**Universidade Federal de Pernambuco**
**Centro de Biociências**
**Programa de Pós-Graduação em Genética**

**Manassés Daniel da Silva**

# Identificação e Validação de Transcritos SuperSAGE de Cana-de-açúcar Diferencialmente Expressos sob Déficit Hídrico

**Recife**
**2017**

**Manassés Daniel da Silva**

# Identificação e Validação de Transcritos SuperSAGE de Cana-de-açúcar Diferencialmente Expressos sob Déficit Hídrico

Tese apresentada ao Programa de Pós-Graduação em Genética, área de concentração Biologia Molecular, da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutor em Genética.

**Orientador**: Prof. Dr. Éderson Akio Kido

**Recife**

**2017**

Catalogação na Fonte:
Bibliotecária Marylu Souza, CRB-4/1564

Silva, Manassés Daniel da
Identificação e validação de transcritos SuperSAGE de cana-de-açúcar diferencialmente expressos sob déficit hídrico / Manassés Daniel da Silva – Recife: O Autor, 2017.

176 f.: il.

Orientador: Éderson Akio Kido
Tese (doutorado) – Universidade Federal de Pernambuco. Centro de Biociências. Doutorado em Genética, 2017.
Inclui referências e anexos

1. Expressão gênica 2. Cana-de-açúcar – Melhoramento genético 3. Plantas – Efeito da seca I. Kido, Éderson Akio II. Título.

572.865 CDD (22.ed.) UFPE/CB-2017-339

**Manassés Daniel da Silva**

# Identificação e Validação de Transcritos de Cana-de-açúcar Diferencialmente Expressos sob Déficit Hídrico

Tese apresentada ao Programa de Pós-Graduação em Genética, área de concentração Biologia Molecular, da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutor em Genética.

**Aprovado em: 23/05/2017**

**Banca Examinadora:**

_____

**Dr. Éderson Akio Kido**
**Universidade Federal de Pernambuco**

_____

**Dra. Ana Maria Benko Iseppon**
**Universidade Federal de Pernambuco**

_____

**Dra. Valesca Pandolfi**
**Universidade Federal de Pernambuco**

_____

**Dr. Paulo Paes de Andrade**
**Universidade Federal de Pernambuco**

_____

**Dr. Edson Luiz Folador**
**Universidade Federal da Paraíba**

# AGRADECIMENTOS

Agradeço primeiramente ao meu pai, **Tito Daniel da Silva**, por ter dado todo apoio psicológico, financeiro e toda motivação necessários para conclusão desta etapa acadêmica.

Agradeço à minha mãe, **Zuleide Daniel de Oliveira**, pela paciência em me lembrar de todos os compromissos e horários, ajudando-me a ser sempre pontual.

Agradeço aos meus irmãos, **Ceel Daniel da Silva** e **Lucas Daniel da Silva** pelos incentivos, brincadeiras e estresses, que só os irmãos podem fornecer.

Agradeço ao **Dr. Prof. Éderson Akio Kido**, por me iniciar na ciência, elevar meu pensamento crítico e ter me orientado até o momento da entrega desta Tese.

Agradeço ao **Dr. José Ribamar da Costa Ferreira Neto**, pela orientação em temas científicos e no desenho de *primers*.

Agradeço a todos os membros do Laboratório de Genética Molecular, em especial para Marislane, Amanda Thamires, George, Amanda Souza e Valesca, companheiros de bancada e da vida.

Agradeço à minha noiva **Rahisa Helena da Silva**, pela paciência, apoio, amor e carinho dados a mim durante essa empreitada.

Agradeço ao suporte financeiro oferecido pelas agências de fomento: **CNPq**, **FINEP** e especialmente à **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior** - **CAPES**.

*"You have to do your own growing, no matter how tall your grandfather was."*

Abraham Lincoln, 1858.

# Resumo

A cana-de-açúcar é uma cultura de importância industrial, cultivada em países tropicais e subtropicais. Dentre os fatores abióticos, a seca é o que mais negativamente impacta a produtividade vegetal. Portanto, o desenvolvimento de variedades tolerantes à seca está relacionado com a sustentabilidade e a viabilidade econômica da cultura. O presente trabalho pretende contribuir para um melhor entendimento dos processos envolvidos na resposta da cana-de-açúcar à seca, através da identificação e validação de alvos moleculares diferencialmente expressos após tratamento de supressão de rega (24 h), com vistas ao desenvolvimento futuro de marcadores moleculares funcionais. Para tanto, análises *in silico* de bibliotecas HT-SuperSAGE (de raízes de variedades tolerantes e sensíveis sob estresse e sem estresse) possibilitaram identificar alvos moleculares relacionados a diferentes estratégias de resposta e a genes de fatores de transcrição (FTs). Destes, 26 alvos foram validados via RT-qPCR (dentre os quais, duas aquaporinas e cinco fatores de transcrição), com ao menos dois dos cinco genes de referência padronizados. Adicionalmente, a análise *in silico* identificou 94 genes FTs enriquecidos que controlam a expressão dos transcritos HT-SuperSAGE mais induzidos nos dois conjuntos de cultivares contrastantes em resposta ao estresse. Também foram identificados os elementos regulatórios atuantes em *Cis* (*cis-Acting regulatory elements* - CARE) mais prevalentes em regiões promotoras de nove dos 146 FTs induzidos. Dessa forma, os resultados aqui apresentados fornecem recursos valiosos de genômica funcional para aplicação no melhoramento da cana-de-açúcar frente ao déficit hídrico, cujo desdobramento biotecnológico auxiliará na sustentabilidade da manutenção e expansão dessa cultura.



**Palavras-chave:** Transcriptômica. RT-qPCR. Expressão gênica. Promotores.

## Abstract

Sugarcane is a significant industrial crop and it is grown in tropical and subtropical countries. Among the abiotic factors, drought is the one that most negatively impacts plant productivity. Therefore, the development of drought-tolerant varieties is related to the sustainability and economic viability of the crop. This paper aims to contribute to a better understanding of the processes involved in the response of sugarcane to drought, through the identification and validation of differentially expressed molecular targets after irrigation suppression (24 h) treatment, with a view to the future development of functional molecular markers. To this end, *in silico* analyzes of HT-SuperSAGE libraries (from roots of tolerant and sensitive cultivars under stress and without stress) made it possible to identify molecular targets related to different response strategies and genes of transcription factors (TFs). Of these, 26 targets were validated via RT-qPCR, (including two aquaporins and five transcription factors), with at least two of the five standardized reference genes. In addition, *in silico* analysis identified 94 enriched TFs genes that control the expression of the most induced HT-SuperSAGE transcripts in the two sets of contrasting cultivars in response to stress. The most prevalent cis-Acting regulatory elements (CARE) of nine of the 146 induced TFs were also identified. Thus, the results presented here provides valuable functional genomics resources for application in sugarcane breeding against water deficit, whose biotechnological development will help in the sustainability of the maintenance and expansion of this crop.



**Keywords:** Transcriptomics. RT-qPCR. Gene expression. Promoters.

# Lista de Ilustrações

## Revisão de Literatura

**Figura 1:** Fluxograma da técnica HT-SuperSAGE. Traduzido de Matsumura *et al.* (2008).....38

## Artigo I

**Figure 1:** Average gene expression stability values ($M$) of six sugarcane potential reference genes ($\alpha$TUB: alpha-tubulin; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; H1: histone H1; SAMDC: S-adenosylmethionine decarboxylase; UBQ: ubiquitin; 25S rRNA: 25S ribosomal RNA) based on the GeNorm analysis. .....49

**Figure 2:** Pairwise variation ($V$) analysis for six potential reference genes of sugarcane ($\alpha$-tubulin, glyceraldehyde 3 phosphate dehydrogenase, histone H1, S-adenosylmethionine decarboxylase, ubiquitin, and 25S rRNA) based on the GeNorm analysis [17]. The addition to the two most stable genes ($\alpha$TUB and GAPDH) of a random third gene ($V$ 2/3 = 0.15), a fourth gene ($V$ 3/4 = 0.14), or even more ($V$ 4/5 and $V$ 5/64) still exhibited desirable values (basically below than 0.15).....49

**Figure 3:** Relative expression of glutamine-dependent asparagine synthetase (AS), pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP$\alpha$1), plasma membrane intrinsic protein 1-1 (PIP1-1), and 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase) by the REST software v. 2.0.13 (after the ΔΔCq method) in cDNAs of sugarcane roots under abiotic stress (24 h of continuous dehydration or salt stress; 100 mM NaCl), normalized by the reference genes H1, $\alpha$TUB, and GAPDH. (a) Tolerant bulk (CTC6, CTC15, SP83-

2847, and SP83-5073 accessions) compared to its negative control. (b) Susceptible bulk (CTC9, CTC13, SP90-1638, and SP90-3414) compared to its negative control. (c) Salt-tolerant accession (RB931011) after 30 min of salt stress compared to its negative control. (d) RB931011 after 90 min of salt stress compared to its negative control. Relative expression with the median value (horizontal dotted line at the colored box) and range comprising 100% of the observations (horizontal bars), being 50% of them in the confidence interval at 95% (colored box).....50

## Artigo II

**Figure 1:** RT-qPCR of candidate genes carried out by sugarcane root cDNAs after 24 h of irrigation supression and negative control. The expression values were normalized by the reference genes glyceraldehyde 3 phosphate dehydrogenase (GAPDH) and 25S ribosomal RNA (25S rRNA) and relative expression data calculated by REST software (v.2.0.13) [37]. Tolerant cultivars: CTC6, CTC15, SP83-2847 and SP83-5073. Sensitive cultivars: CTC9, CTC13, SP90-1638 and SP90-3414. ACCD: 1-aminocyclopropane-1-carboxylate deaminase; AP: aspartic proteinase; AAT: Na+ dependent neutral amino acid transporter; TRX: thioredoxin-like 1-2, cloroplast; PFK-2: 6-phosphofructo-2-kinase; ß-EXP: beta-expansin 8 precursor; ACC: acetyl-CoA carboxylase; E2, ubiquitin conjugating enzyme; 20SPß4, 20S proteasome beta 4 subunit; TPI, triosephosphate isomerase; FBA, fructose-bisphosphate aldolase; TKT, transketolase; C13, legumain cysteine-type endopeptidase; LDP, dihydrolipoyl dehydrogenase; MDH, malate dehydrogenase.....92

## Artigo III

**Figure 1:** Percentage of HT-SuperSAGE unitags from sugarcane plants (24 h of

continuous dehydration or regular daily irrigation) identified as potential aquaporins after keyword searches in the EST annotation (“aquaporin,” “tonoplast intrinsic protein,” “PIP,” “TIP,” “NIP,” “SIP”) or in the GO terms (“water transport”). Total of unitags:1,579.....96

**Figure 2:** Neighbor Joining dendrogram (MEGA v.5.2.1 software [ 27]) and integrated heat map (bootstrap values of 1,000 replications), showing the phenetic grouping of 24 potential aquaporin amino acid sequences identified by HT-SuperSAGE unitags from sugarcane accessions after 24 h of continuous dehydration (with their respective EST and unitag identifiers), and aquaporins sequencesof *Arabidopsis thaliana*, human and *Yersinia pseudotuberculosis* (all these labeled by an asterisk). Nomenclature: Isoforms are preceded by the abbreviated species name (At - *Arabidopsis thaliana*; Hs - *Homo sapiens*; Hv - *Hordeum vulgare*; Os - *Oryza sativa*; Pv - *Panicum virgatum*; Sb - *Sorghum bicolor*; So - *Saccharum officinarum*; Yp - *Yersinia pseudotuberculosis*; Zm - *Zea mays*).....98

**Figure 3:** Representation of aquaporin subfamilies expressed in sugarcane after 24 h of continuous dehydration ((a), (b), and (c)) and during normal irrigation (d), involving bulks of genotypes tolerant and sensitive to stress. HT-SuperSAGE libraries: SD24T (bulk of tolerant genotypes under stress; SD24S (bulk of sensitive genotypes under stress); SDTC (control bulk of tolerant genotypes); SDSC (control bulk of sensitive genotypes). .....100

**Figure 4:** Relative quantification of SsPIP1 aquaporin-1 (unitag SD264077) and SoPIP1-3/PIP1-4 (unitag SD231548) with sugarcane cDNAs [tolerant genotypes: (a) CTC6, (b) CTC15, (c) SP83-2847, (d) SP83-5073; sensitive genotypes [(e) CTC9, (f) CTC13, (g) SP90-1638, (h) SP90-3414] under stress (24 h of

continuous dehydration). The expression ratios of the same genotypes under control condition were normalized by reference genes GAPDH and 25S rRNA. #Relative expression (REST© v. 2.0.13 software) with the mean value (horizontally dotted line) and range of the observations (100%, horizontal bars); in red: confidence interval at 95%.....103

## Artigo IV

**Figura 1:** Distribuição de FTs por família no transcriptoma de *Saccharum* spp. (A) e respectivas distribuições de FTs por família em acessos tolerantes (B) e acessos sensíveis (D). Os números das 20 maiores famílias FT foram contados separadamente, enquanto que as demais famílias e os membros Órfãos foram integrados em outros.....120

**Figura 2:** Fatores de transcrição diferencialmente expressos sobre supressão de rega por 24 h.....121

**Figura 3:** Diagrama de Venn evidenciando as respostas dos diferentes bulks de cultivares frente ao estresse de supressão de rega por 24 h, a partir de *unitags* HT-SuperSAGE diferencialmente expressas (UR, *upregulated*; DR, *downregulated* em teste Audic-Claverie, $p < 0,05$).....122

**Figura 4:** Clusterização hierárquica[1] em relação a sequências mais diferencialmente expressas ($p < 0,0001$) em 24TxC (contraste dos acessos tolerantes, estresse de 24 horas e controles controlados) e suas respectivas respostas em 24SxC (contraste dos sensíveis sob estresse de 24 h de supressão de rega e respectivo controle).....124

**Figura 5:** Os dez mais *Cis-Acting regulatory elements* de cana-de-açúcar mais prevalentes entre os 186 motivos identificados na região promotora (2,000 pb

*upstream*) de nove fatores de transcrição de cana-de-açúcar induzidos (p<0.05) em cultivares tolerantes sob 24 h de supressão de rega.....125

**Figura 6:** Famílias identificadas na análise de *TF Enrichment* em transcritos HT-SuperSAGE exclusivamente induzidos (*p-value* < 0,0001) para cada *bulk* de cultivares de cana-de-açúcar contrastantes e suas espectivas quantidades de regulações.....129

**Figura 7:** Diagrama de Venn destacando o número de FTs com interações com alvos supra-representados, identificados a partir dos transcritos induzidos de cada *bulk* de cultivares após a aplicação do estresse de supressão de rega de 24 h.....130

**Figura 8:** RT-qPCR dos genes candidatos performada em cDNAs raízes de cana-de-açúcar sobre déficit hídrico (24 h de desidratação contínua). Os valores de expressão foram normalizados por αTubulina GAPDH e Histona H1. Dados relativos à expressão calculados pelo software REST (v.2.0.13).....132

# Lista de Tabelas

## Artigo I

**Table 1:** Potential reference genes and target genes with the individual accession numbers, annotations, and primer sequences.....46

**Table 2:** Parameters derived from RT-qPCR data analysis, including PCR amplification efficiency ($E$) established by calibration curves for each quantified target.....48

**Table 3:** Expression stability values for sugarcane candidate calculated by NormFinder and BestKeeper software. .....49

**Table 4:** Relative expression rates of target genes (PFP$\alpha$1, AS, PIP1-1, and ACC oxidase) based on RT-qPCR with roots, cDNAs of sugarcane accessions under abiotic stress, and respective unitag regulation by SuperSAGE analysis covering drought A stress (24 h of continuous dehydration) or salt B stress (100 mM NaCl).....51

**Table 5:** Potential reference gene combinations (and number of genes involved in each comparison) used in gene expression normalization of glutamine-dependent asparagine synthetase (AS), with cDNAs of sugarcane accessions (root under drought stress, 24 h of continuous dehydration).....52

## Artigo II

**Table 1:** The 50 most induced HT-SuperSAGE unitags expressed by the tolerant bulk (four sugarcane cultivars) and at the same time repressed (or n.s. at p-value < 0.05 in Audic-Claverie test) by the sensitive bulk after 24 h of irrigation suppression.

Also, their respective EST anchoring the unitag, the EST annotation and the fold change values (FC) considering the tolerant and the sensitive contrast (FC tol , FC sens , respectively).....89

**Table 2:** Proposed candidate genes and their results based on in silico analysis of HT-SuperSAGE unitags expressed by drought-tolerant or -sensitive bulks (four cultivars each one) and their relative expression by RT-qPCR, using sugarcane cDNAs from roots of each cultivar after 24 h of irrigation suppression.....90

## Artigo III

**Table 1:** EST databases used for unitag annotations from sugarcane HT-SuperSAGE libraries (tolerant/sensitive genotypes after 24 h of continuous dehydration or regular daily irrigation) and BLASTn nonredundant results.....97

**Table 2:** Putative sugarcane aquaporin isoforms (42) based on unitags of root dehydration (24 h) observed in the two main contrasts of HT-SuperSAGE libraries.....99

**Table 3:** The 30 HT-SuperSAGE unitags most expressed and annotated as aquaporins from sugarcane libraries with contrasting genotypes under stress (24 h of continuous dehydration) or normal daily irrigation.....99

**Table 4:** Aquaporin unitags with distinct expression rates in root HT-SuperSAGE libraries from contrasts of tolerant T (SD24T versus SDTC) and sensitive S (SD24S versus SDSC) sugarcane genotypes and after continuous dehydration (24 h).....102

**Table 5:** Primers sequences for RT-qPCR of SsPIP1-1, SoPIP1-3/PIP1-4 (designed

from sugarcane ESTs), 25S rRNA, and GAPDH (as reference genes).....104

**Table 6:** Relative expression rates of aquaporins PIP1-1 (SD264077 unitag) and PIP1-3/PIP1-4 (SD231548 unitag) in bulks of tolerant or sensitive genotypes, respectively, and RT-qPCR data (both in bulks and each genotype).....104

## Artigo IV

**Tabela 1:** Famílias de fatores de transcrição identificadas a partir de alinhamentos de unitags expressas em genótipos tolerantes de cana-de-açúcar (Saccharum spp.) sobre déficit hídrico (24h de supressão de rega) com sequências nucleotídicas de fatores de transcrição* e suas respectivas regulações ($p$ > 0,05).....118

**Tabela 2:** Fatores de transcrição UR (*p-value* <0,05) que permitiram extrair 2.000 pares de bases upstream do início do transcrito, em sequências genômicas de cana-de-açúcar (BACs - *bacterial artificial chromossome*), seus *Cis-Acting regulatory elements* (CARE) mais prevalentes, suas respectivas funções e respectivas referências.....126

**Tabela 3:** Genes candidatos propostos e seus respectivos resultados baseados nas análises *in silico* de *unitags* HT-SuperSAGE expressas por expressed *bulks* tolerantes ou sensíveis (quatro cultivares em cada) e suas expressões relativas adiquiridas via RT-qPCR, utilizando-se cDNAs de raízes de cada cultivar extraídos após estresse (24 h de supressão de rega) ou sob condições de controle.....132

# Lista de Abreviaturas, Siglas e Símbolos

| Item | Definição |
|---|---|
| 20SPß4 | 20S *proteasome beta 4 subunit* (subunidade beta 5 do proteassoma 20S) |
| 25S rRNA | 25S *ribosomal* RNA (RNA ribossomal 25S) |
| AAT | *Na+-dependent neutral amino acid transporter* (Transportador de aminoácido neutro dependente de Na+) |
| ABA | Ácido abscísico |
| ACC | *Acetyl-CoA carboxylase* (Acetil CoA carboxilase) |
| ACC oxidase | *1-Aminocyclopropane-1-carboxylate oxidase* (1-Aminociclopropano-1-carboxilato oxidase) |
| ACCD | *1-aminocyclopropane-1-carboxylate deaminase* (1-aminociclopropano-1-carboxilato deaminase) |
| ALD | *Fructose-bisphosphate aldolase* (frutose bisfosfato aldolase) |
| AP | *Aspartic proteinase* (proteinase ácido aspártico) |
| AP2/EREB | *APETALA 2 /ethylene-responsive element binding factor* (APETALA 2 / fator de ligação do elemento sensível ao etileno) |
| APX | Ascorbato peroxidase |
| ARF | *Auxin response factor* (fator de resposta à auxina) |
| AS | *Glutamine-dependent asparagine synthetase* (Asparagina sintetase dependente de glutamina) |
| BAC | *Bacterial artificial chromosome* (cromossomo bacteriano artificial) |
| bHLH | *Basic helix-loop-helix* (hélice-volta-hélice básica) |
| BLAST | *Basic Local Alignment Search Tool* (Ferramenta básica de busca de alinhamento local) |
| BLASTn | *Nucleotide BLAST* (BLAST de nucleotídeos) |
| BLASTx | *Search protein database using a translated nucleotide* |

| | |
|---|---|
| | BLAST (BLAST com busca por proteínas utilizando nucleotídeos traduzidos em aminoácidos) |
| Bp | *Base pairs* (pares de bases) |
| C | Conteúdo de um conjunto haplóide de cromossomos |
| C13 | Legumain cysteine-type endopeptidase (endopeptidase tipo legumina-cisteína) |
| cAPX | Cytoplasmic ascorbate peroxidase (ascorbato peroxidase cistoplasmática) |
| CARE | *Cis-Acting regulatory elements* (elementos regulatórios atuantes em *cis*) |
| CAT | Catalase |
| cDNA | *Complementary DNA* (DNA complementar) |
| CDS | *Coding Sequence* (Sequência codificante) |
| CONAB | Companhia Nacional de Abastecimento |
| Cq | *Quantification cycle value* (ciclos de quantificação) |
| CTC | Centro de Tecnologia Canavieira |
| DE | *Differentially expressed* (diferencialmente expressos) |
| DHAR | Dehidroascorbato redutase |
| DNA | *Desoxiribonucleic Acid* (ácido desoxirribonucléico) |
| DPBB_1 | *Double-psi beta-barrel 1* (folha-beta de duplo psi) |
| DR | *Down regulated* (reprimido) |
| E | *Efficiencie of amplification* (eficiência da aplificação) |
| E2 | *Ubiquitin conjugating enzyme* (enzima da conjugação da ubiquitina) |
| EC | *Enzyme Commission Number* (Números da Comissão para Enzimas) |
| ENO | Enolase |
| EST | *Expressed Sequence Tag* (Etiquetas de sequências expressas) |
| FAO | *Food and Agriculture Organization of the United Nations* (Organização da comida e agricultura das nações unidas) |
| FBA | *Fructose-bisphosphate aldolase* (aldolase frutose- |

| | |
|---|---|
| | bisfosfato) |
| FC | *Fold Change* (Modulação da expressão) |
| GAPDH | *Glyceraldehyde-3-phosphate dehydrogenase* (Gliceraldeído-3-fosfato desidrogenase) |
| gDNA | *Genomic desoxiribonucleic acid* (ácido desoxirribonucleico genômico) |
| GO | *Gene Ontology* (Ontologia gênica) |
| GOPX | Guaicol peroxidase |
| GPI | *Glucose 6-phosphate isomerase* (glicose 6-fosfato isomerase) |
| GPX | *Glutathione peroxidase* (glutationa peroxidase) |
| GR | *Glutathione reductase* (glutationa redutase) |
| GRAS | *Gibberellic acid insensitive (GAI), repressor of GA1 (RGA), and scarecrow (SCR)* (Insensível ao ácido giberélico (GAI), repressor de GA1 (RGA) e espantalho (SCR) |
| GRASSIUS | *Grass Regulatory Information Services* (serviços de informação regulatória de gramíneas) |
| GST | *Glutathione-S-transferase* (glutationa-S- transferase) |
| H1 | *Histone H1* (histona H1) |
| HD-Zip | *Leucine zipper homeodomain* (homeodomínio zíper de leucina) |
| Hsps | *Heat shock proteins* (proteínas de choque térmico) |
| HT-SuperSAGE | *High-Throughput Super Serial Analysis of Gene Expression* (super análise serial da expressão gênica de alto rendimento) |
| IAC | Instituto Agronômico de Campinas |
| IPCC | *Intergovernmental Panel on Climate Change* (Painel intergovernamental para as mudanças climáticas) |
| iPGM | *Phosphoglycerate mutase* (fosfoglicerato mutase) |
| LDP | *Dihydrolipoyl dehydrogenase* (dihidrolipoyl desidrogenase) |
| LEA | *Late-Embryogenesis-Abundant* (abundantes no final da |

| | |
|---|---|
| | embriogênese) |
| MAS | *Marker-assisted selection* (seleção assistida por marcadores) |
| Mb | Mega bases |
| MDH | *Malate dehydrogenase* (malato desidrogenase) |
| MDHAR | monodehydroascorbate reductase (monodehidroascorbato redutase) |
| MIP | *Major intrinsic protein* (principais proteínas intrínsecas) |
| MIQE | *The Minimum Information for Publication of Quantitative Real-Time PCR Experiments* (Informação mínima para publicação da quantificação relativa de experimentos de PCR em tempo real) |
| MLIP15 | *Low temperature induced 15 kDa maize protein* (proteína do milho de 15 kDa induzida por baixa temperatura) |
| mRNA | *Messenger RNA* (RNA mensageiro) |
| NAC | *NAM [No Apical Meristem], ATAF1-2, and CUC2 [Cup-Shaped Cotyledon]* (NAM [sem meristema apical], ATAF1-2, e CUC2 [cotilédone em forma de taça]) |
| NaCl | Cloreto de sódio |
| NIP | *Nodulin-26–like intrinsic protein* (Proteínas de 26 kDa intrínsecas ao nódulo) |
| NTC | *No template controls* (controles sem fita molde) |
| ORF | *Open read frame* (quadro aberto de leitura) |
| *Os* | *Oryza sativa* |
| PCR | *Polymerase chain reaction* (reação em cadeia da polimerase) |
| PEG | *Polyethylene glycol* (polietileno glicol) |
| PFK-2 | *6-phosphofructo-2-kinase* (6-fosfofrutocinase) |
| PFP$\alpha$1 | *Pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit* (subunidade alfa da pirofosfato frutose-6-fosfato 1-fosfotransferase) |
| PGK | *Phosphosglycerate kinase* (fosfoglicerato quinase) |
| PGM | *Phosphoglucomutase* (fosfoglucomutase) |

| | |
|---|---|
| PGPB | *Plant growth-promoting bacteria* (bactérias endofíticas promotoras de crescimento) |
| PIP1-1 | *Plasma membrane intrinsic protein* 1-1 |
| PYK | *Pyruvate kinase* (piruvato kinase) |
| QR | *Quinone reductase* (quinona redutase) |
| R | *Correlation coefficient* (coeficiente de correlação) |
| REST | *Relative Expression Software Toll* (Programa-ferramenta para expressão relativa) |
| RIDESA | Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético |
| RNA | *Ribonucleic acid* (Ácido ribonucléico) |
| RNAi | *Interference RNA* (RNA de interferência) |
| ROS | *Reactive oxygen species* (Espécies reativas de oxigênio) |
| RT-qPCR | *Quantitative reverse transcription polymerase chain reaction* (reação quantitativa da transcrição reversa em cadeia da polimerase) |
| RWC | *Relative water content* (conteúdo relativo de água) |
| SAGE | *Serial Analysis of Gene Expression* (análise serial da expressão gênica) |
| SAM | *Shoot apical meristem* (meristema apical do caule) |
| SAMDC | *S-Adenosylmethionine decarboxylase* (S-Adenosilmetionina descarboxilase) |
| SD24S | Biblioteca HT-SuperSAGE do *bulk* de genótipos sensíveis ao estresse sob condições de estresse (supressão de rega de 24 h) |
| SD24T | Biblioteca HT-SuperSAGE do *bulk* de genótipos tolerantes ao estresse sob condições de estresse (supressão de rega de 24 h) |
| SDSC | Biblioteca HT-SuperSAGE do *bulk* de genótipos sensíveis ao estresse sob condições controle |
| SDTC | Biblioteca HT-SuperSAGE do *bulk* de genótipos tolerantes ao estresse sob condições controle |
| Sens | *Sensitive cultivar* (cultivar sensível) |

| | |
|---|---|
| *So* | *Saccharum officinarum* |
| SOD | *Superoxide dismutase* (superóxido dismutase) |
| *Ss* | *Saccharum* spp. |
| ß-EXP | *Beta-expansin 8 precursor* (Precusor da beta-expansina 8) |
| SuperSAGE | *Super Serial Analysis of Gene Expression* (super análise serial da expressão gênica) |
| TALE | *Three amino acid loop extension* (volta de três aminioácidos de extensão) |
| tAPX | *Thylakoid ascorbate peroxidase* |
| TCA | *Tricarboxylic acid* (ácido tricarboxílico) |
| TF | *Transcription factor* (fator de transcrição) |
| TIM | *Triosephosphate isomerase* (triose fosfato isomerase) |
| TIP | *Tonoplast intrinsic protein* (proteínas intrínsecas ao tonoplasto) |
| TKT | *Transketolase* (trascetolase) |
| Tol | *Tolerant cultivar* (cultivar tolerante) |
| TPI | *Triosephosphate isomerase* (triose fosfato isomerase) |
| TRX | *Thioredoxin* (Tioredoxina) |
| UBQ | *Ubiquitin*(ubiquitina) |
| UNICA | União da Indústria da Cana-de-Açúcar |
| *unitags* | *Unique tags* (tags únicas) |
| UR | *Up regulated* (superexpresso) |
| WRKY | motivo W R K Y |
| ZIM | *Zinc-finger protein expressed in Inflorescence Meristem (proteína de dedo de zinco expressa em meristema de Inflorescência)* |
| *Zm* | *Zea mays* |
| αTUB | *Alpha-tubulin* (alfa-tubulina) |

# Sumário

**1 INTRODUÇÃO.....23**

**2 REVISÃO DA LITERATURA.....26**

**2.1 A cultura da cana-de-açúcar.....26**

**2.2 Déficit hídrico em vegetais: Causas, consequências e respostas moleculares.....28**

**2.3 Análises globais da expressão gênica.....30**
2.3.1 Genômica estrutural e transcriptômica na cana-de-açúcar: ferramentas, disponibilidade de dados e aplicações.....31
2.3.2 A tecnologia SuperSAGE.....36

**2.4 Programas de melhoramento da cana-de-açúcar no Brasil.....39**

**3 OBJETIVOS.....42**

**3.1 Objetivo geral.....42**

**3.2 Objetivos específicos.....42**

**4 ARTIGO I** "Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-Time PCR in Drought-Stressed Sugarcane".....43

**5 ARTIGO II** "Unraveling Drought-tolerant and -sensitive Sugarcane Cultivars by HT-SuperSAGE Analysis".....56

**6 ARTIGO III** "Expression Analysis of Sugarcane Aquaporin Genes under Water Deficit".....93

**7 ARTIGO IV** "Análise da Expressão de Genes de Fatores de Transcrição de Cana-de-açúcar sob déficit hídrico ".....108

**8 DISCUSSÃO GERAL.....151**

**9 CONCLUSÕES.....156**

**REFERÊNCIAS.....158**

**ANEXO A** “Normas para submissão (Genetics and Molecular Biology - ISSN 1678-4685)”.....165

**ANEXO B** “Curriculum vitae (Lattes)”.....172

# 1 INTRODUÇÃO

A cana-de-açúcar (*Saccharum* spp. híbrido) é uma importante gramínea de propagação vegetativa pertencente ao gênero *Saccharum*, tribo Andropogoneae e família Poaceae. Seus principais produtos são açúcar e o biocombustível etanol. Segundo a *Food and Agriculture Organization of the United Nations* (FAO), a produção mundial dessa cultura totaliza aproximadamente dois bilhões de toneladas de biomassa anuais, originando um faturamento bruto de mais de 48 bilhões de dólares. A cana-de-açúcar é cultivada em mais de 100 países ocupando aproximadamente 26 milhões de hectares e o Brasil, com 684,77 milhões de toneladas estimadas pela Companhia Nacional de Abastecimento (CONAB) para safra 2016/2017, desponta como maior produtor mundial dessa cultura, sendo também o principal processador da cana-de-açúcar em açúcar e etanol.

A demanda crescente por açúcar e biocombustíveis tem favorecido a expansão do setor, exigindo maior produtividade e condições de crescimento em regiões antes consideradas adversas, em decorrência de fatores climáticos e/ou de solos. Entre os estresses abióticos, o déficit hídrico é o mais prejudicial e afeta constantemente as áreas de expansão de cultivo. Portanto, o desenvolvimento de genótipos de cana-de-açúcar tolerantes à seca está estritamente relacionado à sustentabilidade e à viabilidade econômica da cultura.

A complexidade do caráter tolerância à seca dificulta a obtenção de genótipos tolerantes ao déficit hídrico a partir de métodos convencionais, como o melhoramento genético clássico. A seca provoca um complexo conjunto de

respostas moleculares nas plantas, que se iniciam com a percepção do estresse – culminando em cascatas de transdução de sinais – e se manifestam como alterações aos níveis celular, fisiológico e morfológico. Essas alterações incluem: fechamento dos estômatos, a repressão do crescimento celular e da fotossíntese, além da ativação da respiração celular. As plantas em condições de seca também acumulam osmólitos e sintetizam proteínas especificamente envolvidas na tolerância ao estresse.

Além das dificuldades para analisar a complexidade das respostas das plantas à seca, a cana-de-açúcar é um desafio do ponto de vista genético: o complexo *Saccharum* spp. híbrido foi originado a partir de cruzamentos interespecíficos que resultaram em um dos genomas mais complexos do reino vegetal. Assim, é um grande desafio tentar melhorar a cana-de-açúcar para o déficit hídrico e isso tem incentivado pesquisas de genômica funcional, principalmente a partir de transcriptomas, que vêm obtendo sucesso em identificar novos genes, gerar dados de regulação gênica, além de possibilitar a descoberta de genes candidatos a marcadores moleculares funcionais.

Como milhares de etiquetas (*tags*) de 26 pares de bases podem ser sequenciadas em um mesmo experimento HT-SuperSAGE (*Super Serial Analysis of Gene Expression*), um compreensivo padrão do transcriptoma pode ser gerado, permitindo a detecção de milhares de transcritos diferencialmente expressos. Essa tecnologia possibilita análises qualitativas e quantitativas do transcriptoma de um ou mais genomas de eucariotos, a partir da mudança da frequência de, teoricamente, todo mRNA possuidor de cauda poli-A.

O presente trabalho, a partir da metodologia de transcriptômica HT-SuperSAGE, apresenta uma contribuição para um maior entendimento dos

processos envolvidos na resposta da cana-de-açúcar à seca através da identificação e validação de alvos moleculares de diferentes estratégias de tolerância à seca e de genes responsivos de fatores de transcrição, visando o desenvolvimento de marcadores moleculares funcionais para uso em seleção assistida ou mesmo uso para transformação genética de cultivares. Para tanto, foram selecionados e disponibilizados seis genes de (αTUB, GAPDH, H1, SAMDC, UBQ, 25S rRNA), utilizados na validação de um total de 26 alvos. Os alvos validados neste trabalho distribuem-se em um painel de 15 diferentes componentes de sete estratégias de resposta à seca (atenuação do estresse de excesso de etileno; crescimento radicular; degradação proteica; detoxificação oxidativa; biossíntese de ácidos graxos; transporte de aminoácidos, e metabolismo de carboidratos), quatro alvos validados na padronização de genes de referência (PIP1-1 [estresse salino], ACC oxidase [estresse salino], AS [déficit hídrico] e PFP$\alpha$1 [déficit hídrico]), duas aquaporinas (PIP1-1 e PIP 1-3/PIP1-4) de um conjunto de 54 aquaporinas diferencialmente expressas (DE) e cinco fatores de transcrição (ScEREB58, ScMYB103, ScARF43, ScNAC5 e ScZIM14) de um painel de 541 TFs DE. Também foram identificados os elementos regulatórios atuantes em *Cis* (*cis-Acting regulatory elements* - CARE) mais prevalentes em regiões promotoras de nove dos TFs induzidos. Esses dados fornecem pistas valiosas das respostas de variedades tolerantes e sensíveis ao estresse, bem como são fontes de alvos interessantes para estudos de transgenia e desenvolvimento de marcadores moleculares para seleção assistida.

# 2 REVISÃO DA LITERATURA

## 2.1 A cultura da cana-de-açúcar

A cana-de-açúcar (*Saccharum* spp. híbrido) é uma importante gramínea C4, de propagação vegetativa, com relevante habilidade em acumular elevados níveis de sacarose nos colmos (RAE *et al.*, 2005; DILLON *et al.*, 2007). As cultivares comerciais modernas dessa gramínea apresentam um dos genomas mais complexos do reino vegetal, com cariótipo 2n variando de 100 a 130 cromossomos (D'HONT *et al.*, 1996) e um tamanho 2C estimado em 10.000 Mb (D'HONT e GLASZMANN, 2001). Apresenta também elevados níveis de ploidia (D'HONT e GLASZMANN, 2001), além de frequentes eventos de aneupliodia (GRIVET e ARRUDA 2001), o que resultou em cultivares comerciais com grau de ploidia estimado em aproximadamente 12X (WANG *et al.*, 2010). Essa complexidade pode ser explicada pelos cruzamentos interespecíficos entre *Saccharum officinarum*, *S. barbieri*, *S. sinense* e as espécies selvagens *S. spontaneum* e *S. robustum* (D'HONT *et al.*,1996; GRIVET e ARRUDA 2001).

O valor dessa cultura é dado pelos seus principais produtos, o açúcar e o biocombustível etanol (MENOSSI *et al.* 2008). Também é utilizada *in natura*, como ração animal; ou como matéria-prima para a fabricação de rapadura, melado e aguardente. Seus resíduos também são aproveitados na agricultura e indústria – o vinhoto é transformado em adubo e o bagaço é queimado nas caldeiras para geração de energia térmica ou convertido em etanol de segunda geração (CAPUTO *et al.*, 2008; MENOSSI *et al.* 2008; CARDONA, QUINTERO e PAZ, 2010; HOFSETZ e SILVA, 2012).

No ano de 2013, a produção mundial dessa cultura totalizou aproximadamente dois bilhões de toneladas de biomassa e resultou em um

faturamento bruto de mais de 48 bilhões de dólares (FAO, 2015). A cana-de-açúcar é cultivada em mais de 100 países, ocupando mais de 26 milhões de hectares. O Brasil, com 684,77 milhões de toneladas estimadas para safra 2016/2017 que estão sendo cultivadas em 8,9 milhões de hectares, expansão de 318,4 mil hectares, ou 3,7% em comparação a safra anterior, desponta como maior produtor mundial dessa cultura, como também o principal processador da cana-de-açúcar em açúcar e etanol (FAO, 2015; CONAB, 2017). Segundo a Companhia Nacional de Abastecimento (CONAB, 2017), a produção nacional está concentrada no estado de São Paulo responsável pela produção de 381,70 milhões de toneladas cultivadas em 4,65 milhões de hectares, que corresponde a 52,3% de toda a área cultivada com cana-de-açúcar no país. Em seguida estão os estados de Goiás (com 10,4% da área total cultivada), Minas Gerais (9,2%), Mato Grosso do Sul (6.9%), Paraná com (7,0%), Alagoas (3,8%), Pernambuco (3%) e Mato Grosso (2,6%), com os demais estados representando (4,8%).

Apesar da expansão em sua área de cultivo, a lavoura brasileira de cana-de-açúcar apresentou na safra 2012/2013 um desenvolvimento aquém do estimado e 8,3% inferior ao da safra 2011/2012. O clima foi o principal agente da queda na produtividade, consequência das chuvas escassas em 2011 e 2012 para as regiões Sudeste e Centro-Oeste, que também afetaram o final da colheita de 2012 da região Nordeste, reduzindo a brotação e o desenvolvimento da cultura em diversas regiões do país, tanto nas áreas antigas como nas áreas de renovação e expansão (CONAB, 2012). Somando-se às intepéries inesperadas, a cultura da cana-de-açúcar está se expandindo

para regiões não tradicionais de cultivo, onde o déficit hídrico é o maior desafio à sua produtividade (VIANNA e SENTELHAS, 2014).

## 2.2 Déficit hídrico em vegetais: Causas, consequências e respostas moleculares

Os estresses abióticos, principalmente o déficit hídrico e alta salinidade, afetam negativamente o crescimento e desenvolvimento vegetal, representando uma das principais forças seletivas durante a evolução vegetal (ELSAYED, RAFUDEEN e GOLLDACK, 2014). Estima-se que os estresses abióticos possam reduzir o potencial de produção das plantas cultivadas em até 70% e, dentre esses estresses, a seca é considerada o mais prejudicial (GOSAL *et al.*, 2009). A deficiência hídrica, devido à sua capacidade de afetar simultaneamente um número variado de características morfológicas e funcionais nas plantas, acaba se tornando uma limitação importante para a produção agrícola sustentável em todo o mundo (SILVA *et al.*, 2008). As principais causas da seca na cultura da cana-de-açúcar estão associadas à variabilidade e má distribuição da pluviosidade nas áreas de cultivo (FARIAS *et al.*, 2008), que tendem a aumentar à medida que a cultura é expandida para regiões não tradicionais de plantio. Além disso, as previsões climáticas até o ano de 2030 do *Intergovernmental Panel on Climate Change* (IPCC) não são animadoras; são esperados períodos maiores de calor e seca, resultantes do aquecimento global (BATES *et al.*, 2008) e essas condições climáticas devem afetar negativamente a produção de áreas de cultivo da cana-de-açúcar.

Os principais prejuízos causados pela seca nas regiões de plantio de cana-de-açúcar são a redução de produtividade (CONAB, 2012) e o aumento do índice de mortalidade das soqueiras, reduzindo assim o tempo de

renovação dos canaviais (FARIAS *et al.*, 2008). Como a maior parte das plantações brasileiras de cana-de-açúcar é feita em regime de sequeiro (SILVA *et al.*, 2002), essa cultura é extremamente exposta às intempéries climáticas.

A redução de conteúdo hídrico celular em plantas provoca um complexo conjunto de respostas moleculares, que partem da percepção do estresse – ativando cascatas de transdução de sinas – e se manifesta como alterações aos níveis celular, fisiológico e morfológico (BRAY, 1993). Essas alterações incluem: fechamento dos estômatos, a diminuição do crescimento celular e fotossíntese, além da ativação da respiração celular. As plantas em condições de seca ainda respondem e se aclimatam acumulando osmólitos e proteínas especificamente envolvidas na tolerância ao estresse (SHINOZAKI e YAMAGUCHI-SHINOZAKI, 2007).

Os genes expressos durante a seca podem ser classificados em dois grupos funcionais (SHINOZAKI e YAMAGUCHI-SHINOZAKI, 2007):

- O primeiro grupo codifica proteínas associadas à tolerância da planta ao estresse, incluindo aquaporinas (Maurel *et al.*, 2008), proteínas transportadoras de aminoácidos (TEGEDER, 2012), proteinases (KIDRIČ *et al.*, 2014) e enzimas de remoção de espécies reativas de oxigênio (GILL e TUTEJA 2010), todas com função protetora nas células. Esse grupo ainda contém as enzimas requeridas na biossíntese de osmólitos (KIDO *et al.*, 2013), como prolina (Verbruggen e Hermans, 2008), açúcares (FERNANDEZ *et al.*, 2010) e dehidrinas (ABEDINI *et al.*, 2017)

- O segundo grupo inclui várias proteínas componentes de vias regulatórias, como fatores de transcrição (SINGH *et al.*, 2002), proteínas quinases (BARTELS *et al.*, 2005), fosfatases (SINGH *et al.*, 2015), enzimas

envolvidas no metabolismo de fosfolipídios (FOCKE *et al.*, 2003) e na biossíntese do ABA (GUO *et al.*, 2016).

Como indicado por Hu e Xiong (2014), uma maior elucidação dos complexos mecanismos de resistência à seca acelerará o desenvolvimento de novas variedades com maior resistência ao estresse. Boa parte desse conhecimento pode ser obtido através da genômica funcional, a partir de transcriptomas expressos em diferentes condições de déficit hídrico.

### 2.3 Análises globais da expressão gênica

O sequenciamento total de um genoma não esclarece muitos dos mecanismos moleculares complexos envolvidos na regulação da expressão gênica. Essa etapa, conhecida como genômica estrutural, é apenas o início de um processo que visa compreender a função das estruturas ativas dos genomas. A segunda etapa, denominada genômica funcional, consiste em utilizar as informações providas pela genômica estrutural para o desenvolvimento de abordagens metodológicas aplicadas na caracterização das funções gênicas (HIETER e BOGUSKI, 1997).

Entre as tecnologias mais utilizadas para análises de expressão gênica, destacam-se os métodos baseados em hibridação – por exemplo: *complementary* DNA (cDNA) *microarrays* (SAHA *et al.*, 1995) – e, especialmente, os métodos baseados em sequenciamento – *Expressed Sequence Tag* (EST) (ADAMS *et al.*, 1991); *Serial Analysis of Gene Expression* (SAGE) (VELCULESCU *et al.,* 1995) e suas versões aprimoradas – *Long Serial Analysis of Gene ExpressionI* (LongSAGE) (SAHA *et al.*, 2002) e *Super Serial Analysis of Gene Expression* (SuperSAGE) (MATSUMURA *et al.*, 2003); RNA-

seq (MORTAZAVI, 2008) e mais recentemente a *Massive analysis of cDNA ends* (MACE) (KAHL *et al.*, 2012).

As técnicas baseadas em sequenciamento utilizam a frequência observada de fragmentos dos transcritos expressos na amostra, obtidos a partir de clones de cDNA, que devem ser sequenciados, agrupados, contados e identificados. A abundância de fragmentos encontrados para um determinado gene pode fornecer uma estimativa de sua atividade na amostra estudada, além de possibilitar a descoberta de novos genes e possíveis transcritos alternativos (HARBERS e CARNINCI, 2005). Essas técnicas são consideradas técnicas de “plataforma aberta”, ou seja, permitem ter uma visão geral de quais são os genes que estão sendo expressos, sem a necessidade de seleção prévia que existe nos métodos baseados em hibridação (PINHEIRO, 2009).

### 2.3.1 Genômica estrutural e transcriptômica na cana-de-açúcar: ferramentas, disponibilidade de dados e aplicações

Na última década, o conhecimento nas áreas de genética, genômica estrutural e transcriptômica tem avançado consideravelmente com o aumento do processamento de dados e das novas plataformas de sequenciamento de alta performance (MARDIS, 2008; GLENN, 2011). Todavia, apesar da importância da cana-de-açúcar para a economia mundial (FAO, 2015) e das novas plataformas de sequenciamento, seu genoma ainda não está completamente sequenciado (SETTA *et al.*, 2014), estando disponíveis o seu DNA plastidial (GenBank: AE009947.2; CALSA JÚNIOR *et al.*, 2004), mitocondrial (GenBank: PRJNA380328) e trechos genômicos sequenciados a partir de BACs (REZENDE *et al.*, 2012; SETTA *et al.*, 2014; OKURA *et al.*, 2016).

Apesar das dificuldades quanto à ausência de genoma, foram realizados mais de 60 ensaios, segundo busca recente (03/2017) no *pubmed* (termos de busca: *sugarcane AND transcriptome*), em genômica funcional a partir de diferentes tecnologias transcriptômicas que vêm obtendo sucesso em gerar dados de regulação gênica, além de possibilitar a descoberta de genes candidatos a marcadores moleculares funcionais. Abaixo ensaios que relacionam as principais metodologias de obtenção de transcriptoma da cana-de-açúcar sob déficit hídrico:

1- Macroarranjos e derivados:

Rodrigues *et al.* (2009) monitoraram a resposta transcricional das cultivares SP83-5073 (tolerante a seca) e SP90-1638 (sensível a seca) frente ao déficit hídrico de supressão de rega (2, 8 e 10 dias), via *chip* de macroarranjo composto de 3.575 clones de ESTs de cana-de-açúcar. Essa equipe identificou uma grande diferença nos perfis transcricionais das variedades contrastantes, onde 165 genes foram diferencialmente expressos no acesso tolerante, enquanto que 432 genes foram diferencialmente expressos nos cultivares sensíveis, sendo identificada a repressão de importantes vias relatadas ao estresse nas plantas sensíveis. Lembke *et al.* (2012), por sua vez, analisaram, a partir de um arranjo de 21.901 sondas *sense* e *antisense* de oligonucleotideos de cana-de-açúcar, a resposta da cultivar sensível SP90-1638 sob déficit hídrico (24, 72 e 120 h de supressão de rega) e detectaram 987 sondas diferencialmente expressas (928 *sense* e 59 *antisense*) e por fim determinaram que os genes *antisense* apresentavam, em sua maioria, o mesmo padrão transcricional de seu respectivo gene *sense* (i.e. quando um transcrito *sense* era induzido o seu respectivo *antisense* também

aumentava sua transcrição). Recentemente, Li *et al* (2016) performaram um microarranjo, (4 × 44 K) com 15.593 sondas de cana-de-açúcar, na cultivar chinesa GT21 sobre três níveis de estresse hídrico (leve , moderado e severo; alcançados no terceiro, sétimo e nono dia se supressão de rega, respectivamente) e identificaram 1.501 genes diferencialmente expressos (821 induzidos e 680 reprimidos), 901 dos quais foram anotados em 36 categorias GO e 325 foram mapeados em 101 vias metabólicas importantes como biossíntese de metabolitos secundários, ribossomos, metabolismo de carbono.

2- *High-Throughput Super Serial Analysis of Gene Expression* (HT-SuperSAGE):

Kido *et al.* (2012), a partir de quatro bibliotecas HT-SuperSAGE [tolerante estressada (quatro cultivares tolerantes CTC15, CTC6, SP83-2847 e SP83-5073 sob supressão de rega por 24 h), tolerante controle (mesmos cultivares com rega mantida), sensível estressada (quatro cultivares sensíveis (CTC9, CTC13, SP90-3414, e SP90-1638 sob supressão de rega de 24 h) e sensível controle (mesmos cultivares com rega mantida)], sequenciaram mais de oito milhões de transcritos que corresponderam a 205,975 *tags* não redundantes (*unitags*) sequenciadas, com 75.404 *unitags* mapeadas em 164.860 ESTs, disponibilizando um painel de dados valioso para estudos posteriores.

3- Sequenciamento de alta performance de microRNAs:

Ferreira e colaboradores (2012) realizaram um sequenciamento de alta performance de oito bibliotecas de microRNAs (miRNAs) oriundos de tecido foliar de dois cultivares, RB867515 (tolerante) e RB855536 (sensível), submetidos a quatro tratamentos de seca (2, 4, 6 e 8 dias de supressão de

rega). Esse ensaio permitiu identificar 18 famílias de miRNA e, dentre todos os 21 miRNAs unicos (sem redundância), sete foram diferencialmente expressos. Posteriormente essa equipe validou os níveis de expressão, via RT-qPCR, de cinco miRNAs (ssp-miR164, ssp-miR394, ssp-miR397, ssp-miR399-seq 1 e miR528). Esses dados constituíram um aumento significante no número de miRNAs identificados na cana-de-açúcar.

Posteriormente, Thiebaut *et al* (2014) analisaram a expressão de sRNA em folhas e raízes de oito cultivares de cana-de-açúcar (CTC15, CTC6, SP83-2847 SP83-5073 e CTC9, CTC13, SP90-1638 e SP90-3414, tolerantes e sensíveis a seca, respectivamente), em resposta ao tratamento com déficit hídrico (24 horas de supressão de rega), através da análise de sequenciamento de alta performance oito bibliotecas de sRNAs, cada uma composta de 4 cultivares em *bulk*, totalizando 35 milhões de *reads* nas bibliotecas de RNA foliar e 24 milhões de reads (raízes), permitindo detectar 28 (folha) e 36 (raíz) famílias conservadas.

Thiebaut *et al.* (2014) constataram que genes da via de biosíntese de sRNAs foram reprimidos nas bibliotecas de cultivares tolerantes e induzidos nas compostas por sensíveis em resposta ao tratamento de supressão de rega. Porém, ao se validar os dados de expressão por genótipo em três dos oito acessos (SP90-1638 - sensível; SP83-2847 e SP83-5073 -tolerantes), via RT-qPCR, observaram respostas divergentes entre acessos tolerantes e dentro das réplicas biológicas desses acessos (apenas o perfil de expressão de miR159 foi confirmada no cultivar tolerante SP83-2847, mas não pelo SP83-5073, também tolerante). Esses resultados indicaram que ao menos uma

fração dos microRNAs tem uma regulação similar em cultivares de tolerância contrastante ao déficit hídrico.

4- RNA-seq:

Vargas *et al.* (2014) analizaram dados de sequenciamento de oito bibliotecas RNA-seq de RNAs extraídos de raízes e brotos de plantas da cultivar SP70-1143 inoculadas com a bactéria diazotrófica (mutualista) *Gluconacetobacter diazotrophicus* PAL5 e submetidas à supressão de rega (3 dias) para caracterizar os efeitos da associação entre a cana-de-açúcar e este organismo diazotrófico. Essa estratégia resultou em um transcriptoma de referência composto de mais de 160 milhões de *reads* (19 a 26 milhões de *reads* por biblioteca), representando 119.225 transcritos únicos, dentre os quais, 20.103 foram diferencialmente expressos nas bibliotecas não inoculadas com *G. diazotrophicus* e 25.781 nas inoculadas. Os ensaios de RT-qPCR revelaram que as raízes inoculadas possuíam maiores níveis de colonização por *G. diazotrophicus* sob tratamento de supressão de rega. Também observaram que as plantas colonizadas pela bactéria apresentaram vários genes que responderam à seca de maneira oposta às plantas não inoculadas, especialmente em raízes, indicando que essa interação pode disparar a ativação de genes que aumentam a tolerância à seca na planta hospedeira.

Quanto à disponibilidade de dados de genes expressos, destacam-se os seguintes bancos de dados públicos:

1- NUCEST *ou entrez EST* (https://www.ncbi.nlm.nih.gov/nucest):

Banco integrante ao portal do *National Center for Biotechnology Information* (NCBI), que apresenta para *Saccharum hybrid cultivar* 284.441

*Expressed sequence tags* (EST) geradas em diferentes tecidos, condições de tratamento e estágios de desenvolvimento.

2- *Gene index project* (ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/):

Este banco do *Computational Biology Laboratory* (Universidade de Harvard) inclui um total de 121.370 *unigenes* de *Saccharum officinarum* consistindo de 42.377 TCs (sequencias do tipo "*Tentative Consensus*") e 78.993 *singletons* (sequências não clusterizadas).

3- SUCEST (http://sucest-fun.org/):

Um banco originado de um grande consórcio de pesquisadores brasileiros que originou a maior coleção de ESTs de cana-de-açúcar do mundo, com aproximadamente 238,000 ESTs originadas de 26 diversas bibliotecas de cDNA (PINTO *et al.*, 2004).

4- GRASSIUS (Grassius; http://grassius.org/):

Base de dados que apresenta o conjunto completo de fatores de transcrição (FT) de cinco gramíneas, incluindo a cana-de-açúcar, com 1.354 FTs (YILMAZ *et al.*, 2009).

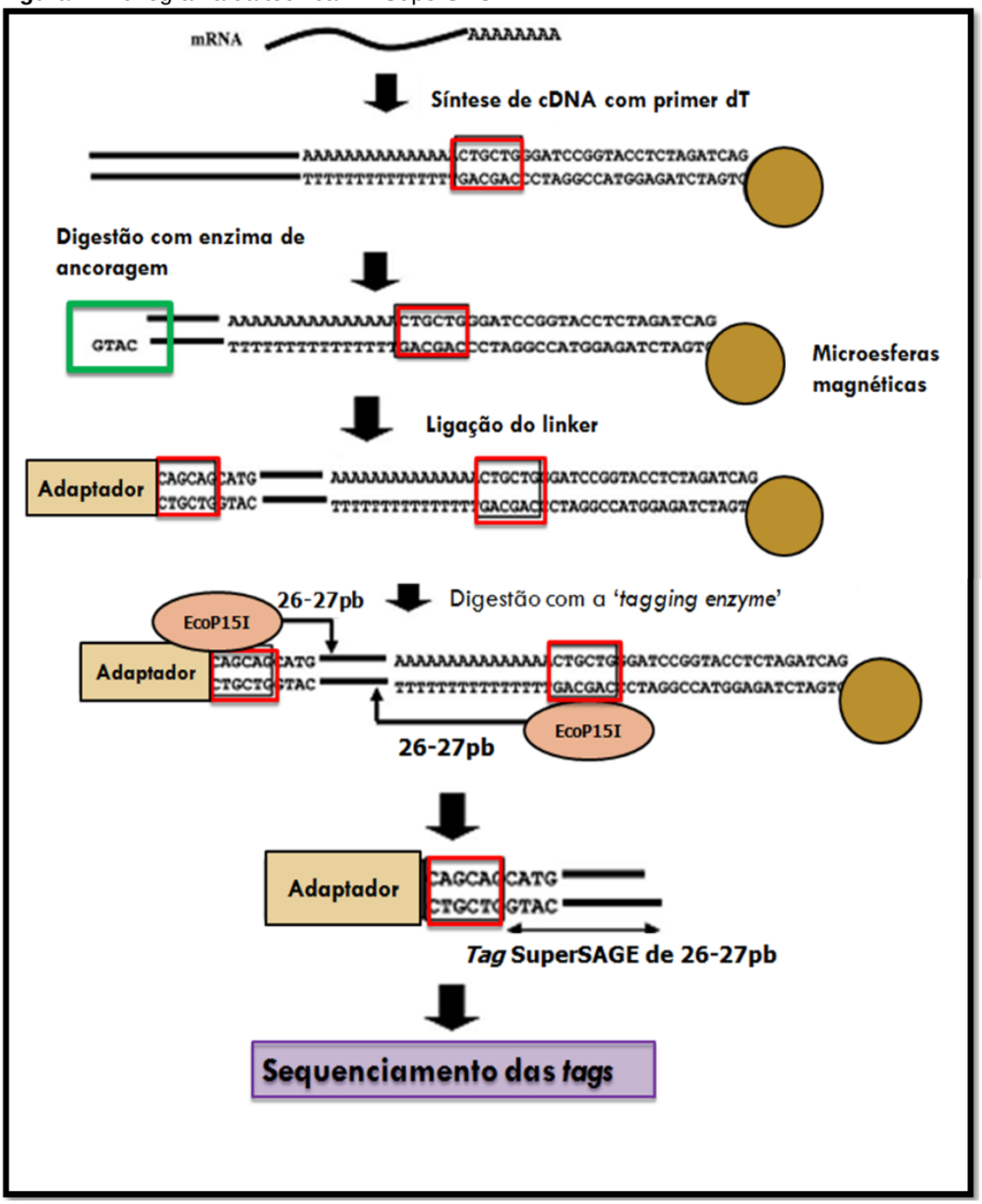
### 2.3.2 A tecnologia SuperSAGE

A SuperSAGE (*Super Serial Analysis of Gene Expression*), gera um abrangente perfil de transcrição, a partir de análises qualitativas e quantitativas de padrões de transcrição de milhões de *tags* de 26 pares de bases (pb) sequenciadas em um mesmo experimento, oriundas de um ou mais genomas de eucariotos por ensaio; baseando-se nas mudanças de frequência em RNA mensageiros possuidores de cauda poliA (MATSUMURA *et al.*, 2004). A técnica se torna ainda mais valiosa quando combinada com tecnologias de sequenciamento (de nucleotídeos) de alto desempenho (MATSUMURA *et al.*

2010; MATSUMURA *et al.*, 2011; MATSUMURA *et al.*, 2012). Essa combinação, denominada *High-Throughput*-SuperSAGE ou *deep*SuperSAGE, fornece muito mais dados (três a quatro ordens de magnitude), com custos relativamente baixos, redução de tempo e esforços quando comparados com a utilização do sequênciamento tradicional (MATSUMURA *et al.*, 2011).

Nesta metodologia, o cDNA é sintetizado a partir de um *primer* / adaptador dT. Após a digestão com a enzima de corte frequente, *Nla*III, um segundo adaptador é ligado ao final do cDNA clivado. Posteriormente, os fragmentos com os adaptadores são digeridos com *Eco*P15I (enzima de restrição do tipo III de corte a 26-27 pb da extremidade 3' do local de reconhecimento, demoninada *tagging* enzime), e purificados via eletroforese de gel de poliacrilamida para posterior sequenciamento (MATSUMURA *et al.*, 2008) (Figura 1).

**Figura 1**. Fluxograma da técnica HT-SuperSAGE.




**Fonte:** Traduzido de Matsumura *et al.* (2008).

O aumento no tamanho da *tag* melhora significativamente a eficiência da identificação dos genes correspondentes às *tags* em relação as duas metodologias anteriores - SAGE, com 14 pb (VELCULESCU *et al.,* 1995) e LongSAGE com 22 pb (MATSUMURA *et al.*, 2008), proporcionando novas aplicações à HT-SuperSAGE, dentre as quais estão:

1- Observação simultânea de perfis da expressão gênica de duas ou mais espécies que estão interagindo (como por exemplo as interações de parasita e hospedeiro descritas por Matsumura *et al.*, 2003);

2- Utilização de *tags* para produção de RNA de interferência *iRNA* (Matsumura *et al.*, 2004);

3- Utilização de *tags* como *primers* para 3'*RACE* PCR (MATSUMURA *et al.*, 2003, MATSUMURA *et al.*, 2004, MATSUMURA *et al.*, 2008);

4- Confecção de *chips microarray* (MATSUMURA *et al.*, 2004; MATSUMURA *et al.*, 2008).

Essa técnica tem sido utilizada com sucesso para estudar respostas em plantas frente a estresses bióticos e abióticos, identificando dezenas de milhares de genes diferencialmente expressos, tais como, arroz (MATSUMURA *et al.*, 2003), banana (COEMANS *et al.*, 2005), grão-de-bico (MOLINA *et al.*, 2008; MOLINA *et al.*, 2011), feijão-caupi (KIDO *et al.*, 2011), cana-de-açúcar (KIDO *et al.*, 2012), ambrósia americana (KELISH *et al.*, 2014), lima mexicana (CANALES *et al* 2016).

## 2.4 Programas de melhoramento da cana-de-açúcar no Brasil

Devido ao seu papel na economia mundial, a cana-de-açúcar está constantemente inserida em programas de melhoramento que visam à introdução de características de interesse agronômico, como resistência a pragas e patógenos, tolerância a herbicidas (CIDADE *et al.*, 2006), incluindo também aumento no teor de sacarose (BARBOSA *et al.*, 2015), tolerância a ambientes alagados (DEREN *et al.*, 1991), seca e frio (IAC, 2015) e maturação precoce (CARNEIRO *et al.*, 2015).

Atualmente existem no Brasil três grandes programas de melhoramento genético de cana-de-açúcar: (I) Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro (RIDESA, sigla RB), constituída de um convênio entre nove universidades federais (UFPR, UFSCar, UFV, UFRRJ, UFS, UFAL, UFRPE, UFG, UFMT e UFPI). Trata-se do maior programa de melhoramento da cana-de-açúcar brasileiro e de maior sucesso, sendo responsável por 62% das variedades cultivadas no país (RIDESA, 2015); (II) Centro de Tecnologia Canavieira (CTC), fundado em 1969, em uma iniciativa de um grupo de usinas da região de Piracicaba, SP, com o objetivo de investir no desenvolvimento de variedades mais produtivas e agregar qualidade à produção de açúcar e álcool. Reestruturado em 2004, com o objetivo de se tornar o principal centro mundial de desenvolvimento e integração de tecnologias da indústria sucroenergética, o CTC possui o maior e mais completo banco de germoplasma de cana-de-açúcar do mundo com mais de cinco mil variedades, resultando na capacidade de dobrar a taxa de inovação do setor (CTC, 2015); (III) Instituto Agronômico de Campinas (IAC), precursor da pesquisa agrícola no Brasil, iniciou suas atividades em 1892, com Franz W. Dafert, cientista austríaco que desenvolveu o primeiro estudo envolvendo 42 variedades nobres de cana (*Saccharum officinarum*) em duas condições de cultivo. Nas décadas de 40 e 50, foram avaliados os primeiros genótipos desenvolvidos em Campos e Piracicaba, que resultaram nas primeiras variedades criadas no Brasil. Atualmente, o programa foi responsável pelo lançamento de 39 variedades de cana-de-açúcar (IAC, 2015).

A despeito do sucesso de produção obtido em variedades de cana-de-açúcar advindas de programas de melhoramento genético clássico, o

lançamento de uma nova variedade de cana-de-açúcar demanda de oito a 12 anos de trabalho (MORAIS, 2015). Nesse sentido, uso da engenharia genética se mostra como uma excelente oportunidade de obtenção de uma agricultura mais eficiente e com resultados obtidos em curto prazo.

# 3 OBJETIVOS

## 3.1 Objetivo geral

Identificar e validar alvos moleculares de cana-de-açúcar com base em *tags* HT-SuperSAGE diferencialmente expressas sob condições de déficit hídrico (resposta imediata após 24 h de supressão de rega).

## 3.2 Objetivos específicos

1. Identificar novos alvos moleculares responsivos (induzidos e reprimidos) ao déficit hídrico.

2. Identificar, em regiões promotoras de transcritos de cana-de-açúcar e sequências genômicas de sorgo, sequências regulatórias em *cis*.

3. Validar as expressões, via RTqPCR, de ao menos 25 alvos selecionados com base nas *tags* HT-SuperSAGE.

# 4 ARTIGO I

## Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-time PCR in Drought-stressed Sugarcane

Artigo científico publicado no periódico The ScientificWorld Journal (ISSN: 1537-744X)

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 357052, 12 pages
http://dx.doi.org/10.1155/2014/357052

*Research Article*

# Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-Time PCR in Drought-Stressed Sugarcane

**Roberta Lane de Oliveira Silva,[1] Manassés Daniel Silva,[1] José Ribamar Costa Ferreira Neto,[1] Claudia Huerta de Nardi,[2] Sabrina Moutinho Chabregas,[3] William Lee Burnquist,[3] Günter Kahl,[4] Ana Maria Benko-Iseppon,[1] and Ederson Akio Kido[1]**

[1] *Federal University of Pernambuco (UFPE/CCB/Genética), 50670-420 Recife, PE, Brazil*
[2] *ABC Medical School, 09060-650 São Paulo, SP, Brazil*
[3] *Sugarcane Technology Center (CTC), 13400-970 Piracicaba, SP, Brazil*
[4] *Institute of Molecular Biosciences, Goethe University Frankfurt am Main, 60438 Frankfurt am Main, Germany*

Correspondence should be addressed to Ederson Akio Kido; ederson.kido@gmail.com

Received 14 January 2014; Revised 20 April 2014; Accepted 21 April 2014; Published 2 June 2014

Academic Editor: Juan J. Loor



One of the most challenging aspects of RT-qPCR data analysis is the identification of reliable reference genes. Ideally, they should be neither induced nor repressed under different experimental conditions. To date, few reference genes have been adequately studied for sugarcane (*Saccharum* spp.) using statistical approaches. In this work, six candidate genes (αTUB, GAPDH, H1, SAMDC, UBQ, and 25S rRNA) were tested for gene expression normalization of sugarcane root tissues from drought-tolerant and -sensitive accessions after continuous dehydration (24 h). By undergoing different approaches (GeNorm, NormFinder, and BestKeeper), it was shown that most of them could be used in combinations for normalization purposes, with the exception of SAMDC. Nevertheless three of them (H1, αTUB, and GAPDH) were considered the most reliable reference genes. Their suitability as reference genes validated the expression profiles of two targets (AS and PFPα1), related to SuperSAGE unitags, in agreement with results revealed by previous *in silico* analysis. The other two sugarcane unitags (ACC oxidase and PIP1-1), after salt stress (100 mM NaCl), presented their expressions validated in the same way. In conclusion, these reference genes will be useful for dissecting gene expression in sugarcane roots under abiotic stress, especially in transcriptomic studies using SuperSAGE or RNAseq approaches.

## 1. Introduction

Sugarcane (*Saccharum* spp.) is a major crop with vegetative propagation and capacity to accumulate high levels of sucrose in the culms [1–3]. World production of this crop in 2011 resulted in about two billion tons of raw material, which corresponded to a gross income of more than US$ 52 billion [4]. Despite the economic importance of sugarcane, the knowledge of relevant genetic mechanisms remains challenging, due to the fact that this crop presents one of the largest and most intricate genomes of the plant kingdom, with diploid numbers ranging from 100 to 130 chromosomes, indicating a high ploidy level, as well as regular aneuploidy events [5–7]. Because of this complexity, the use of molecular tools represents an attractive approach to the improvement of sugarcane breeding programs. Moreover, transcriptomic studies have been prioritized, allowing identification of candidate genes involved in developmental processes and plant responses to environmental cues, which have eventually led to the discovery of functional molecular markers [8]. Reverse transcription quantitative real-time PCR (RT-qPCR) is based on a high specific polymerase chain reaction associated with sensitive fluorescence, allowing the detection of variations in gene expression, including discreetly transcribed genes [9, 10]. This technology has been used as a diagnostic tool for identification of plant pathogens, transgene expression [9],

and human diseases [11, 12] and confirmation of transcriptional profiles generated by different methodologies, such as EST libraries [13], Microarray [14], HT-SuperSAGE [15], and RNAseq [16]. The reliability of RT-qPCR data based on relative quantification is depending upon comparative transcription of target genes to stable reference genes [17, 18]. The use of reference genes that undergo changes in transcription across experimental groups can dramatically alter the conclusions on targeted gene expression [18]. In order to test for normalization of reference gene expression, several statistical algorithms, such as GeNorm [17], NormFinder [19], and BestKeeper [20], have recently been developed. Despite the importance of proper reference gene selection for reliable and accurate RT-qPCR assays, most reports involving sugarcane have not described, or compared, methods in order to determine the efficiency of reference genes [21–23], suggesting arbitrary criteria for this selection. To our knowledge, only one systematic study aimed to assess, standardize, and validate reference genes (GAPDH, $\beta$-tubulin, $\beta$-actin, and 25S rRNA) for tissue and genotype-specific gene expression analysis in sugarcane [24]. Additionally, this kind of study has not been carried out before under stress conditions which often alter the behavior of some genes. Thus, a rigorous selection of reference genes for expression profiling validation in sugarcane under biotic or abiotic stress was yet to be described. In the present work, screening and validation of new reliable reference genes for expression analysis in sugarcane roots were carried out. Besides, additional resources for target validation were evaluated, especially considering comprehensive transcription profiling, like those provided by HT-SuperSAGE [25], in sugarcane, revealing hundreds of candidate genes responsive against drought stress, requiring subsequent validation.

## 2. Materials and Methods

*2.1. Plant Material and Treatments.* Sugarcane drought assay: accessions were previously evaluated in a glasshouse trial conducted by the Center for Sugarcane Technology (CTC) in Piracicaba, Brazil (22° 41′ S; 47° 33′ W), aiming to identify drought-tolerant and drought-sensitive ones. Four accessions were selected as drought-tolerant (CTC6, CTC15, SP83-2847, and SP83-5073) and another four as drought-sensitive (CTC9, CTC13, SP90-3414, and SP90-1638). Some of them were previously reported as tolerant and sensitive to drought based on chlorophyll and water content measurements together with field observations [28]. Briefly, for the drought stress assay, plants of each accession were grown under glasshouse conditions (30.2 ± 5.7°C (maximum), 16.8 ± 1.9°C (average), and 9.3 ± 3.0°C (minimum) and 71.5 ± 5.1% (average) relative humidity under natural photoperiod) in 40 L pods in a randomized experimental design (comprising six repetitions) under daily irrigation (4 L·h$^{-1}$) until reaching three months of age. Plants were submitted to drought conditions by continuous dehydration caused by the interruption of irrigation during 24 hours. Roots of both, stressed and unstressed plants, were collected, immediately frozen in liquid nitrogen, and stored at −80°C until total RNA extraction. On the other hand, the salinity stress assay was performed as follows: acclimated plants of the RB931011 clone *in vitro* cultivated (referred to as salt tolerant by the Brazilian RIDESA program of sugarcane breeding) were grown in a greenhouse (UFPE, Recife, PE, Brazil; 8° 04′S, 34° 55′W) in pots containing washed sand (washed) and they were watered daily with Hoagland solution, throughout three months. Later, NaCl (100 mM) was added to the nutritive solution as the salt stress. Roots from both stressed and nonstressed (negative control) plants were collected after stress induction (30 and 90 min) and immediately frozen in liquid nitrogen until a total RNA extraction was performed.

*2.2. Total RNA Isolation, Purification, cDNA Synthesis, and HT-SuperSAGE Libraries.* Total RNA was extracted with RNeasy Mini kit (Qiagen) according to the manufacturer's instructions, treated with DNAse (Qiagen), and purified with RNeasy Mini kit (Qiagen). RNA samples were quantified using Quant-iT RNA assay kit (Invitrogen) with the Qubit fluorometer (Invitrogen). RNA integrity was verified in 1.5% agarose gel electrophoresis with blue-green loading dye (LGC Biotechnology) staining. The purified RNA (1 $\mu$g) of each sugarcane accession was reverse-transcribed using QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions and resuspended in a final volume of 20 $\mu$L. The cDNA synthesis reaction was incubated at 42°C for 2 minutes (genomic DNA digestion), 42°C for 15 minutes (reverse transcription), and 95°C for 3 minutes (enzyme inactivation) and stored at −20°C. The procedures for HT-SuperSAGE library generation followed Matsumura et al. [29], including the attachment of library-specific adaptors carried out by GenXPro GmbH (Frankfurt, Germany), allowing the identification of library-specific reads after SOLEXA sequencing. Concerning the drought stress, four libraries were generated as described by Kido et al. [25]: the bulk of tolerant accessions under stress and the respective negative control and the bulk of sensitive accessions after stress and the respective negative control. The bulks were composed by equimolar amounts of poly-A$^+$ mRNA from all accessions comprising the respective library. In relation to the salt stress, equimolar amounts of total RNA from each sample/time were assembled to compose the two bulks used to generate the SuperSAGE libraries (stressed and control).

*2.3. Primer Design, Amplification Efficiency, and RT-qPCR Analysis.* Sugarcane ESTs (Table 1) from dbEST database (http://www.ncbi.nlm.nih.gov/nucest), related to independent pathways as an attempt to minimize the effects of coregulation, were used for primers design with the Primer 3 software [30] with minor modifications: the amplicon length range was set to 70–200 bp, the melting temperatures were between 40°C (minimum), 50°C (optimum), and 60°C (maximum), and the CG content ranged from 45 to 55% (optimum of 50%). In relation to H1 and SAMDC, the primers sequences were obtained from the literature [23]. All primer pairs were synthesized by Bioneer Corporation (South Korea) and some details of these primers are given

Table 1: Potential reference genes and target genes with the individual accession numbers, annotations, and primer sequences.

| Gene (accession numbers) | Source* | Predicted function | Description | Primer sequences (forward/reverse) |
|---|---|---|---|---|
| Reference genes | | | | |
| αTUB (CN607271) | dbEST | Structural constituent of cytoskeleton | Alpha-tubulin | (F) CCATTGGCAAGGAGATTGTT<br>(R) TCCACCAACTGCATTGAAGA |
| GAPDH (CA254672) | dbEST | Glycolysis, gluconeogenesis | Glyceraldehyde 3 phosphate dehydrogenase | (F) GGTTCACTTGAAGGGTGGTG<br>(R) TGAGGTGTACCTGTCCTCGTT |
| H1 (CA116806) | [23] | Chromatin condensation | Histone H1 | (F) CGCACACGCACACTGAAAG<br>(R) CGGTGGCCATGATCAAAAA |
| SAMDC (CA127376) | [23] | Polyamine and triamine biosynthesis | S-Adenosylmethionine decarboxylase | (F) TGCTGCTGAAGACGCTGTTG<br>(R) TCGCCTTCAAAGCAGTGTAGAAC |
| UBQ (CA077378) | dbEST | Protein degradation | Ubiquitin | (F) ACCGAAGGTTGCATTCAAGAC<br>(R) GGGTTTGGGTCCGTTAGAAG |
| 25S rRNA (BQ536525) | dbEST | Translation | 25S ribosomal RNA | (F) GCAGCCAAGCGTTCATAG<br>(R) CGGCACGGTCATCAGTAG |
| Target genes | | | | |
| PFPα1 (XM_004973200.1) | dbEST | Carbohydrate metabolism, glycolysis | Pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit | (F) TTATGAGTTGCTGCGAGAGAAG<br>(R) TATCTCAATGTCGCCCATGTAG |
| AS (FM212633.1) | dbEST | Asparagine synthesis | Glutamine-dependent asparagine synthetase | (F) CCAGAGAACACACCCACAAC<br>(R) ATGCCACACTAGGACCTCCA |
| PIP1-1 (CF572112) | dbEST | Integral to membrane, water transport | Plasma membrane intrinsic protein 1-1 | (F) GTTCCTATCCTTGCCCCACT<br>(R) AGGCGTGATCCCTGTTGTAG |
| ACC oxidase (TC127289) | SoGI | Ethylene biosynthesis | 1-Aminocyclopropane-1-carboxylate oxidase | (F) GGGACCTCTTGCAGATAATGTC<br>(R) CTCTGGCAATGGTCCATAGAA |

*Databases: dbEST (NCBI; http://www.ncbi.nlm.nih.gov/) and Gene Index (SoGI; http://compbio.dfci.harvard.edu/tgi/); Rodrigues et al. [23]. αTUB: alpha-tubulin; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; H1: histone H1; SAMDC: S-adenosylmethionine decarboxylase; UBQ: ubiquitin; 25S rRNA: 25S ribosomal RNA; PFPα1: pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit (EC 2.7.1.90); AS: glutamine-dependent asparagine synthetase (EC 6.3.5.4); PIP1-1: plasma membrane intrinsic protein; ACC oxidase: 1-aminocyclopropane-1-carboxylate oxidase (EC 1.14.17.4).

in Table 1. An initial standard PCR was performed with the potential reference genes, using the sugarcane cDNA samples (bulk of tolerant and sensitive accessions), in order to investigate the PCR products. Amplicons were analyzed on 1.5% agarose gel electrophoresis followed by blue-green loading dye staining (LGC Biotechnology). Additionally, a dissociation curve analysis was made, in RT-qPCR assay, to confirm the specificity of the amplification by the candidate genes. Calibration curves using a dilution series of the cDNA pool (1, $10^{-1}$, $10^{-2}$, $10^{-3}$, and $10^{-4}$) were made to calculate the PCR amplification efficiencies ($E = 10^{-1/\text{slope}}$) [27] for each quantified candidate gene, their respective correlation coefficients ($R$), and $y$ interceptors. The RT-qPCR amplifications were performed on LineGene 9660 model FQD-96A (Bioer), using SYBR Green detection. Each reaction mixture comprised 1 $\mu$L of template cDNA (diluted 5-fold), 5 $\mu$L of HotStart-IT SYBR Green qPCR Master Mix 2x (USB), 0.05 $\mu$L of ROX (normalization dye), 1.95 $\mu$L of ultrapure water, and 1 $\mu$L of each primer (0.05 $\mu$M), forming a final volume of 10 $\mu$L. Three biological and three technical replicates were used in each run for RT-qPCR analysis, and a no template control (NTC) was also included. The reactions were subjected to an initial denaturation step of 95°C for 2 min, followed by 40 cycles at 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s in 96-well reaction plates. The dissociation curves were analyzed at 65–95°C for 20 min after 40 cycles. The baseline and quantification cycle (Cq) were determined using the LineGene 9600 software (version 1.1.10).

*2.4. Data Analysis.* The potential reference genes were ranked, and the number of candidate genes required for an optimal normalization was indicated according to their gene expression stability using sugarcane cDNA samples, after being analyzed by the GeNorm (version 3.5) [17], NormFinder (version 0.953) [19], and BestKeeper (version 1) [20] software. The GeNorm and NormFinder input data were based on relative quantities applying the ΔΔCq method [31]. The GeNorm software determines the reference gene stability measurement ($M$) as the average pairwise variation of each reference gene with all the other reference genes and enables the elimination of the least stable gene and the recalculation of the $M$ values, resulting in the ranking of the most stable genes. The average expression stability value ($M$-value) was a parameter for quantification of stable reference gene candidates, in which a low $M$-value indicated a more stable expression [17]. The NormFinder tool was applied to identify and rank the most suitable genes for RT-qPCR normalization from the set of candidates, considering intragroup and intergroup variations, in a model-based approach of mixed linear effect modeling [19]. The BestKeeper software, using raw Cq values as input, was applied to identify the most stable expressed genes by a Pearson correlation coefficient (geometric mean of Cq values of candidate genes), calculating the standard deviation (SD) of Cq values among the entire data set. The relative gene expression levels (based on the relative quantities after the ΔΔCq method) were evaluated with the REST© tool (Relative Expression Software Tool, version 2.0.13), which bases its performance on pairwise comparisons using randomization and bootstrapping techniques (Pairwise Fixed Reallocation Randomization Test©) [32]. The normalization of the RT-qPCR was performed by taking the geometric averages of the combined reference genes, using the negative control to normalize this relative expression, and testing the hypothesis of significant differences between the control and treatment. With the input of multiple target and reference genes and based on the normalized values of the target genes, the software indicates the direction of the difference between the groups, as well as their $P$ value. Also, the MIQE guidelines (The Minimum Information for Publication of Quantitative Real-Time PCR Experiments) [26] were followed in order to increase transparency and reliability of the results obtained.

## 3. Results

*3.1. RNA Integrity, Specificity, and Efficiency Amplifications.* All the reference candidate genes (αTUB, GAPDH, H1, SAMDC, UBQ, and 25S rRNA) amplified the cDNAs generated from the RNAs samples (Supplemental Figure 1(a) in Supplementary Material available online at http://dx.doi.org/10.1155/2014/357052) using the proposed primers. Based on standard PCR amplifications, only a single product was observed with a specific primer pair (Supplemental Figure 1(b)) and these results were supported by the dissociation curve analysis (Supplemental Figure 1(c)). Based on the standard curves using a serial dilution of the cDNA pool (Supplemental Figure 2), the real-time PCR amplification efficiency ($E$), considering the selected six candidate genes, ranged from 98.34% to 100.89%, with correlation coefficients ($R$) varying from 0.984 to 0.999, while the slopes ranged from −3.50 to −3.21 (Table 2). Considering the efficiency of 100%, the value of the expected slope should be −3.32, while slopes ranging from −3.10 to −3.58 would represent efficiency comprising 90% to 110%, thereby characterizing acceptable reactions. These parameters derived from the RT-qPCR analysis, and others in accordance with the MIQE Guidelines, are shown in Tables 2 and S2. The results showed favorable conditions for amplification, efficiency in successive dilutions, and acceptable variations in gene expression across samples, representing potential for choosing a suitable reference gene. Thus, in gene expression studies, fluctuations due to pipetting errors, variations in the quantification of samples, or the concentration of reagents could be normalized with the aid of these suitable reference genes [33].

*3.2. Gene Expression Stability of the Reference Gene Candidates.* The six candidate genes selected for normalization (αTUB, GAPDH, H1, SAMDC, UBQ, and 25S rRNA) in RT-qPCR tests showed Cq values ranging from 13.06 to 28.00 (Supplemental Table S1). Most of the candidate genes presented Cq values with slight variations (below one cycle), except UBQ and SAMDC. Based on these values, 25S rRNA was the most abundantly transcribed gene (average Cq = 14.00), while αTUB was the least abundant (average Cq = 27.48). These data, in order to assess the gene expression stability of the reference gene candidates, were used in GeNorm [17], NormFinder [19], and BestKeeper [20] analysis.

Table 2: Parameters derived from RT-qPCR data analysis*, including PCR amplification efficiency ($E$) established by calibration curves for each quantified target.

| Gene | Tm (°C) | Product size (bp) | Average Cq | $E$ (%) | NTC (Cq) | Correlation coefficient ($R$) | Slope | $Y$ intercept |
|---|---|---|---|---|---|---|---|---|
| αTUB | 75.9 | 104 | 27.48 | 99.53 | 35.82 | 0.998 | −3.33 | 37.34 |
| GAPDH | 81.8 | 100 | 23.69 | 100.89 | N.D. | 0.984 | −3.30 | 41.40 |
| H1 | 78.0 | 57 | 24.76 | 97.41 | 33.59 | 0.999 | −3.39 | 37.92 |
| SAMDC | 79.0 | 60 | 22.74 | 99.66 | 36.62 | 0.992 | −3.33 | 34.36 |
| UBQ | 81.2 | 153 | 24.64 | 98.34 | N.D. | 0.999 | −3.36 | 37.86 |
| 25S rRNA | 82.9 | 108 | 14.00 | 99.82 | N.D. | 0.999 | −3.33 | 35.54 |
| PFPα1 | 83.6 | 151 | 21.92 | 92.99 | 32.90 | 0.988 | −3.50 | 40.44 |
| AS | 79.8 | 112 | 24.94 | 105.05 | N.D. | 0.990 | −3.21 | 40.26 |
| PIP1-1 | 84.6 | 134 | 24.47 | 95.55 | N.D. | 0.995 | −3.43 | 37.50 |
| ACC oxidase | 82.2 | 152 | 30.21 | 93.52 | N.D. | 0.990 | −3.49 | 44.65 |

*Based on MIQE Guidelines [26]. Tm: melting temperature (°C); N.D.: not detected; NTC: no template control; αTUB: alpha-tubulin; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; H1: histone H1; UBQ: ubiquitin; SAMDC: S-adenosylmethionine decarboxylase; 25S rRNA: 25S ribosomal RNA; PFPα1: pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit (EC 2.7.1.90); AS: glutamine-dependent asparagine synthetase (EC 6.3.5.4); PIP1-1: plasma membrane intrinsic protein; ACC oxidase: 1-aminocyclopropane-1-carboxylate oxidase (EC 1.14.17.4).

Table 3: Expression stability values for sugarcane candidate calculated by NormFinder and BestKeeper software.

| Ranking | NormFinder analysis | | BestKeeper analysis | |
|---|---|---|---|---|
| | Gene name | $M$ | Gene name | CV ± SD |
| 1 | αTUB | 0.11 | H1 | (1.06 ± 0.26) |
| 2 | H1 | 0.16 | αTUB | (1.06 ± 0.29) |
| 3 | GAPDH | 0.19 | GAPDH | (1.33 ± 0.31) |
| 4 | 25S rRNA | 0.28 | 25S rRNA | (3.89 ± 0.55) |
| 5 | UBQ | 0.31 | UBQ | (2.04 ± 0.50) |
| 6 | SAMDC | 0.33 | SAMDC | (2.53 ± 0.57) |

$M$: average expression stability value; CV: coefficient of variance; SD: standard deviation; αTUB: alpha-tubulin; H1: histone H1; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; 25S rRNA: 25S ribosomal RNA; UBQ: ubiquitin; SAMDC: S-adenosylmethionine decarboxylase.

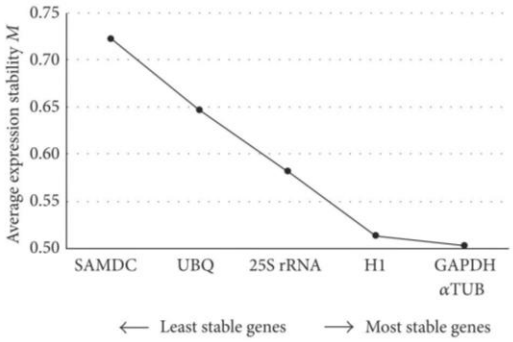



Figure 1: Average gene expression stability values ($M$) of six sugarcane potential reference genes (αTUB: alpha-tubulin; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; H1: histone H1; SAMDC: S-adenosylmethionine decarboxylase; UBQ: ubiquitin; 25S rRNA: 25S ribosomal RNA) based on the GeNorm analysis [17].

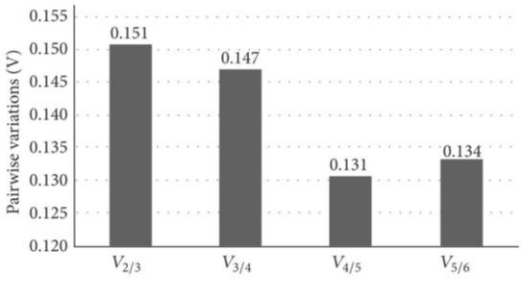



Figure 2: Pairwise variation ($V$) analysis for six potential reference genes of sugarcane (α-tubulin, glyceraldehyde 3 phosphate dehydrogenase, histone H1, S-adenosylmethionine decarboxylase, ubiquitin, and 25S rRNA) based on the GeNorm analysis [17]. The addition to the two most stable genes (αTUB and GAPDH) of a random third gene ($V_{2/3}$ = 0.15), a fourth gene ($V_{3/4}$ = 0.14), or even more ($V_{4/5}$ and $V_{5/64}$) still exhibited desirable values (basically below than 0.15).

*3.2.1. GeNorm Analysis.* Considering the average expression stability values ($M$-value), αTUB ($M$ = 0.61), GAPDH ($M$ = 0.62), and histone H1 ($M$ = 0.63) were the most stable genes while SAMDC represented the most variable ($M$ = 0.87) gene. However, all of them showed an expressive high stability with $M$-values below 1 (Figure 1), suggesting that all the six candidates may be adequate for normalizing gene expression data under the conditions used in the present work. Besides, based on the pairwise variation ($V$) data (Figure 2), it was possible to determine the optimal number of reference genes required for the relative quantification analysis and to investigate the effect of gene addition in this normalization. The data suggested that the addition to the two most stable genes (αTUB and GAPDH) considering a third gene ($V_{2/3}$ = 0.15; Figure 2), a fourth ($V_{3/4}$ = 0.14), or even more genes ($V_{4/5}$ and $V_{5/6}$; Figure 2) still exhibited desired values (below 0.15 as proposed by Vandesompele et al. [17]). To normalize the gene expression in the above mentioned sugarcane samples, αTUB, GAPDH, and H1 seem to be sufficient (Figure 1).

*3.2.2. NormFinder and BestKeeper Analysis.* Basically, the gene expression stability ranking provided by the NormFinder and BestKeeper software exhibited the same order, with only the first two candidates switching places comparing the ranking (Table 3). The two software programs identified histone H1 ($M$ = 0.28; CV ± SD = 1.06 ± 0.26) and αTUB ($M$ = 0.32; CV ± SD = 1.06 ± 0.29) as the most stable genes, followed by GAPDH, 25S rRNA, UBQ, and SAMDC (see respective values in Table 3). Again, SAMDC showed the highest instability, in agreement with the GeNorm results. All the three software programs presented fairly consistent results showing the first three (H1, αTUB, and GAPDH) as the most stable and reliable genes for RT-qPCR data normalization. Two of them (H1 and αTUB) are reported as RT-qPCR normalizing genes suitable for sugarcane roots under abiotic stress for the first time.

*3.3. Normalization of Target Genes.* In order to assess the applicability of the recommended histone H1, αTUB, and GAPDH as reference genes in relative expression studies using RT-qPCR, four targets based on sugarcane SuperSAGE unitags annotated as glutamine-dependent asparagine synthetase (AS, EC 6.3.5.4), pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit (PFPα1, EC 2.7.1.90), plasma membrane intrinsic protein (PIP1-1), and 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase, EC 1.14.17.4) were evaluated (Table 4). HT-SuperSAGE survey

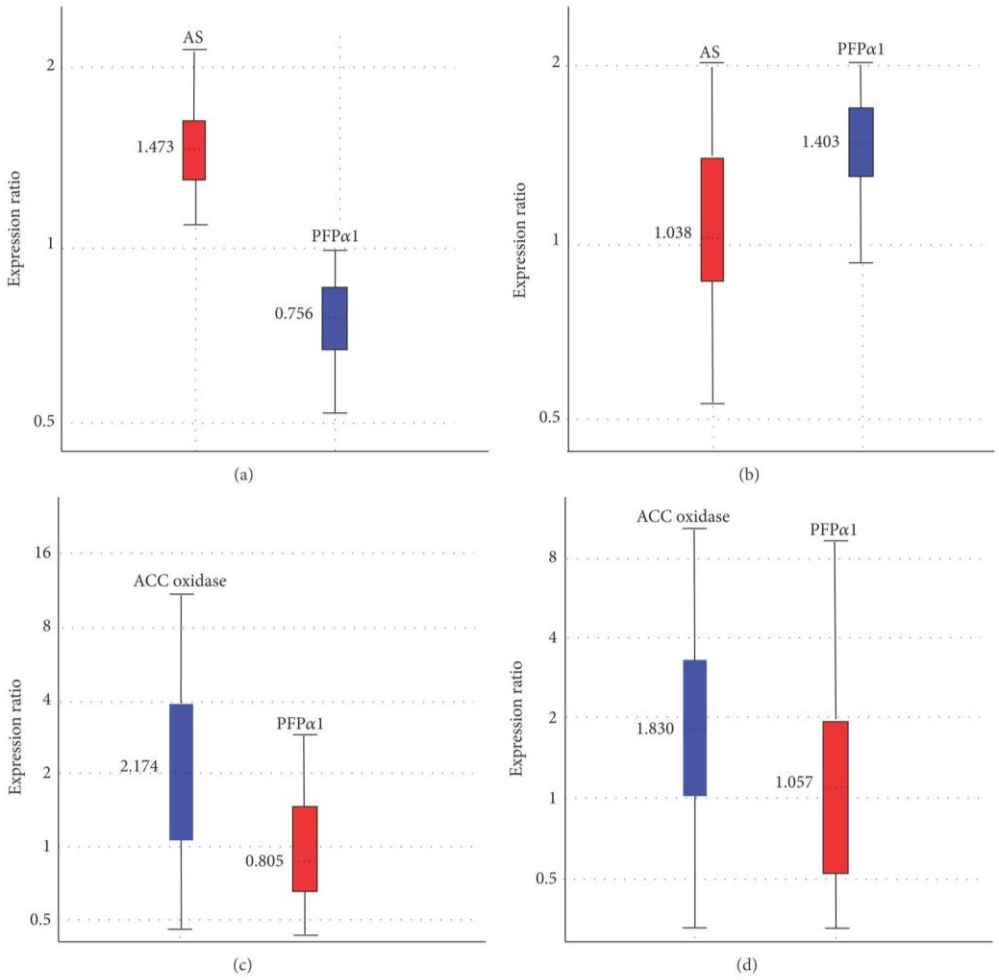



FIGURE 3: Relative expression of glutamine-dependent asparagine synthetase (AS), pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFPα1), plasma membrane intrinsic protein 1-1 (PIP1-1), and 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase) by the REST software v. 2.0.13 (after the ΔΔCq method) in cDNAs of sugarcane roots under abiotic stress (24 h of continuous dehydration or salt stress; 100 mM NaCl), normalized by the reference genes H1, αTUB, and GAPDH. (a) Tolerant bulk (CTC6, CTC15, SP83-2847, and SP83-5073 accessions) compared to its negative control. (b) Susceptible bulk (CTC9, CTC13, SP90-1638, and SP90-3414) compared to its negative control. (c) Salt-tolerant accession (RB931011) after 30 min of salt stress compared to its negative control. (d) RB931011 after 90 min of salt stress compared to its negative control. Relative expression with the median value (horizontal dotted line at the colored box) and range comprising 100% of the observations (horizontal bars), being 50% of them in the confidence interval at 95% (colored box).

pointed SD282748 unitag as a potential AS being upregulated 1.92 times in the drought-tolerant bulk after the stress (24 h of continuous dehydration) as compared to the unstressed control while no relevant unitag frequency change was observed ($P < 0.05$) with the sensitive contrast (Table 4). The RT-qPCR relative quantification results confirmed the overexpressed status in relation to both cDNA bulks with 1.473-fold change for the tolerant bulk compared to its negative control (Figure 3(a)) and no significant change (1.038 times) considering the sensitive bulk in the comparative contrast (Figure 3(b)). In turn, SD179780 unitag (annotated as PFPα1) did not respond ($P < 0.05$) to water deficit stimulus in any contrast analyzed involving sugarcane drought-tolerant or -sensitive accessions, in agreement with the RT-qPCR results, showing constitutive expression of this gene (Figures 3(a) and 3(b), resp.). In an attempt to explore the use of the proposed reference genes, a target relative to the ASS122537 unitag (ACC oxidase) from the salt HT-SuperSAGE libraries,

Table 4: Relative expression rates of target genes (PFPα1, AS, PIP1-1, and ACC oxidase) based on RT-qPCR with roots, cDNAs of sugarcane accessions under abiotic stress, and respective unitag regulation by SuperSAGE analysis covering drought[a] stress (24 h of continuous dehydration) or salt[b] stress (100 mM NaCl).

| Unitag | Annotation | SuperSAGE [FC/Regulation*] | | RT-qPCR*& | |
|---|---|---|---|---|---|
| | | Tolerant | Sensitive | Tolerant | Sensitive |
| SD282748[a] | AS | 1.92#/UR | −1.10#/ns | 1.473#/UR | 1.038#/ns |
| D179780[a] | PFPα1 | 1.99#/ns | −1.07#/ns | 0.756#/ns | 1.403#/ns |
| ASS122537[b] | ACC oxidase | 1.95/UR | — | 2.174/UR (30′)<br>1.830/ns (90′) | — |
| ASS140030[b] | PIP1-1 | −1.31/ns | — | 0.805/ns (30′)<br>1.057/ns (90′) | — |

AS: glutamine-dependent asparagine synthetase (EC 6.3.5.4); PFPα1: pyrophosphate fructose-6-phosphate 1-phosphotransferase alpha subunit (EC 2.7.1.90); ACC oxidase: 1-aminocyclopropane-1-carboxylate oxidase (EC 1.14.17.4); PIP1-1: plasma membrane intrinsic protein. #Bulk with four accessions by each library; FC: fold change [ratio of the frequencies (normalized to 1,000,000) observed in the stressed library in relation to the control library]; &relative expression level by REST software (v.2.0.13) after the ΔΔCq method, *$P < 0.05$ [27]; UR: upregulated; ns: not significant at $P < 0.05$. The time in the parentesis represents the salt stress exposition (min).

which was induced (UR) by the salt-tolerant accession (1.95 times after the salt stress exposition, NaCl 100 mM), showed overexpression after 30 min according to RT-qPCR results (Figure 3(c) and Table 4). Another target relative to the ASS140030 unitag (PIP1-1) from sugarcane after salt stress presented RT-qPCR results confirming the constitutive expression observed *in silico* ($P < 0.05$), even after the bulk has been opened in two times of salt exposition (30 and 90 min, Figures 3(c) and 3(d) and Table 4). The same PIP1-1 had been validated by RT-qPCR with cDNAs from the drought-tolerant and -sensitive sugarcane accessions (24 h of continuous dehydration) using GAPDH and 25S rRNA as the reference genes, in a previous work, showing differential expressions as expected by the SuperSAGE analysis [34].

## 4. Discussion

The understanding of sugarcane physiology under environmental stress remains under intensive research, due to the socioeconomic importance of this crop and the increasing unpredictability of environmental conditions worldwide. In this regard, gene expression analysis is an attractive approach to dissect plant physiological response to stress conditions. Nevertheless, reference gene selection has received limited attention in sugarcane. RT-qPCR is currently one of the most used techniques for gene expression analysis, due to its rapid, specific, and highly sensitive parameters. However, problems with RNA samples variations, standardization, and protocols efficiency (RNA extraction, RT, and qPCR) have routinely been observed [35]. Furthermore, the choice of normalizing genes remains one of the most time consuming and difficult steps in RT-qPCR. It requires reference genes to be constitutively expressed under external stimuli. Additionally, it needs to exhibit little or no behavior change in different cell types or tissues, as well as in specific developmental stages and experimental conditions [36–38]. Stal Papini-Terzi et al. [21] described transcriptional profile of signal transduction events in different sugarcane tissues, using reference genes selected based upon the literature (tubulin and actin), microarray data, and ESTs (polyubiquitin and 14-3-3 proteins). To individually normalize gene expression in sugarcane under certain conditions, Rocha et al. [22] relied upon four reference genes (14-3-3, polyubiquitin, GAPDH, and 25S rRNA). Moreover, Rodrigues et al. [23] used β-tubulin as the reference gene based on previous data [24]. To our knowledge, Iskandar et al. [24] represented an attempt to prospect stable sugarcane reference genes by checking the reliability of four genes (β-actin, β-tubulin, GAPDH, and 25S rRNA) in leaf, root, and internode tissues of some sugarcane cultivars and representatives of *Saccharum* genus, but none of them under abiotic stress. According to the authors, GAPDH was the most stable gene (CV = 51%) comparing different tissues, followed by β-actin and β-tubulin (CV = 81% and 68%, resp.); regarding species, β-actin showed the lowest coefficient of variation (31%) followed by GAPDH (33%). Although these methods are useful for prospecting candidate reference genes [21] or addressing gene expression using validated reference genes for target tissues [22, 23], selection using more appropriate statistical approaches should be the method of choice for identification of new reliable reference genes. In this way, software programs like GeNorm, NormFinder, and BestKeeper have assisted researchers by indicating reference genes suitable for expression profiling normalization studies [39]. In the present study, a group of potential reference genes (αTUB, GAPDH, H1, SAMDC, UBQ, and 25S rRNA) were evaluated by the three software programs, in order to evaluate their reliability for expression profiles normalization in sugarcane roots under abiotic stress (24 h of continuous dehydration). Basically, all three software programs pointed histone H1, αTUB, and GAPDH as the most reliable reference genes, with some of them switching places in the ranking. This set of genes was employed here as reference genes to validate sugarcane cDNAs relative to unitags from SuperSAGE libraries composed of roots of plants after stress exposition. Thus, the gene expression stability ranking provided by NormFinder and BestKeeper software showed the same order after the third place. Besides, based on the GeNorm analysis and *M*-values, all the six candidates may be suitable for normalizing gene expression data

Table 5: Potential reference gene combinations (and number of genes involved in each comparison) used in gene expression normalization of glutamine-dependent asparagine synthetase (AS), with cDNAs of sugarcane accessions (root under drought stress, 24 h of continuous dehydration).

| Comparison | Gene combinations | Number of genes | Expression* value | *P* value | Regulation |
|---|---|---|---|---|---|
| 1 | αTUB, H1, and SAMDC | 3 | 1.431 | 0.112 | ns |
| 2 | GAPDH, 25S rRNA, and SAMDC | 3 | 1.482 | 0.009 | ns |
| 3 | αTUB, H1, GAPDH, 25S rRNA, and SAMDC | 5 | 1.496 | 0.113 | ns |
| 4 | αTUB, H1, UBQ, and SAMDC | 4 | 1.557 | 0.105 | ns |
| 5 | UBQ and SAMDC | 2 | 1.595 | 0.305 | ns |
| 6 | αTUB and H1 | 2 | 1.519 | 0.000 | UR |
| 7 | GAPDH and 25S rRNA | 2 | 1.600 | 0.017 | UR |
| 8 | αTUB, H1, GAPDH, and 25S rRNA | 4 | 1.559 | 0.011 | UR |
| 9 | αTUB, H1, and UBQ | 3 | 1.667 | 0.033 | UR |
| 10 | GAPDH, 25S rRNA, and UBQ | 3 | 1.725 | 0.017 | UR |
| 11 | αTUB, H1, GAPDH, 25S rRNA, and UBQ | 5 | 1.640 | 0.017 | UR |

*REST software analysis after the ΔΔCq method. UR: upregulated; ns: not significant at $P < 0.05$; αTUB: alpha-tubulin; H1: histone H1; SAMDC: S-adenosylmethionine decarboxylase; GAPDH: glyceraldehyde 3 phosphate dehydrogenase; 25S rRNA: 25S ribosomal RNA; UBQ: ubiquitin.

as presented here. But, combining the two best candidates (αTUB and H1) with the most variable one (SAMDC, $M = 0.87$), as reference genes normalizing the target AS (induced in SuperSAGE analysis), the REST software did not detect the expected overexpression (comparison 1, Table 5) due to the SAMDC largest standard deviation (SD) influencing the *P* value calculated, consequently leading to a false negative and possible misinterpretation of data. Alias, any other combination including SAMDC as reference gene did not reveal the alleged overexpression (comparisons 2–5, Table 5), indicating that this gene is not suitable for gene expression normalization in roots of sugarcane accessions under the evaluated stress. However, Hong et al. [40] reported SAMDC as the most reliable reference gene in grass *Brachypodium distachyon* when evaluated under four abiotic stress conditions (high salt (300 mM), cold (4°C for 5 h), heat (42°C for 2 h), and drought (400 mM mannitol)). In addition, Li and Chen [41], when describing SAMDC as a target gene, verified that this gene was induced in roots of rice seedlings at three leaf stages (after application of 171 mM salt for 24 h and 20 mM exogenous abscisic acid (ABA) and dehydration using 15% PEG6000). These results highlight the need to choose appropriate reference genes for each experiment, especially under stress conditions.

On the other hand, the other five genes (αTUB, GAPDH, H1, UBQ, and 25S rRNA) could be successfully employed in the normalization analyses, composing different combinations of reference genes (comparisons 6–11, Table 5), with similar results to that observed for the proposed set (H1, αTUB, and GAPDH). In relation to the 25S rRNA gene, it was the most abundant transcript (Cq 13.06), in agreement with results previously obtained from rice (Cq values of 15 [42]) and sugarcane (Cq values of 16.6 [24]). This can be explained by the fact that rRNA comprises the majority of total RNA present in a cell and, thus, further dilution for its use in RT-qPCR approaches [43] would be required. In the present work, only a 1 : 5 dilution was applied. Furthermore, the abundance of transcripts can affect the stability and, therefore, the normalizing results for the reference gene candidates [44]. In turn, GAPDH was also one of the most stable genes, confirming it to be an appropriate reference gene for experiments involving sugarcane roots under water deficit conditions. Concerning αTUB, it was indicated by GeNorm as the most stable gene using the bulks of accessions, both under regular irrigation and after 24 h of continuous dehydration. By the NormFinder and BestKeeper analysis, this gene was the second most appropriate reference gene. However, Fan et al. [38] assessing the reliability of reference genes in 14 different tissues and developmental stages of *Phyllostachys edulis* observed that αTUB showed a larger variation ($M = 1.94$) among all candidates. Similar results were reported by Zhong et al. [45] with litchi (*Litchi chinensis* Sonn.) under several experimental conditions (tissues, organs, developmental stages, and varieties), showing αTUB as the most variable gene among 10 candidates. Under biotic and abiotic stresses αTUB also displayed instability, as demonstrated by Die et al. [39], Hong et al. [40], and Zhu et al. [46]. Thus, the selection of suitable reference genes to normalize gene expression in sugarcane and other plant species seems to be essential because reference genes may be differently regulated in different species, displaying particular gene expression patterns [43]. The proposed reference genes (H1, αTUB, and GAPDH) validated the gene expression of sugarcane cDNAs related to SuperSAGE unitags, showing upregulation or even constitutive basis, in the mentioned drought assay. Among those targets, AS is a crucial component of the asparagine synthesis, acting as a key member in nitrogen assimilation, recycling, and storage in higher plants [47]. The overexpression results observed in the present study supported those found in a previous microarray [48] showing AS induction in wheat accession considered tolerant to drought after 36 h of irrigation suppression. AS were also detected by RT-PCR, using mRNA samples from roots and shoots (two-week-old plantlets), significantly induced after salinity (250 mM), osmotic stress (using mannitol 5.0% (w/v)), and exogenous abscisic acid (ABA) application (20 mM) [49].

Altogether, these results indicate the involvement of the AS gene in response to several stresses. Considering PFPα1, the constitutive expression observed in the present assay has been reported previously and was in agreement with Lim et al. [50] who demonstrated by semiquantitative RT-PCR that the expression of PFPα1 in wild-type *Arabidopsis* was also constitutive in different tissues (roots, leaves, and flowers) and also in distinct developmental stages (15, 25, and 45 days after planting). PFPα1 is responsible for the addition of phosphate to the second D-fructose 6-phosphate in the glycolysis pathway [51] and is essential for maintenance of carbohydrate metabolism and other processes in plant cells [50]. In sugarcane, PFPs are known to play a prominent role in sucrose accumulation, especially in immature and metabolically active tissues, taking part in glycolysis and in carbon compartmentalization [52].

The effort to explore the use of the proposed reference genes (H1, αTUB, and GAPDH) normalizing the expressions of sugarcane cDNAs (associated with SuperSAGE unitags) also from roots of plants under salt stress (100 mM) was effective in a preliminary study. This way, the induction of ASS122537 unitag (annotated as ACC oxidase, enzyme responsible to convert the ethylene precursor ACC to ethylene, [53]), as revealed by *in silico* SuperSAGE analysis, was confirmed by RT-qPCR after 30 minutes of salt exposition (tolerant accession), suggesting that rapid ethylene production is an adaptive response to the new conditions imposed by the salt stress to the evaluated genotype. Unfortunately this overexpression was not detected during the 90 min of stress exposition. Nevertheless, there is evidence that a variety of stressful conditions trigger the synthesis of ethylene [54]. Regarding salt stress, it has been observed that this hormone signaling may be required for triggering the tolerance process. Yang et al. [55], when evaluating *Arabidopsis* mutants (ein2-5, ein3-1, and ctr1-1) and wild plants ecotype Col-0, found that mutants insensitive to ethylene (ein2-5 or ein3-1) were more sensitive to saline stress when compared to their wild counterpart. The opposite was found in the mutant sensitive to ethylene (ctr1-1), which showed significant tolerance to salt stress. Concerning the PIP1-1, the RT-qPCR results showed no significant differences in any of the two sampled times, confirming the SuperSAGE results with the bulk comprising both sampling times. The same target in RT-qPCR assay confirmed differential expressions expected by the SuperSAGE analysis, with root cDNAs from the drought-tolerant and -sensitive sugarcane accessions (24 h of continuous dehydration) and GAPDH and 25S rRNA as the reference genes [34]. An explanation could be that the time intervals used for stress exposition were not enough for PIP1-1 expression. In rice, Guo et al. [56] reported PIP1-1 expression in response to salt stress (250 mM NaCl) after 2 h of stress exposition.

## 5. Conclusions

The potential of the six proposed reference genes (αTUB, GAPDH, H1, SAMDC, UBQ, and 25S rRNA) was confirmed after they were tested with cDNAs from sugarcane roots under drought stress (24 h of continuous dehydration) and analyzed by three different software programs (GeNorm, NormFinder, and BestKeeper). With the exception of SAMDC, all the other candidate genes seem to be suitable for sugarcane expression profiling normalization, but three of them (αTUB, H1, and GAPDH) were considered as the best reference genes. In this study, two new reference genes were reported for the first time for sugarcane (αTUB and H1), to undergo a RT-qPCR validation study involving expression in roots under abiotic stresses. Also, the present work pointed GAPDH and 25S rRNA genes, both indicated by Iskandar et al. [24], as reference genes in a previous study, also suitable for use with sugarcane root under abiotic stress. Using the proposed set of reference genes (αTUB, H1, and GAPDH), it was confirmed that the relative expression profile, with the aid of the REST software, of cDNAs was associated with unitags (26 bp) and annotated as AS and PFPα1, using a bulk of cDNAs relative to the drought-tolerant sugarcane accessions (four accessions, 24 h of continuous dehydration), in agreement with the HT-SuperSAGE data. Another two unitags (associated with ACC oxidase and PIP1-1) had their expression profiles validated by RT-qPCR, using cDNAs from sugarcane roots after salt stress exposition (100 mM NaCl), in an attempt to explore other possibilities using these reference genes. In conclusion, this set of reference genes will be useful for dissecting gene expression in sugarcane roots, especially in advanced transcriptomic studies using SuperSAGE or RNAseq approaches covering abiotic stresses.

## Conflict of Interests

The authors have declared that they have no conflict of interests.

## Acknowledgments

The authors acknowledge the Brazilian institutions FINEP (Financiadora de Estudos e Projetos), FACEPE (Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco), and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for financial support and fellowships. They also thank Dr. Björn Rotter, Ruth Jungmann, Dr. Peter Winter (GenXPro GmbH, Germany), and Eng. Ernestino Ciambarella (Labtrade do Brasil) for scientific and technical advice.

## References

[1] G. E. Welbaum and F. C. Meinzer, "Compartmentation of solutes and water in developing sugarcane stalk tissue," *Plant Physiology*, vol. 93, no. 3, pp. 1147–1153, 1990.

[2] A. L. Rae, C. P. L. Grof, R. E. Casu, and G. D. Bonnett, "Sucrose accumulation in the sugarcane stem: pathways and control points for transport and compartmentation," *Field Crops Research*, vol. 92, no. 2-3, pp. 159–168, 2005.

[3] S. L. Dillon, F. M. Shapter, R. J. Henry, G. Cordeiro, L. Izquierdo, and L. S. Lee, "Domestication to crop improvement: genetic resources for Sorghum and Saccharum (Andropogoneae)," *Annals of Botany*, vol. 100, no. 5, pp. 975–989, 2007.

[4] FAOSTAT, "Food and Agriculture Organization of the United Nations," in *FAO Statistical Databases*, 2011, http://faostat.fao.org/.

[5] A. D'Hont, L. Grivet, P. Feldmann, S. Rao, N. Berding, and J. C. Glaszmann, "Characterization of the double genome structure of modern sugarcane cultivars (Saccharum spp.) by molecular cytogenetics," *Molecular and General Genetics*, vol. 250, no. 4, pp. 405–413, 1996.

[6] L. Grivet and P. Arruda, "Sugarcane genomics: depicting the complex genome of an important tropical crop," *Current Opinion in Plant Biology*, vol. 5, no. 2, pp. 122–127, 2002.

[7] N. Jannoo, L. Grivet, N. Chantret et al., "Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome," *The Plant Journal*, vol. 50, no. 4, pp. 574–585, 2007.

[8] J. M. Manners and R. E. Casu, "Transcriptome Analysis and Functional Genomics of Sugarcane," *Tropical Plant Biology*, vol. 4, no. 1, pp. 9–21, 2011.

[9] C. Gachon, A. Mingam, and B. Charrier, "Real-time PCR: what relevance to plant studies?" *Journal of Experimental Botany*, vol. 55, no. 402, pp. 1445–1454, 2004.

[10] S. Guénin, M. Mauriat, J. Pelloux, O. van Wuytswinkel, C. Bellini, and L. Gutierrez, "Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references," *Journal of Experimental Botany*, vol. 60, no. 2, pp. 487–493, 2009.

[11] F. Rizzi, L. Belloni, P. Crafa et al., "A novel gene signature for molecular diagnosis of human prostate cancer by RT-qPCR," *PLoS ONE*, vol. 3, no. 10, Article ID e3617, 2008.

[12] G. Gao, R. D. Chernock, H. A. Gay et al., "A novel RT-PCR method for quantification of human papillomavirus transcripts in archived tissues and its application in oropharyngeal cancer prognosis," *International Journal of Cancer*, vol. 132, no. 4, pp. 882–900, 2013.

[13] S. Y. C. Yong, C. Y. Choong, P. L. Cheong et al., "Analysis of ESTs generated from inner bark tissue of an Acacia auriculiformis x Acacia mangium hybrid," *Tree Genetics and Genomes*, vol. 7, no. 1, pp. 143–152, 2011.

[14] A. Golisz, M. Sugano, S. Hiradate, and Y. Fujii, "Microarray analysis of Arabidopsis plants in response to allelochemical L-DOPA," *Planta*, vol. 233, no. 2, pp. 231–240, 2011.

[15] C. Molina, M. Zaman-Allah, F. Khan et al., "The salt-responsive transcriptome of chickpea roots and nodules via deepSuperSAGE," *BMC Plant Biology*, vol. 11, article 31, 2011.

[16] P. M. Bleeker, E. A. Spyropoulou, P. J. Diergaarde et al., "RNA-seq discovery, functional characterization, and comparison of sesquiterpene synthases from Solanum lycopersicum and Solanum habrochaites trichomes," *Plant Molecular Biology*, vol. 77, no. 4-5, pp. 323–336, 2011.

[17] J. Vandesompele, K. de Preter, F. Pattyn et al., "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes," *Genome Biology*, vol. 3, no. 7, 2002.

[18] L. Gutierrez, M. Mauriat, S. Guénin et al., "The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants," *Plant Biotechnology Journal*, vol. 6, no. 6, pp. 609–618, 2008.

[19] C. L. Andersen, J. L. Jensen, and T. F. Ørntoft, "Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets," *Cancer Research*, vol. 64, no. 15, pp. 5245–5250, 2004.

[20] M. W. Pfaffl, A. Tichopad, C. Prgomet, and T. P. Neuvians, "Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: bestKeeper—excel-based tool using pair-wise correlations," *Biotechnology Letters*, vol. 26, no. 6, pp. 509–515, 2004.

[21] F. Stal Papini-Terzi, F. Riso Rocha, R. Zorzetto Nicoliello Vêncio et al., "Transcription profiling of signal transduction-related genes in sugarcane tissues," *DNA Research*, vol. 12, no. 1, pp. 27–38, 2005.

[22] F. R. Rocha, F. S. Papini-Terzi, M. Y. Nishiyama Jr. et al., "Signal transduction-related responses to phytohormones and environmental challenges in sugarcane," *BMC Genomics*, vol. 8, article 71, 2007.

[23] F. A. Rodrigues, M. L. de Laia, and S. M. Zingaretti, "Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants," *Plant Science*, vol. 176, no. 2, pp. 286–302, 2009.

[24] H. M. Iskandar, R. S. Simpson, R. E. Casu, G. D. Bonnett, D. J. Maclean, and J. M. Manners, "Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane," *Plant Molecular Biology Reporter*, vol. 22, no. 4, pp. 325–337, 2004.

[25] E. A. Kido, J. R. C. Ferreira Neto, R. L. O. Silva et al., "New insights in the sugarcane transcriptome responding to drought stress as revealed by Super SAGE," *The Scientific World Journal*, vol. 2012, Article ID 821062, 14 pages, 2012.

[26] S. A. Bustin, V. Benes, J. A. Garson et al., "The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments," *Clinical Chemistry*, vol. 55, no. 4, pp. 611–622, 2009.

[27] R. Rasmussen, "Quantification on the light cycler," in *Rapid Cycle Real-Time PCR, Methods and Applications*, S. Meuer, C. Wittwer, and K. Nakagawara, Eds., pp. 21–34, Springer, Heidelberg, Germany, 2001.

[28] F. Thiebaut, C. Grativol, M. Carnavale-Bottino et al., "Computational identification and analysis of novel sugarcane microRNAs," *BMC Genomics*, vol. 13, no. 2, p. 290, 2012.

[29] H. Matsumura, D. H. Krüger, G. Kahl, and R. Terauchi, "SuperSAGE: a modern platform for genome-wide quantitative transcript profiling," *Current Pharmaceutical Biotechnology*, vol. 9, no. 5, pp. 368–374, 2008.

[30] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," *Methods in Molecular Biology*, vol. 132, pp. 365–386, 2000.

[31] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.

[32] M. W. Pfaffl, G. W. Horgan, and L. Dempfle, "Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR," *Nucleic acids research*, vol. 30, no. 9, p. e36, 2002.

[33] Apllied Biosystems, "Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR," 2004.

[34] M. D. Silva, R. L. O. Silva, J. R. C. Ferreira Neto et al., "Expression analysis of sugarcane aquaporin genes under water deficit," *Journal of Nucleic Acids*, vol. 2013, Article ID 763945, 14 pages, 2013.

[35] T. Nolan, R. E. Hands, and S. A. Bustin, "Quantification of mRNA using real-time RT-PCR," *Nature Protocols*, vol. 1, no. 3, pp. 1559–1582, 2006.

[36] K. Dheda, J. F. Huggett, S. A. Bustin, M. A. Johnson, G. Rook, and A. Zumla, "Validation of housekeeping genes for normalizing RNA expression in real-time PCR," *BioTechniques*, vol. 37, no. 1, pp. 112–119, 2004.

[37] M. Banda, A. Bommineni, R. A. Thomas, L. S. Luckinbill, and J. D. Tucker, "Evaluation and validation of housekeeping genes in response to ionizing radiation and chemical exposure for normalizing RNA expression in real-time PCR," *Mutation Research—Genetic Toxicology and Environmental Mutagenesis*, vol. 649, no. 1-2, pp. 126–134, 2008.

[38] C. Fan, J. Ma, Q. Guo, X. Li, H. Wang, and M. Lu, "Selection of reference genes for quantitative real-time PCR in bamboo (*Phyllostachys edulis*)," *PLoS ONE*, vol. 8, no. 2, Article ID e56573, 2013.

[39] J. V. Die, B. Román, S. Nadal, and C. I. González-Verdejo, "Evaluation of candidate reference genes for expression studies in Pisum sativum under different experimental conditions," *Planta*, vol. 232, no. 1, pp. 145–153, 2010.

[40] S.-Y. Hong, P. J. Seo, M.-S. Yang, F. Xiang, and C.-M. Park, "Exploring valid reference genes for gene expression studies in Brachypodium distachyon by real-time PCR," *BMC Plant Biology*, vol. 8, article 112, 2008.

[41] Z.-Y. Li and S.-Y. Chen, "Differential accumulation of the S-adenosylmethionine decarboxylase transcript in rice seedlings in response to salt and drought stresses," *Theoretical and Applied Genetics*, vol. 100, no. 5, pp. 782–788, 2000.

[42] Q.-F. Li, S. S. M. Sun, D.-Y. Yuan, H.-X. Yu, M.-H. Gu, and Q.-Q. Liu, "Validation of candidate reference genes for the accurate normalization of real-time quantitative RT-PCR data in rice during seed development," *Plant Molecular Biology Reporter*, vol. 28, no. 1, pp. 49–57, 2009.

[43] M. Jain, A. Nijhawan, A. K. Tyagi, and J. P. Khurana, "Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR," *Biochemical and Biophysical Research Communications*, vol. 345, no. 2, pp. 646–651, 2006.

[44] R. Hu, C. Fan, H. Li, Q. Zhang, and Y.-F. Fu, "Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR," *BMC Molecular Biology*, vol. 10, article 1471, p. 93, 2009.

[45] H.-Y. Zhong, J.-W. Chen, C.-Q. Li et al., "Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions," *Plant Cell Reports*, vol. 30, no. 4, pp. 641–653, 2011.

[46] X. Zhu, X. Li, W. Chen et al., "Evaluation of new reference genes in papaya for accurate transcript normalization under different experimental conditions," *PLoS ONE*, vol. 7, no. 8, Article ID e44405, 2012.

[47] L. Gaufichon, M. Reisdorf-Cren, S. J. Rothstein, F. Chardon, and A. Suzuki, "Biological functions of asparagine synthetase in plants," *Plant Science*, vol. 179, no. 3, pp. 141–153, 2010.

[48] M. Mohammadi, N. N. V. Kav, and M. K. Deyholos, "Transcriptional profiling of hexaploid wheat (*Triticum aestivum* L.) roots identifies novel, dehydration-responsive genes," *Plant, Cell and Environment*, vol. 30, no. 5, pp. 630–645, 2007.

[49] H. Wang, D. Liu, J. Sun, and A. Zhang, "Asparagine synthetase gene TaASN1 from wheat is up-regulated by salt stress, osmotic stress and ABA," *Journal of Plant Physiology*, vol. 162, no. 1, pp. 81–89, 2005.

[50] H. Lim, M.-H. Cho, J.-S. Jeon, S. H. Bhoo, Y.-K. Kwon, and T.-R. Hahn, "Altered expression of pyrophosphate: fructose-6-phosphate 1-phosphotransferase affects the growth of transgenic Arabidopsis plants," *Molecules and Cells*, vol. 27, no. 6, pp. 641–649, 2009.

[51] R. E. Reeves, D. J. South, H. J. Blytt, and L. G. Warren, "Pyrophosphate: D-fructose 6 phosphate 1 phosphotransferase. A new enzyme with the glycolytic function of 6 phosphofructokinase," *Journal of Biological Chemistry*, vol. 249, no. 24, pp. 7737–7741, 1974.

[52] J.-H. Groenewald and F. C. Botha, "Down-regulation of pyrophosphate: fructose 6-phosphate 1-phosphotransferase (PFP) activity in sugarcane enhances sucrose accumulation in immature internodes," *Transgenic Research*, vol. 17, no. 1, pp. 85–92, 2008.

[53] H. Kende, "Ethylene biosynthesis," *Annual Review of Plant Physiology and Plant Molecular Biology*, vol. 44, no. 1, pp. 283–307, 1993.

[54] U. Druege, "Ethylene and plant responses to abiotic stress," in *Ethylene Action in Plants*, N. A. Kahn, Ed., pp. 81–118, Springer, New York, NY, USA, 2006.

[55] L. Yang, Y.-G. Zu, and Z.-H. Tang, "Ethylene improves *Arabidopsis* salt tolerance mainly via retaining $K^+$ in shoots and roots rather than decreasing tissue $Na^+$ content," *Environmental and Experimental Botany*, vol. 86, pp. 60–69, 2013.

[56] L. Guo, Y. W. Zi, H. Lin et al., "Expression and functional analysis of the rice plasma-membrane intrinsic protein gene family," *Cell Research*, vol. 16, no. 3, pp. 277–286, 2006.

# 5 ARTIGO II

## Unraveling Drought-tolerant and -sensitive Sugarcane Cultivars by HT-SuperSAGE Analysis

Artigo aceito para publicação no periódico *GENE* (ISSN: 0378-1119)

# Unraveling Drought-tolerant and -sensitive Sugarcane Cultivars by HT-SuperSAGE Analysis

Manassés Daniel da Silva[1], Roberta Lane de Oliveira Silva[1], José Ribamar Costa Ferreira Neto[1], Ana Maria Benko-Iseppon[1] and Ederson Akio Kido[1*]

[1]Federal University of Pernambuco (UFPE), Bioscience Center, Department of Genetics, 50670-420 Recife, PE, Brazil

[*]Corresponding author

Email addresses:
MDS: manassesdaniel@gmail.com
RLOS: lane.roberta@gmail.com
JRCFN: netocostaferreira@gmail.com
AMBI: ana.iseppon@gmail.com
EAK: kido.ufpe@gmail.com

**Abstract:** Drought is the most damaging among the major abiotic stresses. The complexity involving both the quantitative feature of drought tolerance and the sugarcane genome makes breeding difficult. Transcriptomic studies allow a global overview of expressed genes, therefore providing the basis for the identification of functional molecular markers useful in MAS (marker-assisted selection). In this work, we report the analysis of HT-SuperSAGE tags (26 bp) generated by two contrastant sugarcane bulks. Each one of them formed by four cultivars considered either tolerant or sensitive to drought. We identified induced unitags with different expressions in the tolerant and sensitive bulk of cultivars after 24 h of irrigation suppression. The ESTs presenting these unitags allowed a proposal of 15 molecular candidates, which expressed profiles validated by RT-qPCR considering each cultivar separately. The

proposed targets were grouped into seven broad metabolic processes: ethylene stress attenuation; root growth; protein degradation; oxidative detoxification; fatty acid synthesis; amino acid transport; and carbohydrate metabolism. Two drought-tolerant cultivars (SP83-2847 and CTC6) presented, each of them, nine induced targets, sharing six of them. Among the sensitive cultivars, CTC13 induced one (repressed other two), while SP90-1636 induced two and repressed one. These candidates acting as useful markers may be promising in MAS.

**Keywords:** bioinformatic, water deficit, abiotic stress, RT-qPCR, *Saccharum* spp.

## 1. Introduction

Sugarcane (*Saccharum* spp. hybrid.) is a significant grass, with sugar and ethanol as its two major products [1]. It is grown in more than 100 countries occupying approximately 26 million hectares. With an estimate of 654.6 million tons for the 2015/2016 season, Brazil is emerging as the world's largest sugarcane producer, as well as the major manufacturer of sugar and ethanol [2,3]. However, biotic and abiotic stresses negatively influence crops production [4]. Among the abiotic stresses, drought is the most damaging and often affects agriculture areas [5,6]. Water deficit causes a complex set of molecular responses in plants, which begins with stress perception, proceeds with a signal transduction cascade, and culminates with physiological and morphological changes at cellular levels. These responses include the closure of the stomata, the suppression of cell growth and photosynthesis, the activation of cellular respiration, among others [7]. Plants under drought can also synthesize and accumulate some metabolites and proteins distinctively involved in stress tolerance [8]. Besides the complexity of plant responses to drought, the present sugarcane hybrids have one of the most complex genomes of the plant kingdom. They have high polyploidy [9] and frequent aneuploidy events [10], with karyotypes ranging from 100 to 130

chromosomes [11], in addition to a size 2C estimated at 10,000 Mb [9]. These hybrids resulted from the *Saccharum* spp. complex originated from interspecific crosses between *Saccharum officinarum, S. barbieri, S. sinense* and the wild species of *S. spontaneum* and *S. robustum* [9,10]. To sum up, the complexity of the drought tolerance character and the sugarcane genome make it difficult to obtain tolerant plants by conventional plant breeding programs. Therefore the improvement of sugarcane for drought conditions is still a major challenge. This challenge has encouraged the inclusion of new techniques in sugarcane breeding programs. Despite the new platforms for high-performance sequencing [12,13], and the fact that the sugarcane genome is not yet fully resolved [14], functional genomics studies have been carried out. They have been conducted with transcripts generated by oligoarrays [15,16], deep sequencing of small RNAs [17,18], SuperSAGE [19], RNA-seq [20] and microarray technology [21]. Such effort has provided some gene regulation data, useful for identifying suitable functional molecular markers. To this end, the present study exploited the available data from high-throughput (HT-) SuperSAGE libraries of two bulks (of four cultivars each) contrasting in sugarcane responses to the irrigation suppression for 24 h [19]. The SuperSAGE technique with 26-bp tags sequencing has some advantages regarding simplicity of data analysis and cost [22,23], when compared to the microarray methodologies widely applied in sugarcane studies (see [15,16,21]). Researchers have used SuperSAGE to study plant responses to biotic and abiotic stresses, identifying differentially expressed genes in plants such as rice [22], banana [24], chickpea [25,26], cowpea [27], ragweed [28], Mexican lime [29], and also sugarcane [19]. The aim of this work was the analysis of sugarcane transcriptomes in response to irrigation suppression for 24 h, to select molecular targets involved in different drought tolerance strategies.

## 2. Methods

### 2.1. Sugarcane cultivars, HT-SuperSAGE libraries and unitags annotation

Eight sugarcane cultivars from the Sugarcane Research Center (Centro de Tecnologia Canavieira – CTC, Piracicaba, São Paulo, Brazil) formed two distinct pools regarding the applied stress (24 h after irrigation suppression). The tolerant bulk consisted of cultivars CTC6, CTC15, SP83-2847 and SP83-5073, while the sensitive bulk comprised cultivars CTC9, CTC13, SP90-3414 and SP90-1638. These cultivars have been reported as tolerant and sensitive to drought, based on field observations by sugarcane producers, and some physiological analyses, including their relative content of water and chlorophyll [30,31]. Stress was applied to seedlings grown under greenhouse conditions (in 40 L pots, in a completely randomized experimental design with six replicates), under daily irrigation (4 $L.h^{-1}$), during three months. The cultivation conditions indicated temperatures of 30.2 ± 5.7 ºC (maximum), 16.8 ± 1.9 ºC (average) e 9.3 ± 3.0 ºC (minimum), relative humidity of 71.5 ± 5.1% (average) and natural photoperiod. Thereafter, plants were stressed by irrigation suppression, for 24 h, with their roots under stress and those from plants without stress (negative control) collected, immediately frozen in liquid nitrogen and stored at -80 °C until total RNA extraction. Total RNA were isolated by Trizol® (Invitrogen), treated with DNAse (Qiagen®), purified with the RNeasy Mini kit (Qiagen®) and quantified in a spectrophotometer. The bulks of each treatment [stressed and negative control (not stressed)], comprised purified RNAs of each cultivar of the respective bulk (tolerant or sensitive), in an equimolar quantity. The HT-SuperSAGE libraries [tolerant stressed (SD24T), tolerant negative control (SDTC), sensitive stressed (SD24S) and sensitive negative control (SDSC)] followed Matsumura et al. [22] with modifications, including the sequencing by Illumina Genome Analyzer II (SOLEXA technology) [19]. To summarize, the available data comprised 205,975 unitags (unique tags) and their observed frequencies

in the libraries. These frequencies observations allowed unitags to be classified as up-expressed (induced), down-expressed (repressed) or non-significant ($p<0.05$), after Audic-Claverie test [32]. Also for each single unitag, the modulation of its gene expression (FC: Fold Change) was calculated as the frequency observed ratio in one library compared to the other. Unitags induced in the tolerant contrast (SD24T *vs*. SDTC; stressed *versus* respective negative control) and repressed or not significant in the sensitive contrast (SD24S *vs*. SDSC) were aligned (BLASTn) against Expressed Sequence Tags (ESTs) from dbEST (http://www.ncbi.nlm.nih.gov/nucest) and Gene Index (SoGI; ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/). Accepted alignments comprised those with best hit with sugarcane sequences showing score 42 - 52, plus/plus orientation, and with 5'CATG preserved. Also, the ESTs anchoring unitags were aligned against CDSs (BLASTn) from the Phytozome database (http://www.phytozome.org/), to confirm the identities (e-value cutoff $e^{-22}$). Furthermore, these ESTs were categorized by the Blast2GO tool (https://www.blast2go.com).

### 2.2. RNA extraction and cDNA synthesis

Total RNAs from tolerant and sensitive cultivars were extracted and purified as mentioned before. The RNAs integrities were verified on 1.5% (w / v) agarose gel stained with Blue-Green Loading Dye (LGC Biotechnology®). Samples were quantitated using the Quant-iT ™ RNA Assay Kit (Invitrogen®) in Qubit fluorometer (Invitrogen®). Total RNAs (1 μg) were converted into cDNAs using the QuantiTect Reverse Transcription Kit (Qiagen®) and resuspended in a final volume of 20 μL. The cDNA synthesis was incubated at 42 °C for 2 min in gDNA Wipeout Buffer (for removal of contaminating genomic DNA), 42 °C for 15 min (reverse transcription) and 95 °C for 3 min (inactivation of the enzyme), in a thermocycler. Samples were stored at

-20 ° C until the RT-qPCR reactions. RT-qPCR assays followed the guidelines of the MIQE protocol (The Minimum Information for Publication of Quantitative Real-Time PCR Experiments) [33], aiming to increase reliability in the results.

### 2.3. Primer design, amplification efficiency, and RT-qPCR analyses

RT-qPCR validated the gene expression of 15 molecular candidates (Table S1). Meticulous care was taken to confirm the identities of the ESTs used in the primer designs. Initially, they were BLASTn aligned against RefSeq mRNA (NCBI; http://www.ncbi.nlm.nih.gov/) and sorghum CDSs from Phytozome database (http://www.phytozome.org/). Primers were designed based on sugarcane ESTs showing the selected unitags using Primer 3 software [34]. The following parameters were applied: amplicon size (between 70 and 200 bp), melting temperature [between 40ºC (minimum), 50ºC (optimum) and 60ºC ), And GC content [between 45-55% (50% optimum)]. Primers from the scientific literature were those related to 25S rRNA and GAPDH [35]. The designed primers (Table S1) were used to amplify the proposed candidates (AAT, $Na^+$ dependent neutral amino acid transporter; ACC, acetyl-CoA carboxylase; ACCD, 1-aminocyclopropane-1-carboxylate deaminase; AP, aspartic proteinase; ß-EXP, beta-expansin 8 precursor; PFK, 6-phosphofructo-2-kinase; TRX, thioredoxin-like 1-2; E2, ubiquitin conjugating enzyme; 20SPß4, 20S proteasome beta 4 subunit; TPI, triosephosphate isomerase; FBA, fructose-bisphosphate aldolase; TKT, transketolase; C13, legumain cysteine-type endopeptidase; LDP, dihydrolipoyl dehydrogenase e MDH, malate dehydrogenase). All primers were synthesized by Bioneer Corporation (South Korea) and tested in cDNAs amplifications using conventional PCR, under the conditions of 5 min at 95 ºC, followed by 35 cycles of 95 °C for 15 s, 58 °C for 30 s, and 72 °C for 30 s. PCR products were analyzed on 1.5% (w / v) agarose gel stained with Blue-Green Loading Dye (LGC Biotechnology®). The RT-

qPCR promoted dissociation curves to confirm the amplification specificity of the proposed candidates. Also, calibration curves were made using cDNAs samples from the tolerant and sensitive cultivars bulks and serial dilutions (1, $10^{-1}$, $10^{-2}$, $10^{-3}$ e $10^{-4}$). These curves had the purpose of calculating the amplification efficiency (E = $10^{(-1/\text{standard curve slope})}$; [36]) of each candidate (Fig. S.1, Fig. S.2). RT-qPCR reactions were carried out in a final volume of 10 μL, including: 1 μL of the cDNA sample (diluted five times), 5 μL of the HotStart-IT SYBR™ Green qPCR Master Mix 2x (USB®), 0.05 μL of ROX, 1.95 μL of ultrapure water, and 1 μL of each primer (0.05 μM). In all reactions of RT-qPCR (in 96-well plates), performed using three biological and three technique replicates, a negative control (cDNA sample), and two reference genes (25S rRNA and GAPDH) were used. RT-qPCR reactions were performed on LineGene 9660, model FQD-96A (Bioer®), using SYBR Green detection system. These reactions used the following programming: initial denaturation of 95 ° C for 2 min, followed by 40 cycles of 95 ° C for 15 s, 58 ° C for 30 s, and 72 ° C for 30 s. The dissociation curves were analyzed at 65 to 95 ° C for 20 min after 40 cycles. The LineGene 9600 software (version 1.1.10) determined the quantification cycles (Cq), melting temperatures, and absolute and relative quantifications. Relative expression data were calculated by the ΔΔCq method and evaluated using the REST software (Relative Expression Software Toll, version 2.0.13) [37], using randomization test with 2000 permutations.

### 2.4. Unitags mapping in metabolic KEGG maps

Enzymes of the glycolysis pathway (map00010), TCA cycle (map00020) and Pentose phosphate pathway (map00030) according to the KEGG database (www.kegg.jp) and their EC numbers (Enzyme Commission Numbers) were associated with the annotated ESTs anchoring specific unitags. Furthermore, the pathways were mapped using the

Kegg Mapper tool (http://www.genome.jp/kegg/tool/map_pathway2.html; [38]), following the colors: red for induced unitags, green for repressed and yellow for non-significant, based on their expressed regulation.

## 3. Results and discussion

### 3.1. The induced unitags presented by the bulk of tolerant cultivars

From the universe of 205,975 HT-SuperSAGE unitags, 9,831 (Table S.2) were induced in the tolerant cultivars, concerning the stressed treatments *versus* the negative control [SD24T *vs*. SDTC], but showing different expression by the sensitive bulk. This set is potentially relevant, as the unitags observed in the bulk of tolerant cultivars presented induced regulation, while in the bulk of sensitive cultivars, these unitags were absent, repressed or n.s (non-significant) after Audic-Claverie tests ($p<0.05$). Considering these unitags, 1,533 were exclusive in those tolerant cultivars, and from the remainder presented in the sensitive contrast [SD24S *vs*. SDSC], 1,772 were suppressed.

Concerning the BLASTn alignments with those tolerant induced unitags, the acceptable alignments (CATG preserved at the 5' end), involved 5,077 unitags and 30,252 ESTs (*S. officinarum* or *S.* spp.). Concerning these 5,077 unitags identified in ESTs, 1,808 anchored in ESTs with informative annotations (gene or its function) and GO terms available, 1,565 in ESTs with some useful annotations and 614 in ESTs only characterized by GO terms . Unitags presented in ESTs with no annotations or without GO terms counted 1,090 (data not shown).

The 50 unitags that presented greatest changes in their expressions after stress stimulus based on the FC values and the informative annotations (or GO terms) of the respective *Saccharum* ESTs anchoring these tolerant induced unitags are shown in Table 1. This set included some detoxification enzymes, such as thylakoid ascorbate peroxidase [tAPX, EC 1.11.1.11, unitag SD8928 ($FC_{Tol} = 183.99$)], cytoplasmic ascorbate

peroxidase [CAPX, EC 1.11.1.11, unitag SD182788 ($FC_{Tol}$ = 29.41)] and quinone reductase [QR, EC 1.6.5.2, unitag SD108897 ($FC_{Tol}$ = 25.83, $FC_{Sens}$ = - 2.66)]. Even in normal conditions, plants produce reactive oxygen species (ROS) in several metabolic processes that include electron transport. The influence of ROS in the expression of genes involved in growth, cell cycle/ programmed cell death, defense against pathogens, systemic signaling, and responses to abiotic stresses was reported [39-41]. However, some responses to plant abiotic stresses lead to ROS overproduction and once these elements are highly reactive and toxic, they cause damage to proteins, lipids, carbohydrates and DNAs, leading to oxidative stress [42,43]. Plant strategy against oxidative stress is a process known as ROS scavenging. The main enzymes of the ROS scavenging machinery are superoxide dismutase (SOD, EC 1.15.1.1), catalase (CAT, EC 1.11.1.6), ascorbate peroxidase, (APX, EC 1.11.1.11), glutathione peroxidase (GPX, EC 1.11.1.9), glutathione reductase (GR, EC 1.6.4.2), monodehydroascorbate reductase (MDHAR, EC 1.6.5.4), dehydroascorbate reductase (DHAR, EC 1.8.5.1), guaiacol peroxidase (GOPX, EC 1.11.1.7) and glutathione-S-transferase (GST, EC 2.5.1.18) [43,44]. Seven of these nine key enzymes were detected: APX [one of the 50 most induced unitags (Table 1) and observed as isoforms from distinct compartments (tAPX, in the thylakoids and cAPX in the cytoplasm)]; SOD (unitag SD225439, $FC_{Tol}$ = 3.21, $FC_{Sens}$ = -3.66); CAT (unitag SD246162, $FC_{Tol}$ = 3.18); MDHAR (unitag SD108442, $FC_{Tol}$= 3.39); DHAR (unitag SD160961, $FC_{Tol}$ = 2.38, $FC_{Sens}$ = -11.18); GPX [unitag SD109179, $FC_{Tol}$= 4.77); and GST (unitag SD114493, $FC_{Tol}$ = 9.93). In general, the tolerant bulk showed a strong induction of many of these enzymes, when compared to the sensitive bulk. This information suggests a more efficient ROS scavenging machinery at tolerant cultivars, even after 24 h of irrigation suppression, making them good candidates for functional molecular markers in breeding programs.

Moreover, among the 50 most induced unitags (Table 1) were two chaperones: 17.5 kDa heat shock protein of class II (unitag SD113531, $FC_{Tol}$ = 39.34, $FC_{Sens}$ = -13.07) and heat shock protein 81-1 (unitag SD89732, $FC_{Tol}$ = 17.88, $FC_{Sens}$ = -2.28). Chaperones are proteins responsible for the folding, grouping, transporting, as well as for the degradation of several proteins under normal cellular conditions and they assist in the refolding under stress conditions. They participate in the defense line restoring the normal conformation of denatured stressed proteins, being crucial in plant homeostasis. Many of these expressed chaperones stressed proteins correlated with resistance to stress were initially identified as heat shock proteins (Hsps) [45].

Two enzymes of the protein degradation pathway were observed (Table 1): ubiquitin conjugating enzyme E2 (unitag SD214183, $FC_{Tol}$ = 31.39, validated by RT-qPCR) and polyubiquitin (unitag SD15211, $FC_{Tol}$ = 21.46). Ubiquitins are enzymes responsible for the regulation of different proteins through ubiquitination, which is the covalent attachment of an ubiquitin on the substrate protein to be recognized by a proteasome where it is degraded [46,47]. Ubiquitins play a critical role in regulating plant responses to abiotic stresses by facilitating responses to environmental stimuli by modulating the abundance of the major downstream transcriptional factors responsive to stress [48,49]. The overexpression of a single ubiquitin increases multiple stress tolerance with no negative effect on plant growth under favorable development conditions [48,49]. Transgenic tobacco plants overexpressing a wheat ubiquitin gene were more tolerant to saline stress (150 mM NaCl), and dry (200 mM mannitol for 45 days) than the wild tobacco plants [49]. In the same way, accentuation in tolerance to water deficit (20% polyethylene glycol-PEG for 15 days) and saline stress (200 mM NaCl for seven days) was observed in transformed plants of *Arabidopsis* sp overexpressing the ubiquitin-conjugating enzyme gene from soybean [48].

Additionally, two transcription factors (TFs) were present in the set of 50 most tolerant induced unitags (Table 1): the leucine zipper homeodomain 16 (HD-Zip 16, unitag SD78974, $FC_{Tol}$ = 28.61) and the low temperature induced 15 kDa maize protein (MLIP15, unitag SD231907, $FC_{Tol}$ = 18.68). TFs are signal transmitting proteins that regulate growth and development as well as plant adaptation to stresses by temporal and spatial regulation of target genes [50,51]. The protein homeodomain leucine zipper (HD-Zip) is found only in plants and encode more than 47 isoforms divided into four subfamilies HD-Zip I to IV [52,53]. Henriksson et al. [52] analyzed the expression of 17 isoforms of HD-Zip I and verified the induction of six of them against the application of ABA treatment and seven against saline stress.

Other interesting annotated sugarcane ESTs, which expressly modulated their expressions, based on the respective tolerant induced unitags, are available in Table 1.

### 3.2. Gene expression validation of induced unitags by RT-qPCR

The gene expression of 15 molecular candidates (Table 2) were independently validated by RT-qPCR for each cultivar (tolerant or sensitive). All of the designed primers (Table S1) for the proposed candidates (AAT, ACC, ACCD, AP, ß-EXP, PFK, TRX, MDH, LDP, E2, 20SPß4, C13, TPI, FBA, TKT) and for the applied reference genes (GAPDH and 25S rRNA) amplified cDNAs generated from the RNAs samples. Also, the dissociation curves for each primer pair showed specificity, only amplifying one product (Fig. S1). Similarly, there was no amplification in the templates controls (NTCs) (data not shown). Efficiencies of amplification (E), correlation coefficients (R) and slopes, for each candidate (Fig. S.2), stimated from standard curve using serial dilution ($10^{0}$, $10^{-1}$, $10^{-2}$, $10^{-3}$ e $10^{-4}$) and a pool of root cDNAs (tolerant or sensitive) from plants under or without stress (negative control), ranged, respectively, from 92.89 to 100.89%, 0.98 to 1.00 and -3.30 to -3.50 (Table S1). These results presented

favorable amplification conditions and acceptable variations to generate reliable data of relative expression among samples. According to the MIQE protocol [33], acceptable values for efficiency range from 90% to 110% and -3.10 to -3.58 for the slope, reaching 100% slope efficiency when in -3.32.

The proposed candidates embodied diferent metabolic processes: Ethylene stress attenuation; root growth; protein degradation (Ubiquitination pathway and plant proteases); oxidative detoxification; fatty acid synthesis; amino acid transport; carbohydrate metabolism (Glycolysis, TCA cycle, Pentose phosphate pathway). In general, the RT-qPCR results validated the *in silico* data of the selected induced unitags, showing differences among the cultivars, even among the tolerant or the sensitive ones, as shown below (Table 2 and Fig. 1).

### 3.2.1. Ethylene stress attenuation / ACCD

Ethylene is an important phytoregulator, acting in several physiological processes, such as root initiation, fruit ripening, seed germination, foliar abscission, gravitational responses, senescence, phytogulator biosynthesis, and stress signaling [54-56]. Plants under biotic or abiotic stresses present ethylene in higher concentrations, favoring an inhibition of plant growth by a physiological condition known as "ethylene stress" [57-59]. ACCD (1-aminocyclopropane-1-carboxylate deaminase; EC 3.5.99.7) catalyzes the conversion of ACC (1-aminocyclopropane-1-carboxylic acid), the immediate precursor of ethylene biosynthesis, into ammonia and α-ketobutyrate [54]. The reduction of ACC levels through ACC deaminase leads to the decrease of the excessive ethylene levels in plants, thus limiting stress damages [59-66]. *In silico* data for the unitag SD57621, a putative ACCD, showed bulk results of $FC_{Tol}$ = 1.90. The RT-qPCR results for the corresponding ACCD, showed significant induction (UR) of two tolerant cultivars (CTC6 and SP83-2847, respectively 1.93 and 1.73 fold; Table 2; Fig. 1). In turn, all

sensitive cultivars did not show significant differences in expression after the applied stress (Table 2, Fig. 1). Therefore, at least in these two tolerant cultivars, the proposed ACCD could be acting inhibiting ethylene stress and be limiting its damages, thus, favoring these sugarcane cultivars to grow and better tolerate stress. Four induced unitags related to plant growth genes were observed which supports this perspective. For cell division cycle #48, the unitag SD70367 presented $FC_{Tol}$ = 2.05 and $FC_{Sens}$ = -1.95, and for the eukaryotic translation initiation factor 3/subunit 7, three unitags considered n.s. by the sensitive bulk showed positive tolerant FC: SD103709 ($FC_{Tol}$ = 3.39); SD214026 ($FC_{Tol}$ = 1.94), and SD103711 ($FC_{Tol}$ = 1.93).

### 3.2.2. Root growth / ß-expansin 8

Expansins (α- and β-expansins), previously defined as group-1 grass pollen allergen, are the primary mediators of plant cell wall growth at low pH, implicating in cell division and elongation [67-70]. The in silico data for ß-expansin 8 (ß-EXP8), based on the unitag SD240878, showed $FC_{Tol}$ = 2.78 and $FC_{Sens}$ = -1.14. The RT-qPCR data presented diversified results (Table 2, Fig. 1): one tolerant cultivar showing induction (SP83-5073), another tolerant cultivar with repression (CTC15), and two sensitive cultivars (CTC 13 and SP90-1638) also with induction after the stress. Transcription of ExpB8 was induced (5, 10, 15 and 20 h) in the apical part of the maize roots treated with a ψW of -1.6 MPa, in relation to the control [water potential (ψW) of -0.03 Mpa] [71]. Also, the expression of β-expansin EXP B2 was significantly induced in soybean roots of plants simultaneously submitted to drought (0.2% PEG), and hormone (0.5 μM indole acetic acid - AIA), when compared with the control treatment [72]. The authors also observed that soybean EXP B2 increased the cell division and elongation, accentuating the growth and the assimilation of P in transgenic Arabidopsis overexpressing this target. Likewise, sugarcane ß-EXP8 could participate in the

architecture of the root system and contribute to drought tolerance, favoring a better acclimatization to this stress.

### 3.2.3. Protein degradation: Ubiquitination Pathway (E2, 20SPß4) and Plant Proteases (AP, C13)

Protein degradation and cleavage of proenzymes are important events for the development and plant acclimatization to stress conditions. Two main components of the ubiquitination pathway, involved in degradation of ubiquitin-labeled peptides, were here validated by RT-qPCR: ubiquitin conjugating enzyme (E2; EC 6.3.2.19), and 20S proteasome beta 4 subunit (20SPß4; EC 3.4.25.1). The RT-qPCR results for E2 (unitag SD214183, $FC_{Tol}$ = 31.39), one of the 50 most overexpressed, showed induction by the tolerant cultivar SP83-2847 (Table 2, Fig. 1). Further, RT-qPCR results for 20SPß4 (unitag SD230987, $FC_{Tol}$ = 2.23, $FC_{Sens}$ = 1.13; Table 2), showed induction by the tolerant cultivar SP83-5073, and by the sensitive cultivar SP90-3414; but, repression by the sensitive cultivar SP90-1638 (Table 2, Fig. 1). An efficient regulation of the breakdown of irreversibly degraded proteins during stress conditions, accurate processing of stock proteins, activation of functional proteins in the vacuoles by proteases, and degradation of ubiquitin-labeled peptides are events required for adequate initial acclimatization to stress. Presumably, some sugarcane cultivars are taking advantage of these processes.

Proteases are essential for the development and maintenance of plant protein homeostasis [73], catalyzing peptide bonds [74]. The protease of aspartic acid (AP; EC 3.4.23), represented by the unitag SD212448 ($FC_{Tol}$ = 3.39, $FC_{Sens}$ = -1.14; Table 2), is one of the most active endopeptidases in acid pH [75]. APs participate in pre-protein processing and *in vitro* protein degradation [76]. They are also present in the conduction and degradation of storage proteins during seed germination [77,78], degradation of

ribulose-1,5-bisphosphate carboxylase / oxygenase (Rubisco) in senescent leaves [79] and hydrolyzing proteins that are irreversibly damaged in responses to abiotic and biotic stresses [80-82]. The RT-qPCR results for this candidate showed induction of three (CTC6, SP83-2847, SP83-5073) of the four tolerant cultivars, whereas no significant variation was observed for any of the sensitive cultivars, confirming the *in silico* results (Table 2, Fig. 1). AP induction in roots of sugarcane tolerant to water deficit (RB867515), after drought treatments (24 h, three days and five days), was also reported using cDNA-AFLP [83].

In turn, cysteine proteinases have functions similar to AP, being essential in plant development, with involvement in storage, protein degradation, senescence, programmed cell death and in response to biotic and abiotic stresses [84,74]. Likewise, Legumain cysteine proteinases are vacuolar processing enzymes (VPE), functioning as seed storage proteins (legumins) [85], besides responding to several abiotic stresses [86]. The RT-qPCR results (Table 2, Fig. 1) for similar C13 (Legumain cysteine-type endopeptidase; EC 3.4.22.34), represented by the unitag SD34777 ($FC_{Tol}$ = 2.15, $FC_{Sens}$ = -1.55; Table 2) were induced by two tolerant cultivars (CTC6 and SP83-2847) and two sensitive cultivars (CTC9 and SP90-3414).

### 3.2.4. Oxidative detoxification / TRX

Thioredoxin 1-2 disulfide reductase (TRX, EC 1.8.1.9) acts on plastidial detoxification, regulating the redox state of target proteins (NADP-malate dehydrogenase, fructose-1,6-bisphosphatase, and chloroplast 2-Cysteine peroxiredoxin) [87-91]. The RT-qPCR results for putative TRX (Table 2, Fig. 1), represented by the unitag SD125268 ($FC_{Tol}$ = 5.93, $FC_{Sens}$ = -1.14; Table 2), showed induction by CTC15 and SP83-2847 (tolerant cultivars), and CTC9 (sensitive cultivar). Probably, these sugarcane cultivars take advantage of this enzyme activation during plastidial detoxification, and plant

acclimatization to the applied stress. The TRX CDSP 32, a 32 kD chloroplast protein activated by drought conditions in potato plants (*Solanum tuberosum* L.), presented induction after increasing water deficit for 12 days, with the leaf relative water content (RWC) reaching 60%, while the control presented 95% RWC foliar [87]. TRX genes evaluated in rice plants submitted to dehydration (25% de PEG for 1, 9 and 24 h) preponderantly induced most of them [92].

### 3.2.5. Fatty acid synthesis / ACC

Acetil CoA carboxilase (ACC, EC 6.4.1.2) is an enzyme that catalyzes the ATP-dependent carboxylation of acetyl-coenzyme A into malonyl-CoA, which are monomers of the fatty acid biosynthesis pathway, also being the primary regulator of this pathway [93,94]. A potential induced ACC (unitag SD65889, $FC_{Tol}$= 4.70, $FC_{Sens}$ = -1.67; Table 2) presented RT-qPCR results (Table 2, Fig. 1) showing induction by the tolerant cultivars CTC6 and SP83-2847, and repression by the sensitive cultivar CTC13. In these cultivars, the ACC may be involved in the synthesis of flavonoids, signaling responses to stress, or participating in the dryness reduction by epicuticular wax deposition. In peanut, there are reports of ACC overexpression in one of two tolerant genotypes under water suppression (seven days). Under stress, the tolerant genotype COC041 significantly increased leaf epicuticular wax deposition (EWL), while the sensitive genotype decreased the EWL [95]. The ACC locus was mapped in a major QTL (Quantitative Trait Loci) for drought tolerance in the millet grass (*Pennisetum glaucum* (L.) R. Br.), using a segregated population [96], along with 17 other genes. Besides ACC, HT-SuperSAGE unitags were detected for other four genes presented in this QTL: uridylate kinase (unitag SD86688, $FC_{Tol}$ = 4.80, $FC_{Sens}$ = 1.28); serine threonine kinase (unitag SD25475, $FC_{Tol}$ = 8.48, $FC_{Sens}$ = -1.25); chlorophyll a / b binding protein (unitag SD253068, $FC_{Tol}$ = 3.18), and ubiquitin conjugating enzyme

(unitag SD214183, $FC_{Tol}$ = 31.39). The latter was validated here by RT-qPCR (Table 2, Fig. 1), and all of them are promissing for the use in the sugarcane breeding, due to the existing synteny between grasses.

### 3.2.6. Amino acid transport / AAT

Amino acid transporters are responsible for the flow of amino acids across membranes, a secondary active process mediated by a co-transport of a proton and an amino acid, and they are essential for plant growth and development [97, 98]. These carrier proteins respond to environmental stimuli, such as luminosity, abiotic and biotic stresses, and their transcription levels can be regulated even with post-transcriptional regulation [99-101]. The unitag SD252857, a potential $Na^+$-dependent amino acid transporter (AAT; Table 2), showed induction by the tolerant ($FC_{Tol}$ = 6.50) and n.s. by the sensitive bulk ($FC_{Sens}$ = 1.17). Its induction expression was validated by RT-qPCR (Table 2, Fig. 1) for three of the four tolerant cultivars (CTC6, CTC15, and SP83-2847), and only one sensitive cultivar (CTC9). Besides, there is a genome-wide survey covering the gene family of amino acid transporters in rice (*Oryza sativa* L.) [102], but no reports are known addressing the role of the $Na^+$-dependent amino acid transporter in plants under abiotic stresses. Despite this, the AAT induction may indicate that the transport of neutral amino acids (nine of the 20 amino acids) would be more pronounced in the tolerant cultivars under 24 h after irrigation suppression when compared with the sensitive cultivars under the same stress.

### 3.2.7. Carbohydrate metabolism

#### 3.2.7.1. Glycolytic pathway / PFK, TPI, FBA

The 6-phosphofructo-2-kinase (PFK; EC 2.7.1.105), a glycolysis key enzyme [103], was represented here by the tolerant induced unitag SD133647 ($FC_{Tol}$ = 2.67, $FC_{Sens}$ = -1.02; Table 2). The RT-qPCR results (Table 2, Fig. 1) validated the induction

expression of CTC6 (tolerant cultivar; 3.08 fold) and SP90-1638 (sensitive cultivar; 2.58 fold) and the suppression of CTC13 (sensitive cultivar; 0.66 fold). The enzyme triosephosphate isomerase (TPI; EC 5.3.1.1) was another candidate here validated. This protein catalyzes the conversion of dihydroxyacetone phosphate into D-glyceraldehyde-3-phosphate in the glycolysis pathway [104], and it was represented by the unitag SD42156 ($FC_{Tol}$ = 2.10, $FC_{Sens}$ = -1.73). The RT-qPCR results confirmed the induction expression of the tolerant cultivar SP83-2847 (2.44 fold; Table 2, Fig. 1). In turn, fructose-bisphosphate aldolase (FBA; EC 4.1.2.13) catalyzes the reversible cleavage of fructose bisphosphate into glyceraldehyde 3-phosphate and dihydroxyacetone phosphate [105]. This enzyme was represented here by the unitag SD99499 ($FC_{Tol}$ = 2.78 and $FC_{Sens}$ = -1.52), which presented in RT-qPCR assays (Table 2, Fig. 1), induction expression of the tolerant cultivars CTC6 (1.57 fold) and SP83-2847 (1.76 fold), and also of the sensitive cultivar SP90-3414 (1.61 fold). The transcriptional levels of some enzymes of the glycolysis pathway, including PFK, TPI, and FBA, were detected stressed in roots of rice seedlings, in response to dryness (in Whatman paper after six hours) [106]. Furthermore, based on additional enzymes of the glycolysis pathway [glucose 6-phosphate isomerase (GPI, EC 5.3.1.9), enolase (ENO, EC 4.2.1.11), pyruvate kinase (PYK, EC 2.7.1.40) and glyceraldehyde phosphate (GAPDH, EC 1.2.1.12)] facing the same stresses, the authors suggested that glycolysis rates were accentuated in rice cells to compensate for the energy reduction caused by the abiotic stresses. Besides, PFK, TPI, and FBA, other components of the glycolysis pathway, also presented unitags (Table S.3). Those include the previously cited [GPI (unitag SD210623, $FC_{Tol}$ = 3.83, $FC_{Sens}$ = 1.26), ENO (unitag SD135385, $FC_{Tol}$ = 25.43), PYK (unitag SD106518, $FC_{Tol}$ = 5.56, $FC_{Sens}$ = 3.51)] and phosphoglucomutase (PGM, EC 5.4.2.2; unitag SD100485, $FC_{Tol}$ = 2.97, $FC_{Sens}$ = 1.76), Fructose-bisphosphate aldolase

(ALD, unitag SD117579, $FC_{Tol}$ = 3.18), phosphosglycerate kinase (PGK, EC 2.7.2.3, unitag SD140350, $FC_{Tol}$ = 2.38), triosephosphate isomerase (TIM, EC 5.3.1.1, unitag SD167995, $FC_{Tol}$ = 5.96), phosphoglycerate mutase (iPGM, EC 5.4.2.12; unitag SD75391, $FC_{Tol}$ = 1.76, $FC_{Sens}$ = 1.49). All these unitags presented induction by the tolerant bulk and n.s. by the sensitive bulk (Table S.3, Fig. S.3), reinforcing the glycolysis accentuation, after 24 h of irrigation suppression, in tolerance response. Accordingly, PFK, FBA, and TPI are attractive candidates for future use in sugarcane breeding programs.

#### 3.2.7.2. TCA cycle / LDP, MDH

Dihydrolipoyl dehydrogenase (LDP, EC 1.8.1.4) is a crucial component of the tricarboxylic acid (TCA) cycle which is an important aerobic pathway for the carbohydrates and fatty acids oxidations, catalyzing the electron transfer between pyridine nucleotides and disulfide compounds [107]. LPD, represented here by the unitag SD86754 ($FC_{Tol}$ = 5.93, $FC_{Sens}$ = -1.38), presented induction of the tolerant cultivar CTC15 (1.74 fold) in the RT-qPCR assays (Table 2, Fig. 1). This enzyme is involved with some abiotic stresses responses [108-110]. LDP overexpression in mangrove (*Rhizophora apiculata*) plants under hydroponic culture, 24 hours after exposed to 450 mM NaCl, was validated by RT-qPCR [108]. Malate dehydrogenase (MDH; EC 1.1.1.37), another TCA cycle enzyme, catalyzes the conversion (reversible reaction) of oxaloacetate into malate [109]. The unitag SD70910 (Table S.4), a putative MDH, showed induction by the tolerant bulk ($FC_{Tol}$ = 9.61 and $FC_{Sens}$ = -1.33) after the applied stress. The RT-qPCR results presented induction of the tolerant cultivar CTC6 (2.46 fold; Table 2, Fig. 1). Also, MDH has been reported in plant acclimatization to some abiotic stresses [110-112]. MDH overexpression was detected in apple plants with improved cold and salt tolerance when compared to wild-type plants [112]. The authors

concluded that MDH conferred transgenic apple plants greater tolerance to the studied stresses, modifying the redox state and increasing the salicylic acid level.

The mapping of TCA cycle components, according to KEGG Pathway database, considered nine unitags annotated to them (Table S.4), presenting all of them induced by the tolerant bulk but not by the sensitive bulk (Fig. S.4). As pointed to a glycolysis, a better energy efficiency could be expected by the tolerant cultivars in the drought stress tolerance responses.

#### 3.2.7.3. Pentose phosphate pathway / TKT

Transketolase (TKT; EC 2.2.1.1) participates in the second stage of the Pentose phosphate pathway, a process of glucose turnover, and it takes part of the Calvin cycle [113]. The unitag SD13813 ($FC_{Tol}$ = 12.72 and $FC_{Sens}$ = - 1.37; Table 2), a potential TKT, was one of the 50 most induced unitags by the tolerant bulk after the applied stress (Table 1). The RT-qPCR (Table 2, Fig. 1) validated this induction expression of the tolerant cultivars CTC6 (1.74 fold) and SP83-5073 (1.83 fold). TKT responsiveness to environmental factors has been already reported, such as its involvement in rehydration process of *Craterostigma plantagineum*, a resurrection plant, with the induction of two TKT transcripts (in leaves and roots, after six and seven hours, respectively) [114].

Regarding the mapping of unitags of pentose phosphate cycle components (KEGG Pathway database), eight of ten unitags (Table S.5) showed induction by the tolerant bulk, and at the same time n.s. or repressed by the sensitive bulk (Fig. S.5). The data overview suggests that sensitive cultivars are not so efficient in the mobilization of the energy machinery when under stress (24 h after irrigation suppression).

## 3.3. Individualizing sugarcane cultivars tolerant and sensitive to drought

The expression validation by RT-qPCR, of each proposed candidate and each cultivar composing the formed bulk, showed that not all cultivars from the same bulk had the same molecular responses as the results observed *in silico* for the entire bulk (Table 2, Fig. 1). This fact was expected, due to the different genetic backgrounds involved, the several processes acting and reflecting on some tolerance to stress in those cultivars, and particularly due to the intricate nature of the drought tolerance feature [115], culminating with a gradual tolerant phenotype in response to the applied stimulus. Joining genes and transcripts from a variety of genetic backgrounds is a plus when combining different genotypes in bulks. Another positive feature was the financial resources savings due to the lower number of libraries to be generated and sequenced. However, the presence of two or more genotypes in a bulk means a combination of common and distinct genotype transcripts, and if a differentiated expression is detected, it is necessary to find out which genotype potentially transcribed that transcript of interest. Besides, the dilution factor triggered by this mixed sample can interfere with the detection of a particular transcript.

The tolerant cultivars closer to the expected *in silico* inductions presented by the tolerant bulk were SP83-2847 and CTC6, each of them showing nine of the 15 candidates validated by RT-qPCR. From these nine, both tolerant cultivars shared six candidates: ACCD (from the Ethylene stress attenuation); AP and C13 (Protein degradation); ACC (Fatty acid synthesis); FBA (Glycolysis), and AAT (Amino acid transporter). These functional molecular markers are probably acting in the tolerance response, attenuating the ethylene stress and positively contributing to the better energy allocation, protein homeostasis, amino acid transport, and synthesis of flavonoids (signaling responses to stress or favoring the epicuticular wax deposition). In total, SP83-2847 and CTC6 validated 12 markers; except for LDP (from the TCA cycle), ß-

EXP8 (root growth accentuation) and 20SPß4 (Ubiquitination Pathway). In turn, the tolerant cultivar SP83-5073 induced AP (plant protease), ß-EXP8 (root growth accentuation), 20SPß4 (Ubiquitination Pathway), and TKT (Pentose phosphate pathway), while CTC15 induced AAT (Amino acid transporter), LDP (TCA cycle), and TRX (Oxidative detoxification). Some targets validated samples of only one tolerant cultivar, but four of them (ß-EXP8, PFK, MDH, and TPI) showing distinctive induction (Fig. 1). The relative importance of these candidates for the tolerance response needs further studies.

Concerning the enzymatic activities of CAT and APX, two ROS (Reactive oxygen species) detoxification enzymes, estimated in plants under water suppression periods (3, 10 and 20 days), two of those four drought-tolerant sugarcane cultivars reported here (SP83-2847 and SP83-5073) were evaluated and showed different responses [116]. In SP83-2847 an increase in both enzymatic activities occurred in the first treatment time. The same was not observed for SP83-5073, indicating that CAT and APX may have a protective role only in SP83-2847 [116].

Trying to evaluate the same set of eight sugarcane cultivars reported here, Thiebaut et al. [31] analyzed the expression of sRNAs in leaves and roots of plants after 24 h of irrigation suppression. The validation results (RT-qPCR), however, only considered SP83-2847 and SP83-5073 (as drought-tolerant cultivars), and SP90-1638 (-sensitive cultivar), and four targets. Based on those results, only the expression profile of one (miR159) was confirmed by the tolerant cultivar SP83-2847, but not by SP83-5073. These results are in the same direction obtained here since SP83-2847, which the main favorable agronomic trait is rusticity [117], induced nine targets when compared with SP83-5073, which validated only four. Our results also indicate CTC6, which presents

high productivity and production stability as the favorable agronomic traits [118], as another elite drought-tolerant accession.

Regarding the sensitive cultivars, half of them hardly presented induced candidates. The cultivar CTC13 (Table 2) only induced ß-EXP8 (root growth accentuation) but repressed PFK (Glycolysis), and ACC (Fatty acid synthesis). The cultivar SP90-1638 induced PFK (Glycolysis) and ß-EXP8 (root growth accentuation) but repressed 20SPß4 (Ubiquitination Pathway). In turn, the cultivar SP90-3414 induced 20SPß4 (Ubiquitination Pathway), C13 (Protein degradation), FBA (Glycolysis), and ACC (Fatty acid synthesis). In the same way, CTC9 induced AAT (Amino acid transporter), TRX (Pentose phosphate pathway), ACC (Fatty acid synthesis), and C13 (Protein degradation), and repressed ß-EXP8 (root growth accentuation). Despite the fact that drought-sensitive cultivars induced some candidates, their molecular profiles were distinct from those presented by SP83-2847 and CTC6 (both drought-tolerant cultivars). Keeping in mind that these cultivars, although being sensitive to the applied stress, have some favorable agronomic traits, such as ratooning and high productivity (SP90-1638) or high productivity and precocity (CTC9), according to the Brazilian breeding program [118]. Some targets, such as ACCD (Ethylene stress inhibition), MDH and LDP (from the TCA cycle), and AP and E2 (protein degradation) did not present induction by any of the evaluated sensitive cultivars, and their importance in the tolerance response should be verified in additional studies.

## 4. Conclusions

For a transcriptomic study, the choice of suitable cultivars is of great importance, especially when reviewing a complex character such as drought tolerance. Here, the HT-SuperSAGE technique allowed the evaluation of four drought-tolerant cultivars and

four drought-sensitive cultivars, after 24 h of irrigation suppression. Based on the results of 15 proposed and validated candidates, covering wide-ranging metabolic processes, their genetic profiles allowed to discriminate the drought-tolerant and also the -sensitive cultivars. Two drought-tolerant cultivars were considered similar (CTC6 and SP83-2847) and very promising for breeding purposes, showing very distinct profiles from two drought-sensitive cultivars (CTC13 and SP90-1638). The proposed set of molecular candidates could be promising in sugarcane breeding programs.

## 5. Competitive interests

The authors declare that they have no competing interests.

## 6. Authors' contributions

EAK generated the HT-SuperSAGE libraries. MDS and JRCFN performed the in silico analyzes. MDS and RLOS performed the RT-qPCR assays. AMBI and EAK assisted the discussion and critically reviewed the manuscript. All authors read and approved the final manuscript. EAK coordinated the research.

## 7. Acknowledgments

The authors are thankful to Centro de Tecnologia Canavieira (CTC, Piracicaba, SP, Brazil) and Ruth Jungmann, Dr. Björn Rotter, Dr. Peter Winter (GenXPro GmbH, Germany) and Dr. Guenter Kahl (*in memorium*, Frankfurt University, Germany) by the scientific advice and technical support. This work was supported by FINEP (Financiadora de Estudos e Projetos), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Ciência e Tecnologia do

Estado de Pernambuco (FACEPE) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## 8. References

[1] M. Menossi, M.C. Silva-Filho, M. Vincentz, M. a. Van-Sluys, G.M. Souza, Sugarcane functional genomics: Gene discovery for agronomic trait development, Int. J. Plant Genomics 2008 (2008) 1-11. doi:10.1155/2008/458732.

[2] FAOSTAT - Food and Agriculture Organization of the Unite Nations: FAO Statistical Databases. http://faostat.fao.org (2015). Accessed 18 Jun 2015.

[3] CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. www.conab.gov.br/ (2015). Accessed 18 Dec 2015.

[4] J.S. Boyer, Plant Productivity and Environment, Science, 218 (1982) 443-448.

[5] S.S. Gosal, S.H. Wani, M.S. Kang, Biotechnology and Drought Tolerance, J. Crop Improv. 23 (2009) 19–54. doi:10.1080/15427520802418251.

[6] S.M. Zingaretti, F.A. Rodrigues, J.P. Graça, L.M. Pereira, M.V. Lourenço, Sugarcane Responses at Water Deficit Conditions. In: I.M.M. Rahman, H. Hasegawa, editors. Water Stress. Croatia: Intech.2012. p. 256-276.

[7] E. A. Bray, Molecular Responses to Water Deficit., Plant Physiol. 103 (1993) 1035–1040. doi:10.1093/jxb/ern087.

[8] K. Shinozaki, K. Yamaguchi-Shinozaki, Gene networks involved in drought stress response and tolerance, J. Exp. Bot. 58 (2007) 221–227. doi:10.1093/jxb/erl164.

[9] A. D'Hont, GlaszmanJ.C., Sugarcane Genome Analysis With Molecular Markers: A First Dacade of Research, Proc. Int. Soc. Sugar Cane Technol. 24 (2001) 556–559. 10. Grivet L, Arruda P. Sugarcane genomics: depicting the complex genome of an important tropical crop. Curr Opin Plant Biol. 2002;5.

[10] L. Grivet, P. Arruda, Sugarcane genomics: Depicting the complex genome of an important tropical crop, Curr. Opin. Plant Biol. 5 (2002) 122–127. doi:10.1016/S1369-5266(02)00234-0.

[11] A. D'Hont, L. Grivet, P. Feldmann, S. Rao, N. Berding, J.C. Glaszmann, Characterization of the double genome structure of sugarcane cultivars (Saccharum ssp.) by molecular cytogenetics., Mol. Gen. Genet. 250 (1996) 405–413.

[12] Mardis, Next-generation DNA sequencing methods., Annu. Rev. Genomics Hum. Genet. 9 (2008) 387–402. doi:10.1146/annurev.genom.9.081307.164359.

[13] T.C. Glenn, Field guide to next-generation DNA sequencers, Mol. Ecol. Resour. 11 (2011) 759–769. doi:10.1111/j.1755-0998.2011.03024.x.

[14] N. de Setta, C.B. Monteiro-Vitorello, C.J. Metcalfe, G.M.Q. Cruz, L.E. Del Bem, R. Vicentini, et al., Building the sugarcane genome for biotechnology and identifying evolutionary trends., BMC Genomics. 15 (2014) 540. doi:10.1186/1471-2164-15-540.

[15] F.A. Rodrigues, M.L. de Laia, S.M. Zingaretti, Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants, Plant Sci. 176 (2009) 286–302. doi:10.1016/j.plantsci.2008.11.007.

[16] C.G. Lembke, M.Y. Nishiyama, P.M. Sato, R.F. de Andrade, G.M. Souza, Identification of sense and antisense transcripts regulated by drought in sugarcane, Plant Mol. Biol. 79 (2012) 461–477. doi:10.1007/s11103-012-9922-1.

[17] T.H. Ferreira, A. Gentile, R.D. Vilela, G.G.L. Costa, L.I. Dias, L. Endres, et al., microRNAs Associated with Drought Response in the Bioenergy Crop Sugarcane (Saccharum spp.), PLoS One. 7 (2012). doi:10.1371/journal.pone.0046703.

[18] A. Gentile, T.H. Ferreira, R.S. Mattos, L.I. Dias, A.A. Hoshino, M.S. Carneiro, et al., Effects of drought on the microtranscriptome of field-grown sugarcane plants, Planta. 237 (2013) 783–798. doi:10.1007/s00425-012-1795-7.

[19] E.A. Kido, J.R.C. Ferreira Neto, R.L.D.O. Silva, V. Pandolfi, A.C.R. Guimarães, D.T. Veiga, et al., New Insights in the Sugarcane Transcriptome Responding to Drought Stress as Revealed by Supersage, Sci. World J. 2012 (2012) 1–14. doi:10.1100/2012/821062..

[20] L. Vargas, A.B.S. Brigida, J.P. Mota, T.G. Carvalho, C.A. Rojas, D. Vaneechoutte, et al., Drought tolerance conferred to sugarcane by association with gluconacetobacter diazotrophicus: A transcriptomic view of hormone pathways, PLoS One. 9 (2014) 1–37. doi:10.1371/journal.pone.0114744.

[21] C. Li, Q. Nong, M.K. Solanki, Q. Liang, J. Xie, X. Liu, et al., Differential expression profiles and pathways of genes in sugarcane leaf at elongation stage in response to drought stress, Sci. Rep. 6 (2016) 25698. doi:10.1038/srep25698.

[22] H. Matsumura, S. Reich, A. Ito, H. Saitoh, S. Kamoun, P. Winter, et al., Gene expression analysis of plant host-pathogen interactions by SuperSAGE., Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 15718–15723. doi:10.1073/pnas.2536670100.

[23] H. Matsumura, D.H. Krüger, G. Kahl, R. Terauchi, SuperSAGE: A Modern Platform for Genome-Wide Quantitative Tran- script Profiling, Curr. Pharm. Biotechnol. 9 (2008) 368–374. doi:10.2174/138920108785915157.

[24] B. Coemans, H. Matsumura, R. Terauchi, S. Remy, R. Swennen, L. Sági, SuperSAGE combined with PCR walking allows global gene expression profiling of banana (Musa acuminata), a non-model organism, Theor. Appl. Genet. 111 (2005) 1118–1126. doi:10.1007/s00122-005-0039-7.

[25] C. Molina, B. Rotter, R. Horres, S.M. Udupa, B. Besser, L. Bellarmino, et al., SuperSAGE: the drought stress-responsive transcriptome of chickpea roots., BMC Genomics. 9 (2008) 553. doi:10.1186/1471-2164-9-553.

[26] C. Molina, M. Zaman-Allah, F. Khan, N. Fatnassi, R. Horres, B. Rotter, et al., The salt-responsive transcriptome of chickpea roots and nodules via deepSuperSAGE., BMC Plant Biol. 11 (2011) 31. doi:10.1186/1471-2229-11-31.

[27] E.A. Kido, P. Kelle, D.A. Barbosa, J. Ribamar, C. Ferreira, Identification of Plant Protein Kinases in Response to Abiotic and Biotic Stresses Using SuperSAGE Identification of Plant Protein Kinases in Response to Abiotic and Biotic Stresses Using SuperSAGE, (2011). doi:10.2174/1389203711109070643.

[28] A. El Kelish, F. Zhao, W. Heller, J. Durner, J. Winkler, H. Behrendt, et al., Ragweed (Ambrosia artemisiifolia) pollen allergenicity: SuperSAGE transcriptomic analysis upon elevated CO2 and drought stress, BMC Plant Biol. 14 (2014) 176. doi:10.1186/1471-2229-14-176.

[29] E. Canales, Y. Coll, I. Hernández, R. Portieles, M.R. García, Y. López, et al., "Candidatus Liberibacter asiaticus", causal agent of citrus huanglongbing, is reduced by treatment with brassinosteroids, PLoS One. 11 (2016) 1–16. doi:10.1371/journal.pone.0146223.

[30] F. Thiebaut, C. Grativol, M. Carnavale-Bottino, C.A. Rojas, M. Tanurdzic, L. Farinelli, et al., Computational identification and analysis of novel sugarcane microRNAs., BMC Genomics. 13 (2012) 290. doi:10.1186/1471-2164-13-290.

[31] F. Thiebaut, C. Grativol, M. Tanurdzic, M. Carnavale-Bottino, T. Vieira, M.R. Motta, et al., Differential sRNA regulation in leaves and roots of sugarcane under water depletion, PLoS One. 9 (2014). doi:10.1371/journal.pone.0093822.

[32] S. Audic, J.M. Claverie, The significance of digital gene expression profiles, Genome Res. 7 (1997) 986–995. doi:10.1101/gr.7.10.986.

[33] S. a Bustin, V. Benes, J. a Garson, J. Hellemans, J. Huggett, M. Kubista, et al., The MIQE guidelines:Minimum Information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (2009) 611–622. doi:10.1373/clinchem.2008.112797.

[34] S. Rozen, H. Skaletsky, Primer3 on the WWW for Biologist Programmers, Totowa, NJ. Humana Press. 132 (2000) 365–386. doi:10.1385/1-59259-192-2:365.

[35] R. L. O. Silva, M.D. Silva, J.R.C. Ferreira Neto, C.H. Nardi, S.M. Chabregas, W.L. Burnquist, G. Kahl, A.M. Benko-Iseppon, E.A. Kido, Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-Time PCR in Drought-Stressed Sugarcane. Scientific World Journal. 2014 (2014) 1-13.

[36] R. Rasmussen, Quantification on the LightCycler. In. S. Meuer, C. Wittwer, K. Nakagawara, editors: Rapid cycle real-time PCR, methods and applications. Heidelberg:Springer Press. 2001, 2001.

[37] M.W. Pfaffl, G.W. Horgan, L. Dempfle, Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR., Nucleic Acids Res. 30 (2002) e36. doi:10.1093/nar/30.9.e36.

[38] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets, Nucleic Acids Res. 40 (2012) 1–6. doi:10.1093/nar/gkr988..

[39] K. Apel, H. Hirt, REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction, Annu. Rev. Plant Biol. 55 (2004) 373–399. doi:10.1146/annurev.arplant.55.031903.141701.

[40] R. Mittler, S. Vanderauwera, M. Gollery, F. Van Breusegem, Reactive oxygen gene network of plants, Trends Plant Sci. 9 (2004) 490–498. doi:10.1016/j.tplants.2004.08.009.

[41] J.M. Van Norman, N.W. Breakfield, P.N. Benfey, Intercellular communication during plant development, Plant Cell. 23 (2011) 855–864. doi:10.1105/tpc.111.082982\rtpc.111.082982 [pii].4.

[42] S. Shigeoka, T. Ishikawa, M. Tamoi, Y. Miyagawa, T. Takeda, Y. Yabuta, et al., Regulation and function of ascorbate peroxidase isoenzymes, J. Exp. Bot. 53 (2002) 1305–1319. doi:10.1093/jexbot/53.372.1305.

[43] S.S. Gill, N. Tuteja, Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants, Plant Physiol. Biochem. 48 (2010) 909–930. doi:10.1016/j.plaphy.2010.08.016.

[44] G. Noctor, C.H. Foyer, ASCORBATE AND GLUTATHIONE: Keeping Active Oxygen Under Control, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49 (1998) 249–279. doi:10.1146/annurev.arplant.49.1.249..

[45] W. Wang, B. Vinocur, O. Shoseyov, A. Altman, Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response, Trends Plant Sci. 9 (2004) 244–252. doi:10.1016/j.tplants.2004.03.006..

[46] A. Ciechanover, The ubiquitin – proteasome pathway : on protein death and cell life, EMBO17 (1998) 7151–7160.

[47] W.J. Lyzenga, S.L. Stone, Abiotic stress tolerance mediated by protein ubiquitination, J. Exp. Bot. 63 (2012) 599–616. doi:10.1093/jxb/err310.

[48] G.A. Zhou, R.Z. Chang, L.J. Qiu, Overexpression of soybean ubiquitin-conjugating enzyme gene GmUBC2 confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in Arabidopsis, Plant Mol. Biol. 72 (2010) 357–367. doi:10.1007/s11103-009-9575-x..

[49] Q. Guo, J. Zhang, Q. Gao, S. Xing, F. Li, W. Wang, Drought tolerance through overexpression of monoubiquitin in transgenic tobacco, J. Plant Physiol. 165 (2008) 1745–1755. doi:10.1016/j.jplph.2007.10.002.

[50] B. Lemon, R. Tjian, Orchestrated response: A symphony of transcription factors for gene control, Genes Dev. 14 (2000) 2551–2569. doi:10.1101/gad.831000.

[51] J. Jin, H. Zhang, L. Kong, G. Gao, J. Luo, PlantTFDB 3.0: A portal for the functional and evolutionary study of plant transcription factors, Nucleic Acids Res. 42 (2014) 1182–1187. doi:10.1093/nar/gkt1016.

[52] E. Henriksson, A.S.B. Olsson, H. Johannesson, H. Johansson, J. Hanson, P. Engström, et al., Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships., Plant Physiol. 139 (2005) 509–518. doi:10.1104/pp.105.063461.et.

[53] F.D. Ariel, P. a. Manavella, C. a. Dezar, R.L. Chan, The true story of the HD-Zip family, Trends Plant Sci. 12 (2007) 419–426. doi:10.1016/j.tplants.2007.08.003..

[54] M. Honma, T. Shimomura, Metabolism of 1-Aminocyclopropane-1-carboxylic Acid, Agric. Biol. Chem. 42 (1978) 1825–1831. doi:10.1080/00021369.1978.10863261.

[55] F.B. Abeles, P.W. Morgan, M.E. Saltveit. Ethylene in Plant Biology. 2nd ed. San Diego, CA: Academic;1992

[56] A.B. Bleecker, H. Kende, Ethylene: a gaseous signal molecule in plants., Annu. Rev. Cell Dev. Biol. 16 (2000) 1–18. doi:10.1146/annurev.cellbio.16.1.1.

[57] J.C. Stearns, B.R. Glick, Transgenic plants with altered ethylene biosynthesis or perception, Biotechnol. Adv. 21 (2003) 193–210. doi:10.1016/S0734-9750(03)00024-7.

[58] V. Govindasamy, M. Senthilkumar, P. Bose, L.D. Kumar, A. Kannepalli, ACC Deaminase Containing PGPR for Potential Exploitation in Agriculture. In: D.K. Maheshwari, editor. Bacteria in Agrobiology: Plant Nutrient Management. Springer Berlin Heidelberg; 2011. p. 183-208.

[59] H. Hyodo, Stress/wound ethylene. In:A.k. Mattoo, Suttle JC, editors. The plant hormone ethylene. Boca Raton CRC; 1991. p. 65–80.

[60] B. Glick, D. Penrose, J. Li, A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria, J. Theor. Biol. 190 (1998) 63–8. doi:10.1006/jtbi.1997.0532..

[61] B.R. Glick, Z. Cheng, J. Czarny, J. Duan, Promotion of plant growth by ACC deaminase containing soil bacteria, Eur J Plant Pathol. 119 (2007) 329–339.

[62] N. Hontzeas, S.S. Saleh, B.R. Glick, Changes in gene expression in canola roots induced by ACC-deaminase-containing plant-growth-promoting bacteria., Mol. Plant. Microbe. Interact. 17 (2004) 865–871. doi:10.1094/MPMI.2004.17.8.865..

[63] L. McDonnell, J.M. Plett, S. Andersson-Gunneras, C. Kozela, J. Dugardeyn, D. Van Der Straeten, et al., Ethylene levels are regulated by a plant encoded 1-aminocyclopropane-1-carboxylic acid deaminase, Physiol. Plant. 136 (2009) 94–109. doi:10.1111/j.1399-3054.2009.01208.x.

[64] B.R. Glick, Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase, FEMS Microbiol. Lett. 251 (2005) 1–7. doi:10.1016/j.femsle.2005.07.030.

[65] J.-H. Lim, S.-D. Kim, Induction of Drought Stress Resistance by Multi-Functional PGPR Bacillus licheniformis K11 in Pepper., Plant Pathol. J. 29 (2013) 201–8. doi:10.5423/PPJ.SI.02.2013.0021.

[66] S. Ali, T.C. Charles, B.R. Glick, Amelioration of high salinity stress damage by plant growth-promoting bacterial endophytes that contain ACC deaminase., Plant Physiol. Biochem. 80 (2014) 160–7. doi:10.1016/j.plaphy.2014.04.003.

[67] D.J. Cosgrove, Update on Cell Growth Cell Wall Loosening by Expansins 1, Plant Physiol. 118 (1998) 333–339. doi:10.1104/pp.118.2.333.

[68] D.J. Cosgrove, Enzymes and other agents that enhance cell wall extensibility., Annu. Rev. Plant Physiol. Plant Mol. Biol. 50 (1999) 391–417. doi:10.1146/annurev.arplant.50.1.391.

[69] A.C. Li, D.J. Cosgrove, Grass group I pollen allergens (b-expansins) lack proteinase activity and do not cause wall loosening via proteolysis, Eur. J. Biochem. 268 (2001) 4217–4226. doi:10.1046/j.1432-1327.2001.02336.x.

[70] B. Muller, G. Bourdais, B. Reidy, C. Bencivenni, A. Massonneau, P. Condamine, et al., Association of specific expansins with growth in maize leaves is maintained under environmental, genetic, and developmental sources of variation., Plant Physiol. 143 (2007) 278–290. doi:10.1104/pp.106.087494..

[71] Y. Wu, E.T. Thorne, R.E. Sharp, D.J. Cosgrove, Modification of expansin transcript levels in the maize primary root at low water potentials., Plant Physiol. 126 (2001) 1471–1479. doi:10.1104/pp.126.4.1471.

[72] W. Guo, J. Zhao, X. Li, L. Qin, X. Yan, H. Liao, A soybean β-expansin gene GmEXPB2 intrinsically involved in root system architecture responses to abiotic stresses, Plant J. 66 (2011) 541–552. doi:10.1111/j.1365-313X.2011.04511.x.

[73] M. Kidrič, J. Kos, J. Sabotič, Proteases and their endogenous inhibitors in the plant response to abiotic stress, Bot. Serbica. 38 (2014) 139–158..

[74] I. Vaseva, J. Sabotič, J. Šuštar-Vozlič , et al. The Response ofPlants to Drought Stress: The Role of Dehydrins, Chaperones, Proteases and Protease Inhibitors in Maintain Cellular Protein Function. In: Droughts New Research. Edited by D.F. J.D.S. Neves. 2012 (January).

[75] J.S. Bond, P.E. Butler, Intracellular proteases, Annu Rev Biochem. 56 (1987) 333–364.

[76] E.P. Beers, B.J. Woffenden, C. Zhao, Plant proteolytic enzymes: possible roles during programmed cell death. Plant Mol Biol. 44 (2000) 399–415.

[77] A. Mutlu 1, S. Gal, Plant aspartic proteinases: enzymes on the way to a function, Physiol. Plant. 105 (1999) 569–576. doi:10.1034/j.1399-3054.1999.105324.x.

[78] P. Runeberg-Roos, M. Saarma, Phytepsin, a barley vacuolar aspartic proteinase, is highly expressed during autolysis of developing tracheary elements and sieve cells. Plant J. 15 (1998) 139–145.

[79] Y. Kato, S. Murakami, Y. Yamamoto, H. Chatani, Y. Kondo, T. Nakano, et al., The DNA-binding protease, CND41, and the degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase in senescent leaves of tobacco, Planta. 220 (2004) 97–104. doi:10.1007/s00425-004-1328-0.

[80] X. Yao, W. Xiong, T. Ye, Y. Wu, Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in Arabidopsis, J. Exp. Bot. 63 (2012) 2579–2593. doi:10.1093/jxb/err433.

[81] R. Guo, X. Xu, B. Carole, X. Li, M. Gao, Y. Zheng, et al., Genome-wide identification, evolutionary and expression analysis of the aspartic protease gene superfamily in grape., BMC Genomics. 14 (2013) 554. doi:10.1186/1471-2164-14-554.

[82] G.S. Timotijević, M.D. Milisavljević, S.R. Radović, M.M. Konstantinović, V.R. Maksimović, Ubiquitous aspartic proteinase as an actor in the stress response in buckwheat, J. Plant Physiol. 167 (2010) 61–68. doi:10.1016/j.jplph.2009.06.017.

[83] J.S. Vantini, G.C. Dedemo, D.F.J. Gimenez, L.F.S. Fonseca, Differential gene expression in drought-tolerant sugarcane roots, 14 (2015) 7196–7207.

[84] M. Grudkowska, B. Zagdańska, Multifunctional role of plant cysteine proteinases, Acta Biochim. Pol. 51 (2004) 609–624. doi:045103609.

[85] K. Müntz, A.D. Shutov, Legumains and their functions in plants, Trends Plant Sci. 7 (2002) 340–344. doi:10.1016/S1360-1385(02)02298-7.

[86] T. Kinoshita, K. Yamada, N. Hiraiwa, M. Kondo, M. Nishimura, I. Hara-Nishimura, Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions, Plant J. 19 (1999) 43–53. doi:10.1046/j.1365-313X.1999.00497.x.

[87] P. Rey, G. Pruvot, N. Becuwe, F. Eymery, D. Rumeau, G. Peltier, A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in Solanum tuberosum L. plants, Plant J. 13 (1998) 97–107. doi:10.1046/j.1365-313X.1998.00015.x..

[88] C. Vieira Dos Santos, P. Rey, Plant thioredoxins are key actors in the oxidative stress response, Trends Plant Sci. 11 (2006) 329–334. doi:10.1016/j.tplants.2006.05.005..

[89] V. Collin, E. Issakidis-Bourguet, C. Marchand, M. Hirasawa, J.M. Lancelin, D.B. Knaff, et al., The Arabidopsis plastidial thioredoxins. New functions and new insights into specificity, J. Biol. Chem. 278 (2003) 23747–23752. doi:10.1074/jbc.M302077200.

[90] E. Gelhaye, N. Rouhier, N. Navrot, J.P. Jacquot, The plant thioredoxin system, Cell. Mol. Life Sci. 62 (2005) 24–35. doi:10.1007/s00018-004-4296-4.

[91] M. Broin, P. Rey, Potato plants lacking the CDSP32 plastidic thioredoxin exhibit overoxidation of the BAS1 2-cysteine peroxiredoxin and increased lipid Peroxidation in thylakoids under photooxidative stress., Plant Physiol. 132 (2003) 1335–1343. doi:10.1104/pp.103.021626..

[92] M. Nuruzzaman, A.M. Sharoni, K. Satoh, T. Al-Shammari, T. Shimizu, T. Sasaya, et al., The thioredoxin gene family in rice: Genome-wide identification and expression profiling under different biotic and abiotic treatments, Biochem. Biophys. Res. Commun. 423 (2012) 417–423. doi:10.1016/j.bbrc.2012.05.142.

[93] D. Post-Beitenmiller, G. Roughan, J.B. Ohlrogge, Regulation of plant fatty acid biosynthesis: analysis of acyl-coenzyme A and acyl-acyl-carrier protein substrate pools in spinach and pea chloroplasts. Plant Physiol. 100 (1992) 923–930.

[94] M. Focke, E. Gieringer, S. Schwan, L. Jänsch, S. Binder, H.-P. Braun, Fatty acid biosynthesis in mitochondria of grasses: malonyl-coenzyme A is generated by a

mitochondrial-localized acetyl-coenzyme A carboxylase., Plant Physiol. 133 (2003) 875–884. doi:10.1104/pp.103.027375.

[95] K.R. Kottapalli, R. Rakwal, J. Shibato, G. Burow, D. Tissue, J. Burke, et al., Physiology and proteomics of the water-deficit stress response in three contrasting peanut genotypes, Plant, Cell Environ. 32 (2009) 380–407. doi:10.1111/j.1365-3040.2009.01933.x.

[96] D. Sehgal, V. Rajaram, I.P. Armstead, V. Vadez, Y.P. Yadav, C.T. Hash, et al., Integration of gene-based markers in a pearl millet genetic map for identification of candidate genes underlying drought tolerance quantitative trait loci, BMC Plant Biol. 12 (2012) 9. doi:10.1186/1471-2229-12-9.

[97] Ortiz-Lopez, H. Chang, D.R. Bush, Amino acid transporters in plants., Biochim. Biophys. Acta. 1465 (2000) 275–280.

[98] D.R. Bush, Proton-Coupled Sugar and Amino Acid Transporters in Plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 44 (1993) 513–542. doi:10.1146/annurev.pp.44.060193.002501..

[99] M. Tegeder, Transporters for amino acids in plant cells: Some functions and many unknowns, Curr. Opin. Plant Biol. 15 (2012) 315–321. doi:10.1016/j.pbi.2012.02.001.

[100] S. Delrot, R. Atanassova, L. Maurousset, Regulation of sugar, amino acid and peptide plant membrane transporters, Biochim. Biophys. Acta - Biomembr. 1465 (2000) 281–306. doi:10.1016/S0005-2736(00)00145-0.

[101] X. Liu, D.R. Bush, Expression and transcriptional regulation of amino acid transporters in plants, Amino Acids. 30 (2006) 113–120. doi:10.1007/s00726-005-0248-z..

[102] H. Zhao, H. Ma, L. Yu, X. Wang, J. Zhao, Genome-Wide Survey and Expression Analysis of Amino Acid Transporter Gene Family in Rice (Oryza sativa L.), PLoS One. 7 (2012). doi:10.1371/journal.pone.0049210.

[103] W.C. Plaxton, the Organization and Regulation of Plant Glycolysis., Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 185–214. doi:10.1146/annurev.arplant.47.1.185.

[104] E. Pichersky, L.D. Gottlieb, Evidence for Duplication of the Structural Genes Coding Plastid and Cytosolic Isozymes of Triose Phosphate Isomerase in Diploid Species of Clarkia,116 Genetics (1983) 421–436.

[105] J.J. Marsh, H. G. Lebherz, Fructose-bisphosphate aldolases: an evolutionary history, Trends Biochem Sci. 17 (1992) 110–113.

[106] D. Minhas, A. Grover, Transcript levels of genes encoding various glycolytic and fermentation enzymes change in response to abiotic stresses, Plant Sci. 146 (1999) 41–51. doi:10.1016/S0168-9452(99)00092-8.

[107] D.J. Carothers, G. Pons, M.S. Patel, Dihydrolipoamide dehydrogenase: Functional similarities and divergent evolution of the pyridine nucleotide-disulfide oxidoreductases, Arch. Biochem. Biophys. 268 (1989) 409–425. doi:10.1016/0003-9861(89)90309-3.

[108] M. Sahebi, M. Hanafi, Extraction of total RNA from mangrove plants to identify different genes involved in its adaptability to the variety of stresses, Pakistan J. 50 (2013) 529–536.

[109] R.A. Musrati, M. Kollárová, N. Mernik, D. Mikulášová, Malate Dehydrogenase: Distribution, Function and Properties, Gen. Physiol. Biophys. 17 (1998) 193–210.

[110] Z. Chen, L. Sun, P. Liu, G. Liu, J. Tian, H. Liao, Malate Synthesis and Secretion Mediated by a Manganese-Enhanced Malate Dehydrogenase Confers Superior Manganese Tolerance in Stylosanthes guianensis., Plant Physiol. 167 (2015) 176–88. doi:10.1104/pp.114.251017..

[111] Q. Li, J. Zhao, J. Zhang, Z. Dai, L. Zhang, Ectopic Expression of the Chinese Cabbage Malate Dehydrogenase Gene Promotes Growth and Aluminum Resistance in Arabidopsis, Front. Plant Sci. 7 (2016) 1–11. doi:10.3389/fpls.2016.01180.

[112] Q.J. Wang, H. Sun, Q.L. Dong, T.Y. Sun, Z.X. Jin, Y.J. Hao, et al., The enhancement of tolerance to salt and cold stresses by modifying the redox state and salicylic acid content via the cytosolic malate dehydrogenase gene in transgenic apple plants, Plant Biotechnol. J. 14 (2016) 1986–1997. doi:10.1111/pbi.12556.

[113] T. ap Rees, The organisation of glycolysis and the oxidative pentose phosphate pathway in plants, In Encyclopaedia of Plant Physiology, NS, vol 18. Edited by R. Douce, D.A. Day, Berlin Springer (1985) 391–417.

[114] G. Bernacchia, G. Schwall, F. Lottspeich, F. Salamini, D. Bartels, The transketolase gene family of the resurrection plant Craterostigma plantagineum: differential expression during the rehydration phase., EMBO J. 14 (1995) 610–618.

[115] H. Hu, L. Xiong, Genetic engineering and breeding of drought-resistant crops., Annu. Rev. Plant Biol. 65 (2014) 715–41. doi:10.1146/annurev-arplant-050213-040000.
[116] M.C. Cia, A.C.R. Guimarães, L.O. Medici, S.M. Chabregas, R.A. Azevedo, Antioxidant responses to water deficit by drought-tolerant and -sensitive sugarcane varieties, Ann. Appl. Biol. 161 (2012) 313–324. doi:10.1111/j.1744-7348.2012.00575.x. [117] agrobyte. http://www.agrobyte.com.br/variedade_cana.htm/ (2017). Acessed 18 Jan 2017.
[118] CTC - CENTRO DE TECNOLOGIA CANAVIEIRA. http://www.ctcanavieira.com.br/ (2017). Accessed 18 Jan 2017.

**Table(s)**

**Table 1.** The 50 most induced HT-SuperSAGE unitags expressed by the tolerant bulk (four sugarcane cultivars) and at the same time repressed (or n.s. at p-value < 0.05 in Audic-Claverie test) by the sensitive bulk after 24 h of irrigation suppression. Also, their respective EST anchoring the unitag, the EST annotation and the fold change values (FC) considering the tolerant and the sensitive contrast ($FC_{tol}$, $FC_{sens}$, respectively).

| Unitag | $FC_{Tol}$ | $FC_{Sens}$ | EST ID | Annotation |
|---|---|---|---|---|
| SD110234 | 209.02 | -1.09 | CO373199 | S-adenosyl-L-methionine decarboxylase |
| SD148347 | 192.33 | **-27.14** | TC114152 | 0 |
| SD8928 | 183.99 | n.d. | TC131103 | Thylakoid-bound ascorbate peroxidase |
| SD244046 | 106.50 | n.d. | CF574491 | ribosomal protein |
| SD15507 | 87.82 | 1.69 | TC144209 | Cysteine proteinase inhibitor |
| SD151012 | 78.28 | n.d. | TC141418 | Hypoxia induced conserved protein region containing protein |
| SD177657 | 59.61 | n.d. | CF573548 | Profilin |
| SD113531 | 39.34 | **-13.07** | TC137214 | 17.5 kDa class II heat shock protein |
| SD247649 | 31.79 | -1.63 | TC133246 | 0 |
| SD214183 | 31.39 | n.d. | CF573288 | ubiquitin-conjugating enzyme |
| SD182788 | 29.41 | n.d. | TC120640 | Cytosolic ascorbate peroxidase |
| SD78974 | 28.61 | n.d. | TC135929 | Homeodomain leucine zipper protein 16 |
| SD21711 | 28.21 | 1.40 | CA251528 | 3-hydroxydecanoyl-ACP dehydratase |
| SD108897 | 25.83 | **-2.66** | TC144102 | Quinone reductase 2 |
| SD135385 | 25.43 | n.d. | TC132317 | Enolase |
| SD9805 | 25.04 | n.d. | TC139637 | Glycine-rich RNA binding protein |
| SD271576 | 24.24 | **-1.18** | TC145473 | O-methyltransferase 2 |
| SD974 | 23.05 | **-2.11** | TC117869 | 60S ribosomal protein L27 |
| SD97410 | 22.25 | n.d. | TC113567 | 0 |
| SD15211 | 21.46 | n.d. | TC116548 | Polyubiquitin |
| SD181737 | 19.47 | n.d. | TC146390 | Elongation factor 1-alpha |
| SD71948 | 19.07 | 1.50 | TC116208 | NADH ubiquinone oxidoreductase subunit |
| SD231907 | 18.68 | n.d. | TC113683 | MLIP15 |
| SD20245 | 18.68 | n.d. | CF576896 | histone protein |
| SD260174 | 18.68 | **-2.17** | TC137231 | Acid phosphatase |
| SD146732 | 18.28 | 1.01 | TC117661 | 0 |
| SD89732 | 17.88 | **-2.28** | CF571626 | heat shock protein 81-1 |
| SD130835 | 17.88 | **-20.22** | TC128723 | Adhesive/proline-rich |
| SD97673 | 17.09 | **-19.16** | TC113567 | 0 |
| SD59386 | 16.96 | n.d. | TC128702 | 0 |
| SD170716 | 16.69 | n.d. | TC133350 | Reticulon |
| SD81084 | 16.69 | **-22.88** | CA236418 | Short-chain dehydrogenase/reductase SDR |
| SD20852 | 16.69 | -2.19 | TC145671 | ENOD40 |
| SD67697 | 15.9 | -2.63 | CA123780 | Protein translation factor SUI1 homolog |
| SD216657 | 15.83 | 1.32 | CA154394 | 0 |
| SD28082 | 15.1 | **-6.39** | TC143544 | Aquaporin PIP2-4 |
| SD233168 | 15.1 | n.d. | CA293374 | 22 kDa drought-inducible protein |
| SD9802 | 14.7 | n.d. | TC122856 | Glycine-rich protein |
| SD66227 | 14.41 | -1.54 | TC133157 | Protein kinase domain |

Cont.

| Unitag | $FC_{Tol}$ | $FC_{Sens}$ | EST ID | Annotation |
|---|---|---|---|---|
| SD202531 | 14.31 | n.d. | CF576381 | Glucose-6-phosphate/phosphate antiporter |
| SD212840 | 14.31 | 1.27 | CF575729 | stress-related protein |
| SD160238 | 13.91 | n.d. | TC120160 | 0 |
| SD275064 | 13.91 | n.d. | TC144939 | Histone H4 |
| SD212934 | 13.91 | n.d. | TC148493 | Metallothionein-like protein 1B |
| SD243866 | 13.51 | **-15.97** | TC119187 | Aquaporin PIP2-1 |
| SD82594 | 13.51 | **-9.58** | CA135157 | Elongation factor 1-alpha |
| SD230319 | 13.28 | 2.20 | TC153087 | 60S ribosomal protein L23A |
| SD9806 | 13.11 | n.d. | TC153183 | Glycine-rich RNA binding protein |
| SD193697 | 13.11 | 1.69 | TC114241 | 0 |
| SD13813 | 12.72 | -1.37 | TC134850 | Transketolase. chloroplast |

*In bold: repressed (p value < 0.05)

**Table 2.** Proposed candidate genes and their results based on in silico analysis of HT-SuperSAGE unitags expressed by drought-tolerant or -sensitive bulks (four cultivars each one) and their relative expression by RT-qPCR, using sugarcane cDNAs from roots of each cultivar after 24 h of irrigation suppression.

| | **Candidate genes** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Method** | ACCD | TRX | LDP | ß-EXP8 | E2 | 20SPß4 | AP | C13 |
| HT-SuperSAGE[&] | | | | | | | | |
| Tolerant bulk | 1.90[#] / UR | 5.94[#] / UR | 5.93[#] / UR | 2.78[#] / UR | 31.39[#] / UR | 2.23[#] / UR | 3.39[#] / UR | 2.15[#] / UR |
| Sensitive bulk | 1.09[#] / n.s. | -1.14[#] / n.s. | -1.38[#] / n.s. | -1.14[#] / n.s. | not detected | 1.13[#] / n.s. | -1.14[#] / n.s. | -1.55[#] / DR |
| RT-qPCR[**] | | | | | | | | |
| Tolerant cultivar | | | | | | | | |
| CTC6 | 1.93 / UR | 1.05 / n.s. | 1.23 / n.s. | 1.87 / n.s. | 0.89 / n.s. | 0.77 / n.s. | 2.18 / UR | 3.71 / UR |
| CTC15 | 0.64 / n.s. | 3.32 / UR | 1.74 / UR | 0.27 / DR | 0.85 / n.s. | 1.12 / n.s. | 1.16 / n.s. | 0.93 / n.s. |
| SP83-2847 | 1.73 / UR | 1.74 / UR | 0.86 / n.s | 0.97 / n.s. | 1.77 / UR | 1.00 / n.s. | 1.64 / UR | 2.23 / UR |
| SP83-5073 | 1.20 / n.s. | 1.32 / n.s. | 1.07 / n.s. | 3.87 / UR | 0.91 / n.s. | 1.72 / UR | 2.01 / UR | 1.34 / n.s. |
| Sensitive cultivar | | | | | | | | |
| CTC9 | 1.20 / n.s. | 2.34 / UR | 1.00 / n.s. | 0.44 / DR | 1.47 / n.s. | 1.47 / n.s. | 1.15 / n.s. | 2.90 / UR |
| CTC13 | 0.95 / n.s. | 1.13 / n.s. | 0.83 / n.s. | 4.57 / UR | 0.96 / n.s. | 0.92 / n.s. | 0.77 / n.s. | 1.29 / n.s. |
| SP90-1638 | 1.63 / n.s. | 1.27 / n.s. | 1.06 / n.s. | 2.16 / UR | 0.88 / n.s. | 0.48 / DR | 2.01 / n.s. | 1.19 / n.s. |
| SP90-3414 | 1.17 / n.s. | 0.94 / n.s. | 1.27 / n.s. | 0.99 / n.s. | 0.97 / n.s. | 1.71 / UR | 1.03 / n.s. | 2.40 / UR |

Cont.

| | **Candidate genes** | | | | | | |
|---|---|---|---|---|---|---|---|
| **Technics and Accesses** | PFK | TPI | FBA | TKT | MDH | ACC | AAT |
| HT-SuperSAGE | | | | | | | |
| Tolerant bulk | 2.67[#] / UR | 2.10[#] / UR | 2.78[#] / UR | 12.72[#] / UR | 9.61[#] / UR | 4.70[#] / UR | 6.50[#] / UR |
| Sensitive bulk | -1.02[#] / n.s. | -1.73[#] / DR | -1.52[#] / n.s. | -1.37[#] / n.s. | -1.33[#] / n.s. | -1.66[#] / DR | 1.17[#] / n.s. |
| RT-qPCR[**] | | | | | | | |
| Tolerant cultivar | | | | | | | |
| CTC6 | 3.08 / UR | 0.70 / n.s. | 1.57 / UR | 1.74 / UR | 2.46 / UR | 1.62 / UR | 2.37 / UR |
| CTC15 | 1.11 / n.s. | 0.97 / n.s. | 1.30 / n.s. | 1.23 / n.s. | 1.14 / n.s. | 1.56 / n.s. | 1.89 / UR |
| SP83-2847 | 1.13 / n.s. | 2.44 / UR | 1.76 / UR | 1.23 / n.s. | 0.83 / n.s. | 2.23 / UR | 1.64 / UR |
| SP83-5073 | 1.01 / n.s. | 0.93 / n.s. | 1.31 / n.s. | 1.83 / UR | 1.25 / n.s. | 1.70 / n.s. | 0.90 / n.s. |
| Sensitive cultivar | | | | | | | |
| CTC9 | 0.69 / n.s. | 0.87 / n.s. | 1.47 / n.s. | 1.48 / n.s | 1.09 / n.s. | 2.46 / UR | 3.40 / UR |
| CTC13 | 0.66 / DR | 0.88 / n.s. | 0.78 / n.s. | 0.80 / n.s | 0.73 / n.s | 0.51 / DR | 0.97 / n.s. |
| SP90-1638 | 2.58 / UR | 1.02 / n.s. | 1.09 / n.s. | 0.79 / n.s. | 1.52 / n.s. | 0.72 / n.s. | 1.27 / n.s. |
| SP90-3414 | 1.08 / n.s. | 0.96 / n.s. | 1.61 / UR | 1.17 / n.s. | 0.70 / n.s | 1.61 / UR | 1.42 / n.s. |

[#]Fold change [FC: ratio of frequencies (normalized to 1,000,000) observed in the stressed library in relation to the respective control library]; ****** Relative expression based on the REST software (v.2.0.13) [37] after the ΔΔCq method; UR: induced; DR: repressed. ns: not significant at $p < 0.05$ [32]. ACCD: 1-aminocyclopropane-1-carboxylate deaminase; AP: aspartic proteinase; AAT: Na+ dependent neutral amino acid transporter; TRX: thioredoxin-like 1-2, cloroplasto; PFK-2: 6-phosphofructo-2-kinase; ß-EXP: beta-expansin 8 precursor; ACC: acetyl-CoA carboxylase; E2, ubiquitin conjugating enzyme; 20SPß4, 20S proteasome beta 4 subunit; TPI, triosephosphate isomerase; FBA, fructose-bisphosphate aldolase; TKT, transketolase; C13, legumain cysteine-type endopeptidase; LDP, dihydrolipoyl dehydrogenase; MDH, malate dehydrogenase; reference genes: glyceraldehyde 3 phosphate dehydrogenase (GAPDH) and 25S ribosomal RNA (25S rRNA).

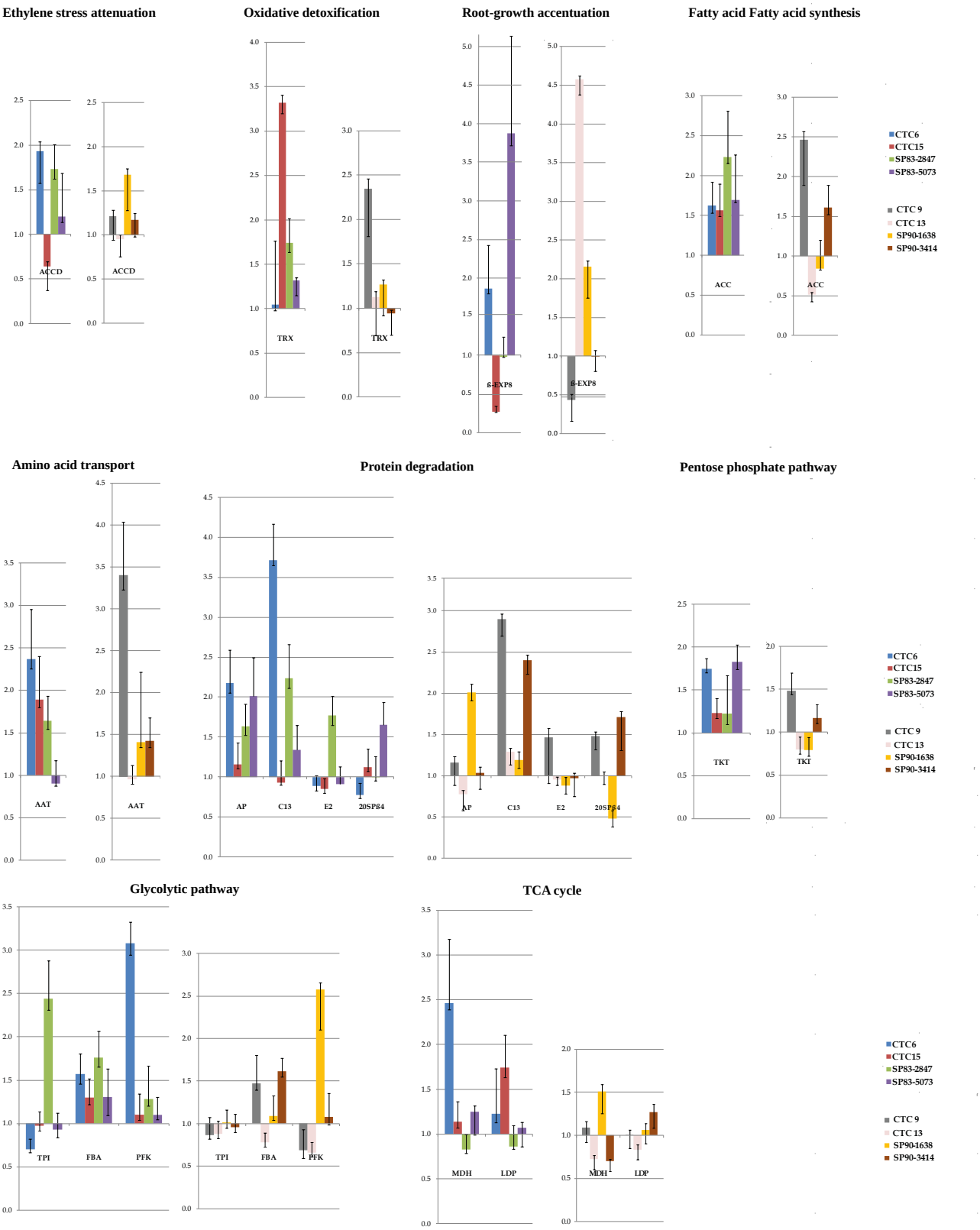



**Fig. 2.** RT-qPCR of candidate genes carried out by sugarcane root cDNAs under 24 h of continuous dehydration and negative control. The expression values were normalized by the reference genes GAPDH (Glyceraldehyde 3 phosphate dehydrogenase) and 25S rRNA (25S ribosomal RNA) and Relative expression data calculated by REST software (v.2.0.13) [37].Tolerant accesses: CTC6, CTC15, SP83-2847 and SP83-5073; Sensitive accesses CTC9, CTC13, SP90-1638 and SP90-3414). ACCD: 1-aminocyclopropane-1-carboxylate deaminase; AP: aspartic proteinase; AAT: Na+ dependent neutral amino acid transporter; TRX: thioredoxin-like 1-2, cloroplasto; PFK-2: 6-phosphofructo-2-kinase; ß-EXP: beta-expansin 8 precursor; ACC: acetyl-CoA carboxylase; E2, ubiquitin conjugating enzyme; 20SPß4, 20S proteasome beta 4 subunit; TPI, triosephosphate isomerase; FBA, fructose-bis-phosphate aldolase; TKT, transketolase; C13, legumain cysteine-type endopeptidase; LDP, dihydrolipoyl dehydrogenase; MDH, malate dehydrogenase.

# 6 ARTIGO III

## Expression Analysis of Sugarcane Aquaporin Genes under Water Deficit

Artigo científico publicado no periódico *Journal of Nucleic Acids* (ISSN: 2090-021X)

Hindawi Publishing Corporation
Journal of Nucleic Acids
Volume 2013, Article ID 763945, 14 pages
http://dx.doi.org/10.1155/2013/763945

Hindawi

*Research Article*

# Expression Analysis of Sugarcane Aquaporin Genes under Water Deficit

**Manassés Daniel da Silva,[1] Roberta Lane de Oliveira Silva,[1] José Ribamar Costa Ferreira Neto,[1] Ana Carolina Ribeiro Guimarães,[2] Daniela Truffi Veiga,[2] Sabrina Moutinho Chabregas,[2] William Lee Burnquist,[2] Günter Kahl,[3] Ana Maria Benko-Iseppon,[1] and Ederson Akio Kido[1]**

[1] *Department of Genetics, Federal University of Pernambuco (UFPE), 50670-901 Recife, PE, Brazil*
[2] *Biotechnology Division, Sugarcane Technology Center (CTC), 13400-970 Piracicaba, SP, Brazil*
[3] *Institute of Molecular Biosciences, Frankfurt University, 60438 Frankfurt am Main, Germany*

Correspondence should be addressed to Ederson Akio Kido; kido.ufpe@gmail.com

Received 30 September 2013; Accepted 12 November 2013

Academic Editor: Shigenori Iwai



The present work is a pioneer study specifically addressing the aquaporin transcripts in sugarcane transcriptomes. Representatives of the four aquaporin subfamilies (PIP, TIP, SIP, and NIP), already described for higher plants, were identified. Forty-two distinct aquaporin isoforms were expressed in four HT-SuperSAGE libraries from sugarcane roots of drought-tolerant and -sensitive genotypes, respectively. At least 10 different potential aquaporin isoform targets and their respective unitags were considered to be promising for future studies and especially for the development of molecular markers for plant breeding. From those 10 isoforms, four (*So*PIP2-4, *So*PIP2-6, *Os*PIP2-4, and *Ss*PIP1-1) showed distinct responses towards drought, with divergent expressions between the bulks from tolerant and sensitive genotypes, when they were compared under normal and stress conditions. Two targets (*Ss*PIP1-1 and *So*PIP1-3/PIP1-4) were selected for validation via RT-qPCR and their expression patterns as detected by HT-SuperSAGE were confirmed. The employed validation strategy revealed that different genotypes share the same tolerant or sensitive phenotype, respectively, but may use different routes for stress acclimation, indicating the aquaporin transcription in sugarcane to be potentially genotype-specific.

## 1. Introduction

Sugarcane (*Saccharum* spp.) is a valuable crop once it accumulates high levels of sucrose in the stems [1, 2]. In 2011, the twenty largest sugarcane producers generated about 1.7 billion tons of sucrose worldwide, valued about 52.5 billion dollars [3]. However, abiotic stresses can reduce the potential yield of these cultivated plants by 70%, with drought being the most dangerous one [4]. Water deficit, and its influence onto a variable number of morphological and functional characters in plants, eventually becomes one of the main obstacles to sustainable agricultural production worldwide [5].

The reduction of the water content in a plant cell provokes a complex network of molecular responses, involving stress perception, signal transmission in a transduction cascade and physiological, cellular, and morphological changes [6], including stomatal closure, suppression of cell growth and photosynthesis, and activation of cellular respiration. Plants under drought still respond to it and adapt by accumulating specific osmolytes and proteins for stress tolerance [7].

Genes expressed during drought can be classified into two functional groups. The first group encodes proteins that increase plant tolerance to stress, such as water channels proteins (aquaporins), proteases, and detoxification enzymes, all having a protective function. To this group belong enzymes catalyzing the biosynthesis of osmolytes, like derivatives of amino acids, sugars and various LEA (Late-Embryogenesis-Abundant) proteins. The second group of

genes encodes various proteins, such as transcription factors, kinases, phosphatases, and enzymes involved in regulatory pathways, as phospholipid metabolism and ABA biosynthesis [7]. The aquaporins or MIPs (Major Intrinsic Proteins) are proteins assembling into water channels of cell membrane and facilitate osmosis for rapid bidirectional transport of water [8]. Besides, these proteins are also involved in many plant metabolic processes, including acquisition of nutrients, cell growth, carbon fixation, cell signaling, and various stress responses [9, 10]. The aquaporins also allow permeation of small molecules such as glycerol [11], urea [12] CO2 [13], ammonia [14], boric acid [15], H2O2 [16], and even arsenic [17]. According to the phylogenetic analysis of Johanson and Gustavsson [18], plant aquaporins are classified into four main subfamilies, widely distributed among higher plants: PIPs (plasma membrane intrinsic proteins), TIPs (tonoplast intrinsic proteins), SIPs (small basic intrinsic proteins), and NIPs (26 kDa intrinsic proteins). The aquaporins are presently and extensively studied, since their importance spans from animal [19] and human physiology [20] to osmoadaptation of microorganisms [21] and vegetables [22, 23]. The transcripts encoding sugarcane aquaporins have only marginally been described, despite their significant physiological influence and participation in several processes during plant growth and acclimation against biotic and abiotic stresses [24, 25].

The present study is a first attempt to derive expression markers (functional molecular markers) from HT-SuperSAGE transcriptional profiles in contrasting sugarcane genotypes, in particular addressing specific sugarcane aquaporins, with the aim of better understanding the molecular processes occurring during drought response of the plant. HT-SuperSAGE, among all the genome-wide transcriptome profiling techniques was chosen for its efficiency to generate highly reliable transcription profiles. The increase in the size of the tag to 26 bp, the characteristic of SuperSAGE, drastically improves the annotation of the tag to the corresponding gene [28], allowing to establish genome-wide gene expression profiles of two or more species in one sample (e.g., host-parasite interactions [29, 30]). Besides, SuperSAGE combined high-throughput next-generation sequencing [31, 32], designated DeepSuperSAGE or HT-SuperSAGE, provides even more informations (three to four orders of magnitude) at relatively low cost compared to traditional Sanger sequencing.

## 2. Methodology

*2.1. Unitags Annotation, GO Categorization of ESTs, and Aquaporin Isoforms Identification.* Bioinformatics analyses covered the 8,787,315 tags (26 bp) described by Kido et al. [33] from four root HT-SuperSAGE libraries [SD24T (the bulk of the tolerant genotypes CTC6, CTC15, SP83-2847, SP83-5073, under stress (24 h of continuous dehydration), totalizing 2,542,552 tags); SDTC (the tolerant bulk with daily irrigation, comprising 1,909,543 tags); SD24S (the sensitive bulk of stressed genotypes CTC9, CTC13, SP90-3414 and SP90-1638, with 2,170,998 tags) and SDSC (the bulk of sensitive genotypes without stress, covering 2,164,222 tags)]. After singlet exclusion (tags sequenced only once per library), the unique tags (unitags) were classified as up- (UR) or down-regulated (DR), based on the Audic and Claverie test ($P < 0.05$; [34]), using the DiscoverySpace 4.0 software [35]. The unitag frequencies normalized to a million per library allowed the evaluation of the unitag expression modulation by fold change values (FC) comparing two frequencies. The unitags were aligned by BLASTn with expressed sequence tags (ESTs) from nine public databases, comprising sugarcane ESTs from NCBI (http://www.ncbi.nlm.nih.gov/nucest), and grass ESTs (Poaceae family) from Gene Index (http://compbio.dfci.harvard.edu/tgi/plant.html), including *Saccharum officinarum* (SoGI 3.0), *Sorghum bicolor* (SbGI 9.0), *Zea mays* (ZmGI 19.0), *Oryza sativa* (OsGI 18.0), *Panicum virgatum* (PaviGI 1.0), *Triticum aestivum* (TaGI 12.0), *Hordeum vulgare* (HvGI 11.0), and *Festuca arundinacea* (FaGI 3.0). Only BLASTn alignments (*e value* < 0.0001) with scores 42 to 52 (100% identity), plus/plus orientation and a preserved 5′CATG were accepted, and the best tag - hit was selected prioritizing sugarcane sequences or sequences from closely related species with adequate annotation. ESTs anchoring unitags were then categorized via Gene Ontology (GO; http://www.geneontology.org/GO.doc.shtml), using the Blast2GO tool [36].

Potential ESTs from the MIP gene superfamily were identified using the keywords "aquaporin," "major intrinsic protein," "PIP," "TIP," "PIN," "SIP," "plasma membrane intrinsic protein" "tonoplast intrinsic protein," "nodulin-26-like intrinsic protein" and "small basic intrinsic protein" in the EST annotations, or "water transporter" in the GO terms. These ESTs were classified into the plant aquaporin subfamilies (TIP, NIP, SIP, PIP) and analyzed with the NCBI Conserved Domain Search tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) in an effort to confirm their conserved domains. Also ESTs were BLASTx aligned with proteins sequences from the UniProtKB/Swiss-Prot database (http://www.uniprot.org/help/uniprotkb), trying to confirm the isoform identity by using curated sequences (*e*-value cutoff $e^{-10}$).

*2.2. Comparative and Phylogenetic Analysis of the Putative Aquaporin Isoforms Based on Unitag Expressions.* The predicted peptides from ESTs related to the tags after translation with the ORF finder tool (http://www.ncbi.nlm.nih.gov/projects/gorf/) and BLASTp analysis were subsequently aligned via Clustal W [37]. A dendrogram of the aligned sequences was generated using the MEGA v.5.2.1 software [27], according to the following parameters: Neighbor Joining tree method, pairwise deletion option, and 1000 bootstrap replicates. For a better assignment of the isoforms into the aquaporin subfamilies, 15 *A. thaliana* aquaporin protein sequences were included in addition to the predicted proteins. Also, two sequences served as outgroups, one from humans and one from *Yersinia pseudotuberculosis*. In addition to the phylogenic tree, a heat map was established based on fold changes of the unitags responding to the applied stress.

*2.3. cDNA Synthesis, Primer Design, and RT-qPCR Analysis.* The RNA of each genotype sample constituting the bulks (tolerant stressed, tolerant control, sensitive stressed, and sensitive control) was isolated from sugarcane roots using the RNAeasy Plus Micro Kit (Qiagen) and DNase treatment. The cDNA synthesis reaction was performed with the SuperScriptTM First Strand kit Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's protocol, using 1 g of RNA quantified by the Qubit fluorometer (Invitrogen). Primer pairs were designed from ESTs anchoring unitags, tolerating a maximum of one mismatch, and using the default parameters of the *Primer3* software [38] with some minor modifications [amplicon size: 70 to 200 bp; primer Tm: 40 to 60°C; GC content: 45 to 55%]. These ESTs anchored unitags differentially regulated in the tolerant and sensitive genotype bulks. Prior to the validation of unitags by RT-qPCR, standard curves using a dilution series of the cDNA pool were made to calculate the gene-specific PCR efficiency and regression coefficient ($R^2$) for each gene (Table 6). The RT-qPCR amplifications were performed on the LineGene 9660 model (Bioer), using SYBR Green detection. Each reaction mixture comprised 1 $\mu$L of template cDNA (diluted 5-fold), 5 $\mu$L of *HotStart-IT SYBR Green qPCR Master Mix 2x* (USB), 0.05 $\mu$L of ROX, 1.95 $\mu$L of water, and 1 $\mu$L primer (500 nM each) to a final volume of 10 $\mu$L. The reactions were denatured at 95°C for 2 min, followed by 40 cycles of 95°C for 15 s, 58°C for 30 s, and 72°C for 15 s in 96-well reaction plates, with the detection of the fluorescence signal at the end of each extension step. The melting curves were analyzed at 65–95°C for 20 min after 40 cycles. Three biological replicates and three technical replicates were used for RT-qPCR analysis. The relative quantification data were analyzed with the REST© v.2.0.13 software [26].

## 3. Results and Discussion

*3.1. Aquaporin Data-Mining and Categorization of EST Anchoring Unitags.* The universe of 8,787,315 tags (26 bp) generated from the four HT-SuperSAGE libraries presented 205,975 unique tags (unitags) [33], from which 289 anchored in 484 putative aquaporin ESTs, distributed in nine different databases (details in Table 1), totaling 1,579 BLASTn results with scores of 42 to 52 (100% identity). This set of 484 ESTs anchoring unitags (Table 1) represented the basis for the analysis of transcript profiles based on the respective unitags.

The keyword ("aquaporin," "major intrinsic protein," "tonoplast intrinsic protein," "plasma membrane intrinsic protein," "small basic intrinsic protein," "nodulin-26-like intrinsic protein," "PIP," "TIP," "NIP," "SIP") searches in the EST annotation identified 1,347 ESTs, while the "water transport" GO expression identified 342 ESTs (Figure 1). The searches in the GO terms increased the aquaporin identifications by almost 15%, representing 230 alignments of the total (1,579; Figure 1).

The unitag annotation efficiency relied on the used EST database. As mentioned by Kido et al. [39], Gene Index is a good source for unitag annotation, as it displays adequate gene or protein function descriptions. In the present case,

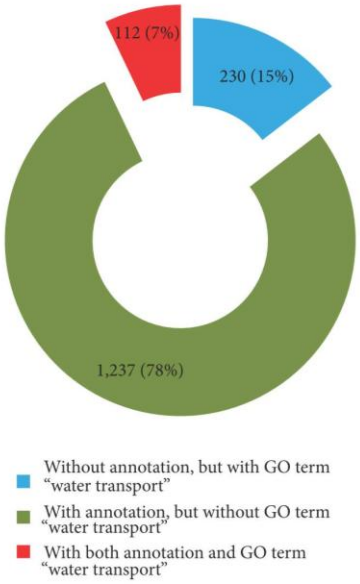



Figure 1: Percentage of HT-SuperSAGE unitags from sugarcane plants (24 h of continuous dehydration or regular daily irrigation) identified as potential aquaporins after keyword searches in the EST annotation ("*aquaporin*," "*tonoplast intrinsic protein*," "PIP," "TIP," "NIP," "SIP") or in the GO terms ("*water transport*"). Total of unitags: 1,579.

the SoGI (*Saccharum officinarum* L.) dataset representing 282,683 ESTs that resulted in 121,342 unique sequences [42,377 TC (Tentative Consensus clusters) plus 78,965 singletons) after clustering. This species took part in the sugarcane breeding programs [40, 41] performed around the world.

Almost all unitags related to the aquaporin annotations (260 of 289; Table 1) anchored in SoGI sequences, which exceeded those obtained with the partial dbEST dataset (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) by almost six times. This dbEST is composed of 277,266 ESTs, mostly from Brazilian sugarcane hybrids (SUCEST-FAPESP project) [42] (Table 1). This may mainly be due to the sizes of the SoGI sequences, as most alignments occurred in the CTs. Nevertheless, the partial dbEST dataset was the second best source for mapping unitags, but its real annotation power was affected by the nonadequate descriptions of the cDNAs (many "unknown" hits). As Sorghum bicolor is the most closely related diploid of *S. officinarum* [2], this species could contribute to the identification of aquaporin isoforms. However, after redundancy exclusion, only 19 unitags anchored in seven unique ESTs based on the best hits (Table 1). This poor performance may be explained by the low number of ESTs available in the SbGI dataset (46,043), being the second smallest databank used (Table 1), in contrast with the high number of available sugarcane ESTs, reinforced by the high homology between sorghum and sugarcane.

Table 1: EST databases used for unitag annotations from sugarcane HT-SuperSAGE libraries (tolerant/sensitive genotypes after 24 h of continuous dehydration or regular daily irrigation) and BLASTn nonredundant results.

| Species | Database | Annot.*/ESTs | PIP | TIP | NIP | SIP | Total BlastN Align. | Score 52 | Align. (up to one mismatch) | Unitags | Unique ESTs+ | Unitags with GO[1] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Saccharum* spp. | dbEST | 24/256,636 | 0 | 0 | 0 | 0 | 154 | 41 | 105 | 43 | 34 | 23 |
| *Saccharum officinarum* | SoGI | 265/121,342 | 157 | 85 | 21 | 11 | 683 | 158 | 425 | 260 | 127 | 50 |
| *Sorghum bicolor* | SbGI | 68/46,043 | 34 | 23 | 5 | 6 | 74 | 11 | 63 | 19 | 7 | 74 |
| *Zea mays* | ZmGI | 347/315,134 | 142 | 154 | 35 | 19 | 49 | 7 | 29 | 13 | 15 | 20 |
| *Oryza sativa* | OsGI | 283/201,220 | 121 | 95 | 33 | 7 | 119 | 26 | 72 | 20 | 86 | 3 |
| *Panicum virgatum* | PaviGI | 147/85,244 | 58 | 45 | 28 | 6 | 174 | 31 | 116 | 22 | 57 | 0 |
| *Triticum aestivum* | TaGI | 542/222,152 | 138 | 128 | 32 | 14 | 253 | 20 | 197 | 23 | 123 | 5 |
| *Hordeum vulgare* | HvGI | 110/83,101 | 32 | 33 | 12 | 9 | 34 | 2 | 30 | 16 | 21 | 7 |
| *Festuca arundinacea* | FaGI | 27/30,244 | 5 | 1 | 1 | 1 | 39 | 10 | 27 | 17 | 14 | 5 |
| Total | | 1,913/1,361,144 | 687 | 564 | 167 | 73 | 1,579 | 306 | 1,064 | 289# | 484 | 45# |

* Aquaporin, tonoplast intrinsic protein and major intrinsic protein, membrane integral protein (PIP, TIP, NIP e SIP); + number of nonredundant ESTs (putative aquaporins) anchoring unitags, [1]"water transporter"; # number without redundancy among the nine databanks.

The unitags proved to be highly specific for aquaporins. A total of 263 unitags (91% of 289) were associated with aquaporin isoforms (189 unitags anchored in just a single EST from a unique database); 19 unitags (7%) were not isoform-specific but comprised the same subclass (PIP1 or PIP2) and only seven (2%) were not specific to any subclass.

*3.2. Comparative Analysis of the Putative Aquaporin Isoforms Identified by the Unitags.* Regarding the total of 484 annotated aquaporin-ESTs anchoring unitags (Table 1), 470 of them (97.10%) aligned with aquaporin-proteins isoforms with an $e$ value $< 10^{-20}$ (BLASTx), from the UniProtKB/Swiss-Prot database, a high-quality annotated databank. This fact confirmed the isoforms identities (data not shown). All of them represented 42 distinct isoforms belonging to the four subfamilies (PIP, TIP, SIP, and NIP) based on the ESTs annotations. The 42 isoforms and their respective number of unitags [considered UR, DR or n.s. ($P < 0.05$)] in the two main HT-SuperSAGE libraries contrasts are shown in Table 2. According to this table, one unitag or more could be associated with a specific isoform. In some cases, two or more ESTs from one database present the same isoform annotation. The Gene Index databases used throughout this work minimized this situation due to the assembled TC (Tentative Consensus) clusters. Besides, unitags aligning more than one locus in the same EST could be resulted by partial *Nla*III digestions. In an attempt to avoid this situation, it was performed double digestions. Additionally, this event could be resulted by sister-tags anchoring one specific EST and isoform. In this case, tags showing a single base substitution (sister-tags) were considered as two different unitags. On the other hand, alternative transcripts could anchor varied unitags. Also, specific isoforms could be mapped in several loci (in the same or in different chromosomes). In addition, *Saccharum* hybrids show complex genomes, as a result of polyploidy and aneuploidy events [40, 41]. Therefore, this diversity of unitags (UR and DR) associated to aquaporin isoforms could allow identify biotechnologically interesting candidates.

From the 71 unique ESTs involved in perfect unitag-EST BlastN alignments (score 52; Table 1), 24 putative aquaporins showed ORFs with over 180 amino acids in size, and these sequences, together with MIP protein sequences from *Arabidopsis thaliana*, *Homo sapiens*, and *Yersinia pseudotuberculosis* were compared in a phenetic analysis. The resulting tree confirmed that putative aquaporins clearly divide into four major clusters, representing the PIP, TIP, SIP, and NIP subfamilies (Figure 2). This tree was consistent with a previous analysis of aquaporin phylogeny in higher plants [10, 18, 43, 44]. As expected, the human HsAPQ1 isoform grouped with the PIP subfamily, since the human APQ1 subfamily was recently recognized to be phylogenetically more similar to the PIP subfamily than to other plant subfamilies [45]. Also, YpGIpF grouped with the NIP aquaporin subfamily. The YpGIpF isoform belongs to a MIP family related to the bacterial GlpF protein glycerol uptake facilitator, classically associated with aquaglyceroporins from NIP and APQ3 subfamilies [45]. Therefore, this tree, which is supported by the scientific literature, presents the 24 aquaporin isoforms identified by HT-SuperSAGE unitags expressed after 24 h of continuous dehydration stress. Moreover, considering only the nine isoforms identified from *S. officinarum* ESTs, this smaller set was also distributed across the four aquaporin subfamilies described for higher plants.

Additionally, the heat maps (Figure 2) revealed by the expression modulation of the unitags (FCs) in the tolerant or sensitive bulks (both with their respective unstressed controls) show that some PIP isoforms are divergently regulated in the bulks of genotypes. Thus, from the 12 PIP transcripts, eight were repressed in the tolerant genotypes under stress. At the same time, eight of those transcripts were induced in the sensitive bulk. Furthermore, the majority of the PIP transcripts showed divergent modulations (contrasting

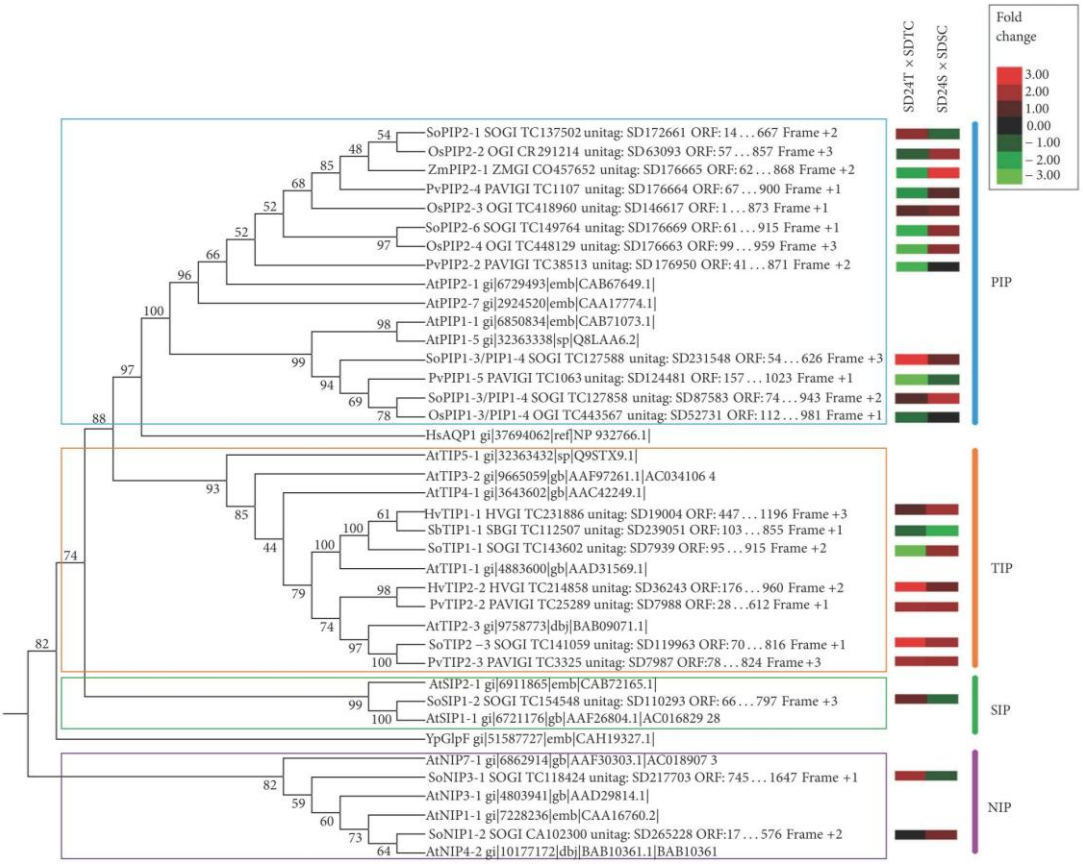



FIGURE 2: Neighbor Joining dendrogram (MEGA v.5.2.1 software [27]) and integrated heat map (bootstrap values of 1,000 replications), showing the phenetic grouping of 24 potential aquaporin amino acid sequences identified by HT-SuperSAGE unitags from sugarcane accessions after 24 h of continuous dehydration (with their respective EST and unitag identifiers), and aquaporins sequences of *Arabidopsis thaliana*, human and *Yersinia pseudotuberculosis* (all these labeled by an asterisk). Nomenclature: Isoforms are preceded by the abbreviated species name (*At-Arabidopsis thaliana; Hs-Homo sapiens; Hv-Hordeum vulgare; Os-Oryza sativa; Pv-Panicum virgatum; Sb-Sorghum bicolor; So-Saccharum officinarum; Yp-Yersinia pseudotuberculosis; Zm-Zea mays*).

results) when the response of both bulks of stressed genotypes is compared. Taken together, different genotypes may have developed different survival strategies.

On the other hand, the TIP transcripts similarly responded to the stress (comparing the modulation between both bulks of genotypes, Figure 2). Of the seven TIP subfamily isoforms studied (Figure 2), five were induced in both bulks of genotypes responding to stress, suggesting the participation of these isoforms in water transport. Finally, the only SIP subfamily representative studied here showed distinct regulation between the analyzed bulks, whereas the two NIP subfamily representatives distinctly responded: NIP3-1 was induced in the tolerant bulk and suppressed in its sensitive counterpart, while NIP3-2 was not modulated in the tolerant bulk, but was induced in the sensitive bulk (Figure 2). For these subfamilies, a larger amount of data is required for further analysis.

*3.3. Transcriptional Profile of Putative Aquaporins Based on Unitags.* The 30 most expressed unitags, based on their normalized frequencies (tpm) in the HT-SuperSAGE libraries, associated to the aquaporin subfamilies PIP (15), TIP (10), SIP (3) and NIP (2) are displayed in Table 3. According to Kjellbom et al. [46] many aquaporin genes are constitutively expressed, with a large number of transcripts (as presented in Table 3), while others are temporally and spatially regulated during plant development or stress responses, as is, for example, the case with unitag SD173276 (Table 3).

After necessary redundancy exclusion, we identified 42 potential aquaporin isoforms. The contribution of each aquaporin subfamily is presented in Figure 3. In each comparison (SD24T vs SDTC: 26; SD24S vs SDSC: 28; SD24T vs SD24S: 28; SDTC vs SDSC: 28), 26–28 isoforms were identified as being expressed in sugarcane roots after onset of stress (24 h of continuous dehydration) or under normal daily irrigation

Table 2: Putative sugarcane aquaporin isoforms (42) based on unitags of root dehydration (24 h) observed in the two main contrasts of HT-SuperSAGE libraries.

| Aquaporin isoform | SD24T versus SDTC | | | SD24S versus SDSC | | |
|---|---|---|---|---|---|---|
| | UR | DR | n.s. | UR | DR | n.s. |
| *Ss*PIP1-1 | 1 | — | 2 | — | 1 | 1 |
| *So*PIP1-2 | 1 | 2 | 3 | 2 | 1 | — |
| *Pv*PIP1-2 | — | — | 1 | — | — | 1 |
| *So*PIP1-3/PIP1-4 | 3 | 24 | 26 | 7 | 7 | 25 |
| *Os*PIP1-3/PIP1-4 | 1 | — | 2 | — | 1 | — |
| *So*PIP1-5 | 1 | 2 | 2 | 3 | — | 2 |
| *Pv*PIP1-5 | — | — | 2 | — | — | 1 |
| *So*PIP2-1 | 2 | 14 | 23 | 4 | 24 | 19 |
| *Zm*PIP2-1 | — | 2 | 4 | 1 | — | 4 |
| *Os*PIP2-1 | 1 | — | 2 | — | 1 | 2 |
| *Pv*PIP2-1 | — | 1 | 9 | 1 | — | 6 |
| *Ta*PIP2-1 | — | 2 | 4 | 1 | — | 2 |
| *Os*PIP2-2 | — | — | 2 | — | — | 3 |
| *Pv*PIP2-2 | — | 2 | 6 | 1 | — | 4 |
| *Os*PIP2-3 | — | — | 1 | — | — | 1 |
| *Zm*PIP2-3 | — | — | 1 | — | — | 1 |
| *So*PIP2-4 | 3 | 4 | 6 | 2 | 3 | 6 |
| *Os*PIP2-4 | — | 2 | — | 1 | 2 | — |
| *Pv*PIP2-4 | — | 4 | 2 | 1 | — | 5 |
| *Fa*PIP2-4 | — | 3 | 7 | 3 | — | 3 |
| *So*PIP2-5 | 1 | 1 | — | — | — | 2 |
| *Ta*PIP2-5 | — | 1 | 4 | 1 | — | 2 |
| *So*PIP2-6 | 2 | — | 2 | 2 | — | 2 |
| *Ta*PIP2-6 | — | 1 | 3 | 1 | — | 2 |
| *Ss*TIP1-1 | 4 | 9 | 14 | 6 | 5 | 18 |
| *So*TIP1-1 | 3 | 7 | 10 | 6 | 7 | 11 |
| *Sb*TIP1-1 | 1 | — | 1 | — | 1 | 1 |
| *Zm*TIP1-1 | — | — | 1 | — | — | — |
| *Hv*TIP1-1 | 1 | — | — | — | 1 | — |
| *Ta*TIP1-1 | — | 1 | 4 | — | 1 | 4 |
| *Ta*TIP1-2 | — | 1 | — | — | — | 1 |
| *So*TIP2-2 | 2 | 9 | 9 | 1 | 8 | 11 |
| *Hv*TIP2-2 | 1 | — | 1 | — | — | 1 |
| *Pv*TIP2-2 | — | — | 2 | — | — | 2 |
| *So*TIP2-3 | 5 | 5 | 10 | 3 | 3 | 5 |
| *Pv*TIP2-3 | — | — | 3 | — | — | 3 |
| *So*TIP4-2 | — | — | 1 | — | — | 1 |
| *Pv*NIP1-1 | — | 1 | — | — | 1 | 1 |
| *So*NIP1-2 | — | — | 1 | — | — | 1 |
| *So*NIP3-1 | 1 | 2 | 12 | — | 1 | 10 |
| *So*SIP1-1 | 1 | — | 2 | 1 | — | 1 |
| *So*SIP1-2 | 3 | 5 | 2 | 2 | 1 | 10 |

Isoforms are preceded by the abbreviated species name (*Fa*: *Festuca arundinacea*; *Hv*: *Hordeum vulgare*; *Os*: *Oryza sativa*; *Pv*: *Panicum virgatum*; *Sb*: *Sorghum bicolor*; *So*: *Saccharum officinarum*; *Ss*: *Saccharum* spp.; *Ta*: *Triticum aestivum*; *Zm*: *Zea mays*). HT-SuperSAGE libraries: SD24T (bulk of tolerant genotypes under stress; SD24S (bulk of sensitive genotypes under stress); SDTC (control bulk of tolerant genotypes); SDSC (control bulk of sensitive genotypes). DR: down-regulated; UR: up-regulated. n.s.: not significant at $P < 0.05$.

conditions. This number of isoforms is close to that of other higher plants (31 aquaporin isoforms in maize [43, 47], 35 in *A. thaliana* [48], 39 in rice [47]), and more than twice the amount predicted for vertebrates (11 to 13 isoforms) [49, 50]. The number of aquaporin isoforms in sugarcane may be even higher, since some isoforms respond only in specific tissues [51] or after - salinity [52], freezing [53], mycorrhization [54], light [55, 56], and cell growth stresses [10]. Therefore, the real number of aquaposin isoforms can only be estimated approximatively by whole genome sequencing. Since sugarcane has one of the most complex genomes of the plant kingdom, with a diploid number of chromosomes ranging from 100 to 130 as a result of aneuploidy and polyploidy events [41], this approach would require significant efforts and investments.

Table 3: The 30 HT-SuperSAGE unitags most expressed and annotated as aquaporins from sugarcane libraries with contrasting genotypes under stress (24 h of continuous dehydration) or normal daily irrigation.

| Unitag id | Aquaporin | Tags per million (tpm) | | | |
|---|---|---|---|---|---|
| | | SD24T | SDTC | SD24S | SDSC |
| SD173282 | *So*TIP2-2 | 1,096 | 3,784 | 1,643 | 1,816 |
| SD231437 | *So*PIP1-3/PIP1-4 | 819 | 1,551 | 1,140 | 990 |
| SD87583 | *So*PIP1-3/PIP1-4 | 964 | 956 | 1,186 | 520 |
| SD119746 | *So*TIP2-3 | 1,162 | 1,041 | 771 | 530 |
| SD173276 | *So*TIP2-2 | 564 | 1,879 | 496 | 0 |
| SD182865 | *So*TIP2-2 | 535 | 876 | 501 | 674 |
| SD87593 | *So*PIP1-3 /PIP1-4 | 579 | 601 | 750 | 377 |
| SD80613 | *So*TIP1-1 | 318 | 234 | 775 | 571 |
| SD80612 | *So*TIP1-1 | 437 | 393 | 423 | 453 |
| SD250744 | *So*PIP2-1 | 321 | 577 | 265 | 508 |
| SD19004 | *Hv*TIP1-1 | 422 | 395 | 411 | 210 |
| SD176669 | *So*PIP2-6 | 275 | 591 | 310 | 183 |
| SD243880 | *So*PIP2-1 | 176 | 326 | 334 | 496 |
| SD28080 | *So*PIP2-4 | 406 | 329 | 224 | 312 |
| SD176663 | *So*PIP2-6 | 135 | 340 | 227 | 136 |
| SD241279 | *So*PIP1-5 | 151 | 184 | 257 | 107 |
| SD84960 | *Ss*TIP1-1 | 202 | 216 | 29 | 136 |
| SD243849 | SoPIP2-1 | 108 | 123 | 113 | 40 |
| SD54852 | *So*PIP2-1 | 107 | 143 | 22 | 52 |
| SD96918 | *So*SIP1-2 | 106 | 66 | 96 | 49 |
| SD96922 | *So*SIP1-2 | 89 | 43 | 114 | 49 |
| SD202395 | *Pv*NIP1-1 | 39 | 53 | 40 | 98 |
| SD243867 | *So*PIP2-1 | 37 | 42 | 79 | 43 |
| SD87586 | *So*PIP1-3/PIP1-4 | 31 | 54 | 28 | 55 |
| SD250859 | *So*PIP2-1 | 39 | 96 | 0 | 32 |
| SD198883 | *So*PIP2-5 | 39 | 27 | 30 | 37 |
| SD21811 | *So*SIP1-1 | 26 | 16 | 26 | 26 |
| SD84958 | *Ss*TIP1-1 | 29 | 15 | 9 | 17 |
| SD217703 | *So*NIP3-1 | 29 | 15 | 13 | 13 |
| SD36243 | *Hv*TIP2-2 | 31 | 9 | 12 | 9 |

Isoforms are preceded by the abbreviated species name (*Hv*: *Hordeum vulgare*; *Pv*: *Panicum virgatum*; *So*: *Saccharum officinarum*; *Ss*: *Saccharum* spp). SuperSAGE libraries: SD24T (bulk of tolerant genotypes under stress; SD24S (bulk of sensitive genotypes under stress); SDTC (control bulk of tolerant genotypes); SDSC (control bulk of sensitive genotypes).

The most transcribed 19 aquaporin unitags belonged to the PIP and TIP subfamilies (Table 3), which matches a report by Alexandersson et al. [51], who analyzed the transcriptional

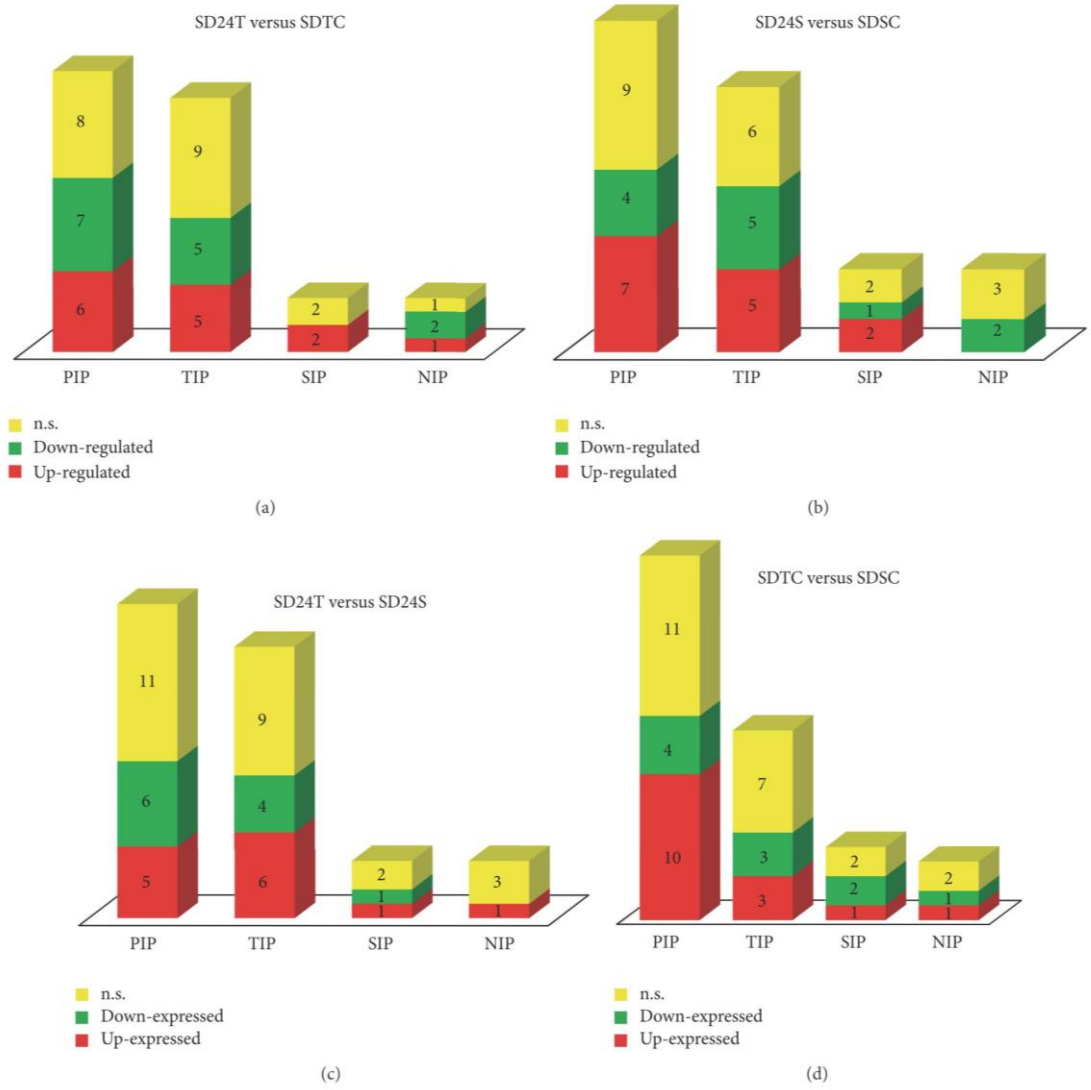



FIGURE 3: Representation of aquaporin subfamilies expressed in sugarcane after 24 h of continuous dehydration ((a), (b), and (c)) and during normal irrigation (d), involving bulks of genotypes tolerant and sensitive to stress. HT-SuperSAGE libraries: SD24T (bulk of tolerant genotypes under stress; SD24S (bulk of sensitive genotypes under stress); SDTC (control bulk of tolerant genotypes); SDSC (control bulk of sensitive genotypes).

profile of 35 *Arabidopsis* aquaporins in three different tissues (roots, leaves, and flowers) during water deficit stress (watering suppression). These authors concluded that in all the studied tissues, the PIP, and TIP aquaporins showed higher expression levels, whereas NIPs aquaporins exhibited particularly low transcriptional levels under stress. Zhu et al. [52] also confirmed a lower amount of NIP and SIP in corn under controlled conditions (continuously aerated hydroponic medium, and parameters described by Gibeaut et al. [57]), as compared to the PIP and TIP, which could be related to the aquaporin transport specificity [58]. NIPs are related to the transport of small solutes [31], whereas the physiological functions of SIPs, in addition to water transport [59], still remain unclear. Otherwise, PIPs form primary channels mediating efficient water uptake and thereby control plasma membrane potentials of permeability, while TIPs, in addition to their high water transport capacity in tonoplasts [60], also transport CO2 [13] and urea [12].

In the present work, SIP and NIP subfamilies were less responsive to the applied stress. We noticed, that NIPs were not up-regulated in the sensitive bulks t SD24S versus SDSC (Figure 3(b)), as well as among the down-regulated unitags

in the tolerant bulks SD24T versus SD24S (Figure 3(c)). On the other hand, the SIP subfamily also harbored no isoform among the down-regulated unitags in the contrast SD24T versus SDTC (Figure 3(a)). Alexandersson et al. [51] also confirmed that some SIP isoforms presented little expression variation in *Arabidopsis* plants under drought stress (watering suppression extended until 12 days). Therefore, AtSIP1-1 was considered as constitutively expressed. This low responsiveness to water deficit can be explained by the unique location of these aquaporins in the endoplasmic reticulum [59], an organelle with tortuous structure and high surface-to-volume ratio with high demand for osmotic balance volume and, therefore, may not require the water transport mediated by aquaporins [59]. Thus, further studies are necessary to define SIP functions more clearly.

The level of aquaporin transcripts varied less than 10 times based on the unitags in the contrasts, except for the SD173276 unitag (a potential SoTIP2-2), which was almost two thousand times higher in the tolerant SDTC versus SDSC contrast, and almost 500 times in the sensitive contrast SD24S versus SDSC (data not shown).

Alexandersson et al. [51] also observed that most aquaporins do not alter their expression under water deficit stress, and no *Arabidopsis* aquaporin isoforms varied their expression more than twice until the seventh day of stress treatment.

The contrast analysis of the tolerant bulks defined four possible targets: one exclusively up-regulated PIP1-1 isoform (SD264077 unitag, FC 3.58), and three exclusively down-regulated PIP2-2s (SD176950 unitag, FC -2.34), PIP2-1 (SD176664 unitag, FC -1.73) and NIP1-1 (SD202395 unitag, FC -1.36).

In rice, the PIP1-1 isoform promoted salt stress tolerance [23], and it was involved in the rehydration after cooling stress in tolerant genotypes [61]. PIP1-1 overexpression conferred tolerance to water deficit in rice and to salt stress in transgenic *Arabidopsis* [62]. This isoform also responded to drought and daytime in grapevine [56]. The up-regulated SD264077 unitag, as a possible PIP1-1 isoform, was validated in the present work by RT-qPCR analysis, as detailed in the next chapter, and represents a potential target for further studies, including the development of molecular markers for marker-assisted selection in breeding (real-time PCR-assisted selection) [63] or cis-genesis (insertion of genes in different accessions of the same species [64]), already successfully applied by Joshi et al. [65]. These authors inserted resistance genes to apple scab under the control of the RubisCO promoter in varieties known to be susceptible to the pathogen.

Isoform PIP2-2 is down-regulated over four times in *Arabidopsis* under 12 days of drought [51] and in barley under salt stress [66]. As expected, it was also observed in the tolerant bulk analysis of the present study, showing FC -2.34. The subsequent PIP2-2 monitoring revealed that aquaporin expression increased sensitivity to salt stress in transgenic rice [67]. This point is relevant, since crosstalks involving shared pathways in response to drought and salinity stress are regular [68, 69]. Thus, this isoform, after appropriate RT-qPCR validation, could be useful as stress-indicator in breeding programs.

On the other hand, the potential usefulness of unitags related to PIP2-1, PIP2-5, PIP2-6, TIP1-1, TIP2-2, SIP1-1, and SIP1-2 in breeding programs still need to be confirmed. In relation to the PIP2-5 isoform (up- and down-regulated simultaneously in the present study, depending on unitags), Jang et al. [70] observed that overexpressing this aquaporin reduced drought tolerance of transgenic *Arabidopsis* and tobacco. The same group proposed that PIP2-5 expression influenced the transcription levels of other PIPs and H+-ATPases (enzymes that regulate the cytoplasmic pH in which levels of H+ interfere with the control of the opening and closing of the aquaporins channels known as the aquaporin gating [71]). Lembke et al. [25] also observed this isoform to be down-regulated under water deficit (72 hours of watering suppression), despite the detected induction via oligonucleotide array hybridization.

Therefore, for the tolerant bulk of genotypes, this isoform is expected to restrain its expression under root dehydration (24 h).

Basically, up- or down-regulation and constitutive expression were all observed in the contrast analysis of tolerant bulks, (except that down-regulation was not observed in the SIP subfamily; Figure 3(a)).

The sensitive bulk of genotypes also presented all three expression levels for each aquaporin subfamily (with the exception of the NIP subfamily, in which up-regulation was not observed; Figure 3(b)). The analysis of the sensitive bulks allowed the identification of only up-regulated [PIP2-2 (SD176665 unitag, FC 4.83), PIP2-4 (SD176663 unitag, FC 1.66), PIP2-6 (SD176669 unitag, FC 1.70)] or only down-regulated [PIP1-1 (SD264077 unitag, FC -4.56), NIP1-1 (SD202395 unitag, FC -2.00)] aquaporin isoforms. These exclusively up- or down-regulated isoforms, respectively, may represent a panel of markers based on real-time PCR, and suggesting high stress sensitivity. In this way, at least two candidates are particularly appealing: (a) PIP1-1, that was up-regulated (SD264077 unitag, FC 3.58) in the tolerant bulks and entirely differently regulated (SD264077 unitag, FC -4.56) in the sensitive bulks; (b) PIP2-2 isoform, which was exclusively down-regulated in the tolerant bulks (SD176950 unitag, FC -2.34) and up-regulated in the sensitive bulks (SD176665 unitag, FC 4.83). Thus, both isoforms are strong candidates for further research aiming at molecular marker development and cis-genesis. Finally, further studies are needed to determine the true meaning of each stress-responsive isoform.

When comparing both genotype bulks under stress (SD24T versus SD24S), all three expression levels (up- or down-regulation and constitutive expression) were observed for each aquaporin isoform subfamily. Notably, a specific isoform in the SIP subfamily was down-regulated in the tolerant bulk, but not in the sensitive bulk (Figure 3(c)). Two more isoforms are worth mentioning: PIP2-4 (SD176664 unitag, FC 3.64), which was more transcribed in the stressed tolerant bulk than in the stressed sensitive one, and PIP2-1 (SD176669 unitag, FC -1.13), which is being less transcribed

Table 4: Aquaporin unitags with distinct expression rates in root HT-SuperSAGElibraries from contrasts of tolerant$^{T}$ (SD24T versus SDTC) and sensitive$^{S}$ (SD24S versus SDSC) sugarcane genotypes and after continuous dehydration (24 h).

| Unitag | Anotation | FC$^{T}$ | Unitag regulation$^{T}$ | FC$^{S}$ | Unitag regulation$^{S}$ |
|---|---|---|---|---|---|
| SD264077 | *Ss*PIP1-1 | 3.58 | UR | −4.56 | DR |
| SD2444 | *So*PIP1-3/PIP1-4 | 5.93 | UR | −2.66 | DR |
| SD231548 | *So*PIP1-3/PIP1-4 | 3.18 | UR | 1.28 | n.s. |
| SD243866 | *So*PIP2-1 | 13.51 | UR | −15.97 | DR |
| SD243874 | *So*PIP2-1 | 3.18 | UR | −2.14 | DR |
| SD28082 | *So*PIP2-4 | 15.10 | UR | −6.39 | DR |
| SD28080 | *So*PIP2-4 | 1.23 | UR | −1.39 | n.s. |
| SD198883 | *So*PIP2-5 | 1.46 | UR | −1.25 | n.s. |
| SD36536 | *So*TIP1-1 | 2.97 | UR | −1.90 | n.s. |
| SD80612 | *So*TIP1-1 | 1.11 | UR | −1.07 | n.s. |
| SD84958 | *Ss*TIP1-1 | 2.02 | UR | −1.82 | DR |
| SD36243 | *Hv*TIP2-2 | 3.53 | UR | 1.34 | n.s. |
| SD182891 | *So*TIP2-2 | 2.78 | UR | −5.85 | DR |
| SD119963 | *So*TIP2-3 | 5.56 | UR | 1.93 | n.s. |
| SD119859 | *So*TIP2-3 | 3.18 | UR | −2.66 | DR |
| SD217703 | *So*NIP3-1 | 1.93 | UR | −1.01 | n.s. |
| SD21811 | *So*SIP1-1 | 1.65 | UR | 1.01 | n.s. |
| SD96919 | *So*SIP1-2 | 2.38 | UR | −1.06 | n.s. |
| SD233575 | *So*PIP1-3/PIP1-4 | −6.56 | DR | 4.21 | UR |
| SD231438 | *So*PIP1-3/PIP1-4 | −5.16 | DR | 1.38 | n.s. |
| SD231437 | *So*PIP1-3/PIP1-4 | −1.89 | DR | 1.15 | UR |
| SD205705 | *So*PIP1-3/PIP1-4 | −1.87 | DR | −1.14 | n.s. |
| SD231440 | *So*PIP1-3/PIP1-4 | −2.79 | DR | −2.14 | n.s. |
| SD241279 | SoPIP1-5 | −1.22 | DR | 2.40 | UR |
| SD91837 | SoPIP2-4 | −17.81 | DR | 1.32 | n.s. |
| SD243847 | *So*PIP2-1 | −7.50 | DR | −1.60 | n.s. |
| SD243911 | *So*PIP2-1 | −5.62 | DR | −1.33 | n.s. |
| SD54851 | *So*PIP2-1 | −2.53 | DR | 1.87 | n.s. |
| SD176663 | *Os*PIP2-4 | −2.51 | DR | 1.66 | UR |
| SD176664 | *So*PIP2-4 | −1.73 | DR | 1.02 | n.s. |
| SD84616 | *So*PIP2-5 | −4.42 | DR | 1.32 | n.s. |
| SD176669 | *So*PIP2-6 | −2.15 | DR | 1.70 | UR |
| SD19005 | *Hv*TIP1-1 | −1.87 | DR | 1.84 | n.s. |
| SD19006 | *Hv*TIP1-1 | −5.16 | DR | 4.67 | UR |
| SD7939 | *Ta*TIP1-1 | −2.95 | DR | 1.76 | n.s. |
| SD80616 | *Ss*TIP1-1 | −2.81 | DR | 2.34 | UR |
| SD182871 | *So*TIP2-2 | −6.09 | DR | −1.63 | n.s. |
| SD173276 | *So*TIP2-2 | −3.33 | DR | 496.29 | UR |
| SD119919 | *So*TIP2-3 | −6.56 | DR | 1.61 | n.s. |
| SD194892 | *So*NIP3-1 | −2.34 | DR | −1.60 | n.s. |

Isoforms are preceded by the abbreviated species name (*Hv*: *Hordeum vulgare*; *Os*: *Oryza sativa*; *So*: *Saccharum officinarum*; *Ss*: *Saccharum* spp. and *Ta*: *Triticum aestivum*). HT-SuperSAGE libraries: SD24T (bulk of tolerant genotypes under stress; SD24S (bulk of sensitive genotypes under stress); SDTC (control bulk of tolerant genotypes); and SDSC (control bulk of sensitive genotypes). FC: ratio of the frequencies (normalized to 1,000,000) observed in the stressed library in relation to the control library. DR: down-regulated; UR: up-regulated; n.s.: not significant at $P < 0.05$.

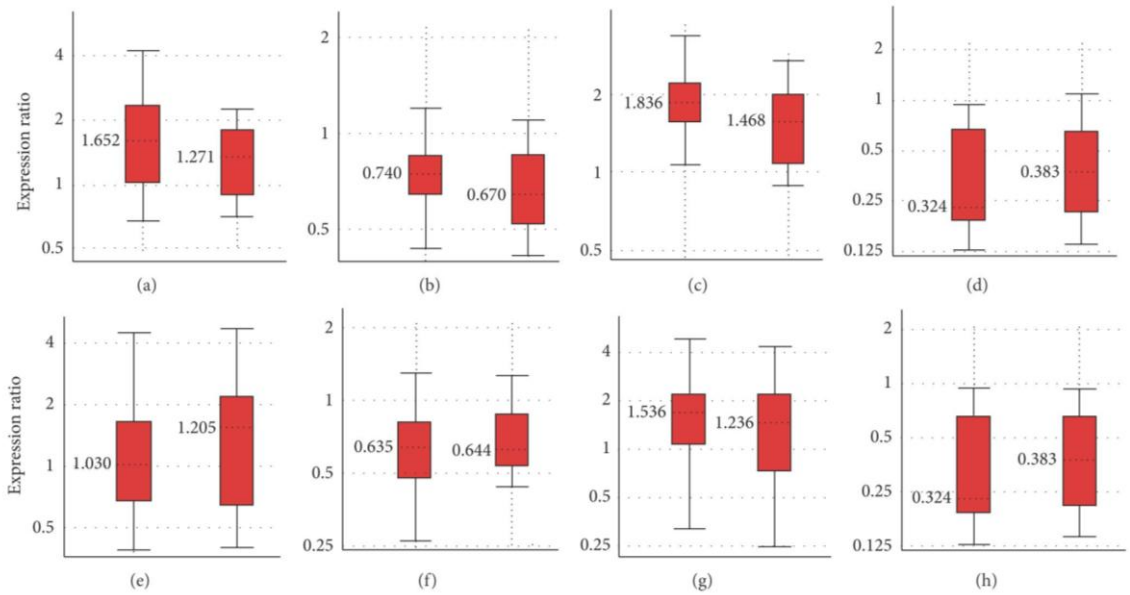



FIGURE 4: Relative quantification of SsPIP1 aquaporin-1 (unitag SD264077) and SoPIP1-3/PIP1-4 (unitag SD231548) with sugarcane cDNAs [tolerant genotypes: (a) CTC6, (b) CTC15, (c) SP83-2847, (d) SP83-5073; sensitive genotypes [(e) CTC9, (f) CTC13, (g) SP90-1638, (h) SP90-3414] under stress (24 h of continuous dehydration). The expression ratios of the same genotypes under control condition were normalized by reference genes GAPDH and 25S rRNA. #Relative expression (REST© v. 2.0.13 software) with the mean value (horizontally dotted line) and range of the observations (100%, horizontal bars); in red: confidence interval at 95%.

in the stressed tolerant bulk as compared to the stressed sensitive bulk.

The unitag related to PIP2-4 (SD176664) was down-regulated in the tolerant bulks after onset of the stress, and it had no relevant expression changes in the sensitive bulks. Nevertheless, it was more expressed in the tolerant bulk when compared with the sensitive bulk, upon stress or even under control conditions. Thus, the tolerant genotypes seemed to produce more PIP2-4 transcripts than the sensitive genotypes. In maize, this aquaporin isoform was up-regulated after only two hours of salt stress, in which time the recovery phase of the osmotic potential falls [52].

In turn, the PIP2-1-related unitag (SD176669) behaved differently under stress and its reaction depends on the genotype (it was down-regulated as compared to the tolerant bulks, and up-regulated as compared to the sensitive bulks). When considering the contrast between both control and stressed bulks of genotypes, this unitag was better expressed in the sensitive bulk than in the tolerant one. By taking into account that this aquaporin isoform increases insensitivity to salinity [67] in transgenic rice and the fact that salinity and drought share many response pathways [68, 69], this isoform deserves further investigation.

Analysis of controls of the two different genotype bulks s showed that the aquaporin isoforms are present in all subfamilies and expressed in the three studied levels during normal daily irrigation (Figure 3(d)). Four isoforms [PIP1-2 (SD92576 unitag); PIP2-1 (SD176664 unitag); PIP2-4 (SD176663 and SD176950 unitags); PIP2-6 (SD176669 unitag)] presented significantly higher abundance in the tolerant bulk, while only one (NIP1-1) was less expressed in relation to the sensitive bulk. In turn, five isoforms (PIP2-2, PIP2-3, PIP2-5, PIP2-6, and TIP4-2) were similarly transcribed in both bulks, while another five isoforms (PIP2-1, PIP2-4, TIP2-2, TIP2-3, and SIP1-2) presented all the three expression levels. Differences in the transcriptional profiles of both controls bulks reinforce the expression modulation of genes presenting in the genotypes composing the bulks.

The two main comparisons SD24T versus SDTC and SD24S versus SDSC revealed that from a total of 18 up-regulated unitags in the tolerant bulks, eight were down-regulated in sensitive bulks, while from 22 other unitags, down-regulated in the tolerant bulks, eight were up-regulated in the sensitive bulks (Table 4). The same isoforms showing different expression levels (Table 4) can be explained by the similarities between aquaporins sequences, in part a consequence of the high level of duplicated plant MIP genes, which is higher than that observed in vertebrates, possibly reflecting the environmental pressures plants are exposed to [45, 49], and also the aneuploidy and polyploidy events observed in the *Saccharum* complex [41].

*3.4. Unitag Expression Validation by RT-qPCR.* The use of RT-qPCR for the confirmation of aquaporin gene expression changes in grass (maize [72] and sugarcane [25]) has already been reported. In the present work we attempted to

Table 5: Primers sequences for RT-qPCR of *Ss*PIP1-1, *So*PIP1-3/PIP1-4 (designed from sugarcane ESTs), 25S rRNA, and GAPDH (as reference genes).

| Unitag | EST/cluster | Database* | Gene | Primers | Tm (°C) | Amplicon (bp) | Regression coefficient ($R^2$) | Amplification efficiency (%) |
|---|---|---|---|---|---|---|---|---|
| SD264077 | gi\|35203438 | dbEST | *Ss*PIP1-1 | 5′-GTTCCTATCCTTGCCCCACT-3′<br>3′-AGGCGTGATCCCTGTTGTAG-5′ | 84.6 | 134 | 0.995 | 95.55 |
| SD231548 | TC127588 | SoGI | *So*PIP1-3/PIP1-4 | 5′-GACTCCCATGTTCCTATCCTTG-3′<br>3′-CGTGATCCCTGTTGTAGATGAT-5′ | 84.3 | 142 | 0.992 | 93.47 |
| — | gi\|33464288 | dbEST | 25S rRNA | 5′-GCAGCCAAGCGTTCATAG-3′<br>3′- CGGCACGGTCATCAGTAG-5′ | 82.9 | 172 | 0.999 | 99.82 |
| — | TC531505 | SoGI | GAPDH | 5′-GGTTCACTTGAAGGGTGGTG-3′<br>3′- TGAGGTGTACCTGTCCTCGTT-5′ | 81.8 | 100 | 0.984 | 100.89 |

Isoforms are preceded by the abbreviated species name (*So*: *Saccharum officinarum*; *Ss*: *Saccharum* spp.).
*Databases: dbEST (NCBI; http://www.ncbi.nlm.nih.gov/), Gene Index (SoGI; http://compbio.dfci.harvard.edu/tgi/)].

determine which genotype was responsible for the bulk of expression in tolerant or sensitive genotypes. To that end, each genotype composing that bulk was independently tested by RT-qPCR analysis. Thus, two unitags [SD264077 (PIP1-1) and SD231548 (PIP1-3/PIP1-4)] considered UR in the tolerant bulk as well as DR and n.s. in the sensitive were selected for expression validation using two reference genes (25S rRNA and GAPDH), both reported to be suitable for sugarcane (Table 5). The relative expression results of the tolerant and sensitive genotypes for the two target genes are shown in Table 6, together with their respective unitag results.

PIP1-1 (SD264077 unitag) was induced by stress in two of the tolerant genotypes (CTC6 and SP83-2847), in comparison to the respective controls (Table 6, Figure 4). Nevertheless, in the remaining tolerant genotypes (CTC15 and SP83-5073) both PIP1-1 genes were down-regulated under the same conditions (Table 6, Figure 4). Thus it can be concluded that CTC6 and SP83-2847 were responsible for the unitag up-regulation. The overexpression of rice PIP1-1 in root and leaf (within 24 h) enhanced the tolerance to drought (200 mM mannitol) and salt stress (100 mM NaCl) in transgenic *Arabidopsis* [62]. Also, PIP1-1 aquaporin isoforms in grapevine were highly expressed in roots (RT-qPCR) in response to water deficit (8 days of constant dehydration [56]).

PIP1-3/PIP1-4 (SD231548 unitag), were stress-induced in genotype SP83-2847 (Table 6, Figure 4), in agreement with the HT-SuperSAGE data. It should be noted that in phylogenetic analyses PIP1-3 and PIP1-4 are highly similar with barley PIP amino acid sequences, being grouped together as one isoform, while they are phylogenetically more distant from PIP1-1 (from barley and rice, [73]). *A. thaliana* PIP1-3 and PIP1-4 isoforms had their transcription level increased more than five times, covering the first 48 h of drought stress (250 mM mannitol), in leaves and roots, as well as in response to salt (150 mM NaCl) and cold stresses [48]. In turn, PIP1-3 overexpression in transgenic rice, combining aquaporin coding sequence with a constitutive corn promoter, showed enhanced stress tolerance to cold [74].

However, in relation to water transport by PIP1-3, which appears to be less permeable to water [74], this isoform could work best in conjunction with PIP2 subgroup members, in which *in silico* analysis showed them to be mostly DR. Considering the remaining components of the tolerant bulk, the genotype CTC6 did not show significant differences in transcript levels, while CTC15 and SP83-5073 presented down-regulated transcription at the onset of root dehydration (Table 6, Figure 4).

The strategy of opening bulks in the RT-qPCR validation reinforced the transcription modulation of sugarcane aquaporins and gave hints to genotype-specific expression. Thus, plants considered physiologically tolerant or sensitive to root

Table 6: Relative expression rates of aquaporins PIP1-1 (SD264077 unitag) and PIP1-3/PIP1-4 (SD231548 unitag) in bulks of tolerant or sensitive genotypes, respectively, and RT-qPCR data (both in bulks and each genotype).

| Technique and genotypes | Target gene | |
|---|---|---|
| | PIP1-1 | PIP1-3/PIP1-4 |
| | Modulation of gene expression[&] | |
| HT-SuperSAGE | | |
| Tolerant bulk[1] | 3.580*/UR | 3.180*/UR |
| Sensitive bulk[1] | −4.560*/DR | 1.280*/ns |
| RT-qPCR and Tolerant Genotypes | | |
| CTC6 | 1.652#/UR | 1.271#/ns |
| CTC15 | 0.740#/DR | 0.670#/DR |
| SP83-2847 | 1.836#/UR | 1.468#/UR |
| SP83-5073 | 0.324#/DR | 0.383#/DR |
| RT-qPCR and sensitive Genotypes | | |
| CTC9 | 1.030#/ns | 1.205#/ns |
| CTC13 | 0.635#/DR | 0.644#/DR |
| SP90-1638 | 1.536#/ns | 1.236#/ns |
| SP90-3414 | 0.324#/DR | 0.383#/DR |

[1]Bulk of the four tolerant or sensitive genotypes; [&]$P < 0.05$;**fold change* [FC: ratio of the frequencies (normalized to 1,000,000) observed in the stressed library in relation to the control library]; #relative expression level using the REST software (v. 2.0.13) [26]; DR: down-regulated; UR: up-regulated; ns: not significant at $P < 0.05$.

dehydration (24 h) varied in the expression of aquaporin isoforms. The same was observed with up (*O. sativa* L. cv. Zhonghan 3) and lowland (cv. Xiushui) rice under water deficit [75]. The RT-qPCR results revealed genotype-specific differences for PIP1-2, PIP1-3, PIP2-1, and PIP2-5 isoforms in roots, and PIP1-2 and PIP1-3 in leaves. The above mentioned isoforms were up-regulated in upland rice, whereas they remained unchanged or DR in lowland rice [75].

Finally, the RT-qPCR protocol, in the present work applied for unitags validation, as well as the identified unitags for PIP1-1 and PIP 1-3/PIP1-4, define a set of functional molecular markers based on the expression profiles validated with appropriate genotypes. This expression marker set will assist breeders in marker-assisted selection of elite genotypes more tolerant to abiotic stresses.

## 4. Conclusions

The present work is a pioneer study specifically addressing the aquaporin transcripts in sugarcane transcriptomes established from HT-SuperSAGE transcription profiles from roots of tolerant and sensitive genotypes after 24 h of continuous dehydration. Almost all 26 bp unitags were annotated using a public sugarcane EST databases, especially S. by *S.*, allowing the identification of potential aquaporins. Categorizing the EST-anchored unitags by Gene Ontology (GO) enhanced the annotation efficiency by almost 15%. These procedures identified potential isoforms of the four aquaporin subfamilies (PIP, TIP, NIP, and SIP) already described for higher plants, together with their respective expression profiles in sugarcane under abiotic stress. Moreover, an efficient protocol for RT-qPCR was developed, enabling gene expression validation of SuperSAGE unitags related to PIP aquaporins (PIP1-1 and PIP1-3/PIP1-4) and involving reference genes encoding GAPDH and 25S rRNA, testing each genotype individually the employed, validation strategy revealed genotype-specificity of the response to the applied stress.

## Conflict of Interests

The authors declare that there is no conflict of interests.

## Acknowledgments

The authors thank Dr. Björn Rotter, Ruth Jungmann, Nicolas Krezdorn, and Dr. Peter Winter (GenXPro GmbH, Germany) for scientific and technical advices. This work has been funded by Brazilian institutions Financiadora de Estudos e Projetos (FINEP), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## References

[1] G. E. Welbaum and F. C. Meinzer, "Compartmentation of solutes and water in developing sugarcane stalk tissue," *Plant Physiology*, vol. 93, no. 3, pp. 1147–1153, 1990.

[2] S. L. Dillon, F. M. Shapter, R. J. Henry, G. Cordeiro, L. Izquierdo, and L. S. Lee, "Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae)," *Annals of Botany*, vol. 100, no. 5, pp. 975–989, 2007.

[3] FAOSTAT, "Food and Agriculture Organization of the Unite Nations. In: FAO statistical databases," 2011, http://faostat.fao.org/.

[4] S. S. Gosal, S. H. Wani, and M. S. Kang, "Biotechnology and drought tolerance," *Journal of Crop Improvement*, vol. 23, no. 1, pp. 19–54, 2009.

[5] M. D. A. Silva, J. A. G. Da Silva, J. Enciso, V. Sharma, and J. Jifon, "Yield components as indicators of drought tolerance of sugarcane," *Scientia Agricola*, vol. 65, no. 6, pp. 620–627, 2008.

[6] E. A. Bray, "Molecular responses to water deficit," *Plant Physiology*, vol. 103, no. 4, pp. 1035–1040, 1993.

[7] K. Shinozaki and K. Yamaguchi-Shinozaki, "Gene networks involved in drought stress response and tolerance," *Journal of Experimental Botany*, vol. 58, no. 2, pp. 221–227, 2007.

[8] S. D. Tyerman, C. M. Niemietz, and H. Bramley, "Plant aquaporins: multifunctional water and solute channels with expanding roles," *Plant, Cell and Environment*, vol. 25, no. 2, pp. 173–194, 2002.

[9] C. Maurel, "Plant aquaporins: novel functions and regulation properties," *FEBS Letters*, vol. 581, no. 12, pp. 2227–2236, 2007.

[10] M. Besse, T. Knipfer, A. J. Miller, J.-L. Verdeil, T. P. Jahn, and W. Fricke, "Developmental pattern of aquaporin expression in barley (*Hordeum vulgare* L.) leaves," *Journal of Experimental Botany*, vol. 62, no. 12, pp. 4127–4142, 2011.

[11] P. Gerbeau, J. Güçlü, P. Ripoche, and C. Maurel, "Aquaporin Nt-TIPa can account for the high permeability of tobacco cell vacuolar membrane to small neutral solutes," *Plant Journal*, vol. 18, no. 6, pp. 577–587, 1999.

[12] L.-H. Liu, U. Ludewig, B. Gassert, W. B. Frommer, and N. Von Wirén, "Urea transport by nitrogen-regulated tonoplast intrinsic proteins in arabidopsis," *Plant Physiology*, vol. 133, no. 3, pp. 1220–1228, 2003.

[13] N. Uehleln, C. Lovisolo, F. Siefritz, and R. Kaldenhoff, "The tobacco aquaporin NtAQP1 is a membrane $CO_2$ pore with physiological functions," *Nature*, vol. 425, no. 6959, pp. 734–737, 2003.

[14] D. Loqué, U. Ludewig, L. Yuan, and N. Von Wirén, "Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate $NH_3$ transport into the vacuole," *Plant Physiology*, vol. 137, no. 2, pp. 671–680, 2005.

[15] C. Dordas, M. J. Chrispeels, and P. H. Brown, "Permeability and channel-mediated transport of boric acid across membrane vesicles isolated from squash roots," *Plant Physiology*, vol. 124, no. 3, pp. 1349–1361, 2000.

[16] T. Henzler and E. Steudle, "Transport and metabolic degradation of hydrogen peroxide in chara corallina: model calculations and measurements with the pressure probe suggest transport of $H_2O_2$ across water channels," *Journal of Experimental Botany*, vol. 51, no. 353, pp. 2053–2066, 2000.

[17] G. P. Bienert, M. Thorsen, M. D. Schüssler et al., "A subgroup of plant aquaporins facilitate the bi-directional diffusion of $As(OH)_3$ and $Sb(OH)_3$ across membranes," *BMC Biology*, vol. 6, article 26, 2008.

[18] U. Johanson and S. Gustavsson, "A new subfamily of major intrinsic proteins in plants," *Molecular Biology and Evolution*, vol. 19, no. 4, pp. 456–461, 2002.

[19] M. S. Lucero, F. Mirarchi, J. Goldstein, and C. Silberstein, "Intraperitoneal administration of Shiga toxin 2 induced neuronal alterations and reduced the expression levels of aquaporin 1 and aquaporin 4 in rat brain," *Microbial Pathogenesis*, vol. 53, no. 2, pp. 87–94, 2012.

[20] G. Di Giusto, P. Flamenco, V. Rivarola et al., "quaporin 2-increased renal cell proliferation is associated with cell volume regulation," *Journal of Cellular Biochemistry*, vol. 113, no. 12, pp. 3721–3729, 2012.

[21] C. Geijer, D. Ahmadpour, M. Palmgren et al., "Yeast aquaglyceroporins use the transmembrane core to restrict glycerol transport," *The Journal of Biological Chemistry*, vol. 287, no. 28, pp. 23562–23570, 2012.

[22] M. Ayadi, D. Cavez, N. Miled, F. Chaumont, and K. Masmoudi, "Identification and characterization of two plasma membrane aquaporins in durum wheat (*Triticum turgidum* L. subsp. durum) and their role in abiotic stress tolerance," *Plant Physiology and Biochemistry*, vol. 49, no. 9, pp. 1029–1039, 2011.

[23] C. Liu, T. Fukumoto, T. Matsumoto et al., "Aquaporin OsPIP1;1 promotes rice salt resistance and seed germination," *Plant Physiology and Biochemistry*, vol. 63, pp. 151–158, 2013.

[24] C. Maure, L. Verdoucq, D.-T. Luu, and V. Santoni, "Plant aquaporins: membrane channels with multiple integrated functions," *Annual Review of Plant Biology*, vol. 59, pp. 595–624, 2008.

[25] C. G. Lembke, M. Y. Nishiyama Jr., P. M. Sato, R. F. Andrade, and G. M. Souza, "Identification of sense and antisense transcripts regulated by drought in sugarcane," *Plant Molecular Biology*, vol. 79, no. 4-5, pp. 461–477, 2012.

[26] M. W. Pfaffl, G. W. Horgan, and L. Dempfle, "Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR," *Nucleic Acids Research*, vol. 30, no. 9, p. e36, 2002.

[27] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," *Molecular Biology and Evolution*, vol. 28, no. 10, pp. 2731–2739, 2011.

[28] H. Matsumura, D. H. Krüger, G. Kahl, and R. Terauchi, "SuperSAGE: a modern platform for genome-wide quantitative transcript profiling," *Current Pharmaceutical Biotechnology*, vol. 9, no. 5, pp. 368–374, 2008.

[29] H. Matsumura, S. Reich, A. Ito et al., "Gene expression analysis of plant host-pathogen interactions by SuperSAGE," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15718–15723, 2003.

[30] H. Matsumura, S. Reich, M. Reuter et al., "perSAGE: a potent transcriptome tool for eukaryotic organisms," in *Current Technologies and Applications*, S. M. Wang, Ed., pp. 77–90, Horizon Scientific Press, Norwich, UK, 2004.

[31] H. Matsumura, K. Yoshida, S. Luo et al., "High-throughput SuperSAGE," *Methods in Molecular Biology*, vol. 687, pp. 135–146, 2011.

[32] H. Matsumura, N. Urasaki, K. Yoshida, D. H. Kruger, G. Kahl, and R. Terauchi, "SuperSAGE: powerful serial analysis of gene expression," *Methods in Molecular Biology*, vol. 883, pp. 1–17, 2012.

[33] E. A. Kido, J. R. C. Ferreira Neto, R. L. O. Silva et al., "New insights in the sugarcane transcriptome responding to drought stress as revealed by SuperSAGE," *The Scientific World Journal*, vol. 2012, pp. 1–14, 2012.

[34] S. Audic and J.-M. Claverie, "The significance of digital gene expression profiles," *Genome Research*, vol. 7, no. 10, pp. 986–995, 1997.

[35] N. Robertson, M. Oveisi-Fordorei, S. D. Zuyderduyn et al., "DiscoverySpace: an interactive data analysis application," *Genome Biology*, vol. 8, no. 1, article R6, 2007.

[36] A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles, "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research," *Bioinformatics*, vol. 21, no. 18, pp. 3674–3676, 2005.

[37] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.

[38] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," *Methods in Molecular Biology*, vol. 132, pp. 365–386, 2000.

[39] E. A. Kido, P. K. de Araújo Barbosa, J. R. C. F. Neto et al., "Identification of plant protein kinases in response to abiotic and biotic stresses using superSAGE," *Current Protein and Peptide Science*, vol. 12, no. 7, pp. 643–656, 2011.

[40] A. D'Hont, L. Grivet, P. Feldmann, S. Rao, N. Berding, and J. C. Glaszmann, "Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics," *Molecular and General Genetics*, vol. 250, no. 4, pp. 405–413, 1996.

[41] L. Grivet and P. Arruda, "Sugarcane genomics: depicting the complex genome of an important tropical crop," *Current Opinion in Plant Biology*, vol. 5, no. 2, pp. 122–127, 2002.

[42] A. L. Vettore, F. R. Da Silva, E. L. Kemper, and P. Arruda, "The libraries that made SUCEST," *Genetics and Molecular Biology*, vol. 24, no. 1–4, pp. 1–7, 2001.

[43] F. Chaumont, F. Barrieu, E. Wojcik, M. J. Chrispeels, and R. Jung, "Aquaporins constitute a large and highly divergent protein family in maize," *Plant Physiology*, vol. 125, no. 3, pp. 1206–1215, 2001.

[44] M. X. Nguyen, S. Moon, and K. H. Jung, "Genome-wide expression analysis of rice aquaporin genes and development of a functional gene network mediated by aquaporin expression in roots," *Planta*, vol. 238, no. 4, pp. 669–681, 2013.

[45] G. Soto, K. Alleva, G. Amodeo, J. Muschietti, and N. D. Ayub, "New insight into the evolution of aquaporins from flowering plants and vertebrates: orthologous identification and functional transfer is possible," *Gene*, vol. 503, no. 1, pp. 165–176, 2012.

[46] P. Kjellbom, C. Larsson, I. Johansson, M. Karlsson, and U. Johanson, "Aquaporins and water homeostasis in plants," *Trends in Plant Science*, vol. 4, no. 8, pp. 308–314, 1999.

[47] A. Bansal and R. Sankararamakrishnan, "Homology modeling of major intrinsic proteins in rice, maize and *Arabidopsis*: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters," *BMC Structural Biology*, vol. 7, article 27, 2007.

[48] J. Y. Jang, D. G. Kim, Y. O. Kim, J. S. Kim, and H. Kang, "An expression analysis of a gene family encoding plasma membrane aquaporins in response to abiotic stresses in *Arabidopsis thaliana*," *Plant Molecular Biology*, vol. 54, no. 5, pp. 713–725, 2004.

[49] R. Zardoya, "Phylogeny and evolution of the major intrinsic protein family," *Biology of the Cell*, vol. 97, no. 6, pp. 397–414, 2005.

[50] F. Magni, C. Sarto, D. Ticozzi et al., "Proteomic knowledge of human aquaporins," *Proteomics*, vol. 6, no. 20, pp. 5637–5649, 2006.

[51] E. Alexandersson, L. Fraysse, S. Sjövall-Larsen et al., "Whole gene family expression and drought stress regulation of aquaporins," *Plant Molecular Biology*, vol. 59, no. 3, pp. 469–484, 2005.

[52] C. Zhu, D. Schraut, W. Hartung, and A. R. Schäffner, "Differential responses of maize MIP genes to salt stress and ABA," *Journal of Experimental Botany*, vol. 56, no. 421, pp. 2971–2981, 2005.

[53] R. Aroca, G. Amodeo, S. Fernández-Illescas, E. M. Herman, F. Chaumont, and M. J. Chrispeels, "The role of aquaporins and membrane damage in chilling and hydrogen peroxide induced changes in the hydraulic conductance of maize roots," *Plant Physiology*, vol. 137, no. 1, pp. 341–353, 2005.

[54] N. Uehlein, K. Fileschi, M. Eckert, G. P. Bienert, A. Bertl, and R. Kaldenhoff, "Arbuscular mycorrhizal symbiosis and plant aquaporin expression," *Phytochemistry*, vol. 68, no. 1, pp. 122–129, 2007.

[55] F. Lopez, A. Bousser, I. Sissoëff et al., "Diurnal regulation of water transport and aquaporin gene expression in maize roots: contribution of PIP2 proteins," *Plant and Cell Physiology*, vol. 44, no. 12, pp. 1384–1395, 2003.

[56] R. K. Vandeleur, G. Mayo, M. C. Shelden, M. Gilliham, B. N. Kaiser, and S. D. Tyerman, "The role of plasma membrane intrinsic protein aquaporins in water transport through roots: diurnal and drought stress responses reveal different strategies between isohydric and anisohydric cultivars of grapevine," *Plant Physiology*, vol. 149, no. 1, pp. 445–460, 2009.

[57] D. M. Gibeaut, J. Hulett, G. R. Cramer, and J. R. Seemann, "Maximal biomass of *Arabidopsis thaliana* using a simple, low-maintenance hydroponic method and favorable environmental conditions," *Plant Physiology*, vol. 115, no. 2, pp. 317–319, 1997.

[58] I. S. Wallace, W.-G. Choi, and D. M. Roberts, "The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins," *Biochimica et Biophysica Acta*, vol. 1758, no. 8, pp. 1165–1175, 2006.

[59] F. Ishikawa, S. Suga, T. Uemura, M. H. Sato, and M. Maeshima, "Novel type aquaporin SIPs are mainly localized to the ER membrane and show cell-specific expression in *Arabidopsis thaliana*," *FEBS Letters*, vol. 579, no. 25, pp. 5814–5820, 2005.

[60] C. Maurel, F. Tacnet, J. Güclü, J. Guern, and P. Ripoche, "Purified vesicles of tobacco cell vacuolar and plasma membranes exhibit dramatically different water permeability and water channel activity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 13, pp. 7103–7108, 1997.

[61] X. Yu, Y. H. Peng, M. H. Zhang, Y. J. Shao, W. A. Su, and Z. C. Tang, "Water relations and an expression analysis of plasma membrane intrinsic proteins in sensitive and tolerant rice during chilling and recovery," *Cell Research*, vol. 16, no. 6, pp. 599–608, 2006.

[62] L. Guo, Y. W. Zi, H. Lin et al., "Expression and functional analysis of the rice plasma-membrane intrinsic protein gene family," *Cell Research*, vol. 16, no. 3, pp. 277–286, 2006.

[63] V. Terzi, G. Pastori, P. R. Shewry, N. Di Fonzo, A. Michele Stanca, and P. Faccioli, "Real-time PCR-assisted selection of wheat plants transformed with HMW glutenin subunit genes," *Journal of Cereal Science*, vol. 41, no. 1, pp. 133–136, 2005.

[64] H. J. Schouten, F. A. Krens, and E. Jacobsen, "Do cisgenic plants warrant less stringent oversight?" *Nature Biotechnology*, vol. 24, no. 7, p. 753, 2006.

[65] S. G. Joshi, J. G. Schaart, R. Groenwold, E. Jacobsen, H. J. Schouten, and F. A. Krens, "Functional analysis and expression profiling of HcrVf1 and HcrVf2 for development of scab resistant cisgenic and intragenic apples," *Plant Molecular Biology*, vol. 75, no. 6, pp. 579–591, 2011.

[66] M. Katsuhara, Y. Akiyama, K. Koshio, M. Shibasaka, and K. Kasamo, "Functional analysis of water channels in barley roots," *Plant and Cell Physiology*, vol. 43, no. 8, pp. 885–893, 2002.

[67] M. Katsuhara, K. Koshio, M. Shibasaka, Y. Hayashi, T. Hayakawa, and K. Kasamo, "Over-expression of a barley aquaporin increased the shoot/root ratio and raised salt sensitivity in transgenic rice plants," *Plant and Cell Physiology*, vol. 44, no. 12, pp. 1378–1383, 2003.

[68] S. Mahajan and N. Tuteja, "Cold, salinity and drought stresses: an overview," *Archives of Biochemistry and Biophysics*, vol. 444, no. 2, pp. 139–158, 2005.

[69] G.-T. Huang, S.-L. Ma, L.-P. Bai et al., "Signal transduction during cold, salt, and drought stresses in plants," *Molecular Biology Reports*, vol. 39, no. 2, pp. 969–987, 2012.

[70] J. Y. Jang, S. H. Lee, J. Y. Rhee, G. C. Chung, S. J. Ahn, and H. Kang, "Transgenic *Arabidopsis* and tobacco plants overexpressing an aquaporin respond differently to various abiotic stresses," *Plant Molecular Biology*, vol. 64, no. 6, pp. 621–632, 2007.

[71] C. Tournaire-Roux, M. Sutka, H. Javot et al., "Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins," *Nature*, vol. 425, no. 6956, pp. 393–397, 2003.

[72] R. Gu, X. Chen, Y. Zhou, and L. Yuan, "Isolation and characterization of three maize aquaporin genes, ZmNIP2;1, ZmNIP2;4 and ZmTIP4;4 involved in urea transport," *BMB Reports*, vol. 45, no. 2, pp. 96–101, 2012.

[73] T. Horie, T. Kaneko, G. Sugimoto et al., "Mechanisms of water transport mediated by PIP aquaporins and their regulation via phosphorylation events under salinity stress in barley roots," *Plant and Cell Physiology*, vol. 52, no. 4, pp. 663–675, 2011.

[74] T. Matsumoto, H.-L. Lian, W.-A. Su et al., "Role of the aquaporin PIP1 subfamily in the chilling tolerance of rice," *Plant and Cell Physiology*, vol. 50, no. 2, pp. 216–229, 2009.

[75] H.-L. Lian, X. Yu, D. Lane, W.-N. Sun, Z.-C. Tang, and W.-A. Su, "Upland rice and lowland rice exhibited different PIP expression under water deficit and ABA treatment," *Cell Research*, vol. 16, no. 7, pp. 651–660, 2006.

# 7 ARTIGO IV

## Análise da Expressão de Genes de Fatores de Transcrição de Cana-de-açúcar sob Déficit Hídrico

Artigo a ser submetido à revista *Genetics and Molecular Biology* (ISSN: 0100-8455)

**Análise da Expressão de Genes de Fatores de Transcrição de Cana-de-açúcar sob Déficit Hídrico**

Manassés Daniel da Silva[1], José Ribamar Costa Ferreira Neto[1], George André de Lima Cabral[1], Ana Maria Benko-Iseppon[1] e Ederson Akio Kido[1*]

**Resumo**

Este trabalho representa uma análise transcricional global de fatores de transcrição (FTs) expressos 24 h após o início da supressão de rega, em raízes de híbridos modernos de cana-de-açúcar. Para tanto, *tags* HT-SuperSAGE (205.975) de raízes de dois *bulks* de quatro cultivares cada, um considerado sensível e outro tolerante, foram alinhadas contra 1.354 sequências nucleotídicas de FTs de *Saccharum* spp. A partir da expressão dos cultivares tolerantes foram identificados 534 FTs e a partir de cultivares sensíveis, 526. Nos cultivares tolerantes, 264 FTs foram diferencialmente expressos (DE), sendo 146 induzidos e 118 reprimidos; enquanto que nos sensíveis, dos 277 FTs DE, 133 foram induzidos e 144 foram reprimidos. Dentre os FTs DE, as famílias mais abundantes foram AP2/EREB, bZIP e ZIM. A análise *in silico* também identificou prováveis FTs que controlariam a expressão das *unitags* mais induzidas nos dois *bulks*, em resposta ao estresse aplicado. Adicionalmente, expressões de cinco FTs foram validadas via RT-qPCR, a partir de cada cultivar (separadamente), identificando-se elementos regulatórios atuantes em *Cis*, mais prevalentes em regiões promotoras de nove FTs induzidos. Os dados fornecem indícios da resposta imediata de cultivares contrastantes ao déficit hídrico, sendo candidatos naturais com potencial para estudos futuros envolvendo seleção assistida por marcadores (SAM) e transgenia.

**Palavras-chave:** transcriptoma, *Saccharum* spp., regulação gênica, HT-SuperSAGE

**Introdução**

Ao longo da história evolutiva as plantas adquiriram mecanismos de tolerância a estresses, cujos processos envolvem mudanças fisiológicas e bioquímicas que resultam em mudanças adaptativas ou morfológicas (Boyko e Kovalchuk, 2008, Urano *et al.*,

2010). Assim, a partir da percepção de condições não favoráveis, ocorre uma reprogramação transcricional direcionada a minimizar os efeitos deletérios nas células vegetais (Han e Wagner, 2013; Ferreira Neto *et al.*, 2016).

Neste contexto, avanços em análises genômicas de larga escala revelaram complexas redes de regulação gênica (*gene regulatory networks*, GRNs) que controlam a expressão gênica global, modificações pós-traducionais (alosteria) e a composição metabólica (Urano *et al.*, 2010). As GRNs são compostas por diferentes regulações genéticas e epigenéticas, que acarretam mudanças na expressão de fatores de transcrição (FTs), na distribuição de nucleossomos, nas modificações de histonas, na metilação do DNA, e nos RNAs não codificadores de proteínas (*non-protein-coding* RNAs, npcRNAs), conduzindo a importantes rotas que respondem a estresses abióticos (Urano *et al.*, 2010).

A maior classe de genes reguladores da expressão gênica, com suas 58 famílias e abrangendo mais de 5% de todo conjunto de genes no grupo de vegetais superiores (Jin *et al.*, 2014), envolve os TFs (proteínas compostas basicamente de quatro domínios: ligação ao DNA, oligomerização, sinal de localização nuclear e regulação da transcrição) (Liu *et al.*, 1999). O domínio de ligação ao DNA é o principal domínio, snedo responsável pela nomenclatura das diferentes famílias de TFs (Yamasaki *et al.*, 2013). A família *APETALA 2 /ethylene-responsive element binding factor* (AP2/EREB) é a maior família de TFs específicas de plantas, com isoformas de várias subfamílias envolvidas em respostas a estresses abióticos, tais como frio, desidratação, calor, bem como na resposta ao regulador de crescimento etileno e no desenvolvimento de flores, raízes, embriões e sementes (Yamasaki *et al.*, 2013). A família bHLH (*basic helix-loop-helix*), também uma das principais famílias de FTs em plantas superiores, apresenta representantes com atuação na sinalização de fitocromos, na biossíntese de flavonoides,

na sinalização hormonal e também nas respostas a estresses (Feller *et al.*, 2011). A família WRKY, nomeada em razão da presença do motivo WRKY, apresenta TFs importantes na regulação da reprogramação transcricional em resposta a estresses em plantas (Chen *et al.*, 2012). Os FTs da família MYB, por sua vez, têm recebido bastante atenção em virtude de serem ABA responsivos, em decorrência de déficit hídrico (Yamaguchi-Shinozaki e Shinozaki, 2006). A família NAC, um anacronismo aos três primeiros genes descritos [NAM (No Apical Meristem), ATAF1-2, e CUC2 (Cup-Shaped Cotyledon)] representa TFs específicos de plantas com mais de 140 transcritos identificados em arroz (Xiong *et al.*, 2005; Nakashima *et al.*, 2012). TFs desta família foram reportados em respostas a estresses bióticos e abióticos (Nakashima *et al.*, 2012), bem como em aumentos na tolerância a estresses abióticos (principalmente seca e salinidade) e em trabalhos com gramíneas transgênicas superexpressando genes NAC (Gao *et al.*, 2010; Jeong *et al.*, 2010; Xue *et al.*, 2011; Saad *et al.*, 2013).

Dentre as gramíneas, a cana-de-açúcar (*Saccharum* spp. híbrida) é uma importante cultura cultivada em mais de 100 países, ocupando aproximadamente 26 milhões de hectares. O Brasil destaca-se como o maior produtor mundial dessa cultura, bem como o principal processador e produtor de seus principais produtos (açúcar e etanol) (FAO, 2016). Dentre os estresses abióticos, o déficit hídrico é o mais prejudicial, afetando periodicamente determinadas áreas de cultivo (Gosal *et al.*, 2009; Zingaretti *et al.*, 2012). Apesar da importância da cultura e das perdas ocasionadas por déficit hídrico ainda não há na literatura um estudo de expressão em larga escala das diferentes famílias de TFs (TFoma) em genótipos considerados tolerantes ou sensíveis de cana-de-açúcar em resposta à seca. No presente trabalho, a partir da metodologia de transcriptômica HT-SuperSAGE, dois grupos de quatro cultivares cada, sendo um

considerado tolerante e outro sensível, foram associados com respostas imediatas (24 h) de TFs ao estresse de supressão de rega.

**Material e Métodos**

**Cultivares de cana-de-açúcar, descrição das bibliotecas HT-SuperSAGE**

Oito cultivares de cana-de-açúcar do Centro de Tecnologia Canavieira – CTC, Piracicaba, São Paulo, Brasil - formaram dois conjuntos de acessos contrastantes para a condição de déficit hídrico. O grupo (*bulk*) tolerante consistiu dos cultivares CTC6, CTC15, SP83-2847 e SP83-5073, enquanto que o *bulk* sensível foi composto por CTC9, CTC13, SP90-3414 e SP90-1638. Estes cultivares têm sido reportados como tolerantes e sensíveis à seca, com base em observações de campo por produtores de cana-de-açúcar e análises fisiológicas, incluindo seu conteúdo relativo de água e clorofila (Thiebaut *et al.*, 2012; Thiebaut *et al.*, 2014).

O estresse (24 h de supressão de rega) foi aplicado conforme descrito por Kido *et al.*, (2012). Após o período de tratamento, ambos os grupos tiveram suas raízes coletadas, imediatamente congeladas e estocadas a -80 °C até a extração de RNA total. RNA totais foram isolados via Trizol® (Invitrogen), tratados com DNAse (Qiagen®), purificados com *RNeasy Mini kit* (Qiagen®) e quantificados em espectrofotometro. *Bulks* de cada tratamento [estressados e controles negativos (não estressados)], foram constituídos de RNAs purificados de cada cultivar de seu respectivo *bulk* (tolerante ou sensível), em uma quantidade equimolar.

O sequenciamento das bibliotecas HT-SuperSAGE [tolerante estressada (SD24T), tolerante controle negativo (SDTC), sensível estressada (SD24S) e sensível controle negativo (SDSC)] baseou-se na metodologia de Matsumura *et al.* (2003) com modificações, incluindo-se o sequenciamento via Illumina Genome Analyzer II (tecnologia SOLEXA) (Kido *et al.*, 2012).

Em resumo, os dados disponíveis compreendem a 205,975 *unitags* (*tags* únicas) e suas respectivas frequências observadas nas oito bibliotecas. As frequências observadas permitiram classificar as *unitags* como *up-regulated* (UR), *down-regulated* (DR) ou não significativa n.s. ($p<0.05$), após teste de Audic-Claverie (Audic e Claverie, 1997). Também foi estimada, para cada *unitag*, a modulação de sua expressão gênica (FC: *Fold Change*), calculada como a razão da frequência observada em uma biblioteca comparada com outra.

**Identificação de fatores de transcrição e análise de expressão**

*Unitags* oriundas dos contrastes tolerante (SD24T vs. SDTC) e sensível (SD24S *vs*. SDSC) foram alinhadas (BLASTn) contra sequências de fatores de transcrição (FTs) do GRASSIUS (http://grassius.org/), um banco de dados com sequências de FTs curadas de várias espécies de gramíneas (Yilmaz *et al.*, 2009). Os alinhamentos classificados como aceitáveis compreenderam aqueles com *best hit* (melhor alinhamento) entre sequências de cana-de-açúcar, *score* > 44 (52 correspondendo a 100% de identidade), permitindo até um erro em um alinhamento de 26pb, e com 5'CATG preservado. Também foram prospectados domínios conservados nos peptídeos traduzidos (Bioedit tool version 7.2.5, Hall, 1999) de transcritos FTs com *unitags* identificadas, com o uso da ferramenta *NCBI Batch Web Conserved Domain Search* (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi; Marchler-Bauer *et al.*, 2015). A análise de cluterização hierárquica foi realizada no *software* Cluster 3.0 (v.1.52) (Eisen *et al.*, 1999) utilizando as *unitags* mais diferencialmente expressas (*p-value* < $1e^{-4}$) sobre parâmetros padrão e valores FC como dados de entrada. Os dendrogramas laterais foram gerados usando o programa *TreeView* (Page, 1996). Finalmente, os diagramas de Venn foram gerados com o auxílio da ferramenta *Venny* (Oliveros *et al.*, 2017).

**Prospecção de Elementos regulatórios em *Cis* (*Cis-Acting regulatory elements* - CARE)**

TFs induzidos foram posicionados em BACs de cana-de-açúcar (Setta *et al.*, 2014) usando-se a ferramenta BLASTn (*e-value* < $1e^{-20}$), com posterior identificação de potenciais regiões promotoras, delimitando-se 2000 pb *upstream* do início do alinhamento feito pelo software BioEdit (Hall, 1999). Elementos regulatórios em *Cis* (*Cis-Acting regulatory elements*) foram minerados *in silico* com auxílio da ferramenta newPLACE (https://sogo.dna.affrc.go.jp; Higo *et al.*, 1999).

**Análise de enriquecimento de FTs**

Com o intuito de identificar os TFs cujos alvos tivereram maior expressão, *unitags* significativamente induzidas (*p-value* < 0.0001) e exclusivas de cada *bulk* de cultivares foram mapeadas em ESTs de cana-de-açúcar oriundas da base de dados NCBI. Sequencias contendo as *unitags* foram alinhadas (BLASTx, *e-value cutoff* $e^{-10}$) com proteínas curadas da base de dados UniProtKB/SuissProt. As proteínas relacionadas foram consideradas alvos diferencialmente induzidos e foram consideradas no enriquecimento de TFs (*TF Enrichment*) a partir da ferramenta PlantRegMap (*Plant Transcriptional Regulatory Map*; http://plantregmap.cbi.pku.edu.cn/tf_enrichment.php; Jin *et al* 2016), considerando um *threshold* limite de p-value ≤ 0.05.

**Extração de RNA e síntese de cDNA para os ensaios RT-qPCR**

RNAs totais de cultivares tolerantes e sensíveis de cana-de-açúcar foram extraídos e purificados como mencionado anteriormente. A integridade dos RNAs foi verificada em gel de agarose 1.5% (w / v) corado com *Blue-Green Loading Dye* (LGC Biotecnologia®). As amostras foram quantificadas com o uso do *Quant-iT ™ RNA Assay Kit* (Invitrogen®) em fluorometro Qubit (Invitrogen®). Alicotas de RNAs totais (1 μg) foram convertidos em cDNAs usando o QuantiTect Reverse Transcription Kit

(Qiagen®) e resuspensos em um volume final de 20 µL. A síntese de cDNA foi incubada a 42 °C por 2 min em tampão gDNA Wipeout (para a remoção de contaminantes de DNA genomico), 42 °C for 15 min (transcrição reversa) e 95 °C por 3 min (inativação da enzima), em termociclador. As amostras obtidas foram estocadas a -20 ° C até os as reações RT-qPCR.

Os ensaios de RT-qPCR seguiram as diretrizes do protocolo MIQE (*The Minimum Information for Publication of Quantitative Real-Time PCR Experiments*) (Bustin *et al.*, 2009), com o objetivo de aumentar a confiabilidade nos resultados.

**Desenho de primers, eficiência da amplificação e análises de RT-qPCR**

Transcritos TFs foram alinhados (BLASTn) contra RefSeq mRNA (NCBI; http://www.ncbi.nlm.nih.gov/) para reconfirmação individual da anotação. *Primers* foram desenhados com base em FTs de cana-de-açúcar que continham *unitags* utilizando-se a ferramenta Primer 3 (Rozen e Skaletsky, 2000). Os seguintes parâmetros foram aplicados: tamanho de *amplicon* (entre 70 e 200 bp), temperatura de dissociação [entre 50ºC (mínima), 70ºC (ótima) e 80ºC (máxima), conteúdo GC [entre 45-55% (50% ótimo)], e um grampo CG (*CG clamp*). Os *primers* desenhados para amplificar os candidatos propostos (ScARF43, ScEREB58 ScMYB103, ScNAC5 e ScZIM14) estão disponibilizados na Tabela S2. Os *primers* utilizados como genes de referência foram *αTubulin*, *Histone* H1 e GAPDH, genes de referência disponibilizados pela literatura (Silva *et al.*, 2014). Todos os *primers* foram sintetizados pela Bioneer Corporation (Coreia do Sul) e testados inicialmente em amplificações de cDNAs em PCR convencional, com a programação de ciclagem de 5 min a 95 ºC, seguido de 35 ciclos de 95 °C por 15 s, 58 °C por 30 s, e 72 °C por 30 s. Produtos PCR foram analisados em géis de agarose a 1,5% (w / v) corado com Blue-Green Loading Dye (LGC Biotecnologia®). Curvas de dissociação também foram analisadas em RT-qPCR para

confirmação de especificidade da amplificação dos candidatos propostos. Além disso, curvas de calibração foram feitas utilizando amostras de cDNAs das cultivares tolerantes e sensíveis e de diluições seriadas (1, $10^{-1}$, $10^{-2}$, $10^{-3}$ e $10^{-4}$). Estas curvas tinham o propósito de calcular a eficiência de amplificação ($E = 10^{(-1/standard\ curve\ slope)}$; Rasmussen, 2001) de cada candidato(Figura S1).

Reações RT-qPCR foram realizadas em volume final de 10 μL, incluindo-se: 1 μL da amostra de cDNA (diluída cinco vezes nas absolutas e a $10^{-1}$ nas relativas), 5 μL de *GoTaq® qPCR Master Mix* (Promega), 2 μL de água ultrapura, e 1 μL de cada primer (0.05 μM). Todas as reações RT-qPCR (em placas de 96 poços) foram performadas em três réplicas biológicas e três técnicas, um controle negativo (cDNA sample), e normalizados por três genes de referência (αTubulin, Histone H1 e GAPDH). Reações de RT-qPCR foram realizadas em termociclador LineGene 9660, modelo FQD-96A (Bioer®), utilizando-se o sistema de detecção de SYBR Green. Estas reações seguiram a seguinte programação: desnaturação inicial de 95 ° C por 2 min, seguida por 40 ciclos de 95 ° C por 15 s e 60 ° C por 60 s. As curvas de dissociação foram analisadas entre 65 - 95 ° C por 20 min após os 40 ciclos RT-qPCR. O programa LineGene 9600 (versão 1.1.10) determinou os ciclos de quantificação (Cq), temperatura de dissociação, e quantificações absolutas e relativas. Os dados da expressão relativa foram calculados com o método ΔΔCq e avaliados usando o programa REST software (version 2.0.13) (Pfaffl, Horgan e Dempfle, 2002), utilizando-se teste de randomização com 2000 permutações.

## Resultados

### Identificação e classificação de TFs em cana-de-açúcar sob estresse de seca.

A partir de 2.264 alinhamentos das *unitags* HT-SuperSAGE com transcritos FTs da base de dados GRASSIUS, compreendendo 1.833 *unitags* de cana-de-açúcar após 24 h de supressão de rega, expressas em raízes dos cultivares tolerante (*bulk* T) e/ou sensível (*bulk* S), foram identificadas 582 potenciais FTs. Esses FTs cobriram 38 das 41 famílias descritas para cana-de-açúcar (Tabela 1, Fig. 1), não sendo constatadas componentes somente das famílias BBR/BPC, C2C2-CO e C2C2-YABBY.

Com base no *bulk* T foram detectados 534 potenciais FTs (Fig. 1B), e com base no *bulk* S foram detectados 526 (Fig. 1C). As famílias FTs mais abundantes em isoformas presentes nos *bulk*s (Fig. 1A) ou isoladamente no *bulk* T (Fig. 1B) e no *bulk* S (Fig. 1C) foram AP2/EREB (*bulk* T: 46 isoformas; *bulk* S: 44), bZIP (*bulk* T: 42 isoformas; *bulk* S: 43) e bHLH (*bulk* T: 36 isoformas). De modo geral, a ordem em abundância de isoformas nas famílias FTs foi similar nos dois *bulks* de cultivares. Variações mais evidentes foram observadas na família MYB *Related*, com o *bulk* T apresentando cinco isoformas a mais que o *bulk* S (Fig 1B e 1C).

**Tabela 1.** Famílias de fatores de transcrição identificadas a partir de alinhamentos de *unitags* expressas em genótipos tolerantes de cana-de-açúcar (*Saccharum* spp.) sobre déficit hídrico (24 h de supressão de rega) com sequências nucleotídicas de fatores de transcrição* e suas respectivas regulações ($p > 0,05$).

| **Família** | **Alinhamentos** | **Fatores de transcrição** | ***Unitags***** | ***Unitags*** **UR TExC** | ***Unitags*** **DR TExC** | ***Unitags*** **n.s. TExC** | ***Unitags*** **UR SExC** | ***Unitags*** **DR SExC** | ***Unitags*** **n.s. SExC** |
|---|---|---|---|---|---|---|---|---|---|
| ABI3 | 55 | 24 | 45 | 3 | 3 | 31 | 5 | 3 | 25 |
| Alfin (ALF) | 38 | 7 | 34 | 0 | 3 | 19 | 2 | 1 | 27 |
| AP2/EREB | 293 | 49 | 249 | 16 | 20 | 127 | 26 | 36 | 124 |
| ARF | 51 | 22 | 33 | 1 | 6 | 17 | 1 | 1 | 26 |
| ARID | 14 | 5 | 13 | 0 | 0 | 9 | 0 | 1 | 3 |
| ARR-B | 5 | 3 | 3 | 0 | 0 | 2 | 0 | 0 | 2 |
| bHLH | 120 | 38 | 112 | 6 | 10 | 73 | 13 | 8 | 61 |
| bZIP | 187 | 48 | 161 | 12 | 16 | 86 | 18 | 22 | 81 |
| BZR | 10 | 2 | 10 | 0 | 0 | 8 | 2 | 0 | 4 |
| C2C2-Dof | 16 | 5 | 15 | 0 | 0 | 11 | 3 | 2 | 3 |
| C2C2-GATA | 31 | 8 | 23 | 2 | 6 | 9 | 0 | 6 | 13 |
| C2H2 | 27 | 6 | 23 | 1 | 0 | 19 | 0 | 3 | 11 |
| C3H | 129 | 31 | 89 | 6 | 6 | 58 | 8 | 9 | 48 |
| CCAAT | 47 | 15 | 36 | 7 | 1 | 21 | 3 | 4 | 20 |
| CPP | 13 | 6 | 6 | 1 | 0 | 6 | 0 | 1 | 4 |
| SBP | 11 | 4 | 11 | 1 | 2 | 5 | 0 | 2 | 8 |
| E2F-DP | 5 | 1 | 5 | 1 | 0 | 1 | 0 | 0 | 5 |
| EIL | 12 | 3 | 11 | 0 | 2 | 6 | 0 | 0 | 3 |
| G2 like | 35 | 14 | 25 | 7 | 0 | 12 | 2 | 1 | 17 |
| GeBP | 15 | 3 | 15 | 0 | 0 | 13 | 0 | 1 | 12 |
| GRAS | 131 | 32 | 109 | 7 | 8 | 64 | 11 | 6 | 70 |

**Continuação**

| **Família** | **Alinhamentos** | **Fatores de transcrição** | ***Unitags***** | ***Unitags* UR TExC** | ***Unitags* DR TExC** | ***Unitags* n.s. TExC** | ***Unitags* UR SExC** | ***Unitags* DR SExC** | ***Unitags* n.s. SExC** |
|---|---|---|---|---|---|---|---|---|---|
| GRF | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| Homeobox | 125 | 28 | 122 | 8 | 8 | 79 | 10 | 12 | 64 |
| HSF | 19 | 5 | 18 | 1 | 4 | 11 | 1 | 0 | 9 |
| MADS | 41 | 12 | 33 | 4 | 0 | 20 | 3 | 1 | 19 |
| MYB Related | 91 | 23 | 76 | 2 | 5 | 52 | 8 | 9 | 42 |
| MYB Family | 71 | 17 | 69 | 3 | 4 | 41 | 8 | 2 | 41 |
| NAC | 124 | 28 | 98 | 17 | 5 | 53 | 9 | 8 | 53 |
| NLP | 52 | 2 | 52 | 7 | 6 | 33 | 10 | 8 | 23 |
| Orphans | 214 | 52 | 188 | 17 | 10 | 125 | 14 | 20 | 95 |
| TCP | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| Trihelix | 3 | 3 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| TUB | 70 | 23 | 49 | 4 | 6 | 31 | 5 | 5 | 23 |
| VOZ | 3 | 3 | 2 | 0 | 0 | 1 | 0 | 0 | 2 |
| WHIRLY | 3 | 1 | 3 | 0 | 0 | 2 | 0 | 0 | 2 |
| WRKY | 58 | 20 | 57 | 4 | 9 | 31 | 8 | 5 | 28 |
| ZF-HD | 4 | 1 | 4 | 0 | 0 | 2 | 1 | 0 | 3 |
| ZIM | 139 | 36 | 60 | 11 | 5 | 30 | 8 | 5 | 32 |
| Totais | 2264 | 582 | 1862** | 149 | 145 | 1111 | 179 | 184 | 1004 |

* sequências de *Saccharum* spp. híbrido oriundas do banco de dados GRASSIUS (grassius.org). **com redundância de mesma *unitag* mapeada em trancrito diferente.

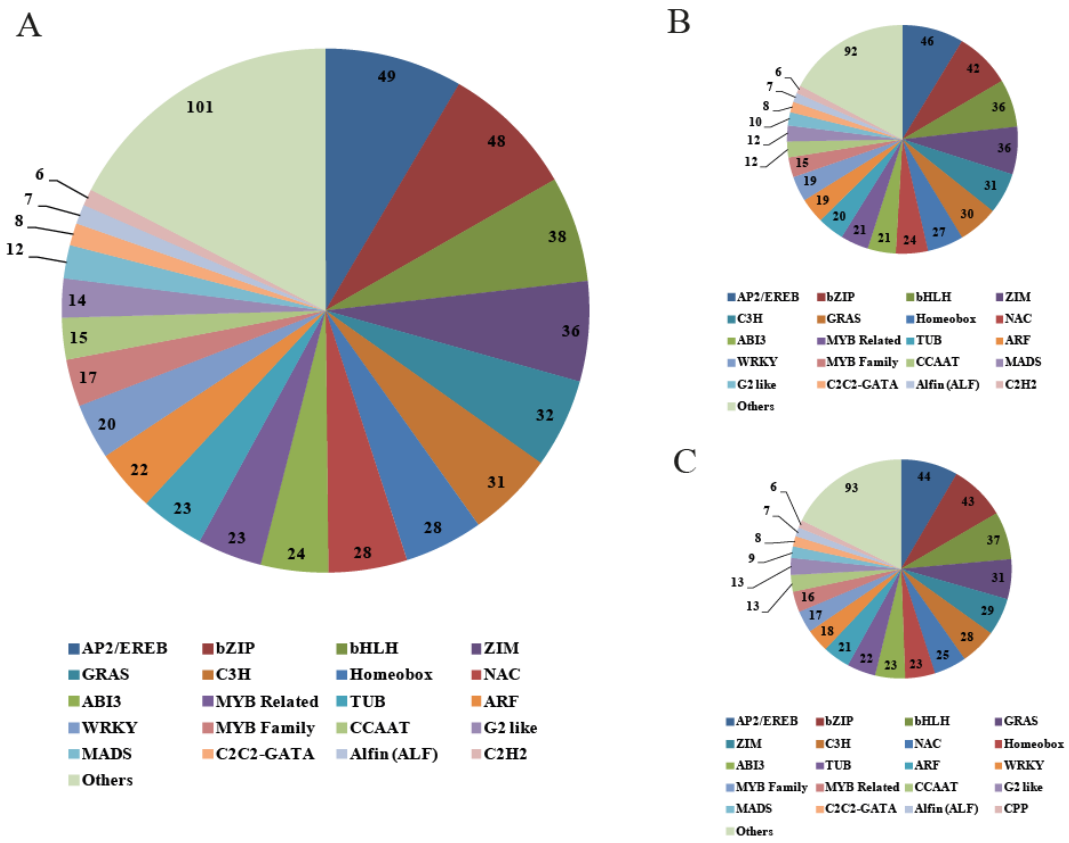



**Fig. 1** Distribuição de FTs por família no transcriptoma de *Saccharum* spp. (A) e respectivas distribuições de FTs por família em acessos tolerantes (B) e acessos sensíveis (C). Os números das 20 maiores famílias FT foram contados separadamente, enquanto que as demais famílias e os membros Órfãos foram integrados em outros.

A curagem das anotações das TFs contendo *unitags* alinhadas foi feita por análise de domínio em sequências traduzidas *in silico* (dados não mostrados), que detectaram domínios em 303 polipeptídios, sendo os mais prevalentes SANT *superfamily* (em 43 peptídeos), ZIM/tify *superfamily* (em 31), bZIP *superfamily* (em 28) e zf-CCCH *superfamily* (em 28).

**Perfis de expressão de fatores de transcrição de cana-de-açúcar em acessos contrastantes**

Com base nas *unitags* expressas pelo *bulk* T sob estresse (24TS) em relação ao respectivo controle negativo (TC), dos FTs identificados, 146 foram considerados induzidos (*unitags* UR, *upregulated*) e 118 reprimidos (*unitags* DR, *downregulated*; $p < 0.05$). Com base no *bulk* S sob estresse (24SS), em relação ao controle negativo sensível

(CS), 133 FTs foram considerados induzidos e 144 reprimidos. A distribuição das famílias diferencialmente expressas nas comparações estudadas [contrastes T (24TS *vs* TC) e S (24SS *vs* SC)] está representada na Fig. 2 De acordo com esta figura, as famílias mais abundantes em FTs diferencialmente expressos pelo *bulk* T foram AP2/EREB (30 FTs), bZIP (22), ZIM (20) e NAC (19); pelo *bulk* S foram AP2/EREB (28), bZIP (26), ZIM (21) e bHLH (19). Diferentemente da distribuição geral, com todas as *unitags* expressas, que apresentou perfis semelhantes nos dois grupos de cultivares, quando se considera as respostas diferencialmente expressas se observam diferenças evidentes na modulação transcricional de parte das famílias em resposta ao estresse aplicado. Por exemplo, a família NAC apresentou mais que o dobro (19 contra oito) em FTs diferencialmente expressos pelo *bulk* T em relação ao *bulk* S, enquanto que a família ARF mais que quintuplicou (11 contra duas) na mesma comparação, ao passo que a família C2C2-GATA, quase dobrou (10 contra seis) (Fig. 2).

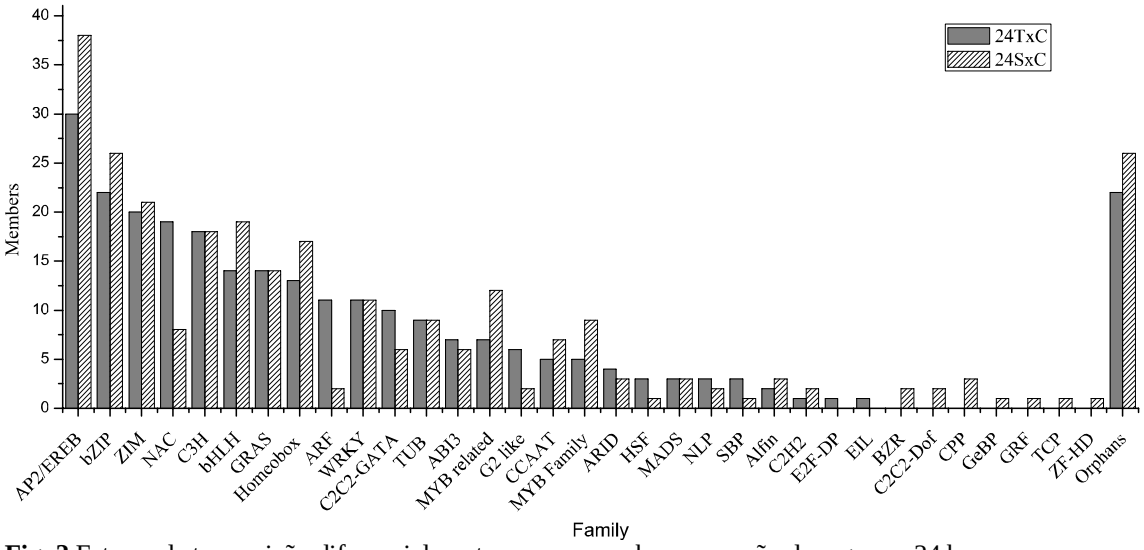



**Fig. 2** Fatores de transcrição diferencialmente expressos sobre supressão de rega por 24 h.

Na tentativa de identificar FTs de respostas expressivas pelo *bulk* T frente ao estresse, os conjuntos de *unitags* UR e DR de cada *bulk* de cultivares foram comparados

em diagrama de Venn (Fig. 3). Assim, 57 FTs apresentando *unitags* UR no contraste T não foram detectadas (n.d.) ou foram consideradas não diferencialmente expressas (n.s.) no contraste S. Deste grupo de 57 FTs, merecem destaques as famílias NAC, GRAS e ARF, com oito, oito e sete isoformas, respectivamente. Outro grupo formado por 31 *unitags* DR no contraste T ou foram n.d. ou n.s. no contraste S, com destaques para as famílias ZIM, C3H e WRKY com 5, 3 e 3 isoformas, respectivamente. Outro grupo de FTs apresentou *unitags* com regulações opostas nos contrastes estudados: 28 FTs com *unitags* UR no contraste T foram consideradas DR no contraste S (destaques para as famílias bZIP, TUB e AP2/EREB, todos com 4 isoformas), ao passo que 16 FTs de *unitags* DR no contraste T foram consideradas UR no contraste S (destaque para a família ZIM com 4 isoformas). Esses conjuntos são particularmente interessantes para estudos futuros, visando desenvolvimento de marcadores moleculares funcionais ou mesmo como alvos moleculares para estudos de transgenia (Tabela S2).

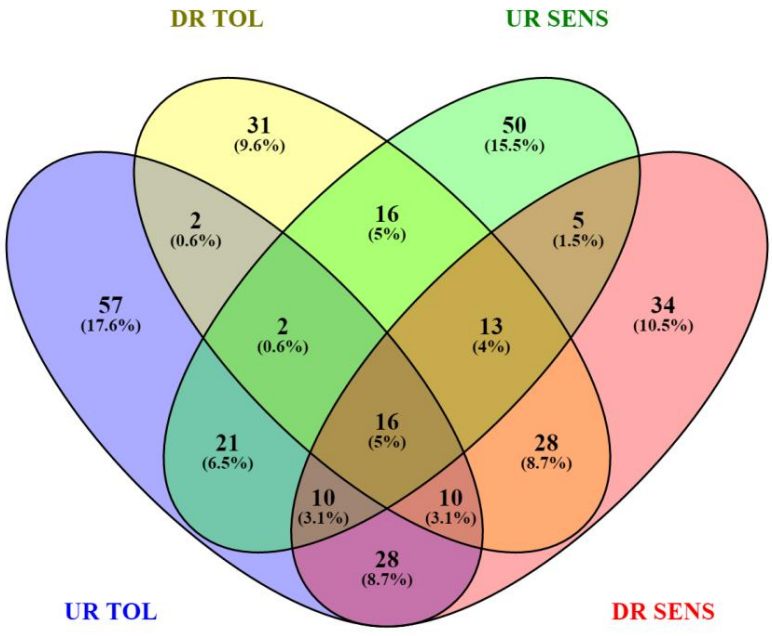



**Fig. 3** Diagrama de Venn evidenciando as respostas dos diferentes *bulks* de cultivares frente ao estresse de supressão de rega por 24 h, a partir de *unitags* HT-SuperSAGE diferencialmente expressas (UR, *upregulated*; DR, *downregulated* em teste Audic-Claverie, $p < 0,05$).

A análise de clusterização hierárquica para os FTs diferencialmente expressos mais expressivos no contraste T (*p-value* < $1e^{-4}$) mostrou a formação de quatro grandes grupos: *Cluster* 1, composto de 16 FTs de *unitags* DR no contraste T e consideradas UR (12 FTs) ou n.s. (1) ou n.d. (2) no contraste S; *Cluster* 2, com 12 TFs de *unitags* DR em ambos os contrastes; *Cluster* 3, com 23 FTs com *unitags* UR no contraste T e DR (13) ou n.s. (6) ou n.d. (4) no contraste S; *Cluster* 4, com 19 FTs de *unitags* UR em ambos os contrastes (Fig. 4). Detalhes envolvendo as 70 *unitags* (composição, regulação, *fold change* e *p-value*) e os alinhamentos com transcritos da base de dados GRASSIUS, estão disponíveis na Tabela S2.

Os *Clusters* 1 e 3, por representarem *unitags* de regulação oposta nas repostas dos *bulks* de cultivares são particularmente interessantes. Dentre os transcritos do *Cluster* 1, que foram expressivamente reprimidos no contraste T e de resposta induzida no contraste S, merecem destaques ScGRAS71, família GRAS (*unitag* SD57840: $FC_{Tol}$ = -1,69 e $FC_{Sens}$ = 1,25), ScNAC58, família NAC (*unitag* SD62818: $FC_{Tol}$ = -15.94 e $FC_{Sens}$ = 2,98), ScEREB23 família AP2/EREB (*unitag* SD8726: $FC_{Tol}$ = -2.33 e $FC_{Sens}$ = 2,08) e ScARF34, família ARF (*unitag* SD271598: $FC_{Tol}$ = -1,74 e $FC_{Sens}$ = 1,59). Dentre os transcritos do *Cluster* 3, merecem destaque os FTs expressivamente induzidos no contraste T e reprimidos no contraste S: ScZIM43, família ZIM (*unitag* SD98547: $FC_{Tol}$ = 5,66 e $FC_{Sens}$ = -1,33), ScZIM46, também da família ZIM (*unitag* SD103749: $FC_{Tol}$ = 2,17 e $FC_{Sens}$ = -1,29), ScEREB104, família AP2/EREB (*unitag* SD148064: $FC_{Tol}$ = 2.68 e $FC_{Sens}$ = -1,62) e ScCA2P3, família CCAAT (*unitag* SD191687: $FC_{Tol}$ = 2,15 e FCSens = -3,18).

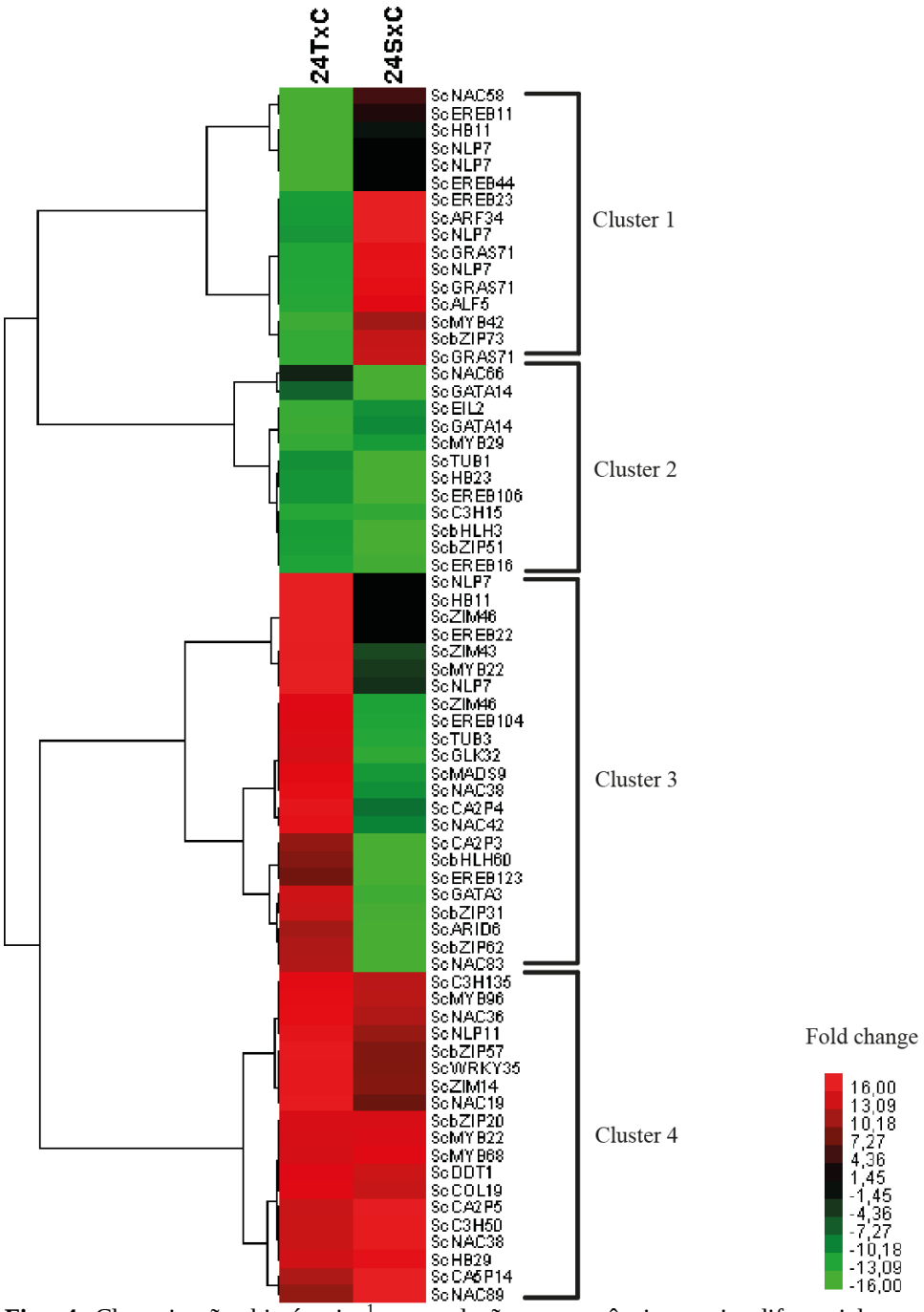



**Fig. 4** Clusterização hierárquica[1] em relação a sequências mais diferencialmente expressas (p <0,0001) em 24TxC (contraste dos acessos tolerantes, estresse de 24 horas e controles controlados) e suas respectivas respostas em 24SxC (contraste dos sensíveis sob estresse de 24 h de supressão de rega e respectivo controle). 1 Pontos vermelhos: *unitags* induzidas; Verde: *unitags* reprimidas; Preto: *unitag* constitutiva ou não detectada.

## Prospecção de *Cis-Acting regulatory elements* (CARE) em regiões promotoras de FTs induzidos

BACs disponíveis de cana-de-açúcar serviram para mapear elementos regulatórios atuando em *Cis* (CARE) de nove FTs de expressão induzida no contraste T.

Foram identificados 3.705 CAREs compreendendo 186 motivos, de um total de 469 da base de dados newPLACE (https://sogo.dna.affrc.go.jp; Higo *et al.*, 1999). Aqueles com prováveis regiões promotoras, delimitadas em 2.000 pb no sentido *upstream* do início dos alinhamentos, envolveram nove FTs (ScMYB103, ScMYB106, ScMADS39, ScbZIP15, ScCOL31, ScGRAS76, ScMYB88, ScARF33 e ScMADS9) e sete BACs (KF184822.1, KF184682.1, KF184769.1, KF184713.1, KF184667.1, KF184668.1 e KF184716.1) (Tabela 2). Os motivos CAREs mais prevalentes detectados em cada região promotora dos nove FTs analisados, bem como as funções previstas para estes motivos, estão dispostos na Tabela 2. Os dez motivos CAREs mais prevalentes dentre os 186 motivos detectados estão dispostos na Figura 5.

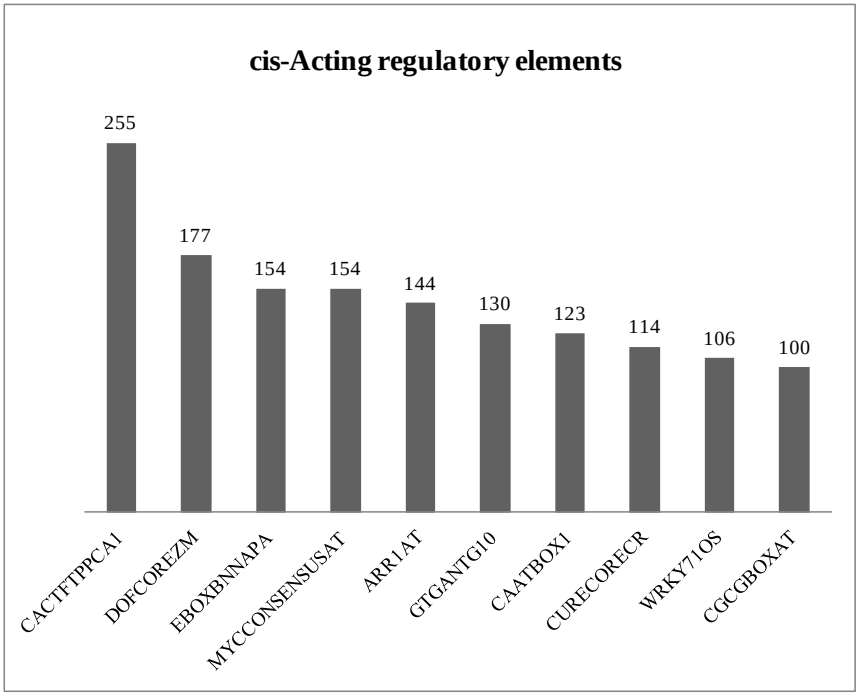



**Fig. 5** Os dez mais *Cis-Acting regulatory elements* de cana-de-açúcar mais prevalentes entre os 186 motivos identificados na região promotora (2,000 pb *upstream*) de nove fatores de transcrição de cana-de-açúcar induzidos ($p$<0.05) em cultivares tolerantes após 24 h de supressão de rega.

**Tabela 2.** Fatores de transcrição UR (*p-value* <0,05) que permitiram extrair 2.000 pares de bases upstream do início do transcrito, em sequências genômicas de cana-de-açúcar (BACs - *bacterial artificial chromossome*), seus *Cis-Acting regulatory elements* (CARE) mais prevalentes, suas respectivas funções e respectivas referências

| Fator de transcrção | *Unitag* (s) | BAC id | Os cinco mais CAREs | Respective quantities | Funções | Referências |
|---|---|---|---|---|---|---|
| | | | CACTFTPPCA1 | 25 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | ACGTATERD1 | 20 | early responsive to dehydration | Simpson *et al*., 2003 |
| ScMYB103 | SD123185 | KF184822.1 | ARR1AT | 18 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | ROOTMOTIFTAPOX1 | 17 | root specific promoter | Elmayan e Tepfer 1995 |
| | | | GTGANTG10 | 16 | regulation of late pollen genes | Rogers *et al*., 2001 |
| | | | CGCGBOXAT | 30 | ethylene signaling, abscisic acid signaling, and light signal perception | Yang e Poovaiah, 2002 |
| | | | GTGANTG10 | 18 | regulation of late pollen genes | Rogers *et al*., 2001 |
| ScMYB106 | SD123185 | KF184822.1 | CACTFTPPCA1 | 16 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | SD123186 | | ARR1AT | 16 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | ACGTATERD1 | 14 | early responsive to dehydration | Simpson *et al*., 2003 |
| | | | DOFCOREZM | 33 | core site required for binding of Dof proteins | Yanagisawa e Schmidt, 1999 |
| | | | CACTFTPPCA1 | 30 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| ScMADS39 | SD25691 SD123660 | KF184682.1 | ARR1AT | 22 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | EBOXBNNAPA | 22 | control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes | Hartmann et., 2005 |
| | | | MYCCONSENSUSAT | 22 | site found in the promoters of the dehydration-responsive gene rd22 | Abe *et al*., 2003 |
| | | | ARR1AT | 24 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | CACTFTPPCA1 | 23 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| ScbZIP15 | SD20988 | KF184769.1 | DOFCOREZM | 22 | core site required for binding of Dof proteins | Yanagisawa e Schmidt, 1999 |
| | | | GTGANTG10 | 21 | regulation of late pollen genes | Rogers *et al*., 2001 |
| | | | WRKY71OS | 19 | binding site of rice WRKY71 | Zhang *et al*., 2004 |

**Continuação**

| Fator de transcrção | *Unitag* (s) | BAC id | Os cinco mais CAREs | Respective quantities | Funções | Referências |
|---|---|---|---|---|---|---|
| | | | CGCGBOXAT | 28 | ethylene signaling, abscisic acid signaling, and light signal perception | Yang e Poovaiah, 2002 |
| | | | CAATBOX1 | 17 | seed storage protein gene | Shirsat *et al*., 1989 |
| ScCOL31 | SD255166 | KF184713.1 | ACGTATERD1 | 16 | early responsive to dehydration | Simpson *et al*., 2003 |
| | | | MYCCONSENSUSAT | 16 | site found in the promoters of the dehydration-responsive gene rd22 | Abe *et al*., 2003 |
| | | | CACTFTPPCA1 | 16 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | CACTFTPPCA1 | 19 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | GTGANTG10 | 18 | regulation of late pollen genes | Rogers *et al*., 2001 |
| ScGRAS76 | SD29759 | KF184667.1 | DOFCOREZM | 17 | core site required for binding of Dof proteins | Yanagisawa e Schmidt, 1999 |
| | | | CGACGOSAMY3 | 16 | cis-elements required for rice alpha-amylase Amy3D expression | Hwang *et al*., 1998 |
| | | | ACGTATERD1 | 16 | early responsive to dehydration | Simpson *et al*., 2003 |
| | | | CACTFTPPCA1 | 51 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | CURECORECR | 28 | copper and hypoxic response | Quinn e Merchant, 1995 |
| ScMYB88 | SD5407 | KF184668.1 | DOFCOREZM | 26 | core site required for binding of Dof proteins | Yanagisawa eSchmidt, 1999 |
| | | | CAATBOX1 | 24 | seed storage protein gene | Shirsat *et al*., 1989 |
| | | | ARR1AT | 19 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | CACTFTPPCA1 | 44 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | DOFCOREZM | 26 | core site required for binding of Dof proteins | Yanagisawa e Schmidt, 1999 |
| ScARF33 | SD88130 | KF184716.1 | MYCCONSENSUSAT | 22 | site found in the promoters of the dehydration-responsive gene rd22 | Abe *et al*., 2003 |
| | | | EBOXBNNAPA | 22 | control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes | Hartmann et., 2005 |
| | | | GATABOX | 21 | conserved GATA motif | Reyes *et al*., 2004 |
| | | | CACTFTPPCA1 | 31 | mesophyll-specific gene expression in the C4 plant | Gowik *et al*., 2004 |
| | | | MYCCONSENSUSAT | 28 | site found in the promoters of the dehydration-responsive gene rd22 | Abe *et al*., 2003 |
| ScMADS9 | SD98266 | KF184682.1 | EBOXBNNAPA | 28 | control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes; | Hartmann et., 2005 |
| | | | ARR1AT | 19 | response regulators operate as transcriptional activators | Sakai *et al*., 2000 |
| | | | CAATBOX1 | 19 | seed storage protein gene | Shirsat *et al*., 1989 |

**Enriquecimento de FTs**

Considerando os transcritos exclusivamente induzidos pelo *bulk* T após o estresse aplicado foram identificadas 8.198 interações entre 235 FTs e 636 genes, sendo que os alvos supra-representados foram detectados para 58 destes FTs (Tabela S3). A partir dos transcritos exclusivamente induzidos pelo *bulk* S, foram detectadas 9.681 interações entre 235 FTs e 778 genes, sendo que os alvos supra-representados foram detectados para 44 dos FTs (Tabela S3). Os alvos supra-representados dos transcritos induzidos pelo *bulk* T foram regulados principalmente por FTs das famílias Dof (5 FTs e 770 interações), MIKC_MADS (5 FTs e 347 interações) e C2H2 (2 FTs e 314 interações) (Figura 7). Os TFs que mais regularam os alvos individualmente foram Sobic.001G106200 (família TALE, apresentando 232 regulações nos tolerantes); Sobic.001G526800 (família MIKC_MADS, com 195 regulações); Sobic.009G014400, Sobic.001G420300, Sobic.006G267900, Sobic.008G136100 e Sobic.004G266200 (todos da família Dof, apresentando 187, 187, 156, 121 e 119, respectivamente). Por sua vez, alvos supra-representados dos transcritos induzidos pelo *bulk* S após o estresse foram regulados principalmente por FTs das famílias Dof (três FTs com 557 interações), C2H2 (dois FTs e 380 interações), AP2/EREB (três FTs e 323 interações) e MYB (nove FTs e 304 interações) (Figura 7). Os TFs que mais regularam os alvos individualmente no *bulk* S foram Sobic.001G106200 (família TALE, apresentando 249 interações; sendo este o mesmo TF responsável por 232 interações nos tolerantes), Sobic.001G420300, Sobic.009G014400 e Sobic.008G136100 (família Dof, com 211, 211 e 135 interações, respectivmente), Sobic.003G058200 (família AP2/EREB, com 194 interações), Sobic.009G024400 e Sobic.008G014801 (família C2H2, ambos regulando 190 alvos).

Assim, diferenças nas respostas transcricionais dos *bulks* de cultivares foram observadas, tais como as ausências de interações de FTs na família B3 e níveis menores de

interação de membros bZIP, Dof, GRAS, G2-like, MIKC_MADS e WRKY, advindas da expressão induzida do *bulk* S em comparação ao *bulk* T (Figura 6). Diferenças foram observadas também ao nível de FTs exclusivos de cada *bulk* , conforme visto no diagrama de Venn (Figura 7). Para os 110 FTs enriquecidos a partir dos dois contrastes, que resultou em 94 FTs distintos (Tabela S3), 16 FTs foram comuns aos dois *bulks* de cultivares (compreendendo as famílias AP2/EREB, bHLH, C2H2, GATA, MYB, MYB_*related*, NAC, TALE e TCP); 57 foram enriquecidos a partir de expressão induzida do *bulk* T (15 famílias: AP2/EREB, ARF, B3, BBR-BPC, bHLH, bZIP, Dof, , GRAS, MIKC_MADS, MYB, NAC, TCP e WRKY) e 21 foram enriquecidos a partir do *bulk* S (11 famílias: AP2/EREB, ARF, ARR-B, BBR-BPC, CAMTA, GATA, HSF, MYB, NAC, TCP e Trihelix).

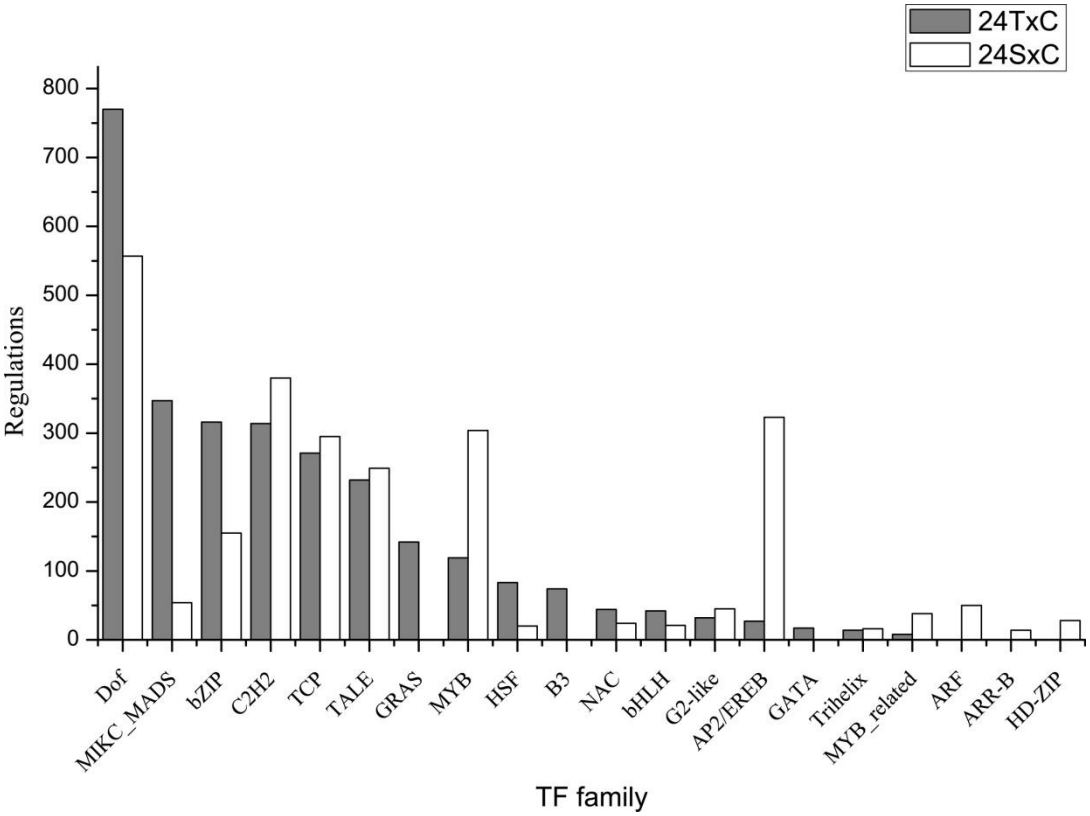



**Fig. 6** Famílias identificadas na análise de *TF Enrichment* em transcritos HT-SuperSAGE exclusivamente induzidos (*p-value* < 0,0001) para cada *bulk* de cultivares de cana-de-açúcar contrastantes e suas espectivas quantidades de interação.

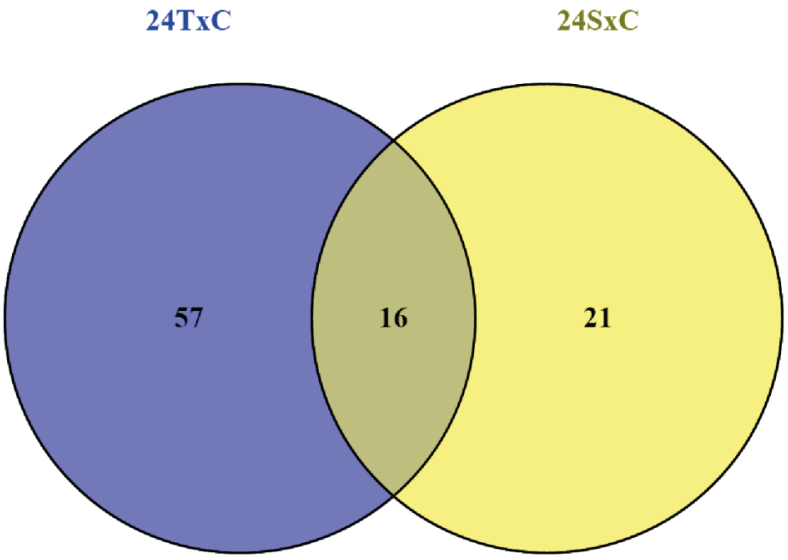



**Fig. 7** Diagrama de Venn destacando o número de FTs com interações com alvos supra-representados, identificados a partir dos transcritos induzidos de cada *bulk* de cultivares após a aplicação do estresse de supressão de rega de 24 h.

**Validação por PCR em tempo real quantitativa de FTs diferencialmente expressos nos dados HT-SuperSAGE**

Para confirmar a acurácia e reprodutibilidade dos dados HT-SuperSAGE, bem como a contribuição individual de cada um dos cultivares, na expressão diferencial detectada a partir dos *bulks*, cinco FTs induzidos após o estresse, de famílias variadas, e que apresentaram *primers* eficientes em análises de RT-qPCR, foram validados seguindo recomendações do protocolo MIQE (Pfaffl *et al.*, 2009). Os FTs selecionados foram *Sc*ARF43 (família ARF), *Sc*EREB58 (família AP2/EREB), *Sc*NAC5 (família NAC), *Sc*MYB103 (G2 like family) e *Sc*ZIM14 (família ZIM) (Tabela 3), cujos resultados de RT-qPCR mostraram concordância geral com as mudanças nas frequências dos transcritos determinados pela análise HT-SuperSAGE (Figura 8 e Tabela 3), sugerindo confiabilidade aos perfis gerados. Assim, o alvo FT *Sc*ARF43, representado pela *unitag* SD88130, considerada UR no contraste T (FCTol = 3,39) e n.s. no contraste S (FCSens = 1,05),

apresentou nos ensaios de RT-qPCR, indução pelo acesso tolerante SP83-2847 (2,41 vezes), repressão pelo cultivar sensível SP90-1638 (0,23 vezes), sendo considerada n.s. para os demais. O alvo FT *Sc*EREB58 [*unitag* SD151691, UR no contraste T (FCTol = 1,81) e DR no contraste S (FCSens = -1,96)] apresentou nos ensaios de RT-qPCR uma indução pelo cultivar tolerante CTC15 (1,91 vezes), repressão pelo cultivar tolerante SP83-2847 (0,35 vezes), sendo considerada n.s. para os demais. O alvo FT *Sc*MYB103 (*unitag* SD123186, anotada como ScGLK32, membro da família G2 like), considerada UR no contraste T (FCTol = 1,73) e n.s. no contraste S (FCSens = -1,16), apresentou nas análises de RT-qPCR indução pelos cultivares tolerantes CTC15 (2,40 vezes) e SP83-2847 (1,50 vezes), repressão (0,61vezes) pelo cultivar tolerante CTC6, sendo considerada n.s. para os demais. Por sua vez, o alvo FT *Sc*NAC5 [*unitag* SD13308, UR no contraste T (FCTol = 4,60) e n.s. no contraste S (FCSens = 1,31)], apresentou nas análises de RTqPCR indução pelos cultivares tolerantes CTC15 (1,76 vezes) e CTC6 (1,53 vezes), sendo considerada n.s. para os demais. Por fim, o alvo FT *Sc*ZIM14 [*unitag* SD258836, UR no contraste T (FCTol = 3,11) e n.s. no contraste S (FCSens = 1,10)], apresentou nas análises de RT-qPCR indução expressiva pelos cultivares tolerantes SP83-5073 (3,34 vezes) e CTC15 (2,03), repressão pelo cultivar tolerante CTC6 (0,46 vezes), indução pelo cultivar sensível SP90-3414 (1,53 vezes), e expressão n.s. para os demais.

**Tabela 3** Genes candidatos propostos e seus respectivos resultados baseados nas análises *in silico* de *unitags* HT-SuperSAGE expressas por expressed *bulks* tolerantes ou sensíveis (quatro cultivares em cada) e suas expressões relativas adiquiridas via RT-qPCR de cDNAs de raízes extraídos após estresse (24 h de supressão de rega) ou sob condições de controle.

| **Técnica e Acessos** | **Genes alvos** | | | | |
|---|---|---|---|---|---|
| | **ScEREB58** | **ScMYB103** | **ScARF43** | **ScNAC5** | **ScZIM14** |
| HT-SuperSAGE[&] | | | | | |
| *Tolerant bulk* | 1,81[#] / UR | 1,73[#] / UR | 3,39[#] / UR | 4,60[#] / UR | 3,11[#] / UR |
| *Sensitive Bulk* | -1,96[#] / DR | -1,16[#] / n.s. | 1,05[#] / n.s. | 1,31[#] / n.s. | 1,10[#] / n.s. |
| RT-qPCR[**] | | | | | |
| *Tolerant accession* | | | | | |
| CTC6 | 1,30 / n.s. | 0,61 / DR | 1,07 / n.s. | 1,53 / UR | 0,46 / DR |
| CTC15 | 1,91 / UR | 2,40 / UR | 1,18 / n.s. | 1,76 / UR | 2,03 / UR |
| SP83-2847 | 0,35 / DR | 1,50 / UR | 2,41 / UR | 0,70 / n.s. | 0,68 / n.s. |
| SP83-5073 | 1,08 / n.s. | 0,69 / n.s. | 0,98 / n.s. | 1,36 / n.s. | 3,34 / UR |
| *Sensitive accession* | | | | | |
| CTC9 | 0,95 / n.s. | 1,81 / n.s. | 1,35 / n.s. | 1,11 / n.s. | 1,17 / n.s. |
| CTC13 | 1,07 / n.s. | 1,47 / n.s. | 0,93 / n.s. | 1,06 / n.s. | 1,42 / n.s. |
| SP90-1638 | 0,76 / n.s. | 0,45 / n.s. | 0,23 / DR | 1,42 / n.s. | 0,87 /n.s. |
| SP90-3414 | 0,77 / n.s. | 0,64 / n.s. | 1,53 / UR | 0,76 / n.s. | 1,53 / n.s. |

#Fold change [FC: ratio of frequencies (normalized to 1,000,000) observed in the stressed library in relation to the respective control library]; ** Relative expression based on the REST software (v.2.0.13) [37] after the ΔΔCq method (expression values were normalized by αTubulin, GAPDH and Histone H1); UR: induced; DR: repressed. ns: not significant at $p < 0.05$ [32]

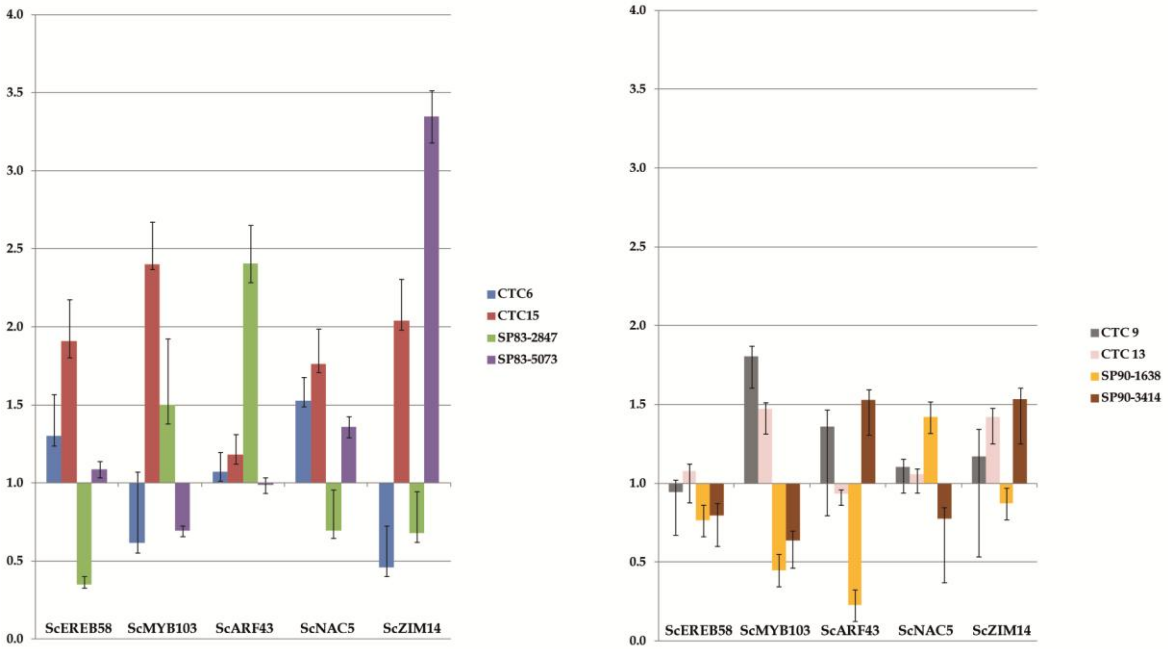



**Fig. 8** RT-qPCR dos genes candidatos performada em cDNAs raízes de cana-de-açúcar sobre déficit hídrico (24 h de desidratação contínua). Os valores de expressão foram normalizados por αTubulina GAPDH e Histona H1. Dados relativos à expressão calculados pelo software REST (v.2.0.13).

**Discussão**

Fatores de transcrição são proteínas que orquestram uma ampla gama de processos celulares através da regulação da expressão de genes alvos, em resposta a diferentes fatores de desenvolvimento e ambientais (Lemon *et al.*, 2000; Jin *et al.*, 2014). Com base nas relevantes funções biológicas, os FTs têm sido extensivamente estudados nas mais importantes culturas, incluindo gramíneas (Burdo *et al.*, 2014; Brito *et al.*, 2015; Liu *et al.*, 2015; Poersch-Bortolon *et al.*, 2016). Apesar da importância econômica da cana-de-açúcar e dos FTs serem excelentes candidatos para modificar caracteres complexos nas plantas (Century *et al.*, 2008), não há na literatura especializada nenhum trabalho apresentando o perfil transcricional do conjunto de TFs expressas por híbridos modernos de cana-de-açúcar, de diferentes respostas ao déficit hídrico, sob este estresse abiótico. Há na literatura, entretanto, relato de FTs de cana-de-açúcar relativos à biossíntese de lignina (Brito *et al.*, 2015) e de eventos de transformação gênica em plantas de tabaco, por exemplo, com TF individual (SodERF3 de *Saccharum officinarum* L. cv Ja60-5), que culminaram acentuando a tolerância das plantas transgênicas para seca e salinidade (Trujillo *et al.*, 2008). Há também evento de transformação de plantas de cana-de-açúcar com TF da família AP2/EREB de *Arabidopsis thaliana* (AtDREB2A CA), acarretando maior tolerância da planta transgênica ao déficit hídrico (Reis *et al.*, 2014). Logo, o trabalho aqui proposto é pioneiro em abordar a resposta imediata (24 h após supressão de rega) do conjunto de FTs expressas em raízes de híbridos comerciais de cana-de-açúcar, com histórias distintas de respostas ao déficit hídrico, conforme informações disponibilizadas pelo programa de melhoramento genético da cana-de-açúcar do Centro de Tecnologia Canavieira (CTC, Piracicaba, SP). Os resultados disponibilizados cobrem um conjunto de FTs e de elementos regulatórios atuantes em *Cis*, e podem ser uteis para estudos de respostas estresse-específica ou mesmo para desenvolvimento de futuros

marcadores moleculares funcionais, com potencial para uso em seleção assistida (SAM) e/ou transgenia.

A anotação funcional do conjunto reportado de FTs (Figura 1, Tabela 1) seguiu o padrão esperado para outras gramíneas (Xiong *et al.*, 2005; Burdo *et al* 2014), estando as famílias AP2/EREB, bZIP e bHLH dentre aquelas mais abundantes em isoformas de FTs observadas em raízes de cana-de-açúcar, tanto no conjunto (Fig. 1A) quanto nos contrastes T (Fig. 1B) e S (Fig. 1C), em separado. As análises de domínio, por sua vez, confirmaram para parte dos polipeptídeos as anotaçoes fornecidas pelo banco GRASSIUS (Figura 2), a qual é uma importante fonte de sequências nucleotídicas e peptídicas para FTs de gramíneas, incluindo milho (Zea mays L.), sorgo (*Sorghum bicolor* (L.) Moench), cana-de-açúcar (*Saccharum* spp. híbrida), arroz (*Oryza sativa* L.) e *Brachipodium* (*Brachypodium distachyon* (L.) Beauv.), sendo muito util para identificação e classificação de FTs neste grupo vegetal (Yilmaz *et al.*, 2009). De modo geral, diferenças na ordem das famílias de FTs expressas pelos *bulks* T e S não foram observadas, com base na abundância de isoformas, indicando que para genes FTs as diferenças são mais ao nível de regulação, conforme verificados na Figs. 3 e 4. Ambos os contrastes (T e S) apresentaram números semelhantes de FTs diferencialmente expressos (264 e 277 isoformas DE, respectivamente), porém as famílias diferiram em abundância de isoformas DE (Fig. 3), sendo muitas delas exclusivas de cada *bulk* de cultivares (Fig. 4), e portando padrões variados de expressão (Fig. 5).

Em relação à família NAC (28 isoformas detectadas; 24 expressas no contraste T e 25 no S), a comparação entre os dois *bulks* mostrou 2,37 vezes mais transcritos DE expressos nos tolerantes que nos sensíveis. Foram oito TFs exclusivamente induzidos pelo *bulk* T (*Sc*NAC5, *Sc*NAC19, *Sc*NAC22, *Sc*NAC23, *Sc*NAC42, *Sc*NAC43, *Sc*NAC46 e *Sc*NAC47), sendo que para *Sc*NAC5 (*unitag* SD13308) as análises de RT-qPCR, com cada

cultivar isoladamente, validou a expressão UR para dois cultivares tolerantes [CTC15 (1,76 vezes); CTC6 (1,53 vezes)]. Os FTs NAC apresentam uma variedade de funções no desenvolvimento de plantas e em respostas a estresses abióticos (Nakashima *et al.*, 2012), sendo que a expressão induzida de genes NAC em resposta a estresses abióticos fora reportada para frio (4$^{o}$C), seca, salinidade (250 mM NaCl) e também para ação de fitoreguladores, tais como ABA e JA, basicamente em arroz e *A. thaliana* (Rabbani *et al.*, 2003;Tran *et al.*, 2004; Ohnishi *et al.*, 2005). A superexpressão de gene NAC estresse responsivo (SNAC10) em raízes de plantas transgênicas de arroz acentuaram a tolerância à seca (Jeong *et al.*, 2010).

Recentemente, plantas de trigo expressando o gene SNAC1 de arroz mostraram significava tolerância a seca e salinidade, em múltiplas gerações, tendo controlado de modo eficiente a expressão de genes envolvidos na sinalização via ABA, incluindo PI3K (*1-phosphatidylinositol-3-phosphate-5-kinase*), SPS (*sucrose phosphate synthase*), PP2C3 (*type 2C protein phosphatases*) e RCAR (*regulatory components of ABA receptor*) (Saad *et al.*, 2013). Transcritos destes genes também apresentaram *unitags* HT-SuperSAGE do presente ensaio sendo expressas, por exemplo: PI3K (*unitag* SD241907; $FC_{tol}$ = 3,18 e $FC_{sens}$ = -1,42), SPS (*unitag* SD80522; $FC_{tol}$ = 3,39 e $FC_{sens}$ = 1,88), PP2C3 [duas *unitags*: SD86612 ($FC_{tol}$ = 2,97 e $FC_{sens}$ = 1,46) e SD42438 ($FC_{tol}$ = -8,65 e $FC_{sens}$ = -1,14)] e RCAR [duas *unitags*: SD122705 ($FC_{tolerante}$ = 1,97 e $FC_{sensível}$ = -1.70) e SD122727 ($FC_{tolerante}$ = 2,20 e $FC_{sensível}$ = -1,42)]. De acordo com Saad *et al.*, 2013, resultados de RT-qPCR nas plantas de trigo transgênicas, mostraram para PI3K, SPS e RCAR, indução (de 1,3 a 3,4 vezes) em comparação com plantas controle não transgênicas, e para PP2C3, a expressão mostrou-se claramente reprimida em dois de três eventos transgênicos, estando estes resultados em acordo com as análises HT-SuperSAGE do presente trabalho. Assim, os FTs NAC DE identificados ou seus alvos validados podem ser úteis como marcadores

moleculares funcionais para ensaios de SAM ou para estudos de transgenia, visando um aumento da tolerância ao estresse hídrico.

Em relação à família ARF (*auxin response factor*) [22 isoformas detectadas; 19 expressas pelo *bulk* T e 21 pelo *bulk* S], a comparação entre os *bulks* mostrou 5,50 vezes mais isoformas DE pelo *bulk* T que pelo S (11 *versus* dois), sendo que sete apresentaram apenas *unitags* UR no contraste T (*Sc*ARF9, *Sc*ARF12, *Sc*ABI17, *Sc*ARF28, *Sc*ARF31, *Sc*ARF33 e *Sc*ARF43). Esta última (*unitag* SD88130) teve a expressão UR validada via RT-qPCR para o cultivar tolerante SP83-2847 e de expressão oposta (DR) para o cultivar sensível SP90-1638 (e n.s. para os demais). Transcritos da família ARF são proteínas de ligação a sequências especificas de DNA, que disparam mudanças transcricionais em resposta à auxina, através da seleção de genes-alvo, atuando como FTs (Li *et al.*, 2016; Weijers e Wagner 2016). Análises genômicas globais de genes da família ARF foram conduzidas em importantes gramíneas cultivadas, como arroz (Wang *et al* 2007), milho (Xing *et al.*, 2011) e sorgo (Wang *et al.*, 2010), o qual é considerado o diplóide mais próximo da cana-de-açúcar (Dillon *et al*, 2007 ). Genes ARF foram implicados na sinalização para IAA e BR, e em respostas para frio, sal e seca (Wang *et al.*, 2010; Hu *et al.*, 2015; Xu *et al.*, 2016), estando envolvidos nas respostas aos estresses abióticos em múltiplas vias de transdução de sinais. Análises de RT-qPCR também detectaram a superexpressão de genes *Sb*ARF10, 16, e 21 em raízes de plantas de sorgo sob estresse de seca (Wang *et al.*, 2010), estando concordantes com as análises HT-SuperSAGE e de validação de *Sc*ARF43, logo, há potencial para exploração destes alvos pelo melhoramento genético.

Em relação à família AP2/EREB foram detectadas 49 isoformas [46 a partir do contraste T (30 consideradas DE) e 44 a partir do S (38 DE)]. Essa família específica de plantas é uma das mais ricas em FTs (Yamasaki *et al.*, 2013), apresentando, em gramíneas,

de 121 (em *Hordeum vulgare* L.; Guo *et al.*, 2016) a 163 genes (em arroz; Sharoni *et al.*, 2011). A isoforma *Sc*EREB58 (*unitag* SD151691) teve sua expressão validada por RT-qPCR, mostrando-se UR pelo cultivar tolerante CTC15. Entretanto, outro cultivar tolerante (SP83-2847) mostrou regulação oposta (DR), enquanto os demais não modularam significativamente a expressão (Tabela 3, Figura 8). Essa família de FTs conduz importantes papéis na regulação transcricional, em diferentes fases do desenvolvimento vegetal: na germinação (Wang *et al.*, 2008), no desenvolvimento embrionário (Mantiri *et al.*, 2008), no florescimento (Krizek e Fletcher, 2005), na frutificação (Karlova *et al.*, 2011), bem como em respostas a estresses bióticos (Sharoni *et al.*, 2011) e abióticos, incluindo salinidade (Guo *et al.*, 2016), seca (Guo *et al.*, 2016; Sharoni *et al.*, 2011), frio (Du *et al.*, 2016), entre outros. A expressão de TF SodERF3 de *Saccharum officinarum* L. cv Ja60-5 em plantas transgênicas de tabaco, quando tratadas com ABA (0,1 mM) ou sob estresse salino (350 mM de NaCl) acentuou a tolerância destas plantas à seca (30 dia de supressão de rega) e salinidade (Trujillo *et al.*, 2008). Ainda, a transformação com FT *At*DREB2A CA (de *A. thaliana*) em plantas de cana-de-açúcar levou a uma maior tolerância ao déficit hídrico (2 – 4 h após supressão de rega), sem prejuízos na biomassa, como resultado da ativação de genes envolvidos em respostas ao déficit hídrico, tais como *galactinol synthase* (metabolismo da rafinose), P5CS (*delta-1-pyrroline-5-carboxylate synthetase*), *late embryogenesis abundant proteins* (LEA), DHY (*dehydrins*), e os genes *Saccharum Drought Responsive* 2 e 4 (SCDRs) (Reis *et al.*, 2014). Transcritos destes genes contendo *unitags* HT-SuperSAGE do presente ensaio sendo expressas foram detectadas para P5CS (síntese de prolina; SD130985; $FC_{tol}$ = 5,81 e $FC_{sens}$ = -1,10) e DHY (SD83601; $FC_{tol}$ = 5,17). Logo, FTs DE da família AP2/EREB podem também ser valiosas para estudos visando o entendimento das respostas transcricionais da cana-de-açúcar frente a estresses abióticos, com aplicações nos programas de melhoramento.

Em relação à família ZIM (*Zinc-finger protein expressed in Inflorescence Meristem*) foram detectadas 36 isoformas [36 detectadas no contraste T (20 TFs DE) e 29 no contraste S (21 TFs DE)]. A família ZIM, também definida como família gênica TIFY, em razão da conservação do padrão de aminoácidos TIF[F/Y]XG (Vanholme *et al.*, 2007), também participa de respostas a fitoreguladores (Shi *et al.*, 2016), estresses bióticos (Zhang *et al.*, 2015) e abióticos (Zhang *et al.*, 2015; Huang *et al.*, 2016; Zhao *et al.*, 2016). As análises *in silico* mostraram ao menos três FTs com destacadas modulações da expressão pelo *bulk* T: *Sc*ZIM43 (*unitag* SD98547; $FC_{Tol}$ = 5,66 e $FC_{Sens}$ = -1,33), *Sc*ZIM46 (*unitag* SD103749; $FC_{Tol}$ = 2,17 e $FC_{Sens}$ = -1.29) e *Sc*ZIM14 (*unitag* SD258836; $FC_{Tol}$ = 3,11 e $FC_{Sens}$ = 1,10). A validação via RT-qPCR deste último mostrou ser UR para os cultivares tolerantes SP83-5073 (3,34 vezes) e CTC15 (2,03 vezes) e para o cultivar sensível SP90-3414 (1,53 vezes); expressão DR foi observada para o cultivar tolerante CTC6 (0,46 vezes) e n.s. para os demais. A regulação UR detectada também em acesso sensível parece não ser suficiente para um fenótipo tolerante devido à complexidade do caráter (Hu *et al.*, 2014). A expressão UR para sete genes ZIM/TIFY foi reportada em plantas de algodão após uma hora sob estresse hídrico, em condições de 17 % PEG 6000 (Zhao *et al* 2016). Os mesmos autores reportaram ainda a superexpressão do gene GaJAZ5 (família TIFY) em *A. thaliana*, culminando com o decréscimo da perda de água, da abertura estomatal e do acúmulo de $H_2O_2$, sugerindo ser a família vital, por serem reguladores importantes na sinalização de JA, nas respostas ao estresse de déficit hídrico e frio (Chini *et al.*, 2007).

A família CCAAT, cujos membros são conhecidos como Fator Y ou HAP, estão presentes em eucariotos superiores (Laloum *et al.*, 2013) conferindo funções importantes na tolerância a estresses (Nelson *et al.*, 2007) e em vias de respostas transcricionais (Park *et al.*, 2010). Das 15 isoformas detectadas (cinco delas DE) mereceram destaques as

isoformas ScCA2P4 (*unitag* SD272530; $FC_{Tol}$ = 3,90 e $FC_{Sens}$ = -1,46) e ScCA2P3 (*unitag* SD191687; $FC_{Tol}$ = 2,15 e $FC_{Sens}$ = -3,18). Plantas transgênicas de milho expressando ZmNF-YB2 (membro da família CCAAT) mostraram tolerância aumentada à seca, proporcionando melhor conteúdo de clorofila, condutância estomática e manutenção da fotossíntese, com reflexos também em condições de campo (Nelson *et al.* 2007). Tal fato é indicativo de que ScCA2P3 e ScCA2P4 são alvos interessantes para aumentar a tolerância da cana-de-açúcar, com aplicações em programas de melhoramento.

Fatores de transcrição da família G2 *like* (Golden2 *like*) foram inicialmente caracterizados em milho como reguladores de processos de diferenciação do cloroplasto (Rossini *et al.*, 2001). Estudos posteriores confirmaram essa participação (Chen *et al.*, 2016), bem como em vias de resistência ao *Fusarium graminearum* em *A. thaliana* (Savitch, *et al* 2007), na regulação do desenvolvimento e maturação de frutos (Nguyen *et al.*, 2014), e em resposta a estresses abióticos, dentre os quais frio e seca (Liu *et al.*, 2016). Um total de 59 genes G2 *like* foram identificados no genoma do milho e 20 transcritos G2 *like* tiveram suas expressões monitoradas em tecido foliar de plantas sob tratamento de frio (4 °C por 6 h, 12 h e 24 h) e seca (20% PEG6000, no intervalo de 30 min - 6 h), sendo que 10 destes transcritos foram UR em ambos os tratamentos (Liu *et al.*, 2016). Na presente análise *in silico* foram detectadas 14 isoformas [10 expressas no contraste T (seis TFs DE) e 13 no contraste S (2 TFs DE)]. O alvo FT *Sc*MYB103 (ou ScGLK32), representado pela *unitag* SD123186 (FCTol = 1,73 e FCSens = -1,16), apresentou nas análises de RT-qPCR indução por dois cultivares tolerantes [CTC15 (2,40 vezes) e SP83-2847 (1,50 vezes)], sendo que outro cultivar tolerante (CTC6) apresentou expressão divergente (0,61vezes), enquanto os demais foram de expressão n.s. A provável região promotora deste alvo também foi caracterizada (ver item a seguir). As análises *in silico*, juntamente com a validação em raízes de ao menos dois cultivares tolerantes, bem como a caracterização de

uma região promotora, aumentam o repertório informativo da família frente ao déficit hídrico, em um tecido ainda pouco estudado (raiz), talvez em razão da origem dos tecidos (foliar) das primeiras caracterizações (Rossini *et al.*, 2001), logo, se constituem em bons candidatos também para caracterização funcional em tecido radicular.

A transformação genética de plantas com promotores indutíveis fornece uma regulação mais precisa da expressão trangenica através de um controle externo (Hernandez-Garcia e Finer, 2014), aumentando, assim, as chances de que a produtividade da planta transformada não seja afetada em condições normais de cultivo. Logo, a busca por esses promotores em genoma de referência constitui uma estratégia importante a ser perseguida. Entretanto, a complexidade do genoma da cana-de-açúcar, de natureza poliplóide e aneuplóide (Wang *et al.*, 2010), fazem com que a montagem de um genoma de referência seja um desafio maior. Há, no entanto, fragmentos genômicos disponíveis em cromossomos artificiais de bactérias (BACs; Rezende *et al.*, 2012; Setta *et al.*, 2014; Visendi *et al.*, 2016). Assim, dos 312 BACs disponibilizados por Setta *et al.* (2014), após prospecção por região promotora, definida como 2.000 pb *upstream* ao início dos alinhamentos com transcrito TF, em somente nove foram detectadas regiões maiores (o tamanho médio da maioria dos BACs foi de 115 kb).

Dentre os CAREs identificados a partir destas regiões promotoras merecem destaques, pela prevalência e por estarem relacionados diretamente com o estresse de seca, os elementos indutiveis MYCCONSENSUSAT e ACGTATERD1, (Abe *et al.*, 2003; Simpson *et al.*, 2003), além de DOFCOREZM (o segundo mais prevalente), requerido para ligação de proteínas Dof (Yanagisawa e Schmidt, 1999). Envolvimento de TFs Dof tem sido reportados em respostas de plantas a estresses abióticos (Gupta *et al.*, 2016) e também bióticos (Kang *et al.*, 2016). Desta forma, a presença destes CAREs em promotores candidataria estes promotores a uma possível natureza estresse-específica em cana-de-

açúcar, para o estresse em questão, sendo necessários, no entanto, estudos complementares que confirmem essa regulação/ indução, bem como validem os resultados em plantas transgências devidamente avaliadas.

Experimentos de alto desempenho como microarranjos e de sequenciamentos de segunda geração, têm gerado quantidades enormes de dados de expressão, depositados em bancos públicos, podendo estes ser reutilizados ou reanalisados por diferentes grupos de pesquisa. Nestas análises, novas ferramentas estão sempre buscando comparações com os dados depositados em situações específicas ainda não investigadas (Rung e Brazma, 2012). A ferramenta *TF enrichment* da base de dados PlantTFDB 4.0 (Jin *et al.*, 2016; http://plantregmap.cbi.pku.edu.cn/tf_enrichment.php) é uma delas, ao proporcionar encontrar FTs que possuam alvos significativamente supra-representados nos genes de entrada (*input genes*), se utilizando de dados de ChIP-seq e de regulações transcricionais identificadas na literatura.

Considerando os transcritos contendo *unitags* HT-SuperSAGE exclusivamente induzidas por cada *bulk* após o estresse, a ferramenta identificou 94 TFs distintos envolvidos em 17.879 regulações (8.198 no *bulk* T e 9.681 no *bulk* S). Na expressão dos cultivares tolerantes, FTs das famílias Dof e TALE (*Three Amino acid Loop Extension*) seriam responsáveis individualmente por mais de 100 regulações. Um deles, FT Sobic.001G106200 (TALE), responsável por mais de 200 regulações em ambos os *bulks*, participaria da ativação de diversos genes em diferentes estratégias de respostas e mitigação ao déficit hídrico, sendo um bom candidato para transformação gênica devido ao amplo expectro de atuação após o curto tempo de exposição ao estresse. TALE é uma superclasse de genes *homeobox* conservados em animais, plantas e fungos, sendo subdividida em duas classes (*KNOTTED1-LIKE homeobox - KNOX and BEL1-like homeobox - BLH or BELL*) envolvidas principalmente no controle do desenvolvimento de

tecidos vegetais, como meristemas apicais, folhas, colmos e inflorescências (Di Giacomo, Iannelli e Frugis, 2013). Membros das duas classes respondem a estresses abióticos, dentre os quais: desidratação, sal (200 mm de NaCl) e frio (4 ± 1 $^{o}$C) (Jain *et al.*, 2008). Estes FTs seriam também candidatos reguladores de amplo espectro para ensaios futuros de transgenia.

Outro conjunto potencialmente valioso compreende os FTs Dof (*DNA-binding One Zinc Finger*; 614 interações / cinco FTs: Sobic.009G014400, Sobic.001G420300, Sobic.006G267900, Sobic.008G136100 e Sobic.004G266200). *Unitags* HT-SuperSAGE identificaram cinco FTs Dof distintos, mas de expressão DE nos cultivares sensíveis (três UR; duas DR), e n.s. ou não detectada nos tolerantes. Vale resaltar que o elemento regulatório requerido para ligação de proteínas Dof (DOFCOREZM) foi o segundo CARE mais prevalente no conjunto das nove potenciais regiões promotoras, reforçando a participação deste grupo de TFs na transcrição de genes de cana-de-açúcar. A família DOFREZM é exclusiva de plantas, com atuações nas respostas de diversos genes envolvidos em variados processos biológicos, que incluem produção de fitoreguladores e adaptação ambiental (Yanagisawa, 2004), participando das respostas a estresses abióticos (Gupta *et al.*, 2016). Logo, proteínas Dof e sítios de regulação DOFCOREZM podem ser alvos interessantes em ensaios de transformação estresse-específica, visando ativações de genes de amplo espectro.

**Considerações Finais**

O presente estudo permitiu verificar a resposta global de FTs expressos em raízes de cana-de-açúcar, de híbridos modernos, subdivididos em dois grupos de cultivares (tolerantes ou sensíveis ao déficit hídrico), frente a um estresse imediato de supressão de rega (após 24 h). Representantes de quase todas as famílias de FTs foram detectados a partir de *unitags* HT-SuperSAGE, com destaques para as famílias mais ricas em FTs

responsivos a fatores ambientais e que, segundo a literatura, controlam ampla gama de genes relacionados com respostas a estresses abióticos. Foram identificados prováveis genes FTs que controlam a expressão dos transcritos HT-SuperSAGE mais superexpressos nos dois *bulks* de cultivares contrastantes, em resposta ao estresse aplicado, com destaques para TFs responsáveis por amplo número de regulações, em termos de alvos supra-representados nos genes induzidos de cada *bulk*. Estes FTs são candidatos potenciais para transformação gênica de plantas com genes FTs de amplo espectro, visando aumentar a resposta tolerante ao défict hídrico. Adicionalmente, elementos regulatórios atuantes em *Cis*, prevalentes em nove regiões promotoras de FTs de cana-de-açúcar, também foram identificados e controlariam genes relacionados com a resposta a estresses abióticos e com a sinalização em respostas a fitoreguladores. Estas regiões promotoras também são candidatos naturais para ensaios de transgenia, na indução de uma resposta estresse-específica, o que em tese diminuiria potenciais prejuízos à produtividade da planta transformada em situações de não estresse. Logo, o esforço aqui apresentado disponibiliza valiosos recursos, representados pelo conjunto de FTs expressos em raízes de plantas de cana-de-açúcar, sugerindo grande potencial biotecnológico, tanto como candidatos naturais para estudos futuros de SAM, quanto em estudos de transgenia, na espera de uma resposta tolerante especifica ao estresse estudado, em condições de estresse, com uso dos elementos regulatórios aqui identificados.

**References**

Abe H, Urao T, Ito T, Seki M and Shinozaki K (2003) Transcriptional Activators in Abscisic Acid Signaling. Society 15:63–78. doi: 10.1105/tpc.006130.salt

Audic S and Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7:986–995. doi: 10.1101/gr.7.10.986

Boyko A and Kovalchuk I (2008) Epigenetic Control of Plant Stress Response. Environ Mol Mutagen 49:61-72.

Brito MS, Nobile PM, Bottcher A, dos Santos AB, Creste S, de Landell MGA, Vincentz M, Vicentini R and Mazzafera P (2015) Expression Profile of Sugarcane

Transcription Factor Genes Involved in Lignin Biosynthesis. Trop Plant Biol 8:19–30. doi: 10.1007/s12042-015-9147-y

Burdo B, Gray J, Goetting-Minesky MP, Wittler B, Hunt M, Li T, Velliquette D, Thomas J, Gentzel I, Dos Santos Brito M *et al.* (2014) The Maize TFome - Development of a transcription factor open reading frame collection for functional genomics. Plant J 80:356–366. doi: 10.1111/tpj.12623

Bustin S a, Benes V, Garson J a, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL *et al.* (2009) The MIQE guidelines:Minimum Information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622. doi: 10.1373/clinchem.2008.112797

Century K, Reuber TL and Ratcliffe OJ (2008) Regulating the regulators: the future prospects for transcription-factor-based agricultural biotechnology products. Plant Physiol 147:20–29. doi: 10.1104/pp.108.117887

Chen L, Song Y, Li S, Zhang L, Zou C and Yu D (2012) The role of WRKY transcription factors in plant abiotic stresses. Biochim Biophys Acta - Gene Regul Mech 1819:120–128. doi: 10.1016/j.bbagrm.2011.09.002

Chen M, Ji M, Wen B, Liu L, Li S, Chen X, Gao D and Li L (2016) GOLDEN 2-LIKE Transcription Factors of Plants. Front Plant Sci 7:1–5. doi: 10.3389/fpls.2016.01509

Chini A, Fonseca S, Fernandez G, Adie B, Chico JM, Lorenzo O, Garcia-Casado G, Lopez-Vidriero I, Lozano FM, Ponce MR, Micol JL, Solano R. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–671. doi: 10.1038/nature06006.

Di Giacomo E, Iannelli M and Frugis G (2013) TALE and Shape: How to Make a Leaf Different. Plants 2:317–342. doi: 10.3390/plants2020317

Dillon SL, Shapter FM, Henry RJ, Cordeiro G, Izquierdo L and Lee LS (2007) Domestication to crop improvement: Genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). Ann Bot 100:975–989. doi: 10.1093/aob/mcm192.

Du C, Hu K, Xian S, Liu C, Fan J, Tu J and Fu T (2016) Dynamic transcriptome analysis reveals AP2/EREB transcription factors responsible for cold stress in rapeseed (Brassica napus L.). Mol Genet Genomics 291:1053–1067. doi: 10.1007/s00438-015-1161-0

Eisen MB, Spellman PT, Brown PO, Botstein D (1999) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci 95: 12930–12933.

Elmayan T and Tepfer M (1995) Evaluation in tobacco of the organ specificity and strength of the rolD promoter, domain A of the 35S promoter and the 35S2 promoter. Transgenic Res 4:388–396. doi: 10.1007/BF01973757

FAOSTAT - Food and Agriculture Organization of the Unite Nations: FAO Statistical Databases. http://faostat.fao.org (2016). Accessed 18 Dec 2016.

Feller A, MacHemer K, Braun EL and Grotewold E (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J 66:94–116. doi: 10.1111/j.1365-313X.2010.04459.x

Ferreira Neto JR, Silva MD, Pandolfi V, Crovella S, Benko-Iseppon AM, Kido EA (2016) Epigenetic Signals on Plant Adaptation: A Biotic Stress Perspective. Current Protein & Peptide Science 18:352–367. doi: 10.2174/1389203717666160724194653

Gao F, Xiong A, Peng R, Jin X, Xu J, Zhu B, Chen J and Yao Q (2010) OsNAC52, a rice NAC transcription factor, potentially responds to ABA and confers drought tolerance in transgenic plants. Plant Cell Tissue Organ Cult 100:255–262. doi: 10.1007/s11240-009-9640-9

Gosal SS, Wani SH and Kang MS (2009) Biotechnology and Drought Tolerance. J Crop Improv 23:19–54. doi: 10.1080/15427520802418251

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M and Westhoff P (2004) cis-Regulatory Elements for Mesophyll-Specific Gene Expression in the C 4 Plant Flaveria trinervia , the Promoter of the C 4 Phosphoenolpyruvate Carboxylase Gene. Plant Cell 16:1077–1090. doi: 10.1105/tpc.019729.C

Guo B, Wei Y, Xu R, Lin S, Luan H, Lv C, Zhang X, Song X and Xu R (2016) Genome-wide analysis of APETALA2/ethylene-responsive factor (AP2/EREB) Gene family in barley (*Hordeum vulgare* L.). PLoS One 11:1–17. doi: 10.1371/journal.pone.0161322

Gupta S, Arya GC, Malviya N, Bisht NC and Yadav D (2016) Molecular cloning and expression profiling of multiple Dof genes of Sorghum bicolor (L) Moench. Mol Biol Rep 43:767–774. doi: 10.1007/s11033-016-4019-6

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98. doi: citeulike-article-id:691774

Han S-K and Wagner D (2013) Role of chromatin in water stress responses in plants. J Exp Bot 65:2785–2799. doi: 10.1093/jxb/ert403

Hartmann U, Sagasser M, Mehrtens F, Stracke R and Weisshaar B (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. Plant Mol Biol 57:155–171. doi: 10.1007/s11103-004-6910-0

Hernandez-Garcia CM and Finer JJ (2014) Identification and validation of promoters and cis-acting regulatory elements. Plant Sci 217–218:109–119. doi: 10.1016/j.plantsci.2013.12.007

Higo K, Ugawa Y, Iwamoto M and Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27:297–300. doi: 10.1093/nar/27.1.297

Hu H and Xiong L (2014) Genetic engineering and breeding of drought-resistant crops. Annu Rev Plant Biol 65:715–41. doi: 10.1146/annurev-arplant-050213-040000

Hu W, Zuo J, Hou X, Yan Y, Wei Y, Liu J, Li M, Xu B and Jin Z (2015) The auxin response factor gene family in banana: genome-wide identification and expression analyses during development, ripening, and abiotic stress. Front Plant Sci 6:742. doi: 10.3389/fpls.2015.00742

Huang Z, Jin S-H, Guo H-D, Zhong X-J, He J, Li X, Jiang M-Y, Yu X-F, Long H, Ma M-D *et al.* (2016) Genome-wide identification and characterization of TIFY family genes in Moso Bamboo (*Phyllostachys edulis*) and expression profiling analysis under dehydration and cold stresses. PeerJ 4:e2620. doi: 10.7717/peerj.2620

Hwang YS, Karrer EE, Thomas BR, Chen L and Rodriguez RL (1998) Three cis-elements required for rice α-amylase Amy3D expression during sugar starvation. Plant Mol Biol 36:331–341. doi: 10.1023/A:1005956104636

Jain M, Tyagi AK and Khurana JP (2008) Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. FEBS J 275:2845–2861. doi: 10.1111/j.1742-4658.2008.06424.x

Jeong JS, Kim YS, Baek KH, Jung H, Ha S-H, Do Choi Y, Kim M, Reuzeau C and Kim J-K (2010) Root-Specific Expression of OsNAC10 Improves Drought Tolerance and Grain Yield in Rice under Field Drought Conditions. Plant Physiol 153:185–197. doi: 10.1104/pp.110.154773

Jin J, Zhang H, Kong L, Gao G and Luo J (2014) PlantTFDB 3.0: A portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res 42:1182–1187. doi: 10.1093/nar/gkt1016

Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J and Gao G (2016) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res gkw982. doi: 10.1093/nar/gkw982

Kang WH, Kim S, Lee HA, Choi D and Yeom SI (2016) Genome-wide analysis of Dof transcription factors reveals functional characteristics during development and response to biotic stresses in pepper. Sci Rep 6:33332. doi: 10.1038/srep33332

Karlova R, Rosin FM, Busscher-Lange J, Parapunova V, Do PT, Fernie AR, Fraser PD, Baxter C, Angenent GC and de Maagd R a (2011) Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. Plant Cell 23:923–41. doi: 10.1105/tpc.110.081273

Kido ÉA, Ferreira Neto JRC, Silva RLDO, Pandolfi V, Guimarães ACR, Veiga DT, Chabregas SM, Crovella S and Benko-Iseppon AM (2012) New Insights in the Sugarcane Transcriptome Responding to Drought Stress as Revealed by Supersage. Sci World J 2012:1–14. doi: 10.1100/2012/821062

Krizek BA and Fletcher JC (2005) Molecular mechanisms of flower development: An armchair guide. Nat Rev Genet 6:688–698. doi: 10.1038/nrg1675

Laloum T, De Mita S, Gamas P, Baudin M and Niebel A (2013) CCAAT-box binding transcription factors in plants: Y so many? Trends Plant Sci 18:157–166. doi: 10.1016/j.tplants.2012.07.004

Lemon B and Tjian R (2000) Orchestrated response: A symphony of transcription factors for gene control. Genes Dev 14:2551–2569. doi: 10.1101/gad.831000

Li S-B, Xie Z-Z, Hu C-G and Zhang J-Z (2016) A Review of Auxin Response Factors (ARFs) in Plants. Front Plant Sci 7:47. doi: 10.3389/fpls.2016.00047

Liu L, White MJ and Macrae TH (1999) Transcription factors and their genes in higher plants: Functional domains , evolution and regulation. Eur j Biochem 257:247–257. doi: 10.1046/j.1432-1327.1999.00349.x

Liu Z, Xin M, Qin J, Peng H, Ni Z, Yao Y and Sun Q (2015) Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). BMC Plant Biol 15:152. doi: 10.1186/s12870-015-0511-8

Liu F, Xu Y, Han G, Zhou L, Ali A, Zhu S and Li X (2016) Molecular Evolution and Genetic Variation of G2-Like Transcription Factor Genes in Maize. PLoS One 11:e0161763. doi: 10.1371/journal.pone.0161763

Mantiri FR, Kurdyukov S, Lohar DP, Sharopova N, Saeed N a, Wang X-D, Vandenbosch K a and Rose RJ (2008) The transcription factor MtSERF1 of the ERF subfamily identified by transcriptional profiling is required for somatic embryogenesis induced by auxin plus cytokinin in *Medicago truncatula*. Plant Physiol 146:1622–1636. doi: 10.1104/pp.107.110379

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI *et al.* (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226. doi: 10.1093/nar/gku1221

Nakashima K, Takasaki H, Mizoi J, Shinozaki K and Yamaguchi-Shinozaki K (2012) NAC transcription factors in plant abiotic stress responses. Biochim Biophys Acta - Gene Regul Mech 1819:97–103. doi: 10.1016/j.bbagrm.2011.10.005

Nelson DE, Repetti PP, Adams TR, Creelman RA, Wu J, Warner DC, Anstrom DC, Bensen RJ, Castiglioni PP, Donnarummo MG *et al.* (2007) Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on

water-limited acres. Proc Natl Acad Sci U S A 104:16450–16455. doi: 10.1073/pnas.0707193104

Nguyen C V, Vrebalov JT, Gapper NE, Zheng Y, Zhong S, Fei Z and Giovannoni JJ (2014) Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. Plant Cell 26:585–601. doi: 10.1105/tpc.113.118794

Ohnishi T, Sugahara S, Yamada T, Kikuchi K, Yoshiba Y, Hirano H-Y and Tsutsumi N (2005) OsNAC6, a member of the NAC gene family, is induced by various stresses in rice. Genes Genet Syst 80:135–139. doi: 10.1266/ggs.80.135

Oliveros JC (2017) VENNY: An interactive tool for comparing lists with Venn diagrams. Available: http://bioinfogp.cnb.csic.es/tools/venny/index.html. Accessed 15 February 2017

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358. doi: 10.1093/bioinformatics/12.4.357

Park MR, Yun KY, Mohanty B, Herath V, Xu F, Wijaya E, Bajic VB, Yun SJ and de Los Reyes BG (2010) Supra-optimal expression of the cold-regulated OsMyb4 transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. Plant, Cell Environ 33:2209–2230. doi: 10.1111/j.1365-3040.2010.02221.x

Pfaffl MW, Horgan GW and Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30:e36. doi: 10.1093/nar/30.9.e36

Poersch-bortolon LB, Pereira JF, Junior AN, Hernán H, Gonzáles S, Abigail G, Torres M, Consoli L, Arenhart RA, Bodanese-zanettini MH *et al.* (2016) Gene expression analysis reveals important pathways for drought response in leaves and roots of a wheat cultivar adapted to rainfed cropping in the Cerrado Biome. 645:629–645. doi: 10.1590/1678-4685-GMB-2015-0327

Quinn JM and Merchant S (1995) Two copper-responsive elements associated with the Chlamydomonas Cyc6 gene function as targets for transcriptional activators. Plant Cell 7:623–8. doi: 10.1105/tpc.7.5.623

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K and Shinozaki KY (2003) Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and AbscisicAcid Application Using cDNA Microarray and RNAGel-Blot Analyses. Plant Physiol 133:1755–1767. doi: 10.1104/pp.103.025742.genes

Rasmussen R (2001), Quantification on the Light Cycler. In: Rapid cycle real-time PCR, methods and applications. editors: Meuer S, Wittwer C, Nakagawara K Heidelberg:Springer Press.

Reis RR, Andrade Dias Brito da Cunha B, Martins PK, Martins MTB, Alekcevetch JC, Chalfun-Júnior A, Andrade AC, Ribeiro AP, Qin F, Mizoi J *et al.* (2014) Induced over-expression of AtDREB2A CA improves drought tolerance in sugarcane. Plant Sci 221–222:59–68. doi: 10.1016/j.plantsci.2014.02.003

Reyes JC, Muro-Pastor MI and Florencio FJ (2004) The GATA family of transcription factors in Arabidopsis and rice. Plant Physiol 134:1718–1732. doi: 10.1104/pp.103.037788

Rezende T, Figueira S, Okura V, Rodrigues da Silva F, Jose da Silva M, Kudrna D, Ammiraju JSS, Talag J, Wing R and Arruda P (2012) A BAC library of the SP80-3280 sugarcane variety (saccharum sp.) and its inferred microsynteny with the sorghum genome. BMC Res Notes 5:185. doi: 10.1186/1756-0500-5-185

Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM and Twell D (2001) Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene g10. Plant Mol Biol 45:577–585. doi: 10.1023/A:1010695226241

Rossini L, Cribb L, Martin DJ and Langdale JA (2001) The maize golden2 gene defines a novel class of transcriptional regulators in plants. Plant Cell 13:1231–44. doi: 10.1105/TPC.13.5.1231

Rozen S and Skaletsky H (2000) Primer3 on the WWW for General Users and for Biologist Programmers. Totowa, NJ Humana Press 132:365–386. doi: 10.1385/1-59259-192-2:365

Rung J and Brazma A (2012) Reuse of public genome-wide gene expression data. Nat Rev Genet 1–11. doi: 10.1038/nrg3394

Saad ASI, Li X, Li HP, Huang T, Gao CS, Guo MW, Cheng W, Zhao GY and Liao YC (2013) A rice stress-responsive NAC gene enhances tolerance of transgenic wheat to drought and salt stresses. Plant Sci 203–204:33–40. doi: 10.1016/j.plantsci.2012.12.016

Sakai H, Aoyama T and Oka A (2000) Arabidopsis ARR1 and ARR2 response regulators operate as transcriptional activators. Plant J 24:703–711. doi: 10.1046/j.1365-313X.2000.00909.x

Savitch L V., Subramaniam R, Allard GC and Singh J (2007) The GLK1 "regulon" encodes disease defense related proteins and confers resistance to Fusarium graminearum in Arabidopsis. Biochem Biophys Res Commun 359:234–238. doi: 10.1016/j.bbrc.2007.05.084

Setta N, Monteiro-Vitorello CB, Metcalfe CJ, Cruz GMQ, Del Bem LE, Vicentini R, Nogueira FTS, Campos RA, Nunes SL, Turrini PCG et al. (2014) Building the sugarcane genome for biotechnology and identifying evolutionary trends. BMC Genomics 15:540.

Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi IR, Omura T and Kikuchi S (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. Plant Cell Physiol 52:344–360. doi: 10.1093/pcp/pcq196

Shi M, Zhou W, Zhang J, Huang S, Wang H and Kai G (2016) Methyl jasmonate induction of tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots is mediated by JASMONATE ZIM-DOMAIN repressor proteins. Sci Rep 6:20919. doi: 10.1038/srep20919

Shirsat A, Wilford N, Croy R and Boulter D (1989) Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. MGG Mol Gen Genet 215:326–331. doi: 10.1007/BF00339737

Silva RLO, Silva MD, Nardi CH, Chabregas SM, Burnquist WL, Kahl G, Benko-Iseppon AM, Kido EA (2014) Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-time PCR in drought- stressed Sugarcane. The Sci World J 2014:1-13 doi: dx.doi.org/10.1155/2014/357052

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K and Yamaguchi-Shinozaki K (2003) Two different novel cis-acting elements of erd1, a clpA homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. Plant J 33:259–270. doi: 10.1046/j.1365-313X.2003.01624.x

Thiebaut F, Grativol C, Carnavale-Bottino M, Rojas CA, Tanurdzic M, Farinelli L, Martienssen R a, Hemerly AS and Ferreira PCG (2012) Computational identification and analysis of novel sugarcane microRNAs. BMC Genomics 13:290. doi: 10.1186/1471-2164-13-290

Thiebaut F, Grativol C, Tanurdzic M, Carnavale-Bottino M, Vieira T, Motta MR, Rojas C, Vincentini R, Chabregas SM, Hemerly AS *et al.* (2014) Differential sRNA regulation in leaves and roots of sugarcane under water depletion. PLoS One. doi: 10.1371/journal.pone.0093822

Tran LP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K and 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. Plant Cell 16:2481–2498. doi: 10.1105/tpc.104.022699.Shinozaki

Trujillo LE, Sotolongo M, Menéndez C, Ochogavía ME, Coll Y, Hernández I, Borrás-Hidalgo O, Thomma BPHJ, Vera P and Hernández L (2008) SodERF3, a novel sugarcane ethylene responsive factor (ERF), enhances salt and drought tolerance when overexpressed in tobacco plants. Plant Cell Physiol 49:512–525. doi: 10.1093/pcp/pcn025

Urano K, Kurihara Y, Seki M and Shinozaki K (2010) "Omics" analyses of regulatory networks in plant abiotic stress responses. Curr Opin Plant Biol 13:132–138. doi: 10.1016/j.pbi.2009.12.006

Vanholme B, Grunewald W, Bateman A, Kohchi T and Gheysen G (2007) The tify family previously known as ZIM. Trends Plant Sci 12:239–244. doi: 10.1016/j.tplants.2007.04.004

Visendi P, Berkman PJ, Hayashi S, Golicz AA, Bayer PE, Ruperao P, Hurgobin B, Montenegro J, Chan C-KK, Staňková H et al. (2016) An efficient approach to BAC based assembly of complex genomes. Plant Methods 12:2. doi: 10.1186/s13007-016-0107-9

Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Han B and Tao Y (2007) Genome-wide analysis of the auxin response factors (ARF) gene family in rice (Oryza sativa). Gene 394:13–24. doi: 10.1016/j.gene.2007.01.006

Wang C, Wang H, Zhang J and Chen S (2008) A seed-specific AP2-domain transcription factor from soybean plays a certain role in regulation of seed germination. Sci China, Ser C Life Sci 51:336–345. doi: 10.1007/s11427-008-0044-6

Wang J, Roe B, Macmil S, Yu Q, Murray JE, Tang H, Chen C, Najar F, Wiley G, Bowers J *et al.* (2010) Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. BMC Genomics 11:261. doi: 10.1186/1471-2164-11-261

Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, Guilfoyle TJ, Chen M and Qi Y (2010) Auxin-related gene families in abiotic stress response in *Sorghum bicolor*. Funct Integr Genomics 10:533–546. doi: 10.1007/s10142-010-0174-3

Weijers D and Wagner D (2016) Transcriptional Responses to the Auxin Hormone. Annu Rev Plant Biol 67:539–574. doi: 10.1146/annurev-arplant-043015-112122

Xing H, Pudake RN, Guo G, Xing G, Hu Z, Zhang Y, Sun Q and Ni Z (2011) Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. BMC Genomics 12:178. doi: 10.1186/1471-2164-12-178

Xiong Y, Liu T, Tian C, Sun S, Li J and Chen M (2005) Transcription factors in rice: A genome-wide comparative analysis between monocots and eudicots. Plant Mol Biol 59:191–203. doi: 10.1007/s11103-005-6503-6

Xu W, Chen Z, Ahmed N, Han B, Cui Q and Liu A (2016) Genome-wide identification, evolutionary analysis, and stress responses of the GRAS gene family in castor beans. Int J Mol Sci 17:1–16. doi: 10.3390/ijms17071004

Xu Y-X, Mao J, Chen W, Qian T-T, Liu S-C, Hao W-J, Li C-F and Chen L (2016) Identification and expression profiling of the auxin response factors (ARFs) in the

tea plant (Camellia sinensis (L.) O. Kuntze) under various abiotic stresses. Plant Physiol Biochem 98:46–56. doi: 10.1016/j.plaphy.2015.11.014

Xue GP, Way HM, Richardson T, Drenth J, Joyce PA and McIntyre CL (2011) Overexpression of TaNAC69 leads to enhanced transcript levels of stress up-regulated genes and dehydration tolerance in bread wheat. Mol Plant 4:697–712. doi: 10.1093/mp/ssr013

Yamaguchi-Shinozaki K and Shinozaki K (2006) Transcriptional Regulatory Networks in Cellular Responses and Tolerance To Dehydration and Cold Stresses. Annu Rev Plant Biol 57:781–803. doi: 10.1146/annurev.arplant.57.032905.105444

Yamasaki K, Kigawa T, Seki M, Shinozaki K and Yokoyama S (2013) DNA-binding domains of plant-specific transcription factors: Structure, function, and evolution. Trends Plant Sci 18:267–276. doi: 10.1016/j.tplants.2012.09.001

Yanagisawa S and Schmidt RJ (1999) Diversity and similarity among recognition sequences of Dof transcription factors. Plant J 17:209–214. doi: 10.1046/j.1365-313X.1999.00363.x

Yang T and Poovaiah BW (2002) A calmodulin-binding/CGCG box DNA-binding protein family involved in multiple signaling pathways in plants. J Biol Chem 277:45049–45058. doi: 10.1074/jbc.M207941200

Yilmaz A, Nishiyama MY, Fuentes BG, Souza GM, Janies D, Gray J and Grotewold E (2009) GRASSIUS: a platform for comparative regulatory genomics across the grasses. Plant Physiol 149:171–180. doi: 10.1104/pp.108.128579

Zhang Z-L, Xie Z, Zou X, Casaretto J, Ho TH and Shen QJ (2004) A Rice WRKY Gene Encodes a Transcriptional Repressor of the Gibberellin Signaling Pathway in Aleurone Cells. Plant Physiol 134:1500–1513. doi: 10.1104/pp.103.034967

Zhang Z, Li X, Yu R, Han M and Wu Z (2015) Isolation, structural analysis, and expression characteristics of the maize TIFY gene family. Mol Genet Genomics 290:1849–1858. doi: 10.1007/s00438-015-1042-6

Zhao G, Song Y, Wang C, Butt HI, Wang Q, Zhang C, Yang Z, Liu Z, Chen E, Zhang X *et al.* (2016) Genome-wide identification and functional analysis of the TIFY gene family in response to drought in cotton. Mol Genet Genomics 291:2173–2187. doi: 10.1007/s00438-016-1248-2

Zhao H, Dong L, Sun H, Li L, Lou Y, Wang L, Li Z and Gao Z (2016) Comprehensive analysis of multi-tissue transcriptome data and the genome-wide investigation of GRAS family in Phyllostachys edulis. Sci Rep 6:1–15. doi: 10.1038/srep27640

Zingaretti SM, Rodrigues FA, Graça JP, Pereira LM, Lourenço MV. Sugarcane Responses at Water Deficit Conditions. In: Rahman IMM, Hasegawa H, editors. Water Stress. Croatia: Intech.2012. p. 256-276.

# 8 DISCUSSÃO GERAL

A análise da estabilidade de seis potenciais genes de referência (αTUB, GAPDH, H1, SAMDC, UBQ, 25S rRNA) feita a partir de metodologias estatísticas de três *softwares* diferentes (GeNorm, Normfinder e Bestkeeper) teve como objetivo ampliar o número de genes disponíveis para a normalização de ensaios de validação via RT-qPCR de perfis transcricionais de raízes de cana-de-açúcar sobre déficit hídrico. Com pequenas mudanças na ordem de estabilidade, todas as três metodologias estatísticas selecionaram Histona H1, αTubulina e GAPDH como o conjunto de genes normalizadores mais estável. Dentre estes genes, a literatura relata GAPDH, αTUB, UBQ e 25S rRNA estão sendo rotineiramente estudados na normalização de dados de RT-qPCR em várias espécies vegetais, incluindo gramíneas como arroz (JAIN *et al.*, 2006; LI *et al.*, 2009), trigo (PAOLACCI *et al.*, 2009), *Brachypodium distachyon* (HONG *et al.*, 2008) e bambu (FAN *et al.*, 2013). Por outro lado, não encontramos estudos indicando Histona H1 como normalizador e sim alguns trabalhos reportando a participação de histona H3 na normalização de eudidicotiledôneas como ervilha (DIE *et al.*, 2010), algodão (TU *et al.*, 2007) e chicória (MAROUFI *et al.*, 2010), sem ensaios prévios com histonas para cana-de-açúcar.

Posteriormente foi conduzida a identificação de alvos a partir da técnica HT-SuperSAGE que monitorou a expressão de transcritos de quatro acessos previamente considerados tolerantes (THIEBAUT *et al.*, 2014) ao déficit hídrico *versus* outro grupo de quatro acessos considerados sensíveis (THIEBAUT *et al.*, 2014), onde a validação dos 26 alvos propostos via RT-qPCR permitiu discriminar cada acesso. Estes candidatos são agrupados em sete processos metabólicos envolvidos com a resposta à seca: atenuação do estresse de excesso de etileno;

crescimento radicular; degradação proteica; detoxificação oxidativa; biossíntese de ácidos graxos; transporte de aminoácidos, e metabolismo de carboidratos e duas classes de famílias gênicas: aquaporinas e fatores de trancrição. Com a abertura do *bulk*, para 22 dos 26 alvos validados, destacaram-se os acessos tolerantes CTC6 e SP83-2847, por induzirem 11 e 15 dos 22 candidatos, respectivamente e ambos compartilham sete destes alvos. No total, esses dois candidatos validaram 16 candidatos. Para o grupo de acessos sensíveis, dois deles (CTC9 e SP90-3414) induziram quatro e seis candidatos, respectivamente. Outro acesso sensível, CTC13, induziu um candidato (e reprimiu quatro), enquanto que SP90-1638, induziu dois (e reprimiu dois). Baseados nesses resultados, os acessos CTC6 e SP83-2847 foram considarados igualmente tolerantes ao estresse aplicado, enquanto que CTC13 SP90-1636 foi considerado igualmente sensíveis ao mesmo estresse. Portanto, o conjunto de candidatos propostos pode funcionar como marcadores moleculares funcionais para a seleção de acessos tolerantes ao déficit hídrico e podem ser promissores nos programas de melhoramento da cana-de-açúcar.

Partindo-se para análises de famílias específicas, foram idientificadas 42 isoformas distintas de aquaporinas, membros das quatro subfamílias (PIP, TIP, SIP e NIP) de aquaporinas descritas em vegetais superiores (MAUREL *et al.*, 2008), estando de acordo com a média encontrada em outras espécies vegetais (JANG *et al.*, 2004; BANSAL e SANKARARAMAKRISHNAN, 2007). Ao final, foram destacados 40 (18 induzidos nos tolerantes e ao mesmo tempo reprimidas ou não significativas nos sensíveis; 22 reprimidas nos tolerantes e ao mesmo tempo induzidas ou não significativas nos sensíveis) potenciais alvos distintos de isoformas de aquaporinas e suas respectivas *unitags* para estudos futuros de

validação do perfil transcricional detectado. Entre estes alvos, dois (PIP1-1 e PIP1-3/PIP1-4) foram validados via RT-qPCR e apresentaram potencial para desenvolvimento de marcadores moleculares para uso no melhoramento genético.

Apesar da importância econômica da cana-de-açúcar e dos TFs serem excelentes candidatos para melhorar características complexas nas plantas cultivadas (CENTURY *et al.*, 2008), nenhum trabalho apresentou o perfil transcricional de todo o conjunto de FTs, em cana-de-açúcar frente a estresses abióticos, incluindo-se o déficit hídrico. Sendo relatados para essa cultura os FTs associados a biossíntese de lignina (BRITO *et al.*, 2015), e de trabalhos com transformação gênica com foco em um FT (TRUJILLO *et al.*, 2008; REIS *et al.*, 2014). Portanto, esse trabalho é, desta forma, pioneiro em abordar a resposta do conjunto total de fatores de transcrição de raízes de cana-de-açúcar frente ao déficit hídrico em estágio inicial (24 h de supressão de rega), possibilitando a disponibilização de um conjunto de 541(264 nos tolerantes e 277 nos sensíveis) FTs diferencialmente expressos para SAR ou mesmo para transgenia, que pode ser estresse-específica com uso de alguns dos *Cis-Acting Regulatory Elements* de cana-de-açúcar aqui identificados.

Devido à possibilidade de se utilizar sequências genômicas provenientes de montagens de BACs de cana-de-açúcar (REZENDE *et al.*, 2012; SETTA *et al.*, 2014; VISENDI *et al.*, 2016), descartou-se a utilização do genoma de sorgo na identificação de CARE. Foi possível a identificação de sequências CARE em nove regiões promotoras de 2.000 pb upstream do início dos transcritos TFs de cana-de-açúcar induzidos, a partir do uso de 312 montagens de BAC de Setta *et al.*, (2014), destacam-se os promotores MYCCONSENSUSAT e ACGTATERD1

(quarto e 13° mais prevalentes, respectivamente), por serem relacionados diretamente à seca (ABE *et al.*, 2003; SIMPSON *et al.*, 2003) e DOFCOREZM (segunda mais prevalente), um motivo requerido para ligação de proteínas Dof (YANAGISAWA e SCHMIDT, 1999), sendo reportados TFs Dof envolvidos em respostas vegetais a estresses bióticos (KANG *et al.*, 2016) e abióticos (GUPTA *et al.*, 2016). Portanto esses CAREs são fortes candidatos a promotores estresse-específicos de cana-de-açúcar para o estresse de seca, sendo necessários futuros ensaios de transformação ectópica em organismos modelos para confirmação de seu controle estresse-específico. Experimentos microarranjo de alta performance e sequenciamento de nova geração têm gerado uma quantidade substancial de dados de análises amplas da expressão do genoma, onde boa parte desses dados sendo rotineiramente depositadas em banco de dados públicos e podem ser reutilizados por outros grupos de pesquisa para comparar dados com os obtidos em seus ensaios e/ou minerar situações específicas não investigadas em dados publicados (RUNG e BRAZMA, 2012). Este trabalho recai sobre a segunda alternativa utilizando-se da ferramenta de enriquecimento de FTs (*TF enrichment*) da base de dados PlantTFDB 4.0 (JIN *et al.*, 2016; http://plantregmap.cbi.pku.edu.cn/tfenrichment.php) para identificação de FTs responsáveis pela indução de exclusivamente UR para dois *bulks* contrastantes de cana-de-açúcar. Essa ferramenta utiliza dados de regulação transcricional obtidos da literatura e de dados da metodologia ChIP-seq, reconhecendo os mais presentes nas regulações (JIN *et al.*, 2016). Ao fim, destacou-se o FT Sobic.001G106200 (TALE), responsável mais de 200 regulações em ambos os bulks, indicando uma participação na ativação de diversos genes em diferentes estratégias de respostas e mitigação do déficit hídrico, sendo um bom candidato

para transformação gênica em ensaios futuros que visem uma resposta de amplo espectro em tempo curto. *Three Amino acid Loop Extension* (TALE) é uma superclasse de TFs homeobox conservada em animais, plantas e fungos e é subdividida em duas classes (*KNOTTED1-LIKE homeobox* - KNOX e *BEL1-like homeobox* - *BLH* ou *BELL*) envolvida principalmente no controle do desenvolvimento de diversos tecidos vegeteais, como gema apical, folhas, caules e inflorescências (DI GIACOMO, IANNELLI e FRUGIS, 2013) e membros das duas classes respondem a diferentes estresses abióticos (JAIN *et al.*, 2008).

# 9 CONCLUSÕES

A estratégia de se utilizar a tecnologia HT-SuperSAGE no monitoramento da transcrição de quatro acessos considerados tolerantes ao déficit hídrico *versus* outro grupo de quatro acessos considerados sensíveis, permitiu identificar alvos moleculares relacionados a diferentes mecanismos de resposta e a genes pertecentes a duas importantes classes de proteínas, aquaporinas e fatores de transcrição.

Quanto às aquaporinas, apresenta-se o primeiro painel de transcritos do conjunto total (com respectivas isoformas das famílias PIP, TIP, SIP e NIP) dessa classe de proteínas na cana-de-açúcar e a sua resposta ao déficit hídrico.

Também é disponibilizada a resposta do FToma [conjunto total de fatores de transcrição (FTs)] da cana-de-açúcar frente ao déficit hídrico.

Os BACs disponíveis para cana-de-açúcar permitiram identificar os elementos regulatórios atuantes em *Cis* (*cis-Acting regulatory elements* - CARE) mais prevalentes em regiões promotoras de nove dos FTs induzidos. Os CARE mais prevalentes são controlados por FTs relacionados a estresses abióticos e vias de sinalização hormonal e podem ser usados em ensaios de transgenia estresse-específica, diminuindo potenciais prejuízos à produtividade da planta transformada;

A padronização dos genes normalizadores indicou H1, αTUB e GAPDH como os HKGs mais adequados para estudos de RT-qPCR em raízes de genótipos de cana-de-açúcar submetidos ao déficit hídrico, sendo H1 e αTUB relatados pela primeira vez para a espécie.

Ao todo foram validados via RT-qPCR 26 alvos relacionados a diferentes estratégias de resposta à seca (15 alvos validados), e genes pertecentes a duas grandes classes de proteínas: aquaporinas (dois alvos validados) e fatores de transcrição (cinco alvos validados), com ao menos dois genes de referência padronizados e quatro outros alvos (PIP1-1 [estresse salino], ACC oxidase [estresse salino], AS [déficit hídrico] e PFP$\alpha$1 [déficit hídrico]) utilizados na validação da estabilidade do melhor conjunto de genes de referência (GAPDH, $\alpha$TUB e H1).

As cultivares CTC6 e SP83-2847 foram separadas, a partir da validação de 15 alvos, relacionados a sete estratégias de aclimatação à seca, como os mais tolerantes (cada um induzindo nove dos 15 alvos),

CTC13 (induziu apenas um alvo e reprimiu dois) e SP90-1636 (induziu apenas dois dos 15 alvos e reprimiu um) foram destacadas como as cultivares mais sensíveis e essas cultivares podem ser utilizadas como padrão de novos ensaios de transcriptômica frente ao estresse hídrico.

Todos esses dados fornecem pistas valiosas das respostas de variedades tolerantes e sensíveis ao estresse, bem como são fontes de alvos e promotores interessantes para estudos de transgenia e desenvolvimento de marcadores moleculares para seleção assistida.

# REFERÊNCIAS


ABE, H.; URAO, T.; ITO, T.; SEKI, M.; SHINOZAKI, K. Transcriptional Activators in Abscisic Acid Signaling. **Society**, v. 15, n. January, p. 63–78, 2003.

ABEDINI, R.; GHANEGOLMOHAMMADI, F.; PISHKAMRAD, R.; et al. Plant dehydrins: shedding light on structure and expression patterns of dehydrin gene family in barley. **Journal of Plant Research**, v. 0, n. 0, p. 0, 2017. Springer Japan.

ADAMS MD *et al.* Complementary dna sequencing: expressed sequence tags and human genome project. **Science** 252:1651-1656, 1991.

BANSAL, A.; SANKARARAMAKRISHNAN, R. Homology modeling of major intrinsic proteins in rice, maize and Arabidopsis: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. **BMC Structural biology**, v. 7, n. 27, p. 1–17, 2007.

BARTELS, D.; SUNKAR, R. Drought and Salt Tolerance in Plants. **Critical Reviews in Plant Sciences**, v. 24, n. 1, p. 23–58, 2005.

BATES B.C., KUNDZEWICZ Z.W., WU S.; PALUTIKOF J.P. (2008) Climate change and water resources in systems and sectors. In: BATES B.C.; KUNDZEWICZ Z.W., WU S.; PALUTIKOF J.P. (eds) **Climate Change and Water**. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 2008, 76p.

BRAY, E. A. Molecular Responses to Water Deficit. **Plant physiology**, v. 103, p. 1035–1040, 1993.

SANTOS BRITO, M.; NOBILE, P. M.; BOTTCHER, A.; et al. Expression Profile of Sugarcane Transcription Factor Genes Involved in Lignin Biosynthesis. **Tropical Plant Biology**, v. 8, n. 1–2, p. 19–30, 2015.

BARBOSA G.V. *et al*. CULTIVAR RELEASE RB99395 : Sugarcane cultivar with high sucrose content. **Crop Breeding and Applied Biotechnology**, p. 187–190, 2015.

CALSA JÚNIOR, T.; CARRARO, D. M.; BENATTI, M. R.; et al. Structural features and transcript-editing analysis of sugarcane (Saccharum officinarum L.) chloroplast genome. **Current Genetics**, v. 46, n. 6, p. 366–373, 2004.

CANALES, E.; COLL, Y.; HERNÁNDEZ, I.; *et al*. “Candidatus Liberibacter asiaticus”, causal agent of citrus huanglongbing, is reduced by treatment with brassinosteroids. **PLoS ONE**, v. 11, n. 1, p. 1–16, 2016.

CAPUTO M.M., BEAUCLAIR E.G.F., SILVA M.A. AND PIEDADE S.M.S. Response of sugarcane genotypes to the application of ripeners. **Bragantia**,v 67, p.15-23, 2008.

CARDONA, C. A.; QUINTERO, J. A.; PAZ, I. C. Production of bioethanol from sugarcane bagasse: Status and perspectives. **Bioresource Technology**, v. 101, n. 13, p. 4754–4766, 2010.

CARNEIRO M.S., CHAPOLA R.G., JÚNIOR A.R.F., CURSI D.E., BARRETO F.Z., BALSALOBRE W.A., HOFFMANN H.P. (2015). RB975952 – Early maturing

sugarcane cultivar. **Crop Breeding and Applied Biotechnology**. v. 15, p. 193-196, 2015.

CENTURY, K.; REUBER, T. L.; RATCLIFFE, O. J. Regulating the regulators: the future prospects for transcription-factor-based agricultural biotechnology products. **Plant physiology**, v. 147, n. 1, p. 20–29, 2008.

COEMANS, B.; MATSUMURA, H.; TERAUCHI, R.; et al. SuperSAGE combined with PCR walking allows global gene expression profiling of banana (Musa acuminata), a non-model organism. **Theoretical and Applied Genetics**, v. 111, n. 6, p. 1118–1126, 2005.

CIDADE, D.A.P., GARCIAI, R.O., DUARTEI, A.C., SACHETTO-MARTINSII G., MANSURI, E. Morfogênese in vitro de variedades brasileiras de cana-de-açúcar. **Pesq agropec bras**, v. 40, p. 385-391, 2006.

Companhia Nacional de Abastecimento (CONAB). **Acompanhamento de safra brasileira:** cana-de-açúcar, segundo levantamento, agosto/2012. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12_09_05_09_11_59_boletim_cana_portugues_-_agosto_2012_2o_lev.pdf> Acesso em: 22 nov, 2012.

Companhia Nacional de Abastecimento (CONAB). **Acompanhamento de safra brasileira:** cana-de-açúcar, segundo levantamento, agosto/2016. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/16_08_17_10_07_35_boletim_cana_portugues_-_2o_lev_-_16-17.pdf> Acesso em: 21 de mar, 2017.

Centro de Tecnologia Canavieira (CTC). **Histórico**. Disponível em: <http://www.ctcanavieira.com.br> Acesso em: 12 de jul, 2015.

DEREN, C. W.; SNYDER, G. H.; MILLER, J. D.; PORTER, P. S. Screening for and heritability of flood-tolerance in the Florida (CP) sugarcane breeding population. **Euphytica**, v. 56, n. 2, p. 155–160, 1991.

D'HONT, A.; GRIVET, L.; FELDMANN, P.; *et al*. Characterization of the double genome structure of sugarcane cultivars (Saccharum ssp.) by molecular cytogenetics. **Mol. Gen. Genet.**, v. 250, p. 405–413., 1996.

D'HONT A.; GLASZMANN J.C. Sugarcane Genome Analysis With Molecular Markers: A First Decade of Research. **Proc. Int. Soc. Sugar Cane Technol.**, v. 24, p. 556-559, 2001.

DIE, J. V.; ROMÁN, B.; NADAL, S.; GONZÁLEZ-VERDEJO, C. I. Evaluation of candidate reference genes for expression studies in Pisum sativum under different experimental conditions. **Planta**, v. 232, n. 1, p. 145–153, 2010.

DILLON, S. L.; SHAPTER, F. M.; HENRY, R. J.; *et al*. Domestication to crop improvement: Genetic resources for Sorghum and Saccharum (Andropogoneae). **Annals of Botany**, v. 100, n. 5, p. 975–989, 2007.

FAN, C.; M.A., J.; GUO, Q.; *et al*. Selection of Reference Genes for Quantitative Real-Time PCR in Bamboo (Phyllostachys edulis). **PLoS ONE**, v. 8, n. 2, 2013.

FARIAS, C. H. D. A.; FERNANDES, P. D.; AZEVEDO, H. M.; DANTAS NETO, J. Índices de crescimento da cana-de-açúcar irrigada e de sequeiro no Estado da Paraíba. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 12, n. 4, p. 356–362, 2008.

FERNANDEZ O., BÉTHENCOURT L., QUERO A., SANGWAN R.S., CLÉMENT C. Trehalose and plant stress responses: friend or foe? **Trends Plant Sci**, v. 15, p. 409-417, 2010.

FERREIRA, T. H.; GENTILE, A.; VILELA, R. D.; *et al*. microRNAs Associated with Drought Response in the Bioenergy Crop Sugarcane (*Saccharum* spp.). **PLoS ONE**, v. 7, n. 10, 2012.

FOCKE, M.; GIERINGER, E.; SCHWAN, S.; *et al*. Fatty acid biosynthesis in mitochondria of grasses: malonyl-coenzyme A is generated by a mitochondrial-localized acetyl-coenzyme A carboxylase. **Plant physiology**, v. 133, n. 2, p. 875–884, 2003.

Food and Agriculture Organization of the United Nations (FAO), **FAO STAT, Download:** Production / Crops. Disponível em: <http://faostat3.fao.org/download/Q/QC/E> Acesso em: 18 de jul, 2015.

GIACOMO, E. D.I.; IANNELLI, M.; FRUGIS, G. TALE and Shape: How to Make a Leaf Different. **Plants**, v. 2, n. 2, p. 317–342, 2013.

GILL, S.S.; TUTEJA, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant Physiology and Biochemistry**, v. 48, n. 12, p. 909–930, 2010. Elsevier Masson SAS.

GLENN, T.C. Field guide to next-generation DNA sequencers. **Molecular Ecology Resources**, v. 11, n. 5, p. 759–769, 2011.

GRIVET, L.; ARRUDA, P. Sugarcane genomics: Depicting the complex genome of an important tropical crop. **Current Opinion in Plant Biology**, v. 5, n. 2, p. 122–127, 2002.

GOSAL, S. S.; WANI, S. H.; KANG, M. S. Biotechnology and Drought Tolerance. **Journal of Crop Improvement**, v. 23, n. 1, p. 19–54, 2009.

GUPTA, S.; ARYA, G.C.; MALVIYA, N.; BISHT, N.C.; YADAV, D. Molecular cloning and expression profiling of multiple Dof genes of Sorghum bicolor (L) Moench. **Molecular Biology Reports**, v. 43, n. 8, p. 767–774, 2016. Springer Netherlands.

GUO, H.; SUN, Y.; PENG, X.; et al. Up-regulation of abscisic acid signaling pathway facilitates aphid xylem absorption and osmoregulation under drought stress. **Journal of Experimental Botany**, v. 67, n. 3, p. 681–693, 2016.

HARBERS, M.; CARNINCI, P. Tag-based approaches for transcriptome research and genome annotation. **Nat Methods**, v. 2, p. 495-502, 2005.

HIETER, P.; BOGUSKI, M. Functional genomics: it's all how you read it. **Science** v. 278, p. 601-602, 1997.

HOFSETZ, K.; SILVA, M. A. Brazilian sugarcane bagasse: Energy and non-energy consumption. **Biomass and Bioenergy**, v. 46, p. 564–573, 2012.

HONG, S.-Y.; SEO, P. J.; YANG, M.-S.; XIANG, F.; PARK, C.-M. Exploring valid reference genes for gene expression studies in Brachypodium distachyon by real-time PCR. **BMC plant biology**, v. 8, p. 112, 2008.

HU, H.; XIONG, L. Genetic engineering and breeding of drought-resistant crops. **Annual review of plant biology**, v. 65, n. November 2013, p. 715–41, 2014.

JAIN, M.; NIJHAWAN, A.; TYAGI, A. K.; KHURANA, J. P. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. **Biochemical and Biophysical Research Communications**, v. 345, n. 2, p. 646–651, 2006.

JAIN, M.; TYAGI, A. K.; KHURANA, J. P. Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. **FEBS Journal**, v. 275, n. 11, p. 2845–2861, 2008.

JIN, J.; TIAN, F.; YANG, D.-C.; et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. **Nucleic Acids Research**, p. gkw982, 2016.

JANG, J. Y.; KIM, D. G.; KIM, Y. O.; et al. An expression analysis of a gene family encoding plasma membrane aquaporins in response to abiotic stresses in Arabidopsis thaliana. **Plant molecular biology**, v. 54, n. 5, p. 713–725, 2004.

KAHL, G.; MOLINA, C.; ROTTER, B; JÜNGLING, R.; FRANK, A.; KREZDORN, N.; HOFFMEIER, K.; WINTER, P. Reduced representation sequencing of plant stress transcriptomes. **J. Plant Biochem. Biotechnol.** DOI:10.1007/s13562-012-0129-y, 2012.

KANG, W.H.; KIM, S.; LEE, H.A.; CHOI, D.; YEOM, S.I. Genome-wide analysis of Dof transcription factors reveals functional characteristics during development and response to biotic stresses in pepper. **Sci Rep**, v. 6, 2016.

KELISH, A. EL; ZHAO, F.; HELLER, W.; et al. Ragweed (Ambrosia artemisiifolia) pollen allergenicity: SuperSAGE transcriptomic analysis upon elevated CO2 and drought stress. **BMC Plant Biology**, v. 14, n. 1, p. 176, 2014.

KIDO, É.A.; BARBOSA, P.K.; NETO, J.R.; PANDOLFI, V.; HOULLOU-KIDO, L.M.; CROVELLA, S.; MOLINA, C.; KAHL, G.; BENKO-ISEPPON A.M. Identification of Plant Protein Kinases in Response to Abiotic and BioticStresses Using SuperSAGE. **Curr Protein Pept Sci**, v12, p. 643-656, 2011.

KIDO, É. A.; FERREIRA NETO, J. R. C.; SILVA, R. L. D. O.; *et al*. New Insights in the Sugarcane Transcriptome Responding to Drought Stress as Revealed by Supersage. **The Scientific World Journal**, v. 2012, p. 1–14, 2012.

KIDO, E. A; FERREIRA NETO, J. R. C.; SILVA, R. L. O.; *et al*. Expression dynamics and genome distribution of osmoprotectants in soybean: identifying important components to face abiotic stress. **BMC bioinformatics**, v. 14, n. Suppl 1, p. S7, 2013.

KIDRIČ, M.; KOS J.; SABOTIČ, J. Proteases and their endogenous inhibitors in the plant response to abiotic stress. **Botanica Serbica**, v. 38, n. 1, p. 139–158, 2014.

LEMBKE, C. G.; NISHIYAMA, M. Y.; SATO, P. M.; ANDRADE, R. F. DE; SOUZA, G. M. Identification of sense and antisense transcripts regulated by drought in sugarcane. **Plant Molecular Biology**, v. 79, n. 4–5, p. 461–477, 2012.

LI, C.; NONG, Q.; SOLANKI, M. K.; et al. Differential expression profiles and pathways of genes in sugarcane leaf at elongation stage in response to drought stress. **Scientific reports**, v. 6, n. December 2015, p. 25698, 2016.

LI, Q. F.; SUN, S. S. M.; YUAN, D. Y.; et al. Validation of candidate reference genes for the accurate normalization of real-time quantitative RT-PCR data in rice

during seed development. **Plant Molecular Biology Reporter**, v. 28, n. 1, p. 49–57, 2009.

MAROUFI, A.; BOCKSTAELE, E. VAN; LOOSE, M. DE. Validation of reference genes for gene expression analysis in chicory (Cichorium intybus) using quantitative real-time PCR. **BMC molecular biology**, v. 11, p. 15, 2010.

MATSUMURA, H.; REICH, S.; ITO, A.; et al. Gene expression analysis of plant host-pathogen interactions by SuperSAGE. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 26, p. 15718–15723, 2003.

MATSUMURA, H. REICH S; REUTER, M; KRUGER, D.H.; WINTER, P.; KAHL, G.; TERAUCHI, R. SuperSAGE: A Potent Transcriptome Tool for Eukaryotic Organisms Tool for Eukaryotic Organisms. In: SAN, M.W. (ed) **Current Technologies and Applications**. Horizon Scientific Press, Norwich, United Kingdom, pp 77-90, 2004.

MATSUMURA, H.; KRÜGER, D. H.; KAHL, G.; TERAUCHI, R. SuperSAGE: A Modern Platform for Genome-Wide Quantitative Tran- script Profiling. **Current Pharmaceutical Biotechnology**, v. 9, p. 368–374, 2008.

MATSUMURA, H.; YOSHIDA, K.; LUO, S.; *et al*. High-throughput superSAGE for digital gene expression analysis of multiple samples using next generation sequencing. **PLoS ONE**, v. 5, n. 8, 2010.

MATSUMURA H, YOSHIDA K, LUO S, KRÜGER DH, KAHL G, SCHROTH GP, TERAUCHIA R (2011) High-Throughput SuperSAGE. In: PARK, D.J. (ed) **Methods in Molecular Biology**, PCR Protocols, v. 687, p. 135-146, 2011.

MATSUMURA, H.; MOLINA, C.; KRÜGE, D.H.; TERAUCHI, R.; KAHL, G. DeepSuperSAGE: High-Throughput Transcriptome Sequencing with Now- and Next-Generation Sequencing Technologies. In: HARBERS, M.; KAHL, G. (eds) **Tag-Based Next Generation Sequencing.** Wiley, pp 1-21, 2012.

MAUREL, C.; VERDOUCQ, L.; LUU, D.-T.; SANTONI, V. Plant aquaporins: membrane channels with multiple integrated functions. **Annual review of plant biology**, v. 59, p. 595–624, 2008.

MENOSSI, M.; SILVA-FILHO, M. C.; VINCENTZ, M.; VAN-SLUYS, M. A.; SOUZA, G. M. Sugarcane functional genomics: Gene discovery for agronomic trait development. **International Journal of Plant Genomics**, v. 2008, 2008.

Molina C *et al.* The drought stress-responsive transcriptome of chickpea roots. **BMC Genomics**. v. 9, p. 1-28, 2008.

MOLINA, C.; ZAMAN-ALLAH, M.; KHAN, F.; et al. The salt-responsive transcriptome of chickpea roots and nodules via deepSuperSAGE. **BMC plant biology**, v. 11, 2011.

MORAIS LK *et al.* (2015) Breeding of Sugarcane. In: VonMark VC; David AD (eds) **Handbook of Plant Breeding**. vol 9. Springer, New York, pp 29-42.

MORTAZAVI, A.; WILLIAMS, B. A.; MCCUE, K.; SCHAEFFER, L.; WOLD, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nat Methods**, v. 5, n. 7, p. 1–8, 2008.

OKURA, V. K.; SOUZA, R. S. C. DE; SIQUEIRA TADA, S. F. DE; ARRUDA, P. BAC-Pool Sequencing and Assembly of 19 Mb of the Complex Sugarcane Genome. **Frontiers in plant science**, v. 7, n. March, p. 342, 2016.

PAOLACCI, A. R.; TANZARELLA, O. A.; PORCEDDU, E.; CIAFFI, M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. **BMC molecular biology**, v. 10, n. 1, p. 11, 2009.

PINHEIRO, D.G. Desenvolvimento de uma Plataforma Integrativa para Depuração e Análise de Dados de Expressão Gênica. Faculdade de Medicina de Ribeirão Preto, **Tese de doutorado**, Ribeirão Preto, p. 1-125, 2009.

PINTO, L.R.; OLIVEIRA, K.M.; ULIAN, E.C.; GARCIA, A.A.F.; SOUZA, A.P. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome**, v 47, p 795–804, 2004.

RAE, A. L.; GROF, C. P. L.; CASU, R. E.; BONNETT, G. D. Sucrose accumulation in the sugarcane stem : pathways and control points for transport and compartmentation. Field **Crops Research**, v. 92, p. 159–168, 2005.

Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (RIDESA), **Apresentação**. Disponível em: <http://www.ridesa.agro.ufg.br/>, Acesso em: 12 jul, 2015.

REIS, R. R.; ANDRADE DIAS BRITO DA CUNHA, B.; MARTINS, P. K.; *et al*. Induced over-expression of AtDREB2A CA improves drought tolerance in sugarcane. **Plant Science**, v. 221, p. 59–68, 2014.

REZENDE, T.; FIGUEIRA, S.; OKURA, V; RODRIGUES, SILVA, F.; *et al*. A BAC library of the SP80-3280 sugarcane variety (*Saccharum* sp.) and its inferred microsynteny with the sorghum genome. **BMC research notes**, v. 5, n. 1, p. 185, 2012.

RODRIGUES, F. A.; LAIA, M. L. DE; ZINGARETTI, S. M. Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants. **Plant Science**, v. 176, n. 2, p. 286–302, 2009.

SAHA, S.; SPARKS, A.B.; RAGO, C.; *et al.* Quantitative monitoring of gene expression patterns with a complementary dna microarray. **Science**, v. 270 p. 467-470, 1995.

SAHA S., SPARKS, A.B., RAGO, C., AKMAEV, V.; WANG, C.J.; VOGELSTEIN, B.; KINZLER, K.W.; VELCULESCU, V.E. **Nat Biotechnol** v. 20, p. 508-512, 2002.

SETTA, N. DE; MONTEIRO-VITORELLO, C. B.; METCALFE, C. J.; *et al*. Building the sugarcane genome for biotechnology and identifying evolutionary trends. **BMC genomics**, v. 15, n. 1, p. 540, 2014.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. **Journal of Experimental Botany**, v. 58, p. 221–227, 2007.

SILVA, M. D. A.; ALBERTO, J.; ENCISO, J.; et al. YIELD COMPONENTS AS INDICATORS OF DROUGHT TOLERANCE OF SUGARCANE. , **Sci. Agric**., p. 620–627, 2008.

SILVA, V. DE P. R. DA; GUEDES, M. J. F.; LIMA, W. F. A.; CAMPOS, J. H. B. C. Modelo de previsão de rendimento de culturas de sequeiro, no semi-árido do

Nordeste do Brasil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, p. 83–87, 2002.

SINGH, A.; JHA, S. K.; BAGRI, J.; PANDEY, G. K. ABA inducible rice protein phosphatase 2C confers ABA insensitivity and abiotic stress tolerance in arabidopsis. **PLoS ONE**, v. 10, n. 4, p. 1–24, 2015.

SINGH, K. B.; FOLEY, R. C.; OÑATE-SÁNCHEZ, L. Transcription factors in plant defense and stress responses. **Current Opinion in Plant Biology**, v. 5, p. 430–436, 2002.

TEGEDER, M. Transporters for amino acids in plant cells: Some functions and many unknowns. **Curr Opin Plant Biol**, v. 15, p. 315–321, 2012.

THIEBAUT, F.; GRATIVOL, C.; TANURDZIC, M.; *et al*. Differential sRNA regulation in leaves and roots of sugarcane under water depletion. **PLoS ONE**, v. 9, n. 4, 2014.

TU, L.; ZHANG, X.; LIU, D.; et al. Suitable internal control genes for qRT-PCR normalization in cotton fiber development and somatic embryogenesis. **Chinese Science Bulletin**, v. 52, n. 22, p. 3110–3117, 2007.

TRUJILLO, L. E.; SOTOLONGO, M.; MENÉNDEZ, C.; *et al.* SodERF3, a novel sugarcane ethylene responsive factor (ERF), enhances salt and drought tolerance when overexpressed in tobacco plants. **Plant and Cell Physiology**, v. 49, p. 512–525, 2008.

VARGAS, L.; BRÍGIDA, A.B.S.; MOTA FILHO, J. P.; *et al*. Drought tolerance conferred to sugarcane by association with gluconacetobacter diazotrophicus: A transcriptomic view of hormone pathways. **PLoS ONE**, v. 9, p. 1–37, 2014.

VELCULESCU, V.E.; ZHANG, L.; VOGELSTEIN, B.; KINZLER, K.W. Serial analysis of gene expression. **Science**, v. 270, p. 484-487, 1995.

VERBRUGGEN N, HERMANS C Proline accumulation in plants: a review.

**Amino Acids**, v 35 p. 753-759, 2008.

VIANNA, M. DOS S.; SENTELHAS, P. C. Simulation of the water deficit risk in sugarcane-crop expansion regions in Brazil. **Pesquisa Agropecuaria Brasileira**, v. 49, n. 4, p. 237–246, 2014.

VISENDI, P.; BERKMAN, P. J.; HAYASHI, S.; et al. An efficient approach to BAC based assembly of complex genomes. **Plant Methods**, v. 12, n. 1, p. 2, 2016.

WANG, J.; ROE, B.; MACMIL, S.; et al. Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. **BMC genomics**, v. 11, p. 261, 2010.

YANAGISAWA, S.; SCHMIDT, R. J. Diversity and similarity among recognition sequences of Dof transcription factors. **Plant Journal**, v. 17, n. 2, p. 209–214, 1999.

YILMAZ, A.; NISHIYAMA, M. Y.; FUENTES, B. G.; et al. GRASSIUS: a platform for comparative regulatory genomics across the grasses. **Plant physiology**, v. 149, n. 1, p. 171–180, 2009.

# ANEXO A

## Normas para submissão (Genetics and Molecular Biology - ISSN 1678-4685)

### INSTRUCTIONS TO AUTHORS

### Scope and policy

**Genetics and Molecular Biology** (formerly named Revista Brasileira de Genética/Brazilian Journal of Genetics - ISSN 0100-8455) is published by the Sociedade Brasileira de Genética (Brazilian Society of Genetics).

The journal considers contributions that present the results of original research in genetics, molecular biology, evolution and related scientific disciplines. Manuscripts presenting methods and applications only, without an analysis of genetic data, will not be considered.

It is a fundamental condition that submitted manuscripts have not been published or are not under consideration for publication elsewhere. With the acceptance of a manuscript for publication, the publishers acquire full and exclusive copyright for all languages and countries.

Manuscripts considered in conformity with the scope of the journal, as evaluated by the Editor, are reviewed by an Associate Editors and two or more external reviewers. Acceptance by the Editor is based on the quality of the work as substantial contribution to the field and on the overall presentation of the manuscript.

The official abbreviation for Genetics and Molecular Biology is **Genet. Mol. Biol.**

**Open-Access policy**
Genetics and Molecular Biology articles are made available in full content at SciELO (Scientific Library Online) hosted at http://www.scielo.br/gmb. Back issues dating until 1998 are available through this site.

GMB articles published since 2009 are also indexed at PubMed Central and available in full text version at: http://www.ncbi.nlm.nih.gov/pmc/journals/1440/.

Back issues of earlier titles (Brazilian Journal of Genetics and Revista Brasileira de Genética) are hosted at GMB's own site: http://www.gmb.org.br

Membership to the Brazilian Society of Genetics entitles subscription to Genetics and Molecular Biology.

For nonmembers and institutions, the annual subscriptions rates (four issues/year) are informed at the journal's

website http://www.gmb.org.br.

## Submission of papers

**1. Manuscripts have to be submitted through our online submission platform:**

https://mc04.manuscriptcentral.com/gmb-scielo

A cover letter addressed to the Editor-in-Chief is required

**2. For submission the following instructions must be observed:**

**a)** The manuscript must be submitted by the Corresponding Author, identified as such in the title page of the manuscript. This is the person who will also check the page proofs, and arranges for any payment that may incur during the editorial process.

**b)** Entering the following metadata is required: (i) the manuscript title, (ii) a short running title (max. 35 characters), (iii) the Abstract, and (iv) up to five keywords. All these items must be exactly the same as those figuring in the first two pages of the manuscript file.

**c)** Statements are required informing that the data have not been published and are not under consideration elsewhere, and that all authors have approved the submission of the manuscript. Furthermore, possible conflicts of interest (e.g. due to funding, consultancies) must also be disclosed. For statements on ethical issues in research see below (3.1.m).

**d)** The names of all co-authors, including institutional affiliations and e-mail addresses must be entered, as contact information for the Editorial Office.

**e)** In the referee suggestions field, up to five reviewer names can be entered by the author(s); valid e-mail contact addresses for these are required, in case they are selected by the editor. These suggestions can be made separately as preferred and opposed reviewer(s).

**f)** Files must be uploaded separately and identified according to file types, respecting the following sequence: main text document (title page as page 1), tables, figures and, if applicable, supplementary material. The main text file must include the title page, Abstract, References and, if applicable, figure legends, which must be typed on a separate page following the References and Internet Resources sections. Each table, figure and element containing supplementary material must be saved and uploaded in a separate file. Formats for text and tables are Word or RTF in Windows platform. Figures should be in TIFF or JPEG formats (see detailed instructions in 3.1.i).

**g)** Manuscripts including photos or any other identifiable data of human subjects must be accompanied by a copy of the signed consent by the individual or his/her guardian.

Failure to adhere to these guidelines can delay the handling of your contribution and manuscripts may be returned before being reviewed.

Special attention should be given to the structuring of the manuscript and correct language usage. These are important factors in the smooth running of the editorial

and peer-review process, and can result in faster publication.

**3. Categories of Contribution**
**3.1. Research Articles**

Manuscripts must be written in English in double-spaced, 12-point type throughout; marked with consecutive line and page numbers, beginning with the title page.

The following elements must start on a new page and be ordered as they are listed below:

a) **The Title Page** must contain: a concise and informative title; the authors' names (first name at full length); the authors' institutional affiliation, including department, institution, city, state or province, and country; different affiliations must be indicated with superscript Arabic numbers; a short running title of up to 35 characters (including spaces); up to five key words; the corresponding author's name, full postal, and email address.

b) **The Abstract** must be a single paragraph that does not exceed 200 words and summarizes the main results and conclusions of the study. It should not contain references.

c) **The text** must be as succinct as possible. *Text citation*s: articles should be referred to by authors' surnames and date of publication; citations with two authors must include both names separated by "and"; in citations with three or more authors, name the first author and use *et al.* List two or more references in the same citation in chronological order, separated by semi-colons. When two or more works in a citation were published in the same year, list them alphabetically by the first author surname. For two or more works by the same author(s) in a citation, list them chronologically, with the years separated by commas. (Example: Freire-Maia *et al.*, 1966a, 1966b, 2000). Only articles that are published or in press should be cited. In the case of personal communications or unpublished results, all contributors must be listed by initials and last name (*et al.* should not be used). *Number*s: In the text, numbers nine or less must be written out except as part of a date, a fraction or decimal, a percentage, or a unit of measurement. Use Arabic numerals for numbers larger than nine. Do not start a sentence with an Arabic numeral. *Binomial Name*s: Latin names of genera, species and infraspecific taxa must be printed in italics; we also recommend to present names of orders or families in the Title and/or when first mentioned in the text. URLs for programs, data or other sources should be listed in the Internet Resources Section, immediately following the References Section, not in the text.

The text includes the following elements:

*Introduction* - Description of the background that led to the study.

*Material (or Subjects) and Methods* - Details relevant to the conduct of the study. Statistical methods should be explained at the end of this section.

*Results* - Undue repetition in text and tables should be avoided. Statistical analyses should be presented as complete as possible, i.e. not only *P*-values should be shown, but also all other test variables required for full appreciation of the results. Comments on relevance of results are appropriate but broader discussion should be part of the Discussion section.

*Discussion* - The findings of the study should be placed in context of relevant

published data. Ideas presented in other publications should not be discussed solely to make an exhaustive presentation.

Some manuscripts may require different formats appropriate to their content.

d) **The Acknowledgments** must be a single paragraph that immediately follows the discussion and includes references to grant support.

e) **Conflict of Interest:** Any possible conflict of interest must be disclosed here. If there is none, please state: The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

f) **The References Section:** Only articles that are published or in press should be included in this section. Works submitted for publication but not yet accepted, personal communications and unpublished data must be cited within the text. "Personal communication" refers to information obtained from individuals other than the authors of the manuscript being submitted; "unpublished data" refers to data produced by at least one of the authors of the manuscript under consideration. Works of restricted circulation (e.g., theses not available in public databases, congress abstracts not published in regular journals or public databases) should not be listed in this section.

References must be ordered alphabetically by the first author surname; references with the same first author should be ordered as follows: first, as single author in chronological order; next, with only one more co-author in alphabetical order by the second author; and finally followed by references with more than two co-authors, in chronological order, independent of the second author surnames. In references with more than 10 authors only the first ten should be listed, followed by *et al.* Use standard abbreviations for journal titles as suggested by NCBI (http://www.ncbi.nlm.nih.gov/journals/) or Thomson Reuters Web of Science.

*Sample journal article citation:*

Breuer ME and Pavan C (1955) Behaviour of polytene chromosomes of *Rhynchosciara angelae* at different stages of larval development. Chromosoma 7:371-386.

Yonenaga-Yassuda Y, Rodrigues MT and Pellegrino KCM (2005) Chromosomal banding patterns in the eyelid-less microteiid lizard radiation: The X1X1X2X2:X1X2Y sex chromosome system in *Calyptommatus* and the karyotypes of *Psilophtalmus* and *Tretioscincus* (Squamata, Gymnophthalmidae). Genet Mol Biol 28:700-709.

*Sample book citation:*

Dobzhansky T (1951) Genetics and Origin of Species. 3rd edition. Columbia University Press, New York, 364 p.

*Sample chapter-in-book citation:*

Crawford DC and Howard-Peebles PN (2005) Fragile X: From cytogenetics to molecular genetics. In: Gersen SL and Keagle MB (eds) The Principles of Clinical Cytogenetics. 2nd edition. Humana Press, New Jersey, pp 495-513.

*Sample electronic article citation:*

Gotzek D, Ross KG (2009) Current status of a model System: The gene Gp-9 and its association with social organization in fire ants. PLoS One 4:e7713.

g) **Internet Resources Section:** this section should contain a list of URLs referring to data presented in the text, as well as software programs and other Internet resources used during data processing. Date of consultation must be stated.

*Sample Internet resource citation:*

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/OMIM (September 4, 2009)

LEM
Software, http://dir.niehs.nih.gov/dirbb/weinbergfiles/hybrid_design.htm (September 4, 2009)

h) **Tables: Formats for tables are Word or RTF in Windows platform. They** must be prepared with the table tool (do not use space bar or tabulator) and must be numbered consecutively in Arabic numerals). A concise title should be provided above the table. Each column should have a title in the box head. Footnotes typed directly below the table should be indicated in lowercase superscript letters. Tables that are to appear in the printed version must be saved in Word format and not as figures, so that they can later be fitted during typesetting. Each table must be saved and uploaded as a separate file.

i) **Figures: Formats for figures are TIFF or JPEG. They** must be numbered consecutively using Arabic numerals. Figures in Word, PowerPoint or Excel format cannot be published. Only sequence data can be presented in Word format. Journal quality reproduction will require resolution yielding 300 dpi for grayscale and color figures. These resolutions refer to the output size of the file, that being the size in which it will appear printed in the journal; if it is anticipated that images will be enlarged or reduced, the resolutions should be adjusted accordingly. Figures composed of several elements should be sent as a single panel, obeying the print size definitions of the journal (single or two columns width). Scanned figures should not be submitted. Color illustrations are accepted and will be reproduced free of chrge in the electronic and printed versions. Figure legends must be included at the end of the main text file and should be typed on a new page. When uploading, identify each illustration by the first author name and the number of the respective figure. Each figure/panel must be saved and uploaded as a separate file.

j) **Nomenclature:** Taxonomic names should be in accordance with current international standards. For rules concerning gene names and gene symbols, please see separate Instruction form.

k) **Sequences** may appear in text or in figure. DNA, RNA and protein sequences equal to or greater than 50 units must be entered into public databases and accession numbers must be provided upon acceptance of the article. Failure to do so will inadvertently delay publication.

l) **Data access:** reference should be made to availability of detailed data and materials used for reported studies.

m) **Ethical issues:** Reports of experiments on live vertebrates must include a statement in the text that the institutional review board approved the work and the protocol number must be provided. For experiments involving human subjects, a statement must be provided that informed consent was obtained from all subjects. If photos or any other identifiable data are included, a copy of the signed consent must

be uploaded during manuscript submission.

n) **Supplementary Material**: Data that the authors consider of importance for completeness of a study, but which are too extensive to be included in the print version, can be submitted as Supplementary Material. At publication, this material will be made available together with the electronic version. In case a manuscript contains such material, it should be appropriately identified within the text file. Supplementary material in tables should be identified as Table S1, Table S2, etc., in case of figures they should be named accordingly, Figure S1, Figure S2. In addition, a list of this material should be presented at the end of the manuscript text file, containing the following statement:

*Supplementary material - the following online material is available for this article:*

- *Table S1 – < short title >*
- *Figure S1 – < short title >*

### 3.2 Short Communications

Short Communications present brief observations that do not warrant full-length articles. They should not be considered preliminary communications. They should be 15 or fewer typed pages in double spaced 12-point type, including literature cited, include an Abstract no longer than five percent of the paper's length, but no further subdivision, with introduction, material and methods, results and discussion in a single section and without headers. Up to four items (tables and/or figures) may be submitted. The title page and reference section format is that of a full-length Research Article. For Supplementary Material see instructions in item 3.1.n

### 3.3 Genome Insight

Genome Insight is for focused papers, usually of approximately 1500 words (up to four tables or figures), that publish new genome data as they become submitted to GenBank. This section is the premier forum to deliver that information directly to the genome community in a rapid and efficient publication of the genome. Data must be related to a complete (or nearly complete) and fully annotated genome for prokaryote or viruses, but a draft may be accepted for an eukaryote genome. While the focus of Genome Insight is necessarily involved in novel sequences, the manuscript must contain specifically novel biological, evolutive, biotechnological and/or metabolic insights revealed by data. The work may provide comparative analyses of previously published genomes that contain a substantial and novel insight of broadest biological and genetic significance.

Submitted manuscript must contain an abstract, which should be a brief report on the organism as well as its relevance and the main insight revealed by the genome. The text (approximately 1500 words- excluding abstract, references and acknowledgements) should not contain subdivisions, but must contain the rational for the selection of such organism as well as organism information (including taxonomy, natural habitat, phylogenetic position, eventual pathogenicity, symbiotic, biotechnological use, etc), methodology (genome sequencing and assembly; reference number at GenBank), genome relevance (which should indicate the main insights revealed by the data analysis and main conclusions. Acknowledgements and References (up to 20 references) headings should be included. Figure Legends should be provided at the end of the manuscript.

Metagenome, transcriptome as well as epigenome data may also be considered for publication, but prior submission of the abstract to the Editor is necessary.

Note: The title page, abstract and reference section format is that of a full-length Research Article. For Supplementary Material see item 3.1.n. in our Instructions to Authors.

**3.4 Letters to the Editor**

Relate or respond to recent published items in the journal. Discussions of political, social and ethical issues of interest to geneticists are also welcome in this form.

**3.5 Review Articles**

Review Articles are welcome. The Editor must be contacted prior to submission. Please, provide an Abstract and a list of your recent publications in the area.

**3.6 Book Reviews**

Publishers are invited to submit books on Genetics, Evolution and related disciplines, for review in the journal. Aspiring reviewers may propose writing a review.

**3.7 History, Story and Memories**

These are accounts on historical aspects of Genetics relating to Brazil.

**4. Articles accepted for publication**

Once an article is accepted, the Editorial Office will send it to copy editor for language and technical corrections. If major corrections were proposed, the manuscript with the highlighted corrections will be returned to the corresponding author for approval. The final version approved by the authors must be free of any text/correction markings when returned to the Editorial Office.

After typesetting, page proofs will be sent to the corresponding author. Changes made to page proofs, apart from typesetting errors, will be charged to the authors. Notes added in proof require Editorial approval.

Together with the proofs, a form of consent to publish and transfer of copyright is sent to the corresponding author, who will have sign this form, also on behalf of any co-authors, and send it by e-mail to the Editorial Office.

**5. Availability of articles and deposition in databases**

Article copies are provided as PDF-files. Authors may deposit these in their personal or institutional homepage, as well as in public databases.

**6. Publication charge**

There is a publication charge for manuscripts once they are accepted. For price information, exemptions and waiver policies, please consult the journal homepage http://www.gmb.org.br.

# ANEXO B

## *Curriculum vitae* (Lattes)

**Manasses Daniel da Silva**
Curriculum Vitae

---

### Dados pessoais

**Nome** Manasses Daniel da Silva
**Filiação** Tito Daniel da Silva e Zuleide Daniel de Oliveira
**Nascimento** 16/12/1987 - Camaragibe/PE - Brasil
**Carteira de Identidade** 6964972 sds - PE - 17/08/2011
**CPF** 074.151.494-08

**Endereço residencial** Rua Cassimiro de Abreu, 18
Timbi - Camaragibe
54768-420, PE - Brasil
Telefone: 81 34845401

**Endereço profissional** Universidade Federal de Pernambuco, Centro de Ciências Biológicas
Av. Prof. Moraes Rego, 1235
Cidade Universitária - Recife
50670-901, PE - Brasil
Telefone: 81 21268000

**Endereço eletrônico**
E-mail para contato : manassesdaniel@gmail.com
E-mail alternativo danielmann_7@hotmail.com

---

### Formação acadêmica/titulação

**2013** Doutorado em Genética.
Universidade Federal de Pernambuco, UFPE, Recife, Brasil
Título: Identificação e Validação de Transcritos SuperSAGE de Cana-de-açúcar Diferencialmente Expressos sob Déficit Hídrico
Orientador: Éderson Akio Kido
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

**2011 - 2013** Mestrado em Genética.
Universidade Federal de Pernambuco, UFPE, Recife, Brasil
Título: Análise in Silico de Subfamílias de Aquaporinas em Cana-de-açúcar (Saccharum spp.) sob Condições de Déficit Hídrico via Tecnologia SuperSAGE, Ano de obtenção: 2013
Orientador: Éderson Akio Kido
Bolsista do(a): Conselho Nacional de Desenvolvimento Científico e Tecnológico

**2006 - 2010** Graduação em Ciências Biológicas- Bacharelado.
Universidade Federal de Pernambuco, UFPE, Recife, Brasil
Título: IDENTIFICAÇÃO DE ALVOS MOLECULARES EM CANA-DE-AÇÚCAR PARA TOLERÂNCIA AO DEFICIT HÍDRICO
Orientador: Éderson Akio Kido
Bolsista do(a): Fundação de Amparo à Ciência e Tecnologia do Estado de

Pernambuco

## Produção

## Produção bibliográfica

### Artigos completos publicados em periódicos

**1.** NETO, JOSÉ; **DA SILVA, MANASSÉS**; PANDOLFI, VALESCA; CROVELLA, SÉRGIO; BENKO-ISEPPON, ANA; KIDO, EDERSON
Epigenetic Signals on Plant Adaptation: A Biotic Stress Perspective. Current Protein and Peptide Science. , v.17, p.1 - 1, 2016.

**2.** PANDOLFI, VALESCA; NETO, JOSÉ; **SILVA, MANASSÉS**; AMORIM, LIDIANE; WANDERLEY-NOGUEIRA, ANA; DE OLIVEIRA SILVA, ROBERTA; KIDO, EDERSON; CROVELLA, SERGIO; ISEPPON, ANA
Resistance (R) Genes: Applications and Prospects for Plant Biotechnology and Breeding. Current Protein and Peptide Science. , v.17, p.1 - 1, 2016.

**3.** SILVA, ROBERTA LANE DE OLIVEIRA; **SILVA, MANASSÉS DANIEL**; FERREIRA NETO, JOSÉ RIBAMAR COSTA; NARDI, CLAUDIA HUERTA DE; CHABREGAS, SABRINA MOUTINHO; BURNQUIST, WILLIAM LEE; KAHL, GÜNTER; BENKO-ISEPPON, ANA MARIA; KIDO, EDERSON AKIO
Validation of Novel Reference Genes for Reverse Transcription Quantitative Real-Time PCR in Drought-Stressed Sugarcane. The Scientific World Journal. , v.2014, p.1 - 12, 2014.

**4. Silva, M. D.**; Silva, R. L. O.; FERREIRA NETO, J. R. C.; GUIMARAES, A. C. R.; VEIGA, D. T.; Chabregas, S. M.; Burnquist, W. L.; Kahl, G.; BENKO-ISEPPON, A. M.; Kido, E. A.
Expression Analysis of Sugarcane Aquaporin Genes under Water Deficit. Journal of Nucleic Acids. , v.2013, p.1 - 14, 2013.

**5.** Kido, E. A.; FERREIRA NETO, J. R. C.; Silva, R. L. O.; BELARMINO, L. C.; BEZERRA NETO, J. P.; SOARES-CAVALCANTI, N. M.; PANDOLFI, V.; **Silva, M. D.**; NEPOMUCENO, A. L.; BENKO-ISEPPON, A. M.
Expression dynamics and genome distribution of osmoprotectants in soybean: identifying important components to face abiotic stress. BMC Bioinformatics. , v.14, p.1 - 1, 2013.

### Capítulos de livros publicados

**1.** FERREIRA NETO, J. R. C.; SOUZA, A. C. M.; **Silva, M. D.**; BENKO-ISEPPON, A. M.; PANDOLFI, V.; Costa, A. F.; Kido, E. A.
The Transcriptional Modulation of Inositols and Raffinose Family Oligosaccharides Pathways in Plants — An (A)Biotic Stress Perspective In: Abiotic and Biotic Stress in Plants - Recent Advances and Future Perspectives.1 ed. : InTech, 2016, p. 81-99.

### Trabalhos publicados em anais de eventos (resumo)

1. SILVA, R. H.; **Silva, M. D.**; CARNEIRO, K. R. S.; SILVA FILHO, J. L. B.; Kido, E. A.

ANÁLISE DE EST-SSR DE SACCHARUM SPP. E TRANSFERIBILIDADE DE MARCADORES PARA PENNISETUM PURPUREUM SCHUM In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

2. **Silva, M. D.**; SOUZA, M. C. P.; SOUZA, A. C. M.; FARIAS, J. L. S.; SILVA, V.; Costa, A. F.; Kido, E. A.
Análise In Silico do Kinoma do Feijão-caupi In: 4º Congresso Nacional de Feijão-caupi - CONAC, 2016, Sorriso.
**Anais do 4º Congresso Nacional de Feijão-caupi - CONAC**. , 2016.

3. **Silva, M. D.**; CABRAL, G. A. L.; SILVA, R. H.; SILVA-FILHO, J. L. B.; Kido, E. A.
ANÁLISE TRANSCRIPTÔMICA DE FATORES DE TRANSCRIÇÃO EXPRESSOS SOB ESTRESSE DE DÉFICIT HÍDRICO EM CANA-DE-AÇÚCAR In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

4. NASCIMENTO, A. T. B.; **Silva, M. D.**; SILVA, S. A.; CARNEIRO, K. R. S.; FERREIRA NETO, J. R. C.; Costa, A. F.; Kido, E. A.
Expressão in silico na via de sinalização por ABA de transcritos de feijão-caupi sob estresse de seca. In: 4º Congresso Nacional de Feijão-caupi - CONAC, 2016, Sorriso.
**Anais do 4º Congresso Nacional de Feijão-caupi - CONAC**. , 2016.

5. **Silva, M. D.**; CABRAL, G. A. L.; SILVA, R. H.; SILVA-FILHO, J. L. B.; Costa, A. F.; Kido, E. A.
Identificação de Fatores de Transcrição em Feijão-caupi [Vigna unguiculata (L.) Walp.] In: 4º Congresso Nacional de Feijão-caupi - CONAC, 2016, Sorriso.
**Anais do 4º Congresso Nacional de Feijão-caupi - CONAC**. , 2016.

6. NASCIMENTO, A. T. B.; **Silva, M. D.**; FERREIRA-NETO, J. R. C.; SILVA, S. A.; Kido, E. A.
IDENTIFICAÇÃO IN SILICO DA EXPRESSÃO DE TRANSCRITOS DE CANA-DE-AÇÚCAR SOB ESTRESSE DE SECA NAS VIAS DE SINALIZAÇÕES POR FITORMÔNIOS In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

7. Silva, R. L. O.; **Silva, M. D.**; SILVA, J. B.; REGO, M. S.; BENKO-ISEPPON, A. M.; Kido, E. A.
IDENTIFICAÇÃO IN SILICO DE TRANSPORTADORES DE AMINOÁCIDOS EM CANA-DE-AÇÚCAR SUBMETIDA À DESIDRATAÇÃO In: XXI Encontro de Genética do Nordeste, 2016, Recife.
**Anais**. , 2016. p.1 - 421

8. FERREIRA NETO, J. R. C.; **Silva, M. D.**; PIRES, C. J.; BORGES, A. N. C.; VIANA, J. B. V.; BENKO-ISEPPON, A. M.; Kido, E. A.
IMPORTÂNCIA DA REANÁLISE DE DADOS DISPONÍVEIS DE TRANSCRITÔMICA: EXEMPLO DE SOJA (GLYCINE MAX) SOB ESTRESSES ABIÓTICOS In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016.

9. **Silva, M. D.**; FERREIRA-NETO, J. R. C.; CARNEIRO, K. R. S.; SILVA, R. H.; Kido, E. A.
MAPEAMENTO DE UNITAGS HT-SUPERSAGE DE CANA-DE-AÇÚCAR (SACCHARUM SPP.) EM SEQUÊNCIAS CODIFICANTES (CDS) DE SORGO (SORGHUM BICOLOR (L.) MOENCH) E IDENTIFICAÇÃO DE GRUPOS DE LIGAÇÃO. In: XXI Encontro de Genética do Nordeste, 2016, Recife.
**Anais 2016**. , 2016.

10. SOUZA, A. C. M.; SOUSA, K. P.; GUERRA, V. T.; SILVA, E. W.; SILVA, V.; **Silva, M. D.**; Kido, E. A.
MICROSSINTENIA E MICROCOLINEARIDADE DE QTL DE MILHETO RELACIONADA À TOLERÂNCIA À SECA EM GRAMÍNEAS In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

11. SANTO, S. S. R.; PANDOLFI, V.; RIVAS, R.; LIMA, M. O.; BELARMINO, L. C.; FERREIRA-NETO, J. R. C.; **Silva, M. D.**; BENKO-ISEPPON, A. M.
PROSPECÇÃO DE DEFENSINAS EM ESPÉCIES MEDICINAIS In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

12. CABRAL, G. A. L.; **Silva, M. D.**; ARAUJO JUNIOR, N. R.; Kido, E. A.
PROSPECÇÃO IN SILICO DE ELEMENTOS CIS-REGULATÓRIOS EM PROMOTORES DE GENES DA FAMÍLIA DE FATORES DE TRANSCRIÇÃO DE ESTRESSE TÉRMICO (HEAT STRESS TRANSCRIPTION FACTOR- HSF) EM PINHÃO MANSO (JATROPHA CURCAS L.) In: XXI Encontro de Genética do Nordeste - ENGENE, 2016, Recife.
**Anais 2016**. , 2016. p.1 - 421

13. BEZERRA NETO, J. P.; FERREIRA-NETO, J. R. C.; Kido, E. A.; **Silva, M. D.**; BENKO-ISEPPON, A. M.; Guida, M.
Understanding transcriptional strategy for Inositol pathway in soybean root dehydration stress tolerance In: X-Meeting 2016 - 12th International Conference of the AB3C, 2016, Belo Horizonte.
**Anais do X-Meeting 2016**. , 2016.

14. SILVA, J. B.; REGO, M. S.; Silva, R. L. O.; **Silva, M. D.**; BENKO-ISEPPON, A. M.; Kido, E. A.
ANÁLISE IN SILICO DE EXPANSINAS EM CANA-DE-AÇÚCAR EM RESPOSTA À SECA In: II Congresso Internacional Multiprofissional em Saúde, 2015, Recife.
**Anais do II Congresso Internacional Multiprofissional em Saúde**. , 2015.

15. ARAUJO, A. C.; LEMOS, A. B.; **Silva, M. D.**; Silva, R. L. O.; BENKO-ISEPPON, A. M.; Kido, E. A.
IDENTIFICAÇÃO DE TRANSCRITOS SUPERSAGE DE ACETYL-COA CARBOXYLASE EM CANA-DE-AÇÚCAR SOB DÉFICIT HÍDRICO In: II Congresso Internacional Multiprofissional em Saúde, 2015, Recife.
**Anais do II Congresso Internacional Multiprofissional em Saúde**. , 2015.

16. SILVA, R. H.; **Silva, M. D.**; SILVA, E. W.; FERREIRA NETO, JOSÉ RIBAMAR COSTA; Kido, E. A.
MINERAÇÃO DE MARCADORES SSR EM ESTs DE Saccharum SPP. RELACIONADAS A ESTRESSES ABIÓTICOS In: XV Jornada de Ensino, Pesquisa e Extensão da UFRPE, 2015, Recife.
**Anais :: JEPEX 2015**. , 2015.

17. MATOS, M. K. S.; **Silva, M. D.**; BEZERRA NETO, J. P.; ARAUJO, F. T.; AMORIM, L. L. B.; Benko, A. M. B.
Diversidade estrutural 'in silico' de candidatos a fatores de transcrição da família MYB no transcriptoma do feijão-caupi In: 20º Encontro de Genética do Nordeste - ENGENE, 2014, Campina Grande.
**Anais do 20º Encontro de Genética do Nordeste**. , 2014.

18. SILVA, R. H.; LOCATELLI, A. C. P.; SILVA, W. M.; SILVA-FILHO, J. L. B.; CASTRO, N. M. C. R.; POMPELLI, M. F.; **Silva, M. D.**; Kido, E. A.
Diversidade genética de acessos de Pinhão-manso (Jatropha curcas L.) através de marcadores moleculares SSR In: 20º Encontro de Genética do Nordeste - ENGENE, 2014, Campina Grande.
**Anais do 20º Encontro de Genética do Nordeste**. , 2014.

19. **Silva, M. D.**; Silva, R. L. O.; FERREIRA-NETO, J. R. C.; Kido, E. A.
Identificação e validação de transcritos SuperSAGE de cana-de-açúcar diferencialmente expressos sob déficit hídrico In: 20º Encontro de Genética do Nordeste - ENGENE
**Anais do 20º Encontro de Genética do Nordeste**. , 2014.

20. Silva, R. L. O.; **Silva, M. D.**; BENKO-ISEPPON, A. M.; Kido, E. A.
Validação de genes diferencialmente expressos em cana-de-açúcar em condições de déficit hídrico através da técnica de SuperSAGE In: XX Encontro de Genética do Nordeste (ENGENE),

2014, Campina Grande.
**Anais do 20º Encontro de Genética do Nordeste**. , 2014.

**Trabalhos publicados em anais de eventos (resumo expandido)**

1. SILVA, J. B.; REGO, M. S.; Silva, R. L. O.; **Silva, M. D.**; BENKO-ISEPPON, A. M.; Kido, E. A. AVALIAÇÃO DE TRANSCRITOS DE PROTEASES ASSOCIADOS À TOLERÂNCIA AO DÉFICIT HÍDRICO EM CANA DE AÇUCAR In: II Congresso Internacional Multiprofissional em Saúde, 2015, Recife.
**Anais do II Congresso Internacional Multiprofissional em Saúde**. , 2015.

2. ARAUJO, A. C. C.; LEMOS, A. B.; **Silva, M. D.**; Silva, R. L. O.; Benko, A. M. B.; Kido, E. A. IDENTIFICAÇÃO DE UMA 1-AMINOCICLOPROPANO-1-CARBOXILATO DEAMINASE COM RESPOSTAS CONTRASTANTES ENTRE ACESSOS TOLERANTES E SENSÍVEIS DE Saccharum spp. SOB DÉFICIT HÍDRICO In: II Congresso Internacional Multiprofissional em Saúde, 2015, Recife.
**Anais do II Congresso Internacional Multiprofissional em Saúde**. , 2015.

**Artigos aceitos para publicação**

1. **Silva, M. D.**; Silva, R. L. O. ; FERREIRA NETO, J. R. C. ; BENKO-ISEPPON, A. M. ; Kido, E. A. Unraveling Drought-tolerant and -sensitive Sugarcane Cultivars by HT-SuperSAGE Analysis. GENE, 2017.

## Bancas

## Bancas
**Participação em banca de trabalhos de conclusão**

**Graduação**

1. Kido, E. A.; MAIA, M. M. D.; FERREIRA NETO, J. R. C.; **Silva, M. D.**
Participação em banca de Monique Mayara Rodrigues Costa. **MAPEAMENTO IN SILICO DE TRANSCRITOS SUPERSAGE DE TRANSCRITOS DE FEIJÃO CAUPI [Vigna unguiculata (L.). Walp.] EXPRESSOS SOB DESIDRATAÇÃO RADICULAR EM GENOMA DE FEIJÃO-COMUM (Phaseolus vulgaris L.).**, 2014
(Ciências Biológicas) Universidade Federal Rural de Pernambuco

2. Kido, E. A.; MAIA, M. M. D.; **Silva, M. D.**; FERREIRA NETO, J. R. C.
Participação em banca de Carolina Basante Silva. **MAPEAMENTO IN SILICO DE TRANSCRITOS (tags SuperSAGE) DE RAÍZES DE CANA-DE-AÇÚCAR (Saccharum ssp.) SOB ESTRESSE ABIÓTICO (100 mM DE NaCl), EM GENOMA DE SORGO (Sorghum bicolor).**, 2014
(Ciências Biológicas) Universidade Federal Rural de Pernambuco

---

## Citações

**Web of Science** Total de citações : 10;Total de trabalhos : 5;Data : 10/11/2016; Fator H: 2;
Nome(s) do autor utilizado(s) na consulta para obter o total de citações:
Silva MD