

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA – MELHORAMENTO
GENÉTICO DE PLANTAS**

RUANA CAROLINA CABRAL DA SILVA

**DIVERSIDADE E EXPRESSÃO DIFERENCIAL DE NODULINAS PRECOSES
NO TRANSCRIPTOMA DE *Stylosanthes scabra* Vogel SUBMETIDA A
ESTRESSE ABIÓTICO**

**RECIFE – PE
2019**

RUANA CAROLINA CABRAL DA SILVA

DIVERSIDADE E EXPRESSÃO DIFERENCIAL DE NODULINAS PRECOCES
NO TRANSCRIPTOMA DE *Stylosanthes scabra* Vogel SUBMETIDA A
ESTRESSE ABIÓTICO

Dissertação apresentada ao Programa de Pós-Graduação em Agronomia, na área de concentração em Melhoramento Genético de Plantas (PPGAMGP) da Universidade Federal Rural de Pernambuco como parte dos requisitos exigidos para obtenção do título de Mestre em Melhoramento Genético de Plantas.

Orientador: Reginaldo de Carvalho

Coorientadores: João Pacífico Bezerra Neto

Roberta Lane de Oliveira Silva

RECIFE – PE
2019

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema Integrado de Bibliotecas da UFRPE
Biblioteca Central, Recife-PE, Brasil

S586d Silva, Ruana Carolina Cabral da
Diversidade e expressão diferencial de nodulinas precoces no
transcriptoma de *Stylosanthes scabra* Vogels submetida a estresse
abiótico/ Ruana Carolina Cabral da Silva. - 2019.
87 f. : il.

Orientador: Reginaldo de Carvalho.

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

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

1. Bioinformática 2. Metabolismo - Regulação 3. Estilosante
4. Regulação de expressão gênica I. Carvalho, Reginaldo de, orient.
II. Bezerra Neto, João Pacífico, coorient. III. Silva, Roberta Lane de
Oliveira, coorient. IV. Título

CDD 581.1

RUANA CAROLINA CABRAL DA SILVA

**DIVERSIDADE E EXPRESSÃO DIFERENCIAL DE NODULINAS PRECOCES
NO TRANSCRIPTOMA DE *Stylosanthes scabra* Vogel SUBMETIDA A
ESTRESSE ABIÓTICO**

Aprovada em: ___/___/___

COMISSÃO EXAMINADORA

Dr. Reginaldo de Carvalho

(Orientador - Universidade Federal
Rural de Pernambuco)

Dr. José Ribamar Costa Ferreira Neto

(Membro externo - Universidade
Federal de Pernambuco)

Dr. Nataniel Franklin de Melo

(Membro externo - Embrapa
Semiárido)

RECIFE – PE

2019

*Aos meus pais e todos aqueles que
participaram direta e indiretamente
dessa conquista*

AGRADECIMENTOS

A princípio, gostaria de agradecer a minha mãe **Rita de Cássia**, pela pessoa incrível e maravilhosa que é, pelo apoio incondicional e por ter me ajudado em todos os aspectos, bem como ao meu pai **José Tavares**, por ambos sempre estarem lá, por mim. Sou grata pelo “serrumaninho”, dono do meu amor, **Apollo Jr.**

À minha vó **Rita Fernandes**, por todo apoio e carinho e pela pessoa de luz que é, me ajudando e incentivando sempre, assim como na torcida para o meu melhor. Agradeço também aos demais familiares, pelo incentivo e ajuda em tudo que precisei, tenho por todos um carinho especial.

Ao meu orientador, Prof. Dr. **Reginaldo de Carvalho**, pois através dele tive a oportunidade de conhecer novos lugares, assim como o **Laboratório de Genética e Biotecnologia Vegetal (LGBV)** coordenado pela Prof.^a Dr.^a **Ana Maria Benko Iseppon**, que me recebeu de braços abertos, agradeço imensamente a oportunidade que me foi concebida de aprender novas técnicas, ter novas visões no meio científico e por todo apoio que continua me dando.

Aos meus coorientadores Dr.^a **Roberta Lane** e Dr. **João Pacífico**, pelos ensinamentos, além de por toda ajuda e incentivo. A todos que fazem parte do LGBV - graduandos, mestrandos, doutorandos, pós-doutorandos e professoras, por todo apoio durante esta caminhada, pela ajuda, palavras de incentivo e ensinamentos, tenho um carinho especial por cada um.

A pessoa maravilhosa que tive o prazer de conhecer **Thialisson Caaci**, jamais esquecerei seus ensinamentos, agradeço pela paciência, ajuda e conhecimentos compartilhados, você foi essencial nesse processo. Agradeço também a **Jéssica Barbosa**, por toda ajuda concedida, o meu muito obrigada.

Aos queridos amigos, **Elenilson Souza**, **Joelson Crispim** e **Wilson Dias**, por todo incentivo, carinho e conversas compartilhadas, vocês vieram para somar e deixar tudo melhor. Agradeço pelos conhecimentos compartilhados, jamais esquecerei, assim como por todos os momentos de descontrações e saídas, tenho um carinho enorme por vocês.

Agradeço imensamente aos amigos que tive o prazer de conhecer durante esta jornada, **Andreia Cristiny**, **Ayug Lemos**, **Bruna Rafaella**, **Marianne Firmino**, **Suyane Melo** e **Vanessa Emanuelle**, por todas as conversas e

momentos compartilhados, vocês foram essenciais.

E sem esquecer todos os meus amigos, os que vieram para perto e os que continuam me dando todo o apoio, independente da distância, em especial: **Amanda Dias, Átila Virgínia, Dioginys César, Fernanda Freitas, Jael Lima, Kelly Moreira, Letícia Macedo, Luana Magalhães, Lucas Araújo, Maisa Lucena** e tantos outros, cada um tem um lugar especial em meu coração.

Ao meu amor, **Rebeca Andrade**, pela pessoa incrível que é, sempre me ouvindo e apoiando, agradeço a paciência e conforto em todos os momentos difíceis, bem como pelos momentos de alegrias compartilhados, és mais que especial em minha vida.

Aos meus amigos e colegas de curso, pelos momentos compartilhados, sempre um ajudando o outro, vocês são especiais em minha vida.

À **Universidade Federal Rural de Pernambuco**, por toda estrutura, ao **Departamento de Agronomia**, e ao programa de pós-graduação, no qual faço parte, como também ao **Departamento de Genética** da **Universidade Federal de Pernambuco**, por todo trabalho oferecido para que tudo desse certo, assim como a **CAPES** pela bolsa concebida e incentivo a pesquisa científica.

GRATIDÃO!

*“A mente que se abre a uma nova ideia jamais
volta ao seu tamanho original”*

- Albert Einstein

RESUMO

Novos avanços nas tecnologias relacionadas às “ômicas” vêm favorecendo a identificação de mecanismos de sinalização que regulam as respostas de vegetais aos estresses. A nodulação e a capacidade de fixar nitrogênio em associação com bactérias do gênero *Rhizobium* spp. podem determinar a sobrevivência de plantas em ambientes adversos. Entre as diversas classes de genes que atuam em situações de estresse, destacam-se as nodulinas, que são genes do hospedeiro expressos durante o processo de nodulação, também associadas ao funcionamento dos nódulos. Estudos recentes relatam que *Stylosanthes scabra* Vogel tem sido bastante utilizada como forrageira, apresentando elevada tolerância aos ambientes com solos pobres e baixa disponibilidade hídrica. O presente projeto teve como objetivo avaliar os mecanismos de resposta utilizados por *S. scabra* sob estresse abiótico (supressão de rega), especificamente a expressão gênica diferencial de nodulinas precoces em diferentes momentos após o início do estresse. Para tanto, foi realizada uma seleção de sequências-sonda correspondentes a genes codificadores de nodulinas curadas de angiospermas disponíveis em bancos de dados e na literatura. As sequências candidatas foram alinhadas via tBLASTn contra o transcriptoma de *S. scabra* visando selecionar candidatos para sua caracterização estrutural e validação da expressão. Para avaliar a diversidade estrutural, foram geradas árvores fenéticas. Paralelamente, o perfil de expressão *in silico* das nodulinas precoces foi avaliado, desenhando-se *primers* para validação de sua expressão. Foram detectados dois grupos gênicos de nodulinas precoces ENOD e NIN, ambos apresentando dois domínios conservados. Para ambos os grupos foi possível identificar diferentes resíduos conservados. A predição da localização subcelular indicou que os transcritos do grupo ENOD estão associados à membrana celular, enquanto a maioria dos transcritos NIN foi endereçada ao núcleo. Em relação ao perfil de expressão, 33 transcritos ENOD apresentaram dados estatisticamente suportados, sendo 16 induzidos e 17 reprimidos. Para NIN, do total de 55 que apresentaram expressão válida, 21 apresentaram-se como induzidos e 34 como reprimidos. A partir dos dados aqui apresentados e considerando-se a importância agrônômica de *S. scabra*, nota-se a relevância do estudo, com transcritos-candidatos modulados que têm potencial para estudos de validação, além de auxiliar no entendimento de mecanismos moleculares associados à nodulação sob condições de estresse.

Palavras-chave: Ômicas; Bioinformática; Estilosantes; Regulação gênica.

ABSTRACT

New advances in technologies related to omics favored the identification of signaling mechanisms that regulate plant responses to stress. Nodulation and the ability to fix attach in association with bacteria of the genus *Rhizobium* spp. may determine plants survival in adverse environments. Among the several classes of genes that act under stress, the nodulins deserve mentioning, as host genes expressed during the nodulation process, being also associated with nodule functioning. Recent studies report that *Stylosanthes scabra* Vogel has been widely used as forage, presenting high tolerance to environments with poor soils and low water availability. The present project aimed to evaluate the response mechanisms used by *S. scabra* under abiotic stress (watering suppression), specifically the differential gene expression of early nodulins at different moments after the onset of stress. For this purpose, a selection of seed-sequences was performed, corresponding to cured angiosperm nodulin coding genes available in databases and in the literature. Selected nodulin seed sequences were aligned via tBLASTn against *S. scabra* transcriptome to select candidates for structural characterization and expression validation. To evaluate the structural diversity, phenological trees were generated. In parallel, the in silico expression profile of the early nodulins was evaluated, and primers were designed for validation of their expression. Two groups of early nodulins have been detected: ENOD and NIN, both exhibiting two conserved domains. For both groups it was possible to identify different conserved residues. Prediction of subcellular localization indicated that ENOD transcripts are associated with the cell membrane, whereas most NIN transcripts were addressed to the nucleus. Regarding the expression profile, 33 ENOD transcripts presented statistically supported data, 16 of them induced and 17 repressed. For NIN, of the total of 55 that presented valid expression, 21 were presented as inductees and 34 as repressed. From the data presented here and considering the agronomic importance of *S. scabra*, the relevance of the study is noteworthy, with modulated candidate transcripts that have potential for validation studies, besides helping to understand the molecular mechanisms associated with nodulation under stress conditions.

Keywords: Omics; Bioinformatics; *Stylosanthes*; Gene regulation.

LISTA DE FIGURAS

CAPÍTULO I – Revisão da literatura

Figura 1. Esquema simplificado do mecanismo de nodulação em leguminosas. 24

Figura 2. Representação da percepção do fator Nod pela respectiva dos receptores 37

CAPÍTULO II – Artigo

Figura 1. Fluxograma das etapas presentes da pesquisa para identificação e caracterização das sequências 53

Figura 2. Representação dos domínios conservados dos grupos de nodulinas precoces. 55

Figura 3. Motivos conservados dos grupos gênicos de nodulinas precoces. 57

Figura 4. Alinhamento das sequências do grupo ENOD de nodulinas precoces. 58

Figura 5. Alinhamento das sequências do grupo NIN de nodulinas precoces. 59

Figura 6. Representação esquemática dos 23 transcritos do grupo ENOD ancorados em cinco cromossomos utilizando genoma de referência de *Arachis duranensis*. 61

Figura 7. Representação esquemática dos 36 transcritos do grupo *NIN* ancorados em seis cromossomos utilizando genoma de referência de *Arachis duranensis*. 62

Figura 8. Perfil de expressão *in silico* de nodulinas precoces no transcriptoma de *S. scabra*. 64

Figura 9. Dendrograma refletindo padrões de distribuição do domínio osENOD1_like e demais espécies. 67

Figura 10. Dendrogramas refletindo padrões de distribuição do grupo NIN.	68
Figura 11. Estruturas tridimensionais das nodulinas	70
Figura 12. Análise estereoquímicas dos modelos tridimensionais das nodulinas modeladas	71
Figura 13. Análise da estrutura tridimensional das nodulinas após simulação em dinâmica molecular.	72
Figura 14. Análise do espalhamento dos átomos dos modelos tridimensionais das nodulinas modeladas.	73

LISTA DE TABELAS

CAPÍTULO II – Artigo

Tabela 1. Dados da mineração dos candidatos de nodulinas precoces.	55
Tabela 2. Dados de caracterização estrutural e funcional do grupo ENOD de nodulinas precoces.	55
Tabela 3. Dados de caracterização estrutural e funcional do grupo NIN de nodulinas precoces.	56
Tabela 4. Desenho de primers dos candidatos de nodulinas precoces.	65

LISTA DE ABREVIATURAS E SIGLAS

Item	Definição
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
cDNA	<i>Complementary DNA</i> (DNA complementar)
DNA	<i>Deoxyribonucleic acid</i> (Ácido desoxirribonucleico)
FBN	Fixação biológica do nitrogênio
FTs	Fatores de transcrição
ID	Identidade
KDa	Kilodalton
mRNA	RNA mensageiro
NCBI	<i>National Center for Biotechnology Information</i> (Centro Nacional de Informação Biotecnológica)
NF	Fatores Nod
pb	Pares de base
PCR	Reação em cadeia da polimerase
P.I	Ponto Isoelétrico
p.M	Peso Molecular
RNA	<i>Ribonucleic Acid</i> (Ácido ribonucleico)
RNA-Seq	Sequenciamento de RNA
RT-PCR	Reação em Cadeia da Polimerase com transcrição reversa
<i>S. scabra</i>	<i>Stylosanthes scabra</i> Vogel
Tm	Temperatura de melting
UniProt	<i>Universal Protein Resource</i> (Recurso Universal de Proteínas)

SUMÁRIO

1 INTRODUÇÃO	16
2 OBJETIVOS	18
2.1 Geral	18
2.2 Específicos	18
3 REVISÃO DA LITERATURA	19
3.1 Resposta dos vegetais frente a condições de estresse	19
3.1.1. Estresses abióticos	20
3.1.2. Estresse salino	20
3.1.3. Déficit hídrico.....	21
3.1.4. Plantas tolerantes e seus mecanismos de sobrevivência	22
3.1.4.1. <i>Nodulação</i>	23
3.2. Família Fabaceae	25
3.2.1. Importância econômica	26
3.2.2. Gênero <i>Stylosanthes</i>	26
3.3 Ômicas	30
3.3.1. Plataformas de sequenciamento e geração de dados	30
3.3.2. Transcriptômica de espécies vegetais.....	31
3.3.4. Ômicas no entendimento da fixação de nitrogênio.....	34
3.3.5. Nodulinas: estrutura e função molecular	35
3.4 Bioinformática na caracterização de sequências	38
REFERÊNCIAS	40
CAPÍTULO II	49
5 CONSIDERAÇÕES FINAIS	80
REFERÊNCIAS	82
ANEXO	85
Material suplementar	86

1 INTRODUÇÃO

Sabe-se que o processo de nodulação é extremamente importante, particularmente para leguminosas, incluindo o arbusto forrageiro *Stylosanthes scabra* Vogel (Fabaceae). Tem-se conhecimento de que essa espécie forrageira é especialmente tolerante à seca e à salinidade, compreendendo um recurso genético com potencial para doador de genes para uso biotecnológico em outras angiospermas, com ênfase para leguminosas.

Em geral, variações climáticas, bem como diferentes condições de solo, acabam afetando o desenvolvimento dos vegetais, e comprometendo seu crescimento. Entretanto, algumas espécies apresentam elevada tolerância a ambientes com pouca disponibilidade hídrica, como é o caso de *S. scabra*, sendo bastante utilizadas como forrageiras em diferentes regiões. Essa espécie – como outras do gênero *Stylosanthes* – apresenta tais características, sendo adequada para cultivo em regiões semiáridas e com solos com baixa disponibilidade de nutrientes.

A capacidade de realizar fixação biológica do nitrogênio (FBN) está entre os principais atributos de leguminosas, destacando-se que espécies desse grupo mantêm essa capacidade mesmo sob condições de estresse abiótico, incluindo seca e salinidade moderada. Dentre os diferentes genes associados à nodulação e que atuam em situações de estresse as nodulinas merecem menção, sendo expressas pelo hospedeiro durante a formação do nódulo nas raízes das leguminosas e no processo da fixação biológica do nitrogênio (FBN). Como esperado, a nodulação em *S. scabra* é afetada pelos estresses, havendo redução no crescimento. Apesar disso, o fato de continuar nodulando e de se adaptar à falta de água, justifica uma análise, identificação e validação da expressão de genes relacionados ao processo de simbiose, como as nodulinas, de modo a gerar um perfil mais completo sobre a participação destes genes na FBN, visto que estes são responsáveis pelo início e manutenção da nodulação, sendo estes importantes para as espécies leguminosas, como o caso de *Estilosantes*.

Genes estresse-induzidos compreendem moléculas indutoras de tolerância, assim como reguladores da expressão, ativos na transdução de

sinais de diversas vias metabólicas (NAKASHIMA; YAMAGUCHI-SHINOZAKI, 2014). O conjunto de todas as interações entre fatores de transcrição (FTs) e seus genes-alvo em uma célula pode ser representado como uma rede de interação (CHEN; ZHU, 2004).

Pesquisas de bioinformática associadas à biologia molecular visando identificar tais genes vêm permitindo a geração de dados acerca dos mecanismos moleculares frente às condições de estresses, favorecendo os estudos em melhoramento genético, no que se refere a aplicações de novas técnicas que visem à criação de plantas mais resistentes ou tolerantes a diversas condições.

Considerando o exposto, o presente trabalho visa identificar, caracterizar e validar a expressão diferencial de genes codificantes de nodulinas precoces no transcriptoma de *S. scabra* em diferentes momentos após a suspensão da irrigação, a fim de propiciar um melhor entendimento sobre o processo de nodulação no âmbito de um estresse abiótico (supressão de rega) em tempo precoce. Além disso, pretende-se identificar genes candidatos promissores para aplicação em inferências biotecnológicas voltadas para a FBN em leguminosas e em programas de melhoramento genético vegetal da espécie.

Vale destacar, que estudos transcriptômicos visando entender a participação dos genes de nodulinas, tanto na nodulação, quanto na FBN, ainda são escassos, especialmente, tratando-se de *S. scabra*, ressaltando a importância de estudos que visam esta compreensão, tendo em vista a relevância das leguminosas forrageiras para desenvolvendo de cultivares cada vez mais adaptadas em diferentes situação e ambientes.

2 OBJETIVOS

2.1 Geral

Avaliar os mecanismos de resposta utilizados por *Stylosanthes scabra* Vogel sob déficit hídrico, no que se refere à identificação, caracterização e expressão gênica diferencial de nodulinas precoces em diferentes momentos após a suspensão da irrigação.

2.2 Específicos

- Identificar a diversidade estrutural de famílias gênicas de nodulinas precoces em *S. scabra* bem como os seus domínios funcionais, visando à determinação de suas características comparativamente a outras angiospermas, através de análise *in silico*;
- Expressão das nodulinas na resposta ao déficit hídrico (supressão de rega) em *S. scabra* comparativamente a leguminosas taxonomicamente próximas, como o amendoim (*Arachis hypogaea* L.) e outras espécies de *Arachis*;
- Selecionar genes candidatos atuantes no processo de nodulação em *S. scabra* para validação de sua expressão diferencial via RT-qPCR e posterior aplicação em programas de melhoramento da espécie.

3 REVISÃO DA LITERATURA

3.1 Resposta dos vegetais frente a condições de estresse

Alterações climáticas afetam diretamente o desenvolvimento e sobrevivência dos vegetais. Neste sentido, estudos que envolvem condições de estresse são utilizados para avaliar as respostas das plantas e sua tolerância sob condições adversas. Em ambientes onde muitas vezes os vegetais estão expostos à seca ou a salinidade durante o seu período de crescimento, os indivíduos aclimatados frequentemente conseguem sobreviver e completar o seu ciclo reprodutivo (WALTER et al., 2013).

Foi definida como “marca de estresse” uma modificação genética ou bioquímica que ocorre na planta após a exposição a um determinado tipo de estresse, demandando da planta respostas mais rápidas em eventos subsequentes de estresse (BRUCE et al., 2007; WALTER et al., 2011). Dessa forma, os vegetais desenvolveram mecanismos específicos que lhes permitem detectar mudanças ambientais, desencadeando respostas moleculares visando minimizar danos e conservar recursos valiosos para o seu crescimento (MITTLER; BLUMWALD, 2010).

De uma forma geral, o mecanismo de defesa se inicia com a percepção do estresse, desencadeando uma cascata de eventos moleculares, levando à transmissão e processamento de sinais para geração de respostas mediadas por fitormônios, mensageiros secundários, fatores de transcrição (FTs, *Transcription Factors*), além de genes e suas proteínas sinalizadoras codificadas (PANDEY, 2008). A ativação dos genes envolvidos nesses mecanismos é minuciosamente controlada por um grande número de proteínas reguladoras, as quais constituem a segunda estratégia de defesa e podem atuar em dois níveis da expressão gênica: (1) nível transcricional (quantidade de mRNA produzido) e/ou (2) nível traducional (quantidade de proteínas produzidas) (ALVES et al., 2014).

As respostas vegetais aos estresses são complexas, envolvendo mudanças nos níveis fisiológico, bioquímico e molecular, ativando níveis de expressão gênica referente à condição ambiental que lhes está sendo imposta (ATKINSON; URWIN, 2012). Para a maioria dos genes, essa regulação é realizada em nível transcricional, garantindo um menor gasto energético da

célula, além de níveis adequados de proteínas (MATOS, 2015). Dentre os mecanismos regulatórios transcricionais, estão incluídos: a remodelagem da cromatina, a organização/reorganização dos nucleossomos, a metilação do DNA e as interações envolvendo TFs (FELLER et al., 2011).

Frio, seca e alta salinidade são condições de estresse comuns que afetam negativamente o crescimento da planta e sua produção. Compreender os mecanismos pelos quais as plantas transmitem os sinais para a maquinaria celular a fim de ativar respostas adaptativas é de fundamental importância para desenvolver culturas mais tolerantes, melhorando a eficiência da produção (BEZERRA-NETO, 2016).

3.1.1. Estresses abióticos

Como organismos sésseis, as plantas estão mais expostas a diversas condições de estresse, tais como: variações de temperatura e intensidade da luz, inundações, seca, salinidade e presença de metais pesados no solo (CRAMER et al., 2011). As plantas requerem, fundamentalmente, energia proveniente da luz, água, carbono e nutrientes minerais para seu crescimento. Estresses abióticos são definidos como condições ambientais que reduzem o crescimento e a produção vegetal abaixo dos níveis ótimos (CRAMER et al. 2011) e se constituem em severos limitantes da produção vegetal. Quando tais estresses ocorrem, várias respostas bioquímicas e fisiológicas são induzidas nas plantas, de forma a propiciar a tolerância ou aumentar as chances de sobrevivência às condições adversas (BENKO-ISEPPON et al., 2011).

3.1.2. Estresse salino

O efeito da salinidade sobre o desenvolvimento das plantas é um assunto discutido em vários países, principalmente, nos que apresentam regiões áridas e semiáridas. A salinidade pode ser definida como a situação de excesso de sais solúveis, sódio trocável ou ambos em horizontes ou camada superficiais, afetando o desenvolvimento e o metabolismo vegetal desde a sua germinação, provocando redução no crescimento e conseqüentemente na produtividade vegetal (RIBEIRO et al., 2009; SCHOSSLER et al., 2012).

Sabe-se que o estresse salino inibe o crescimento das plantas, por aumentar o potencial osmótico da solução do solo, restringindo a disponibilidade

da água e/ou pela acumulação excessiva de íons nos tecidos vegetais, podendo ocasionar toxicidade iônica e desequilíbrio nutricional ou ainda ambas as situações. Além disso, é responsável por grande impacto negativo no crescimento vegetal, sendo capaz de promover alterações em funções metabólicas, fisiológicas e anatômicas (DUTRA, et al., 2017).

O conhecimento sobre a influência do estresse na germinação de sementes tem grande importância, especialmente para a ecofisiologia, pois tem como objetivo avaliar os limites de tolerância e a capacidade de adaptação de espécies ou acessos em estudo. Adicionalmente, a capacidade de germinação das sementes de algumas espécies ou acessos sob condições de estresse confere vantagens ecológicas em relação a outras que são sensíveis (LAVEZO et al., 2015).

3.1.3. Déficit hídrico

A água é um elemento fundamento para crescimento e desenvolvimento vegetal e a diminuição da disponibilidade de água no solo leva ao déficit hídrico. Segundo Ferreira (2017), este déficit pode causar desbalanço nutricional, redução do potencial hídrico, redução na absorção de água do solo e redução na fotossíntese. Entre os fatores ambientais, a água é o recurso mais limitante e compõe a maior proporção do volume celular nas plantas, além de representar um fator fundamental na produção vegetal (BITTENCOURT; SILVA, 2018). Ainda segundo estes autores, a diminuição da disponibilidade hídrica causa um efeito deletério no crescimento e desenvolvimento da planta, influenciando seu ciclo de vida diretamente, com evidentes em todos os estágios fenológicos do crescimento vegetal. Quando em contato com o estresse hídrico, o fechamento estomático figura como uma das primeiras respostas da planta, provocando uma redução na disponibilidade de dióxido de carbono (CO_2) nas folhas, além de inibir a fixação de carbono com consequente redução das taxas fotossintéticas (LISAR et al., 2012; OSAKABE et al., 2014).

Algumas plantas desenvolveram mecanismos capazes de diminuir os efeitos da falta de água no solo, que podem ser transmitidos geneticamente, tais como: aumento do sistema radicular, redução da área foliar, ajuste osmótico e fechamento estomático, entre outros (TAIZ; ZEIGER, 2009). A tolerância à seca resulta de várias características, as quais podem variar dependendo da

severidade e da taxa de imposição do déficit hídrico, da idade e das condições nutricionais da planta, do tipo e da profundidade do solo, da carga pendente de frutos, da demanda evaporativa da atmosfera e da face de exposição do terreno (MATOS et al., 2014).

3.1.4. Plantas tolerantes e seus mecanismos de sobrevivência

A adaptação das plantas ao estresse se dá de diversas maneiras, como por exemplo, a manutenção da homeostase celular, detoxificação de compostos nocivos e alterações no crescimento. Estes mecanismos são ativados em função do tempo de duração e da intensidade do estresse e processam-se nos variados níveis de complexidade (SILVEIRA; PANTALIÃO; BRONDAN, 2015).

Algumas plantas apresentam adaptações para ambientes com condições abióticas extremas como, por exemplo, seca severa e altas temperaturas. Plantas adaptadas a ambientes de seca e calor extremos apresentam adaptações específicas sem as quais sua ocorrência e sobrevivência seriam ameaçadas. Juntos, os estresses abióticos são responsáveis por desencadear uma série de respostas das plantas, que podem ser percebidas através de modificações morfológicas, fisiológicas, moleculares e metabólicas, a fim de tolerar tais estresses (BENKO-ISEPPON et al., 2011).

Dentre as mudanças morfológicas e fisiológicas em condição de déficit hídrico, o número e tamanho das folhas reduzido são as mais expressivas e indispensáveis para retardar a desidratação (SANTANA et al., 2011). O ajustamento da área foliar total e do número de folhas são necessários para suportar o estresse imposto e aclimatar-se à nova condição (MATOS et al., 2014), bem como o aumento do sistema radicular (volume e/ou profundidade) visando à mais eficiente captação de água (SILVEIRA; PANTALIÃO; BRONDAN, 2015).

Com relação aos aspectos moleculares, de acordo com Bittencourt e Silva (2018) durante o estresse hídrico, algumas proteínas podem ser reguladas negativamente (reprimidas) ou positivamente (induzidas). As proteínas induzidas protegem as plantas na adaptação ao estresse e aumentam sua resistência. Dessa forma, muitos estudos têm se dedicado a caracterizar e compreender os genes que codificam tais proteínas, incluindo nodulinas, estas associadas aos processos fisiológicos, assim como à sua expressão, podem facilitar a

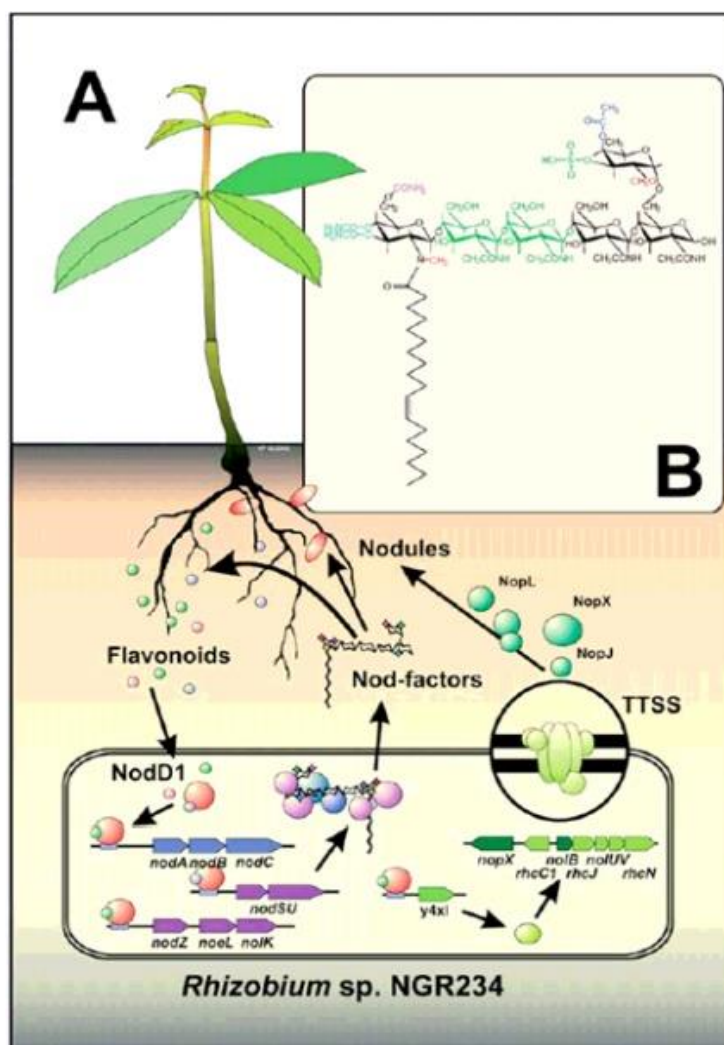
identificação de genes reguladores e funcionais, responsivos ao estresse, e com possibilidade de aplicação em eventos de transformação genética ou como marcadores moleculares de déficit hídrico. Ainda segundo os mesmos autores, uma maior compreensão das bases moleculares de resistência à seca associada a combinações de desempenho ecológico, anatômico, morfológico e fisiológico possibilita avanços na bioengenharia visando à manipulação genética para obtenção de culturas tolerantes à seca.

3.1.4.1. Nodulação

As plantas da família Fabaceae possuem um mecanismo de extrema importância para o desenvolvimento e crescimento vegetal devido à sua capacidade de realizar simbiose com bactérias, propiciando o desenvolvimento de nódulos nas raízes das plantas e a fixação biológica do nitrogênio (FBN). A capacidade de realizar essa fixação é de grande importância, culminando com o uso de rizóbios na agricultura (BEZERRA-NETO, 2016).

A formação de um simples nódulo resulta de um processo altamente complexo, envolvendo uma cascata metabólica com múltiplas etapas. De forma geral, o sistema radicular das plantas é altamente influenciado pelos exudatos radiculares, que são compostos orgânicos capazes de atrair alguns microrganismos que, por sua vez, podem ter seu crescimento favorecido. Os genes do hospedeiro quando ativados, induzem a bactéria a produzir outras moléculas. Tais bactérias pertencem ao gênero *Rhizobium*, sendo conhecidas como rizóbios, destacando-se por sua capacidade fixadora de nitrogênio e pela capacidade de comunicação com as plantas através de moléculas denominadas de fatores Nod (NFs), associadas ao processo de iniciação da nodulação (HUNGRIA; CAMPO; MENDES, 2007). Tal processo resulta na formação de uma nova membrana dentro das células do hospedeiro, advinda da troca de sinalização e desenvolvimento de metabólitos. Nesse cenário, a planta fornece carboidratos aos bacteroides, que por sua vez, respondem através de diferenciação e aumento do tamanho das colônias de bactérias – provenientes de sua multiplicação – as quais oxidam os carboidratos e os utilizam para obter energia (**Figura 1**). Trata-se de um processo altamente benéfico para a planta e para a bactéria, sendo limitado a alguns grupos vegetais (Revisado por SANTOS; REIS, 2008).

Figura 1. Esquema simplificado do mecanismo de nodulação em leguminosas. (A) Flavonóides excretados pelas raízes ativam a expressão de genes do rizóbio requeridos para o processo de nodulação. (B) Fatores Nod são do lipo-quitooligosacarídeos modificados como os β -1,4-oligômeros conectados de N-acetil-D-glucosamina, com um ácido graxo substituindo o grupo N-acetil em suas terminações não redutoras (BROUGHTON et al., 2003).



Dentre o processo de nodulação, destaca-se as nodulinas, que são genes expressos pelo hospedeiro, nos quais são responsáveis por dar início e participar da manutenção da formação dos nódulos, assim como participam da FBN. Apesar do processo inicial da nodulação ainda não ser bem elucidado, vários genes já foram descritos como nodulinas, sendo divididos em função do processo em que participam, por exemplo, no reconhecimento e transdução dos sinais dos fatores Nod rizobiais, culminando na expressão gênica de outros

genes requeridos na nodulação e na formação do nódulo radicular primordial (MELLO, 2009).

3.2. Família Fabaceae

A família Fabaceae ou Leguminosae é considerada uma das mais importantes dentre as angiospermas (Eudicotiledôneas), sendo a terceira maior desse grupo, conta com 727 gêneros e cerca de 19.325 espécies, distribuídas em seis subfamílias, a saber: Faboideae, Duparquetioideae, Caesalpinioideae, Cercidoideae, Detarioideae e Dialioideae (AZANI, 2017). Compreende um dos grupos mais diversificados da flora mundial, com distribuição cosmopolita e espécies nativas em todos os continentes, com exceção apenas da Antártida (LEWIS et al., 2005).

No Brasil, a família Fabaceae possui 222 gêneros e 2.837 espécies, incluindo 1535 endêmicas (LIMA et al., 2017). São consideradas responsáveis pela grande diversidade em florestas tropicais, tendo papel de destaque como elemento florístico nas principais formações florestais brasileiras, estima-se que o Cerrado seja o bioma que abriga maior diversidade de espécies já descritas (28,68%), seguido pela Amazônia (27,57%), Mata Atlântica (23,25%), Caatinga (15,28%), Pampa (2,69%) e Pantanal (2,54%) (LIMA et al., 2012). Os representantes desta família botânica possuem hábitos de crescimento variados, existindo árvores de pequeno, médio e grande porte, ervas anuais ou perenes, arbustos e lianas, entre outros (SOUZA; LORENZI, 2005; SOUZA; SOUZA, 2011).

Do ponto de vista ecológico, de acordo com Lewis (1987), as Fabaceae constam como elemento principal de diferentes formações vegetais, desde os picos das serras montanhosas até o litoral arenoso, e da floresta úmida tropical até desertos, havendo também espécies aquáticas. Seus representantes são frequentemente bem adaptados à primeira colonização, considerados pioneiros e aptos à exploração de ambientes abertos e perturbados, devido, em parte, às suas associações com ectomicorrizas ou bactérias fixadoras de nitrogênio do gênero *Rhizobium* (que são capazes de converter nitrogênio atmosférico em amônia), sendo tais leguminosas extremamente valiosas como fornecedoras de adubos naturais (LEWIS, 1987). Adicionalmente, o cultivo de leguminosas forrageiras proporciona grandes benefícios ao sistema de produção, pois

contribui com a diversificação das pastagens e possibilita o cultivo mesmo em solos mais pobres (GAMA 2013).

3.2.1. Importância econômica

A família das Fabaceae constitui um dos grupos botânicos de maior importância econômica. A maioria de suas espécies é cultivada desde a antiguidade devido principalmente ao seu potencial alimentício e medicinal, embora existam várias outras utilidades diretas, como: forrageiras, melíferas, ornamentais, madeiras, como fornecedoras de celulose, óleos, adubo verde, carvão, lenha, resinas, tintas, vernizes e cortiça (MIOTTO; LÜDTKE; OLIVEIRA, 2008; SOUZA; SOUZA, 2011).

Dentre as forrageiras, as maiores contribuições referem-se ao seu potencial de incorporar nitrogênio ao solo e disponibilizar melhor desenvolvimento das pastagens, possibilitando a redução do uso de fertilizantes nitrogenados e elevando o potencial produtivo (LOPES et al., 2011). As leguminosas também se destacam por aumentar o carbono orgânico disponível no solo, além de suprimir o crescimento de plantas daninhas (WARWICK et al. 2016).

Destaca-se que, devido à importância agrônômica, a família é superada apenas pelas gramíneas (QUEIROZ, 2009). De acordo Lewis (1987), as leguminosas possuem maior potencial para a produção de proteína vegetal, prevendo-se que futuramente será consumida em quantidades cada vez maiores pela população mundial. No entanto, segundo Costa (2017), a adoção de leguminosas nativas em pastagens ainda é muito baixa, estando aquém de seu potencial. Embora já existam progressos nesse sentido, o custo de sementes das variedades mais bem-sucedidas ainda limita o seu uso extensivo.

3.2.2. Gênero *Stylosanthes*

O gênero *Stylosanthes* é originário das Américas Central e do Sul e representa um importante grupo de leguminosas (revisado por GALDINO, 2014). Em sua maioria, são plantas perenes, com sistema radicular bem desenvolvido, tolerantes à seca, com grande capacidade de adaptação a solos de baixa fertilidade sendo que a maioria das espécies apresenta excelente capacidade de fixar nitrogênio atmosférico (UBIALI et al., 2013).

As espécies deste gênero chamam atenção devido ao seu potencial de forragem, com excelente qualidade nutricional e fácil adaptação a diferentes condições ambientais (CALLES; ESCHULTZE-KRAFT, 2016). Além da reconhecida capacidade produtiva, *Stylosanthes* é citado por sua capacidade de restaurar a fertilidade (FABRICE, 2015) e melhorar as propriedades físicas do solo, desta forma fornecendo ótima cobertura vegetal (NEGREIROS et al. 2010).

A maioria das espécies de *Stylosanthes* é autógama. Porém, podem apresentar taxas de fecundação cruzada variáveis para cada espécie (VIEIRA et al., 2007). O número de cromossomos também é relativamente variável, com diferentes níveis de ploidia, como por exemplo $2n=40$ (*S. scabra*) e $2n=20$ (*S. seabrana*) (ANDRADE LIRA, 2015). As abordagens citogenéticas convencionais demonstraram que a maioria das espécies do gênero são diploides com $2n = 20$ e alguns são tetraploide ($2n = 40$) ou hexaploide ($2n = 60$) (MAASS; SAWKINS, 2004).

Estudos moleculares prévios mostraram que os poliploides são provavelmente formados por hibridação interespecífica (alopoliploidia) (MARQUES et al., 2018).

- Taxonomia

O gênero *Stylosanthes* é bastante complexo taxonomicamente, sendo composto por 50 espécies, com 25 delas de ocorrência no Brasil (SANTOS-GARCIA, 2011). Levantamentos para a ocorrência desse gênero no semiárido Nordeste determinaram a ocorrência de quatro espécies (*S. capitata* Vogel, *S. humilis* Kunth, *S. scabra* Vogel e *S. viscosa* (L.) W.), revelando grande diversidade em diferentes locais e mostrando que existe variabilidade significativa no gênero a ser explorada (KARIA, 2008; OLIVEIRA, 2015). A distinção das espécies dentro do gênero tem sido realizada predominantemente a partir de caracteres morfológicos (COSTA, 2017).

O gênero é dividido em duas seções, baseadas na presença/ausência de eixo plumoso na base de flores e/ou frutos e no número de bractéolas. A primeira seção é caracterizada por apresentar eixo rudimentar na base e três bractéolas; entre as espécies dessa seção estão *S. macrocephala*, *S. hamata* e *S. scabra*. Já a segunda seção não possui eixo rudimentar na base das flores e frutos e apresenta duas bractéolas, como *S. humilis*, *S. angustifolia*, *S. guianensis*, entre

outras (COSTA et al. 2008).

- Diversidade

As leguminosas forrageiras do gênero *Stylosanthes* apresentam elevada diversidade morfológica e agronômica, são nativas das regiões tropical e subtropical, onde são conhecidas 48 espécies, com 44 delas ocorrentes apenas na América. No Brasil, já foram catalogadas 29 espécies, onde 13 delas são endêmicas (Revisado por COSTA, 2006).

A espécie *S. scabra* (Vogel) possui grande destaque no âmbito do gênero *Stylosanthes*, ocorrendo espontaneamente em florestas abertas, pastagens, várzeas, bancos de diques, estradas, locais perturbados e áreas de resíduos, figurando entre culturas de regiões tropicais e subtropicais. Essa espécie tem sido bastante utilizada como forrageira, por apresentar elevada tolerância em ambientes com solos pobres e com baixa disponibilidade hídrica. Desta forma, constitui-se uma alternativa para a melhoria, contribuindo para estruturação (DEDECK; GALDINO; VIERA, 2006), e descompactação do solo (NEGREIROS NETO et al., 2010). Em países de climas semelhantes ao do semiárido brasileiro, como por exemplo, a Austrália, países africanos e asiáticos, as espécies desse gênero estão sendo empregadas em pastagens, devido à alta produção de massa verde e a tolerância a estresses abióticos (RESENDE; RESENDE; LAURA 2006).

Nagaich et al. (2013) avaliaram diferentes acessos de *S. scabra* (Vogel) em condições climáticas de semiárido da Índia e encontraram indivíduos tolerantes à seca que apresentam alta correlação positiva com a produção de biomassa. Entretanto, no Brasil, poucos estudos vêm sendo realizados, embora os resultados alcançados até o momento demonstrem uma riqueza e alta diversidade das espécies desse gênero em diferentes regiões do país (OLIVEIRA, 2015).

- Bancos Genéticos

Diversos bancos de germoplasma das espécies do gênero *Stylosanthes* são encontrados principalmente fora do Brasil, como é o caso da Austrália, por exemplo, onde a espécie é bastante cultivada. No Brasil, existem programas de melhoramento genético em forrageiras tropicais (gramíneas e leguminosas

forrageiras) que apresentam resultados satisfatórios. A Embrapa lidera as pesquisas com estas espécies tropicais há mais de três décadas (JANK et al., 2011) e tem contribuído com mais de 70 % das cultivares em uso no país, com destaque para os programas desenvolvidos com *Arachis* spp., *Brachiaria* spp., *Panicum maximum* Jacq., *Pennisetum* spp., além de espécies de *Stylosanthes* (OLIVEIRA; QUEIRÓZ, 2016).

Diversos trabalhos de coleta e conservação de germoplasma de *Stylosanthes* spp. são relatados na literatura envolvendo a região semiárida do Brasil. Romão et al. (2008), coletaram e identificaram diversos acessos e espécies de estilosantes na região semiárida da Bahia, sendo elas: *S. scabra* Vogel, *S. pilosa* M.B. Ferreira & N.M. de Sousa Costa, *S. capitata* Vogel, *S. angustifolia* e *S. viscosa* S.W.A. Mistura et al. (2008) relataram um grande percentual de ocorrência (77,3 %) de *Stylosanthes* em dez municípios da microrregião do Juazeiro (BA), demonstrando a importância de coletas visando conservar os recursos genéticos presentes nessa região.

- Forragem e biorremediação

Como informado anteriormente, a maioria das leguminosas estabelece simbiose com bactérias que crescem em nódulos que se formam nas raízes. Por essa razão, as leguminosas são componentes valiosos em misturas forrageiras com gramíneas e na rotação de culturas visando diminuir a dependência de fertilizantes nitrogenados (BARNES et al., 2003). Vale salientar que os adubos nitrogenados possuem custo econômico e ambiental elevado (KHOUZANI; NEJAD, 2015).

A Austrália foi o primeiro país a reconhecer o valor de *Stylosanthes* como forrageira, onde através de programas de melhoramento genético produziu espécies mais adaptadas a diferentes ambientes e com grande potencial econômico (EDYE, 1997).

A fitorremediação é uma tecnologia relativamente nova que contempla o uso de plantas para transformar ou eliminar o contaminante acumulado nos tecidos vegetais. Essa tecnologia tem diversas vantagens, incluindo baixo custo, se comparado a técnicas tradicionais de descontaminação, além de manter as propriedades biológicas e físicas do solo (revisado por LEAL, 2014). Tendo conhecimento de que as espécies do gênero *Stylosanthes* possuem a

característica de associação simbiótica, a utilização da fitorremediação pode ser uma alternativa para auxiliar na absorção e transformação de contaminantes (DE MELO et al., 2009).

3.3 Ômicas

Novos avanços nas tecnologias relacionadas às “ômicas” estão favorecendo a identificação de mecanismos de sinalização transcricional, traducional e pós-traducional, que regulam as respostas vegetais ao estresse (AHUDA et al., 2010). Nesse sentido, uma melhor caracterização e compreensão das respostas vegetais frente ao estresse contribuirão para o desenvolvimento de novas tecnologias e cultivares mais adaptadas ou tolerantes a períodos de déficit hídrico e salinidade, ocasionando impacto direto na manutenção e incremento da produção agrícola mundial (BEZERRA-NETO, 2016).

A análise de dados transcriptômicos em associação com modelos de redes metabólicas vem auxiliando na exploração de dados de maneira global, revelando genes relacionados à tolerância em condições de estresses e sua participação em processos metabólicos importantes, identificando-se pontos de convergência de regulação (GALHARDO et al., 2013).

3.3.1. Plataformas de sequenciamento e geração de dados

As novas tecnologias de sequenciamento, denominadas em inglês de “*Next Generation Sequencing*” (NGS, em português Sequenciamento de Nova Geração), começaram a ser comercializadas em 2005 e vêm evoluindo rapidamente. Tais essas tecnologias promovem o sequenciamento de DNA em plataformas capazes de gerar informação sobre milhões de pares de bases em uma única corrida (CARVALHO; SILVA, 2010). Atualmente, os sequenciadores estão divididos em 1ª geração (Degradação química e Interrupção da cadeia-Sanger), 2ª geração (454-Roche, Illumina Genome Analyser-HiSeq/MiSeq e SOLID), 3ª geração (Ion Torrent e PacBio RS) e 4ª geração (Nanopore) (SANTOS et al., 2013).

O sequenciamento de RNA (RNA-Seq), também denominado “*Whole Transcriptome Shotgun Sequencing*” (WTSS), engloba um conjunto de técnicas experimentais e computacionais que permitem identificar a abundância de

sequências de RNA em amostras biológicas em um determinado estágio de desenvolvimento ou sobre determinada condição fisiológica (KORPELAINEN et al., 2015; WANG; GERSTEIN; SNYDER, 2009). Compreender o transcriptoma de um organismo é essencial para identificar e interpretar os elementos genômicos, como, por exemplo, genes e RNAs não codificantes, bem como na forma como estes atuam nos processos biológicos (WANG; GERSTEIN; SNYDER, 2009).

A informação gerada por projetos de sequenciamento necessita ser armazenada e organizada em bancos de dados disponíveis *online* para que outros pesquisadores tenham acesso à informação. Bancos de dados genômicos armazenam sequências de DNA, RNA, proteínas e suas anotações (SELZER; MARHÖFER; ROHWER, 2008b). Os bancos de dados do Centro Nacional de Informação de Biotecnologia (NCBI, do inglês *National Center for Biotechnology Information*), do *European Molecular Biology Laboratory* (EMBL, Laboratório Europeu de Biologia Molecular) e Banco de Dados de DNA do Japão (DDBJ, *DNA Data Bank of Japan*) são os maiores bancos de dados disponíveis atualmente (FERREIRA, 2017).

3.3.2. Transcriptômica de espécies vegetais

O estudo do transcriptoma, que por sua vez, corresponde ao conjunto total de transcritos expressos num dado tecido e num dado momento, irá determinar a frequência de transcritos em vários tipos de células, tecidos, e situações, incluindo estágios de desenvolvimento ou situações de estresse biótico ou abiótico. Tais informações, quando disponíveis com réplicas e quando comparadas duas ou mais situações controladas, permitem inferir sobre o nível de expressão de um dado gene ou isoformas, com base no transcriptoma (ADAMS, 2008).

De maneira generalizada, o estudo de regulação de genes envolve uma comparação de populações de RNAs, como por exemplo: tratado (contendo estresse em função de algum evento) e controle (sem evento estressante). Estas informações, são obtidas a partir de uma variedade de medidas realizadas por meio de diferentes tecnologias. E, ao serem comparadas entre si, podem apresentar variações, favorecendo estudos envolvendo atividade gênica em

diferentes situações (FERREIRA-NETO; BENKO-ISEPPON; KIDO, 2018).

O conhecimento destas informações biológicas complexas figura entre os principais objetivos para a área de bioinformática, que está relacionada à aplicação e desenvolvimento de métodos de armazenamento, análise e modelagem de dados biológicos e médicos (KASABOV, 2007).

Avanços no conhecimento dos mecanismos moleculares de resposta à seca são de grande importância para entender como tais mecanismos afetam o crescimento e o desenvolvimento das plantas (SEKI, et al., 2007). Segundo Tuberosa e Salvi (2006) as novas plataformas genômicas para sequenciamento de DNA em associação com análises de bioinformática adicionaram novos *insights* sobre a base genética da tolerância à seca. No entanto, o desafio atual é descobrir como combinar essas abordagens para explicar a resposta fisiológica de acessos tolerantes. De acordo com Athar e Ashraf (2009) a exploração desses conceitos levará a nossa uma melhor compreensão de como as plantas podem tolerar adequadamente essa condição. Nesse sentido, a transcriptômica seria o primeiro passo, pois a abundância de um determinado gene que provavelmente responde à condição estressante, poderá fornecer um melhor entendimento sobre esse processo (KIDO et al., 2016; FERREIRA-NETO et al., 2018).

Em projetos que envolvem o transcriptoma de determinadas espécies sob situações específicas, são utilizadas informações oriundas do mRNA, e dessa forma, o sequenciamento de RNA mensageiro (RNAm) maduro é priorizado. Após o isolamento do RNAm, o mesmo é convertido para DNA por meio da técnica de Reação em Cadeia da Polimerase (PCR, *Polymerase Chain Reaction*) com uso de uma enzima do tipo 'transcriptase reversa' (RT-PCR, Reverse Transcriptase - *Polymerase Chain Reaction*). Com esta técnica pode-se gerar uma biblioteca de cDNA (DNA complementar ao RNAm) que representa todo o conjunto de RNAm existente na célula ou tecido no momento da extração (revisado por SILVA, 2012).

O RNA-Seq, apresenta-se como a tecnologia mais atual de sequenciamento em estudos transcriptômicos. A informação do sequenciamento é reconstruída com auxílio de ferramentas de bioinformática, que avaliarão os dados gerados e montarão as *reads* (WANG; GERSTEIN; SNYDER, 2009).

Assim, o transcriptoma é o reflexo direto da expressão gênica e avalia o

conjunto de transcritos de um organismo em determinadas situações celular, como por exemplo, quais genes estão sendo expressos em momentos de estresse no vegetal (MIR, 2005).

3.3.3. Transcriptômica e genômica de Fabaceae

Especificamente, muito pouco tem sido realizado para a construção de redes de regulação gênica para leguminosas, com exceção das espécies economicamente mais relevantes, como a soja, apesar da grande quantidade de dados de expressão gênica acumulados durante os últimos anos (BENKO-ISEPPON et al., 2012).

Com a disponibilidade do genoma completo da soja (SCHMUTZ et al., 2010) e de numerosas anotações de genes e proteínas (WANG et al., 2010) torna-se importante desenvolver e integrar os métodos de bioinformática para a construção de redes gênicas confiáveis, sejam elas por meio dos dados de expressão ou com os dados de genômica funcional (BEZERRA-NETO, 2016). Zhu et al. (2012) desenvolveram um protocolo para integrar diversas ferramentas e abordagens de bioinformática, o qual permite (a partir de dados de expressão de larga escala de soja) elucidar os mecanismos e processos biológicos por meio de redes regulatórias de desenvolvimento e crescimento celular sob determinada condição de estresse. Tal metodologia pode ser adotada para outras leguminosas, reduzindo a problemática de anotação de dados e sendo suficientemente robusta para a geração de redes a partir de dados do transcriptoma (BEZERRA-NETO, 2016).

Alguns estudos voltados para o sequenciamento de genomas e transcriptomas de espécies de Fabaceae, já vêm sendo realizados, no entanto, o quantitativo de trabalhos e a abrangência de espécies estudadas ainda são muito reduzidos, comparativamente ao número de espécies da família e à sua importância econômica, bem como ao potencial dessas leguminosas, especialmente no que se refere às espécies forrageiras.

No Brasil, em estudos de transcriptômicas de Fabaceae, a Universidade Federal de Pernambuco está no centro de uma rede de pesquisa (*The Cowpea Genome Consortium* (Cp-GC), antigo 'Projeto Brasileiro do Transcriptoma do Feijão-Caupi' (rede NordEST); <http://bioinfo03.ibi.unicamp.br/vigna/>) que

gerenciou uma gama de laboratórios para executar ensaios de genômica, transcriptômicas, em diferentes espécies sob situações de estresse.

O Cp-GC também gerou 49.820 ESTs de folhas e raízes obtidas a partir da exposição da cultura a diferentes estresses bióticos e abióticos. Essas foram agrupadas em conjunto com 264.945 ESTs depositados no banco de dados Cowpea HarvEST V.1.33 (<http://harvest.ucr.edu/>) e GenBank (NCBI), estando disponível para ancorar e anotar transcritos em análise.

Dentre os estudos realizados, Kido et al. (2011) mineraram dados em bibliotecas de SuperSAGE para a produção de um catálogo de quinases expressas em feijão-caupi sob diferentes condições de estresses (biótico e abióticos). Ferreira-Neto et al. (2018) disponibilizaram os primeiros genes de referência estatisticamente elencados para normalização de ensaios de expressão gênica em feijão-caupi sob déficit hídrico e alta salinidade.

Atualmente, o Cp-GC possui significativa quantidade de dados moleculares associados a análises globais dos transcriptomas de feijão-caupi submetidos a diferentes condições, permitindo estudar, especificamente, diversos temas ainda pouco analisados pela ótica da transcriptômica, mesmo em espécies cultivadas de importância econômica.

3.3.4. Ômicas no entendimento da fixação de nitrogênio

O desenvolvimento de técnicas de sequenciamento é de extrema importância para a ciência, no âmbito da genômica, transcriptômica e proteômica, pois através do sequenciamento de DNA, RNA e proteínas, além de dados de bioinformática, novos genes e vias metabólicas têm sido descobertos em microrganismos (NEVES, 2015; BENKO-ISEPPON et al., 2017). Segundo os mesmos autores, os estudos de ômicas e biologia de sistemas podem permitir um melhor entendimento dos aspectos integrados da biologia dos organismos, assim como inter-relações de sequências, estudos de estrutura, padrões de expressão gênica em diferentes ambientes. Tais estudos podem envolver relações simbiote-hospedeiro, incluindo interações entre bactéria e planta hospedeira durante a formação de nódulo nas raízes e conseqüentemente FBN (NEVES, 2015).

Dessa maneira, as ômicas abrangem a análise global dos sistemas

biológicos, identificando, quantificando e caracterizando os componentes da maquinaria celular (BENKO-ISEPPON et al., 2017). Nesta perspectiva, pode-se ter, por meio das ômicas, um conhecimento maior sobre o processo de FBN através da formação dos nódulos nas raízes de leguminosas, sendo possível identificar outras atividades importantes com o auxílio de ferramentas biotecnológicas associadas a estas.

3.3.5. Nodulinas: estrutura e função molecular

Mais de 40 genes simbióticos envolvidos na formação de nódulos e no processo de FBN, denominados de nodulinas já foram identificados (SEIDO, 2017). As nodulinas são genes do hospedeiro expressos durante o processo de nodulação ou envolvidos no funcionamento dos nódulos. A superfamília das nodulinas, por sua vez, vem sendo subdivida tradicionalmente em dois subgrupos: precoces e tardias. Como o nome indica, as precoces são aquelas expressas imediatamente após o contato da bactéria com a planta, estando diretamente envolvidas na formação da estrutura nodular. Nesse grupo destacam-se ENOD11, ENOD12, ENOD40, NIN, DMI e SYM (MACHADO, 2014).

Segundo Mohanta e Bae (2015), associações simbióticas em leguminosas envolvem troca de sinais permitindo o reconhecimento mútuo. No caso da associação rizobial, lipo-quito-oligossacarídeos chamados de fatores Nod (NF) desempenham um papel crucial. Acredita-se que sua ligação a receptores de plantas ativa uma cascata de sinalização envolvendo a geração de uma resposta oscilatória de cálcio intracelular sustentada e dependente de várias proteínas-chave, incluindo receptores quinases contendo motivos LysM (NFP / NFR1 / NFR5) ou LLR (DMI2 / SYMRK), canais iônicos putativos (DMI1 / CASTOR / POLLUX), dependente de cálcio e calmodulina proteína quinase (DMI3 / CCaMK), bem como vários transcritos reguladores (NSP1 / NSP2 / Ets2 / ERN).

A ativação da via de sinalização de NF resulta na expressão de um conjunto de genes envolvido no programa de desenvolvimento simbiótico (OLDROYD; DOWNIE 2004). Curiosamente, mutantes defeituosos em certos genes envolvidos nessa via (por exemplo, DMI1, DMI2 e DMI3 em *Medicago truncatula*) também são afetados em sua capacidade de estabelecer a

associação simbiótica com fungos micorrízicos arbusculares (FMA) (CATOIRA et al., 2000). Portanto, componentes específicos da via NF são necessários para ambos os tipos de endossimbiose e são considerados como integrantes de um antigo caminho simbiótico envolvido na acomodação de micróbios pela célula vegetal (KISTNER; PARNISKE, 2002) (Figura 2).

Marsh et al. (2007) sugeriram em seu trabalho que os genes NIN funcionam acima da sinalização inicial do fator Nod levando à organogênese do nódulo além de um proposto papel na entrada bacteriana.

Os genes primitivos de nodulina (ENOD) correspondem a genes de nodulina cuja expressão pode ser induzida apenas por fatores Nod, tais como: ENOD12 em *Pisum sativum* e *M. truncatula* (SCHERES et al. 1990; PICHON et al., 1992). O ENOD40 pode funcionar na nodulação como uma molécula de sinalização; apesar de não possuir um quadro de leitura aberta, codificam dois pequenos peptídeos (SOUSA et al., 2001; RÖHRIG et al., 2002). Notavelmente, os peptídeos ENOD40 podem se ligar e melhorar a estabilidade da síntese de sacarose (HARDIN et al., 2003; RÖHRIG et al., 2004). Por sua vez, genes ENOD11 estudados em *M. truncatula* codificam um gene repetitivo de uma proteína rica em prolina (JOURNET et al., 2001; CHARRON et al., 2004).

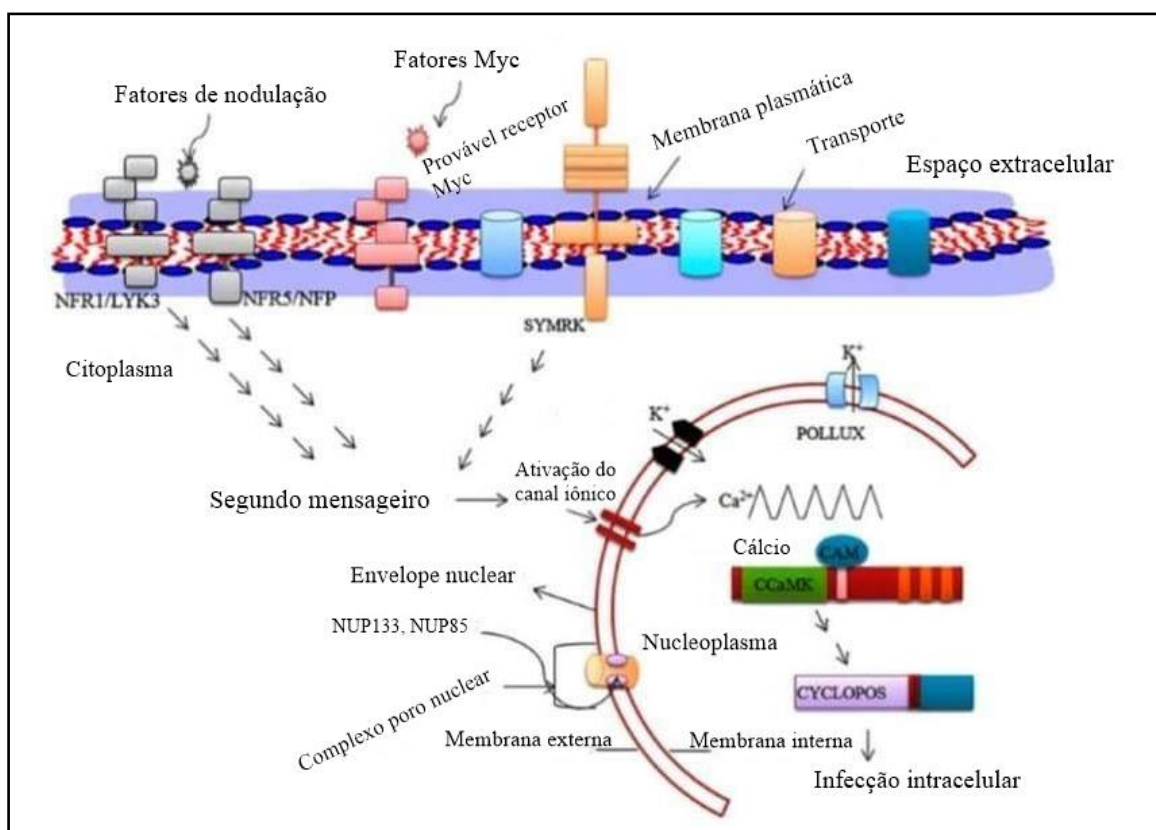
Estudos recentes apontaram a importância de NOD1 e NOD2 (receptores do tipo NOD) na defesa do hospedeiro contra patógenos a partir de ensaios realizados com sistemas celulares *in vitro*, modelos murinos *in vivo*, bem como em estudos de suscetibilidade genética em humanos (revisado por NASCIMENTO, 2016). Além do papel das proteínas NOD1 e NOD2 na resposta imune inata às infecções bacterianas e virais, vários relatos têm sugerido sua função no estímulo das respostas imunes adaptativas (MUKHERJEE et al., 2018).

A expressão das nodulinas precoces está relacionada com vários aspectos funcionais do nódulo, sendo correlacionada com o início da fixação de N₂. A importância da nodulação nas plantas é enfatizada por Vollmann et al. (2011), que evidenciou sua relação com o aumento de teor de clorofila das folhas.

De acordo com Mohanta e Bae (2015), ao perceber o sinal Nod, as moléculas receptoras são transportadas para o espaço intercelular onde, por sua vez, ocorre ativação de mensageiros secundários no citoplasma os quais

levam à ativação de moléculas semelhantes incluindo os genes DMI1. Após a ativação dos DMI1, o íon K^+ liberado no nucleoplasma torna o citoplasma hiperpolarizado. Uma vez hiperpolarizado, o íon Ca^{2+} perinuclear entra no nucleoplasma resultando em pico de Ca^{2+} . Este se liga ao CcaMK no qual passa o sinal ao gene CYCLOPOS, responsável pela regulação do processo de micorrização (**Figura 2**).

Figura 2. Representação da percepção do fator Nod (NF) pela respectiva dos receptores. Figura adaptada de Mohanta e Bae, 2015.



Em relação as nodulinas tardias, estas são principalmente expressas nas estruturas dos nódulos maduros ao redor do local da FBN, assim como participa do processo. Nodulinas tardias codificam proteínas que realizam papéis na transformação de bactérias em bacteroides e sua colonização dentro dos nódulos. São exemplos de nodulinas tardias: leghemoglobina, NOD35 e transportadores do simbiossomo (Revisado por SEIDO, 2017).

De acordo com Schröder et al., 1997 a expressão das nodulinas tardias, que participam da troca metabólica entre a planta e o microssimbionte está associada ao processo de FBN, no qual é acionada pela nitrogenase expressa

no rizóbio, dentre as proteínas participantes incluem-se as transportadoras de membrana e proteínas associadas especificamente com a membrana do simbiossomo (KAISER et al., 2003; JEONG et al., 2004; WIENKOOOP; SAALBACH, 2003; CATALANO; LANE; SHERRIER, 2004), possibilitando a formação e a manutenção do processo simbiótico por meio fluxo de inúmeras moléculas e íons requeridos para o processo (ROBERTS; TYERMAN, 2002).

3.4 Bioinformática na caracterização de sequências

Em vista do crescente acúmulo e demanda por dados moleculares sobre plantas, métodos computacionais adotados na bioinformática apresentam potencial para produzir conhecimentos valiosos sobre a regulação da transcrição em resposta a pressões ambientais. Para entender a regulação e os mecanismos de adaptação aos estresses ambientais, se torna fundamental a identificação dos componentes que compõem estes sistemas regulatórios, incluindo tanto os genes codificantes de fatores de transcrição (FTs), como aqueles codificantes de seus efetores. Com base no conhecimento destas interações é possível a criação de modelos computacionais que auxiliem na compreensão e elucidação da arquitetura de uma determinada rede em células quando sob condições de estresse (CHEN; ZHU, 2004; LAI et al., 2011).

Compreender os mecanismos pelos quais as plantas transmitem os sinais para a maquinaria celular de modo a ativar respostas adaptativas é de fundamental importância para desenvolver culturas mais tolerantes, melhorando a eficiência da produção (BEZERRA-NETO, 2016). De forma geral, a percepção da planta ao estresse, desencadeia uma cascata de eventos moleculares, levando à transmissão e processamento de sinais para geração de respostas mediadas por fitormônios, mensageiros secundários, fatores de transcrição (FTs) e proteínas sinalizadoras (PANDEY, 2008).

As ferramentas de bioinformática são programas de computador que auxiliam na interpretação dos resultados gerados pelos diferentes sequenciamentos. Atualmente centenas de bancos de dados vêm disponibilizando informações específicas para determinadas espécies ou grupos de organismos, incluindo sequências nucleotídicas ou proteicas, de grupos de moléculas, de elementos gênicos ou genômicos (FERNANDEZ-SUAREZ; GALPERIN, 2012).

Dentre os bancos de dados públicos que armazenam uma grande quantidade de dados de sequências de nucleotídeos e de proteínas, destacam-se o GenBank do NCBI e o UniProt. O UniProt é um banco de dados que integra, interpreta e padroniza dados de inúmeras fontes para formar o maior e mais completo catálogo de sequências proteicas e de anotação funcional. Assim, o UniProt disponibiliza uma base de conhecimento estável e abrangente de sequências proteicas curadas, com livre acesso, compreendendo sequências totalmente classificadas e ricamente anotadas e curadas (THE UNIPROT CONSORTIUM, 2012).

REFERÊNCIAS

- ADAMS, J. Transcriptome: connecting the genome to gene function. **Nature Education**, v. 1, p. 1, 2008.
- AHUJA, I.; DE VOS, R. C. H.; BONES, A. M. Plant molecular stress responses face climate change. **Trends in Plant Science**, v. 15, p. 664-674, 2010.
- ALVES, M. S. et al. Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. **Proteomes**, v. 2, p. 85-106, 2014.
- ARGENTA, G. et al. Relação entre teor de clorofila extraível e leitura do clorofilômetro na folha de milho. **Revista Brasileira de Fisiologia Vegetal**, v. 1, p. 158-167, 2001.
- ATHAR H. R.; ASHRAF, M. Strategies for crop improvement against salinity and drought stress: an overview (2009). In: ASHRAF, M.; OZTURK, M.; ATHAR, H. R. (eds) *Salinity and water stress: improving crop efficiency*, 3rd edn. Dodrecht, The Netherlands, p. 1-16.
- ATKINSON, N. J.; URWIN P. E. The interaction of plant biotic and abiotic stresses: from genes to the field. **Journal of Experimental Botany**, v. 63., n. 10, p. 3523-3544, 2012.
- AZANI, N. et al. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny The Legume Phylogeny Working Group (LPWG). **Taxon**, v. 66, n.1, p. 44-77, 2017.
- BARNES, R. F.; NELSON, C. J.; COLLINS, M.; MOORE, K. J. *Forages an introduction to grassland agriculture*. Blackwell Publishing Professional, Iowa, USA, 6th edition, volume I, 2003.
- BENKO-ISEPPON, A. M. et al. Mendel e suas exceções à luz das ômicas e da biologia de sistemas. In: ARAGÃO, F. J. L.; MOREIRA, J. R. *Mendel: 150 anos depois*. Embrapa Edições (Brasília, DF), p. 502, 2017.
- BENKO-ISEPPON, A. M.; NEPOMUCENO, A. L.; ABDELNOOR, R. V. Genosoja - The Brazilian Soybean Genome Consortium: High Throughput Omics and Beyond. **Genetics and Molecular Biology**, v. 35, p. 1-5, 2012.
- BENKO-ISEPPON, A. M. et al. Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. **Revista Brasileira de Geografia Física**, v. 6, p. 1112-1134, 2011.
- BEZERRA NETO, J. P. *Repostas transcricionais e estresse-induzidas em genótipos contrastantes de Glycine max (Soja) e Vigna unguiculata (Feijão-caupi)*. Tese (Doutorado em Ciências Biológicas) – Centro de Biociências, Universidade Federal de Pernambuco. Recife, p. 206, 2016.
- BITTENCOURT, P. P.; SILVA, L. N. N. S. Estresse hídrico em plantas: aspectos morfofisiológicos, adaptações e mecanismos de resposta. In: VIII Botânica

no Inverno 2018/ Org. Aline Possamai Della [et al.]. – São Paulo: Instituto de Biociências da Universidade de São Paulo, Departamento de Botânica, p. 275, 2018.

- BRUCE, T. J. A.; MATTHES, M. C.; NAPIER, J. A.; PICKETT, J. A. Stressful “memories” of plants: Evidence and possible mechanisms. **Plant Science**, v.173, p.603-608, 2007.
- CALLES, T.; SCHULTZE-KRAFT, R. New species, nomenclatural changes and recent taxonomic studies in the genus *Stylosanthes* (Leguminosae): An update. **Tropical Grasslands-Forrajões Tropicais**. v. 4, n. 2, p. 122-128, 2016.
- CARVALHO, C. C. G.; SILVA, D. C. G. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**, v.40, n.3, p.735-744, 2010.
- CATALANO, C. M.; LANE, W. S.; SHERRIER, D. J. Biochemical characterization of symbiosome membrane proteins from *Medicago truncatula* root nodules. **Electrophoresis**, v. 25, p. 519-531, 2004.
- CATOIRA, R. et al. Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. **Plant Cell**, v. 2, p. 1647-1666, 2000.
- CHEN, W. J.; ZHU, T. Networks of transcription factors with roles in environmental stress response. **Trends Plant Science**, v. 9, n. 12, p. 591-596, 2004.
- COSTA, N. M. S. Revisão do gênero *Stylosanthes* Sw. Tese (Doutorado em Engenharia Agrônoma), Universidade Técnica de Lisboa, p. 469, 2006.
- CRAMER, G. R.; URANO, K.; DELROT, S.; PEZZOTTI, M.; SHINOZAKI, K. Effects of abiotic stress on plants: a systems biology perspective. **BMC Plant Biology**, v. 11, n. 163, p. 1471-2229, 2011.
- DEDECEK, R. A.; GALDINO, S.; VIEIRA, L. M. Perdas de solo e água em pastagens cultivadas em solo arenoso da Bacia do Alto Taquari, MS. Corumbá: Embrapa Pantanal; [Colombo]: Embrapa Florestas, 2006.
- DE MELO, R. F. et al. Potencial de quatro espécies herbáceas forrageiras para fitorremediação de solo contaminado por arsênio. **Revista Brasileira de Ciência do Solo**, v. 33, n. 2, p. 455-465, 2009.
- DUTRA, T. R.; MASSAD, M. D.; MOREIRA, P. R.; RIBEIRO, E. S. M. Efeito da salinidade na germinação e crescimento inicial de plântulas de três espécies arbóreas florestais. **Pesquisa florestal brasileira**, Colombo, v. 37, n. 91, p. 323-330, 2017.
- EDYE, L. A. Commercial development of *Stylosanthes* pastures in Northern Australia. I. Cultivar development within *Stylosanthes* in Australia. **Tropical Grasslands**, v.31, n.5, p. 503-508, 1997.

- EMBRAPA CAMPO GRANDE. Cultivo e uso do Estilosantes Campo Grande. Campo Grande, MS: **Embrapa Gado de Corte**, p. 11, 2007 (Comunicado técnico, 105).
- FABRICE, C. E. S.; SOARES-FILHO, C. V.; PINTO, M. F.; PERRI, S. H. V. Recuperação de pastagens de *Brachiaria decumbens* degradada com introdução de *Stylosanthes* e adubação fosfatada. **Revista Brasileira de Saúde e Produção Animal**, v. 16, n. 4, p. 758-771, 2015.
- FELLER, A.; MACHEMER, K.; BRAUN, E. L.; GROTEWOLD, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. **The Plant Journal**, v. 66, p. 94-116, 2011.
- FERNANDEZ-SUAREZ, X. M.; GALPERIN, M. Y. The 2013 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection. **Nucleic Acids Research**, v. 41, n. D1, p. D1-D7, 2012.
- FERREIRA, M. A. M. Bioinformática como ferramenta no melhoramento genético de plantas. (Trabalho de Conclusão de Curso) – Departamento de Ciências Biológicas, Centro de Ciências Humanas e Naturais, Universidade Federal do Espírito Santo, Vitória, 2017.
- FERREIRA-NETO, J. R. C.; BENKO-ISEPPON, A. M.; KIDO, E. A. Transcriptômica: por quê, como e para que? Uma Abordagem para Iniciantes. 1ª ed. Beau Bassin: Novas Edições Acadêmicas, v. 1, p. 68, 2018.
- FERREIRA-NETO, J. R. C. et al. Cowpea and abiotic stresses: identification of reference genes for transcriptional profiling by qPCR. **Plant Methods**, v.14, n.88, p.1-17, 2018.
- GALDINO, A. C. Ocorrência e multiplicação de *Stylosanthes* em Pernambuco. Tese (Doutorado Integrado em Zootecnia) – Departamento de Zootecnia, Universidade Federal Rural de Pernambuco, Recife, p. 94, 2014.
- GALHARDO, M.; SINKKONEN, L.; BERNINGER, P.; LIN, J.; SAUTER, T.; HEINÄNIEMI, M. Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network. **Nucleic Acids Research**, v. 42, n. 3, p. 1474-1496, 2013.
- GAMA, T. C.; VOLPE, E.; LEMPP, B.; GALDEIA, E. C. Recuperação de pasto de capim-braquiária com correção e adubação de solo e estabelecimento de leguminosas. **Revista Brasileira de Saúde Produção Animal**, v. 14, n. 4, p. 635-647, 2013.
- HARDIN, S. C. et al. Phosphorylation of sucrose synthase at serine 170: Occurrence and possible role as a signal for proteolysis. **The Plant Journal**, v. 35, p. 588-603, 2003.
- HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C. A importância do processo de fixação biológica do nitrogênio para a cultura da soja: componente essencial para a competitividade do produto brasileiro. Londrina: Embrapa Soja, p. 80, 2007 (Embrapa Soja. Documentos, 283).

- JANK, L; VALLE, C. B.; RESENDE, R. M. S. Breeding tropical forages. **Crop Breeding and Applied Biotechnology**, v. 1, p. 27-34, 2011.
- JEONG J, et al. A nodule-specific dicarboxylate transporter from alder is a member of the peptide transporter family. **Plant Physiology**, v. 134, p. 969–978, 2004.
- JUDD, W. S.; CAMPBELL, C. S.; KELLOGG, E. A.; STEVENS, P. F.; DONOGHUE, M. J. Sistemática vegetal: Um enfoque filogenético. 3.Ed. Porto alegre: Artmed, p. 612, 2008.
- KAISER, B. N. et al. The soybean NRAMP homologue, GmDMT1, is a symbiotic divalent metal transporter capable of ferrous iron transport. **The Plant Journal**, v. 35, p. 295–304, 2003.
- KASABOV, N. Global, local and personalised modeling and pattern discovery in bioinformatics: An integrated approach. **Pattern Recognition Letters**, v. 28, n. 6, p. 673-685, 2007.
- KHOUZANI, M. R. Z.; NEJAD, T. S. Evaluation the effect of bio fertilizer and nitrogen fertilizer on rate of biological nitrogen fixation (BNF) and nodulation of cowpea varieties in Ahwaz, Iran. **Advances in Environmental Biology**, v. 9, n. 4, p. 407-414, 2015.
- KIDO, E. A. et al. Identification of plant protein kinases in response to abiotic and biotic stresses using SuperSAGE. **Current Protein & Peptide Science**, v. 12, n. 7, p. 643-656, 2011.
- KIDO, E. A.; FERREIRA-NETO, J. R. C.; PANDOLFI, V.; SOUZA, A. C. M.; BENKO-ISEPPON, A. M. Drought Stress Tolerance in Plants: Insights from Transcriptomic Studies. In: HOSSAIN, M. A.; WANI, S. H.; BHATTACHARJEE, S.; BURRITT, D. J.; TRAN, L-S. P. (Eds.). Drought Stress Tolerance in Plants, Vol 2 Molecular and Genetic Perspectives. Springer International Publishing Switzerland. ISBN 978-3-319-32421-0. Chapter 6: p. 153-185, 2016.
- KISTNER, C.; PARNISKE, M. Evolution of signal transduction in intracellular symbiosis. **Trends in Plant Science**, v. 7, n. 11, p. 511-518, 2002.
- KORPELAINEN, E. et al. RNA-seq Data Analysis A Practical Approach. [s.l.] CRC Press, 2015.
- LAI, Z.; et al. Arabidopsis sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. **Plant Cell**, v. 23, p. 3824–3841, 2011.
- LAVEZO, L.; BRAGA, L. F.; BASTISTÃO, A. C.; BONFANTE, L. V. Estresse osmótico na germinação de sementes de *Petiveria alliacea*. **Revista Brasileira de Plantas Mediciniais**, v.17, n.4, p.622-630, 2015.
- LEAL, E. F. Plantas nativas e fontes de fosfato na remoção e estabilização química de metais traços de um vertissolo. Dissertação (Mestrado em Solos

- e Qualidade de Ecossistemas), Centro de Ciências Agrárias, Ambientais e Biológicas, Universidade Federal de Recôncavo da Bahia, p. 72, 2014.
- LEWIS, G. P. Legumes of Bahia. Kew: Royal Botanic Gardens. p. 369, 1987.
- LEWIS, G. P.; SCHRIRE, B.; MACKINDER, B.; LOCK, M. Legumes of the world. Royal Botanical Gardens, Kew. p. 592, 2005.
- LIMA, H. C. et al. (orgs) Lista de espécies da flora do Brasil. Jardim Botânico do Rio de Janeiro. (<http://floradobrasil.jbrj.gov.br/2012/FB000115>).
- LIMA; H. C. et al. 2010. Fabaceae. In: Lista de Espécies da Flora do Brasil. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/2010/FB000115>>. Acesso em: 12 agosto de 2017.
- LISAR, S. Y. S. et al. Water stress in plants: causes, effects and responses Seyed. In: RAHMAN, I. M. M.; HASEGAWA, H. Water stress. Croatia: Intech, cap. 1, p. 1-14, 2012.
- MACHADO, L. R. D. Desenvolvimento inicial de leguminosa forrageira em simbiose com rhizobium em substrato tratado com biossólido da indústria de laticínios. Dissertação (Mestrado em Engenharia Civil e Ambiental)- Universidade Estadual Paulista “Júlio de Mesquita Filho”, p. 73, 2014;
- MARQUES, A. et al. Origin and parental genome characterization of the allotetraploid *Stylosanthes scabra* Vogel (Papilionoideae, Leguminosae), an important legume pasture crop. **Annals of botany**, v. 122, n. 7, p. 1143-1159, 2018.
- MARSH, J. F et al. *Medicago truncatula* NIN is essential for Rhizobial independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. **Plant Physiology**, v. 35, p; 144-324, 2007.
- MATOS, F. S. et al. Estratégia morfofisiológica de tolerância ao déficit hídrico de mudas de pinhão manso. **Magistra**, n. 26 v.1 p. 19-27, 2014.
- MATOS, M. K. S. Expressão diferencial e diversidade de fatores de transcrição da família MYB em feijão-caupi. Dissertação (Mestrado em Genética) – Centro de Ciências Biológicas, Universidade Federal de Pernambuco. Centro de Ciências Biológicas. Recife, p. 137, 2015.
- MELLO, G. S. V. Análise computacional de genes associados ao metabolismo de fixação de nitrogênio no feijão-caupi (*Vigna unguiculata*) e cana-de-açúcar (*Saccharum* spp.). Dissertação (Mestrado em Genética) – Departamento de Genética, Universidade Federal de Pernambuco, Recife, p. 160, 2009.
- MIOTTO, S. T. S.; LÜDTKE, R.; OLIVEIRA, M. L. A. A. A família Leguminosae no Parque Estadual de Itapuã, Viamão, Rio Grande do Sul, Brasil. **Revista Brasileira de Biociências**, v. 6, n. 3, 2008.

- MIR, L. *Gênômica*. São Paulo: Atheneu, 2005.
- MISTURA, C. et al. Frequência do gênero *Stylosanthes* e atributos físico do solo na microrregião de Juazeiro-BA. In: Encontro da rede de recursos genéticos vegetais da Bahia, Vitória da Conquista - BA. Rede de Recursos Genéticos Vegetais-RGVs. Recurso Genéticos Vegetais (CD-ROM), 2008.
- MITTLER, R.; BLUMWALD, E. Genetic engineering for modern agriculture: challenges and perspectives. **Annual Review of Plant Biology**, v. 61, p. 443-462, 2010.
- MOHANTA, T. K.; BAE, H. Functional Genomics and Signaling Events in Mycorrhizal Symbiosis. **Journal of Plant Interactions**, v.10 p. 21–40, 2015.
- NAGAICH, D.; TIWARI, K. K.; SRIVASTVA, N.; CHANDRA, A. Assessment of genetic diversity and morphophysiological traits related to drought tolerance in *Stylosanthes scabra*. **Acta Physiologia and Plantarum**. v.35, p. 3127-3136, 2013.
- NAKASHIMA, K.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. **Frontiers in Plant Science**, v. 5, n. 170, 2014.
- NASCIMENTO, M. S. L. A sinalização via NOD2-RIP2 modula a imunidade adaptativa contra *Leishmania infantum*. Tese (Doutorado em Ciências) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo. São Paulo, p. 168, 2016.
- NEGREIROS, J. V.; SANTOS, A. C.; SANTOS, P. M.; SANTOS, T. M.; FARIA, A. F. G. Atributos físicos de solos sob a consorciação gramíneas-leguminosas no norte do estado do Tocantins. **Engenharia na Agricultura**, v. 18, n. 2, p. 140-150, 2010.
- NEVES, A. A. C. Análises genômicas de *Methylobacterium mesophilicum* SR1.6/6 com ênfase na interação com a planta hospedeira. Tese (Doutorado em Biotecnologia) – Programa de pós-graduação Interunidades em Biotecnologia, Universidade de São Paulo, São Paulo, p. 45, 2015.
- OLDROYD, G. E. D.; DOWNIE, J. A. Calcium, kinases and nodulation signaling in legumes. **Natura Reviews Molecular Cell Biology**, v. 5, n, 7, p. 566-576, 2004.
- OLIVEIRA, R. S.; QUEIRÓZ, M. A. Banco Ativo de Germoplasma de *Stylosanthes* spp. da Universidade Estadual de Feira de Santana – BA. **Revista RG News**, v.2, n. 2, 2016.
- OSAKABE, Y. et al. Response of plants to water stress. **Frontiers in Plant Science**, v. 86, p. 1-8, 2014.
- PANDEY, G. K. Emergence of a novel calcium signaling pathway in plants: CBL-CIPK signaling network. **Physiology and Molecular Biology of Plants**, v.

14, p. 51-68, 2008.

- PICHON, M. et al. *Rhizobium meliloti* elicits transient expression of the early nodulin gene ENOD12 in the differentiating root epidermis of transgenic alfalfa. **Plant Cell**, v. 4, p. 1199-1211, 1992.
- RESENDE, R. M. S.; RESENDE, M. D. V.; LAURA, V. A. Genotypic evaluation of accessions and individual selection in *Stylosanthes* spp. by simulated BLUP method. **Crop Breeding and Applied Biotechnology**, v.6, p. 253-260, 2006.
- RIBEIRO, M. R.; et al. Química dos solos salinos e sódicos. In: MELO, V. F.; ALLEONI, L. R. F. (ed.). Química e mineralogia do solo. Parte II – Aplicações. Enciclopédia biosfera, Centro Científico Conhecer, Goiânia. Viçosa: Sociedade Brasileira de Ciência do Solo, v. 8, n. 15, p. 449-484, 2009.
- ROBERTS, D. M.; TYERMAN, S. D. Voltage-dependent cation channels permeable to NH_4^+ , K^+ , and Ca^{2+} in the symbiosome membrane of the model legume *Lotus japonicus*. **Plant Physiology**, v. 128, p. 370–378, 2002.
- RÖHRIG, H.; JOHN, M.; SCHMIDT, J. Modification of soybean sucrose synthase by S-thiolation with ENOD40 peptide A. **Biochemical and Biophysical Research Communications**, v. 325, p. 864-870, 2004.
- RÖHRIG, H.; SCHMIDT, J.; MIKLASHEVICH, E.; SCHELL, J.; JOHN, M. Soybean ENOD40 encodes two peptides that bind to sucrose synthase. **Proceedings of the National Academy of Sciences**, v. 99, p. 1915-1920, 2002.
- ROMAO, R. L. et al. Resgate de germoplasma de *Stylosanthes* em área do semiárido Baiano. In: Encontro da rede de recursos genéticos vegetais da Bahia, III., 2008, Vitória da Conquista- BA. Rede de Recursos Genéticos Vegetais da Bahia (CR-ROM), 2008.
- SANTANA, M. J.; CARVALHO, J. A.; SILVA, E. L.; MIGUEL, D. S. Efeito da irrigação com água salina em um solo cultivado com o feijoeiro (*Phaseolus vulgaris* L.). **Ciência e Agrotecnologia**, v.27, p.443-450, 2011.
- SANTOS, L. A.; REIS, V. M. A formação do nódulo em leguminosas. Seropédica: Embrapa Agrobiologia, 14 p. Documentos / Embrapa Agrobiologia, n. 251, 2008.
- SANTOS, W. F.; OLIVEIRA, M. S.; ROMEIRO, L. C. T. L.; CUNHA, M. A. A. Sequenciamento de DNA: métodos e aplicações. **Proceedings of Safety, Health and Environment World Congress**, v. 13, p. 139-141, 2013.
- SCHAUSER, L.; WIELOCH, W.; STOUGAARD, J. **Journal of Molecular Evolution**, v. 60, n. 229, 2005.
- SCHERES, B. et al. Sequential induction of nodulin gene expression in the developing pea nodule. **Plant Cell**, v. 2, p. 687-700, 1990.

- SCHMUTZ, J. et al. Genome sequence of the palaeopolyploid soybean. **Nature**, v. 463, n. 7278, p. 178-183, 2010.
- SCHOSSLER, T. R.; MACHADO, D. M.; ZUFFO, A. M.; ANDRADE, F. R.; PIAUILINO, A. C. Salinidade: efeitos na fisiologia e na nutrição mineral de plantas. **Enciclopédia biosfera**, v.8, n.15, 2012.
- SCHRÖDER, G.; FRÜHLING, M.; PÜHLER, A.; PERLICK, A. M. The temporal and spatial transcription pattern in root nodules of *Vicia faba* nodulin genes encoding glycinerich proteins. **Plant Molecular Biology**, v. 33, p. 113–123, 1997.
- SEIDO, S. L. Estudos biométricos e de QTL associados à fixação biológica de nitrogênio em feijão-caupi. Tese (Doutorado em Melhoramento Genético de Plantas) – Departamento de Agronomia, Universidade Federal Rural de Pernambuco, Recife, p. 139, 2017.
- SEKI, M.; UMEZAWA, T.; URANO, K.; SHINOZAKI, K. Regulatory metabolic networks in drought stress responses. **Current Opinion in Plant Biology**, n.10, p. 296-302, 2007.
- SILVA, S. C. Caracterização *in silico* e análise da expressão de microRNAs e genes alvo em diferentes espécies de *Coffea*. Dissertação (Mestrado em Biotecnologia Vegetal), Universidade Federal de Lavras, Minas Gerais, p. 103, 2012.
- SILVEIRA, R. D. D.; PANTALIÃO, G. F.; BRONDANI, C. Estudos genômicos de tolerância à seca em arroz: uma breve revisão. **Multi-Science Journal**, v.1, n.1, p.62-69, 2015.
- SOUSA, C. et al. Translational and structural requirements of the early nodulin gene enod40, a short-open reading frame-containing RNA, for elicitation of a cell-specific growth response in the alfalfa root cortex. **Molecular and Cellular Biology**, v.21, p. 354–366, 2001.
- SOUZA, N. M.; SOUZA, L. A. G. Levantamento do potencial de aproveitamento das leguminosas no distrito da Barreira do Andirá, Barreirinha, AM. **Enciclopédia Biosfera-Centro Científico Conhecer**, v. 7, n.12, 2011.
- SOUZA, V. C.; LORENZI, H. Botânica sistemática: Guia ilustrado para identificação das famílias de Angiospermas da flora brasileira, baseado em APG II. Plantarum, Nova Odessa, 2005.
- TAIZ, L.; ZEIGER, E. Fisiologia vegetal. 5. ed. Porto Alegre: Artmed, 2009. 719 p.
- THE UNIPROT CONSORTIUM. Update on activities at the Universal Protein Resource (UniProt) in 2013. **Nucleic Acids Research**, v. 41, n. D1, p. D43–D47, 2012.
- TUBEROSA, R.; SALVI, S. Genomics-based approaches to improve drought tolerance of crops. **Trends in Plant Science**, v. 11, p. 405–412, 2006

- UBIALI, D. G.; RAFAEL, G. F. S.; OLIVEIRA, L. P. Obstrução intestinal em bovinos associada ao consumo de *Stylosanthes scabra* Vogel (Fabaceae Papilionoideae). **Pesquisa Veterinária Brasileira**, v. 33 n. 2, 2013.
- VOLLMANN, J. et al. Digital image analysis and chlorophyll metering for phenotyping the effects of nodulation in soybean. **Computers and Electronics in Agriculture**, v. 75, p.190-195, 2011.
- WALTER, J.; JENTSCH, A.; BEIERKUHNLEIN, C.; KREYLING, J. Ecological stress memory and cross stress tolerance in plants in the face of climate extremes. **Environmental and Experimental Botany**, v.94, p.3– 8, 2013.
- WALTER, J. et al. Do plants remember drought? Hints towards a drought-memory in grasses. **Environmental and Experimental Botany**, v.71, p.34–40, 2011.
- WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57–63, 1 jan. 2009.
- WANG, Z. et al. SoyDB: a knowledge database of soybean transcription factors. **BMC Plant Biology**, v. 10, p. 14, 2010.
- WIENKOOP, S.; SAALBACH, G. Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. **Plant Physiology**, v.131, p. 1080–1090, 2003.

CAPÍTULO II

Artigo a ser submetido à Revista
Bioinformatics and biology insights

IDENTIFICAÇÃO E VALIDAÇÃO DE NODULINAS PRECOSES NO TRANSCRIPTOMA DE *Stylosanthes scabra* Vogel SOB DÉFICIT HÍDRIDO

Ruana Carolina Cabral da Silva¹; Roberta Lane de Oliveira Silva²; Jéssica Barboza da Silva²; João Pacífico Bezerra Neto²; Thialisson Caaci do Vale Amorim²; Carlos André dos Santos Silva²; Ana Maria Benko-Iseppon^{2#}; Reginaldo de Carvalho¹

¹Universidade Federal Rural de Pernambuco, Departamento de Biologia, Laboratório de Citogenética Vegetal, Recife, PE.

²Universidade Federal de Pernambuco, Departamento de Genética, Laboratório de Genética e Biotecnologia Vegetal, Recife, PE.

#Autor para correspondência:

Ana Maria Benko-Iseppon, Universidade Federal de Pernambuco, Departamento de Genética, Centro de Ciências Biológicas, Av. Prof. Moraes Rego 1235, CEP 50.670-420, Recife, PE, Brasil. Telefone: 55-81-2126-7816; Fax: 55-82-2126-8522; E-mail: ana.benko.iseppon@pq.cnpq.br

INTRODUÇÃO

Sabe-se que as mudanças climáticas se intensificaram, apresentando-se como uma condição problemática em nível global. Tais alterações afetam diretamente o desenvolvimento e sobrevivência dos vegetais, incluindo a disponibilidade de recurso hídrico para a planta, advindo da seca, gerada pela falta de chuva (JÚNIOR et al., 2018). Em resposta ao déficit hídrico as plantas desencadeiam inúmeros mecanismos de respostas geneticamente determinadas. Para a maioria dos genes, essa regulação é realizada em nível transcricional, garantindo desta forma um menor gasto energético da célula, além de níveis adequados de proteínas (MATOS, 2015; KUERT et al., 2016).

Em geral, as nodulinas precoces são expressas imediatamente após o contato da bactéria com a planta e estão diretamente envolvidas na formação da estrutura nodular, enquanto as tardias são ativadas posteriormente (FERGUSON et al., 2018). Levando em consideração que o processo de duração é um processo longo, pretende-se entender este tendo como base as nodulinas precoces, visto que estas se apresentam nos tempos iniciais, assim

como levando em consideração os tempos transcriptômicos.

A análise de dados transcriptômicos em associação com modelos de redes metabólicas vem auxiliando na identificação de novos genes relacionados à defesa vegetal frente a situações de estresse bem como em inferências sobre a participação desses genes em processos metabólicos importantes, identificando-se pontos de convergência de regulação (GALHARDO et al., 2013).

Como a *S. scabra* é especialmente tolerante à seca e à salinidade, torna-se uma potencial espécie doadora de genes para uso biotecnológico, especialmente em Fabaceae. Sabe-se que a estilósantes continua nodulando sob os estresses citados, embora seu crescimento seja reduzido. Sua capacidade de tolerar tais estresses e de nodular, merece uma análise detalhada, especialmente visando à identificação e validação da expressão de genes relacionados ao processo de simbiose, como as nodulinas.

Neste cenário, é de suma importância entender como essa família gênica se composta transcricionalmente perante estresses, destacando que não há relatos na literatura de estudos voltados a estas análises, especialmente, tratando-se de estilósantes, sendo assim, os dados produzidos agregam valores fisiológicos molecular dos genes associados a espécie perante estresse abiótico.

Vale ressaltar, que as leguminosas, como *S. scabra* são componentes valiosos em consórcios de plantas forrageiras (especialmente com gramíneas) e na rotação de culturas, contribuindo para a diminuição e dependência de fertilizantes nitrogenados (BARNES et al., 2003) que apresentam alto custo, tanto econômico como ambiental (KHOUZANI; NEJAD, 2015).

Além da importância socioeconômica da espécie, o presente estudo apresenta algumas questões endereçadas, incluindo: (1) Quantos transcritos candidatos de nodulinas (TCNs) são modulados (induzidos ou reprimidos) em *S. scabra* sob desidratação radicular comparativamente ao controle não estressado? (2) O painel de expressão dos TCNs se mantém ao longo de diferentes tempos de coleta após a imposição do estresse? (3) Há conservação ou diversificação estrutural quando se comparam os TCNs de *S. cabra*, àqueles identificados nos genomas sequenciados de espécies de *Arachis*? (4) Qual a distribuição de homólogos de TCNs de *S. scabra* no genoma das *Arachis* L.?

Neste sentido, o presente estudo objetivou avaliar os mecanismos

utilizados por *S. scabra* na resposta ao déficit hídrico, bem como sobre a expressão gênica diferencial de nodulinas precoces em diferentes momentos após a suspensão da irrigação.

2 METODOLOGIA

2.1 Obtenção do material vegetal

O experimento com *S. scabra* foi conduzido nos telados da Embrapa Semiárido (Petrolina, PE) e utilizou plantas propagadas por estaquia, a partir de uma matriz, o acesso 85/UNEB, proveniente do banco Ativo de Germoplasma da Universidade do Estado da Bahia. As plantas foram submetidas ao déficit hídrico em quatro tratamentos distintos. Foram utilizadas três plantas por vaso em todos os tratamentos. As plantas controles foram mantidas nas condições de disponibilidade de água, sendo irrigadas apenas uma vez por dia no mesmo horário. Para os tratamentos, a irrigação foi suspensa, com material vegetal (folhas e raízes) coletados após seis 24 e 48 horas da suspensão de rega, bem como um grupo com reidratação após 6 horas. O material vegetal foi armazenado em ultra freezer -80 °C até a extração do RNA total. Foram utilizadas três réplicas biológicas para cada grupo, sendo cada réplica biológica constituída por três plantas em cada vaso. O tecido radicular foi coletado, imediatamente congelado em nitrogênio líquido e armazenado em freezer -80 °C até a extração de RNA.

A extração do RNA total foi realizada a partir do protocolo CTAB-Acetato utilizando 200 mg de tecido foliar e tratamento de DNase. As amostras foram analisadas quanto a sua pureza, concentração e integridade em espectrofotômetro NanoDrop 2000c (ThermoFisher Scientific), fluorímetro Qubit® (Life Technologies) e Bioanalyser (Agilent Technologies), respectivamente.

2.2 Sequenciamento e Montagem do Transcriptoma

As amostras nos tempos seis e 24 horas após imposição do estresse com os tratamentos irrigados e sob déficit hídrico foram enviadas para sequenciamento pela tecnologia de RNA-Seq (Illumina HiSeq2500). Em seguida,

foi realizada a montagem das sequências para obtenção do transcriptoma da espécie via ferramenta Trinity (GRABHERR et al., 2011), seguindo o protocolo de Haas e colaboradores (2013), e a análise estatística dos genes diferencialmente expressos através do programa Edge-R e DESeq 2.

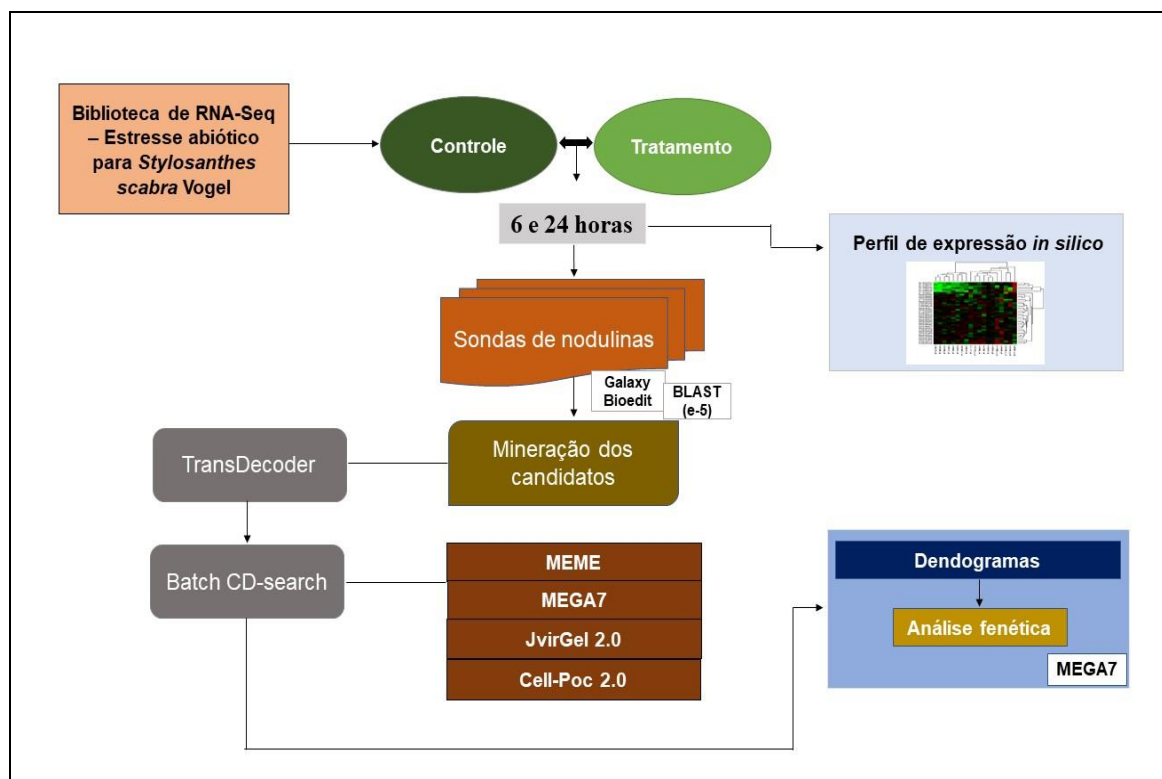
2.3 Identificação e caracterização das sequências

Inicialmente, por meio de palavras-chaves, foi realizada a busca por sequências sonda de nodulinas nos bancos de dados do UniProt (<https://www.uniprot.org/>) e NCBI (<https://www.ncbi.nlm.nih.gov/>). As sondas foram utilizadas para busca de candidatos no transcriptoma de estilosantes via tBLASTn, utilizando um ponto de corte de e^{-05} . De posse dos candidatos, foi realizada a tradução das sequências no TransDecoder e posterior busca por domínios conservados no Batch CD-search. A predição do ponto isoelétrico (p.I) e peso molecular (w.M) foi realizada através do JVirGel 2.0, bem como a localização subcelular das sequências por meio do Cell-PLoc 2.0 e WoLF PSORT.

Adicionalmente, foi determinada a posição dos loci associados aos transcritos obtidos, utilizando o genoma de *Arachis duranensis*, depositado no banco de dados PeanutBase (<https://www.peanutbase.org/>). Para verificar o comportamento transcricional dos candidatos nos tempos experimentais, foi realizada a análise de expressão *in silico* dos transcritos caracterizados, com heatmaps gerados por meio do cluster 3.0 e visualização no javaTreeView. A análise de expressão permitiu selecionar candidatos diferencialmente expressos para o desenho dos *primers*, realizado pelo programa Primers3plus e analisado quando a sua especificidade no Primer-Blast do NCBI, utilizando o genoma de duas espécies de *Arachis* (*A. hypogaea* e *A. duranensis*).

As sequências proteicas que apresentaram domínio conservado foram submetidas a um alinhamento múltiplo, por meio do MEGA7, estas sendo caracterizadas pelo o fato do grupo compartilhar domínios completos em comum, incluído nestes identificou-se os motivos conservados característicos de cada classe das nodulinas, os quais foram também analisados pelo MEME. Além disso, o alinhamento permitiu gerar as árvores fenéticas, pelo método de Neighbor-Joining (**Figura 1**).

Figura 1. Fluxograma das etapas presentes da pesquisa para identificação e caracterização das sequências, envolvendo perfil de expressão gênica e prospecção de genes candidatos por meio de ferramentas de análise *in silico*.



2.4 Modelagem tridimensional das sequências

Dois sequências de nodulinas foram selecionadas para modelagem tridimensional, a escolha dessas sequências foi baseada no seu nível de expressão, sendo uma representando o domínio PB1_NLP (>Ss-NIN-14) e outra RWP_RK (>Ss-NIN-30). Devido à falta de uma estrutura resolvida e depositadas no banco de dados de proteínas similar as sequências candidatas a modelagem, foi necessário utilizar um protocolo híbrido utilizando o algoritmo RosettaCM (SONG et al. 2013) no qual se utiliza de modelagem comparativa e nas regiões desalinhadas utiliza a biblioteca de fragmentos de novo. Cem modelos foram gerados e tiveram suas estruturas relaxadas com o protocolo relax (NIVÓN, MORETTI; BAKER 2013), em seguida feito uma análise de cluster e obtido o modelo de consenso. Posteriormente, PROCHECK (LASKOWSKI et al. 1993) foi usado para a análise estereoquímica do modelo, usando o gráfico de Ramachandran, observando a posição dos aminoácidos nas regiões mais favoráveis e observado o Z-Score.

2.5 Simulação em dinâmica molecular

As simulações foram conduzidas com o pacote GROMACS 2019.3. As proteínas foram colocadas no centro de uma caixa cúbica de dimensões em torno de 4 x 4 x 4 nm e solvatadas usando o modelo de água SPC (BERENDSEN et al. 1981). O sistema incluiu solvatação em solução de NaCl em concentração fisiológica (0,15M) substituindo 46 moléculas de soluto por 25 NA e 21 íons CL para PB1_NLP (>Ss-NIN-14) e 56 moléculas de soluto por 24 NA e 32 íons CL para RWP_RK (>Ss-NIN-30), com energia minimizada. A temperatura foi mantida em 300 K no conjunto NvT com restrição de átomos de soluto na posição inicial. O método LINCS (HESS et al. 1997) foi usado para restringir ligações envolvendo átomos de hidrogênio. A integração foi realizada pelo algoritmo *leapfrog* (HOCKNEY 1970) usando um passo de tempo de integração 2-fs. Os sistemas foram inicialmente otimizados energeticamente usando 50.000 passos do algoritmo de descida mais íngreme. Todas as simulações atomísticas foram realizadas por 100 ns usando os campos de força GROMOS 53A6 (OOSTENBRINK et al. 2004) e as condições de contorno periódicas nas direções x, y e z usando o conjunto NpT. Então, a dinâmica molecular foi realizada sem restrições a pressão e temperatura constantes de 1 atm e 300 K.

3 RESULTADOS

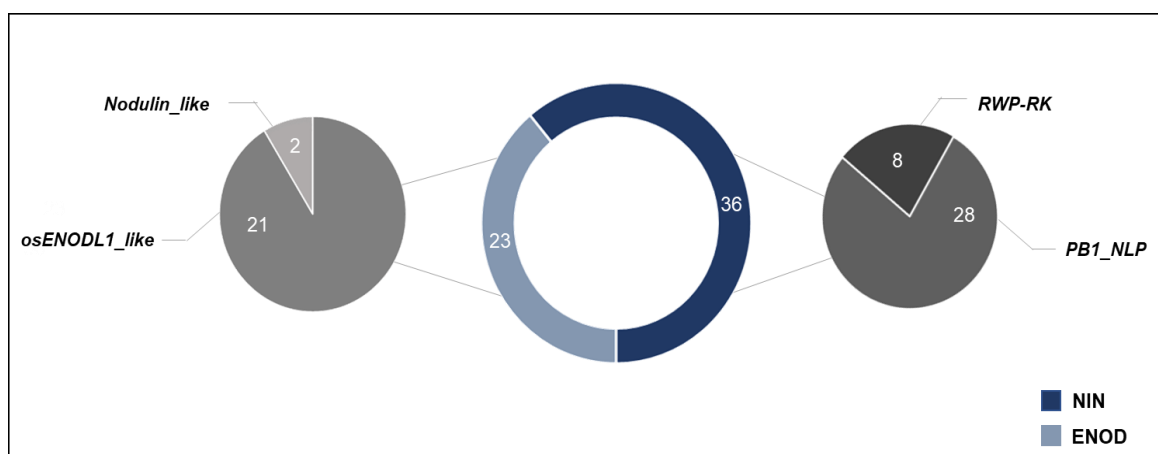
3.1 Identificação e estrutura de nodulinas precoces

A busca por sequências sondas (*Seed sequences*) correspondentes a genes codificadores de nodulinas, disponíveis no banco de dados UniProt e NCBI, retornou dados para dois grupos gênicos de nodulinas precoces presentes no transcriptoma de *S. scabra*, o ENOD, com 26916 candidatos, e o NIN, com 70 candidatos (**Tabela 1**). Após a retirada da redundância, retornaram 4968 candidatos para o grupo ENOD e 53 para o grupo NIN, contanto 23 domínios conservados ENOD (11 isoformas) e 36 domínios conservados NIN (12 isoformas) (**Tabela 1**).

Tabela 1. Dados da mineração dos candidatos de nodulinas precoces.

Grupos gênicos	Quantidade total	Sem redundância	ORF's	Domínios conservados
ENOD	26916	4968	154-208	23
NIN	70	53	102-825	36

Foram identificadas 23 sequências com domínios conservados para o grupo ENOD. Desse total, 21 sequências estão associadas ao domínio osENODL1_like e duas ao domínio nodulin-like (**Figura 2**). Todas as proteínas apresentaram localização subcelular na membrana, com ponto isoelétrico variando de 4,87 a 9,98 e peso molecular de 17,12 a 67,42 kD (**Tabela 2**).

Figura 2. Representação dos domínios conservados dos grupos de nodulinas precoces. *Nodulin_like* e *osENODL1_Like* são domínios pertencentes a ENOD, e *RWP-RK* e *PB1_NLP* a NIN.**Tabela 2.** Dados de caracterização estrutural e funcional do grupo ENOD de nodulinas precoces. Sendo ID = Identidade, coluna de domínio identificando o mesmo para cada transcrito detectado, p.I = ponto isoelétrico; p.M = peso molecular, e a localização subcelular referente a compartimentação em que os transcritos se encontram.

ID	Domínio	p.I	p.M	Localização Subcelular
>Ss-ENOD-01	Nodulin-like	8.88	67.426	Membrana celular
>Ss-ENOD-02	Nodulin-like	9.468	31.915	Membrana celular
>Ss-ENOD-03	OsENODL1_like	9.779	28.695	Membrana celular
>Ss-ENOD-04	OsENODL1_like	9.965	28.284	Membrana celular
>Ss-ENOD-05	OsENODL1_like	9.965	27.949	Membrana celular
>Ss-ENOD-06	OsENODL1_like	9.415	64.349	Membrana celular
>Ss-ENOD-07	OsENODL1_like	9.245	64.399	Membrana celular
>Ss-ENOD-08	OsENODL1_like	9.149	64.672	Membrana celular
>Ss-ENOD-09	OsENODL1_like	4.871	27.426	Membrana celular

>Ss-ENOD-10	OsENODL1_like	5.117	27.914	Membrana celular
>Ss-ENOD-11	OsENODL1_like	5.812	20.781	Membrana celular
>Ss-ENOD-12	OsENODL1_like	6.077	20.961	Membrana celular
>Ss-ENOD-13	OsENODL1_like	5.815	20.902	Membrana celular
>Ss-ENOD-14	OsENODL1_like	6.473	19.952	Membrana celular
>Ss-ENOD-15	OsENODL1_like	6.469	19.985	Membrana celular
>Ss-ENOD-16	OsENODL1_like	6.95	21.358	Membrana celular
>Ss-ENOD-17	OsENODL1_like	8.076	21.468	Membrana celular
>Ss-ENOD-18	OsENODL1_like	8.076	21.102	Membrana celular
>Ss-ENOD-19	OsENODL1_like	4.973	17.128	Membrana celular
>Ss-ENOD-20	OsENODL1_like	6.733	21.782	Membrana celular
>Ss-ENOD-21	OsENODL1_like	7.213	21.876	Membrana celular
>Ss-ENOD-22	OsENODL1_like	7.813	20.083	Membrana celular
>Ss-ENOD-23	OsENODL1_like	7.332	17.47	Membrana celular

Por sua vez, para o grupo NIN, foram obtidas 36 sequências que apresentaram domínios conservados, sendo 28 associadas ao domínio PB1_NLP e oito ao domínio RWP-NLP (**Figura 2**). Para os transcritos desse grupo que continham o domínio PB1_NLP, a localização subcelular predita foi o núcleo, enquanto que para os transcritos com domínio RWP-NLP, a maioria estava associada ao cloroplasto. No que se refere ao ponto isoelétrico e peso molecular, os candidatos apresentaram variação de 4,50 a 12,36 e de 11,13-87,07 kD, respectivamente (**Tabela 3**).

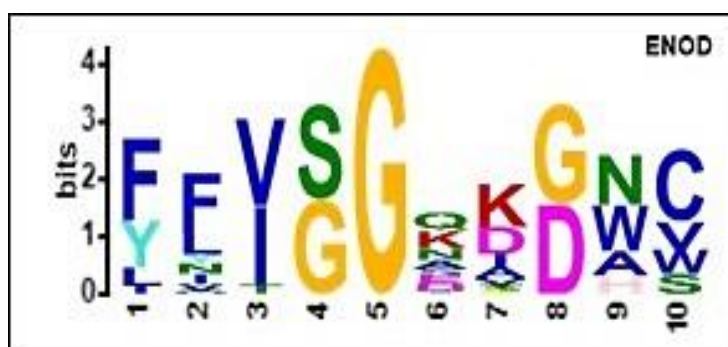
Tabela 3. Dados de caracterização estrutural e funcional do grupo NIN de nodulinas precoces. Sendo ID = Identidade, coluna de domínio identificando o mesmo para cada transcrito detectado, p.I = ponto isoelétrico; p.M = peso molecular, e a localização subcelular referente a compartimentação em que os transcritos se encontram.

ID	Domínio	p.I	p.M	Localização subcelular
>Ss-NIN-01	PB1_NLP	6.636	83.972	Núcleo
>Ss-NIN-02	PB1_NLP	6.432	73.231	Núcleo
>Ss-NIN-03	PB1_NLP	7.742	35.114	Núcleo
>Ss-NIN-04	PB1_NLP	6.611	86.856	Núcleo
>Ss-NIN-05	PB1_NLP	6.68	86.889	Núcleo
>Ss-NIN-06	PB1_NLP	6.611	86.91	Núcleo
>Ss-NIN-07	PB1_NLP	6.695	86.812	Núcleo
>Ss-NIN-08	PB1_NLP	6.519	87.069	Núcleo
>Ss-NIN-09	PB1_NLP	6.421	58.574	Núcleo
>Ss-NIN-10	PB1_NLP	6.127	57.38	Núcleo
>Ss-NIN-11	PB1_NLP	8.561	28.987	Núcleo
>Ss-NIN-12	PB1_NLP	6.675	27.928	Cloroplasto
>Ss-NIN-13	PB1_NLP	8.664	45.749	Núcleo
>Ss-NIN-14	PB1_NLP	4.625	37.992	Núcleo
>Ss-NIN-15	PB1_NLP	4.637	24.375	Núcleo
>Ss-NIN-16	PB1_NLP	4.999	11.128	Cloroplasto
>Ss-NIN-17	PB1_NLP	4.497	35.554	Cloroplasto

>Ss-NIN-18	PB1_NLP	8.862	42.753	Núcleo
>Ss-NIN-19	PB1_NLP	12.12	13.293	Cloroplasto
>Ss-NIN-20	PB1_NLP	7.919	49.44	Núcleo
>Ss-NIN-21	PB1_NLP	9.931	35.187	Núcleo
>Ss-NIN-22	PB1_NLP	7.898	13.933	Núcleo
>Ss-NIN-23	PB1_NLP	7.898	13.891	Núcleo
>Ss-NIN-24	PB1_NLP	5.912	12.039	Núcleo
>Ss-NIN-25	PB1_NLP	8.515	29.863	Núcleo
>Ss-NIN-26	PB1_NLP	8.515	29.782	Núcleo
>Ss-NIN-27	PB1_NLP	4.573	21.32	Núcleo
>Ss-NIN-28	PB1_NLP	6.036	12.463	Núcleo
>Ss-NIN-29	RWP-RK	7.742	35.114	Núcleo
>Ss-NIN-30	RWP-RK	5.017	32.326	Cloroplasto
>Ss-NIN-31	RWP-RK	12.358	11.182	Cloroplasto
>Ss-NIN-32	RWP-RK	6.007	51.87	Extracelular
>Ss-NIN-33	RWP-RK	5.045	33.675	Cloroplasto
>Ss-NIN-34	RWP-RK	9.817	42.181	Cloroplasto
>Ss-NIN-35	RWP-RK	9.964	41.977	Cloroplasto
>Ss-NIN-36	RWP-RK	10.078	20.25	Núcleo

Para analisar os motivos conservados dentro dos domínios das famílias gênicas de nodulinas precoces e as frequências mais prevalentes, foi realizado um alinhamento com as sequências que apresentaram domínios completos. Através da busca por motivos conservados no MEME, foi obtido um padrão associado com cada uma das famílias identificadas (**Figura 3**). Adicionalmente, os motivos também foram identificados através do alinhamento das sequências utilizando o MEGA7, sendo identificados cinco motivos conservados nas sequências pertencentes ao domínio osENODL1_like e três a nodulin-like. (**Figura 4**).

Figura 3. Motivos conservados dos grupos gênicos de nodulinas precoces. Cada letra pertence a um motivo detectado de cada grupo, nas quais letras de mesma cor apresentam motivos em comum. Letras em tamanho maior estão mais frequentes que letras de tamanho menores.



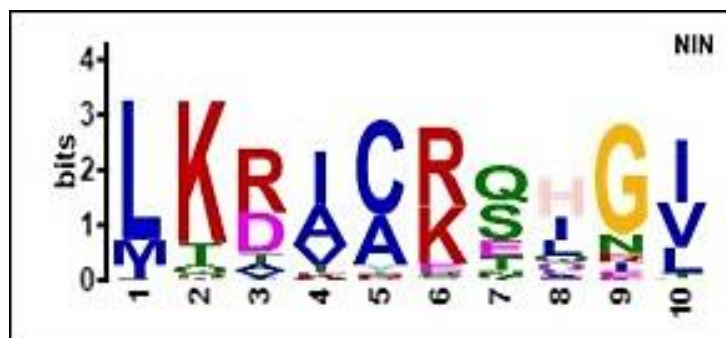
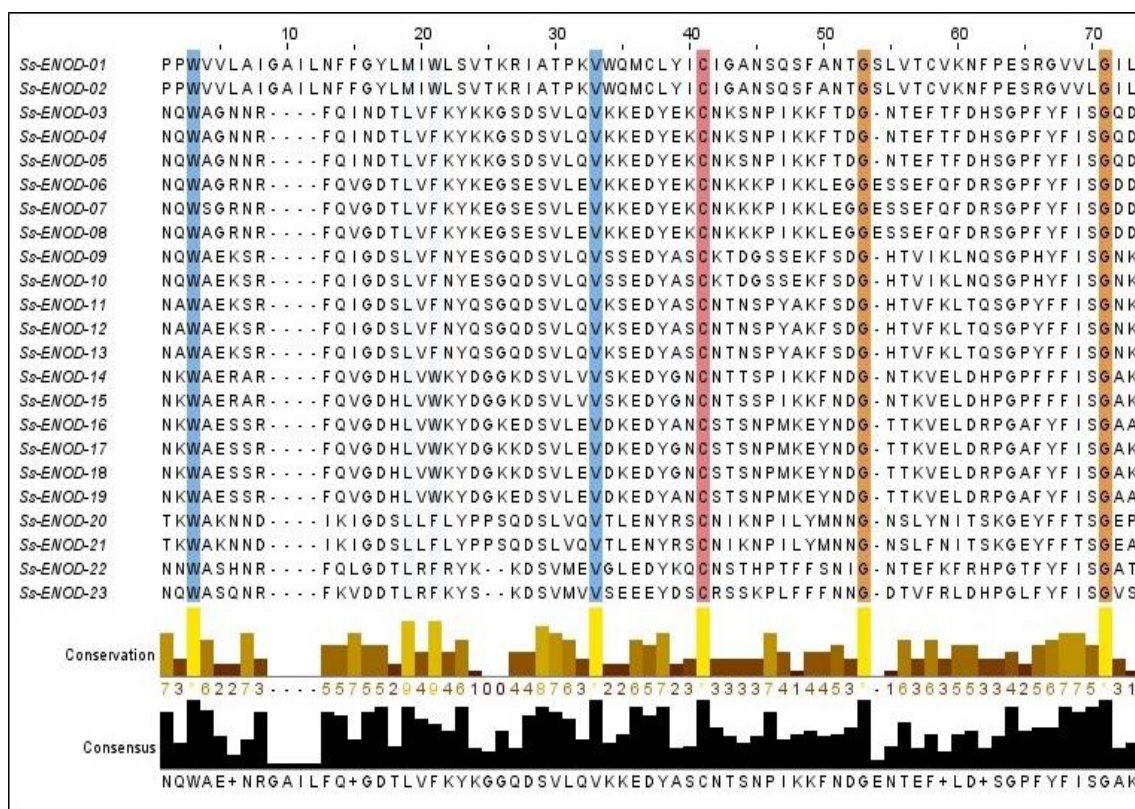


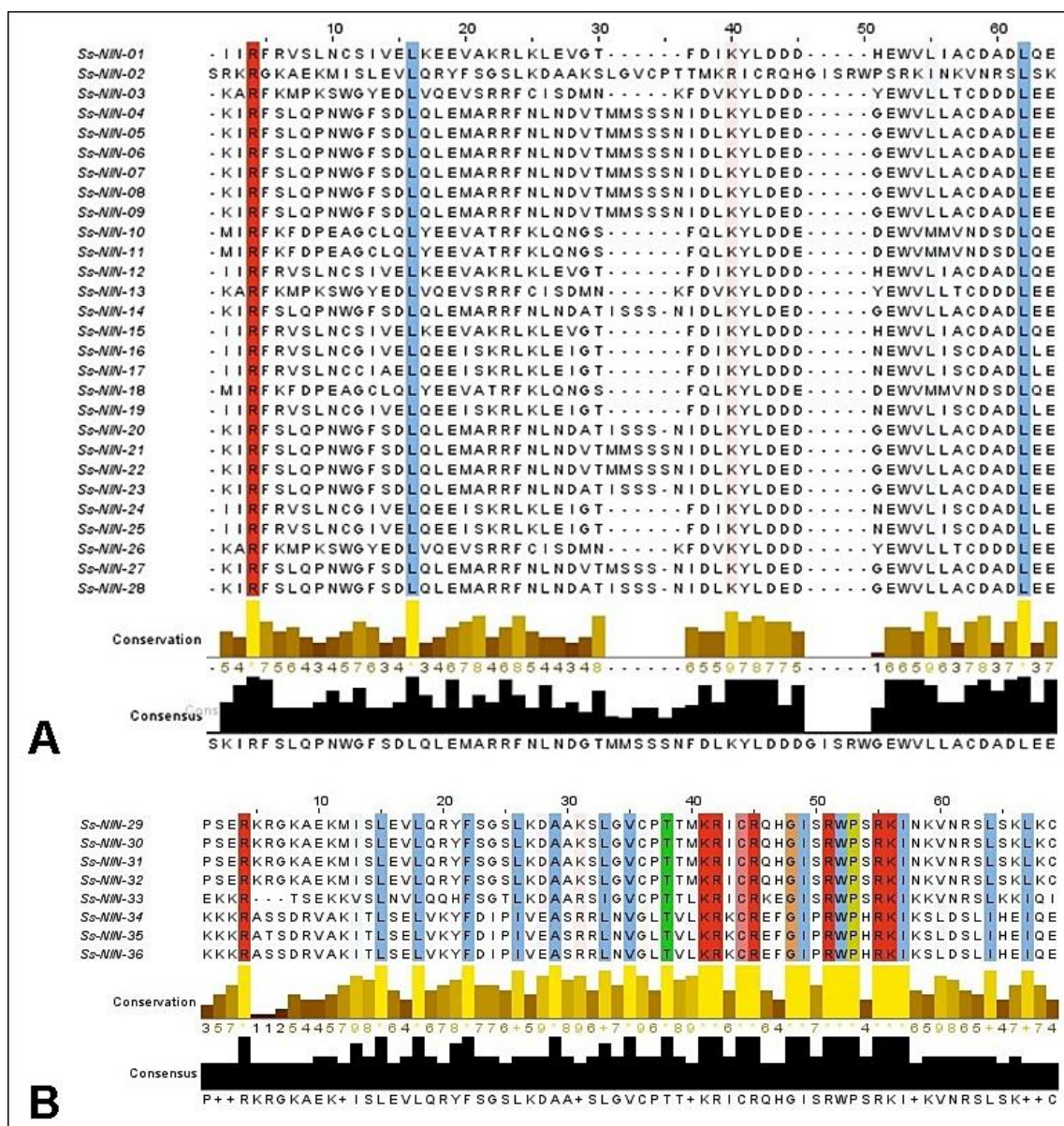
Figura 4. Alinhamento das sequências do grupo ENOD de nodulinas precoces. Colunas coloridas no alinhamento demonstram os motivos conservados detectados no alinhamento múltiplo do domínio. Conservation = Conservação (nível em que se encontra o grau de conservação dos motivos); Consensus = Consenso (nível do resíduo conservado).



Apesar de pouco citado na literatura, o padrão de alinhamento para o grupo ENOD revelou regiões importantes, contendo altos níveis de conservação, indicando possivelmente uma função importante dos transcritos no meio celular. Além disso, também foi possível observar um nível máximo do padrão de conservação para os aminoácidos Triptofano (W), Valina (V), Cisteína (C) e Glicina (G), os quais conferem estabilidade à proteína.

Para o grupo NIN, foram identificados três motivos para o domínio PB1_NLP e 23 motivos para o domínio RWP-RK (Figura 5).

Figura 5. Alinhamento das sequências do grupo NIN de nodulinas precoces. Colunas coloridas no alinhamento demonstram os motivos conservados detectados no alinhamento múltiplo do domínio. Conservation = Conservação (nível em que se encontra o grau de conservação dos motivos); Consensus = Consenso (nível do resíduo conservado). Legenda: A – Motivos conservados do domínio PB1_NLP; B – Motivos conservados do domínio RWP-RK.



Os transcritos que contém o domínio PB1_NLP se apresentam em algumas sequências com resíduos conservados semelhantes às encontradas

nos transcritos do domínio RWP-RK. No entanto, para este segundo domínio, é observada variação e um número maior de resíduos conservados presentes no alinhamento (**Figura 5**). No alinhamento A, observa-se o nível máximo do padrão de conservação para Arginina (R) e Leucina (L), como também é verificado no alinhamento B. Entretanto, no alinhamento B estão incluídos ainda os aminoácidos Fenilalanina (F), Isoleucina (I), Valina (V), Treonina (T), Lisina (K), Cisteína (C), Glicina (G), Triptofano (W) e Prolina (P) (**Figura 5**).

Vale destacar ainda os resíduos conservados que permaneceram tanto em ENOD quanto no alinhamento de RWP-RK do grupo NIN, sendo estes, triptofano, cisteína e glicina.

3.2 Distribuição cromossômica de genes para nodulinas precoces em *Arachis L.*

A ancoragem das sequências revelou que os transcritos da família gênica ENOD distribuíram-se em cinco cromossomos da espécie *Arachis duranensis*, sendo os cromossomos 3 e 10 os mais representativos, com seis genes em cada um. Os cromossomos 8 e 9 ancoraram cinco e quatro genes, respectivamente. Por sua vez, apenas dois genes ancoraram no cromossomo 6, sendo o cromossomo com o menor número de genes ancorados da análise (**Figura 6**).

O grupo de genes codificantes para NIN se distribuíram por seis cromossomos, sendo o cromossomo 7, com 13 genes, o mais representativo, seguido pelo cromossomo 5, com 10 genes ancorados, e pelos cromossomos 1, 8 e 9, com três genes ancorados em cada (**Figura 7**).

É interessante ressaltar que os dois grupos apresentaram transcritos ancorados nos cromossomos 3, 8 e 9, o que indica semelhança entre essas regiões. Além disso, a maioria dos transcritos ancorou em regiões intercaladas e nas porções terminais dos braços cromossômicos, próximas às regiões teloméricas.

Figura 6. Representação esquemática dos 23 transcritos do grupo ENOD ancorados em cinco cromossomos utilizando genoma de referência de *Arachis duranensis*. Note a quantidade de genes e sua localização ao longo dos cromossomos.

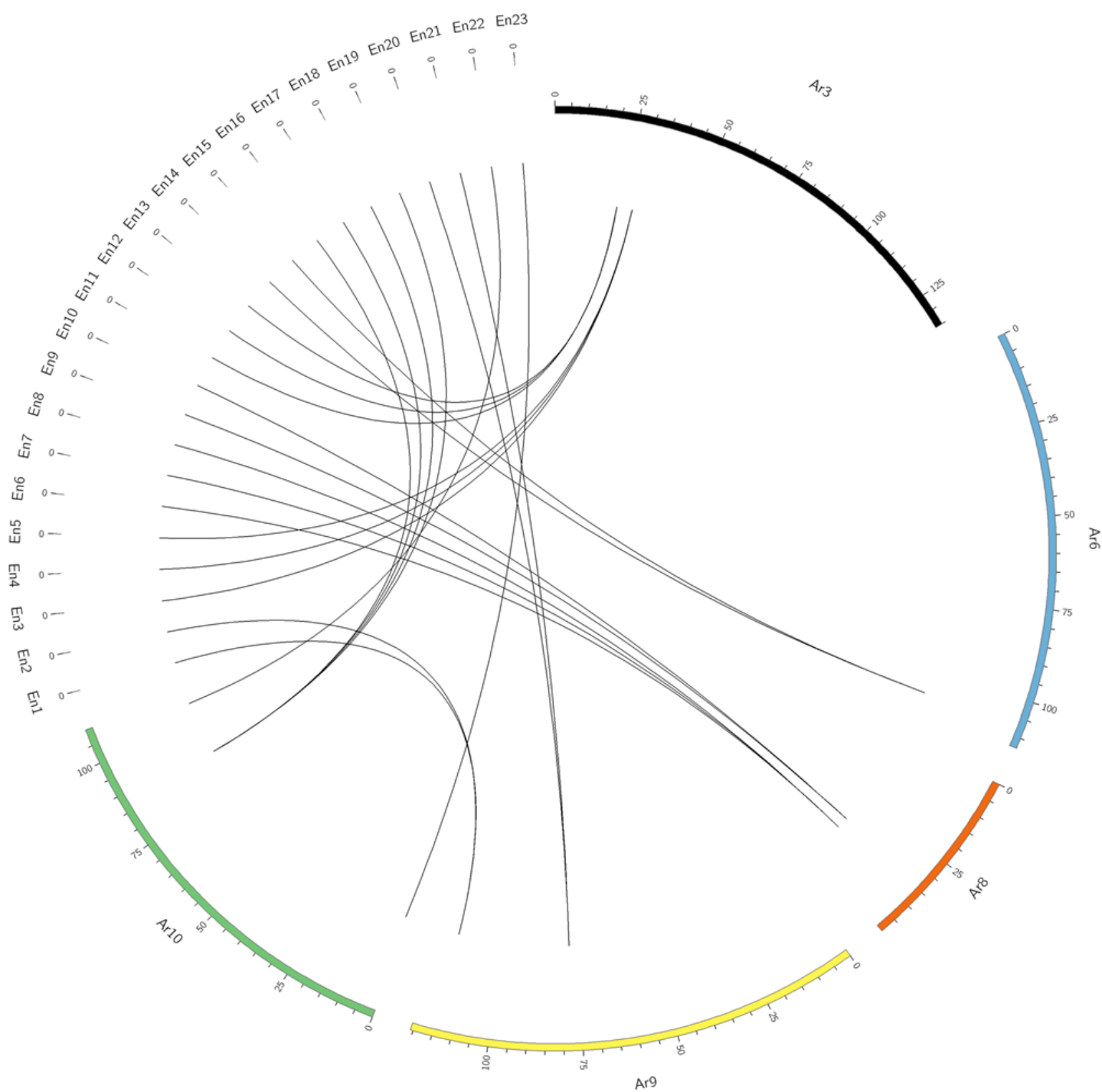
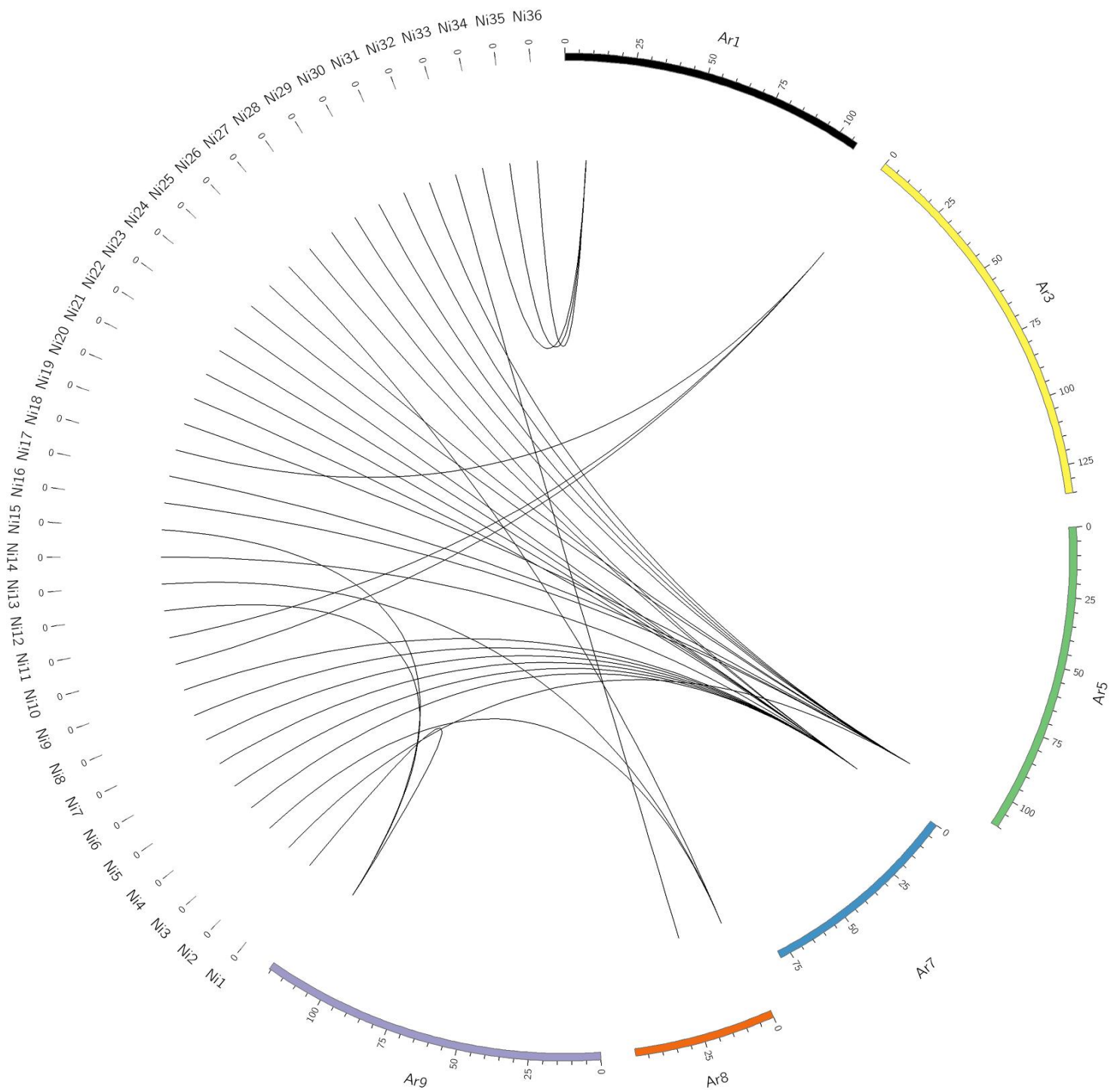


Figura 7. Representação esquemática dos 36 transcritos do grupo NIN ancorados em seis cromossomos utilizando genoma de referência de *Arachis duranensis*. Note a quantidade de genes e sua localização ao longo dos cromossomos.



3.3 Expressão diferencial dos grupos gênicos e desenho de *primers*

No perfil de expressão gerado através da clusterização hierárquica com os dados de RNA-Seq obtidos no transcriptoma de *S. scabra* nos tempos de 6 e 24 horas após suspensão de rega foram identificados, nos qual os parâmetros utilizados foi o log₂ de Fold Change, sendo este o quantitativo dos valores de expressão de um gene ou transcrito em duas condições diferentes. Identificou-se um total de 33 transcritos com expressão estatisticamente suportada, destes 16 transcritos foram induzidos, dos quais seis tiveram sua expressão no tempo de 6 h e dez tiveram no tempo de 24 h. Em relação aos reprimidos foram identificados 17, sendo oito no tempo de 6 h e nove no tempo de 24 h. Já para NIN, observou-se um total de 55 transcritos com valores de expressão, desse total 21 transcritos foram induzidos, sendo 14 no tempo de 6 h e 7 no tempo de 24 h e, com relação aos reprimidos, obteve-se um total de 34 transcritos, sendo dez no tempo de 6 h e 24 em 24 h. Observa-se que a maioria dos genes foi reprimida, no caso de ENOD apresentando o mesmo valor que o total de induzidos. No entanto, algumas sequências foram induzidas nos dois tempos, como por exemplo, ENOD-07, ENOD-08 e ENOD22 (**Figura 8**).

Para o grupo ENOD, a maioria das sequências induzidas foi identificada no tempo de 24 horas, em contraste com o grupo NIN, que teve em sua maioria genes reprimidos, principalmente com 24 horas de estresse, indicando um possível silenciamento na atividade desses genes em tempos mais prologados de estresse. Por exemplo, NIN-14 e NIN-30, apresentaram indução nos tempos iniciais, mas no tempo tardio, foram reprimidos (**Figura 8**).

Após a análise de expressão *in silico*, foram desenhados oito pares de *primers* para candidatos dos dois grupos de nodulinas precoces visando sua posterior utilização nas validações por RT-qPCR. Os candidatos que apresentaram indução em pelo menos um tempo de estresse foram priorizados para o desenho (**Tabela 4**).

Figura 8. Perfil de expressão *in silico* de nodulinas precoces no transcriptoma de *S. scabra*. A) Genes ENOD; B) Genes NIN. Vermelho = induzidos; Verde = reprimidos.

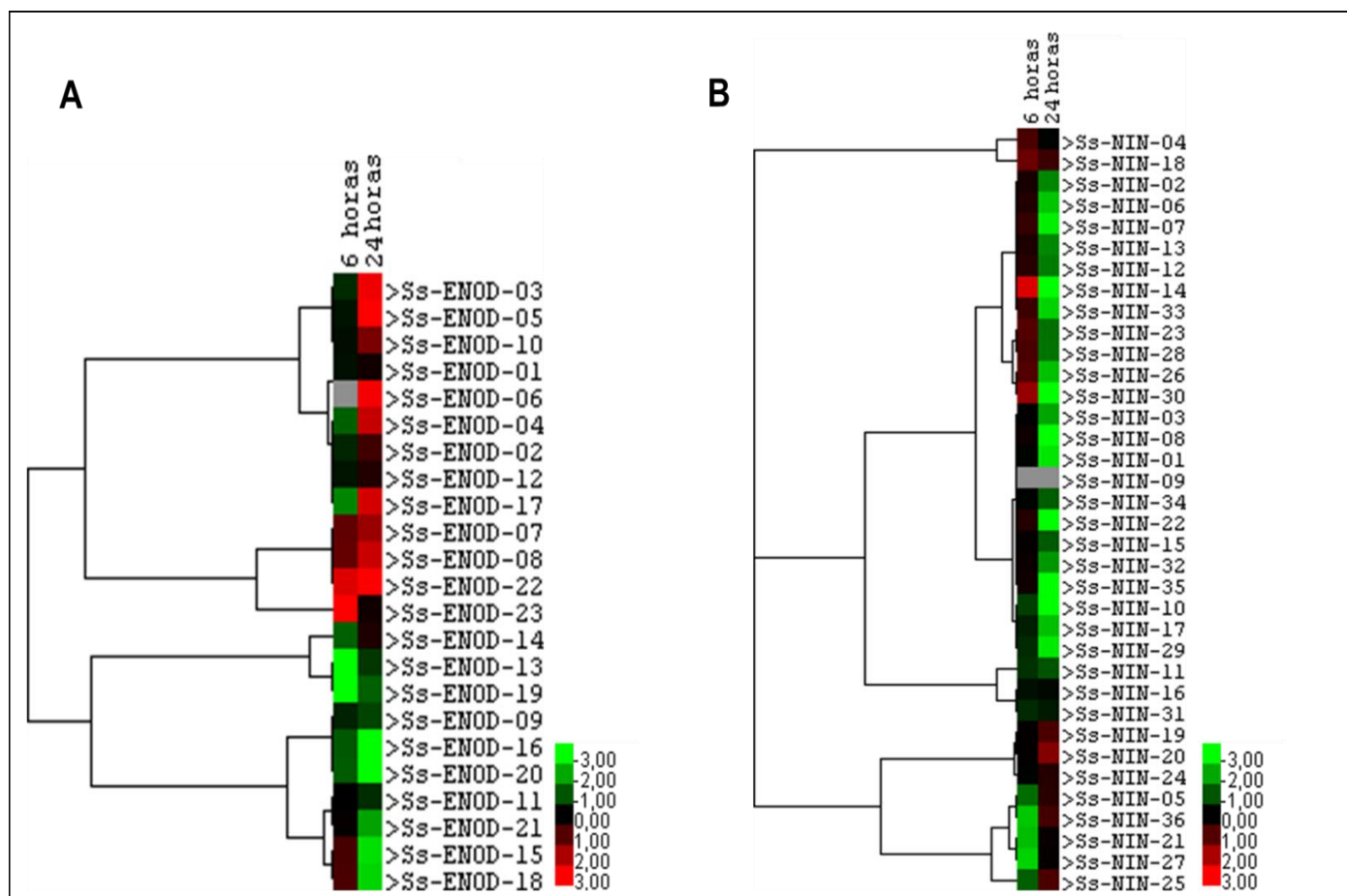


Tabela 4. Desenho de *primers* dos candidatos de nodulinas precoces. ID = identidade; Tm = temperatura de melting; GC = Guanina e Citosina (%); Self complementarity = Auto-complementaridade; pb = pares de base.

ID	Primers	Length	Tm	GC%	Self complementarity	Self 3' complementarity	Produto (pb)
TRINITY_DN101118_c0_g1_i1_F	ACCCCAGCAAACGGTGAAGCTG	22	65.8	59.09	4	3	107
TRINITY_DN101118_c0_g1_i1_R	ACGGGGCACTGCAAGAAAGGAC	22	65.54	59.09	5	1	
TRINITY_DN119623_c0_g1_i3_F	TGGAACCTGTTGTTCCCAGCCC	22	64.99	59.09	5	0	77
TRINITY_DN119623_c0_g1_i3_R	AACGTTGGTGGAAAGCGCAGG	20	63.89	60	6	0	
TRINITY_DN129730_c0_g1_i13_3_F	TGCAGCAGCAACGAAAATGGCG	22	65.51	54.55	6	2	102
TRINITY_DN129730_c0_g1_i13_3_R	ACACCAGGAACGCCAAACGG	20	63.56	60	3	1	
TRINITY_DN118017_c0_g1_i1_2_F	AGCCAACACTCCCAGAGCCATC	22	64.57	59.09	3	1	79
TRINITY_DN118017_c0_g1_i1_2_R	TGAAGGTCCAGCTGTTGCTCCG	22	65.51	59.09	6	2	
TRINITY_DN129730_c0_g1_i9_2_F	TGCAGCAGCAACGAAAATGGCG	22	65.51	54.55	6	2	97
TRINITY_DN129730_c0_g1_i9_2_R	AGGAACGCCAAACGGGAAGC	20	63.63	60	2	2	
TRINITY_DN127032_c0_g1_i5_3_F	GCGCAAGCATCATGGAAACCCC	22	65.14	59.09	5	0	88
TRINITY_DN127032_c0_g1_i5_3_R	ACTGCATGGGGCGAGTTTGC	20	63.99	60	4	2	
TRINITY_DN130946_c0_g1_i8_4_F	TCAAGCACGTCGATTTGGCCG	21	64.09	57.14	4	2	142
TRINITY_DN130946_c0_g1_i8_4_R	AAGGGGTGTGCGTGTCTGTTG	20	63.54	60	2	0	
TRINITY_DN129755_c0_g1_i5_2_F	AAGCTCCGAGAGCCAGATGAGC	22	64.37	59.09	6	2	101
TRINITY_DN129755_c0_g1_i5_2_R	AGGCCCGGCGAGCATAACAATC	22	65.51	59.09	5	0	

3.4 Análise fenética

Os dendrogramas foram gerados utilizando sequências já caracterizadas e disponíveis no NCBI, do grupo ENOD de nodulinas precoces de diferentes espécies, incluindo não leguminosas. Após o alinhamento com as sequências identificadas em *S. scabra* verificou-se que as leguminosas se agruparam, em contraste com espécies não leguminosas, e subdividiram-se em dois grupos. É importante ressaltar que transcritos que possuíam o domínio nodulin_like não se agruparam com transcritos que carregam outro domínio, indicando uma especificidade de domínio entre os transcritos do grupo ENOD (Figura 9).

Para o grupo NIN, foram gerados dois dendrogramas de acordo com o tipo de domínio apresentado no grupo. O primeiro com transcritos que carregam em sua sequência o domínio PB1_NLP (**Figura 10A**) e o segundo para transcritos que possuíam o domínio RWP-RK (**Figura 10B**). No dendrograma contendo sequências com o domínio PB1_NLP, foi possível observar a formação de *clusters* com sequências de outras espécies, sendo formados três grupos distintos que não se agruparam com transcritos que apresentavam o domínio RWP-RK (**Figura 10A**). Por sua vez, na árvore gerada com sequências que possuíam o domínio RWP-RK, foi verificada a formação de dois grupos, com o grupo maior possuindo sequências de Estilosantes e outras espécies (leguminosas e não leguminosas), e com o segundo grupo composto apenas por sequências de Estilosantes.

Figura 9. Dendrograma refletindo padrões de distribuição do domínio *osENOD1_like* e demais espécies. *Ss* = *Stylosanthes scabra*.

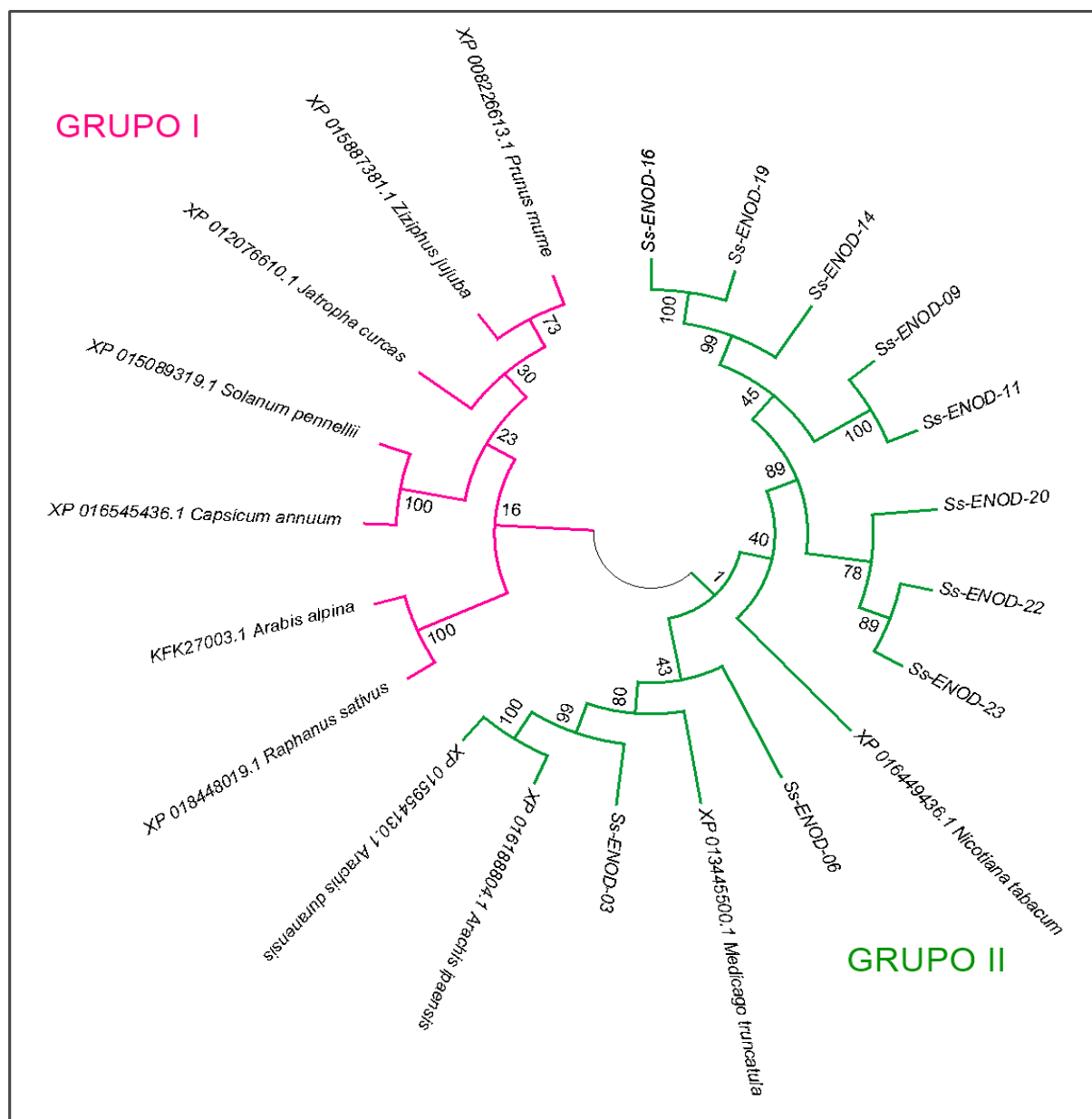
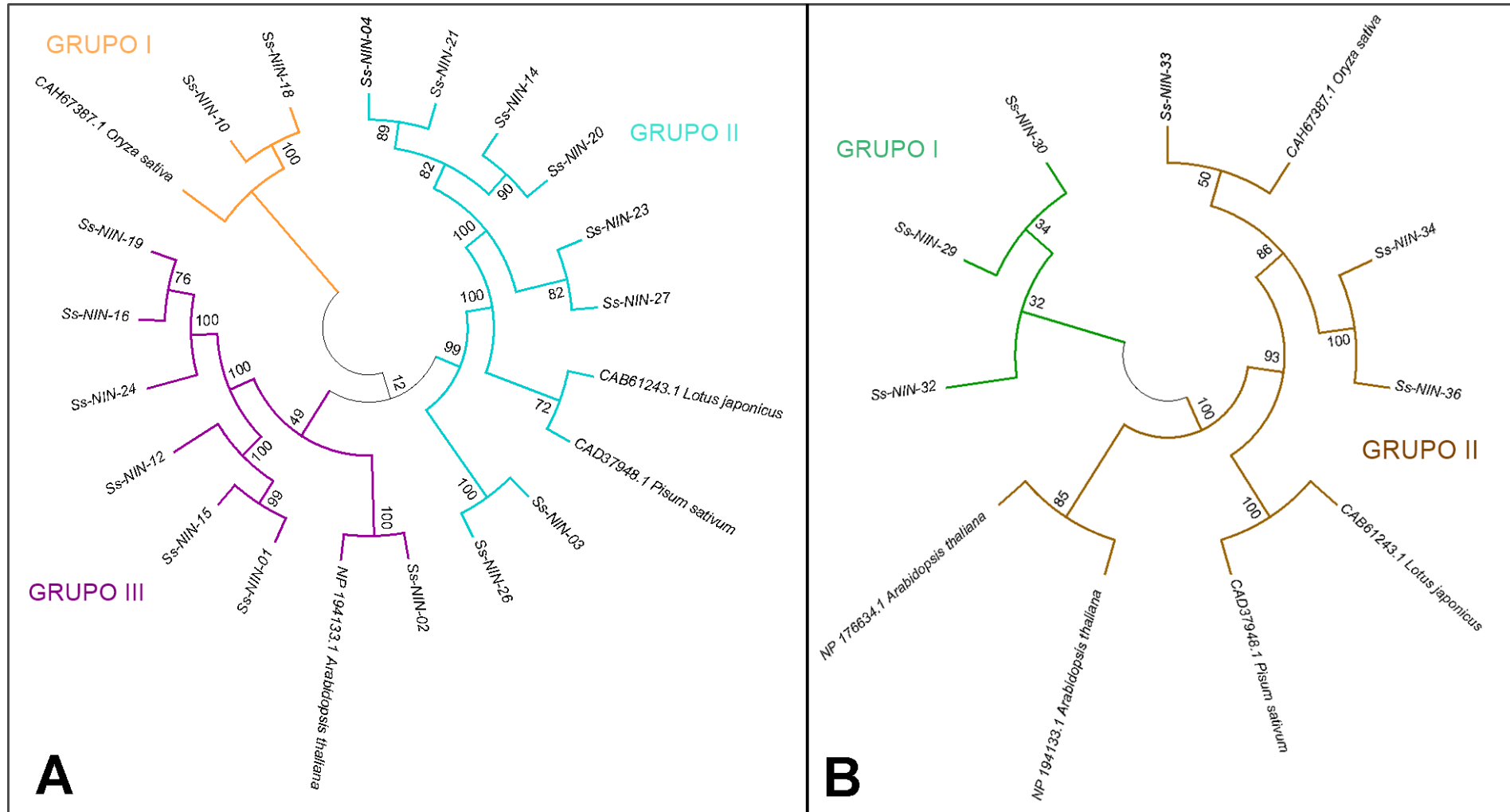


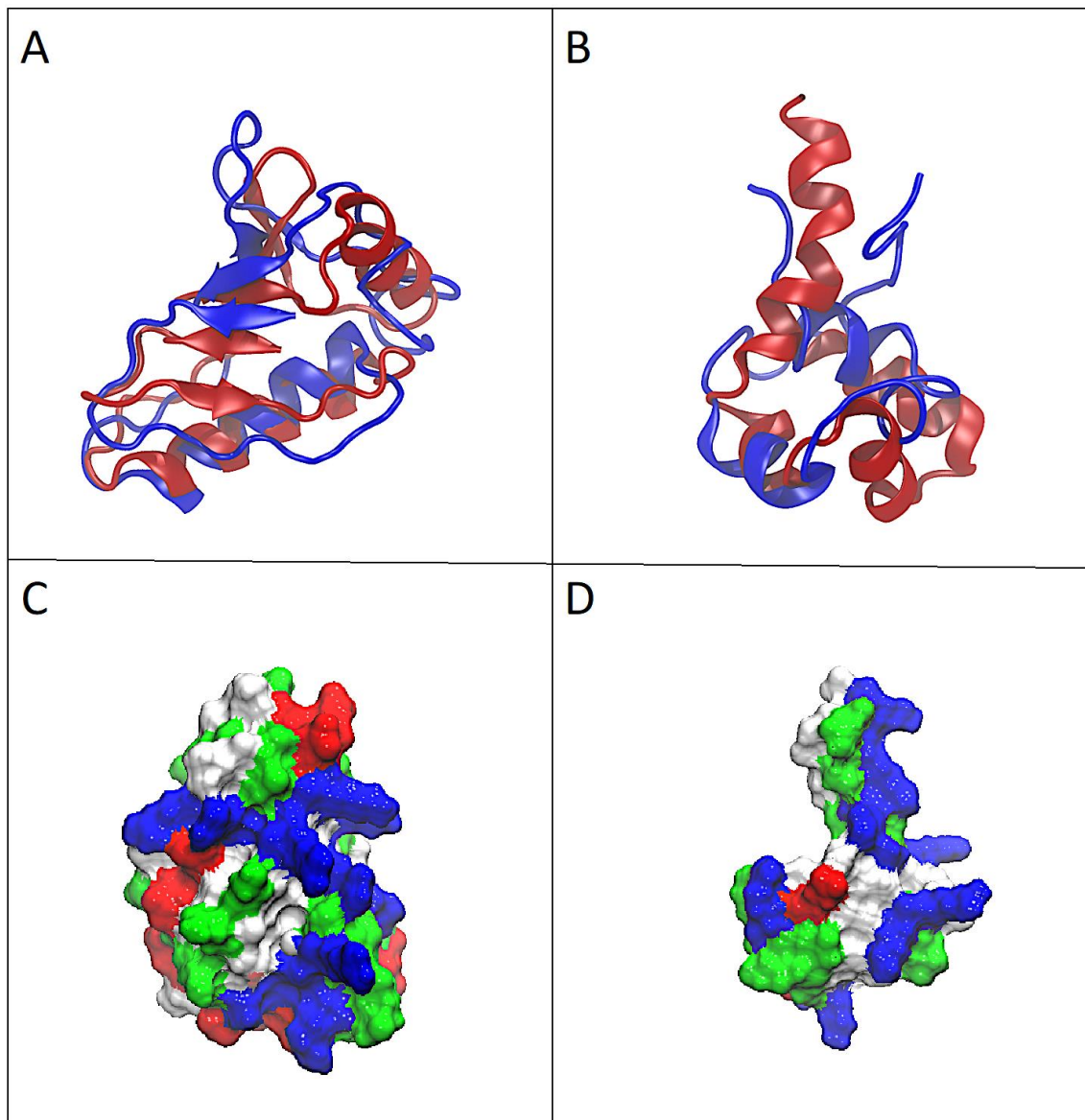
Figura 10. Dendrogramas refletindo padrões de distribuição do grupo NIN. A) Dendrograma do domínio PB1_NLP; B) Dendrograma do domínio RWP-RK. Ss = *Stylosanthes scabra*.



3.5 Modelagem tridimensional

As modelagens das sequências foram inicialmente baseadas em estruturas resolvidas disponíveis no Banco de Dados de Proteínas (PDB), usando o algoritmo do RosettaCM (**Figura 11**). Para a sequência com o domínio PB1_NLP, os modelos selecionados foram do domínio PDB1 de *Saccharomyces cerevisiae* (código PDB 1Q1O), *Homo sapiens* (código PDB 2G4S) e uma outra cepa de *Saccharomyces cerevisiae*, ATCC 204508 / S288c (código PDB 2KFJ), compartilhando 20,72%, 26,32% e 22,22% de identidade, respectivamente. Já para o domínio RWP_RK, os modelos selecionados foram uma proteína de ligação com o DNA de *Escherichia phage* (código PDB 1NER), um repressor de proteína de *Escherichia phage* (código PDB 2HIN) e a estrutura do complexo Cro N15 de *Escherichia phage* (código PDB 6ON0), compartilhando 23,68%, 27,78% e 27,78% de identidade, respectivamente. No entanto, nenhum dos modelos ofereceu cobertura total para as sequências candidatas, sendo também necessário utilizar as bibliotecas de fragmento para modelagem *de novo*.

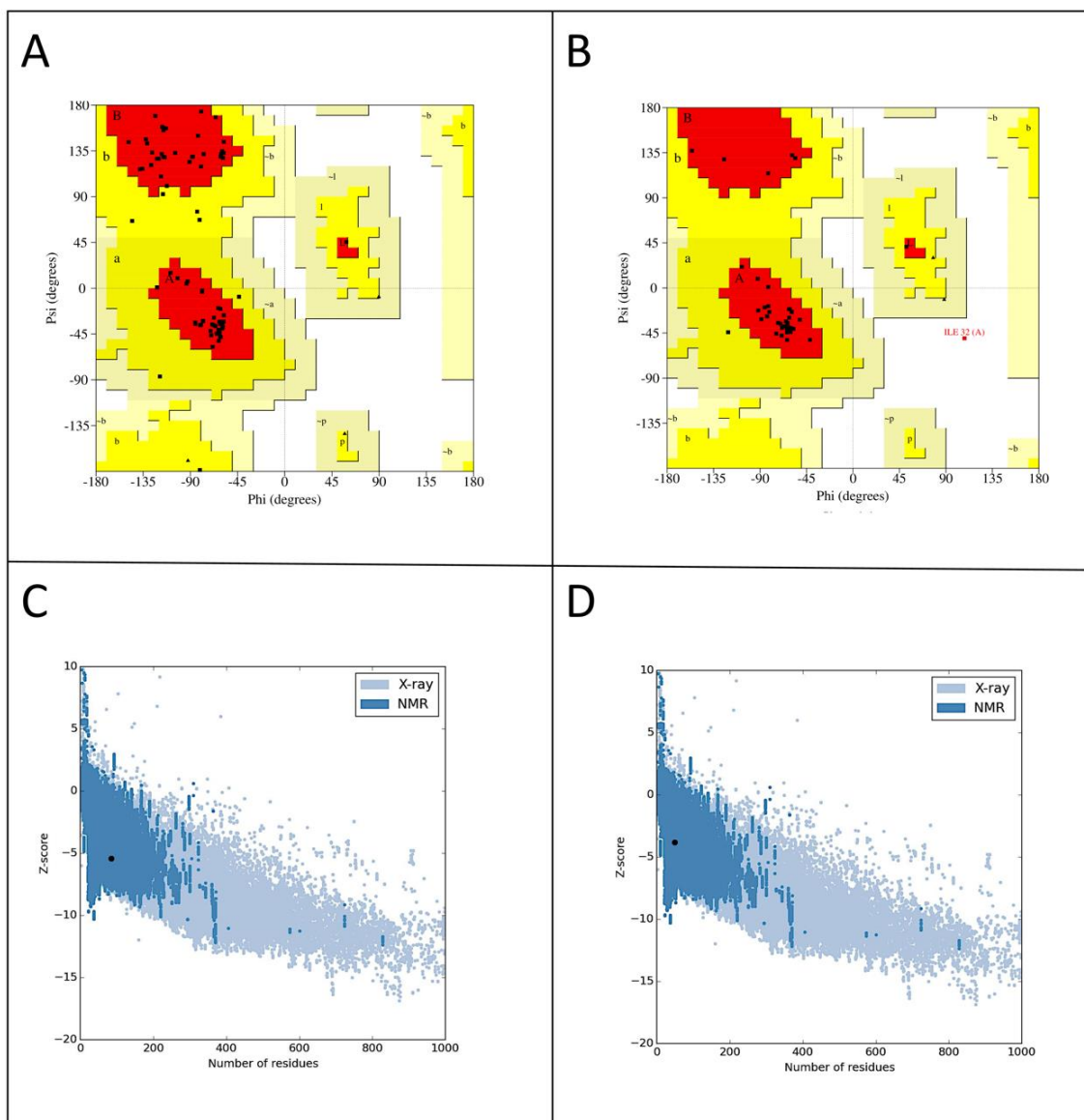
Figura 11. Estruturas tridimensionais das nodulinas antes (em vermelho) e depois (em azul) da simulação em dinâmica molecular. Em **A**- as estruturas de PB1_NLP (RMSD 4,5 nm) apresentando duas α -hélices uma β -folha antiparalela. Em **B**- as estruturas de RWP_RK (RMSD 9,5 nm) apresentando entre duas a três α -hélices. **C** e **D**- energia superficial das respectivas estruturas (PB1_NLP e RWP_RK), resíduos não polares (brancos), resíduos básicos (azul), resíduos ácidos (vermelho) e resíduos polares (verde).



Após as análises estereoquímicas dos aminoácidos, foi possível observar que todos os modelos gerados possuíam >90% dos aminoácidos nas posições mais favoráveis para todos os modelos (**Figura 12**). Além disso, o Z-score indicou que todos os modelos estavam dentro do intervalo de pontuação

tipicamente encontrado para estruturas de tamanhos similares.

Figura 12. Análise estereoquímicas dos modelos tridimensionais das nodulinas modeladas. Em A e B mostram os gráficos de Ramachandran mostrando que os dois modelos apresentam >90% dos aminoácidos nas regiões mais favoráveis de PB1_NLP e RWP_RK, respectivamente. Em C e D mostra cada modelo dentro do intervalo de pontuação indicada de estruturas similares.

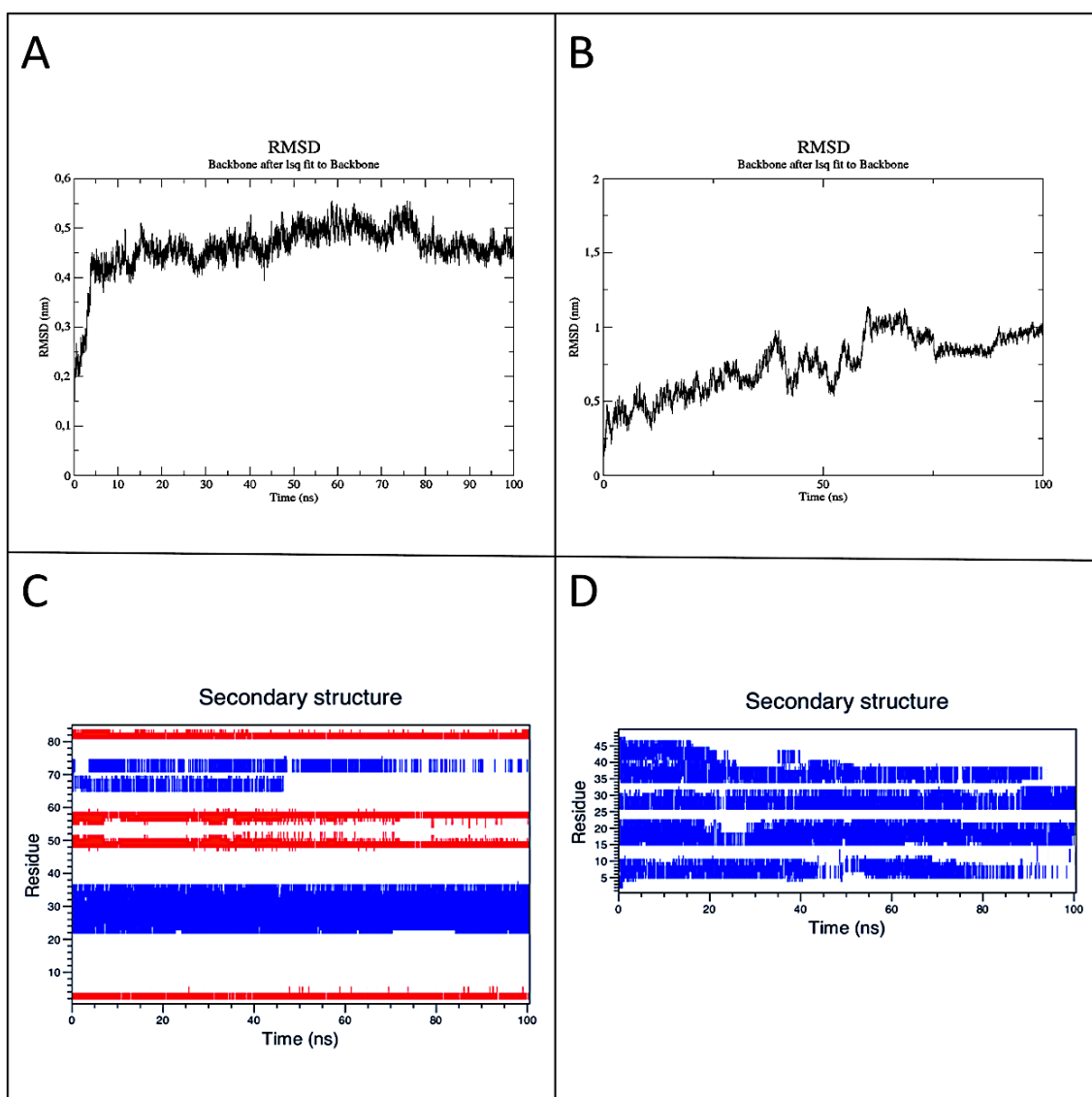


3.6 Simulação em dinâmica molecular

Os dois modelos de nodulinas foram submetidos com as mesmas condições na simulação de dinâmica molecular, a fim de observar as possíveis

modificações que os modelos poderiam sofrer em contato com os íons e o solvente. Portanto, alguns parâmetros foram avaliados como desvio quadrático médio (RMSD), flutuação quadrada média da raiz (RMSF) e estrutura secundária do peptídeo pelo gráfico DSSP (**Figura 13**).

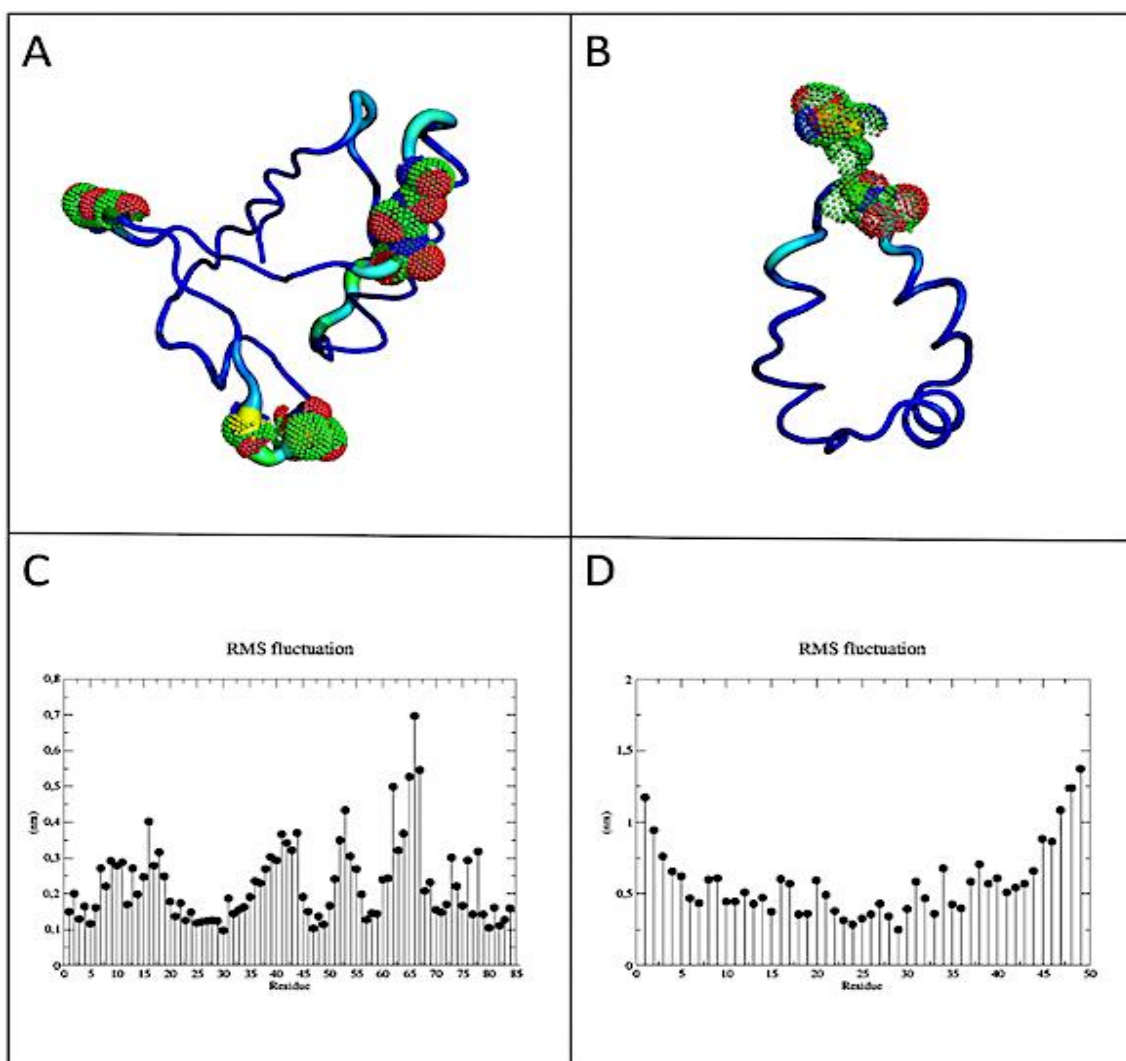
Figura 13. Análise da estrutura tridimensional das nodulinas após simulação em dinâmica molecular. Em A e B o RMSD de PB1_NLP e RWP_RK, respectivamente. Em C e D o gráfico da estrutura secundária ao longo dos 100 ns de simulação. Azul = α -hélices; Vermelho = β -folha.



Analisando os gráficos, é possível observar que os dois modelos deixaram a conformação original. O modelo de PB1_NLP exibiu a menor variação na

análise do RMSD (**14A**; **14B**). Em relação à estrutura secundária, observou-se que o ambos os modelos perderam estrutura compatíveis com α -hélices (**Figura 14C e 14D**). Ao observar a dinâmica de espalhamento dos átomos em conjunto com a flutuação dos aminoácidos, verificou-se que as posições que sofreram mais saíram da conformação estavam localizadas em regiões de alça e das extremidades dos modelos, sendo PB1_NLP o aquele apresentou aminoácidos mais flexíveis (**Figura 14**).

Figura 14. Análise do espalhamento dos átomos dos modelos tridimensionais das nodulinas modeladas. Em **A e B** o gráfico da temperatura tridimensional (b-factor) de PB1_NLP e RWP_RK, respectivamente. Em **C e D** o gráfico de flutuação dos aminoácidos dos respectivos modelos ao longo dos 100 ns da simulação em dinâmica molecular.



4 DISCUSSÃO

4.1 Identificação de genes de nodulinas

A análise do transcriptoma de *S. scabra* revelou a presença de dois grupos de genes de nodulinas, corroborando ao verificado na literatura, que relata a presença destes grupos em outras leguminosas, e em algumas não leguminosas, como: *Cucurbita pepo*, *Oryza sativa* e *Arabidopsis thaliana* (VIEIRA-MELLO, 2011). É importante ressaltar que sequências oriundas de abordagens transcriptômicas apresentam-se como um valioso recurso para espécies que não possuem o genoma completo sequenciado, como é o caso da espécie em estudo, pois auxiliam na descoberta de novos genes, no desenvolvimento de marcadores moleculares e de novas cultivares tolerantes/resistentes, além de permitirem análises evolutivas e de expressão de famílias gênicas específicas, como o caso de nodulinas precoces detectadas no presente estudo, reforçando que estas são de extrema importância no que se refere a nodulação e fixação biológica do nitrogênio (GUIMARÃES et al., 2012).

4.2 Domínios e motivos conservados

De acordo com dados da literatura, o domínio conservado ENODLs, pertencente ao grupo ENOD das nodulinas precoces, desempenha um papel na célula de diferenciação e reorganização da parede celular durante a nodulação em leguminosas. Algumas nodulinas, incluindo ENOD16 e ENOD8, em *Medicago truncatula*, atuam na manutenção da integridade do simbiossomo, sendo fundamentais para manter a FBN durante estresses. As alterações ocasionadas no metabolismo dos lipídeos constituintes da membrana do simbiossomo estão provavelmente envolvidas na proteção contra estresses abióticos, uma vez que alterações estruturais nesta membrana podem afetar a atividade de enzimas importantes (HIRSCH et al., 2009).

Em relação ao grupo NIN, diversos estudos relatam sua atuação como reguladores centrais da nodulação, desempenhando papéis essenciais na epiderme e no córtex radicular, sendo necessários no início de infecções bacterianas e na promoção da organogênese dos nódulos (SCHAUSER et al., 1999; MARSH et al., 2007; SOYANO et al., 2013). Tal descrição é essencial para entender a função do gene na maquinaria celular, visto que os domínios e

motivos conservados indicam que estes desempenham papéis importantes. Neste sentido, tem-se que esses genes também podem atuar como reguladores negativos, inibindo respostas adicionais de NF após a ativação inicial desta via de sinalização na epiderme da raiz e promovendo a autorregulação da nodulação, que limita o número final de nódulos que se formam (MARSH et al., 2007; SOYANO et al., 2014; YORO et al., 2014). Conseqüentemente, este mecanismo pode indicar um dos motivos pelos quais os genes NIN se mostraram mais reprimidos em nossas análises de regulação de expressão.

4.3 Localização subcelular e atividade biológica

Todas as sequências presentes no grupo ENOD foram localizadas na membrana celular. É importante ressaltar que as membranas celulares estão entre as principais estruturas afetadas pelo estresse hídrico. De acordo com Gigon et al., (2004), em condições de déficit hídrico, o metabolismo de lipídeos da membrana é alterado e apresenta funções importantes na aclimatação ao estresse, o que pode sugerir que *S. scabra* pode estar utilizando este mecanismo de ativação da defesa quando submetida ao déficit hídrico. Neste cenário é possível entender a localização do grupo, pois esta pode indicar sua intensa participação da modulação fisiológica da planta em resposta ao estresse.

As sequências do grupo NIN estavam direcionadas em sua maioria direcionadas para o núcleo, confirmando a sua associação na ativação da expressão de outros genes, no que se refere a cascata de genes envolvidos na nodulação. Neste aspecto, Soyano et al. (2014) descreve que, no córtex, genes NIN ativam a expressão dos peptídeos CLE, que, por sua vez, podem estar envolvidos na regulação da entrada de nutrientes para a célula.

4.4 Distribuição cromossômica no genoma de *Arachis duranensis*

As porções terminais dos braços cromossômicos, próximas às regiões teloméricas, onde a maioria das sequências foram encontradas são regiões compostas por DNA repetitivo, conferindo estabilidade aos cromossomos. As sequências teloméricas, por sua vez, são responsáveis por proteger as extremidades dos cromossomos, formando um complexo de dobramento, fazendo com que as extremidades se tornem protegidas (BERNAL; TUSELL,

2018).

A ancoragem também demonstra as regiões conservadas em diferentes cromossomos, nas quais podem ter se originado por diferentes processos. As análises realizadas por Bertoli et al. (2016) estudando sequência do genoma de espécies de *Arachis duranensis* e *Arachis ipaensis* sugerem que as linhagens de *Arachis* vêm acumulando mutações de forma relativamente rápida desde a divergência do clado.

Os genomas da maioria dos organismos apresentam diversas cópias de genes que estão envolvidos tanto na estrutura, quanto função, bem como apresentam uma história evolutiva comum, entre si (SANTOS, 2007). Neste cenário, entendendo que Estilosantes tem *Arachis duranensis*, espécie utilizada para ancoragem, como seu ancestral mais antigo no caldo, observa-se indicativos que a presença dos genes encontrados em diferentes cromossomos se dá por meio da conservação apresentada entre estas espécies em função do tempo.

Segundo Santos (2007), a larga distribuição de motivos e proteínas conservadas nos filós atuais indica que a diversidade atual existente, provém em grande parte, de novas combinações e modificações dos caracteres moleculares preexistentes, o que pode-se observar na ancoragem para os genes de nodulinas precoces (ENOD e NIN) descritos nas análises, visto que apresentam regiões com genes bem próximos, podendo ter ocorrido o processo de recombinação, de caracteres moleculares entre eles. Ainda segundo a autora, o aparecimento de novas funções gênicas nos diferentes organismos por duplicação gênica provém, em grande parte, da reutilização do material genético preexistente. De acordo com os dados apresentados no estudo, observa-se a possibilidade de ocorrência de eventos de translocações recíprocas, ou ainda a presença do grupo em distintos cromossomos ter se dado por meio de processos de duplicação genômica seguidos de diploidização originando genes com novas funções dentro de um mesmo clado.

4.5 Análise fenética

Neste estudo, observou-se a divisão de três grandes grupos no dendrograma gerado com sequências do grupo ENOD de *S. cabra* e diferentes plantas presentes no NCBI, incluindo as espécies dicotiledôneas e

monocotiledôneas. Os grupos foram formados com espécies mais próximas, indicando além de sua presença e possível função das raízes de *S. cabra* apresentando semelhanças com espécies bastante próximas, como o caso de *Arachis L.* e *Medicago spp.* Estas semelhanças e diferenças podem indicar que essas sequências possivelmente apresentam diferenças estruturais ou funcionais, o que de fato pode ser observado nas análises, visto que é visualização uma especificidade no que se refere ao grupo ENOD em leguminosas.

Para NIN, as sequências com domínio PB1_NLP, formaram três grupos, observou-se que um agrupou com outras leguminosas *Lotus japoninus* e *Psium sativum*, e outro grupo maior que se agrupou apenas com *Arabidopsis thaliana*, bem como duas sequências específicas que se agruparam apenas com *Orzya sativa*, indicando uma possível especificidade das sequências no grupo confirmando as conclusões de Riechman et al. (2000) de que não há parentes próximos para as proteínas de leguminosas NIN, em arroz ou *Arabidopsis*, indicando uma possível funcionalidade compartilhada, anterior à separação das leguminosas. Entretanto, para o domínio RWP-RK, observou-se dois grandes grupos, no qual um agrupou-se apenas com sequências do próprio grupo e o segundo com outras leguminosas e também *A. thaliana*, sugerindo uma funcionalidade exclusiva dentro da espécie, intimamente associada com a evolução do processo de nodulação.

Bezerra-Neto et al. (2019) também identificou esta abundância no seu estudo, analisando diversidade e evolução de aquaporinas, nas quais também são tidas como nodulinas, contendo grupos gênicos presentes em diferentes espécies, sendo identificadas em mono e dicotiledôneas, bem como *Arabidopsis*.

4.6 Expressão gênica diferencial de nodulinas precoces

A indução de genes que pertencem a diversas categorias funcionais indica a existência de uma complexa resposta nos nódulos das raízes de leguminosas. Neste sentido, vários genes foram identificados como induzidos tanto em células corticais micorrizadas como também durante a colonização de raiz rizobiana em espécies de leguminosas, e dentre estes, encontra-se os genes de nodulinas precoces ENOD2, ENOD40 (VAN RHIJN et al., 1997) e ENOD11 (CHABAUD et al., 2002).

A expressão de genes NIN também é induzida nas células epidérmicas

em resposta ao NF. Dentro da epiderme, genes NIN promovem a expressão do gene NPL e reprimem a expressão do gene ENOD11 ao competir com o gene ERN1. O gene ERN1 promove o sinal da epiderme para o córtex, que por sua vez, inicia a sinalização da citocinina, promovendo a expressão de NIN nas células corticais. Genes NIN e aqueles da sinalização da citocinina atuam dentro de um ciclo de retroalimentação positiva, com NIN promovendo a expressão de CRE1, enquanto a sinalização de citocinina promove a expressão de NIN, expandindo o grau de respostas de nodulação dentro do córtex da raiz. Níveis elevados de genes NIN no córtex direcionam a organogênese do nódulo através da promoção de NF-YA1 e NF-YA2 (SOYANO et al., 2013).

Cheng-Wu et. al., (2019) relatam que os genes NIN desempenham papel central na regulação do crescimento e desenvolvimento celular em infecções, como também apontam os genes NIN como sendo necessários para a remodelação da parede celular, no ciclo celular e na expressão de genes transportadores.

Viera-Mello et al. (2011) relatou em seu estudo que genes NIN são pouco expressos, ao estudar expressão gênica de nodulinas em cana-de-açúcar, corroborando com os resultados apresentados no presente trabalho, onde a maioria das sequências NIN foi reprimida, o que pode indicar que a sua repressão se deva a uma possível diminuição de gasto energético em termos de produção de RNA e/ou proteínas, visto que a maioria das sequências foi encontrada no núcleo. Resultados semelhantes foram reportados por Ferguson et al., (2018), analisando a expressão de genes NIN no processo de regulação da nodulação em leguminosas, onde tiveram expressão reprimida por conta de alterações de alguns compostos.

Tais resultados demonstram que os genes NIN podem ser genes de iniciação responsivos à sinalização de outros genes que possivelmente estão ligados ao controle e regulação do rizóbio, indicando que possuem papel fundamental nesta interação e participação frente a condições adversas.

Por outro lado, alguns estudos apontam a superexpressão de NIN após a percepção do fator NOD, reduzida em mutantes CYCLOPS, no qual corresponde a um substrato de fosforilação de CCaMK (YANO et al., 2008). Singh et al. (2014) mostraram que CYCLOPS é um ativador da transcrição de ligação ao DNA regulado por CCaMK, o qual inicia a expressão do gene NIN levando à

organogênese do nódulo. Tais resultados revelam um mecanismo direto de decodificação do sinal de cálcio nuclear pelo complexo CCaMK/CYCLOPS. A percepção de sinais de cálcio estimula a fosforilação de CYCLOPS, que por sua vez, ativa a expressão gênica para o início da divisão celular e desenvolvimento de órgãos simbióticos.

Em nosso estudo, os dados de expressão encontrados para NIN estão contrastantes aos observados para o grupo ENOD, que em sua maioria foi induzido, principalmente no tempo de 24 horas, sugerindo sua atividade intensa ao longo do estresse. Os genes ENOD foram essencialmente encontrados na membrana celular, sugerindo sua participação direta na defesa e proteção celular. Entretanto, é necessário enfatizar que genes que codificam proteínas e transportadores de membrana também representam uma porção relativamente grande de todos os genes sob o controle de NIN (CHENG-WU, et. al., 2019).

A sinalização do fator NOD induz a expressão de NIN, ERN1 e de outros fatores de transcrição. O ERN1 é necessário para a regulação de alguns genes com expressão simbiótica específica, enquanto o NIN controla uma diversidade de funções, incluindo modificações da parede celular via NPL. Neste sentido, destaca-se que a relação entre NIN, NF-YA1 e ERN1 ajudam a compreender melhor como as redes transcricionais complexas atuam.

Em geral, genes induzidos participam ativamente durante estresse, o que pode indicar que de fato ENOD participa intensamente do processo nodulação e possivelmente de defesa quando Estilosantes foi submetida a estresse hídrico, em contraste com NIN que houve um recuo, entretanto, isto pode explicado pelo o fato de NIN participar da ativação de outras cascata, por isso este pode não ter uma expressão tão representativa ao avançar do tempo do estresse.

4.7 Modelagem e dinâmica estrutural

Observou-se após a modelagem que a posição dos aminoácidos estava na região mais favorável (acima de 90%) demonstrando que os modelos foram validados. Após a simulação destas verificou-se que para o domínio conservado PB1_NLP ocorreu a perda de uma α -hélices ao longo das análises, enquanto que a β -folha se manteve, bem como para domínio RWP-RK apresentou perda das estruturas compatíveis com α -hélices. Para o primeiro domínio verificou-se

que os aminoácidos 15, 53, 62, 64, 65 e 66, foram os que apresentaram maior flutuação ou dinâmica de dispersão dos átomos em relação aos demais aminoácidos, enquanto que para RWP-RK, as extremidades foram as que apresentaram maior nível de variação, como por exemplo os aminoácidos 1, 49 e 50. Após as análises de modelagem e simulação, verificou-se que a proteína NIN-14 com domínio PB1_NLP foi mais estável, esta estabilidade pode ser explicada devido as ligações de menor força, além de apresentar uma ponte dissulfeto, enquanto que NIN-30 com domínio RWP-RK, por ser uma estrutura menor e não ser estabilidade por pontes dissulfeto apresenta maior flexibilidade consequentemente menor estabilidade.

Vale destacar que estes dados revelam o primeiro modelo de proteínas pertencentes ao grupo NIN de nodulinas precoces, nos quais não se há relatos na literatura, assim como não foram resolvidos nem por difração de raio-X e ressonância magnética. Reforçando a ideia de conservação Schauser, Wieloch e Stougaard (2005) verificando o domínio RWP-RK em *Arabidopsis*, *Oryza sativa* e *Lotus japonicus*, relatou que regiões correspondentes a NIN de Lotus e ervilha precisava de mais resultados comparativos, entretanto nenhuma estrutura destas é descrita. Estes ainda sugerem que ocorre uma característica especial das proteínas NIN de leguminosas em comparação com não leguminosas, assim como o domínio RWP-RK é encontrado em uma variedade de proteínas, indicando um módulo funcional com estrutura e função conservadas.

5 CONSIDERAÇÕES FINAIS

A presente pesquisa ajudou a ampliar os conhecimentos existentes sobre o processo de nodulação de *S. scabra*, principalmente no que se refere aos mecanismos de resposta desse vegetal frente a estresses abióticos. Foi possível a geração de um banco de dados curado com sequências de transcritos relacionados à nodulinas diferencialmente expressas nos tempos experimentais, permitindo a caracterização da orquestração transcricional da via de nodulação em *S. scabra*. Adicionalmente, permitiu o entendimento do padrão de distribuição cromossômico de duas categorias de nodulinas precoces (ENOD e NIN) no genoma da espécie aparentada *A. duranensis*, como padrão para futuras análises de sintenia. A ancoragem indica a existência de regiões conservadas

em cromossomos distintos, possivelmente devido a duplicações segmentais seguida de translocações recíprocas ou ainda a processos de duplicação genômica seguidos de diploidização. Tais regiões redundantes provavelmente permitiram a neofuncionalização de genes codificantes de nodulinas precoces em *A. duranensis*, o que poderá ser comparado em *S. scabra* caso seu genoma completo seja sequenciado no futuro. Os dados ainda revelam dois modelos de proteínas do grupo NIN de nodulinas precoces, destacando, que estes são os primeiros, visto que não se tem depositado no banco de dados de proteínas ou relatados na literatura.

REFERÊNCIAS

- BERENDSEN, H. J.; POSTMA, J. P.; VAN GUSTEREN, W. F.; HERMANS, J. Interaction models for water in relation to protein hydration. **Intermolecular forces**, p. 331-342, 1981.
- BERNAL, A.; TUSELL, L. Telomeres: implications for Cancer development International. **Journal of Molecular Sciences**, v. 19, n. 294, p. 1-21, 2018.
- BERTIOLO, D. J. et al. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. **Nature Genetics**, v. 48, p. 438-446, 2016.
- BEZERRA-NETO et al. Plant Aquaporins: Diversity, Evolution and Biotechnological Applications. **Current Protein and Peptide Science**, v. 20, p. 368-395, 2019.
- BUSTIN, S. A; et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. **Clinical Chemistry**, v. 55 n. 4, p. 611-22, 2009.
- CHABAUD, M.; VENARD, C.; DEFAUX-PETRAS, A.; BÉCARD, G; BARKER, D. G. Targeted inoculation of *Medicago truncatula* in vitro root cultures reveals MtENOD11 expression during early stages of infection by arbuscular mycorrhizal fungi. **New Phytologist**, v. 156, p. 265-273, 2002.
- CHENG-WU, L et al. NIN atua como um hub de rede controlando um módulo de crescimento necessário para a infecção Rhizobial. **Plant physiology**, v.179, n. 4, p.1704-1722, 2019.
- FERGUSON, B. J. et al. Legume nodulation: the host controls the party. **Plant, Cell & Environment**, p. 1–11, 2018.
- GALHARDO, M. et al. Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network. **Nucleic Acids Research**, v. 42, p. 1474-1496, 2013.
- GIGON, A.; MATOS, A.-R.; LAFFRAY, D.; ZUILY-FODIL, Y.; PHAM-THI, A. Effect of drought stress on lipid metabolism in the leaves of *Arabidopsis thaliana* (Ecotype Columbia). **Annals of Botany**, v.94, p.345-351, 2004.
- GRABHERR, M. G et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. **Nature biotechnology**, v. 29, n. 644, 2011.
- GUIMARÃES, A. A. et al. Genetic and symbiotic diversity of nitrogenfixing bacteria isolated from agricultural soils in the western amazon by using cowpea as the trap plant. **Applied and Environmental Microbiology**. v. 78, n.18, p. 6726- 6733, 2012.
- HAAS, B. J. et al. De novo transcript sequence reconstruction from RNA-Seq

- using the Trinity platform for reference generation and analysis. **Nature protocols**, v. 8, n.1494, 2013.
- HESS, B.; BEKKER, H.; BERENDSEN, H. J.; FRAAIJE, J. G. LINCS: a linear constraint solver for molecular simulations. **Journal of computational chemistry**, v. 18, n. 12, p. 1463-1472, 1997.
- HOCKNEY, R. W. Methods in Computational Physics. **Alder**, p. 136-211, 1970.
- JÚNIOR, G. N. A. et al. Estresse hídrico em plantas forrageiras: Uma revisão. **PUBVET**, v. 13, p. 148, 2018.
- KUNERT, K. J. et al. Drought stress responses in soybean roots and nodules. **Frontiers in Plant Science**, v. 7, n. 1015, 2016.
- LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of applied crystallography**, v. 26, n. 2, p. 283-291.
- MARSH, J. F, et al. Medicago truncatula NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. **Plant physiology**, v. 144, n. 1, p. 324–335, 2007.
- MATOS, M. K. S. Expressão Diferencial e Diversidade de Fatores de Transcrição da Família MYB em Feijão-Caupi. Tese (Doutorado em Genética) – Centro de Ciências Biológicas, Universidade Federal de Pernambuco. Recife, p. 137, 2015.
- NIVÓN, L. G.; MORETTI, R.; BAKER, D. A Pareto-optimal refinement method for protein design scaffolds. **PloS one**, v. 8, n.4, e59004, 2013.
- OOSTENBRINK, C.; SOARES, T. A.; VAN DER VEGT, N. F; VAN GUNSTEREN, W. F. Validation of the 53A6 GROMOS force field. **European Biophysics Journal**, v. 34, n. 4, p. 273-284, 2005.
- PFÄFFL, M. W.; HORGAN, G. W.; DEMPFFLE, L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. **Nucleic Acids Research**, v.30, 2002.
- RIECHMAN, J. L. et al. Fatores de transcrição de Arabidopsis: análise comparativa à escala do genoma entre eucariotas: análise comparativa à escala do genoma entre eucariotas. **Sei**, v. 290, n. 2, p.105-110, 2000.
- SAFFER, A. M. Expansão de papéis para pectinas no desenvolvimento de plantas. **Journal of Integrative Plant Biology**, v. 60, p. 910–923, 2018.
- SANTOS, M. E. P. Evolução por duplicação gênica: estrutura e função da classe FoxP. Dissertação (Mestrado em Biologia Evolutiva e do Desenvolvimento) – Departamento de Biologia Animal, Universidade de Lisboa, Lisboa, p. 73, 2007.

- SCHAUSER, L. et al. A plant regulator controlling development of symbiotic root nodules. **Nature**, v. 402, n. 6758, p. 191-195, 1999.
- SINGH, S.; KATZER, K.; LAMBERT, J.; CERRI, M.; PARNISKE, M. CYCLOPS, um ativador de transcrição de ligação a DNA, orquestra o desenvolvimento de nódulos radiculares simbióticos. **Hospedeiro Celular e Micróbio**, v. 15, n. 2, p. 127-129, 2014.
- SONG, Y. et al. High-resolution comparative modeling with RosettaCM. **Structure**, v. 21, n. 10, p. 1735-1742, 2013.
- SOYANO, T.; HIRAKAWA, H.; SATO, S.; HAYASHI, M.; KAWAGUCHI, M. Nodule Inception creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. **Proceedings of the National Academy of Sciences**, v.111, p. 14607-14612, 2014.
- SOYANO, T.; KOUCHI, H.; HIROTA, A.; HAYASHI, M. Nodule inception Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in *Lotus japonicus*. **PLOS Genetics**, v. 9, p. e1003352, 2013.
- VAN, R. P. et al. Expression of early nodulin genes in alfalfa mycorrhizae indicates that signal transduction pathways used in forming arbuscular mycorrhizae and Rhizobium-induced nodules may be conserved. **Proceedings of the National Academy of Sciences**, v. 94, p. 253-265, 1997.
- VIEIRA-MELLO, G. S.; SANTOS, P. B.; SOARES-CAVALCANTI, N. M.; BENKO-ISEPPON, A. M. Identification and Expression of Early Nodulin in Sugarcane Transcriptome Revealed by in Silico Analysis. **Comput Intel Meth Bioinfo Biostat**, v. 6685, p. 72–85, 2011.
- YANO, K. et al. CYCLOPS, a mediator of symbiotic intracellular accommodation. **Proceedings of the National Academy of Sciences**, v. 105, p. 20540-20545, 2008.
- YORO, E, et al. A positive regulator of nodule organogenesis, nodule inception, acts as a negative regulator of rhizobial infection in *Lotus japonicus*. **Plant Physiology**, v. 165, n. 2, p. 747–758, 2014.

ANEXO

1) Normas da ABNT 2019. Disponível em: <http://www.abnt.org.br/>.

Material suplementar

A) Tratamentos empregados nos ensaios de transcriptômica, fisiologia e proteômica de *S. scabra* sob déficit hídrico – Embrapa Semiárido, Petrolina, PE

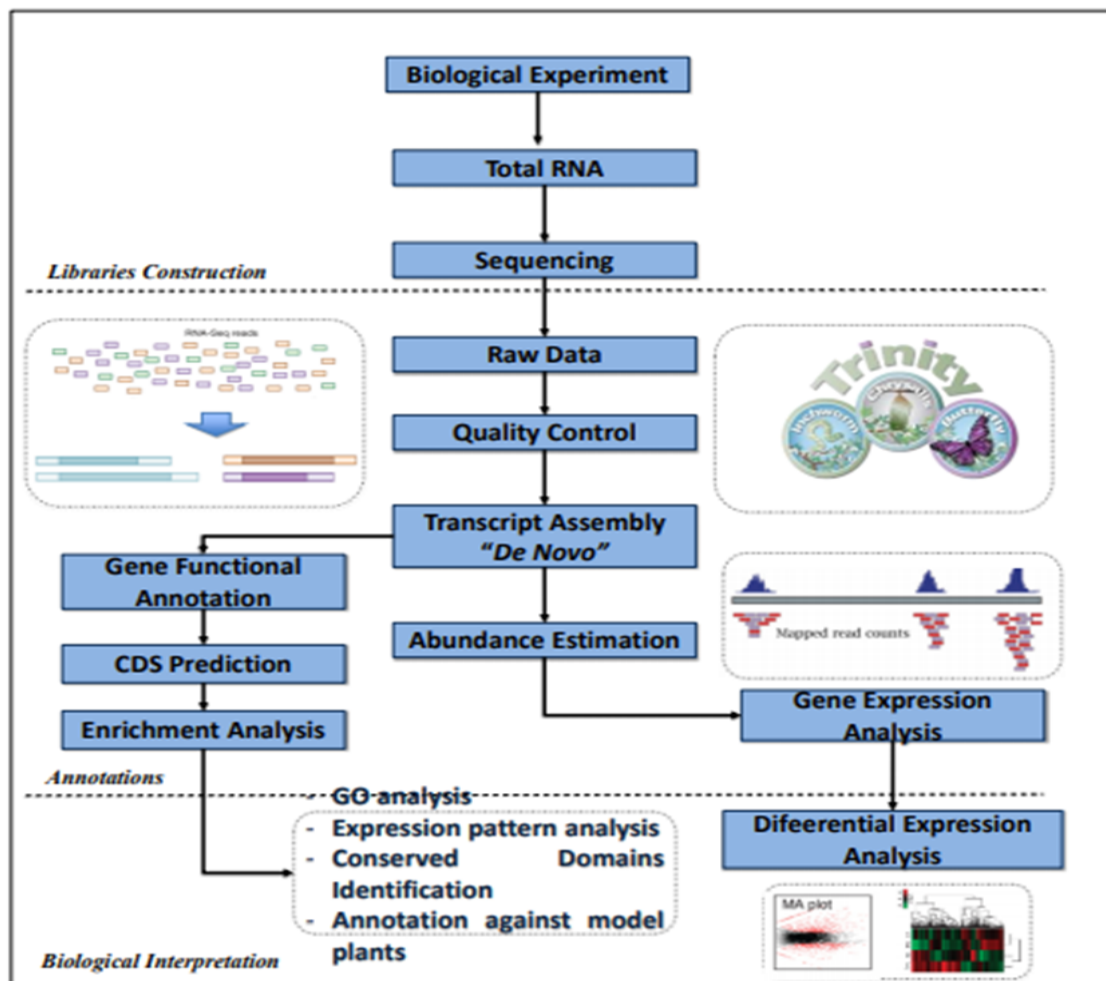
TRATAMENTOS		TRANSCRIPTÔMICA		FISIOLOGIA	PROTEÔMICA
		Vasos (Plantas)	RB	Vasos (Plantas)	Vasos (Plantas)
T1 (6 h DH)	<i>Controle</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
	<i>Tratamento</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
T2 (24 h DH)	<i>Controle</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
	<i>Tratamento</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
T3 (48 h DH)	<i>Controle</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
	<i>Tratamento</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
T4 (6 h após RI)	<i>Controle</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
	<i>Tratamento</i>	07 (21)	RB1	03 (09)	09 (27)
		07 (21)	RB2		
		07 (21)	RB3		
TOTAL		168 (504)		24 (72)	72 (216)

*DH: Déficit hídrico por suspensão de rega; RI: Reidratação; RB, Réplicas biológicas. Consideradas 3 plantas por vaso.



Fonte: Laboratório de Genética e Biotecnologia Vegetal (LGBV).

B) Fluxograma com as etapas de triagem e montagem das sequências RNA-Seq.



Fonte: Laboratório de Genética e Biotecnologia Vegetal (LGBV).