712.145.471 leituras de 50 pb e 655.534.080 leituras de 100 pb), excluindo aquelas referentes a amostra SM6_rep3 (770 sequências), possivelmente contaminada, totalizando 2.367.678.781 leituras.

4.2 Análise do transcriptoma das duas cultivares, aos 30, 60 e 90 dias de déficit hídrico

4.2.1 Processamento das leituras brutas

As leituras brutas geradas a partir do sequenciamento (2.367.678.781) foram pré-processadas, gerando um número final total de 1.798.266.112 leituras. Esses valores revelam que a quantidade de sequências de baixa qualidade removida foi pequena, fazendo com que o número de sequências utilizadas na análise não diminuísse muito quando comparado às sequências brutas, evidenciando a boa qualidade do sequenciamento.

4.2.2 Nova montagem *de novo* e avaliação do transcriptoma

Após o pré-processamento das leituras obtidas do sequenciamento das bibliotecas na plataforma Illumina HiScanSQ System (2x50 pb) e HiSeq2500 (2x100 pb) e remoção de adaptadores e contaminantes, a montagem *de novo* foi feita utilizando-se o programa Trinity (v.2.6.6) (Grabherr et al., 2011), a qual resultou em um total de 251.653 transcritos montados, sendo 105.522 genes distintos, com tamanho médio de 1.089

pares de bases e um total de 278.721.096 bases. Essa montagem resultou em um N50 dos genes igual a 1.565 pares de bases com uma porcentagem de GC de 49,90%, apresentados em detalhe na Tabela 3.

Tabela 3: Parâmetros e estatísticas da montagem *de novo* do transcriptoma de folhas de cana-de-açúcar.

Descritivo	Número
Número total de genes	105.522
Número de transcritos	251.653
Proporção GC	49,90%
N50 (pb)	1.565
Mediana (pb)	738
Média (pb)	1.089
Número total de bases	278.721.096
Maior transcrito (pb)	15.275
N50: valor que divide a montagem pela metade em relação ao comprimento dos transcritos	

Além disso, foi feita uma comparação da pontuação das duas montagens, utilizando dois avaliadores. O resultado dos dois programas avaliadores, Detonate (RSEM) e TransRate, pontuaram que a nova montagem seria mais adequada.

A Tabela 4 apresenta os principais resultados da avaliação das montagens. O maior número da razão entre mapeamentos e "bons mapeamentos", o maior tamanho da menor sequência e o maior N50, definiu a escolha pela nova montagem *de novo*. O parâmetro de "mapeamentos bons/ruins" do Transrate considera a proporção de ligações entre as sequências contíguas e os transcritos, que chama de **pontes**.

Tabela 4: Principais métricas de avaliação das montagens *de novo* do transcriptoma, geradas pelos programas RSEM-EVAL e Transrate.

Métrica	Avaliador	1ª montagem	Nova montagem
Nº seqs.	TransRate	362662	251653
Menor seq.	TransRate	201	301
Fragmentos mapeados	TransRate	7.17E+07	1.98E+08
Bons mapeamentos	TransRate	6.06E+07	1.69E+08
% bons mapeamentos / total	TransRate	0.067	0.187
N50	Assemblathon	1247	1565
Contigs sem leituras alinhadas	RSEM	2486	176
Pontuação	RSEM	-9.06E+10	-7.13E+11

N50: valor que divide a montagem pela metade em relação ao comprimento dos transcritos

Os resultados obtidos na nova montagem *de novo* foram semelhantes aos valores de N50 (1.367 pb) e porcentagem de GC (46,39%) obtidos na montagem *de*

novos de transcriptomas de 6 genótipos de cana-de-açúcar que são empregados como progenitores em programas de melhoramento de cana-de-açúcar no Brasil, encontrados no trabalho de Cardoso-Silva et al. (2014), que também utilizou o programa Trinity para a montagem. Estes resultados evidenciam a boa qualidade da montagem *de novo* dos dados obtidos via RNA-Seq.

4.2.3 Identificação dos transcritos diferencialmente expressos (DEs) por cultivar, déficit hídrico e tempo de estresse

A análise de expressão diferencial foi realizada para cada cultivar - tolerante (T) e sensível (S) - separadamente para os dois níveis de déficit hídrico aplicados - estresse moderado (M) e estresse severo (S), em relação ao controle (não estressadas - N), aos 30, 60 e 90 dias após a aplicação do tratamento (DAT). Aqui, foram considerados transcritos diferencialmente expressos pelo programa DESeq2, aqueles com o valor de "Fold Change" maior do que o módulo de 2 ($FC \geq 2$) e valor de significância estatística (p -valor ajustado pelo teste de falsos positivos de Benjamini-Hochberg) abaixo de 0,1.

Foram identificados 352 transcritos diferencialmente expressos (DEs) entre as plantas sob déficit hídrico moderado em, no mínimo, um dos tempos avaliados (30, 60 e 90 dias). Desses transcritos DEs, 177 foram expressos exclusivamente na cultivar tolerante, apenas três compartilhados entre as duas cultivares e 172 expressos exclusivamente da cultivar sensível.

Entretanto, o número de transcritos DEs nas plantas sob déficit hídrico severo foi bem maior quando comparada ao moderado, 3837 transcritos DEs foram identificados em, no mínimo, um dos tempos (30, 60 e 90 dias), sendo 1006 exclusivos da cultivar tolerante, 168 compartilhados entre as duas cultivares, e 2663 exclusivos da cultivar sensível à deficiência hídrica.

A relação entre os transcritos diferencialmente expressos que são comuns ou exclusivos dentro de cada cultivar, nas três diferentes épocas amostrais (30, 60 e 90 dias), para os diferentes níveis de déficit hídrico separadamente, estão representados nos Diagramas de Venn abaixo (Figura 9 e 10).

Na cultivar tolerante sob déficit moderado, o número de transcritos DEs aumentou ao longo do tempo, sendo maior aos 90 DAT (51.6%) (Figura 9a). Já para a cultivar tolerante sob déficit hídrico severo (Figura 9b), o maior número de transcritos DEs foi observado aos 60 DAT (43.4%).

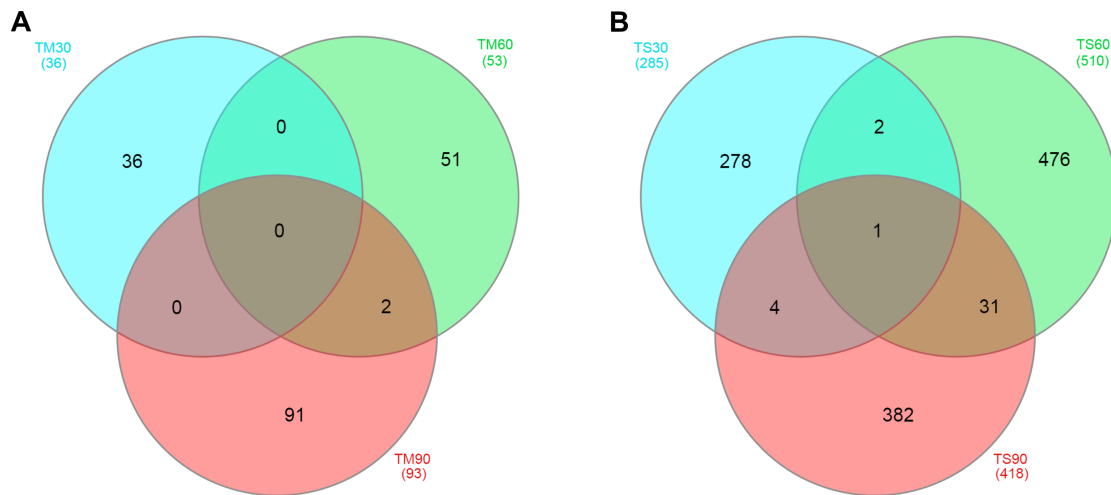


Figura 9: Diagrama de Venn do número de transcritos diferencialmente expressos na cultivar de cana-de-açúcar tolerante sob déficit hídrico moderado (A) e severo (B), aos 30, 60 e 90 dias após o início dos tratamentos (p-valor ajustado < 0.1). TM: cultivar tolerante sob déficit hídrico moderado. TS: cultivar tolerante sob déficit hídrico severo.

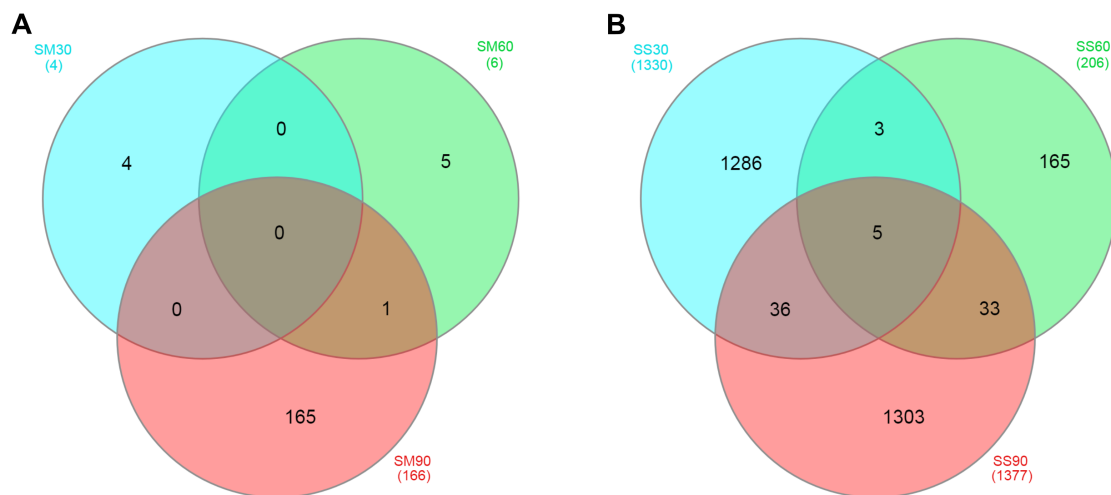


Figura 10: Diagrama de Venn do número de transcritos diferencialmente expressos na cultivar sensível sob déficit hídrico moderado (A) e severo (B), aos 30, 60 e 90 dias após o início dos tratamentos (p-valor ajustado < 0.1). SM: cultivar sensível sob déficit hídrico moderado. SS: cultivar sensível sob déficit hídrico severo.

Neste tratamento, apenas um transcrito apresentou expressão diferencial nas três datas de avaliação. Este transcrito (30431_c0_g1_i1) não recebeu nenhuma anotação conhecida de acordo com todos os bancos de dados utilizados, podendo fazer

parte do grupo de transcritos/genes que ainda não possuem anotação e que podem estar envolvidos na resposta de defesa ao déficit hídrico prolongado na cana-de-açúcar.

Na cultivar sensível sob déficit moderado, o número de transcritos DEs aumentou consideravelmente ao longo do tempo: aos 30 DAT, apenas quatro transcritos (2.3%) apresentaram expressão diferencial, enquanto que aos 90 DAT, o número de DEs foi de 166 transcritos (94.8%). Neste tratamento, nenhum transcrito apresentou expressão diferencial nas três datas de avaliação (Figura 10a). Sob déficit hídrico severo (Figura 10b), o número de transcritos DEs também foi maior aos 90 DAT (48.6%). Poucos transcritos foram DEs em comum às três datas de avaliação - apenas cinco.

Entre os cinco transcritos DEs nas três datas avaliadas, observamos que o transcrito 34543_c0_g1_i1 apresentou expressão diferencial reprimida na cultivar sensível sob déficit hídrico severo, aos 30, 60 e 90 dias, sendo que não apresentou expressão diferencial em nenhuma outra situação avaliada. Este transcrito foi anotado como uma proteína contendo domínio NAC.

Fatores de transcrição fazem parte de um grupo de proteínas chamadas proteínas reguladoras, que atuam na transdução de sinais e controlam processos celulares, pois regulam a expressão de genes responsivos ao estresse, se ligando diretamente aos promotores dos genes alvos em sequências específicas do DNA (Shinozaki et al., 2003; Ferro, 2008).

Em relação a família NAC de fatores de transcrição, *Arabidopsis thaliana* apresenta 106 membros e arroz apresenta 149 membros (Lia-Jia, Yu-Xian, 2006). Fujita et al. (2004), em experimento com plantas transgênicas, mostrou uma proteína NAC induzida por desidratação, alta salinidade e tratamentos com ABA e JA (proteína RD26 NAC), fazendo parte da via ABA dependente tipo II. Além disso, outro fator de transcrição do tipo NAC, da via ABA independente tipo IV, se liga a elementos *cis* do promotor do gene ERD1, ativando a sua expressão. Este gene ERD1, que atua diretamente na resposta a diferentes estresses abióticos, está descrito como induzido por desidratação (Tran et al., 2004).

A análise de expressão diferencial indicou que o transcrito que codifica uma proteína contendo domínio NAC apresenta expressão diferencial reprimida na cultivar sensível sob déficit hídrico severo em relação ao controle, nas três datas avaliadas. Este comportamento pode estar relacionado ao caráter de sensibilidade atribuído a esta cultivar, uma vez que pode reprimir a ativação de genes responsáveis por uma resposta frente a uma situação de estresse hídrico.

O demais transcritos DEs comuns as três datas de avaliação apresentaram alguma porcentagem de identidade com domínios proteicos com função desconhecida/hipotética ou não receberam nenhuma anotação conhecida ("no hits"), de acordo

com todos os bancos de dados utilizados, também podendo fazer parte do grupo de transcritos/genes que podem estar envolvidos na resposta de defesa ao déficit hídrico prolongado na cana-de-açúcar.

4.2.4 Anotação dos transcritos

O conjunto total de 251.653 transcritos montados, receberam um número de anotações pelos critérios descritos no material e métodos, conforme a Tabela 5, na tentativa de caracterizar toda a expressão gênica a partir da nova montagem.

Tabela 5: Anotações dos transcritos por referência.

Referência, programa ou "pipeline" de anotação	Transcritos anotados
Genoma de <i>A. tauschii</i>	77.512
Genoma de <i>A. thaliana</i>	33.084
Genoma de <i>H. vulgare</i>	71.671
Genoma de <i>O. sativa</i>	73.424
Genoma de <i>S. bicolor</i>	101.876
Genoma de <i>Z. mays</i>	90.613
Genoma monoploide de <i>S. officinarum</i>	101.128
InterProScan	104.534
Eggnog-mapper	62.087
Diamond (blastX) NR	228.272
DroughtDB (proteínas)	4.722
Alter et al. (2015), Garsmeur et al. (2018)	

A partir desses resultados, observamos que a anotação utilizando os genomas mais próximos do gênero *Saccharum*, como gramíneas (*Zea mays* e *Sorgum bicolor*), foram os que mais corresponderam a alinhamentos que atendessem os critérios de "hits" válidos, valor este que diminui conforme aumenta a distância filogenética do organismo anotado para a cana-de-açúcar.

O transcriptoma de referência também foi alinhado em relação ao genoma monoploide mosaico de cana-de-açúcar recém-publicado (Garsmeur et al., 2018), gerando um número de alinhamentos um pouco inferior àqueles relativos ao *S. bicolor*. A explicação para esse número inferior de alinhamentos é que este genoma, apesar de importante para unificar as anotações e referências de cana-de-açúcar, ainda não é suficiente para capturar toda a diversidade dessa cultura, especialmente levando em consideração a alta complexidade e ploidia do seu genoma.

4.2.5 Categorização dos transcritos DEs das cultivares a partir de ortólogos (KEGG) e ontologia (GO)

Os transcritos DEs sob déficit hídrico moderado (Figura 11), e severo (Figura 12), separadamente, em comparação a ausência de estresse, em ao menos um tempo, na cultivar sensível (A) e tolerante (B), receberam termos de ontologia gênica (GO) relativos a processos biológicos, a partir do resultado da anotação no Egnog-mapper (Huerta-Cepas et al., 2017). O "Gene Ontology" permite classificar a anotação em três categorias distintas: processos biológicos, função molecular e componente celular (Ashburner et al., 2000).

Aqui serão apresentados os resultados das três categorias para os transcritos DEs sob déficit hídrico moderado, pois o número de transcritos DEs neste estresse foi menor, e ainda receberam poucos termos em cada uma delas, e apenas os resultados de processos biológicos para os transcritos DEs sob déficit hídrico severo, uma vez que apresentaram várias respostas relacionadas à estresses abióticos. Os transcritos da Figura 11a não foram estatisticamente DEs segundo o teste de p–valor ajustado $<0,05$.

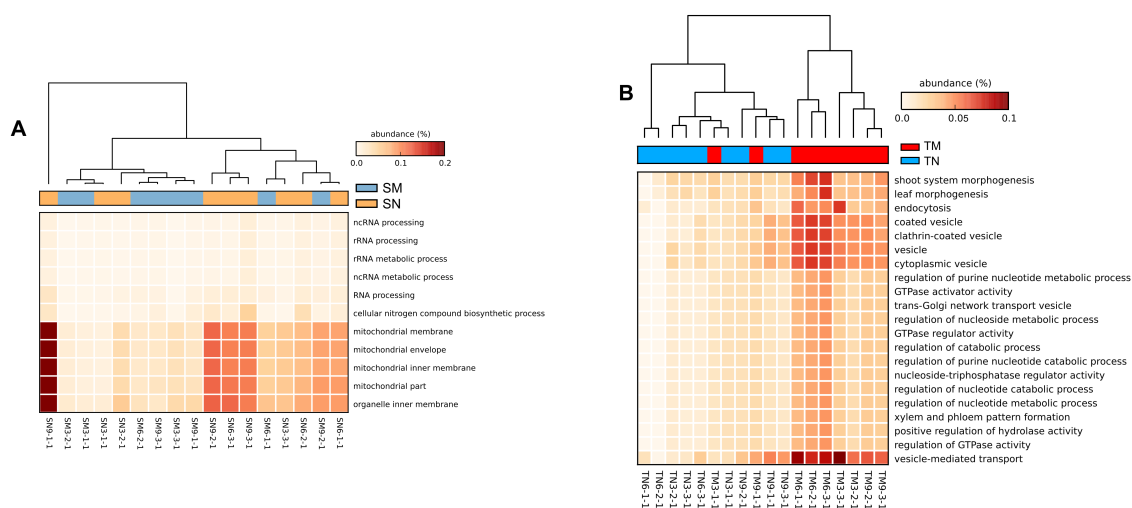


Figura 11: Mapa de calor para termos de ontologia gênica (GO) relativos a processos biológicos, função molecular e componente celular, atribuídos aos transcritos de folhas de cana-de-açúcar diferencialmente expressos sob déficit hídrico moderado em comparação à ausência de estresse, em ao menos um tempo (p -valor ajustado $<0,05$). A: cultivar sensível ao déficit hídrico; B: cultivar tolerante ao déficit hídrico. Porção A: não-significativos segundo o teste de p -valor ajustado $<0,05$.

Na cultivar sensível sob déficit hídrico moderado, observa-se uma abundância de transcritos relacionados a mitocôndria, responsável pela respiração celular e conse-

quente produção de energia (Figura 11a). A sua função é vital para a célula, sem a qual há morte celular.

As amostras não apresentaram divergência perfeita entre as plantas sob déficit hídrico moderado e as plantas controle, mas, ainda assim, é possível observar que esses processos estão mais abundantes nas plantas não estressadas, evidenciando alterações no rendimento de processos vitais como a respiração celular, quando a planta se encontra sob uma situação de déficit hídrico.

Sob uma condição de déficit hídrico moderado, a cultivar tolerante apresentou, entre outros, mecanismos relacionados ao transporte mediado por vesículas, manutenção de vesículas citoplasmáticas, processo de endocitose e morfogênese de folhas.

Foi observada uma abundância maior de transcritos relacionados a estes processos e componentes nas plantas sob déficit hídrico moderado, em comparação as plantas não estressadas, e também, a abundância aumentou de acordo com a passagem do tempo, estando mais evidentes nas plantas avaliadas aos 60 e 90 DAT (Figura 11b).

Uma vesícula pode ser definida como um pequeno saco membranoso que pode armazenar ou transportar substâncias. Além disso, as vesículas, como os vacúolos, são utilizados nas células vegetais principalmente para o controle osmótico e o armazenamento de nutrientes, podendo o aumento destes componentes celulares indicar uma resposta inicial das plantas tolerantes frente à uma situação de déficit hídrico moderado.

Entre as plantas sob déficit hídrico severo, nota-se, na cultivar tolerante, vários mecanismos implicados na resposta ao déficit hídrico, como respostas a estresses oxidativos, espécies reativas de oxigênio, estresse (geral), além de diversas respostas ao calor e luminosidade.

Estão também mais abundantes os termos de GO associados a metabolismo secundário (caroteno, tetraterpenoides, terpeno), indicando que fazem parte da resposta da planta tolerante ao déficit hídrico. Ainda, esses processos estão muito mais abundantes na plantas sob déficit hídrico severo, quando comparadas às plantas não estressada (Figura 12b).

Já na cultivar sensível, outras categorias apareceram, sendo que todas estão mais abundantes nas plantas não afetadas pelo déficit hídrico, como regulação de processos biológicos e celulares, desenvolvimento pós-embrionário e de estruturas anatômicas e reprodutivas, processos envolvidos com reprodução, transporte de fluidos, entre outras (Figura 12a).

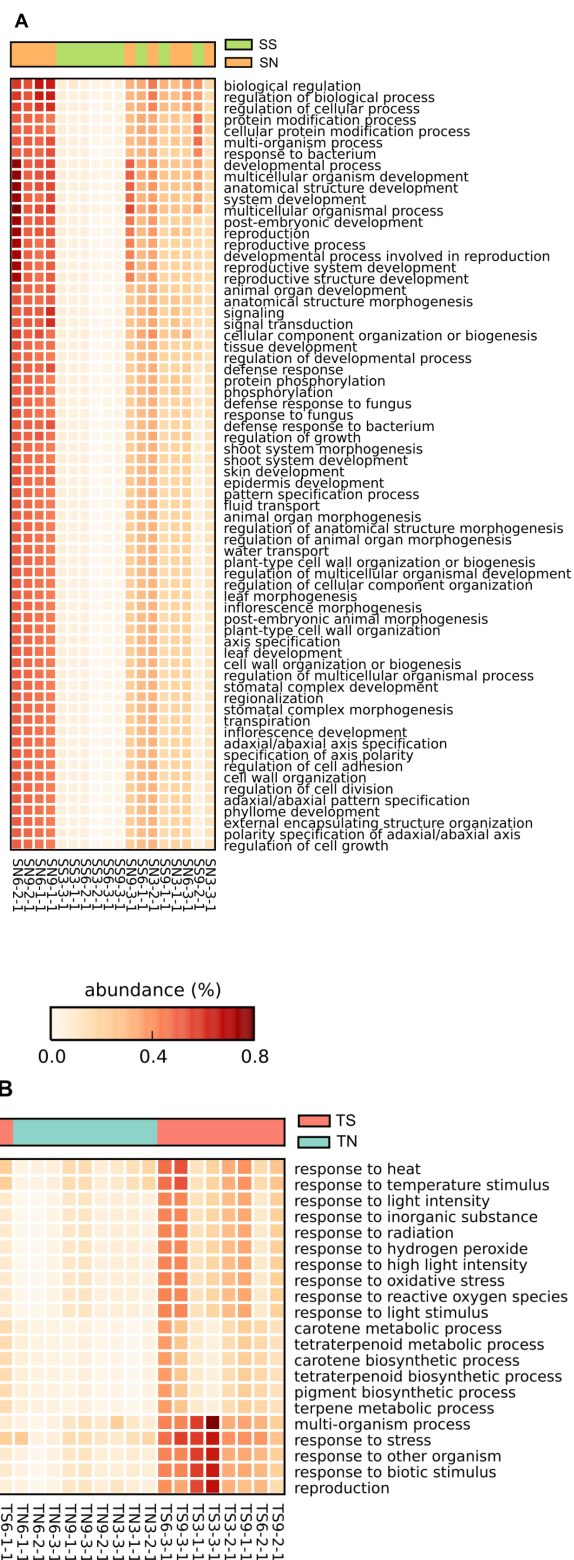


Figura 12: Mapa de calor para termos de ontologia gênica (GO) relativos a processos biológicos, atribuídos aos transcritos de folhas de cana-de-açúcar diferencialmente expressos sob déficit hídrico severo, em comparação à ausência de estresse, em ao menos um tempo (p -valor ajustado $< 0,05$). A: cultivar sensível ao déficit hídrico. B: cultivar tolerante ao déficit hídrico.

Essas observações evidenciam a resposta da cultivar tolerante ao déficit, mantendo seu crescimento, sem que o rendimento de outros processos vitais sejam diminuídos.

Além dos termos GO gerados pelas anotações com InterProScan e Egnog-mapper, os transcritos também receberam termos KEGG para rotas metabólicas desse segundo protocolo. Considerando os transcritos estatisticamente DEs segundo o teste de p-valor ajustado $<0,1$, apenas os transcritos DEs da cultivar tolerante sob déficit hídrico severo, na comparação com o controle, receberam classificação KEGG, apresentada na Figura 13.

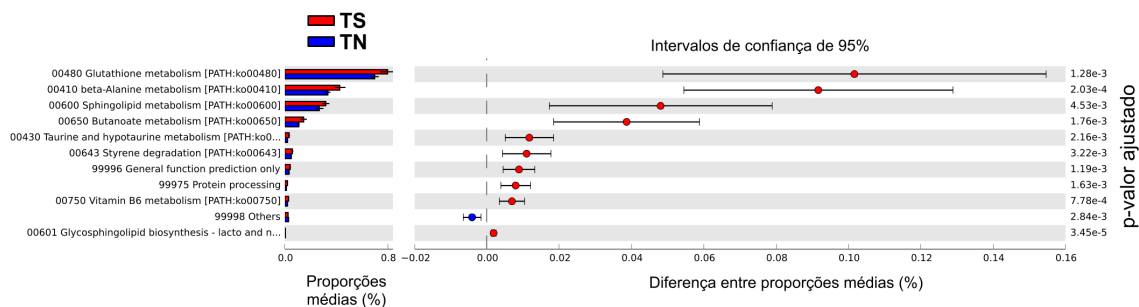


Figura 13: Transcritos diferencialmente expressos na cultivar tolerante, entre as plantas controle e sob estresse severo, agrupados por rotas metabólicas do KEGG (p-valor ajustado $<0,1$). TN: plantas controle da cultivar tolerante. TS: plantas da cultivar tolerante sob estresse severo.

Observa-se que vários transcritos DEs sob déficit hídrico severo receberam termos KEGG relacionados ao metabolismo da Glutathione (GSH) (Tabela 6). Muitos estudos tem identificado o comportamento da GSH e das enzimas e produtos da sua rota metabólica nas respostas das plantas à seca, uma vez que o estresse osmótico é um dos efeitos primários do estresse hídrico e a glutathione e antioxidantes relacionados são bem conhecidos por sua função de eliminação de espécies reativas de oxigênio (ERO), desempenhando um papel protetor no estresse biótico (Kocsy et al., 2002; Cheng et al., 2015).

A glutathione é um antioxidante não enzimático que regula múltiplas funções metabólicas. Esta molécula tem chamado atenção devido a sua interação com outras moléculas sinalizadoras e fito-hormônios. Neste sentido, Hasanuzzaman et al. (2017) discutiram o progresso recente dos estudos da biossíntese e metabolismo da GSH, e seus papéis em plantas na tolerância ao estresse abiótico. Em relação ao papel da GSH e suas enzimas metabólicas em uma situação de seca, elas interagem como reguladores de crescimento das plantas, tanto em condições normais quanto estressantes. A rota do metabolismo da glutathione pode ser observada na Figura 14.

Tabela 6: Perfil de expressão ao longo do tempo dos transcritos DEs (p-valor ajustado < 0,1) em ambas as cultivares e níveis de déficit hídrico, que receberam termos do KEGG Pathways da rota do metabolismo da glutatona (ko00480). Exibidos os valores de L2FC. TM: cultivar tolerante sob déficit hídrico moderado. TS: cultivar tolerante sob déficit hídrico severo. SM: cultivar sensível sob déficit hídrico moderado. SS: cultivar sensível sob déficit hídrico severo. 30, 60 e 90: datas de avaliação após a aplicação dos tratamentos.

TRANSCRITO	ANOTAÇÃO	TM30	TM60	TM90	TS30	TS60	TS90	SM30	SM60	SM90	SS30	SS60	SS90
40836_c2.g3.i4	isocitrate dehydrogenase	-	-	-	-	-	-	-	-	-2.732	-	-	-
40836_c2.g3.i2	isocitrate dehydrogenase	-	2.196	-	-	2.802	-	-	-	-	-	-	-
40789_c1.g1.i1	isocitrate dehydrogenase	-	-	-	-2.119	-	-	-	-	-	-	-	-
40836_c2.g3.i1	isocitrate dehydrogenase	-	-	-	-	2.221	-	-	-	-	-	-	-
36333_c3.g1.i20	glucose-6'P 1-dehydrogenase	-	-	-	-	-	-3.339	-	-	-	-	-	-
36333_c3.g1.i21	glucose-6'P 1-dehydrogenase	-	-	-	-2.849	-	-	-	-	-	-	-	-
35874_c0.g2.i2	glucose-6'P 1-dehydrogenase	-	-	-	-	-	-	-	-	-2.271	-	-	-
45361_c0.g1.i3	L-ascorbate peroxidase	-	-	-	-	3.014	-3.217	-	-	-	-	-	-
35239_c1.g1.i12	L-ascorbate peroxidase	-	-	-	-	-3.531	-	-	-	-	-	-	-
45361_c0.g1.i17	L-ascorbate peroxidase	-	-2.506	-	-	-	-	-	-	-	-	-	-
37377_c1.g1.i1	spermidine synthase	-	-	-	-	-	-	-	-	-	-2.218	-	-
40969_c0.g1.i1	glutathione S-transferase	-2.675	-	-2.605	-	-	-	-	-	-	-	-	-
26373_c0.g1.i1	glutathione S-transferase	-	-	-	-	-2.345	-	-	-	-	-	-	-
44184_c2.g2.i3	leucyl aminopeptidase	-2.353	-	-	-	-	-	-	-	-	-	-	-
33746_c1.g1.i1	leucyl aminopeptidase	-3.070	-	-	-	-	-	-	-	-	-	-	-
44184_c2.g2.i16	leucyl aminopeptidase	-	-	-	-	-	-	-	-	-4.185	-	-	-
44184_c2.g2.i5	leucyl aminopeptidase	-	-	-	-	-	-	-	-	-4.201	-	-	-
46191_c1.g2.i3	aminopeptidase N	-	-	-	-	-	-	-	-	-	-2.017	-	-
46191_c1.g1.i2	aminopeptidase N	-	-	-	-	-	-	-	-	-	-2.838	-	-
46191_c1.g2.i1	aminopeptidase N	-	-	-	-	-	-	-	-	-	-2.031	-	-
48040_c0.g1.i2	5-oxoprolinase	-	-	-	-	-2.204	-	-	-	-	-	-	-
48765_c2.g1.i12	5-oxoprolinase	-	-	-	-	-	-	-	-	-	-3.500	-	-
48765_c2.g1.i5	5-oxoprolinase	-	-4.875	-	-	-4.357	-	-	-	-	-	-	-
48040_c0.g1.i4	5-oxoprolinase	-2.050	-	-	-	-	-	-	-	-	-	-	-
48765_c2.g1.i6	5-oxoprolinase	3.874	-	3.129	3.808	-	-	-	-	-	-	-	-
40886_c1.g1.i8	glutamate-cysteine ligase	-	-	-	-	-	-	-	2.663	-	-	-	-
47654_c1.g1.i10	glutathione synthase	-	-	-	-	-	-	-	-	2.926	-	-	-
47654_c1.g1.i1	glutathione synthase	-	-	2.948	-	-	-	-	-2.960	-	-	-3.346	-
20249_c0.g1.i2	glutathione hydrolase	-	-2.755	-	-	-	-	-	-	-	-	-	-

(-): o transcrito não apresentou expressão diferencial nesta data de avaliação.

Labudda e Azam (2014) também revisaram a função da glutatona e suas enzimas relacionadas nas respostas de uma variedade de plantas submetidas à seca, confirmando que a GSH desempenha um papel central no metabolismo das células vegetais durante o estresse abiótico, desempenhando um papel muito importante na proteção das plantas contra o estresse oxidativo induzido pelo déficit hídrico nos tecidos.

Dois transcritos foram anotados como Glutathione S-transferase (GST) (EC:2.5.1.18). Eles apresentaram expressão diferencial exclusivamente na cultivar tolerante. O transcrito 40969_c0_g1_i1 foi reprimido aos 30 e 90 DAT sob déficit hídrico moderado, e o outro transcrito (26373_c0_g1_i1) foi reprimido aos 60 DAT sob déficit hídrico severo.

A expressão reprimida da GST em uma condição de estresse hídrico, na cultivar tolerante, concorda com o que já foi descrito na literatura, visto que Chen et al. (2012) forneceram evidências para apoiar o papel negativo do AtGSTU17 na tolerância à seca e ao estresse salino: as plantas mutantes (*atgstu17*) foram mais tolerantes à seca e aos estresses salinos em comparação as plantas selvagens, além de acumular maiores níveis de GSH e ácido abscísico (ABA), menores aberturas estomáticas e menores taxas de transpiração.

O estudo propôs que a GST (AtGSTU17) possui um papel nas respostas adaptativas às secas e aos estresses do sal, funcionando como um componente negativo das vias de transdução de sinal mediadas pelo estresse.

Estudos confirmam que níveis elevados de GSH endógeno podem conferir tolerância ao estresse hídrico e melhorar danos causados pelo aumento de espécies reativas de oxigênio (ERO) (Hasanuzzaman et al., 2017).

A enzima aminopeptidase N (EC:3.4.11.2) na cultivar tolerante não apresentou expressão diferencial em nenhum dos tempos e níveis de déficit hídrico. Entretanto, na cultivar sensível sob déficit hídrico severo, três transcritos (46191_c1_g2_i1, 46191_c1_g2_i2 e 46191_c1_g2_i3) apresentaram expressão reprimida aos 30 DAT. Esta enzima converte L-Cisteinil-glicina em Glicina.

Desta maneira, a cultivar tolerante manteve a produção de glicina, que pode ser convertida em GSH através da enzima Glutathione sintase (EC:6.3.2.3), a qual apresentou-se induzida no transcrito 47654_c1_g1_i1 na cultivar tolerante sob déficit hídrico moderado, aos 90 DAT, e este mesmo transcrito, na cultivar sensível, apresentou-se reprimido aos 60 DAT em ambos os níveis de déficit hídrico.

O estresse hídrico moderado causou um aumento significativo na atividade da enzima que produz GSH na cultivar tolerante aos 90 DAT, enquanto na cultivar sensível, a sua atividade foi reprimida, o que pode representar uma resposta tardia da planta tolerante frente ao estresse osmótico, efeito primário do estresse causado pelo déficit hídrico.

Na cultivar tolerante a expressão induzida do transcrito 40836_c2_g3_i2 e 40836_c2_g3_i1, que codificam a enzima isocitrato desidrogenase (EC:1.1.1.42) aos 60 DAT sob déficit hídrico moderado e severo, pode ter relação na resposta desta cultivar frente ao estresse hídrico, considerando que esta enzima converte NADP⁺ em NADH, o qual é utilizado pela glutathiona redutase na conversão de glutathiona oxidada em GSH.

Além disso, a expressão do transcrito 20249_c0_g1_i2 que codifica a enzima glutathiona hidrolase (EC:3.4.19.13), que converte a GSH para L-Cisteinil-glicina e L-Glutamato, foi reprimida aos 60 DAT na cultivar tolerante sob déficit hídrico moderado, podendo indicar que a cultivar tolerante, sob déficit hídrico moderado mantém os níveis de GSH através da repressão da expressão de uma das enzimas que promove a hidrólise da glutathiona.

A enzima ascorbato peroxidase (APX) faz parte da via Asada-Halliwel para a eliminação de espécies reativas de oxigênio (ERO). A eliminação do peróxido de hidrogênio (H₂O₂) pela APX é a primeira etapa do ciclo ascorbato-glutathiona (ASC-GSH) contra o estresse oxidativo (Foyer, Halliwel, 1976; Asada, 1994).

Três transcritos da enzima L-ascorbato peroxidase (EC:1.11.1.11) apresentaram diferentes padrões de expressão diferencial na cultivar tolerante: um deles foi reprimido aos 60 DAT sob déficit hídrico moderado (453601_c0_g1_i17), e o outro (35239_c1_g1_i12) foi reprimido aos 60 DAT sob déficit hídrico severo. Enquanto um terceiro transcrito (45361_c0_g1_i3) foi induzido aos 60 DAT e reprimido aos 90 DAT sob déficit hídrico severo. A cultivar sensível não apresentou expressão diferencial destes transcritos.

Ferro (2008), avaliando a expressão gênica no palmito de cana-de-açúcar submetidas à restrição hídrica por 17 dias, encontrou duas isoformas de ascorbato peroxidase (APX) diferencialmente expressas, uma cuja expressão foi reprimida do 5º ao 9º dia e induzida do 9º ao 17º dia, e a outra que apresentou indução do 5º ao 9º dia e uma repressão do 9º ao 13º dia.

Um estudo das alterações no conteúdo de ascorbato e glutathiona durante estresse hídrico em *Arabidopsis thaliana* selvagem e mutantes deficientes em ascorbato e glutathiona, durante um período de 10 dias, mostrou um forte aumento de peróxido de hidrogênio e ascorbato em vacúolos nas plantas selvagens, indicando o papel importante do ascorbato na desintoxicação de ERO em vacúolos durante o estresse hídrico (Koffler et al., 2014).

A expressão diferencial da enzima ascorbato peroxidase, sugere que as plantas da cultivar tolerante ativaram uma resposta de defesa contra o estresse oxidativo, provavelmente ocasionado pelo estresse hídrico, a qual não foi observada na cultivar sensível. Além disso, como foi anteriormente observado (Figura 11b), as plantas da cultivar tolerante sob déficit hídrico moderado apresentaram mecanismos relacionados ao transporte mediado por vesículas, manutenção de vesículas citoplasmáticas e processos de endocitose, sendo que, vesículas como vacúolos são utilizados nas células vegetais principalmente para o controle osmótico, estando relacionados com uma resposta inicial frente aos estresse hídrico, ainda mais evidentes nas plantas avaliadas aos 60 e 90 DAT.

Sabe-se que a glutatona ocupa uma posição central na defesa antioxidante, tornando-se uma molécula importante para a proteção das plantas contra vários estresses abióticos, incluindo o déficit hídrico. A função da glutatona como antioxidante que confere tolerância ao estresse abiótico está sendo bem estabelecida, com evidências de muitos estudos de plantas (Hasanuzzaman et al., 2017).

Porém, resultados com relação às respostas gerais ao estresse ainda são extremamente variáveis entre as diferentes plantas, sendo necessário que estudos futuros identifiquem os genes responsáveis pela biossíntese de GSH e os mecanismos celulares subjacentes à regulação de enzimas relacionadas ao metabolismo ou degradação de GSH, além dos eventos de interconversão de vias entre a glutatona e outras moléculas de sinalização durante uma condição de estresse (Hasanuzzaman et al., 2017).

Como apenas os transcritos DEs (p -valor ajustado $<0,1$) na cultivar tolerante sob déficit hídrico severo receberam termos no KEGG, foram caracterizados os mecanismos mais abundantes atribuídos à estes transcritos DEs, relacionados à metabolismos, e, ao mesmo tempo, tentamos compreender como eles se comportam na planta sensível (Figura 15).

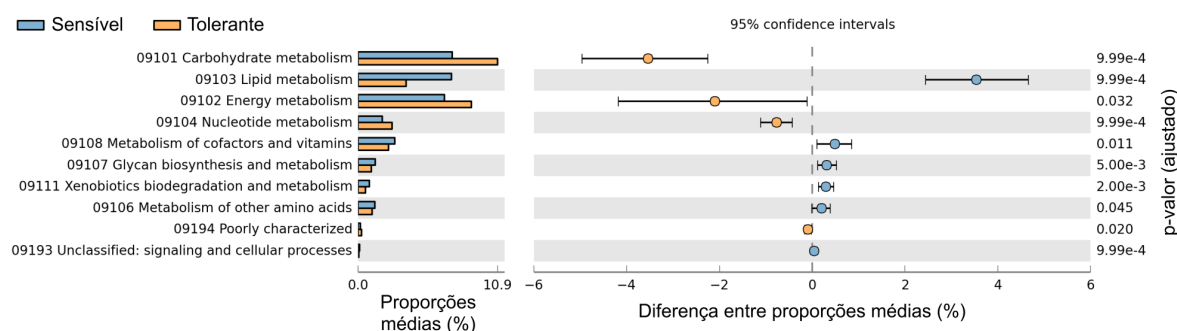


Figura 15: Comparação dos metabolismos entre as cultivares através dos transcritos DEs exclusivos da cultivar tolerante (somente aqueles com p -valor ajustado $<0,05$).

Observa-se que o metabolismo da planta tolerante nas condições de déficit hídrico é mais voltado para os carboidratos e gerenciamento de energia, enquanto a cultivar sensível apresentou maior abundância de termos em se tratando de lipídeos e metabolismo secundário, como cofatores e vitaminas.

Este cenário condiz com o comportamento avaliado da cultivar tolerante, que manteve sua produtividade apesar do déficit hídrico. Por outro lado, a reação da planta sensível sob déficit hídrico pode ser pela deposição de ceras, que são basicamente lipídeos (Shepherd, Griffiths, 2006; Macková et al., 2013), e podem estar mais abundantes na cultivar sensível.

4.2.6 Enzimas diferencialmente expressas no metabolismo da sacarose envolvidas na resposta ao déficit hídrico prolongado

Visto que, na análise de termos do KEGG, a cultivar tolerante apresentou maior abundância de termos para o metabolismo de carboidrato em comparação à sensível, e considerando a importância da sacarose, principal carboidrato de interesse no ciclo produtivo, para a cultura da cana-de-açúcar e para ambas as cultivares, consideradas potenciais em produtividade (Pincelli, Silva, 2012), a dinâmica dos transcritos DEs relacionados ao metabolismo da sacarose, da celulose e do amido, foi comparada entre as duas cultivares, e em ambos os níveis de déficit hídrico separadamente. A Figura 16 destaca as enzimas com expressão diferencial (considerando p -valor $< 0,05$).

O transcrito que codifica a enzima ADPG sintase (EC:2.7.7.27), que converte a α -D-glicose-1P em ADP-glicose, apresentou expressão diferencial induzida aos 30 e 60 DAT na cultivar tolerante sob déficit hídrico severo, e aos 60 e 90 DAT na cultivar sensível. Sob déficit hídrico moderado, apenas a cultivar sensível apresentou expressão diferencial reprimida aos 90 DAT.

A enzima amido sintase (EC:2.4.1.21) converte a ADP-glicose em amilose, e esta só apresentou expressão diferencial aos 90 DAT na cultivar sensível sob déficit hídrico severo, sendo esta expressão induzida. A síntese do amido a partir da amilose foi favorecida pela indução do transcrito que codifica a enzima amilose isomerase (EC:2.4.1.18) aos 60 e 90 DAT na cultivar sensível sob déficit hídrico severo, e reprimida aos 30 e 90 DAT na cultivar tolerante sob o mesmo nível de estresse. Sob déficit hídrico moderado, apenas a cultivar tolerante apresentou expressão diferencial do transcrito que codifica a enzima, aos 90 DAT, sendo a expressão reprimida.

Em geral, o amido não é interessante para o processamento comercial da cana-de-açúcar, apesar de consistir no máximo até 0.05% do conteúdo de açúcar da planta, além de servir como um açúcar de reserva (Santos et al., 2015).

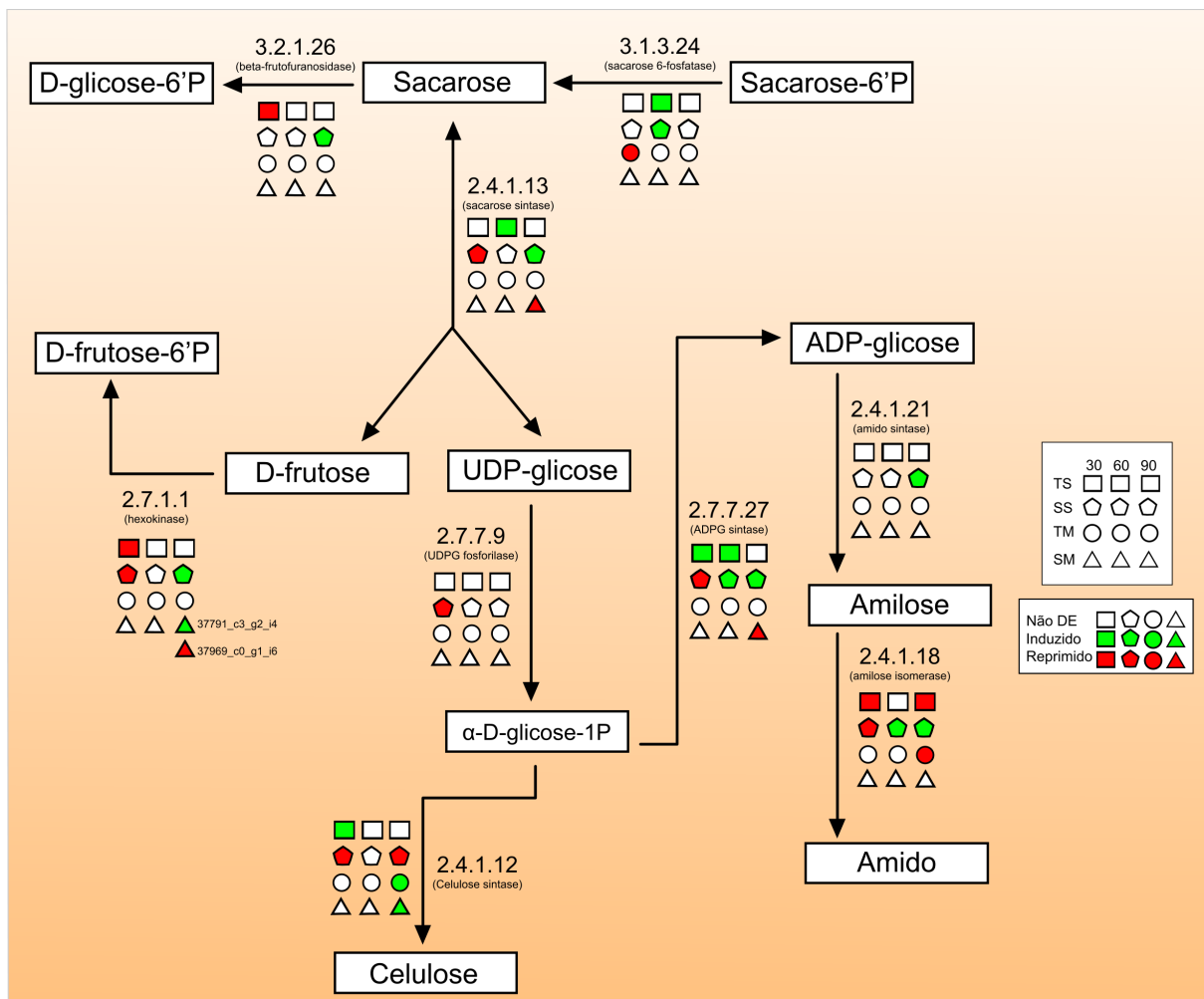


Figura 16: Enzimas (número EC.) DEs nas cultivares tolerante e sensível sob déficit hídrico moderado e severo, mapeadas na via de metabolismo da sacarose (ko00500), adaptadas do KEGG Pathways (p-valor<0,05). TS: cultivar tolerante sob déficit hídrico severo. SS: cultivar sensível sob déficit hídrico severo. TM: cultivar tolerante sob déficit hídrico moderado. SM: cultivar sensível sob déficit hídrico moderado.

Ainda pouco se sabe sobre o papel do amido em uma situação de estresse hídrico. Encontramos que o balanço de íons K^+ e Cl^- e de amido, dentro das células-guarda, regulam a abertura/fechamento dos estômatos (Bianchi et al., 2016).

A manutenção da síntese do amido, que é acompanhada da redução simultânea da concentração de glicose e frutose, resulta em um aumento no potencial osmótico da célula, que está associado ao processo de fechamento dos estômatos (Taiz, Zeiger, 2004). A resposta da planta sensível ao induzir aos 60 e 90 DAT a enzima que sintetiza o amido pode representar um mecanismo de defesa tardio frente ao déficit hídrico, que não é utilizado na resposta da planta tolerante ao déficit hídrico.

A fonte de energia que mais compete com a sacarose são as fibras (celulose), que chegam a compor até 14% em concentração nas cultivares voltadas à produção

de açúcar (Santos et al., 2015). Na rota metabólica da sacarose para a conversão em celulose, encontramos que a cultivar tolerante sob déficit hídrico severo induz a transcrição da enzima celulose sintase (EC:2.4.1.12), responsável pela síntese da celulose, aos 30 DAT, enquanto a cultivar sensível reprime a transcrição dessa enzima aos 30 e 90 DAT. Na condição de déficit hídrico moderado, ambas as cultivares apresentam expressão induzida aos 90 DAT desta mesma enzima.

Além disso, o transcrito que codifica a enzima UDP-glicose fosforilase (EC:2.7.7.9), responsável pela modificação da UDP-glicose em -D-glicose-1P, o precursor da celulose, tem sua expressão reprimida aos 30 DAT na cultivar sensível sob déficit hídrico severo, e não apresentou expressão diferencial na cultivar tolerante em nenhuma das datas avaliadas.

Uma revisão sobre as respostas da parede celular de plantas a vários estresses abióticos, incluindo a seca, mostrou que plantas tendem a aumentar a expressão de genes que codificam a UDP-glicose fosforilase em resposta ao déficit hídrico (Gall et al., 2015; Zheng et al., 2014), diferentemente do resultado aqui observado, no qual a cultivar tolerante não apresentou expressão diferencial para a UDP-glicose fosforilase e, portanto, deve apresentar outros mecanismos para tolerância ao déficit hídrico, e a cultivar sensível apresentou a expressão reprimida aos 30 DAT, podendo este comportamento estar relacionado à maior sensibilidade desta cultivar ao déficit hídrico.

Ainda na via metabólica da sacarose, observamos que aos 60 DAT, sob déficit hídrico severo, ambas as cultivares apresentaram transcritos DEs induzidos que codificam a enzima sacarose 6-fosfatase (EC:3.1.3.24), responsável pela conversão da Sacarose-6'P em Sacarose, confirmando o fato de que ambas as cultivares são consideradas potenciais em produtividade (Pincelli, Silva, 2012).

Além disso, a expressão do transcrito que codifica a enzima sacarose sintase (EC:2.4.1.13) está induzida aos 60 DAT na cultivar tolerante sob déficit hídrico severo, e sob o mesmo estresse, na cultivar sensível, a expressão está reprimida aos 30 DAT e induzida aos 90 DAT, podendo estar atuando tanto para a produção da celulose, como da sacarose. Esta enzima também foi observada com aumento da sua expressão em resposta ao estresse hídrico (Gall et al., 2015; Zheng et al., 2014).

Esses resultados podem sugerir que uma das respostas da planta tolerante ao déficit hídrico é a síntese da celulose, mantendo a integridade da parede celular e a pressão de turgência, permitindo o crescimento contínuo das células mesmo sob déficit (Ricardi et al., 2014), assim mantendo a síntese de sacarose, que a torna viável economicamente (produtiva).

4.2.7 Relação dos transcritos DEs na resposta ao déficit hídrico prolongado

Para compreender melhor o comportamento da cana-de-açúcar sob déficit hídrico, além da análise de expressão diferencial das enzimas das rotas metabólicas da sacarose/celulose/amido, que podem estar envolvidas na resposta ao déficit hídrico, foi realizada uma outra análise exploratória dos transcritos DEs (p -valor $<0,05$) através da investigação dos principais mecanismos de resposta à seca, especialmente aqueles que aparecem DEs em mais de uma das datas avaliadas.

Para isso, os transcritos DEs foram anotados diretamente nas proteínas (ortólogos) do DroughtDB (Alter et al., 2015), banco de dados de referência curada que disponibiliza sequências de ortólogos associados experimentalmente a diversos papéis na resposta ao déficit hídrico, em diversas plantas.

Assim, o número de transcritos DEs, induzidos ou reprimidos pelo estresse hídrico, que receberam anotação no DroughtDB, de ambas as cultivares, datas de avaliação e níveis de déficit hídrico em relação as plantas controle, estão apresentados na Tabela 7 e na Tabela 8.

Entre os principais ortólogos encontrados, tanto para a cultivar tolerante, como para a sensível, está o receptor do tipo quinase RPK1, apresentando um total de 98 transcritos DEs que receberam essa anotação em ambos os níveis de déficit hídrico e nas três datas de avaliação, na cultivar tolerante, e 90 transcritos DEs na sensível. Este receptor, que possui repetições ricas em leucina (LRR), está envolvido nas vias de transdução de sinal, regulando as respostas do ácido abscísico (ABA) (Alter et al., 2015).

Sob uma condição de déficit hídrico moderado, a cultivar tolerante apresentou números de transcritos DEs induzidos e reprimidos semelhantes aos 30 e 60 DAT, sendo que aos 90 DAT apenas apresentou transcritos reprimidos. Já sob déficit hídrico severo, um maior número de transcritos apresentaram-se DEs induzidos aos 30 DAT, e aos 60 e 90 DAT o número de transcritos DEs induzidos e reprimidos, anotados como RPK1, foi igual.

A cultivar sensível, sob déficit hídrico moderado, apresentou um maior número de transcritos DEs induzidos aos 30 dias, que foi diminuindo ao longo do tempo e houve um aumento o número de transcritos DEs reprimidos, principalmente aos 90 DAT. Sob déficit hídrico severo, o maior número de transcritos DEs reprimidos ocorreu aos 30 DAT.

Hong et al. (1997) descreveu que o RPK1 é um receptor do tipo quinase, que possui repetições ricas em leucina, isolado de *Arabidopsis thaliana*, e que a sua expressão é induzida por ABA, desidratação, alto teor de sal e baixa temperatura.

Tabela 8: Número de transcritos DEs na cultivar sensível sob déficit hídrico moderado e severo, em relação ao controle, que receberam anotação nos mecanismos de resposta à seca do DroughtDB (Alter et al., 2015) (p-valor<0,05). I: número de transcritos induzidos, R: número de transcritos reprimidos. M: déficit hídrico moderado. S: déficit hídrico severo.

Mecanismo DE	30I		30R		60I		60R		90I		90R	
	M	S	M	S	M	S	M	S	M	S	M	S
AAO3	0	0	0	1	0	0	0	0	0	0	0	0
ABCG25	0	0	0	1	0	0	0	0	0	0	0	0
ABCG40	1	1	0	4	0	1	0	0	3	4	2	2
ABO1/ELO1	0	0	0	1	0	0	0	0	0	0	0	0
ADAP	1	0	0	0	0	0	0	0	0	0	0	0
AGO1	0	0	0	1	2	1	2	3	0	1	1	1
AHK1	0	0	0	2	0	0	0	0	0	0	2	1
ALDH22A1	1	1	0	0	0	0	0	0	1	2	1	0
ALDH311	0	0	0	1	0	0	0	0	0	0	0	0
AtBG1	1	2	0	2	0	1	0	0	2	1	2	0
AtrbohD	0	0	0	1	0	0	0	1	0	0	1	0
AtrbohF	0	0	0	0	0	0	0	0	0	0	0	0
AVP1	0	0	0	0	0	0	0	0	0	0	1	2
CAU1	0	0	0	1	0	0	0	0	0	0	0	0
CBP20	0	0	0	0	0	0	1	0	1	0	0	0
CDPK7	1	0	0	0	0	0	0	0	0	0	0	1
CIPK03	0	0	1	4	0	0	0	0	0	3	4	1
CIPK12	0	0	0	0	0	0	0	0	1	0	0	2
CIPK15	0	0	0	0	0	0	0	1	0	0	0	0
CLCc	2	1	0	0	0	0	0	0	0	0	0	0
CPK21	0	0	0	2	1	0	0	0	1	0	0	0
CPK23	1	0	0	0	0	0	0	0	0	0	0	1
CYP707A3	0	1	0	2	0	1	0	0	2	0	0	1
EDT1/HDG11	0	0	0	1	0	0	0	0	0	0	0	0
ERD1	0	0	0	1	0	1	0	0	3	1	1	1
FAR1	0	3	0	1	0	0	0	2	2	0	2	1
FHY3	0	0	0	0	0	0	0	0	0	0	0	1
GF14λ	0	0	0	0	1	1	0	0	0	0	0	0
GL1-2	0	0	0	0	0	1	0	0	0	3	0	0
GORK	0	0	0	1	0	0	0	1	1	0	0	0
GTG1	0	0	1	1	0	0	0	0	0	0	0	0
HD2C	0	0	0	0	0	0	0	0	0	1	1	0
HDA19	0	0	0	0	0	0	0	0	1	0	0	0
HSFA1b	0	0	0	0	0	0	0	0	0	2	1	0
KAT2	0	0	0	2	0	0	0	0	0	0	0	0
LEW1	0	0	0	0	0	0	0	1	0	0	0	0
MRP4	0	0	0	8	0	3	0	0	1	1	1	3
MYB2	0	0	0	0	0	0	1	0	0	0	0	0
NADP-ME	1	0	0	5	0	0	0	0	1	3	0	1
NCED	1	0	0	0	0	0	0	0	0	0	0	0
NFYA5	0	0	0	0	0	0	0	1	0	0	0	0
NHO1	0	0	1	1	0	0	0	0	0	0	0	0
OsbZIP71	0	0	0	0	0	0	0	0	0	1	1	0
OST1/SRK2E	0	0	0	0	0	0	0	0	1	1	1	0
OST2	0	0	0	2	0	0	0	4	0	1	0	1
PARP1	0	0	0	2	0	0	1	0	0	0	0	0
PP2C	0	1	0	0	0	0	0	0	0	0	0	0
RPK1	9	5	1	17	1	7	3	3	4	13	13	14
Sl0GS1	0	0	0	0	0	0	0	0	0	0	0	1
SRK2C	0	0	0	0	0	0	0	0	0	1	0	0
TPS1	0	0	0	0	0	0	0	0	0	0	0	2
UGE-1	0	0	1	0	0	0	0	0	0	1	1	0
ZmCPK4	0	0	0	0	0	1	0	0	1	0	0	0

RPK1 também foi descrito como molécula reguladora chave da membrana, envolvido na principal via de sinalização do ABA e na percepção precoce de ABA, regulada

positivamente, sendo esta regulação demonstrada em testes em que a repressão da expressão de RPK1 na *A. thaliana* diminuiu a sensibilidade ao ABA durante a germinação, crescimento e fechamento estomático (Osakabe et al., 2005).

Desta maneira, na cana-de-açúcar, o maior número de transcritos DEs induzidos aos 30 DAT sob déficit hídrico severo, e reprimidos aos 90 DAT sob déficit hídrico moderado, demonstra um papel importante da RPK1 na resposta ao déficit hídrico para o genótipo tolerante de forma precoce, enquanto para a cultivar sensível, apresentou-se reprimido, podendo indicar que a expressão deste receptor não faz parte da resposta desta cultivar frente ao déficit hídrico, ou então que a expressão reprimida deste receptor torna o genótipo mais sensível ao déficit hídrico.

Um transportador dependente de ATP, que no banco DroughtDB recebeu a identificação ABCG40, também foi anotado em ambas as cultivares. Este transportador foi relacionado à importação de ABA para as células em *Arabidopsis thaliana*.

Segundo Kang et al. (2010), a captação de ABA em células de levedura e BY2 expressando ABCG40 de *A. thaliana* foi aumentada, enquanto a captação de ABA em protoplastos de plantas em que a expressão de ABCG40 foi nocauteada, foi diminuída em comparação às células de controle. Os estômatos dos mutantes *atabcg40* com perda de função fecharam mais lentamente em resposta ao ABA, resultando na redução da tolerância à seca, ou seja, quando este gene é nocauteado, a regulação positiva de genes responsivos ABA é atrasada, os estômatos se fecham mais lentamente.

A presença de ABCG40 está associada a um fechamento mais rápido dos estômatos, e assim, a sua expressão encontra-se em maior número de transcritos DEs induzidos aos 30 e 60 DAT na cultivar tolerante sob déficit hídrico severo, enquanto que sob déficit hídrico moderado apresentou indução aos 90 DAT, possivelmente quando as plantas desta cultivar começam a sentir os efeitos do estresse moderado e a gerar mais respostas de proteção contra a desidratação.

Já na planta sensível sob déficit hídrico severo, os transcritos DEs apresentaram-se reprimidos aos 30 DAT. A repressão da expressão deste gene pode estar relacionada com o fechamento mais lento dos estômatos, e conseqüentemente maior sensibilidade deste genótipo ao déficit hídrico. O fechamento estomatal precoce parece ser mais interessante para a planta, visto que o déficit hídrico é mantido por todo o tratamento.

Outro gene com mais transcritos anotados foi o MRP4, 13 transcritos DEs na cultivar tolerante e 17 transcritos DEs na cultivar sensível. As plantas da cultivar sensível sob déficit hídrico severo apresentaram o maior número de transcritos DEs anotados como MRP4, aos 30 DAT, sendo a expressão reprimida. Já as plantas da cultivar tolerante não apresentaram grande variação entre o número de transcritos DEs induzidos e reprimidos em ambos os níveis de déficit hídrico.

Este gene MRP4 codifica proteínas transmembrana transportadoras dependente de ATP. Sua ausência (nockout) foi documentada como causadora da perda de água pela abertura estomática desregulada e consequente aumento da sensibilidade à seca em *A. thaliana* (Klein et al., 2004), indicando o envolvimento do AtMRP4 na complexa regulação da abertura estomática.

No caso das plantas de cana-de-açúcar, o genótipo tolerante parece não precisar induzir a expressão gênica do MRP4 para uma resposta frente ao déficit hídrico, possivelmente ativando outras vias de resposta, enquanto no genótipo sensível, em que a expressão é reprimida já aos 30 DAT de estresse, este mecanismo pode estar associado à maior sensibilidade da cultivar RB855453 frente ao estresse pela abertura estomática desregulada.

Além dos transcritos discutidos, as tabelas apresentam vários outros ortólogos anotados que apresentam expressão reprimida ou induzida em ambas as cultivares avaliadas e que podem estar envolvidos com as respostas das plantas de cana-de-açúcar frente as situações de déficit hídrico.

4.3 Validação dos resultados do RNA-Seq pela técnica de PCR quantitativa em tempo real (RT-qPCR)

A boa qualidade e integridade dos "pools" de RNA total (RINs entre 6,8 e 7,8) (Figura 17), e das amostras de cDNA obtidas através do NanoDrop (Tabela 9) permitiu dar início à técnica de RT- qPCR.

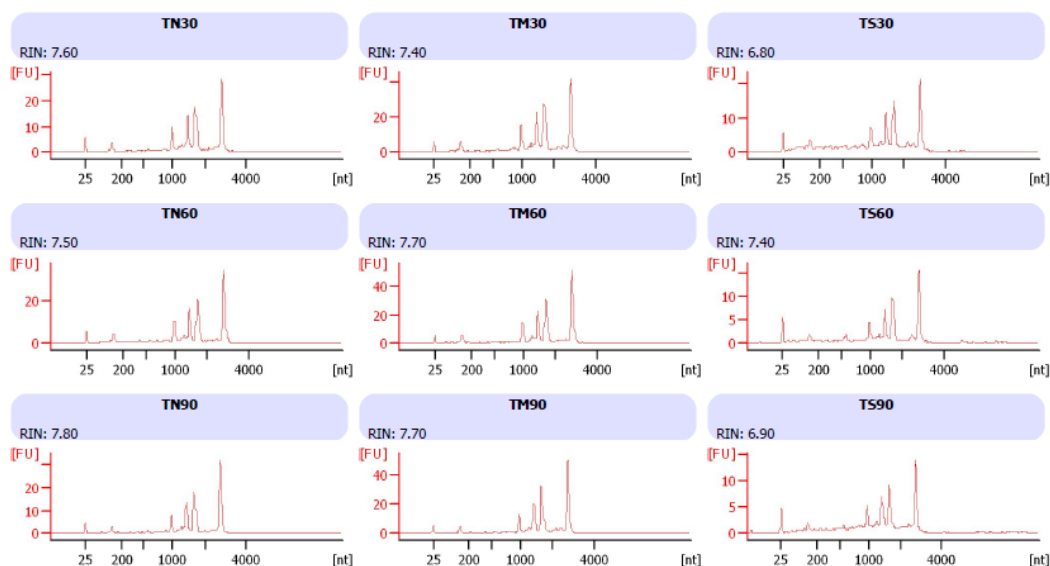


Figura 17: Eletroferograma e valor de RIN dos "pools" de RNA total das folhas de cana-de-açúcar da cultivar tolerante, em todos os tempos e níveis de déficit hídrico, geradas pelo Bioanalyzer (Agilent 2100 Bioanalyzer).

Tabela 9: Resultados obtidos através da leitura das amostras de cDNA no equipamento NanoDrop. TC: plantas controle da cultivar tolerante. TM: cultivar tolerante sob déficit hídrico moderado. TS: cultivar tolerante sob déficit hídrico severo. 30, 60 e 90: datas de avaliação após a aplicação dos tratamentos.

Amostras	Concentração (ng/ μ L)	A260/A280nm	A260/230nm
TC30	732,2	1,56	1,94
TM30	743,4	1,58	1,89
TS30	727,9	1,58	1,73
TC60	733,4	1,57	1,94
TM60	778,0	1,58	1,88
TS60	715,6	1,56	1,89
TC90	748,3	1,57	1,88
TM90	778,1	1,59	1,75
TS90	729,0	1,56	1,89

A titulação dos oligonucleotídeos iniciadores, de todos os genes (transcritos alvo e endógenos), mostrou-se mais eficiente para determinadas concentrações. O programa utilizado para analisar os dados obtidos gerou gráficos das curvas de dissociação (“Melt curve”) para as concentrações (Figura 18), e assim, foi determinado a melhor concentração de oligonucleotídeos para cada um dos genes, indicadas na Tabela 10. Esta etapa de titulação dos oligonucleotídeos iniciadores é necessária para a confirmação da especificidade do nucleotídeo à molécula alvo.

Após a análise e determinação da melhor concentração de oligonucleotídeos iniciadores para cada um dos genes, foi realizada a construção de uma curva padrão através de diluições seriadas de cDNA (“pool” de amostras) para verificação da eficiência de amplificação da reação.

Todos os genes analisados apresentaram valores de R² igual a 0,99 e EFF%: entre 90 e 110%, mostrando-se eficientes (Figura 19 e Figura 20). A concentração de 100 ng apresentou o melhor resultado para todos os genes em questão.

Tabela 10: Concentrações dos oligonucleotídeos dos transcritos-alvo e genes de referência usados para validação dos resultados do RNA-Seq por meio de RT-qPCR.

Oligonucleotídeos titulados	Concentração (F/R, nM)
Endógeno: 14-3-3	300/300
Endógeno: Actina	600/600
Endógeno: β -Tubulina	600/600
Alvo: Peroxidase extracelular	300/300
Alvo: Glutathiona S-transferase	300/300
Alvo: Poligalacturonase	600/600
Alvo: Receptor do tipo quinase (GbRLK)	300/300

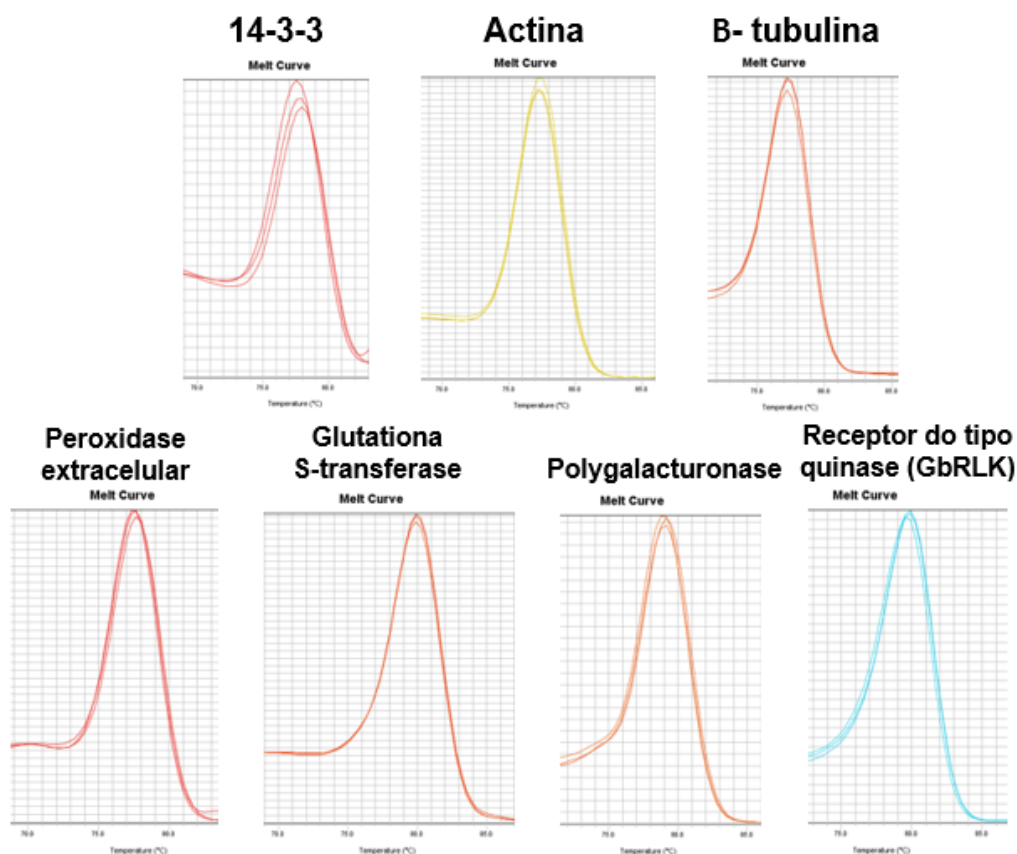


Figura 18: Curva de “melting” dos genes 14-3-3, Actina, β -Tubulina, Peroxidase extracelular, Glutathione S-transferase, Poligalacturonase e Receptor do tipo quinase (GbRLK). A presença de um único pico após a curva de dissociação confirmou a especificidade de cada oligonucleotídeo.

O programa Suite Software v1.01 (“Applied Biosystems”, Foster, CA, USA), utilizado para as análises, gera valores de “score” para os genes, o qual corresponde à medida da estabilidade da expressão de um gene de referência em comparação a todos os genes de referência utilizados na análise (MANUAL EXPRESSION SUITE SOFTWARE v 1.0.3 *apud* (Fonseca, 2016).

Estes valores de “score”, quando menores que 1,5, indicam que não houve diferença na expressão desses genes entre os diferentes tratamentos e, por isso, podem ser utilizados como genes de referência para normalização dos resultados de RT- qPCR (Vandesompele et al., 2002).

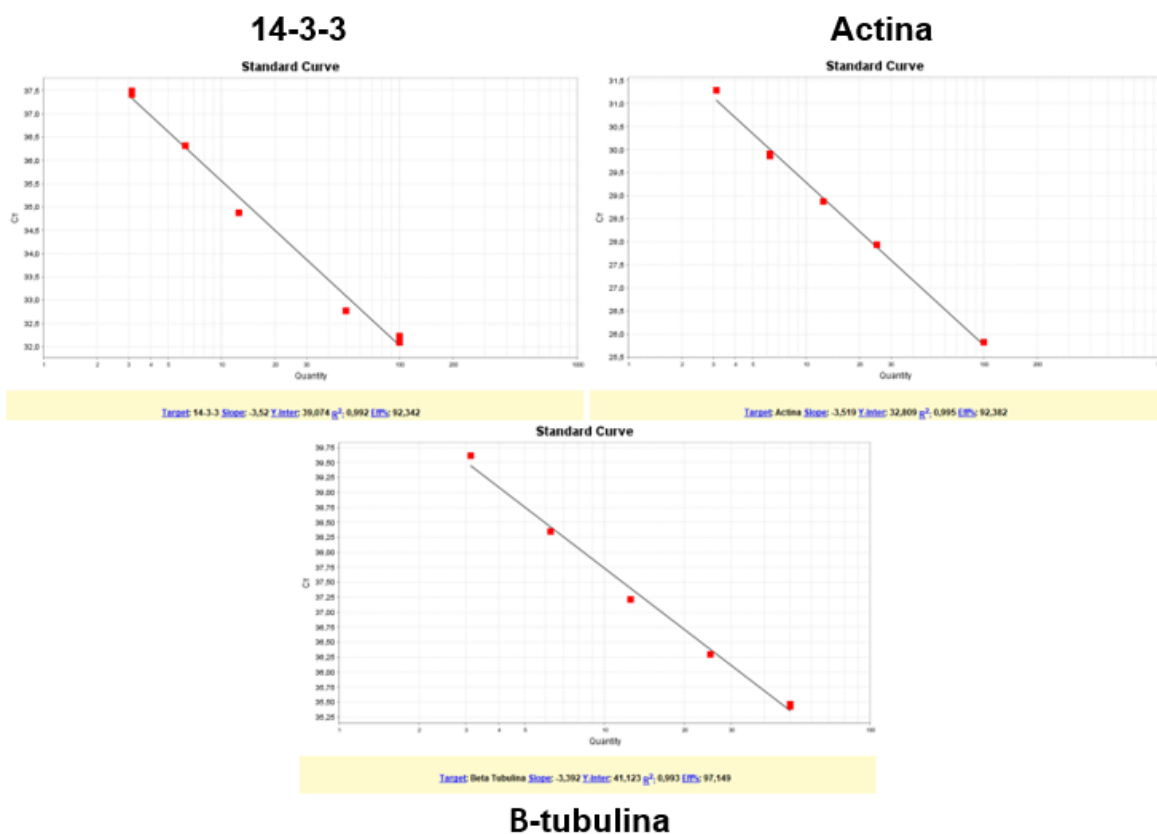


Figura 19: Curva de diluição dos genes endógenos.

Para a normalização da expressão dos genes alvos, foram testados como controles endógenos os genes codificadores da proteína 14-3-3, da actina e da β -Tubulina. Dos três genes testados para esta finalidade, a actina, apesar de apresentar o maior valor de estabilidade entre os três genes ("score": 1,42) foi indicado como melhor endógeno. Isso aconteceu porque, além do score inferior à 1,5, ao verificarmos a amplificação do delta Ct, foi observado que o seu Ct foi próximo à 27 ciclos, sendo que a proteína 14-3-3 e a β -Tubulina apresentaram Ct superiores a 35 ciclos, o que não é considerado favorável para a normalização das análises de RT-qPCR.

A proteína 14-3-3 e a β -Tubulina, apesar de apresentarem "scores" inferiores à actina ("score" β -Tubulina: 1,22 e 14-3-3: 1,21), e de serem de muito utilizadas na literatura para normalização de expressão gênica, não se mostraram muito adequadas neste caso, pois apresentaram variação de expressão ao longo do experimento.

Os resultados obtidos pela RT-qPCR apresentaram algumas diferenças em relação aos resultados obtidos pela técnica de RNA-Seq. Foram testados transcritos (genes) que apresentaram expressão diferencial exclusivamente na cultivar tolerante (p -valor < 0,05).

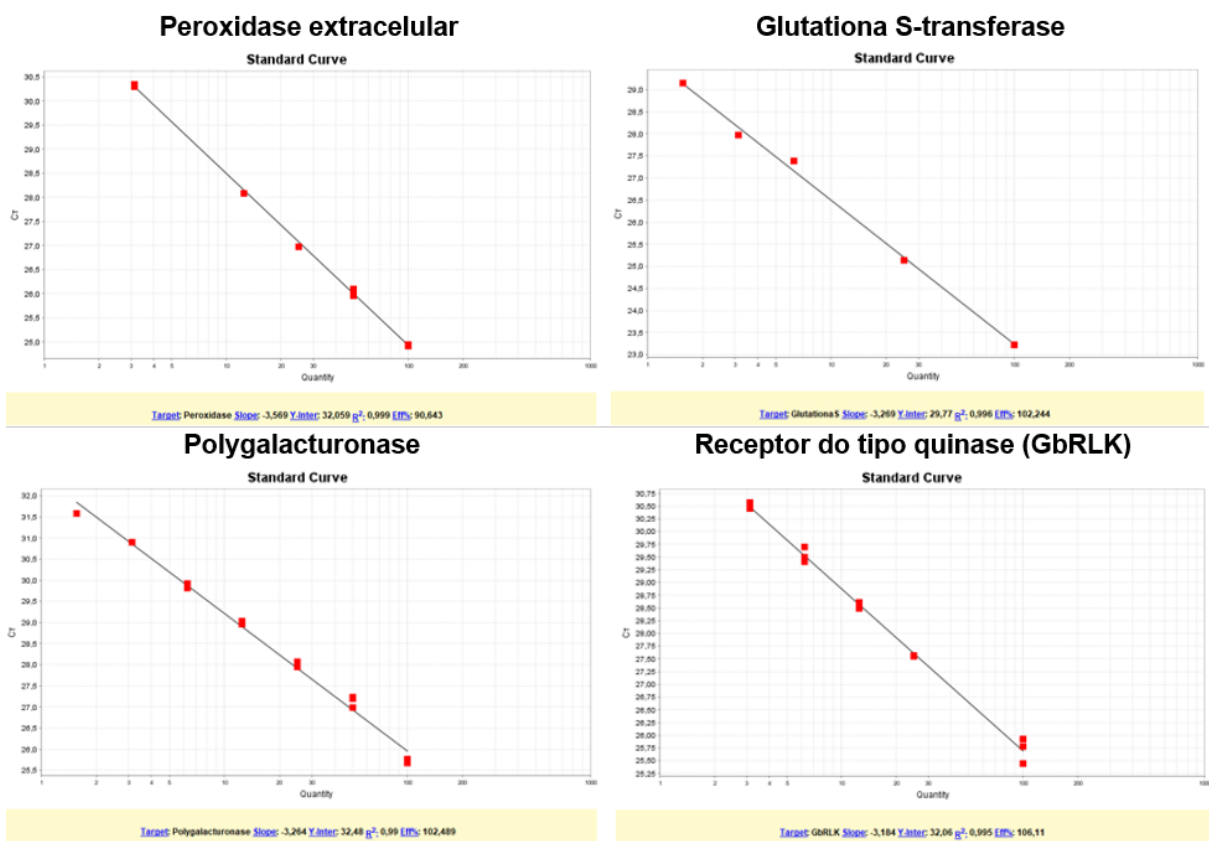


Figura 20: Curva de diluição dos transcritos-alvo.

A expressão do transcrito que codifica a Peroxidase extracelular apresentou-se DE induzida aos 90 DAT sob déficit hídrico severo em ambas as análises. Além disso, a expressão do transcrito Poligalacturonase, apresentou-se reprimida aos 30 DAT sob déficit hídrico severo também em ambas as análises. Os valores de L2FC obtidos pelas técnicas de RNA-Seq e RT-qPCR estão apresentados na Figura 21.

Já a expressão dos transcritos que codificam a expressão de um Receptor do tipo quinase, e da enzima Glutaciona S-transferase, apresentou variações entre as duas técnicas (Figura 22 e 23).

Na análise de expressão diferencial utilizando o programa DESeq2, a expressão do transcrito que codifica um Receptor do tipo quinase apresentou-se reprimida aos 30 DAT sob déficit hídrico moderado, e aos 60 DAT sob ambos os níveis de déficit hídrico. Contudo, na RT-qPCR, este transcrito apresentou-se também reprimido aos 30 DAT sob déficit hídrico moderado, porém induzido aos 60 DAT sob ambos os níveis de déficit hídrico, apresentando perfis de expressão divergentes entre as técnicas (Figura 22).

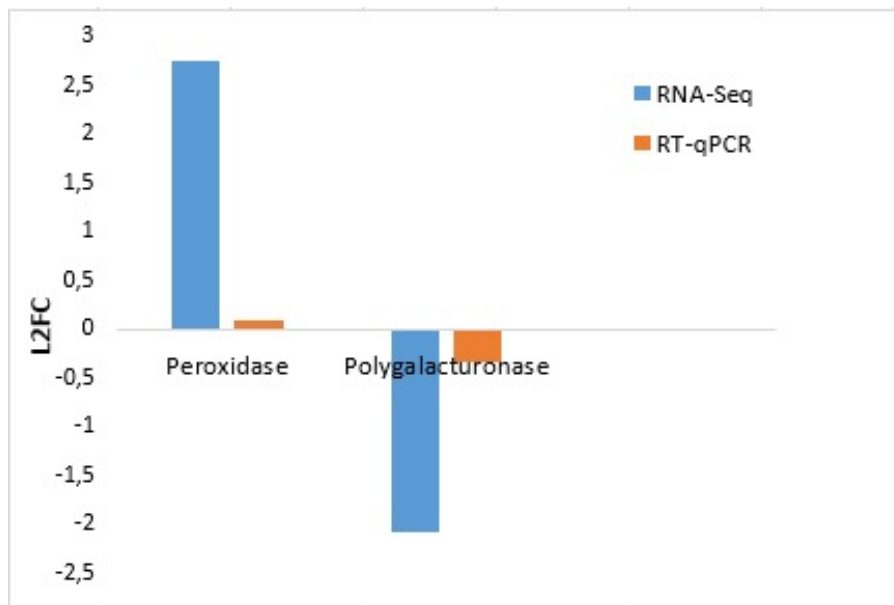


Figura 21: Comparação dos valores de expressão, obtidos por RNA-Seq e RT-qPCR em dois transcritos diferencialmente expressos.



Figura 22: Comparação dos valores de expressão diferencial do transcrito Receptor do tipo quinase, obtidos por RNA-Seq e RT-qPCR.

A análise de RNA-Seq do transcrito que codifica a enzima Glutathione S-transferase mostrou expressão diferencial reprimida os 60 DAT sob déficit hídrico severo, enquanto que a RT-qPCR mostrou sua expressão levemente induzida (Figura 23).

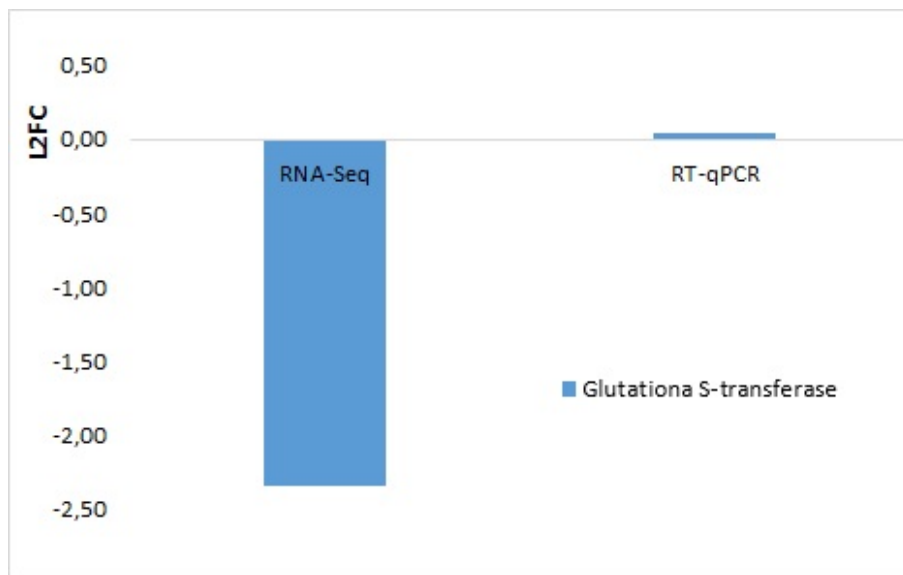


Figura 23: Comparação dos valores de expressão diferencial do transcrito Glutathione S-transferase, obtidos por RNA-Seq e RT-qPCR.

Ambas as técnicas, RNA-Seq e RT-qPCR, são confiáveis. Entretanto, neste trabalho elas utilizaram análises diferentes: a RT-qPCR foi realizada utilizando um pool das amostras, enquanto que as análises geradas a partir do RNA-Seq foram feitas usando as amostras individuais. Essa diferença pode levar à variações entre os resultados finais. Além disso, o teste estatístico utilizado para a escolha dos transcritos a serem validados (p -valor) menor que 0,05, pode não ser confiável na eliminação de falsos positivos.

Outra explicação para essas variações deve-se ao fato de estarmos trabalhando com a cultura da cana-de-açúcar, cujo genoma é altamente complexo e poliploide. Assim, após o desenho dos oligonucleotídeos iniciadores, os mesmos foram alinhados na montagem *de novo* realizada com o transcriptoma, e apresentaram alinhamentos únicos nas sequências alvo. Porém, nos casos em que houve variação entre os resultados, pode ter ocorrido a amplificação de sequências de cromossomos diferentes em que a expressão daquele gene variou em comparação ao que foi observado na análise de RNA-Seq.

Conclusões

Com base nos resultados obtidos neste trabalho, pode-se concluir que:

A análise de RNA-Seq foi utilizada para identificar um grande número de transcritos nas folhas das cultivares de cana-de-açúcar contrastantes à seca, frente ao déficit hídrico prolongado, permitindo a avaliação dos perfis de expressão gênica em resposta à seca.

Foi observado que as cultivares utilizam mecanismos distintos para lidar com o déficit hídrico:

1. apresentam muito mais transcritos DEs exclusivos em comparação ao número de DEs compartilhados;
2. isso é ainda mais acentuado com ao aumento da intensidade do déficit hídrico;
3. esses diferentes mecanismos foram observados a nível de processos biológicos, rotas metabólicas e no metabolismo de carboidratos.

Os resultados de expressão diferencial obtidos também revelam a indução de enzimas que podem estar relacionadas com as respostas da cultivar de cana-de-açúcar tolerante (SP81-3250) frente ao déficit hídrico, através da síntese da sacarose e celulose, e da produção da glutathione, sendo que a glutathione e antioxidantes relacionados já foram observados desempenhando um papel protetor no estresse biótico em várias plantas.

Os transcritos diferencialmente expressos que não foram classificados ("no hits" - sem anotação) são importantes e podem estar envolvidos no entendimento dos mecanismos moleculares de tolerância de plantas e no desenvolvimento de novas variedades mais tolerantes à deficiência hídrica.

6. Referências bibliográficas

Adams J (2008) Transcriptome: connecting the genome to gene function. **Nature Education** 1:195.

Aitken KS, McNeil MD, Berkman PJ, Hermann S, Kilian A, Bundock PC, Li J (2014) Comparative mapping in the Poaceae family reveals translocations in the complex polyploid genome of sugarcane. **BMC Plant Biology** 14:190.

Alter S, Bader KC, Spannagl M, Wang Y, Bauer E, Schön CC, Mayer KF (2015) DroughtDB: an expert-curated compilation of plant drought stress genes and their homologs in nine species. **Database** 2015:1-7.

Ambrosone A, Giacomo M D, Leone A, Grillo MS, Costa A (2013) Identification of early induced genes upon water deficit in potato cell cultures by cDNA-AFLP. **Journal of Plant Research** 126:169–178.

Asada K (1994) Production and action of active oxygen species in photosynthetic tissues. In: Foyer CH, Mullineaux PM (Eds.) **Causes of photooxidative stress and amelioration of defense systems in plants**. New York: CRC Press, p. 77–104.

Ashburner M, Ball CA, et al. (2000) Gene ontology: tool for the unification of biology. **Nature Genetics** 25:25.

Basnayake J, Jackson P, Inman-Bamber N, Lakshmanan P (2012) Sugarcane for water-limited environments. Genetic variation in cane yield and sugar content in response to water stress. **Journal of Experimental Botany** 63:6023–6033.

Belesini A, Carvalho F, Telles B, Castro G D, Giachetto P, Vantini J, Carlin S, Cazetta J, Pinheiro D, Ferro M (2017) De novo transcriptome assembly of sugarcane leaves submitted to prolonged water-deficit stress. **Genetic Molecular Research** 16:10–4238.

Belesini AA (2015) **Análise do transcriptoma de folhas de cana-de-açúcar submetidas à prolongada limitação hídrica usando RNA-Seq**. 140 f. Dissertação (Mestrado em Agronomia – Genética e Melhoramento de Plantas) – Unesp, Jaboticabal.

Bianchi L, Germino GH, Silva MdeA (2016) Adaptação das Plantas ao Déficit Hídrico. **Acta Iguazu** 5:15–32.

Biazuzo MMdeA (2013) **Análise da expressão diferencial de genes relacionados à tolerância ao déficit hídrico em feijoeiro comum**. 179 f. Tese (Doutorado em Ciências) — Universidade de São Paulo, Piracicaba.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics** 30:2114–2120.

Bray EA (1997) Plant responses to water deficit. **Trends in plant science** 2:48–54.

Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. **Nature Biotechnology** 34:525.

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. **Nature methods** 12:59.

Bustin SA, Mueller R (2005) Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. **Clinical Science** 109:365–379.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. **BMC Bioinformatics** 10:421.

Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, Hub A, Group WPW (2008) AmiGO: online access to ontology and annotation data. **Bioinformatics** 25:288–289.

Cardoso-Silva CB, Costa EA, Mancini MC, Balsalobre TWA, Canesin LEC, Pinto LR, Carneiro MS, Garcia AAF, Souza AP de, Vicentini R (2014) De novo assembly and transcriptome analysis of contrasting sugarcane varieties. **PloS one** 9:e88462.

Carvalho MCdCG, Silva DCGd et al. (2010) Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural** 735–744.

Cattivelli L, Rizza F, Badeck FW, Mazzucotelli E, Mastrangelo AM, Francia E, Marè C, Tondelli A, Stanca AM (2008) Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. **Field Crops Research** 105:1–14.

Chaves MM, Maroco JP, Pereira JS (2003) Understanding plant responses to drought—from genes to the whole plant. **Functional Plant Biology** 30:239–264.

Chen JH, Jiang HW, Hsieh EJ, Chen HY, Chien CT, Hsieh HL, Lin TP (2012) Drought and salt stress tolerance of an *Arabidopsis* glutathione S-transferase U17 knockout mutant are attributed to the combined effect of glutathione and abscisic acid. **Plant Physiology** 158:340–351.

Cheng MC, Ko K, Chang WL, Kuo WC, Chen GH, Lin TP (2015) Increased glutathione contributes to stress tolerance and global translational changes in *Arabidopsis*. **The Plant Journal** 83:926–939.

Cia M, Guimarães A, Medici L, Chabregas S, Azevedo R (2012) Antioxidant responses to water deficit by drought-tolerant and-sensitive sugarcane varieties. **Annals of Applied Biology** 161:313–324.

Collins FS, Green ED, Guttmacher AE, Guyer MS (2003) A vision for the future of genomics research. **Nature** 422:835.

Conab (2019) **Acompanhamento da safra brasileira de Cana-de-açúcar, v. 5 - Safra 2018/19, n. 4 - Quarto levantamento, abril/2019**. Brasília: CONAB, 1-75.

Costa AM (2011) Prospecção gênica e bioinformática. In: Faleiro FG, Andrade SRM, Junior FBR (Eds.) **Biociologia: estado da arte e aplicações na agropecuária**. Planaltina: Embrapa Cerrados, p. 730.

Costa V, Casamassimi A, Ciccodicola A (2010) Nutritional genomics era: opportunities toward a genome-tailored nutritional regimen. **The Journal of Nutritional Biochemistry** 21:457–467.

Delgado AA, Cesar MAA (1977) **Elementos de tecnologia e engenharia do açúcar de cana**. Piracicaba: Esalq – Departamento de Tecnologia Rural, p. 364.

Devarumath RM, Kalwade SB, Kulkarni PA, Sheelavanthmath SS, Suprasanna P. (2013) Integrating OMICS approaches in sugarcane improvement. In: (Ed.) **OMICS Applications in Crop Science**. Florida: CRC Press, p. 191–250.

D'Hont A, Grivet L, Feldmann P, Glaszmann J, Rao S, Berding N (1996) Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. **Molecular and General Genetics** 250:405– 413.

D'Hont A, Lu YH, León DGd, Grivet L, Feldmann P, Lanaud C, Glaszmann JC (1994) A molecular approach to unraveling the genetics of sugarcane, a complex polyploid of the Andropogoneae tribe. **Genome** 37:222–230.

D'Hont A, Rao P, Feldmann P, Grivet L, Islam-Faridi N, Taylor P, Glaszmann JC (1995) Identification and characterisation of sugarcane intergeneric hybrids, *Saccharum officinarum* x *Erianthus arundinaceus*, with molecular markers and DNA in situ hybridization. **Theoretical and Applied Genetics** 91:320–326.

Duvick J, Fu A, Muppirala U, Sabharwal M, Wilkerson MD, Lawrence CJ, Lushbough C, Brendel V (2007) PlantGDB: a resource for comparative plant genomics. **Nucleic Acids Research** 36:D959–D965.

D'Hont A (2005) Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. **Cytogenetic and Genome Research** 109:27–33.

Embrapa - Empresa Brasileira de Pesquisa Agropecuária. **Fenologia da Cana-de-açúcar**. Disponível em: <<http://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG016822122006154840.html>>. Acesso em: 30 jun. 2019.

Embrapa. Empresa Brasileira de Pesquisa Agropecuária. **Agroecologia da Cana-de-Açúcar**. Disponível em: <http://www.cana.cnpm.embrapa.br/agroeco.html>. Acesso em: 25 jun. 2019.

Fragoso RR, Nakasu YET, Rocha TL, Rosa AJM (2011) Genômica Funcional. In Faleiro FG, Andrade SRM de, Junior FBdosR (Eds.) **Biociologia: estado da arte e aplicações na agropecuária**. Planaltina: Embrapa Cerrados, p. 730.

Ferreira TH, Gentile A, Vilela RD, Costa GGL, Dias LI, Endres L, Menossi M (2012) microRNAs associated with drought response in the bioenergy crop sugarcane (*Saccharum* spp.) **PLoS One** 7:e46703.

Ferreira TH, Tsunada MS, Bassi D, Araújo P, Mattiello L, Guidelli GV, Righetto GL, Gonçalves VR, Lakshmanan P, Menossi M (2017) Sugarcane water stress tolerance mechanisms and its implications on developing biotechnology solutions. **Frontiers in Plant Science** 8:1077.

Ferro MIT (2008) **Estudo do estresse hídrico em cana-de-açúcar utilizando genômica funcional**. 168 f. Tese (Livre Docência) – Unesp, Jaboticabal.

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M et al. (2016) InterPro in 2017—beyond protein family and domain annotations. **Nucleic Acids Research** 45:D190–D199.

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. **Nucleic Acids Research** 39:W29–W37.

Fonseca LFS (2016) **Expressão gênica diferencial em tecido muscular de bovinos Nelore divergentes para qualidade de carne**. 99 f. Tese (Doutorado em Genética e Melhoramento Animal) – Unesp, Jaboticabal.

Foyer CH, Halliwell B (1976) The presence of glutathione and glutathione reductase in chloroplasts: a proposed role in ascorbic acid metabolism. **Planta** 133:21–25.

Freitas Ld (2011) **Influência de fragmentos florestais nativos sobre os parâmetros químicos, físicos e microbiológicos de solos cultivados com cana-de-açúcar**. 115 f. Dissertação (Mestrado em Ciências Biológicas – Biologia Vegetal) – Unesp, Rio Claro.

Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LSP, Yamaguchi-Shinozaki K, Shinozaki K (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. **The Plant Journal** 39:863–876.

Furlan LR, Ferraz ALJ, Bortolossi JC (2007) A genômica funcional no âmbito da produção animal: estado da arte e perspectivas. **Revista Brasileira de Zootecnia** 331–341.

Gachon C, Mingam A, Charrier B (2004) Real-time PCR: what relevance to plant studies? **Journal of Experimental Botany** 55:1445–1454.

Gall H L, Philippe F, Domon JM, Gillet F, Pelloux J, Rayon C (2015) Cell wall metabolism in response to abiotic stress. **Plants** 4:112–166.

Gao R, Duan K, Guo G, Du Z, Chen Z, Li L, He T, Lu R, Huang J (2013) Comparative transcriptional profiling of two contrasting barley genotypes under salinity stress during the seedling stage. **International Journal of Genomics** 2013:19.

Garsmeur O, Droc G, Antonise R, Grimwood J, Potier B, Aitken K, Jenkins J, Martin G, Charron C, Hervouet C et al. (2018) A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications** 9:2638.

Gentile A, Ferreira TH, Mattos RS, Dias LI, Hoshino AA, Carneiro MS, Souza GM, Calsa T, Nogueira RM, Endres L et al. (2013) Effects of drought on the microtranscriptome of field-grown sugarcane plants. **Planta** 237:783–798.

Giachetto PF, Higa RH (2014) Bioinformática aplicada à agricultura. In: Embrapa (Ed.) **Embrapa Informática Agropecuária**. Brasília: EMBRAPA, p. 67–85.

Gonçalves ER (2008) **Fotossíntese, osmorregulação e crescimento inicial de quatro variedades de cana-de-açúcar submetidas à deficiência hídrica**. 66 f. Dissertação (Mestrado em Agronomia – Produção Vegetal e Proteção de Plantas) – Universidade Federal de Alagoas, Rio Largo.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature biotechnology** 29:644.

Graça JPd (2009) **Avaliação de parâmetros fisiológicos em cultivares de cana-de-açúcar submetidas ao déficit hídrico**. 52 f. Dissertação (Mestrado em Agronomia – Genética e Melhoramento de Plantas) – Unesp, Jaboticabal.

Grivet L, Arruda P (2002) Sugarcane genomics: depicting the complex genome of an important tropical crop. **Current opinion in Plant Biology** 5:122–127.

Gupta V, Raghuvanshi S, Gupta A, Saini N, Gaur A, Khan M, Gupta R, Singh J, Duttamajumder S, Srivastava S et al. (2010) The water-deficit stress-and red-rot-related genes in sugarcane. **Functional & Integrative Genomics** 10:207–214.

Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J (2012) How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? **BMC Genomics** 13:734.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis, **Nature Protocols** 8:1494.

Hall A, Vilella F, Trapani N, Chimenti C (1982) The effects of water stress and genotype on the dynamics of pollen-shedding and silking in maize. **Field Crops Research** 5:349–363.

Hasanuzzaman M, Nahar K, Anee TI, Fujita M (2017) Glutathione in plants: biosynthesis and physiological role in environmental stress tolerance. **Physiology and Molecular Biology of Plants** 23:249–268.

Hassan NM, El-Bastawisy ZM, El-Sayed AK, Ebeed HT, Alla MMN (2015) Roles of dehydrin genes in wheat tolerance to drought stress. **Journal of Advanced Research** 6:179–188.

- Heberle H, Meirelles GV, Silva FR da, Telles GP, Minghim R (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics** 16:169.
- Hong SW, Jon JH, Kwak JM, Nam HG (1997) Identification of a receptor-like protein kinase gene rapidly induced by abscisic acid, dehydration, high salt, and cold treatments in *Arabidopsis thaliana*. **Plant physiology** 113:1203–1212.
- Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Mering C von, Bork P (2017) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. **Molecular Biology and Evolution** 34:2115–2122.
- Inman-Bamber N, Smith D (2005) Water relations in sugarcane and response to water deficits. **Field Crops Research** 92:185–202.
- Invitrogen. **Manuals and Protocols**. Vienna, Austria. Disponível em: <<http://www.lifetechnologies.com/global/en/website-overview/abwelcome.html>>.
- Jain M (2011) Next-generation sequencing technologies for gene expression profiling in plants. **Briefings in functional genomics** 11:63–70.
- Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, Lee Y (2010) PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. **Proceedings of the National Academy of sciences** 107:2355–2360.
- Klein M, Geisler M, Suh SJ, Kolukisaoglu HÜ, Azevedo L, Plaza S, Curtis MD, Richter A, Weder B, Schulz B et al. (2004) Disruption of AtMRP4, a guard cell plasma membrane ABCC-type ABC transporter, leads to deregulation of stomatal opening and increased drought susceptibility. **The Plant Journal** 39:219–236.
- Kocsy G, Szalai G, Galiba G (2002) Induction of glutathione synthesis and glutathione reductase activity by abiotic stresses in maize and wheat. **The Scientific World Journal** 2:1699–1705.
- Koffler BE, Luschin-Ebengreuth N, Stabentheiner E, Müller M, Zechmann B (2014) Compartment specific response of antioxidants to drought stress in *Arabidopsis*. **Plant Science** 227:133–144.
- Kogenaru S, Yan Q, Guo Y, Wang N (2012) RNA-Seq and microarray complement each other in transcriptome profiling. **BMC genomics** 13:629.
- Labudda M, Azam FMS (2014) Glutathione-dependent responses of plants to drought: a review. **Acta Societatis Botanicorum Poloniae** 83:3–12 .
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. **Nature Methods** 9:357.
- Lestari P, Van K, Kim MY, Hwang CH, LEE BW, LEE SH (2006) Differentially expressed genes related to symbiotic association in a supernodulating soybean mutant and its wild-type. **Molecular Plant Pathology** 7:235–247.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC bioinformatics** 12:323.

Li C, Nong Q, Solanki MK, Liang Q, Xie J, Liu X, Li Y, Wang W, Yang L, Li Y (2016) Differential expression profiles and pathways of genes in sugarcane leaf at elongation stage in response to drought stress. **Scientific Reports** 6:25698.

Lia-Jia Q, Yu-Xian Z (2006) Transcription factor families in Arabidopsis: major progress and outstanding issues for future research. **Current opinion in Plant Biology** 9:544–549.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta CT$ method. **Methods** 25:402–408.

Lopes MS, Araus JL, Heerden PD V, Foyer CH (2011) Enhancing drought tolerance in C4 crops. **Journal of Experimental Botany** 62:3135–3153.

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. **Genome Biology** 15:550.

Macková J, Vašková M, Macek P, Hronková M, Schreiber L, Šantrček J (2013) Plant response to drought stress simulated by ABA application: changes in chemical composition of cuticular waxes. **Environmental and Experimental Botany** 86:70–75.

Mahajan S, Tuteja N (2005) Cold, salinity and drought stresses: an overview. **Archives of Biochemistry and Biophysics** 444:139–158.

Marchiori PER (2014) **Fisiologia de cana-de-açúcar sob déficit hídrico: plasticidade fenotípica, transporte de água, metabolismo antioxidante e fotossíntese**. 102 f. Tese (Doutorado em Agricultura Tropical e Subtropical) – IAC, Campinas.

Marguerat S, Bähler J (2010) RNA-Seq: from technology to biology. **Cellular and molecular life sciences**, Springer, 67:569–579.

Marin FR. **Variedades (de cana-de-açúcar)**. *Ageitec*: (http://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG01_42_1110200717570.html).

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. **Genome Research** 18:1509–1517.

Marques MO, Mutton MA, Azania AAPM, Jr LC T, Nogueira GA, Vale D (2006) **Tópicos em tecnologia sucroalcooleira**. Jaboticabal: Gráfica Multipress Ltda.

Martins AM (2013) **Sequenciamento de DNA, montagem de novo do genoma e desenvolvimento de marcadores microssatélites, indels e SNPs para uso em análise genética de *Brachiaria ruziziensis***. 198 f. Tese (Doutorado em Biologia Molecular) – Universidade de Brasília, Brasília.

Menossi M, Silva-Filho MdC, Vincentz M, Van-Sluys MA, Souza GM (2008) Sugarcane functional genomics: gene discovery for agronomic trait development. **International Journal of Plant Genomics** 2008:11.

Metzker ML (2010) Sequencing technologies—the next generation. **Nature Reviews Genetics** 11:31.

Mitra J (2001) Genetics and genetic improvement of drought resistance in crop plants. **Current Science** 80:758–763.

Mittler R, Merquiol E, Hallak-Herr E, Rachmilevitch S, Kaplan A, Cohen M (2001) Living under a ‘dormant’ canopy: a molecular acclimation mechanism of the desert plant *Retama raetam*. **The Plant Journal** 25:407–416.

Molina C, Rotter B, Horres R, Udupa SM, Besser B, Bellarmino L, Baum M, Matsu-mura H, Terauchi R, Kahl G et al. (2008) SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. **BMC Genomics** 9:553.

Mozambani AE, Pinto AdS, Segato S, Mattiuz C (2006) História e morfologia da cana-de-açú car. In: Segato SV, Pinto AS, Jendiroba E, Nóbrega JCM (Eds.) **Atualização em produção de cana-de-açúcar**. Piracicaba, ESALQ/USP, p. 11-18.

Nelson DL, Cox MM (2004) Lehninger Principles of Biochemistry. New York: W. H. Freeman, 1120p.

Nepomuceno A, Neumaier N, Farias J, Oya T (2001) Tolerância à seca em plantas: mecanismos fisiológicos e moleculares. **Biotecnologia Ciência & Desenvolvimento** 23:12–18.

Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K (2005) Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis, **The Plant Cell**, Am Soc Plant Biol, 17:1105–1119.

Oshlack A, Robinson MD, Young MD (2010) From RNA-Seq reads to differential expression results. **Genome Biology** 11:1–10.

Palhares AC (2010) **Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar (*Saccharum spp.*)**. 98 f. Dissertação (Mestre em Ciências – Genética e Melhoramento de Plantas) – ESALQ, Piracicaba.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. **Bioinformatics** 30:3123-3124.

Paula Md, Pereira FAR, Arias ERA, Scheeren BR, Souza CCd, Mata DSd (2010) Fixação de carbono e a emissão dos gases de efeito estufa na exploração da cana-de-açúcar. **Ciênc. Agrotec. (Impr.)** 34:633–640.

Pincelli RP, Silva MdA (2012) Alterações morfológicas foliares em cultivares de cana-de-açúcar em resposta à deficiência hídrica. **Bioscience Journal** 546–556.

Pinheiro C, Chaves M (2010) Photosynthesis and drought: can we make metabolic connections from available data? **Journal of Experimental Botany** 62:869–882.

Pinto A, Melo-Barbosa H, Miyoshi A, Silva A, Azevedo V (2011) Application of RNA-Seq to reveal the transcript profile in bacteria. **Genetic Molecular Research** 10:1707–18.

R Core Team (2011) **R: A Language and Environment for Statistical Computing**. Vienna, Austria. Disponível em: (<http://www.R-project.org/>).

Ricardi MM, González RM, et al. (2014) Genome-wide data (ChIP-seq) enabled identification of cell wall-related and aquaporin genes as targets of tomato ASR1, a drought stress-responsive transcription factor. **BMC Plant Biology** 14:29.

Rodrigues CM (2011) **Expressão diferencial de genes em laranja doce (*Citrus sinensis* L. Osb) em tangerina (*Citrus reticulata* blanco) em resposta à infecção por *Xylella fastidiosa***. 140 f. Tese (Doutorado em Genética – Genética Vegetal) – Unesp, Botucatu.

Rodrigues F, Graça J D, Laia M D, Nhani-Jr A, Galbiati J, Ferro M, Ferro J, Zingaretti S (2011) Sugarcane genes differentially expressed during water deficit. **Biologia Plantarum** 55:43–53.

Rodrigues FA, Laia ML de, Zingaretti SM (2009) Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants. **Plant Science** 176:286–302.

Rodrigues JD (1995) **Fisiologia da cana-de-açúcar**. 101 f. Unesp, Botucatu.

Rodríguez M, Canales E, Borrás-Hidalgo O (2005) Molecular aspects of abiotic stress in plants. **Biotecnologia Aplicada** 22:1–10.

Rodríguez M, Canales E, Borroto CJ, Carmona E, Lopez J, Pujol M, Borrás-Hidalgo O (2006) Identification of genes induced upon water-deficit stress in a drought-tolerant rice cultivar. **Journal of Plant Physiology** 163:577–584.

SANTANA MO (2011) **Expressão gênica da resposta à seca em cana-de-açúcar (*Saccharum* spp.)** 92 f. Dissertação (Mestrado em Genética) — Universidade Federal de Pernambuco, Recife.

Santos F, Borém A, Caldas C (Eds.) (2015) Sugarcane: agricultural production, bioenergy and ethanol. EUA: Academic Press, 467p.

Scarpore FV (2011) **Simulação do crescimento da cana-de-açúcar pelo modelo agro-hidrológico SWAP/WOFOST**. 164 f. Tese (Doutorado em Ciências – Física do Ambiente Agrícola) – Esalq, Piracicaba.

Shendure J (2008) The beginning of the end for microarrays? **Nature Methods** 5:585.

Shepherd T, Griffiths D W (2006) The effects of stress on plant cuticular waxes. **New Phytologist** 171:469–499.

Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. **Journal of experimental botany** 58: 221–227.

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. **Current opinion in plant biology** 6:410–417.

Silva ALCde, Costa WAJMda (2004) Varietal variation in growth, physiology and yield of sugarcane under two contrasting water regimes. **Tropical Agricultura Research** 16:1-12.

Silva PBda (2010) Aspectos fisiológicos de seis genótipos de cana-de-açúcar submetidos a estresse hídrico. 96 f. Dissertação (Mestrado em Agronomia – Produção Vegetal) – Universidade Federal de Alagoas, Rio Largo.

Smith-Unna R, Bournnell C, Patro R, Hibberd JM, Kelly S (2016) TransRate: reference-free quality assessment of de novo transcriptome assemblies. **Genome research** 26:1134-1144 .

Soneson C, Love MI, Robinson MD (2015) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences, **F1000Research** 4:1521.

Souza CBde (2012) “**Biotechnologia de cana-de-açúcar (*Saccharum* spp.) para tolerância a estresse hídrico.**” 156 f. Tese (Doutorado em Genética e Biologia Molecular) – Unicamp, Campinas.

Spannagl M, Nussbaumer T, Bader KC, Martis MM, Seidel M, Kugler KG, Gundlach H, Mayer KF (2015) PGSB PlantsDB: updates to the database framework for comparative plant genome research. **Nucleic acids research** 44:D1141– D1147.

Stolf R, **Identificação e análise da expressão de genes relacionados com tolerância à seca em soja através de microarranjos de DNA e PCR em tempo real. 2007.** 125 f. Tese – (Doutorado em Agronomia – Genética e Melhoramento de Plantas) – Unesp, Jaboticabal.

Taiz L, Zeiger E (Eds.) (2004) Fisiologia vegetal. Porto Alegre: Artmed, 719 p.

Telles BR (2016) **Características fisiológicas e moleculares de uma cultivar de cana-de-açúcar tolerante à seca e submetida ao déficit hídrico prolongado.** Dissertação (Mestrado em Agronomia – Genética e Melhoramento de Plantas) – Unesp, Jaboticabal.

Telles BR, Carvalho FM de S, et al. (2018) Prolonged water deficit reveals new profile of sugarcane gene expression and metabolic pathway related to tolerance. **Sugar Tech** 1:1–11.

Tran LSP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. **The Plant Cell** 16:2481–2498.

Tzvelev N (1989) The system of grasses (Poaceae) and their evolution. **The Botanical Review** 55:141–203.

Unica - União da Indústria de Cana-de-açúcar. **Setor sucroenergético – Açúcar**. Disponível em: <https://www.unica.com.br/setor-sucroenergetico/acucar/> Acesso em: 30 jun. 2019.

Unica- União da Indústria de Cana-de-açúcar. **UnicaData – Produção – Histórico de Produção e Moagem - Por Safra**. Disponível em: <http://www.unicadata.com.br/historico-de-producao-e-moagem.php?idMn=32&tipoHistorico=4&acao=visualizar&idTabela=2334&safra=2018%2F2019&estado=RS%2CSC%2CPR%2CSP%2CRJ%2CMG%2CES%2CMS%2CMT%2CGO%2CDF%2CBA%2CSE%2CAL%2CPE%2CPB%2CRN%2CCE%2CPI%2CMA%2CTO%2CPA%2CAP%2CRO%2CAM%2CAC%2CRR>. Acesso em: 30 jun. 2019.

Vandesompele J, Preter K D, Pattyn F, Poppe B, Roy N V, Paepe A D, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome biology** 3:research0034–1.

Vantini JdaS (2013) **Expressão gênica diferencial por cDNA-AFLP em raízes de cana-de-açúcar submetida ao déficit hídrico** (2013) 118 f. Tese (Doutorado em Agronomia – Genética e Melhoramento de Plantas) – Unesp, Jaboticabal.

Vantini JS, Dedemo G, Gimenez DFJ, Fonseca L, Tezza R, Mutton M, Ferro JA, Ferro MI (2015) Differential gene expression in drought-tolerant sugarcane roots. **Genetic Molecular Research** 14:7196–7207.

Vargas L, Brigida ABS, et al. (2014) Drought tolerance conferred to sugarcane by association with *Gluconacetobacter diazotrophicus*: a transcriptomic view of hormone pathways. **PLoS One** 9:e114744.

Vettore AL, Silva FRda, et al. (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome research** 13:2725–2735.

Vettore AL, Silva FRd, Kemper EL, Arruda P (2001) The libraries that made SUCEST, **Genetics and Molecular Biology** 24:1–7.

Vliet AH V (2010) Next generation sequencing of microbial transcriptomes: challenges and opportunities. **FEMS Microbiology letters** 302:1–7.

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta** 218:1–14.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics, **Nature Reviews Genetics** 10:57.

Watson JD, Crick FH et al. (1953) Molecular structure of nucleic acids. **Nature** 171:737–738.

Wilhelm BT, Marguerat S, Goodhead I, Bähler J (2010) Defining transcribed regions using RNA-Seq. **Nature protocols** 5:255.

Wilkinson S, Davies WJ (2002) ABA-based chemical signalling: the co-ordination of responses to stress in plants. **Plant, cell & Environment** 25:195–210.

Xiang CC, Chen Y (2000) cDNA microarray technology and its applications. **Biotechnology Advances** 18:35–46.

Xiong J (Ed.) (2006) Essential bioinformatics. San Francisco: Cambridge University Press, 339 p.

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic-and cold-stress-responsive promoters. **Trends in Plant Science** 10:88–94.

Zheng M, Meng Y, Yang C, Zhou Z, Wang Y, Chen B (2014) Protein expression changes during cotton fiber elongation in response to drought stress and recovery. **Proteomics** 14:1776–1795.

Zhu JK (2001) Cell signaling under salt, water and cold stresses. **Current opinion in Plant Biology** 4:401–406.