

AGNES ANGÉLICA GUEDES DE BARROS

**CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE
OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes scabra* (Vogel)
SOB DÉFICIT HÍDRICO**

Recife – PE

2019

AGNES ANGÉLICA GUEDES DE BARROS

**CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE
OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes scabra* (Vogel)
SOB DÉFICIT HÍDRICO**

**Dissertação apresentada ao
Programa de Pós-Graduação em
Agronomia – Melhoramento
Genético de Plantas, da
Universidade Federal Rural de
Pernambuco como parte dos
requisitos para obtenção do título
de Mestre.**

COMITÊ DE ORIENTAÇÃO:

Orientador: Prof. Dr Reginaldo de Carvalho - UFRPE

Coorientadores: Dr^a Roberta Lane de Oliveira Silva - UNIVASF

Dr João Pacífico Bezerra-Neto - UFPE

Recife – PE

2019

Ficha catalográfica

B277c Barros, Agnes Angélica Guedes de
Caracterização estrutural, funcional e diversidade de
osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel)
sob déficit hídrico/ Agnes Angélica Guedes de Barros.– 2019.
135f. : il.

Orientador: Reginaldo de Carvalho.

Coorientadores: Roberta Lane de Oliveira Silva, João Pacífico
Bezerra-Neto.

Dissertação (Mestrado) – Universidade Federal Rural de
Pernambuco, Programa de Pós-Graduação em Agronomia –
Melhoramento Genético de Plantas, Recife, BR-PE, 2019.

Inclui referências.

1. Leguminosa 2. Expressão gênica 3. Sequência de bases
4. Recursos hídricos 5. Escassez I. Carvalho, Reginaldo, orient.
II. Silva, Roberta Lane de Oliveira, coorient. III. Bezerra-Neto, João
Pacífico, coorient. IV. Título

CDD 581.1

**CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE
OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes scabra* (Vogel)
SOB DÉFICIT HÍDRICO**

AGNES ANGÉLICA GUEDES DE BARROS

Dissertação defendida e avaliada pela Banca Examinadora em: 31/07/2019

Orientador:

Prof. Dr. Reginaldo de Carvalho - UFRPE

Banca examinadora:

Dr. Nataniel Franklin de Melo – Pesquisador EMBRAPA Semiárido

Dr. José Ribamar Costa Ferreira Neto - UFPE

Recife – PE

2019

À memória do meu querido pai, Adilson Barros

Dedico.

Agradecimentos

À Universidade Federal Rural de Pernambuco, pelo apoio institucional.

À Universidade Federal de Pernambuco, pela disponibilidade de infraestrutura e equipamentos utilizados na execução das minhas atividades.

À Pós-Graduação em Agronomia da Universidade Federal Rural de Pernambuco pela obtenção do título de Mestre.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo apoio financeiro e à EMBRAPA Semiárido pela concessão do material vegetal utilizado na pesquisa.

À Prof^a Dr^a Ana Maria Benko-Iseppon pela oportunidade de trabalhar com seu grupo de pesquisa no Laboratório de Genética e Biotecnologia Vegetal (LGBV-UFPE) o que me proporcionou grandes ensinamentos enriquecendo minha carreira profissional.

Ao meu orientador Prof. Dr. Reginaldo de Carvalho e aos meus coorientadores Dr^a Roberta Lane e Dr. João Pacífico pelos ensinamentos e apoio prestados.

À minha mãe Maria Correia, por ser minha fortaleza.

À minha amiga da graduação e da vida Larissa Pereira, pela paciência e palavras de conforto nos momentos que mais precisei.

À Jessica Barboza por todo conhecimento compartilhado comigo além do apoio emocional e amizade nos momentos difíceis.

A José Ribamar Neto por compartilhar sua experiência na qual contribuiu imensamente no desenvolvimento no meu trabalho.

Aos membros do LGBV pelo suporte e convivência, em especial a Mariane, Hugo e Rômulo pela amizade e momentos de descontração e risadas mais que necessárias, além de Thialisson pelo imenso suporte técnico.

"O nitrogênio em nosso DNA, o cálcio em nossos dentes, o ferro em nosso sangue, o carbono em nossas tortas de maçã... Foram feitos no interior de estrelas em colapso, agora mortas há muito tempo. Nós somos poeira das estrelas."

Carl Sagan

LISTA DE ABREVIATURAS

AA_Kinase	<i>Amino Acid Kinase Family</i>
ABA	Ácido abscísico
Aldedh	Família de Aldeídos Dehidrogenases, <i>Aldehyde dehydrogenase Family</i>
ALDH F10	Aldeídos Dehidrogenases Familia 10, <i>Aldehyde dehydrogenase Family 10</i>
AMADH	Amino aldehyde dehydrogenase
APG	Grupo de Filogenia de Angiospermas, <i>Angiosperm Phylogeny Group</i>
ATP	Trifosfato de adenosina
BADH	Betaína aldeído dehidrogenase, <i>Betaine-aldehyde dehydrogenase</i>
BFG	Grupo de Flora do Brasil, <i>Brazilian Flora Group</i>
BLAST	Ferramenta Básica de Busca por Alinhamento Local, <i>Basic Local Alignment Search Tool</i>
BLOSSUM	Matriz de substituição de blocos de aminoácidos, <i>BLOCKS of Amino Acid SUBstitution Matrix</i>
cDNA	DNA Complementar, Complementary DNA
CMO	Colina Monooxigenase, <i>Choline monooxygenase</i>
CRIA	Centro de Referência em Informação Ambiental
DNA	Ácido Desoxirribonucleico, Deoxyribonucleic Acid
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EPAGRI	Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina
FC	<i>Fold Change</i>
FIG Superfamily	Superfamília de fosfatases dependentes de metal
FISH	Hibridização Fluorescente in situ, <i>Fluorescence in situ Hybridization</i>
GB	Glicina betaína
GC%	Porcentagem de Guanina e Citosina
GISH	Hibridação Genômica in situ, <i>Genomic in situ Hybridization</i>

Glyco Transf 20	Família glicosiltransferase 20, <i>Glycosyltransferase family 20</i>
GSA-DH	Dehidrogenase Glutâmica Semialdeído, <i>Glutamic Semialdehyde Dehydrogenase</i>
HAD-like Superfamily	Superfamília de Haloácidos Dehidrogenases, <i>Haloacid Dehydrogenase Superfamily</i>
HT-SuperSAGE	<i>High-Throughput Super Serial Analysis of Gene Expression</i>
IMP	Inositol monofosfatase, <i>Inositol monophosphatase</i>
LEA	Proteínas Abundantes de Embriogênese Tardia, <i>Late Embryogenesis Abundant proteins</i>
LWGP	Grupo de Trabalho de Filogenia de Leguminosas, <i>Leguminosae Working Group Phylogeny</i>
MEGA v.7	<i>Molecular Evolutionary Genetics Analysis</i> versão 7
MEME	<i>Multiple EM for Motif Elicitation</i>
MIPS	Mio inositol fosfato sintase, <i>Myo inositol Phosphate Synthase</i>
miRNAs	Micro RNAs
mRNA	RNA mensageiro
NAD(P)H	Fosfato de dinucleotídeo de adenina e nicotinamida
NCBI	Centro Nacional para Informação Biotecnológica, <i>National Center for Biotechnology Information</i>
NGS	Sequenciamento de Nova Geração, <i>Next Generation Sequencing</i>
ORF	Fase de Leitura Aberta, <i>Open Reading Frame</i>
P5CR	Pirrolina 5 carboxilato redutase, <i>Pyrroline-5-carboxylate reductase</i>
P5CS	Pirrolina 5 carboxilato sintase, <i>Pyrroline-5-carboxylate synthase</i>
PAM	Mutação Pontual Aceita, <i>Point Accepted Mutation</i>
PCR	Reação em Cadeia da Polimerase, <i>Polymerase Chain Reaction</i>
QAC	Composto Amônio Quaternário
QTL	Locus de Característica Quantitativa, <i>Quantitative Trait Loci</i>

rADN	DNA ribossomal
RNA	Ácido Ribonucleico, <i>Ribonucleic Acid</i>
RNA-Seq	Sequenciamento do RNA; RNA Sequencing
ROS	Espécie Reativa de Oxigênio, <i>Reactive Oxygen Species</i>
RT-qPCR	PCR quantitativa em Tempo Real, <i>Quantitative Real Time PCR</i>
Tm	Temperatura de melting
TPP	Trealose 6 fosfato fosfatase, <i>Trehalose 6-phosphate phosphatase</i>
TPS	Trealose 6 fosfato sintase, <i>Trehalose-6-phosphate synthase</i>
UFLA	Universidade Federal de Lavras
UNEB	Universidade Estadual da Bahia
UniProt	Banco de Dados Universal de Proteínas, <i>Universal Protein</i>

LISTA DE ILUSTRAÇÕES

CAPÍTULO I: REFERENCIAL TEÓRICO

	Página
Figura 1: Resposta do vegetal a estresses abióticos com ênfase no déficit hídrico e salinidade.	25
Figura 2: Ilustração esquemática da transdução de sinais induzidas pelo calor e desenvolvimento da tolerância ao estresse nas plantas.	26
Figura 3: Rota de sinalização de ABA-dependente e ABA-independente durante estresse hídrico.	29
Figura 4: Via metabólica de Prolina em vegetais durante estresse abiótico	33
Figura 5: Via metabólica de Glicina-betaína.	35
Figura 6: Via metabólica de Trealose.	37
Figura 7: Representação do sequenciamento pelo método de Sanger.	39
Figura 8: Componentes a nível celular, abordagens multiômicas e estratégias para identificar o resultado do impacto de estresses abióticos e/ou interações com microrganismos.	41
Figura 9: Representação esquemática da técnica de RNA-Seq.	44
Figura 10: Uma visão do processamento da linguagem natural biológica (BioNLP) e aplicações de text mining na biologia.	46

CAPÍTULO II: CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes Scabra* (Vogel) SOB DÉFICIT HÍDRICO.

Figura 1. Alinhamento das sequências de P5CS de <i>S. scabra</i> com identificação de motivos conservados.	84
Figura 2. Alinhamento das sequências de P5CR de <i>S. scabra</i> com identificação de motivos conservados.	85
Figura 3a. Alinhamento das sequências de BADH de <i>S. scabra</i> com identificação de motivos e resíduos conservados.	86
Figura 3b. Alinhamento das sequências ortólogas de BADH utilizadas na fenética para identificação do aminoácido crítico para produção de GB.	87
Figura 4. Alinhamento das sequências de CMO de <i>S. scabra</i> com identificação do motivo mononuclear non-heme Fe-binding.	88
Figura 5. Alinhamento das sequências de MIPS de <i>S. scabra</i> com identificação de motivos conservados.	90
Figura 6. Alinhamento das sequências de IMP de <i>S. scabra</i> com identificação de motivos conservados.	91
Figura 7. Alinhamento das sequências com apenas domínio *TPS de <i>S. scabra</i> com identificação de resíduos conservados.	92
Figura 8. Alinhamento das sequências de TPP de <i>S. scabra</i> com identificação de motivos conservados.	93
Figura 9. Alinhamento das sequências de TPS de <i>S. scabra</i> com identificação de resíduos conservados.	94
Figura 10: Heatmap construindo a partir da contagem de reads dos osmoprotetores em duas bibliotecas de RNA-Seq sob condição de estresse hídrico nos tempos de 6 hrs e 24 hrs.	97

Figura 11: Ancoragem dos transcritos de <i>S. scabra</i> em cromossomos virtuais de <i>A. duranensis</i> (n=10) onde cores diferentes representam famílias gênicas diferentes.	98
Figura 12. Árvore fenética construída usando o método de Neighbor-Joining para inferência sobre a diversidade de candidatos a genes de osmoprotetores em <i>S. scabra</i> .	100
Figura 13a. Dendrograma mostrando as relações entre ortólogos aos candidatos de P5CS.	103
Figura 13b. Dendrograma mostrando as relações entre ortólogos aos candidatos de P5CR.	104
Figura 14a. Dendrograma mostrando as relações entre ortólogos aos candidatos de MIPS.	105
Figura 14b. Dendrograma mostrando as relações entre ortólogos aos candidatos de IMP.	106
Figura 15a. Dendrograma mostrando as relações entre ortólogos aos candidatos de CMO.	107
Figura 15b. Dendrograma mostrando as relações entre ortólogos aos candidatos de BADH.	108
Figura 16a. Dendrograma mostrando as relações entre ortólogos aos candidatos de TPS.	109
Figura 16b. Dendrograma mostrando as relações entre ortólogos aos candidatos de TPP.	110

LISTA DE TABELAS

CAPÍTULO II: CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes Scabra* (Vogel) SOB DÉFICIT HÍDRICO.

	Página
Tabela 1: Sequências sondas buscadas nos bancos de dados utilizadas na mineração.	80
Tabela 2: Classes de osmoprotetores encontrados no transcriptoma de <i>S. scabra</i> . *TPS corresponde a sequências com apenas domínio TPS.	81
Tabela 3: Ponto isoelétrico, peso molecular e localização subcelular dos candidatos à osmoprotetores em <i>S. scabra</i> .	81

SUMÁRIO

CAPÍTULO I: CONSIDERAÇÕES GERAIS.....	18
1. INTRODUÇÃO GERAL.....	19
2. REFERENCIAL TEÓRICO.....	22
2.1 Estresses abióticos	22
2.2 Seca e déficit hídrico: efeitos sobre as plantas	23
2.3 Mecanismos de resistência ao déficit hídrico	24
2.3.1 Percepção do estresse	24
2.3.2 Mecanismos moleculares	26
2.3.3 Ajuste osmótico.....	29
2.4 Osmoprotetores	30
2.4.1 Diversidade, classificação e atividade	30
2.4.1.1 Prolina	32
2.4.1.2 Glicina-betaína	34
2.4.1.3 Mio-inositol	35
2.4.1.4 Trealose.....	36
2.5 Prospecção e análise de genes na era do sequenciamento de nova geração	38
2.5.1 Método Sanger e Sequenciamento de nova geração.....	38
2.5.2 Ômicas.....	41
2.6 Bioinformática	45
2.7 Família Fabaceae	47
2.7.1 Classificação, origem, distribuição e importância agrônômica.....	47
2.7.2 <i>Stylosanthes scabra</i>	49
2.7.2.1. Classificação e distribuição geográfica	49
2.7.2.2. Importância econômica.....	49
2.7.2.3 Bancos de germoplasma	50
2.7.2.4. Melhoramento genético	51
2.7.2.5 Análises genéticas e moleculares	52

3. REFERÊNCIAS BIBLIOGRÁFICAS	54
CAPÍTULO II: CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE OSMOPROTETORES NO TRANSCRIPTOMA DE <i>Stylosanthes scabra</i> (Vogel) SOB DÉFICIT HÍDRICO	74
1. INTRODUÇÃO	75
2. METODOLOGIA	77
2.1 Material vegetal, imposição do estresse, construção das bibliotecas de RNA-Seq e sequenciamento.....	77
2.2 Identificação e caracterização de transcritos em <i>S. scabra</i>	77
2.3 Ancoragem dos osmoprotetores de <i>S. scabra</i> no genoma de <i>A. duranensis</i> ..	78
2.4 Análises fenéticas dos osmoprotetores em <i>S. scabra</i>	79
2.5 Perfil de expressão <i>in silico</i>	79
3. RESULTADOS.....	80
3.1 Mineração das sequências candidatas e osmoprotetores em <i>S. scabra</i>	80
3.2 Análise de motivos	82
3.4 Ancoragem de sequências de <i>S. scabra</i> no genoma de <i>A. duranensis</i> :	97
3.5 Análises fenéticas:	99
4. DISCUSSÃO	111
4.1 Avaliação dos domínios e motivos conservados.....	111
4.2 Genes de osmoprotetores.....	115
4.3 Análise Fenética e de Classificação dos Genes	117
4.4 Expressão diferencial	120
5. CONCLUSÕES	122
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	123

RESUMO

Estresses abióticos, como o déficit hídrico, prejudicam o desenvolvimento das plantas causando mudanças morfológicas, fisiológicas, bioquímicas e moleculares, alterando a expressão de genes que ativam respostas bioquímicas e fisiológicas, como por exemplo, o acúmulo de solutos compatíveis. Os osmoprotetores são moléculas altamente solúveis que atuam principalmente no ajuste osmótico celular mantendo a sua turgidez (volume celular) mesmo com o potencial hídrico baixo. Abordagens moleculares e computacionais têm possibilitado novas perspectivas sobre a estrutura e função de proteínas e genes já conhecidos e se mostrado eficientes em estudos com espécies de genoma conhecido, no entanto, poucos estudos têm sido conduzidos com espécies nativas. *Stylosanthes scabra*, leguminosa com alto potencial forrageiro, é nativa do semi-árido brasileiro e apresenta tolerância à seca. O presente projeto teve como objetivo identificar e caracterizar transcritos candidatos pertencentes à diferentes classes de osmoprotetores no transcriptoma de *S. scabra* sob déficit hídrico. Inicialmente, sequências candidatas para genes osmoprotetores foram obtidas no banco de dados de *S. scabra*, procedendo-se com a anotação dos dados, tradução e identificação dos domínios conservados por meio de ferramentas *in silico*. Um total de 74 sequências de osmoprotetores com domínios conservados foram identificadas em *S. scabra*, sendo 40 de Trealose (TPS e TPP), 17 de Mio inositol (MIPS e IMP), 10 de Glicina betaína (BADH e CMO) e 7 de Prolina (P5CS e P5CR) resultados equivalentes ao descrito em outras espécies vegetais. Estes foram ancorados no genoma de *Arachis duranensis* (n=10) revelando um total de 19 *loci* ancorados similar ao encontrado em soja, outra espécie do gênero Fabaceae. Árvores fenéticas de ortólogos de genes de osmoprotetores demonstraram o agrupamento das sequências de *S. scabra* com outras espécies da família Fabaceae, comprovando sua similaridade e apontando uma conservação destas proteínas. A análise de expressão gênica diferencial *in silico* foi realizada utilizando dados de RNA-Seq onde dos 74 transcritos 35 mostraram-se diferencialmente expressos, apresentando indução e repressão em ambos os tempos de 6 h e 24 h de estresse indicando que estes participam da resposta do vegetal ao déficit hídrico. Esses resultados mostram-se valiosos para o melhoramento genético da espécie *S. scabra* e de outras leguminosas de importância agrícola.

Palavras-chave: Estresse abiótico; RNA-Seq; Fabaceae; Expressão gênica.

ABSTRACT

Abiotic stress conditions like water deficit affect plant growth causing morphological, physiological, biochemical and molecular changes, changing genes expression that activate biochemical and physiological responses, such as compatible solutes accumulation. Osmoprotectants are highly soluble compounds that act primarily on cellular osmotic adjustment maintaining their turgidity (cell volume) even with low water potential. Molecular and computational approaches have made possible new perspectives on the structure and function of proteins and genes already known and have proven efficient in studies with known genome species, however, few studies have been conducted with native species. *Stylosanthes scabra*, a legume with high forage potencial, it is Brazilian semiarid native specie and has drought tolerance. The present work aimed to identify and characterize transcripts belonging to different classes of osmoprotectants in *S. scabra* transcriptome of under water deficit as well as validate its expression in RT-qPCR. Initially, candidate sequences for osmoprotectants genes were obtained from *S. scabra* database proceeding with data annotation, translation and identification of conserved domains through *in silico* methods. A total of 74 osmoprotectants sequences with conserved domains were identified for *S. scabra*: 40 of Trehalose (TPS and TPP), 17 of Myo inositol (MIPS and IMP), 10 of Glycine betaine (BADH and CMO) and 7 of Proline (P5CS and P5CR) equivalent results described in other plants. These were anchored in the *A. duranensis* genome (n = 10) revealing a total of 19 anchored *loci* similar to the results that has been found in soybean, another Fabaceae specie. Phenetic orthologous trees of osmoprotectants genes demonstrated the clustering of *S. scabra* sequences with other Fabaceae species, proving their similarity and pointing out to these protein conservation. The *in silico* differential gene expression analysis was performed using RNA seq data, where the 35 transcripts were differentially expressed, presenting induction and repression in both 6 h and 24 h stress periods showing that these genes participate in *S. scabra* water deficit response. These results prove to be valuable for genetic breeding of *S. scabra* and other legumes of agricultural importance.

Keywords: Abiotic stress; RNA-Seq; Fabaceae; Gene expression.

CAPÍTULO I

CONSIDERAÇÕES GERAIS

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

1. INTRODUÇÃO GERAL

Estresses abióticos, como seca, baixas temperaturas, baixa fertilidade do solo e salinidade figuram entre as principais causas de perdas na qualidade e produtividade agrícola no mundo (Carmo-Silva et al. 2012; Awasthi et al. 2014). Estas adversidades ambientais causam efeitos nas plantas que levam a mudanças morfológicas, fisiológicas, bioquímicas e moleculares. As células vegetais percebem a condição estressante através de vias sensoriais que sinalizam mensagens secundárias, transdução de sinais hormonais e reguladores transcricionais (Cvikrova et al. 2013; Danquah et al. 2014; Gilroy et al. 2014). Muitos desses sinais convergem, induzindo genes que codificam proteínas e enzimas envolvidas diretamente no metabolismo, por exemplo sob estresse (Casaretto et al. 2016). O acúmulo de solutos compatíveis (osmoprotetores) é uma das estratégias utilizadas pelas plantas para suportar condições adversas que limitam ou impedem o seu desenvolvimento.

Osmoprotetores são moléculas de baixo peso molecular, altamente solúveis e não tóxicas mesmo em altas concentrações. Seu acúmulo permite que o indivíduo sobreviva em situações extremas. São moléculas eletricamente neutras que não inibem a ação de enzimas e aumentam a capacidade da célula de preservar o balanço osmótico em situações estressantes, mantendo sua estrutura e função. Destacam-se também como potentes estabilizadores de proteínas, protegendo a membrana celular dos efeitos do estresse. Estão distribuídos em vários grupos taxonômicos, desde algas até plantas superiores, incluindo álcoois de açúcar (ex. mio-inositol), açúcares complexos (ex. trealose e frutanos) e metabólitos (ex. glycina-betaina, prolina e ectoína) (Yancey 2005; Geeta 2014).

A prolina é caracterizada como um aminoácido, essencial no metabolismo primário, atuando na proteção celular e agindo como um agente osmótico nas plantas superiores. Duas enzimas participam da biossíntese de prolina via glutamato (que é a maior fonte de prolina encontrada nas plantas), a PC5S (Pyrroline- 5-Carboxylase Synthetase) e a PC5R (Pyrroline-5-Carboxylase Reductase) (Sengupta et al. 2016). O acúmulo de prolina é diretamente proporcional à atividade da enzima P5CS (Delauney e Verma 1993; Szekely et al. 2008) codificada por dois genes: P5CS1 e P5CS2. O gene P5CS1 é induzido por condições de estresse abiótico, como a seca, salinidade e baixas temperaturas (Strizhov et al. 1997; Ginzberg et al. 1998; Abraham et al. 2003). Zhang et al.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

(2013) demonstraram a conexão entre a metilação do DNA e o acúmulo de prolina durante estresse osmótico no arroz.

Classificado como um composto amônio-quaternário (QAC), a glicina-betaína (GB) está presente na maioria dos organismos. Está contida nos cloroplastos e plastídios em condições de estresse hídrico, salino e de baixas temperaturas (Jagendorf e Takabe 2001). Genes que codificam as enzimas envolvidas na biossíntese da glicina-betaína têm sido identificados, clonados e empregados em plantas que não acumulam GB, como *Arabidopsis thaliana*, *Brassica napus*, *Nicotiana tabacum* e *Oryza sativa*, tendo em vista que sua superexpressão confere resistência a estresses abióticos (McCue e Hanson 1990).

O mio-inositol é um álcool cíclico importante no crescimento e desenvolvimento dos vegetais. Em condições de estresse salino, facilita o sequestro de Na⁺ protegendo o aparato fotossintético das plantas (Nelson et al. 1998). Nisa et al. (2016) analisando o gene *GsMIPS2*, que codifica a enzima mio-inositol-1-phosphate synthase (MIPS EC 5.5.1.4) participante da biossíntese de mio-inositol em *Glycine soja* (soja selvagem), concluíram que sua expressão favorece a tolerância à salinidade.

A trealose é um dissacarídeo composto de duas moléculas de glicose, unidas por uma ligação alfa-alfa (Sengupta 2016), sendo descrita inicialmente em plantas revivescentes ou de ressurreição (*resurrection plants*) que conseguem sobreviver à extrema desidratação, 'revivendo' após receberem água, como, por exemplo, em *Myrothamnus lepidophylla* e na *Selaginella tamariscina* (Godijn e Van Dun 1999).

Essas proteínas importantes em tais vias metabólicas provêm do conjunto de transcritos produzidos sob essas condições de estresse, definido como transcriptoma sendo extremamente importante na compreensão de variados processos biológicos (Alba et al 2004; Subramanian et al 2005). Seu estudo é essencial na interpretação dos elementos funcionais do genoma, tendo como principais objetivos: (a) catalogar os transcritos de um tecido/organismo (incluindo mRNA, RNAs não codificantes e pequenos RNAs); (b) definir a estrutura dos genes; (c) quantificar os níveis de expressão dos genes (Morozova - Marra 2008). Nesse contexto, a tecnologia RNA-Seq permite sequenciar o transcriptoma total de um indivíduo com alta qualidade e profundidade na detecção dos transcritos (Cardoso 2015). Possui algumas vantagens em comparação com outras: (a) quantifica os níveis de expressão dos genes com maior precisão; (b) possibilita o desenvolvimento de SNPs, permitindo estimar a expressão de alelos específicos; (c) identifica *splicing* alternativo ou sequências

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

rearranjadas quando se tem um genoma de referência; (d) permite a detecção de transcritos mesmo em níveis baixos; (e) não necessita de conhecimento prévio do genoma do organismo (Marguerat - Bahler 2010).

Para compreender os efeitos causados pela transcrição e tradução de um gene é necessário entender sua expressão local e seus efeitos em nível dos processos celulares e impactos nos sistemas regulatórios e no fenótipo da planta de forma geral e em uma visão espaço-temporal. Com isso, a bioinformática se apresenta como uma importante ferramenta na análise da expressão diferencial *in silico*, através da construção de bibliotecas de dados em diferentes tecidos e condições, permitindo a identificação de genes envolvidos nas respostas a fatores bióticos e abióticos (Leal-Bertinoli et al. 2003).

As espécies do gênero *Stylosanthes* Swartz. figuram como importantes leguminosas forrageiras para utilização em áreas tropicais e semiáridas, especialmente em solos ácidos e com baixa fertilidade natural (Guodao et al. 1997). A família Leguminosae ou Fabaceae é dividida em três grandes subfamílias, Mimosoideae, Caesalpinioideae e Papilionoideae (Judd et al. 2002). Tanto o gênero *Arachis* quanto o gênero *Stylosanthes* estão incluídos na subfamília Papilionoideae, na tribo Dalbergieae, confirmando assim, a proximidade dos gêneros (Klitgaard e Lavin 2005).

A espécie *Stylosanthes scabra* (Vogel), nativa do semiárido brasileiro, apresenta tolerância à seca e à baixa fertilidade do solo, além de ter um grande potencial forrageiro (Stace e Edye 1984), sendo bem difundidas em pastagens de regiões de clima temperado, a exemplo da Austrália, que foi o primeiro país a reconhecer o potencial das espécies do gênero como forrageiras (Edye 1997). Apesar desse potencial, as espécies do gênero ainda são pouco utilizadas em regiões semiáridas de clima tropical (Karia 2008).

O presente estudo pretende avaliar o perfil de expressão de genes candidatos a produção de osmoprotetores nas condições déficit hídrico em *S. scabra*, inferindo sobre seu papel e estrutura, comparando-os com a espécie relacionada *Arachis duranensis* (Krapov. & W. C. Greg.).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

2. REFERENCIAL TEÓRICO

2.1 Estresses abióticos

Fatores externos que resultam em algum tipo de influência negativa sobre a planta, podem ser caracterizados como estresses e comumente induzem respostas que podem ser reversíveis ou não em todos os níveis do vegetal (Salamoni 2008). Segundo Linchtenthaler (1996) o estresse, a depender da situação imposta, é definido como um agente favorável ou desfavorável, sendo caracterizado como: eu-estresse, sendo um fator ativador, estimulante e positivo para o desenvolvimento; e dis-estresse, que seria um caso severo de dano permanente, que irá afetar negativamente o vegetal.

Estresses abióticos, como seca, altas e baixas temperaturas, baixa fertilidade do solo e salinidade figuram entre as principais causas de quedas na qualidade e na produtividade agrícola no mundo (Benko-Iseppon et al. 2011, Carmo-Silva et al. 2012, Awasthi et al. 2014). Tais condições afetam o desenvolvimento vegetal, impedindo que o mesmo utilize todo o seu potencial genético e resulte, em caso extremo, na morte da planta (Mahajan e Tuteja 2005). Quando a demanda de energia pelo vegetal é maior que a sua produção ocorre inicialmente uma desestabilização de suas funções, que podem ou não ser restauradas através da ativação de processos fisiológicos de adaptação (Larcher 2000).

As plantas desenvolveram mecanismos essenciais ao longo de seu processo evolutivo, de maneira a se adaptar a estresses ambientais (Wang et al. 2001, Borghetti et al. 2005). Sendo organismos sésseis, ou seja, que geralmente não se deslocam de seu local de fixação, os vegetais são expostos constantemente a condições ambientais adversas, devendo promover processos para “tolerar” tais estresses (Taiz e Zeiger 2004, Shao et al. 2007). A tolerância está ligada à capacidade da planta em conseguir manter seu metabolismo em níveis consideráveis, não prejudicando o seu desenvolvimento, mesmo sob condições desfavoráveis, onde tal quadro pode ser irreversível, dependendo do estágio de desenvolvimento do vegetal, genótipo (Willadino e Camara 2010), do tempo de exposição a tal condição estressante e também da sua severidade (Chaves et al. 2003, Ghelfi et al. 2011).

Tais estresses desencadeiam respostas bioquímicas e fisiológicas nas plantas, de modo a proporcionar uma tolerância, aumentando as chances de sua

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

sobrevivência (Benko-Iseppon et al. 2012). As perdas causadas por tais estresses ocasionam quedas na produção das colheitas em níveis preocupantes, podendo chegar a mais de 50 %, desta forma compreendendo um enorme desafio a ser enfrentado, principalmente devido à crescente população mundial e às mudanças climáticas (Bray et al. 2000, Mahajan e Tuteja 2005). Com o aumento dos níveis de dióxido de carbono na atmosfera e a consequente elevação na temperatura média do ar, haverá uma maior demanda atmosférica pela água, além da ocorrência, em várias regiões, de uma frequência de chuvas cada vez mais espaçada levando à ocorrência de períodos prolongados de seca (Dore 2005).

2.2 Seca e déficit hídrico: efeitos sobre as plantas

Sendo definida como um estado onde a demanda por energia da planta supera a sua produção para manter a sua sobrevivência, a seca é classificada como um estresse ambiental que leva inicialmente a um imediato desequilíbrio das funções do vegetal, posteriormente normalizado, levando a um decréscimo em seu crescimento, seguido pela indução de processos fisiológicos visando à adaptação da planta a tal condição (Larcher 2006). Por sua vez, o déficit hídrico pode ser caracterizado como um decréscimo da disponibilidade de água para o vegetal, em relação ao que é necessário à sua sobrevivência. À medida que o solo seca, a retenção aumenta e a disponibilidade de água cai, tornando-se mais difícil para a planta absorvê-la (Bergamaschi 1992). Para Fan et al. (2006) o déficit hídrico é determinado quando a planta não consegue absorver água suficiente pelo sistema radicular de modo a suprir as demandas do vegetal. Tais demandas, principalmente a fotossíntese e a transpiração, são essenciais para o desenvolvimento vegetal (Davies e Zhang 1991).

Além de afetar a fotossíntese e a transpiração, a baixa disponibilidade de água influencia outros processos metabólicos da planta tais como: fechamento estomático e redução da condutância estomática além de afetar o crescimento do vegetal (Silva et al. 2002, Gomes et al. 2004, Portes et al. 2006). Segundo Lima et al. (2003), a capacidade dos estômatos em controlar a perda de água influencia diretamente no crescimento e desenvolvimento de plantas que passaram por períodos de déficit hídrico. Sendo um dos primeiros mecanismos de resposta vegetal para evitar a perda de água, o fechamento dos estômatos promove uma redução na taxa fotossintética em situações de déficit hídrico no solo ou na atmosfera. Em condições moderadas de

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

estresse hídrico, os estômatos se fecham, a fim de reduzir a perda de água através da transpiração e conseqüentemente eleva a temperatura das folhas, por fim, causando uma perda na eficiência de transformação de energia luminosa para energia química (Chaves et al. 2002, 2003).

A absorção de água pelas plantas ocorre quando o potencial hídrico das raízes é mais negativo que o potencial hídrico do solo (Larcher 1995). Por sua vez, o déficit hídrico moderado afeta o desenvolvimento do sistema radicular. A razão da biomassa de raiz/parte aérea é controlada por um balanço funcional entre a absorção de água pelas raízes e a fotossíntese, sendo tal balanço afetado negativamente em condições de baixa disponibilidade hídrica. Inicialmente, a expansão foliar é afetada e (à medida que o déficit hídrico avança) há um crescimento de raízes mais profundas, buscando zonas ainda úmidas do solo, visto que as primeiras camadas do solo secam rapidamente (Hoogenbomm et al. 1987). Estudos comprovam que o déficit hídrico do solo aumenta a concentração de ácido abscísico (ABA) na raiz, provocando o fechamento dos estômatos e a redução da expansão foliar (Taiz e Zeiger 2010).

O déficit hídrico, além de afetar o crescimento da planta, afeta principalmente a produtividade do vegetal e pode ocorrer em qualquer região, até mesmo em regiões úmidas (Cushman e Bohnert 2000). Estas adversidades causam efeitos nas plantas que resultam em mudanças morfológicas, fisiológicas, bioquímicas e moleculares. As células vegetais percebem a condição estressante através de vias sensoriais que sinalizam mensagens secundárias, hormônios, transdutores de sinais e reguladores transcricionais (Cvikrova et al. 2013, Danquah et al. 2014, Gilroy et al. 2014), sendo de suma importância o entendimento de como essas respostas adaptativas ocorrem e são ativadas. O reconhecimento de tais processos é fundamental, principalmente para o desenvolvimento de novas cultivares comerciais tolerantes à seca (Assad 2002). A compreensão de como o déficit hídrico prejudica o desenvolvimento das culturas é de grande importância para o desenvolvimento de cultivares tolerantes à seca, seja através da transformação genética ou pelo melhoramento genético convencional (Fageria 1989).

2.3 Mecanismos de resistência ao déficit hídrico

2.3.1 Percepção do estresse

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

De maneira geral, os estresses abióticos se iniciam com a percepção do estresse pelo vegetal, que irá desencadear uma série de eventos moleculares que culminarão na transdução de sinais e posteriores respostas intermediadas por fitormônios, fatores de transcrição (FTs), mensageiros secundários, ativando genes e/ou proteínas sinalizadoras (Pandey 2008). Estudos demonstraram que respostas a estresses do tipo hídrico ou salino geram mecanismos de percepção que englobam vários genes, os quais estão divididos em duas categorias: (a) os que possuem relação na sinalização e regulação transcricional, como HSF, CBF/DREB, bHLH, bZIP, ERF (Figura 1 A) e (b) os que atuam na defesa ou preservação das membranas celulares e de proteínas – como como proteínas de choque térmico e chaperonas; proteínas LEA, osmoprotetores; removedores de radicais livres; aquaporinas; trocadores e transportadores de íons (Figura 1B), que atuam no transporte de água e íons (Seki et al. 2002, Benko-Iseppon et al. 2011). Uma vez ativados, os genes de resposta ao estresse detoxificam a ROS, ativando enzimas de detoxificação e limpadores de radicais livres, reativando enzimas essenciais e proteínas estruturais que, por fim, reestabelecem a homeostase celular resultando no desenvolvimento da tolerância ao estresse imposto (Figura 2) (Hasanuzzaman et al. 2013).

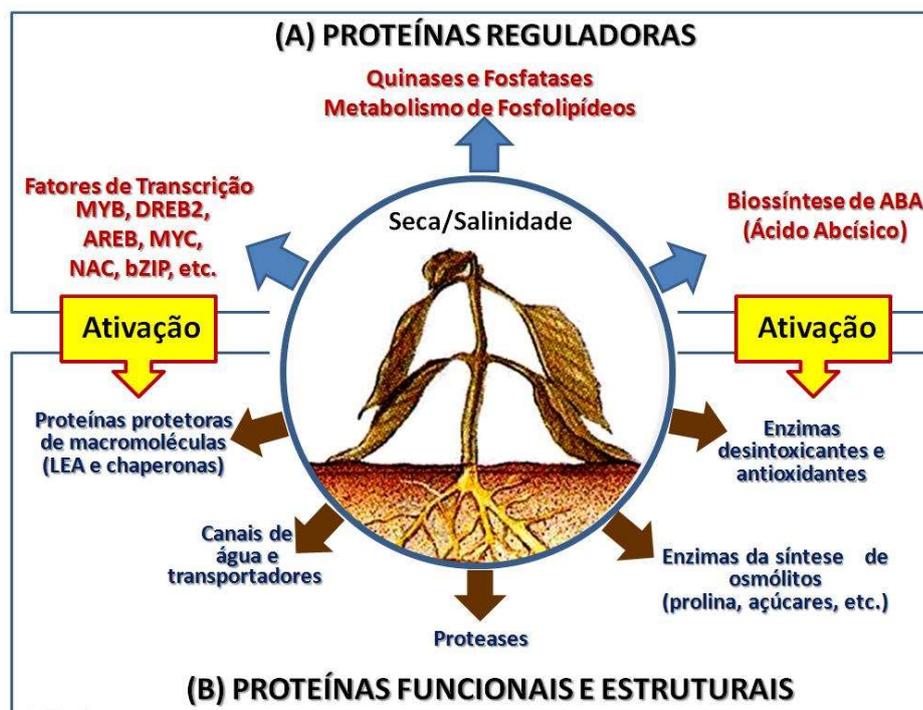


Figura 1. Resposta do vegetal a estresses abióticos com ênfase no déficit hídrico e salinidade, onde o vegetal ao perceber a situação estressante, envia sinais que são traduzidos por meio de fitormônios, tradutores de sinais e reguladores de transcrição,

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

provocando a indução de genes que codificam para proteínas funcionais e estruturais que estão diretamente envolvidas na tolerância ao estresse (Fonte: Benko-Iseppon et al. 2011).

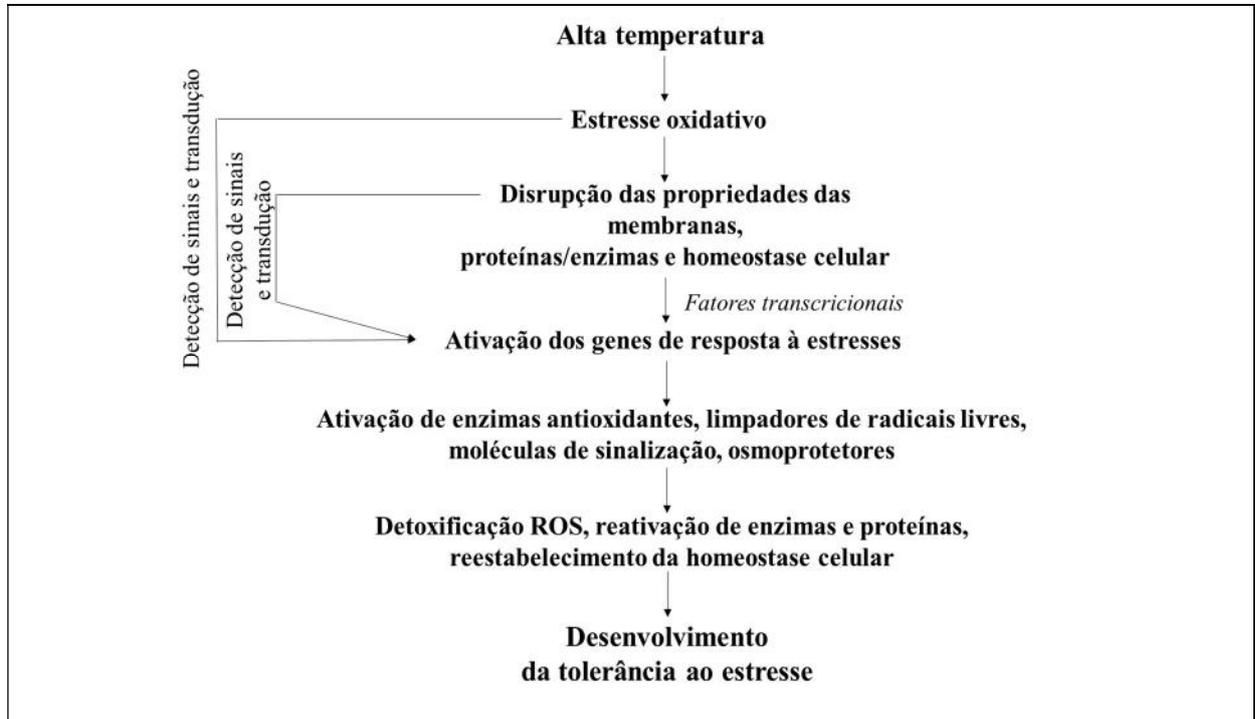


Figura 2. Esquema da transdução de sinais induzidas pelo calor e desenvolvimento da tolerância ao estresse nas plantas. Adaptado de Hasanuzzaman et al. (2013).

A análise de genes que são induzidos sob condições de estresse figura entre as mais efetivas abordagens tanto para o entendimento dos mecanismos moleculares frente à tolerância a tais estresses, como também à resposta do vegetal aos mesmos para identificar/introduzir genes de tolerância em cultivares suscetíveis por inferências biotecnológicas. Há vários exemplos de genes envolvidos na resposta a estresses abióticos com enorme potencial biotecnológico (Seki 2003, Avni Öktem et al. 2008). A expressão desses genes tem sido identificada, promovendo a adaptação da planta e sua sobrevivência sob condições de estresse (Tuteja 2009). Plantas estressadas percebem os sinais externos e internos através de diferentes vias interligadas entre si, que são usadas na regulação das respostas e consequente no desenvolvimento da tolerância (Kaur e Gupta 2005).

2.3.2 Mecanismos moleculares

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Os estímulos externos ocasionados pela seca são percebidos e capturados por sensores nas membranas celulares, onde esses sinais são transmitidos por diversas rotas que resultam na expressão de genes responsáveis pela adaptação ao estresse hídrico. Uma variedade de mensageiros secundários como Ca^{2+} , ROS, ABA, fosfogliceróis, diacilglicerol e fatores de transcrição executam um importante papel em várias rotas de transmissão de sinais (Kaur e Asthir 2017).

A ocorrência de ROS nas células das plantas contém o ânion superóxido (O_2^-), o radical hidroxila ($\cdot\text{OH}$) e o peróxido de hidrogênio (H_2O_2), que são produzidos em inúmeros compartimentos celulares, tais como cloroplastos, mitocôndrias, peroxissomos, retículo endoplasmático, parede celular e apoplasto (Gill e Tuteja 2010, Sharma et al. 2012). Seu nível é regulado através de vias que metabolizam genes denominados recicladores (Bhattacharjee 2012). Uma infinidade de fatores do ambiente pode prejudicar o balanço entre a produção e a eliminação de ROS, ocasionando quase sempre em dano oxidativo devido ao aumento dos agentes oxidantes dentro das células. Os vegetais então fabricam uma série de compostos antioxidantes de modo a impedir tais danos, como por exemplo, a superóxido dismutase (SOD), a catalase (CAT), a ascorbato peroxidase (APX), a monodehidroascorbato redutase (MDHAR) e a glutatona redutase (GR) (Miller et al. 2010, Shama et al. 2010, Krishnamurthy e Rathinasabapathi 2013). A sinalização de ROS em resposta a estresses, especialmente o abiótico, interage com fitormônios reguladores de resposta ao estresse, com destaque para o ácido abscísico (ABA). A biossíntese do ácido abscísico (ABA) ocorre logo após a percepção do estresse pelo vegetal, sendo incitada nas raízes, mas podendo também ser produzida nos mesófilos das folhas. No entanto, é no tecido radicular que o ABA é encontrado em grandes quantidades, sendo transportado via xilema, para a parte aérea da planta (Milosavljević et al. 2012). Este fitormônio está ligado a respostas fisiológicas importantes, tais como o fechamento dos estômatos, ativação de genes e o acúmulo de metabólitos, íons e solutos compatíveis que irão conferir à planta uma proteção frente ao déficit hídrico (Fujita et al. 2011).

De acordo com Kim (2014), os genes de resposta à seca podem ser divididos em dois grupos: ABA-dependente e ABA-independente (Figura 3). Os genes induzidos via ABA-dependente durante o estresse hídrico são controlados por cinco classes diferentes de fatores de transcrição. O elemento responsivo ao ABA (ABRE) com o consenso ACGTGG/TC é ligado ao domínio básico tipo zíper leucina (bZIP). Três

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

fatores de transcrição bZIP em *A. thaliana* (AREB1/ABF2, AREB2/ABF4 e ABF3) são expressos em resposta à deficiência de água e a ativação destes requer o acúmulo de ABA e a indução de uma proteína quinase ABA-responsiva que irá ativar os fatores de transcrição através da fosforilação. Outros fatores de transcrição também estão envolvidos na expressão de genes via ABA durante o déficit hídrico celular. Três genes que codificam uma classe de FT exclusiva de plantas, domínio proteico NAC (ANAC019, ANAC055 e ANAC072), sendo seus membros induzidos pelo déficit hídrico e tratamento com ABA. O gene RD22 de *Arabidopsis*, induzido pelo déficit hídrico, foi descoberto sendo regulado pelo ABA. A região promotora de RD22 contém também MYC (CANNTG) e MYB (C/TAACNA/G) que são FTs acumulados após um aumento na concentração de ABA. A superexpressão desses FTs resulta em uma melhora na sensibilidade ao ABA e na tolerância à seca (Abe et al. 2003)

Em relação à via ABA-independente, a família gênica DREB figura como um importante FT que induz uma série de genes relativos a estresses abióticos, conferindo assim, resistência às plantas. Estes são divididos em dois grupos: DREB1, que está envolvido na via de transdução de sinais relativos a baixas temperaturas e DREB2, relacionado à via de transdução de sinais pela desidratação (Agarwal et al. 2006). Outra via independente de ABA foi identificada após a observação do gene de resposta primária à desidratação (ERD1), sendo verificado que seus transcritos acumulavam antes de qualquer aumento de ABA em resposta à desidratação e alta salinidade (Nakashima et al. 1997). Análises dos promotores de ERD1 revelaram que fatores de transcrição pertencentes à família NAC e dedo de zinco homeodomínio (ZF-HD) são essenciais na ativação do gene ERD1. O aumento da tolerância à seca pode estar relacionado com a redução da taxa de transpiração (aumento do fechamento estomático) e ao aumento da sensibilidade ao ABA (Tran et al. 2007).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

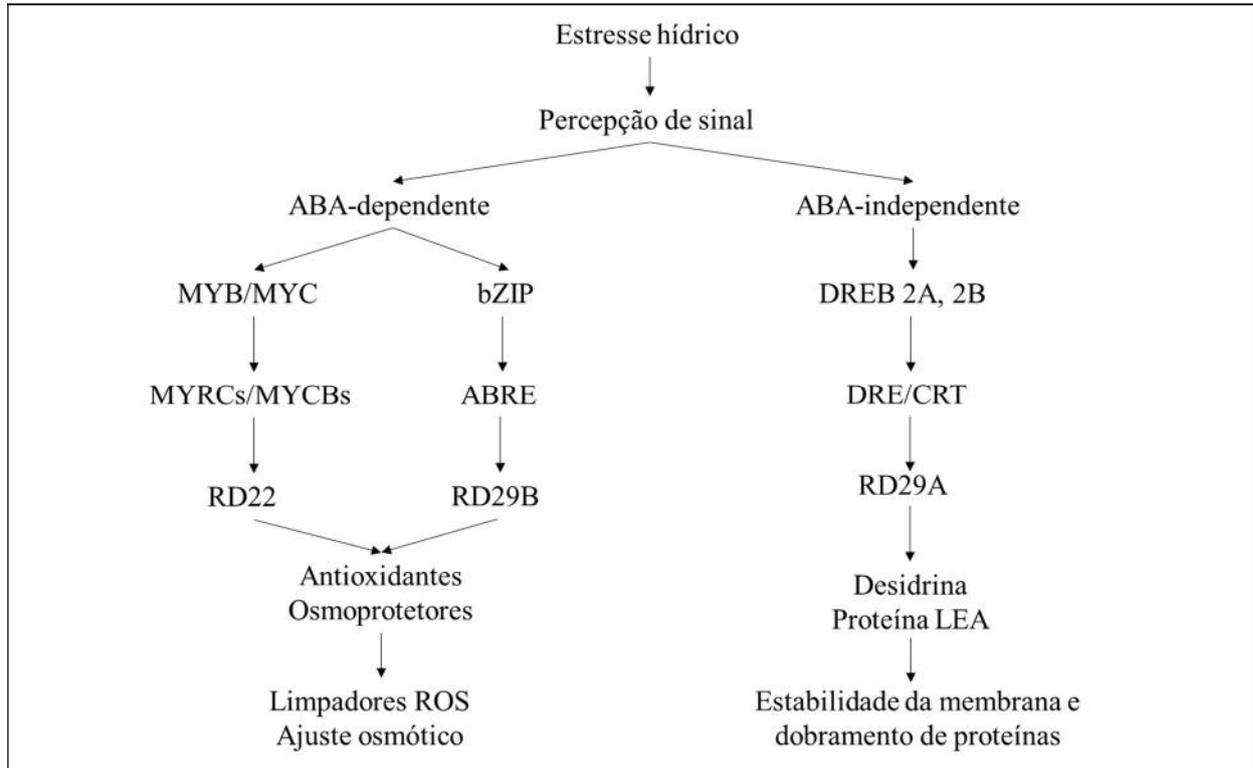


Figura 3. Rota de sinalização de ABA-dependente e ABA-independente durante estresse hídrico. bZIP - proteínas de dedo de zinco, ABRE - Elementos responsivos a ABA, DREB - Elemento responsivo à desidratação, DRE / CRT - Repetição C (CRT) / responsivo à desidratação fator de ligação a elementos (GRE), LEA - proteínas abundantes na embriogênese tardia, MYB - oncogene mieloblastose, MYCR / MYCBs - sequência de reconhecimento MYC / MYB, ROS - espécies reativas de oxigênio, RD - responsivo à desidratação. Adaptado de Kaur e Asthir (2017).

2.3.3 Ajuste osmótico

O ajustamento osmótico pode ser definido como o processo de acúmulo de solutos nas células nas situações em que seu potencial hídrico está reduzido, ajudando na manutenção do turgor celular (Chaves e Oliveira 2004). Desta forma, contribui ativamente na sustentação de processos fisiológicos, tais como abertura estomática, fotossíntese e crescimento do vegetal (Condon 1982, Blum et al. 1983, Morgan 1984, Ludlow e Muchow 1990, Blum 1996; Kamoshita et al. 2001). O aumento na concentração de solutos também é uma consequência da redução na quantidade de água durante o déficit hídrico (Sauter et al. 2001) e o acúmulo de compostos osmóticos como aminoácidos (como a prolina), glicina-betaína e açúcares (como a trealose) auxilia na proteção da planta contra efeitos deletérios deste estresse, não apenas pelo ajustamento osmótico mas também pela detoxificação de ROS, proteção

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

da integridade das membranas e a estabilização de enzimas ou proteínas (Ashraf e Foolad 2007). O impedimento da produção de ROS, realizado pela mitocôndria durante o estresse hídrico, compreende uma estratégia importante que permite que a planta lide com a escassez de água sem sofrer muitos danos (Xu et al. 2011).

Plantas transgênicas superexpressando a enzima P5CS (pirroline-5-carboxilase sintetase) demonstraram o acúmulo de prolina durante estresse hídrico, configurando assim, uma resistência ao estresse osmótico (Khan et al. 2015). Glicina-betaína, um amônio-quaternário, também desenvolve um importante papel no estresse abiótico em plantas, como soluto compatível, interagindo diretamente com macromoléculas ou formando paredes de hidratação em torno de complexos macromoleculares para prevenir a sua desnaturação (Giri 2011). Estudos recentes mostram que a trealose, desenvolve um papel significativo no crescimento de plantas sobre estresse hídrico (Nuccio et al. 2015), observando-se que a superexpressão de diferentes isoformas da enzima trealose-6-fosfato sintase conferiu resistência à seca em arroz (Li et al. 2011).

2.4 Osmoprotetores

2.4.1 Diversidade, classificação e atividade

O conceito de solutos compatíveis foi introduzido inicialmente por Brown (1976). São caracterizados como compostos solúveis de baixo peso molecular, comumente neutros e atóxicos, que não interferem no funcionamento celular mesmo em grandes concentrações. Seu acúmulo favorece a adaptação dos organismos a situações ambientais adversas (Sengupta et al. 2016). Tais osmólitos estão presentes em quase todos os organismos, desde bactérias até plantas, que sintetizam esses compostos em resposta ao estresse osmótico (Yancey 2001).

O acúmulo desses compostos solúveis provoca um decréscimo no potencial osmótico celular, promovendo a manutenção do turgor da célula e a absorção de água, contribuindo para que os processos fisiológicos essenciais do vegetal sejam mantidos, como a fotossíntese, abertura estomática e crescimento (Blum, Mayer e Gozlan 1983, Morgan 1984, Ludlow e Muchow 1990, Blum 1996). Os osmoprotetores atuam estabilizando as macromoléculas da célula e conseqüentemente diminuindo perdas de atividades enzimáticas, bem como da integridade da membrana celular.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Além disso, possibilitam que as células recobrem seu volume e concentrações internas de sal (Génard et al. 1991). Esses compostos estão divididos quimicamente em quatro grupos: (a) metilaminas, (b) polióis, (c) açúcares (como a glicina-betaína e a trealose) e (d) aminoácidos e derivados (como a prolina). A glicina-betaína, a trealose e a prolina são os principais osmoprotetores que os vegetais superiores acumulam durante situações de estresse hídrico ou salino, ajudando na sobrevivência da planta em determinada condição adversa (Rhodes e Hanson 1993, Brigotti et al. 2003). Diversos estudos vêm sendo desenvolvidos a partir da utilização desses solutos compatíveis, visando aumentar a produtividade de culturas em situações de déficit hídrico e sua posterior aplicação em programas de melhoramento vegetal (Blum et al. 1983, Ludlow e Muchow 1990, Tangprensri et al. 1991, Belhassem et al. 1995, Zhang, Nguyen e Blum 1999, Claussen 2005).

A abundância de osmólitos proporciona um aumento na osmolaridade celular e do fluxo de água, proporcionando o turgor necessário para que a célula possa se expandir. Outra característica importante desses solutos compatíveis é a habilidade de promover estabilidade estrutural a proteínas, prevenindo seu enovelamento (*fold*ing) incorreto durante períodos de estresse (Lamitina et al. 2006). Alterações metabólicas que resultam de longa exposição a estresses abióticos variam de espécie para espécie. Alguns vegetais acumulam aminoácidos, aminas, açúcares e álcoois de açúcar em altos níveis quando expostos à seca, salinidade ou baixas temperaturas por vários dias (Sengupta et al. 2016).

A prolina compreende um conhecido soluto compatível, sendo acumulada quando as plantas são submetidas a situações de seca, salinidade ou baixas temperaturas. A alta concentração dessa substância desempenha um papel crucial na proteção da atividade fotossintética durante o estresse, suprimindo a atividade oxigenase (Sengupta et al. 2016). Outro composto que também tem sido amplamente estudado no que diz respeito a estresses abióticos é a glicina-betaína, a qual possui um papel de proteção contra a seca, salinidade e baixas temperaturas. Plantas que conseguem acumular naturalmente glicina-betaína demonstraram um aumento na tolerância a estresses, entretanto, nem todas as plantas conseguem acumular esse composto, incluindo algumas com grande importância econômica como o arroz, tomate, batata e *Arabidopsis*, as quais não conseguem acumular glicina-betaína em situações normais e de estresse (Yang e Lu 2005), sendo essas espécies um potencial alvo na introdução de genes que sintetizam esse composto.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Durante o estresse hídrico, o açúcar pode ser substituído por água e, dessa forma, a superfície da membrana é hidratada e protegida de fusão, além de formar uma barreira entre a membrana e a molécula de fosfolípido (Sengupta et al. 2016). O acúmulo de trealose, um dissacarídeo composto por duas moléculas de glicose, foi inicialmente observado em plantas revivescentes ou de ressurreição. Em *Arabidopsis*, a quantidade de trealose duplica quando a planta é exposta por quatro horas a altas temperaturas e seu conteúdo aumenta em quase oito vezes quando a planta é exposta a baixas temperaturas por quatro dias (Kaplan et al. 2004).

O aspecto em comum a esses solutos compatíveis citados é o acúmulo dos mesmos em altas concentrações sem prejudicar o metabolismo (Bohnert e Jensen 1996). O acúmulo desses compostos está associado a altos níveis de tolerância a estresses em várias espécies plantas, não sendo seus benefícios limitados a uma determinada espécie (Rontein et al. 2002) A superexpressão de genes que sintetizam esses compostos em plantas transgênicas produz com sucesso variedades resistentes a diversos estresses abióticos. Por exemplo, plantas de tabaco (*N. tabacum*) foram transformadas com o gene BADH (que sintetiza glicina-betaína) de *Spinacia oleracea* (Yang et al. 2008, Hasthanasombut et al. 2010), enquanto o arroz (*O. sativa*) foi transformado com o gene P5CS (que sintetiza prolina) de *Vigna aconitifolia* (Kishor et al. 1995) e com o gene TPS (que sintetiza trealose) de *Saccharomyces cerevisiae* (Holmstrom et al. 1996, Karim et al. 2007).

2.4.1.1 Prolina

A prolina é caracterizada como um dos aminoácidos presentes nas proteínas de todos os organismos vivos. Sua síntese ocorre em todo o reino vegetal em consequência do estresse osmótico, sendo essencial no metabolismo primário, protegendo a célula de danos e agindo como agente osmótico (McCue e Hanson 1990, Delauney et al. 1993). Em plantas, a principal fonte de biossíntese de prolina é a via do glutamato (Glu), principalmente quando se trata de estresse osmótico (Delauney e Verma 1993). Duas enzimas catalisam esta via, a Δ 1-pirrolina-5-carboxilato-sintase (P5CS) na primeira etapa e a Δ 1-pirrolina-5-carboxilato-redutase (P5CR) na etapa final (Figura 4). Em geral, existem dois genes que codificam a pyrroline-5-carboxylate synthetase (P5CS) e um gene que codifica para pyrroline-5-carboxylate redutase (P5CR), em plantas. As enzimas envolvidas na biossíntese da

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

prolina via glutamato estão localizadas no citoplasma e nos cloroplastos (Sengupta et al. 2016).

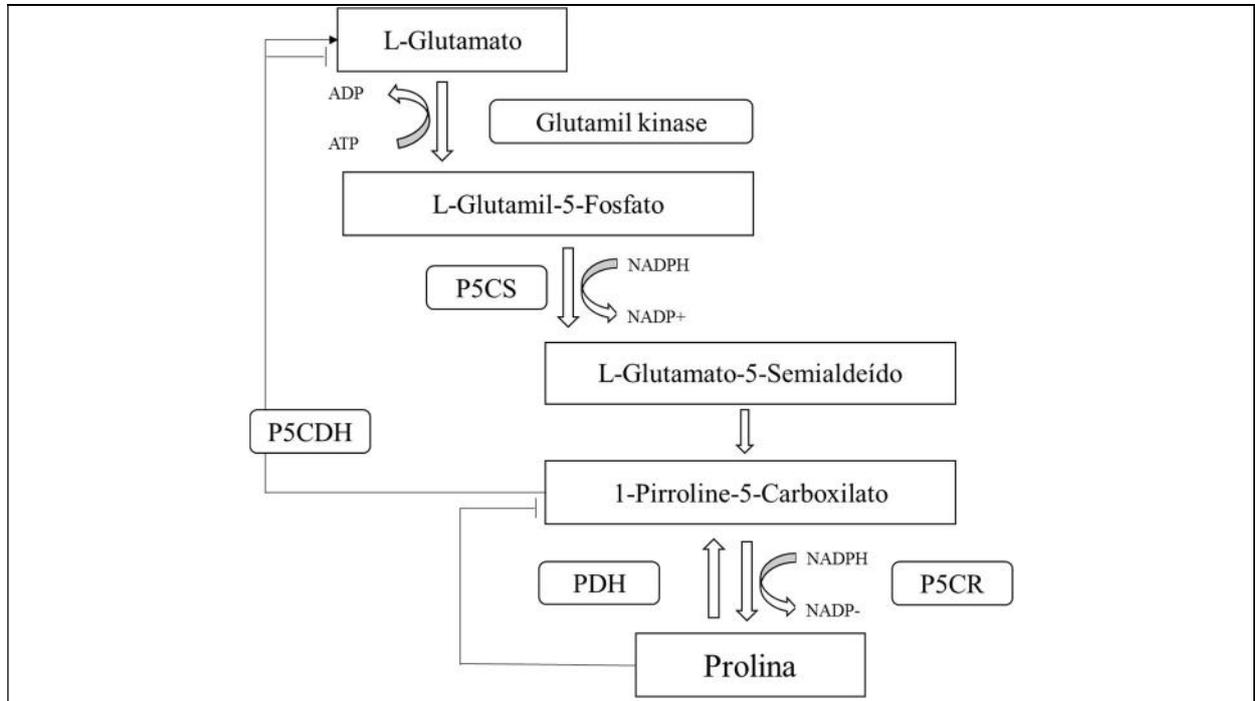


Figura 4. Via metabólica de Prolina em vegetais durante estresse abiótico. [Pyrroline-5 carboxylate synthetase (P5CS), pyrroline-5-carboxylate redutase (P5CR), ornithine-delta-aminotransferase (OAT), proline dehydrogenase (PDH), pyrroline-5-carboxylate dehydrogenase (P5CDH)]. Adaptado de Sengupta et al. (2016).

O acúmulo de prolina está ligado diretamente à atividade da enzima P5CS, a qual é codificada por dois genes denominados P5CS1 e P5CS2 que são expressos, em sua maioria, em células em divisão nas raízes, respectivamente, em resposta a estresses abióticos (Strizhov et al. 1997, Szekely et al. 2008) Já foi demonstrado que a expressão do gene P5CS1 em tecidos vegetais foi induzida em condições de estresse abiótico, como seca, salinidade e baixas temperaturas (Strizhov et al. 1997, Ginzberg et al. 1998, Abraham et al. 2003). Por sua vez, transcritos de P5CR aumentam consideravelmente a sua expressão em resposta ao estresse osmótico, indicando que o gene P5CR é regulado sob tal condição (Delauney e Verma 1990, Williamson e Slocum 1992, Verbruggen et al. 1993). Adicionalmente, Zhang et al. (2013) demonstraram a conexão entre a metilação do DNA e o acúmulo de prolina durante estresse osmótico no arroz. Signorelli e Monza (2017) identificaram dois genes diferentes envolvidos na síntese de prolina em *Lotus japonicus*, LjP5CS1 e

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

LjP5CS2 e analisaram a expressão dos mesmos em uma condição de estresse osmótico, incubando por 24 h as folhas em uma solução de 30% de polyethylene glycol 8000, que simula um potencial hídrico de -1.5 Mpa e constataram que apenas a expressão do gene LjP5CS1 foi induzida durante a condição, concluindo que o mesmo era responsável pelo acúmulo de prolina na espécie especialmente durante o déficit hídrico.

A prolina também possui outras funções importantes além do ajuste osmótico, tais como proteção e manutenção da membrana celular (Mansour 1998), redução de energia (Verbruggen et al. 1996), eliminação de radicais hidroxil (Hong et al. 2000) além de servir também como fonte de carbono e nitrogênio (Ahmad e Hellebust 1988, Peng et al. 1996).

2.4.1.2 Glicina-betaína

Classificado como um composto amônio-quaternário (QAC), a glicina-betaína, está presente na maioria dos organismos. Está contida nos cloroplastos e plastídios, sob condições de estresse hídrico, salino e baixas temperaturas (Jagendorf e Takabe 2001). Em vegetais, está presente em algas e em diversas famílias, incluindo Majeeopodiaceae, Amaranthaceae, Compositae e Malvaceae (Blunden e Gordon 1986, Rhodes e Hanson 1993, Gage e Rathinasabapathi 1999).

A glicina-betaína é sintetizada por duas vias distintas: colina ou glicina (figura) e ambas as enzimas envolvidas na sua biossíntese são estresse-induzidas, aumentando o nível de betaína (McCue e Hanson 1992, Rathinasabapathi et al. 1997). Na maioria das plantas superiores, animais e também microrganismos que sintetizam GB, a oxidação da colina em betaína aldeído é catalisada pela enzima colina mono-oxigenase, codificada pelo gene *CMO*. Por sua vez, a oxidação da betaína aldeído em GB é catalisada pela enzima betaína aldeído desidrogenase codificada pelo gene *BADH* (Rathinasabapathi et al. 1997) (Figura 5).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

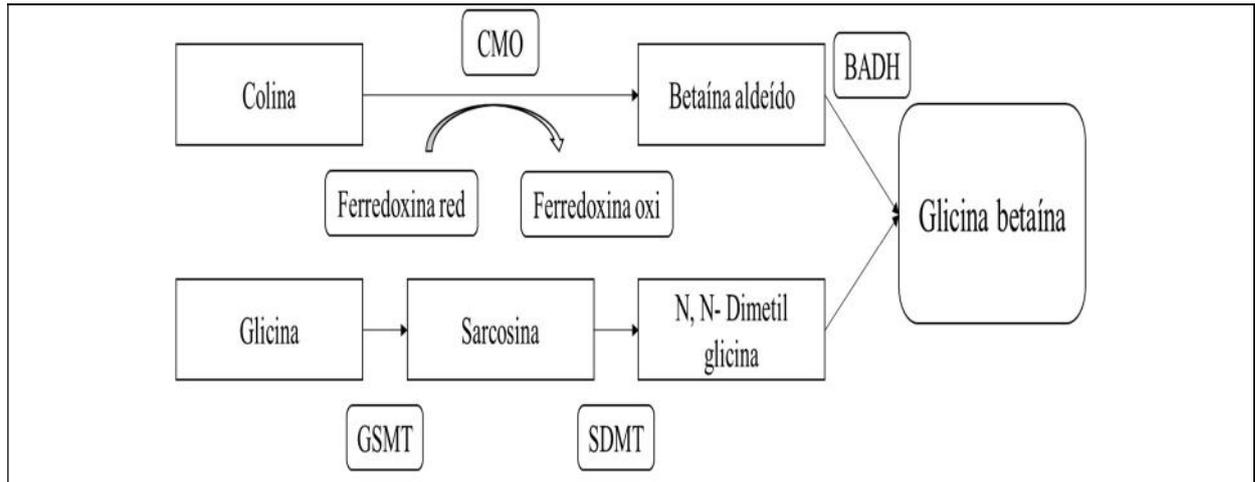


Figura 5. Via metabólica de Glicina-betaína. [Ferredoxin-dependant choline monooxygenase (CMO), NAD⁺-dependant betaine aldehyde dehydrogenase (BADH), glycine sarcosine methyltransferase (GSMT) and sarcosine dimethylglycine methyltransferase (SDMT)]. Adaptado de Sengupta et al. 2016.

Estudos demonstraram que a glicina-betaína é bastante eficaz como um soluto compatível, atuando na estabilização de enzimas e proteínas complexas além de contribuir para a manutenção do potencial osmótico no citoplasma, mantendo a integridade das membranas quando a célula está exposta a altas concentrações de sais ou temperaturas extremas (Papageorgiou e Murata 1995).

Plantas como espinafre e milho acumulam glicina betaína em suas folhas quando expostas a estresses como seca, salinidade e temperaturas baixas, de modo a sobreviver sob condições adversas (Storey et al. 1977, Nyssola et al. 2000). Genes que codificam as enzimas envolvidas na biossíntese da glicina-betaína têm sido identificados, clonados e empregados em plantas que não acumulam GB, como a *A. thaliana*, *B. napus*, *Nicotiana tabacum* e *O. sativa*, pois a superexpressão de tais genes confere resistência a estresses abióticos (McCue e Hanson 1990).

2.4.1.3 Mio-inositol

Sendo de suma importância para o crescimento e desenvolvimento dos vegetais, o mio-inositol é classificado como um álcool cíclico que participa de vários processos metabólicos, como a síntese de ácido fítico, no estoque e transporte de auxinas nas células e na produção de moléculas que estão ligadas à resposta a estresses (Loewus e Loewus 1983).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Distribuído em abundância em todo sistema biológico, o mio-inositol é sintetizado por um caminho de dois passos, denominado de "caminho Loewus". O primeiro passo é a conversão de Dglucose 6-P para D-mio-inositol(1)-Monofosfato, 1D-MI-1-P, que é catalisado por um Lmyo- inositol 1-fosfato sintase (MIPS) (Majumder et al. 1997), seguido por sua desfosforilação para libertar o mio-inositol pelo L-Myo-inositol 1-fosfato dependente de Mg⁺⁺ fosfatase (IMP) (Parthasarathy et al. 1994). Por causa do potencial do mio-inositol, já foram geradas algumas plantas transgênicas que expressam esta substância, de preferência usando a enzima MIPS ou inositol derivado de outras enzimas (Silva et al. 2011). Zhai et al. (2015). Analisando a expressão do gene *lbMIPS1* em plantas transgênicas de batata doce (*Ipomoea batatas*) verificou que a superexpressão do mesmo aumentou significativamente a resistência ao nematóide (*Ditylenchus destructor*) assim como a tolerância à seca e salinidade em condições de campo. A análise morfológica e fisiológica de plantas transgênicas de *A. thaliana* superexpressando o gene *TaMIPS2* que sintetiza mio-inositol em trigo reduziu a sensibilidade da planta ao calor (Khurana et al. 2017).

Além de osmoprotetor, o mio-inositol também possui importante papel na germinação de sementes, transporte de açúcar, nutrição mineral, metabolismo de carboidratos, estrutura de membranas, formação de parede celular e homeostase de hormônios (Styer 2000). Ma et al. (2019) demonstraram que a expressão do gene *GhMIPS1D*, que sintetiza mio inositol em algodão em plantas transgênicas de *Arabidopsis thaliana* além de melhorar fenotipicamente as plantas promoveu significativamente um alongamento das células das raízes, levando aos autores à conclusão que este gene é um regulador positivo envolvido no alongamento de células vegetais.

2.4.1.4 Trealose

A trealose é classificada como um dissacarídeo composto de duas moléculas de glicose ligadas por uma ligação do tipo alpha-alpha. Foi inicialmente descoberta no centeio em 1832, sendo posteriormente detectada em vários organismos, tais como bactérias, fungos, invertebrados e plantas (Elbein 1974). A princípio, pensava-se que a trealose era restrita a plantas revivescentes ou de ressurreição (*resurrection plants*) que conseguem sobreviver à desidratação extrema, 'revivendo' após receberem água, como por exemplo, a angiosperma *Myrothamnus lepidophylla* e a pteridófita

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Selaginella tamariscina (Goddijn e Van Dun 1999). Entretanto, experimentos posteriores provaram a existência da substância em outras plantas, como tabaco, arroz e *Arabidopsis* (Sengupta et al. 2016).

Em plantas, a biossíntese desse açúcar ocorre normalmente pela formação de trealose-6-fosfato (T6P) do UDP-glicose e glicose-6-fosfato, uma reação catalisada pela trehalose 6-fosfato sintase (TPS). Posteriormente, a T6P é desfosforilada pela fosfatase da trehalose-6-fosfato (TPP), resultando na formação de trealose livre (Wingler 2002) (Figura 6). A trealose sintetizada pode ser hidrolisada em dois monômeros pela enzima trehalase (TRE). A maioria dos genomas de plantas contêm famílias gênicas de trealose (Lunn, 2007, Avonce et al. 2010). O genoma de *A. thaliana*, por exemplo, contêm 11 genes TPS e 10 genes TPP (Lunn 2007, Avonce et al. 2010).

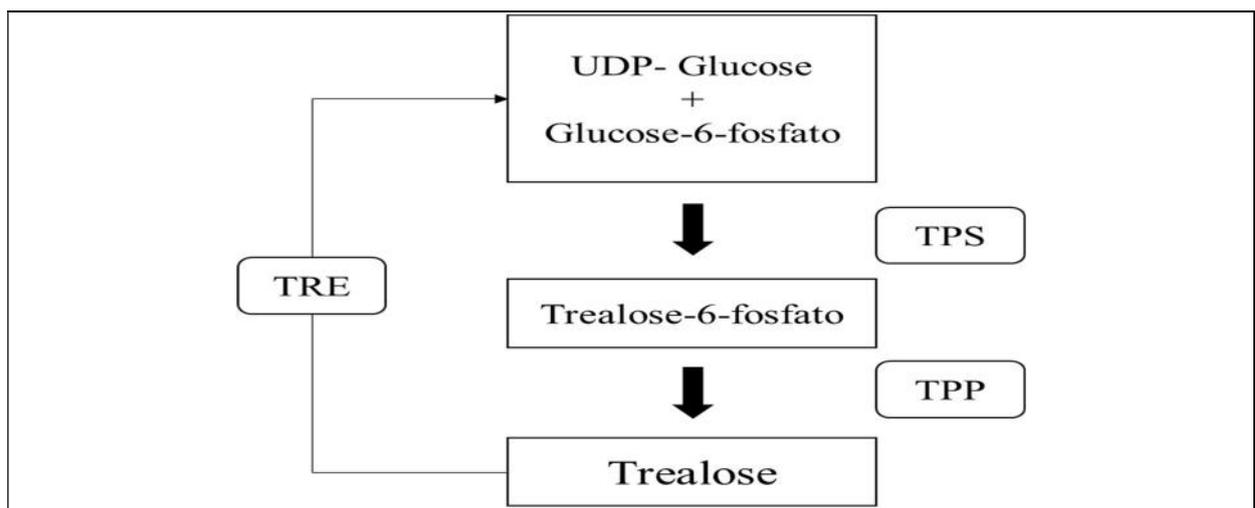


Figura 6. Via metabólica de Trealose. [Trehalose-6-phosphate synthase (TPS), trealose- 6- phosphate phosphatase (TPP), trehalase (TRE), galactinol synthase (GolS), raffinose-synthase (RafS) e stachyose-synthase (STS)]. Adaptado de Sengupta et al. (2016).

Em vegetais, a trealose atua como osmoprotetor, protegendo proteínas e membranas celulares da inativação ou desnaturação que geralmente são causadas em situações de estresse, tais como a seca, salinidade, temperaturas baixas e oxidação. Durante a situação estressante, a trealose mantém os lipídios em uma fase fluida eficientemente e com isso evita a separação de fases e fusão das membranas. Também possibilita que a célula mantenha suas funções vitais por longos períodos,

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

atuando como estabilizador de enzimas e vesículas durante desidratação (Crowe et al. 1987).

Considerando seu papel em relação a estresses abióticos, a trealose vem sendo amplamente estudada. Garg et al. (2002) demonstraram que a superprodução do açúcar trealose tem potencial para aumentar a tolerância a estresses abióticos, como salinidade, seca e baixas temperaturas, em plantas transgênicas de arroz, comparadas aos controles não transgênicos. Introduzindo um gene que sintetiza trealose de *Grifola frondosa* em plantas de tabaco submetidas a déficit hídrico e estresse salino, Zhang et al. (2005) observaram que as plantas transgênicas foram capazes de acumular altos níveis de trealose resultando em um incremento na tolerância à seca e à salinidade, melhorando conseqüentemente o aspecto fisiológico das plantas. Nuccio et al. (2015) superexpressaram um gene codificando trealose-6-fosfato fosfatase (TPP) em arroz em inflorescências de milho em desenvolvimento e em condições de campo, usando um promotor floral. Desta forma, verificaram que a superexpressão do gene incrementou uma melhora de cerca de 9% para 49% em condições de déficit hídrico leve e de 31% para 123% em déficit hídrico severo nas condições das plantas transgênicas em relação as plantas que não tiveram o gene introduzido, visto que a trealose regula o crescimento e desenvolvimento das plantas.

2.5. Prospecção e análise de genes na era do sequenciamento de nova geração

2.5.1 Método Sanger e Sequenciamento de nova geração

Em 1977, Sanger e colaboradores descreveram o primeiro método de sequenciamento de DNA, que consistia em realizar a síntese de uma fita de DNA marcada, complementar à fita molde na qual se desejava realizar o sequenciamento. Para a desnaturação, o DNA é aquecido, quebrando as pontes de hidrogênio que ligam suas bases e – uma vez que a dupla fita dupla tenha sido separada - uma fita simples de DNA é obtida. Através das fitas simples, o oligonucleotídeo sintético, denominado de *primer*, se liga a uma região de homologia em uma das fitas e então o alongamento é realizado, onde a enzima DNA polimerase adiciona desoxirribonucleotídeos (dNTPs) sequencialmente na extremidade 3' do iniciador (*primer*), complementar ao DNA molde (Shendure e Ji 2008). Nesta reação também é adicionada uma pequena quantidade de dideoxinucleotídeos (ddNTPs) marcados

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

com uma molécula fluorescente, que não possuem o grupo OH livre na extremidade 3', e dessa forma, a DNA polimerase não consegue prosseguir com a adição do próximo nucleotídeo, realizando assim uma marcação do último nucleotídeo adicionado ao fragmento. Por fim, os fragmentos são separados de acordo com o seu tamanho por eletroforese em gel de poliacrilamida, onde é possível visualizar a separação de moléculas de DNA que diferem em apenas um nucleotídeo, apresentando assim, a resolução necessária para o sequenciamento do DNA (Sanger 1988, Huang et al. 1992) onde a marcação fluorescente confere uma cor diferente para cada tipo de didesoxinucleotídeos (Carraro e Kitajima 2002) (Figura 7).

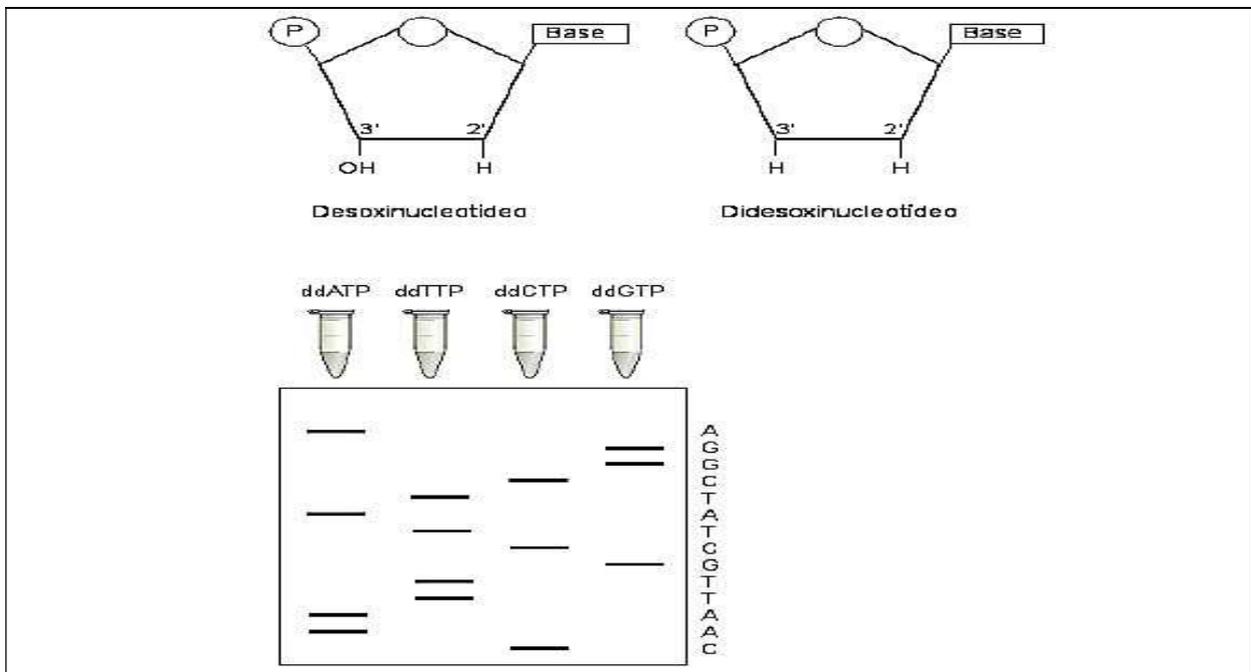


Figura 7. Representação do sequenciamento pelo método de Sanger, onde em cada tubo é adicionado um tipo de didesoxinucleotídeo marcado junto com os demais desoxinucleotídeos e após vários ciclos de reação de amplificação, os produtos dos tubos são submetidos à eletroforese, permitindo a leitura das sequências. Fonte: Giusti, Kettener e Fuchs-Ferraz (2016).

Anos após Sanger desenvolver sua técnica, por volta de 1986, o primeiro sequenciador automático de DNA foi lançado, denominado de ABI 370, seguido pelo primeiro sequenciador de eletroforese capilar, o ABI 3100, lançado em 1998. Através do aprimoramento desses sequenciadores automáticos, em 2005 – com o intuito de acelerar e diminuir os custos do processo – foram desenvolvidos novos métodos de sequenciamento, denominados de tecnologias de nova geração (Fernandes 2015).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Os protocolos do NGS (*Next Generation Sequencing*) iniciam-se com a construção de uma biblioteca, onde ocorre a fragmentação do DNA inicial e a sua ligação a adaptadores (sequências pequenas com bases conhecidas) nas extremidades 3' e 5'. Estes adaptadores são utilizados como *primers* (Goodwin et al. 2016) na amplificação através de PCR (*Polymerase Chain Reaction*) produzindo *clusters* de sequências idênticas. Além disso, podem ser adicionadas sequências curtas chamadas de '*barcodes*' que permitirão o sequenciamento de múltiplas amostras ao mesmo tempo. O sequenciamento é realizado através de diversos ciclos de incorporação, onde as sequências dos *clusters* são adquiridas por meio de *softwares* plataforma-específicos, através de um processo definido como '*base calling*' que converte em bases o sinal adquirido em cada etapa de incorporação (Pfeifer 2017).

Dentre as plataformas de sequenciamento mais utilizadas, a que mais se destaca atualmente é a Illumina, devido ao grande volume de sequências obtidas, à sua qualidade, espectro de utilidade e também por possuir o menor custo por base sequenciada (Glenn 2011). Similarmente ao sequenciamento Sanger, o Illumina é realizado por síntese usando a DNA polimerase e nucleotídeos terminadores marcados. O DNA a ser sequenciado é fragmentado, por meio de um processo de nebulização (a nebulização é uma forma mecânica de fragmentar o DNA, onde a quebra é realizada através da ação do gás nitrogênio), onde uma seleção dos fragmentos de tamanho apropriados é realizada e posteriormente, ligam-se adaptadores em ambas as extremidades. Tais fragmentos são postos em uma lâmina (*flowcell*) que possuem adaptadores complementares aos que estão contidos nos fragmentos, para que estes possam se ligar à placa. Posteriormente, ocorre a incorporação de nucleotídeos não marcados com fluorescência, até que toda a amplificação do fragmento esteja realizada, produzindo *clusters*, que irão conter mais de um milhão de cópias do mesmo fragmento. Após a amplificação, as imagens de cada ciclo são decodificadas, determinando-se a sequência de bases de cada *cluster* presente na *flowcell* (Fietto e Maciel 2016).

O NGS possibilita inúmeras aplicações em áreas como: expressão gênica, biologia do RNA, regulação do genoma, sequenciamento do genoma, organização do genoma, transcrição, replicação, tradução e outras (Reuter et al. 2015), podendo ser utilizado para: (a) construir um novo genoma através de um organismo desconhecido; (b) avaliar a variação genética de um organismo utilizando um genoma de referência; (c) observar a transcrição de um organismo ou célula a partir do cDNA; (d) estudar

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

mecanismos regulatórios e epigenomas; (e) analisar a diversidade de microrganismos em amostras ambientais (Park e Kim 2016).

2.5.2 Ômicas

As abordagens multiômicas podem ser utilizadas para decifrar mudanças em níveis de genômica, proteômica ou metabolômica nas plantas (Figura 8) e em outros organismos. Juntamente com os avanços em bioinformática, os dados científicos das multiômicas têm melhorado o conhecimento e entendimento das respostas das plantas frente a estresses (Meena et al. 2017).

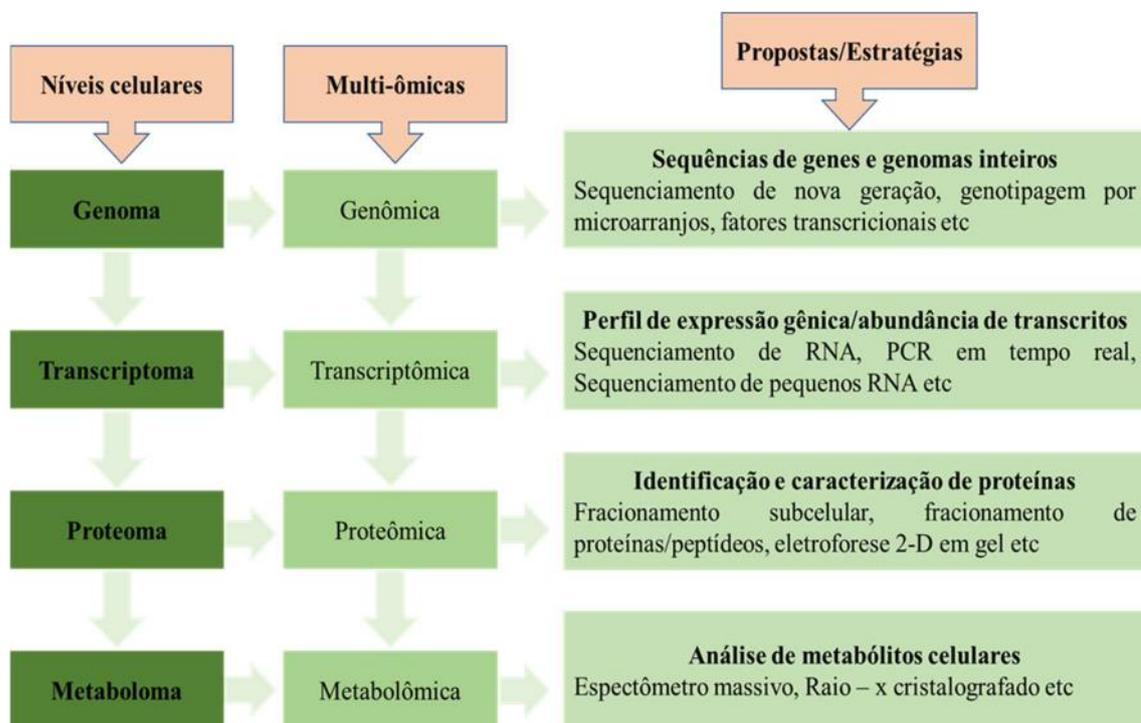


Figura 8. Componentes a nível celular, abordagens multiômicas e estratégias para identificar o resultado do impacto de estresses abióticos e/ou interações com microrganismos. Adaptado de Meena et al. (2017).

O DNA é o ponto de partida das evidências moleculares ligadas à resistência ou tolerância a estresses, sendo que os vegetais contêm em seu genoma (genômica) vários genes atuantes em situações desafiantes. O transcriptoma pode ser definido como o conjunto de transcritos (moléculas de RNA) de uma célula, tecido ou organismo e, diferentemente do genoma, varia segundo o tecido e função celular e do estágio de desenvolvimento do organismo, sendo afetado por fatores externos

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

(bióticos ou abióticos) e internos (como mudanças bioquímicas e fisiológicas) (Wang et al. 2009).

Em situações adversas, como as mudanças ambientais, quando transcritos os genes produzem RNAs mensageiros (mRNAs) que são traduzidos em proteínas funcionais, parte delas responsiva a estresses, produzindo o transcriptoma (transcriptômica) e o proteoma (proteômica). A integração da genômica, transcriptômica e proteômica acordam o estudo completo de um dado organismo, através do entendimento dos mecanismos moleculares que permitem que o mesmo se desenvolva e se adapte a diversas condições (Ferreira-Neto et al. 2018). Sabe-se que o vegetal emprega uma regulação pós-transcricional da expressão gênica por RNAs ou microRNAs codificadores não proteicos (miRNAs) (Chinnusamy et al. 2007). Estes microRNAs têm grande importância no estudo da resistência e vários deles já foram identificados em vegetais, estabelecendo seu papel na acumulação de osmólitos (osmoprotetores) e na resposta à falta de nutrientes (Hasanuzzaman et al. 2013). Mi RNAs regulam a expressão gênica direcionando a clivagem ou repressão translacional de mRNAs complementares ou induzindo o silenciamento transcricional de genes alvo. MiRNAs induzidos por estresses abióticos regulam negativamente genes alvo que provavelmente iriam codificar reguladores negativos de resposta ao estresse, por outro lado, a baixa regulação desses miRNAs resultam no acúmulo de mRNAs alvo que codificam reguladores positivos da tolerância aos estresses (Chinnusamy et al. 2007).

Além disso, proteomas estão interligados em vários processos bioquímicos produzindo vários metabólitos (metaboloma). Redes metabólicas de solutos compatíveis (como a prolina, glicina-betaína, glicose, frutose, galactinol, rafinose, entre outras) possuem papéis importantes acerca da tolerância ao estresse térmico. Por exemplo, a proteína de ligação ao RNA (*ATGRP7*) aumenta em resposta à baixa temperatura e diminui sob alta temperatura, como demonstrado em experimento com *A. thaliana*, analisando temperaturas de 4, 20 e 32°C. Sua abundância foi significativamente correlacionada com a concentração de prolina e glutamina (Wienkoop et al. 2008). Nesse sentido, as ômicas possibilitam a descoberta de genes cruciais na cascata de respostas a estresses de modo que possam contribuir para a aquisição de tolerância (Vinocur e Altman 2005).

Várias tecnologias visando ao entendimento e quantificação do transcriptoma têm sido desenvolvidas, as quais estão fundamentadas principalmente na hibridização

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

e no sequenciamento de cDNA, como as tecnologias de microarranjos e de RNA-Seq. O microarranjo é uma ferramenta para análise que gera perfis de expressão de um grande número de genes que podem ser identificados como induzidos ou reprimidos em determinada situação, incluindo estresses, como o térmico, por exemplo. Em um estudo com *N. tabacum*, Rizhsky et al. (2002) examinaram 170 amostras de cDNA de plantas de tabaco (*N. tabacum*) submetidas a estresse hídrico (70 % do teor de água) ou térmico (alta temperatura – 38 °C) e verificaram que muitos genes não foram expressos apenas nos estresses individuais, mas sim na combinação de ambos os estresses. Posteriormente, em outro estudo, Rizhsky et al. (2004) utilizaram o transcriptoma extraído de tecido foliar de *Arabidopsis* para analisar alterações na regulação de transcritos em resposta aos mesmos estresses e na combinação de ambos e observaram 1075 transcritos induzidos na condição de déficit hídrico, 262 na condição de estresse térmico (calor) e 1057 na combinação das duas, concluindo um aumento em 262 % de transcritos relacionados à resposta à alta temperatura e assim como no estudo em tabaco, muitos transcritos foram aumentados apenas na situação dos estresses combinados.

Conforme revisado por Silva-Júnior et al. (2018), as técnicas de microarranjo ganharam recentemente um novo impulso por sua capacidade de detecção de mutações no genoma de um ou vários indivíduos, avaliando alterações de sítio único (SNPs, *Single Nucleotide Polymorphisms*) ou variações no número de cópias de alguns genes (CNVs, *Copy Number Variation*), com aplicações no diagnóstico de doenças hereditárias ou adquiridas. Porém, conforme indicado pelos mesmos autores, as desvantagens mais citadas incluem o alto custo de um único experimento, a necessidade de um grande número sondas de DNA e a falta de controle sobre o conjunto de transcritos analisados. Além disso, o uso preferencial de sondas curtas (embora em grande número) diminui a precisão e a especificidade relativa das análises de expressão diferencial usando microarranjos, demandando uma validação dos resultados de expressão com outras técnicas moleculares (Silva-Júnior et al. 2018).

Também denominado de *Whole Transcriptome Shotgun Sequencing*, o RNA-Seq é uma das tecnologias de transcriptômica que envolve Sequenciamento de Nova Geração (NGS), sequenciando os RNAm convertidos em cDNA (DNA complementar) com o objetivo de detectar informações do transcriptoma de um organismo. Ao contrário da técnica de microarranjos, que possui a limitação de avaliar apenas

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

sequências de genes conhecidos, o RNA-Seq não precisa de conhecimento prévio dos genes que se deseja detectar, qualquer transcrito que esteja sendo expresso no determinado momento e condição, poderá ser detectado, além de também detectar eventos de *splicing* (Varuzza 2013; Ferreira-Neto et al. 2018). RNA-Seq sendo uma técnica bem moderna e eficiente, consegue sequenciar todos os todos os fragmentos de RNAs (com cauda poli A ou não) de centenas de pares de bases aleatoriamente (Ferreira-Neto et al. 2018). Na técnica, os RNAm são isolados e convertidos em cDNA. Posteriormente, são conectados a adaptadores nas duas extremidades (5' e 3') e submetidos à amplificação por PCR. A seguir, as moléculas são sequenciadas a partir de ambas as extremidades (*paired-end*) ou em apenas uma só (*single-end*) onde o tamanho dos fragmentos sequenciados (*reads*) pode variar em tamanho de 35 a 400 pb a depender da plataforma utilizada. Por fim, os *reads* que resultaram do sequenciamento são alinhados entre si ou com um genoma de referência, onde irá permitir o mapeamento da estrutura e análise da expressão gênica ou podem ser utilizados na montagem *de novo* (do inglês, *de novo assembly*) do transcriptoma (Holt e Jones 2008) (Figura 9).

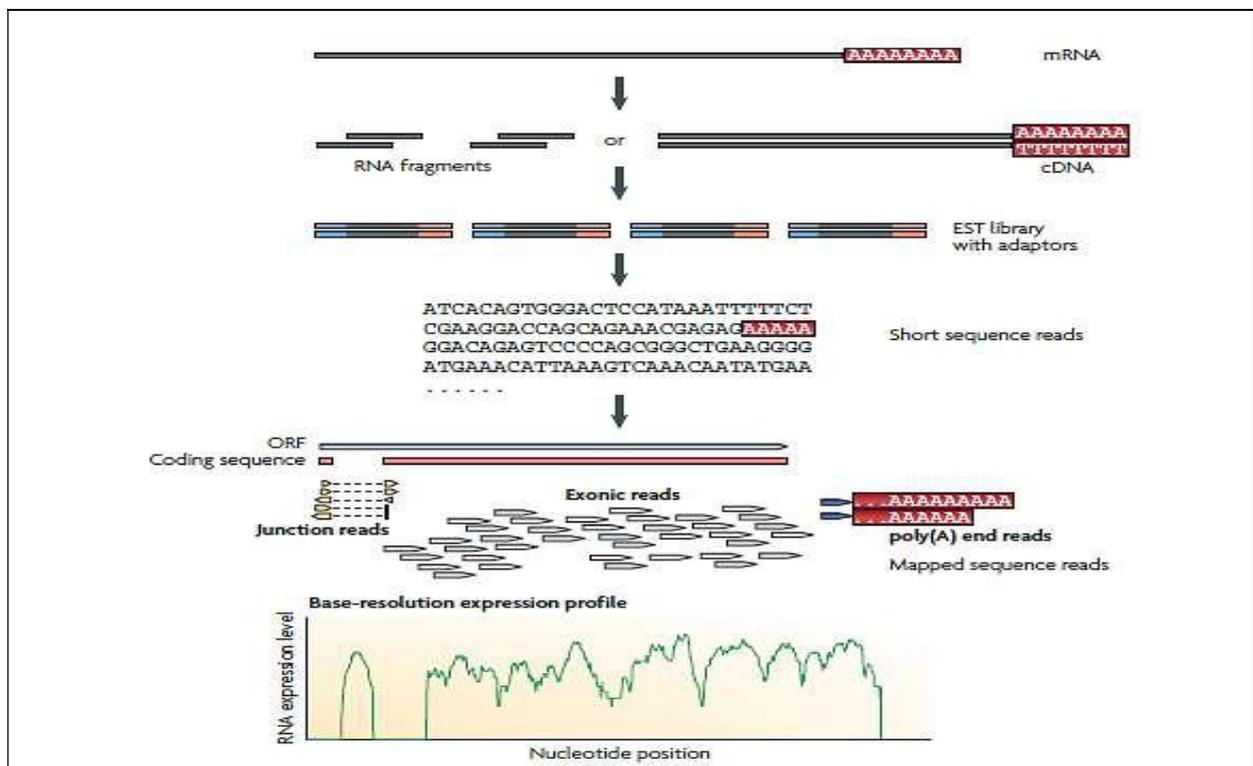


Figura 9. Representação esquemática da técnica de RNA-Seq. Primeiramente, o RNAm é capturado ou separado dos RNAs estruturais, sendo convertido para cDNA por uma transcriptase reversa. A seguir ocorre uma fragmentação, sendo adicionados

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

adaptadores (em azul) em ambas as extremidades. Os adaptadores servem como primers no sequenciamento de alto rendimento. Os *reads* resultantes são alinhados a um genoma de referência ou transcriptoma, podendo ser classificados em três tipos: exonic reads (reads de saída), junction reads (reads de junção) e poly(A) end-reads (reads de terminação poli A) nos quais são usados para gerar uma base de dados de perfil de expressão para cada um dos genes. Fonte: Wang, Gerstein e Snyder (2009).

O processamento dos dados gerados pela técnica de RNA-Seq envolve várias técnicas computacionais que estimam e comparam os transcritos de diferentes amostras biológicas em um determinado momento e condição. Assim como em qualquer outra tecnologia de sequenciamento, o RNA-Seq lida com vários desafios computacionais, como o desenvolvimento de métodos eficientes para armazenar, recuperar e processar uma grande quantidade de dados, que devem ser superados para que erros na análise de imagens e na remoção de *reads* de baixa qualidade possam ser reduzidos (Wang, Gerstein e Snyder 2009; Ferreira-Neto et al. 2018).

A análise molecular de características importantes, como tolerância a estresses, é de extrema importância para o desenvolvimento de estratégias que possam combater essas adversidades. Nesse contexto, ferramentas de transcriptômica se destacam uma vez que os RNAs codificam as proteínas e enzimas que participam diretamente destes processos bem como na regulação de vias metabólicas ligadas a eles. Com isso, os estudos nessa área permitem além da descoberta de novos genes, a correção através de engenharia genética de eventuais “defeitos” em vias metabólicas importantes o que irá resultar em uma melhor apresentação de uma característica do organismo (Ferreira-Neto et al. 2018).

2.6 Bioinformática

A bioinformática pode ser caracterizada como uma ciência multidisciplinar que engloba a ciência da computação, estatística, matemática e a tecnologia de informação na análise de dados biológicos. No final da década de 1950 e início dos anos 1960 os pesquisadores começaram a utilizar computadores nas suas pesquisas acadêmicas. Em 1957, a IBM Corporation produziu a primeira linguagem de programação, denominada Fortran, que era de fácil aprendizado, sendo melhor adequada em aplicações científicas (Ledley 1959). Porém, foi só no início da década de 1960 que os computadores começaram a fornecer verdadeiras contribuições, com

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Margaret Oakley Dayhoff, cujo trabalho era formular programas Fortran com o objetivo de determinar sequências de aminoácidos (Dayhoff 1965).

Com o surgimento e a disseminação da internet, nas décadas de 1980 e 1990, a biologia computacional foi crescendo. Novos métodos computacionais favoreceram a criação de bancos de dados de nucleotídeos e proteínas que possibilitam análises de similaridade, entre essas sequências, permitindo inferências estruturais e funcionais (Kanehisa e Bork 2003). Entre os bancos de dados públicos que armazenam sequências de nucleotídeos e proteínas, destacam-se o GenBank do NCBI (*National Center for Biotechnology Information*; <http://www.ncbi.nlm.nih.gov/>) e o UNIPROT (*Universal Protein Resource*; <http://www.uniprot.org>) que disponibiliza um catálogo de sequências proteicas curadas com sua anotação funcional, integrando, interpretando e catalogando dados de várias fontes, disponibilizando assim, uma ampla base de sequências proteicas de livre acesso, ricamente anotadas e curadas (The Uniprot Consortium 2012). A figura 10 apresenta diversos bancos de dados disponíveis atualmente agrupados de acordo com a sua aplicação biológica, indicando ferramentas disponíveis online ou para download (Krallinger e Valencia 2005).

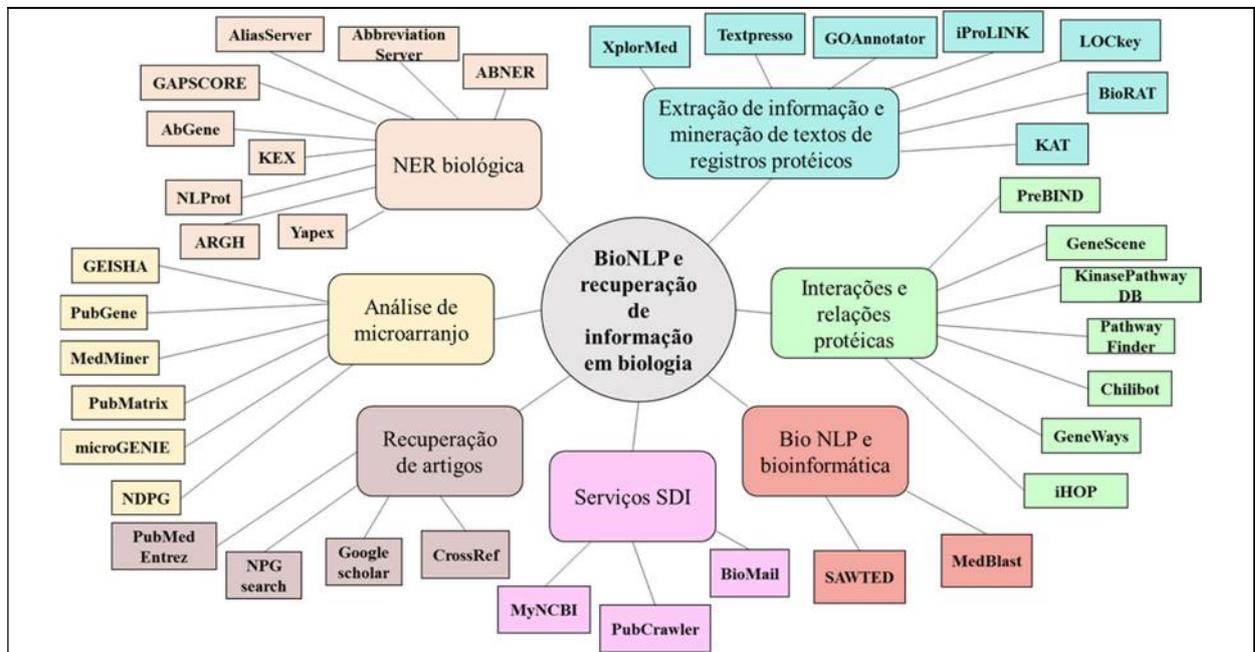


Figura 10. Uma visão do processamento da linguagem natural biológica (BioNLP) e aplicações de *text mining* na biologia. O tópico central é envolvido por sete retângulos com cantos arredondados com suas aplicações correspondentes, dadas por outros retângulos mais externos com ferramentas de anotação estão disponíveis online ou

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

para download: NER (nome da identidade reconhecida), SDI (informação seletiva disseminada). Adaptado de Krallinger e Valencia (2005).

Paralelamente à criação dos bancos de dados, outras ferramentas possibilitam o acesso a análises *in silico* de sequências desconhecidas e novos genes, gerando um alinhamento comparativo com sequências já existentes na literatura, além de detectar a regulação da expressão de diversos genes que fazem parte de vias metabólicas ou processos específicos (Kent et al. 2002, Saeys et al. 2007).

A mineração de texto (*text mining*) é uma ferramenta de obtenção de dados não estruturados a partir de um banco de dados estruturado utilizando algoritmos de análise de dados envolvendo três subáreas: recuperação, extração e questões e respostas da informação, assim permitindo a identificação de identidades biológicas e suas interações (Espíndola et al. 2010). A relação entre genes e proteínas é uma importante área da mineração de texto, pois a combinação de métodos estatísticos com a aprendizagem de máquina em capturar as interações dessa relação disponível na literatura cria sistemas utilizáveis de alta precisão (Rhee et al. 2006). Portanto, a combinação de múltiplas plataformas integrando resultados é eficaz no esclarecimento dos sistemas moleculares, havendo a possibilidade de que se desenvolvam estratégias para o melhoramento de plantas cultivadas, aumentando a produtividade (Edwards e Batley 2004, Mochida e Shinozaki 2010).

O grande volume de dados depositados diariamente em bancos genômicos, torna a biologia cada vez mais dependente das ferramentas computacionais. Tais ferramentas possibilitaram diversos estudos, como por exemplo, o alinhamento e a montagem *de novo* de sequências, a detecção de polimorfismos e de outras estruturas em genomas, a criação de marcadores moleculares e análises de transcriptomas, entre outras (Giachetto e Higa 2015), tornando a bioinformática uma ferramenta indispensável para a biologia molecular (Petrey e Honig 2014).

2.7 Família Fabaceae

2.7.1 Classificação, origem, distribuição e importância agronômica

Fabaceae Lindl. compreende uma das maiores famílias das Angiospermas, a qual compreende cerca de 727 gêneros e 19.327 espécies (LWGP 2017).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Anteriormente, essa família era classificada em três subfamílias distintas, Papilionoideae ou Faboideae, Caesalpinioideae e Mimosoideae (Cronquist 1981), no entanto, após estudos filogenéticos recentes, a família passou por uma reclassificação onde agora compreende seis subfamílias: Duparquetioideae, Cercidoideae, Detarioideae, Dialioideae, Caesalpinioideae e Papilionoideae, sendo a Papilionoideae a subfamília que abrange o maior número de gêneros e espécies (LWGP 2017).

A família Fabaceae possui distribuição cosmopolita, sendo seus registros fósseis datados de mais de 60 milhões de anos, encontrados na Europa, América do Norte, Ásia e África (Lavins et al. 2004). Acredita-se que tal distribuição se deva ao fato da sua habilidade em fixar nitrogênio atmosférico por meio de associação simbiótica com bactérias, característica peculiar desta família (Lavins et al. 2004, Amorim 2014) As plantas dessa família são encontradas na maioria dos biomas existentes no planeta, desde florestas secas, savanas (Dryflor 2016, APG IV 2016), desertos, regiões temperadas e também altas altitudes (LPWG 2017). No Brasil, estão registrados 222 gêneros, sendo 16 deles endêmicos, com 2.837 espécies, das quais 1.535 são endêmicas. A região Nordeste, possui 170 gêneros e 1.108 espécies, sendo 128 gêneros e 606 espécies endêmicas da Caatinga (BFG 2015), confirmando a importância do gênero neste bioma.

Fabaceae é a segunda família de maior importância econômica mundial – sendo superada pela família Poaceae – visto que algumas espécies são apontadas como importantes fontes de proteína vegetal, destacando-se principalmente na produção de grãos, de forragem e na adubação verde, entre outros (Judd et al. 1999). (Lewis et al. 2005). Seus frutos e sementes são utilizados para a produção de corantes, óleos, perfumes, inseticidas, na alimentação humana, animal e como forrageira. Além disso, possui uso medicinal, ornamental, na produção de madeiras nobres (consideradas as mais valiosas do mundo), além de sua grande importância agrônômica, devido ao enriquecimento de solos com nitrogênio, pela sua capacidade de fixação biológica (Lima et al. 1994).

Entre as espécies com maior relevância alimentícia (produção de grãos), destacam-se as dos gêneros *Arachis* L. (amendoim), *Cajanus* Adans. (feijão-guandu), *Glycine* L. (soja), *Cicer* L. (grão-de bico), *Pisum* L. (ervilha), *Phaseolus* L. (feijão) e *Vigna* Savi (feijão-caupi), entre outros (Judd et al. 2009, Sousa et al. 2011). Os gêneros que compreendem a maioria das espécies de Fabaceae utilizadas na alimentação animal como pastagens são: *Stylosanthes*, *Arachis*, *Centrosema*,

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Macroptilium, *Desmodium*, *Leucaena*, *Zornia* e *Calopogonium* (Pereira et al. 2001, Borges do Valle 2002).

2.7.2. *Stylosanthes scabra*

2.7.2.1 Classificação e distribuição geográfica

Em 1788, Swartz descreveu o gênero *Stylosanthes* com duas espécies, *S. procumbens* Swartz (= *S. hamata* (L.) Taubert) e *S. viscosa* Swartz. Em 1838, Vogel, estudando o gênero, o dividiu em duas seções, *Eustylosanthes* e *Styposanthes* (Brandão e Costa 1982). Mohlenbrock (1963) revisando o gênero reconheceu 25 espécies e as classificou em dois grupos: *Stylosanthes* (14 spp.) e *Astyposanthes* (11 spp.) e, posteriormente, acrescentou cinco novos táxons, aumento o total de espécies para 30. Dentre essas, nove espécies de *Stylosanthes* possuem importância econômica, *Stylosanthes capitata* Vogel., *S. fruticosa* (Retz.) Alston, *S. guianensis* (Aubl.) Sw., *S. hamata* (L.) Taub., *S. humilis* Kunth, *S. macrocephala*, *S. scabra* Vogel, *S. seabrana* e *S. viscosa* (L.) Sw. (Calles e Schultze-Kraft 2016, Vanni e Fernandes 2011).

A distribuição do gênero *Stylosanthes* é bem ampla, podendo ser encontrada naturalmente desde a América Central até a América do Sul, tendo como centro de origem o Brasil (Willians et al. 1984). Foram registradas espécies do gênero na Venezuela, Guiana Francesa, Suriname e Paraguai, sendo 43 espécies exclusivas do continente americano. No Brasil, sua ocorrência é registrada em todos os estados, possuindo 29 espécies, sendo 13 endêmicas, entre essas a *S. scabra*, nativa do semiárido brasileiro (Ferreira e Costa 1979, Stace e Edye 1984).

2.7.2.2 Importância econômica

A utilização de leguminosas em consórcio com gramíneas melhora a diversificação da dieta do animal e aumenta a massa de forragem devido à capacidade de fixação de nitrogênio no solo exercida pelas leguminosas, beneficiando a gramínea (Vitor et al. 2008). O nitrogênio é um composto altamente importante para os vegetais, pois aumenta o crescimento e a massa foliar e melhora sua qualidade, devido ao aumento no teor proteico, levando conseqüentemente ao aumento da produção de

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

biomassa (Barbero et al. 2010) com diminuição considerável na necessidade de adubação desse elemento para a gramínea (Almeida et al. 2003). O nitrogênio atmosférico é fixado pela leguminosa através da simbiose com bactérias do gênero *Rhizobium*, sendo disponibilizado no solo pelo desprendimento dos nódulos, com possibilidade de ser utilizado pela gramínea, aumentando sua produção (Barbero et al. 2009). Segundo dados da EMBRAPA (2007), a utilização da variedade Estilosantes Campo Grande (*S. macrocephala* x *S. capitata*) ocupa de 20 % a 40 % da área total na produção consorciada em solos arenosos e com baixa fertilidade, sendo que a fixação biológica aumentou os valores de nitrogênio, que variavam de 60 kg/ha/ano a 80 kg/ha/ano.

A espécie *S. scabra*, que apresenta tolerância à seca e à baixa fertilidade do solo, apresentando importante potencial forrageiro (Stace e Edye 1984), sendo bem difundido em pastagens de regiões de clima temperado, a exemplo da Austrália, primeiro país a reconhecer o potencial das espécies do gênero como forrageiras (Edye 1997). Por sua vez, a espécie *S. humilis*, que foi introduzida acidentalmente, se adaptou e se expandiu nas savanas tropicais da Austrália no início do século 20. Na década de 70, na região semiárida, a espécie ocupava cerca de dezenas de milhares de pastagens cultivadas. No entanto, com a chegada do fungo *Colletotrichum gloeosporioides*, causador da antracnose, grandes áreas cultivadas com *S. humilis* foram devastadas. Dessa forma, o país ampliou as pesquisas com outras espécies (como a *S. scabra*), e desenvolveu diversas cultivares, permitindo que o gênero permanecesse em seu sistema de produção (Miller et al. 1997). Uma dessas cultivares é a cv. Seca (Oran 1990), desenvolvida através de acessos de *S. scabra* coletados em Pernambuco – Brasil, que é resistente à seca e à antracnose, sendo bem difundida e utilizada na Austrália.

2.7.2.3 Bancos de germoplasma

Apesar de ter uma grande ocorrência no Brasil, sendo registrada em 22 estados brasileiros, a maioria do Nordeste – com 300 ocorrências na Bahia e 181 ocorrências em Pernambuco (CRIA, 2018) – o potencial de uso de *S. scabra* ainda é pouco explorado (Lira 2015). Entretanto, na Austrália há cultivares comerciais derivadas de acessos de germoplasma coletados no semiárido brasileiro (Stace e Edye 1984).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Em 2016, foi criado o Banco Ativo de Germoplasma de *Stylosanthes* spp., localizado na Universidade Estadual de Feira de Santana (Bahia), com o objetivo de auxiliar programas de melhoramento genético de espécies do gênero no desenvolvimento de cultivares adaptadas às condições de semiárido brasileiro. Nesse banco, já existem 25 espécies catalogadas, sendo a *S. scabra* a que apresenta o terceiro maior número de exsicatas (101) e o maior número de locais de ocorrência (Oliveira e Queiroz 2016).

2.7.2.4 Melhoramento genético

O melhoramento de leguminosas para forragem nos países tropicais, apesar de recente, tem sido fortalecido nas últimas décadas (Valle et al. 2009) com destaque para *Stylosanthes* spp., a leucena (*Leucaena* spp.) e o amendoim forrageiro (*Arachis pintoï*) (Herling e Pereira 2016). Os objetivos dos programas de melhoramento de leguminosas forrageiras concentram-se principalmente no aumento da produtividade, em melhorias da persistência no pasto, produção maior de sementes e resistência a doenças (Vieira et al. 2007). Costa (2006) destacou que as espécies do gênero *Stylosanthes* por demonstrarem uma grande diversidade na região de semi-árido brasileiro, boas características bromatológicas além da tolerância à solos degradados e secos, são ótimas candidatas a programas de melhoramento genético. Para *Stylosanthes* spp. as metas incluem uma maior produção de sementes, persistência, produção de matéria seca (Valle et al. 2009; Oliveira et al. 2016) resistência à antracnose, principal doença que acomete a cultura (Falco et al. 2016), além de melhor adaptação a estresses abióticos, como a seca e tolerância ao alumínio, devido às mudanças climáticas recentes (Jank et al. 2011). Uma vez que predomina a autogamia na maioria das espécies do gênero, a coleta de genótipos de ocorrência natural e a obtenção de linhas puras por meio de sucessivas autofecundações ainda é o método de melhoramento mais utilizado para *Stylosanthes* (Euclides et al. 2013).

Os programas de melhoramento genético de *Stylosanthes* começaram na década de 60 (Andrade e Karia 2000). Esse gênero detém o maior número de cultivares forrageiras tropicais lançadas no mundo, com cerca de 29 cultivares liberadas decorrentes de trabalhos de avaliação e seleção realizados na Austrália (Karia et al. 2002). Através de marcadores morfológicos e agrônômicos foi observada uma grande variabilidade intraespecífica em 72 acessos de *S. guianensis*, 44 de *S.*

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

capitata e 26 de *S. scabra*, o que ressalta a importância de bancos de germoplasma no auxílio dos programas de melhoramento (Karia et al. 2002), Esses bancos, além de gerarem informações sobre características morfológicas, fisiológicas e agrônômicas, também fornecem dados acerca de tolerância a estresses nos acessos de cada espécie, contribuindo para a seleção de acessos mais adequados em programas de melhoramento (Idris e Saad 2001).

No Brasil, os programas de melhoramento do gênero *Stylosanthes* se restringem às ações da EMBRAPA, mais precisamente com as espécies *S. capitata*, *S. guianensis* e *S. macrocephala*, visando a produção de sementes e matéria seca e resistência à antracnose (Falco et al. 2016, Simeão et al. 2006), A cultivar mais importante no Brasil, a Estilosantes Campo Grande, foi lançada em 2000, sendo produto do cruzamento das espécies *S. capitata*, e *S. macrocephala* (Karia et al. 2002). No Nordeste, já foram realizadas coletas em expedições para regiões semiáridas da Bahia e de Pernambuco, onde há uma grande ocorrência das espécies do gênero (Oliveira 2015). Oliveira et al. (2016) destacou as espécies *S. scabra*, *S. humilis*, *S. capitata* e *S. viscosa* são as mais promissoras para a produção de forragem, observando a superioridade dos acessos pertencentes à espécie *S. scabra*.

2.7.2.5 Análises genéticas e moleculares

Em *Stylosanthes*, os poucos dados referentes à citogenética relatam a existência de acessos diploides, tetraploides e hexaploides em várias espécies (Cameron 1967, Goldblatt 1981, Battistin 1981, Vieira 1988). *S. scabra* é um alotetraploide com as espécies *S. viscosa* e *S. sp. aff. S. scabra* como seus progenitores diploides (Liu e Musial 1997). Lira (2015) analisando 17 acessos identificados inicialmente como *S. scabra*, coletados no semiárido brasileiro e mantidos no Banco Ativo de Germoplasma da Embrapa Cerrados (Planaltina-DF), observaram que três acessos apresentaram cariótipos com $2n=20$ e 14 com $2n=40$, demonstrando a existência de acessos pertencentes a espécie *S. seabrana* ($2n=20$), segundo Date (2010) e *S. scabra* ($2n=40$), de acordo com os trabalhos de Cameron (1967) e Liu e Musil (1997). Em outro estudo, Liu (1997) demonstrou que a variação genética entre os acessos de *S. scabra* coletados no Brasil, era bem menor que a dos acessos coletados na Colômbia e Venezuela. Adicionalmente, o autor verificou que a maioria dos acessos presentes no estudo foram coletados nas regiões em que se

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

originaram, uma vez que os acessos da Colômbia foram agrupados ao *cluster* de dois grupos brasileiros, sugerindo que a variação genética entre as coleções brasileiras não foi contínua e que os diferentes grupos de *S. scabra* podem ter evoluído de várias plantas tetraploides formadas localmente por eventos de hibridização entre os dois diploides progenitores, (*S. viscosa* e *S. sp. aff. S. scabra*).

Em estudo posterior, Liu et al. (1999) propuseram que os progenitores do alotetraploide AABB *S. scabra* evoluíram de dois genomas distintos das espécies ‘*S. seabrana*’/ *S. hamata* (A) e *S. viscosa* (B), hipótese posteriormente confirmada por Marques et al. (2018). Neste estudo, Marques et al. (2018) utilizaram as técnicas de hibridização *in situ* fluorescente (FISH) e hibridização genômica *in situ* (GISH) e sequenciamento Illumina de DNA genômico de *S. scabra*, *S. hamata* e *S. viscosa*, visando montar e comparar unidades completas de DNA ribossômico (rADN) e genomas de cloroplasto e verificaram que o doador do genoma maternal de *S. scabra* era mais similar ao do genoma de *S. hamata*. Dados deste mesmo estudo também sugerem que a variação encontrada no DNA ribossomal em *S. scabra* e seus progenitores podem ser explicadas pelos processos de homogeneização, amplificação e deleção que ocorreram desde suas origens. Além disso, também determinaram a idade da espécie *S. scabra*, datando de aproximadamente 0.63 a 0.52 milhões de anos atrás (meio do Pleistoceno).

Costa et al. (2018) avaliando a diversidade genética de quatro populações naturais de *S. scabra* localizadas no semiárido do estado de Pernambuco, demonstraram que os indivíduos coletados em Petrolina e Santa Cruz do Capibaribe possuíam a maior variabilidade genética, o que ressalta a importância da coleta de amostras de diferentes localidades para construção de bancos de germoplasma.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

2. REFERÊNCIAS BIBLIOGRÁFICAS

- Abe H, Urao T, Ito T, Seki M, Shinozaki K e Yamaguchi-Shinozaki K (2003) *Arabidopsis* *AtMYC2* (bHLH) and *AtMYB2* (MYB) function as transcriptional activators in abscisic acid signaling. **The Plant Cell** **15**: 63–78.
- Abrahám E, Rigó G, Székely G, Nagy R, Koncz C e Szabados L (2003) Light-dependent induction of proline biosynthesis by abscisic acid and salt stress is inhibited by brassinosteroid in *Arabidopsis*. **Plant Molecular Biology** **51**: 363–372.
- Abrahão JJS e Roma CFC (2010) Produção animal e valor nutritivo da forragem de pastagem de coast cross consorciada com amendoim forrageiro. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** **62**: 645-653.
- Agarwal PK, Agarwal P, Reddy MK e Sopory SK (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. **Plant Cell Rep** **25**: 1263–1274.
- Agboma PC, Sinclair TR, Jokinen K, Peltonen-Sainio P e Pehu E (1997) An evaluation of the effect of exogenous glycine betaine on the growth and yield of soybean: timing of application, watering regimes and cultivars. **Field Crops Research** **54**: 51–64.
- Ahmad I e Hellebust JA (1988) The relationship between inorganic nitrogen metabolism and proline accumulation in osmoregulatory responses of two Euryhaline microalgae. **Plant Physiol.** **88**: 348–354.
- Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JKC, Martin G, Tanksley SD, Bouzayen M, Jahn MM e Giovannoni J (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. **The Plant Journal: For Cell and Molecular Biology** **39**: 697–714.
- Almeida RG, Nascimento JRD, Euclides VPB, Macedo MCM, Fonseca DM, Brâncio PA e Barbosa RA (2003) Disponibilidade, composição botânica e valor nutritivo da forragem de pastos consorciados, sob três taxas de lotação. **Revista Brasileira de Zootecnia** **32**: 36-46.
- Amorim LDM (2014) **Fabaceae Lindl. da Floresta Nacional de Assú, semiárido do Rio Grande do Norte, Brasil**. 111 p.
- Andrade RP e Karia CT (2000) **O uso de *Stylosanthes* em pastagens no Brasil**. Em: **Simpósio de forragicultura e pastagem: Temas em evidência**. Anais... Lavras: UFLA: 273-309p.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- APG IV (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. **Botanical Journal of the Linnean Society** **181**: 1–20.
- Ashraf M e Foolad MR (2007) Roles of glycine betaine and proline in improving plant abiotic stress resistance. **Environmental and Experimental Botany** **59**: 206–216.
- Assad ED, Pinto HS, Zullo-Junior J e Brunini O (2002) Mudanças climáticas: o aquecimento global e a agricultura. **Reportagens - Com Ciência**.
- Avni Öktem H, Eyidoğan F, Selçuk F, Tufan Öz M, da Silva JAT e Yücel M (2008) Revealing Response of Plants to Biotic and Abiotic Stresses with Microarray Technology. **Genes, Genomes and Genomics**: 14-48
- Avonce N, Wuyts J, Verschooten K, Vandesteene L e Van Dijck P (2010) The *Cytophaga hutchinsonii* ChTPSP: First characterized bifunctional TPS-TPP protein as putative ancestor of all eukaryotic trehalose biosynthesis proteins. **Molecular Biology and Evolution** **27**: 359–369.
- Awasthi R, Kaushal N, Vadez V, Turner NC, Berger J, Siddique KHM e Nayyar H (2014) Individual and combined effects of transient drought and heat stress on carbon assimilation and seed filling in chickpea. **Functional Plant Biology** **41**: 1148.
- Barbero LM, Cecato U, Lugão SMB, Gomes JAN, Limão VA e Basso KC (2009) Produção de forragem e componentes morfológicos em pastagem de coastcross consorciada com amendoim forrageiro. **Revista Brasileira de Zootecnia** **38**: 788-795.
- Battistin A (1981) **Estudo biosistemático de diferentes taxons do gênero Stylosanthes Sw. (Leguminosae-Papilionoideae)**. 106p.
- Belhassen E, This D e Monneveux P (1995) L'adaptation génétique face aux contraintes de sécheresse. **Cahiers Agricultures** **4**: 251-261.
- Benko-Iseppon AM, Soares-Cavalcanti NM, Berlarmino LC, Bezerra-Neto JP, Amorim LLB, Ferreira Neto JRC, Pandolfi V, Azevedo HMA, Silva RLO, Santos MG, Alves MVS, Kido EA (2011) Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. Volume Especial do III Simpósio de Mudanças Climáticas e Desertificação no Semiárido Brasileiro. **Revista Brasileira de Geografia Física** **4**: 1112-1134.
- Bergamaschi H (1992) Desenvolvimento de déficit hídrico em culturas. In Bergamaschi, H. (eds) **Agrometeorologia Aplicada à Irrigação**. UFRGS Editora, Porto Alegre, p. 25-32.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- BFG – The Brazil Flora Group (2015) Growing knowledge: an overview of Seed Plant diversity in Brazil. **Rodriguésia** **66**: 1085–1113.
- Bhattacharjee S (2005) Reactive oxygen species and oxidative burst: roles in stress, senescence and signal transduction in plants. **Current Science** **89**: 1113–1121.
- Blum A (1996) Crop responses to drought and the interpretation of adaptation. **Plant Growth Regul.** **20**: 135–48.
- Blum A, Mayer J, Gozlan G (1983) Associations between plant production and some physiological components of drought resistance in wheat. **Plant Cell Environ** **6**: 219–25.
- Blumwald E (2000). Sodium transport and salt tolerance in plants. **Curr Opin Cell Biol** **12**: 431-4.
- Blunden G e Gordon SM (1986) Betaines and their sulphonio analogues in marine algae. **Prog Phycol Res** **4**: 39–80.
- Bohnert HJ e Jensen RG (1996) Strategies for engineering water stress tolerance in plants. **Trends Biotechnol** **14**: 89–97.
- Borém A e Santos FR (2001) **Biotecnologia simplificada**. Viçosa. Editora Suprema.
- Borghetti F, Nogueira RJMC, Nogueira R, Takaki M, Cambraia J, Araújo LS, Willadino L, Cavalcante UMT, Araújo EM, Albuquerque MEBC, Silva C, Cavalcante M, Albuquerque MDFPMD, Silva W da S e, Pimentel C, Borghetti AF and Araújo EL (2005) Estresses ambientais: danos e benefícios em plantas. **ScienceOpen**.
- Brandão MB e Costa NMS (1982) **O gênero *Stylosanthes* Swartz no Estado de Minas Gerais**. Epamig, 1982.
- Bray EA, Bailey-Serres J, Weretilnyk E (2000) Responses to abiotic stresses. In Gruissem W Buchannan B Jones R (Eds) **Biochemistry and molecular biology of plants**. American Society of Plant Physiologists. Rockville, p. 1158-1249.
- Brigotti M, Petronini PG, Carnicelli D, Alfieri RR, *et al.* (2003) Effect of osmolarity, ions and compatible osmolytes on cell-free protein synthesis. **Biochem. J.** **369**: 369–374.
- Brown AD (1976) Microbial water stress. **Bacteriol Rev** **40**: 803–846.
- Calles T e Schultze-kraft R (2016) New species, nomenclatural changes and recent taxonomic studies in the genus *Stylosanthes* (Leguminosae): An update. **Tropical Grasslands-Forrajões Tropicales** **4**: 122–128.
- Cameron DF (1967) Chromosome number and morphology of some introduced *Stylosanthes* species. **Aust. J. Agric. Res.** **18**: 375-379.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Cardoso-Silva CB (2015) **Análise do transcriptoma e de sequências genômicas de variedades comerciais de cana-de-açúcar.** 104p.
- Carmo-Silva AE, Gore MA, Andrade-Sanchez P, French AN, Hunsaker DJ e Salvucci ME (2012) Decreased CO₂ availability and inactivation of Rubisco limit photosynthesis in cotton plants under heat and drought stress in the field. **Environmental and Experimental Botany 83**: 1–11.
- Carraro DM e Kitajima, JP (2002) Sequenciamento e Bioinformática de Genomas Bacterianos. **Biotecnol Ciênc Des 28**: 16-20.
- Casaretto JA, El-Kereamy A, Zeng B, Stiegelmeyer SM, Chen X, Bi Y-M e Rothstein SJ (2016) Expression of *OsMYB55* in maize activates stress-responsive genes and enhances heat and drought tolerance. **BMC genomics 17**: 312.
- Chaves MM, Maroco J, Pereira JS (2003) Understanding plant responses to drought from genes to the whole plant. **Functional Plant Biology 30**: 239-264.
- Chaves MM e Oliveira MM (2004) Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. **J Exp Bot. 55**: 2365–84.
- Chaves MM, Pereira JS, Rodrigues ML, Ricardo CPP, Osório ML, Carvalho I, Faria T, Pinheiro C (2002) How plants cope with water stress in the Field. Photosynthesis and growth. **Annals of Botany 89**: 907-916.
- Chinnusamy V, Zhu J, Zhou T e Zhu J-K (2007) Small Rnas: Big Role In Abiotic Stress Tolerance Of Plants. Springer Netherlands, Dordrecht, p. 223–260. In Jenks MA, Hasegawa PM, and Jain SM (eds) **Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops.**
- Claussen W (2005) Proline as a measure of stress in tomato plants. **Plant Science 168**: 241-248.
- Colaco C, Sen S, Thangavelu M, Pinder S e Roser B (1992) Extraordinary Stability of Enzymes Dried in Trehalose: Simplified Molecular Biology. **Biotechnology (Nature Publishing Company) 10**: 1007–11.
- Condon AG (1982) **Water relations and osmotic adjustment in wheat : as influenced by water stress history and plant development.**
- Costa JC, Fracetto GGM, Fracetto FJC, Souza TC, Santos MVF, Lira Júnior MA (2018) Genetic diversity in natural populations of *Stylosanthes scabra* using ISSR markers. **Genetics and Molecular Research. 17.**
- Costa NMS (2006) **Revisão do gênero *Stylosanthes* Sw.** 469p.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- CRIA, Centro de Referência em Informação Ambiental, Fundação de Amparo à Pesquisa do Estado de São Paulo *Stylosanthes*. In: **speciesBase**. Disponível em: <<http://www.splink.org.br/index?lang=pt>>. Acesso em: Dez 2018.
- Cronquist A (1981) **An Integrated System of Classification of Flowering Plants**. New York, Columbia Univ. Press. 1262.
- Crowe JH, Crowe LM, Carpenter JF e Wistrom AC (1987) Stabilization of dry phospholipid bilayers and proteins by sugars. **Biochem. J.** **242**: 1–10.
- Cushman JC e Bohnert HJ (2000) Genomic approaches to plant stress tolerance. **Current Opinion in. Plant Biology** **3**: 117-124.
- Cvikrova M, Gemperlova L, Martincova O, Vankova R (2013) Effect of drought and combined drought and heat stress on polyamine metabolism in proline-over-producing tobacco plants. **Plant Physiol Biochem** **73**: 7–15.
- Danquah A, de Zelicourt A, Colcombet J, Hirt H (2014) The role of ABA and MAPK signaling pathways in plant abiotic stress responses. **Biotechnology Advances** **32**: 40–52.
- Date RA, Jansen PI, Messer B, Eagles DA (2010) — Morphological variation and classification of field-grown *Stylosanthes seabrana* and *S. scabra*. *Tropical Grasslands* 44.
- Davies WJ e Zhang J (1991) Root signals and the regulation of growth and development of plants in drying soil. **Annual Review of Plant Physiology and Molecular Biology** **42**: 55-76.
- Dayhoff MO (1965) Computer aids to protein sequence determination. **J. Theor. Biol** **8**: 97–112.
- Delauney AJ e Verma DPS (1993) Proline biosynthesis and osmoregulation in plants. **The Plant Journal** **4**: 215–223.
- Dore MHI (2005) Climate change and changes in global precipitation patterns: What do we know? **Environmental International** **31**: 1167-1181.
- Dryflor (2016) Plant diversity patterns in Neotropical dry forests and their conservation implications. **Science** **353**: 1383–1387.
- Edwards D e Batley J (2004) Plant bioinformatics: from genome to Phenome. **Trends in Biotechnology** **22**: 232-237.
- Edye LA (1997) Commercial development of *Stylosanthes* pastures in Northern Australia. I. Cultivar development within *Stylosanthes* in Australia. **Tropical Grasslands** **31**: 503-508.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Elbein AD (1974) The metabolism of alpha, alpha-trehalose. **Adv. Carb. Chem. Biochem.** **30**: 227-256.
- EMBRAPA (2007) **Cultivo e uso do Estilosantes Campo-Grande**. Embrapa Gado de Corte. Comunicado Técnico, 105. Campo Grande, MS. 11p.
- Espindola FS, Calábria LK, Rezende AAA, Barbosa PB, Assumpção SF, Marques RAI *et al.* (2010) Recursos de bioinformática aplicados às ciências Ômicas como genômica, transcriptômica, proteômica, interatômica e metabolômica. **Bioscience J** **26**: 463-77.
- Euclides VPB, Valle CB, Macedo MCM, Almeida RG, Montagner DB e Barbosa RA (2013) Brazilian scientific progress in pasture research during the first decade of XXI century. **Revista Brasileira de Zootecnia** **39**: 151-168.
- Fageria, NK (1989) **Solos tropicais e aspectos fisiológicos das culturas**. EMBRAPA - DPU. Brasília, 425p.
- Falco JS, Fernandes CD, Verzignassi JR, Mallmann G, Queiróz CA, Chagas HA, Batista MV e Quetez FA (2016) Reação de genótipos de *Stylosanthes capitata* à antracnose. **Summa Phytopathologica** **42**:140-148.
- Fallon KM e Phillips R (1989) Responses to Water Stress in Adapted and Unadapted Carrot Cell Suspension Cultures. **Journal of Experimental Botany** **40**: 681–687.
- Fan L, Linker R, Gepstein S, Tanimoto E, Yamamoto R, Neumann PM (2006) Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increase lignin metabolism and progressive stellar accumulation of wall phenolics. **Plant Physiology** **140**: 603-612.
- Ferreira MB e Costa NMS (1979) **O gênero *Stylosanthes* Sw. no Brasil**. Belo Horizonte : Empresa de Pesquisa Agropecuária de Minas Gerais, 116p.
- Ferreira-Neto JRC, Benko-Iseppon AM, Kido EA (2018) **Transcriptômica: por quê, como e para que? Uma Abordagem para Iniciantes**. Beau Bassin: Novas Edições Acadêmicas, 68p.
- Fietto J e Maciel T (2013) Sequenciando genomas. Em: Leandro Márcio Moreira (Org.) **Ciências Genômicas: fundamentos e aplicações**. Sociedade Brasileira de Genética. Ribeirão Preto, p. 65-80.
- Fujita Y, Fujita M, Shinozaki K e Yamaguchi-Shinozaki K (2011) ABA-mediated transcriptional regulation in response to osmotic stress in plants. **Journal of Plant Research** **124**: 509–525.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- G. Grabherr M, Haas B, Yassour M, Levin JZ, Thompson DA and Amit I (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. **Nat Biotechnol** **29**: 644–652.
- Gage DA, Rathinasabapathi B (1999) Role of glycine betaine and dimethylsulfoniopropionate in water stress tolerance. In Shinozaki K, Yamaguchishinozaki K (eds) **Molecular response to cold, drought, heat and salt stress in higher plants**. RG Landes Company. Georgetown, p. 125–152.
- Garg AK, Ju-Kon K, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu RJ (2002) Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. **Proceedings of National Academy of Sciences of the United States of America** **99**: 15898-15903.
- Geeta (2014) Genetic engineering for osmolyte overproduction for abiotic stress tolerance in plants. **Int J Curr Res** **6**: 6130–6135.
- Génard H, Le-Saos J, Billard JP, Trémolières A et al. (1991) Effect of salinity on lipid composition, glycine betaine content and photosynthetic activity in chloroplasts of *Suaeda maritima*. **Plant Physiol. Biochem.** **29**: 421-427.
- Ghelfi A, Gaziola SA, Cia MC, Chabregas SM, Falco MC, Kuser-Falcão PR, Azevedo RA (2011) Cloning, expression, molecular modeling and docking analysis of glutathione transferase from *Saccharum officinarum*. **Annals of Applied Biology** **159**: 267–280.
- Giachetto PF e Higa RH (2015) **Bioinformática Aplicada na Agricultura**. In: Embrapa Informática Agropecuária, Capítulo em livro técnico-científico. Capítulo 4, parte II.
- Gill SS e Tuteja, N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant Physiology and Biochemistry** **48**: 909-930.
- Gilroy S, Suzuki N, Miller G, Choi W-G, Toyota M, Devireddy AR e Mittler R (2014) A tidal wave of signals: calcium and ROS at the forefront of rapid systemic signaling. **Trends in Plant Science** **19**: 623–630.
- Ginzberg I, Stein H, Kapulnik Y, Szabados L, Strizhov N, Schell J, Koncz C e Zilberstein A (1998) Isolation and characterization of two different cDNAs of delta1-pyrroline-5-carboxylate synthase in alfalfa, transcriptionally induced upon salt stress. **Plant Molecular Biology** **38**: 755–764.
- Giri J (2011) Glycine betaine and abiotic stress tolerance in plants. **Plant Signal. Behav.** **6**: 1746-1751.
- Giusti J, Kettener K e Fuchs-Ferraz MCP (2016) Influência do sequenciamento de nova geração no futuro da genética de conservação. **Revista RG News** **2(2)**: 88-99.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Glenn TC (2011) Field guide to next-generation DNA sequencers. **Mol. Ecol. Resour** **11**: 759-69.
- Goddijn OJM e Van Dun K (1999) Trehalose metabolism in plants. **Trends Plant Sci** **4**: 315–319.
- Goldblatt P (1981) Cytology and phylogeny of leguminosae. In: Polhill RM e Raven RH **Advances in legume systematics**. Kew: Royal Botanic Garden. 427- 463.
- Gomes MMA et al. (2004) Interactions between leaf water potential, stomatal conductance and abscisic acid content of orange trees submitted to drought stress. **Brazilian Journal of Plant Physiology** **16**: 155-161.
- Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. **Nat Rev Genet** **17**: 333-51.
- Guodao L, Phaikaew C e Stür WW (1997) Status of *Stylosanthes* development in other countries.II. *Stylosanthes* development and utilisation in China and southe-east Asia. **Tropical Grasslands**.
- Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M. (2013) Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. **Int J Mol Sci**. **14**: 9643-84.
- Hasthanasombut S, Ntui V, Supaibulwatana K, Mii M, Nakamura I (2010) Expression of Indica rice *OsBADH1* gene under salinity stress in transgenic tobacco. **Plant Biotechnol Rep** **4**: 75–83.
- Herling VR e Pereira LET (2016) **Leguminosas forrageiras de clima tropical e temperado**. **Universidade de São Paulo**. Pirassununga, 103p.
- Holmstrom KO, Mantyla E, Welin B, Mandal A, Palva ET (1996) Drought tolerance in tobacco. **Nature** **379**: 683–684.
- Hong Z, Lakkineni K, Zhang Z e Verma DPS (2000). Removal of feedback inhibition of Δ -pyrroline-5-carboxylate synthetase results in increased proline accumulation and protection of plants from osmotic stress. **Plant Physiol**. **122**: 1129-1136.
- Hoogenboom G, Huck MG, Peterson CM (1987) Root growth rate of soybean as affected by drought stress. **Agron. J**. **79**: 607–614.
- Huang XC, Quesada MA, Mathies RA (1992) DNA sequencing using capillary array electrophoresis. **Anal. Chem**. **64**: 2149–2154.
- Idris S e Saad MS (2001) Characterization of plant genetic resources. In Saad MS e Rao VR (Eds.). **Establishment and management of field genebank, a training manual**. Serdang: IPGRI-APO. p. 81-86.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Ishitani M, Majumder AL, Bornhouser A, Michalowski CB, Jensen RG e Bohnert HJ (1996) Coordinate transcriptional induction of myo-inositol metabolism during environmental stress. **The Plant journal : for cell and molecular biology** **9**: 537–548.
- Jagendorf AT e Takabe T (2001) Inducers of Glycinebetaine Synthesis in Barley. **Plant Physiology** **127**: 1827–1835.
- Jank L, Valle CB e Resende RMS (2011) Breeding tropical forages. **Crop Breeding and Applied Biotechnology** **11**: 27-34.
- Judd WS, Campbell CS, Kellogg EA, Stevens PF (1999) **Plant Systematics: a phylogenetic approach**. Sinauer Associates. Massachusetts, 464p.
- Judd WS, Campbell CS, Kellogg EA, Stevens PF e Donoghue MJ (2009) **Sistemática Vegetal** - 3.ed.: Um Enfoque Filogenético. Artmed Editora, 631p.
- Kamoshita A, Rodriguez R, Yamauchi A, Wade LJ (2001) Response of rainfed-lowland rice genotypes to prolonged drought and rewatering during the vegetative stage. In: Fukai S, and Basnayake J (eds) **Increased lowland rice production in the Mekong region**. ACIAR Proceedings No 101. Canberra, p. 78–85.
- Kanehisa M e Bork P (2003) Bioinformatics in post-sequencing era. **Nature Genetics** **33**: 305-310.
- Kaplan F, Kopka J, Dale WH, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL (2004) Exploring the temperature-stress metabolome of Arabidopsis. **Plant Physiol** **136**: 4159–4168.
- Karia CT (2008) **Caracterização genética e morfoagronômica de germoplasma de *Stylosanthes guianensis* (Aubl.) Sw.**
- Karia CT, Andrade RP, Charchar MJDA, Gomes AC (2002) **Caracterização morfológica de acessos do gênero *Stylosanthes* no banco ativo de germoplasma da Embrapa Cerrados - coleção 1994/1995**. Embrapa Cerrados. Planaltina, 24p.
- Karim S, Aronsson H, Ericson H, Pirhonen M, Leyman B, Welin B, Mantyla E, Palva ET, Van Dijck P, Holmstrom KO (2007) Improved drought tolerance without undesired side effects in transgenic plants producing trehalose. **Plant Mol Biol** **64**: 371–386.
- Kaur G e Asthir B (2017). Molecular responses to drought stress in plants. **Biologia Plantarum**. **61**: 201-209.
- Kaur N e Gupta AK (2005) Signal transduction pathways under abiotic stresses in plants. **Curr. Sci.** **88**: 1771–1780.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM *et al.* (2002) The human genome browser at UCSC. **Genome Research** **12**: 996-1006.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Khan MS, Ahmad D, Khan M.A (2015) Utilization of genes encoding osmoprotectants in transgenic plants for enhanced abiotic stress tolerance. **Electron. J. Biotech. 18:** 257-266.
- Khurana N, Sharma N e Khurana P (2017) Overexpression of a heat stress inducible, wheat myo-inositol-1-phosphate synthase 2 (*TaMIPS2*) confers tolerance to various abiotic stresses in *Arabidopsis thaliana*. **Agri Gene 6:** 24-30.
- Kim TH (2014) Mechanism of ABA signal transduction: agricultural highlights for improving drought tolerance. **J. Plant Biol. 57:** 1-8.
- Kishor PBK, Hong Z, Miao GH, Hu CAA, Verma DPS (1995) Overexpression of 1-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. **Plant Physiol 108:** 1387–1394
- Klitgaard BB e Lavin M (2005) Tribe Dalbergieae sens. lat. In Lewis GP, Schrire B, Mackinder B e Lock M (eds.) **Legumes of the world**. Royal Botanic Gardens, Kew. Pp. 307-335.
- Krallinger M e Valencia A (2005) Text-mining and information-retrieval services for molecular biology. **Genome Biology 6:** 224.
- Krishnamurthy A e Rathinasabapathi B (2013) Novel interplay between auxin and reactive oxygen species signaling Oxidative stress tolerance in plants. **Plant Signaling and Behavior 8:** 1–5.
- Lamitina T, Huang CG, Strange K (2006) Genome-wide RNAi screening identifies protein damage as a regulator of osmoprotective gene expression. **Proc Natl Acad Sci U S A 103:** 12173–12178.
- Larcher W (1995) **Physiological Plant Ecology, Ecophysiology and Stress Physiology of Functional Groups**. Springer. Berlin, Heidelberg, New York. 506p.
- Larcher W (2000) **Ecofisiologia Vegetal**. RiMa Artes e Texto. São Paulo, 319p.
- Larcher W (2006) **Ecofisiologia vegetal**. Rima. São Carlos, 550p.
- Lavin M, Schrire BP, Lewis G, Pennington RT, Delgado-Salinas A, Thulin M, Hughes CE, Matos AB e Wojciechowski MF (2004) Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. **Philosophical Transactions of the Royal Society B: Biological Sciences 359:** 1509–1522.
- Leal-Bertioli SC, Messenberg P, Fávero GAP, Moretzsohn MC et al. (2003) Amendoim Selvagem - Uma fonte de resistência a pragas. **Ver. Biotec. Cien. Desenv. 31:** 116-119.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Ledley RS (1959) Digital Electronic Computers in Biomedical Science: Computers make solutions to complex biomedical problems feasible, but obstacles curb widespread use. **Science 130**: 1225–1234.
- Lewis GP (1987) **Legumes of Bahia**. Royal Botanic Gardens, Kew. 369.
- Li HW, Zang BS, Deng XW, Wang XP (2011) Overexpression of the trehalose-6-phosphate synthase gene *OsTPS1* enhances abiotic stress tolerance in rice. **Planta 234**: 1007-1018.
- Lima HC, Correia CMB, Farias DS (1994) Leguminosae. In: Lima MPM, Guedes-Bruni RR (orgs.). **Reserva Ecológica de Macaé de Cima: Nova Friburgo - RJ: Aspectos florísticos das espécies vasculares**. Jardim Botânico do Rio de Janeiro, Rio de Janeiro.
- Lima WP, Jarvis P, Rhizopoulou S (2003) Stomatal responses of *Euzalyptus* species to elevated CO₂ concentration and drought stress. **Scientia Agricola 60**: 231-238.
- Linchtenthaler HK (1996) Vegetation stress: an introduction to the stress concept in plants. **Journal of Plant Physiology 148**: 4-14
- Lira IC de SA (2015) **Caracterização citogenética e morfoagronômica de acesso de *Stylosanthes* spp. (Fabaceae – Papilionoideae) coletados no Nordeste brasileiro**.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH e Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. **Cell 133**: 523–536.
- Liu C (1997) Geographical distribution of genetic variation in *Stylosanthes scabra* revealed by RAPD analysis. **Euphytica 98**: 21.
- Liu CJ e Musial JM (1997) *Stylosanthes* sp. aft. *S. scabra*: a putative diploid progenitor of *Stylosanthes scabra* (Fabaceae). **Plant Systematics and Evolution. 208**: 99-105.
- Loewus FA e Loewus MW (1983) Myo-inositol: its biosynthesis and metabolism. **Plant Physiol. 34**: 137-161.
- Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome biology 15**: 550.
- Ludlow MM e Muchow RC (1990) A critical evaluation of traits for improving crop yields in water-limited environments. **Adv Agron. 43**: 107–53.
- LWPG (2017) Leguminosae Working Group Phylogeny. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. **Taxon 66**: 44-77.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Ma R, Song W, Wang F, Cao A, Xie S, Chen X, Jin X e Li H (2019) A Cotton (*Gossypium hirsutum*) Myo-Inositol-1-Phosphate Synthase (*GhMIPS1D*) Gene Promotes Root Cell Elongation in *Arabidopsis*. **International journal of molecular sciences** **20**.
- Mahajan S e Tuteja N (2005) Cold salinity and drought stresses: an overview. **Archives of Biochemistry and Biophysics** **444**: 139-158.
- Majumder AL, Johnson MD e Henry SA (1997) 1L-myo-inositol-1-phosphate synthase. **Biochimica et biophysica Acta** **1348**: 245–256.
- Mansour MMF (1998) Protection of plasma membrane of onion epidermal cells by glycinebetaine and proline against NaCl stress. **Plant Physiol. Biochem.** **36**: 767-772.
- Marguerat S e Bähler J (2010) RNA-seq: from technology to biology. **Cellular and molecular life sciences** **67**: 569–579.
- Marioni JC, Mason CE, Mane SM, Stephens M e Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. **Genome Research** **18**: 1509–1517.
- Marques A, Moraes L, Aparecida Dos Santos M, Costa I, Costa L, Nunes T, Melo N, Simon MF, Leitch AR, Almeida C e Souza G (2018) Origin and parental genome characterization of the allotetraploid *Stylosanthes scabra* Vogel (Papilionoideae, Leguminosae), an important legume pasture crop. **Annals of Botany** **122**: 1143–1159.
- McCue KF e Hanson AD (1990) Drought and salt tolerance: towards understanding and application. **Trends in Biotechnology** **8**: 358–362.
- McCue KF e Hanson AD (1992). Effects of soil salinity on the expression of betaine aldehyde dehydrogenase in leaves: investigation on hydraulic, ionic and biochemical signals. **Aust. J. Plant Physiol.** **19**: 555–564.
- Meena KK, Sorty AM, Bitla UM, Choudhary K, Gupta P, Pareek A, Singh DP, Prabha R, Sahu PK, Gupta VK, Singh HB, Krishanani KK e Minhas PS (2017) Abiotic Stress Responses and Microbe-Mediated Mitigation in Plants: The Omics Strategies. **Front. Plant Sci.** **8**: 172.
- Miller CP, Rains JP, Shaw KA, Middleton CH (1997) Commercial development of *Stylosanthes* pastures in Northern Australia. II. *Stylosanthes* in the 119 northern Australian Beef Industry. **Tropical Grasslands** **31**: 509- 514.
- Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R (2010) Reactive oxygen species homeostasis and signalling during drought and salinity stresses. **Plant Cell Environmental** **33**: 453– 67.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Milosavljević A, Prokić L, Marjanovic M, Stikic R e Sabovljevic A (2012) The effects of drought on the expression of *TAO1*, *NCED* and *EIL1* genes and ABA content in tomato wild-type and flacca mutant. **Archives of Biological Sciences** **64**: 297–306.
- Mochida K, Yoshida T, Sakurai T, Yamaguchi-Shinozaki K, Shinozaki K, Tran LSP (2010) Legume TFDB: an integrative database of *Glycine max*, *Lotus japonicus* and *Medicago truncatula* transcription factors. **Bioinformatics** **26**: 290–291.
- Mohlenbrock RH (1963) Further consideration in *Stylosanthes* (Leguminosae). **Rhodora**, **63**: 245 – 258.
- Morgan JM (1984) Osmoregulation and water stress in higher plants. **Annu Rev Plant Physiol** **35**: 299–319.
- Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S e Marra M (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. **BioTechniques** **45**: 81–94.
- Morozova O e Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. **Genomics** **92**: 255–264.
- Müller J, Aeschbacher RA, Wingler A, Boller T, Wiemken A (2001) Trehalose and trehalase in *Arabidopsis*. **Plant Physiol.** **125**: 1086-1093.
- Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M e Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. **Science** **320**: 1344–1349.
- Nakashima K, Kiyosue T, Yamaguchi-Shinozaki, KShinozaki K (1997) A nuclear gene, *erd1*, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*. **The Plant Journal** **12**: 851–861.
- Nelson DE, Rammesmayer G e Bohnert HJ (1998) Regulation of cell-specific inositol metabolism and transport in plant salinity tolerance. **The Plant Cell** **10**: 753–764.
- Nisa Z, Chen Chen, Yu Y, Chen Chao, Mallano All, Xiang-bo D, Xiao-li S e Yan-ming Z (2016) Constitutive Overexpression of Myo-inositol-1-Phosphate Synthase Gene (*GsMIPS2*) from Glycine soja Confers Enhanced Salt Tolerance at Various Growth Stages in *Arabidopsis*. **Journal of Northeast Agricultural University** **23**: 28–44.
- Nuccio ML, Wu J, Mowers R, Zhou HP, Meghji M, Primavesi LF, Paul MJ, Chen X, Gao Y, Haque E, Basu SS, Lagrimini LM (2015) Expression of trehalose-6- phosphate phosphatase in maize ears improves yield in well watered and drought conditions. **Natur. Biotechnol.** **33**: 862- 869.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Nyysola A, Kerovuo J, Kaukinen P, von Weymarn N, Reinikainen T (2000) Extreme halophiles synthesize betaine from glycine by methylation. **J Biol Chem** **275**: 22196–22201.
- Oliveira RS de (2015) **Coleta, caracterização e avaliação preliminar de acessos de Stylosanthes spp.**
- Oliveira RS e Queiroz MA (2016) Banco Ativo de Germoplasma de *Stylosanthes* spp. da Universidade Estadual de Feira de Santana – BA. **Revista RG News** **2**: 119-128.
- Oliveira RS, Queiróz MA, Romão RS, Silva GC, Brasileiro BP (2016) Genetic diversity in accessions of *Stylosanthes* spp. using morphoagronomic descriptors. **Revista Caatinga** **29**: 101-112.
- Oran RN (1990) **Register of Australian Herbage Plants Cultivars**. CSIRO, Melbourne.
- Pandey GK (2008) Emergence of a novel calcium signaling pathway in plants: CBL-CIPK signaling network. **Physiology and Molecular Biology of Plants** **14**: 51- 68.
- Papageorgiou GC e Murata N (1995). The unusually strong stabilizing effects of glycine betaine on the structure and function of the oxygen-evolving photosystem II complex. **Photosynth. Res.** **44**: 243-252.
- Park ST e Kim J (2016) Trends in Next-Generation Sequencing and a New Era for Whole Genome Sequencing. **Int Neurorol J** **20(Suppl 2)**: S76-83.
- Parthasarathy L, Vadnal RE, Parthasarathy R, Devi CS (1994) Biochemical and molecular properties of lithium-sensitive myo-inositol monophosphatase. **Life Sciences** **54**: 1127–1142.
- Peng Z, Lu Q e Verma DP (1996). Reciprocal regulation of Δ^1 -pyrroline-5-carboxylate synthetase and proline dehydrogenase genes control proline levels during and after osmotic stress in plants. **Mol. Gen. Genet.** **253**: 334-341.
- Pereira AV, Valle CB, Ferreira RDEP, Milles JW (2001) Melhoramento de forrageiras tropicais. In: Nass LL, Valois ACC, Melo IS. Inglis-Valadares MC (eds) **Recursos Genéticos & Melhoramento de Plantas**. Fundação MT, Rondonópolis, p. 549-601.
- Petrey D e Honig B (2014) Structural Bioinformatics of the Interactome. **Annu Rev Biophys** **43**: 193–210.
- Pfeifer SP (2017) From next-generation resequencing reads to a high-quality variant data set. **Heredity** **118**: 111–124.
- Portes MT, Alves TH e Souza GM (2006) Water deficit affects photosynthetic induction in *Bauhinia forficata* Link (Fabaceae) and *Esenbeckia leiocarpa* Engl. (Rutaceae)

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

growing in understorey and gap conditions. **Brazilian Journal of Plant Physiology** **18**: 491–502.

Rathinasabapathi B, Burnet M, Russell BL, Gage DA, Liao PC, Nye GJ, Scott P, Golbeck JH e Hanson AD (1997) Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: prosthetic group characterization and cDNA cloning. **Proceedings of the National Academy of Sciences of the United States of America** **94**: 3454–3458.

Reuter JA, Spacek DV, Snyder MP (2015) High-Throughput Sequencing Technologies. **Mol Cell** **58**: 586–597.

Rhodes D e Hanson AD (1993) Quaternary Ammonium and Tertiary Sulfonium Compounds in Higher Plants. **Annual Review of Plant Physiology and Plant Molecular Biology** **44**: 357–384.

Rizhsky L, Hongjian L, Mittler R (2002) The combined effect of drought stress and heat shock on gene expression in tobacco. **Plant Physiol** **130**: 1143–1151.

Rizhsky L, Liang HJ, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress. **Plant Physiology** **134**: 1683–1696.

Robinson MD, McCarthy DJ and Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics (Oxford, England)** **26**: 139–140.

Rontein D, Basset G, Hanson AD (2002) Metabolic engineering of osmoprotectant accumulation in plants. **Metabolic engineering** **4**: 49-56.

Rozen S e Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. **Methods in Molecular Biology** **132**: 365–386.

Saeys Y, Inza I, Larrañaga P (2007) A review of feature selection techniques in bioinformatics. **Bioinformatics** **23**: 2507–1257.

Salamoni AT (2008) **Apostila de aulas teóricas de Fisiologia Vegetal**. Departamento de Engenharia Florestal, Universidade Federal de Santa Maria. sem.2.

Sanger F (1988) Sequences, sequences and sequences. **Annu. Rev. Bioch** **57**: 1 – 28.

Sauter A, Davies WJ, Hartung W (2001) The long-distance abscisic acid signal in the droughted plant: the fate of the hormone on its way from root to shoot. **J Exp Bot.** **52(363)**: 1991–7.

Saxena SC, Salvi P, Kaur H, Verma P, Petla BP, Rao V, Kamble N e Majee M (2013) Differentially expressed myo-inositol monophosphatase gene (*CaIMP*) in chickpea

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

(*Cicer arietinum* L.) encodes a lithium-sensitive phosphatase enzyme with broad substrate specificity and improves seed germination and seedling growth under abiotic stresses. **Journal of experimental botany** **64**: 5623–5639.

Schmidt T e Heslop-Harrison JS (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. **Trends in Plant Science** **3**: 195–199.

Seki M, Kamei A, Yamaguchi-Shinozaki K, Shinozaki K (2003) Molecular responses to drought, salinity and frost: common and different paths for plant protection. **Current Opinion in Biotechnology** **14**: 194-199.

Sengupta A, Chakraborty M, Saha J, Gupta B e Gupta K (2016) Polyamines: Osmoprotectants in Plant Abiotic Stress Adaptation. Springer India, New Delhi, p. 97–127. In Iqbal N, Nazar R, and A. Khan N (eds) **Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies**. Springer, New Delhi.

Shao HB, Guo QJ, Chu LY, Zhao XN, Su ZL, Hu YC, Cheng JF (2007) Understanding molecular mechanism of higher plant plasticity under abiotic stress. **Colloids and Surfaces B: Biointerfaces** **54**: 37-45.

Sharma P, Jha A e R. S. D (2010) **Oxidative Stress and Antioxidative Defense Systems in Plants Growing under Abiotic Stresses**. p. Handbook of Plant and Crop Stress, 3rd Edition, Revised and Expanded, CRC Press, Taylor & Francis Publishing Company, Florida.2010. In Handbook of Plant and Crop Stress.

Sharma P, Jha AB, Dubey RS e Pessarakli M (2012) Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. **Journal of Botany**.

Shendure J e Ji H (2008) Next generations DNA sequencing. **Nature Biotechnology** **26**: 1135 – 1145.

Signorelli S e Monza J (2017) Identification of $\Delta 1$ -pyrroline 5-carboxylate synthase (P5CS) genes involved in the synthesis of proline in *Lotus japonicus*. **Plant Signaling & Behavior**. **12**.

Silva RLO, Ferreira JRCN, Pandol V, Chabregas SM, Burnquist WL, Benko-Iseppon AM, Kido EA (2011) Transcriptomic of sugarcane osmoprotectant during drought stress. In: Vasnthaiah H (ed) **Plants and environment**. InTech Publishers, p 272.

Silva RSS, Jacinto Demuner A, Wagner V, Casali D, Nascimento E, Antônio E, Pinheiro L and Barbosa LC (2002) Efeito do estresse hídrico sobre características de crescimento

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

e a produção de óleo essencial de *Melaleuca alternifolia* Cheel. **Acta Scientiarum (UEM)**. **24**: 1363–1368.

Silva-Junior AC da, Belarmino LC e Benko-Iseppon AM (2018) DNA microarray platform: prospection technological and scientific demand. **Revista GEINTEC - Gestão, Inovação e Tecnologias 8**: 4369–4380.

Simeão RM, Resende MDV, Laura VA, Jank L e Valle CB (2006) Genotypic evaluation of accessions and individual selection in *Stylosanthes* spp. by simulated BLUP method. **Crop Breeding and Applied Biotechnology 6**: 253- 260.

Stace HM (ed) e Edye LA (ed) (1984) **The biology and agronomy of Stylosanthes**. 654p.

Storey R, Ahmad N, Wyn Jones RG (1977) Taxonomic and ecological aspects of the distribution of glycinebetaine and related compounds in plants. **Oecologia 27**: 319–332

Strizhov N, Abraham E, Okresz L, Blickling S, Zilberstein A, Schell J, Koncz C, Szabados L (1997) Differential expression of two P5CS genes controlling proline accumulation during salt-stress requires ABA and is regulated by ABA1, ABI1 and AXR2 in *Arabidopsis*. **Plant J 12**: 557–569.

Styer JC (2000) **Regulating inositol biosynthesis in plants: myo-Inositol phosphate synthase and myo-Inositol monophosphate**. 690p.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES e Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. **Proceedings of the National Academy of Sciences of the United States of America 102**: 15545–15550.

Taiz L e Zeiger E (2010) **Plant physiology**. 3rd edn. Annals of Botany 91: 750–751.

Taiz L e Zeiguer E (2004) **Fisiologia Vegetal**. Tradução de: Santarém ER et al. Artimed. Porto Alegre, 719p.

Tangpremsri T, Fukai S, Fischer KS, Henzell RG (1991) Genotypic variation in osmotic adjustment in grain sorghum Relation with some growth attributes. **Australian Journal of Agricultural Research 42**: 759-767.

The Uniprot Consortium (2012) Update on activities at the Universal Protein Resource (UniProt) in 2013. **Nucleic Acids Research 41**: D43– D47.

Tran L-SP, Nakashima K, Sakuma Y, Osakabe Y, Qin F, Simpson SD, Maruyama K, Fujita Y, Shinozaki K e Yamaguchi-Shinozaki K (2007) Co-expression of the stress-inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

of the ERD1 gene in *Arabidopsis*. **The Plant Journal: For Cell and Molecular Biology** **49**: 46–63.

- Tuteja N (2009) Integrated Calcium Signaling in Plants. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 29–49. In Mancuso S and Baluka F (eds) **Signaling in Plants**.
- Valle CB (2002) Recursos genéticos de forrageiras para áreas tropicais. In: Conferência Virtual Global sobre Produção Orgânica de Bovinos de Corte (eds) **Embrapa-Pantanal**. Corumbá, MS.
- Valle CB, Jank L, Resende RMS (2009) O melhoramento de forrageiras tropicais no Brasil. **Revista Ceres** **56**: 460-472.
- Vanni RO e Fernandes A (2011) The true identity of *Stylosanthes seabrana* B.L. Maass & L. 't Mannetje (Leguminosae Papilionoideae). **Revista Caryologia** **64**: 247-250.
- Varuzza L (2013) **Introdução à análise de dados de sequenciadores de nova geração**. Versão 2.0.1.
- Verbruggen N, Hua XJ, May M e Van Montagu M (1996) Environmental and developmental signals modulate proline homeostasis: evidence for a negative transcriptional regulator. **Proc. Natl. Acad. Sci. USA**. **93**: 8787–8791.
- Verbruggen N, Villarroel R e Van Montagu M (1993) Osmoregulation of a pyrroline -5-carboxylate reductase gene in *Arabidopsis thaliana*. **Plant Physiology** **103**: 771-781.
- Vieira EA, Charchar MJA, Silva MS, Anjos JRN (2007) Virulência de isolados de *Colletotrichum gloeosporioides* de populações selvagens de *Stylosanthes* spp. **Pesquisa Agropecuária Brasileira** **42**: 661-667.
- Vieira MLC (1988) **Estudo citotaxonômico de espécies brasileiras do gênero *Stylosanthes* Sw.** 135p.
- Vinocur B e Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. **Curr. Opin. Biotechnol.** **16**: 123–132.
- Vitor CMT, Fonseca DM, Moreira LM, Nascimento Júnior D, Ribeiro Junior JI, Pereira AL (2008) Rendimento e composição química do capim braquiária introduzido em pastagem degradada de capim-gordura. **Revista Brasileira de Zootecnia** **37**: 2107-2114.
- Wang WX, Vinocur B, Shoseyov O e Altman A (2001) Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. **Acta Horticulturae** **560**: 285-292.
- Wang Z, Gerstein M e Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. **Nature reviews Genetics** **10**: 57–63.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Weckwerth W (2008) Integration of metabolomic and proteomic phenotypes: analysis of data covariance dissects starch and RFO metabolism from low and high temperature compensation response in *Arabidopsis thaliana*. **Mol. Cell Proteom** **7**: 1725–1736.
- Weimberg R, Lerner HR e Poljakoff-Mayber A (1984) Changes in growth and water-soluble solute concentrations in *Sorghum bicolor* stressed with sodium and potassium salts. **Physiologia Plantarum** **62**: 472–480.
- Wienkoop S, Morgenthal K, Wolschin F, Scholz M, Selbig J e Weckwerth W (2008) Integration of metabolomic and proteomic phenotypes: analysis of data covariance dissects starch and RFO metabolism from low and high temperature compensation response in *Arabidopsis thaliana*. **Molecular & cellular proteomics** **7**: 1725–1736.
- Willadino L e Camara TR (2010) Tolerância das plantas à salinidade: aspectos fisiológicos e bioquímicos. **Enciclopédia Biosfera - Centro Científico Conhecer** **6**: 1-23.
- Williams RJ, Reid R, Schultze-Kraft R, Costa NMS e Thomas BD (1984) Natural distribution of *Stylosanthes*. In: Stace HM e Edye LA (Ed.) **The biology and agronomy of Stylosanthes**. North Ryde: Academic Press Australia, 73-101.
- Williamson CL e Slocum RD (1992) Molecular cloning and evidence for osmoregulation of the delta1-pyrroline -5-carboxylate reductase (proC) gene in pea (*Pisum sativum* L.). **Plant Physiology** **100**: 1464-1470.
- Wingler A (2002) The function of trehalose biosynthesis in plants. **Phytochemistry** **60**: 437–440.
- Xu L, Han L e Huang B (2011) Antioxidant enzyme activities and gene expression patterns in leaves of Kentucky bluegrass in response to drought and post-drought recovery. **J. amer. Soc. hort. Sci.** **136**: 247-255.
- Yancey PH (2001). Water stress, osmolytes and proteins. **Amer. Zool.** **41**: 699–709.
- Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. **The Journal of Experimental Biology** **208**: 2819–2830.
- Yang X, Liang Z, Wen X e Lu C (2008) Genetic engineering of the biosynthesis of glycine betaine leads to increased tolerance of photosynthesis to salt stress in transgenic tobacco plants. **Plant Mol Biol** **66**: 73–86.
- Yang X e Lu C (2005) Photosynthesis is improved by exogenous glycine betaine in salt-stressed maize plants. **Physiol Plant** **124**: 343–352.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Zhai H, Wang F, Si Z, Huo J, Xing L, An Y, He S e Liu Q (2016) A myo-inositol-1-phosphate synthase gene, *lbMIPS1*, enhances salt and drought tolerance and stem nematode resistance in transgenic sweet potato. **Plant Biotechnology Journal** **14**: 592–602.
- Zhang CY, Wang NN, Zhang YH, Feng QZ, Yang CW e Liu B (2013) DNA methylation involved in proline accumulation in response to osmotic stress in rice (*Oryza sativa*). **Genetics and molecular research** **12**: 1269–1277.
- Zhang J, Nguyen HT e Blum A (1999) Genetic analysis of osmotic adjustment in crop plants. **Journal of Experimental Botany** **50**: 291-302.
- Zhang SZ, Yang BP, Feng CL e Tang HL (2005) Genetic Transformation of Tobacco with the Trehalose Synthase Gene from *Grifola frondosa* Fr. Enhances the Resistance to Drought and Salt in Tobacco. **Journal of Integrative Plant Biology** **47**: 579-587.
- Zhao L, Ding Q, Zeng J, Wang F-R, Zhang J, Fan S-J and He X-Q (2012) An improved CTAB-ammonium acetate method for total RNA isolation from cotton. **Phytochemical analysis : PCA** **23**: 647–650.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

CAPÍTULO II

CARACTERIZAÇÃO ESTRUTURAL, FUNCIONAL E DIVERSIDADE DE OSMOPROTETORES NO TRANSCRIPTOMA DE *Stylosanthes scabra* (Vogel) SOB DÉFICIT HÍDRICO

1. INTRODUÇÃO

Déficit hídrico são uma das principais causas na perda de produtividade agrícola (Benko-Iseppon et al. 2011) pois causa grande impacto no desenvolvimento do vegetal podendo resultar até mesmo na sua morte (Mahajan e Tuteja 2005). Estes estresses desencadeiam mudanças físicas, químicas e moleculares que irão funcionar como respostas do vegetal de modo a tolerar tais condições (Benko-Iseppon et al. 2011). Uma destas respostas vegetais a essa condição é o acúmulo de solútos compatíveis, denominados de osmoprotetores que atuam estabilizando moléculas, diminuindo perdas de atividade enzimática e mantendo a estabilidade da membrana celular permitindo a célula manter seu volume e concentração interna de sal (Génard et al. 1991). Esses compostos estão divididos quimicamente em quatro grupos: (a) metilaminas, (b) polióis, (c) açúcares (como a glicina-betaína e a trealose) e (d) aminoácidos e derivados (como a prolina) (Brigotti et al. 2003).

Há vários exemplos de genes envolvidos na resposta a estresses abióticos com enorme potencial biotecnológico (Seki 2003, Avni Öktem et al. 2008). A expressão desses genes tem sido identificada, promovendo a adaptação da planta e sua sobrevivência sob condições de estresse (Tuteja 2009). Plantas transgênicas superexpressando a enzima P5CS (pirroline-5-carboxilase sintetase) demonstraram o acúmulo de prolina durante estresse hídrico, configurando assim, uma resistência ao estresse osmótico (Khan et al. 2015). Estudos recentes mostram que a trealose, desenvolve um papel significativo no crescimento de plantas sobre estresse hídrico (Nuccio et al. 2015), observando-se que a superexpressão de diferentes isoformas da enzima trealose-6-fosfato sintase conferiu resistência à seca em arroz (Li et al. 2011).

Apesar de alguns estudos relacionando osmoprotetores ao déficit hídrico em espécies da família Fabaceae, como a soja (Kido et al. 2013) poucos têm sido conduzidos para espécies nativas com elevada tolerância a condições ambientais adversas, em especial para plantas forrageiras, como a *Stylosanthes scabra*. *Stylosanthes scabra* (Vogel) é uma leguminosa com grande potencial forrageiro, apresenta tolerância à seca e à baixa fertilidade do solo e é nativa do semiárido brasileiro (Ferreira e Costa 1979, Stace e Edye 1984).

Assim, o presente estudo teve como o objetivo a caracterização de osmoprotetores em *S. scabra*, o que permitirá a seleção e validação de genes

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

promissores que podem vir a contribuir no melhoramento genético desta cultura e de outras leguminosas de importância agrícola, auxiliando no futuro desenvolvimento de variedades geneticamente modificadas tolerantes a fatores abióticos, como a seca.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

2. METODOLOGIA

2.1 Material vegetal, imposição do estresse, construção das bibliotecas de RNA-Seq e sequenciamento

O experimento de estresse (supressão de rega) para obtenção dos tecidos vegetais para a geração das bibliotecas de RNA-Seq de *S. scabra* foi realizado em casa de vegetação na Embrapa Semiárido (Petrolina – Brasil), utilizando como matriz o acesso 85/UNEB, proveniente do Banco Ativo de Germoplasma da Universidade do Estado da Bahia. A propagação das plantas foi pelo método de estaquia, sendo as mesmas foram cultivadas em potes de plástico de 5 L em condições de temperatura e umidade controladas. Plantas saudas e com crescimento homogêneo foram selecionadas e divididas em três grupos: (1) Controle – plantas mantidas com disponibilidade de água, e com irrigação uma vez ao dia; (2) Tratamento 1 - plantas com irrigação suspensa por 6 horas (h); (3) Tratamento 2 - plantas com irrigação suspensa por 24 h. Foram utilizadas três réplicas biológicas para cada grupo, sendo cada réplica biológica constituída por três plantas em cada vaso. O tecido radicular foi coletado nos tempos estabelecidos, sendo imediatamente congelado em nitrogênio líquido e armazenado em freezer -80 °C até a extração de RNA.

Realizou-se a extração do RNA total a partir do protocolo CTAB-Acetato (Zhao *et al* 2012) utilizando o tecido radicular e tratamento de DNase e foram analisadas quanto a sua pureza, concentração e integridade em espectrofotômetro NanoDrop 2000c (ThermoFisher Scientific), fluorímetro Qubit® (Life Technologies) e Bioanalyser (Agilent Technologies), respectivamente. Posteriormente, os RNAs purificados foram enviados para a construção das bibliotecas de RNA-Seq e sequenciamento no Centro de Genômica Funcional ESALQ/USP utilizando a plataforma V4 2 X 100 HiSeq2500 (Illumina), *paired end*. Por fim, a montagem do transcriptoma foi feita pelo programa Trinity (Grabherr *et al.* 2011) e a análise estatística dos genes diferencialmente expressos através do programa Edge-R (Robinson *et al* 2010) e DESeq 2 (Love *et al* 2014).

2.2 Identificação e caracterização de transcritos em *S. scabra*

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

A busca pelas sequências-sonda de osmoprotetores se deu inicialmente no banco de dados UniProt, sendo selecionadas sequências com anotação curada, usadas para análise de BLASTp (proteína-proteína) no banco do NCBI de modo a adquirir sondas adicionais de osmoprotetores vegetais. Foi realizada uma busca no Batch CD-search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) para selecionar sondas com os domínios de interesse das famílias de osmoprotetores, sendo estes: AA_Kinase + Aldedh (P5CS), P5CR_Dimer + F420_oxidored (P5CR), ALDH_F10_BADH (BADH), Rieske_RO_Alpha_CMO + RHO_alpha_C_CMO-like (CMO), FIG Superfamily (IMP), NAD_binding_5 superfamily (MIPS), Glyco Transf 20 (TPS) e HAD-like Superfamily (TPP)

A mineração das sequências alvo de osmoprotetores, foi realizada na ferramenta Bioedit, através de um tBLASTn (proteína-nucleotídeo) contra o transcriptoma de *S. scabra*, de modo a obter osmoprotetores candidatos com *e-value* igual ou inferior a e^{-05} . Osmoprotetores candidatos de *S. scabra* foram organizados em formato tabular na ferramenta Galaxy (<https://usegalaxy.org/>), gerando um arquivo multifasta de nucleotídeos com integrantes das quatro famílias de interesse.

As sequências de nucleotídeos foram traduzidas com o auxílio do programa Transdecoder. No entanto, algumas sequências não obtiveram tradução neste programa e foram traduzidas individualmente pela ferramenta ORF-finder (<https://www.ncbi.nlm.nih.gov/orffinder/>), selecionando as ORF's mais completas. As sequências foram traduzidas de nucleotídeo para proteína para posterior busca pelos domínios proteicos de interesse nos osmoprotetores candidatos.

Após a tradução, os candidatos foram submetidos a uma análise de domínios no CD-search e de motivos pela ferramenta MEME (<http://meme-suite.org/tools/meme>) para selecionar apenas aqueles que possuíssem os domínios proteicos referentes às famílias de osmoprotetores. Adicionalmente, os candidatos foram submetidos a análises secundárias, de ponto isoelétrico (para observar suas características físico-químicas), peso molecular a partir do JVirGel 2.0 (Hiller et al. 2006) e a localização subcelular por meio do Cell-PLoc 2.0 (Chou e Shen 2010).

2.3 Ancoragem dos osmoprotetores de *S. scabra* no genoma de *A. duranensis*

A ancoragem dos osmoprotetores candidatos foi realizada no genoma de *Arachis duranensis* de ($2n=20$) uma vez que *S. scabra* não possui genoma conhecido,

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

sendo ambas as espécies muito próximas taxonomicamente, sendo possível inferir sobre a identificação de sua distribuição, posição relativa e abundância nos cromossomos. O genoma foi obtido no site PeanutBase (<https://www.peanutbase.org>) onde foi realizado um BLASTn (nucleotídeo-nucleotídeo) contra as sequências de *S. scabra* com critério utilizado de menor valor de *e-value*.

2.4 Análises fenéticas dos osmoprotetores em *S. scabra*

Foi realizada a análise fenética dos transcritos, separando-os de acordo com suas famílias gênicas, de modo a observar como as sequências irão se agrupar com base na sua similaridade. A análise fenética foi realizada no programa MEGA v.7 (Kumar et al. 2016) pelo método Neighbor-Joining, iniciando com um alinhamento das sequências de cada família na ferramenta ClustalW, utilizando a matriz BLOSSUM, por se tratar de sequências de mesma espécie e testado com Bootstrap (Felsenstein 1985) de 2000 replicações.

Uma análise fenética adicional foi realizada utilizando sequências ortólogas a cada uma das subfamílias, já registradas em banco de dados e buscadas no GenBank e no Phytozome. Também foi realizada no programa MEGA v.7 (Kumar et al. 2016) pelo método Neighbor-Joining, iniciando com um alinhamento das sequências de cada família na ferramenta ClustalW com a exceção do gene TPP no qual foi utilizado o alinhamento pelo MUSCLE, utilizando a matriz PAM, por se tratarem de sequências mais variáveis entre si. A distância genética foi calculada pelo método de calibração Dayhoff (Dayhoff, Schwartz e Orcutt 1978) e testada com Bootstrap (Felsenstein 1985) de 2000 replicações.

2.5 Perfil de expressão *in silico*

Para construção do perfil de expressão *in silico*, foram calculadas as frequências dos transcritos isolados em cada biblioteca, submetendo os dados a uma análise de agrupamento hierárquico por meio do software CLUSTER 3.0 (Eisen et al. 1998). Posteriormente, foram construídos *heatmaps* com base nos valores de modulação (Log₂FC, *Fold Change*) das bibliotecas de RNA-Seq, os quais foram visualizados pelo programa TreeView (Page 1996).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

3. RESULTADOS

3.1 Mineração das sequências candidatas e osmoprotetores em *S. scabra*

Para mineração das sequências candidatas no RNA-Seq de *S. scabra*, foi utilizado um total de 836 sequências sondas (*seed sequences*), de espécies próximas, identificadas nos bancos de dados UniProt e NCBI, contemplando as quatro classes de osmoprotetores (Tabela 1).

Tabela 1: Sequências sondas buscadas nos bancos de dados utilizadas na mineração.

Osmoprotetor	Seed sequences	Variação de tamanho (aa)
Prolina	133	20-752
Glicina-betaína	240	71-624
Mio-Inositol	204	44-537
Trealose	220	305-1441

aa = aminoácidos.

A busca por ortólogos no transcriptoma revelou 46.875 sequências de osmoprotetores candidatos em *S. scabra*. Após a retirada de redundâncias e corte por *e-value* $e^{-0.5}$ foram retornados os seguintes candidatos: 69 para prolina, 592 para glicina-betaína, 130 para Myo-inositol e 278 para trealose. Com posterior tradução das sequências para análise de domínio esse número reduziu, restando 74 sequências com domínio completo, sendo sete candidatos para prolina (com domínios AA_Kinase + Aldedh – P5CS e P5CR_Dimer + F420_oxidored – P5CR), 10 para glicina-betaína (ALDH_F10_BADH – BADH e Rieske_RO_Alpha_CMO + RHO_alpha_C_CMO-like – CMO), 17 para mio-inositol (FIG Superfamily – IMP e NAD_binding_5 superfamily – MIPS) e 42 para trealose (Glyco Transf 20 – TPS, HAD-like Superfamily – TPP), como pode ser observado na Tabela 2.

Tabela 2: Classes de osmoprotetores encontrados no transcriptoma de *S. scabra*.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Osmoprotetor	Subclasse	Sequências candidatas	Sequências com Domínio completo	Tamanho (aa)
Prolina	P5CS	52	02	613-755
	P5CR	17	05	255-319
Glicina-betaina	BADH	511	04	375-527
	CMO	81	06	248-264
Myo-Inositol	IMP	89	13	313-474
	MIPS	41	4	437-463
Trealose	*TPS	199	06	443-567
	TPP	79	25	299-390
	TPS		09	719-939

*TPS corresponde a sequências exclusivas do domínio TPS; aa = aminoácidos.

As sequências candidatas de CMO apresentaram apenas o domínio da região C-terminal RHO_alpha_C_CMO-like completo. Para o grupo TPS da trealose, 09 transcritos apresentaram os domínios TPS e TPP, sendo o primeiro situado na região N-terminal e o segundo na região C-terminal, ressaltando que as sequências Ss-TPSP-04, 05 e 15 apresentaram o domínio TPP HAD-like Superfamily com quebra na sua região C-terminal e 05 transcritos apresentaram apenas o domínio Glyco Transf 20.

Na análise de ponto isoelétrico, peso molecular e localização subcelular observou-se que os candidatos apresentaram variação de pH e localização predominantemente no citoplasma e cloroplasto (Tabela 3). É importante destacar que para a prolina, dois candidatos de P5CR (Ss-P5CR-04 e 05) apresentaram pH ácido enquanto os demais apresentaram pH básicos. Para o Mio-inositol, dois candidatos (Ss-IMP-11 e 12) apresentaram pH básico e os demais indicaram pH ácidos. Para a trealose, o candidato Ss-TPS-01 foi o único que apresentou pH ácido e localização no citoplasma.

Tabela 3: Ponto isoelétrico, peso molecular e localização subcelular dos candidatos à osmoprotetores em *S. scabra*.

Osmoprotetor	Subclasse	p.I. (pH)	M.W. (kDa)	Localização subcelular
Prolina	P5CS	6.0	66.5 a 82	Citoplasma
	P5CR	6.4 a 9.6	28.5 a 33.7	Citoplasma

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Glicina- betaina	BADH	5.8 a 5.5	51.1 a 53.6	Cloroplasto/ Mitocôndria/ Peroxissomo
	CMO	5.1 a 5.5	28.5 a 30.1	Cloroplasto
Mio-inositol	IMP	4.7 a 8.3	33.9 a 51.8	Cloroplasto
	MIPS	5.0 a 5.7	48.4 a 51.4	Citoplasma
Trealose	*TPS	5.5 a 8.6	48.2 a 63.1	Citoplasma/ Cloroplasto/ Vacúolo
	TPP	7.3 a 10.1	33.8 a 44.1	Cloroplasto/ Vacúolo
	TPS	5.7 a 6.8	86.1 a 106.0	Cloroplasto/ Vacúolo

3.2 Análise de motivos

A partir do alinhamento múltiplo das sequências foi possível observar os motivos característicos de cada família de osmoprotetores e suas respectivas subclasses. Para a prolina (P5CS), os candidatos apresentaram os seguintes motivos: Região N-Terminal - Leucina zíper - onde no transcrito Ss-P5CS-01 apresentou uma troca de aminoácidos do S₂₁₈ e A₂₁₉ por AG; Glu-5-Kinase - havendo substituição nas posições V₂₈₂ por N (Ss-P5CS-01) e K (Ss-P5CS-02), A₂₈₄ por S e T₂₉₂ por I (Ss-P5CS-02) e V₂₇₂ por L, H₂₈₆ por C e E₂₈₉ por Y (Ss-P5CS-01), em relação ao motivo ATP-Sítio de ligação, este apenas foi identificado no transcrito Ss-P5CS-01. Na região C-Terminal os seguintes motivos foram identificados: NAD(P)H sítio de ligação - com as substituições I₄₉₁ por N (Ss-P5CS-01) e T (Ss-P5CS-02), N₄₉₇ por T em ambos transcritos e I₄₈₅, H₄₈₇, E₄₉₂, D₄₉₆ e G₅₀₀ por A, Y, S, E e D em Ss-P5CS-01 e também GSA-DH apresentando trocas nos aminoácidos Y₇₀₀ por F em ambos transcritos e S₇₀₁ por C em Ss-P5CS-01 e E₇₂₂ por D em Ss-P5CS-02 (Figura 1).

Em P5CR todos os transcritos apresentaram os motivos conservados característicos GxxGxG (Rossmann fold), com exceção de Ss-P5CR-04 e Ss-P5CR-05 apresentando uma C (cisteína) no lugar da primeira G (glicina), GSxPA (marca registrada do núcleo de dimerização) e o motivo SP-A/G-GTT (os candidatos Ss-P5CR-04 e 05 apresentaram trocas de aminoácidos S pelo T, T pelo C (04), G pelo E e T pelo C (05) (Figura 2).

Para glicina-betaina (BADH), todos os transcritos apresentaram o decapeptídeo conservado VTLELGGKSP, havendo apenas uma troca de aminoácido na posição T₂₈₂ por S, o tripeptídeo SKL na região C-Terminal além do heptapeptídeo

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

QLFIDGE na região N-Terminal. Além disso, apresentaram os seguintes resíduos conservados: PWNNYP, ELGKSP e GQI(V)CSATSR e o aminoácido I (Isoleucina) na posição crítica para o acúmulo de GB em vegetais identificado através do alinhamento de sequências ortólogas da proteína com este aminoácido já identificado na literatura (Figuras 3a e 3b).

Para CMO, os candidatos apresentaram apenas a sequência motivo mononuclear não heme de ligação ao ferro $G/Dx_{3-4}Dx_2Hx_{4-5}H$, característico do domínio RHO_alfa_C_CMO-like e não apresentaram o motivo característico tipo Rieske ($2Fe - 2S$) $CxHx_{15-17}Cx_2H$ do domínio Rieske_RO_Alpha_CMO, pois o mesmo apresentou-se incompleto nas sequências (Figura 4).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

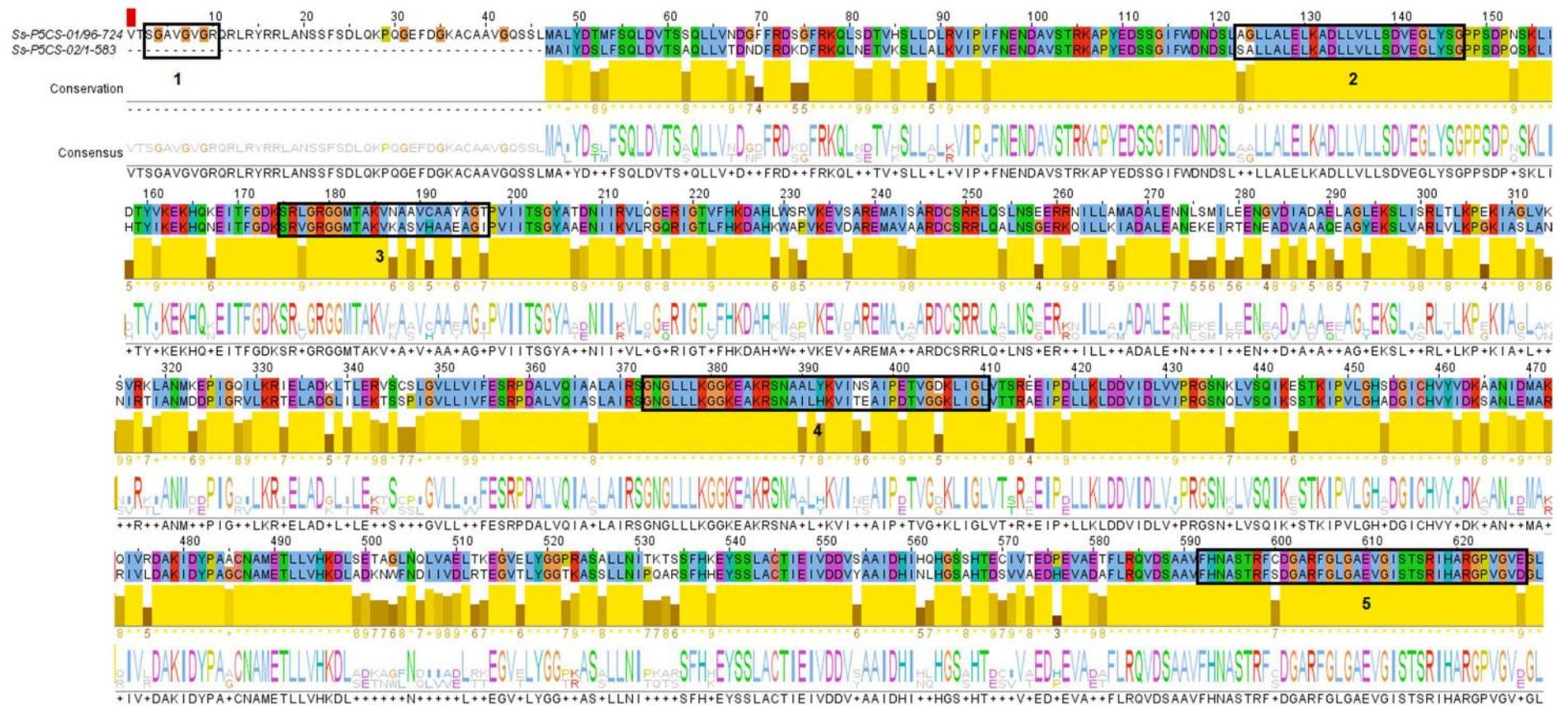


Figura 1. Alinhamento das sequências de P5CS de *S. scabra*. As caixas pretas indicam os motivos conservados: 1- ATP Sítio de Ligação, 2- Leucina Zipper, 3- Glu- 5-Kinase, 4- NAD(P)H Sítio de ligação e 5- GSA-DH.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

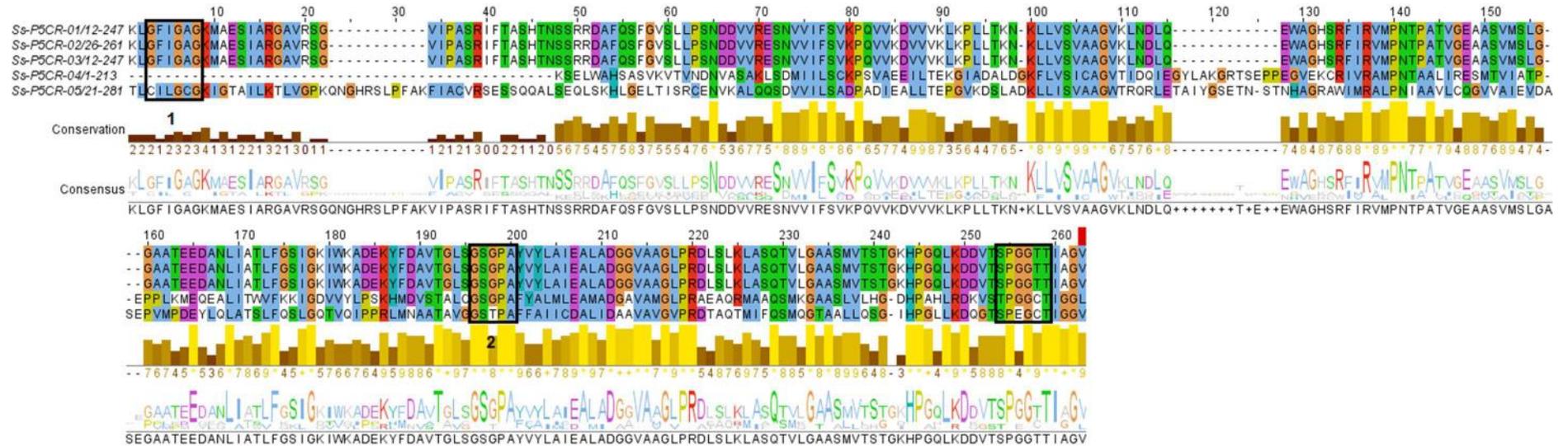


Figura 2. Alinhamento das sequências de P5CR de *S. scabra*. As caixas pretas indicam os motivos conservados: 1- Rossmann fold e 2- marca registrada de dimerização.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

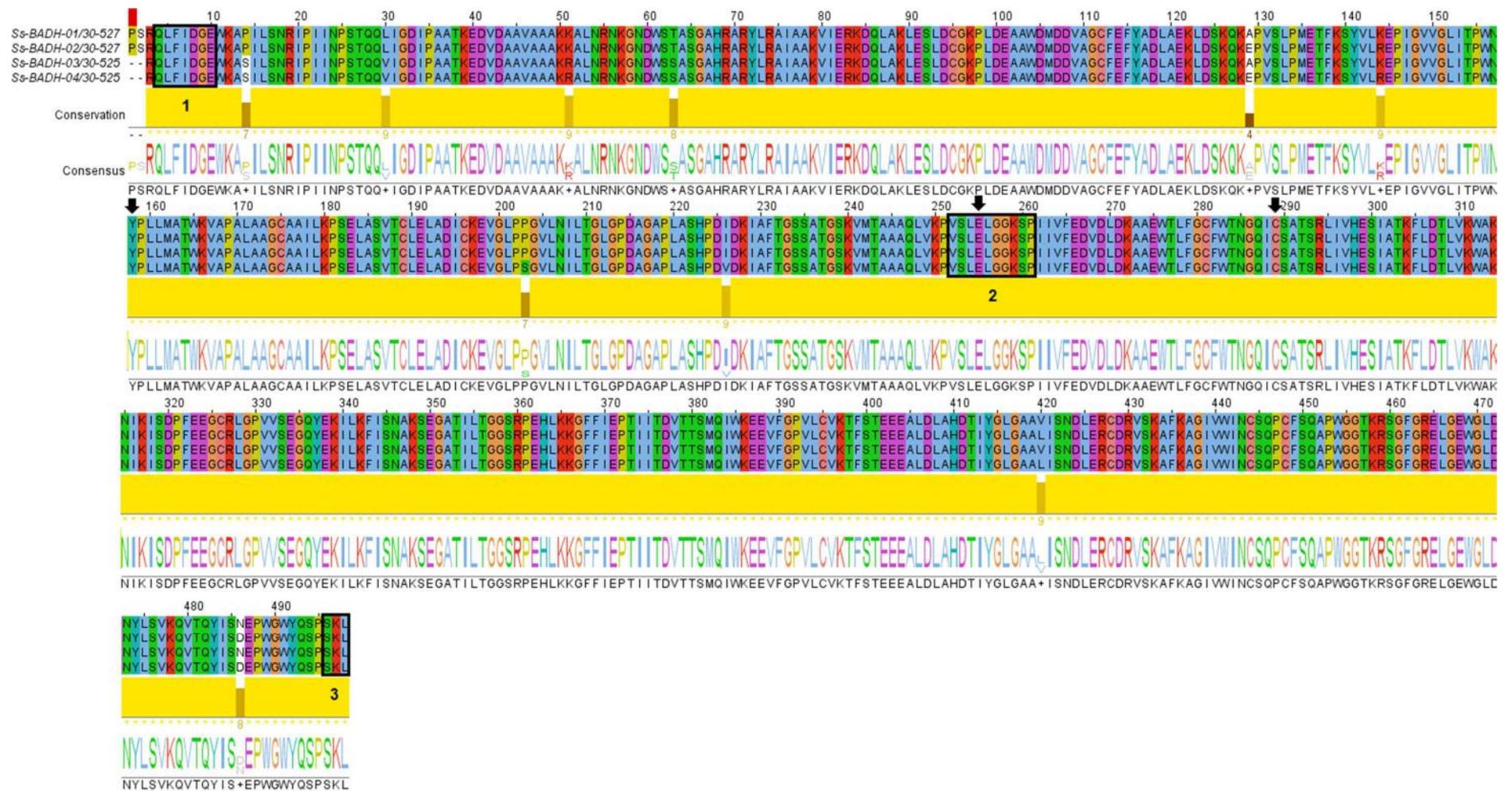


Figura 3a. Alinhamento das sequências de BADH de *S. scabra*. As caixas pretas indicam os motivos: 1- Sinal cloroplasto, 2- decapeptídeo conservado e 3- Sinal peroxissomal. As setas indicam resíduos conservados.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

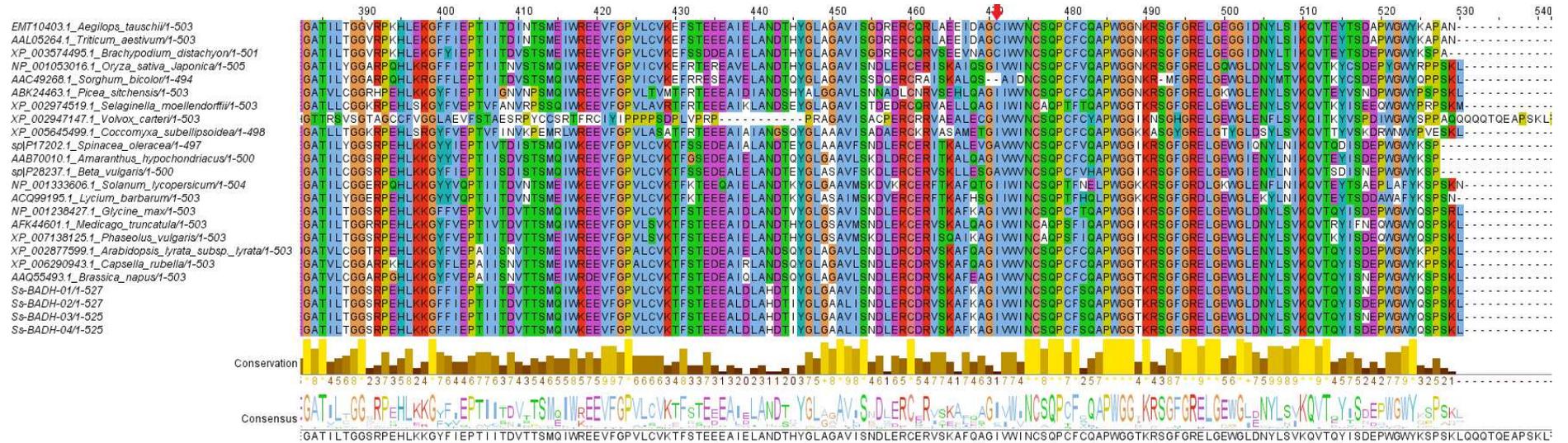


Figura 3b. Alinhamento das sequências ortólogas de BADH utilizadas na fenética para identificação do aminoácido crítico (indicado pela seta vermelha) na produção de GB.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.



Figura 4. Alinhamento das seqüências de CMO de *S. scabra* com identificação do motivo mononuclear não heme de ligação ao ferro.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

Para mio-inositol (MIPS), os quatro candidatos apresentaram os motivos conservados característicos GWGGNNG (havendo troca do aminoácido W₇₁ por L em Ss-MIPS-01 e 02), LWTANTERY (também havendo troca do aminoácido L₂₃₃ por F em Ss-MIPS-01 e 02), NGSPQNTFVPGL (havendo troca nas posições S₂₈₇ e L₂₉₆ por A e C em Ss-MIPS-01 e 02) e SYNHLGNNDG (Figura 5).

Para IMP, foram localizados dois motivos conservados: DPLDGT (havendo troca do aminoácido L₂₃₉ por I em Ss-IMP-03 e 04 e pelo aminoácido V em Ss-IMP-05,06,07,08,09 e 10) não sendo observado nas sequências Ss-IMP-01, 02, 11, 12 e 13 e WDXAAG, porém o motivo característico GEET não foi localizado (Figura 6).

Para trealose, as sequências com apenas domínio *TPS apresentaram todos os resíduos conservados (Figura 7): VANRLP (com exceção de Ss-*TPS-02 e 03), GGL, YYN, LWP, HDYHL, FLHT, GFQ/HT, RLD, SKG, SRG e LRDGMNLVSYE. Os candidatos com apenas o domínio TPP apresentaram todos os motivos Dxxx(T/V), (S/T)Gx e Kx₁₆₋₃₀(G/S)(D/S)xxx(D/N) característicos da família gênica (Figura 8). Já as sequências de TPS apresentaram a maioria dos resíduos conservados característicos do domínio TPS, porém apenas os candidatos Ss-TPS-05 e 06 apresentaram os resíduos ANR₁₁₂LP (Arginina), GG₁₃₄LV (Glicina), DR₃₈₉LD (Arginina) e YE₄₉₆FV (Glutamato) e em contrapartida, estes não apresentaram nenhum motivo característico do domínio TPP: Dxxx(T/V), (S/T)Gx e Kx₁₆₋₃₀(G/S)(D/S)xxx(D/N) - com a exceção do candidato Ss-TPS-05 que apresentou o último motivo - que por sua vez foram identificados nas outras sequências, com a exceção do último que foi encontrado apenas nos candidatos Ss-TPS-01, 02, 07 e 08 (Figura 9).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.



Figura 6. Alinhamento das sequências de IMP de *S. scabra* onde as caixas pretas indicam os motivos conservados.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

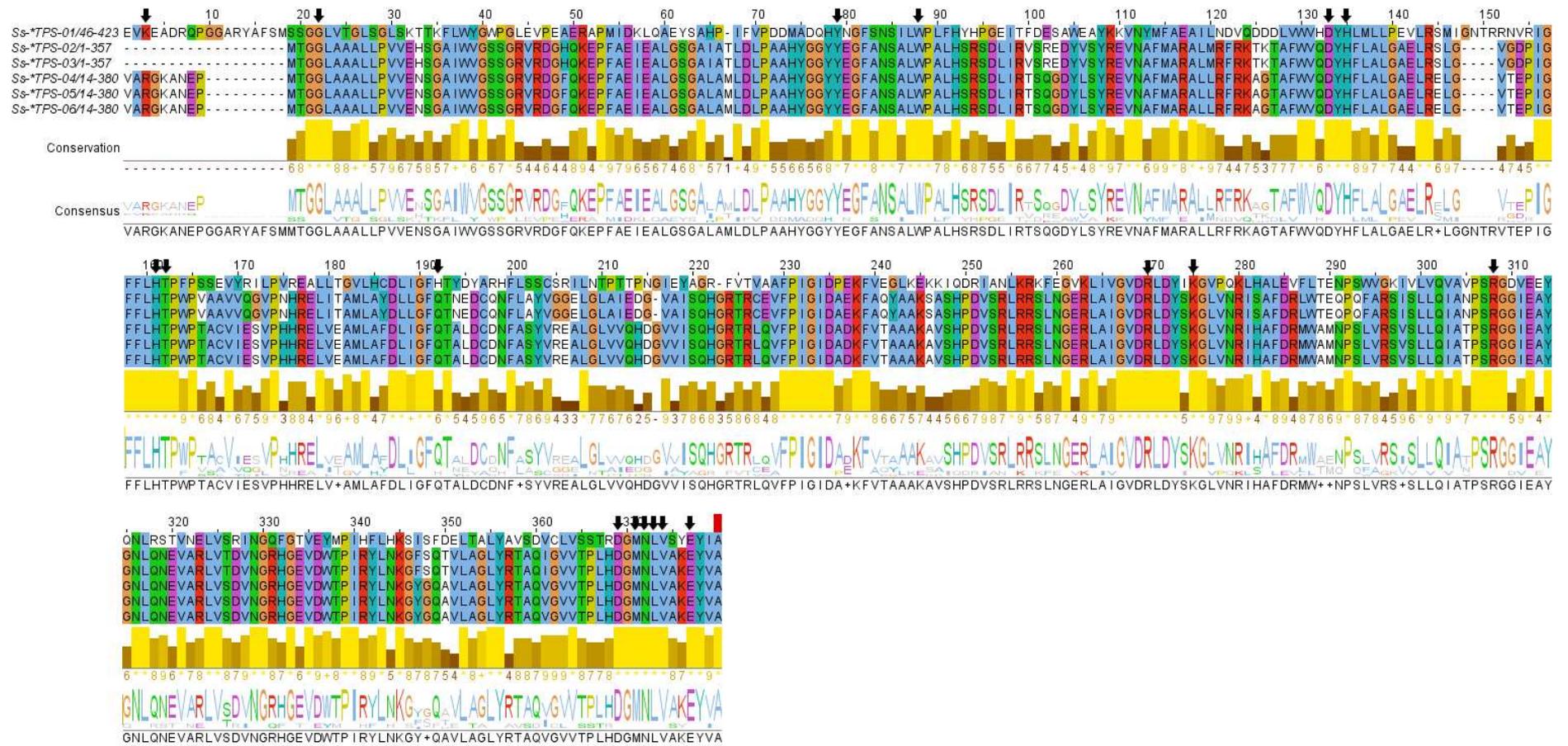


Figura 7. Alinhamento das sequências com apenas domínio *TPS de *S. scabra*. As setas pretas indicam os resíduos conservados destas proteínas.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

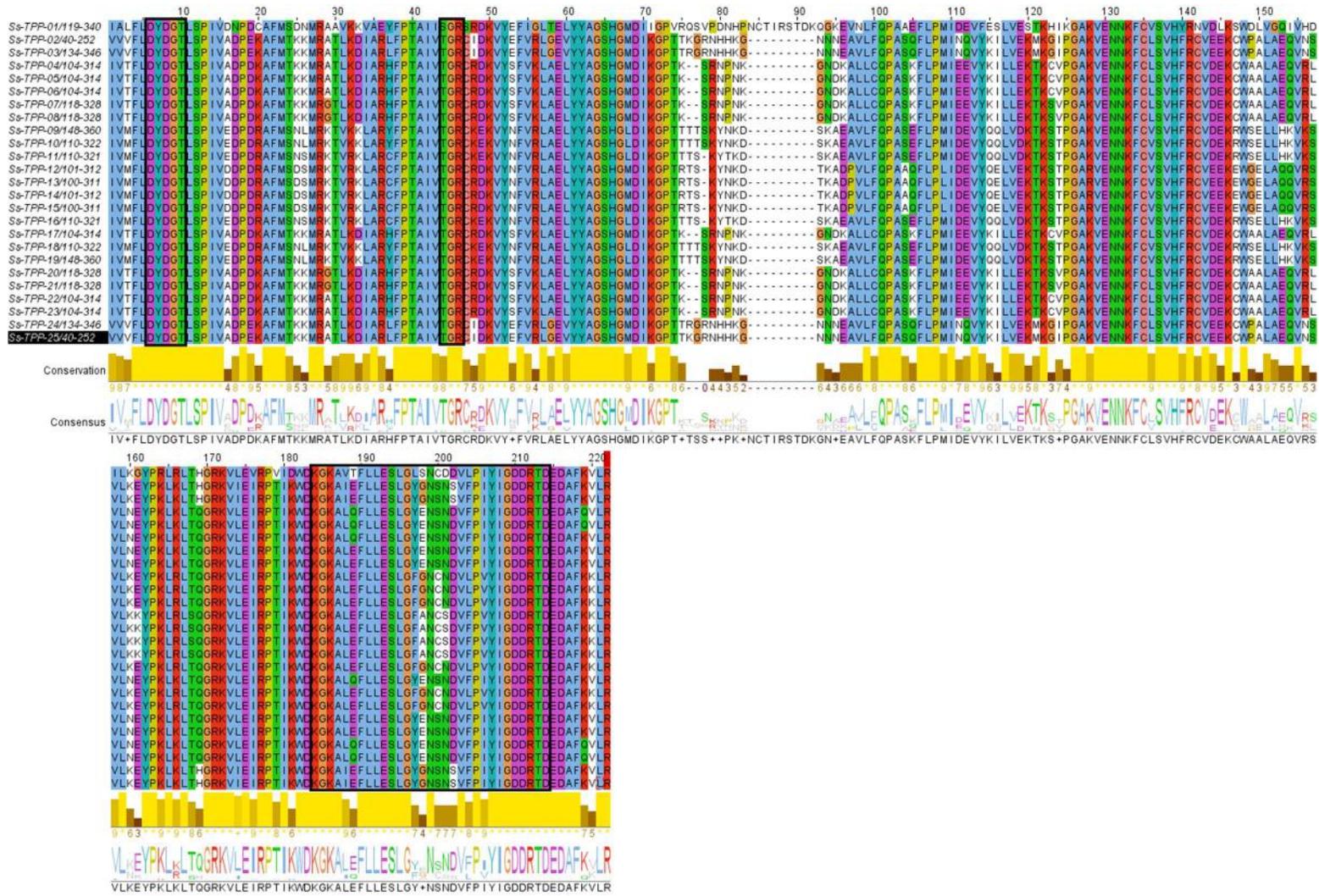
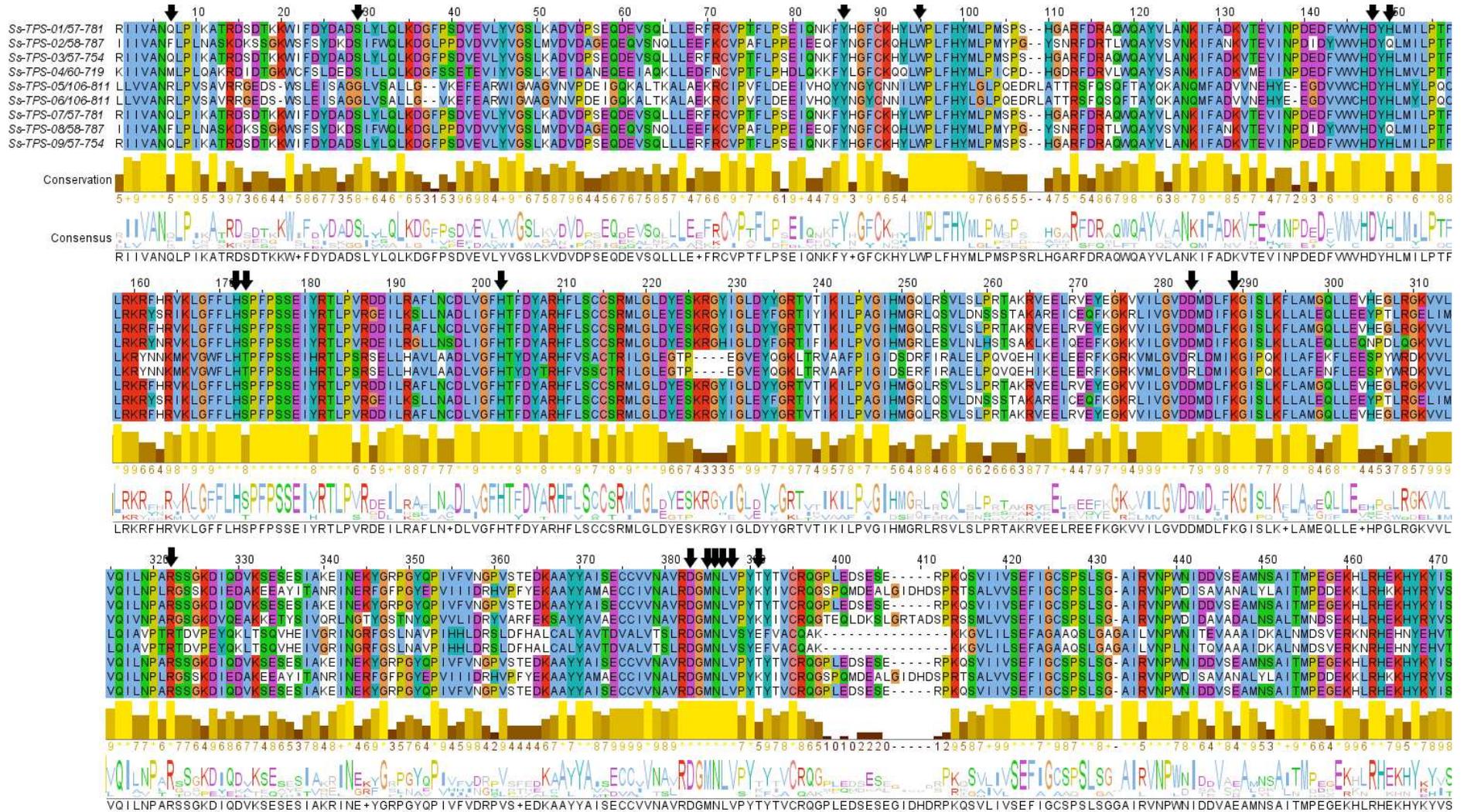


Figura 8. Alinhamento das sequências de TPP de *S. scabra* onde as caixas pretas indicam os motivos conservados.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.



Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

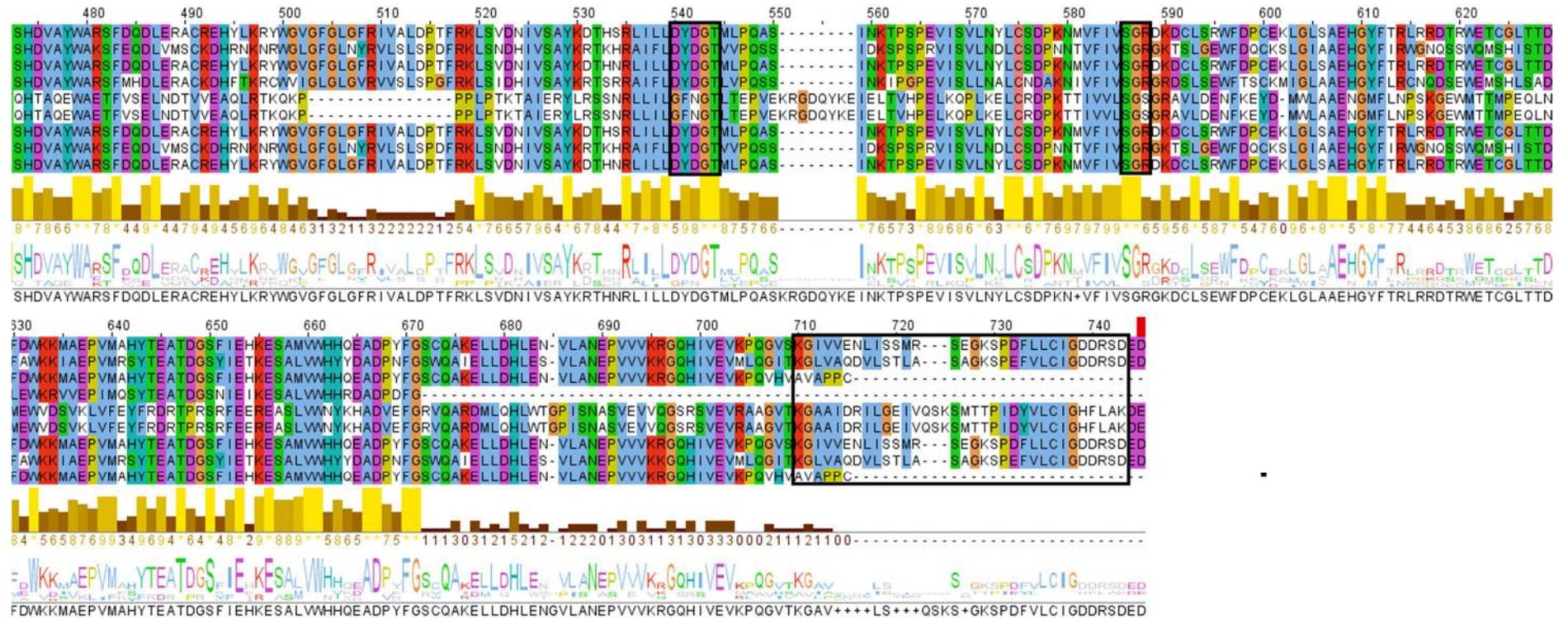


Figura 9. Alinhamento das sequências de TPS de *S. scabra*. As setas indicam os resíduos conservados presentes no domínio TPS e as caixas pretas indicam os motivos conservados presentes no domínio TPP.

3.3 Análise de expressão gênica *in silico*

O perfil de expressão *in silico* foi gerado por meio da utilização dos dados de expressão obtidos nas duas bibliotecas (6 e 24 h) mediante clusterização hierárquica (Figura 10). De forma geral, os osmoprotetores candidatos apresentaram expressão diferencial (induzido ou reprimido) nos dois tempos de estresse hídrico, com poucos candidatos não significativos.

Foi possível observar a formação de três *clusters* bem definidos, onde o primeiro *cluster* apresenta uma repressão mais acentuada no tempo de 6 h e à medida que o tempo de estresse aumenta para 24 h, a repressão vai diminuindo. O segundo *cluster* apresenta uma leve indução com de 6 h de estresse, e com 24 h, essa indução fica mais evidente. O transcrito Ss-*TPS-01 destaca-se nesse *cluster* pois se apresenta reprimido no tempo de 6 h e induzido com 24 h. O terceiro *cluster* apresenta um padrão oposto ao primeiro *cluster*, com uma repressão menos acentuada no tempo de 6 h e que se torna mais evidente com 24 h.

Em relação à expressão observada no tempo de 6 h, nota-se que a quantidade de transcritos induzidos e reprimidos é pequena onde apenas o transcrito Ss-*TPS-06 apresenta uma expressão significativa entre $\log_2FC=2,00$ e $\log_2FC=4,00$. Por outro lado, entre os transcritos que apresentaram repressão, destaca-se o transcrito Ss-IMP-09, que apresentou um alto nível de repressão ($\log_2FC=6,00$). Já no tempo de 24 h, houve uma maior quantidade de transcritos induzidos e reprimidos, embora o número de transcritos reprimidos seja maior. Os transcritos Ss-TPS-01 e Ss-IMP-07 apresentaram o maior nível de indução e repressão, respectivamente.

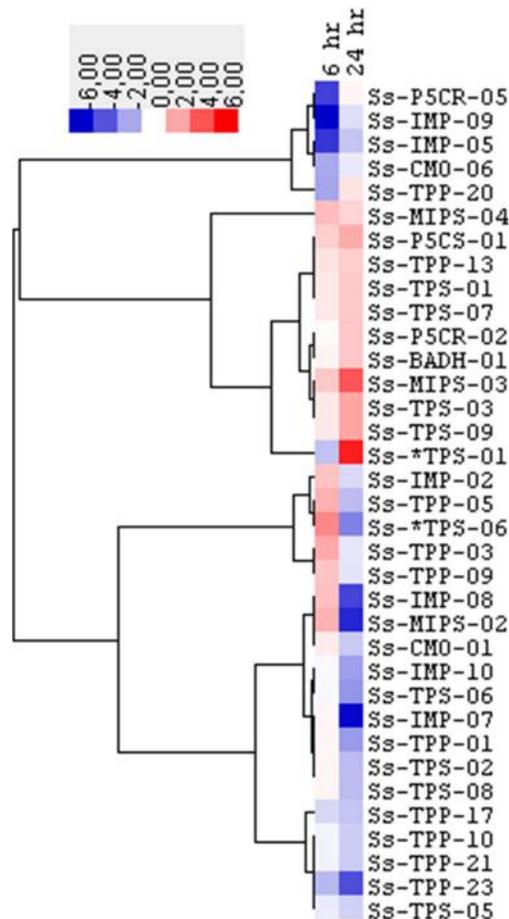


Figura 10: *Heatmap* construindo a partir da contagem de reads dos osmoprotetores em duas bibliotecas de RNA-Seq sob condição de estresse hídrico nos tempos de 6 e 24 h. A cor branca indica a ausência de expressão (transcrito não modulado), tons de azul indicam os níveis de repressão e tons de vermelho indicam os níveis de indução (vide escala de cores). Quanto maior a intensidade da cor (azul ou vermelha), maior o nível de repressão/indução do gene analisado.

3.4 Ancoragem de sequências de *S. scabra* no genoma de *A. duranensis*:

Os transcritos de *S. scabra* apresentaram no genoma de *A. duranensis* (n=10) um total de 19 loci ancorados (Figura 11) agrupados em sua maioria nas regiões terminais, com os cromossomos 1 e 3 apresentando 28 transcritos ancorados. Todos os candidatos de glicina-betaína ancoraram nos cromossomos 1. Por sua vez, os transcritos de trealose ancoraram em sete cromossomos distintos (1, 3, 5, 6, 7, 8 e 10). Foi realizada uma segunda ancoragem no genoma de *Arachis ipaensis* (dados não mostrados), onde foi possível observar que apesar de se localizarem em posições distintas, os genes de Prolina, Mio inositol e Glicina betaína de *S. scabra* apresentaram uma conservação, não observada em Trealose, se distribuindo

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

exatamente nos mesmos cromossomos que *A. duranensis*, indicando a existência de uma sintenia para estes genes entre as espécies. Devido a aproximação taxonômica destas com *S. scabra*, provavelmente esta conservação será mantida. Os candidatos como domínio TPS exclusivo, Ss-P5CR-04 e 05 e Ss-MIPS-01 e 02 não apresentaram resultados na análise.

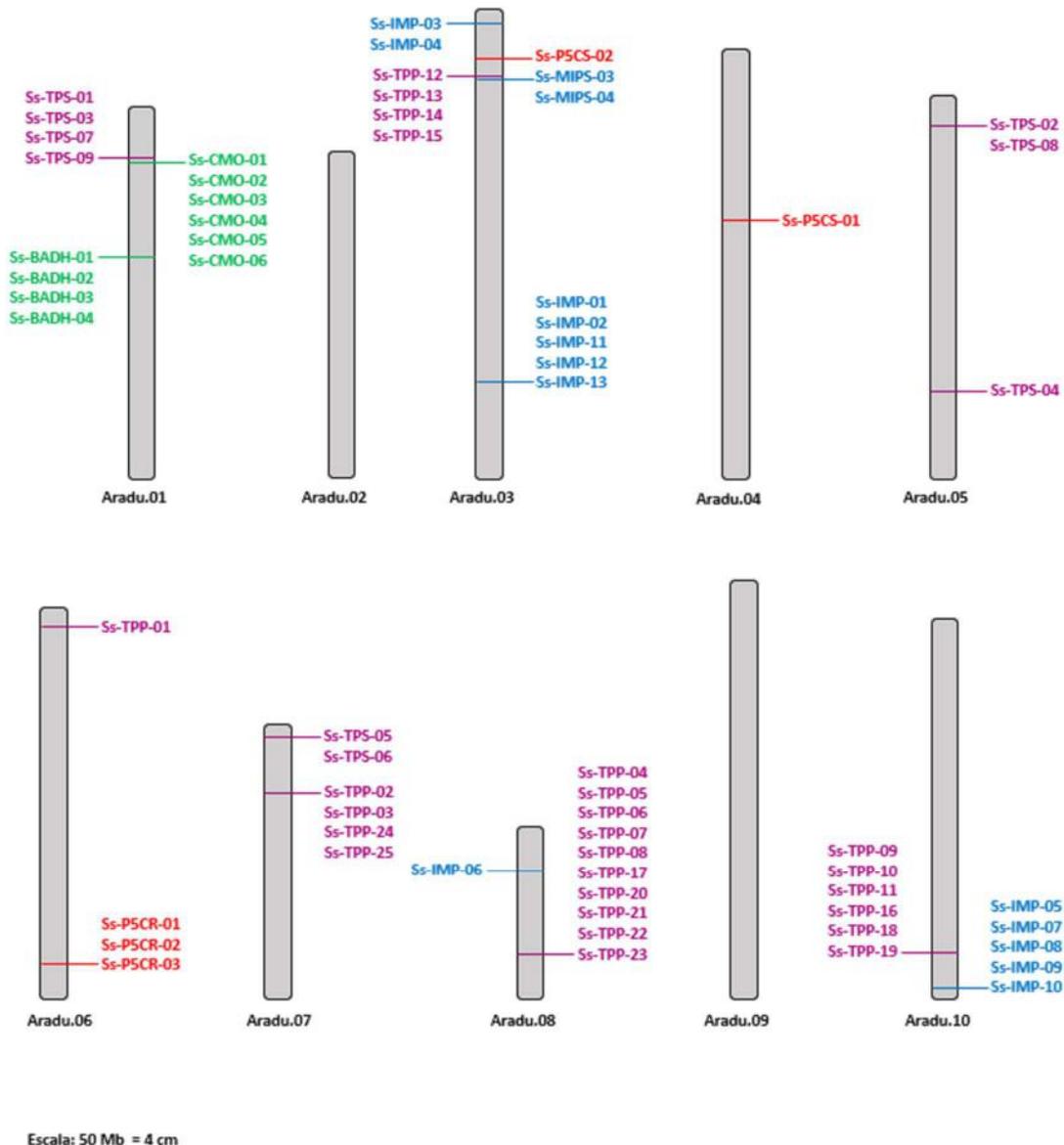


Figura 11: Ancoragem dos transcritos de *S. scabra* em cromossomos virtuais de *A. duranensis* (n=10). Em lilás = Trealose; Verde = Glicina betaína; Azul = Mio inositol; Vermelho = Prolina. Sequências ancoradas na mesma posição indica isoformas.

3.5 Análises fenéticas:

As árvores fenéticas (Figura 12) foram construídas a partir de 74 sequências proteicas candidatas a osmoprotetores em *S. scabra*, sendo 40 de Trealose (TPS e TPP), 17 de Mio inositol (MIPS e IMP), 10 de Glicina betaína (BADH e CMO) e sete de Prolina (P5CS e P5CR). Foi possível observar que em cada família proteica as sequências dividiram-se em dois grandes grupos, em alguns casos com alto valor de confiabilidade (valores de bootstrap >95) numerados em algarismos romanos I e II, separando cada uma em suas devidas subfamílias. No caso da Trealose, compreendeu uma exceção, com a formação de três grupos, onde as sequências que continham apenas o domínio TPS (representadas por um asterisco - *TPS), identificadas com um asterisco, formaram um grupo separado. Também pôde ser notado que as sequências Ss-TPS-05 e 06, que possuem ambos domínios TPS e TPP, agruparam-se juntamente com as sequências do grupo III, o que se deve provavelmente ao fato dessas sequências possuírem maior similaridade no domínio TPS com as sequências do grupo III do que com as do grupo II, que também são proteínas com ambos domínios (TPS e TPP). Esta semelhança pode ser observada na ausência do motivo 10 mostrado no dendrograma das sequências de TPS.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

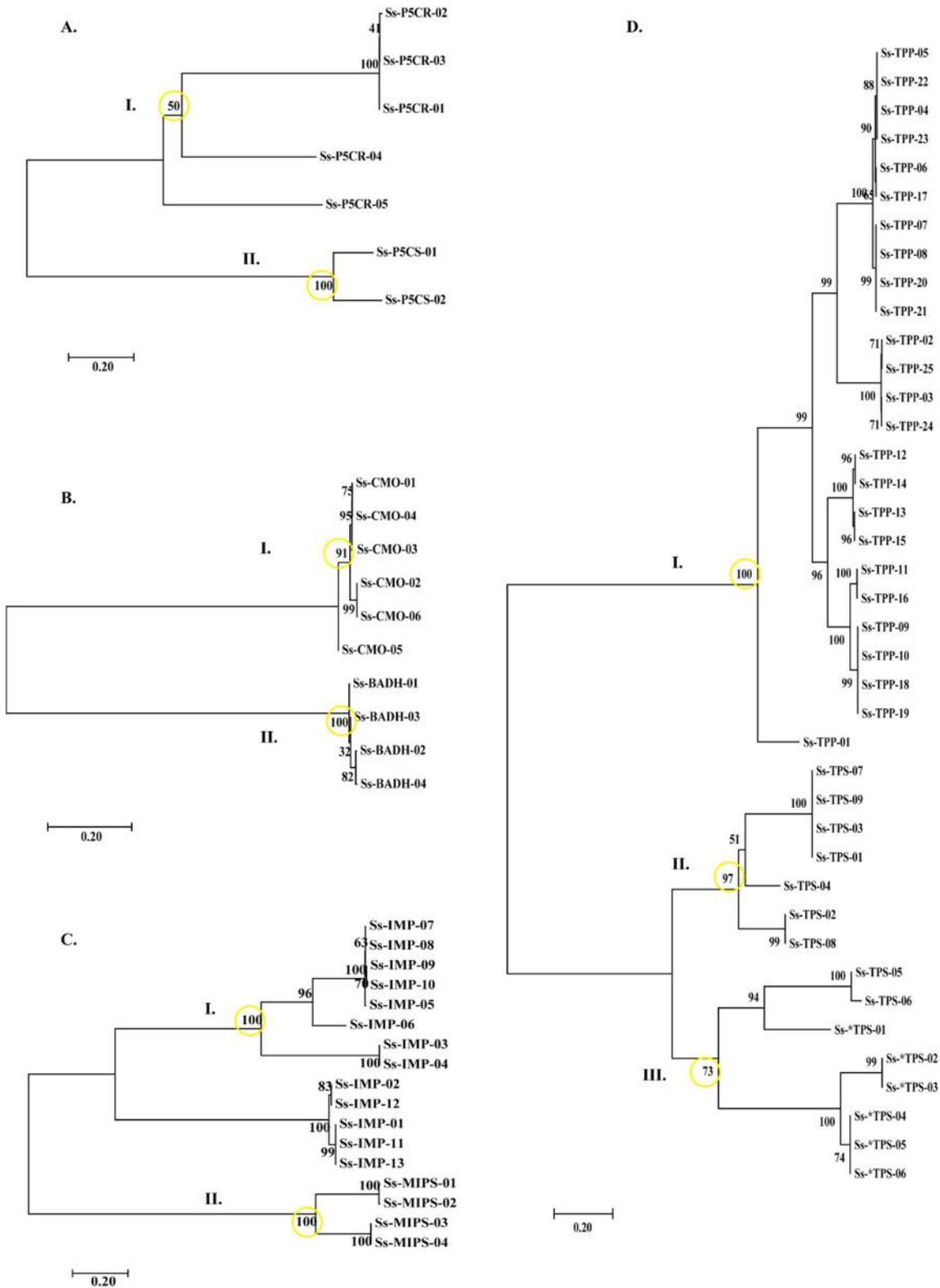


Figura 12. Árvores fenéticas construída usando o método de *Neighbor-Joining* para inferência de diversidade de osmoprotetores candidatos em *S. scabra*. Os números na base dos ramos (alguns circulosados em amarelo) referem-se aos valores de bootstrap (2.000 repetições). A) Prolina, B) Glicina betaina, C) Mio inositol e D) Trealose.

Para análise de classificação dos genes de osmoprotetores foram utilizadas sequências descritas e disponíveis na literatura de diferentes espécies. A relação entre os ortólogos em Prolina, Mio inositol e Glicina betaina foram refletidas em grupamentos conforme a taxonomia dos grupos em monocotiledôneas e dicotiledôneas. Em P5CS e P5CR (Figuras 13a e 13b) o ramo principal das dicotiledôneas é iniciado pela sequência de *Glycine max* enquanto o das monocotiledôneas pela sequência de *Musa acuminata*, com a exceção das sequências Ss-P5CR-04 e 05 que formaram um grupo separado provavelmente devido ao fato de as mesmas apresentarem apenas dois dos motivos identificados no MEME, demonstrando assim a baixa similaridade das mesmas em relação as outras sequências do grupo de dicotiledôneas. Em MIPS (Figura 14a) o ramo das dicotiledôneas é iniciado pelas sequências de *S. scabra* enquanto o ramo das monocotiledôneas é encabeçado pela sequência de *Zea mays*. Porém, Ss-MIPS-01 e 02 separaram-se do grupo e formaram um grupo separado, como pode ser observado, pela ausência do Motivo 9 que está presente nas outras sequências. Já em IMP (Figura 14b) as sequências formaram três grupos com baixo nível de confiabilidade devido aos valores baixos de bootstrap (indicados pelos círculos amarelos) sendo um grupo de monocotiledôneas e dois grandes grupos de dicotiledôneas, onde cinco das 13 sequências de *S. scabra* se agruparam. Como ocorrido em P5CR e MIPS, oito sequências de *S. scabra* se separaram das demais e formaram um grupo a parte devido à baixa similaridade destas sequências em relação às demais como comprovado pela ausência de motivos. Em CMO (Figura 15a) houve separação das sequências taxonomicamente em quatro ramos distintos, com dois de dicotiledôneas, um de monocotiledôneas, sendo a sequência de *Selaginella moellendorffii* a única de gimnosperma.

Em BADH (Figura 15b) a alta similaridade das sequências notada pela análise dos motivos, distinguiu os grupos taxonômicos em 3 grandes grupos de monocotiledôneas, dicotiledôneas e de espécies de gimnospermas. Em Trealose (Figuras 16a e 16b), diferentemente das outras famílias, a separação das sequências não seguiu um padrão taxonômico, mas formou dois grandes grupos (Classe 1 e Classe 2) para TPS e (Grupo A e Grupo B) para TPP de acordo com descrição na literatura. Entretanto em TPP pôde ser observada a formação de um grupamento a parte, abrigando três sequências de *Glycine max*, quatro de *Zea mays* e duas de *Oryza sativa* que – em comparação com as outras sequências presentes no grupo –

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

pode-se notar a ausência de alguns motivos identificados, o que pode justificar sua menor similaridade comparativamente aos demais grupos formados.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

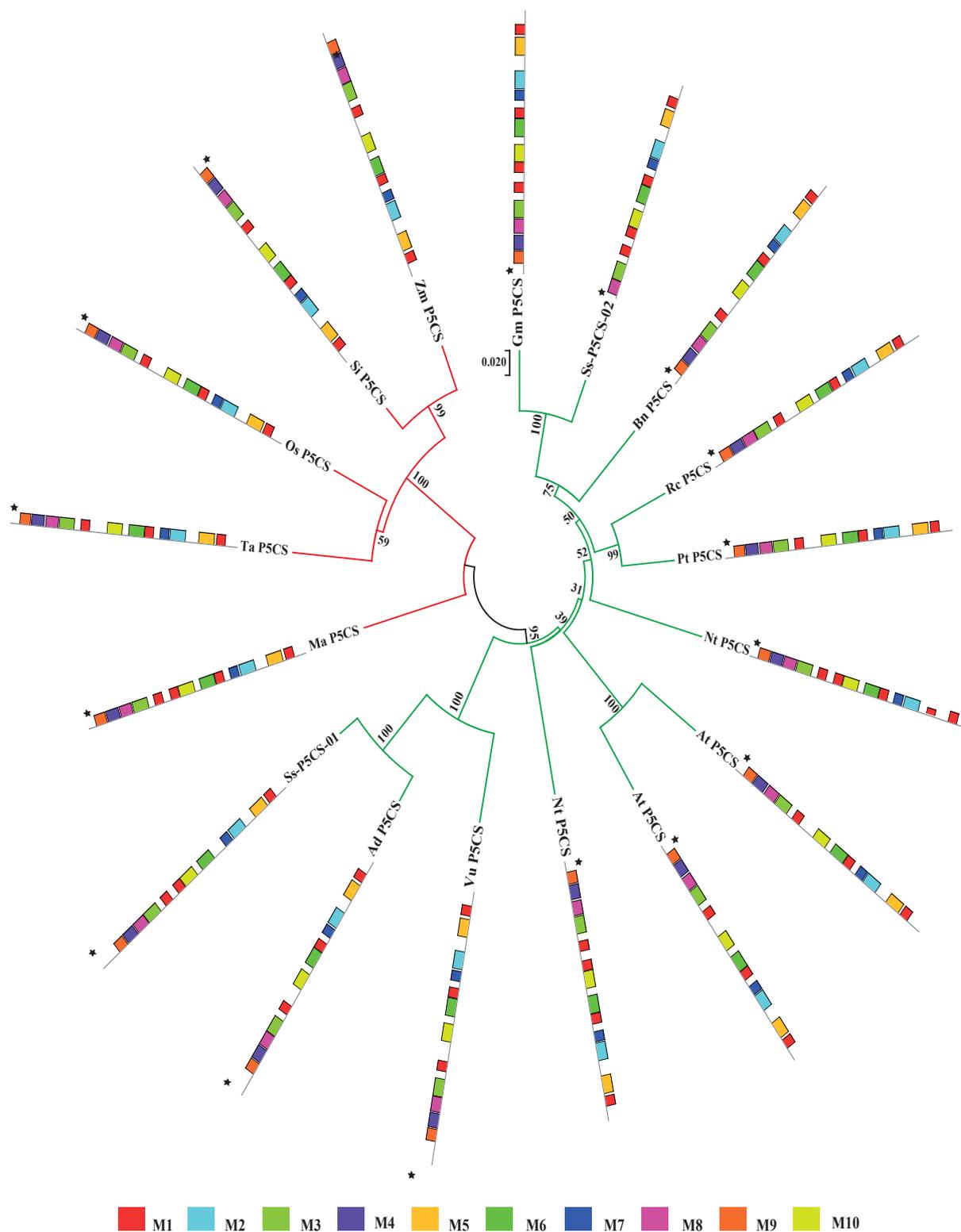


Figura 13a. Dendrograma mostrando as relações entre ortólogos aos candidatos de P5CS pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas e nos ramos verdes espécies dicotiledôneas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Gm (*Glycine max*); Ss (*Stylosanthes scabra*); Bn (*Boehmeria nivea*); Rc (*Ricinus communis*); Pt (*Populus trichocarpa*); Nt (*Nitraria tangutorum*); At (*Arabidopsis thaliana*); Nt (*Nicotiana tabacum*); Vu (*Vigna unguiculata*); Ad (*Arachis duranensis*); Ma (*Musa acuminata*); Ta (*Triticum aestivum*); Os (*Oryza sativa*); Si (*Setaria italica*); Zm (*Zea mays*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

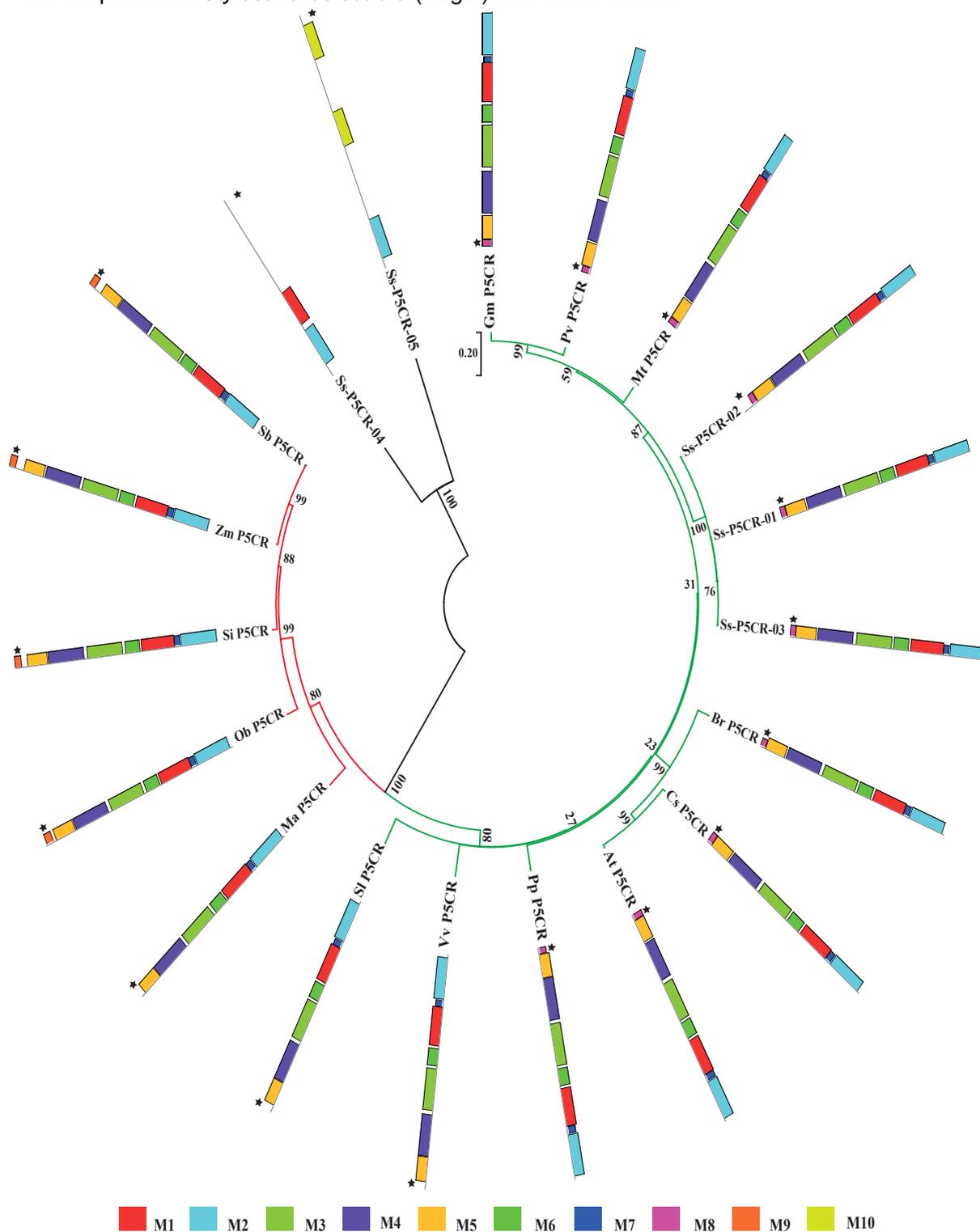


Figura 13b. Dendrograma mostrando as relações entre ortólogos aos candidatos de P5CR pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas e nos ramos verdes espécies dicotiledôneas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); Pv (*Phaseolus vulgaris*); Mt (*Medicago truncatula*); Br (*Brassica rapa*); Cs (*Camelina sativa*); At (*Arabidopsis thaliana*); Pp (*Prunus persica*); Vv (*Vitis vinifera*); Sl (*Solanum lycopersicum*); Ma (*Musa acuminata*); Ob (*Oryza brachyantha*); Si (*Setaria italica*); Zm (*Zea mays*); Sb (*Sorghum bicolor*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

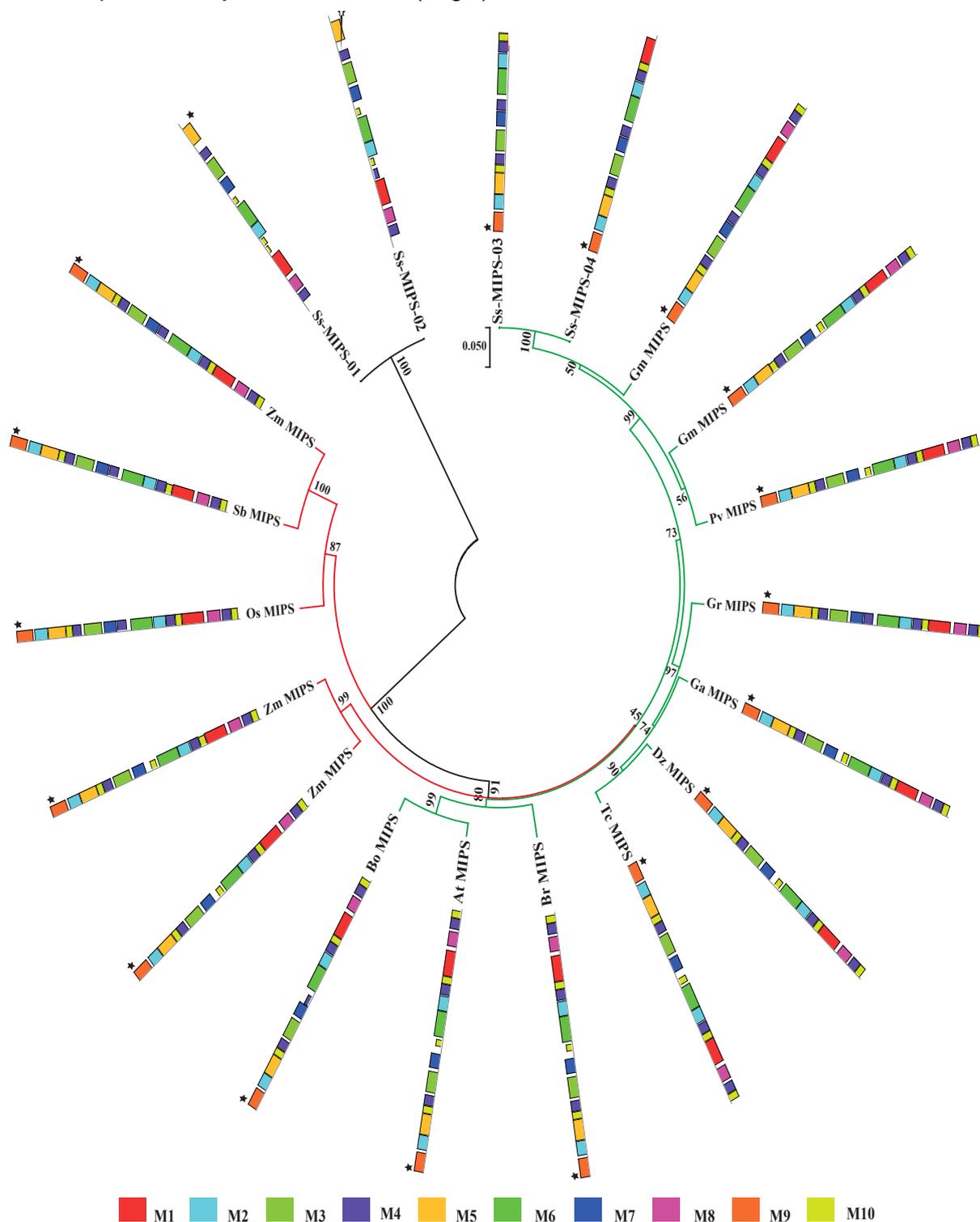


Figura 14a. Dendrograma mostrando as relações entre ortólogos aos candidatos de MIPS pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas e nos ramos verdes espécies dicotiledôneas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); Pv (*Phaseolus vulgaris*); Gr (*Gossypium raimondii*); Ga (*Gossypium arboreum*); Dz (*Durio zibethinus*); Tc (*Theobroma cacao*); Br (*Brassica rapa*); At (*Arabidopsis thaliana*); Bo (*Brassica oleracea*); Zm (*Zea mays*); Os (*Oryza sativa*); Sb (*Sorghum bicolor*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

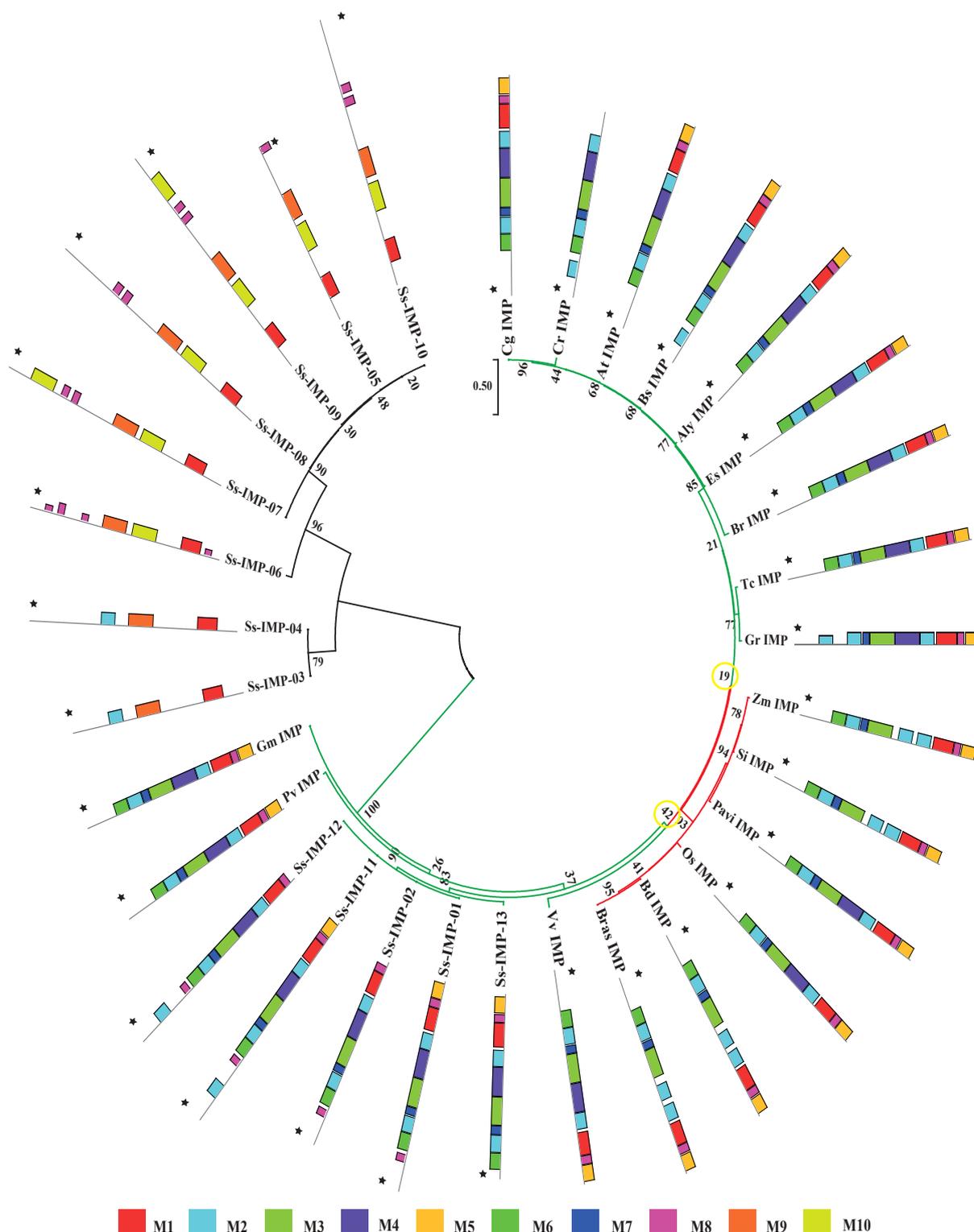


Figura 14b. Dendrograma mostrando as relações entre ortólogos aos candidatos de IMP pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas e nos ramos verdes espécies dicotiledôneas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Cg (*Capsella grandiflora*); Cr (*Capsella rubella*); At (*Arabidopsis thaliana*); Bs (*Boechera stricta*); Aly (*Arabidopsis lyrata*); Es (*Eutrema salsugineum*); Br (*Brassica rapa*); Tc (*Theobroma cacao*); Gr (*Gossypium raimondii*); Zm (*Zea mays*); Si (*Setaria itálica*); Pv (*Panicum virgatum*); Os (*Oryza sativa*); Bd (*Brachypodium distachyon*); Bs (*Brachypodium stacei*); Vv (*Vitis vinífera*); Pv (*Phaseolus vulgaris*); Gm (*Glycine max*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

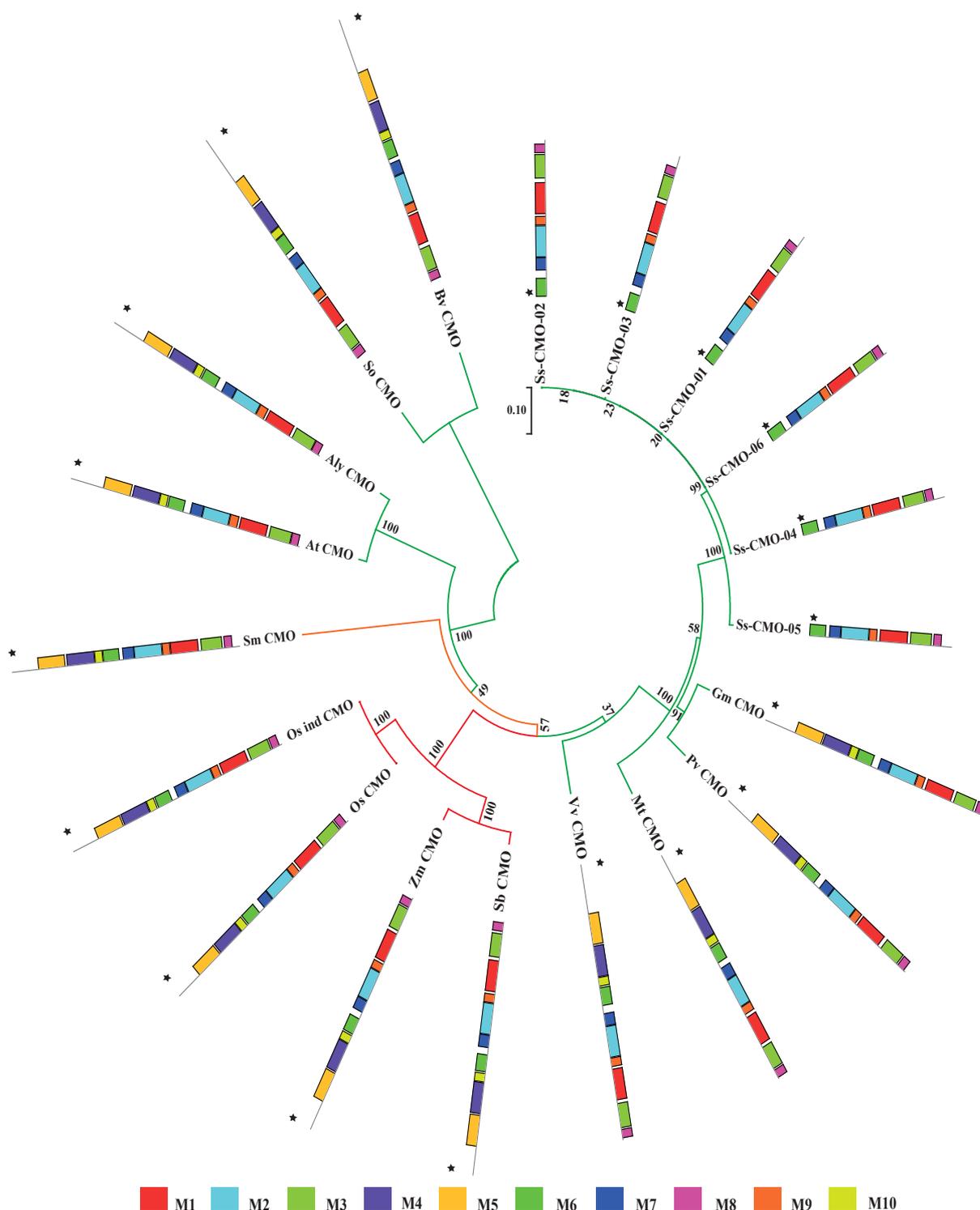


Figura 15a. Dendrograma mostrando as relações entre ortólogos aos candidatos de CMO pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas, nos ramos verdes espécies dicotiledôneas e nos ramos laranjas espécies gimnospermas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); Pv (*Phaseolus vulgaris*); Mt (*Medicago truncatula*); Vv (*Vitis vinífera*); Sb (*Sorghum bicolor*); Zm (*Zea mays*); Os (*Oryza sativa*); Os ind (*Oryza sativa Indica*); Sm (*Selaginella moellendorffii*); At (*Arabidopsis thaliana*); Aly (*Arabidopsis lyrata*); So (*Spinacia oleracea*); Bv (*Beta vulgaris*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

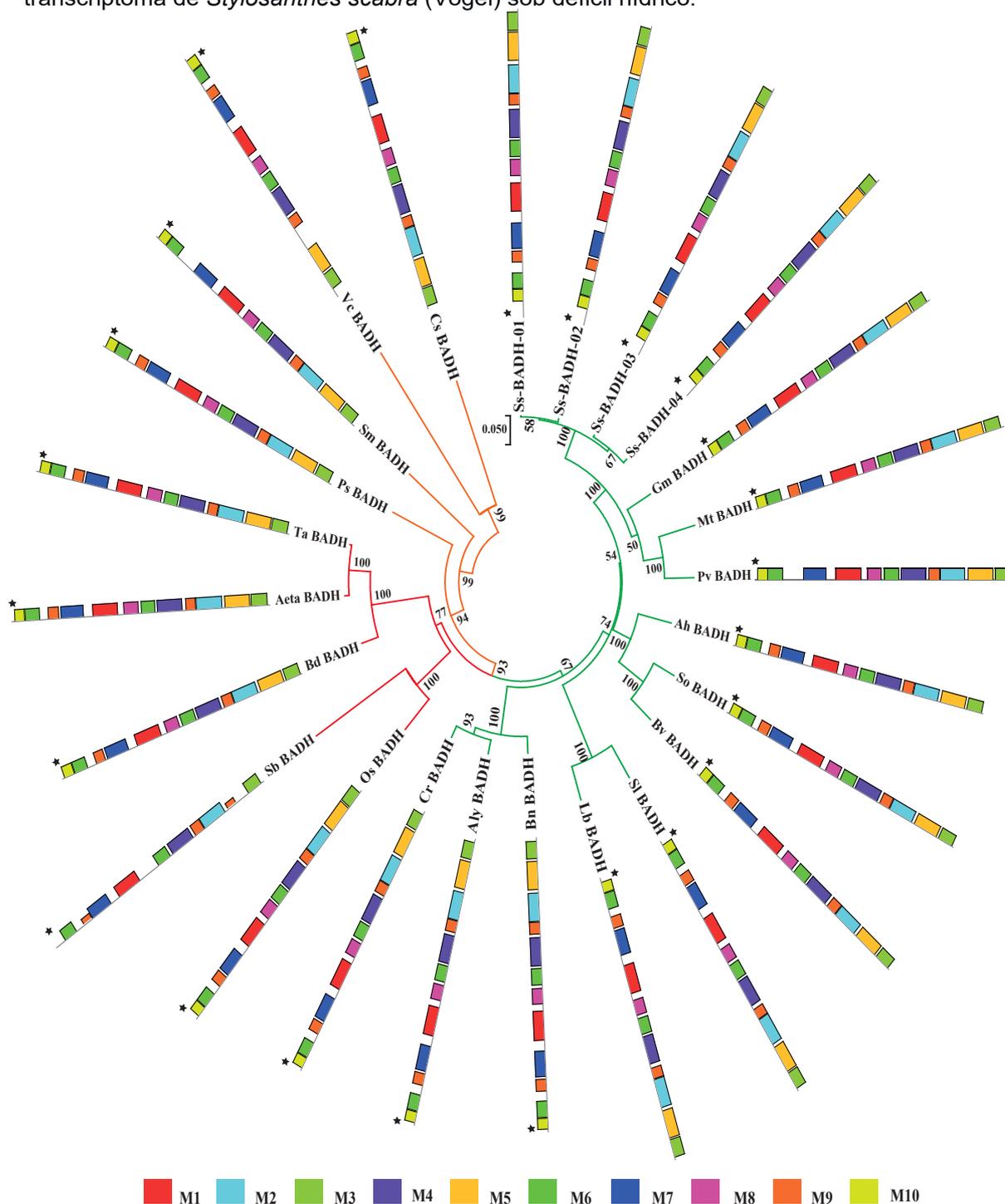


Figura 15b. Dendrograma mostrando as relações entre ortólogos aos candidatos de BADH pelo método neighbor joining onde nos ramos vermelhos estão presentes espécies monocotiledôneas, nos ramos verdes espécies dicotiledôneas e nos ramos laranjas espécies gimnospermas. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); Mt (*Medicago truncatula*); Pv (*Phaseolus vulgaris*); Ah (*Amaranthus hypochondriacus*); So (*Spinacea oleracea*); Bv (*Beta vulgaris*); Sl (*Solanum lycopersicum*); Lb (*Lycium barbarum*); Bn (*Brassica napus*); Aly (*Arabidopsis lyrata*); Cr (*Capsella rubella*); Os (*Oryza sativa*); Sb (*Sorghum bicolor*); Bd (*Brachypodium distachyon*); Aeta (*Aegilops tauschii*); Ta (*Triticum aestivum*); Ps (*Picea sitchensis*); Sm (*Selaginella moellendorffii*); Vc (*Volvox carteri*); Cs (*Coccomyxa subellipsoidea*). Posição da estrela indica início da sequência. M= Motivo.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

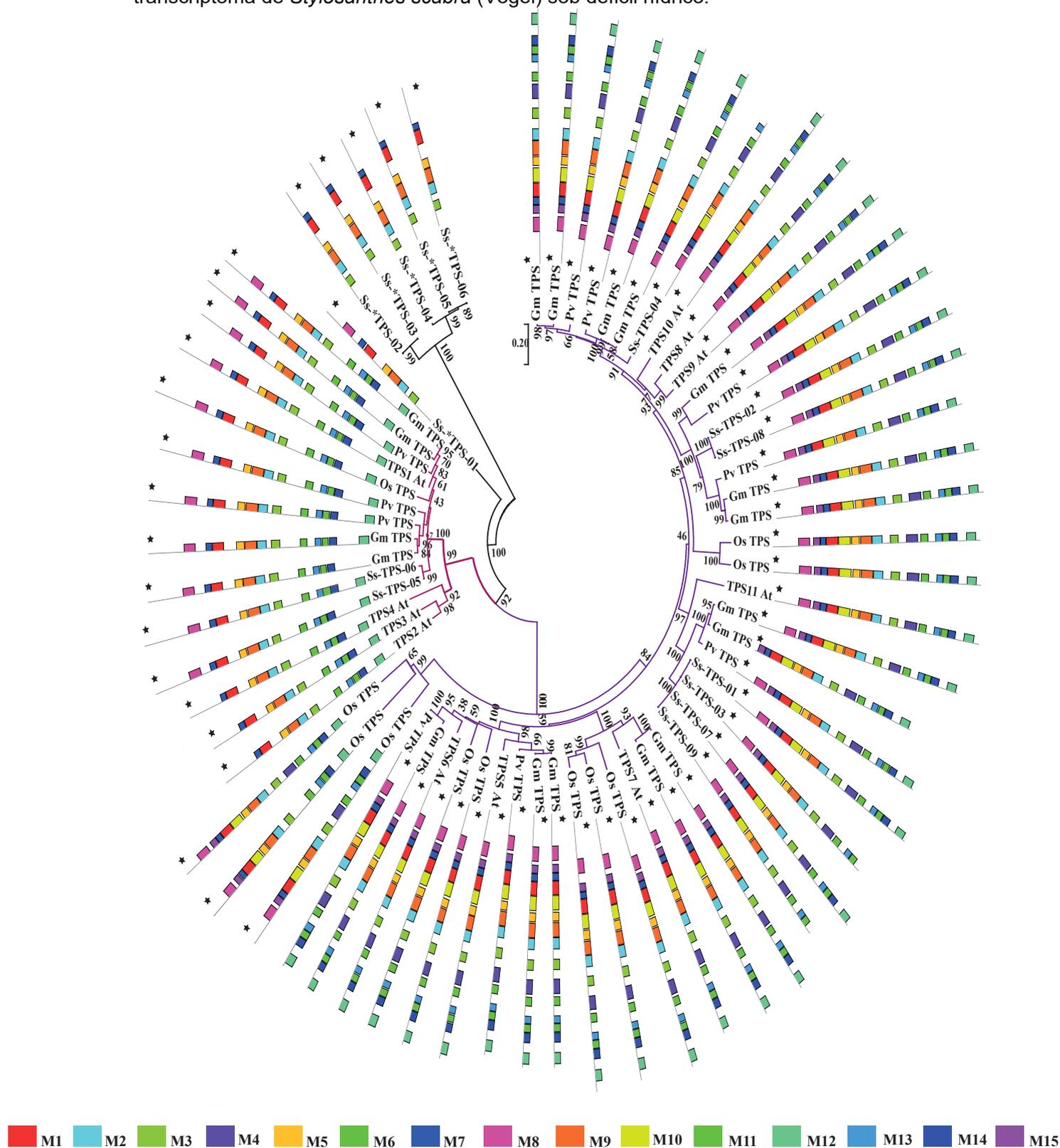


Figura 16a. Dendrograma mostrando as relações entre ortólogos aos candidatos de TPS pelo método neighbor joining onde nos ramos rosas estão presentes as sequências da Classe 1 e nos ramos roxos as sequências da Classe 2. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); At (*Arabidopsis thaliana*); Pv (*Phaseolus vulgaris*); Os (*Oryza sativa*). Posição da estrela indica início da sequência. (M4, M6, M11, M12, M13, M14) motivos domínio TPP e (M1, M2, M3, M5, M7, M8, M9, M10, M15) motivos domínio TPS.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

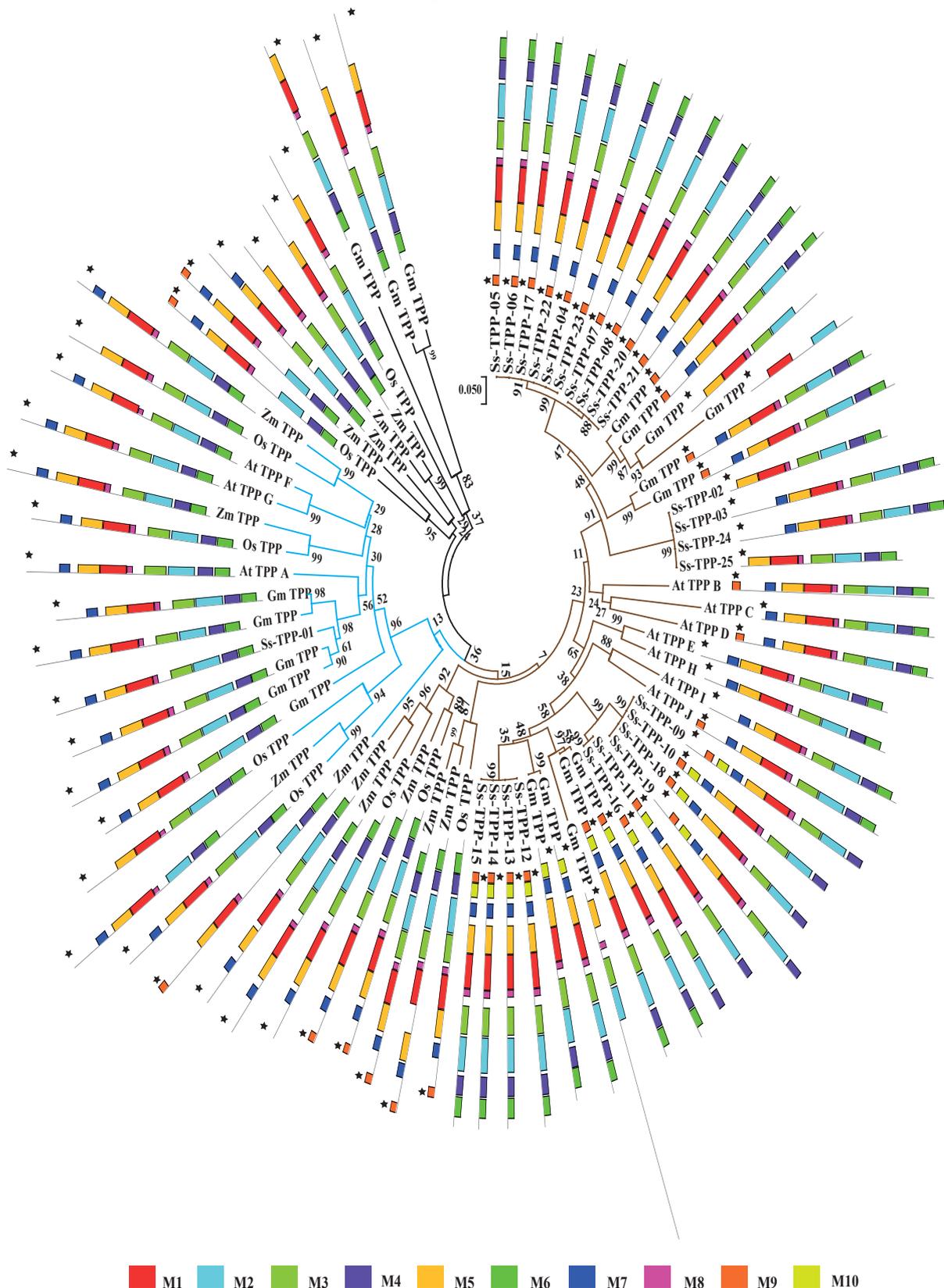


Figura 16b. Dendrograma mostrando as relações entre ortólogos aos candidatos de TPP pelo método neighbor joining onde nos ramos azuis estão presentes as sequências do Clado A e nos ramos marrons as sequências do Clado B. Número nas bases dos ramos referem-se aos valores de bootstrap. **Legenda:** Ss (*Stylosanthes scabra*); Gm (*Glycine max*); At (*Arabidopsis thaliana*); Os (*Oryza sativa*); Zm (*Zea mays*). Posição da estrela indica início da sequência. M= Motivo.

4. DISCUSSÃO

4.1 Avaliação dos domínios e motivos conservados

A variedade de osmoprotetores identificados no transcriptoma de *S. scabra* assemelham-se aos resultados de trabalhos anteriores como o de dos Santos Barros *et al.* (2009), que utilizaram a bases de dados EST e identificaram 3 (P5CS), 4 (P5CR), 10 (TPS), 8 (TPP), 1 (CMO), 3 (BADH) e 3 (MIPS), analisando 19 bibliotecas de diferentes tecidos e condições de cinco espécies do gênero *Eucalyptus*. Por sua vez, Kido *et al.* (2013) utilizaram robusta dados de bibliotecas HT-SuperSAGE verificaram 10.546 transcritos de osmoprotetores em *G. max*, sendo 36 diferencialmente expressos sob estresse hídrico.

O acúmulo de aminoácidos tem sido observado em vegetais durante períodos de estresse abiótico (Rhodes *et al.* 1986; Lugan *et al.* 2010). Como um aminoácido funcional, a prolina possui diversos papéis sob condições de estresse, estabilizando proteínas, membranas e estruturas subcelulares, além de atuar na proteção das funções da célula, fazendo a limpeza das espécies reativas de oxigênio (ROS) (Kaur e Asthir 2015). P5CS [EC:2.7.2.11 1.2.1.41] é uma enzima bifuncional em plantas que catalisa os dois primeiros passos da via glutamato na biossíntese de prolina (Delauney e Verma 1993; Yoshida *et al.* 1995) com aproximadamente 717 aminoácidos e peso molecular estimado de 77.7 kDa (Savouré *et al.* 1995; Strizhov *et al.* 1997). Os resíduos conservados nos domínios Glu-5-Kinase e GSA-DH localizados nos transcritos de *S. scabra*: ATP e NAD(P)H sítios de ligação e a região rica em Leucina foram também encontrados por Zhuang *et al.* (2011), Wang *et al.* (2014) e Li *et al.* (2014) em estudos com *Jatropha curcas* L., Milho e *Lolium perenne* L., respectivamente. O motivo Leucina Zipper está envolvido na preservação da estrutura terciária da enzima (Hu *et al.* 1992) e está presente em ambos domínios, porém em *S. scabra* só foi possível localizá-lo no domínio da região N-Terminal. O domínio Glu-5-Kinase é responsável pela fosforilação do glutamato para formar glutamil fosfato que é reduzido para glutâmico semi-aldeído pelo domínio GSA-DH (Wang *et al.* 2014) formando pirroline-5-carboxylato (P5C), que por sua vez, é reduzido a prolina pela enzima P5CR, caracterizada por dois domínios [Domínio de ligação NAD(P) Rossmann, CATH3.40.50.720; e um domínio C-Terminal ProC, CATH1.10.3730.10], com aproximadamente 280 aminoácidos, 310 kDa em espinafre e localizada no citosol em

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

raízes de soja, pode ser reconhecida através dos motivos conservados GxxGxG (Rossmann fold), GSxPA (marca registrada do núcleo de dimerização) e SP-A/G-GTT (Szoke *et al.* 1992; Murahama *et al.* 2001; Forlani *et al.* 2015), presentes em *S. scabra*. O domínio da região C-Terminal em P5CR, denominado de domínio de dimerização, está envolvido no arranjo do sítio ativo da proteína e baseado em estudos estruturais essa dimerização é essencial na formação de uma enzima completamente funcional (Forlani *et al.* 2015).

Açúcares além de funcionarem como osmoprotetores, também desempenham papel de substrato para o desenvolvimento vegetal e são reguladores da expressão gênica (Koch 1996). Dissacarídeos, como a trealose, estão essencialmente envolvidos na tolerância à estresses em plantas (Keunen *et al.* 2013) e em vários organismos. Este açúcar é capaz de estabilizar e proteger membranas e proteínas, permitindo-os sobreviver a ciclos de desidratação e reidratação (Weisburd 1988; Colaço *et al.* 1992). A estrutura das proteínas TPS em plantas contém dois domínios, TPS na região N-Terminal e TPP na região C-Terminal (Goddijn e Van Dun 1999; Mu *et al.* 2016) enquanto que as proteínas TPP apresentam apenas o domínio TPP (Yang *et al.* 2012). *S. scabra* apresentou candidatos com ambos os domínios (TPS e TPP) além de candidatos com os domínios TPS ou TPP exclusivos.

Os domínios TPS (Glyco_transf_20) e TPP (Trehalose_PPase / HAD-Like Superfamily) além de possuírem resíduos conservados presentes em ambos e encontrados na maioria das sequências de *S. scabra* (com raras exceções) também foram identificados por Avonce *et al.* (2006), Rao *et al.* (2006), Lei *et al.* (2017) e Wang *et al.* (2019). Os resíduos presentes no domínio TPS são importantes para a ligação glucose-6-fosfato (Glc6P) e uridina difosfoglucoose (UDP-Glc) (Avonce *et al.* 2006; Wang *et al.* 2019). HAD-like Superfamily compreende várias fosfatases e hidrolases e estruturas cristalinas dessas proteínas mostraram que os três motivos conservados identificados no presente trabalho, Dxxx(T/V), (S/T)Gx e Kx16-30(G/S)(D/S)xxx(D/N) formam o sítio ativo das enzimas (Fieulaine *et al.* 2005; Burroughs *et al.* 2006).

Álcoois de açúcar também protegem moléculas durante a desidratação e reidratação contra mudanças osmóticas rápidas e estresse oxidativo (Tarczynski *et al.* 1993). Em plantas superiores, mio-inositol possui diversos papéis importantes como o desenvolvimento e crescimento, manutenção dos componentes celulares além de desempenhar o papel de osmoprotetor e fazer parte da resposta à estresses ambientais (Bohnert *et al.* 1995; Valluru e Ende 2011).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

MIPS (EC 5.5.1.4) é a enzima chave da biossíntese de mio-inositol, possuindo cerca de 510 aa em plantas superiores (Ma *et al.* 2019) pois converte D-glucose 6-fostato em myo-inositol-1-fosfato (Ins1P) que em seguida é desfosforilado pela enzima IMP (EC 3.1.3.25) para formar inositol livre (Abreu e Aragão 2007). Resíduos conservados de ambas as proteínas foram encontrados em *S. scabra*: quatro sequências de aminoácidos altamente conservadas nas proteínas MIPS (GWGGNNG; LWTANTERY; NGSPQNTFVPGL e SYNHLGNNDG), descritas por Majumder *et al.* (2003) e por Ma *et al.* (2019), envolvidas na ligação das proteínas para catalisar a reação enzimática e que são essenciais para as funções da mesma (Majumder *et al.* 1997), além de motivos conservados nas proteínas IMP (DPLDGT; WDXAAG e GEET, sendo este último não localizado nas sequências de *S. scabra*) também identificados por Song *et al.* (2016) e Zhang *et al.* (2017) que são descritos como indispensáveis para a ligação de metais e substratos para IMP (Zhang *et al.* 2017).

Análises mutagênicas mostraram que o Triptofano (W) e o Aspartato (D) no motivo WDXAAG, além do Aspartato (D) em DPLDGT, são resíduos chave para a afinidade do lítio e do magnésio e por IMP ser uma proteína com sensibilidade ao lítio (Gillaspy *et al.* 1995), a mutação nesses resíduos revoga quase que completamente a atividade da enzima (Nigou *et al.* 2002). *S. scabra* não apresentou mutações nesses resíduos demonstrando que as proteínas são de fato funcionais na espécie.

Glicina-betaína é um composto amônio quaternário presente em diversos organismos como bactérias, mamíferos, invertebrados marinhos e plantas (Ahmad *et al.* 2008; Chen e Murata 2008). Tem sido sugerido que a glicina-betaína estabiliza membranas, protege proteínas e o fotossistema 2 e reduz o estresse oxidativo (Chen e Murata 2011). Diversas halófitas de interesse agrícola, incluindo algumas Poaceae como o arroz, sintetizam e acumulam glicina-betaína em resposta a estresses abióticos (Weretilnyk *et al.*, 1989). Em vegetais, a glicina-betaína é sintetizada via colina, onde a enzima CMO (EC:1.14.15.7) converte a colina em betaína aldeído (BAL), que é convertida em glicina betaína pela enzima BADH (EC:1.2.1.8).

CMO é uma proteína que composta por aproximadamente 410 aminoácidos em arroz (Luo *et al.* 2012), localizada no cloroplasto (Rathinasabapathi *et al.* 1997) e possui dois motivos conservados: um Rieske-type (2Fe-2S) que corresponde a sequência CxHx15–17Cx2H, onde X representa qualquer aminoácido e uma sequência mononuclear não heme de ligação ao ferro G/Dx3–4Dx2Hx4–5H e tais

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

motivos são considerados essenciais para o funcionamento das proteínas CMO (Mason e Cammack 1992; Rathinasabapathi *et al.* 1997; Hibino *et al.* 2002). Em *S. scabra* só foi possível identificar a presença do motivo G/Dx3–4Dx2Hx4–5H devido ao fato das proteínas apresentarem apenas cerca de 264 aa em suas sequências, provavelmente devido ao domínio Rieske_RO_Alpha_CMO apresentar-se incompleto, indicando que talvez essas proteínas não sejam funcionais na espécie. Em plantas, BADH é uma enzima NAD⁺ dependente que pertence à superfamília ALDH (Aldeído Dehidrogenases), mais precisamente a família ALDH10 (Vasiliou *et al.* 1999), que contém o decapeptídeo (VTLELGGKSP) altamente conservado em suas sequências (Lamark *et al.* 1991).

Proteínas ALDH10 de plantas são aminoaldeídos dehidrogenases (AMADHs) que também possuem três resíduos altamente conservados (PWNYP, GQI(V)CSATSR e ELGGKSP) essenciais para a sua catálise, que situam-se no fundo do canal de substrato e formam conjuntamente o sítio ativo da proteína (Perez-Miller e Hurley, 2003; Wymore *et al.* 2004). O sinal SKL na região C-Terminal indica o direcionamento das proteínas ao peroxissomos (Reumann 2004) e o motivo QLFIDGE indica sua localização no cloroplasto (Weretilnyk e Hanson, 1990). Algumas ALDH10 de plantas são capazes de utilizar eficientemente BAL como substrato (Fujiwara *et al.* 2008; Hanson *et al.* 1985; Hibino *et al.* 2001), e dessa forma, participar da síntese de glicina-betaína.

Análises de cristalografia da proteína de espinafre (SoBADH) moldada *in silico* possibilitou a descoberta de que um aminoácido na posição 441 era crítico para que a enzima aceitasse ou rejeitasse BAL como substrato e sintetizar a glicina-betaína (Díaz-Sánchez *et al.* 2012). Este resíduo, identificado como uma Alanina (A) ou Cisteína (C), está localizado na segunda esfera da interação com o substrato, sendo anterior ao triptofano (W) na posição 456 em SoBADH, determinando o tamanho do “bolso” formado por esse triptofano e uma Tirosina localizada na posição 160, onde o substrato irá se ligar. Nas ALDH10 vegetais com baixa atividade BADH, esse “bolso” tem seu espaço reduzido, pois nessa posição crítica, ao invés de uma Ala ou Cis, terá uma isoleucina (I) que irá puxar o Triptofano contra a Tirosina, diminuindo o espaço e dificultando a ligação com o substrato (Díaz-Sánchez *et al.* 2012).

Com base nesses fatos, Díaz-Sánchez *et al.* (2012) chegaram a conclusão de que as proteínas que possuísem uma isoleucina nesta posição crítica (I441 em SoBADH), denominadas de “Tipo I441”, teriam apenas atividade AMADH enquanto as

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

que possuíssem uma cisteína ou alanina teriam também atividade BADH, sendo um fator limitante para a síntese de glicina-betaína em vegetais. Esta conclusão foi posteriormente confirmada por Kopěčný *et al.* (2013), que relataram as estruturas cristalinas de ALDH10 do milho, que contém uma cisteína na posição equivalente à 441 em SoBADH e possui atividade BADH (ZmAMADH1a; código PDB 4I8P), além de uma isoenzima ALDH10 de tomate que contém uma isoleucina nesta posição e não possui atividade BADH (SIAMADH1; PDB código 4I9B). Muñoz-Clares *et al.* (2014) concluíram que de fato uma atividade de BADH não só iria prover níveis significantes de glicina-betaína mas que também iria evitar o acúmulo de BAL em níveis tóxicos.

As sequências de proteínas identificadas como BADH através da análise de domínio em *S. scabra* apresentaram o decapeptídeo, além de todos os resíduos conservados já descritos anteriormente. Através do alinhamento comparativo entre as sequências de aminoácidos antes da posição crítica de espécies já identificadas por Muñoz-Clares *et al.* (2014) explorando a evolução das ALDH10 em plantas e Jiang *et al.* (2016) em estudos com *Chenopodium quinoa* foi possível determinar uma isoleucina (I) na posição crítica (I₄₇₁) em *S. scabra* indicando que provavelmente a espécie não acumule glicina-betaína. No entanto, análises posteriores de medição de conteúdo do composto, modelagem das proteínas bem como de atividade enzimática poderão confirmar tal hipótese. Em nosso estudo, o trípeptídeo SKL na região C-Terminal e o motivo QLFIDGE na região N-Terminal também foram identificados, o que é consistente com observações anteriores de que a maioria das isoformas do tipo I441 possuem o sinal peroxisomal do tipo 1 (PST1), que se encaixa na sequência (S/A)-(K/R)-(L/M/I) (Lingner *et al.* 2011; Chowdhary *et al.* 2012), enquanto as isoformas do tipo C ou A441 não possuem, exceto em milho, sorgo, *Setaria italica* e *Zoysia tenuifolia* (Muñoz-Clares *et al.* 2014).

4.2 Genes de osmoprotetores

Os osmoprotetores identificados no transcriptoma de *S. scabra*, apresentaram um total de 19 loci ancorados no genoma de *A. duranensis* (n=10), resultados próximos aos encontrados em outra espécie da família Fabaceae por Kido *et al.* 2013 que identificou um total de 25 loci em 19 cromossomos de *G. max* (2n=20).

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

A ancoragem dos transcritos de *S. scabra* no genoma de *A. duranensis* identificou a presença de dois genes P5CS e um gene P5CR, resultados verificados na maioria das plantas superiores, onde P5CS é codificada por dois genes (P5CS1 e P5CS2) enquanto P5CR é codificada por apenas um gene (Verbruggen *et al.* 1993; Su *et al.* 2011; Zhang *et al.* 2014). Análise dos genes SbP5CS1 e SbP5CS2 em sorgo, localizados nos cromossomos 3 e 9 respectivamente, demonstraram uma presença superior de mRNAs de SbP5CS1 sob imposição de estresses salino e déficit hídrico do que de SbP5CS2, concluindo que SbP5CS2 é um gene “*housekeeping*” envolvido no metabolismo da prolina, não possuindo relação com resposta a estresses (Su *et al.* 2011). Resultados semelhantes foram reportados por Székely *et al.* (2008) analisando os genes AtP5CS1 e AtP5CS2 de *A. thaliana*. Em nosso estudo, apenas o transcrito Ss-P5CS-01, ancorado no cromossomo 4 de *A. duranensis*, mostrou-se induzido sob o estresse aplicado, indicando que provavelmente apenas este gene de P5CS esteja de fato relacionado à resposta ao estresse na espécie.

Plantas superiores compreendem uma série de genes de Trealose, por exemplo, *O. sativa* possui 11 TPS genes (Yang *et al.* 2012) e 9 TPP genes (Avonce *et al.* 2006) e *A. thaliana* contém 11 membros TPS (Vandesteene *et al.* 2010) e 10 membros TPP (Avonce *et al.* 2006). Segundo Fernandez *et al.* (2010), essa redundância é um indicador que a trealose possui um importante papel metabólico. O total de genes TPS e TPP identificado no presente trabalho foi similar ao reportado na literatura, onde a ancoragem dos transcritos de *S. scabra* no genoma de *A. duranensis* apontaram uma abundância de quatro genes *TPS (com ambos os domínios) e seis genes TPP. No entanto, os candidatos com o domínio TPS exclusivo não obtiveram resultados, o que reforça a conclusão de Avonce *et al.* (2006) que mostraram que em procariotos os genes TPS e TPP possuem duas proteínas distintas nos seus genomas, já em eucariotos, tais genes são fundidos em uma única proteína, sugerindo que as proteínas bifuncionais dos procariotos devem ser os ancestrais evolutivos das enzimas que participam da biossíntese da trealose em eucariotos (Avonce *et al.* 2010).

Estudos prévios apontam que a família de genes MIPS é pequena e em geral, é composta geralmente de um ou dois membros na maioria dos vegetais, exceto *A. thaliana* com 3 genes e por *Z. mays* (micotiledônea diplóide), *Brassica rapa*, *B. oleracea*, *G. max* (dicotiledôneas diploides) e *Gossypium hirsutum* e *Gossypium barbadense* (dicotiledôneas tetraplóides), que possuem quatro genes cada (Ma *et al.*

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

2019). Resultados consistentes com os verificados em *S. scabra*, que apresentou apenas um gene MIPS.

4.3 Análise Fenética e de Classificação dos Genes

Todas as monocotiledôneas possuem dois genes codificando proteínas ALDH10, exceto pelo milho, que possui três genes. Por sua vez, as dicotiledôneas possuem um número variado de genes que codificam proteínas ALDH10, desde um (*Ricinus communis*) gene apenas até três (*Vitis vinifera*), sendo *G. max* um caso especial, pois também possui três genes, porém, acredita-se que um deles seja um pseudogene (Muñoz-Clares *et al.* 2014). Análises de sequências de aminoácidos demonstraram que a maioria das plantas possuem duas isoformas ALDH10, provavelmente devido a eventos de duplicação gênica, sendo o tipo I441 o mais comum (Díaz-Sánchez *et al.* 2012). *S. scabra* apresentou quatro isoformas de BADH, sendo estas duplicadas pois a ancoragem no genoma de *A. duranensis* revelou apenas um gene codificando a proteína BADH e um gene codificando CMO.

O alinhamento das sequências de P5CS e P5CR demonstrou que ambas são bem conservadas. O cladograma claramente separou as classes em dois grupos distintos de monocotiledôneas e dicotiledôneas, resultados também reportados por Wang *et al.* (2017), Li *et al.* (2014) e Forlani *et al.* (2015). As sequências de *S. scabra* se alinharam com o grupo de dicotiledôneas, em P5CS a separação das sequências Ss-P5CS-01 e 02 na árvore provavelmente se deve ao fato de ambas as sequências se tratarem de genes distintos. Tal separação também foi descrita por Li *et al.* (2014), analisando genes de P5CS (*Pr5CS1* e *Pr5CS2*) de *L. perenne* L, que observaram que ambos os genes se separaram em dois clados distintos de monocotiledôneas, que continham sequências de P5CS1 e P5CS2 de várias espécies. A duplicação de genes desempenha um papel importante na evolução, podendo resultar em novas funções (Pérez-Pérez *et al.* 2009) e também é descrita como uma fonte importante para a variedade de genes P5CS (Turchetto-Zolet *et al.* 2009).

Em plantas superiores, TPS e TPP são comumente classificadas em três subgrupos distintos de acordo com a similaridade com as proteínas TPS e TPP microbianas e/ou a presença de motivos específicos (Thaller *et al.* 1998; Leyman *et al.* 2001; Eastmond *et al.* 2003). Em *A. thaliana* as proteínas da classe 1 (*AtTPS1-4*) apresentaram maior similaridade com TPS1 de levedura e desse grupo a *AtTPS1*

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

acreditava-se ser a única que exibia atividade enzimática de TPS em levedura (Blázquez *et al.* 1998; Vandesteene *et al.* 2010). No entanto, estudos posteriores demonstraram que *AtTPS2* e *AtTPS4* também exibiram atividade enzimática de TPS, exceto *AtTPS3*, que possivelmente é um pseudogene (Delorge *et al.* 2015). As proteínas da classe 2 (*AtTPS5-11*) não possuem atividade enzimática TPS e TPP (Avonce *et al.* 2010), porém ainda contém motivos conservados característicos da HAD-Like Superfamily (TPP) (Thaller *et al.* 1998). Por sua vez, as proteínas da classe 3 (*AtTPPA-J* com cerca de 320–385 aminoácidos), correspondem a enzimas funcionais com domínio TPP exclusivo (Vogel *et al.* 1998; Vandesteene *et al.* 2010). Apesar de não possuírem domínio ativo de TPS, as proteínas da classe 2 tiveram suas expressões relatadas em diferentes tecidos de *Arabidopsis*, arroz e *Populus* sugerindo que as mesmas devem possuir outras funções. *AtTPS5* em *Arabidopsis*, por exemplo, tem um importante papel na termotolerância, a qual é possibilitada através de sua interação com o co-ativador transcricional MBF1c (Suzuki *et al.* 2008).

Tem sido sugerido que a origem dos genes TPP (Classe 3) nas plantas é diferente da origem das classes que contém domínio TPS (Classe 1 e classe 2) (Avonce *et al.* 2010). Em *S. scabra*, as sequências de TPS se alinharam formando dois grandes *clusters*, conforme relatado por (Avonce *et al.* 2010) em *Arabidopsis* e por Yang *et al.* (2012) em *Populus*, *Arabidopsis* e arroz, em *Phaseolus vulgaris* por Barraza *et al.* (2016), em algodão por Mu *et al.* (2016), em *Nelumbo nucifera* por Jin *et al.* (2016) e em *Moringa oleifera*. Lin *et al.* (2018), separando-se em classe 1 e classe 2. As sequências Ss-TPS-05 e 06 de *S. scabra* alinharam-se com as sequências pertencentes a classe 1 de *Arabidopsis*, além disso, apresentaram a maioria dos resíduos conservados do domínio TPS e nenhum motivo conservado do domínio TPP (com exceção de Ss-TPS-05 que apresentou apenas um), o que sugere que estas sequências provavelmente possuem o domínio TPS ativo e o domínio TPP inativo.

As sequências Ss-TPS-01, 02, 03, 04, 07, 08 e 09 por não terem apresentado a maioria dos resíduos conservados do domínio TPS juntaram-se com as sequências da classe 2 e provavelmente não possuem domínio TPS ativo. Já em relação ao domínio TPP nas sequências da classe 2, as mesmas apresentaram dois dos três domínios conservados, com Ss-TPS-01, 02, 07 e 08 apresentando todos os motivos conservados, indicando que estas sequências provavelmente possuem o domínio TPP ativo. No entanto, o processo e função específica de cada domínio necessita de

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

análises posteriores para confirmar essa hipótese. As sequências de *S. scabra* da classe 2 dividiram-se em várias subfamílias contendo membros tanto de monocotiledôneas quanto de dicotiledôneas, sugerindo que esta classe de genes TPS surgiu a partir de antigos eventos de duplicação gênica seguidos de especialização, como proposto em *Arabidopsis* (Ramon *et al.* 2009; Avonce *et al.* 2010).

Diversos autores (Henry *et al.* 2014; Xie *et al.* 2015; Han *et al.* 2016) classificam as sequências TPP conforme divisão proposta em *Arabidopsis* (Vandesteene *et al.* 2010), que se dá em dois clados, Clado A (AtTPPA, AtTPPF, e AtTPPG) e Clado B (AtTPPB, AtTPPC, AtTPPD, AtTPPE, AtTPPH, AtTPPI, e AtTPPJ). As sequências de *S. scabra* apresentaram a mesma divisão, com a sequência Ss-TPP-01 agrupando-se no Clado A e as demais no Clado B. As sequências do Clado B apresentaram uma maior diversidade, formando subgrupos de mono e dicotiledôneas. Paul *et al.* (2018) também descreveu a presença de um clado isolado, com sequências de *G. max* e trigo, o que também foi observado no presente trabalho, com três sequências de *G. max*, quatro de *Z. mays* e três de *O. sativa*.

A árvore fenética das sequências da proteína MIPS, mostraram uma clara segregação destes genes em monocotiledôneas e dicotiledôneas, resultados também reportados por Abreu e Aragão (2007) e Ma *et al.* (2019), o que indica que é possível que a divergência dos genes MIPS tenha ocorrido após a divergência entre mono e dicotiledôneas (Abreu e Aragão 2007). A análise de distribuição de motivos mostrou que estes genes estão localizados similarmente em todas as espécies, além da presença de resíduos altamente conservados em MIPS de eucariotos (Abreu e Aragão 2007), sugerindo que MIPS é uma família proteica bastante conservada em diversas espécies de plantas.

O dendrograma gerado para IMP mostrou que algumas sequências de *S. scabra* estavam mais relacionadas com sequências de outras espécies da família Fabaceae, resultado semelhante ao obtido por Zhang *et al.* (2017), que constatou que o IMP de *O. sativa* estava relacionado com o IMP de *O. brachyantha* (arroz selvagem).

Análises fenéticas de ALDH10 em vegetais identificaram grupos distintos de monocotiledôneas e dicotiledôneas, assim como um grupo monofilético formado por sequências de espermatófitas (plantas com sementes) (Kopěčný *et al.* 2013; Muñoz-Clares *et al.* 2014), e o fato destas plantas primitivas possuírem apenas uma isoforma de ALDH10 com o I441 suporta a afirmação anteriormente reportada por Díaz-Sánchez *et al.* (2012), de que as isoformas do tipo I441 correspondem a proteína

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

ancestral da família ALDH10. Este mesmo resultado foi encontrado no presente trabalho, onde foi possível verificar a separação dos ramos de acordo com a diferenciação taxonômica, com as sequências de *S. scabra* agrupando-se com sequências de Fabaceae. O mesmo ocorreu na árvore fenética das sequências de CMO, como também observado por Luo *et al.* (2012).

4.4 Expressão diferencial

A expressão de P5CS1 em diferentes tecidos já se mostrou induzida em vários tipos de estresses abióticos, incluindo déficit hídrico, salinidade e baixas temperaturas (Strizhov *et al.* 1997; Yoshida *et al.* 1995; Abraham *et al.* 2003). Per *et al.* (2017), analisando a expressão de genes P5CS em tratamentos com déficit hídrico e estresse salino em *A. thaliana* e *O. sativa*, demonstraram que houve uma maior abundância de transcritos nos brotos das plantas em ambos os estresses, indicando um alto nível de expressão nas primeiras horas de aplicação do estresse, com decréscimo gradual ao longo das 12 horas de estresse total. Em contraste, no presente estudo, foi observada uma leve indução no tempo de 6 horas, que foi acentuada com 24 horas de estresse.

Em relação à expressão nas raízes, os autores concluíram que *AtP5CS* e *OsP5CS* mostraram maior resposta ao estresse salino em relação ao déficit hídrico. Estas observações os levaram a conclusão que a biossíntese de prolina era de fato afetada significativamente sob condições de estresse abiótico, conclusão esta que também pode ser atribuída ao presente estudo, sugerindo que os genes P5CS desempenhavam papéis distintos a depender do tipo de estresse. Os transcritos de P5CR em *S. scabra* também se mostraram induzidos sob déficit hídrico, resultado semelhante ao verificado por Ma *et al.* (2008) e Kido *et al.* (2013), indicando que de fato a prolina possa desempenhar um papel importante para a espécie diante da condição de estresse.

A trealose pode proteger substâncias bioativas e estruturas celulares, como proteínas e ácidos nucleicos, sobre condições ambientais adversas, como alta salinidade e temperatura, seca, oxidação e baixas temperaturas (Elbein *et al.* 2003; Jang *et al.* 2003). Neste estudo, após submeter *S. scabra* ao déficit hídrico por dois intervalos de tempo (6 e 24 horas), 20 transcritos das classes TPS e TPP apresentaram expressão diferencial em raízes, corroborando com os trabalhos de Mu *et al.* (2016), Lin *et al.* (2018) e Jiang *et al.* (2019). No entanto, os genes TPS da classe

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

1 (Ss-TPS-05 e 06) mostraram-se reprimidos na condição de estresse, em contraste ao observado em outras espécies, como *Arabidopsis* (Avonce *et al.* 2004), *M. olerifera* (Lin *et al.* 2018) e *Ipomoea batatas* (Wang *et al.* 2019), nas quais *AtTPS1*, *MoTPS1* e 2 e *IbTPS1* foram induzidos.

Por outro lado, os genes da classe 2 (Ss-TPS-01, 03, 07 e 09) apresentaram indução, resultados similares aos encontrados por Mu *et al.* (2016), que analisando o padrão de expressão em desses genes em três cultivares de algodão (*G. raimondii*, *G. arboreum* L. e *G. hirsutum* L.) detectaram que os mesmos foram induzidos ou reprimidos em ambas as classes em tecidos radiculares sob déficit hídrico.

Jiang *et al.* (2019) relataram que a superexpressão do gene *OsTPP3* conferiu tolerância à seca em *O. sativa*. Em nosso estudo, os transcritos da classe 3 (TPP) também se apresentaram induzidos em raízes de *S. scabra* sob déficit hídrico, indicando que o metabolismo da trealose pode possuir um importante papel para a espécie diante da condição de seca.

Diversos autores (Das-Chatterjee *et al.* 2006; Goswami *et al.* 2014; Joshi *et al.* 2013; Kaur *et al.* 2013; Tan *et al.* 2013) já relataram que os genes MIPS participam da tolerância a estresses abióticos, como salinidade, desidratação e altas temperaturas em várias espécies. Poucas plantas tolerantes a estresses que codificam proteínas MIPS são capazes de deter sua atividade catalítica na presença de altas temperaturas ou em concentrações elevadas de sal para sintetizar inositol na condição do estresse (Chen *et al.* 2000; Majee *et al.* 2004; Dastidar *et al.* 2006). É sugerido que essa capacidade ocorra devido a mudanças na sua codificação de genes a depender do ambiente no qual as mesmas estejam inseridas. O arroz selvagem, por exemplo, codifica MIPS que conferem tolerância à salinidade enquanto que o arroz cultivado codifica MIPS sensível ao sal (Majee *et al.* 2004). Neste estudo, *S. scabra*, nós reportamos a indução de transcritos MIPS em dois tempos de estresse, apresentando um nível de expressão mais acentuado no tempo de 6 horas comparado ao de 24 horas, em especial no transcrito Ss-MIPS-03 em contraste aos resultados reportados por Kaur *et al.* (2013), que demonstraram que o gene *CaMIPS2*, um dos dois genes que codifica isoformas de MIPS em grão de bico, é um gene de resposta imediata à desidratação além de ser rapidamente induzido por aplicação exógena de ABA.

5. CONCLUSÕES

- Foram caracterizados 74 osmoprotetores em *S. scabra* gerando um banco de sequências protéicas curadas para estes genes, disponíveis para uso em futuros projetos visando o melhoramento da espécie. Considerando cada subfamília, especialmente nas regiões dos domínios e motivos a maioria das proteínas apresentaram um alto grau de conservação.
- Em Glicina betaína, as proteínas CMO apresentaram o domínio da região N-Terminal incompleto e com ausência do motivo conservado CxHx15–17Cx2H sendo provável que estas não sejam funcionais. Em relação as proteínas BADH, foi identificado uma Isoleucina na posição crítica para integração do substrato BAL utilizado pela enzima para a produção do composto, o que indica que provavelmente em *S. scabra* não ocorra a síntese de GB.
- A análise fenética de classificação dos ortólogos de genes de Prolina, Mio inositol e Glicina betaína apontou para o agrupamento das sequências de *S. scabra* com outras espécies do gênero Fabaceae, comprovando a similaridade e conservação das mesmas.
- O perfil transcricional dos osmoprotetores mostrou-se induzido in silico para as quatro famílias nos dois tempos de estresse aplicado implicando sua importante função na osmoregulação e tolerância à desidratação radicular em *S. scabra*.
- Ss-*TPS-01 que possui apenas o domínio TPS, apresentou alto nível de indução na condição do estresse, merecendo atenção para posteriores estudos, pois pode se tratar de uma provável variante vegetal com atividade específica.
- Estes resultados contribuíram para um melhor entendimento das características moleculares dos genes de osmoprotetores bem como sua participação na defesa da espécie *S. scabra* contra o déficit hídrico apresentando potencial para o melhoramento genético da cultura como também de outras leguminosas de importância agrícola sensíveis ao estresse.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu EFM e Aragão FJL (2007) Isolation and characterization of a myo-inositol-1-phosphate synthase gene from yellow passion fruit (*Passiflora edulis* f. *flavicarpa*) expressed during seed development and environmental stress. **Annals of botany** **99**: 285–292.
- Ahmad R, Kim MD, Back K-H, Kim H-S, Lee H-S, Kwon S-Y, Murata N, Chung W-I e Kwak S-S (2008) Stress-induced expression of choline oxidase in potato plant chloroplasts confers enhanced tolerance to oxidative, salt, and drought stresses. **Plant cell reports** **27**: 687–698.
- Avni Öktem H, Eyidoğan F, Selçuk F, Tufan Öz M, da Silva JAT e Yücel M (2008) Revealing Response of Plants to Biotic and Abiotic Stresses with Microarray Technology. **Genes, Genomes and Genomics**: 14-48
- Avonce N, Mendoza-Vargas A, Morett E e Iturriaga G (2006) Insights on the evolution of trehalose biosynthesis. **BMC evolutionary biology** **6**: 109.
- Avonce N, Wuyts J, Verschooten K, Vandesteene L e Van Dijck P (2010) The Cytophaga hutchinsonii ChTPSP: First characterized bifunctional TPS-TPP protein as putative ancestor of all eukaryotic trehalose biosynthesis proteins. **Molecular Biology and Evolution** **27**: 359–369.
- Barraza A, Contreras-Cubas C, Estrada-Navarrete G, Reyes JL, Juárez-Verdayes MA, Avonce N, Quinto C, Díaz-Camino C e Sanchez F (2016) The Class II Trehalose 6-phosphate Synthase Gene PvTPS9 Modulates Trehalose Metabolism in *Phaseolus vulgaris* Nodules. **Frontiers in Plant Science** **7**.
- Benko-Iseppon AM, Soares-Cavalcanti NM, Berlarmino LC, Bezerra-Neto JP, Amorim LLB, Ferreira Neto JRC, Pandolfi V, Azevedo HMA, Silva RLO, Santos MG, Alves MVS, Kido EA (2011) Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. Volume Especial do III Simpósio de Mudanças Climáticas e Desertificação no Semiárido Brasileiro. **Revista Brasileira de Geografia Física** **4**: 1112-1134.
- Blázquez MA, Santos E, Flores CL, Martínez-Zapater JM, Salinas J and Gancedo C (1998) Isolation and molecular characterization of the Arabidopsis TPS1 gene, encoding trehalose-6-phosphate synthase. **The Plant Journal: For Cell and Molecular Biology** **13**: 685–689.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Bohnert H, Nelson D e Jensen R (1995) Adaptations to Environmental Stresses. **The Plant Cell** **7**: 1099–1111.
- Brigotti M, Petronini PG, Carnicelli D, Alfieri RR, *et al.* (2003) Effect of osmolarity, ions and compatible osmolytes on cell-free protein synthesis. **Biochem. J.** **369**: 369–374.
- Burroughs AM, Allen KN, Dunaway-Mariano D e Aravind L (2006) Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. **Journal of Molecular Biology** **361**: 1003–1034.
- Chen L, Zhou C, Yang H e Roberts MF (2000) Inositol-1-phosphate synthase from *Archaeoglobus fulgidus* is a class II aldolase. **Biochemistry** **39**: 12415–12423.
- Chen THH e Murata N (2008) Glycine betaine: an effective protectant against abiotic stress in plants. **Trends in plant science** **13**: 499–505.
- Chen THH e Murata N (2011) Glycine betaine protects plants against abiotic stress: mechanisms and biotechnological applications. **Plant, cell & environment** **34**: 1–20.
- Chou KC e Shen HB (2010) Cell-PLoc 2.0: An improved package of web-servers for predicting subcellular localization of proteins in various organisms. **Natural Science** **2010**: 1090–1103.
- Chowdhary G, Kataya AR, Lingner T e Reumann S (2012) Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. **BMC plant biology** **12**: 142.
- Das-Chatterjee A, Goswami L, Maitra S, Dastidar KG, Ray S e Majumder AL (2006) Introgression of a novel salt-tolerant L-myo-inositol 1-phosphate synthase from *Porteresia coarctata* (Roxb.) Tateoka (PcINO1) confers salt tolerance to evolutionary diverse organisms. **FEBS letters** **580**: 3980–3988.
- Dayhoff MO e Schwartz RM (1978) Chapter 22: **A model of evolutionary change in proteins**. In in Atlas of Protein Sequence and Structure.
- Delauney AJ e Verma DPS (1993) Proline biosynthesis and osmoregulation in plants. **The Plant Journal** **4**: 215–223.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Delorge I, Figueroa CM, Feil R, Lunn JE e Van Dijck P (2015) Trehalose-6-phosphate synthase 1 is not the only active TPS in *Arabidopsis thaliana*. **The Biochemical Journal** **466**: 283–290.
- Díaz-Sánchez ÁG, González-Segura L, Mújica-Jiménez C, Rudiño-Piñera E, Montiel C, Martínez-Castilla LP e Muñoz-Clares RA (2012) Amino acid residues critical for the specificity for betaine aldehyde of the plant ALDH10 isoenzyme involved in the synthesis of glycine betaine. **Plant Physiology** **158**: 1570–1582.
- dos Santos Barros P, Soares-Cavalcanti NM, Vieira-Mello GS, Wanderley-Nogueira AC, Calsa-Junior T and Benko-Iseppon AM (2009) In Silico Evaluation of Osmoprotectants in *Eucalyptus* Transcriptome. Springer Berlin Heidelberg, p. 66–77. In Masulli F, Tagliaferri R, and Verkhivker GM (eds) **Computational Intelligence Methods for Bioinformatics and Biostatistics**.
- Eastmond PJ, Li Y e Graham IA (2003) Is trehalose-6-phosphate a regulator of sugar metabolism in plants? **Journal of Experimental Botany** **54**: 533–537.
- Eisen MB, Spellman PT, Brown PO e Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. **Proceedings of the National Academy of Sciences of the United States of America** **95**: 14863–14868.
- Elbein AD, Pan YT, Pastuszak I e Carroll D (2003) New insights on trehalose: a multifunctional molecule. **Glycobiology** **13**: 17R-27R.
- Felsenstein J (1985) Confidence Limits on Phylogenies: An Approach Using The Bootstrap. **International Journal of Organic Evolution** **39**: 783–791.
- Fernandez O, Béthencourt L, Quero A, Sangwan RS e Clément C (2010) Trehalose and plant stress responses: friend or foe? **Trends in Plant Science** **15**: 409–417.
- Ferreira MB e Costa NMS (1979) **O gênero *Stylosanthes* Sw. no Brasil**. Belo Horizonte : Empresa de Pesquisa Agropecuária de Minas Gerais, 116p.
- Fieulaine S, Lunn JE, Borel F e Ferrer J-L (2005) The structure of a cyanobacterial sucrose-phosphatase reveals the sugar tongs that release free sucrose in the cell. **The Plant Cell** **17**: 2049–2058.
- Forlani G, Makarova KS, Ruszkowski M, Bertazzini M e Nocek B (2015) Evolution of plant δ 1-pyrroline-5-carboxylate reductases from phylogenetic and structural perspectives. **Frontiers in Plant Science** **6**.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Fujiwara T, Hori K, Ozaki K, Yokota Y, Mitsuya S, Ichiyanagi T, Hattori T e Takabe T (2008) Enzymatic characterization of peroxisomal and cytosolic betaine aldehyde dehydrogenases in barley. **Physiologia Plantarum 134**: 22–30.
- Ghosh Dastidar K, Maitra S, Goswami L, Roy D, Das KP e Majumder AL (2006) An insight into the molecular basis of salt tolerance of L-myo-inositol 1-P synthase (PcINO1) from *Porteresia coarctata* (Roxb.) Tateoka, a halophytic wild rice. **Plant physiology 140**: 1279–1296.
- Gillaspy GE, Keddie JS, Oda K e Gruissem W (1995) Plant inositol monophosphatase is a lithium-sensitive enzyme encoded by a multigene family. **The Plant cell 7**: 2175–2185.
- Goddijn OJM e Van Dun K (1999) Trehalose metabolism in plants. **Trends Plant Sci 4**: 315–319.
- Goswami L, Sengupta S, Mukherjee S, Ray S, Mukherjee R e Lahiri Majumder A (2014) Targeted expression of L-myo- inositol 1-phosphate synthase from *Porteresia coarctata* (Roxb.) Tateoka confers multiple stress tolerance in transgenic crop plants. **Journal of Plant Biochemistry and Biotechnology 23**: 316–330.
- Han B, Fu L, Zhang D, He X, Chen Q, Peng M e Zhang J (2016) Interspecies and Intraspecies Analysis of Trehalose Contents and the Biosynthesis Pathway Gene Family Reveals Crucial Roles of Trehalose in Osmotic-Stress Tolerance in Cassava. **International Journal of Molecular Sciences 17**.
- Hanson AD, May AM, Grumet R, Bode J, Jamieson GC e Rhodes D (1985) Betaine synthesis in chenopods: Localization in chloroplasts. **Proceedings of the National Academy of Sciences of the United States of America 82**: 3678–3682.
- Henry C, Bledsoe SW, Siekman A, Kollman A, Waters BM, Feil R, Stitt M e Lagrimini LM (2014) The trehalose pathway in maize: conservation and gene regulation in response to the diurnal cycle and extended darkness. **Journal of Experimental Botany 65**: 5959–5973.
- Hibino T, Meng YL, Kawamitsu Y, Uehara N, Matsuda N, Tanaka Y, Ishikawa H, Baba S, Takabe T, Wada K, Ishii T e Takabe T (2001) Molecular cloning and functional characterization of two kinds of betaine-aldehyde dehydrogenase in betaine-accumulating mangrove *Avicennia marina* (Forsk.) Vierh. **Plant Molecular Biology 45**: 353–363.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Hibino T, Waditee R, Araki E, Ishikawa H, Aoki K, Tanaka Y e Takabe T (2002) Functional Characterization of Choline Monooxygenase, an Enzyme for Betaine Synthesis in Plants. **Journal of Biological Chemistry** **277**: 41352–41360.
- Hiller K, Grote A, Maneck M, Münch R e Jahn D (2006) JVirGel 2.0: computational prediction of proteomes separated via two-dimensional gel electrophoresis under consideration of membrane and secreted proteins. **Bioinformatics** **22**: 2441–2443.
- Hu CA, Delauney AJ and Verma DP (1992) A bifunctional enzyme (delta 1-pyrroline-5-carboxylate synthetase) catalyzes the first two steps in proline biosynthesis in plants. **Proceedings of the National Academy of Sciences of the United States of America** **89**: 9354–9358.
- Jang I-C, Oh S-J, Seo J-S, Choi W-B, Song SI, Kim CH, Kim YS, Seo H-S, Choi YD, Nahm BH and Kim J-K (2003) Expression of a bifunctional fusion of the *Escherichia coli* genes for trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase in transgenic rice plants increases trehalose accumulation and abiotic stress tolerance without stunting growth. **Plant Physiology** **131**: 516–524.
- Jiang D, Chen W, Gao J, Yang F e Zhuang C (2019) Overexpression of the trehalose-6-phosphate phosphatase OsTPP3 increases drought tolerance in rice. **Plant Biotechnology Reports** **13**: 285–292.
- Jiang Y, Zhu S, Yuan J, Chen G and Lu G (2016) A betaine aldehyde dehydrogenase gene in quinoa (*Chenopodium quinoa*): structure, phylogeny, and expression pattern. **Genes & Genomics** **38**: 1013–1020.
- Jin Q, Hu X, Li X, Wang B, Wang Y, Jiang H, Mattson N e Xu Y (2016) Genome-Wide Identification and Evolution Analysis of Trehalose-6-Phosphate Synthase Gene Family in *Nelumbo nucifera*. **Frontiers in Plant Science** **7**: 1445.
- Joshi R, Ramanarao MV e Baisakh N (2013) *Arabidopsis* plants constitutively overexpressing a myo-inositol 1-phosphate synthase gene (SalNO1) from the halophyte smooth cordgrass exhibits enhanced level of tolerance to salt stress. **Plant physiology and biochemistry** **65**: 61–66.
- Kaur G e Asthir B (2015) Proline: a key player in plant abiotic stress tolerance. **Biologia Plantarum** **59**: 609–619.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Kaur H, Verma P, Petla BP, Rao V, Saxena SC e Majee M (2013) Ectopic expression of the ABA-inducible dehydration-responsive chickpea L-myo-inositol 1-phosphate synthase 2 (CaMIPS2) in *Arabidopsis* enhances tolerance to salinity and dehydration stress. **Planta** **237**: 321–335.
- Keunen E, Peshev D, Vangronsveld J, Van Den Ende W e Cuypers A (2013) Plant sugars are crucial players in the oxidative challenge during abiotic stress: extending the traditional concept. **Plant, Cell & Environment** **36**: 1242–1255.
- Kido EA, Ferreira Neto JR, Silva RL, Belarmino LC, Bezerra Neto JP, Soares-Cavalcanti NM, Pandolfi V, Silva MD, Nepomuceno AL e Benko-Iseppon AM (2013) Expression dynamics and genome distribution of osmoprotectants in soybean: identifying important components to face abiotic stress. **BMC Bioinformatics** **14**: S7.
- Koch KE (1996) Carbohydrate-modulated gene expression in plants. **Annual Review of Plant Physiology and Plant Molecular Biology** **47**: 509–540.
- Kopečný D, Končítíková R, Tylichová M, Vigouroux A, Moskalíková H, Soral M, Šebela M e Moréra S (2013) Plant ALDH10 family: identifying critical residues for substrate specificity and trapping a thiohemiacetal intermediate. **The Journal of Biological Chemistry** **288**: 9491–9507.
- Kumar S, Stecher G e Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution** **33**: 1870–1874.
- Lamark T, Kaasen I, Eshoo MW, Falkenberg P, McDougall J e Strøm AR (1991) DNA sequence and analysis of the bet genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. **Molecular Microbiology** **5**: 1049–1064.
- Lei M, Wu X, Zhang J, Wang H e Huang C (2017) Gene cloning, expression, and characterization of trehalose-6-phosphate synthase from *Pleurotus ostreatus*. **Journal of Basic Microbiology** **57**: 580–589.
- Leyman B, Van Dijck P e Thevelein JM (2001) An unexpected plethora of trehalose biosynthesis genes in *Arabidopsis thaliana*. **Trends in Plant Science** **6**: 510–513.
- Li H, Guo H, Zhang X e Fu J (2014) Expression profiles of Pr5CS1 and Pr5CS2 genes and proline accumulation under salinity stress in perennial ryegrass (*Lolium perenne* L.). **Plant Breeding** **133**: 243–249.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Li HW, Zang BS, Deng XW, Wang XP (2011) Overexpression of the trehalose-6-phosphate synthase gene *OsTPS1* enhances abiotic stress tolerance in rice. **Planta** **234**: 1007-1018.
- Lin M, Jia R, Li J, Zhang M, Chen H, Zhang D, Zhang J e Chen X (2018) Evolution and expression patterns of the trehalose-6-phosphate synthase gene family in drumstick tree (*Moringa oleifera* Lam.). **Planta** **248**: 999–1015.
- Lingner T, Kataya AR, Antonicelli GE, Benichou A, Nilssen K, Chen X-Y, Siemsen T, Morgenstern B, Meinicke P e Reumann S (2011) Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and in vivo subcellular targeting analyses. **The Plant Cell** **23**: 1556–1572.
- Lugan R, Niogret M-F, Leport L, Guégan J-P, Larher FR, Savouré A, Kopka J e Bouchereau A (2010) Metabolome and water homeostasis analysis of *Thellungiella salsuginea* suggests that dehydration tolerance is a key response to osmotic stress in this halophyte. **The Plant Journal: For Cell and Molecular Biology** **64**: 215–229.
- Luo D, Niu X, Yu J, Yan J, Gou X, Lu B-R and Liu Y (2012) Rice choline monooxygenase (*OsCMO*) protein functions in enhancing glycine betaine biosynthesis in transgenic tobacco but does not accumulate in rice (*Oryza sativa* L. ssp. japonica). **Plant cell reports** **31**: 1625–1635.
- Ma R, Song W, Wang F, Cao A, Xie S, Chen X, Jin X e Li H (2019) A Cotton (*Gossypium hirsutum*) Myo-Inositol-1-Phosphate Synthase (*GhMIPS1D*) Gene Promotes Root Cell Elongation in Arabidopsis. **International journal of molecular sciences** **20**.
- Mahajan S e Tuteja N (2005) Cold salinity and drought stresses: an overview. **Archives of Biochemistry and Biophysics** **444**: 139-158.
- Majee M, Maitra S, Dastidar KG, Pattnaik S, Chatterjee A, Hait NC, Das KP e Majumder AL (2004) A novel salt-tolerant L-myo-inositol-1-phosphate synthase from *Porteresia coarctata* (Roxb.) Tateoka, a halophytic wild rice: molecular cloning, bacterial overexpression, characterization, and functional introgression into tobacco-conferring salt tolerance phenotype. **The Journal of biological chemistry** **279**: 28539–28552.
- Majumder AL, Chatterjee A, Ghosh Dastidar K e Majee M (2003) Diversification and evolution of L-myo-inositol 1-phosphate synthase. **FEBS letters** **553**: 3–10.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Majumder AL, Johnson MD e Henry SA (1997) 1L-myo-inositol-1-phosphate synthase. **Biochimica et biophysica Acta 1348**: 245–256.
- Mason JR e Cammack R (1992) The electron-transport proteins of hydroxylating bacterial dioxygenases. **Annual review of microbiology 46**: 277–305.
- Mu M, Lu X-K, Wang J-J, Wang D-L, Yin Z-J, Wang S, Fan W-L e Ye W-W (2016) Genome-wide Identification and analysis of the stress-resistance function of the TPS (Trehalose-6-Phosphate Synthase) gene family in cotton. **BMC genetics 17**: 54.
- Muñoz-Clares RA, Riveros-Rosas H, Garza-Ramos G, González-Segura L, Mújica-Jiménez C e Julián-Sánchez A (2014) Exploring the evolutionary route of the acquisition of betaine aldehyde dehydrogenase activity by plant ALDH10 enzymes: implications for the synthesis of the osmoprotectant glycine betaine. **BMC plant biology 14**: 149.
- Murahama M, Yoshida T, Hayashi F, Ichino T, Sanada Y and Wada K (2001) Purification and characterization of Δ 1-pyrroline-5-carboxylate reductase isoenzymes, indicating differential distribution in spinach (*Spinacia oleracea* L.) leaves. **Plant & cell physiology 42**: 742–50.
- Nigou J, Dover LG e Besra GS (2002) Purification and biochemical characterization of *Mycobacterium tuberculosis* SuhB, an inositol monophosphatase involved in inositol biosynthesis. **Biochemistry 41**: 4392–4398.
- Nuccio ML, Wu J, Mowers R, Zhou HP, Meghji M, Primavesi LF, Paul MJ, Chen X, Gao Y, Haque E, Basu SS, Lagrimini LM (2015) Expression of trehalose-6- phosphate phosphatase in maize ears improves yield in well watered and drought conditions. **Natur. Biotechnol. 33**: 862- 869.
- Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. **Computer applications in the biosciences 12**: 357–358.
- Paul MJ, Gonzalez-Uriarte A, Griffiths CA e Hassani-Pak K (2018) The Role of Trehalose 6-Phosphate in Crop Yield and Resilience1. **Plant Physiology 177**: 12–23.
- Per TS, Khan NA, Reddy PS, Masood A, Hasanuzzaman M, Khan MIR e Anjum NA (2017) Approaches in modulating proline metabolism in plants for salt and drought stress tolerance: Phytohormones, mineral nutrients and transgenics. **Plant physiology and biochemistry 115**: 126–140.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Perez-Miller SJ e Hurley TD (2003) Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. **Biochemistry** **42**: 7100–7109.
- Pérez-Pérez JM, Candela H e Micol JL (2009) Understanding synergy in genetic interactions. **Trends in genetics** **25**: 368–376.
- Ramon M, De Smet I, Vandesteene L, Naudts M, Leyman B, Van Dijck P, Rolland F, Beeckman T e Thevelein JM (2009) Extensive expression regulation and lack of heterologous enzymatic activity of the Class II trehalose metabolism proteins from *Arabidopsis thaliana*. **Plant, Cell & Environment** **32**: 1015–1032.
- Rao KN, Kumaran D, Seetharaman J, Bonanno JB, Burley SK e Swaminathan S (2006) Crystal structure of trehalose-6-phosphate phosphatase-related protein: Biochemical and biological implications. **Protein Science : A Publication of the Protein Society** **15**: 1735–1744.
- Rathinasabapathi B, Burnet M, Russell BL, Gage DA, Liao PC, Nye GJ, Scott P, Golbeck JH e Hanson AD (1997) Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: prosthetic group characterization and cDNA cloning. **Proceedings of the National Academy of Sciences of the United States of America** **94**: 3454–3458.
- Reumann S (2004) Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. **Plant Physiology** **135**: 783–800.
- Rhodes D, Handa S e Bressan RA (1986) Metabolic changes associated with adaptation of plant cells to water stress. **Plant Physiology** **82**: 890–903.
- Savouré A, Jaoua S, Hua XJ, Ardiles W, Van Montagu M e Verbruggen N (1995) Isolation, characterization, and chromosomal location of a gene encoding the delta 1-pyrroline-5-carboxylate synthetase in *Arabidopsis thaliana*. **FEBS letters** **372**: 13–19.
- Seki M, Kamei A, Yamaguchi-Shinozaki K, Shinozaki K (2003) Molecular responses to drought, salinity and frost: common and different paths for plant protection. **Current Opinion in Biotechnology** **14**: 194-199.
- Song F, Su H, Yang N, Zhu L, Cheng J, Wang L e Cheng X (2016) Myo-Inositol content determined by myo-inositol biosynthesis and oxidation in blueberry fruit. **Food chemistry** **210**: 381–387.
- Stace HM (ed) e Edye LA (ed) (1984) **The biology and agronomy of *Stylosanthes***. 654p.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Strizhov N, Abraham E, Okresz L, Blickling S, Zilberstein A, Schell J, Koncz C, Szabados L (1997) Differential expression of two P5CS genes controlling proline accumulation during salt-stress requires ABA and is regulated by ABA1, ABI1 and AXR2 in *Arabidopsis*. **Plant J** **12**: 557–569.
- Su M, Li X-F, Ma X-Y, Peng X-J, Zhao A-G, Cheng L-Q, Chen S-Y e Liu G-S (2011) Cloning two P5CS genes from bioenergy sorghum and their expression profiles under abiotic stresses and MeJA treatment. **Plant Science: An International Journal of Experimental Plant Biology** **181**: 652–659.
- Suzuki N, Bajad S, Shuman J, Shulaev V e Mittler R (2008) The transcriptional co-activator MBF1c is a key regulator of thermotolerance in *Arabidopsis thaliana*. **The Journal of Biological Chemistry** **283**: 9269–9275.
- Székely G, Abrahám E, Csépló A, Rigó G, Zsigmond L, Csiszár J, Ayaydin F, Strizhov N, Jásik J, Schmelzer E, Koncz C e Szabados L (2008) Duplicated P5CS genes of *Arabidopsis* play distinct roles in stress regulation and developmental control of proline biosynthesis. **The Plant Journal: For Cell and Molecular Biology** **53**: 11–28.
- Szoke A, Miao G-H, Hong Z and Verma DPS (1992) Subcellular Location of Δ 1-Pyrroline-5-Carboxylate Reductase in Root/Nodule and Leaf of Soybean. **Plant Physiology** **99**: 1642–1649.
- Tan J, Wang C, Xiang B, Han R e Guo Z (2013) Hydrogen peroxide and nitric oxide mediated cold- and dehydration-induced myo-inositol phosphate synthase that confers multiple resistances to abiotic stresses. **Plant, cell & environment** **36**: 288–299.
- Tarczynski MC, Jensen RG e Bohnert HJ (1993) Stress protection of transgenic tobacco by production of the osmolyte mannitol. **Science** **259**: 508–510.
- Thaller MC, Schippa S and Rossolini GM (1998) Conserved sequence motifs among bacterial, eukaryotic, and archaeal phosphatases that define a new phosphohydrolase superfamily. **Protein Science: A Publication of the Protein Society** **7**: 1647–1652.
- Turchetto-Zolet AC, Margis-Pinheiro M e Margis R (2009) The evolution of pyrroline-5-carboxylate synthase in plants: a key enzyme in proline synthesis. **Molecular genetics and genomics** **281**: 87–97.
- Tuteja N (2009) Integrated Calcium Signaling in Plants. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 29–49. In Mancuso S and Baluka F (eds) **Signaling in Plants**.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Valluru R e Van den Ende W (2011) Myo-inositol and beyond--emerging networks under stress. **Plant Science: An International Journal of Experimental Plant Biology** **181**: 387–400.
- Vandesteene L, Ramon M, Le Roy K, Van Dijck P e Rolland F (2010) A single active trehalose-6-P synthase (TPS) and a family of putative regulatory TPS-like proteins in *Arabidopsis*. **Molecular Plant** **3**: 406–419.
- Vasiliou V, Bairoch A, Tipton KF e Nebert DW (1999) Eukaryotic aldehyde dehydrogenase (ALDH) genes: human polymorphisms, and recommended nomenclature based on divergent evolution and chromosomal mapping. **Pharmacogenetics** **9**: 421–434.
- Verbruggen N, Villarroel R e Van Montagu M (1993) Osmoregulation of a pyrroline -5-carboxylate reductase gene in *Arabidopsis thaliana*. **Plant Physiology** **103**: 771-781.
- Vogel G, Aeschbacher RA, Müller J, Boller T e Wiemken A (1998) Trehalose-6-phosphate phosphatases from *Arabidopsis thaliana*: identification by functional complementation of the yeast *tps2* mutant. **The Plant Journal: For Cell and Molecular Biology** **13**: 673–683.
- Wang Gang, Zhang J, Wang Guifeng, Fan X, Sun X, Qin H, Xu N, Zhong M, Qiao Z, Tang Y e Song R (2014) Proline responding1 Plays a Critical Role in Regulating General Protein Synthesis and the Cell Cycle in Maize. **The Plant Cell** **26**: 2582–2600.
- Wang L, Guo Z, Zhang Y, Wang Y, Yang G, Yang L, Wang R e Xie Z (2017) Characterization of LhSorP5CS, a gene catalyzing proline synthesis in Oriental hybrid lily Sorbonne: molecular modelling and expression analysis. **Botanical Studies** **58**.
- Wang W, Yu H, Kim HS, Yang Y, Qiu X e Kwak S-S (2019) Molecular characterization of a sweet potato stress tolerance-associated trehalose-6-phosphate synthase 1 gene (IbTPS1) in response to abiotic stress. **Plant Biotechnology Reports** **13**: 235–243.
- Weisburd S (1988) Death-Defying Dehydration. **Sci News** **13**: 107-110.
- Weretilnyk EA e Hanson AD (1990) Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. **Proceedings of the National Academy of Sciences of the United States of America** **87**: 2745–2749.

Barros, AAG. Caracterização estrutural, funcional e diversidade de osmoprotetores no transcriptoma de *Stylosanthes scabra* (Vogel) sob déficit hídrico.

- Weretilnyk EA, Bednarek S, McCue KF, Rhodes D e Hanson AD (1989) Comparative biochemical and immunological studies of the glycine betaine synthesis pathway in diverse families of dicotyledons. **Planta 178**: 342–352.
- Wymore T, Hempel J, Cho SS, Mackerell AD, Nicholas HB e Deerfield DW (2004) Molecular recognition of aldehydes by aldehyde dehydrogenase and mechanism of nucleophile activation. **Proteins 57**: 758–771.
- Xie DW, Wang XN, Fu LS, Sun J, Zheng W e Li ZF (2015) Identification of the trehalose-6-phosphate synthase gene family in winter wheat and expression analysis under conditions of freezing stress. **Journal of Genetics 94**: 55–65.
- Yang H-L, Liu Y-J, Wang C-L e Zeng Q-Y (2012) Molecular Evolution of Trehalose-6-Phosphate Synthase (TPS) Gene Family in Populus, Arabidopsis and Rice. **PLoS ONE 7**.
- Yoshiba Y, Kiyosue T, Katagiri T, Ueda H, Mizoguchi T, Yamaguchi-Shinozaki K, Wada K, Harada Y e Shinozaki K (1995) Correlation between the induction of a gene for delta 1-pyrroline-5-carboxylate synthetase and the accumulation of proline in Arabidopsis thaliana under osmotic stress. **The Plant Journal: For Cell and Molecular Biology 7**: 751–760.
- Zhang M, Huang H e Dai S (2014) Isolation and expression analysis of proline metabolism-related genes in Chrysanthemum lavandulifolium. **Gene 537**: 203–213.
- Zhang R-X, Qin L-J e Zhao D-G (2017) Overexpression of the OsIMP Gene Increases the Accumulation of Inositol and Confers Enhanced Cold Tolerance in Tobacco through Modulation of the Antioxidant Enzymes' Activities. **Genes 8**.
- Zhuang GQ, Li B, Guo HY, Liu JL e Chen F (2011) Molecular cloning and characterization of P5CS gene from Jatropha curcas L. **African Journal of Biotechnology 10**: 14803-14811.