

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE BIOCÊNCIAS  
BACHAREL EM BIOTECNOLOGIA MOLECULAR**

**EDGAR LUIS WASCHBURGER**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DOS FATORES DE TRANSCRIÇÃO DO  
DE PITANGUEIRA (*Eugenia uniflora* L.)**

Porto Alegre

2019

**EDGAR LUIS WASCHBURGER**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DOS FATORES DE TRANSCRIÇÃO DO  
DE PITANGUEIRA (*Eugenia uniflora L.*)**

Trabalho de conclusão de curso apresentado como requisito parcial para obtenção do título de Bacharel em Biotecnologia com ênfase em Biotecnologia Molecular na Universidade Federal do Rio Grande do Sul.

Orientadora: Profa. Dra. Andreia Carina Turchetto Zolet

Co-orientadora: Ma. Drielli Canal

Porto Alegre

2019

### CIP - Catalogação na Publicação

Waschburger, Edgar Luis  
Identificação e caracterização dos fatores de  
transcrição DOF de pitangueira (*Eugenia uniflora* L.) /  
Edgar Luis Waschburger. -- 2019.  
41 f.  
Orientadora: Andreia Carina Turchetto-Zolet.

Coorientadora: Drielli Canal.

Trabalho de conclusão de curso (Graduação) --  
Universidade Federal do Rio Grande do Sul, Instituto  
de Biociências, Curso de Biotecnologia: Biotecnologia  
Molecular, Porto Alegre, BR-RS, 2019.

1. Filogenia. 2. Pitangueira. 3. Fatores de  
Transcrição. 4. DOF. 5. in silico. I. Turchetto-Zolet,  
Andreia Carina, orient. II. Canal, Drielli, coorient.  
III. Título.

## RESUMO

A pitangueira (*Eugenia uniflora* L.) é uma planta frutífera da família *Myrtaceae*, nativa do Brasil e apresenta ampla distribuição ao longo do território nacional. Suas populações estão adaptadas a condições climáticas contrastantes, como as condições secas e solo arenoso da restinga, bem como a mata fechada e ombrófila da mata ciliar. Já caracterizados no processo de adaptação e desenvolvimento, os fatores de transcrição DOF são proteínas reguladoras da expressão gênica do tipo dedo de zinco específicas de plantas. Essas características os tornam excelentes candidatos para o estudo da adaptação das populações de *E. uniflora*, visto que provavelmente contribuem para isso.

Neste trabalho, foram identificados um total de 17 genes DOF no transcrito de *E. uniflora*, com dois deles possuindo o domínio incompleto e sendo retirados de análises subsequentes. A partir da construção das árvores filogenéticas de aminoácidos e nucleotídeos, as sequências DOF se agruparam em 11 grupos funcionais distintos, possibilitando a predição funcional de 10 genes de *E. uniflora*. O resultado da análise de motivos conservados corroborou com os das análises evolutivas e a análise de pressão seletiva confirmou a ampla seleção purificadora atuante no domínio DOF.

Em conclusão, esse estudo possibilitou a identificação de genes DOF em pitangueira ainda não identificados. Enquanto isso, as análises evolutivas traçaram um rumo para estudos funcionais futuros a partir da predição de grupos funcionais conservados, além de elucidar questões ainda não respondidas.

## ABSTRACT

The Brazilian-cherry tree (*Eugenia uniflora* L.) is part of the *Myrtaceae* family of plants. It's a native plant with wide distribution among the Brazilian territory where its different populations are adapted to contrasting climatic conditions such as the salty and dry soil of the restinga, as well as the ombrophilous Atlantic forest. Thoroughly characterized in many development and adaptation processes, DOF transcription factors are plant specific regulatory proteins that have a zinc finger like domain. These characteristics make them well suited candidates for a study on the adaptation of the different Brazilian-cherry tree populations.

In this work, a total of 17 genes were identified in *E. uniflora*'s transcriptome, two of which had partial DOF domains and were excluded from further analysis. Upon construction of the amino acid and nucleotide phylogenetic trees DOF sequences were grouped in 11 distinct functional groups allowing a functional prediction of 10 of the 17 *E. uniflora*'s genes. Conserved motifs analysis results corroborate with the evolutionary analysis made, while the selective pressure analysis confirms a purifying selective pressure.

In conclusion, this study allowed the identification of novel DOF genes in the *E. uniflora* species. The evolutionary studies performed here paved a way for future functional studies through the prediction of conserved functional groups and elucidating previous unanswered questions.

## LISTA DE FIGURAS

<b>Figura 1.</b> Variação fenotípica entre pitangueiras .....	2
<b>Figura 2.</b> Representação da estrutura primária do domínio .....	3
<b>Figura 3.</b> Logo do domínio DOF .....	15
<b>Figura 4.</b> Árvore filogenética de aminoácidos.....	17
<b>Figura 5.</b> Árvore filogenética de nucleotídeos.....	18
<b>Figura 6.</b> Identificação dos motivos conservados nas sequências DOF.....	24
<b>Figura 7.</b> Representação dos motivos nas sequências DOF de Eugenia uniflora....	26

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>1</b>
1.1 <i>EUGENIA UNIFLORA</i> L. COMO ESPÉCIE DE ESTUDO.....	1
1.2 A FAMÍLIA DE FATORES DE TRANSCRIÇÃO <i>DOF</i> .....	2
1.2.1 <i>Participação de DOF na regulação da germinação</i> .....	3
1.2.2 <i>Participação de DOF no desenvolvimento vegetal</i> .....	5
1.2.3 <i>Participação de DOFs na elaboração de tecidos vasculares</i> .....	7
1.2.4 <i>Participação de DOFs na fisiologia de estômatos</i> .....	8
1.2.5 <i>Participação de DOFs na produção de metabólitos secundários</i> .....	9
1.2.6 <i>Participação dos DOFs na regulação do florescimento</i> .....	10
<b>2. OBJETIVOS .....</b>	<b>11</b>
2.1 OBJETIVO GERAL .....	11
2.2 OBJETIVOS ESPECÍFICOS.....	11
<b>3. MATERIAIS E MÉTODOS .....</b>	<b>11</b>
3.1 IDENTIFICAÇÃO DE SEQUÊNCIAS <i>DOF</i> .....	11
3.2 PROPRIEDADES PROTEICAS, MOTIVOS CONSERVADOS E LOGO .....	12
3.3 ALINHAMENTO E ANÁLISE FILOGENÉTICA .....	12
3.4 ANÁLISE DE PRESSÃO SELETIVA.....	13
<b>4. RESULTADOS E DISCUSSÃO.....</b>	<b>13</b>
4.1 IDENTIFICAÇÃO DOS GENES <i>DOF</i> EM <i>EUGENIA UNIFLORA</i> E CONSTRUÇÃO DO LOGO .....	13
4.2 ANÁLISE FILOGENÉTICA .....	15
4.3 ANÁLISE DE MOTIVOS CONSERVADOS E DOMÍNIOS TRANSMEMBRANA .....	23
4.4 ANÁLISE DE PRESSÃO SELETIVA.....	26
<b>5. CONCLUSÕES E PERSPECTIVAS .....</b>	<b>27</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>28</b>

## 1. INTRODUÇÃO

### 1.1 *Eugenia uniflora* L. como espécie de estudo

*Eugenia uniflora* L., conhecida como “pitangueira”, é uma espécie de frutífera nativa da Mata Atlântica, não endêmica e pertencente à família Myrtaceae. Também fazem parte dessa família uma diversidade de outras espécies frutíferas nativas, como a Cereja-do Rio Grande (*E. involucrata*), uvaia (*E. pyriformis*) jabuticada (*Plinia cauliflora*), araçá (*Psidium cattleianum*), guabiroba (*Campomanesia xanthocarpa*), dentre outras. Além das frutíferas, que apresentam grande potencial econômico, a espécie arbórea *Eucalyptus grandis* (Eucalipto), também faz parte dessa família. *E. uniflora* possui interesse científico e alto potencial para exploração comercial pela atividade antibacteriana de seus metabólitos secundários e extratos brutos (FALCÃO et al., 2018), possuindo ainda atividades antifúngicas (SILVA-ROCHA et al., 2015), anti-helmínticas (SANTOS et al., 2012), anti-inflamatórias (RATTMANN et al., 2012), antioxidantes, analgésicas e espaço para a sua utilização no tratamento de doenças crônicas, como diabetes (SOBEH et al., 2019). Sua distribuição abrange desde a região costeira do nordeste até a região sul do Brasil, leste do Paraguai e norte da Argentina e Uruguai (OLIVEIRA-FILHO & FONTES, 2000)(SCARANO, 2002)(ROESCH et al., 2009). Possui uma grande área de dispersão, e está adaptada a diferentes ecoregiões com características contrastantes. Apresenta porte arbustivo na restinga brasileira, localidade costeira de terreno arenoso e salino, forte insolação e ventos fortes, enquanto possui porte arbóreo na mata atlântica no interior da região sul, caracterizada por vegetação fechada, húmida e ombrófila (Figura 1). Essa disparidade entre populações também foi confirmada a nível genético, onde indivíduos da restinga possuem uma variação genética muito menor, quando comparada com indivíduos da mata atlântica no sul da distribuição (mata ciliar e estepes no Bioma Pampa), possivelmente resultado de uma expansão populacional recente de plantas adaptadas às condições da restinga (TURCHETTO-ZOLET et al., 2016).

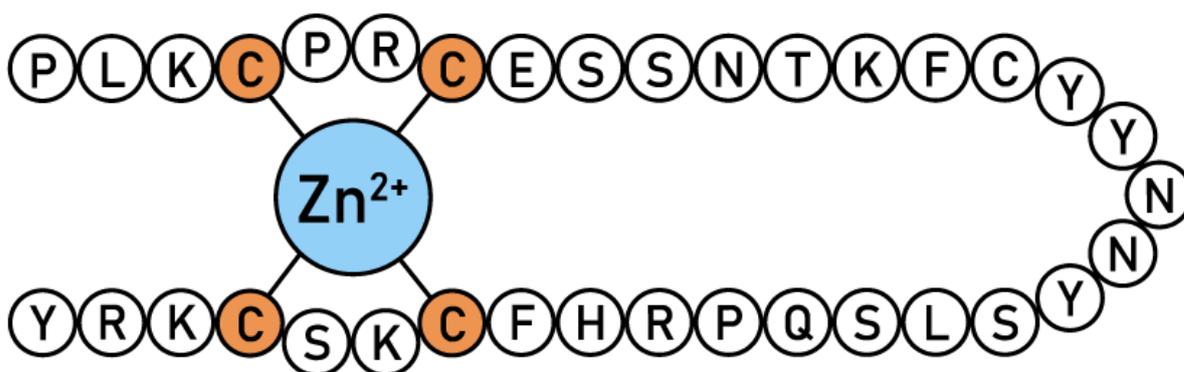


**Figura 1. Variação fenotípica entre pitangueiras.** Plantas de pitangueira presentes na ecoregião da restinga apresentam porte arbustivo (esquerda), indicadas por setas, enquanto plantas na mata ciliar apresentam porte arbóreo (direita).

## 1.2 A família de fatores de transcrição DOF

A produção de compostos secundários e o desenvolvimento da planta são mecanismos importantes para a adaptação e manutenção do organismo, entretanto esses mecanismos são dispendiosos de energia e necessitam ser precisamente regulados. Um dos principais mecanismos de regulação de processos fisiológicos é a nível transcricional, onde os principais atuantes são os fatores de transcrição. Muitos fatores de transcrição já foram caracterizados como atuantes em processos fisiológicos de desenvolvimento e adaptação (HICHRI et al., 2014; LASHBROOKE et al., 2015; VIMOLMANGKANG et al., 2013). Em meio a tantos fatores de transcrição caracterizados, está a família de fatores de transcrição ligantes a DNA com um dedo (do inglês *DNA-binding with one finger*, ou apenas DOF). É exclusiva do clado *Viridiplantae* e foram identificadas em uma ampla gama de espécies, incluindo monocotiledôneas, como o arroz, milho e sorgo (CHEN; CAO, 2014; KUSHWAHA et al., 2011; LIJAVETZKY; CARBONERO; VICENTE-CARBAJOSA, 2003), eudicotiledôneas, tais como *Arabidopsis*, algodão e tomate (CAI et al., 2013; LI et al., 2018; YANAGISAWA, 2002), licofitas, como *Selaginella moellendorffii* (MORENO-RISUENO et al., 2007), musgos como *Physcomitrella patens* (SHIGYO et al., 2007)

e algas, tais como *Chlorella vulgaris* e *Chlamydomonas reinhardtii* (MORENO-RISUENO et al., 2007; TOKUNAGA et al., 2019). Possuem um número altamente variável de representantes, desde 1 gene em *Chlamydomonas reinhardtii* até mais de 40 genes em espécies poliploides, como *Zea mays* (MORENO-RISUENO et al., 2007; CHEN; CAO, 2015). Essa diversidade é resultado de múltiplos eventos de duplicação a nível genético, assim como genômico (MORENO-RISUENO et al., 2007). Suas sequências variam desde 200 até 400 aminoácidos e possuem um domínio de ligação ao DNA do tipo dedo de zinco CX<sub>2</sub>CX<sub>2</sub>1CX<sub>2</sub>C (Figura 2) (UMEMURA et al., 2004), específico às sequências 5'-(A/T)AAAG-3'(YANAGISAWA; SCHMIDT, 1999), assim como pelo menos um domínio para interação com outras proteínas. Tanto estudos *in silico* quanto *in natura* demonstraram a capacidade das proteínas DOF de serem expressas em tecidos vasculares, principalmente no floema (HIR; BELLINI, 2013). Mesmo tendo sido caracterizadas em muitos outros papéis fisiológicos além do desenvolvimento vascular, a manutenção da sua expressão em tecidos vasculares pode ser crítica para a atuação dos genes DOF.



**Figura 2. Representação da estrutura primária do domínio.** Os resíduos de C ligantes a zinco estão coloridos de laranja, enquanto o átomo de zinco está representado em azul claro. Linhas pretas indicam as ligações de hidrogênio.

### 1.2.1 Participação de DOF na regulação da germinação

O processo de germinação é regulado pelos níveis fisiológicos de hormônios como ácido giberélico (GA) e ácido abscísico (ABA), onde maiores níveis de GA induzem a germinação, enquanto maiores níveis de ABA induzem a dormência da semente. Sob as condições certas, a incidência de luz no espectro do vermelho-distante contribui para o processo de germinação. A percepção da luz no comprimento vermelho e vermelho-distante pelas células vegetais é mediada por

proteínas denominadas fitocromos. A absorção de luz vermelha por fitocromos muda seu estado conformacional, convertendo-os do estado *Phytochrome red* (Pr) para o estado *Phytochrome far-red* (Pfr), possibilitando a sua translocação para o núcleo e interação com fatores de transcrição. Alguns fitocromos apresentam cinéticas de degradação proteolítica diferentes quando comparados. Por exemplo, o fitocromo A (PhyA) é degradado rapidamente quando convertido ao seu estado Pfr, enquanto o fitocromo B (PhyB) permanece no estado Pfr por mais tempo, tornando-os adaptados a orquestrar vias de sinalização diferentes baseadas no equilíbrio entre suas formas. Os genes *AthDOF2.5* (*DAG2*) e *AthDOF3.7* (*DAG1*), foram caracterizados em mutantes que possuíam tempos de germinação alterados. Mutantes para *DAG1* apresentam uma menor necessidade de incidência de luz vermelha para a germinação, enquanto mutantes para *DAG2* demonstram o oposto (GUALBERTI et al., 2002). Um dos fatores de transcrição que interage com PhyB é *Phytochrome Interacting Factor3-Like 5* (PIL5), responsável por inibir o processo de germinação. PhyB induz a degradação de PIL5 via ubiquitinação e encaminhamento para o proteassomo 26S. PIL5, por sua vez, induz a expressão de fatores de transcrição da família DELLA, como *GA Insensitive* (GAI) e *Repressor of ga1-3* (RGA), inibidores da via de síntese de GA e essenciais para a manutenção dos níveis endógenos de ABA (OH et al., 2007)(PISKUREWICZ et al., 2009). Quanto ao gene *DAG1*, foi mostrado que seus níveis estão relacionados com a expressão de PIL5, sugerindo uma indução de *DAG1* por PIL5 (GABRIELE et al., 2010). Sabe-se que é necessária a interação entre *DAG1* e GAI para a repressão de *AtGA3ox1*, gene da via de biossíntese de GA (BOCCACCINI et al., 2014). *DAG1* também revelou inibir a expressão de *CYP707A2*, gene da via catabólica de ABA, demonstrando semelhança com as funções das proteínas DELLA (BOCCACCINI et al., 2016). Em contrapartida, *DAG2* apresenta-se inibido pela expressão de PIL5, assim como inibição direta por *DAG1*, comprovada por imunoprecipitação de cromatina acoplada à PCR quantitativa (ChIP-qPCR). Quanto aos genes regulados por *DAG2*, RGA aparenta estar negativamente regulado, enquanto os genes catabólitos de GA: *GA3ox1* e *GA3ox2* demonstram-se induzidos, entretanto não se sabe se essa interação é direta ou não (SANTOPOLO et al., 2015).

Outro gene envolvido na regulação da germinação é *AthDOF3.2* (*DOF6*). *DOF6* é principalmente expresso em sementes de *A. thaliana*, postergando o período de germinação, mas possuindo pouco efeito após o amadurecimento da

semente. Sua superexpressão é claramente evidenciada pelo fenótipo de inibição do crescimento da planta (RUEDA-ROMERO et al., 2012). Tal fenótipo é dado pela expressão de genes relacionados com a biossíntese de ABA, como o gene *ABA1*. Entretanto, não se sabe se a interação entre a região promotora de *ABA1* e *DOF6* é direta, ou se é necessária a dimerização com outra proteína para a devida interação. Um ligante de *DOF6* é RGA-Like 2 (*RGL2*), membro da família de fatores de transcrição DELLA também responsável pela síntese de ABA e regulação da germinação (RAVINDRAN et al., 2017). O complexo *DOF6-RGL2* demonstrou-se necessário para a regulação de *GATA12*, também atuante na dormência da semente. Em geral, os genes *DAG1* e *DAG2* atuam na inibição da germinação pela catálise de GA e inibição da catálise de ABA, enquanto *DOF6* atua na mesma função, mas, complementarmente, pela síntese de ABA.

### 1.2.2 Participação de DOF no desenvolvimento vegetal

Após a germinação, a plântula é muito sensível a fatores externos, precisando responder a múltiplos estímulos para se desenvolver normalmente e atingir seu estágio de reprodução. Dois genes DOF responsáveis pelos primeiros estágios de desenvolvimento da planta são os genes *AthDOF3.4* e *AthDOF3.6* ou *Ocs Binding Element Protein 1* e *3* (*OBP1* e *OBP3*). Assim como outras proteínas OBP, *OBP1* e *OBP3* são fatores de transcrição responsivos a elementos ocs, sequências presentes na região promotora de genes relacionados a estresse oxidativo. Proteínas OBP, por serem fatores de transcrição DOF, reconhecem sequências que não são ocs, por isso necessitam da heterodimerização com outros fatores de transcrição específicos para a sua devida resposta. Ambos os genes demonstram ser responsivos aos níveis hormonais de auxina e ácido salicílico (KANG; SINGH, 2000). Estudos funcionais demonstraram que plantas superexpressando *OBP1* apresentam um maior número de células, mas com tamanho reduzido. Um experimento de CHIP-qPCR demonstrou que *OBP1* atua na divisão celular pela regulação de *CYCD3;3* e da proteína DOF *CDF4*, enquanto uma análise de microarranjo indicou que o fenótipo de células curtas pode ser explicado pela inibição da expressão de enzimas responsáveis pela expansão da parede celular (SKIRYCZ et al., 2008). *OBP1* também demonstrou ser responsivo a estresses

oxidativos ocasionados pela presença de espécies reativas de oxigênio (ROS), aumentando a expressão de GST6, gene da família das glutathione S-transferases (CHEN; CHAO; SINGH, 1996). A superexpressão de *OBP3*, por sua vez, demonstrou um retardo no crescimento vegetal e redução da viabilidade de sementes (KANG; SINGH, 2000). Um estudo seguinte revelou que plantas transgênicas *knockdown* para *OBP3* possuíam deficiências na via de sinalização por PhyB, responsivo à luz vermelha/vermelha-distante, para a inibição do alongamento do hipocótilo (WARD et al., 2005). Outra via alterada foi a via de desenvolvimento do cotilédone mediada por criptocromo 1, responsivo à luz azul (WARD et al., 2005). Sabendo disso, *OBP3* aparenta atuar como um mediador das vias fotorreceptoras e hormonais para o desenvolvimento vegetal. Por último, *AthDOF5.4* (*OBP4*) está envolvido na inibição da progressão do ciclo celular e expansão da parede celular (XU et al., 2016). Apresenta expressão em vários tecidos da planta, entretanto seu maior caso de estudo é a expressão na raiz e a regulação do desenvolvimento da mesma e de raízes laterais em resposta a diversos estresses (RAMIREZ-PARRA et al., 2017; RYMEN et al., 2017; XU; CAI, 2019).

Além dos genes *OBP*, os níveis de expressão de *AthDOF2.3* ou *CYCLING DOF FACTOR 4* (*CDF4*), ligante de *OBP1*, foram encontrados associados à fase S do ciclo celular, sugerindo uma atuação de *CDF4* na divisão celular, entretanto nenhum estudo funcional mais aprofundado foi feito (SKIRYCZ et al., 2008). Evidências filogenéticas dos genes *DOF* mostram *CDF4* e *AthDOF1.5* ou *COGWHEEL 1* (*COG1*) como parálogos próximos e especula-se que possam ter funcionalidades semelhantes. A expressão de *COG1* demonstrou-se como supressora das vias de resposta por *phyA* e *phyB* para incidência de luz no espectro do vermelho e vermelho-distante (PARK et al., 2003). Outro supressor da via de resposta por fitocromo é a proteína *PHYTOCHROME INTERACTING FACTOR 4* (*PIF4*), responsável pela regulação da síntese de brassinosteróides. Ensaios de ChIP-RT-qPCR e EMSA mostraram que *COG1* regula diretamente a expressão de *PIF4* ao interagir com sua região promotora (ZHUOYUN et al., 2017). Os níveis de brassinosteróides estão sabidamente envolvidos com o crescimento da planta (NOLAN; CHEN; YIN, 2018), logo a regulação de *PIF4* explicaria o fenótipo de inibição do crescimento do hipocótilo exposto à luz vermelha e vermelho-distante observado por mutantes caracterizados previamente (PARK et al., 2003).

### 1.2.3 Participação de DOFs na elaboração de tecidos vasculares

Diferentes hormônios regem o planejamento e construção dos tecidos vasculares. Auxinas difundem-se ao longo do comprimento do protoxilema/xilema e regulam a expressão gênica por sua concentração hormonal. Contrariamente, citocininas apresentam atividades mais localizadas, atuando no desenvolvimento do protofloema/floema e suas células vizinhas. Muitos genes são responsivos à expressão de auxina, mas algumas proteínas atuam em funções-chave na transdução deste sinal. *MONOPTEROS (MP)*, também conhecido como *AUXIN RESPONSIVE FACTOR 5 (ARF5)*, é uma proteína *hub* responsável por orquestrar a formação de tecidos vasculares principalmente na raiz. MP possui genes alvo, nomeados de *TARGET OF MONOPTEROS (TMO)*. Entre eles, *AthDOF5.3* ou *TMO6* é regulado diretamente por MP (SCHLERETH et al., 2010). Sua expressão mostrou-se envolvida no desenvolvimento e regeneração vascular, assim como em processos de enxertia (MELNYK et al., 2018). Outros de seus alvos são *TMO5* e *LONESOME HIGHWAY (LHW)*, fatores de transcrição reguladores da expressão de genes, como *LONELY GUY 3 (LOG3)* e *LONELY GUY 4 (LOG4)*, encarregados da ativação de citocininas na raiz. *AthDOF2.1* mostrou-se regulada pelos níveis endógenos de citocinina de maneira dependente da via de biossíntese *TMO5/LHW*, indicando que sua função ocorre *downstream* da via de desenvolvimento primário da raiz mediada por MP (SMET et al., 2019). Além de alvos de MP, o fator de transcrição DOF *AthDOF5.8* atua por interação direta com MP. Ao contrário de casos anteriores, a atuação de *AthDOF5.8* aparenta estar na parte aérea da planta, onde a sua superexpressão acarretou a má formação de nervuras nas folhas (KONISHI et al., 2015; KONISHI; YANAGISAWA, 2015).

Outros genes *DOF* induzidos por citocininas são *AthDOF2.4* e *AthDOF5.1* ou *PHLOEM EARLY DOF 1 e 2 (PEAR1 e 2)*. Seus maiores níveis de expressão ocorrem nos elementos de tubo crivado do protofloema e, a partir da sua difusão por células adjacentes, promovem divisões celulares periclinais na raiz. Ao chegar nas células do procâmbio interno e metaxilema, onde há prevalência de auxinas, sua expressão é inibida pela atuação de fatores de transcrição HD-ZIP III. A expressão dos genes *PEAR* também mostrou induzir os níveis transcricionais de proteínas HD-ZIP III, promovendo um *loop* de *feedback* negativo para o controle da atuação das proteínas *PEAR* (MIYASHIMA et al., 2019). Além do desenvolvimento da vasculatura

da raiz, linhagens superexpressando o gene *AthDOF5.6* ou *HIGH CAMBIAL ACTIVITY 2 (HCA2)*, apresentam má formação do câmbio interfascicular em todos os tecidos da planta, indicando uma atuação mais generalizada desse fator de transcrição (GUO et al., 2009). Sua ação é provavelmente dada em etapas iniciais do desenvolvimento, a partir das células do parênquima, e fundamentais para a correta construção do tecido (GUO et al., 2009).

#### 1.2.4 Participação de DOFs na fisiologia de estômatos

Estômatos são estruturas vegetais presentes principalmente em folhas e responsáveis por controlar a troca gasosa de O<sub>2</sub>, CO<sub>2</sub> e vapor de água com a atmosfera, o último evento é conhecido como transpiração. Plantas com menores taxas de transpiração retêm água mais eficientemente, tornando a otimização de estômatos uma abordagem para a concepção de plantas resistentes à seca (LAPORTE et al., 2002). A formação de estômatos é estabelecida pela divisão sucessiva e maturação de linhagens celulares indiferenciadas presentes nas folhas. As primeiras células que apresentam a forma circular, característica de estômatos, são as células-guarda mães, que, em seguida, se dividem simetricamente em duas células-guarda, formando o estômato. A principal via de desenvolvimento estomatal envolve a expressão sequencial de SPEACHLESS (SPCH), MUTE e FAMA, fatores de transcrição da família bHLH, assim como *loops* de regulação negativa, envolvendo os genes regulados por SPCH, como *EPF1/2*, *TMM* e o próprio *SPCH*. Em *Arabidopsis thaliana*, foi demonstrado que um de seus genes DOF está envolvido com a fisiologia estomatal: *AthDOF5.7* nomeado *STOMATAL CARPENTER 1 (SCAP1)* (NEGI et al., 2013). A superexpressão de SCAP1 leva ao aumento do número de estômatos por área, assim como a inibição dos níveis transcricionais de *EPF2*, *MUTE* e *TMM*, todos alvos de SPCH (CASTORINA et al., 2016). Um modelo para a atuação de SCAP1 seria como um competidor para genes alvos de SPCH, onde SCAP1 atua na sua inibição (CASTORINA et al., 2016). Mutantes nulos para SCAP1 demonstraram alterações na movimentação e morfogênese dos estômatos, evidenciando o papel de SCAP1 na função estomatal (NEGI et al., 2013). Um experimento de ChIP-qPCR revelou que sua atuação ocorre pela regulação de genes como *MYB60*, essencial para a movimentação dos

estômatos, e *GORK*, codificador de um canal exportador de íons de potássio (NEGI et al., 2013).

#### 1.2.5 Participação de DOFs na produção de metabólitos secundários

Metabólitos secundários são compostos orgânicos que não têm implicações diretas no crescimento e desenvolvimento da planta. Não obstante, a sua produção é importante para a adaptação do organismo. Alguns também possuem propriedades fisiológicas de interesse para a medicina humana, tornando a sua pesquisa uma área de atuação extremamente importante. Dentre os genes *DOF* relacionados com a produção de metabólitos secundários, o gene *AthDOF4.2* já foi caracterizado como regulador negativo do metabolismo de fenilpropanoides (SKIRYCZ et al., 2007), condizente com fenótipos de falta de mucilagem observados em plantas transgênicas superexpressando o gene (ZOU et al., 2013). Outro fenótipo presente é a inibição da bifurcação de ramos, compartilhada com seu parálogo *AthDOF4.4* (ZOU et al., 2013). Não se sabe ao certo como esse fenótipo é dado, entretanto artigos prévios já mostraram a interação das vias de produção de fenilpropanoides com a síntese de hormônios como auxina e citocinina, conectando as vias de biossíntese ao processo de ramificação (KUREPA et al., 2018).

Outro grupo de compostos secundários regulados por proteínas *DOF* é o grupo dos glicosinolatos, mais especificamente os indolglicosinolatos (IGS), sintetizados a partir de triptofano. IGS possuem propriedades tóxicas para animais, atuando na defesa da planta contra insetos e animais herbívoros. IGS também fazem parte da via de biossíntese de auxinas, como ácido indolacético (IAA), podendo atuar na regulação do crescimento vegetal e outros processos fisiológicos. *OBP2* foi identificada como reguladora positiva da biossíntese de IGS. *OBP2* apresenta expressão no tecido vascular, atuando na síntese de IGS a partir de danos físicos na vasculatura do tecido vegetal (SKIRYCZ et al., 2006). Assim como os níveis de IGS, a expressão de IAA também foi aumentada em plantas transgênicas superexpressando *OBP2*, entretanto há pouca evidência relacionando os níveis de auxina sintetizada a partir de IGS com outros processos fisiológicos (SKIRYCZ et al., 2006).

### 1.2.6 Participação dos DOFs na regulação do florescimento

O florescimento é um processo muito complexo que envolve diversos sinais externos, como a incidência de luz e sua duração, e internos, como o estágio de desenvolvimento e condições fisiológicas do organismo. A subfamília de genes *DOF CYCLING DOF FACTORS (CDF)* composta por *AthDOF1.3 (CDF6)*, *AthDOF1.10 (CDF5)*, *AthDOF2.3 (CDF4)*, *AthDOF3.3 (CDF3)*, *AthDOF5.2 (CDF2)* e *AthDOF5.5 (CDF1)*, é profundamente estudada e está amplamente relacionada com o florescimento. Mesmo nomeada como CDF, CDF4 apresenta COG1 como homólogo mais próximo, possivelmente atuando em uma função que seja semelhante a de COG1 (BUESO et al., 2016). Todas outras CDFs foram confirmados experimentalmente como ligantes a GIGANTEA (GI) (FORNARA et al., 2009; KRAHMER et al., 2019), uma proteína envolvida com o relógio circadiano. GI tem seu pico de expressão em torno de 12 horas após o amanhecer durante dias longos (DL, 16 horas de luz/8 horas de escuro) e 8 horas após o amanhecer em dias curtos (DC, 8 horas de luz/16 horas de luz) (RYNER et al., 1999). GI atua na regulação de supressores do florescimento principalmente pela interação com proteínas da família *F-box*, como ZEITLUPE (ZTL), FLAVIN-BINDING, KELCH REPEAT, F-BOX1 (FKF1) e LIGHT KELCH PROTEIN 2 (LPK2), que mediam a degradação de proteínas por ubiquitinação e encaminhamento para a via proteassomo 26S (DAVID et al., 2006). Apenas o gene *CDF6* não teve seu papel elucidado. As outras CDFs foram caracterizadas como supressoras do florescimento pela inibição da transcrição de *CONSTANS (CO)*, que por sua vez promove a transcrição de *FLOWERING LOCUS T (FT)* (FORNARA et al., 2009). As CDFs aparentam atuar pleiotropicamente e possivelmente possuem afinidades diferentes para cada proteína da família *F-box* (IMAIZUMI et al., 2005).

A hipótese do nosso trabalho é que o conjunto de funções dos fatores de transcrição DOF pode estar envolvido na adaptação das populações de *Eugenia uniflora* da restinga e mata ciliar para diferentes condições climáticas e outros fatores ecológicos. Portanto, identificar e caracterizar os genes DOF nessa espécie, assim como analisar possíveis forças de seleção atuantes em suas sequências por ferramentas *in silico* podem auxiliar na compreensão dos mecanismos de adaptação local na mesma.

## 2. OBJETIVOS

### 2.1 Objetivo Geral

Identificar e caracterizar a família de fatores de transcrição DOF em pitanga (*Eugenia uniflora* L.), assim como avaliar o seu possível envolvimento na adaptação de populações de pitangueira a diferentes condições climáticas.

### 2.2 Objetivos Específicos

1. Identificar, no transcrito de *E. uniflora*, sequências homólogas à família de fatores de transcrição DOF.
2. Realizar uma análise filogenética com as sequências identificadas em *E. uniflora* e sequências de outras espécies de plantas que apresentam esses genes já caracterizados para identificar os grupos ortólogos.
3. Caracterizar as sequências preditas das proteínas através da identificação de domínios e motivos conservados usando uma abordagem *in silico*.
4. Verificar o tipo de seleção agindo sobre as formas parálogas dos genes DOF de *E. uniflora*

## 3. MATERIAIS E MÉTODOS

### 3.1 Identificação de sequências DOF

O perfil HMM do domínio DOF (PF02701) foi baixado do banco de dados do Pfam (<http://www.pfam.xfam.org>) e utilizado para uma análise HMMER contra o transcrito de *E. uniflora* L. (GUZMAN et al., 2012), assim como transcritomas disponíveis no banco de dados do Phytozome v12.1.6 ([phytozome.jgi.doe.gov/pz/portal.html](http://phytozome.jgi.doe.gov/pz/portal.html)), como *Arabidopsis thaliana* (TAIR10), *Eucalyptus grandis* (v2.0), *Oryza sativa* (v7\_JGI), *Populus trichocarpa* (v3.0) e *Zea mays* (Ensembl-18). Foram aceitas, para análises seguintes, apenas as sequências

com *E-value* abaixo de 0,01. Os códigos de acesso das sequências foram utilizados para recuperar as sequências de nucleotídeos de seus respectivos genomas. As fases de leitura dos transcritos de *Eugenia uniflora* foram verificadas com a ferramenta ORFfinder ([www.ncbi.nlm.nih.gov/orffinder](http://www.ncbi.nlm.nih.gov/orffinder)) para a correta anotação.

### 3.2 Propriedades proteicas, motivos conservados e logo

A massa molecular (MW) e o ponto isoelétrico (pI) foram determinadas pela ferramenta ProtParam do ExPASy ([web.expasy.org/protparam](http://web.expasy.org/protparam)). A localização celular foi predita utilizando a ferramenta DeepLoc-1.0 ([www.cbs.dtu.dk/services/DeepLoc/](http://www.cbs.dtu.dk/services/DeepLoc/)). Para a predição de domínios conservados, a ferramenta MEME ([meme-suite.org](http://meme-suite.org)) (BAILEY; ELKAN, 1994) foi utilizada procurando por 15 motivos, com um comprimento entre 10 e 100 caracteres e qualquer número de repetições por sequência. Os genes DOF das seis espécies foram usados como entrada. A predição de domínios transmembranas foi realizada no programa Protter ([wlab.ethz.ch/protter/](http://wlab.ethz.ch/protter/)). A construção do logo do domínio DOF foi feita no program WebLogo ([weblogo.berkeley.edu/logo.cgi](http://weblogo.berkeley.edu/logo.cgi)) (CROOKS et al., 2004).

### 3.3 Alinhamento e análise filogenética

A análise filogenética foi realizada para identificar potenciais ortólogos e classificar os genes DOF identificados no transcrito de *E. uniflora*. Para essa análise foram utilizadas, além das sequências DOF identificadas em *E. uniflora*, sequências de espécies cujos genes DOF já foram estudados e caracterizados, tais como *Arabidopsis thaliana* (YANAGISAWA, 2002), *Eucaliptus grandis* (D'ALMEIDA; PASQUALI, 2014), *Oryza sativa* (LIJAVETZKY; CARBONERO; VICENTE-CARBAJOSA, 2003), *Zea may* (CHEN; CAO, 2014) e *Populus trichocarpa* (WANG et al., 2017). As sequências dessas espécies foram obtidas dos trabalhos publicados, bem como a partir de BLAST (*Protein Basic Local Align Sequence Tool*) contra o genoma das mesmas no banco de dados Phytozome. As sequências da região codificadora (CDS, *coding sequence*) dos fatores de transcrição DOF das espécies

citadas acima foram inseridas no programa MEGA7 (*Molecular Evolutionary Genetics Analysis*) (KUMAR; STECHER; TAMURA, 2016) (<https://www.megasoftware.net>) para visualização, edição e alinhamento. As sequências codificadoras foram traduzidas para suas respectivas sequências proteicas e alinhadas utilizando a ferramenta MUSCLE (Edgar, 2004), disponível no MEGA 7, com os valores e opções padrão. Por serem sequências com apenas o domínio DOF conservado, os aminoácidos não correspondentes ao domínio foram excluídos e gaps codificados como dado faltante usando o caractere "?". A partir do alinhamento, a determinação de qual o melhor modelo evolutivo foi realizada nos programas JModelTest2 e ProtTest3 (DARRIBA et al., 2015, 2017) para nucleotídeos e amino ácidos respectivamente. O programa Beast (SUCHARD et al., 2018) foi utilizado para a construção da árvore filogenética pelo método Bayesiano segundo o processo de *Birth-Death* por 30 milhões de gerações seguindo o modelo de substituição JTT+I+G para aminoácidos, enquanto a árvore filogenética de nucleotídeos foi gerada durante 50 milhões de gerações seguindo o modelo evolutivo de substituição GTR+G. O TreeAnnotator foi utilizado para excluir um décimo do total de árvores, correspondente às árvores mais discrepantes, construir a árvore consenso e calcular os valores de probabilidade posterior. O FigTree foi utilizado para visualizar a árvore gerada e editar as figuras. As ilustrações das figuras foram realizadas no programa Adobe Illustrator CC.

### 3.4 Análise de pressão seletiva

Análises de seleção foram realizadas com base nos alinhamentos e suas respectivas árvores filogenéticas como dados de entrada. As taxas de substituição não sinônimos para substituições sinônimas ( $dN / dS$ , ou  $\omega$ ) foram computados, e homogeneidade e seleção positiva foram determinados usando modelos de máxima verossimilhança no programa CODEML em PAML (v.4.9) (YANG, 2007).

## 4. RESULTADOS E DISCUSSÃO

### 4.1 Identificação dos genes *DOF* em *Eugenia uniflora* e construção do logo.

Utilizando o perfil HMM do domínio DOF (PF02701), foram encontradas 17 sequências no transcrito de *Eugenia uniflora*.

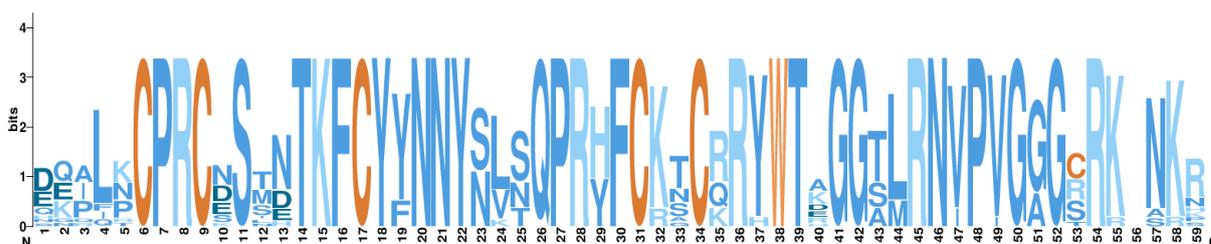
**Tabela 1. Características químicas e propriedades das sequências dos fatores de transcrição DOF recuperadas do transcrito de *E. uniflora*.** Tamanho proteico é dado em número de resíduos de cada proteína (AAs), peso molecular é dado em kiloDaltons (kDa) e ponto isoelétrico é dado em pH.

Gene Accession Number	Protein Length (AAs)	MW (kDa)	pI	Subcellular Location
g22988	109	12,012	9,69	Nucleus
g10632	113	12,966	9,51	Nucleus
g22140	186	20,727	9,62	Nucleus
g11471	242	25,704	9,40	Nucleus
g32847	213	22,901	9,94	Nucleus
g23589	109	11,846	8,59	Nucleus
g42095	297	31,379	9,38	Nucleus
g27855	303	33,576	8,45	Nucleus
g36695	292	31,619	8,84	Nucleus
g17481	332	35,575	9,34	Nucleus
g41142	360	38,370	9,07	Nucleus
g38411	308	32,957	8,71	Nucleus
g17107	307	33,047	8,79	Nucleus
g25930	331	36,445	6,52	Nucleus
g20579	467	50,378	5,23	Nucleus
g22699	507	55,049	6,10	Nucleus
g3653	554	59,573	8,90	Nucleus

Dentre as 17 sequências encontradas, *g17481* possuía códons de parada no meio de sua sequência. Sua fase de leitura foi verificada no programa ORFfinder para sua correta anotação. Após a correção da fase de leitura, sua sequência não possuía nenhuma similaridade com o domínio DOF e foi excluída de análises posteriores. A sequência *g23589* possuía apenas metade do domínio DOF e também foi excluída de análises posteriores. Dentre as 15 sequências restantes, a maioria possui em torno de 300 resíduos, sendo a menor delas *g22988*, com apenas 109 resíduos, e a maior sendo *g3653*, com 554 resíduos (Tabela 1). Podemos levantar uma hipótese de que as sequências *g22988* e *g10632* não estarem corretamente montadas, pois só o domínio DOF possui um comprimento de 60 resíduos, restando 50 resíduos para outros domínios, o que não seria impossível, entretanto, as sequências DOF possuem de 200 a 300 aminoácidos em geral. Seus pontos isoelétricos variam desde 5,23, em *g20579*, até 9,94, em *g32847*, com uma grande parcela variando de 8,40 até 9,94. Quando submetidas à predição da

localização subcelular, todas as sequências demonstraram localização nuclear, típico das outras sequências dos fatores de transcrição DOF.

A construção do logo do domínio DOF nos mostra a frequência dos resíduos de aminoácidos presentes em cada posição (Figura 3). Posições com prevalência de um resíduo podem significar um papel importante na sua função. Ao analisar o logo do domínio DOF das sequências de *E. uniflora*, a conservação dos quatro resíduos de cisteína, típicos do domínio DOF, já é esperada. A conservação do resíduo de triptofano também já foi reportada como importante para a função do domínio DOF (SHIMOFURUTANI et al., 1998). A conservação dos demais resíduos é dada em blocos, como as regiões 14 - 22 e 45 - 52. Outra característica importante é que grande parte dos resíduos conservados possuem propriedades básicas ou polares, possivelmente importantes para a correta interação com a molécula de DNA.

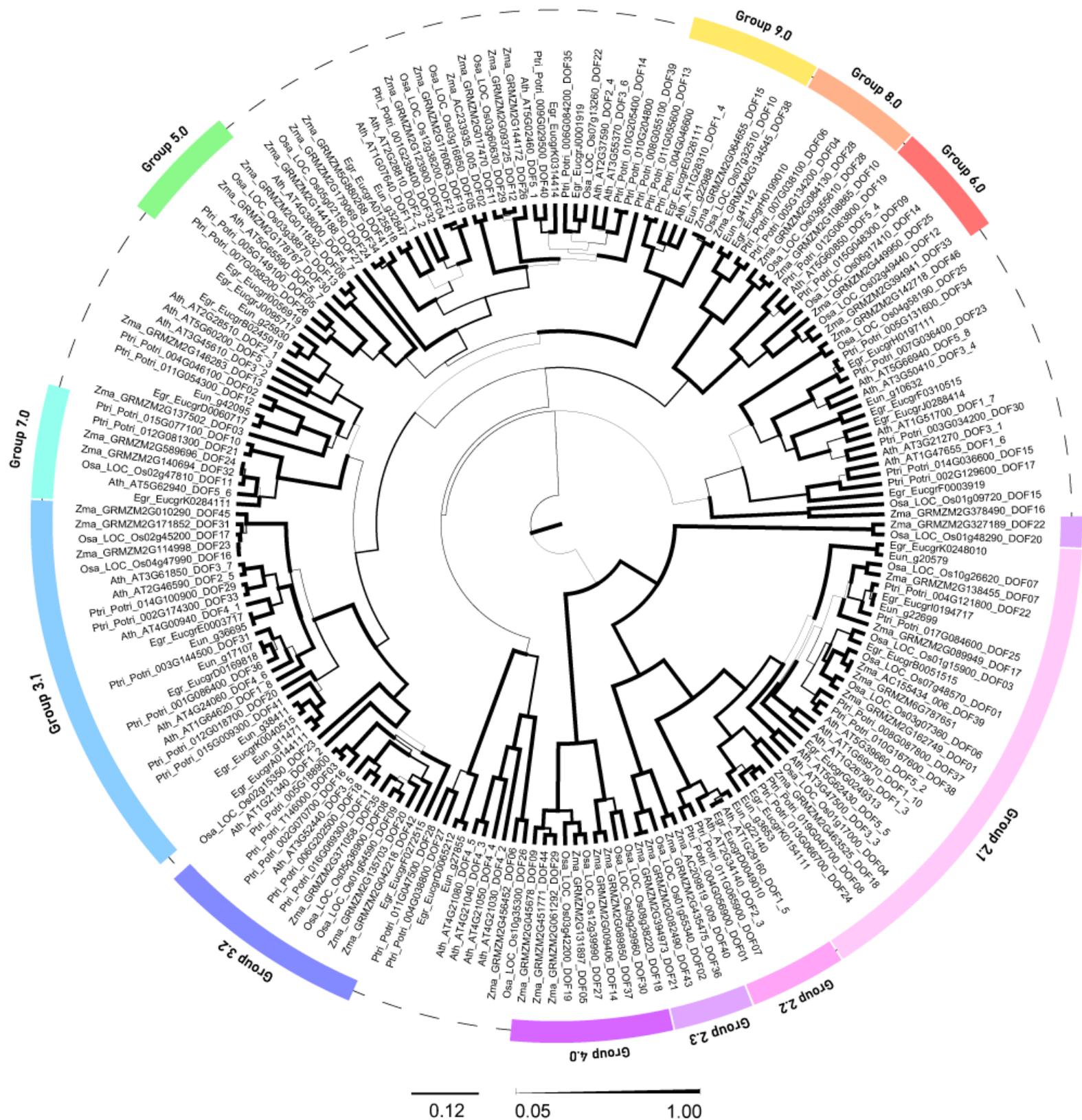


**Figura 3. Logo do domínio DOF.** Elaborado a partir dos 15 genes encontrados em *E. uniflora*. Os resíduos de cisteína (C) e triptofano (W), característicos do domínio DOF estão destacados em laranja. Resíduos ácidos estão em tom azul escuro, resíduos básicos estão em azul claro e todos os outros resíduos estão em tom azul.

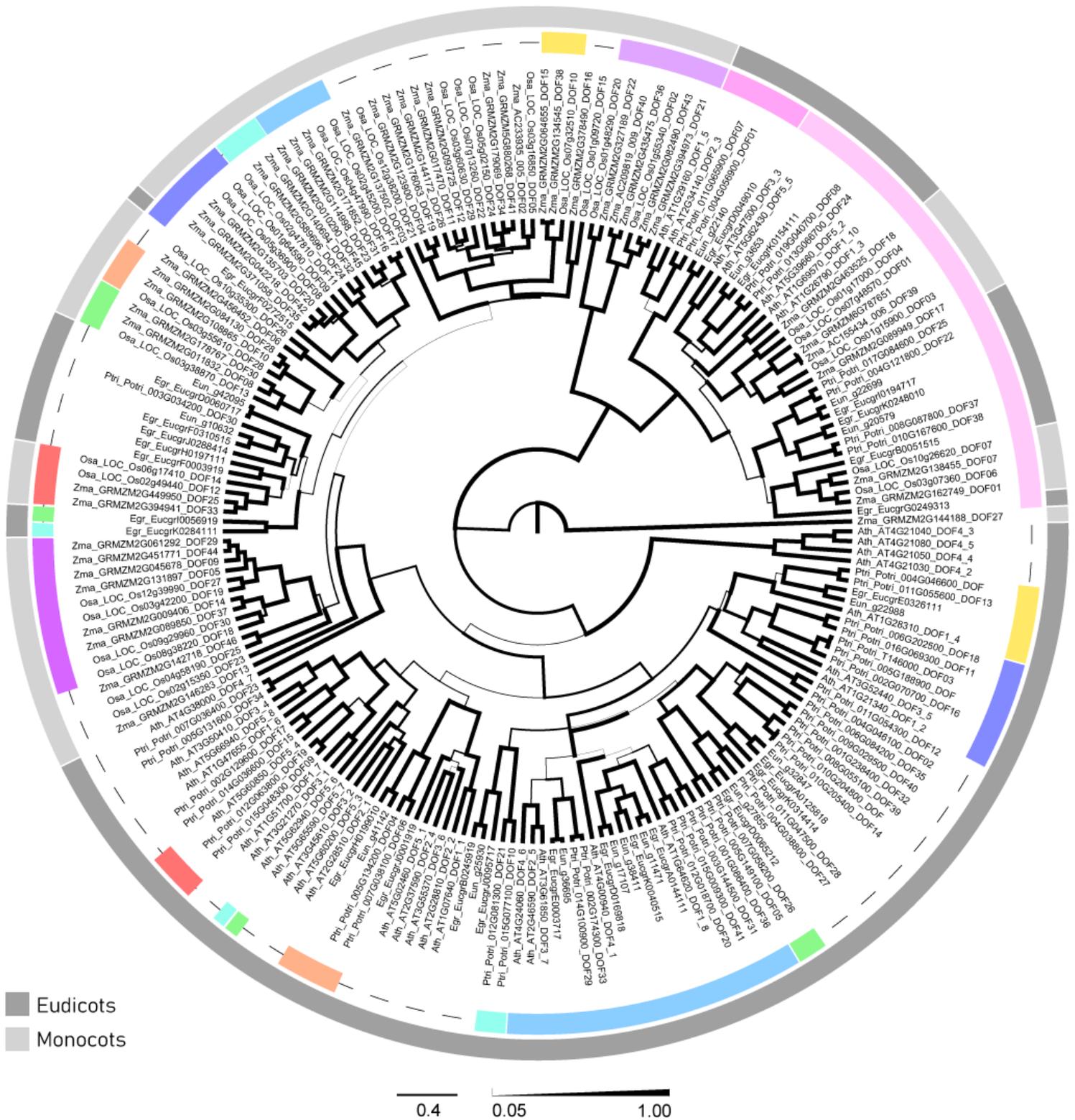
#### 4.2 Análise filogenética

A análise filogenética dos genes DOF foi realizada a partir do alinhamento de um total de 202 sequências, incluindo 17 sequências identificadas em *E. uniflora*, 36 sequências de *Arabidopsis thaliana*, 26 sequências de *Eucalyptus grandis*, 30 sequências de *Oryza sativa*, 46 sequências de *Populus trichocarpa* e 47 de *Zea mays*. Dentre todos os genomas analisados e incluídos nessa análise, apenas o de *Populus trichocarpa* e *Zea mays* proporcionaram sequências ainda não caracterizadas em estudos prévios, totalizando 6 sequências não caracterizadas filogeneticamente (1 em *Z. mays*, GRMZM6G787651, e 5 em *P. trichocarpa*, Potri.004G046600, Potri.010G204800, Potri.005G188900, Potri.011G054400 e

Potri.007G116600). O pseudogene DOF1.9 de *Arabidopsis thaliana* não foi recuperado pelo fato de possuir um códon de parada no meio da sua sequência, impedindo a identificação pela análise HMMER. A partir da construção das árvores filogenéticas de aminoácidos (Figura 4) e de nucleotídeos (Figura 5), observamos que, em geral, a árvore de nucleotídeos obteve uma divisão clara entre monocotiledôneas e eudicotiledôneas, enquanto a árvore de aminoácidos obteve uma distribuição mais homogênea das sequências de cada espécie. Esse fato é explicado pelas substituições silenciosas que se acumularam ao longo do tempo desde a divergência das espécies, enquanto a pressão seletiva atuava na manutenção do domínio DOF e, possivelmente, a função de suas sequências. À luz desse achado, se confirma que a árvore de aminoácidos é a mais adequada para estudos de relação funcional, enquanto a de nucleotídeos é a mais adequada para estudos de relação evolutiva entre sequências DOF. A árvore de nucleotídeos apresenta valores de probabilidade posterior maiores de seus ramos quando comparada à árvore de aminoácidos, visto que possui apenas quatro caracteres.



**Figura 4.** Árvore filogenética de aminoácidos dos genes DOF construída segundo a metodologia bayesiana e o modelo evolutivo JTT, considerando heterogeneidade de substituições dentro das seqüências (JTT+G). Barras coloridas delimitam os grupos conservados entre as seqüências DOF. A linha pontilhada representa seqüências não agrupadas. A espessura dos ramos está correlacionada com o valor da probabilidade posterior de acordo com a legenda.



**Figura 5. Árvore filogenética de nucleotídeos dos genes DOF** construída segundo a metodologia bayesiana e o modelo evolutivo GTR, considerando sítios invariáveis e heterogeneidade de substituições dentro das sequências (GTR+I+G). Barras coloridas delimitam os grupos conservados entre as sequências DOF. A linha pontilhada representa sequências não agrupadas. Barras coradas de cinza claro representam sequências de espécies monocotiledôneas, enquanto barras coradas de cinza escuro representam sequências de espécies eudicotiledôneas. A espessura dos ramos está correlacionada com o valor da probabilidade posterior de acordo com a legenda.

Por serem da mesma família, as sequências de pitangueira sempre se agrupam com pelo menos uma sequência de eucalipto, fato que também é observado pelo agrupamento das sequências de milho e arroz, pertencentes à família *Poaceae*. Baseado nisso, seria plausível dizer que as sequências não agrupadas de eucalipto provavelmente possuem seus ortólogos em pitangueira, totalizando 9 outras sequências, entretanto, não foram recuperadas pelo fato deste estudo utilizar o transcrito de folhas de *E. uniflora* e os genes não identificados podem não ser expressos neste tecido. Devemos lembrar também dos outros 2 genes de pitangueira que foram retirados da análise por estarem com o domínio incompleto. Quanto aos genes de *P. trichocarpa*, todos aparentam estar agrupados com pelo menos outro parálogo em ambas as árvores, devido ao seu evento de poliploidia recente.

A partir da construção da árvore filogenética de aminoácidos (Figura 4), 11 grupos distintos foram identificados. Grande parte dos grupos possui, em média, 7 sequências, mas alguns apresentam números acima de 25. O maior grupo de todos, com 30 sequências, Grupo 2.1 se manteve presente em ambas as árvores com os mesmos integrantes e com uma probabilidade posterior de 0,88 na árvore de aminoácidos e 1 na árvore de nucleotídeos. A diferença presente entre as árvores é a clara divisão entre monocotiledôneas e eudicotiledôneas que ocorreu dentro do Grupo 2.1 na árvore de nucleotídeos, sugerindo que, apesar de sua conservação, ainda há uma variação considerável para discernir os dois grupos. Previamente agrupadas com as sequências CDFs do Grupo 2.1 (WEN et al., 2016; YANAGISAWA, 2002), os genes *CDF4* e *COG1* mostram-se agrupados e isolados desse grupo, portanto caracterizam um novo clado, aqui nomeado de Grupo 2.2. Os Grupos 2.1 e 2.2 possuem um *outgroup*, nomeado de Grupo 2.3, constituído apenas de sequências de monocotiledôneas, com uma probabilidade posterior alta na árvore de nucleotídeos e aminoácidos. Como não há sequências de monocotiledôneas no Grupo 2.2, poderíamos considerar que o Grupo 2.3 representa um agrupamento de genes correspondentes, entretanto, nenhum dos genes do Grupo 2.3 foi caracterizado. Uma observação sobre o Grupo 2.1 e Grupo 2.2 é que suas sequências representativas possuem metioninas conservadas nas posições 13 e 45 do domínio DOF, o que pode contribuir para o seu agrupamento na árvore de nucleotídeos. Isso poderia explicar porque o Grupo 2.1 foi o único grupo que se

conservou nas duas árvores. Por mais interessante que seja essa característica, ainda não é possível dizer com certeza se a influência das metioninas conservadas contribuiu para o seu agrupamento, já que há apenas um códon correspondente a metionina, ou se o Grupo 2.1 e 2.2 são uma das divisões mais antigas dentre as sequências dos fatores de transcrição DOF. O Grupo 2.3, entretanto, não possui resíduos de metionina conservados e provavelmente por isso ficou representado como *outgroup*. Quatro sequências de *E. uniflora* estão presentes nesses grupos: *g20579*, *g22699* e *g3653* no Grupo 2.1 e *g22140* no Grupo 2.2. Conseqüentemente, podemos supor que elas fazem parte de processos fisiológicos como o florescimento e desenvolvimento da plântula respectivamente. Os genes *g20579* e *g22699* estão agrupados perto das sequências CDF de *A. thaliana*, indicando uma função na regulação de genes relacionados ao florescimento. O gene *g3653* está fortemente agrupado com o gene *OsDOF4*, o qual já foi caracterizado funcionalmente como atuante no processo de florescimento (WU et al., 2017a). Por último, o gene *g22140* está agrupado com o gene *COG1*, indicando uma possível função como transdutor de sinais externos mediados por fitocromo e a conversão desses sinais para o desenvolvimento da planta.

Outro grupo que se manteve em ambas as árvores é o Grupo 4.0, representado apenas por sequências de *Zea mays* e *Oryza sativa*. Dentre as sequências presentes no grupo, apenas *OsDOF18* e *OsDOF30* estão caracterizadas. Mutantes para *OsDOF18* mostraram uma subexpressão de transportadores de amônia, uma das principais fontes de nitrogênio inorgânico para a planta, assim como uma prorrogação da senescência foliar e inibição da síntese de ácido jasmônico (SHIM et al., 2019; WU et al., 2017b). Tais características sugerem que esse gene possa estar envolvido na remobilização e absorção de nutrientes, como nitrogênio, em resposta a estresses externos mediados pela via de ácido jasmônico, sejam eles abióticos ou bióticos, entretanto mais estudos precisam ser realizados para a devida caracterização de sua função. Por sua vez, ensaios de mono híbrido e EMSA mostraram interação entre *OsDOF30* e a região promotora de Piruvato Ortofosfato Diquinase (*C4PPDK*), gene codificador da enzima responsável pela conversão de piruvato para fosfoenolpiruvato, etapa essencial da fotossíntese de plantas C4 (ZHANG et al., 2015). Mesmo tendo sequências caracterizadas, ainda não é certo o que se pode especular quanto a função em comum das sequências presentes nesse grupo, assim como a especificidade do grupo em possuir apenas

espécies monocotiledôneas ou se falta representatividade de espécies eudicotiledôneas na árvore.

Uma tendência de estudos prévios é a nomenclatura do agrupamento das sequências *AthDOF4.2*, *AthDOF4.3*, *AthDOF4.4* e *AthDOF4.5* como Grupo I ou C3 em suas árvores (WEN et al., 2016; YANAGISAWA, 2002); esse agrupamento não apresenta nenhuma outra espécie agrupada e possivelmente suporta o fato observado anteriormente da falta de espécies eudicotiledôneas na árvore, principalmente da família das Brassicaceae. Claramente é originado de múltiplos eventos de duplicação, possivelmente após a divergência das Brassicales. Pelo fato de não estarem agrupadas com outra espécie, não há provas que não são específicas apenas de *A. thaliana* e, portanto, não foram consideradas como um grupo funcional conservado das sequências DOF. Ambas as árvores suportam que esse grupo é isolado e não possui nenhuma outra sequência representante, o que indica que essas sequências evoluíram suas funções características isoladamente de outras espécies.

Os grupos restantes não foram mantidos em ambas as árvores, tendo uma clara divisão entre sequências de monocotiledôneas e eudicotiledôneas, mas mantendo as sequências agrupadas dentro dessas divisões. Um deles é o Grupo 3.2, que não possui nenhuma sequência caracterizada, assim como nenhuma sequência de *E. uniflora*, apenas de *E. grandis*, e conseqüentemente não é possível afirmar muito sobre. Entretanto ele possui o gene *OsDOF23* como *outgroup*, já descrito como regulador de genes envolvidos na estocagem de nutrientes na semente (YAMAMOTO et al., 2006). Sequencialmente mais próximo a ele está o Grupo 3.1, agrupando os genes *AthDOF1.8*, *DAG2*, *DAG1*, *AthDOF4.1* (*ITD1*), *AthDOF4.6*, tendo apenas três genes estudados. Isso nos informa que os dois genes atuam de maneiras semelhantes, porém com funções contrastantes. Quanto ao gene *ITD1*, não se sabe qual a sua função, mas seu padrão de expressão mostra localização nas células do córtex e endoderme da raiz (CHEN et al., 2013). Além disso, ele apresenta um domínio de tráfego intercelular com sobreposição ao domínio DOF, permitindo a sua dispersão da região de expressão para tecidos próximos por difusão via plasmodesmos, característica fundamental para a atuação de alguns genes *DOF* envolvidos com o desenvolvimento de tecidos vasculares (MIYASHIMA et al., 2019). Quatro genes de *E. uniflora* estão presentes no Grupo 3.1, sendo eles: *g11471*, *g38411*, *g17107* e *g36695*. Os genes *g11471* e *g38411*

foram provavelmente originados de um evento de duplicação, assim como seus ortólogos em *E. grandis*, evento que é corroborado tanto pela árvore de nucleotídeos quanto a de aminoácidos, sugerindo duplicação antes da divergência das duas espécies, assim como semelhança de função entre os quatro genes. Dentro do Grupo 3.1, não podemos discernir com qual outra sequência caracterizada eles estão agrupados, portanto torna-se difícil inferir exatamente as suas funções. O mesmo vale para o gene *g17107*, que está agrupado a *AthDOF1.8* e *AthDOF4.6*, genes ainda não caracterizados. Um fator importante é reconhecer que a probabilidade posterior relacionada ao ramo de ITD1 e *g36695* não possui uma grandeza relevante, podendo indicar que *g36695* seria ortólogo de DAG1 e DAG2. Em seguida, o Grupo 7.0 não possui nenhuma sequência de *E. uniflora*, entretanto, possui um gene de *E. grandis*, o que poderia ser indício da presença de um gene de pitangueira ortólogo ao de eucalipto que não foi recuperado no transcrito disponível. Dentro do grupo está presente o gene *HCA2* de *A. thaliana*, indicando que os genes relacionados poderiam atuar nas etapas de desenvolvimento do câmbio vascular. O grupo 5.0 apresenta, novamente, apenas uma sequência de *E. grandis* e, pela presença de *SCAP1*, possivelmente faz referência a genes atuantes na maturação e manutenção da atividade de estômatos. O Grupo 9.0 possui apenas o gene *g22988* de *E. uniflora*, sem outros genes previamente descritos. O Grupo 8.0 possui apenas o gene *g41142* de *E. uniflora*, sem outros genes previamente descritos. Não há presença de nenhuma sequência de *A. thaliana*, podendo indicar que houve uma perda durante a história evolutiva da espécie, entretanto não é possível definir em qual momento isso ocorreu. Por não ter sido caracterizado filogeneticamente, é possível que o pseudogene *AthDOF1.9*, de *A. thaliana*, esteja agrupado ao Grupo 8.0, tornando aquisição de mutações *nonsense* o evento de perda.

O último Grupo presente na árvore é o Grupo 6.0, que não possui sequência de *E. uniflora* nem de *E. grandis*. Semelhante ao Grupo 8.0, isso pode ter ocorrido pela pseudogenização de sequências DOF ou perda de material genético pelos membros da família *Myrtaceae*, sendo necessária a busca em bancos de dados genômicos para a elucidação desse questionamento. Presente no Grupo 6.0 está o gene *OBP4*. A ausência de genes ortólogos a *OBP4* é interessante do ponto de vista da família *Myrtaceae*, pois permitiria a expansão da raiz em busca de água em

maiores profundidades, como no caso do eucalipto, e suporte a condições estressantes, como a adaptação da pitangueira ao solo da restinga.

Cinco genes de *E. uniflora* não fazem parte de grupos funcionalmente conservados. Dois genes, *g27855* e *g42095*, agruparam-se apenas com seus ortólogos em eucalipto, *EucgrD0065212* e *EucgrD0060717* respectivamente, e dois outros genes de *P. trichocarpa*. O gene *g25930* possui certa semelhança com *AthDOF2.1*, gene regulador do desenvolvimento primário da raiz. O gene *g25930* provavelmente atua em uma maneira semelhante em *E. uniflora*, sendo dependente tanto dos níveis de auxina como de citocinina para atuação. O gene *g32847* agrupou-se apenas com seu ortólogo em eucalipto, não possuindo nenhum suporte em seus outros agrupamentos. Por último, *g10632* possui certa semelhança com os genes ADOF1 e ADOF2, responsivos a temperaturas baixas (CARVALLO et al., 2011). Por mais que seu mecanismo de atuação não foi elucidado, *g10632* poderia ser um gene candidato a estudos de aclimatação em baixas temperaturas.

Há ainda grupos suportados na árvore de nucleotídeos que estão agrupados na árvore de aminoácidos, entretanto com uma probabilidade posterior muito baixa, como o grupo do gene *g32847*, indicando que seu tempo de divergência não é grande o suficiente para apagar a similaridade entre as sequências de aminoácidos, ou que sua taxa de mutação é muito lenta. O seu agrupamento é resultado de uma história evolutiva comum, ao invés de uma semelhança funcional. Possivelmente representam sequências sobre uma pressão de seleção relaxada quando comparadas a outros clados.

#### 4.3 Análise de motivos conservados e domínios transmembrana

A fim de entender melhor como é conservada a função dos genes DOF de *E. uniflora*, foi realizada uma análise de motivos conservados no programa MEME. Um total de quinze motivos foram buscados nas sequências fornecidas. Um resumo das informações de cada motivo está presente na Figura 6. Os motivos estão ordenados de acordo a estatística de *e-value*, tendo o motivo 15 a menor probabilidade. Todos apresentam relevância estatística, com o motivo 15 possuindo um *e-value* de 1E-80. O menor motivo é o #13, apresentando o parâmetro mínimo de 10 resíduos de aminoácidos, enquanto o maior motivo é o #1, representando parte do domínio DOF

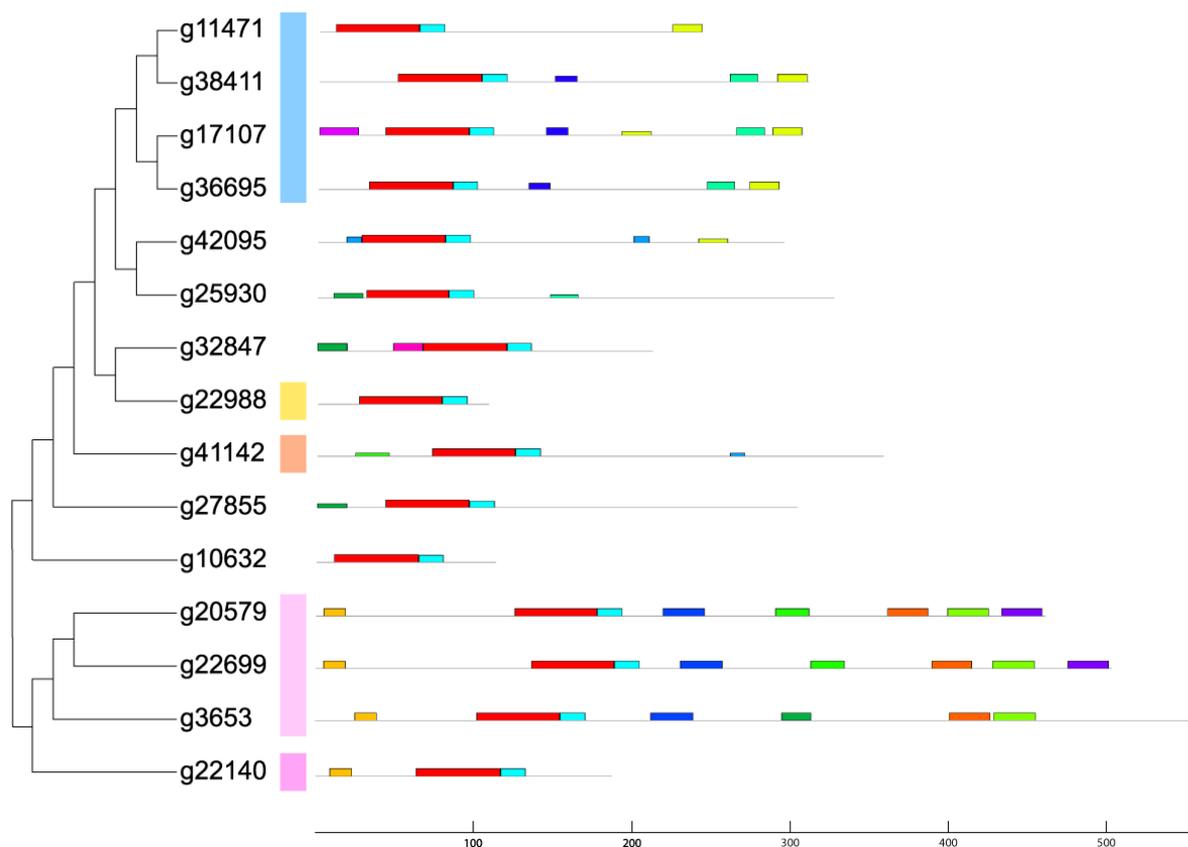
e possuindo 53 resíduos de aminoácido. Em geral, os motivos recuperados apresentam um comprimento médio de 20 aminoácidos.

A estrutura das sequências recuperadas de *E. uniflora* e seus motivos está representada na Figura 7. Todas apresentam o domínio DOF em vermelho, tal como o motivo #2 em ciano. Estudos prévios caracterizaram os resíduos 6 a 10 do motivo #2 como participantes a um sinal de localização nuclear bipartido em genes DOF, composto pela região B1, resíduos 32 a 36, e B2, resíduos 54 a 59 do domínio DOF, pertencente ao motivo #2 (KREBS; MUELLER-ROEBER; RUZICIC, 2009). Não é possível afirmar que o motivo #2 seja um motivo a parte e autônomo, ou se é artefato do algoritmo, afinal ele é uma continuação do domínio DOF. Outra característica interessante do motivo #2 é a presença de vários resíduos de serina em sequência, indicando uma possível região de fosforilação. Pela sua proximidade à região B2, poderíamos questionar se esse sítio de fosforilação atuaria como um interruptor da localização nuclear das proteínas DOF, permitindo a sua difusão via plasmodesmos e regulação da expressão gênica de maneira célula não-autônoma.

Number & Color	Motif Logo	Consensus Sequence
#1		ARPQPEEALKCPRCDSTNTKFCYNNVSLQPRFCKQRRWTGGALRNVP
#2		VGGGCRKNKRSSSSSS
#3		WVPKTLRIDDP EAAKSSIWATLGIK
#4		KNHASETSPVLQANPAALSRSQNFQE
#5		KDPAIKLFGKTIPL
#6		MVFSSLPVYLDPPNWQQQ
#7		NGTVLTFGSDAPLCEMASVNLNLAEEK
#8		GGGASIRPGSMAERARLA
#9		SSGSNSPTLGKHSRDENPLKPEKSEE
#10		GGGDSNGYWNGMLGGGSW
#11		NCNGFPQJPCFPGPPWPYPWN
#12		GRLLFPFDLKPQVSSAS
#13		EQWRLQQQQ
#14		IDLNLAFPKLNHH
#15		MDTAQWPQEIVVKPIEEIVTNTCPK

Figura 6. Identificação dos motivos conservados nas sequências DOF.

Todos os motivos seguintes ao motivo #2 estão presentes em apenas algumas das sequências e há um claro perfil de motivos entre grupos funcionais (Figura 5). O Grupo 2.1 possui, além do domínio DOF e o motivo #2, os motivos #3, #4, #5, #6, #7, #9 e #11, mostrando uma clara conservação de funcionalidade. Estudos prévios já caracterizaram os motivos #3 e #9 como motivos de ligação à FKF1 e GI respectivamente (LUCAS-REINA et al., 2015). O gene *g22140*, pertencente ao Grupo 2.1, também possui o motivo #5, revelando uma possível função em comum com os genes do Grupo 2.1. Além dos motivos #3 e #9, os outros motivos não têm suas funções elucidadas, indicando possíveis interações a serem descobertas. O gene *g3653* é o único que apresenta o motivo #11 dentro do Grupo 2.1, sugerindo uma atuação diferente de seus homólogos e diversificação para interação com outras proteínas. Todos os genes pertencentes ao Grupo 3.1 possuem o motivo #10, sendo o gene *g11471* o único que não possui os motivos #12 e #14. O gene *g17107* é o único dentre todos os outros que apresenta o motivo #15, assim como um motivo repetido, entretanto com um *e-value* maior. Os genes *g42095* e *g25930* são os genes mais próximos agrupados aos genes do Grupo 3.1. A sua ancestralidade funcional em relação aos genes do Grupo 3.1 mostra-se evidente na árvore filogenética, assim como na análise de seus motivos conservados. A presença de sequências semelhantes aos motivos #10 e #12, em *g42095* e *g25930* respectivamente, traz a possibilidade de motivos sobre seleção neutra e adquirindo mutações não sinônimas. Os genes restantes são pobres em motivos C-terminais e possuem um ou dois motivos na sua porção N-terminal. Os genes *g27855* e *g41142* apresentam motivos com maiores valores de *e-value* e poderiam apresentar mais caso o número limite de motivos fosse aumentado. O gene *g32847* é o único que apresenta o motivo #8. Os genes *g10632* e *g22988* não possuem outro motivo além do domínio DOF e motivo #2, possivelmente por erros na montagem do transcrito.



**Figura 7. Representação dos motivos nas sequências DOF de *Eugenia uniflora*.** As sequências estão organizadas de acordo com a sua disposição na árvore filogenética de aminoácidos (Figura 1), entretanto o comprimento dos ramos é apenas ilustrativo, desproporcionais ao tempo de divergência. Barras coloridas representam a quais grupos funcionais cada gene pertence. Linhas pretas indicam regiões não conservadas entre as sequências, enquanto caixas coloridas representam motivos conservados de acordo com a Tabela 2. Caixas finas representam motivos com *e-value* entre  $1e-10$  e a linha de corte de  $1e-5$ . A régua embaixo representa o comprimento da sequência em resíduos de aminoácidos.

Quanto à predição de domínios transmembrana, não houve nenhum resultado positivo, indicando que as proteínas podem estar dispersas, seja no citoplasma ou no núcleo.

#### 4.4 Análise de Pressão Seletiva

Diferentes forças de seleção podem atuar sobre uma mesma sequência. Quando estamos falando de domínios, a principal força atuante é a seleção purificadora, a fim de conservar e manter a funcionalidade do domínio. Mesmo assim, é esperado que haja diversificação da sequência e a possibilidade de outras forças de seleção atuarem no seu lugar a partir disso. Com o propósito de avaliar as forças de seleção atuantes no domínio DOF, uma análise de pressão seletiva foi

realizada no programa CodeML do pacote de programas PamL. A partir da árvore de aminoácidos, e sequência de nucleotídeos, cada valor de  $\ln L$ , para todos os modelos comparados, foi calculado (Tabela 3). Os modelos M2 e M3 foram aceitos enquanto as hipóteses nulas dos modelos M1 e M0 foram rejeitadas, sendo M7 a única hipótese nula aceita. Ao aceitar M2, estamos supondo que há pressão seletiva atuando no domínio e ao conferir os valores de  $\omega$ , todas as sequências apresentaram  $\omega < 1$ , caracterizando a seleção como purificadora. No caso de M3, supomos que a variação de substituições dos sítios presentes no domínio DOF não é homogênea, com códons exibindo maiores padrões de variação que outros, caso também observado na figura do logo do domínio DOF (Figura 3). Por último, a aceitação do modelo M7 representa que não há seleção positiva atuando no domínio DOF. Em conclusão o domínio DOF está sob seleção purificadora atuando de maneira heterogênea em sua sequência, preservando os aminoácidos responsáveis pela sua função.

**Tabela 3. Resultado do software CodeML.**

<b>Model</b>	<b><math>\ln L</math></b>	<b>Comparison</b>	<b> LRT </b>
M1	-13744.20171	M1 x M2	80.96062
M2	-13784.68202		p-value $\ll 10E-3$
M0	-13736.94726	M0 x M3	1357.93976
M3	-13057.97738		p-value $\ll 10E-3$
M7	-13032.45612	M7 x M8	0.0012
M8	-13032.45672		p-value $> 0.90$

## 5. CONCLUSÕES E PERSPECTIVAS

Em conclusão, os genes DOF apresentam vários grupos funcionais conservados, tanto em espécies de eudicotiledôneas como monocotiledôneas. Embora o seu tempo de divergência seja grande, a constante atuação de forças de pressão purificadora atuantes no domínio DOF permitiram o agrupamento de proteínas DOF de ambos os grupos em clados funcionalmente conservados. Em contraste, o acúmulo de mutações sinônimas é evidente quando realizada uma análise filogenética pela sequência nucleotídica dos genes, mostrando uma história

evolutiva muito mais complexa. Dentre os grupos identificados, há a presença de sequências conservadas atuantes em processos distintos, como florescimento, desenvolvimento, germinação, formação de tecidos vasculares, fotorespostas, formação estomatal, entre outros.

A construção de uma árvore filogenética é uma excelente estratégia para a uma caracterização inicial de sequências previamente desconhecidas. Entre os 17 genes DOF identificados no transcrito de *E. uniflora*, 10 foram agrupados aos grupos funcionais descritos anteriormente. A presença de sequências não agrupadas mostra uma grande diversidade de possíveis funções ainda não descobertas e em constante evolução. Mesmo assim, foi possível a identificação de motivos conservados em quase todas as sequências, sendo muitos deles para interações proteína-proteína, essenciais para a sua atuação.

A repetição dos experimentos filogenéticos envolvendo tanto as sequências de nucleotídeos quanto as de aminoácidos é uma perspectiva para garantir a confiabilidade dos resultados obtidos. A identificação de motivos conservados utilizando apenas grupos funcionais pode vir a ser uma perspectiva a fim de identificar regiões conservadas e importantes para a atuação de grupos específicos. Com o final da anotação do genoma de *E. uniflora*, sua sequência genômica poderia ser utilizada para uma varredura e identificação de novas sequências DOF previamente não caracterizadas, assim como análise de suas estruturas gênicas. Entre as sequências identificadas nesse estudo, experimentos funcionais serão importantes para a confirmação e correta caracterização dos dados obtidos, como análise fenotípica de mutantes, expressão tecidual, assim como testes de complementação fenotípica com genes de *E. uniflora* em plantas de *A. thaliana*.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, G. et al. Phylogenetic comparative and expression analysis of genes encoding dof transcription factors from *Eucalyptus grandis*. **BMC Proceedings**, [s. l.], v. 5, n. Suppl 7, p. 1–2, 2011.
- BAILEY, T. L.; ELKAN, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. **Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology**, [s. l.], p. 28–36, 1994.
- BOCCACCINI, A. et al. Independent and interactive effects of DOF affecting

germination 1 ( DAG1 ) and the DELLA proteins GA insensitive ( GAI ) and Repressor of ga1-3 ( RGA ) in embryo development and seed germination. **BMC Plant Biology**, [s. l.], p. 1–11, 2014.

BOCCACCINI, A. et al. The DAG1 transcription factor negatively regulates the seed-to-seedling transition in Arabidopsis acting on ABA and GA levels. **BMC Plant Biology**, [s. l.], p. 1–11, 2016.

BUESO, E. et al. Arabidopsis COGWHEEL1 links light perception and gibberellins with seed tolerance to deterioration. **The Plant Journal**, [s. l.], p. 583–596, 2016.

CAI, X. et al. Genome-wide Analysis of Plant-specific Dof Transcription Factor Family in Tomato. **Journal of Integrative Plant Biology**, [s. l.], v. 55, n. 6, p. 552–566, 2013.

CARVALLO, M. A. et al. A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance : Solanum commersonii , Solanum tuberosum , and Arabidopsis thaliana. **Journal of Experimental Botany**, [s. l.], v. 62, n. 11, p. 3807–3819, 2011.

CASTORINA, G. et al. A novel role for STOMATAL CARPENTER 1 in stomata patterning. **BMC Plant Biology**, [s. l.], p. 1–14, 2016. Disponível em: <<http://dx.doi.org/10.1186/s12870-016-0851-z>>

CHEN, H. et al. Evolutionary and molecular analysis of Dof transcription factors identified a conserved motif for intercellular protein trafficking. **New Phytologist**, [s. l.], v. 1, p. 1250–1260, 2013.

CHEN, W.; CHAO, G.; SINGH, K. B. **The promoter of a H<sub>2</sub>O<sub>2</sub>-inducible, Arabidopsis glutathione S-transferase gene contains closely linked OBF- and OBP-1 binding sites**, The Plant Journal, 1996.

CHEN, Y.; CAO, J. Comparative Analysis of Dof Transcription Factor Family in Maize. **Plant Molecular Biology Reporter**, [s. l.], 2014.

CROOKS, G. E. et al. WebLogo : A Sequence Logo Generator. **Genome Research**, [s. l.], p. 1188–1190, 2004.

DARRIBA, D. et al. Europe PMC Funders Group jModelTest 2 : more models , new heuristics and high- performance computing. **Bioinformatics**, [s. l.], v. 9, n. 8, p. 6–9, 2015.

DARRIBA, D. et al. Europe PMC Funders Group ProtTest 3 : fast selection of best-fit models of protein evolution. **Nature Methods**, [s. l.], v. 27, n. 8, p. 1164–1165, 2017.

DAVID, K. M. et al. Arabidopsis GIGANTEA protein is post-transcriptionally regulated by light and dark. **FEBS Letters**, [s. l.], v. 580, p. 1193–1197, 2006.

FALCÃO, T. R. et al. Crude extract and fractions from Eugenia uniflora Linn leaves showed anti-inflammatory, antioxidant, and antibacterial activities. **BMC Complementary and Alternative Medicine**, [s. l.], v. 18, n. 1, p. 1–12, 2018.

FORNARA, F. et al. Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering

Response. **Developmental Cell**, [s. l.], v. 17, n. 1, p. 75–86, 2009.

GABRIELE, S. et al. The Dof protein DAG1 mediates PIL5 activity on seed germination by negatively regulating GA biosynthesis. **The Plant Journal**, [s. l.], p. 312–323, 2010.

GUALBERTI, G. et al. Mutations in the Dof Zinc Finger Genes DAG2 and DAG1 Influence with Opposite Effects the Germination of Arabidopsis Seeds. **The Plant Cell**, [s. l.], v. 14, n. June, p. 1253–1263, 2002.

GUO, Y. et al. Dof5 . 6 / HCA2 , a Dof Transcription Factor Gene , Regulates Interfascicular Cambium Formation and Vascular Tissue Development in Arabidopsis. **The Plant Cell**, [s. l.], v. 21, n. November, p. 3518–3534, 2009.

GUZMAN, F. et al. Identification of MicroRNAs from *Eugenia uniflora* by High-Throughput Sequencing and Bioinformatics Analysis. **PLoS ONE**, [s. l.], v. 7, n. 11, 2012.

HICHR, I. et al. The *Solanum lycopersicum* zinc finger2 cysteine-2/histidine-2 repressor-like transcription factor regulates development and tolerance to salinity in tomato and arabidopsis. **Plant Physiology**, [s. l.], v. 164, n. 4, p. 1967–1990, 2014.

HIR, R. Le; BELLINI, C. The plant-specific Dof transcription factors family : new players involved in vascular system development and functioning in Arabidopsis. **Frontiers in Plant Science**, [s. l.], v. 4, n. May, 2013.

IMAIZUMI, T. et al. FKF1 F-Box Protein Mediates Cyclic Degradation of a Repressor of CONSTANS in Arabidopsis. **Science**, [s. l.], v. 293, n. 2005, 2005.

KANG, H. G.; SINGH, K. B. Characterization of salicylic acid-responsive, Arabidopsis Dof domain proteins: Overexpression of OBP3 leads to growth defects. **Plant Journal**, [s. l.], v. 21, n. 4, p. 329–339, 2000.

KONISHI, M. et al. MONOPTEROS directly activates the auxin-inducible promoter of the Dof5 . 8 transcription factor gene in Arabidopsis thaliana leaf provascular cells. **Journal of Experimental Botany**, [s. l.], v. 66, n. 1, p. 283–291, 2015.

KONISHI, M.; YANAGISAWA, S. Transcriptional repression caused by Dof5 . 8 is involved in proper vein network formation in Arabidopsis thaliana leaves. **Journal of Plant Research**, [s. l.], 2015.

KRAHMER, J. et al. Time-resolved interaction proteomics of the GIGANTEA protein under diurnal cycles in Arabidopsis. **FEBS Letters**, [s. l.], v. 593, n. December, p. 319–338, 2019.

KREBS, J.; MUELLER-ROEBER, B.; RUZICIC, S. A novel bipartite nuclear localization signal with an atypically long linker in DOF transcription factors. **Journal of Plant Physiology**, [s. l.], v. 167, n. 7, p. 583–586, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.jplph.2009.11.016>>

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7 : Molecular Evolutionary Genetics Analysis Version 7 . 0 for Bigger Datasets Brief communication. **Molecular Biology and Evolution**, [s. l.], v. 33, n. 7, p. 1870–1874, 2016.

KUREPA, J. et al. Modulation of auxin and cytokinin responses by early steps of the phenylpropanoid pathway. **BMC Plant Biology**, [s. l.], p. 1–15, 2018.

KUSHWAHA, H. et al. Genome wide identification of Dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. **Molecular Biology Reports**, [s. l.], v. 38, n. 8, p. 5037–5053, 2011.

LAPORTE, M. M. et al. Engineering for drought avoidance : expression of maize NADP-malic enzyme in tobacco results in altered stomatal function. **Journal of Experimental Botany**, [s. l.], v. 53, n. 369, p. 699–705, 2002.

LASHBROOKE, J. et al. The tomato MIXTA-like transcription factor coordinates fruit epidermis conical cell development and cuticular lipid biosynthesis and assembly. **Plant Physiology**, [s. l.], v. 169, n. 4, p. 2553–2571, 2015.

LI, H. et al. Genome-Wide Identification and Expression Analysis of the Dof Transcription Factor Gene Family in *Gossypium hirsutum* L. **Agronomy**, [s. l.], v. 8, n. 9, p. 186, 2018.

LIJAVETZKY, D.; CARBONERO, P.; VICENTE-CARBAJOSA, J. Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. **BMC Evolutionary Biology**, [s. l.], v. 3, p. 1–11, 2003.

LUCAS-REINA, E. et al. An Evolutionarily Conserved DOF-CONSTANS Module Controls Plant Photoperiodic Signaling. **Plant Physiology**, [s. l.], v. 168, n. June, p. 561–574, 2015.

MELNYK, C. W. et al. Transcriptome dynamics at Arabidopsis graft junctions reveal an intertissue recognition mechanism that activates vascular regeneration. **PNAS PLUS**, [s. l.], v. 115, n. 10, 2018.

MIYASHIMA, S. et al. Mobile PEAR transcription factors integrate positional cues to prime cambial growth. **Nature Letters**, [s. l.], 2019. Disponível em: <<http://dx.doi.org/10.1038/s41586-018-0839-y>>

MORENO-RISUENO, M. Á. et al. The family of DOF transcription factors: From green unicellular algae to vascular plants. **Molecular Genetics and Genomics**, [s. l.], v. 277, n. 4, p. 379–390, 2007.

NEGI, J. et al. A Dof Transcription Factor, SCAP1, Is Essential for the Development of Functional Stomata in Arabidopsis. **Current Biology**, [s. l.], v. 23, n. 6, p. 479–484, 2013.

NOLAN, T.; CHEN, J.; YIN, Y. Cross-talk of Brassinosteroid signaling in controlling growth and stress responses. **Biochemistry Journal**, [s. l.], v. 474, n. 16, p. 2641–2661, 2018.

OH, E. et al. PIL5 , a Phytochrome-Interacting bHLH Protein , Regulates Gibberellin Responsiveness by Binding Directly to the GAI and RGA Promoters in Arabidopsis Seeds. **The Plant Cell**, [s. l.], v. 19, n. April, p. 1192–1208, 2007.

OLIVEIRA-FILHO, ARY T. & FONTES, M. A. L. Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate '. **Biotropica**, [s. l.], v. 32, p. 793–810, 2000.

PARK, D. H. et al. The Arabidopsis COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. **Plant Journal**, [s. l.], v. 34, n. 2, p. 161–171, 2003.

PISKUREWICZ, U. et al. Far-red light inhibits germination through DELLA-dependent stimulation of ABA synthesis and ABI3 activity. **The EMBO Journal**, [s. l.], v. 28, n. 15, p. 2259–2271, 2009.

RAMIREZ-PARRA, E. et al. The transcription factor OBP4 controls root growth and promotes callus formation. **New Phytologist**, [s. l.], p. 1787–1801, 2017.

RATTMANN, Y. D. et al. Analysis of flavonoids from eugenia uniflora leaves and its protective effect against murine sepsis. **Evidence-based Complementary and Alternative Medicine**, [s. l.], v. 2012, 2012.

RAVINDRAN, P. et al. A Novel RGL2 – DOF6 Complex Contributes to Primary Seed Dormancy in Arabidopsis thaliana by Regulating a GATA Transcription Factor. **Molecular Plant**, [s. l.], v. 10, n. 10, p. 1307–1320, 2017. Disponível em: <<https://doi.org/10.1016/j.molp.2017.09.004>>

ROESCH, LUIZ FERNANDO WURDIG; VIEIRA, FREDERICO COSTA BEBER; PEREIRA, VILMAR ALVES; SCHÜNEMANN, ADRIANO LUIS; TEIXEIRA, ITALO FILIPPI; SENNA, ANA JULIA TEIXEIRA; STEFENON, V. M. The Brazilian Pampa: A Fragile Biome. **Diversity**, [s. l.], v. 1, n. 2071–1050, p. 182–198, 2009.

RUEDA-ROMERO, P. et al. Arabidopsis thaliana DOF6 negatively affects germination in non-after-ripened seeds and interacts with TCP14. **Journal of Experimental Botany**, [s. l.], v. 63, n. 5, p. 1937–1949, 2012.

RYMEN, B. et al. ABA Suppresses Root Hair Growth via the OBP4. **Plant Physiology**, [s. l.], v. 173, n. March, p. 1750–1762, 2017.

RYNER, L. C. et al. Control of Circadian Rhythms and Photoperiodic Flowering by the Arabidopsis GIGANTEA Gene. **Science**, [s. l.], v. 285, n. September, p. 4–7, 1999.

SANTOPOLO, S. et al. DOF AFFECTING GERMINATION 2 is a positive regulator of light-mediated seed germination and is repressed by DOF AFFECTING GERMINATION 1. **BMC Plant Biology**, [s. l.], p. 1–11, 2015.

SANTOS, K. K. A. et al. Anti-Trypanosoma cruzi and cytotoxic activities of Eugenia uniflora L. **Experimental Parasitology**, [s. l.], v. 131, n. 1, p. 130–132, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.exppara.2012.02.019>>

SCARANO, F. R. Structure , Function and Floristic Relationships of Plant Communities in Stressful Habitats Marginal to the Brazilian Atlantic Rainforest. **Annals of Botany**, [s. l.], v. 90, p. 517–524, 2002.

SCHLERETH, A. et al. MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. **Nature Letters**, [s. l.], v. 464, n. April, p. 2–6, 2010.

SHIGYO, M. et al. Evolutionary Processes During the Formation of the Plant-Specific Dof Transcription Factor Family. **Plant Cell Physiology**, [s. l.], v. 48, n. 1, p. 179–

185, 2007.

SHIM, Y. et al. Rice DNA-Binding One Zinc Finger 24 ( OsDOF24 ) Delays Leaf Senescence in a Jasmonate-Mediated Pathway. **Plant & cell physiology**, [s. l.], v. 24, n. May, p. 1–12, 2019.

SHIMOFURUTANI, N. et al. Functional analyses of the Dof domain, a zinc finger DNA-binding domain, in a pumpkin DNA-binding protein AOBP. **FEBS Letters**, [s. l.], v. 430, p. 0–5, 1998.

SILVA-ROCHA, W. P. et al. Effect of the crude extract of *Eugenia uniflora* in morphogenesis and secretion of hydrolytic enzymes in *Candida albicans* from the oral cavity of kidney transplant recipients. **BMC Complementary and Alternative Medicine**, [s. l.], v. 15, n. 1, p. 1–15, 2015.

SKIRYCZ, A. et al. DOF transcription factor AtDof1 . 1 ( OBP2 ) is part of a regulatory network controlling glucosinolate biosynthesis in *Arabidopsis*. **The Plant Journal**, [s. l.], v. 1, p. 10–24, 2006.

SKIRYCZ, A. et al. Transcription factor AtDOF4 ; 2 affects phenylpropanoid metabolism in *Arabidopsis thaliana*. **New Phytologist**, [s. l.], p. 425–438, 2007.

SKIRYCZ, A. et al. The DOF transcription factor OBP1 is involved in cell cycle regulation in *Arabidopsis thaliana*. **Plant Journal**, [s. l.], v. 56, n. 5, p. 779–792, 2008.

SMET, W. et al. DOF2.1 Controls Cytokinin-Dependent Vascular Cell Proliferation Downstream of TMO5/LHW. **Current Biology**, [s. l.], v. 3, p. 1–26, 2019.

SOBEH, M. et al. Chemical profiling of secondary metabolites of *Eugenia uniflora* and their antioxidant, anti-inflammatory, pain killing and anti-diabetic activities: A comprehensive approach. **Journal of Ethnopharmacology**, [s. l.], v. 240, n. April, p. 111939, 2019. Disponível em: <<https://doi.org/10.1016/j.jep.2019.111939>>

SUCHARD, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1 . 10. **Virus Evolution**, [s. l.], v. 4, n. 1, p. 1–5, 2018.

TOKUNAGA, S. et al. Overexpression of the DOF-Type Transcription Factor Enhances Lipid Synthesis in *Chlorella vulgaris*. **Applied Biochemistry and Biotechnology**, [s. l.], 2019.

TURCHETTO-ZOLET, A. C. et al. Phylogeography and ecological niche modelling in *Eugenia uniflora* (Myrtaceae) suggest distinct vegetational responses to climate change between the southern and the northern Atlantic Forest. **Botanical Journal of the Linnean Society**, [s. l.], v. 182, n. 3, p. 670–688, 2016.

UMEMURA, Y. et al. The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. **Plant Journal**, [s. l.], v. 37, n. 5, p. 741–749, 2004.

VIMOLMANGKANG, S. et al. An apple MYB transcription factor, MdMYB3, is involved in regulation of anthocyanin biosynthesis and flower development. **BMC Plant Biology**, [s. l.], v. 13, n. 1, p. 1, 2013. Disponível em: <BMC Plant Biology>

- WANG, H. et al. Characterization of dof transcription factors and their responses to osmotic stress in poplar (*populus trichocarpa*). **PLoS ONE**, [s. l.], v. 12, n. 1, p. 1–19, 2017.
- WARD, J. M. et al. The Dof Transcription Factor OBP3 Modulates Phytochrome and Cryptochrome Signaling in Arabidopsis. **The Plant Cell**, [s. l.], v. 17, n. February, p. 475–485, 2005.
- WEN, C. L. et al. Identification and characterisation of Dof transcription factors in the cucumber genome. **Scientific Reports**, [s. l.], v. 6, n. March, p. 1–11, 2016. Disponível em: <<http://dx.doi.org/10.1038/srep23072>>
- WU, Q. et al. Constitutive expression of OsDof4 , encoding a C<sub>2</sub>-C<sub>2</sub> zinc finger transcription factor , confers its distinct flowering effects under long- and short-day photoperiods in rice ( *Oryza sativa* L .). **BMC Plant Biology**, [s. l.], p. 1–13, 2017. a.
- WU, Y. et al. Molecules and Cells Transcription Factor OsDOF18 Controls Ammonium Uptake by Inducing Ammonium Transporters in Rice Roots. **Molecules and Cells**, [s. l.], v. 40, n. 3, p. 178–185, 2017. b.
- XU, P. et al. Transcription Factor Gene that Negatively Regulates Cell Cycle Progression and Cell Expansion in Arabidopsis thaliana. **Scientific Reports**, [s. l.], n. February, p. 1–13, 2016. Disponível em: <<http://dx.doi.org/10.1038/srep27705>>
- XU, P.; CAI, W. Nitrate-responsive OBP4-XTH9 regulatory module controls lateral root development in Arabidopsis thaliana. **PLOS Genetics**, [s. l.], v. 9, p. 1–27, 2019. Disponível em: <<http://dx.doi.org/10.1371/journal.pgen.1008465>>
- YAMAMOTO, M. P. et al. Synergism between RPBF Dof and RISBZ1 bZIP Activators in the Regulation of Rice Seed Expression Genes. **Plant Physiology**, [s. l.], v. 141, n. 4, p. 1694–1707, 2006.
- YANAGISAWA, S. The Dof family of plant transcription factors. **Trends in plant science**, [s. l.], v. 7, n. 12, p. 555–60, 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12475498>>
- YANAGISAWA, S.; SCHMIDT, R. J. Diversity and similarity among recognition sequences of Dof transcription factors. **Plant Journal**, [s. l.], v. 17, n. 2, p. 209–214, 1999.
- YANG, Z. PAML 4: Phylogenetic analysis by maximum likelihood. **Molecular Biology and Evolution**, [s. l.], v. 24, n. 8, p. 1586–1591, 2007.
- ZHANG, Y. et al. Functions of OsDof25 in regulation of OsC4PPDK. **Plant Molecular Biology**, [s. l.], v. 89, n. 3, p. 229–242, 2015.
- ZHUOYUN, W. et al. Brassinosteroid Biosynthesis Is Modulated via a Transcription Factor Cascade of COG1 , PIF4 , and PIF5 1. **Plant Physiology**, [s. l.], v. 174, n. June, p. 1260–1273, 2017.
- ZOU, H. et al. The transcription factor AtDOF4 . 2 regulates shoot branching and seed coat formation in Arabidopsis. **Biochemistry Journal**, [s. l.], v. 388, p. 373–388, 2013.

