

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
CURSO DE GRADUAÇÃO BACHARELADO EM CIÊNCIAS BIOLÓGICAS

ESTUDO DA FUNÇÃO DO GENE CYP705A2 E SEUS IMPACTOS
FISIOLÓGICOS EM *Arabidopsis thaliana*

Raul Simon Batista

Orientador: Felipe dos Santos Maraschin

Porto Alegre

2024

RAUL SIMON BATISTA

ESTUDO DA FUNÇÃO DO GENE CYP705A2 E SEUS IMPACTOS
FISIOLÓGICOS EM ARABIDOPSIS THALIANA

Trabalho de conclusão de curso como
requisito parcial à obtenção do título de
Bacharel em Ciências Biológicas da
Universidade Federal do Rio Grande do Sul.

Orientador: Felipe dos Santos Maraschin.

Porto Alegre

2024

FICHA CATALOGRÁFICA

CIP - Catalogação na Publicação

Simon Batista, Raul
ESTUDO DA FUNÇÃO DO GENE CYP705A2 E SEUS IMPACTOS
FISIOLÓGICOS EM ARABIDOPSIS THALIANA / Raul Simon
Batista. -- 2024.
42 f.
Orientador: Felipe dos Santos Maraschin.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Biociências, Bacharelado em Ciências Biológicas,
Porto Alegre, BR-RS, 2024.

1. CRISPR/Cas. 2. AT4G15350. 3. Fisiologia vegetal.
4. Edição gênica. 5. Metabólitos secundários. I. dos
Santos Maraschin, Felipe, orient. II. Título.

FOLHA DE APROVAÇÃO

Raul Simon Batista

ESTUDO DA FUNÇÃO DO GENE CYP705A2 E SEUS IMPACTOS FISIOLÓGICOS EM ARABIDOPSIS THALIANA

Trabalho de conclusão de curso como
requisito parcial à obtenção do título de
Bacharel em Ciências Biológicas da
Universidade Federal do Rio Grande do Sul.

Orientador: Felipe dos Santos Maraschin.

Porto Alegre, Agosto de 2024.

BANCA EXAMINADORA:

Prof. Dr. Felipe dos Santos Maraschin, professor Associado da Universidade Federal do Rio Grande do Sul (UFRGS).

Prof. Dr. Felipe Klein Ricachenevsky, professor Ajunto da Universidade Federal do Rio Grande do Sul (UFRGS).

Dr. Thomaz Stumpf Trenz; pesquisador de pós-doutorado do departamento de Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul.

AGRADECIMENTOS

Primeiramente, quero agradecer aos meus pais, que me ensinaram a ser como sou e sempre me incentivaram a persistir estudando, apesar da distância e saudade. A meus tios, Patrícia e Gilberto, e meu primo Bruno, por todo amor e carinho que me deram durante todos esses anos de convivência, serei eternamente grato por tudo que fizeram por mim.

Aos meus amigos de longa data, Índio, Vic, Gabi e Rafa, pelos encontros e reencontros que aqueceram meu coração, permitindo continuar mesmo que tão longe. Aos muitos amigos que fiz durante a graduação, em especial à Bianca, por ser minha inspiração, como pessoa e profissional, sendo uma das razões pelos últimos anos terem sido tão divertidos. Também quero agradecer a Laís, Marina, Mariana, Andressa, Luíza, Maicon e Eduarda, por toda a amizade.

Aos meus amigos do laboratório de microbiologia celular, aos quais me acolheram pelos primeiros dois anos da graduação, em especial à minha primeira orientadora, Fabiana, e à Simone, minha mãe científica, te agradeço pelo apoio e companheirismo. Aos muitos amigos do laboratório de fisiologia vegetal, que tornam a rotina mais leve e enriquecedora. Gostaria de agradecer, em especial, ao Jonata, Cibele e Fábio, por toda a ajuda, convivência e amizade, mas fundamentalmente por serem meus exemplos profissionais.

Agradeço imensamente ao melhor orientador que podia ter, Felipe dos Santos Maraschin, por ser uma pessoa extremamente paciente, dedicada e presente. Por ter me ensinado pessoalmente grande parte do que sei, por ter me dado a oportunidade de ter um projeto e fazer ciência ao seu lado, sem sua presença e liberdade, eu não seria o que sou hoje.

Gostaria de agradecer aos pesquisadores da Rovensa Next, com quem tive a oportunidade de conviver e trabalhar durante meu estágio. A experiência que tive, como parte da equipe, moldou minha maneira de trabalhar e observar a profissão de pesquisador, refletindo na construção deste trabalho e da minha vida profissional como cientista.

Este trabalho, assim como toda minha trajetória durante a graduação, só foi possível através das pessoas que me acompanharam, tanto ao meu lado quanto aquelas que estiveram distantes.

RESUMO

O desenvolvimento vegetal é amplamente influenciado por interações com fatores bióticos e abióticos. Nosso grupo de pesquisa estudou o papel de genes envolvidos na síntese de flavonoides do desenvolvimento de *Arabidopsis thaliana*. Comparações entre mutantes de perda de função revelaram que a linhagem mutante SALK093990, anteriormente anotada como isoflavona sintase putativa, mostrou aumento significativo no tamanho da parte aérea, comprimento da raiz e número de raízes laterais. Este mutante possui uma inserção de T-DNA no gene AT4G15350, que codifica a enzima da classe das P450 monooxigenases denominada da classe das P450 monooxigenases denominada CYP705A2. Embora o gene AT4G15350 tenha sido predito como uma isoflavona sintase, análises indicam que ele está mais associado à síntese de compostos terpênicos. Este gene faz parte de um cluster gênico, possivelmente envolvido na síntese de arabidiol e baruol, cujas funções são desconhecidas. Para elucidar a função deste gene, este estudo reuniu informações sobre a proteína e seu perfil de expressão utilizando bases de dados online. Foram construídos vetores para superexpressão e obtenção de alelos de perda de função via CRISPR/Cas9. Nossos resultados, até o momento, ainda não confirmam a participação do gene no fenótipo do mutante. Acredita-se que a alteração na linhagem SALK093990 possa impactar outros transcritos, como um lincRNA vizinho, ou processos não diretamente relacionados ao gene AT4G15350.

Palavras chave: *Arabidopsis thaliana*; Edição gênica; CRISPR/Cas; Citocromo monooxigenases; Terpenos.

ABSTRACT

Plant development is largely influenced by interactions with biotic and abiotic factors. Our research group studies the role of genes involved in the synthesis of flavonoids in the development of *Arabidopsis thaliana*. Comparisons between loss-of-function mutants revealed that the mutant line SALK093990, previously noted as a putative isoflavone synthase, showed significant increases in shoot size, root length, and number of lateral roots. This mutant has a T-DNA insertion in the AT4G15350 gene which encodes a P450 monooxygenase enzyme named CYP705A2. Although the AT4G15350 gene was predicted to be an isoflavone synthase, analyses indicate that it is more associated with the synthesis of terpene compounds. This gene is part of a gene cluster, possibly involved in the synthesis of arabidiol and baruol, whose functions are unknown. To elucidate the function of this gene, this study gathered information about the protein and its expression profile using online databases. Vectors were constructed for overexpression and obtaining loss-of-function alleles via CRISPR/Cas9. The results presented in this study, still do not confirm the gene's participation in the mutant's phenotype. It is believed that the change in the SALK093990 lineage may impact other transcripts, such as a neighboring lincRNA, or processes not directly related to the AT4G15350 gene.

Key words: *Arabidopsis thaliana*; Gene editing; CRISPR/Cas; Cytochrome monooxygenases; Terpenes.

LISTA DE FIGURAS

Figura 1: Locus do gene AT4G15350 explicitando o local do T-DNA e do lincRNA.....	13
Figura 2: Representação das regiões proteicas comuns as CYP.....	14
Figura 3: Amplitude de produtos sintetizados por uma mesma CYP.....	15
Figura 4: Metabolização do oxidosqualeno nos triterpenos a partir das enzimas..... presentes nos MGCs.....	17
Figura 5: Diferenciação dos produtos do oxidosqualeno em compostos com estrutura CBC e CCC.....	18
Figura 6: Representação do vetor pKIR1.1 de maneira simplificada.....	21
Figura 7: Representação dos passos usados para gerar plantas transformadas com o vetor pKIR1.1-Cas9sgRNA.....	22
Figura 8: Esquema representando a estrutura dos vetores pENTR3C e pEarleyGate103.....	23
Figura 9. Comparação plantas de <i>A. thaliana</i> do tipo selvagem e o mutante.....	25
Figura 10: Comprimento da raiz principal e número médio de raízes laterais por planta.....	26
Figura 11. Genotipagem de plantas mutante quanto à presença do T-DNA.....	26
Figura 12. Locus do gene AT4G15350.....	27
Figura 13. Eletroforese de DNA confirmando a clonagem do sgRNA1 no plasmídeo pKIR1.1-Cas9sgRNA1.....	28
Figura 14. Eletroforese de DNA confirmando a transformação do vetor pENTR3C- CDS4 por digestão.....	28
Figura 15. Eletroforese de DNA confirmando a presença do CDS vetor pENTR3C-CDS por PCR.....	29
Figura 16. Eletroforese de DNA com digestão do vetor pENTR3C-CDS4 usando enzima <i>Apal</i>	30
Figura 17. Eletroforese de DNA da confirmação a identidade do vetor pEG103- CDS4.....	30
Figura 18. Eletroforese de DNA das confirmações da presença das construções por PCR em <i>A. tumefaciens</i>	31
Figura 19. Seleção dos eventos transformados com a construção pKIR1.1-sgRNA1...31	31
Figura 20. Padrão de expressão AT4G15350 segundo Arabidopsis eFP Browser.....	32
Figura 21. Padrão de expressão do AT4G15350 com outros genes.....	33
Figura 22. Padrão de expressão do AT4G15350 de acordo com o Expression Atlas.....	33
Figura 23. Padrão de expressão tecidual do AT4G15350 de acordo com o trabalho de Huang <i>et al.</i> , 2019.....	34
Figura 24. Domínios proteicos preditos pelo InterPro.....	35
Figura 25. Representação da atuação do DMNT.....	37
Figura 26. Representação das rotas metabólicas dos 4 MGCs presentes em <i>A.</i> <i>thaliana</i>	38

LISTA DE ABREVIACÕES:

UTR: Untranslated Region; CYP: Cytochrome p450; NADH: Nicotinamide Adenine Dinucleotide; NADPH: Nicotinamide Adenine Dinucleotide Phosphate; OSCs: Oxidosqualeno ciclasas; PEN1: Pentacyclic Triterpene Synthase 1; BARS1: Baruol Synthase 1 ;IPP: Isopentenil difosfato; DMAPP: Difosfato de dimetilalilo; FPP: Farnesil pirofosfato; K: Aminoácido Lisina; E: Aminoácido Ácido glutâmico; R: Aminoácido Arginina; TPS: Terpeno sintase; MVA: Mevalonate; MEP: 2C-methyl-D-erythritol 4-phosphate; CCC: chair-chair-chair; pKIR: pKAMA-ITACHI Red; peg103: pEarleyGate 103; RFP: Red fluorescent protein; GRP: Green Fluorescent Protein; PCR: Polymerase Chain Reaction; LB: Luria Bertani; T-DNA: Transfer DNA; ul: Microlitro; CaMV: Cauliflower Mosaic Vírus; JÁ: Ácido Jasmônico; DMNT: (3E)-4,8-dimethyl-1,3,7-nonatriene; MGCs: Clusters de genes metabólicos; lincRNA: long intergenic non-coding RNA; ORF: Open Reading Frame; mRNA: RNA mensageiro;

SUMÁRIO

1. INTRODUÇÃO.....	11
1.1. Justificativa.....	11
2.1. DESENVOLVIMENTO.....	12
2.1.1. <i>Arabidopsis thaliana</i> e seus impactos na pesquisa.....	12
2.1.2. Relação do Fenótipo do mutante com a fotomorfogênese.....	12
2.1.3. Definição e funções das enzimas citocromos monooxigenases.....	13
2.1.4. Classificação das citocromo monooxigenases.....	15
2.1.5. Terpenos e triterpenos.....	15
2.2. OBJETIVOS.....	19
2.2.1. Objetivos Gerais.....	19
2.2.2. Objetivos específicos.....	19
2.3. MATERIAIS E MÉTODOS.....	20
2.3.1. Material vegetal e condições de cultivo.....	20
2.3.2. Genotipagens da linhagem SALK093990.....	20
2.3.3. Determinação da sequência do RNA guia e construção do vetor pKIR1.1-Cas9sgRNA1.....	20
2.3.4. Procedimentos para clonagem do gene AT4G15350.....	22
2.3.5. Transformação em <i>Arabidopsis thaliana</i>	23
2.3.6. Seleção e Genotipagem do material transformado.....	23
2.3.7. Análises de expressão.....	23
2.4. RESULTADOS.....	25
2.4.1. Experimento preliminar mutante SALK093990.....	25
2.4.2. Genotipagem SALK093990.....	26
2.4.3. Sequência RNAs guias.....	27
2.4.4. Construção do vetor pKIR1.1-Cas9sgRNA1.....	27
2.4.5. Construção do vetor pENTR3C-CDS4.....	28
2.4.6. Construção do vetor pEG103p103-CDS4.....	29
2.4.7. Transformação <i>Arabidopsis thaliana</i> e seleção do material transformado.....	31
2.4.8. Análise do padrão de expressão do gene AT4G15350.....	31
2.4.9. Análise da estrutura proteica do gene AT4G15350.....	32
2.4.10. Revisão dos integrantes do cluster de síntese do arabidiol/baruol.....	34
2.4.11. Revisão dos genes integrantes do cluster de síntese do arabidiol e baruol.....	35
2.5. DISCUSSÃO E PERSPECTIVAS.....	39
3. CONCLUSÃO.....	40
4. REFERÊNCIAS.....	41

1.INTRODUÇÃO

O desenvolvimento vegetal faz parte de uma complexa área de estudo, sendo altamente impactado por interações bióticas e abióticas. Com intuito de melhor compreender o processo, trabalhos anteriores objetivaram estudar o impacto de algumas mutações de genes envolvidos na rota de síntese de flavonóides no desenvolvimento de *Arabidopsis thaliana*.

Resultados anteriores do nosso grupo obtiveram uma série de comparações entre mutantes perda de função para diversos pontos da rota de síntese dos flavonóides. Uma das linhagens mutantes, que estava anotada como uma isoflavona sintase putativa, apresentou plântulas com um expressivo aumento do tamanho tanto da parte aérea, quanto do comprimento da raiz, além de um maior número de raízes laterais. Esse mutante, denominado de SALK093990, possui uma inserção de T-DNA na região 3'UTR do gene AT4G15350, que codifica para uma enzima da superfamília das P450 monooxigenases possivelmente envolvida na síntese de compostos triterpênicos, denominada de CYP705A2.

Muitas das enzimas da superfamília das P450 monooxigenases participam da síntese e modificação de compostos fenólicos, hormônios, terpenos, além de outras moléculas de natureza distinta. Com mais de 16000 integrantes somente no reino vegetal, elas são capazes de oxigenar uma ampla gama de substratos, devido a complexação do oxigênio com o átomo de ferro do grupo heme que as compõem (NELSON, 2018).

Apesar da anotação preditiva do gene AT4G15350 como uma isoflavona sintase, nossas análises identificaram que esta enzima parece estar mais associada à síntese de compostos terpênicos. Além disso, este gene faz parte de um cluster gênico sintênico, co-expresso, possivelmente envolvido na rota de síntese de dois compostos dessa mesma classe: arabidiol e o baruol, com funções ainda desconhecidas.

1.1. Justificativa

O fenótipo drástico de maior crescimento observado no mutante SALK093990 sugere uma importante função do gene AT4G15350 com a produtividade vegetal. Sendo assim, objetivamos reunir as informações existentes relacionadas às características da proteína, assim como do perfil de expressão do gene por meio de bases de dados online, além de realizar experimentos de genômica funcional, com o objetivo de elucidar o papel desta mutação no desenvolvimento vegetal.

2. DESENVOLVIMENTO

2.1.1. *Arabidopsis thaliana* e seus impactos na pesquisa

Apesar de não apresentar importância agrônômica, o estudo em *Arabidopsis thaliana* permitiu avanços na ciência básica, sendo a partir dos anos 1980 amplamente utilizada como modelo de estudo para dicotiledônias (MEINKE *et al.*, 1998). Diferentes características fisiomorfológicas e genéticas corroboraram para sua disseminação como planta modelo dentro da comunidade científica, sendo alguns deles: o desenvolvimento em padrões regulares e conhecidos (BORON; VISSEBERG, 2014), ciclo de vida curto, a possibilidade de cultivo em meio estéril bem como em substrato simples, a possibilidade de autofecundação, o tamanho pequeno para cultivo *in vitro* dos indivíduos adultos, bem como de suas sementes, produzidas em grandes quantidades (MEYEROWITZ, 1989; TAIZ *et al.*, 2017).

Em 1996, múltiplos grupos de pesquisa na área da genômica vegetal estabeleceram esforços que culminaram no surgimento da Arabidopsis Genome Initiative (AGI), que se propôs a identificar e caracterizar todos os genes de *A. thaliana*. Após 4 anos, todos os genes tinham sido sequenciados, fazendo com que *A. thaliana* se tornasse a primeira planta a ter seu genoma completamente sequenciado (MEINKE *et al.*, 1998; TAGI, 2000).

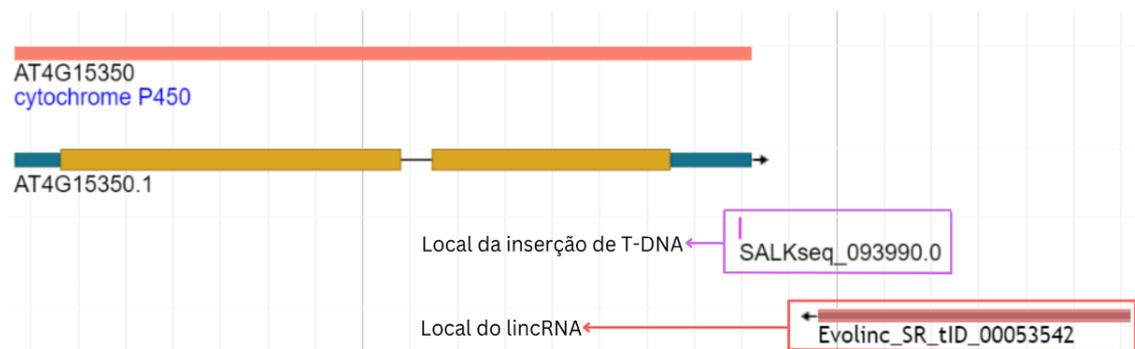
O conhecimento da identidade dos genes é fundamental, mas não suficiente, para o entendimento completo da sua função (ALONSO *et al.*, 2003). Para tanto, técnicas que modulam sua expressão, são essenciais para entendimento da participação fisiológica de seus produtos. Uma estratégia utilizada para ampliar o conhecimento sobre a função de seus genes foi o desenvolvimento de metodologias que nos permitiram gerar eventos mutagênicos em grande escala, sendo construídas coleções de mutantes para uma vasta quantidade de genes, como a mantida pelo *Salk Institute for Biological Studies*. A grande maioria dos genes mutados por tal metodologia tiveram suas funções inicialmente inferidas ao relacionar suas sequências a genes homólogos já caracterizados (O'MALLEY; BARRAGAN; ECKER, 2015; O'MALLEY; ECKER, 2010; MEINKE *et al.*, 1998; ALONSO *et al.*, 2003).

2.1.2. Relação do Fenótipo do mutante com a fotomorfogênese

Plântulas crescidas sob condições escuras apresentam pouco desenvolvimento radicular quando comparadas com plântulas crescidas na luz. Entretanto, quando elas se desenvolvem em condições naturais, com somente a parte aérea exposta à luz, um sistema radicular desenvolve-se normalmente abaixo do solo, na escuridão. A dinâmica entre diferentes regimes luminosos e suas consequentes respostas fisiológicas determinam o desenvolvimento vegetal, sendo aquele em resposta à exposição luminosa da parte aérea denominado de fotomorfogênico. Nosso grupo identificou, utilizando mutantes perda-de-função em *A. thaliana*, genes que potencialmente atuam como reguladores do desenvolvimento vegetal em resposta a luz, sendo vários deles relacionados também a rota de síntese de flavonóides. A análise fenotípica de plântulas mutantes para genes específicos desta rota indica que a síntese de flavonóides é muito importante para o desenvolvimento das raízes (CHRISTINO; MARASCHIN, 2023; MIOTTO *et al.*, 2019).

Parte destas análises consistiram na caracterização de mutantes perda-de-função para genes codificadores de enzimas-chave da rota de síntese dos flavonóides. Dentre as linhagens selecionadas, a SALK093990, associado a uma mutação no gene AT4G15350, foi selecionada devido a sua inicial anotação incomum de ‘isoflavona sintase putativa’, uma vez que *A. thaliana* não possui essa classe de enzimas (Lapcik *et al.*, 2006). Logo de início, este mutante se destacou por apresentar um aumento do tamanho tanto da parte aérea quanto da radicular em comparação com o selvagem (Col-0). Todavia, a razão pela qual tal mutante apresentou seu respectivo fenótipo ainda é desconhecida.

Figura 1. Lócus do gene AT4G15350 explicitando o local da inserção T-DNA e do lincRNA.



Esquema ilustrativo adaptado (jbrowse.arabidopsis.org), representando o locus do gene com a posição da inserção do T-DNA no mutante SALK093990 e de um lincRNA logo em seguida da posição do gene.

2.1.3. Definição e funções das enzimas citocromos monooxigenases

O produto proteico do gene associado à inserção do mutante SALK093990 está relacionado à superfamília das citocromo P450s monooxigenases, ou também chamadas de CYPs. Essa classe enzimática está presente em todos os domínios da vida, catalisando modificações químicas em uma ampla gama de compostos fisiologicamente diversos, como na formação de terpenos e seus derivados, alcaloides, fenilpropanoides (incluindo flavonas e isoflavonas), hormônios vegetais e fúngicos, estibenoides e compostos de origem complexa (Ortiz *et al.*, 2019). Em plantas, que possuem um grande número desses compostos e seus derivados, tal classe é a mais numerosa dentre os genes codificantes de enzimas (Nelson; Werck-Reichhart, 2011). Em humanos, esse grupo de enzimas é responsável por aproximadamente 95% das reações de oxidação e redução, atuando majoritariamente como monooxigenases com uso de NADH ou NADPH como cofatores (Rendic, G., 2015).

Todas as P450 apresentam um grupo heme, sendo que a estrutura conformacional que o cerca tende a ser mais conservada quanto maior sua relação filogenética (Ortiz *et al.*, 2019). No entanto, seus sítios catalíticos apresentam grande variação dependendo dos substratos com os quais elas interagem. Esse grupo de enzimas apresenta um dobramento proteico conservado, não existindo metaloproteínas de outras classes com tal estrutura (Ortiz *et al.*, 2019). Tais enzimas compartilham 4 elementos estruturais comuns: a hélice 1, responsável pela interação com átomo de oxigênio, uma dobra rica em prolina, uma hélice K, e uma sequência ‘PERF’, composta por uma ‘tríade ERR’ (ácido glutâmico, e duas argininas consecutivas) envolvida na estabilização estrutural do grupo heme (Du *et al.*, 2016). A ligação entre o resíduo de cisteína e o átomo de ferro do grupo heme, formando uma estrutura rígida conhecida como ‘protuberância beta’ (*Beta-bulge*) assim

como a hélice 1 anexa ao grupo heme, importante para que mais moléculas de água interajam com o sítio ativo enzimático, estão presentes em todas as CYPs, havendo pequenas variações entre alguns integrantes (Ortiz *et al.*, 2019).

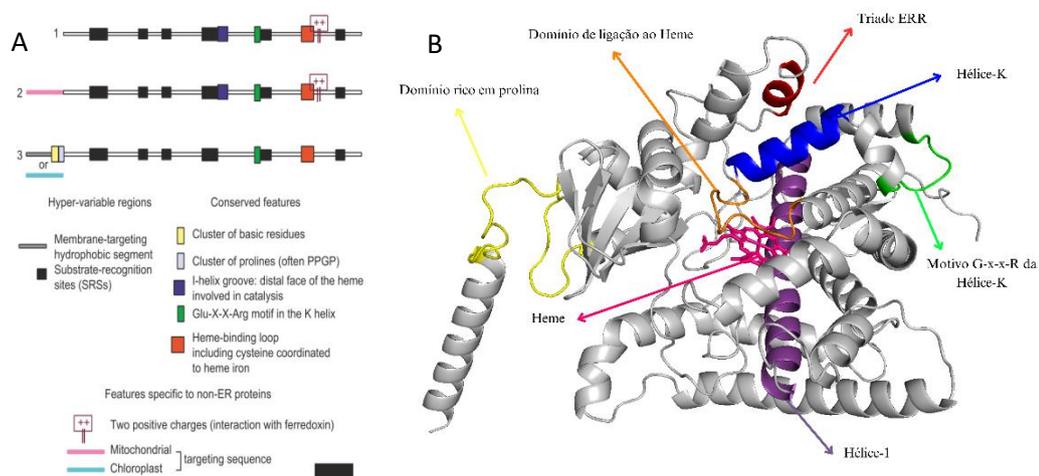


Figura 2: Representação das regiões proteicas comuns as CYPs. A-) Esquema representativo das regiões conservadas das P450, explicitando a dobra rica em prolinas em cinza, a hélice 1 em roxo, a hélice K em verde além da região de ligação ao Heme em laranja, adaptado de Werck-Reichhart *et al.*, 2000 (Werck-Reichhart, 2000). B-) Representação da proteína CYP705A2 com as respectivas seções: Amarelo: Domínio rico em prolina; Laranja: Domínio de ligação ao Heme; Vermelho: Triade ERR; Azul: Hélice-K; Verde: Motivo G-X-X-R da Hélice-K; Roxo: Hélice-1; Rosa: Agrupamento Heme, estrutura proteica obtida a partir do *AlphaFold* (<https://alphafold.ebi.ac.uk/>), com implementação de seus domínios estruturais a partir do software *PyMOL*.

Comumente, P450s eucarióticas tendem a apresentar domínios de interação com membranas, majoritariamente do retículo endoplasmático. Além disso, muitas P450s requerem grandes mudanças conformacionais para que seus substratos possam interagir com seus sítios catalíticos (Ortiz *et al.* 2019).

A grande plasticidade de reações catalisadas pelas citocromo P450, em parte, deriva da capacidade dessas metaloproteínas em complexarem um átomo de oxigênio com o ferro de seu grupo heme. Dessa forma, essas proteínas são capazes de oxidar átomos de carbono, heteroátomos e até mesmo outros elementos metálicos (Ortiz *et al.* 2019) fazendo parte da rota de síntese de uma ampla gama de metabólitos primários e secundários. Outra característica marcante dessa classe enzimática é a sua multiplicidade de substratos e produtos, variando tanto em formatos, quanto em tamanhos. Um exemplo é a baruol sintase, codificada pelo BARS1, que é capaz de catalisar a síntese de 23 compostos, tendo tal capacidade justificada pela gama de conformações estruturais dessas proteínas (Lodeiro *et al.*, 2007) (Guengerich, 2018; Nelson, 2018). Acredita-se que essa grande variação aumenta a capacidade de seres sésseis como as plantas de sobreviver às pressões ambientais (Bathe; Tissier, 2019). A promiscuidade das P450s também está relacionada à processividade de alguns integrantes dessa classe, ou seja, devido à atuação em múltiplas reações subsequentes dentro de uma rota metabólica. Assim, uma única CYP pode ser fundamental para a síntese de múltiplos produtos e, conseqüentemente, ter vários efeitos fisiológicos (Guengerich, 2018).

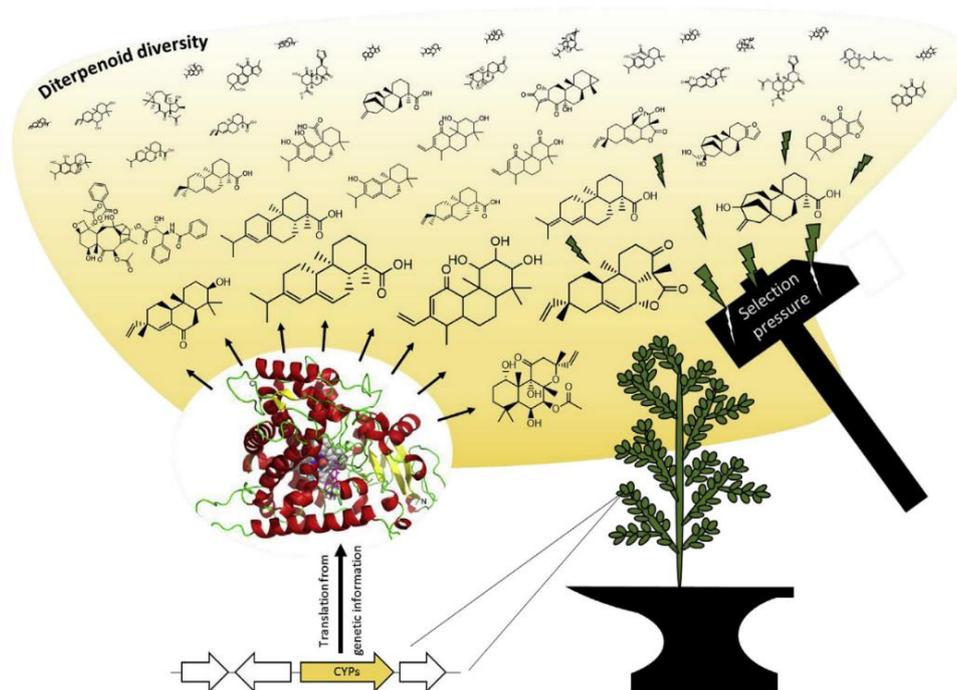


Figura 3: Amplitude de produtos sintetizados por uma mesma CYP. Esquema representativo da amplitude de produtos diterpênicos sintetizados por algumas citocromo monooxigenases, evidenciando o efeito indutor da expressão de enzimas da classe das P450 monooxigenases, com intuito de aumentar e diversificar a produção de compostos diterpênicos, por exemplo. Adaptado de Bathe e Tissier, 2019 (Bathe; Tissier, 2019).

2.1.4. Classificação das citocromo monooxigenases

Com 245 integrantes em *A. thaliana*, as CYPs estão classificadas em clãs, famílias, subfamílias de acordo com as suas sequências de aminoácidos e suas relações filogenéticas (Hansen *et al.*, 2021). Para as plantas vasculares, existem atualmente 11 clãs divididos em dois grupos: aqueles compostos por uma única família como o CYP51 e, aqueles que abrangem múltiplas famílias como CYP71. Além disso, em ambos os casos, sua nomenclatura abrange a sua isoforma, correspondendo a letra final de sua sigla (Werck-Reichhart; Bak; Paquette, 2002). O gene CYP705A2 pertence à família 705, exclusiva de *Brassicaceae* (Sohrabi *et al.*, 2015), inclusa dentro do clã 71, este conglomerado o maior número de CYPs vegetais, com mais de 60% do total (Nelson; Werck-Reichhart, 2011). Este clã contém diversas subfamílias envolvidas na síntese de compostos triterpênicos, incluindo o: CYP71A, CYP71D, CYP81Q, CYP93E e CYP705A. Alguns integrantes também estão envolvidos na modificação de monoterpenos e flavonóides (Ghosh, 2017; Hamberger; Bak, 2013).

2.1.5. Terpenos e triterpenos

Por meio da região genômica, bem como do conhecimento de genes vizinhos e seu padrão de expressão, a proteína CYP705A2, codificada pelo gene AT4G15350, tende a estar relacionada à síntese ou modificação de compostos terpênicos (Marszalek-Zenczak *et al.*, 2023, Wada *et al.*, 2012). Tais compostos, também chamados de isoprenóides, abrangem a maior classe de metabólitos secundários, estando presentes em todos os organismos. Compreendendo mais de 80.000 compostos, tais substâncias não

estão envolvidas em rotas metabólicas diretamente relacionadas ao desenvolvimento e crescimento vegetal, participando primariamente de respostas a sinais ambientais como: atração de polinizadores, proteção contra patógenos ou herbivoria e interações com indivíduos próximos (Ghosh, 2016; Bennett; Bentley, 1989; Zhong; Yue, 2005). Para além dos terpenos, existem outras duas grandes classes de metabólitos especializados também sintetizados por CYPs: os fenilpropanóides e os compostos nitrogenados, estes compreendendo alcalóides, glicosídeos cianogênicos e glicosinolatos (Hamberger; Bak, 2013). A grande diversidade dessas substâncias está intrinsecamente relacionada ao sucesso e ramificação das espécies, uma vez que determinam processos adaptativos a diferentes condições bióticas e abióticas espécie-específicas (Hamberger; Bak, 2013).

Através do trabalho de Wada em 2012, sabe-se que o gene AT4G15350 pertence a um dos quatro *clusters* de genes metabólicos (MGCs) (Wada *et al.*, 2012; Mizutani; Ohta, 2010). Tal classe é definida por apresentar pelo menos três genes, localizados de forma adjacente, que sintetizam enzimas envolvidas na síntese de produtos subsequentes. O AT4G15350 faz parte do cluster responsável pela síntese de dois compostos triterpênicos: o arabidiol e o baruol. Além disso, os outros três *clusters* presentes em *A. thaliana*, englobam a síntese do thalianol, marneral e do tirucalladienol, compostos derivados da modificação do 2,3-oxidosqualeno por enzimas, estruturalmente muito semelhantes, pertencentes ao grupo das P450, como ilustrado na Figura 4.

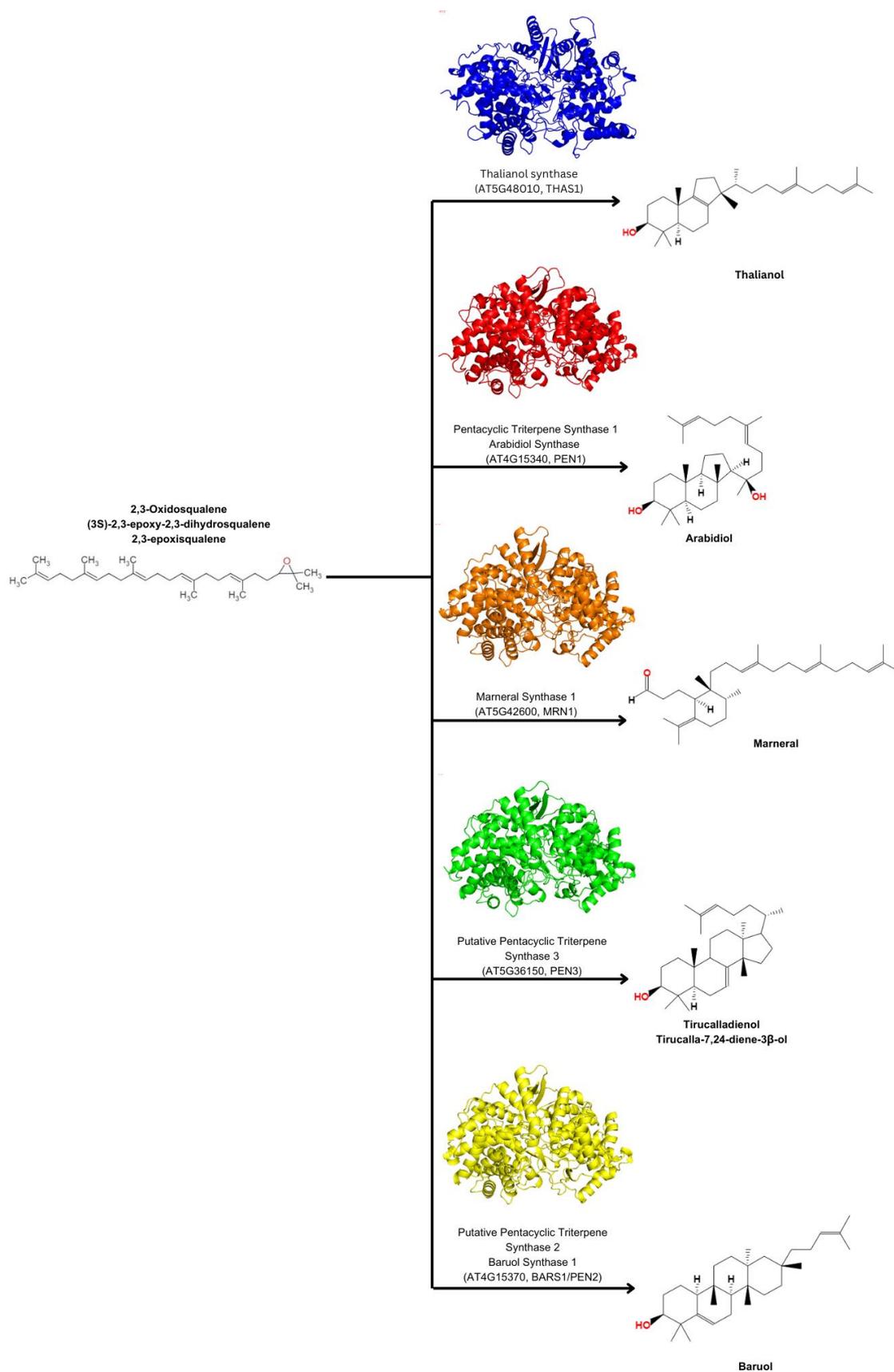
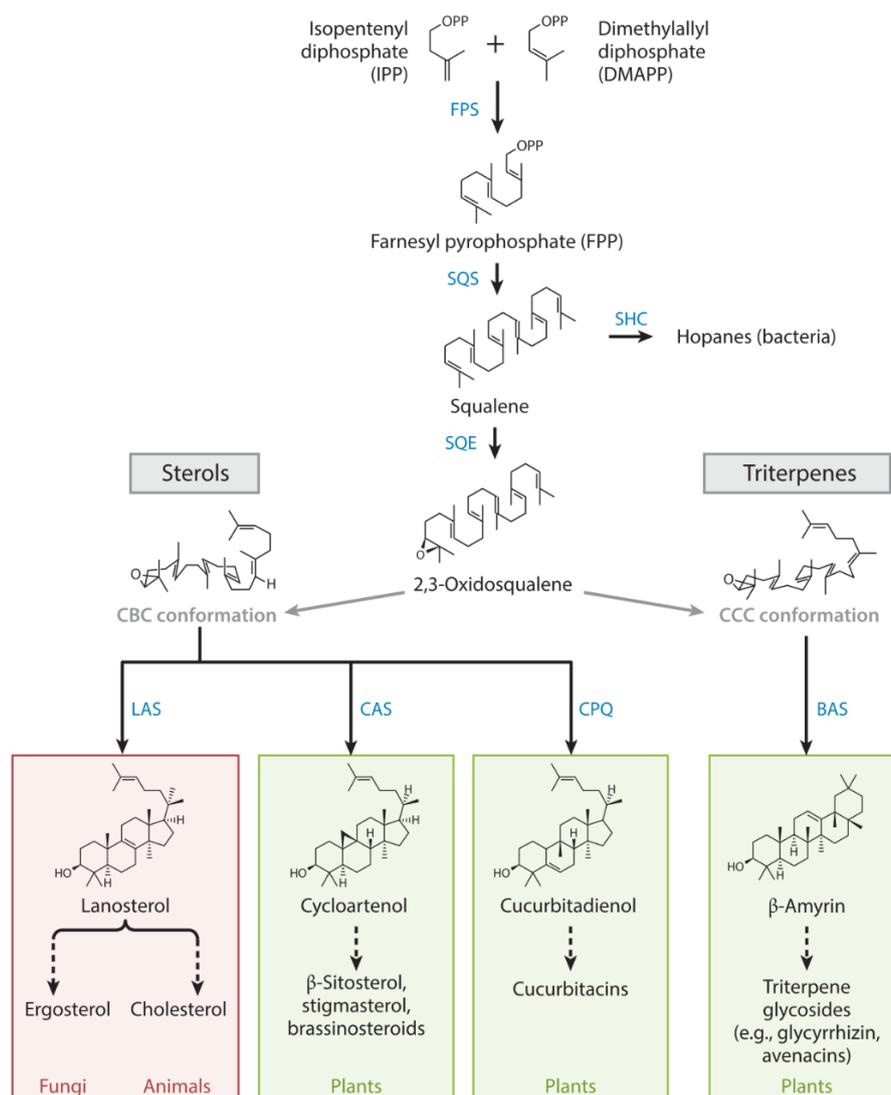


Figura 4. Metabolização do oxidosqualeno nos triterpenos a partir das enzimas presentes nos MGCs. Figura representativa das reações bioquímicas envolvidas na síntese de triterpenos agrupados em cluster de genes encontrados no trabalho de Wada, em 2012 (Wada, *et al.*, 2012; Marszalek-Zenczak *et al.*, 2023; Huang *et al.*, 2019; Bai *et al.*, 2021).

A partir do 2,3-oxidosqualeno, há uma ramificação da rota de síntese entre os triterpenos, com dobramento bioquímico denominado CCC (*Chair-Chair-Chair*), e os



esteróis com dobramento denominado CBC (*Chair-Boat-Chair*), como representado na figura 5. (Thimmappa *et al.*, 2014).

Figura 5. Diferenciação dos produtos do oxidosqualeno em compostos com estrutura CBC e CCC. Rota de síntese dos esteróis e triterpenos em plantas, explicitando os dois tipos de estruturas: CBC, CCC. Adaptado de Thimmappa *et al.*, 2014 (Thimmappa *et al.*, 2014).

A partir do fenótipo observado no mutante SALK093990, juntamente com o desconhecimento acerca de seu impacto fisiológico e, considerando seu possível envolvimento com a também inexplorada rota de síntese do arabidiol e baruol, compostos também pouco caracterizados, o trabalho busca reunir e aprofundar os conhecimentos específicos relacionados ao gene AT4G15350, tendo como finalidade a compreensão do fenômeno biológico observado.

2.2. OBJETIVOS

2.2.1. Objetivos gerais

Caracterizar o mutante SALK093990 através do estudo das possíveis funções do gene AT4G15350, codificante da proteína CYP705A2, no desenvolvimento de plântulas *Arabidopsis thaliana*.

2.2.2. Objetivos específicos

- Avaliar o fenótipo do mutante SALK093990.
- Gerar clones de entrada *Gateway* para a região codificadora e promotora do gene AT4G15350
- Desenhar e clonar RNAs guias (sgRNAs) visando a deleção do gene AT4G15350 pela metodologia de CRISPR/Cas9.
- Construção de vetores binários visando a superexpressão do gene.
- Transformação de *Arabidopsis thaliana* via *Agrobacterium tumefaciens* utilizando os vetores gerados.
- Genotipagem e seleção do material transformado.

2.3. MATERIAIS E MÉTODOS

2.3.1. Material vegetal e condições de cultivo

Plantas do ecótipo Col-0 de *Arabidopsis thaliana* foram usadas como linhagem controle, sendo obtida a partir da *The European Arabidopsis Stock Centre* (NASC, <http://arabidopsis.info/>). Já o mutante SALK093990, foi obtido a partir do *Arabidopsis Biological Resource Center* (ABRC, <https://abrc.osu.edu/>).

Para o cultivo *in vitro*, as sementes foram esterilizadas com álcool 70% por dois minutos, e posteriormente com uma solução de 2,5% hipoclorito e 0,04% de Tween20, seguida de cinco lavagens com água estéril. Após o processo de esterilização, as sementes foram estratificadas por frio (4°C) por dois dias no escuro completo anteriores ao plantio.

Para as plantas crescidas em condições *in vitro*, foi utilizado o meio de cultura adaptado de Murashige and Skoog (MURASHIGE; SKOOG, 1962) meia força, com 1% de sacarose quando indicado, sendo suplementado com 0.8% de ágar (w/v; Kasvi, K25-1800) e 0.5% de MES hidratado (w/v; Sigma Aldrich, M8250), pH 5.7. Todas as plantas foram crescidas a 21 °C ± 2 °C em fotoperíodo de dias longos (16 horas de luz e 8 horas de escuro) sob iluminação com LEDs brancas, resultando em uma intensidade luminosa de 90 μmol m⁻² s⁻¹ (Miotto *et al.*, 2019).

2.3.2. Genotipagens da linhagem SALK093990

Para os experimentos de genotipagem, realizou-se reações de PCR com *primers* que amplificariam na marca de resistência a Canamicina, presente em todos os T-DNAs das linhagens SALK093990, com o T-DNA idealmente presente nas linhagens mutantes. Os *primers* usados foram: NPT2 F e NPT2 R para a marca de resistência e, Lp_SALK093990, Rp_SALK093990 e Lbi1.3 para o T-DNA. As sequências dos *primers* podem ser encontradas na Tabela 2.

2.3.3. Determinação da sequência do RNA guia e construção do vetor pKIR1.1-Cas9sgRNA1

Os sgRNAs foram desenhados através do software *CRISPOR* (<http://crispor.tefor.net/>), tendo suas sequências listadas na Tabela 1. O vetor pKIR1.1 (Tsutsui H.; Higashiyama T., 2017) possui sítios de reconhecimento da enzima de restrição *AarI* (*Thermo Fisher Scientific*, #ER1581) adjacentes ao sítio de inserção do sgRNA. Este vetor se utiliza da tecnologia FAST (*Fluorescence-Accumulating Seed Technology*) que permite a expressão da proteína fluorescente OLE1-RFP (Oleosina 1 - Red Fluorescent Protein), permitindo a expressão específica da proteína repórter na semente (Shimada; Shimada; Hara-Nishimura, 2010). Tal vetor possibilita a visualização do material transformado através da emissão fluorescente ou também da utilização da resistência à higromicina como marca de seleção.

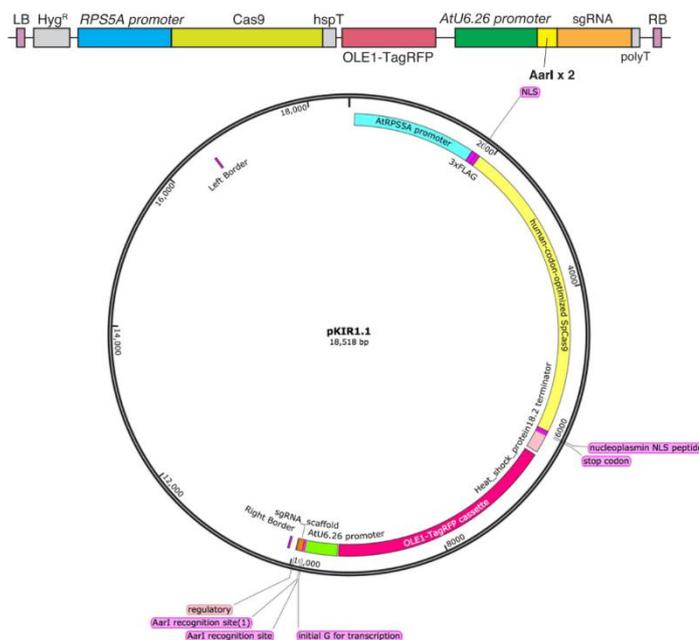


Figura 6: Representação do vetor pKIR1.1 de maneira simplificada. No esquema se evidencia a estrutura do T-DNA. O promotor RPS5A controla a expressão da Cas9, OLE1-TagRFP - oleosina fusionada a RFP (*Red Fluorescent Protein*) e o cassette de expressão do sgRNA a partir do promotor AtU6.26 com os sítios *AarI* adjacentes. Figura obtida do site *Addgene* (www.addgene.org) e Tsutsui H.; Higashiyama T., 2017.

A partir da digestão do pKIR1.1 com a enzima *AarI*, ocorre a formação de um *overhang* de quatro bases, permitindo a utilização dos oligonucleotídeos sgRNA1_Foward e sgRNA1_Reverse como uma extensão capaz de hibridizar nessa sequência. Após a clivagem do pKIR1.1 com *AarI* e sua subsequente purificação, os oligonucleotídeos hibridizados foram fosforilados com a enzima PNK (*Thermo Fisher Scientific*) e ligados usando a enzima T4 ligase (NEB), utilizado 150 ng do vetor e 50 ng do respectivo inserto por quatro horas à temperatura ambiente. Em seguida, este vetor foi transformado em células competentes de *Escherichia coli* da linhagem DH5 α (*Invitrogen*).

As colônias resultantes da transformação foram verificadas quanto a presença do vetor por meio da PCR, usando os *primers* PFGsgRNAf e o sgRNA1_Reverse, presentes na Tabela 2. Em seguida, foram realizadas purificações desses plasmídeos usando o *Kit PureLink™ Quick Plasmid Miniprep Kit (Invitrogen)*, sendo posteriormente sequenciados, utilizando 4,5 pMol do primer T7G RNA R, 150 ng do vetor, no equipamento *ABI-Prism 3500 Genetic Analyzer (Applied Biosystems)* pela empresa ACTGene, Porto Alegre/RS.

O plasmídeo confirmado por sequenciamento (pKII.1 sgRNA1), foi transformado em *A. tumefaciens* EHA105 por eletroporação (Main; Reynolds; Gartland, 1995). As colônias transformadas foram confirmadas por PCR utilizando os *primers* PFGsgRNAf e sgRNA.1_Reverse.

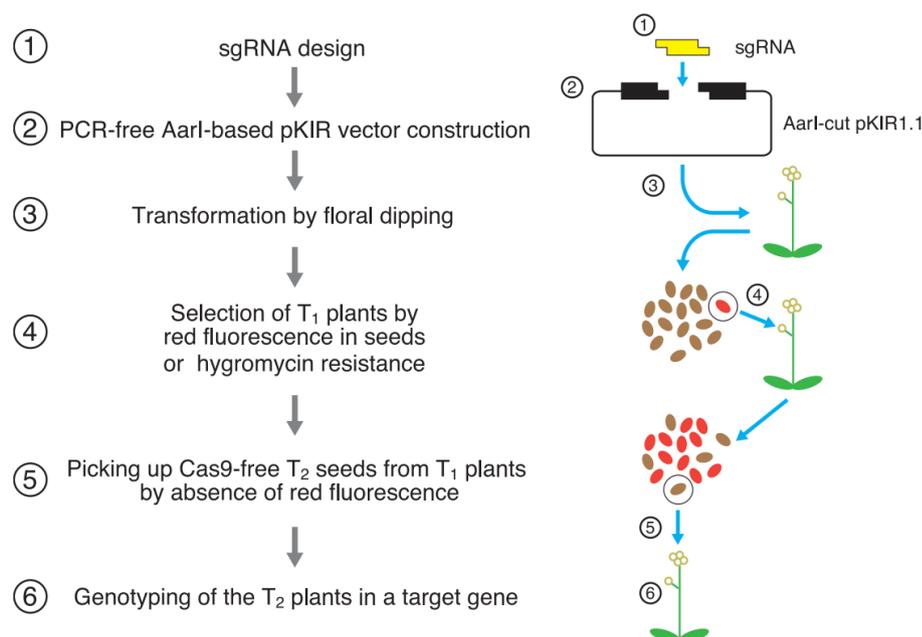


Figura 7: Representação dos passos usados para gerar plantas transformadas com o vetor pKIR1.1-Cas9sgRNA. Esquema representativo dos passos realizados para a construção das plantas mutadas por CRISPR/Cas9 utilizando o vetor pKIR1.1. Adaptado de Tsutsui H.; Higashiyama T., 2017 (Tsutsui H.; Higashiyama T., 2017).

2.3.4. Procedimentos para clonagem do gene AT4G15350

Visando gerar um clone de entrada para a superexpressão do gene CYP705A2, um fragmento contendo a região genômica contendo éxons e íntrons, removendo o stop códon no último éxon do gene AT4G15350, foi obtido a partir da PCR utilizando o DNA genômico de *A. thaliana* usando os *primers* CYP705A_CDS_No_Stop_F e CYP705A_CDS_No_Stop_R, exibidos na Tabela 2. Em seguida o produto foi submetido a uma reação de *Gibson Assembly* (Gibson *et al.*, 2009) com o plasmídeo pENTR3C (*Invitrogen*) previamente linearizado com *EcoRI* (*Thermo Fisher Scientific*). As quatro colônias resultantes da transformação foram inoculadas em meio LB líquido, sofrendo uma posterior purificação de seus plasmídeos de acordo com o kit: *PureLink™ Quick Plasmid Miniprep Kit* (*Invitrogen*; K210010).

A verificação da construção foi feita inicialmente a partir de uma reação de PCR usando os *primers*: AT4G15350 RTqPCR Forward e AT4G15350 RTqPCR Reverse. A identidade do plasmídeo recombinante foi posteriormente confirmada pelo sequenciamento utilizando os *primers*: CYP705A_CDS_No_Stop_F e CYP705A_CDS_No_Stop_R, todos os *primers* se encontram na Tabela 2.

O cassete de expressão foi construído a partir do plasmídeo binário pEarleyGate103 (Earley *et al.*, 2006) através da reação com LR Clonase com o clone de entrada pENTR3C CDS. Esta construção permite expressar a proteína CYP705A2 fusionada a GFP e a uma *tag* de histidina na porção C-terminal, sob o controle do promotor constitutivo CaMV 35S. Para essa construção, o vetor ‘pENTR3C-CDS4’ foi previamente clivado pela enzima *ApaI* para liberar um fragmento de aproximadamente 2.2kb que contém o CDS e os sítios de clonase attL1 e attL2. Em seguida, esse fragmento foi submetido a uma reação de clonase com a LR clonase II (*Invitrogen*; 11791020),

seguindo o protocolo de *Gateway® pENTR™ Dual Selection Vectors* (Invitrogen, 2008). A construção resultante foi confirmada a partir de uma digestão com a enzima *BamHI*.

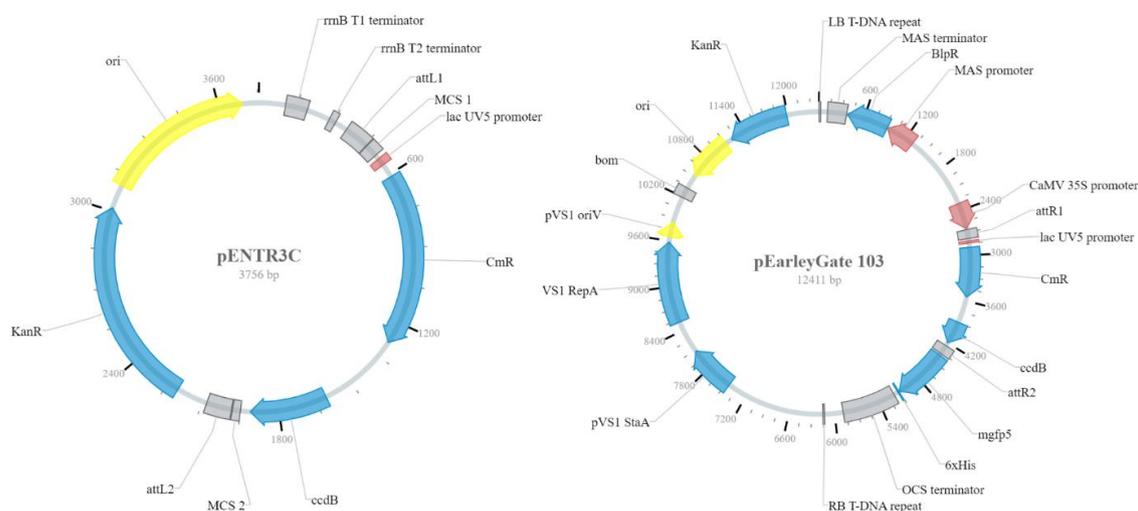


Figura 8: Esquema representando a estrutura dos vetores pENTR3C e pEarleyGate 103. Na figura é possível visualizar seus elementos constitutivos como suas marcas de resistências e regiões necessárias para a reação de clonagem (sítios attL e attR) de ambos os vetores. Imagem obtida a partir do site *Novoprolabs* (www.novoprolabs.com).

2.3.5. Transformação em *Arabidopsis thaliana*

Plantas de *A. thaliana* linhagens Col-0 foram transformadas pelo método de Floral Dip (Zhang *et al.*, 2006), com *A. tumefaciens EHA105*, contendo a construção pKIR1.1-Cas9sgRNA1.

2.3.6. Seleção e Genotipagem do material transformado

As sementes transformadas foram selecionadas pela monitoração pela fluorescência da Oleosina-RFP em microscópio invertido de fluorescência Olympus CKX41 com o conjunto de filtros G (Ex/Em, 510/590).

2.3.7. Análises de expressão

Os perfis de expressão dos genes de interesse foram avaliados nos bancos de dados online: *eFP Browser* (www.bar.utoronto.ca) e *Expression Atlas* (www.ebi.ac.uk).

Tabela 1: Sequência RNAs guias

Nome RNA guia	Sequência do RNA guia
sgRNA1	GGGGCCACTCTTACCAAGAA
sgRNA2	TCTGATTCTCTCCGCTCA

Tabela 2: Sequência dos Primers

Nome dos Primers utilizados	Sequência dos Primers
NPT2_Foward	CCAAGATTCTGCAGTTAGATCGGGTTTCTGGAGT TTAATGA
NPT2_Reverse	TCCTACGTACTIONCGAGGAATTCTCGATCTAGTAAC ATAGATGAC
RP SALK_093990.54.25.x	TCTTTGGAAAAATCAGCCTCC
LP SALK_093990.54.25.x	CTATACAATGGATCATGGCGG
LBb1.3	ATTTTGCCGATTTCGGAAC
PFGsgRNAf	ACGCGATAGGAAAACAAAATATAGC
sgRNA1_Foward	<u>ATTGGGGCCACTCTTACCAAGAA</u>
sgRNA1_Reverse	<u>AAACTTCTTGTAAGAGTGGCCC</u>
T7G RNA R	AAAAAAAGCACCGACTCGGTGCCA
AT4G15350 RTqPCR Foward	GCTTATCGAGACGAAAACGCTG
AT4G15350 RTqPCR Reverse	ACCAGCAATGAAAAGATCCACAAA
CYP705A_CDS_No_Stop_F	GTCGACTGGATCCGGTACCGCAATGGCAGTATTG ATCATC
CYP705A_CDS_No_Stop_R	ATATCTCGAGTGCGGCCGCGGGGAGGTA GGTTTAG

2.4. RESULTADOS

2.4.1. Experimento preliminar mutante SALK093990

O mutante SALK093990 possui uma inserção de T-DNA na porção 3'-UTR do locus AT4G15350. Inicialmente, um experimento comparativo preliminar foi elaborado, comparando o crescimento do mutante SALK093990 com o controle selvagem (Col-0). Em condições em que plantas foram crescidas em meio MS *in vitro*, o mutante apresentou plantas visivelmente maiores (Figura 9), juntamente com um aumento significativo na densidade de raízes laterais (Figura 10). Porém, em experimentos subsequentes, tal fenótipo não somente deixou de ser verificado, como também demonstrou uma severa diminuição na taxa de germinabilidade das sementes. Neste ponto, notamos que o lote de sementes utilizado poderia estar segregando para a mutação, uma vez que o banco de sementes do ABRC disponibiliza somente sementes heterozigotas para este mutante. Sendo assim, é possível que este lote de sementes não represente sementes homogeneamente homozigotas para a mutação.

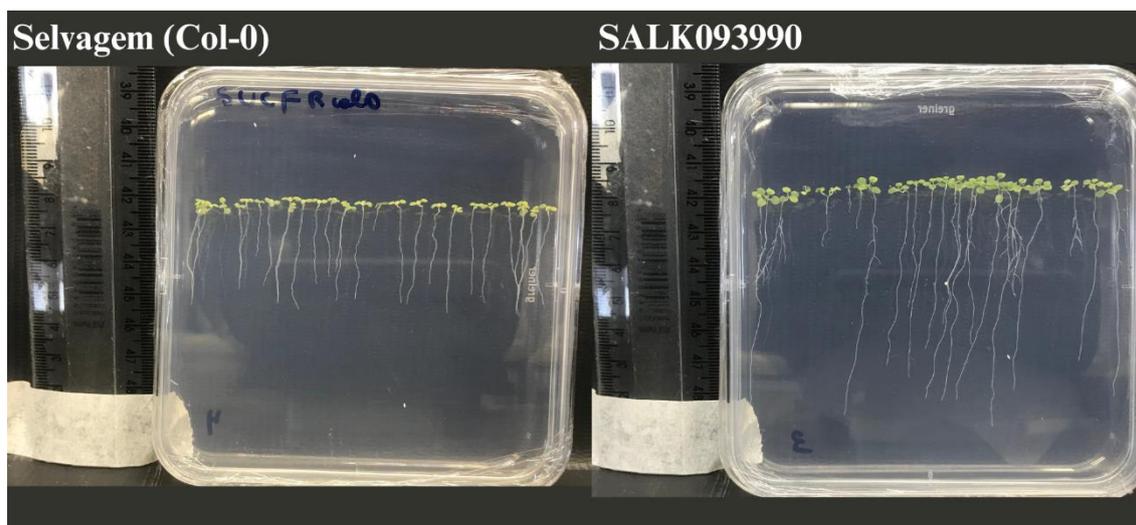


Figura 9. Comparação plantas de *A. thaliana* do tipo selvagem e o mutante. Placas de *A. thaliana* linhagens SALK093990 e Selvagem (Col-0) com 9 dias, crescidas em meio 0.5x MS, sem sacarose, a $21\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}\text{C}$ com fotoperíodo de dias longos, sob iluminação branca de $90\text{ }\mu\text{mol m}^{-2}\text{ s}^{-1}$, e baixas razões R:FR

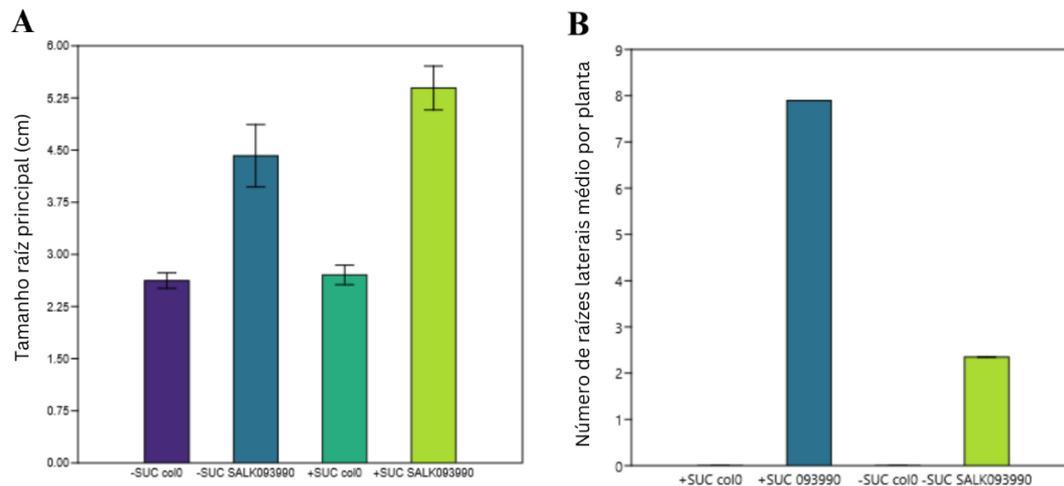


Figura 10: Comprimento da raiz principal e número médio de raízes laterais por planta. A-) gráfico que compara o crescimento da raiz principal da planta mutante SALK093990 (N=29) com a selvagem (Col-0, N=23), com baixa razão R:FR. B-) Gráfico que mostra a diferença entre o número de raízes laterais entre a planta mutante SALK093990 e a selvagem em baixa razão R:FR.

2.4.2. Genotipagem SALK093990

O processo de genotipagem do mutante SALK093990 foi necessário, pois o lote original obtido era heterozigoto para a mutação. Além disso, as plantas mutantes perderam sua viabilidade, gerando dificuldade não somente para adequada genotipagem, mas também para a verificação de seu fenótipo. Assim, para confirmar a presença da mutação, todas as plantas mutantes encontradas nos estoques do grupo de pesquisa foram genotipadas. A partir da extração de DNA, realizou-se reações de PCR com *primers* que amplificariam na marca de resistência a Canamicina (NPT2_Foward e NPT2_Reverse), presente em todos os T-DNAs das linhagens SALK. Como nenhuma linhagem mutante foi encontrada, os estudos fenotípicos foram suspensos até a obtenção de plantas mutantes homozigotas novamente.

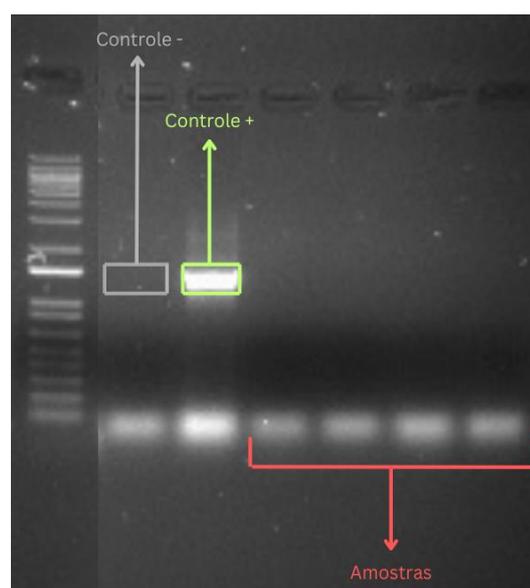


Figura 11. Genotipagem de plantas mutantes quanto à presença do T-DNA. Imagem ilustrativa de uma reação de PCR usando os primers NPT2_Foward e NPT2_Reverse para a detecção do

T-DNA das linhagens SALK. São indicados os controles negativo da reação (sem DNA molde, Controle-), o controle positivo (DNA molde contendo o gene NPTII, Controle+) e reações de 4 indivíduos SALK093990 (Amostras). A ausência de amplificação indica a inexistência do T-DNA. A primeira canaleta indica o marcador 1kb *Plus DNA ladder* (Invitrogen).

2.4.3. Obtenção das sequências RNAs guias

Para o desenho de sgRNAs específicos para o *knockout* do gene AT4G15350, a sequência CDS do gene foi obtida a partir do banco de sequências do *Phytozome* (<https://phytozome-next.jgi.doe.gov/>). Esta sequência foi utilizada para o desenho dos RNAs guias pela plataforma CRISPOR. Foram selecionados dois guias com potencial de gerar mutações em diferentes éxons do gene, sendo eles sgRNA1, presentes na Tabela 1 – GGGGCCACTCTTACCAAGAA (Exon II) e sgRNA2 – TCTGATTCTCTCCGCCTCA (Exon I) (Figura 10).

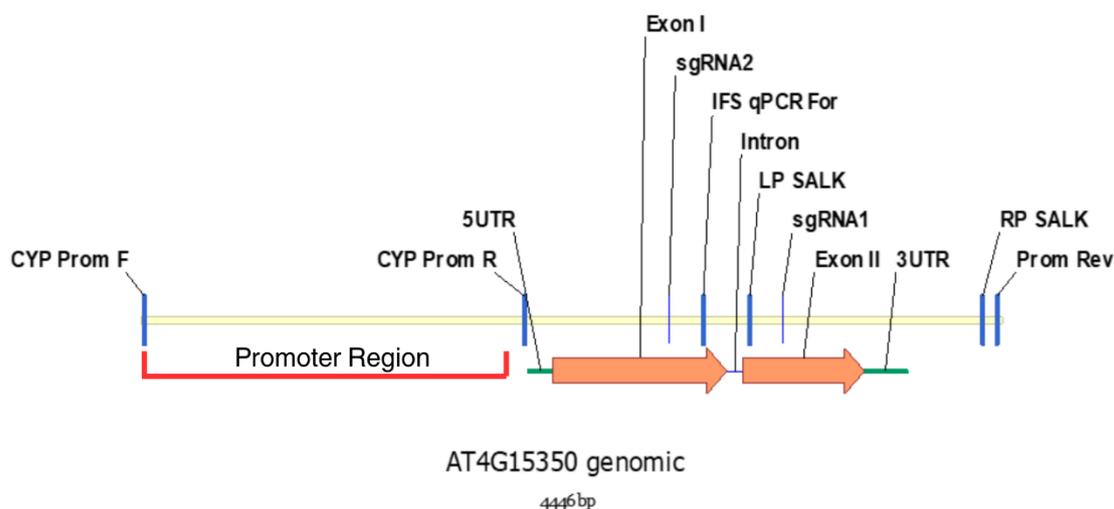


Figura 12. Locus do gene AT4G15350. Esquema mostrando o locus AT4G15350, tendo a região promotora e seus dois éxons em evidência, juntamente com a localização de *primers* e os sgRNA1 e sgRNA2 desenhados. A região RP SALK indica o ponto de inserção do T-DNA no mutante SALK093990.

2.4.4. Construção do vetor pKIR1.1-Cas9 sgRNA1

O sgRNA1 foi clonado no plasmídeo binário pKIR1.1 (Tsutsui e tal., 2017). Obtivemos quatro clones PCR positivos, utilizando os *primers*: PFGsgRNAf e sgRNA1. As colônias 1, 3, 4 e 5, que apresentaram um padrão de bandas aproximadamente 300pb, podendo ser vistas na Figura 13. A identidade da sequência foi confirmada por sequenciamento. Não conseguimos obter clones para o sgRNA2.

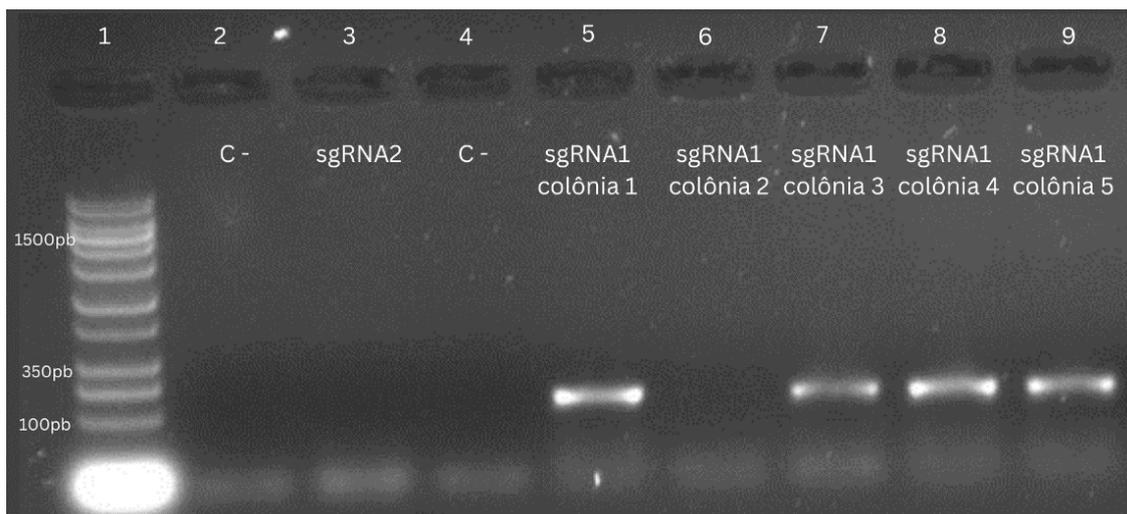


Figura 13. Eletroforese de DNA confirmando a clonagem do sgRNA1 no plasmídeo pKIR1.1-Cas9sgRNA1. Imagem de gel de eletroforese constando o resultado da PCR a partir de colônias de *E. coli* utilizando os *primers* PFGsgRNAf e sgRNA1 de quatro amostras distintas (amostras 5-9). A banda com aproximadamente 300pb (amostras 5, 7, 8 e 9) indicam a presença do inserto. Controles negativos da reação de PCR (*sem DNA molde*, C-) são indicados. A canaleta 1 corresponde ao marcador 1kb *Plus DNA ladder* (*Invitrogen*).

2.4.5. Construção do vetor pENTR3C-CDS4

Um fragmento da região genômica do CYP705A2 foi gerado a partir de uma PCR utilizando o DNA genômico de *A. thaliana*, da linhagem Col-0, com os seguintes *primers*: CYP705A2_CDS_No_Stop_F e CYP705A2_CDS_No_Stop_R, listados na Tabela 2. O fragmento resultante da reação, abrange a sequência dos Exon_I-intron-Exon_II, eliminando o stop códon no final do Exon_II, o que permite a formação de proteínas de fusão na porção C-terminal da proteína alvo.

Em seguida, os produtos foram submetidos a uma reação de *Gibson Assembly*, com o plasmídeo pENTR3C. Como resultado, foram analisadas 6 colônias recombinantes de *E. coli* através da purificação de seus plasmídeos e posterior clivagem com as enzimas *BamHI* e *HindIII*, esperando o padrão 2.87kb e 0.95kb, observado na Figura 14.

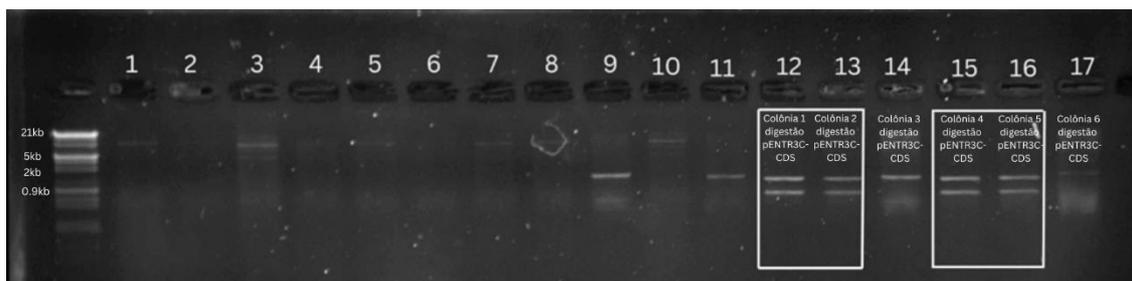


Figura 14. Eletroforese de DNA confirmando a transformação do vetor pENTR3C-CDS4 por digestão. Padrão de clivagem com as enzimas *BamHI* e *HindIII* das seis colônias resultantes da transformação com pENTR3C-CDS4, sendo possível a visualização do padrão 2.87kb e 0.95kb (amostras 12, 13, 15 e 16). A primeira canaleta corresponde ao marcador 1kb *Plus DNA ladder* (*Invitrogen*).

Os plasmídeos presentes nas quatro colônias, que apresentaram o padrão esperado, foram confirmados novamente por uma PCR usando os *primers*: AT4G15350

RTqPCR Foward e AT4G15350 RTqPCR Reverse, sendo visível o padrão de 0.15kb no gel da Figura 15, mostrada abaixo. Em seguida, tais vetores foram sequenciados pela empresa ACTGene, e apenas uma das construções não apresentou mutações em sua sequência e foi denominada de pENTR3C-CDS4.

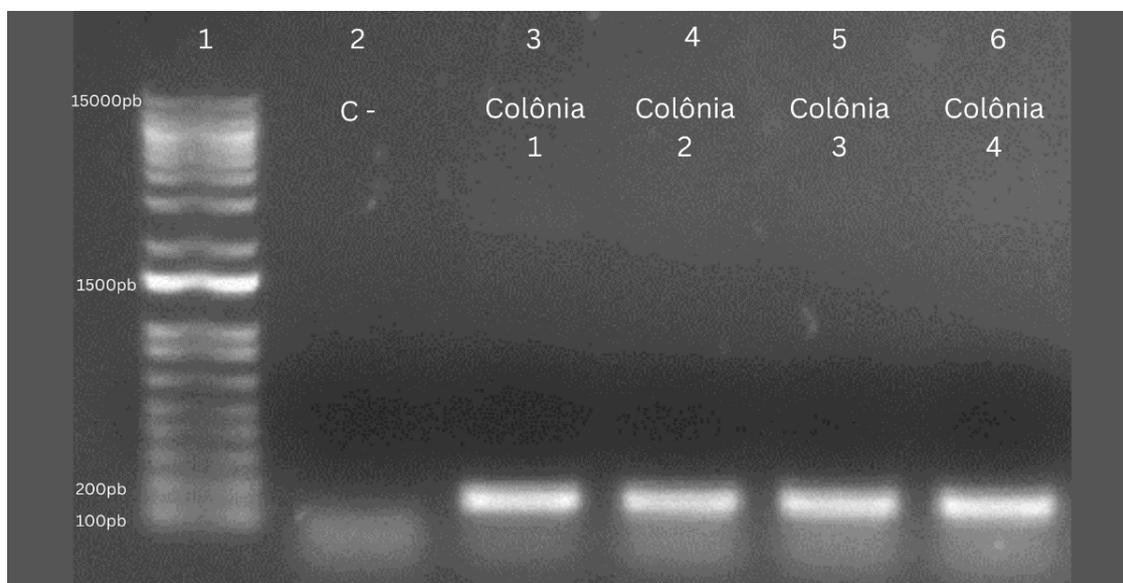


Figura 15. Eletroforese de DNA confirmando a presença do CDS vetor pENTR3C-CDS por PCR. O *amplicon* de aproximadamente 160pb (Amostras 3, 4, 5 e 6) resultante da PCR com os *primers*: AT4G15350_RTqPCR_Foward e AT4G15350_RTqPCR_Reverse confirmam a presença do inserto. Controles negativos da reação de PCR (sem DNA molde, C-) são indicados. A canaleta 1 corresponde ao marcador 1kb *Plus DNA ladder* (Invitrogen).

2.4.6. Construção do vetor pEG103-CDS4

A fim de gerar um cassete para a superexpressão do gene AT4G15350, o clone de entrada pENTR3C-CDS4 foi utilizado em uma reação de recombinação *Gateway* por LR Clonase com o plasmídeo binário pEarleyGate103 (Earley et al., 2006). A construção final permite a superexpressão de uma proteína de fusão CYP705A2-GFP-His controlada pelo promotor CaMV35S.

Inicialmente para gerar tal construção, o pENTR3C-CDS4 foi clivado previamente com a enzima *ApaI*. Dessa forma, a região genômica do CYP705A2 presente no pENTR3C juntamente com as regiões necessárias para a reação de clonase foi excisada, sendo possível a purificação da região de tamanho adequado, +-2kb, na Figura 16.

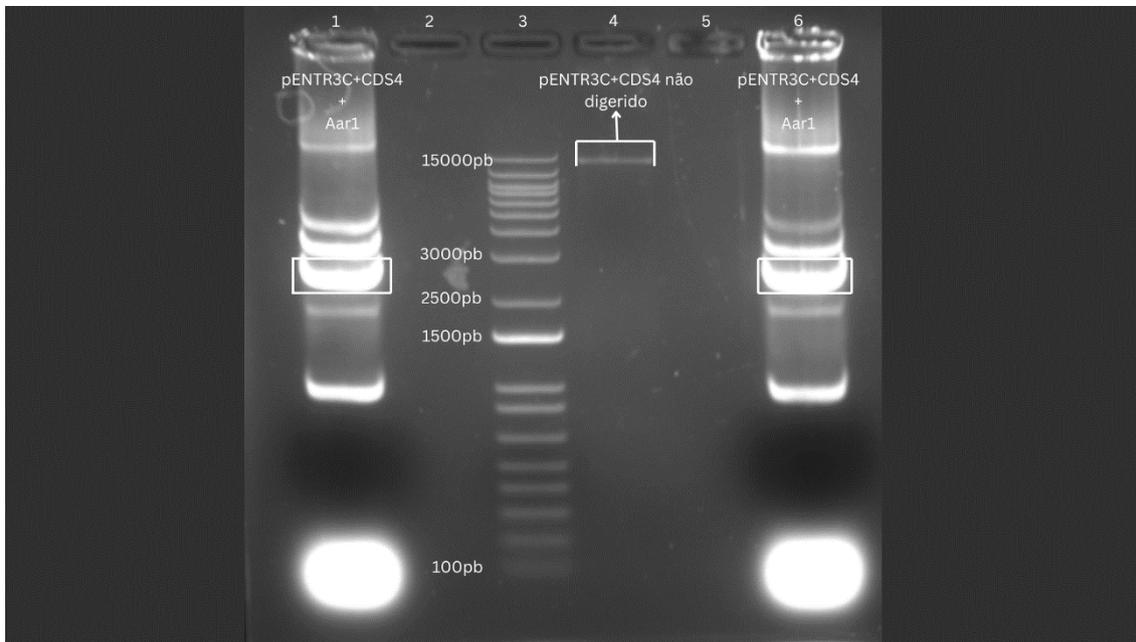


Figura 16. Eletroforese de DNA com digestão do vetor pENTR3C-CDS4 usando enzima *ApaI*. Resultado da digestão do plasmídeo pENTR3C-CDS4 com a enzima *ApaI* (Amostras 1 e 6), as caixas brancas indicam as bandas do vetor que foram purificadas, tais apresentando a região genômica amplificada e os sítios attL1 e attL2. A canaleta 3 corresponde ao marcador 1kb *Plus DNA ladder* (Invitrogen).

Após a purificação dos fragmentos, a reação da LR clonase foi realizada com o vetor de destino pEG103. Os plasmídeos resultantes da reação de LR clonase foram purificados e posteriormente digeridos com a enzima *HindIII*. Um dos clones apresentou o padrão de digestão esperado (10kb e 2kb), confirmando a presença da região CDS no pEG103, como apresentado na coluna 5 da Figura 17.

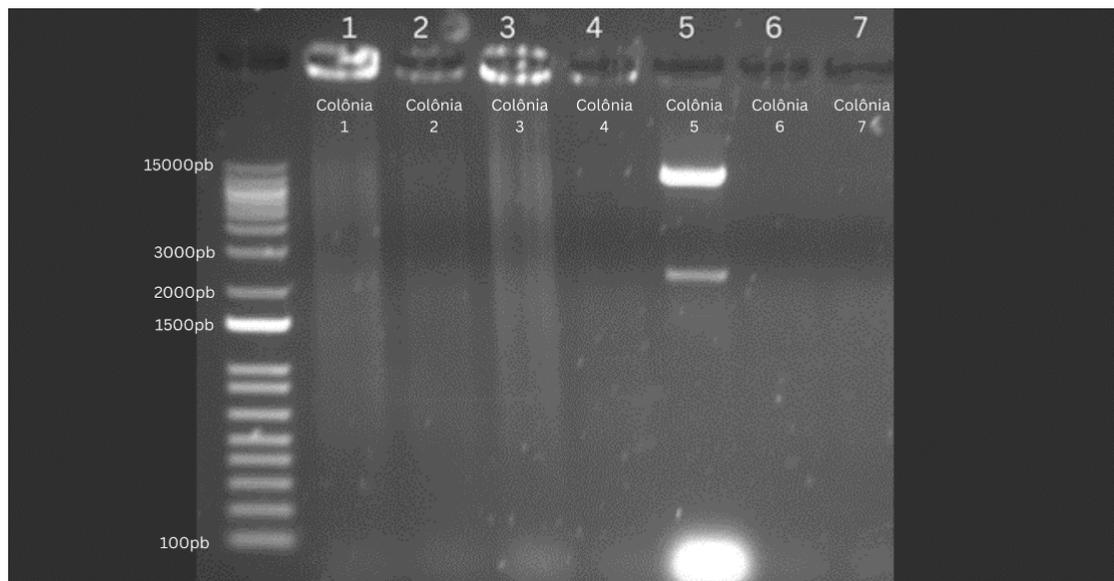


Figura 17. Eletroforese de DNA da confirmação a identidade do vetor pEG103-CDS4. O padrão de 10kb e 2kb da amostra número 5 confirma a identidade da construção. A canaleta à esquerda corresponde ao marcador 1kb *Plus DNA ladder* (Invitrogen).

2.4.7. Transformação de *Agrobacterium tumefaciens* com as construções: pKIR1.1-Cas9sgRNA1 e pEG103-CDS4

Foi realizada a transformação em *A. tumefaciens* EHA105 de ambas as construções, pKIR1.1-Cas9sgRNA1 e pEG103-CDS4, e a presença das mesmas foi confirmada por reações de PCR, em que seus respectivos resultados podem ser vistos na Figura 18 que se encontra abaixo.

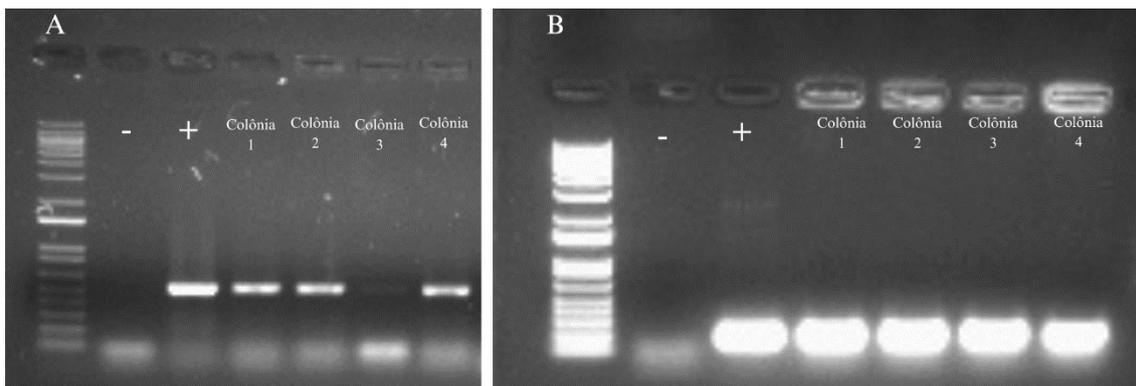


Figura 18. Eletroforese de DNA das confirmações da presença das construções por PCR em *A. tumefaciens*. A) Imagem do gel de eletroforese, utilizando os *primers* PFGsgRNAf e sgRNA1_Reverse, nele podemos ver o *amplicon* de tamanho aproximado de 350pb, esperado para a amplificação. B) Imagem do gel de eletroforese, utilizando os *primers*: AT4G15350 RTqPCR Forward e AT4G15350 RTqPCR Reverse, presentes na Tabela 2, em que é possível visualizar o padrão de 150bp oriundos da amplificação de parte da região codificante do gene AT4G15350 em todas as quatro colônias testadas. (-) e (+) indicam os respectivos controles negativos (sem DNA molde) e positivos das reações. As canaletas à esquerda correspondem ao marcador 1kb Plus DNA ladder (*Invitrogen*).

2.4.8. Transformação *Arabidopsis thaliana* e seleção do material transformado

Para a geração de mutantes editados por CRISPR/Cas9, o plasmídeo pKIR1.1-sgRNA1 foi transformado em *Agrobacterium tumefaciens* EHA105, sendo utilizado para transformação pelo método de Floral Dip (Zhang et al., 2006) em *A. thaliana*. As sementes T1 obtidas foram selecionadas pela expressão da proteína repórter Oleosina-RFP expressa pelas sementes transformadas, como podemos ver imagem de microscopia da Figura 19. Dessas sementes, 19 eventos independentes foram selecionados através dessa metodologia.

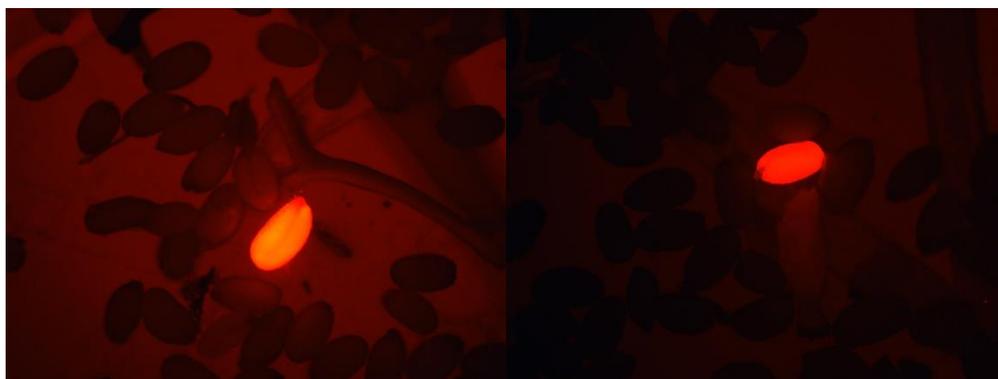


Figura 19. Seleção dos eventos transformados com a construção pKIR1.1-sgRNA1. Identificação dos eventos transformados em sementes fluorescentes T1 oriundas da transformação de *Arabidopsis thaliana* com *A. tumefaciens* EHA105 carregando o plasmídeo pKIR1.1-sgRNA1. Imagens obtidas em microscópio ótico de fluorescência para a detecção do gene repórter OLEOSIN1-RFPtag.

2.4.9. Análise do padrão de expressão do gene AT4G15350

Com o intuito de melhor entender o papel do gene CYP705A2, realizamos uma série de análises a partir de alguns bancos de dados on-line do padrão transcricional do gene em *A. thaliana*. Em dados obtidos na ferramenta *Arabidopsis eFP Browser* (<http://bar.utoronto.ca/>), o tecido que apresentou a maior expressão do gene AT4G15350 foi o meristema apical radicular, mais especificamente a coifa (Figura 20). A expressão do gene tende a estar condicionada a certos hormônios, sendo positivamente regulada na presença de Auxina, Ácido Abscísico, giberelinas, Ácido Jasmônico e brassinosteróides. Em relação a estresses, o gene tende a ser mais expresso quando as plantas são expostas a baixas temperaturas, assim como na presença de elicitores derivados de oomicetos.

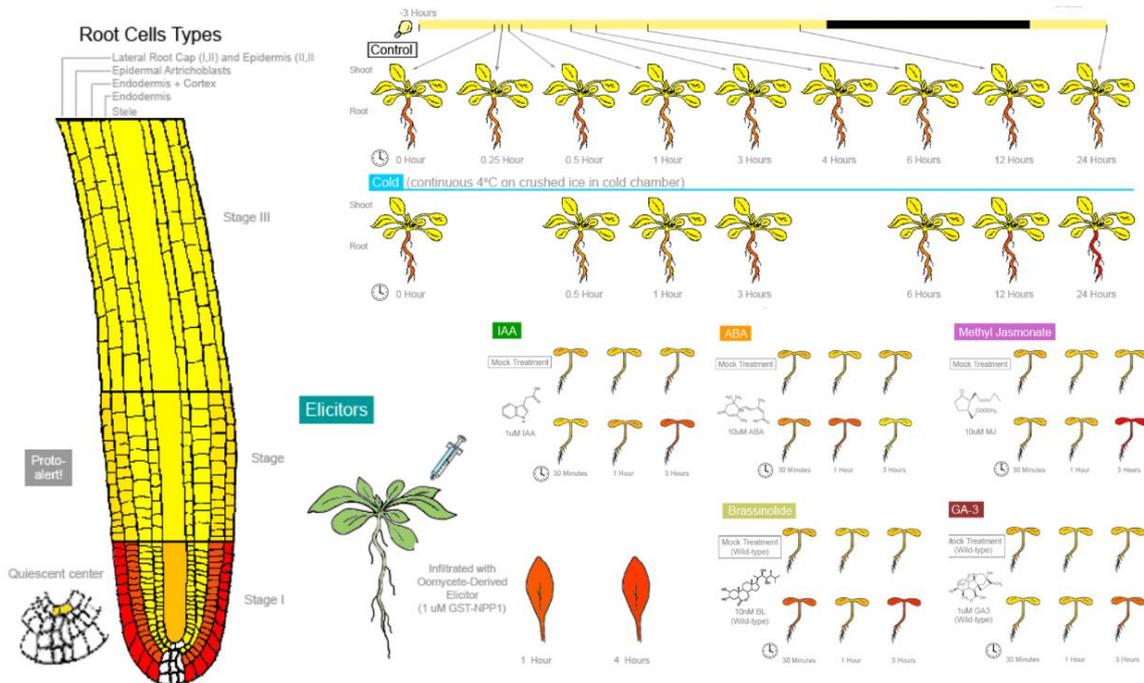


Figura 20. Padrão de expressão de AT4G15350 segundo Arabidopsis eFP Browser. Dados obtidos do grupo de dados *Root* (raiz, com destaque na expressão na coifa), *Abiotic Stresses* (estresses abióticos, com destaque na expressão em raízes expostas ao frio), *Hormone* (hormônios, com destaque à expressão em plântulas induzida por IAA, ABA, MeJA, BR e GA-3) e *Biotic Stress* (estresses bióticos, com destaque à expressão em folhas induzida por elicitores de oomicetos). Na representação é possível observar o padrão de expressão em tons de vermelho.

Segundo este mesmo banco de dados, existe uma variação de sua expressão em relação a alguns mutantes *knockouts* para o gene RHD6 e para ambos CPC (CAPRICE) e TRY (TRIPTYCHON), havendo uma forte indução no caso do primeiro mutante e uma diminuição para o segundo. No caso do mutante *rhdb6*, o mesmo apresenta um menor número de pelos radiculares, também apresentando alterações em relação a emergência das raízes laterais (Masucci; Schiefelbein, 1994). Já no caso do segundo, ambos os genes CPC e TRY atuam no controle negativo do desenvolvimento de tricomas (Kirik *et al.*, 2004).

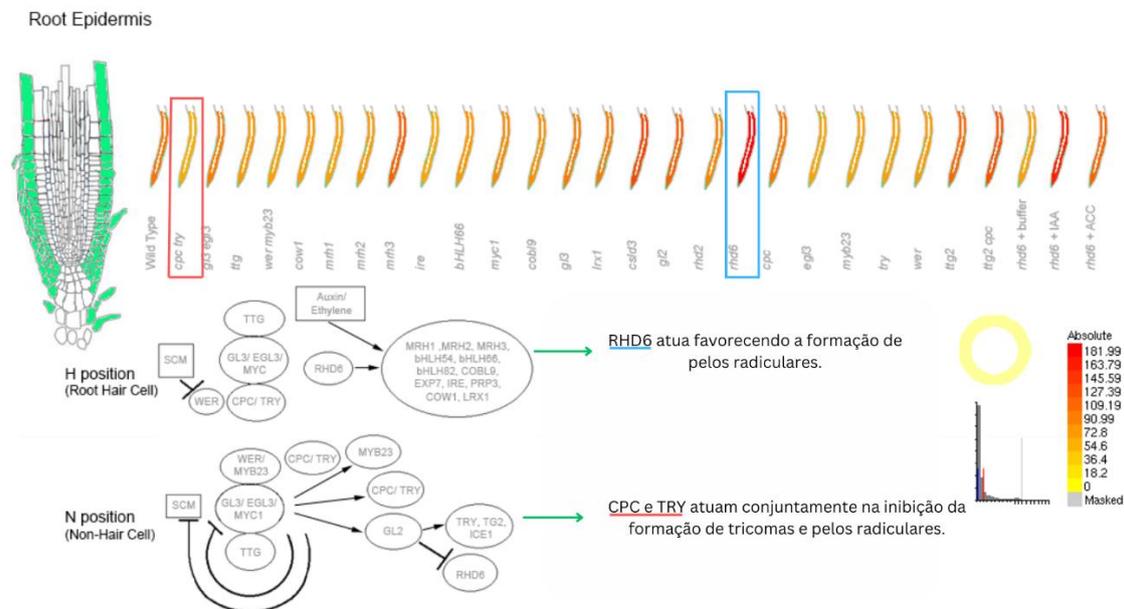


Figura 21. Padrão de expressão de AT4G15350 segundo *Arabidopsis eFP Browser* em backgrounds mutantes para genes específicos da epiderme da raiz. Representação visual do padrão de expressão do gene AT4G15350 em distintos mutantes *knockout*, havendo uma ênfase no mutante mais expresso (*rhd6*) e menos expresso (*cpc try*).

De acordo com o banco de dados *Expression Atlas* (<https://www.ebi.ac.uk/gxa/home>), a expressão do gene ATG415350 é considerada baixa, independente do tecido. Diferentemente do observado no banco de dados *Arabidopsis eFP Browser*, há uma expressão detectável nos tecidos florais e das sementes. Um resultado similar também foi alcançado pelo estudo de Huang *et al.*, 2019, em que o perfil de expressão dos clusters de síntese triterpênicos foram comparados. O perfil de expressão dos genes dos quatro MGCs tende a se concentrar nos tecidos radiculares, incluindo o PEN1 (aqui denominado de ABDS) e o CYP705A1. No entanto, essa tendência não observada para o gene AT4G15350, que apresenta maior expressão nos tecidos florais.

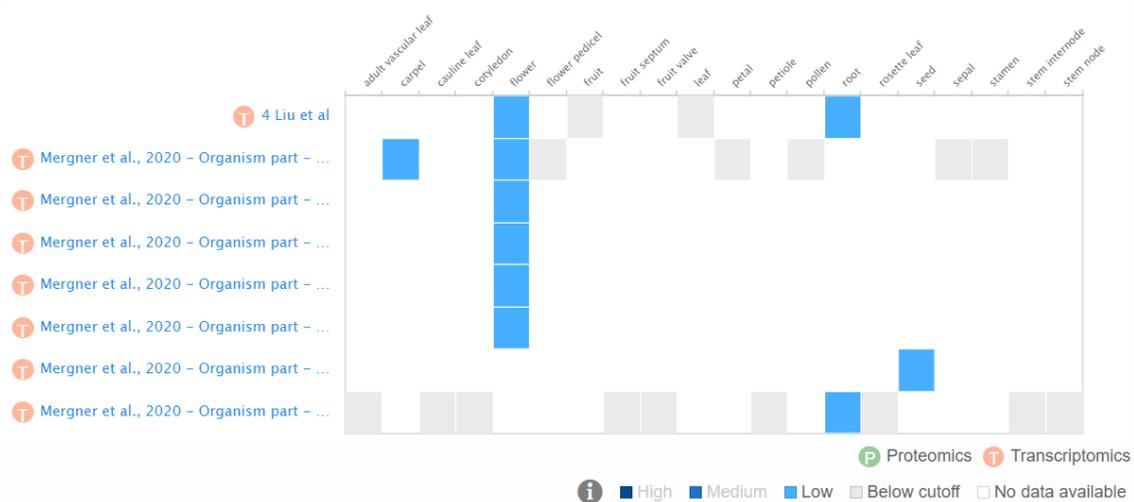


Figura 22. Padrão de expressão de AT4G15350 de acordo com o *Expression Atlas*. A figura mostra o padrão de expressão tecidual de acordo com o banco de dados *Expression Atlas*. Nele visualizamos que o estudo elaborado por Mergner *et al.* visualizou a expressão no tecido floral em 5 estágios florais distintos (Mergner *et al.*, 2020).

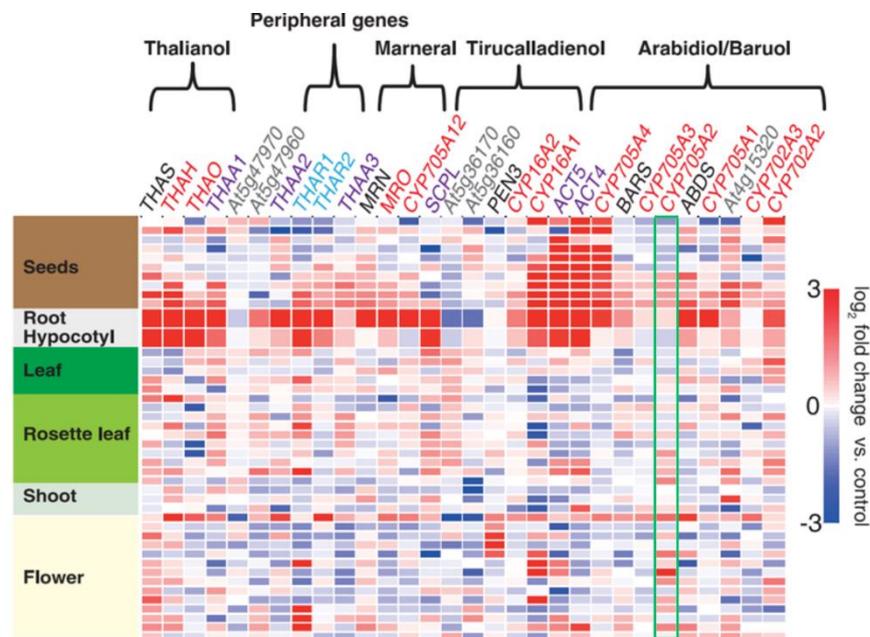


Figura 23. Padrão de expressão tecidual do AT4G15350 de acordo com o trabalho de Huang *et al.*, 2019. Figura adaptada do trabalho de Huang *et al.*, 2019, que compara o perfil de expressão tecidual dos genes envolvidos na síntese de compostos triterpênicos, evidenciando a expressão do CYP705A2 nos tecidos radiculares e florais.

2.4.10. Análise da estrutura proteica do gene AT4G15350

De acordo com os dados gerados pelo *InterPro* (<https://www.ebi.ac.uk/interpro/>), a proteína CYP705A2 apresenta domínios muito similares à família das CYP93, sendo classificada como pertencente ao clã 71. Tal família, exclusiva de angiospermas, possui integrantes envolvidos nos passos iniciais de síntese de isoflavonas, flavonas e compostos triterpênicos, como as saponinas, em outros organismos vegetais. Segundo Du *et al.* 2016, essa família apresenta uma alta similaridade de sequências com as famílias CYP705 e CYP712, indicando uma derivação a partir de um ancestral comum. Por meio dessa semelhança, as primeiras predições relacionadas à estrutura proteica da CYP705A2 podem ter caracterizado sua função como pertencente à família CYP93 e, conseqüentemente, a síntese de flavonóides e isoflavonoídes devido à similaridade entre seus domínios (Du *et al.*, 2016).

O gene AT4G15330, codifica uma enzima da classe das citocromo monooxigenases, classificada como CYP705A1, também pertencente à família 705. Em 2013, Castillo *et al.* constatou que a proteína realiza uma reação de clivagem da cadeia lateral do triterpeno arabidiol, formando o 14-Apo-arabidiol e o DMNT ((3E)-4,8-dimethyl-1,3,7-nonatriene). Tal reação foi visualizada como resposta defensiva ao patógeno *Pythium irregulare*, assim como na presença de Ácido Jasmônico (Castillo; Kolesnikova; Matsuda, 2013; Sohrabi *et al.*, 2015).

Pertencentes a rota de síntese do arabidiol e baruol, os genes PEN1 e BARS1 possuem 91% de similaridade a nível de aminoácidos, codificando duas oxidosqualeno ciclases (OSCs) das 13 encontradas em *A. thaliana* (Lodeiro *et al.*, 2007; Field *et al.*, 2011). Dentre as enzimas pertencente ao grupo das OSCs, alguns integrantes fazem parte de quatro clusters de genes metabólicos (MGCs), consistindo em enzimas da rota de síntese de triterpenos, no caso: thalianol, marneral, tirucalladienol, arabidiol e baruol.

O produto do gene BARS1 sintetiza aproximadamente 90% de baruol, composto ainda não funcionalmente caracterizado. Essa enzima também é capaz de sintetizar outros 22 compostos minoritários a partir do oxidosqualeno (Lodeiro *et al.*, 2007). Apesar da similaridade estrutural e classe bioquímica comum com BARS1, o gene PEN1 sintetiza unicamente o triterpeno arabidiol, também através do oxidosqualeno (Xiang *et al.*, 2006; Ortiz *et al.*, 2019). O arabidiol não possui função biológica caracterizada, porém a clivagem com a cadeia lateral do arabidiol provocada pela enzima CYP705A1, codificada pelo gene AT4G15330, no composto volátil DMNT ((3E)-4,8-dimethyl-1,3,7-nonatriene) e no 14-apo-arabidiol, atuam na defesa contra estresses bióticos (Sohrabi *et al.*, 2015; Castillo; Kolesnikova; Matsuda, 2013).

A degradação do arabidiol é dependente do gene AT4G15330, também pertencente ao cluster de genes co-expressos (Castillo; Kolesnikova; Matsuda, 2013). A expressão desse gene ocorre no meristema radicular, induzida também pela presença do Ácido Jasmônico sendo detectada em resposta defensiva ao patógeno *Pythium irregulare*, (Castillo; Kolesnikova; Matsuda, 2013; Sohrabi *et al.*, 2015).

Um dos efeitos fisiológicos do acúmulo do monoterpene DMNT, visto em *Camellia sinensis*, é a produção de Ácido Jasmônico através do acúmulo SPI (*Saporamin Protease Inhibitor*) em plantas vizinhas ainda não danificadas pelo estresse, como podemos observar na Figura 25 (Meents *et al.*, 2019; Jing *et al.*, 2021). Já o 14-apo-arabidiol, também não possui função específica conhecida, porém em 2017, Sohrabi identificou que ele é convertido, possivelmente por enzimas não pertencentes ao cluster de síntese do arabidiol/baruol, em outros três compostos com função desconhecida: 3-keto-14-apo-arabidiol, alfa-14-acetyl-apo-arabidiol e alfa-14-apo-arabidiol (Sohrabi *et al.*, 2017).

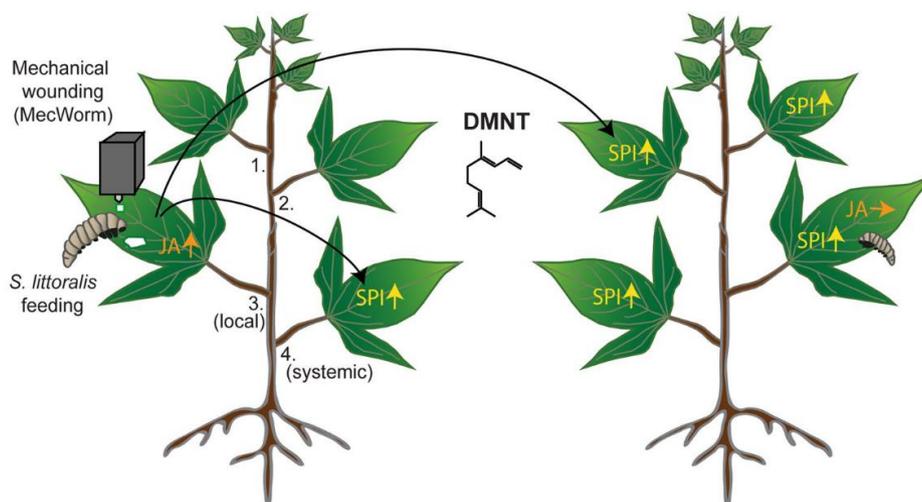


Figura 25. Representação da atuação do DMNT. Esquema representando sistema de defesa a estresse biótico promovido pelo DMNT, adaptada de MEENTS *et al.*, 2019 (Meents *et al.*, 2019). Nessa imagem, a herbívora causada pelas larvas de *S. littoralis* induzem tanto a produção de Ácido Jasmônico, quanto de SPI localmente. Além disso, o DMNT tem sua síntese induzida, sendo emitido para plantas não predadas, de modo a induzir a produção de SPI e Ácido Jasmônico.

A partir do trabalho de Huang *et al.* em 2019, que identificou possíveis novos passos metabólicos na rota de síntese do arabidiol e thalianol, hipotetiza-se atualmente que o 14-apo-arabidiol é transformado em 3-keto-14-apo-arabidiol por ação da enzima THAR1 (At3g29250), que sequencialmente é modificado no alfa-14-acetyl-apo-arabidiol pela THAR2 (At1g66800), sendo finalmente convertido pela enzima THAA2 (At5g47950) em alfa-14-apo-arabidiol renomeado em arabidina (Huang *et al.*, 2019). Esses genes fazem parte de outro cluster co-expresso, envolvido na rota de síntese do thalianol, mas que não possui uma correlação do perfil de expressão de seus integrantes com o cluster de síntese do arabidiol/baruol. Esse composto é expresso nos tecidos radiculares, estando envolvido, juntamente com o segundo composto testado no trabalho de Huang denominado de thalianina, com o estabelecimento da microbiota da rizosfera em *A. thaliana* (Huang *et al.*, 2019).

Devido a conversão do arabidiol em DMNT, e seu respectivo envolvimento com a defesa de estresses bióticos, é possível supor que o gene AT4G15350, co-expresso dentro dessa mesma rota de síntese, possa também estar envolvido. Porém, como o passo enzimático promovido pelo AT4G15330 também produz 14-Apo-arabidiol e, por este ser sequencialmente transformado em outros três compostos também com função desconhecida, não é possível excluir que tais produtos estejam envolvidos em outros processos fisiológicos, aos quais AT4G15350 participe, talvez utilizando algum deles como substrato. Além disso, o triterpeno baruol também não tem sua função descrita, evidenciando a falta de conhecimento a respeito desses cluster de genes e seus produtos.

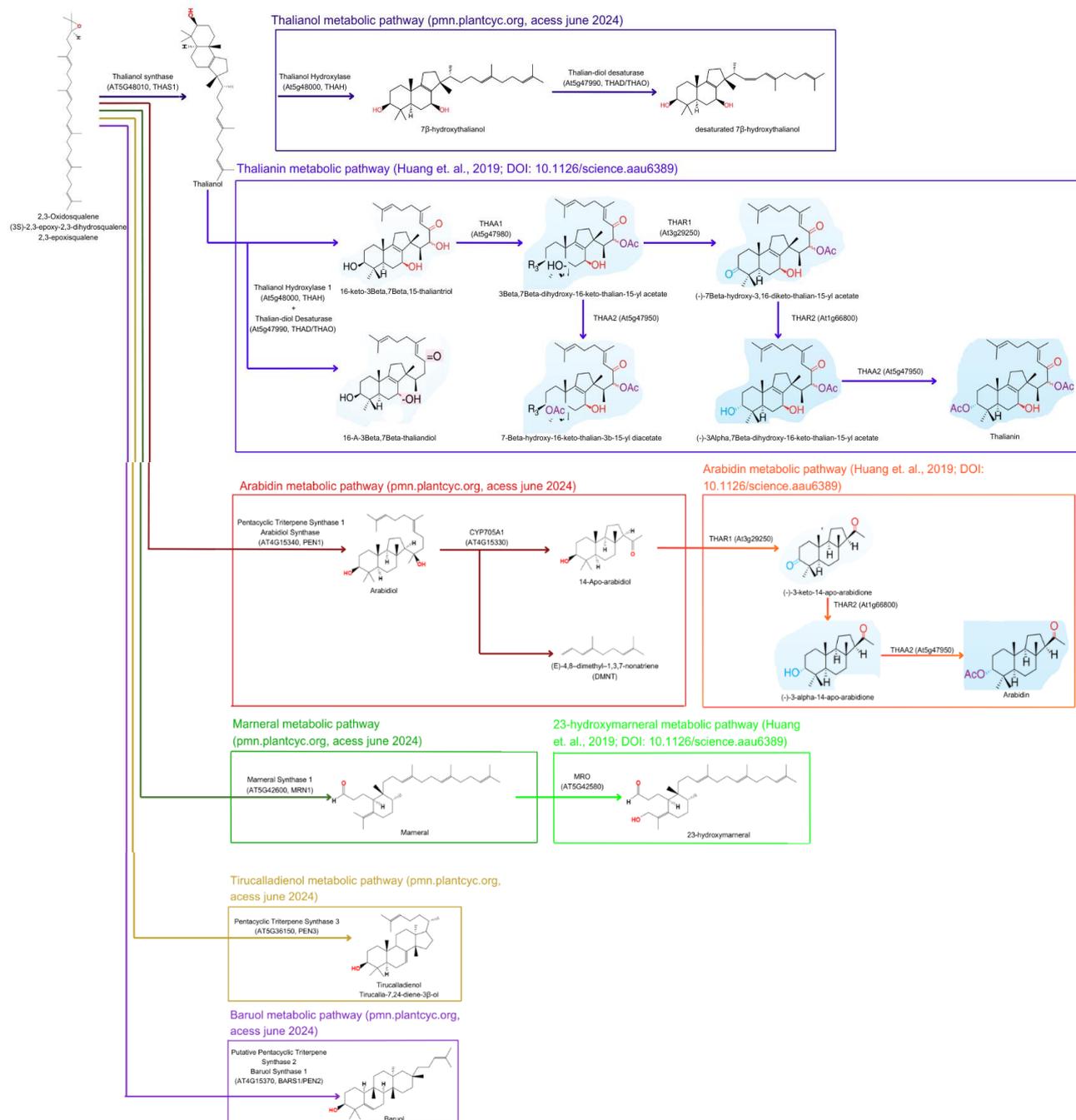


Figura 26. Representação das rotas metabólicas dos quatro MGCs presentes em *A. thaliana*. Figura explicitando as rotas metabólicas de alguns genes envolvidos nos quatro MGCs presentes em *A. thaliana*, nele há rotas presentes no site *Plantcyc* (<https://pmn.plantcyc.org>), assim como predições realizadas no trabalho de Huang et al., 2019 (Huang et al., 2019; Ghosh, 2016; Ghosh, 2017; Lodeiro et al., 2007; Bai et al., 2021; Chen et al., 2023; Sohrabi et al., 2015; Marszalek-Zenczak et al., 2023).

2.5. DISCUSSÃO E PERSPECTIVAS

Segundo o padrão de expressão obtidos através de bancos de dados online, o gene AT4G15350 tende a ser mais expresso no ápice radicular, em resposta a hormônios promotores do crescimento e diferenciação celular, indicando uma possível relação inversa ao fenômeno observado no mutante SALK093990, em que sua baixa expressão geraria um aumento em seu crescimento. No entanto, existe uma correlação entre sua expressão e a presença do Ácido Jasmônico, fitohormônio envolvido na defesa contra patógenos necróticos, que também possui um efeito inibidor do crescimento (Huot *et al.*, 2014).

Em congruência com o desconhecido substrato do gene AT4G15350, além da ausência de caracterização para a maioria dos produtos dos genes pertencentes ao cluster do arabidiol/baruol, com exceção para o DMNT e a arabidina, não é possível correlacionar a função do AT4G15350 com nenhum processo fisiológico, também evidenciando a falta de indícios que o relacione, diretamente, com o fenótipo observado no mutante SALK093990. Mesmo assim, a correlação do seu padrão de expressão tecidual, bem como da co-expressão identificada no trabalho de Wada *et al.*, em 2017, pode indicar uma relação com a modificação de algum dos compostos produzidos na rota, por exemplo o baruol ou os dois derivados do 14-apo-arabidiol (3-keto-14-apo-arabidiol e alfa-14-acetyl-apo-arabidiol) que seguem sem função conhecida (Wada *et al.*, 2012).

No entanto, a partir do trabalho realizado por Nelson *et al.* em 2017, que desenvolveu duas ferramentas capazes de identificar lincRNAs, verificou-se a presença de um lincRNA (*long intergenic non-coding RNA*) com 828pb, 173 bases *upstream* do final do gene AT4G15350 (Nelson *et al.*, 2017), o local desse lincRNA pode ser visto na figura 1. Essa classe de RNAs comumente possuem múltiplas possíveis ORFs (*Open Reading Frames*), apesar de serem traduzidas de forma ineficaz, de modo a produzir um produto proteico disfuncional, ou mesmo sendo incapaz serem traduzidos. Além disso, sua sequência tende a não apresentar semelhanças a domínios proteicos conhecidos, bem como ORFs de pelo menos 300pb (Ulitsky; Bartel, 2013).

Essa classe de RNAs pode ser dividida em dois grandes subgrupos, aqueles capazes de atuar no núcleo, tendo uma ação que abrange desde a modificações epigenéticas, estruturais da cromatina a regulação da transcrição, assim como aqueles que atuam no citoplasma, atuando na regulação da tradução e estabilidade de mRNA (RNA mensageiro) (Wang *et al.*, 2011). Tendo em vista o vasto impacto na regulação gênica, torna-se importante incluir a participação do lincRNA, presente logo após a localidade a qual o gene AT4G15350 se localiza, como possível influenciador do fenótipo observado. Para tanto, torna-se necessário compreender de que modo esse lincRNA está atuando, assim como seus alvos, sendo os estudos de co-expressão uma das metodologias utilizadas (Rinn; Chang, 2012).

3.CONCLUSÃO

Em congruência com o atual desconhecimento do substrato do gene CYP705A2, além da ausência de caracterização para a maioria dos produtos dos genes pertencentes ao cluster do arabidiol/baruol, com exceção para o DMNT e a arabidina, não é possível correlacionar a função do At4g15350 com nenhum processo fisiológico. Os dados dos outros genes do cluster, evidenciam a falta de indícios que o relacione, diretamente, com o fenótipo observado no mutante SALK093990. Mesmo assim, a correlação do seu padrão de expressão tecidual, bem como da co-expressão identificada no trabalho de Marszalek-Zenczak *et al.* pode indicar uma relação com a modificação de algum dos compostos produzidos na rota, por exemplo o baruol ou os dois derivados do 14-apo-arabidiol (3-keto-14-apo-arabidiol, alfa-14-acetyl-apo-arabidiol) que seguem sem função conhecida. Dessa forma, é fundamental considerar que o fenótipo observado no mutante esteja relacionado com alterações provocadas pelo T-DNA no lincRNA.

Dessa forma, os futuros objetivos do projeto buscam descobrir os efeitos do lincRNA presente depois do gene alvo, porém a abordagem a ser adotada ainda não está definida. Além disso, os mutantes de *Arabidopsis* obtidos através da transformação com o vetor pKIR1.1-sgRNA1 e pEG103-CDS4 serão inicialmente genotipados por PCR e, em seguida, terão seus fenótipos caracterizados e comparados com a linhagem selvagem.

REFERÊNCIAS

- Alonso, J. M. et al.** Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, v. 301, n. 5633, p. 653–657, (2003). doi: 10.1126/science.1086391.
- Bai, Y. et al.** (2021). Modulation of *Arabidopsis* root growth by specialized triterpenes. *New Phytologist*, v. 230, n. 1, p. 228–243 (2021), doi: 10.1111/nph.17144.
- Bathe, U.; Tissier, A.** (2019) Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry*, v. 161, n. February, p. 149–162. doi: 10.1016/j.phytochem.2018.12.003.
- Bennett, J. W.; Bentley, R.** (1989) What's in a Name? Microbial Secondary Metabolism. *Advances in Applied Microbiology*, v. 34, n. C, p. 1–28. doi: 10.1016/S0065-2164(08)70316-2.
- Boron, A. K.; Vissenberg, K.** (2014). The *Arabidopsis thaliana* hypocotyl, a model to identify and study control mechanisms of cellular expansion. *Plant Cell Reports*, v. 33, n. 5, p. 697–706, doi: 10.1007/s00299-014-1591-x.
- Castillo, D. A.; Kolesnikova, M. D.; Matsuda, S. P.T.** (2013). An effective strategy for exploring unknown metabolic pathways by genome mining. *Journal of the American Chemical Society*, v. 135, n. 15, p. 5885–5894, doi: 10.1021/ja401535g.
- Chen, H. et al.** (2023). Inactivation of RPX1 in *Arabidopsis* confers resistance to *Plutella xylostella* through the accumulation of the homoterpene DMNT. *Plant Cell and Environment*, v. 46, n. 3, p. 946–961. doi: 10.1111/pce.14528.
- Christino, J. A. R.; Maraschin, F. D. S.** (2023). Avaliação do papel da síntese de flavonóides no desenvolvimento fotomorfogênico de raízes em *Arabidopsis Thaliana*. Disponível em: <https://lume.ufrgs.br/handle/10183/276125>.
- Du, H. et al.** (2016). Genome-wide analysis, classification, evolution, and expression analysis of the cytochrome P450 93 family in land plants. *Plos One*, v. 11, n. 10, p. 1–24. doi: 10.1371/journal.pone.0165020.
- Earley, K. W. et al.** (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant Journal*, v. 45, n. 4, p. 616–629. doi: 10.1111/j.1365-313X.2005.02617.x.
- Field, B. et al.** (2011). Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 38, p. 16116–16121, doi: 10.1073/pnas.1109273108.
- Ghosh, S.** (2016). Biosynthesis of structurally diverse triterpenes in plants: The role of oxidosqualene cyclases. *Proceedings of the Indian National Science Academy*, v. 82, n. 4, p. 1189–1210. doi: 10.16943/ptinsa/2016/48578.
- Ghosh, S.** (2017) Triterpene structural diversification by plant cytochrome P450 enzymes. *Frontiers in Plant Science*, v. 8, n. November, p. 1–15. doi: 10.3389/fpls.2017.01886.
- Gibson, D. G. et al.** (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, v. 6, n. 5, p. 343–345. doi: 10.1038/nprot.2009.77.
- Guengerich, F. P.** (2018). Mechanisms of Cytochrome P450-Catalyzed Oxidations. *ACS Catalysis*, v. 8, n. 12, p. 10964–10976. doi: 10.1021/acscatal.8b03401.
- Hamberger, B.; Bak, S.** (2013). Plant P450s as versatile drivers for evolution of species-

specific chemical diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 368, n. 1612. doi: //dx.doi.org/10.1098/rstb.2012.0426.

Hansen, C. C. et al. (2021). Plant cytochrome P450 plasticity and evolution. *Molecular Plant*, v. 14, n. 8, p. 1244–1265. doi: 10.1016/j.molp.2021.06.028.

Huang, A. C. et al. (2019). A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science*, v. 364, n. 6440. doi: 10.1126/science.aau6389.

Huot, B. et al. (2014). Growth-defense tradeoffs in plants: A balancing act to optimize fitness. *Molecular Plant*, v. 7, n. 8, p. 1267–1287. doi:10.1093/mp/ssu049.

Gateway® pENTR™ Dual Selection Vectors (2008). Disponível em: https://tools.thermofisher.com/content/sfs/manuals/pentrdualselectionvectors_man.pdf.

Gateway® pENTR™ Vectors (2004). Disponível em: https://tools.thermofisher.com/content/sfs/manuals/pentr_man.pdf.

Jing, T. et al. (2021). Herbivore-induced DMNT catalyzed by CYP82D47 plays an important role in the induction of JA-dependent herbivore resistance of neighboring tea plants. *Plant Cell and Environment*, v. 44, n. 4, p. 1178–1191. doi: 10.1111/pce.13861.

Kirik, V. et al. (2004). The ENHANCER of TRY and CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in *Arabidopsis*. *Developmental Biology*, v. 268, n. 2, p. 506–513. doi: 10.1016/j.ydbio.2003.12.037.

Lapcik O. et al. (2006) Isoflavonoids are present in *Arabidopsis thaliana* despite the absence of any homologue to known isoflavonoid synthases. *Plant Physiology and Biochemistry*, v. 44, p106-114. doi: 10.1016/j.plaphy.2005.11.006.

Lodeiro, S. et al. (2007). An oxidosqualene cyclase makes numerous products by diverse mechanisms: A challenge to prevailing concepts of triterpene biosynthesis. *Journal of the American Chemical Society*, v. 129, n. 36, p. 11213–11222. doi: 10.1021/ja073133u.

Main, G. D.; Reynolds, S.; Gartland, J. S. (1995). Electroporation protocols for *Agrobacterium*. *Methods in molecular biology* (Clifton, N.J.), v. 44, p. 405–412. doi:10.1385/0-89603-302-3:405.

Marszalek-Zenczak, M. et al. (2023). Analysis of *Arabidopsis* non-reference accessions reveals high diversity of metabolic gene clusters and discovers new candidate cluster members. *Frontiers in Plant Science*, v. 14, n. January, p. 1–16. doi: 10.3389/fpls.2023.1104303.

Masucci, J. D.; Schiefelbein, J. W. (1994). The *rhd6* mutation of *Arabidopsis thaliana* alters root-hair initiation through an auxin- and ethylene-associated process. *Plant Physiology*, v. 106, n. 4, p. 1335–1346. doi: 10.1104/pp.106.4.1335.

Meents, A. K. et al. (2019). Volatile DMNT systemically induces jasmonate-independent direct anti-herbivore defense in leaves of sweet potato (*Ipomoea batatas*) plants. *Scientific Reports*, v. 9, n. 1, p. 1–13. doi: 10.1038/s41598-019-53946-0.

Meinke, D. W. et al. (1998). *Arabidopsis thaliana*: A model plant for genome analysis. *Science*, v. 282, n. 5389. doi: 10.1126/science.282.5389.662.

Mergner, J. et al. (2020). Mass-spectrometry-based draft of the *Arabidopsis* proteome. *Nature*, v. 579, n. 7799, p. 409–414. doi: 10.1038/s41586-020-2094-2.

Meyerowitz, E. M. (1989). *Arabidopsis*, a useful weed. *Cell*, v. 56, n. 2, p. 263–269, 1989. doi: 10.1016/0092-8674(89)90900-8.

- Miotto, Y. E. et al.** (2019). Identification of root transcriptional responses to shoot illumination in *Arabidopsis thaliana*. *Plant Molecular Biology*, v. 101, n. 4–5, p. 487–498. doi: 10.1007/s11103-019-00918-7.
- Mizutani, M.; Ohta, D.;** (2010). Diversification of P450 genes during land plant evolution. *Annual Review of Plant Biology*; v. 61, p. 291-315. doi: 10.1146/annurev-arplant-042809-112305.
- Murashige, T.; Skoog, F.** (1962). A revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, v. 15, p. 474–497. doi: 10.1111/j.1399-3054.1962.tb08052.x.
- Nelson, D. R.** (2018). Cytochrome P450 diversity in the tree of life. *Biochimica et Biophysica Acta - Proteins and Proteomics*, v. 1866, n. 1, p. 141–154. doi: 10.1016/j.bbapap.2017.05.003.
- Nelson, A. D. L. et al.** (2017). Evolinc: A tool for the identification and evolutionary comparison of long intergenic non-coding RNAs. *Frontiers in Genetics*, v. 8, n. MAY, p. 1–12. doi: 10.3389/fgene.2017.00052.
- Nelson, David; Werck-Reichhart, D.** (2011). A P450-centric view of plant evolution. *Plant Journal*, v. 66, n. 1, p. 194–211. doi: 10.1111/j.1365-313X.2011.04529.x
- O'Malley, R. C.; Barragan, C. C.; Ecker, J. R.** (2015). HHMI Author Manuscript A User's Guide to the *Arabidopsis* T-DNA Insertional Mutant Collections. *Methods in Molecular Biology*, p. 323–342. doi:10.1007/978-1-4939-2444-8_16.
- O'Malley, R. C.; Ecker, J. R.** (2010). Linking genotype to phenotype using the *Arabidopsis* unimutant collection. *Plant Journal*, v. 61, n. 6, p. 928–940. doi: 10.1111/j.1365-313X.2010.04119.x.
- Ortiz, P. R. D. M. et al.** (2019). Cytochrome p450 Structure, Mechanism, and Biochemistry. 4. ed. San Francisco. doi: 10.1007/978-3-319-12108-6.
- Rendic, S.; Guengerich, F. P.** (2015). Survey of human oxidoreductases and cytochrome P450 enzymes involved in the metabolism of xenobiotic and natural chemicals. *Chemical Research in Toxicology*, v. 28, n. 1, p. 38–42. doi: 10.1021/tx500444e.
- Rinn, J. L.; Chang, H. Y.** (2012). Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry*, v. 81, p. 145–166. doi: 10.1146/annurev-biochem-051410-092902.
- Shimada, T. L.; Shimada, T.; Hara-Nishimura, I.** (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*: Technical Advance. *Plant Journal*, v. 61, n. 3, p. 519–528. doi: 10.1111/j.1365-313X.2009.04060.x.
- Sohrabi, R. et al.** (2017). Formation and exudation of non-volatile products of the arabioidiol triterpenoid degradation pathway in *Arabidopsis* roots. *Plant Signaling and Behavior*, v. 12, n. 1, p. 1–6. doi:10.1080/15592324.2016.1265722.
- Sohrabi, R. et al.** (2015). In Planta Variation of Volatile Biosynthesis: An Alternative Biosynthetic Route to the Formation of the Pathogen-Induced Volatile Homoterpene DMNT via Triterpene Degradation in *Arabidopsis* Roots. *Plant Cell*, v. 27, n. 3, p. 874–890. doi: 10.1105/tpc.114.132209.
- Subcloning Efficiency™ DH5α™ Competent Cells Cat. 18265-017. *Molecular and cellular biology*, v. 2, n. 1, p. 1–2, 2006. Disponível em:

https://assets.thermofisher.com/TFS-Assets/LSG/manuals/subcloningefficiencydh5alpha_man.pdf.

TAGI - The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of *Arabidopsis thaliana*. *Nature*, v. 408, n. December, p. 796–815. Disponível em: <https://www.nature.com/articles/35048692>.

Taiz, L. et al. (2017). Fisiologia e desenvolvimento vegetal Diversidade vegetal. ed. 6. Editora Artmed.

Thimmappa, R. et al. (2014). Triterpene biosynthesis in plants. *Annual Review of Plant Biology*, v. 65, p. 225–257. doi: 10.1146/annurev-arplant-050312-120229.

Tsutsui, H.; Higashiyama, T. (2017). PKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in *Arabidopsis thaliana*. *Plant and Cell Physiology*, v. 58, n. 1, p. 46–56. doi: 10.1093/pcp/pcw19.1.

Ulitsky, I.; Bartel, D. P. (2013). XlincRNAs: Genomics, evolution, and mechanisms. *Cell*, v. 154, n. 1, p. 26. doi: 10.1016/j.cell.2013.06.020.

Wada, M. et al. (2012). Prediction of operon-like gene clusters in the *Arabidopsis thaliana* genome based on co-expression analysis of neighboring genes. *Gene*, v. 503, p. 56–64. doi: 10.1016/j.gene.2012.04.043.

Wang, K. C. et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, v. 472, n. 7341, p. 120–126. doi: 10.1038/nature09819.

Werck-Reichhart, D.; Bak, S.; Paquette, S. (2002). Cytochromes P450. *The Arabidopsis Book*, v. 1, n. 1, p. e0028. doi: 10.1199/tab.0028.

Werck-Reichhart, D.; Feyereisen, R. (2000). Cytochromes P450 A Success Story. *Genome Biology*, v. 1, n. 6, p. 1–9. doi: 10.1186/gb-2000-1-6-reviews3003.

Xiang, T. et al. (2006). A new triterpene synthase from *Arabidopsis thaliana* produces a tricyclic triterpene with two hydroxyl groups. *Organic Letters*, v. 8, n. 13, p. 2835–2838. doi: 10.1021/ol060973p.

Zhang, X. et al. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols*, v. 1, n. 2, p. 641–646. doi: 10.1038/nprot.2006.97.

Zhong, J. J.; Yue, C. J. (2005). Plant cells: Secondary metabolite heterogeneity and its manipulation. *Advances in Biochemical Engineering/Biotechnology*, v. 100, n. July, p. 53–88. doi: 10.1007/b136412.