

CONRADO DE CAMPOS GONÇALVES

"CARACTERIZAÇÃO DE GENES DE FATORES DE TRANSCRIÇÃO EXPRESSOS EM CORONA DE *Passiflora*edulis Sims. (PASSIFLORACEAE)"

CAMPINAS 2014



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

CONRADO DE CAMPOS GONÇALVES

"CARACTERIZAÇÃO DE GENES DE FATORES DE TRANSCRIÇÃO EXPRESSOS EM CORONA DE Passiflora edulis Sims. (PASSIFLORACEAE)"

Este exemplar corresponde à redação final da Dissertação defendida pelo candidato

Conrado de Campos Gonçalves

e aprovada pela Comissão Julgadora.

Dissertação apresentada ao Instituto de Biologia da UNICAMP para obtenção do Título de Mestre em Biologia Vegetal.

Orientador: Prof. Dr. Marcelo Carnier Dornelas

CAMPINAS, 2014

Ficha catalográfica Universidade Estadual de Campinas Biblioteca do Instituto de Biologia Mara Janaina de Oliveira - CRB 8/6972

Gonçalves, Conrado de Campos, 1989-

G586c

Caracterização de genes de fatores de transcrição expressos em corona de Passiflora edulis Sims. (Passifloraceae) / Conrado de Campos Gonçalves. -Campinas, SP: [s.n.], 2014.

Orientador: Marcelo Carnier Dornelas.

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Biologia.

1. Fatores de transcrição. 2. Flores - Crescimento e desenvolvimento. 3. Passiflora edulis. 4. Corona (Botânica). I. Dornelas, Marcelo Carnier, 1970-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Characterization of genes of transcription factors expressed in corona in Passiflora edulis Sims. (Passifloraceae)

Palavras-chave em inglês:

Transcription factors

Flowers - Growth and development

Passiflora edulis Corona (Botany)

Área de concentração: Biologia Vegetal Titulação: Mestre em Biologia Vegetal

Banca examinadora:

Marcelo Carnier Dornelas [Orientador]

Gladys Flávia de Albuquerque Melo de Pinna

Claudia Regina Baptista Haddad Data de defesa: 14-02-2014

Programa de Pós-Graduação: Biologia Vegetal

BANCA EXAMINADORA

Prof. Dr. Marcelo Carnier Dornelas (orientador)	Assinatura
Profa. Dra. Gladys Flávia de Albuquerque Melo de Pinna	Assinatura
Profa. Dra. Claudia Regina Baptista Haddad	<u>Maddool</u> Assinatura
Prof. Dr. Marcos José Salvador	Assinatura
Prof. Dr. Carlos Henrique Inácio Ramos	Assinatura

RESUMO

A corona floral é uma estrutura típica das flores da família Passifloraceae. A elucidação dos mecanismos responsáveis pela formação da corona é de extrema importância para a compreensão dos processos evolutivos que permitiram a diversificação das interações com polinizadores nas espécies desta família. Este trabalho teve por objetivo identificar e caracterizar fatores de transcrição (FT) preferencialmente expressos na corona em Passiflora spp. Para tal, o banco de etiquetas de genes expressos (ESTs) PASSIOMA foi investigado e 139 cDNAs foram identificados: 69 derivados de bibliotecas florais de P. edulis, 68 de P. suberosa e 2 de P. pohlii. No intuito de atribuir identidades às sequências realizou-se uma análise de parcimônia, com a obtenção de cladogramas que permitiram a identificação de membros de 31 famílias gênicas de fatores de transcrição em Passiflora spp. Experimentos de macroarranjo mostraram que 9 genes putativos, codificadores de fatores de transcrição, apresentaram expressão nos filamentos da corona. Para a caracterização mais completa desses fatores de transcrição, foram realizados experimentos de RT-PCR em 5 destes 9 genes e hibridização in situ com dois dos 5 genes analisados por RT-PCR. Os resultados obtidos com RT-PCR indicaram que estes genes são expressos nos tecidos da corona mas também em folhas e outros órgãos florais. Os resultados da hibridização in situ confirmaram que os genes encontrados não são especificamente expressos na corona.

ABSTRACT

The corona is a floral structure typical from the Passifloraceae family. The elucidation of the mechanisms responsible for the formation of the corona is of great importance to the comprehension of the evolutionary processes that allowed the diversification of the interactions with pollinators among the species of this family. This work aimed the identification and characterization of transcription factors (FT) preferentially expressed in the corona of *Passiflora* spp. For that, the expressed sequence tags (ESTs) database PASSIOMA was investigated and 139 cDNAs were identified: 69 derived from P. edulis flower libraries, 68 from P. suberosa and 2 from P. pohlii. Aiming to attribute identities to the sequences, a parsimony analysis was performed, with the obtaining of cladograms that allowed the identification of members belonging to 31 different transcription factor gene families in *Passiflora* spp. Macroarray experiments indicated that 9 putative genes coding for transcription factors showed expression in corona filaments. Further characterization of these transcription factors included RT-PCR experiments involving 5 out of the 9 genes and in situ hybridization with 2 out of the 5 genes that were previously analyzed with RT-PCR. The RT-PCR results indicated that these genes are expressed in corona tissues, but also in leaves and other floral organs. The in situ hybridization results confirmed that the genes studied are not specifically expressed in the corona.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Características florais do gênero <i>Passiflora</i>	1
1.2. A origem da corona floral em <i>Passiflora</i> spp.	2
1.3. Coevolução entre flores e seus agentes polinizadores	3
1.4. Inovações morfológicas e fatores de transcrição	4
1.5. As principais famílias de fatores de transcrição	5
1.5.1.Famílias ARF e AUX/IAA	7
1.5.2.Família bHLH	8
1.5.3.Família bZIP	10
1.5.4. Família C ₃ H	11
1.5.5.Família HMG	12
1.5.6.Família MADS-box	13
1.5.7.Família MYB	13
1.5.8.Família PHD	15
1.5.9.Família Trihelix	16
2. OBJETIVOS	17
2.1. Objetivo Geral	17
2.2. Objetivos Específicos	17
3. MATERIAL E MÉTODOS	18
3.1. Identificação de Fatores de Transcrição no banco de dados PASSIOMA.	18
3.2. Análises de similaridade de Sequência	19
3.3. Material Vegetal	20
3.4. Macroarranjo	21
3.5. RT-PCR	24
3.6. Hibridização <i>in situ</i>	25
4. RESULTADOS E DISCUSSÃO	28
4.1. Identificação dos Fatores de Transcrição do PASSIOMA	28
4.2. Análises de Similaridade de Sequenci dos FT de <i>Passiflora</i>	29
4.2.1.Famílias ARF e AUX/IAA	31

4.2.2.Família bHLH	
4.2.3.Família bZIP	
$4.2.4.$ Família C_3H	
4.2.5.Família HMG	
4.2.6.Família MADS-box	
4.2.7.Família MYB	
4.2.8.Família PHD	
4.2.9.Família Trihelix	
4.3. Análise da expressão de fatores de transcrição na corona de <i>P. edulis</i> 54	
4.4. RT-PCR	
4.5. Análise da expressão dos genes PACEPE3010A01.g, PACEPE4001G10.g durante desenvolvimento da corona por hibridização <i>in situ</i>	o
5. CONCLUSÃO	
6. REFERÊNCIAS BIBLIOGRÁFICAS	

AGRADECIMENTOS

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), que tornou possível a elaboração desta dissertação de mestrado, pelo apoio financeiro através da concessão de bolsa de estudo, referente ao processo 2011/14596-7.

À Universidade Estadual de Campinas, em especial ao Instituto de Biologia e ao Programa de Pós-Graduação em Biologia Vegetal pela formação acadêmica e pessoal.

Ao Prof. Marcelo C. Dornelas pela oportunidade, orientação e conselhos durante o projeto e pela ajuda não só na pesquisa, mas também na formação acadêmica em sala de aula.

Às professoras da banca examinadora por aceitarem o convite e contribuírem de forma sempre prestativa para o meu aprendizado e enriquecimento do trabalho.

À Prof^a. Fernanda Ramos Gadelha e ao Eduardo de Figueiredo Peloso, pela orientação, ajuda, paciência e amizade no período da iniciação científica, contribuindo muito no início da minha formação acadêmica.

Aos colegas de trabalho do laboratório, em especial à Natália Moraes Silva, pela parceria, amizade, disposição a ajudar e, principalmente, pelo ótimo convívio durante o mestrado e ao longo da graduação.

Aos meus amigos, hoje dispersos em várias cidades, por dividir comigo momentos de risadas, de conquistas, de frustração e de carinho.

Ao Fabilo pelo apoio neste último ano do mestrado e por me manter firme, ajudando a tomar decisões nos momentos de crise e de alegria e a trilhar bons caminhos no futuro.

Aos meus pais Cláudio e Tânia e aos familiares, em especial minha avó Ercília e tias Maria Inês e Sandra pelo incentivo do estudo, apoio em todos os momentos e amor incondicional.

1. INTRODUÇÃO

1.1. Características florais do gênero Passiflora

O gênero *Passiflora* é o maior da família Passifloraceae e possui aproximadamente 600 espécies distribuídas em regiões tropicais e subtropicais (Cervi, 2005), sendo cerca de 200 delas nativas do Brasil (Viana et al., 2010). Este gênero exibe uma ampla variabilidade genética, fator que é responsável por gerar grande variabilidade morfológica, tanto nos órgãos vegetativos como reprodutivos (Ganga et al., 2004).

As flores de *Passiflora* são exemplos da diversidade e complexidade entre as angiospermas, exibindo características peculiares em sua morfologia floral. Em geral, elas são actinomorfas, bissexuais e pentâmeras (Killip, 1938) com dois verticilos distintos de perianto. Os estames e o ovário encontram-se sobre um cilindro central, o androginóforo. O androceu apresenta cinco estames unidos na base e aderidos ao androginóforo na inserção do ovário. O gineceu é composto de três carpelos fundidos e o ovário é unilocular com placentação parietal.

Outras características particulares da flor de Passifloraceae são: o opérculo, situado no interior do tubo floral e acima da câmara nectarífera; o anel nectarífero, situado abaixo do opérculo; e o límen, uma estrutura em forma de anel ou uma membrana em forma de taça que circunda a base do androginóforo. A presença dos filamentos da corona constitui um verticilo adicional, localizado entre o androginóforo e o perianto, no interior do tubo floral.

1.2. A origem da corona floral em Passiflora spp.

A corona representa uma característica marcante do gênero *Passiflora*. No entanto, estudos recentes indicam que os filamentos evoluíram independentemente múltiplas vezes nas angiospermas, estando presentes também em clados como das Superasterideae (Caryophyllales), monocotiledôneas (Asparagales) e Rosideae (Malphigiales) (Hemingway et al., 2011).

Em *Passiflora*, a corona é uma estrutura altamente desenvolvida e diferenciada (Ulmer e MacDougal, 2004). A complexidade desta estrutura pode variar desde uma fina membrana até uma a algumas séries de filamentos, habitualmente bandeados horizontalmente com diversas cores. Os filamentos podem assumir várias formas e tamanhos, ordenados em séries circulares sucessivas, caracterizando cada espécie (Cervi, 1997; Bernhard, 1999).

Foram propostas diferentes hipóteses sobre a origem evolutiva da corona em Passifloraceae. Puri (1948) e Wilde (1974), do ponto de vista anatômico e sistemático, consideraram os filamentos externos como emergências do perianto e os internos como estaminoidais. Endress (1994), numa interpretação evolutiva, considera que os filamentos da corona podem representar estaminódios de um androceu originalmente poliandro com desenvolvimento centrífugo. Bernhard (1999), em seus trabalhos sobre desenvolvimento floral na família Passifloraceae, discorda da origem estaminoidal porque o desenvolvimento da corona é posterior ao dos estames férteis.

Estudos envolvendo biologia molecular, também buscam explicar a origem da corona através da expressão de fatores de transcrição da família MADS-box, os quais determinam a identidade de órgãos florais nas plantas, de acordo com o modelo ABCDE de desenvolvimento reprodutivo (Theissen e Saedler, 2001). Hemingway et al. (2011), através

de RT-PCR e hibridização *in situ*, conseguiram detectar a expressão de genes MADS-box do grupo B (*PcPI* e *PcTM6*) e C (*PcAG*) no desenvolvimento da corona de *Passiflora caerulea*, sugerindo a homologia com estaminódios proposta por Endress (1994). No entanto, a origem da corona ainda não foi completamente estabelecida uma vez que estes autores não relatam os padrões de expressão dos genes dos grupos A e E, necessários para a formação dos multímeros proteicos funcionais para esta família de fatores de transcrição (Theissen e Saedler, 2001).

1.3. Coevolução entre flores e seus agentes polinizadores

Sistema de polinização é um conceito que envolve o conjunto de caracteres florais adaptados para atrair polinizadores adequados e é um exemplo notável de coevolução entre as flores e seus polinizadores (Faegri e Pijl, 1971; Thomson e Wilson, 2008; Preston et al., 2011). A evolução dos sistemas de polinização é um dos fatores mais importantes para explicar o surgimento da grande diversidade de arquiteturas florais em angiospermas. Diversos caracteres reprodutivos relacionados à atração dos polinizadores foram incorporados às estruturas florais durante a evolução deste grupo, aperfeiçoando a reprodução e possibilitando alta taxa de diversificação (Faegri e Pijl, 1971; Crepet, 1983). Assim, características florais como morfologia, cor, odor e a oferta de recompensas delimitam os sistemas de polinização, restringindo ou orientando a acessibilidade do visitante às flores. Igualmente, as diferenças no tamanho do corpo e comportamento de visita são fatores determinantes para a eficiência das diferentes espécies de polinizadores. Desta forma, processos coevolutivos contribuem para o estabelecimento de sistemas de polinização eficientes, onde características florais e características dos polinizadores são selecionados simultaneamente.

Vários estudos sobre a biologia floral e a polinização em *Passiflora* enfatizam o papel da corona na eficiência da polinização (Knuth, 1904; Janzen, 1968; Semir e Brown, 1975; Benson et al., 1975; Sazima e Sazima, 1978, 1989; Ulmer e MacDougal, 2004). Segundo Ulmer e Macdougal (2004), a corona exerce fundamental importância em sua capacidade de selecionar polinizadores: Coronas grandes e vistosas agem como plataformas de pouso e guias de néctar para abelhas e outros insetos, mas por outro lado filamentos curtos ou incipientes são selecionados por morcegos e beija-flores para que não perturbem a aproximação destes animais às flores durante as visitas de polinização.

1.4. Inovações morfológicas e fatores de transcrição

Estudos sobre a evolução do desenvolvimento reprodutivo em plantas têm se concentrado em desvendar as mudanças gênicas responsáveis por modificações morfológicas, passíveis de seleção natural, que interfiram na organização dos órgãos florais e no surgimento de "novas" estruturas florais (Cubas, 2004; Becker, 2011). A regulação da expressão de genes pelos fatores de transcrição controla muitos dos processos biológicos em uma célula ou organismo, tal como a progressão do ciclo celular, equilíbrio metabólico e fisiológico, bem como respostas ao ambiente (Riechmann et al., 2000; McCarty e Chory 2000; Santos et al., 2011; Moreno-Risueno et al., 2012).

Segundo Mondragon-Palomino e Theissen (2008) os processos de duplicação de genes codificadores de fatores de transcrição, seguidos de sub-funcionalização e/ou neofuncionalização são fundamentais para modular variações que permitam a irradiação das espécies com diferentes morfologias adaptativas. Estudos sugerem que ao menos algumas de tais inovações são derivadas de modificações sutis dos padrões de expressão de

genes responsáveis pela identidade dos órgãos florais. Duplicações e diferenças no padrão de expressão de genes codificadores para fatores de transcrição da família MADS-box da função B, por exemplo, resultaram na formação de estruturas inovadoras no perianto ao longo da evolução de alguns grupos de angiospermas, como as orquídeas, por exemplo (Mondragon-Palomino e Theissen, 2008; Theissen et al., 2000; Melzer et al., 2010).

1.5. As principais famílias de fatores de transcrição

Os fatores de transcrição (FT) são proteínas capazes de se ligar em sequências específicas de DNA, ativando ou reprimindo a transcrição de genes. A maioria dos fatores de transcrição é agrupada em famílias de acordo com seu domínio de ligação ao DNA (Airoldi e Davies, 2012; Moreno-Risueno et al., 2012; Pires e Dolan, 2012). Após o sequenciamento do genoma de *Arabidopsis thaliana*, todos os genes vegetais codificadores de fatores de transcrição puderam ser identificados e descritos. Sabe-se, por exemplo, que as três maiores famílias gênicas de fatores de transcrição em *Arabidopsis thaliana* são AP2/EREBP (APETALA2/ethylene responsive element binding protein), MYB-(R1)R2R3 e bHLH, cujos membros somados representam 9% do total de fatores de transcrição conhecidos em plantas (Figura 1 - Riechmann et al., 2000).

São conhecidas cerca de 46 famílias de genes de fatores de transcrição em *Arabidopsis* (Riechmann et al., 2000). Algumas dessas famílias mais importantes e mais representativas para este trabalho serão descritas a seguir.

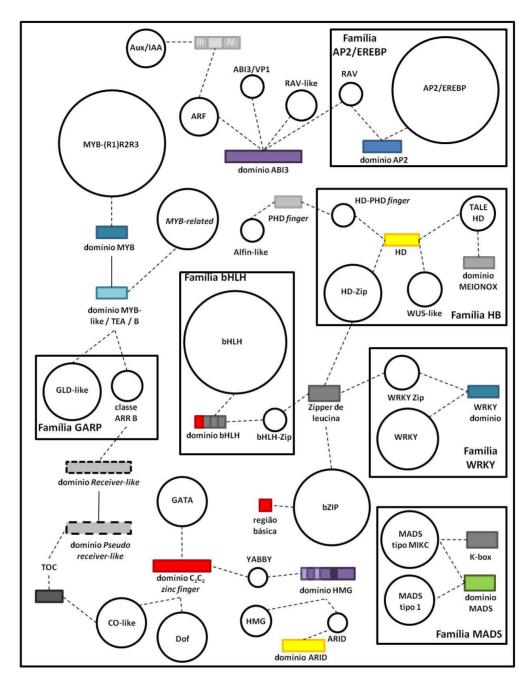


Figura 1: Relações entre diferentes famílias de fatores de transcrição em *Arabidopsis thaliana*. As famílias estão representadas por círculos, cujo tamanho é proporcional ao número de membros na família. Domínios que sofreram misturas e estão conectados a outros grupos são indicados nos retângulos de cores diferentes, cujos tamanhos são proporcionais ao tamanho do domínio. As linhas tracejadas indicam possíveis conexões entre as famílias (adaptado de Riechmann et al., 2000).

1.5.1. Famílias ARF e AUX/IAA

Os membros das famílias AUXIN RESPONSE FACTOR (ARF) e AUX/IAA são genes relativamente bem conhecidos em *Arabidopsis* devido às suas funções na via de sinalização de auxina, em processos de crescimento e desenvolvimento (veja a revisão de Guilfoyle e Hagen, 2007). O controle da expressão dos genes de resposta à auxina depende sempre da ação conjunta de um ARF e um AUX/IAA. Os ARFs funcionam ligados especificamente em sequências de DNA TGTCTC, denominadas elementos de resposta à auxina (AuxREs), nos promotores dos genes de resposta a esse hormônio, ativando-os ou reprimindo-os e recrutando os AUX/IAA (Guilfoyle e Hagen, 2007).

ARFs, como muitos FT, contem domínios modulares que podem funcionar independentes uns dos outros. A maioria dessas proteínas consiste de uma região aminoterminal com um domínio de ligação ao DNA (DBD), uma região mediana de ativação (AD) ou repressão (RD) e uma região carboxi-terminal com um domínio de dimerização (RD) – Figura 2A. O DBD é do tipo B3, semelhante ao encontrado em outras famílias como ABI3/VP1.

Enquanto os membros da família AUX/IAA são geralmente vistos como repressores da expressão de genes responsivos à auxina, os ARF podem ativar ou reprimir a transcrição, dependendo da sequência e estrutura da região mediana da proteína (Guilfoyle e Hagen, 2007). Acredita-se que os AUX/IAA regulam a transcrição induzindo a modificação da atividade dos ARF. Os membros da família AUX/IAA possuem dois motivos de ligação proteína-proteína compartilhados com os ARF, através dos quais as proteínas se ligam (Figura 2A). Uma vez que os FT ARF aparecem normalmente ligados ao DNA em dímeros (ARF-ARF) ou monômeros, a formação do dímero ARF-AUX/IAA pode alterar a conformação anterior e produzir uma regulação diferente (Liscum e Reed, 2002).

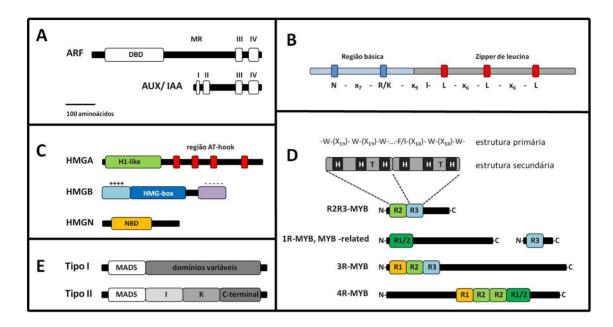


Figura 2: Esquemas dos domínios conservados dos fatores de transcrição de algumas famílias conhecidas em *A. thaliana*. (A) Comparação da estrutura dos genes das famílias ARF e AUX/IAA, mostrando os domínios comuns (III e IV) de ligação heterodimérica (adaptado de Liscum e Reed, 2002). (B) Estrutura primária do domínio bZIP. A região básica esta mostrada em azul e os resíduos conservados são pintados nas caixas azuis e vermelhas (adaptado de Jakoby et al., 2002). (C) Estrutura global das três famílias de proteína HMG e a estrutura do HMG-box (adaptado de Hu et al., 2008). (D) Ilustração mostrando as diferentes classes de proteínas MYB, dependendo do número de repetições do domínio MYB (adaptado de Dubos et al., 2010). (E) Estrutura dos fatores de transcrição MADS-box tipo I e tipo II (adaptado de Masiero et al., 2011).

1.5.2. Família bHLH

As proteínas bHLH (basic helix-loop-helix) são encontradas em animais, fungos e plantas e constituem uma das maiores famílias de fatores de transcrição conhecidas. Muitas dessas proteínas já foram identificadas em plantas, contudo a maioria ainda permanece com função desconhecida (veja a revisão de Carretero-Paulet et al., 2010). bHLHs representam componentes regulatórios chave no controle da rede transcricional de um grande número de processos biológicos. Em plantas, foram descritas atividades no controle da resposta à luz,

sinalização hormonal, ramificação de ramos, desenvolvimento de flores e frutos, micrósporos, tricomas, estômatos e raízes (Carretero-Paulet et al., 2010).

Esse FTs são definidos pelo domínio bHLH, composto de cerca de 60 aminoácidos arranjados em uma estrutura típica. A região N-terminal, de 15-20 resíduos, é normalmente rica em aminoácidos básicos e está envolvida na ligação ao DNA. Alguns aminoácidos conservados nesta região básica determinam o reconhecimento ao chamado núcleo consenso hexanucleotídico E-box, de sequência CANNTG, enquanto que outros resíduos conferem especificidade a um dado tipo de E-box, chamado G-box (CACGTG). A região HLH promove interação proteína-proteína, permitindo a formação de complexos homo- e heterodímeros. Além do domínio bHLH, as proteínas desta família exibem pouca ou nenhuma sequência conservada. Apesar disso, relações funcionais e evolutivas entre as subfamílias podem ser baseadas em padrões de íntrons, motivos de ligação ao DNA e estrutura conservada das proteínas (Carretero-Paulet et al., 2010).

Uma classificação prévia de bHLH de animais levou à definição de 6 grandes linhagens funcionais e evolutivas (grupos A-F). A maioria das proteínas está classificada no grupo A ou B e são as capazes de ligar a sequência do núcleo consenso E-box. Grupo B inclui também membros exibindo motivo de ligação a G-box e proteínas que compartilham um domínio ZIP na porção C-terminal. Fatores de transcrição bHLH do grupo C possuem domínio PAS e se ligam a sequências E-box. As proteínas do grupo E contem um resíduo conservado de prolina ou glicina na região básica. E o grupo D e F é representado por proteínas com a região básica diferente; algumas proteínas do grupo D são incapazes de se ligar ao DNA, funcionando como repressores da atividade de outros bHLH apenas quando forma heterodímeros; e os FT do grupo F incluem as chamadas proteínas COE, nome do domínio que apresentam.

1.5.3. Família bZIP

Os membros da família gênica bZIP foram descritos pela primeira vez em 2002, quando foram descobertos 75 genes que codificavam para fatores de transcrição com a característica dos bZIP. Estudos genéticos e moleculares de alguns desses FT mostraram que eles regulam diversos processos biológicos, como defesa contra patógenos, controle de resposta à luz e estresses, maturação de sementes e desenvolvimento floral (Jakoby et al., 2002).

Essas proteínas caracterizam-se por possuírem o domínio conservado denominado bZIP, que contem duas estruturas essenciais: primeiro, uma região básica de aproximadamente 16 resíduos de aminoácidos com um sinal de localização nuclear, seguida de um motivo conservado N-x7-R/K responsável pela ligação ao DNA; segundo, uma repetição de 7 leucinas, ou outro aminoácido hidrofóbico, localizadas exatamente 9 resíduos em direção à porção C-terminal, criando uma hélice anfipática (Figura 2B). A habilidade de formar homo- e heterodímeros é influenciada pela atração e repulsão eletrostática de resíduos polares flanqueando a superfície de interação hidrofóbica das hélices. As proteínas bZIP de plantas se ligam a sequências de DNA com um núcleo ACGT, preferencialmente a A-box (TACGTA), C-box (GACGTC) e G-box (CACGTG) (Jakoby et al. 2002).

1.5.4. Família C₃H

Os fatores de transcrição *zinc-finger* são muito importantes em muitos processos biológicos como morfogênese, transdução de sinal e resposta a estresses ambientais (Peng et al., 2012). Os domínios *zinc finger*, contidos na família CCCH (ou C₃H), são classificados de acordo com o arranjo dos aminoácidos ligados ao zinco e estão presentes em muitos fatores de transcrição, desempenhando papéis importantes nas interações com outras moléculas. Em plantas, outras famílias de fatores de transcrição contendo *zinc fingers* são RING-finger, C₂H₂, C₂C₂, ERF, WRKY, Dof e LIM (Wang et al., 2008).

As proteínas da família C₃H, no entanto, diferentemente dos membros das famílias irmãs C₂H₂ e C₂C₂, possuem 3 resíduos de cisteína e um resíduo de histidina em seu domínio *zinc finger*. O domínio CCCH de ligação a ácidos nucleicos é muito frequente e foi identificado inicialmente em animais e fungos. Ele consiste de um motivo C-X-C-X-H e, baseado nos diferentes números de aminoácidos entre Cys (C) e His (H), é possível dividir a família em 18 grupos (Wang et al., 2008).

Recentemente, foi descoberto que os FT C₃H, diferiam dos *zinc-finger* de *Arabidopsis* já conhecidos e identificados porque regulam a expressão gênica através da ligação com RNA e não apenas com DNA. Há muitas evidências de que o domínio CCCH pode funcionar do processamento de moléculas de RNA. Poucos FT desta família são caracterizados funcionalmente em plantas. Em *A. thaliana* HUA1 é uma proteína que se liga ao RNA e participa de mecanismos regulatórios do desenvolvimento floral (Wang et al., 2008). A proteína PEI1 é uma CCCH embrião-específica, responsável por atuar durante a embriogênese de *Arabidopsis*, expressando-se na região apical do embrião (Wang et al., 2008).

1.5.5. Família HMG

As proteínas HMG (high mobility group) constituem a segunda mais abundante família de proteínas presentes na estrutura física dos cromossomos. Comparadas com as histonas, elas ocorrem em quantidades variáveis e se ligam a cerca de 10% dos nucleossomos. Baseando-se em sua estrutura primária, as proteínas HMG de mamíferos foram divididas em 3 subfamílias: HMGA, contendo motivo *AT-hook* de ligação ao DNA; HMGB, contendo um HMG-box; e HMGN, contendo um domínio de ligação a nucleossomos (Figura 2C). Proteínas HMGA e HMGB já foram identificadas e caracterizadas em plantas, no entanto, as HMGN parecem ser exclusivas dos vertebrados. As proteínas HMGA são caracterizadas pela presença de um domínio N-terminal com sequência similar ao domínio globular de ligação da histona H1. Em colaboração com fatores de transcrição, elas contribuem para a regulação correta de genes alvos (veja a revisão de Grasser et al. 2007).

Enquanto no genoma dos vegetais há geralmente um ou dois genes codificando HMGAs, dependendo da espécie, para as proteínas HMGB há diversos genes. Elas se ligam ao DNA através do motivo HMG-box, induzindo mudança estrutural da cromatina. Além das propriedades de ligação com a cromatina, essas proteínas também interagem com outras, incluindo histonas e fatores de transcrição e participando da regulação de transcrição de múltiplos processos (Grasser et al. 2007).

Deste modo, nas proteínas HMGB de plantas, o domínio central de ligação ao DNA HMG-box é flanqueado por um domínio básico N-terminal e um domínio ácido C-terminal. O domínio HMG-box contem cerca de 75 resíduos de aminoácidos, com uma dobra conservada em forma de L formada por 3 α-hélices (Grasser et al. 2007).

1.5.6. Família MADS-box

Em *Arabidopsis*, alguns dos 107 membros da família MADS-box são conhecidos por estarem envolvidos predominantemente em processos do controle do desenvolvimento vegetal. Os fatores de transcrição MADS mais bem estudados são aqueles envolvidos na determinação da identidade dos órgãos florais. A análise do estudo de mutantes para genes homeóticos resultou na formulação de um modelo genético, o modelo ABC, que explica como funções combinadas de três classes de genes MADS (A, B e C) determinam a identidade dos quatro verticilos florais em *Arabidopsis* (Parenicová et al. 2003).

Análises detalhadas das proteínas MADS-box são mais restritas ao clado MIKC. Da porção N-terminal até a C-terminal podem ser identificados os domínios: MADS-box (M), de intervenção (I), *keratin-like* (K) e o C-terminal (C). O domínio MADS possui 58 aminoácidos e é o responsável pela ligação ao DNA, reconhecendo sequências especificas denominadas CArG boxes [CC (A/T)₆GG]. O domínio I é menos conservado e contribui para a especificação da dimerização. O domínio K é caracterizado por uma estrutura *coiled-coil* que facilita a dimerização das proteínas MADS-box. E o domínio C é o menos conservado deles e contribui para a formação de complexos multiméricos (Theissen e Saedler, 2001; Parenicová et al., 2003).

1.5.7. Família MYB

A família de proteínas MYB é grande, funcionalmente diversa e presente em todos os eucariotos. É constituída de fatores de transcrição com um número variado de repetições de domínios MYB de ligação ao DNA. Em plantas, esta família foi expandida, principalmente através da grande subfamília R2R3-MYB. Esses FT são responsáveis por uma variedade

enorme de processos vegetais, funcionando como fatores fundamentais em redes regulatórias que controlam desenvolvimento, metabolismo e resposta a estresses bióticos e abióticos (Dubos et al., 2010).

O domínio MYB geralmente consiste de um máximo de 4 sequências repetidas e imperfeitas de cerca de 52 aminoácidos, cada uma formando 3 α-hélices. A segunda e a terceira hélices de cada repetição formam uma estrutura HTH (helix-turn-helix) com 3 triptofanos regularmente espaçados. A terceira hélice de cada repetição é chamada de "hélice de reconhecimento", a qual faz contato direto com DNA e intercala no sulco maior da molécula (Dubos et al., 2010).

Esses fatores de transcrição podem ser divididos em diferentes classes, dependendo do número de repetições de domínios MYB presentes na proteína (1, 2, 3 ou 4), todas as quatro classes são encontradas em plantas. As três repetições de uma proteína MYB são denominados R1, R2 e R3. A menor classe é o grupo 4R-MYB, com apenas um gene no genoma de *Arabidopsis*, e possui 4 repetições R1/R2-like. A segunda classe contém proteínas R1R2R3-type MYB (3R-MYB), tipicamente codificadas por 5 genes no genoma de *Arabidopsis* e relacionadas com controle do ciclo celular. A terceira classe compreende proteínas com uma repetição ou uma parte dela, denominadas proteínas "MYB-related" e são divididas em várias subclasses (Riechmann et al. 2000). Mais distante, mas também relacionado com a superfamília MYB está o grupo denominado "GARP". Estas proteínas parecem estar envolvidas em processos específicos de plantas: a família de FT KANADI pertence a este grupo e interage com os FT do grupo ARF para mediar a sinalização de auxinas e a formação do padrão abaxial-adaxial do embrião (Izhaki e Bowman, 2007; Kelley et al., 2012).

A maioria dos genes MYB vegetais, no entanto, codifica proteínas da classe R2R3-MYB, as quais têm uma estrutura modular, com um domínio N-terminal MYB de ligação ao DNA e um domínio C-terminal de ativação ou repressão. Ao contrário do domínio MYB conservado, as outras regiões das proteínas R2R3-MYB são altamente variáveis (Figura 2D – Dubos et al., 2010). A organização dos domínios R2R3 das proteínas MYB específicas de plantas provavelmente evoluiu a partir de um gene tipo R1R2R3 ancestral cuja primeira repetição foi perdida (Kranz et al., 2000).

Uma vez que as proteínas vegetais do tipo MYB-R1R2R3 estão mais estreitamente relacionadas com as proteínas MYB de origem animal do que com as proteínas do tipo R2R3 de plantas, foi sugerido que elas possam ter funções relacionadas com os das proteínas MYB de animais, tais como o controle da proliferação celular (Cominelli e Tonelli, 2009). Inversamente, as proteínas MYB-R2R3 podem ter evoluído para regular os processos específicos de plantas, incluindo o metabolismo secundário, as respostas a hormônios vegetais, bem como a identidade de tipos de células específicas de plantas, como os estômatos e pelos radiculares (Tominaga-Wada et al., 2012).

1.5.8. Família PHD

A família PHD contém um *zinc finger* modificado (do tipo Cys₄–His–Cys₃ *zinc finger*). Baseado na posição do resíduo único de histidina, o esqueleto de cisteína do domínio PHD é claramente diferente de outros domínios *zinc finger* como RING ou LIM. Trabalhos demonstraram que as proteínas PHD estão associadas com complexos de remodelamento da cromatina ou contribuem para acetilação de histonas (Halbach et al., 2000). Assim, o domínio *zinc finger* modificado encontrado em proteínas PHD também foi

observado em outros fatores de transcrição de plantas e animais, associados com a regulação transcricional mediada pela cromatina. Adicionalmente ao domínio PHD, as proteínas desta família também possuem um homeodomínio e um domínio zíper de leucina (Halbach et al., 2000).

1.5.9. Família Trihelix

Os fatores de transcrição desta família são conhecidos por possuírem um domínio conservado de ligação ao DNA contendo três hélices (helix-loop-helix-loop-helix) e que se liga especificamente à sequência G-R-(T/A)-A-A-(T/A), também denominada "elemento GT", um componente em *cis* que confere respostas à luz. Assim, membros desta família também são conhecidos como fatores GT (Zhou, 1999). Os estudos iniciais sobre os FT desta família mostraram seu papel na regulação de genes responsivos à luz. Entretanto, outros trabalhos indicam que proteínas Trihelix também desempenham papel importante em diferentes processos de crescimento e desenvolvimento de flores, estômatos, tricomas, embriões e sementes (Barr et al., 2012).

2. OBJETIVOS

2.1. Objetivo Geral

O objetivo geral do projeto foi identificar fatores de transcrição preferencialmente expressos nos filamentos da corona de *Passiflora* spp. Com este intuito, objetivos específicos foram estabelecidos:

2.2. Objetivos Específicos

- A. Identificar, na biblioteca de cDNAs PASSIOMA, sequências potencialmente codificadoras de fatores de transcrição. Caracterizar, por meio de análise de similaridade de sequência, as identidades destes cDNAs, atribuindo-os às famílias conhecidas de fatores de transcrição;
- B. Identificar, entre os genes putativos identificados em A, os expressos em tecidos da corona, através de experimentos de macroarranjo.
- C. Caracterizar, com o uso de RT-PCR, o padrão de expressão das sequências identificadas em B, em perianto, corona, androginóforo, estames e pistilo de *Passiflora edulis*;
- D. Caracterizar, com o uso de hibridização in situ, o padrão de expressão das sequências identificadas em B nos tecidos de botões florais em diferentes estágios de desenvolvimento.

3. MATERIAL E MÉTODOS

3.1. Identificação de Fatores de Transcrição no banco de dados PASSIOMA

A identificação dos possíveis FT expressos durante o desenvolvimento reprodutivo em Passiflora foi feita através de busca no banco de dados de etiquetas de genes expressos (ESTs) PASSIOMA (Dornelas et al., 2006; Cutri e Dornelas, 2012). As buscas foram realizadas por palavras-chave, usando o nome das famílias de FT atualmente conhecidas em Arabidopsis thaliana. Para tal, foram utilizadas informações presentes no banco de dados The Arabidopsis Information Resource (TAIR – http://www.Arabidopsis.org), onde foi possível encontrar o nome das principais famílias gênicas conhecidas em vegetais; e também foi utilizado o trabalho de Riechmann et al. (2000) como guia para investigação dos principais domínios de ligação ao DNA que caracterizam os FT vegetais. Adicionalmente realizou-se uma busca por similaridade de sequência, com o uso da ferramenta tBLASTx local do banco de dados PASSIOMA, utilizando como sequência-isca (query) membros representativos de cada família gênica ou o produto consensual de alinhamentos de sequências de FT de A. thaliana (Markov chains). Uma verificação cruzada confirmatória da identificação das sequências encontradas no PASSIOMA foi realizada pelo alinhamento destas com as depositadas em bancos de dados públicos (e.g. NCBI nr, nt, e dbEST; Benson et al., 2002), com o emprego do algoritmo BLASTx (Altschul et al. 1990; http://www.ncbi.nlm.nih.gov/BLAST/).

3.2. Análises de similaridade de Sequência

Com o intuito de determinar as relações de similaridade entre as sequências encontradas no banco de dados e de confirmar se as sequências encontradas pertenciam legitimamente às famílias de FT às quais foram atribuídas inicialmente, foram construídos cladogramas com sequências de aminoácidos dos FT encontrados no banco de dados do PASSIOMA, das espécies *Passiflora edulis, Passiflora suberosa, Passiflora pohli* e dos respectivos membros das famílias correspondentes em *Arabidopsis thaliana*.

Para a análise comparativa, as sequências de aminoácidos de todos os membros de cada família de FT de *A. thaliana* foram obtidas através de bancos de dados públicos (NCBI: www.ncbi.gov.br; e TAIR: http://www.*Arabidopsis.*org/). Com o uso do software ClustalX (Thompson et al., 1994), as sequências de aminoácidos traduzidas das sequências de nucleotídeos dos cDNA encontradas no PASSIOMA (pelo uso do software BioEdit) foram alinhadas às sequências de *Arabidopsis*. Os alinhamentos obtidos foram corrigidos manualmente e utilizados para a obtenção de matrizes de similaridade por parcimônia e cálculos de Bootstrap com 1000 repetições, com o uso do software *The Molecular Genetics Analysis* (MEGA) versão 4.0 (Kumar et al., 2009). As topologias finais das árvores foram obtidas pelo método *Neighbor-joining* (Saitou e Nei, 1987).

Utilizou-se a ferramenta digital "*The Bio-Analytic Resource (BAR) for Plant Biology*" (Winter et al., 2007) com o intuito de obter informações sobre os padrões de expressão conhecidos dos genes codificadores de FT em *Arabidopsis*.

3.3. Material Vegetal

3.3.1. Material para RT-PCR

Órgãos florais de flores em antese de *P. edulis* foram coletados visando à extração de RNAs totais. Conjuntos de órgãos florais (perianto, corona, coluna do androginóforo, estames e pistilo) foram rapidamente separados em diferentes sacos de papel alumínio e imediatamente congelados em N₂ líquido.

3.3.2. Material para Hibridização in situ

Para os experimentos de hibridização in situ foram coletados botões florais de P. edulis em diferentes estágios de desenvolvimento. Durante a coleta, estes botões foram classificados de acordo com o tamanho e a pigmentação da corona em três classes: A, B e C. Os botões de 0,3 a 0,5 cm de comprimento foram classificados como "tipo A". Nestes botões a corona ainda não estava formada (estágio 1) ou estava em estágios iniciais de formação dos primórdios dos filamentos da corona e do opérculo (estágio 2). Os botões com 0,6 a 1,1 cm de comprimento foram classificados como "tipo B". Dentre estes botões, alguns apresentavam a diferenciação apenas da primeira (mais externa) série de filamentos (estágio 3) ou de ambas as séries de filamentos (estágio 4). Tanto no estágio 3 como no 4, os filamentos da corona ainda não apresentam pigmentação visível a olho nu. Os botões com 1,1 a 2 cm de comprimento foram classificados como "tipo C". Estes botões apresentaram filamentos da corona bem desenvolvidos e com pigmentação (estágio 5). Os botões coletados foram fixados em parafomaldeído a 4% e, posteriormente, submetidos a vácuo até o dia seguinte. A seguir, os botões foram desidratados em uma série etanólica (de 30% a etanol absoluto), com trocas de solução a cada 4h. O material foi mantido em etanol absoluto, na geladeira, até o momento do uso.

3.4. Macroarranjo

3.4.1. Confecção da sonda marcada com DIG

Para o experimento de macroarranjo utilizou-se como sonda a população de RNAs mensageiros de corona marcada com digoxigenina. Para tal, RNAs totais foram extraídos de 50 a 100mg de filamentos da corona de flores na antese, seguindo o método do Trizol® (Invitrogen). O RNA total obtido foi ressuspendido em 30 µl de água MILI-Q tratada com dietil pirocarbonato (DEPC). Uma alíquota de 3µl da solução de RNA foi separada por eletroforese em gel de agarose a 1% em TAE (Tris-acetato a 40 mM, EDTA a 1mM) contendo 1,25 µg/mL de brometo de etídeo para verificação da qualidade. As amostras assim avaliadas foram quantificadas utilizando-se Qubit® (Invitrogen) de acordo com as instruções do fornecedor. As amostras foram diluídas com água-DEPC de modo a conterem 5 μg de RNA em 8 μL de solução e tratadas com DNase® (AMBION), de acordo com o protocolo do fabricante. Procedeu-se a síntese da primeira fita do cDNA utilizando-se o kit "Super Script III First Strand Synthase" (Invitrogen), de acordo com as instruções do fabricante, utilizando-se um primer poli-T contendo a sequência do promotor T7 (5' -GGGATATCACTCAGCATAAT₍₁₇₎ – 3'). Cinco microlitros das amostras de cDNA assim obtidas foram diluídos em 45 µL de água MILI-Q autoclavada. Para testar a qualidade dos cDNAs sintetizados e para a normalização das concentrações das amostras, realizou-se uma reação de PCR com primers que amplificam um fragmento de 453 pb do gene constitutivo PeACTINA (Tabela 1), sob as seguintes condições: desnaturação inicial a 95°C por 3 min, seguida de 32 ciclos de: 95°C por 30 s, 60°C por 30 s e 72°C por 1 min e um ciclo de extensão final a 72°C por 5 min.

Utilizando o cDNA de fita simples sintetizado, foi feita a síntese da segunda fita de DNA utilizando parte do kit "CloneMinerTM II cDNA Library Construction" (Invitrogen). O cDNA dupla fita assim obtido foi utilizado como molde para a transcrição *in vitro* da sonda de RNA marcada com digoxigenina (DIG) com o uso da enzima T7 e segundo as instruções do kit "DIG RNA Labeling" (Roche).

3.4.2. Confecção da membrana de macroarranjo

Após a seleção das sequências codificadoras para os fatores de transcrição de Passiflora encontrados no PASSIOMA, clones de cDNA correspondentes às sequências de P. edulis foram utilizados para a montagem da membrana de macroarranjo. Os clones de cDNA contidos em vetores pSPORT1 e armazenados em freezer -80°C, foram transformados em bactérias quimiocompetentes E. coli TOP10 (Invitrogen), que foram transferidas individualmente para poços de placas de 96 posições e crescidas em meio LB contendo 100mg/mL de ampicilina, sob agitação (250 rpm) a 37°C por 18h. Após esse período, as células foram sedimentadas por centrifugação (6 min a 3000 rpm) e ressuspendidas em 240 µL de GTE (Glicose a 20%; Tris a 1M e pH 7,4; EDTA a 0,5M e pH 8,0). Após nova centrifugação (6 min a 4000 rpm), descartou-se o sobrenadante com adição de 62,5 µL de GTE com RNAse A (0,8 mg/mL) em cada amostra. Depois de 5min em agitação em vórtex, foram adicionadas em cada amostra 60 µL de solução de NaOH a 0,2M e SDS a 1%, com um período de incubação de 10 min à temperatura ambiente. Adicionou-se, então, 60 µL de KOAc a 3M e a placa foi incubada a 90°C por 30 min. Seguiu-se o resfriamento da placa em gelo por 10 min e centrifugação por 6 min a 4000 rpm (4°C). O volume total do sobrenadante foi transferido para uma placa filtro (MultiScreen HTS, 22µm) e centrifugado por 4 min a 4000 rpm (20°C). Posteriormente, adicionou-se ao filtrado 90 μ L de isopropanol, e procedeu-se a uma centrifugação por 45 min a 4000 rpm (4°C). O sobrenadante foi descartado e 160 μ L de etanol a 70% gelado foram adicionados às amostras, seguindo-se uma centrifugação por 10 min a 4000 rpm (4°C). Mais uma vez o sobrenadante foi descartado e, após a secagem das placas, o DNA plasmidial foi ressuspendido em 30 μ L de água MilliQ. Para confirmação da extração, as amostras de DNA plasmidial foram separadas por eletroforese em gel de agarose a 1% conforme descrito no item 3.4.1.

Após a extração do DNA plasmidial, foi feita a separação da dupla fita do plasmídeo por aquecimento da placa a 96°C por 5 min e então resfriando no gelo por 10 min. Com o auxilio de um repicador de 96 pinos, alíquotas do DNA plasmidial extraído foram transferidas para uma membrana positivamente carregada (Hybond-N+ Amersham), de forma que cada amostra estivesse presente em quadruplicata na membrana. A segunda amostra foi posicionada à direita da primeira, a terceira réplica, abaixo da primeira e uma quarta réplica foi posicionada abaixo da segunda. As últimas posições em baixo à direita da placa de 96 poços foram ocupadas por: um poço contendo água (controle negativo), e três poços contendo o cDNA utilizado como molde para a síntese da sonda (controle positivo) em três diluições: 1:10, 1:50 e 1:100, respectivamente. A membrana contendo o macroarranjo foi mantida em estufa a 120°C por 30 min para fixação.

3.4.3. Hibridização do macroarranjo

A membrana contendo o macroarranjo foi incubada overnight em estufa a 42°C com 10mL de solução de hibridização (Tris-HCl a 10mM pH7,5; NaCl a 300mM; formamida deionizada a 50%; EDTA a 1mM; 1x Denharts; Dextran Sulfato a 10% e 600 ng de sonda marcada). Após a hibridização, o excesso de sonda não-hibridizada foi retirado com duas

lavagens em solução salina SSC quatro vezes concentrada (NaCl a 0,15M e citrato de sódio a 0,015M) por 30 min cada e duas lavagens em solução salina SSC duas vezes concentrada, também por 30 min cada. Todas as lavagens foram realizadas à temperatura de 42°C.

Para o processo de imunodetecção, a membrana foi enxaguada brevemente com a Solução de Lavagem (Tris-HCl a 0,1M; NaCl a 0,15M e pH 7,5) e em seguida incubada em temperatura ambiente por 30 min em Solução de Bloqueio (1% de *Blocking Reagent* Roche dissolvido em Solução de Lavagem). Adicionou-se então o anticorpo anti-digoxigenina conjugado à fosfatase alcalina (*Fab-Fragments*, Roche) diluído 1:5000 em Solução de Bloqueio. A solução contendo anticorpo foi mantida em contato com a membrana por 30 min e então, descartada. A membrana foi lavada em Solução de Detecção (Tris-HCl 0,1M; NaCl 0,1M; pH 9,5) por 5min, após os quais a solução foi descartada. A seguir a membrana foi tratada com 10 mL de substrato que consistiu em uma solução comercial (Pierce) de NBT (Nitro-Blue Tetrazolium Chloride) / BCIP(5-Bromo-4-Chloro-3'-Indolyphosphatase p-Toluidine) contendo Levamissole (1mM). A reação de detecção aconteceu em temperatura ambiente e no escuro por cerca de 16h.

3.5. RT-PCR

RNAs totais dos diferentes órgãos florais e de folha de *P. edulis* foram isolados pelo método do Trizol® (Invitrogen), seguido de tratamento com DNAse (AMBION) e a síntese do cDNA utilizando-se o kit "Super Script First Strand Synthase" (Invitrogen), como descrito no item 3.4.1. O programa utilizado para a construção dos primers geneespecíficos foi o Primer 3 Input Program (http://frodo.wi.mit.edu/primer3) (Rozen e Skaletsky, 1998). Os pares de primers utilizados estão listados na Tabela 1. Em geral, as

condições de amplificação foram as mesmas utilizadas para amplificação da *PeACTINA* com exceção das temperaturas de anelamento, também listadas na Tabela 1.

Os produtos de amplificação foram separados em gel de agarose 1% em TAE (40 mM Tris-acetato, 1 mM EDTA) contendo 1,25 µg/mL de brometo de etídeo. As bandas foram visualizadas em trans-iluminador de UV e fotodocumentada pelo Gel Doc 2000 *BioRad*®, com o software Quantity One.

Tabela 1: Pares de iniciadores (primers) utilizados para a RT-PCR dos genes codificadores de fatores de transcrição e do gene constitutivo *PeACTINA*.

Nome do primer	Sequência	Tamanho do fragmento
rome do primer	Sequencia	amplificado (pb)
PACEPE3005E04_F	5'- CAATCGATGGATGCAACAGT -3'	178
PACEPE3005E04_R	5'- TTGGGGAATGAGAAATGTCC -3'	170
PACEPE3010A01_F	5'- CGCAAGAATTGAAGGCAACAGA -3'	150
PACEPE3010A01_R	5'- ACAATATCCCAAGCCCACAGAG -3'	
PACEPE3008D10_F	5'- ACGAAAGGCTAAGCAGGAGA -3'	150
PACEPE3008D10_R	5'- GCTTTGCCCACAGCAGAAA -3'	
PACEPE3018E06_F	5'- GCTACTTCCCCTTTGCACAC -3'	145
PACEPE3018E06_R	5'- GTCTTGGGAAGAGGGTTTTGA -3'	110
PACEPE4001G10_F	5'- GAGAATGCTTCCTCCACGTC -3'	207
PACEPE4001G10_R	5'- GCCTTCCCATTCTCCAACTT -3'	
PeACTINA_F	5'- TCTTGATCTTGCTGGTCGTG -3'	453
PeACTINA_R	5'- TTGATCTTCATGCTGCTTGG -3'	

3.6. Hibridização in situ

As hibridizações *in situ* foram realizadas utilizando-se sondas não-radioativas, segundo as técnicas descritas na literatura (Dornelas et al., 1999; 2000). Como molde para as sondas de mRNA foram utilizados fragmentos de cDNAs obtidos da biblioteca do

projeto PASSIOMA. Os fragmentos utilizados encontram-se clonados no vetor pSPORT1 (Invitrogen) e foram obtidos da mesma forma que descrita no item 3.4.2.

Para a confecção das sondas, as culturas de bactérias contendo os respectivos insertos de cDNA clonados no vetor foram reativadas em 3ml de meio LB líquido suplementado com ampicilina (100 mg/L), sob agitação (250 rpm) a 37°C por 16h. Os plasmídeos contendo os clones de cDNAs foram isolados com o kit Pure Link Quick Plasmid Miniprep (Invitrogen). Os plasmídeos assim obtidos linearizados com a enzima EcoR1 a 37°C por 2 horas. Em seguida, realizou-se a precipitação dos plasmídeos linearizados (10% de acetato de sódio 3M e duas vezes e meia o volume de etanol absoluto) por 20 min a 12.000 rpm e ressuspendidos em 10 μl de água tratada com dietil pirocarbonato (DEPC). Os plasmídeos assim obtidos foram dosados utilizando-se Qubit® (Invitrogen) de acordo com as instruções do fornecedor. Um micrograma de plasmídeos linearizados foi utilizado como molde para a síntese de RNA *in vitro*. As sondas antisense foram sintetizadas por transcrição *in vitro* pela enzima SP6 em reação contendo uracila marcada com digoxigenina (DIG-UTP), utilizando o kit *DIG RNA Labeling* (Roche) de acordo com as instruções do fabricante.

Os botões florais fixados e desidratados conforme descrito no item 3.3.2 foram tratados com misturas de álcool: xilol (3:1, 1:1, 1:3 e xilol puro), e posteriormente emblocados em parafina, seccionados em micrótomo (10 µm) e montados em lâminas de vidro pré-tratadas com solução de 3-Aminopropiltrietoxisilana (2%, Pierce) em acetona. Os cortes de material emblocado foram desparafinados com dois banhos de xilol puro de 15 min cada, seguidos por banhos com misturas de xilol: etanol nas proporções 3:1; 1:1 e 1:3 (15 min cada banho), finalizando com duas imersões em etanol absoluto (10 min cada). Após estes tratamentos as lâminas foram secas ao ar.

As amostras foram submetidas ao tratamento de pré-hibridização com proteinase-K (1 µg/mL em Tris-HCl 0.05M; pH7.5) por 12 min a 37°C e lavadas a seguir com água tratada com dietil pirocarbonato (DEPC). Foram adicionados 300 µL de solução de hibridização em cada lâmina (Tris-HCl a 10mM e pH7,5; NaCl a 300mM; formamida deionizada a 50%; EDTA a 1mM; 1x Denharts; Dextran Sulfato a 10% e 600 ng de sonda marcada), cobertas com filme plástico e incubadas em câmara úmida a 42°C por cerca de 16h. Após a hibridização, o excesso de sonda não-hibridizada foi retirado com duas lavagens em solução salina SSC quatro vezes concentrada (NaCl a 0,15M e citrato de sódio a 0,015M) por 30 min cada e duas lavagens em solução salina SSC duas vezes concentrada, também por 30 min cada. Todas as lavagens foram realizadas à temperatura de 42°C. Para o processo de imuno-detecção, as lâminas foram inicialmente mantidas em cubetas em tampão de detecção I (DB1: Tris-HCl a 1M e NaCl a 5M e pH7,5) por 5 min. Em seguida, foram submetidas por 30 min ao tratamento em tampão de bloqueio DB2 (1% Blocking reagent Roche dissolvido em DB1). As lâminas foram incubadas com anticorpo antidigoxigenina conjugado à fosfatase alcalina (Fab-Fragments, Roche) diluído 1:1000 em tampão DB2 a 37°C por uma hora em câmara úmida. Após esse período, foram realizadas duas lavagens seguidas em DB1 de 15 min cada e um banho em DB3 (Tris-HCl a 0,1M; NaCl a 0,1M; MgCl₂. 6H2O a 0,05M e pH9,5) por 5 min. Para finalizar, aplicou-se 300 μL da solução comercial de NBT (Nitro-Blue Tetrazolium Chloride) / BCIP (5-Bromo-4-Chloro-3'-Indolyphosphatase p-Toluidine) contendo o supressor (Levamissole a 1mM) como substrato (Pierce). A reação de detecção aconteceu em temperatura ambiente em câmara úmida e no escuro por 16h. O material hibridizado foi observado e fotodocumentado em microscópio ZEISS modelo AXIOCAM HRc para posterior análise.

4. RESULTADOS E DISCUSSÃO

4.1. Identificação dos Fatores de Transcrição do PASSIOMA

Com o objetivo de identificar e caracterizar fatores de transcrição (FT) preferencialmente expressos na corona de *P. edulis*, utilizou-se a biblioteca de cDNAs gerada pelo projeto PASSIOMA.

Para execução deste trabalho algumas premissas e hipóteses foram levantadas. (1) Entre as sequências contidas nas bibliotecas do PASSIOMA há genes expressos preferencialmente nos tecidos de corona. (2) Os genes responsáveis pela ontogênese da corona são conservados em diferentes espécies de *Passiflora*. (3) Os genes responsáveis pela ontogênese da corona são encontrados em outras espécies de plantas, mas nestas foram recrutados para outras funções.

Para encontrar os FT preferencialmente expressos na corona de *P. edulis*, primeiramente foi feita uma busca por palavras-chave dos nomes das famílias de FT atualmente conhecidas em *A. thaliana*. As famílias gênicas foram selecionadas levando em consideração a premissa (3), de duas maneiras: através do banco de dados *The Arabidopsis Information Resource* (TAIR – http://www.*Arabidopsis*.org); e pelo trabalho de Riechmann et al. (2000) para investigação dos principais domínios de ligação ao DNA que caracterizam os FT. Além da busca por palavra-chave, também foi realizada uma procura utilizando o algoritmo tBLASTx.

As famílias de FT selecionadas são mostradas na Tabela 2, juntamente com o número de genes encontrados no PASSIOMA para cada família. Foram encontrados no total 191 reads que possivelmente codificariam FT. Desses 81,3% (113 reads) corresponderam a singletons, ao passo que o restante 18,7% (78 reads) estão incluídos dentro de 26 clusters,

que formaram 26 contigs. Desta forma, foram consideradas 139 sequências de cDNA presentes no PASSIOMA que putativamente são codificadoras de FT. Esses dados refletem a baixa redundância deste banco de dados, uma vez que menos clusters foram encontrados em relação aos singletons. Esperava-se encontrar um numero de reads próximo ao observado, uma vez que em genomas vegetais cerca de 1% dos genes codificam fatores de transcrição (Riechmann et al. 2000), e há 10.272 sequências de alta qualidade no PASSIOMA (Cutri e Dornelas, 2012), ou seja, 139 corresponderia a cerca de 1 a 2% do número total de sequências presentes no banco de dados. As famílias que possuem o maior número de membros em plantas-modelo forneceram o maior número de sequências no banco de dados PASSIOMA, como bHLH, C₂H₂, C₃H, Homeobox, MADS e MYB.

Dos 139 cDNAs putativos codificadores de FT encontrados, 49,6% (69 cDNAs) pertenciam a *P. edulis*, 48,9% (68 cDNAs) pertenciam a *P. suberosa* e 1,4% (2 cDNAs) foram de *P. pohlii*. O que também está de acordo com as proporções totais de sequências depositadas no PASSIOMA para cada espécie. Apenas as sequências de *P. edulis* foram utilizadas nos experimentos de análise de expressão.

4.2. Análises de Similaridade de Sequência dos FT de Passiflora

Foram elaborados cladogramas utilizando sequências de aminoácidos dos FT encontrados juntamente com a dos FT de *A. thaliana*. Foram construídas árvores para todas as 31 famílias de fatores de transcrição encontradas. Entretanto, aqui só serão mostradas as famílias mais importantes e representativas deste trabalho.

Tabela 2: Famílias de FT selecionadas através do TAIR e do trabalho de Riechamnn et al. (2000),

assim como o número de cDNAs putativos codificadores de FT no PASSIOMA.

Família de FT em A. thaliana	N° de genes encontrados no PASSIOMA	Famílias de FT em A. thaliana	N° de genes encontrados no PASSIOMA
ABI3VP1	1	GRAS	3
ALFIN-LIKE	1	GRF	0
ARF	3	Heat Shock Fators (HSF)	1
ARID	0	HMG	6
AP2-EREBP	3	Homeobox	8
AtRKD	0	HRT	0
AUX/IAA	9	JUMONJI	0
BBR/BPC	0	MADS	7
ЬНЦН	11	МҮВ	14
bZIP	4	NAC	4
BZR	0	NLP	0
C ₂ C ₂ -CONSTANS-like	9	Orphan	0
C ₂ C ₂ -Dof	5	PHD	5
C ₂ C ₂ -Gata	3	RAV	0
C ₂ C ₂ -YABBY	3	REM	0
C_2H_2	10	TCP	1
СЗН	8	Trihelix	7
CAMTA	0	TUB	1
CCAAT-box (DR1/HAP2-3-5)	3	VOZ-9	2
СРР	1	Whirly	2
E2F-DP	0	WRKY	0
EIL	2	ZF-HD	1
G2-LIKE	1		
GeBP	0	TOTAL	139

4.2.1. Famílias ARF e AUX/IAA

Os membros das famílias *AUXIN RESPONSE FACTOR* (ARF) e AUX/IAA são conhecidos por suas funções na via de sinalização de auxina em processos de crescimento e desenvolvimento. O controle da expressão dos genes de resposta à auxina depende sempre da ação conjunta de um ARF e um AUX/IAA. (Guilfoyle e Hagen, 2007).

Foram encontrados 3 cDNAs possivelmente codificadores de ARFs no PASSIOMA, um em *P. suberosa* (PACEPS7014A05.g) e dois em *P. edulis* (PACEPE3017E08.g e PACEPE3010H07.g). A sequência PACEPE3010H07.g não foi utilizada na filogenia devido a sua baixa qualidade. Observou-se que PACEPE3017E08.g é mais similar ao gene *ARF6* (At1g30330) em *Arabidopsis* (Figura 3). Encontrar este gene no PASSIOMA seria esperado porque experimentos de hibridização *in situ* mostraram que *ARF1*, *ARF2* e *ARF6* são expressos em flores em desenvolvimento de *Arabidopsis*, ao passo que a maioria dos outros membros da família se expressa em tecidos vegetativos (Guilfoyle e Hagen, 2007). Além disso, experimentos com mutantes mostram que mutantes simples e duplos, como *arf2*, *arf1/arf2*, *arf3* e *arf6/arf8*, possuem flores com morfologia anormal e/ou baixa fertilidade, sugerindo sua atividade redundante no controle da regulação de genes responsivos à auxina no desenvolvimento reprodutivo (Okushima et al., 2005).

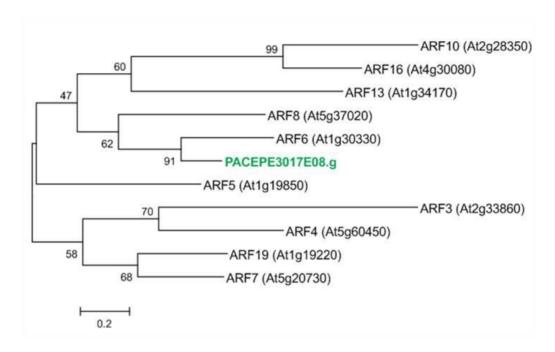


Figura 3: Cladograma com sequências proteicas de ARF de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. A sequência de *P. edulis* encontrada no PASSIOMA aparece destacada. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Foram encontradas 9 sequências pertencentes à família AUX/IAA no PASSIOMA. Três dessas sequências – o singleton encontrado em *P. pohlii* (PACEPP1002F09.g) e as sequências PACEPS7009B06.g e PACEPE3001A08.g – de *P. suberosa* e *P. edulis* respectivamente – mostraram pouca similaridade com as demais sequências de AUX/IAA, sendo encontradas como grupo externo da árvore e, portanto, foram retiradas da análise comparativa. Apesar disso, os resultados de BLASTx contra o banco de dados público NCBI mostraram que se tratam de ortólogos do *IAA9* de outras espécies. Estudos em *V. vinifera* mostraram expressão de um ortólogo de *IAA9* (*VvIAA9*) em folhas e frutos, além de aumento da expressão desde a antese até o inicio do amadurecimento do fruto. Fujita et al. (2012) demonstraram que a superexpressão de *VvIAA9* em *A. thaliana* promove o

crescimento, incluindo rápida transição floral, mas sem alterações da morfologia da planta (Fujita et al., 2012).

As outras 6 sequências encontradas estão mostradas na figura 4. Foi encontrada uma sequência em *P. suberosa* (PACEPS7015H06.g) e outras duas em *P. edulis* (PACEPE3011E05.g e PACEPE3005A01.g), as quais aparentaram ser ortólogas do gene *IAA16* (At3g04730) de *Arabidopsis*. Trabalhos analisando mutantes de inserção para o gene *IAA16* exibem fenótipo alterado relacionado ao aumento da resistência de resposta ao ABA no alongamento das raízes e estabelecimento dos tecidos da semente, além de redução da sinalização de auxina (Rinaldi et al., 2012).

A sequência de *P. edulis*, PACEPE3007E06.g, mostrou semelhança com o clado contendo IAA16, AXR3/IAA17, PAP2/IAA27, SLR/IAA14 e AXR2/IAA7. As alterações observadas nos mutantes *iaa16* já foram identificadas também em mutantes para os genes *IAA7, IAA14* e *IAA17*, demonstrando que o funcionamento normal de todas essas proteínas é necessário para a resposta completa à auxina.

Foi encontrado também outro gene putativo da família IAA em *P. edulis* (PACEPE3027F11.g) e trata-se de um ortólogo de *IAA5* (At1g15580), segundo a análise de sequências. A ferramenta BAR mostrou expressão do *IAA5* principalmente em folhas de *Arabidopsis*. No entanto, esse ortólogo aparece em botões florais de *P. edulis*, indicado uma possível atuação no desenvolvimento floral desta espécie.

Por fim, a sequência PACEPE3020B01.g não mostrou valores de bootstrap satisfatórios para sua identificação através do cladograma. No entanto o alinhamento pelo BLASTx mostrou maior semelhança com o clado contendo SHY/IAA6 e IAA4 (At5g43700).

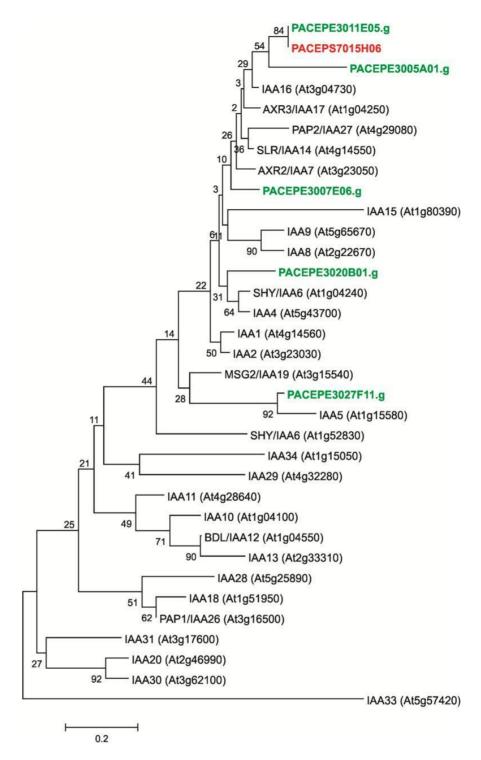


Figura 4: Cladograma com sequências proteicas de AUX/IAA de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e a sequência de *P. suberosa* aparece destacada em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

A função de cada um dos 29 membros da família AUX/IAA ainda não é completamente conhecida. Estudos usando mutantes simples não mostraram defeitos visíveis no desenvolvimento quando comparados com as plantas tipo selvagem, indicando a ação redundante destes fatores de transcrição. Duplos ou triplos mutantes de genes AUX/IAA relacionados, como *iaa8/iaa9 ou iaa5/iaa19/iaa6* também exibiram fenótipo selvagem, apesar de se observar mudança no padrão de expressão global dos genes (Alonso et al., 2005).

4.2.2. Família bHLH

As proteínas bHLH (basic helix-loop-helix) constituem uma das maiores famílias de fatores de transcrição conhecidas. Ainda não são conhecidas as funções de todas as bHLHs, mas já se sabe que representam componentes regulatórios chave no controle da rede transcricional de um grande número de processos biológicos. Já foram descritas em plantas atividades no controle da resposta à luz, sinalização hormonal, ramificação de ramos, desenvolvimento de flores e frutos, micrósporos, tricomas, estômatos e raízes (Carretero-Paulet et al., 2010).

A busca por sequências codificadoras de fatores de transcrição do tipo bHLH na biblioteca do PASSIOMA indicou a existência de 11 cDNAs, 9 foram encontrados em *P. edulis* e 2 em *P. suberosa*. As análises comparativas mostraram que a sequência PACEPE3016G12.g (Figura 5) é um ortólogo putativo do gene *ABORTED MICROSPORES* de *Arabidopsis* (*AMS* – At2g16910). Este gene codifica uma proteína necessária ao desenvolvimento das células do tapete nas anteras e para a formação dos micrósporos pós-meióticos (Xu et al., 2010). A sequência PACEPE3002H02.g encontra-se

no mesmo clado da sequência anterior, mas o valor de bootstrap que sustenta o nó do cladograma (Figura 5) não foi significativo para comprovar sua semelhança com *AMS*.

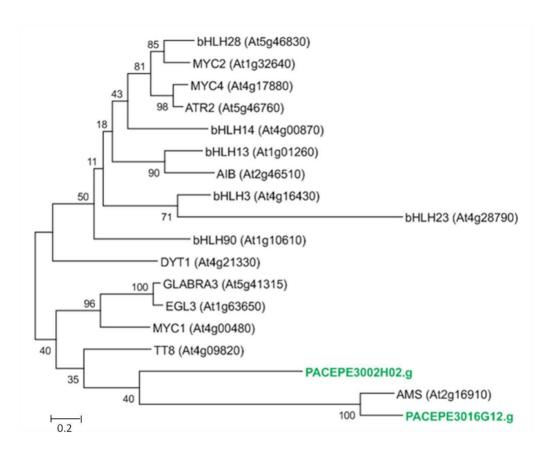


Figura 5: Cladograma com sequências proteicas de bHLH de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Foram encontradas outras duas sequências (PACEPE3019G06.g e PACEPE3022C01.g) que parecem ser parálogas, produtos de uma duplicação recente, ambas semelhantes ao gene *BIGPETAL* (*BPE* – At1g59640) – Figura 6. O gene *BPE* foi descrito como codificador de um fator de transcrição bHLH envolvido no controle do tamanho das pétalas. É um dos genes ativados por AP3, PI e AP1 e reprimido por AG. Seu

nome é devido ao fenótipo do mutante *bpe*, o qual possui pétalas maiores que as do tipo selvagem. A literatura sugere que BPE interfere na expansão celular pós-mitótica (Szécsi et al., 2006).

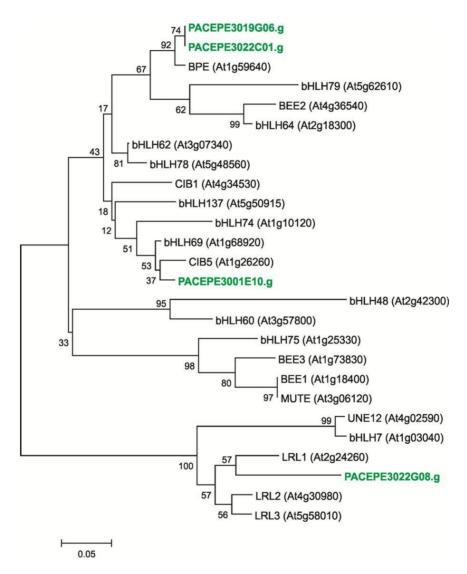


Figura 6: Cladograma com sequências proteicas de bHLH de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Além das sequências já mencionadas, PACEPE3022G08.g (Figura 6) está inclusa no clado contendo os genes *LRL* (*LRL1*, *LRL2 e LRL3* ou *bHLH66*, *bHLH69 e bHLH82*, respectivamente). Mais uma vez, pouca informação sobre esses genes está disponível atualmente. Sabe-se, no entanto, que os genes *AtLRL* são descritos em *Lotus japonicus* como *ROOTHAIRLESS* (LjRHL). Mutações neste locus previnem a formação de pelos radiculares e mostram ação redundante dos membros deste clado (Karas et al., 2009).

A sequência PACEPE3030C06.g foi incluída no clado contendo os genes bHLH27 (At4g29930) e bHLH35 (At5g57150), como mostrado na figura 7A. A sequência de P. suberosa, PACEPS7008G02.g, assim como PACEPE3029B03.g, de P. edulis, são ortólogas putativas do gene BS1 (bHLH135 – At1g74500), como pode ser observado na figura 7B. O gene BS1 codifica um bHLH que não é capaz de se ligar diretamente ao DNA e que está envolvido na regulação da sinalização pela luz. Castelain et al. (2012) relataram a expressão deste gene nos meristemas apical e radicular e em raízes laterais. Experimentos com linhagens que superexpressam AtBS1 em A. thaliana mostraram que a planta se torna menos sensível à luz azul vermelha e vermelho-distante, o que explica o fenótipo de hiposensibilidade do mutante bs1 à luz, mesmo quando as plantas crescem sob condições de luz branca (Castelain et al., 2012).

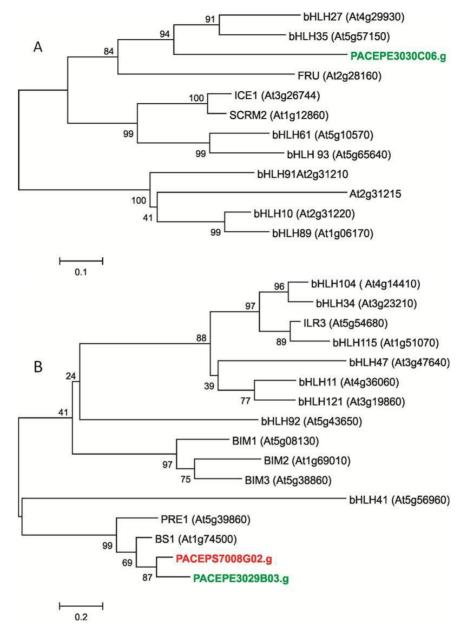


Figura 7: Cladograma com sequências proteicas de bHLH de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e a sequência de *P. suberosa* aparece destacada em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

4.2.3. Família bZIP

São conhecidos cerca de 75 genes que codificam fatores de transcrição com domínio conservado bZIP em *Arabidopsis*. Estudos mostraram que eles são importantes na regulação de diversos processos biológicos como defesa contra patógenos, controle de resposta à luz e estresses, maturação da semente e desenvolvimento floral (Jakoby et al., 2002).

No PASSIOMA foram encontrados 4 sequências que codificam motivos conservados que indicam pertencerem à família bZIP. A primeira sequência de cDNA encontrada, PACETS1002D06.g não foi usada no alinhamento para análise comparativa por apresentar baixa qualidade, o mesmo ocorreu com PACEPS7014D06.g.

Sendo assim, o cladograma mostrou que o singleton PACEPE3018H05.g assemelhase ao gene At2g24340. Segundo BAR (Winter et al., 2007) há expressão fraca deste gene no meristema da inflorescência e em sementes. O outro singleton PACEPE3005E04.g não apresentou bootstrap significativo no cladograma. Logo sua identificação não pode ser feita por esse método. Entretanto, através do alinhamento pelo BLASTx observou-se a semelhança dessa sequência com At1g52320, a qual é identificada como codificadora de uma "putative bZIP protein" e é expressa em folhas e sementes em Arabidopsis (Figura 8). Como essas proteínas ainda não foram bem caracterizadas, não se conhece sua função e o seu papel no desenvolvimento reprodutivo.

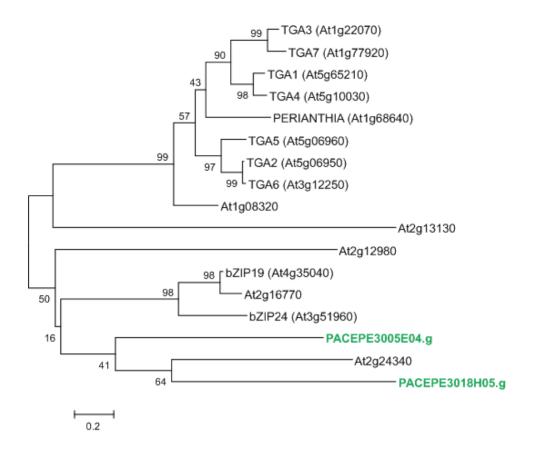


Figura 8: Cladograma com sequências proteicas de bZIP de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

4.2.4. Família C₃H

Foram encontrados 8 genes putativos no PASSIOMA que codificam potencialmente FT do tipo C₃H. Três destas sequências (PACEPS7019G11.g, PACETS1001D03.g e PACEPS7012A11.g) não puderam ser incluídas na análise por causa do tamanho insuficiente de seus *reads*. As sequências PACEPE3010A01.g, PACEPS7014G01.g e PACEPS7018E01.g são semelhantes (Figura 9) e segundo o cladograma, pode-se observar que estas sequências são ortólogas de AtC3H55 (At5g12440) de *Arabidopsis*. A ferramenta

BAR indica que o gene que codifica esta proteína é expresso no meristema apical e no meristema da inflorescência, assim como nos botões florais e nos carpelos (Winter et al., 2007). A sequência PACEPE3007G08.g possivelmente apresenta semelhança com as proteínas AtC3H59 (At5g40880) e AtC3H62 (At5g49200) (Figura 9), para as quais há evidencias de expressão nos tecidos das raízes (Winter et al., 2007).

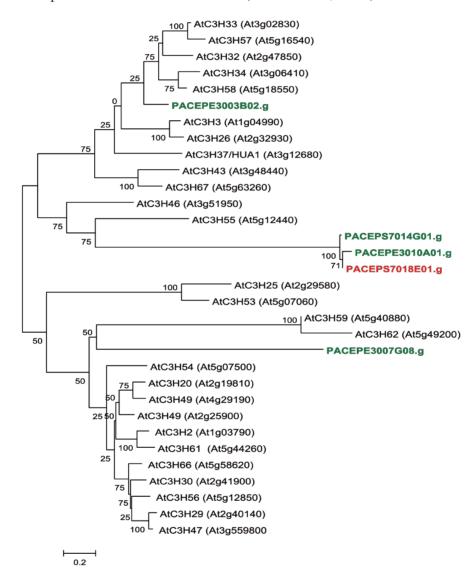


Figura 9: Cladograma com sequências proteicas de C₃H de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Através da análise de similaridade de sequências não foi possível afirmar a identificação de PACEPE3003B02.g. No entanto, esta sequência está presa no clado contendo *HUA1* (Figura 9). Os genes *HUA1* e *HUA2* foram identificados como reguladores da identidade de estames e carpelos (Li et al, 2001). Igualmente, AtC3H43 também presente no mesmo clado, é encontrado em inflorescências (Wang et al. 2008).

4.2.5. Família HMG

As proteínas HMG são abundantes na estrutura física dos cromossomos. Através do controle das interações proteína-proteína e proteína-DNA, essas proteínas auxiliam a formação de complexos de fatores de transcrição, regulando a expressão gênica (Grasser, 2003; Stros, 2010). As proteínas desta família são divididas em proteínas HMGA e HMGB, sendo a subfamília HMGB a mais diversa. (Grasser et al., 2007). HMGB1 a HMGB5 foram caracterizadas em milho e *Arabidopsis* e possuem expressão generalizada (Grasser, 2003).

Seis sequências de cDNA potencialmente codificadoras de HMG foram encontradas no banco de dados do PASSIOMA: 4 em *P. edulis* e 2 em *P. suberosa* (Figura 10). As sequências PACEPE3011A01.g e PACEPE3027H04.g são provavelmente parálogas entre si e ortólogas do clado contendo HMGB2 e HMGB3. Juntamente com este clado encontrase PACEPE3003B04.g e PACEPS7013F10.g, as quais são ortólogas entre si e são semelhantes aos mesmo HMGB2 e HMGB3. Experimentos apontam que a expressão de *HMGB2*, *HMGB3* e *HMGB4* foi aumentada em plantas selvagens sob condições de estresse pelo frio, enquanto que a expressão de *HMGB2* e *HMGB3* foi marcadamente reduzida em estresse salino e por seca. Nesses dois tipos de estresse, plantas que superexpressavam *HMGB2* apresentaram germinação e crescimento retardados, indicando o papel dessas proteínas na regulação de situações de estresse (Kwak et al., 2007).

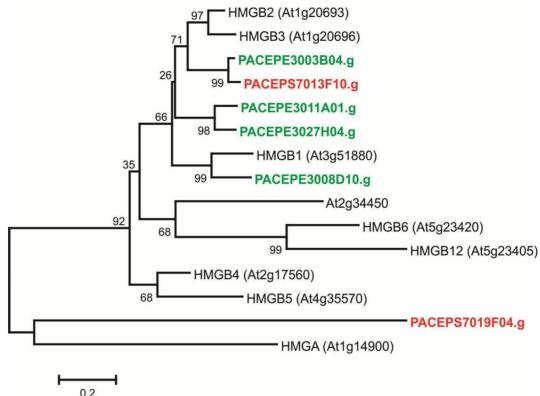


Figura 10: Cladograma com sequências proteicas de HMG de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Identificou-se, também, a sequência PACEPE3008D10.g, que é o ortólogo putativo de HMGB1 de *Arabidopsis*. Experimentos mostraram que a alteração da expressão de *HMGB1* em *Arabidopsis* influencia o crescimento da plantas e sua resposta às situações de estresse. A superexpressão de *HMGB1* gerou raízes primarias curtas e afetou a sensibilidade ao agente tóxico metil-metanosulfonato (Lildballe et al., 2008). A expressão ectópica de um ortólogo de *HMGB1* de *Zea mays* em sementes de tabaco (*Nicotiana tabacum*) reduziu o tamanho da raiz primária. A ausência desta proteína afeta aspectos do

crescimento da planta e tolerância a estresse (Pedersen et al., 2010). Já a sequência PACEPS7019F04.g, de *P. suberosa*, provavelmente codifica uma proteína HMGA.

4.2.6. Família MADS-box

Foram detectados 7 genes putativos codificadores de MADS-box no PASSIOMA. As sequências PACEPS7012E07.g e PACEPE3011A12.g são provavelmente ortólogas dos genes *SHATTERPROF* (*SHP* e *SHP2*; Figura 11), envolvidos na formação da zona de deiscência nas síliquas de *Arabidopsis*, com expressão concentrada em carpelos e óvulos (Ferrandiz, 2000).

A sequência PECEPE3005F02.g, de *P. edulis*, é um ortólogo do gene *FRUITFULL* (*FUL*), conforme mostra a figura 11. O mutante *ful* não é capaz de liberar as sementes do fruto porque apresenta excesso de tecido de replum, comprometendo o tecido que origina as valvas. A superexpressão de *FUL* ocasiona um fenótipo no fruto semelhante ao observado no mutante *shp1shp2*, assim, acredita-se que *SHP* seja regulado negativamente pelo gene *FUL* (Ferrandiz, 2000).

As sequências de *P. suberosa*, PACEPS7009D01.g é ortóloga de alguma proteína SEPATALLATA, uma vez que o bootstrap do nó que a prende no clado contendo essas proteínas é significativo. Os genes que codificam estas proteínas possuem sequências muito semelhantes e atuam redundantemente na determinação da identidade dos órgãos florais. Observou-se que quando os genes *SEPALLATA* possuem expressão reduzida (mutantes triplos *sep*), todos os órgãos florais se convertem em folhas. Isso ocorre porque essas proteínas são essenciais para as interações necessárias para que os MADS-box do grupo ABC se liguem nos promotores dos genes por eles ligados (Pelaz et al., 2000).

A sequência PACETS1001G10.g, também codificadora putativa de um fator de transcrição MADS, não teve sua identidade relatada devido a sua baixa qualidade.

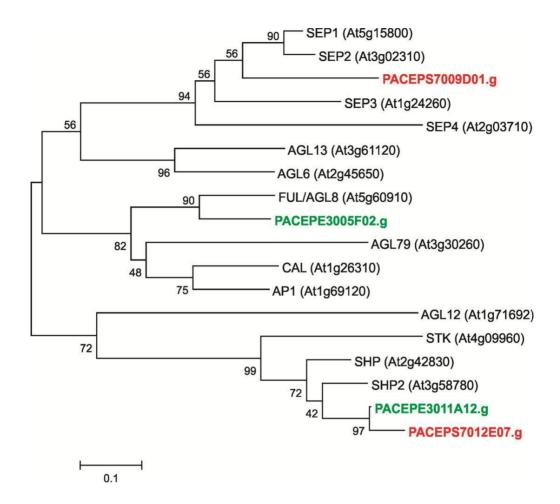


Figura 11: Cladograma com sequências proteicas de MADS de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

Os outros genes MADS encontrados em P. edulis foram PACEPE3001E12.g e PACEPE3002E10.g. O primeiro é um ortólogo do gene AG, um MADS de classe C que determina a formação de estames e carpelos, tendo sua expressão concentrada nos 3° e 4°

verticilos e reprimida nos demais pela ação dos genes de classe A (Figura 12). A segunda sequência mencionada é ortóloga a *PISTILLATA* (Figura 12), um MADS da classe B, expresso nos 2º e 3º verticilos, determinando a formação de pétalas e estames nestes locais (Theissen e Saedler, 2001).

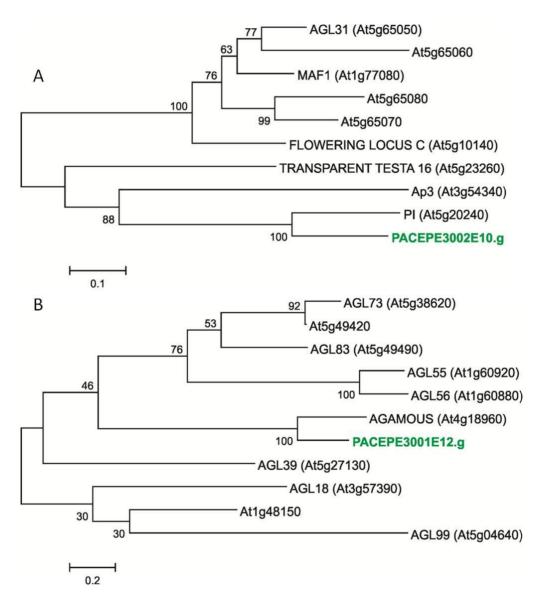


Figura 12: Cladogramas com sequências proteicas de MADS-box de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

4.2.7. Família MYB

Em *Arabidopsis* são conhecidos cerca de 150 genes codificadores de FT da família MYB, principalmente do grupo R2R3-MYB. São responsáveis por uma variedade enorme de processos vegetais, funcionando como fatores fundamentais em redes regulatórias que controlam desenvolvimento, metabolismo e resposta a estresses bióticos e abióticos (Dubos et al., 2010).

Esta família de fatores de transcrição foi a que exibiu maior número de sequências no PASSIOMA, 14 genes putativos foram encontrados, 5 deles em *P. edulis* e 9 em *P. suberosa*. A figura 13 demonstra a dificuldade na identificação de algumas sequências encontradas: PACEPS7023C07.g e PACEPS7023H05.g aparentam ser parálogas e possivelmente fruto de uma duplicação em *P. suberosa*. O mesmo ocorre com as sequências PACEPE3006A01.g e PACEPE3028C10.g, em *P. edulis*. Essas 4 sequências, juntamente com PACEPE3002C03.g e PACEPS7024A12.g estão presentes no mesmo clado, próximo de alguns subgrupos da subfamília R2R3-MYB. O BLASTx dessas sequências contra o NCBI também não mostrou um resultado muito preciso, classificando essas proteínas como "*MYB-related transcription factors*".

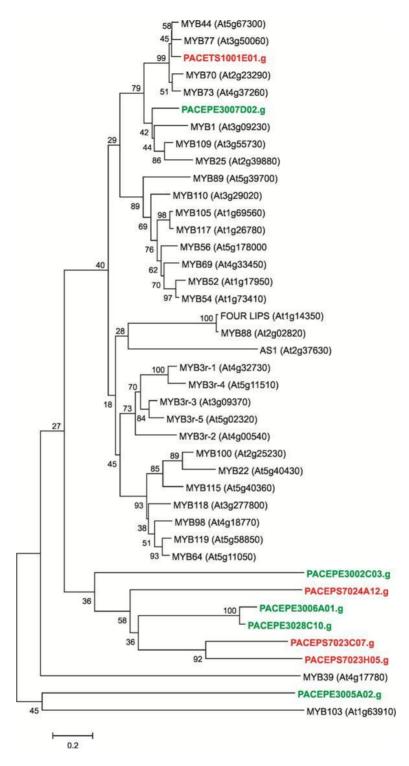


Figura 13: Cladograma com sequências proteicas de MYB de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

A sequência PACEPE3005A02.g é um cluster de dois reads (PACEPE3018E06 e PACEPE3005A02.g) e mostrou a proteína MYB103 como a mais semelhante dos MYB no genoma de *Arabidopsis*, mas o bootstrap não foi significativo. Em *P. edulis*, o singleton PACEPE3007D02.g se assemelha ao grupo que contém as proteínas AtMYB1 (At3g09230), AtMYB109 (At3g55730) e AtMYB25 (At2g39880). Por ser uma família muito grande a maioria dos MYB ainda não possui sua função determinada, sabe-se através da ferramenta BAR que o gene *AtMYB1* tem expressão nas folhas e flores, já o gene *AtMYB25* nos meristemas apical e da inflorescência e *AtMYB109* se expressa preferencialmente nos estames (Winter et al., 2007).

PACETS1001E01.g é uma sequência de *P. suberosa* incluída em um clado contendo quatro possíveis ortólogos em *Arabidopsis*, sobre os quais, no entanto, não há informações funcionais disponíveis.

4.2.8. Família PHD

Os FT da família PHD possuem um domínio PHD, um *zinc finger* encontrado em muitas proteínas regulatórias de plantas e animais, e um domínio HD. Essas proteínas estão frequentemente associadas com a regulação transcricional mediada pela cromatina (Halbach et al., 2000). Já foi mostrado que proteínas PHD de *Arabidopsis* possuem função na fertilidade e florescimento, além de estarem envolvidos na manutenção e estabelecimento dos meristemas apical e radicular (Saiga et al, 2008).

Cinco genes putativos foram encontrados: 4 deles em *P. suberosa*. A sequência PACEPS7004G09.g encontra-se num clado com baixo suporte de boostrap, o que comprometeu a sua identificação (Figura 14).

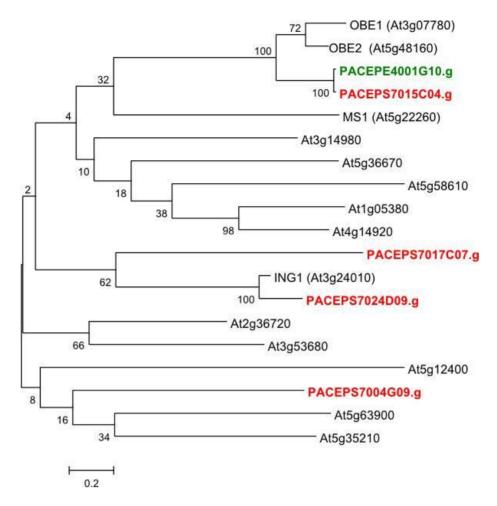


Figura 14: Cladograma com sequências proteicas de PHD de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. A sequência de *P. edulis* encontrada no PASSIOMA aparece destacada em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

As sequências PACEPE7024D09.g e PACEPS7017C07.g, também de *P. suberosa*, codificam FT ortólogos a AtING1 de *Arabidopsis*. No entanto, o bootstrap do nó de ligação a PACEPS7017C07.g possui baixo suporte. A proteína AtING1 é conhecida por ligar-se a histonas metiladas no núcleo e expressa-se por toda a planta (Lee et al., 2009).

As sequências PACEPE4001G10.g e PACEPS7015C04.g, presentes em *P. edulis* e *P. suberosa* respectivamente, são ortólogas entre si e às proteínas OBERON1 (At3g07780) e OBERON2 (At5g48160). Estudos mostraram que esses fatores de transcrição atuam na manutenção e estabelecimento dos meristemas em *Arabidopsis*. O duplo mutante *obe1/obe2* exibe terminação prematura do meristema apical, sugerindo atividade redundante destes genes. Outras análises revelaram que os genes *OBE1* e *OBE2* permitem que as células vegetais adquiram atividade meristemática através da via *WUSCHEL-CLAVATA*, necessária para manutenção da população das células do meristema (Saiga et al., 2008).

4.2.9. Família Trihelix

As proteínas desta família (também conhecidas como proteínas GT) possuem um domínio de ligação ao DNA conservado especifico de plantas, chamado Trihelix. Estas proteínas estão possivelmente envolvidas na regulação de respostas ligadas a estímulos luminosos, assim como processos de crescimento e desenvolvimento de órgãos florais (Zhou, 1999).

Foram descobertas 7 sequências possivelmente codificadoras de FT da família Trihelix em *Passiflora*, sendo 4 de *P. suberosa* e 3 de *P. edulis*. A sequência PACEPE3012F11.g mostrou ser ortóloga do clado contendo GT-4 e GT-1 (Figura 15A). Sabe-se que as proteínas GT-1 e GT-2 são fatores de transcrição que interagem com promotores de genes responsivos a estímulo luminoso (Smalle et al., 1998). A sequência de *P. suberosa* PACEPS7007B11.g também é relacionada a este clado (Figura 15A), mas com um suporte de bootstrap baixo, que não permite a atribuição de sua identidade. As proteínas codificadas por PACEPS7005B01.g e PACEPE3028E10.g são ortólogas entre si (a

primeira de *P. suberosa* e a segunda de *P. edulis*), mas devido aos baixos valores de bootstrap que suportam os ramos externos não puderam ter sua identidade estabelecida. PACEPS7001H10.g é ortóloga de At5g47660 que codifica uma proteína relacionada à estabilidade de telômeros e controle de expressão gênica em regiões de heterocromatina (Bowen et al., 2010).

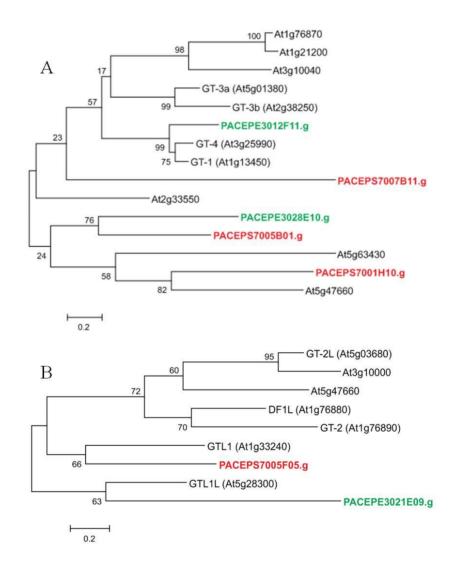


Figura 15: Cladograma com sequências proteicas de Trihelix de *Arabidopsis* e com sequências proteicas deduzidas de *Passiflora*. As sequências de *P. edulis* encontradas no PASSIOMA aparecem destacadas em verde e as sequências de *P. suberosa* aparecem destacadas em vermelho. Os números próximos de cada nó representam os valores bootstrap para 1000 replicações. Barra indica a taxa de substituição.

As sequências PACEPS7005F05.g, de *P. suberosa* e PACEPE3021E09.g de *P. edulis* são semelhantes às proteínas GTL1 e GTL1L (GTL1-like ou GTL2; Smalle et al., 1998) de *Arabidopsis*, embora o suporte de bootstrap seja pouco significativo para ambas. GTL1 esta presente no núcleo de tricomas em desenvolvimento e a perda de função deste gene leva à endorreduplicação em tricomas. Além disso, os mutantes *gtl1* possuem a expressão de vários genes do ciclo celular alterada e alteração da ploidia celular. GTL1L provavelmente possui função redundante à de GTL1 (Breuer et al., 2009).

4.3. Análise da expressão de fatores de transcrição na corona de P. edulis

Uma vez que todas as sequências putativas codificadoras de fatores de transcrição encontradas no PASSIOMA foram identificadas, foi utilizada a técnica do macroarranjo para encontrar aquelas que possuem expressão nos filamentos da corona de *P. edulis*. Dessa forma, apenas as 69 sequências de cDNA encontradas para esta espécie foram utilizadas nos próximos experimentos. O resultado da hibridização pode ser encontrado na figura 16.

Pode-se observar que 9 genes putativos mostraram resultado positivo no macroarranjo. Foram considerados resultados positivos apenas os cDNAs que tiveram as 4 réplicas hibridizadas em 3 repetições do experimento. Dessa maneira, algumas sequências que provavelmente são expressas na corona, pois tiveram algumas réplicas hibridizadas, não foram consideradas. Entre essas sequências estão: PACEPE3011E05.g identificado como um ortólogo de IAA16; PACEPE3007A04.g e PACEPE3012H02.g, sequências putativas codificadoras de FT da família Homeobox; e PACEPE3021G03, uma sequência codificadora de um FT da família TUBBY.

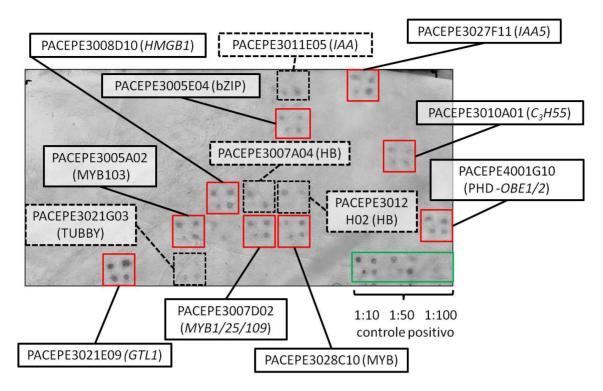


Figura 16: Macroarranjo com amostras de cDNA de *P. edulis*, plotadas em quadruplicata e hibridizado utilizando sonda de RNAm total de corona marcada com digoxigenina. Os resultados consistentes em três experimentos são mostrados nos quadrados vermelhos. O retângulo verde indica o controle positivo (cDNA utilizado para confecção da sonda em três diluições). A posição ao lado esquerdo do controle positivo 1:10 é ocupada por uma amostra de água (controle negativo).

Houve hibridização de um gene putativo da família AUX/IAA, o PACEPE3027F11.g, identificado como um IAA5 (Figura 16). A função específica de IAA5 ainda não é conhecida. A ferramenta BAR (Winter et al., 2007), indicou a expressão deste gene em botões florais e folhas em desenvolvimento. Sabe-se ainda que os FT da família AUX-IAA estão envolvidos na via de sinalização de auxina em processos de crescimento e desenvolvimento (Guilfoyle e Hagen, 2007).

Outra sequência com expressão em corona foi PACEPE3005E04.g, um gene da família bZIP, homólogo a At1g52320 de *Arabidopsis*. Esta proteína é um bZIP putativo e

ainda não se conhece sua função no desenvolvimento das plantas, entretanto, sua expressão ocorre em folhas senescentes, pétalas, sépalas e estames (Winter et al., 2007).

PACEPE3010A01.g, identificado pela análise comparativa como um ortólogo de AtC3H55, de função desconhecida em *Arabidopsis*, também mostrou sinal de hibridização (Figura 16). A presença destes transcritos na corona corrobora a expressão observada para AtC3H55 (ferramenta BAR, Winter et al., 2007) no meristema caulinar vegetativo e da inflorescência, além de botões florais e carpelo.

A família MYB foi a que mostrou o maior número de genes expressos na corona, o que era esperado, uma vez que é a família de FT com o maior número de membros em plantas: PACEPE3005A02.g, PACEPE3007D02.g e PACEPE3028C10.g. A primeira, PACEPE3005A02.g, é similar a *MYB103*, que em *Arabidopsis* está envolvido na regulação da biosíntese de lignina e expressa-se em tecidos provasculares (Ohman et al., 2012). A sequência PACEPE3007D02.g foi identificada como próxima do que contém as proteínas AtMYB1, AtMYB109 e AtMYB25. Sabe-se que esses genes se expressam nos estames (*AtMYB109*) e nos meristemas apical e da inflorescência (*AtMYB25*), assim como nas folhas e flores (*AtMYB1*) (Winter et al., 2007). Já a sequência PACEPE3028C10.g não teve a sua identidade determinada, uma vez que faz parte de um clado que contém apenas sequências de *Passiflora*. É importante observar que a análise de sequências (Figura 13) apontou PACEPE3006A01.g como um parálogo de PACEPE3028C10.g e que, no entanto, não se expressou na corona, apontando uma provável divergência de função entre estas sequências.

A sequência PACEPE3008D10.g, identificada como HMGB1, também mostrou expressão na corona. Sabe-se que os FT da família HMG se ligam ao DNA e induzem mudança estrutural da cromatina, interagindo com histonas e outros fatores de transcrição e

participando da regulação gênica. A proteína HMGB1 está relacionada com certos aspectos do crescimento e tolerância ao estresse. Segundo estudos, tanto a ausência como a superexpressão de *HMGB1* causam raízes mais curtas que o tipo selvagem, além de afetar a sensibilidade a agentes genotóxicos. A superexpressão deste gene ainda diminui a taxa de germinação de sementes na presença de altas concentrações de NaCl (Lildballe et al., 2008).

PACEPE4001G10 foi identificada como um ortólogo de *OBE1/OBE2* de *Arabidopsis*, que estão relacinados à manutenção e estabelecimento dos meristemas caulinar e radicular (Saiga et al., 2012).

Finalmente, PACEPE3021E09.g, que potencialmente codifica uma proteína da família Trihelix semelhante a GTL1L, envolvido no desenvolvimento dos tricomas foliares em *Arabidopsis* (Smalle et al., 1998; Breuer et al., 2009), mas que também se expressa em pétalas (de acordo com a ferramenta BAR, Winter et al., 2007).

4.4. RT-PCR

As sequências que apresentaram expressão consistente na corona, de acordo com os resultados de macroarranjo, foram também investigadas por RT-PCR. Os resultados de RT-PCR indicaram expressão de todos os genes analisados em todos os órgãos florais e em folhas de *P. edulis*. Esse padrão de expressão generalizado para todos os fatores de transcrição pode ser considerado incomum, mas uma análise mais detalhada dos possíveis ortólogos destas sequências, demonstrou que os mesmos possuem expressão documentada em tecidos vegetativos e reprodutivos de *Arabidopsis*.

A expressão em todos os tecidos de genes codificadores para proteínas HMG pode ser explicada porque estas são muito abundantes na estrutura física dos cromossomos, ligandose a cerca de 10% dos nucleossomos (Grasser et al., 2007).

A expressão generalizada do gene da família C₃H (PACEPE3010A01.g) esta de acordo com o seu possível ortólogo, AtC3H55 que é expresso em raízes, inflorescências, folhas e caule de *Arabidopsis* (Wang et al., 2008).

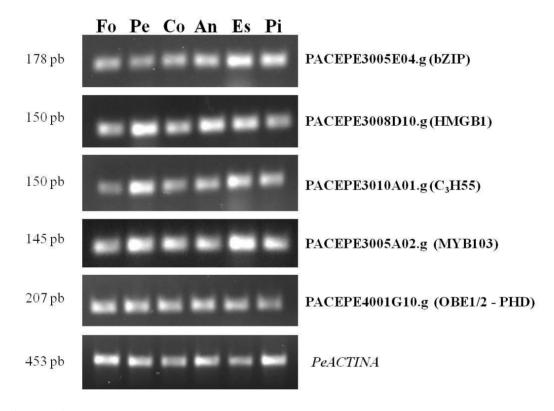


Figura 17: RT-PCR de diferentes tecidos de *P. edulis*. Cada painel mostra as bandas correspondentes à amplificação do fragmento de um dos genes de interesse. No último painel, as bandas são referentes à amplificação do gene constitutivo *PeACTINA*, usado *c*omo controle positivo. **Fo**: folha, **Pe**: perianto, **Co**: corona, **An**: androginóforo, **Es**: estames e **Pi**: pistilo.

A sequência possivelmente codificadora de MYB103, PACEPE3005A02g, também mostrou expressão em todos os tecidos analisados. Estudos mostraram que em *Arabidopsis* está proteína está envolvida na regulação da biosíntese de lignina e expressa-se em tecidos provasculares, o que corrobora com sua expressão generalizada.

Devido à impossibilidade de identificação da sequência potencialmente codificadora de um bZIP, PACEPE3005E04.g (Figura 8), pouco se pode dizer sobre seu padrão de expressão além de que está expresso em vários tecidos de *P. edulis*.

4.5. Análise da expressão dos genes PACEPE3010A01.g, PACEPE4001G10.g durante o desenvolvimento da corona por hibridização *in situ*

Apesar dos resultados de RT-PCR indicarem que nenhum dos genes estudados expressou-se preferencialmente na corona, não havia conhecimento sobre o padrão de expressão espacial e temporal dos mesmos durante o desenvolvimento floral em *P. edulis*. Com essa finalidade, realizamos a análise do padrão de expressão tecidual para dois genes: PACEPE3010A01.g e PACEPE4001G10.g.

A figura 18 ilustra o padrão de expressão do gene PACEPE3010A01.g ao longo do desenvolvimento da corona. Desde o início, quando a corona surge como uma emergência no tubo floral, o sinal de hibridização foi uniforme nas células epidérmicas e subepidérmicas (Figura 18A), permanecendo durante a formação da primeira série de filamentos da corona e no opérculo (Figura 18B). Nos próximos estágios, com o surgimento de emergências de mais séries de filamentos, percebe-se maior concentração de transcritos no ápice da primeira série da corona, no opérculo e no nectário (Figuras 18C e 18D). Durante o desenvolvimento tardio da corona, a expressão do PACEPE3010A10

concentrou-se no ápice dos filamentos mais longos (mais externos) e no nectário, enquanto a expressão no opérculo não foi mais observada (Figura 18E).

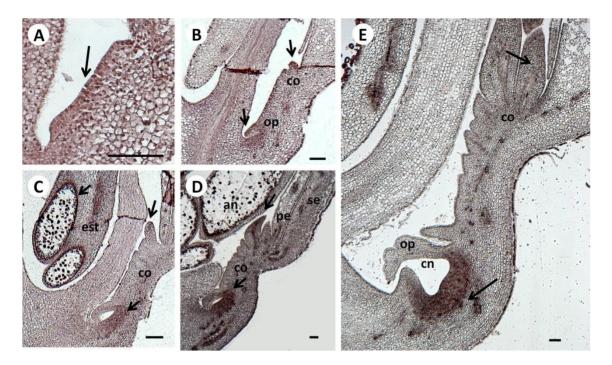


Figura 18: Hibridização *in situ* em botões florais seccionados longitudinalmente, mostrando o padrão de expressão de PACEPE3010A01.g, durante o desenvolvimento da corona em *P. edulis*. (A) Sinal de hibridização uniforme durante emergência da corona no tubo floral. (B) A coloração rosada indica a presença de transcritos no ápice da primeira série de filamentos da corona e do opérculo. (C) Presença de sinal de hibridização no ápice da primeira série de filamentos da corona e no nectário. (D) Continuação do desenvolvimento da corona com precipitado em tecidos similares ao anterior. (E) No desenvolvimento tardio da corona, sinal de hibridização concentrado no nectário e nos filamentos mais desenvolvidos da corona. As setas evidenciam a presença de transcritos. an: antera, cn: câmara nectarífera, co:corona, est: estame, pe: pétala, op: opérculo, se: sépala. Barras = 150 μm.

Adicionalmente, o padrão de expressão observado para PACEPE3010A01.g confirmou os resultados obtidos por RT-PCR, em que o gene expressou-se em todos os tecidos testados de botões florais. Por meio da hibridização *in situ* observou-se expressão

do gene nas camadas epidérmicas e subepidérmicas das sépalas, nas pétalas, nas anteras e na corona (Figura 19A) e no pistilo (Figura 19B), nos tecidos estigmáticos e nos primórdios de óvulos (Figura 19C). A expressão nos primórdios dos óvulos foi observada ainda em estágios mais tardios, quando estes já apresentavam desenvolvimento de tegumentos (Figura 19D). As figuras 19E, 19F e 19G confirmam o que foi visto nos cortes longitudinais, a expressão na corona ocorre nos filamentos mais desenvolvidos.

O gene PACEPE4001G10.g também teve seu padrão de expressão investigado em *P. edulis*, em diferentes estágios de desenvolvimento da corona e outros tecidos florais. A figura 20 ilustra a expressão ao longo do desenvolvimento da corona. Pode-se perceber o início da expressão na emergência da corona (Figura 20A). Em um estágio posterior, o sinal de hibridização concentrou-se na extremidade da primeira série em formação de filamentos e no nectário (Figura 20B). O padrão de expressão nos filamentos mais externos foi observado em estágios subsequentes de desenvolvimento e uma concentração do sinal de hibridização foi observada na região de diferenciação dos tecidos do nectário (Figuras 20C e 20D). Em estágios mais tardios do desenvolvimento da corona, os transcritos de PACEPE4001G10.g concentraram-se na extremidade dos filamentos da corona e no nectário floral, sendo excluídos do opérculo (Figura 20E).

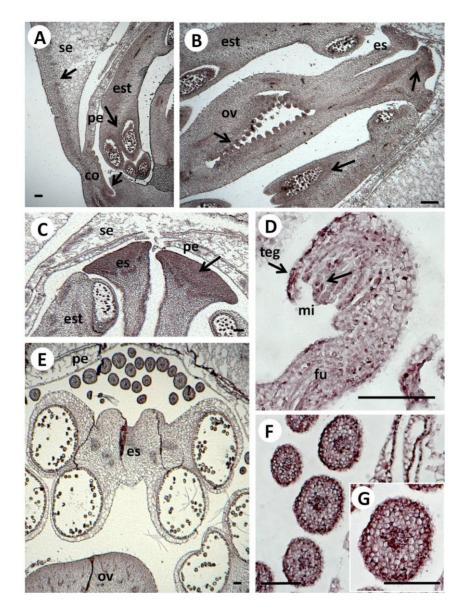


Figura 19: Hibridização *in situ* em botões florais seccionados longitudinal e transversalmente, mostrando o padrão de expressão de PACEPE3010A01.g, durante o desenvolvimento floral de *P. edulis*. (A) Corte longitudinal de botão com 0,3 a 0,5 mm, observa-se sinal de hibridização na sépala, corona e anteras. (B) Corte longitudinal de botão do primeiro estágio com sinal de hibridização nos óvulos e no estigma. (C) Detalhe da presença de transcritos na região do estigma, no carpelo. (D) Detalhe da expressão gênica no óvulo, encontrado num corte transversal de botão floral com 1,1-2,0 cm. (E) Corte transversal de botão do terceiro estágio, no desenvolvimento tardio da corona, observa-se sinal mais evidente nos filamentos mais desenvolvidos da corona. (F) e (G) Detalhes da expressão na corona em corte transversal. As setas evidenciam a presença de transcritos. **co**:corona, **es:** estigma, **est**: estame, **fu**: funículo, **mi**: micrópila, **pe**: pétala, **ov**: avario, **se**: sépala, **teg**: tegumento. Barras = 100 μm.

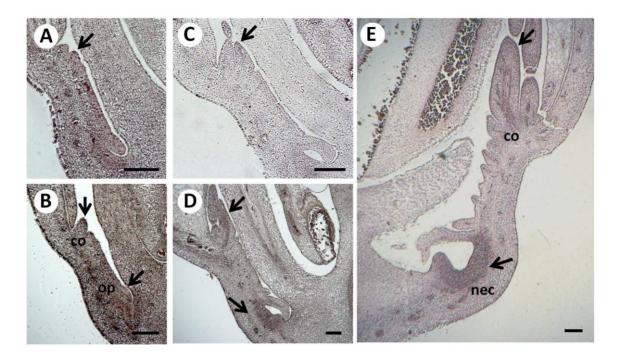


Figura 20: Hibridização *in situ* em botões florais seccionados longitudinalmente, mostrando o padrão de expressão de PACEPE4001G10.g, durante o desenvolvimento da corona em *P. edulis*. (A) Sinal de hibridização durante desenvolvimento da emergência da corona. (B) A coloração rosada indica a presença de transcritos no ápice da primeira série de filamentos da corona em formação, no opérculo, no nectário em formação e nas células epidérmicas da sépala. (C) Terceiro estágio do desenvolvimento com sinal mais fraco, mas concentrado no ápice dos filamentos em desenvolvimento. (D) Continuação do desenvolvimento da corona com precipitado em tecidos similares ao anterior. (E) No desenvolvimento tardio da corona, sinal de hibridização concentrado no nectário e nos filamentos mais desenvolvidos da corona. As setas evidenciam a presença de transcritos. **co**:corona, **op**: opérculo, **nec**: nectário. Barras = 200 μm.

Observou-se ainda a expressão de PACEPE4001G10.g em outros órgãos florais, assim como visto na RT-PCR. Pode-se perceber a presença de transcritos deste gene nos estames e pistilo (Figura 21A). As figuras 21B e 21D mostram cortes longitudinais de botão floral em diferentes estágios do desenvolvimento e evidenciam a expressão nos óvulos e estigma, além do tecido de condução no estilete. Em estágios tardios do desenvolvimento da antera, observou-se a presença de transcritos de PACEPE4001G10.g

nos tecidos do tapete e em grãos de pólen (Figura 21F). Cortes transversais dos filamentos da corona (Figuras 21C e 21E) indicaram a presença de transcritos nas regiões epidérmica e pró-vasculares.

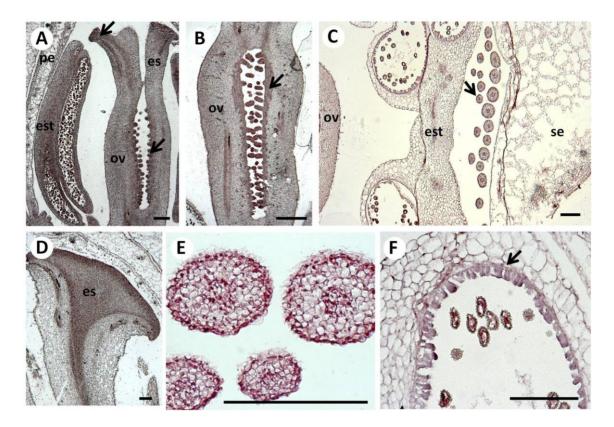


Figura 21: Hibridização *in situ* em botões florais seccionados longitudinal e transversalmente, mostrando o padrão de expressão de PACEPE4001G10.g, durante o desenvolvimento floral de *P. edulis*. (A) Corte longitudinal de botão com 0,3 a 0,5 mm, observa-se sinal de hibridização no ovário, estames e no estigma. (B) Corte longitudinal de botão do primeiro estágio com sinal de hibridização nos óvulos. (C) Corte transversal de botão do terceiro estágio com 1,1-2,0 cm., no desenvolvimento tardio da corona, observa-se sinal mais evidente nos filamentos da corona. (D) Detalhe da presença de transcritos na região do estigma, no carpelo, em corte longitudinal. (E) Detalhes da expressão na corona em corte transversal. (F) Detalhe da expressão do gene nas células do tapete da antera. As setas evidenciam a presença de transcritos. **es:** estigma, **est**: estame, **pe**: pétala, **ov**: ovário, **se**: sépala. Barras = 200 μm.

5. CONCLUSÃO

A identificação dos fatores de transcrição (FTs) presentes na biblioteca de cDNAs gerada pelo projeto PASSIOMA resultou na observação de 139 genes putativos. A triagem dos genes pertencentes a cada espécie de *Passiflora* usada no projeto PASSIOMA mostrou que 49,6% (69 genes) pertencem a *P. edulis*, 48,9% (68 genes) pertencem a *P. suberosa* e 1,4% (2 genes) são de *P. pholi*.

As análises filogenéticas permitiram a identificação da maioria dos fatores de transcrição encontrados. Comparando as sequências encontradas no PASSIOMA com as obtidas do banco de dados TAIR para cada família gênica estudada, foi possível a determinação de ortologias potenciais de FTs descobertos em *Passiflora* com FTs de função conhecida em *Arabidopsis*. Foram identificados ESTs oriundos de botões florais de *Passiflora* potencialmente ortólogos de genes de *Arabidopsis* cujo padrão de expressão só foi descrito em órgãos vegetativos.

O experimento de macroarranjo mostrou que 9 dos 69 genes codificadores de FT existentes no PASSIOMA para *P. edulis* são expressos na corona. A RT-PCR mostrou ainda que para 5 desses genes há expressão não só na corona, mas em todos os órgãos florais, sugerindo várias funções para as proteínas codificadas por eles. Por fim, a hibridização *in situ* realizada para PACEPE3010A01.g e PACEPE4001G10.g mostrou detalhes sobre o padrão de expressão desses genes durante o desenvolvimento da corona e em outros órgãos florais. Isso sugere, que esses FT podem estar envolvidos no desenvolvimento e/ou metabolismo de diferentes órgãos florais, além da corona.

Este trabalho estabelece um primeiro registro da identificação sistemática de fatores de transcrição em *Passiflora* spp. e provê bases para estudos futuros da função de FT no desenvolvimento de *P. edulis*.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Airoldi CA e Davies B. 2012. "Gene duplication and the evolution of plant MADS-box transcription factors". *J Genet Genomics*. 39:157-165.
- Alonso M, Chan A, Chang C, Overvoorde PJ, Okushima Y, Ecker JR, Hughes B. 2005. "Functional Genomic Analysis of the AUXIN/INDOLE-3-ACETIC ACID Gene Family Members in *Arabidopsis thaliana*". *Plant Cel.* 17:3282-3300.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. "Basic local alignment search tool". *Journal Molecular Biology*. 215:403-410.
- Barr MS, Willmann MR, Jenik PD. 2012. "Is there a role for trihelix transcription factors in embryo maturation?" *Plant Signal Behav*. 7:205-209.
- Becker A, Alix K, Damerval C. 2011. "The evolution of flower development: current understanding and future challenges". *Ann Bot*. 107:1427-31.
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2002. "GenBank." *Nucleics Acid Reserch*. 30:17-20.
- Benson WW, Brown Jr KS, Gilbert LE. 1975. "Coevolution of plants and herbivores: passion flower butterflies". *Evolution*. 29:659-680.
- Bernhard A. 1999. "Flower structure, development and systematics in Passifloraceae and *Abatia* (Flacourtiaceae)." *International Journal of Plant Sciences*. 160:135-150.
- Bowen AJ, Gonzalez D, Mullins JG, Bhatt AM, Martinez A, Conlan RS. 2010. "PAH-domain-specific interactions of the *Arabidopsis* transcription coregulator SIN3-LIKE1 (SNL1) with telomere-binding protein 1 and ALWAYS EARLY2 Myb-DNA binding factors". *J Mol Biol*. 395:937-49.

- Breuer C, Kawamura A, Ichikawa T, Tominaga-Wada R, Wada T, Kondou Y, Muto S, Matsui M, Sugimoto K. 2009. "The trihelix transcription factor GTL1 regulates ploidy-dependent cell growth in the *Arabidopsis* trichome". *Plant Cell*. 21:2307-22.
- Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, Bilbao-Castro JR, Robertson DL. 2010. "Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae". *Plant physiology*. 153:1398–412.
- Castelain M, Hir RL, Bellini C. 2012. "The non-DNA-binding bHLH transcription factor PRE3/bHLH135/ATBS1/TMO7 is involved in the regulation of light signaling pathway in *Arabidopsis*". *Physiologia plantarum*. 145:450–60.
- Cervi AC. 1997. "Passifloraceae do Brasil. Estudo do gênero *Passiflora*". *Fontqueria*. 45:1-92.
- Cervi AC. 2005. "Passiflora L. (Passifloraceae) species published and described in the last 55 years (1950-2005) in South America". Estud. Biol. 27: 19-24.
- Cominelli E e Tonelli C. 2009. "A new role for plant R2R3-MYB transcription factors in cell cycle regulation". *Cell Res.* 19:1231-2.
- Crepet WL. 1983. "The role of pollination in the evolution of the angiosperms. In Pollination biology". *Academic Press*. Orlando, p.29-50.
- Cubas P. 2004. "Floral zygomorphy, the recurring evolution of a successful trait". *BioEssays*. 26:1175-1184.
- Cutri L e Dornelas MC. 2012. "PASSIOMA: Exploring Expressed Sequence Tags during Flower Development in Passiflora spp". *Comparative and functional genomics*. 2012:510549.
- Dornelas MC, Wittich P, Von Recklinghausen I, Van Lammeren AAM, Kreis M. 1999. "Characterization of three novel members of the *Arabidopsis* SHAGGY-related protein kinase (ASK) multigene family". *Plant Molecular Biology*. 39:137-147.
- Dornelas MC, Van Lammeren AAM, Kreis M. 2000. "Arabidopsis thaliana SHAGGY related protein kinases (AtSK11 and 12) function in perianth and gynoecium development". *Plant Journal*. 21:419-429.

- Dornelas MC, Tsai SM, Rodriguez APM. 2006. "Expressed sequence tags of genes involved in the flowering process of *Passiflora* spp". *Ornamental and Plant Biotechnology*. 1:483-488.
- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. "MYB transcription factors in *Arabidopsis*". *Trends in plant science*. 15:573–81.
- Endress PK. 1994. "Diversity and evolutionary biology of tropical flowers". *Cambridge University Press*, Cambridge.
- Faegri K e Pijl L. 1971. "The principles of pollination ecology". 2 ed., Oxford, Pergamon.
- Ferrandiz C. 2000. "Negative Regulation of the SHATTERPROOF Genes by FRUITFULL During *Arabidopsis* Fruit Development". *Science*. 289:436–438.
- Fujita K, Horiuchi H, Takato H, Kohno M, Suzuki S. 2012. "Auxin-responsive grape Aux/IAA9 regulates transgenic *Arabidopsis* plant growth". *Molecular biology reports*. 39:7823–9.
- Ganga RMD, Ruggiero C, Lemos EGM, Grili GVG. 2004. "Genetic diversity in yellow passion fruit utilizing fAFLP molecular markers." *Rev. Bras. Frutic.* 26: 494-498.
- Grasser KD. 2003. "Chromatin-associated HMGA and HMGB proteins: versatile coregulators of DNA-dependent processes". *Plant Mol Biol.* 53:281-95.
- Grasser KD, Launholt D, Grasser M. 2007. "High mobility group proteins of the plant HMGB family: dynamic chromatin modulators." *Biochimica et biophysica acta*. 1769:346–57.
- Guilfoyle TJ e Hagen G. 2007. "Auxin response factors". *Current opinion in plant biology*. 10:453–60.
- Halbach T, Scheer N, Werr W. 2000. "Transcriptional activation by the PHD finger is inhibited through an adjacent leucine zipper that binds 14-3-3 proteins." *Nucleic acids research*. 28:3542–50.
- Hemingway CA, Christensen AR, Malcomber ST. 2011. "B- and C-class gene expression during corona development of the blue passionflower (*Passiflora caerulea*, Passifloraceae)". *Am J Bot.* 98:923-34.

- Izhaki A e Bowman JL. 2007. "KANADI and class III HD-Zip gene families regulate embryo patterning and modulate auxin flow during embryogenesis in *Arabidopsis*". *Plant Cell*. 19:495-508.
- Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F. 2002. "bZIP transcription factors in *Arabidopsis*". *Trends in plant science*. 7:106–11.
- Janzen DH. 1968. "Reproductive behavior in the Passifloraceae and some of its pollinators in the Central America". *Behavior*, 32:33-48.
- Karas B, Amyot L, Johansen C, Sato S, Tabata S, Kawaguchi M, Szczyglowski K. 2009. "Conservation of lotus and *Arabidopsis* basic helix-loop-helix proteins reveals new players in root hair development". *Plant physiology*. 151:1175–85.
- Kelley DR, Arreola A, Gallagher TL, Gasser CS. 2012. "ETTIN (ARF3) physically interacts with KANADI proteins to form a functional complex essential for integument development and polarity determination in *Arabidopsis*". *Development*. 139:1105-1109.
- Killip EP. 1938. "The American species of Passifloraceae. Publications of the Field Museum of Natural History". *Botanical Series*. 19:1–613.
- Knuth P. 1904. "Handbuch der Bluetenbiologie III". Engelmann, Leipzig.
- Kranz H, Scholz K, Weisshaar B. 2000. "c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages". *Plant J.* 21:231-235.
- Kumar M, Busch W, Birke H, Kemmerling B, Nürnberger T, Schöffl F. 2009. "Heat shock factors HsfB1 and HsfB2b are involved in the regulation of Pdf1.2 expression and pathogen resistance in *Arabidopsis*". *Molecular plant*. 2:152–65.
- Kwak K, Kim JY, Kim YO, Kang H. 2007. "Characterization of transgenic *Arabidopsis* plants overexpressing high mobility group B proteins under high salinity, drought or cold stress". *Plant & cell physiology*.48:221–31.
- Lee WY, Lee D, Chung WI, Kwon CS. 2009. "Arabidopsis ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers". Plant J. 58:511-24.

- Li J, Jia D, Chen X. 2001. "HUA1, a regulator of stamen and carpel identities in *Arabidopsis*, codes for a nuclear RNA binding protein". *Plant Cell*. 13:2269-81.
- Lildballe DL, Pedersen DS, Kalamajka R, Emmersen J, Houben A, Grasser KD. 2008. "The expression level of the chromatin-associated HMGB1 protein influences growth, stress tolerance and transcriptome in *Arabidopsis*". *J Mol Biol.* 384:9-21.
- Liscum E e Reed JW. 2002. "Genetics of Aux/IAA and ARF action in plant growth and development." *Plant molecular biology*. 49:387–400.
- McCarty DR e Chory J. 2000. "Conservation and innovation in plant signaling pathways". *Cell*. 103: 201-209.
- Melzer R, Wang Y-Q, Theissen G. 2010. "The naked and the dead: The ABCs of gymnosperm reproduction and the origin of the angiosperm flower". *Seminars in Cell & Developmental Biology*. 21:118-128.
- Mondragon-Palomino e Theissen G. 2008. "MADS about the evolution of orchid flowers". *Trends in Plant Science*. 13: 5-19.
- Moreno-Risueno MA, Van Norman JM, Benfey PN. 2012. "Transcriptional switches direct plant organ formation and patterning". *Curr Top Dev Biol.* 98:229-257.
- Ohman D, Demedts B, Kumar M, Gerber L, Gorzsás A, Goeminne G, Hedenström M, Ellis B, Boerjan W, Sundberg B. 2012. "MYB103 is required for *FERULATE-5-HYDROXYLASE* expression and syringyl lignin biosynthesis in *Arabidopsis* stems". *Plant J.* 73: 63–76
- Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui A, Nguyen D, Onodera C, Quach H, Smith A, Yu G, Theologis A. 2005. "Functional genomic analysis of the *AUXIN RESPONSE FACTOR* gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19." 17:444–463.
- Parenicová L, Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L. 2003. "Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: New openings to the MADS world". *Plant Cell*. 15:1538-1551.

- Pedersen DS, Merkle T, Marktl B, Lildballe DL, Antosch M, Bergmann T, Tönsing K, Anselmetti D, Grasser KD. 2010. "Nucleocytoplasmic distribution of the *Arabidopsis* chromatin-associated HMGB2/3 and HMGB4 proteins". *Plant physiology*. 154:1831–41.
- Peng X, Zhao Y, Cao J, Zhang W, Jiang H, Li X, Ma Q, Zhu S, Cheng B. 2012. "CCCH-type zinc finger family in maize: genome-wide identification, classification and expression profiling under abscisic acid and drought treatments". *PLoS One*. 7:e40120.
- Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. 2000. "B and C floral organ identity functions require *SEPALLATA* MADS-box genes". *Nature*. 405(6783):200-3.
- Pires ND e Dolan L. 2012. "Morphological evolution in land plants: new designs with old genes". *Philos Trans R Soc Lond B Biol Sci.* 367:508-518.
- Preston JC, Martinez CC, Hileman LC. 2011. "Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome". *Proc Natl Acad Sci.* 108:2343-8.
- Puri V. 1948. "Studies in floral anatomy. V. On the structure and nature of the corona in certain species of the Passifloraceae". *J. Ind. Bot. Soc.* 27:130-149.
- Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G. 2000. "*Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes". *Science*. 290:2105-10.
- Rinaldi M, Liu J, Enders T, Bartel B, Strader LC. 2012. "A gain-of-function mutation in IAA16 confers reduced responses to auxin and abscisic acid and impedes plant growth and fertility". *Plant molecular biology*. 79:359–73.
- Rozen S e Skaletsky HJ. 1998. "Software: Primer3." http://www.genome.wi.mit.edu/genome_software/other/primers3.html

- Saiga S, Furumizu C, Yokoyama R, Kurata T, Sato S, Kato T, Tabata S, Suzuki M, Komeda Y. 2008. "The *Arabidopsis OBERON1* and *OBERON2* genes encode plant homeodomain finger proteins and are required for apical meristem maintenance". *Development*. 135:1751-9.
- Saiga S, Möller B, Watanabe-Taneda A, Abe M, Weijers D, Komeda Y. 2012. "Control of embryonic meristem initiation in *Arabidopsis* by PHD-finger protein complexes". *Development*. 139:1391-8.
- Saitou N e Nei M. 1987. "The neighbour joining method: a new method for reconstructing phylogenetic trees". *Molecular Biology Evolution*. 4: 406-425.
- Santos AP, Serra T, Figueiredo DD, Barros P, Lourenço T, Chander S, Oliveira MM, Saibo NJ. 2011. "Transcription regulation of abiotic stress responses in rice: a combined action of transcription factors and epigenetic mechanisms". *OMICS*. 15:839-857.
- Sazima M e Sazima I. 1978. "Bat pollination of the passion flower, *Passiflora mucronata*, in southeastern Brazil". *Biotropica*. 10:100-9.
- Sazima I e Sazima M. 1989. "Mamangavas e irapuás (Hymenoptera, Apoidea): visitas, interações e consequências para a polinização do maracujá (Passifloraceae)". *Revista Brasileira de Entomologia*. 33:108-118.
- Semir J e Brown Jr KS. 1975. "Maracujá: a flor da paixão". Revista Geográfica Universal. 2:40-47.
- Smalle J, Kurepa J, Haegman M, Gielen J, Van Montagu M, Van Der Straeten D. 1998. "The trihelix DNA-binding motif in higher plants is not restricted to the transcription factors GT-1 and GT-2". *Proc Natl Acad Sci USA*. 95:3318-22.
- Stros M. 2010. "HMGB proteins: interactions with DNA and chromatin". *Biochim Biophys Acta*. 1799:101-13.
- Szécsi J, Joly C, Bordji K, Varaud E, Cock JM, Dumas C, Bendahmane M. 2006. "BIGPETALp, a bHLH transcription factor is involved in the control of *Arabidopsis* petal size". *The EMBO journal*. 25:3912–20.
- Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H. 2000. "A short history of MADS-box genes in plants". *Plant Molecular Biology*. 42:115-149.

- Theissen G e Saedler H. 2001. "Floral quartets". Nature. 409:469-471.
- Thompson JD, Higgins DG, Gibson TJ. 1994. "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice". *Nucleic Acids Res.* 22:4673-4680.
- Thomson JD e Wilson P. 2008. "Explaining evolutionary shifts between bee and hummingbird pollination: convergence, divergence and directionality". *International Journal Plant Sciences*. 169:23-38.
- Tominaga-Wada R, Nukumizu Y, Sato S, Kato T, Tabata S, Wada T. 2012. "Functional divergence of MYB-related genes, *WEREWOLF* and *AtMYB23* in *Arabidopsis*". *Biosci Biotechnol Biochem*, 76:883-887.
- Ulmer T e MacDougal JM. 2004. "Passiflora, Passion flowers of the world". Timber Press, Cambridge, UK.
- Viana AJC, Souza MM, Araújo IS, Corrêa RX, Ahnert D. 2010. "Genetic diversity in *Passiflora* species determined by morphological and molecular characteristics". *Biologia Plantarum*. 54:535-8.
- Wang D, Guo Y, Wu G, Yang G, Li Y, Zheng C. 2008. "Genome-wide analysis of CCCH zinc finger family in *Arabidopsis* and rice". *BMC genomics*. 9:44.
- Wilde WJJO. 1974. "The genera of tribe *Passiflorae* (Passifloraceae), with special reference to flower morphology". *Blumea*. 22:37-50.
- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. "An 'Electronic Fluorescent Pictograph' browser for exploring and analyzing large-scalebiological data sets. *PLoS One*. 2:e718.
- Xu J, Yang C, Yuan Z, Zhang D, Gondwe MY, Ding Z, Liang W, Zhang D, Wilson ZA. 2010. "The ABORTED MICROSPORES regulatory network is required for postmeiotic male reproductive development in *Arabidopsis thaliana*". *The Plant cell*. 22:91–107.
- Zhou D. 1999. "Regulatory mechanism of plant gene transcription by GT-elements and GT-factors". *Trends in plant science*. 4:210–214.

Anexo

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de	minha dissertação de Mestrado/tese de	e Doutorado
intitulada "Caracterização de genes de fatores d	Commission of the Commission of Commission o	
edulis Sims. (Passifloraceae)"		
) não se enquadra no § 3º do Artigo 1º da I biossegurança. 	Informação CCPG 01/08, referente a bio	oética e
Tem autorização da(s) seguinte(s) Comissão(õe	es):	
(x) CIBio - Comissão Interna de Biosseguran	iça, projeto No. <u>2006/03</u> , In:	stituição:
() CEUA – Comissão de Ética no Uso de An	imais , projeto No	Instituição:
() CEP - Comissão de Ética em Pesquisa, po	rotocolo No, Institu	uição:
Aluno: Conrado de	e Campos Gonaçalves Prof. Dr. Marcelo Carnier Dornelas Prof. Dr. Marcelo Carnier Dornelas	rneia:
Para uso da Comissão ou Comitê pertinente: () Deferido () Indeferido	Martelde	
Carimbo e assinatura	Prof. Dr. MARCELO LANCELLOTTI	
	Presidente da Comissão Interna de Biossegurança Instituto de Biologia - UNICAMP	
Para uso da Comissão ou Comitê pertinente: () Deferido () Indeferido	Instituto se pietogia - UNICAMP	

Carimbo e assinatura